

**Model-data integration for predictive
assessment of groundwater reactive transport
systems**

Luca Carniato

Model-data integration for predictive assessment of groundwater reactive transport systems

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K. C. A. M. Luyben,
voorzitter van het College voor Promoties,

in het openbaar te verdedigen op 11 december 2014 om 12:30 uur

door

Luca Carniato

Environmental engineer, University of Padova
geboren te Treviso, Italië

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. N.C. van de Giesen

Copromotor: Dr.ir. G. H. W. Schoups

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. N.C. van de Giesen,	Technische Universiteit Delft, promotor
Dr.ir. G. H. W. Schoups,	Technische Universiteit Delft, copromotor
Prof. dr. T.N. Olsthoorn,	Technische Universiteit Delft
Prof. dr. T. J. Heimovaara,	Technische Universiteit Delft
Prof. dr. ir. A. W. Heemink,	Technische Universiteit Delft
Prof. dr. ir. R. Sethi,	Polytechnic University of Turin
Prof. dr. ir. P. Seuntjens,	University of Gent
Pror. dr.ir. J.B. van Lier	Technische Universiteit Delft, reservelid

Dr.ir. G. H. W. Schoups heeft als begeleider in belangrijke mate aan de totstandkoming van het proefschrift bijgedragen.

Keywords: Groundwater reactive transport, Permeable reactive barriers, Long term performance, Bayesian inference, Inverse modeling, Model simplification, High performance computing

Front & Back: Elemental analysis of a permeable reactive barrier filling. The parts containing oxygen are shown in violet, such as the sand grains and the mineral layer around the iron particles.

Copyright © 2014 by Luca Carniato

ISBN 978-94-6182-519-3

Summary

Predicting the evolution of groundwater contamination is a major concern for society, in particular when investments are made to remediate the contamination. Groundwater reactive transport models are valuable tools to integrate the available measurements in a consistent framework, improving our understanding of the physical system and potentially revealing aspects of the system that were not considered before. In this thesis reactive transport models are developed for predicting groundwater contamination at a field site where a zero-valent iron permeable reactive barrier (PRB) is installed. In doing so, several general challenges with integrating data into reactive transport models are addressed.

A first challenge is conceptual model uncertainty: available data should provide enough information for a correct conceptualization of the system. In chemical systems the concentration of each component is affected by several reactions and alternative reaction networks or descriptions of the reaction rates might provide similar model results. Moreover, the collection of new data might invalidate previous model conceptualizations and one might expect the conceptual model to be continuously updated as new information is acquired. Conceptual model uncertainties can have a large effect on reactive transport model predictions, yet they often are ignored. This thesis provides examples of quantification and reduction of conceptual uncertainty in reactive transport models. Quantification of conceptual model uncertainty was presented in the first part of this thesis, where a column scale experiment of a permeable reactive barrier is described. In the experiment, a substantial decline of the remediation performance over time was observed, mainly due to the development of a carbonate mineral coating around the iron particles. The column measurements were integrated in a multi-component reactive transport model, which subsequently provided predictions of long-term PRB efficiency under reduced flow conditions representative of the field site of interest. Different models of the deactivation process proposed in the literature were all able to reasonably well reproduce the column experiment measurements. The extrapolated long-term efficiency under different flow rates was however significantly different between the different models. These results highlight significant conceptual model un-

certainties associated with extrapolating long-term PRB performance based on lab-scale column experiments. Nevertheless, possible improvements of the experimental design were suggested. Despite the large amount of data collected in the column experiment, model conceptual uncertainties could not be resolved clearly and simpler deactivation models might be justified. A simplified deactivation model was proposed, where the decline of the barrier reactivity is simply proportional to the effective groundwater velocity, in contrast to the complex geochemical model where detailed mineral precipitation reactions are accounted for. The main advantage of the simplified model over the geochemical model is that it does not require inorganic concentration measurements for the inference of its parameters. However, the simplified model failed in reproducing part of the column measurements as it does not take into account the feedbacks present in the true geochemical system, which might be important for predicting the long term barrier performance. Reduction of conceptual model uncertainty was demonstrated for a field-scale reactive transport model, where alternative descriptions of the groundwater recharge process provided similar simulations of the groundwater levels but different simulations of the contaminant plumes, with different estimates of the hydraulic conductivity fields. In this case, model conceptual uncertainties were partially solved by integrating simultaneously in the model groundwater heads, concentrations and direct estimation of the hydraulic conductivities.

A second challenge concerns the data integration method. Since reactive transport models describe the evolution of multiple species in space and time, multiple data sets are used in parameter inference, for example measurements of dissolved concentrations of various chemical species. A common approach is to optimize parameters using an overall measure of fit, such as the weighted sum of squared residuals, where each data set is given a weight based on prior knowledge of the measurement and model errors. In practice, model errors are difficult to determine a priori and often they are much larger than measurements errors. Moreover, model errors induce correlations between residuals, which are typically ignored. A multivariate approach (MV) was proposed to integrate different data types, which accounts for specific correlations between residuals. The MV method allows integrating out the (co)variances (or weights) from the parameter posterior leading to an efficient estimation of model parameters, with no a priori assumptions about the magnitude of the errors. When applied to inference of the parameters of a column-scale reactive transport model, it was shown that accounting for residual correlation between species provides more accurate parameter estimation for high residual correlation levels, whereas its influence for predictive uncertainty is negligible. A

limitation of the multivariate method is that it cannot be implemented using a full covariance matrix with a dimension equal to the number of all measurements. Instead, residuals must be grouped such that the number of groups does not exceed the number of measurements per group.

A third challenge for field-scale models is the characterization of the spatial heterogeneity of soil properties from the measurements. Generally, little is known about heterogeneity and often soil properties are assumed constant within parts of the model domain or even in the entire model domain. However, heterogeneity has a considerable effect on the spreading of pollutants and must be accounted for to provide reliable predictions of contamination, increasing the number of model parameters to be estimated. Increasing the number of model parameters has two negative effects on parameter inference. First, it increases the computational effort. This is particularly relevant for reactive transport models characterized by long simulation times (ranging from several hours to days for field-scale applications). Second, a large number of model parameters can make the inverse problem ill-posed, where at least one of the criteria for well posedness (existence, uniqueness and stability) is not met. The reactive transport model developed for predicting the evolution of the contamination at the PRB site is affected by both of these two negative effects, since it accounts for spatial heterogeneity and several physical and chemical processes that increase the simulation time. The computational issue is solved using high performance computing. Ill-posedness is alleviated using a regularization procedure and a step-wise approach for integrating different data sets into the model. It is shown that the joint estimation of flow and transport parameters from head and concentration data improves the matching of the simulations to the measurements compared to the separate estimation of flow and transport parameters, in particular for the more mobile contaminants (cis-DCE and VC). Moreover, a more realistic estimation of the effective porosity and a reduction of the dispersion coefficients are obtained in the joint inversion. Additional measurements are required to validate the estimated recharge fractions and hydraulic conductivities.

In the last part of the thesis, the inferred hydraulic and chemical parameters were used to predict future contaminant migration at the field site. These scenarios include all processes considered in the parameter inference, also accounting for the iron deactivation process due to mineral precipitation. The results indicate that the PRB is expected to remain effective at degrading in-situ groundwater contamination until at least 2035. However, existing contamination downstream of the reactive barrier is likely to persist beyond 2035, due to the small dilution effect of infiltrating rain and the small biodegradation

rates at the site.

Samenvatting

Het kwantificeren van de migratie van contaminanten in grondwater is een belangrijke taak, met name bij de ondersteuning van grondwatersanering. Reactieve transportmodellen laten in principe toe om beschikbare kennis en data van de vervuilde site te integreren en te analyseren, zodat inzicht in de lokale grondwaterstroming en contaminantenmigratie verbetert. Deze thesis ontwikkelt reactieve grondwatertransportmodellen voor een case study met een reactieve ijzerwand. Hierbij worden een aantal uitdagingen kenmerkend voor het gebruik van reactieve transportmodellen aangepakt.

Een eerste uitdaging heeft te maken met conceptuele modelonzekerheid: beschikbare data moet voldoende informatie bevatten om een correct conceptueel model van het systeem te maken. In chemische systemen worden de concentraties van de verschillende componenten beïnvloed door een aantal chemische reacties die niet altijd eenduidig bepaald kunnen worden. Het is niet ongebruikelijk dat beschikbaarheid van nieuwe data leidt tot een herziening van het conceptuele model. Conceptuele modelonzekerheden kunnen een grote invloed hebben op de resultaten, en toch worden ze vaak over het hoofd gezien. Deze thesis geeft voorbeelden van hoe conceptuele modelonzekerheid van reactieve transportmodellen kan gekwantificeerd en gereduceerd worden. Kwantificatie van conceptuele modelonzekerheid wordt geïllustreerd in het eerste deel van de thesis aan de hand van een analyse van een kolomexperiment met een reactieve ijzerwand. In dit experiment werd een significante afname waargenomen waarmee de ijzerwand contaminanten in het grondwater afbreekt, veroorzaakt door het chemisch neerslaan van carbonaatmineralen rond de reactieve ijzerdeeltjes. Dit proces van ijzerdeactivatie is onvolledig gekend en vormt dus een bron van conceptuele modelonzekerheid. Verschillende mathematische modellen uit de literatuur werden toegepast op de data. Hieruit blijkt dat alle modellen een gelijkaardige fit geven met de data, maar dat ze sterk verschillende voorspellingen geven van de deactivatie bij een lagere doorstroomsnelheid. Deze resultaten illustreren de onzekerheden bij het extrapoleren van de langetermijnwerking van reactieve ijzerwanden op basis van (relatief korte) kolomexperimenten. Het deactivatieproces werd vervolgens verder onderzocht met een vereenvoudigd screening model dat niet gebaseerd is op com-

plexe geochemische reacties, maar dat de deactivatie voorspelt op basis van de doorstromingsnelheid. Dit model was niet in staat de data te reproduceren en de analyse toont daarmee het nut aan van de complexe geochemische modellen, ondanks de aanwezigheid van significante onzekerheid in het deactivatieproces. Conceptuele modelonzekerheid treedt ook op bij het modelleren van de vervuilingsspluim op de veldlocatie van dit onderzoek. Op deze site speelt grondwatervoeding door neerslag een belangrijke rol, en alternatieve modellen van dit proces leiden tot gelijkaardige simulaties van de grondwaterspiegel, maar grote verschillen in de beweging van de vervuilingsspluim. In dit geval worden observaties van zowel grondwaterniveaus als chemische concentraties gebruikt om de conceptuele modelonzekerheid met betrekking tot de grondwatervoeding te reduceren.

Een tweede uitdaging bij het gebruik van reactieve transportmodellen betreft de integratie van model en data. Aangezien dergelijke modellen de ontwikkeling van meerdere verbindingen in ruimte en tijd beschrijven, worden meerdere datasets gebruikt voor de afleiding van modelparameters, bijvoorbeeld de metingen van opgeloste concentraties en van verscheidene chemische verbindingen. Een gebruikelijke benadering is om parameters te optimaliseren met behulp van een algehele maat van optimalisatie, zoals de gewogen som van de gekwadrateerde afwijkingen, waarbij elke dataset een gewicht toegekend krijgt gebaseerd op voorafgaande kennis van de meet- en modelfouten. In de praktijk zijn modelfouten moeilijk voorafgaand te bepalen en vaak zijn ze veel groter dan meetfouten. Bovendien leiden modelfouten tot correlaties tussen de afwijkingen welke typisch worden genegeerd. Een multivariate benadering (MV) is voorgesteld om verschillende typen data te integreren, die specifieke correlaties tussen de afwijkingen in rekening neemt. De MV-methode stelt men in staat om (co)varianties (of gewichten) door integratie uit de resulterende waarschijnlijkheid van de parameters (posterior) te laten vallen. Het gevolg is een efficiënte schatting van modelparameters zonder voorafgaande aannames over de grootte van de fouten. Door toepassing op de afleiding van de parameters van een reactief transportmodel op kolomschaal is aangetoond dat het in rekening nemen van de correlatie van afwijkingen tussen verbindingen een nauwkeuriger parameterschatting oplevert voor hogere correlatieniveaus van afwijkingen, hoewel de invloed op de voorspelbare onzekerheid verwaarloosbaar is. Een beperking van de multivariate methode is dat deze niet geïmplementeerd kan worden bij het gebruik van een volledige covariantiematrix met een dimensie gelijk aan het aantal metingen. In plaats daarvan moeten afwijkingen gegroepeerd worden zodat het aantal groepen niet meer wordt dan het aantal metingen per groep. Een derde uitdaging voor modellen op veldschaal is

de karakterisering van de ruimtelijke heterogeniteit van bodemeigenschappen vanuit de metingen. Over het algemeen is er weinig bekend over heterogeniteit en vaak worden bodemeigenschappen constant verondersteld binnen delen van het modeldomein of zelfs het gehele modeldomein. Echter, heterogeniteit heeft een behoorlijk effect op de verspreiding van vervuilingen en moet in rekening worden genomen om betrouwbare voorspellingen te verschaffen van de vervuiling, waardoor het aantal te schatten modelparameters toeneemt. Het opschroeven van het aantal modelparameters heeft twee negatieve gevolgen voor de afleiding van parameters. Ten eerste verhoogt het de benodigde rekenkracht. Dit is in het bijzonder van belang bij reactieve transportmodellen die gekarakteriseerd worden door lange simulatietijden (variërend van meerdere uren tot dagen voor toepassingen op veldschaal). Ten tweede kan een groot aantal modelparameters het inversieprobleem slecht geconditioneerd maken, waarbij aan tenminste één van de criteria voor een correcte conditionering (het bestaan, het uniek zijn en stabiliteit) niet wordt voldaan. Het reactieve transportmodel dat ontwikkeld is voor het voorspellen van de verspreiding van de vervuiling op de PRB-locatie heeft te maken met beide negatieve gevolgen, omdat het ruimtelijke heterogeniteit en meerdere fysische en chemische processen in rekening neemt welke de simulatietijd vergroten. Het rekenkundige probleem is opgelost door middel van hogeprestatieberekeningen. De slechte conditionering is verlicht door middel van een reguleringsprocedure en een stapsgewijze benadering voor het integreren van verschillende datasets in het model. Het is aangetoond dat de gezamenlijke schatting van stromings- en transportparameters vanuit waterstands- en concentratiedata de overeenkomst tussen simulaties en metingen verbetert vergeleken met een afzonderlijke schatting van stromings- en transportparameters, in het bijzonder voor de mobielere vervuilingen (cis-DCE en VC). Bovendien worden een realistischere schatting van de effectieve porositeit en een afname van de dispersiecoëfficiënten verkregen in de gezamenlijke inversie. Extra metingen zijn noodzakelijk om de geschatte wegzijgingsfracties en hydraulische geleidbaarheden te valideren. In het laatste deel van het proefschrift zijn de afgeleide hydraulische en chemische parameters gebruikt om de toekomstige vervuilingsmigratie op de veldlocatie te voorspellen. Deze scenario's omvatten alle processen die beschouwd zijn in de parameterafleiding, alsmede het ijzerdeactivatieproces als gevolg van mineraalneerslag.

De resultaten tonen aan dat de PRB naar verwachting effectief blijft voor de degradatie van in situ grondwaterverontreinigingen tot tenminste 2035. De bestaande verontreiniging benedenstrooms van de reactieve ijzerwand duurt echter naar alle waarschijnlijkheid voort na 2035 als gevolg van het geringe

verdunningseffect van infiltrerende regen en de geringe biodegradatiesnelheden ter plekke.

Contents

Summary	v
Samenvatting	ix
List of Symbols	xvii
1. Introduction	1
1.1. Background	1
1.2. Objectives	4
1.3. Outline of the thesis	5
2. Site description	9
2.1. Measurements	10
2.2. Biodegradation lab experiments	14
3. Predicting longevity of iron permeable reactive barriers using multiple iron deactivation models	19
3.1. Introduction	20
3.2. Materials and Methods	22
3.2.1. Column design	22
3.2.2. Mineral analysis and analytical methods	22
3.2.3. Numerical model	24
3.2.4. Flow and transport parameters	27
3.2.5. Parameter inference	28
3.3. Results and discussion	29
3.3.1. Mineral analysis	29
3.3.2. Column experimental results	31
3.3.3. Calibration results	34
3.3.4. Long-term performance prediction	38
3.3.5. Other sources of uncertainty in the long term prediction and strategies to reduce them	41
3.4. Conclusions	42

4. Comparing longevity estimations provided by geochemical and simplified models	45
4.1. Simplified model description	45
4.2. Parameter inference	48
4.3. Comparison between the geochemical and simplified model on the short term column data	50
4.4. Long term predictions	52
4.5. Conclusions	56
5. Inference of reactive transport model parameters using a Bayesian multivariate approach	59
5.1. Introduction	60
5.2. Bayesian framework	62
5.3. Application to reactive transport model	68
5.3.1. Model modifications	68
5.3.2. Case studies	70
5.3.3. Assessing model structural error	71
5.4. Results	72
5.4.1. Effect of residual correlation	72
5.4.2. Effect of model structural error	76
5.4.3. Computational efficiency	87
5.5. Conclusions	88
6. Site model	91
6.1. Introduction	91
6.2. Groundwater flow model	94
6.2.1. Aquifer properties heterogeneity	96
6.2.2. Recharge	97
6.3. Reactive transport model	99
6.3.1. Contaminant dissolution from the source areas	100
6.3.2. Contaminant sorption to aquifer material	101
6.3.3. Abiotic contaminant degradation within the reactive barrier	102
6.3.4. Treatment of contaminant biodegradation and substrate consumption	103
7. Inverse modelling at the site and prediction of the PRB efficiency	113
7.1. Non linear inversion theory	113
7.1.1. Parameter and prediction uncertainty	117
7.2. Inversion procedure	119

7.3. Optimization of the flow parameters	120
7.3.1. Full recharge model results	120
7.3.2. Reduced recharge model results	123
7.3.3. Effect of low regularization constrains	126
7.4. Conservative transport simulations	127
7.4.1. Description of the experiments	127
7.4.2. Results	129
7.5. Optimization of reactive transport parameters	132
7.5.1. Description of the experiments	132
7.5.2. Simulations with degradation rates determined from batch experiments	133
7.5.3. Manual tuning of the degradation rates	133
7.5.4. Inversion with equal weights	134
7.5.5. Inversion with weights proportional to the inverse of the concentrations	139
7.6. Joint optimization of flow and reactive transport parameters .	141
7.6.1. Joint optimization results	144
7.7. Geochemical scenarios	151
7.7.1. Effect of the barrier installation on contamination . . .	156
7.7.2. Effect of the barrier installation on inorganic concentra- tions	157
7.7.3. Long term PRB efficiency	159
7.8. Conclusions	160
8. Conclusions and future research	163
8.1. Conclusions	164
8.2. Future research	165
A. Appendix to chapter 3	169
B. Appendix to chapter 5	179
C. Appendix to chapter 6	185
D. Appendix to chapter 7	187
Bibliography	193
Acknowledgments	209
About the author	211

Publications	213
---------------------	------------

List of Symbols

ϕ	vector of model parameters
α_{CC}	deactivation constant for carbonate
α_{IC}	deactivation constant for iron hydroxy carbonate
β^2	regularization weight factor
μ	maximum degradation rates
Φ_m	measurement objective function
Φ_r	regularization objective function
Σ	residual covariance
σ_{ME_j}	measurement error variance for set j
σ_{MO_j}	model error variance for set j
θ_i	slope of the front for contaminant i (STS model)
$\tilde{C}_k$	solid phase concentration of contaminant k ($mol\ mol^{-1}$)
$\tilde{r}_k$	global reaction rate of contaminant k in the solid phase ($mol\ mol^{-1}\ d^{-1}$)
ε	residual vector or residual matrix
φ_{Fe^0}	iron volume fraction (-)
φ_i	mineral i volume fraction (-)
ϱ_b	dry bulk density ($kg\ L^{-1}$)
C	prior covariance matrix of the model parameters
C_k	concentration contaminant k ($mol\ L^{-1}$)

$C_{k,i}^*$	equilibrium concentration in groundwater for contaminant k at source i ($mol\ L^{-1}$)
C_{k-SOL}	contaminant k solubility
C_{Wk}	concentration of the source/sink term ($mol\ L^{-1}$)
D_{ij}	hydrodynamic dispersion coefficient ($m^2\ d^{-1}$)
DP	deactivation period (STS model)
f_{DEC}	stoichiometric coefficient for hydrogen consumption by dechlorinator bacteria
F_i	scaling factor for contaminant i (STS model)
f_{oc}	carbon content in the soil (-)
$f_{SUB/VOC}$	moles of substrate required for the degradation of one mole of contaminant
H_2	hydrogen concentration ($mol\ L^{-1}$)
H_2^*	hydrogen threshold concentration for dechlorinators ($mol\ L^{-1}$)
H_2^{scale}	inhibitory aqueous hydrogen concentration ($mol\ L^{-1}$)
I_{H_2}	hydrogen inhibition factor
IAP	ion activity product
K_i^m	solubility of mineral i
K_{d-k}	distribution coefficient for contaminant k ($L\ kg^{-1}$)
$k_{eff,i}$	precipitation rate coefficient for mineral i ($mol\ L^{-1}\ s^{-1}$)
$k_{h,CO_3^{2-}}$	half saturation constant for iron corrosion by carbonate
K_{H_2}	half saturation constant for hydrogen (sulfate biodegradation, $mol\ L^{-1}$)
$k_{IrC,TIC}$	rate coefficient for iron corrosion by carbonate ($mol\ m^{-2}\ s^{-1}$)
$k_{IrC,W}$	rate coefficient for iron corrosion by water ($mol\ m^{-2}\ s^{-1}$)
K_{IrC}	equilibrium constant for iron corrosion by water

List of Symbols

$K_{k-PRB\ 1/2}$	half saturation constant for degradation of contaminant k within the PRB ($mol\ L^{-1}$)
k_{la}	contaminant mass transfer coefficient
K_{oc-k}	contaminant k partition coefficient ($L\ kg^{-1}$)
k_{PRB-k}	rate coefficient for degradation of contaminant k within the PRB
$k_{SO_4^{2-}-H_2}$	rate coefficient of sulfate biodegradation ($mol\ L^{-1}\ s^{-1}$)
$K_{SO_4^{2-}}$	half saturation constant for sulfate (sulfate biodegradation, $mol\ L^{-1}$)
por	porosity (-)
q_i	Darcy flow in the i direction ($m\ d^{-1}$)
r_k	global reaction rate of contaminant k in the aqueous phase ($mol\ L^{-1}\ d^{-1}$)
r_{k-DEG}	biodegradation rate of the compound k
r_{k-PRB}	degradation rate of the contaminant k within the PRB
r_{k-SOU}	release rate of the compound k from the sources
R_k	retardation factor for contaminant k (-)
RR_i	residual reactivity for contaminant i (STS model)
S	matrix of the residual products
S_r	iron reactive surface area ($m^2\ L^{-1}$)
SE	standard error
SUB	substrate concentration ($mol_{SUB}\ L^{-1}$)
T_c	average thickness of the mineral coating around the iron particles
$TVOC$	total concentration of volatile organic compounds
TZ	transition zone (STS model)
V_{im}	molar volume of mineral i
V_i	total deactivated iron volume for contaminant i (STS model)

$V_{ref,i}$	iron volume deactivated to the minimum reactivity for contaminant i (STS model)
$V_{TZ,i}$	normalized deactivated volume within the regions of full and minimum reactivity for contaminant i (STS model)
W	source/sink term in the groundwater flow and transport equations (d^{-1})
w_i	weight assigned to observation set i
X	Jacobian matrix of the model η
$x_{max,i}$	distance from the column inlet travelled by the precipitation front for contaminant i (STS model)
x_{ref}	column length (STS model)
y_{ij}	measurement j of the observation set i
CC	calcium carbonate (aragonite)
cis-DCE	cis-1,2-dichloroethene
DCA	1,2-Dichloroethane
DNAPL	dense non aqueous phase liquid
DREAM-ZS	DiffeRential Evolution Adaptive Metropolis algorithm
EMPA	electron micro probe analyzer
GLS	generalized least squares method
IC	iron hydroxy carbonate
IH	iron hydroxide
MAP	maximum a posteriori
MIN3P	saturated/unsaturated groundwater reactive transport model
MODFLOW	modular groundwater flow model
MV	multivariate method
OLS	ordinary least squares method
PCE	tetrachloroethene

PEST	parameter estimation software
PHAST	geochemical reactive transport simulator based on PHREEQC
PHREEQC	model for simulating chemical reactions in natural waters
PHT3D	geochemical reactive transport simulator based on PHREEQC and MODFLOW
ppb	parts per billion
PRB	permeable reactive barrier
PVs	pore volumes
RMSE	root mean squared error
RT3D	reactive transport model based on MODFLOW
SCE-UA	Shuffled Complex Evolution algorithm
STS	spatial temporal simulation model
TCA	1,1,1-trichloroethane
TCE	trichloroethene
TEAPs	terminal electron acceptors
TOC	total organic carbon
trans-DCE	trans-1,2-dichloroethene
VC	vinyl chloride
VITO	Flemish institute for technological research
VOCs	volatile organic compounds
WLS	weighted least squares method
WLS(we)	weighted least squares with weight estimation method
XRD	X-Ray diffraction
BDL	below detection limit
η	model-predicted values
Φ_l^m	user specified target measurement function

h	hydraulic head (m)
K	saturated hydraulic conductivity ($m\ d^{-1}$)
S_s	specific storage (m^{-1})
S_y	specific yield (-)
Z	model inputs (e.g. boundary conditions)

1. Introduction

1.1. Background

Planning and management of groundwater resources often requires mathematical models able to simulate subsurface processes. In the field of groundwater remediation a mathematical model might serve as a tool for predicting the efficiency of a remediation technology. Intuitively, a mathematical model is a set of equations able to simulate the excitation-response relation of a real system [143]. Groundwater models are typically based on partial differential equations describing the subsurface flow and transport phenomena. These equations contain parameters whose values need to be inferred from the data, known as inverse modeling or parameter inference. The solution of the inverse problem plays a key role in groundwater modeling, since no reliable predictions can be made with a model unable to mimic the measurements.

In Fig. 1.1 a common framework for the development of a groundwater reactive transport model from a set of data (e.g. groundwater levels and/or dissolved concentrations), the inference of its parameters and its validation is shown. In the first step, the data are interpreted and used to develop a conceptual model, which qualitatively represents the processes and the relations between the processes considered relevant for the simulation of the natural system. Afterward, the processes and their relations are formalized into equations defining a mathematical model capable of describing the excitation-response relation of the real system. In real applications these equations are not solvable analytically (except for cases where strong simplifying assumption are made) and a numerical model must be developed, where the exact mathematical equations are approximated with discretized equations. The solution of the numerical model requires the specification of boundary and initial conditions (BC and IC), driving forces (e.g. rainfall) and parameter values pertaining to the particular time and space discretization used in the numerical model. The values of these effective parameters are not known and need to be inferred from the measurements, by varying the parameter values until a reasonable fit (commensurate with the estimated measurement and model error) to the measurements is obtained. Besides the fit to the measurements, the model is

deemed calibrated also when the prior assumptions made in the inversion are met (for example normality of the weighted residuals) and reasonable parameter values are obtained. After estimation of model parameters, the model is tested under different conditions (BC, IC and driving forces) to verify its ability to reproduce a different set of measurements. This step, known as validation, is crucial to assess the predictive capabilities of the model and might provide useful indications for further model development. Once the model simulations fit the calibration and validation sets to a reasonable level, model predictions can be made. Besides predictions, the calibrated model can be used to optimize data collection, indicating the measurements that minimize prediction uncertainty.

There are several challenges in applying this framework to develop and infer the parameters of reactive transport models at lab and field scale.

Challenge 1: model conceptual uncertainty A first challenge is the conceptualization step. The available data should provide enough information to allow the correct conceptualization of the natural system. In chemical systems the concentration of each component is affected by several reactions and alternative reaction networks or descriptions of the reaction rates might provide similar model results [102, 108]. Moreover, the collection of new data might invalidate previous model conceptualizations and one might expect the conceptual model to be continuously updated as new information is acquired [16]. To reduce the uncertainty associated with a non-unique conceptualization additional data can be collected, for example measuring oxygen concentration in groundwater to determine if biodegradation occurs under aerobic or anaerobic conditions or performing separate experiments to investigate only a particular process (e.g. biodegradation batch tests).

Another example is the characterization of the contaminant sources. Contaminant sources are commonly treated as zones at constant concentration or zones containing a contaminant mass that dilutes over time. Such conceptual model does not account for a more realistic situation where the contaminants are released over a longer time period. By collecting additional measurements the conceptual model can be improved and prediction uncertainty reduced. When additional data can not be collected, uncertainties in process knowledge can be quantified by performing several predictions with different conceptual models. Conceptual model uncertainties can play a large effect when using reactive transport models for long-term predictions.

Challenge 2: multiple data sources A second challenge concerns the integration of models and data. Since reactive transport models describe the evolution of multiple species in space and time, multiple data sets are used in parameter inference, for example measurements of dissolved concentrations of various chemical species. In principle, this results in a multi-objective optimization problem, where one tries to fit multiple data sets all at once. A common approach is to optimize parameters using an overall measure of fit [2, 32, 31, 9, 57], such as the weighted sum of squared residuals (differences between observed and simulated values), where each data set is given a weight based on prior knowledge of the measurement and model errors (e.g. misspecified BC or IC). In practice, model errors are difficult to determine a priori and often they are much larger than measurements errors. Moreover, model errors induce correlations between residuals, which are typically ignored although it affects the parameter estimation [41]. Therefore, there is a need for efficient inference methods that optimally combine information from various data sources, while accounting for relevant observation and model errors.

Challenge 3: spatial heterogeneity A third challenge for field-scale models is the characterization of the spatial heterogeneity of soil properties (e.g. hydraulic conductivity and specific yield) from available measurements. Generally, little is known about heterogeneity and often soil properties are assumed constant within parts of the model domain or even in the entire model domain [151, 74, 30]. However, heterogeneity has a considerable effect on the spreading of pollutants and must be accounted for to provide reliable predictions of contaminant spreading, thereby increasing the number of model parameters to be estimated from limited data [52, 89, 168, 44, 147].

Increasing the number of model parameters has two effects on parameter inference. First, it increases the computational effort, since more forward model runs are generally required. This is particularly relevant for reactive transport models, which are often characterized by long simulation times (ranging from several hours to days for field-scale applications), requiring the use of high-performance or parallel computing. Second, a large number of model parameters can make the inverse problem ill-posed, where at least one of the criteria for well posedness (existence, uniqueness and stability) is not met. A common technique to solve this problem is to introduce regularization, stabilizing the inversion and limiting the search of the solution within the parameter space defined by the modeller [4, 38, 145]. The strength of the regularization constraints is often determined a priori, for example specifying the desired level of model fit to the measurements. The use of highly parameterized in-

version in reactive transport models remains a challenging problem, and its application on real groundwater reactive transport systems is still limited.

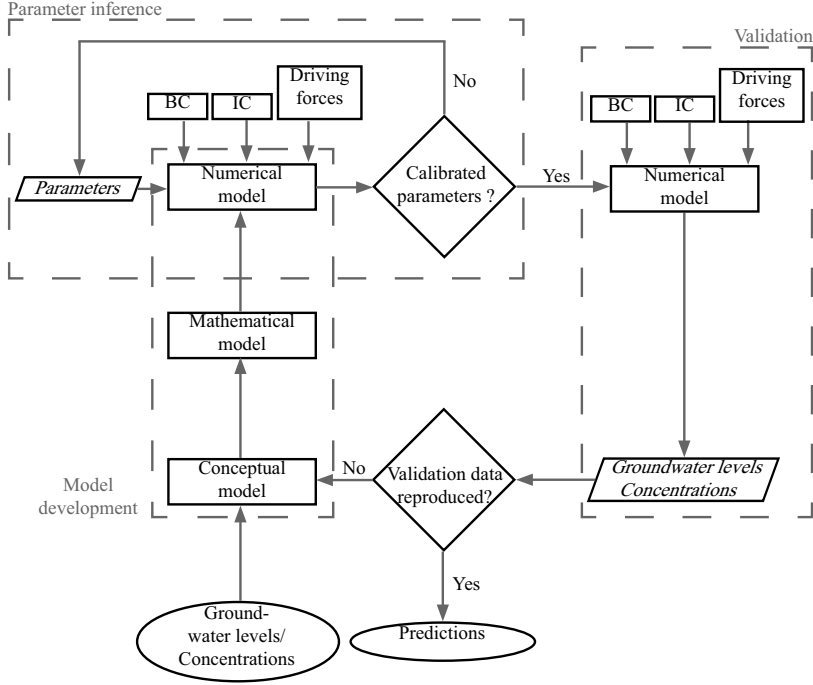


Figure 1.1.: A general framework illustrating the development of a groundwater reactive transport model from the data, the inference of its parameters and its validation. BC and IC indicate the boundary and initial conditions respectively.

1.2. Objectives

The objectives of this thesis are to address the challenges mentioned in the previous section, within the context of remediation of contaminated groundwater using in-situ chemical and biochemical processes. In these applications, groundwater reactive transport models provide a means of integrating site-specific knowledge and monitoring data for evaluating current and predicting

future contaminant plume behavior.

Challenge 1: model conceptual uncertainty The first objective of this thesis is to quantify conceptual uncertainty of groundwater reactive transport models and its effect on predictions of groundwater contaminant concentrations. Instead of using a single conceptual model, multiple a priori plausible process descriptions are considered and evaluated against available data. This strategy is applied to modelling geochemical processes in iron-based permeable reactive barriers (PRB) for cleaning up groundwater contaminated with chlorinated solvents (chapters 3 and 4). Alternatively, a strategy of iterative model refinement is proposed and applied to modelling aquifer recharge and its effect on dilution of groundwater contaminant concentrations (chapters 6 and 7).

Challenge 2: multiple data sources The second objective of this thesis is to develop and apply methods for integrating data from multiple sources (groundwater levels and groundwater concentrations of multiple chemical species) into groundwater reactive transport models. Methods tested and applied range from heuristic techniques for weighting data from different sources (chapters 3 and 7) to a multivariate Bayesian approach that formally integrates multi-species data, while accounting for correlations (chapter 5).

Challenge 3: spatial heterogeneity The third objective of this thesis is to develop and apply groundwater reactive transport models with spatially variable hydraulic parameters and groundwater recharge, based on limited data from multiple sources (groundwater levels, groundwater concentrations of multiple chemical species, point-scale measurements of soil hydraulic conductivity, and lab-scale derived information on chemical and biochemical reactions). This is done for a field-site where groundwater contaminated by chlorinated solvents is being cleaned up using a combination of chemical processes (by means of an iron PRB) and biochemical processes (natural attenuation). The proposed methodology relies on regularization to reconcile the highly parameterized model with the limited data available (chapter 7).

1.3. Outline of the thesis

In chapter 2 the contaminated field site and the data used in this thesis are described. Groundwater at the site is contaminated with chlorinated solvents

and remediation is being done by a combination of an iron PRB and natural attenuation.

In chapter 3 lab-scale experimental data are used to study geochemical processes in iron-based PRB. Specifically of interest is the decline in remediation efficiency of these barriers over time, due to mineral precipitation in the pore space of the PRB, resulting in deactivation of the reactive iron material and potential clogging of the barrier. To account for the conceptual uncertainty of these processes, four different geochemical deactivation models are proposed based on experimental observations (consisting of mineral analysis and geochemical concentration measurements) and results from previous studies [66]. This a priori postulated set of models is evaluated on data from the experiment and used to quantify the effect of model uncertainty on prediction of long-term PRB remediation efficiency.

In chapter 4 an alternative, simplified model of the iron deactivation process in PRB is proposed and compared to the geochemical models of chapter 3. The model describes the deactivation front with geometrical formulas without explicitly accounting for mineral precipitation. In order to predict long-term PRB remediation efficiency, the simplified model was adapted from its original formulation proposed by Kouznetsova et al. [84].

In chapter 5 a multivariate Bayesian approach is presented that formally integrates multi-species data, while accounting for correlations between different data sources. This approach is compared to existing weighted least squares methods, using the models and data reported in chapter 3.

In chapter 6 a three-dimensional groundwater reactive transport model is proposed and developed for predicting groundwater contaminant transport at the field site described in chapter 2. The model includes relevant flow and transport processes, and accounts for contaminant sorption, biodegradation (natural attenuation), and contaminant degradation in the PRB. Special attention is paid to modelling aquifer heterogeneity (spatial variability of hydraulic parameters) and spatial variations and model conceptual uncertainty in recharge entering the aquifer at the site.

In chapter 7 the spatially distributed parameters of the reactive transport model introduced in chapter 6 are inferred from available field measurements. The methodology relies on (i) pilot points to quantify spatial heterogeneity of hydraulic parameters and recharge, (ii) a regularization procedure to reconcile the large number of spatially distributed model parameters with the limited field data, (iii) a step-wise approach for integrating different data sets into the model, and (iv) high performance computing. The strategy of step-wise integration of groundwater levels and groundwater concentration data is com-

pared to direct joint estimation of all model parameters from all data. Finally, the model is used to predict future contaminant migration and remediation efficiency at the site.

In chapter 8 the results are synthesized and future research directions are outlined.

2. Site description

The studied contaminated site, depicted in Fig. 2.1, is located in Wilrijk, just south of Antwerp, Belgium. The factory shown in the figure began production of compressors around 1951 and used chlorinated solvents mainly for painting and degreasing activities. The first use of chlorinated solvents (VOCs) occurred in 1957, including tetrachloroethene (PCE), trichloroethene (TCE) and 1,1,1-trichloroethane (TCA), hereafter referred to as source contaminants. The use of VOC stopped in 1996, when it was banned by Flanders environmental law. Historically, the use of chlorinated solvents at the site has resulted in contamination of the underlying aquifer, generating a contamination plume spreading from the factory to a pasture area next to the factory. The aquifer mainly consists of quaternary sandy clay deposit, underlain by a low-permeable thick clay deposit (“Boom formation”) at an average depth of 4.5 m below land surface (<https://dov.vlaanderen.be>). Groundwater flows laterally on top of the clay layer from factory to pasture, as shown in Fig. 2.1 by contoured average groundwater levels (based on 731 measurements between 2001 and 2012). As can be seen, the hydraulic gradient is higher in the pasture area as compared to the hydraulic gradient in the factory area, indicating a different effective groundwater velocity between the two parts of the site. The average hydraulic conductivity calculated from slug tests is 0.19 m d^{-1} . A cross section of the aquifer of interest is shown in Fig. 2.4.

In order to delineate the extent of the contamination and design a remediation strategy, contaminant concentrations, including source contaminants and their products (mainly 1,2-dichloroethane (DCA), cis-dichloroethane (cis-DCE) and vinyl chloride (VC)), were measured in 123 piezometers between 2000 and 2005. Most of the piezometers were screened one meter above the clay layer, since most of the contaminants were found at that depth. In Fig. 2.2 the resulting average total contaminant concentration map is shown. As can be seen from the figure, multiple separate plumes have developed at the site.

In October 2005, a permeable reactive barrier (PRB) was installed at the site (purple line in Fig. 2.1). The barrier consists of two trenches, separated by an impermeable bentonite section, each approximately 95 meter long and 0.3 meters wide, penetrating the entire depth of the phreatic aquifer. The

trenches were filled with a mixture of fine iron and sand (20% and 80% in volume, respectively). After barrier installation, monitoring of contaminant concentrations focused on areas near the barrier and in the pasture (black piezometers in Fig. 2.2).

Based on the records of the historical activities, the presumed source locations are numbered in Fig. 2.2 and described in Tab. 2.1. The contamination in the northern part of the factory is clearly separated from the contamination close to the PRB and in the pasture area, allowing to focus the modeling effort on the most monitored part. All monitoring data collected close to the barrier were used in the modeling, a selected group of piezometers are labeled and numbered in Fig. 2.4 to simplify the discussion. Their original names are reported in Tab. 2.2.

2.1. Measurements

Most of the concentration and groundwater measurements available for the site were collected as a part of the compulsory monitoring program for the local environmental authorities, which mainly consist on punctual groundwater levels and concentrations. In this thesis more groundwater heads and concentration measurements were collected. A graphical summary of the amount of the data available for the modeling of the site is shown in Fig. 2.3. Below a description of the measurements used in this thesis is provided.

- Groundwater levels: groundwater levels from 2000 to 2012 are available from the monitoring programs. However, in this database only the levels collected between 2001 and 2003 cover the entire site area. In order to provide more data for the model a bimonthly monitoring campaign of the groundwater levels was performed in 2011 in 62 piezometers distributed over the site (Fig. 2.1). This added 372 groundwater levels to the database.
- Concentration data: groundwater quality was monitored in the piezometers shown in Fig. 2.2. Most of the concentration measurements that were collected before the barrier installation consist of individual measurements aimed to distinguish polluted from not polluted areas. After the barrier installation contaminant concentrations were collected only in the black piezometers shown in Fig. 2.2, twice a year during the period 2005-2009 and three times a year during the period 2010-2011. Inorganic concentrations (dissolved carbonate, bicarbonate, chloride, sulphate, calcium, iron and magnesium) were collected in a subset of the

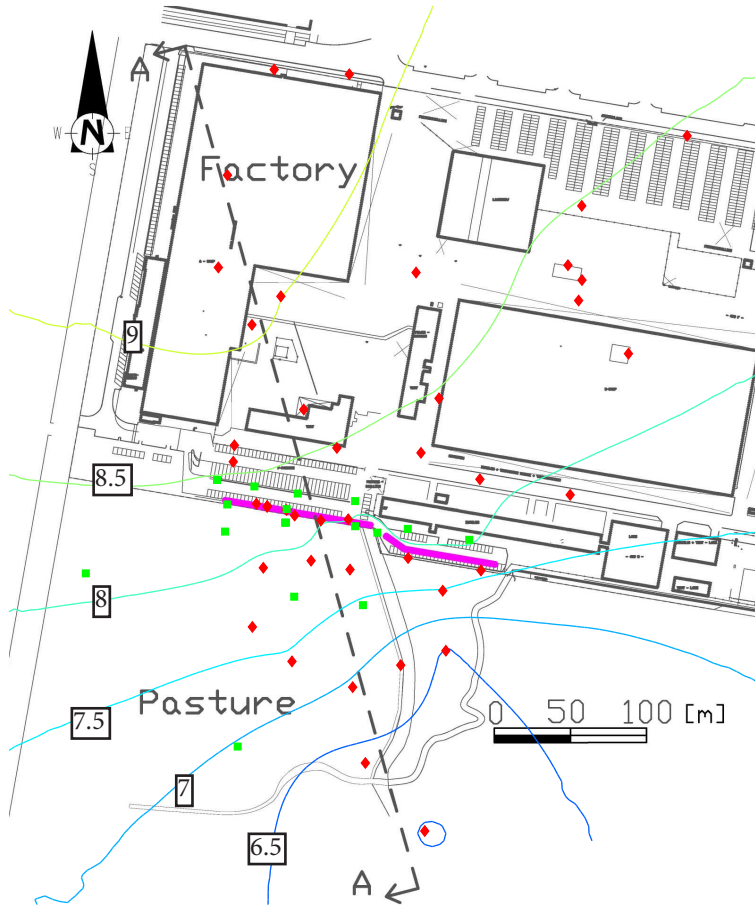


Figure 2.1.: Areal view of the site. In 2011 groundwater levels were collected in all indicated piezometers with a bimonthly frequency. Slug tests were performed in the piezometers indicated by green squares. The purple line shows the location of the PRB.

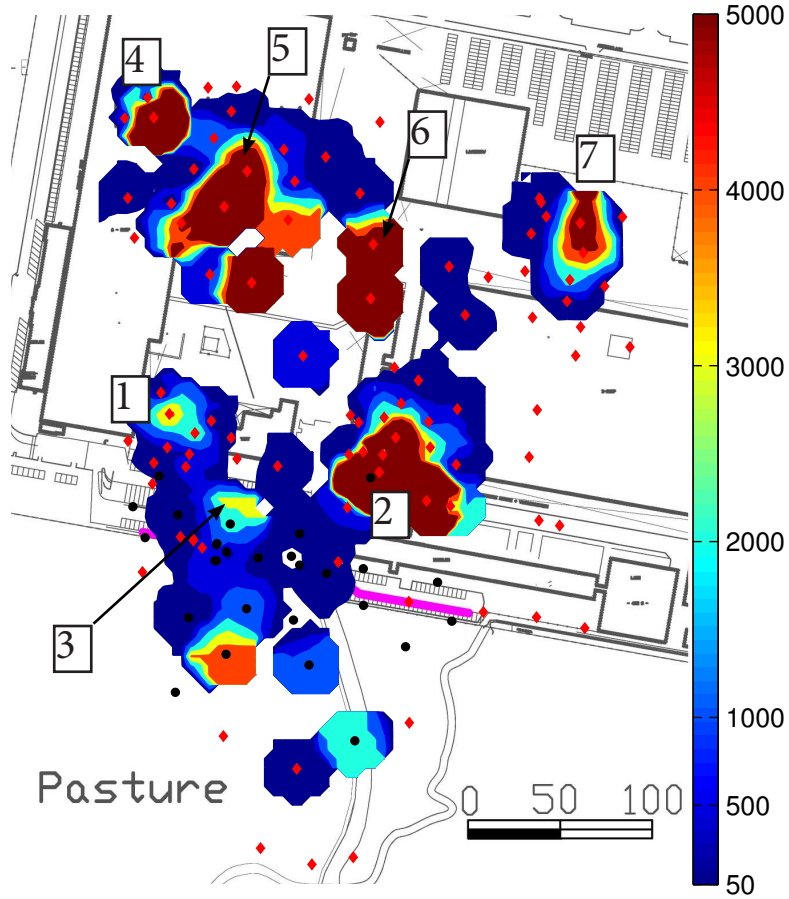


Figure 2.2.: Interpolated average total contaminant concentration in parts per billion (ppb). After the barrier installation, only the piezometers indicated by the black dots were monitored. The presumed source locations are numbered.

Table 2.1.: Presumed release periods and contaminant types for the sources indicated in Fig. 2.2

Source	Period	Activity	Contaminants
1	1957-1994	Compressors tests	PCE, TCE
2	1966-1996	Degreasing, painting	TCA
3	1957-?	Dumping area	PCE, TCE
4	1970-1996	Degreasing, painting	TCE, TCA
5	1970-1992	Degreasing, painting	TCE, TCA
6	1966-1985	Degreasing, painting	TCE
7	1965-1975	Degreasing, painting	TCE, TCA

black piezometers with a similar temporal frequency (two times a year in 2010, 2009 and 2006 and three times a year in 2011).

- Ground surface and clay layer elevations: to define the top and bottom limits of the phreatic aquifer, the ground surface and clay layer elevations are required. Land surface elevation was measured in the piezometers indicated in Fig. 2.1. Based on these measurements, the land surface was reconstructed using kriging interpolation. The clay elevation at the piezometers was obtained by subtracting from the interpolated land surface surface the clay depths measured in the deep piezometers. The clay layer elevation was interpolated using 142 clay depth measurements. As can be seen from the cross section reported in Fig. 2.4, the interpolated clay elevation is irregular, especially in the pasture area where more depth measurements were taken closely together.
- Divers: divers allow for high-frequency measurements of the hydraulic pressure. In this site groundwater levels were measured every 15 minutes in six piezometers: 26, 27, 4, 20, 15, 17. Close to the barrier divers were installed in two shallow piezometers (26 and 27, screened from groundwater surface to 2 meters below it) and in two deep piezometers (4 and 20, screened from 1 meter above the clay bottom to the clay bottom). Other divers were installed in the pasture area. The aim of the diver measurements was to corroborate the hypothesis that groundwater flow direction reverses during strong rain events, due to the different infiltration patterns in the factory area (paved) and in the pasture area (unpaved). This was demonstrated by the inversion of the head gradient between two shallow piezometers (Fig. 2.3(b)). The inversion of the head gradient in the deep piezometers 4 and 20 was negligible, indicating that

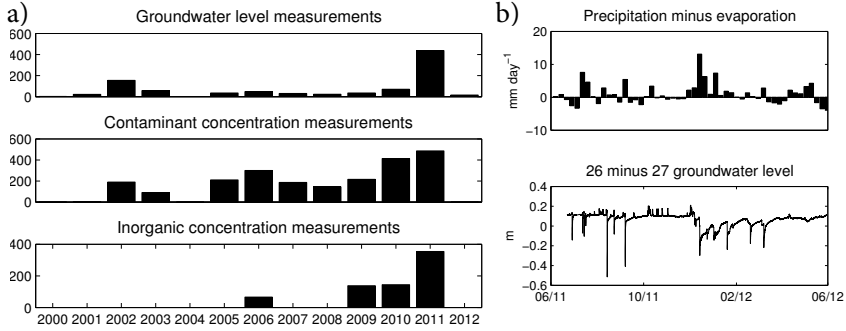


Figure 2.3.: a) Amount of groundwater levels, contaminant concentrations and inorganic concentrations measurements collected at the site. b) Weekly precipitation minus evaporation at the site and differences between the groundwater levels at piezometers 26 and 27 (diver data).

the infiltrating rainwater might dissipate in the shallow aquifer part.

- **Slug tests:** slug tests were performed on 17 piezometers to estimate the hydraulic conductivity under two different conditions. In the first condition water was poured into the piezometer (slug in), while in the second condition water was pumped out from the piezometer (slug out). The drawdown was analyzed with the Bouwer and Rice method [14]. The hydraulic conductivities estimated in the slug in tests were on the average one order of magnitude smaller than the slug out tests. The estimated hydraulic conductivities were used as prior information in the inverse modeling at the site (chapter 7).
- **Lab experiments:** biodegradation experiments were performed to assess the biodegradation potential as described in the following section. Column experiments were performed to investigate the performance of the permeable reactive barrier over time, as described in detail in chapter 3.

2.2. Biodegradation lab experiments

For five piezometers batch experiments containing 68 milliliters (ml) of polluted groundwater and 30 grams of aquifer material in a 160 ml bottles were performed by the Flemish institute for technological research (VITO) to assess the biodegradation potential. The experiments lasted for 384 days. The con-

2.2 Biodegradation lab experiments

taminant concentrations were determined from the gaseous concentrations in the headspace (67 ml), assuming equilibrium between liquid and gas phases. Experiments were performed in triplicates for each piezometer. Samples of aquifer material and polluted groundwater were collected from piezometers 9, 10, 21, 16 and 23, along the presumed direction of the plume originating from source 3.

For the first four piezometers (9, 10, 21, 16) no substantial contaminant degradation was observed (with a clear increase of daughter products concentration after parent compounds degradation). In all these batch experiments the total organic carbon content (TOC) was lower than 0.3%. It is hypothesized that organic carbon can be consumed by biodegradation and after its depletion the biodegradation potential becomes negligible. Fig. 2.5(a) and Fig. 2.5(b) show the contaminant concentrations measured in two batch experiments performed with piezometer 16 groundwater and aquifer material. After 295 days additional groundwater was added, increasing the concentration of all contaminants. As can be seen, the contaminant profiles are flat during the first 295 days. However, in the second experiment (Fig. 2.5(b)), lactate was introduced on day 314, causing an abrupt decline of PCE and TCE concentrations and an increase of cis-DCE and VC.

A similar phenomena was observed for piezometer 10 (Fig. 2.5(c) and Fig. 2.5(d)). In this case polluted groundwater was added after 169 and 290 days. In first batch (Fig. 2.5(c)) no substantial contaminant degradation was observed whereas in the second batch experiment (Fig. 2.5(d)) a decline of PCE and TCE concentrations was observed after the addition of 500 μ M (micro moles)

Table 2.2.: Reference numbers and original names for the piezometers indicated in Fig. 2.4.

Number	Name	Number	Name	Number	Name
1	V2PB102	10	V2PB305	19	M2PB402
2	V2PB103	11	V2PB303	20	V2B6
3	V2PB104	12	V2PB304	21	V2PB402
4	V2B4	13	V2PB400	22	V2PB501
5	V2PB105	14	PB1004	23	V2PB504
6	M2A1	15	V2PB405	24	V2PB503
7	M2A2	16	V2PB404	25	V2PB601
8	V2PB300	17	V2PB1005	26	V2B3
9	V2PB603	18	M2PB20	27	V2B5

of hydrogen on day 302.

Only for piezometer 23 (Fig. 2.5(e-f)) contaminant biodegradation was observed for the entire duration of the experiment without the addition of lactate or hydrogen. The total organic carbon measured for this piezometer was 5.4%, significantly higher than other piezometers; supporting the hypothesis that biodegradation occurs only in presence of an adequate content of organic carbon. This evidence will be used for the conceptualization and the development of the natural attenuation model in chapter 6.

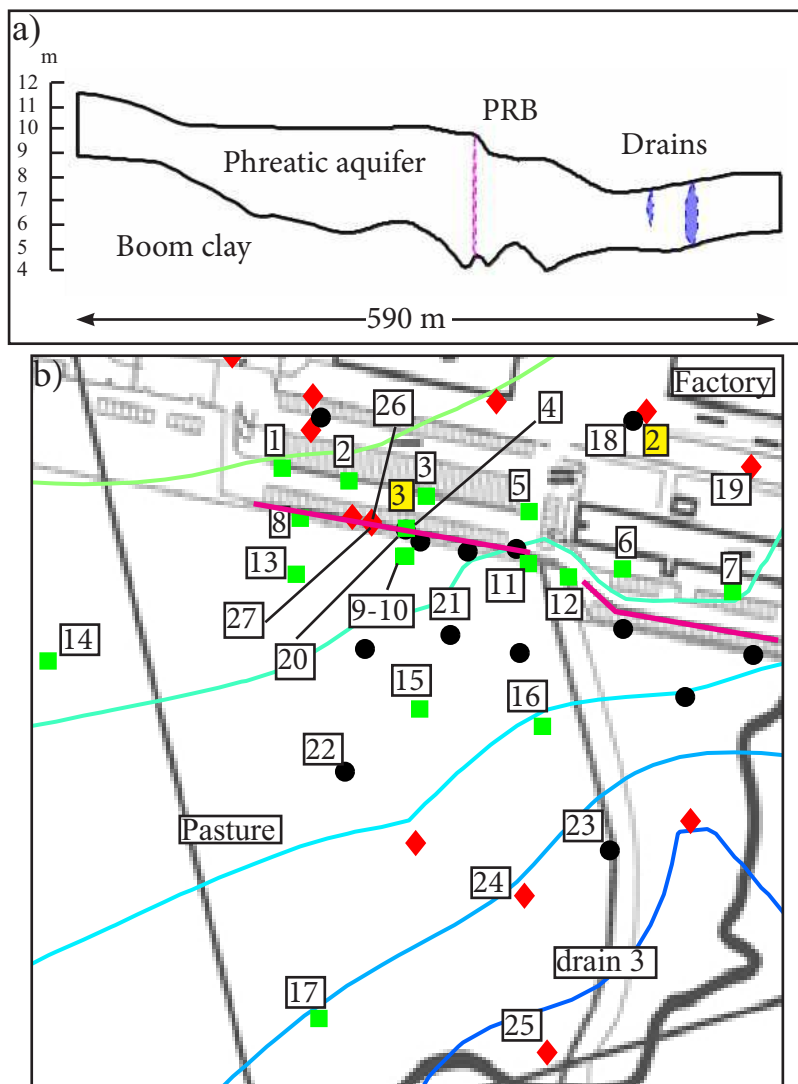


Figure 2.4.: a) A-A cross section (Fig. 2.1) and b) piezometer labels. Original piezometer names are listed in Tab. 2.2.

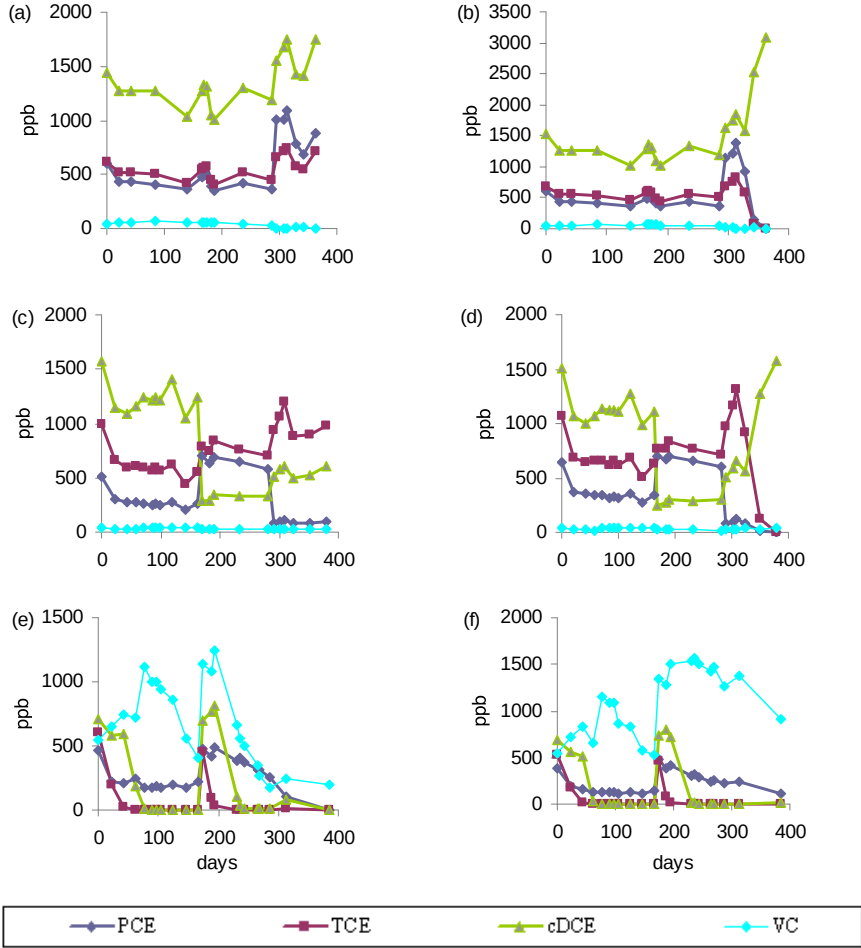


Figure 2.5.: Concentration profiles for the batch experiments: figures (a) and (b) for the batch experiments with piezometer 16 aquifer material (additional groundwater was added after 295 days for both experiments, for the experiment (b) lactate was added after 314 days), (c) and (d) for the batch experiments performed with piezometer 10 aquifer material (additional groundwater was added after 169 and 290 days, for the experiment (d) 500 μM of hydrogen were added on day 302) and (e) and (f) for the batch experiments performed with piezometer 23 aquifer material (additional groundwater was added after 174 days).

3. Predicting longevity of iron permeable reactive barriers using multiple iron deactivation models¹

Abstract: In this study we investigate the model uncertainties involved in predicting long-term permeable reactive barrier (PRB) remediation efficiency based on a lab-scale column experiment under accelerated flow conditions. A PRB consisting of 20% iron and 80% sand was simulated in a laboratory-scale column and contaminated groundwater was pumped into the column for approximately 1 year at an average groundwater velocity of $3.7\text{e-}1\text{ m d}^{-1}$. Dissolved contaminant (PCE, TCE, cis-DCE, trans-DCE and VC) and inorganics (Ca^{2+} , Fe^{2+} , TIC and pH) concentrations were measured in groundwater sampled at different times and at eight different distances along the column. These measurements were used to calibrate a multi-component reactive transport model, which subsequently provided predictions of long-term PRB efficiency under reduced flow conditions (i.e., groundwater velocity of $1.4\text{e-}3\text{ m d}^{-1}$), representative of a field site of interest in this study. Iron reactive surface reduction due to mineral precipitation and iron dissolution was simulated using four different models. All models were able to reasonably well reproduce the column experiment measurements, whereas the extrapolated long-term efficiency under different flow rates was significantly different between the different models. These results highlight significant model uncertainties associated with extrapolating long-term PRB performance based on lab-scale column experiments. These uncertainties should be accounted for at the PRB design phase, and may be reduced by independent experiments and field observations aimed at a better understanding of reactive surface deactivation mechanisms in iron PRBs.

¹Based on: Carniato L., Schoups G., Seuntjens P., Van Nooten T., Simons Q., and Bastiaens L. (2012). Predicting longevity of iron permeable reactive barriers using multiple iron deactivation models, *Journal of Contaminant Hydrology*, 142–143(0), 93–108.

3.1. Introduction

The use of iron PRBs for groundwater remediation is a very active research field, which has grown tremendously during the last 15 years. Anaerobic iron corrosion increases the pH inside iron PRBs and promotes the precipitation of secondary minerals. Typical minerals observed in real site applications are iron oxides (magnetite, hematite), iron (oxy)-hydroxides (ferrous hydroxide, ferric hydroxide, green rust, goethite, lepidocrocite) and carbonates (calcite, aragonite, siderite) [95]. An important issue for long-term operation is the reactivity, porosity and permeability reduction caused by secondary mineral precipitation [161].

Much research has focused on monitoring of pilot studies or actual field sites where PRBs were installed [11]. Based on the study of nine PRB installations, Li et al. [95] reported porosity reductions ranging from 0.0007 to 0.03 per year. The porosity reduction depends on the groundwater geochemistry and the inflow rates [11, 122]. Generally, high flow rates and high concentrations of mineral forming ions enhance mineral precipitation. A similar conclusion has been reported for column studies [73, 152]. Annual rates of hydraulic conductivity reduction have been reported between a factor of 1.5 and 7 [95]. The rate of hydraulic conductivity reduction corresponds to the rate of porosity reduction [116], although in some cases the reduction in hydraulic conductivity is greater than what can be attributed to porosity reduction alone [133].

Much less research has been devoted to the rigorous modeling of these processes to enhance physical understanding and predict long-term efficiency of these systems. Multi-component reactive transport modelling of PRB systems has been reported by Li and Benson [94], Jin suk et al. [68], Jeon et al. [66], Li et al. [96], Mayer et al. [109], Yabusaki et al. [165]. The study of Mayer et al. [109] focused on a PRB at the Coast Guard Support Center near Elizabeth City, North Carolina. Reactive transport modelling was performed using the MIN3P model of Mayer et al. [110]. Their simulations illustrated that secondary mineral precipitation in the upgradient portion of the barrier may decrease barrier porosity in the long-term. In a similar study, Yabusaki et al. [165] reported reactive transport simulation results using the OS3D model for a pilot-scale PRB installed at Moffett Field in Mountain View, California. They confirmed that secondary mineral precipitation in the barrier may impact long-term operation of the PRB. Using data from previous studies and a modelling approach based on MODFLOW[50] and RT3D [27], Li et al. [96] showed that porosity reductions in zero-valent iron PRBs are mostly due to precipitation of calcium and iron carbonates, and therefore are more extensive when ground-

water contains high concentrations of bicarbonate and carbonate ions. More recently, a reactive transport model was applied to simulate iron deactivation along 50 cm long laboratory columns fed with artificial groundwater containing TCE and calcium carbonate [65]. In this study, long-term performance of a hypothetical one-dimensional PRB under different conditions was predicted after manual calibration to laboratory measurements. Results from the study of Jeen et al. [65] suggest that under field conditions iron deactivation due to mineral coating may be more important than porosity reduction. Other studies confirm this finding [54, 36, 152, 79, 115].

Finally, degradation of cis-DCE by different iron materials was studied in several column experiments by Jin suk et al. [68]. Long-term efficiency of a one-dimensional PRB was extrapolated with a reactive transport model calibrated to measurements from the column experiments. Extrapolation included lowering the groundwater velocity from 1 m d^{-1} (column experiment) to 0.1 m d^{-1} (hypothetical field conditions).

The possibility of using the monitoring data of one PRB site to extrapolate the longevity of another PRB is limited by the unique hydrogeological and geochemical conditions of each site, which have a strong effect on PRB longevity [54]. Model predictions based on lab-scale experiments performed under site-specific conditions may help in predicting long-term PRB performance. However, lab-scale column experiments are, for practical reasons, typically performed under accelerated flow conditions and the model uncertainty has not been adequately addressed in the above studies. For example, an open question is whether lab experiments are able to capture all the dominant processes that affect long-term efficiency of PRBs. Lack of information in data from column experiments may lead to significant uncertainty in process representation, leading to highly uncertain long-term extrapolations.

In this study we evaluate and compare four different deactivation models that have been described before [65, 96, 99, 109]. These models were incorporated into a reactive transport model where mineral precipitation and contaminant degradation were formulated as kinetic reactions. Afterwards, the models were calibrated with the data measured under accelerated flow in a laboratory scale PRB column. The calibration provided the values of the deactivation constants, and the rate coefficients of mineral precipitation and contaminant degradation. The models were subsequently used to predict the contaminant concentrations in a long-term numerical experiment representative of flow conditions at a PRB field site of interest in this study. The field conditions were modelled by lowering the effective groundwater velocity about 260 times and by extending the operating time from 1 to 30 years.

3.2. Materials and Methods

3.2.1. Column design

In the experiment, groundwater from an existing contaminated site was pumped through a 25 cm column filled with a mixture of fine zero-valent iron ($1.04 \text{ m}^2 \text{ g}^{-1}$, density 5.6 g cm^{-3}) and sand (Fig. 3.1). The iron grain size distribution was: 20-69 μm (10% by weight), 69-145 μm (40%), 145-282 μm (40%) and 282-400 μm (10%). This was the same mixture as used in the PRB at the field site.

Sampling ports at different levels (0, 0.03, 0.05, 0.075, 0.1, 0.15, 0.2 and 0.25 m) within the column were used to sample aliquots of groundwater and measure organic and inorganic concentrations [150]. Sampling was performed at different times, after various pore volumes (PVs) were pumped through the column (5, 55, 110, 209, 244, 325 and 604 PV corresponding to 6, 34, 63, 118, 152, 202 and 384 days). Ten anaerobic bags containing the contaminated groundwater were used to feed the column throughout the experiment using a peristaltic pump (Pharmacia).

Due to low flow velocities at the site, groundwater was not all collected on the same day, resulting in variable influent concentrations in the column experiment. Changes in influent concentrations and inflow rates at different PVs are listed in Tab. A.1 of the appendix. The average total organic carbon concentration in the liquid phase (TOC) was estimated at 13 mg L^{-1} from independent measurements at the same piezometer where the influent solution was extracted. Organic contaminants account for about 9.3% of TOC. The average hydraulic conductivity of the column barrier was calculated from flow rate and hydraulic gradient measured by manometers connected to the column ports.

3.2.2. Mineral analysis and analytical methods

At the end of the experiment, the column was dismantled and the column material was divided into 5 different sections for mineral characterization using X-Ray diffraction (XRD) and electron-microscopy. For XRD-analysis of mineral precipitate layers, 20 g subsamples were taken from each section under nitrogen atmosphere and the iron particles were separated from the sand after drying using a magnet. To improve XRD detection, the precipitates were detached from Fe^0 filings by sonification for 4 minutes in 40 mL acetone [97]. The fine precipitate fraction was recovered by filtration of the acetone solution and

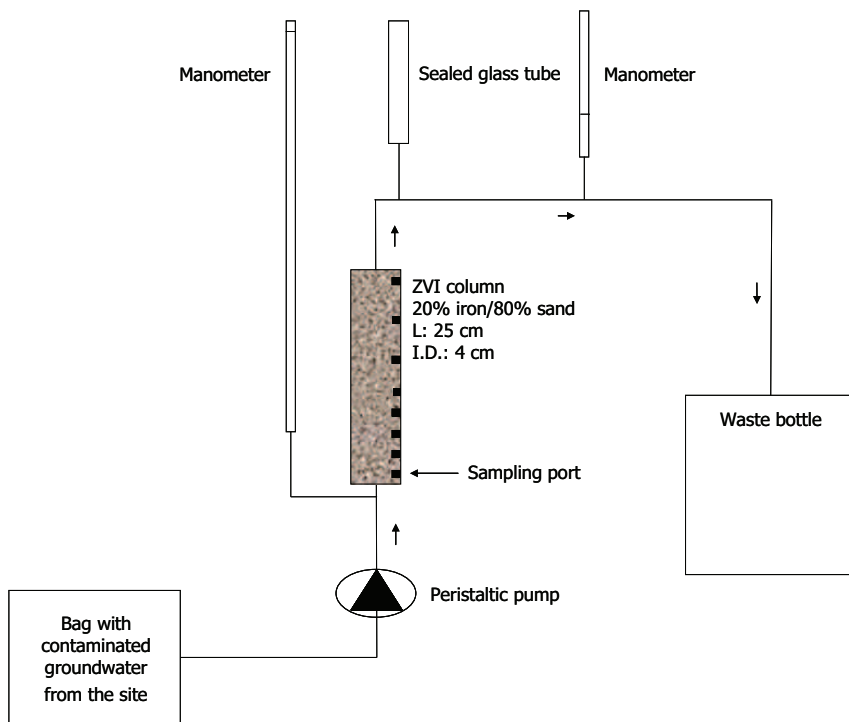


Figure 3.1.: Scheme of the column experiment

stored anaerobically in capped glass vials for no more than one week before analysis. XRD analysis was performed using a PANalytical X-ray diffractometer (X'pert Pro) and $\text{CuK}\alpha$ X-radiation. The sample was scanned from 2° to $120^\circ 2\theta$ (40kV-40mA; step size: 0.04° ; step time: 1 s step^{-1}). For microscopic analysis, approximately 2 g subsamples were taken and dried in the anaerobic glove box for 24 hours. Dry Fe^0 filings were embedded in epoxy resin and polished after hardening for 12 hours. Polished sections were studied using an AXIOPLAN imaging reflected-light microscope (Zeiss) and the selected sections were platina coated for examination with an Environmental Scanning Electron Microscope (JEOL JSM-6340F), fitted with an Omega detector (PGT SPIRIT) for energy dispersive X-ray analysis (EDX). XRD and microscopic analysis were performed only at the end of the experiment as the

passivating minerals are expected to be present mostly at the end of the experiment, after the precipitation of dissolved carbonates. Analytical methods used to determine contaminant and inorganic ions concentrations are reported in Tab. A.2 of the appendix.

3.2.3. Numerical model

The multi-component reactive transport model PHAST [121] was used to predict concentrations of dissolved and solid species in the column experiment. PHAST simulates multicomponent, reactive solute transport in saturated groundwater flow systems. The flow and transport calculations are based on a modified version of HST3D [76] while the geochemical reactions are simulated with the geochemical model PHREEQC [120]. In this study, 36 complexation reactions were taken from the PHREEQC and EQ/6 databases [163], including equilibrium constants for mineral precipitation and gas solubility (Tab. A.3 and Tab. A.4). Three mineral phases were included in the model: aragonite ($CaCO_3$), iron hydroxy carbonate ($Fe_2(OH)_2CO_3$) and amorphous iron hydroxide ($Fe(OH)_{2(am)}$). These minerals were identified by previous literature studies to be the main causes for porosity and reactivity reduction in PRB systems [92, 65, 96, 72, 165]. This choice is also supported by the results of the mineral analysis (sec. 3.3.1).

Fourteen dissolved mobile species were included in the model: PCE, TCE, trans-DCE, cis-DCE, VC, ethene, Ca^{2+} , Cl^- , Na^+ , Fe^{2+} , Fe^{3+} , CO_3^{2-} , H^+ and $H_{2(aq)}$. Chemical degradation of the organic species (PCE, TCE, cis-DCE, trans-DCE and VC) by zero-valent iron is assumed to mainly occur by β -elimination and much less by hydrogenolysis, with a small formation of toxic intermediates (e.g. VC). Therefore, we adopted the average stoichiometric coefficients proposed by Arnold and Roberts [1] and a similar reaction network (Tab. A.5). Individual VOC kinetic rates were described by a mixed-order rate law [164]:

$$r_{k-PRB} = k_{k-PRB} S_r \frac{C_k}{K_{k-PRB 1/2} + C_k} \quad (3.1)$$

where k_{k-PRB} is the rate coefficient per unit iron reactive surface area ($\text{mol m}^{-2} \text{s}^{-1}$), S_r is iron reactive surface area per unit water volume ($\text{m}^2 \text{L}^{-1}$), $K_{k-PRB 1/2}$ is the half-saturation constant (mol L^{-1}) and C_k is the concentration of contaminant k . Iron corrosion by groundwater usually occurs under anaerobic conditions with the consumption of acidity and the production of

3.2 Materials and Methods

hydrogen. A first-order dependence of iron corrosion on reactive surface area as proposed by Mayer et al. [109] was assumed:

$$r_{IrC,W} = \max \left\{ \left[k_{IrC,W} S \left(1 - \frac{IAP_{IrC,W}}{K_{IrC,W}} \right) \right], 0 \right\} \quad (3.2)$$

where $k_{IrC,W}$ is the rate coefficient for iron corrosion by water per unit iron reactive surface area ($\text{mol m}^{-2} \text{s}^{-1}$), IAP_{IrC} is the for iron corrosion and $K_{IrC,W}$ is the equilibrium constant ($\log K_{IrC} = -11.78$ as reported by Mayer et al. [109]). Mineral precipitation following a pH increase was described by the following rate equation [90]:

$$r_i^m = \max \left\{ \left[k_{eff,i} \left(\frac{IAP_i^m}{K_i^m} - 1 \right) \right], 0 \right\} \quad (3.3)$$

where $k_{eff,i}$ is the rate coefficient for precipitation of mineral i ($\text{mol L}^{-1} \text{s}^{-1}$), and IAP_i^m and K_i^m are ion activity product and solubility of mineral i , respectively (the values of $\log K_i^m$ are reported in Tab. A.3). The decrease in porosity and permeability was calculated in both models using the following equation:

$$por_{t+\Delta t} = por_t - \sum_{i=1}^{n_m} V_i^m \int_t^{t+\Delta t} por_t r_i^m dt \quad (3.4)$$

$$K_t = \left[\frac{por_t^3}{(1 - por_t)^2} \right] \left[\frac{(1 - por_0)^2}{por_0^3} \right] K_0 \quad (3.5)$$

where por_t is porosity at time t , and V_{im} is molar volume of mineral i (L mol^{-1}). The summation in Eq.3.4 is over the two mineral phases ($n_m = 2$) that were considered relevant for porosity decline, in this case aragonite and iron hydroxy carbonate. In this scheme the porosity was updated explicitly ($\Delta t = 90$ days for the field simulation), assuming a slow porosity change. Eq.3.5 (Kozeny-Carmen relationship) was used to update the initial permeability K_0 to the value at time t . Iron reactive surface deactivation by mineral coating was implemented using two different relationships. In the first relationship,

iron reactive surface area S_r is represented mathematically by an exponential formula [65]:

$$S_r(x, t) = S_{r0} \exp\left(-\sum_{i=1}^{n_m} \alpha_i \varphi_i(x, t)\right) \quad (3.6)$$

where S_{r0} is the initial iron reactive surface area per unit water volume ($\text{m}^2 \text{L}^{-1}$), φ_i is the volume fraction of mineral i at location x and time t , and α_i is the deactivation constant for mineral i (-). In columns that receive carbonate-rich groundwater, the deactivating effect of amorphous iron hydroxide can be neglected compared to that of carbonate minerals [65]. Therefore, we assume that iron hydroxy carbonate and aragonite precipitation in the column results in a decrease in contaminant degradation rates, as these minerals form carbonate-rich, deactivating coatings on the iron grains [64]. In the second relationship, reactive surface evolution is represented by a linear formula [116]:

$$S_r(x, t) = \max \left[\left(S_0 - \frac{\Delta_n}{T_c} \right), 0 \right] \quad (3.7)$$

where T_c is average thickness of the mineral coatings (m). The porosity reduction Δ_n was calculated based on the simulated accumulation of aragonite and iron hydroxy carbonate. This agrees with Eq.3.4, where aragonite and iron hydroxy carbonate were considered as the only minerals contributing to porosity decrease.

The model did not consider the influence of hydrogen bubbles on porosity and hydraulic conductivity reduction, assuming that produced gas bubbles leave the column instantaneously, similar to previous modelling studies [65]. The assumption of a fully saturated column is justified by the relatively small volume fraction of hydrogen gas bubbles in the column, which was estimated to be around 6%, compared to a total porosity value of 42% (estimate based on corrosion rate in Tab.3.2, iron reactive surface area of 130 m^2 , average residence time in the column of 0.64 d, and known values of hydrogen gas solubility and volume under standard conditions). Furthermore, porosity is expected to recover to some extent during the experiment due to iron passivation and the establishment of preferred pathways of gas migration [170]. The linear and exponential formulas were further modified to consider the effect of

iron dissolution on reactive surface area:

$$S_r(x, t) = S_{r0} \exp\left(-\sum_{i=1}^{n_m} \alpha_i \varphi_i(x, t)\right) \left(\frac{\varphi_{Fe^0}}{\varphi_{Fe^0}^0}\right)^{\frac{2}{3}} \quad (3.8)$$

$$S_r(x, t) = \max \left[\left(S_0 - \frac{\Delta_n}{T_c} \right) \left(\frac{\varphi_{Fe^0}}{\varphi_{Fe^0}^0} \right)^{\frac{2}{3}}, 0 \right] \quad (3.9)$$

where $\varphi_{Fe^0}^0$ and φ_{Fe^0} are the initial and time t molar volume fractions of Fe^0 (-). The “two-thirds” formula has been reported before to consider the change of reactive surface area by dissolution of a mineral phase. Mayer et al. [109] used the two-thirds relation in the reactive transport model of an in-situ PRB. Li et al. [96] used Eq.3.7 to model reactive surface deactivation by iron dissolution and carbonate mineral precipitation in a hypothetical PRB installation. In this study, we applied both exponential and linear formulas with and without the “two-thirds” dissolution factor to evaluate the validity of these different deactivation models and compare their long-term predictions of the PRB performance.

3.2.4. Flow and transport parameters

The column was discretized using a 0.01 m space interval and a 1500 s time step was used to solve the numerical equations. This discretization was a good compromise between CPU time and the results of the numerical solution. Indeed, small differences were found between this discretization and a finer one (0.002 m space interval and 300 s time step), which required more CPU time. Flow boundary conditions were specified flux at the column inlet and zero pressure head at the column outlet. The initial flow condition was zero pressure head in the whole domain. For transport, a Cauchy boundary condition was specified at the column inlet (specified mass flux) and a free exit boundary condition at the column outlet. Initially, the column was assumed to be free of contaminants, and filled with water at pH 7 and low concentrations of all dissolved inorganic species ($1e-10$ mol L^{-1}). Longitudinal dispersivity was fixed at a value of $1e-3$ m, in line with similar column experiments [65].

3.2.5. Parameter inference

The parameter estimation tool PEST Doherty [39] was used to calibrate the model parameters (Tab. 3.2) by minimizing the sum of the squared deviations between all laboratory measurements and their corresponding model-predicted values. Formally, the objective function of this minimization problem is given by:

$$\Phi = \sum_{i=1}^m w_i \sum_{j=1}^{n_i} (y_{ij} - \eta_{ij}(\phi, Z))^2 \quad (3.10)$$

where ϕ is a parameter vector, y_{ij} is the measurement j from observation set i , $\eta_{ij}(\phi, Z)$ is the corresponding model-predicted value, Z represents other model input (e.g. boundary and initial conditions), and w_i is a weight assigned to observation set i . Initial parameter estimates were subsequently refined by automatic calibration with PEST. Due to the long execution time (about 20 minutes for each forward run) and the relatively large number of model parameters, a stepwise procedure was adopted [31]. First, insensitive and/or highly correlated parameters were fixed at their initial values. These included the aragonite and amorphous iron hydroxide precipitation rate coefficients ($k_{eff,CC}$ and $k_{eff,IH}$). The remaining parameters were then split into two groups, each group being estimated using a different set of observations. Specifically, measured calcium, dissolved iron, total inorganic carbon and pH profiles at 5 PV were used to estimate the iron corrosion and iron hydroxy carbonate precipitation rate coefficients (k_{IRC} and $k_{eff,IC}$), as the other parameters only had limited influence on these early inorganic profiles. The remaining parameters (i.e., VOC degradation rate coefficients and iron deactivation constants) were subsequently estimated using measured organic and inorganic profiles at 55, 109 and 209 PV, as these profiles were expected to contain information on iron deactivation and its effect on contaminant degradation. Observation weights in Eq. 3.10 were calculated to take into account the differences in magnitude between different types of measurements (e.g. pH vs. VOCs concentrations), ensuring a balanced contributions of different types of measurements to the overall objective function.

3.3. Results and discussion

3.3.1. Mineral analysis

The results of the XRD analysis are reported in Tab.3.1, confirming the hypothesis about the presence of aragonite and iron hydroxy carbonate in the column. As expected, quartz (SiO_2) was detected in the whole column, as it is part of the filling material. Calcite ($CaCO_3$) and rhodochrosite ($MnCO_3$) were mostly present at the beginning of the column while iron hydroxy carbonate was detected most strongly toward the end of the column. The relatively insoluble carbonate minerals precipitated at the beginning of the column, consuming carbonate and reducing the groundwater pH buffer capacity. As the buffer capacity decreases, and iron corrosion proceeds, pH tends to increase, causing the precipitation of the previously soluble iron hydroxy carbonate. These competitive mechanisms were included in the model through the kinetic formulation of the mineral precipitation rates (Eq.3.3), which consider the dependency of the precipitation rate on the ion activity product. Rhodochrosite, although present in the XRD analysis, was not included in the model because the manganese inlet concentration was very low and the amount of rhodochrosite formed is assumed negligible.

Amorphous iron hydroxide could not be detected by the XRD analysis, because XRD does not allow the identification of amorphous phases. However, precipitation of amorphous iron hydroxide is very likely to occur during anaerobic iron corrosion, as reported by Reardon [129]. Thus, iron hydroxide precipitation was considered into the model.

Goethite ($FeOOH$) was present at the beginning of the column (0-2 cm), probably due to an aerobic zone at the column inlet. This zone was probably caused by the presence of dissolved oxygen in the groundwater feeding the column. Preliminary model simulations including aerobic iron corrosion showed a fast consumption of dissolved oxygen with Fe^{3+} production (Tab. A.3). Fe^{3+} is not stable at the column pH and may precipitate forming goethite and magnetite (Fe_3O_4). The direct precipitation of goethite and magnetite would also result in a pH decrease (Tab. A.3), which was not observed and would furthermore prevent the observed precipitation of carbonate minerals. This suggests that precipitation of Fe^{3+} as goethite and magnetite was not a dominant process during column operation. Therefore the precipitation of these minerals was not included into the model. The presence of magnetite in the deeper part of the column may also be explained by the disproportionation of amorphous iron hydroxide [129]. Magnetite may also be formed by autoredox of Fe^{3+}

Table 3.1.: Height of the main diffraction peaks measured by the X-Ray diffraction analysis (in counts s⁻¹). The analysis was performed after the dismantling of the column.

Section height (cm)	0-2	2-4	6.5-8.5	14-16	21.5-23.5
Quartz	302	1221	1297	877	1888
Iron hydroxy carbonate	97	225	353	473	375
Aragonite	97	138	276	207	760
Calcite	302	349	293	177	200
Rhodochrosite	212	357	254	141	88
Goethite	222	74	73	50	59
Magnetite	212	169	239	217	252

oxides present on the iron surface [12]. From the mineral analysis and the above considerations we concluded that the main minerals for which precipitation reactions need to be included in the model were aragonite, iron hydroxy carbonate and amorphous iron hydroxide.

The precipitation of calcium carbonate minerals was accounted for by including aragonite precipitation in the model, similar to previous studies [65]. The addition of calcite in the model would add an additional precipitation rate that most likely is highly correlated with aragonite precipitation rate and thus difficult to infer using the column measurements.

Fig. 3.2 shows the results of the EMPA (electron micro probe analyzer) analysis at 2 cm from the column inlet. The inner part of the corrosion layers are mostly composed of minerals containing iron and oxygen (most likely amorphous iron hydroxide, iron hydroxy carbonate and magnetite). As can be seen from Fig. 3.2, the presence of manganese minerals (rhodochrosite) and adsorbed silica into the iron grains is negligible. Calcium was detected in the outer parts of the corrosion layers (calcite and aragonite). Interestingly, a sulfur mineral is present in the inner part of the corrosion layer, where also iron is present. Most probably, this mineral is amorphous iron sulfide (FeS), which can be produced by biological transformation of sulfate into sulfide, followed by precipitation of FeS [109]. However, measured sulfate concentrations along the column did not show a clear decreasing trend that would support the inclusion of these processes into the model and the calibration of the relative parameters (see Fig. A.2).

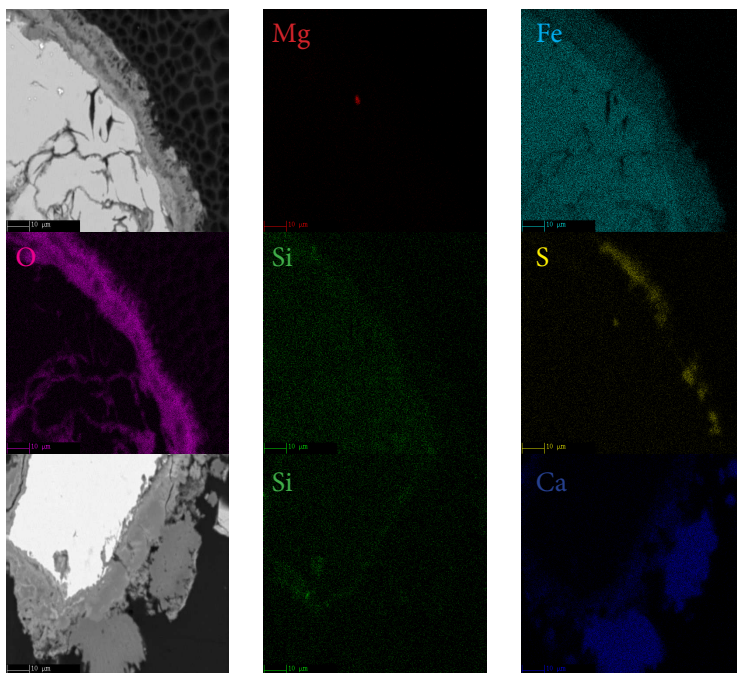


Figure 3.2.: Elements detected in one column section by EMPA analysis at the end of the experiment. This section was at 2 cm from the influent.

3.3.2. Column experimental results

Measured PCE and VC concentrations are shown in Fig. 3.3(a) and Fig. 3.3(b). Upward shifts in the observed PCE profiles through time clearly show a decreased remediation performance of the column, hypothesized to be caused by mineral precipitation resulting in iron deactivation. In some cases the profiles overlap (profiles at 160 PV and 244 PV) or later profiles lie below earlier profiles (profiles at 244 PV and at 209 PV). This seems to contradict the hypothesis of a gradual decrease in contaminant degradation. However, the PCE influent concentrations were variable (Tab. A.1) and PCE desorption may take place from non-reactive iron sites when the column is flushed with a solution with lower PCE concentration [10]. An example is the PCE profile at 209 PV, where the measured PCE concentrations at 0.03, 0.05, 0.075 and 0.1 m were

equal or even higher than the influent PCE concentration. Indeed, PCE concentration of the inlet solution used between 139 to 186 PV was 55 % higher than the PCE concentration of the inlet solution used between 187 to 229 PV (Tab. A.1). Another confirmation of this hypothesis is the similar shape of the profiles at 160 and 209 PV between 0 and 0.1 m.

Unnormalized VC concentration profiles are shown in Fig. 3.3(b), illustrating variations in the influent concentrations. With the exception of profiles at 160 and 604 PV, VC profiles show a systematic VC production within the column (see for example the 55 and 325 PV profiles), suggesting that some portion of cis-DCE was degraded by sequential degradation (Tab. A.5). Values for the kinetic parameters of all VOC degradation reactions (k_{k-PRB} and $K_{k-PRB}^{1/2}$ in Eq.3.1) were determined from the observed 5 PV profiles by fitting the first part with a zero-order rate model (to estimate $k_{k-PRB}S_r$) and the last part of the profile with a first-order rate model (to estimate $K_{k-PRB}^{1/2}$, as suggested by Wüst et al. [164]), assuming that at 5 PV the reactive surface area is close to its initial value.

Dissolved calcium (Fig. 3.3(c)) was rapidly consumed at the beginning of the experiment (5 and 55 PV) and in the first 0.15 m. This suggests accumulation of carbonate minerals (aragonite and calcite) close to the column inlet. At later PV, calcium was less consumed by mineral precipitation and the profiles tend to flatten (for example at 325 and 604 PV), again suggesting a gradual process of iron deactivation and/or porosity reduction. The initial aragonite precipitation rate ($k_{eff,AR}$) was calculated using measured differences between the calcium concentrations at 0.03 m and in the influent at 5 PV, assuming that calcium was removed only by aragonite precipitation.

Dissolved iron concentrations (Fig. 3.3(d)) in the influent are negligible but increase rapidly within the column for all PV. Close to the column effluent, iron concentrations decrease again, showing that the rapid release due to iron corrosion was followed by precipitation of iron minerals (likely iron hydroxy carbonate and iron hydroxide, as suggested by the mineral analysis). At later times, iron concentrations at the column effluent tend to increase (i.e. at 604 PV), indicating a reduction in iron mineral precipitation, as observed for calcium.

For iron hydroxy carbonate it was not possible to estimate the precipitation rate directly from the iron concentrations, because dissolved iron could also precipitate as iron hydroxide. Therefore, initial rate coefficients of iron hydroxide and iron hydroxy carbonate precipitation ($k_{eff,IH}$, $k_{eff,IC}$ in Eq.3.3) were taken from Jeen et al. [64]. The iron corrosion rate by water ($k_{IRC,W}$ in Eq.3.2) was estimated dividing the initial production rate of the hydrogen

3.3 Results and discussion

gas collected at the end of the column (mol s^{-1}) by the initial iron reactive surface area (m^2). This estimated value was found close to batch experiments where the corrosion rate of the same iron type was estimated from hydrogen gas production (giving an average $\log k_{IrC,W}$ equal to -10.12).

The exponential deactivation model has a total of 9 adjustable parameters (5 contaminant degradation rates, the iron hydroxy carbonate precipitation rate $k_{eff,IC}$, the iron corrosion rate by water $k_{IrC,W}$ and two deactivation constants α_{IC} and α_{CC}) while the linear one has one less (only one deactivation constant T_C). The total number of adjustable parameters was assessed through a preliminary sensitivity analysis, which resulted in excluding the rate coefficient of iron hydroxide precipitation (insensitive) and the rate coefficient of calcium carbonate precipitation (which can be estimated from calcium concentrations). Non-uniqueness of the inverse problem is expected [138]. However, a good estimation of the initial parameter values should increase the chances that reasonable estimates of the optimal parameter values are obtained at the end of the inversion process. As described above, initial estimates were obtained for 6 adjustable parameters ($k_{PCE-PRB}$, $k_{TCE-PRB}$, $k_{trans-DCE-PRB}$, $k_{cis-DCE-PRB}$, k_{VC-PRB} and $k_{IrC,W}$), with the remaining parameters initially estimated from literature (α_{CC} , α_{IC} and $k_{eff,IC}$) or empirically (T_C).

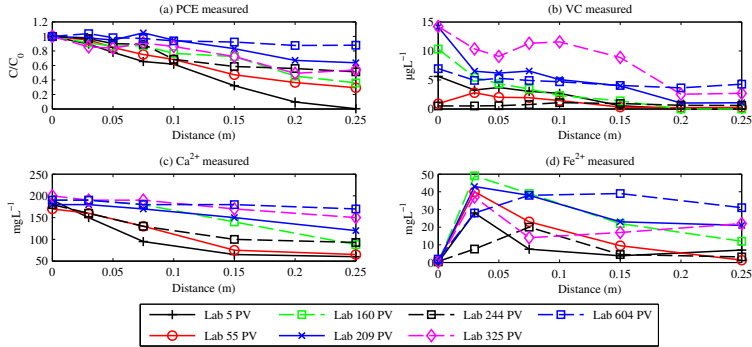


Figure 3.3.: Measured profiles at different pore volumes (PV). (a) PCE, (b) VC, (c) dissolved calcium and (d) dissolved iron.

3.3.3. Calibration results

The results of the deactivation models not considering the dissolution factor (Eq.3.6 and Eq.3.7) will be presented here, as iron dissolution was found to be negligible in the column experiment. Results including iron dissolution (based on Eq.3.8 and Eq.3.9) were very similar (see Fig. A.1).

Fig. 3.4 compares simulated and observed profiles in the column after calibration, using observed profiles at 5, 55, 160 and 209 PV. In general, the gradual decrease in VOC degradation is reproduced by the models, with simulated results in Fig. 3.4 indicating that mineral precipitation results in significant iron deactivation, whereas porosity reduction is much less important, confirming the results reported by Jeon et al. [65]. The calibrated models were able to simulate reasonably well the PCE concentrations at 5, 55 and 160 PV near the influent, but worse results were obtained at 0.15 and 0.25 m for both models. The models tend to better reproduce the measurements between 0 and 0.1 m, because more measurements were taken in this part of the column (the measuring interval decreases after 0.1 m) and their contribution in the objective function was greater.

The simulated PCE profile at 209 PV lies below the 160 PV because the inflow rate between 187 and 229 PV was lower than the inflow rate between 139 and 187 PV (Tab. A.1) and thus the contaminant resides more time within the column. In the real observations the 209 PV PCE concentrations remain high in the first 0.1 m, possibly due to contaminant desorption from non-reactive sites, which was not accounted for in the numerical model. In Fig. 3.4(c) and Fig. 3.4(d) the simulated VC concentrations are also shown. From these Figures it can be seen that the VC produced by sequential degradation was underestimated at 5 PV while the best results were obtained at 55 and 160 PV.

A possible improvement of the model would be to include in the calibration process the stoichiometric coefficients of the VOC degradation network (Tab. A.5), as these stoichiometric coefficients may depend on the iron type. However, the direct estimation of stoichiometric coefficients from our column data would have made the parameter estimation process more difficult.

After calibration, we assessed the validity of the calibrated model by running the model until the end of the experiment and comparing observed and simulated profiles at 244, 325 and 604 PV (Fig. 3.5). These results show that deactivation is somewhat greater with the linear versus the exponential deactivation model. This can be attributed to the differences between the deactivation formulas: in the exponential formula, aragonite causes a larger reactivity

3.3 Results and discussion

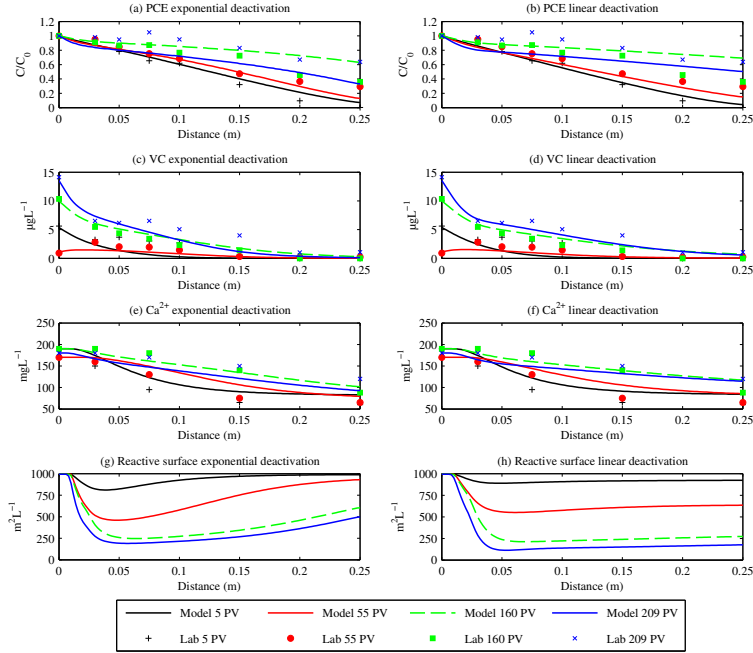


Figure 3.4.: Measured and simulated profiles for the calibration step. (a, c, e, g) PCE, VC, dissolved calcium and reactive surface profiles with the exponential deactivation model and (b, d, f, h) PCE, VC, dissolved calcium and reactive surface profiles with the linear deactivation model.

deactivation than iron hydroxy carbonate, whereas in the linear formula both minerals have the same influence. As a result, the reactive surface at the column outlet was much lower in the linear case as iron hydroxy carbonate had accumulated mostly in this region (compare Fig. 3.5(g) and Fig. 3.5(h)). Overall the fitting to observed concentrations was better in Fig. 3.5 (validation tests) than in Fig. 3.4 (calibration tests) for both models, probably due to the lower variability of the influent concentration and flow rates in this part of the experiment (Tab. A.1).

From the Figures it can also be concluded that the two deactivation models provide similar fits to the observations, although a better fit was obtained with the exponential model. This is likely due to an additional deactivation parameter in the exponential model.

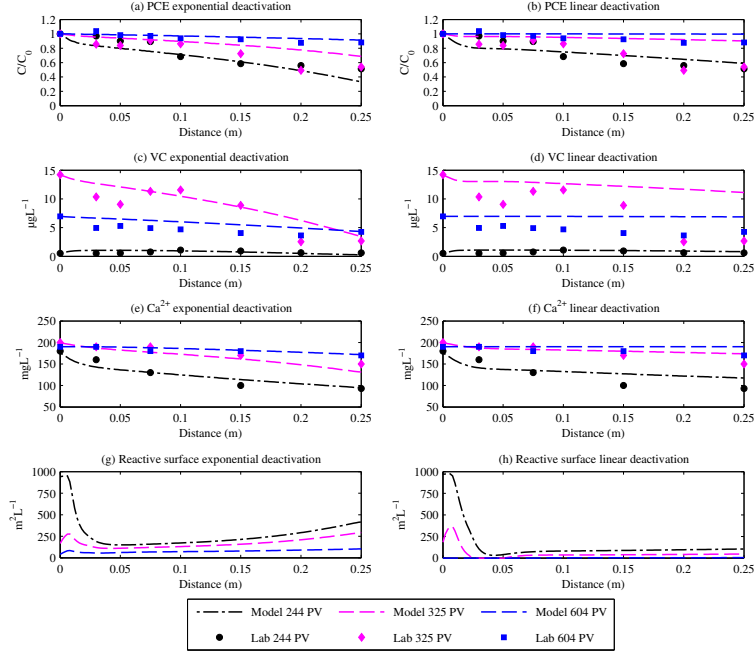


Figure 3.5.: Measured and simulated profiles for the validation step. (a, c, e, g) PCE, VC, dissolved calcium and reactive surface profiles with the exponential deactivation model and (b, d, f, h) PCE, VC, dissolved calcium and reactive surface profiles with the linear deactivation model.

The dissolved iron concentrations are overestimated up to 244 PV (Fig. 3.6(a)) and underestimated at 604 PV. It is hypothesized that an additional iron mineral, which does not consume carbonate, may need to be included in the model without causing a substantial change in simulated TIC, pH and dissolved calcium concentration profiles. A potential candidate mineral is amorphous iron sulfide (FeS), as mentioned in sec. 3.3.1. The biological transformation of sulfate into sulfide and the subsequent FeS precipitation were included in the model in order to test this hypothesis and quantify the amount of iron removed by FeS precipitation. The biological transformation of sulfate was modeled as in Mayer et al. [109] with consumption of dissolved hydrogen produced by iron corrosion. The biological reaction was assumed to be limited by the availability

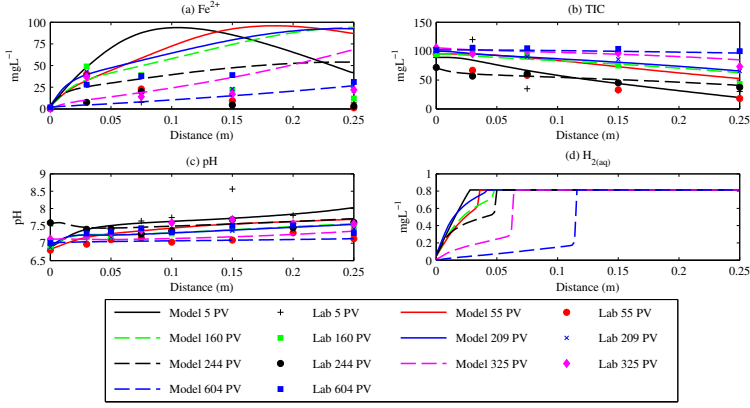


Figure 3.6.: Measured and simulated profiles for the calibration step with the exponential deactivation model. (a) Fe^{2+} , (b) TIC, (c) pH and (d) $\text{H}_{2(\text{aq})}$.

of sulfate and hydrogen gas :

$$\frac{d[\text{SO}_4^{2-}]}{dt} = -k_{\text{SO}_4-\text{H}_2} \left(\frac{[\text{SO}_4^{2-}]}{K_{\text{SO}_4^{2-}} + [\text{SO}_4^{2-}]} \right) \left(\frac{[\text{H}_{(\text{aq})}]}{K_{\text{H}_2} + [\text{H}_{(\text{aq})}]} \right) \quad (3.11)$$

where $k_{\text{SO}_4^{2-}-\text{H}_2}$ is the rate coefficient of sulphate consumption ($2.5\text{e-}8 \text{ mol L}^{-1} \text{ s}^{-1}$) and $K_{\text{SO}_4^{2-}}$ and K_{H_2} are the half saturation constants for sulfate and hydrogen ($10^{-4} \text{ mol L}^{-1}$ and $10^{-7} \text{ mol L}^{-1}$ respectively) . The reaction stoichiometry is reported in Tab. A.5.

The precipitation of amorphous iron sulfide was modeled using Eq.3.3 and a sulfide precipitation rate of $2.51\text{e-}10$, as in Mayer et al. [109]. Considering these high rates of sulfate transformation and sulfide removal, the dissolved iron concentrations at 5 and 55 PV did not decrease more than 25% (Fig. A.2(a)). Moreover, when including FeS precipitation, the dissolved sulfate simulated concentrations were underestimated (Fig. A.2(e)).

Therefore, the exclusion of FeS precipitation from the model can not be the only reason for the overestimation of the iron concentrations. Similar overestimations of the early dissolved iron concentrations were simulated by Jeon [63] on his column C, the most similar to this column for the iron type and carbonate influent concentrations. Better results were obtained for TIC concentrations; where only the 55 PV profile is overestimated. On average the

simulated pH reproduce the measured values but some profiles were overestimated (55 PV) and others were underestimated (325 and 604 PV). For dissolved hydrogen the model simulated a moving front. At the beginning of the experiment almost all the column is saturated with hydrogen while at later pore volumes the first part of the column becomes undersaturated. This is a consequence iron deactivation, which was larger at the column influent due to aragonite precipitation (Fig. 3.4(g) and Fig. 3.5(g)).

Tab. 3.2 summarizes initial parameter estimates before calibration, optimal values after calibration, and confidence limits at the optimal values together with literature ranges for each parameter. As can be seen from the table, for the parameters where a good initial estimate was calculated ($k_{PCE-PRB}$, $k_{TCE-PRB}$, $k_{trans-DCE-PRB}$, $k_{cis-DCE-PRB}$, k_{VC-PRB} and $k_{IRC,W}$) the optimal value is close to the initial estimation and the 95% linear confidence interval is narrow. Furthermore, the differences between the degradation rates in the two models are small, indicating that the models behave similarly, confirming the results of Fig. 3.4 and Fig. 3.5. However, the rate coefficient of iron hydroxy carbonate precipitation was not well identified (wide 95% confidence limit) due to its correlation with aragonite precipitation (both minerals compete for carbonate in solution).

For the exponential deactivation model also the two deactivation constants were not well identified due to high negative correlation between α_{IC} and α_{CC} (-0.77). This part of the model may be more prone to non-uniqueness (a decrease in one deactivation constant can be partially compensated for by an increase in the other deactivation constant to yield similar values of simulated concentrations). The high correlation between the deactivation constants makes the exponential model less identifiable than the linear model using our column measurements. Jeen et al. [64] assigned different deactivation parameters for aragonite and iron hydroxy carbonate because aragonite was supposed to reduce the iron reactivity more than iron hydroxy carbonate. For our column experiment a more appropriate model would be to consider just one deactivation constant for the exponential model to account for the total carbonate mineral fraction. This would eliminate the correlation between the two deactivation constants and would make the single deactivation exponent more identifiable from the column data.

3.3.4. Long-term performance prediction

After model calibration and validation, numerical experiments were conducted to compare model predictions of long-term PRB remediation efficiency under

flow rates that are more representative of field applications. In particular, groundwater velocity was decreased about 260 times from an average value of $3.7\text{e-}1 \text{ m d}^{-1}$ in the column experiment to a value of $1.4\text{e-}3 \text{ m d}^{-1}$, as estimated at the field site of this study. Degradation and reaction parameters determined for the column were assumed to remain valid at this lower flow velocity. This assumption implies that mass transfer does not becoming the limiting step for degradation at these lower flow velocities. This is supported by estimates of diffusive mass transfer that are at least two orders of magnitude larger than VOC degradation rates (see mass transfer considerations in Appendix A). By decreasing flow velocity, travel and residence times greatly increase, resulting in higher simulated pH values (Fig. 3.7), representative of field PRBs. As mineral precipitation and corrosion rates explicitly depend on pH and geochemical conditions (Eq.3.3 and Eq.3.2), simulated iron passivation was also affected by the reduced velocities.

Results with the four different deactivation formulas are shown in Fig. 3.8. As a first case, the exponential and linear formulas without iron dissolution were used (Eq.3.6 and Eq.3.7). In the exponential case, complete PCE degradation was observed even after 30 years of operation. Indeed, in the exponential formula the reactive surface is never zero and a residual reactive surface is always present. This residual reactive surface can completely degrade the contaminants when the residence time is large enough, as was the case here. To test the validity of this interpretation, the same simulation was performed using a larger groundwater velocity (0.01 m d^{-1}). In this case, a further migration of the VOC profiles was simulated, similar to simulation results reported by Jeen et al. [65] (Fig. A.3 of the appendix). Comparing Fig. 3.8(a) and Fig. 3.8(b), it is evident that with the linear formula the reactive surface is much more depleted, mainly because the linear model allows complete coverage of iron grains by carbonate minerals ($S_r = 0$). As a result, much larger contaminant migration was observed in the linear model.

As a second case, we tested the exponential and linear formulas accounting for the dissolution of iron. With iron dissolution included the iron reactive surface decreases not only by mineral precipitation but also by transformation of solid iron into dissolved iron caused by the corrosion reaction. However, it should be noted that most of the dissolved iron precipitates back as amorphous iron hydroxide, mostly at the end of the column. Amorphous iron hydroxide has been proven to play an active role in contaminant degradation since it can adsorb the ferrous iron generated from the corrosion reaction [117], which is still capable of contaminant degradation [107]. When iron dissolution is included into the model the role of amorphous iron hydroxide on contaminant

Table 3.2.: Calibrated model parameters for the column test.

Parameter	Literature range	Initial value ^b	Exponential deactivation		Linear deactivation	
			Optimal value	95 % confidence limit	Optimal value	95 % confidence limit
$\log k_{PCE-PRB}$ (mol m ² s ⁻¹)	NA ^a	-13.07	-13.01	-13.13 to -12.89	-12.97	-13.13 to -12.82
$\log k_{TCE-PRB}$ (mol m ² s ⁻¹)	-12.72 ^c to -10.89 ^d	-12.69	-12.47	-12.56 to -12.38	-12.46	-12.54 to -12.38
$\log k_{cis-DCP-PRB}$ (mol m ² s ⁻¹)	NA	-14.18	-13.66	-13.82 to -13.51	-13.69	-13.80 to -13.59
$\log k_{trans-DCP-PRB}$ (mol m ² s ⁻¹)	-10.60 ^d	-12.66	-12.22	-12.36 to -12.09	-12.21	-12.30 to -12.12
$\log k_{VOC-PRB}$ (mol m ² s ⁻¹)	NA	-14.24	-13.31	-13.50 to -13.11	-13.32	-13.46 to -13.17
$\log k_{TCE,W}$ (mol m ² s ⁻¹)	-12.46 ^e to -7.20 ^e	-10.23	-9.88	-10.04 to -9.72	-9.90	-10.05 to -9.74
$\log k_{eff,IC}$ (mol m ² s ⁻¹)	-10.89 ^f to -8.51 ^e	-10.89	-9.38	-10.50 to -8.26	-9.38	-10.48 to -8.27
α_{IC} (-)	2 ^f to 5 ^f	5	2.1	-85.1 to 89.4	-	-
α_{CC} (-)	5.5 ^f to 90 ^g	85	175	101.9 to 248.0	-	-
T_C (m)	1e-8 ^h to 4e-4 ⁱ	1e-5	-	-	3.1e-8	2.5e-8 to 3.7e-8

^a not applicable. ^b contaminant rate coefficients were estimated from the concentrations at 5 PV, $k_{TCE,W}$ was estimated from the rate of gas generation, $k_{eff,IC}$, α_{IC} and α_{CC} were taken from column C of Jeon et al. [66], T_C from the visual inspection of the SEM images.
^c from Prommer et al. [126]. ^d from Whist et al. [164]. ^e from Li et al. [96]. ^f from Jeon et al. [66]. ^g from Jin suk et al. [68].
^h from Zhang and Gillham [170]. ⁱ from Kamolporwiji et al. [72].

degradation is disregarded. However, if dissolution is ignored (as in the first experiment above), the corrosion reaction continues even when the total mass of solid iron is completely corroded. We evaluate the resulting prediction uncertainty by comparing these two extreme cases, i.e. full deactivation by iron dissolution (Eq.3.8 and Eq.3.9), and no deactivation by iron dissolution (Eq.3.6 and Eq.3.7).

The results of this second numerical experiment are shown in Fig. 3.8(c) and Fig. 3.8(d). Long-term predictions differ from the previous case, as contaminant breakthrough is simulated before 15 years, again with a faster breakthrough for the linear model. From Fig. 3.8 it can be concluded that the linear deactivation model predicts smaller lifetimes than the exponential model, as the former allows complete deactivation of the reactive surface. Hence, the linear model provides conservative estimates of the PRB longevity (with iron dissolution included), predicting in this case contaminant breakthrough (PRB failure) after 10 years. This estimate is site specific, as flow rates and contaminant and inorganic concentrations all influence PRB longevity (compare for example Fig. 3.8(a) with the Fig. A.3 of the appendix where the same model was forced with a different groundwater velocity). Another important result is that the use of a single deactivation model to predict the longevity of permeable reactive barrier might lead to misleading predictions. A better approach when the iron passivation mechanism is not precisely known is to provide longevity estimates using multiple passivation models.

3.3.5. Other sources of uncertainty in the long term prediction and strategies to reduce them

The development of a reactive transport model which can be calibrated to high velocity column experiment and applied to realistic low-velocity conditions should properly account for the main processes which are deemed relevant for the PRB longevity in the real site. The correct description of these processes in a reactive transport model is crucial when the model is applied under low flow velocities and must be supported by experimental data. As shown in the previous section, different estimates of PRB longevity were generated by different descriptions of the deactivation process. Core sampling could be performed during the barrier operation to estimate the iron reactivity at different times. In this manner field data could help to test the validity of the deactivation models.

Since the exponential deactivation constants were difficult to estimate using the stepwise approach mentioned above other approaches could be used. As

proposed by Jeen et al. [64], an exponential function could be directly fitted to the average column reactivity (estimated from the contaminant profiles at different pore volumes) and the total carbonate mineral fractions of the entire column (estimated from column mass balances or aqueous concentrations). In this way a direct estimation of the exponential deactivation constant related to the carbonate fraction could be provided.

Another source of uncertainty concerned the determination of the rate of biological sulfate degradation, which apparently proceeds too slowly to be estimated from our column data. This process may be relevant in field applications due to the active role of amorphous iron sulfide on contaminant degradation [150, 80]. A separate estimation of this degradation rate may be required, for example using an anaerobic batch experiment.

A good characterization of the minerals present at the end of the column experiment (XRD/EMPA analysis) is crucial to discriminate which minerals must be included the reactive transport model (sec.3.3.1). However, some minerals that are found in field PRBs (such as carbonate and sulfate green rusts [162]) may not show up in laboratory experiments that use relatively high flow rates. Rate coefficients for these minerals may then need to be based on values reported in literature, introducing additional uncertainty. Other sources of uncertainties include the effect of hydrogen gas production on the flow through the barrier and the effect of organic matter. Klausen et al. [78] performed column experiments to investigate the effect of solution composition on the reactivity of granular iron media. They concluded that natural organic matter may adsorb onto iron hydroxide surfaces reducing the degradation potential of the reactive media.

Finally, dissolved silica originating from the sand used in the PRB mixture may also form complexes that hinder access of contaminants to the iron surface. Furthermore, the role that silica plays in reduced iron reactivity may also be compound-specific [81].

3.4. Conclusions

This study presented an experimental and simulation study of a laboratory-scale iron PRB, with the aim of studying PRB longevity and long-term efficiency for degrading chlorinated solvents. The column experiment was performed using an accelerated flow rate in order to be able to observe contaminant breakthrough (PRB failure) within a reasonable amount of time (the experiment lasted 384 days). A numerical reactive transport model was de-

3.4 Conclusions

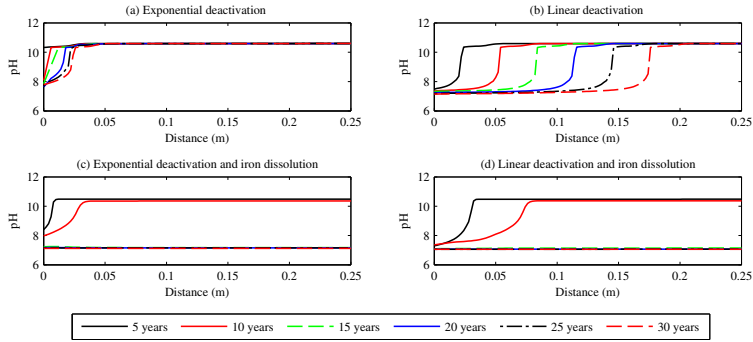


Figure 3.7.: pH profiles predicted by the different deactivation models

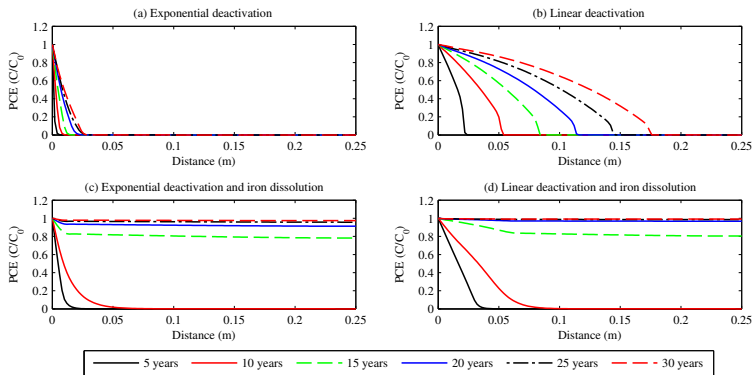


Figure 3.8.: Long-term efficiency under field condition. (a) PCE profiles with exponential deactivation; (b) PCE profiles with linear deactivation; (c) PCE profiles with exponential deactivation and iron dissolution; and (d) PCE profiles with linear deactivation and iron dissolution.

veloped that simulates the main physical and chemical processes in the PRB, including four different iron deactivation models from literature. Model parameters were estimated from literature and from the column data using an inverse approach. The model was subsequently used to predict PRB longevity at smaller flow rates that are more representative of field conditions.

The work presented here contributes in two ways. First, we have shown that care needs to be taken when predicting PRB longevity using currently available iron deactivation models. Column data in this study did not allow identification of the iron deactivation mechanism (linear vs exponential, iron hydroxide reactive or not), resulting in significant uncertainty of predicted longevity. A conservative strategy would for example rely on the shortest lifespan predicted by the various models; in our case this resulted in a predicted lifespan of 10 years. In all reported simulations, the maximum porosity reduction due to carbonate mineral precipitation was less than 25% (0.0083 per year, in the range reported in real studies) and its influence on contaminant degradation was small compared iron deactivation, independent of the deactivation model used.

The second contribution is methodological. We documented and successfully applied a method for estimating the multiple parameters of the reactive transport model used to simulate PRB longevity, thereby largely avoiding problems of poor parameter identifiability commonly encountered in such complex models. The adopted methodology consists of a careful initial and independent estimation of parameters, fixing non-sensitive parameters to their initial estimates, and joint automatic calibration of the remaining sensitive parameters. This procedure avoids over-parameterization and reduces correlation uncertainty of the estimated parameters.

Overall, our results point to a need for further research on iron deactivation mechanisms, the reactivity of iron hydroxide, and other potentially important effects related to hydrogen gas, silica, and organic matter, which were omitted from the current analysis due to a lack of quantitative understanding. Our results indicate that uncertainties in process knowledge should be acknowledged, and if possible, be quantified by including alternative process descriptions in the analysis. As new data and process knowledge become available, these can be incorporated into the models, resulting in a subsequent decrease in prediction uncertainty.

4. Comparing longevity estimations provided by geochemical and simplified models¹

The prediction of the PRB performance using only data from the column experiment was proven to be uncertain and dependent on the choice of the iron deactivation model. In this chapter the possibility of using a simplified model to predict the barrier longevity without explicitly accounting for inorganic geochemistry is explored. Since the barrier failure at the contaminated site is not yet observed, the extrapolations of the simplified model will be compared with those provided by the full geochemical model under different flow rates.

4.1. Simplified model description

The simplified model used in this study is based on the spatial temporal simulation model (STS) of Kouznetsova et al. [84], where the mineral front is described in a simplified way (Fig. 4.1). Compared to the original model formulation, the parameters describing the mineral precipitation front are functions of the effective groundwater velocity to allow the extrapolation from lab condition (fast velocity) to site condition (low velocity). At early times (Fig. 4.1(a)) the mineral front is not fully developed and the iron reactivity is depleted only at the column inlet, forming a transition zone TZ . Inside the transition zone the reactivity has its minimum at the column inlet and gradually increases to the initial value deeper into the column. At later time (Fig. 4.1(b)) the whole front has developed in the column and two distinguished zones forms, one before the deactivation front where the reactivity is depleted to the residual value RR_i and the front itself, where the reactivity increase from RR_i to the initial value. Ahead the front the iron media is fully reactive. At the time

¹Based on: Carniato L., Schoups G., and Bastiaens L. a. Seuntjens P. (2013). Simplified models to predict long term efficiency of Fe^0 permeable reactive barriers, Second European Symposium on Water Technologies & Management, 20-21 November 2013, Leuven, Belgium

equal to the deactivation period DP (Fig. 4.1(c)), the entire column length (x_{ref}) is deactivated.

Accordingly to the STS model, the deactivation front can be fully described by three parameters: the deactivation period DP (which determines the front velocity), the length of the transition zone TZ and the residual reactivity RR_i for contaminant i (which indicates the maximum reactivity depletion). When different residual reactivities RR_i for each contaminant are considered, the effect of iron deactivation on contaminant degradation is compound specific [159].

The effect of the deactivation front is taken into account by multiplying the initial reactive surface S_{r0} by the scaling factors $F_i(x, t)$. Therefore, the reactive surface for the contaminant i at location x and time t is expressed as:

$$S_{ri}(x, t) = S_{r0}F_i(x, t) \quad (4.1)$$

In this model the reactive surface is not an explicit function of the carbonate mineral fractions, where a geochemical model capable of describing mineral precipitation reactions is required (chapter 3). The simplified model does not require inorganic concentration measurements because the parameters DP , TZ and RR_i are directly estimated from the contaminant concentration data. To calculate the scaling factors $F_i(x, t)$ the following quantities normalized by the cross sectional area are required:

$$V_{ref,i} = x_{ref}(1 - RR_i) \quad (4.2)$$

$$V_{TZ,i} = \frac{TZ(1 - RR_i)}{2} \quad (4.3)$$

$$V_i(t) = \frac{t(V_{ref,i} - V_{TZ,i})}{DP} \quad (4.4)$$

where $V_{ref,i}$ is the iron volume deactivated to the minimum reactivity, $V_{TZ,i}$ is the normalized deactivated volume within regions of full and minimum reactivity and $V_i(t)$ is the total deactivated iron volume at time t . At early times the transition zone is not fully developed in the column ($V_{i(t)} \leq V_{TZ,i}$)

4.1 Simplified model description

and the distance from the column inlet up to which the precipitation front has propagated is:

$$x_{max,i}(t) = \sqrt{\frac{2V_i(t)}{\theta_i}} \quad (4.5)$$

Where θ_i is the slope of the front:

$$\theta_i = \frac{1 - RR_i}{TZ} \quad (4.6)$$

$x_{max,i}$ equals to TZ when the transition zone is completely propagated into the column. This occurs at time:

$$t^* = \frac{TZ}{2x_{ref} + TZ} DP \quad (4.7)$$

For times $t > t^*$, when $V_i(t) > V_{TZ}$, x_{max} is:

$$x_{max,i} = TZ + \frac{V_i(t) - V_{TZ,i}}{1 - RR_i} \quad (4.8)$$

Now the scaling factor $F_i(x, t)$ can be expressed as follow:

$$F_i(x, t) = RR_i \text{ for } x \leq x_{max,i} - TZ \quad (4.9)$$

$$F_i(x, t) = 1 - \theta_i(x_{max,i} - x) \text{ for } x_{max,i} - TZ < x \leq x_{max,i} \quad (4.10)$$

$$F_i(x, t) = 1 \text{ for } x > x_{max,i} \quad (4.11)$$

Note that the TZ can be wider than the column length, as might be expected for thin iron walls or high groundwater velocities. As can be seen from

Eq.4.2-Eq.4.4 only DP , TZ and RR_i parameters are required to track the movement of the front and to compute the scaling factors $F_i(L, t)$.

In order to extrapolate the barrier performance at low groundwater flow velocity (site conditions) the deactivation period DP , TZ and the RR_i were divided and multiplied by the effective groundwater velocity v_{eff} :

$$DP = \frac{DPN}{v_{eff}} \quad (4.12)$$

$$TZ = TZN v_{eff} \quad (4.13)$$

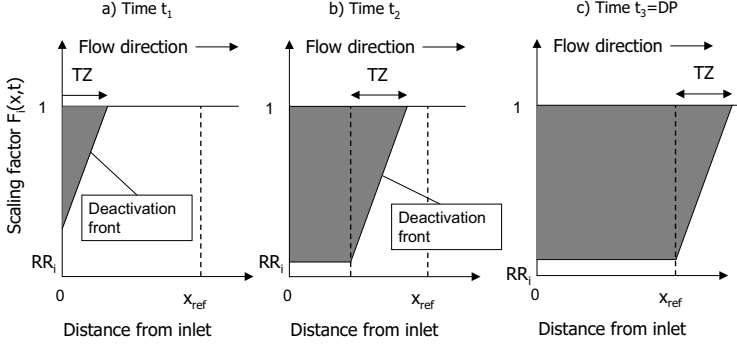
$$RR_i = v_{eff} RRN_i \quad (4.14)$$

the DPN , TZN and RRN_i parameters were calibrated using the contaminant concentration profiles at 5, 55, 160 and 209 PV of the column experiment described in chapter 3. The profiles at 244, 325 and 604 PV were used for model validation. For the site conditions the DP , TZ and RR_i parameters were calculated using a lower v_{eff} in Eq.4.12-Eq.4.14, as estimated at the site.

The STS model was implemented in RT3D [27], which is more efficient than PHAST for cases where modeling aqueous speciation reactions is not required. Indeed, a single STS model simulation completes in 7 seconds contrary to PHAST which requires about 20 minutes to solve the complete set of geochemical reactions. Advection, dispersion and contaminant degradation were accounted in the simplified model. The column physical parameters (porosity, dispersivities) and the boundary conditions (inflow rates and concentrations) were retained as in the full geochemical model (chapter 3). The stoichiometric coefficients of the contaminant degradation reactions were also kept fixed at the values used in the geochemical model (Tab.A.5). The degradation rates of chlorinated ethenes were modeled using a mixed order kinetic (Eq.3.1) and the reactive surface area was updated with the advancement of the mineral front.

4.2. Parameter inference

The calibration was performed in a stepwise fashion to avoid the simultaneous estimation of correlated parameters. The degradation parameters of the



Adapted from Kouznetsova et al. (2007)

Figure 4.1.: Description of the mineral front in the STS model

higher chlorinated ethenes (residual reactivities and degradation rates) were calibrated before the degradation parameters of lower chlorinated ethenes, assuming that the degradation products have no influence on the parent compound degradation rates (i.e. interspecies competition was not modelled). The parameters belonging to each contaminant were calibrated on the corresponding measurement set (e.g. PCE parameters were calibrated using PCE concentrations). An exception regards the *TZN* and *DPN* parameters, which are not compound specific and were estimated directly from PCE concentrations. For each contaminant i (Tab. 4.1) the following objective function was minimized:

$$\Phi_i(\phi) = \sum_{j=1}^{n_i} (y_{ij} - \eta_{ij}(\phi, X))^2 = \sum_{j=1}^{n_i} \varepsilon_{ij}^2 \quad (4.15)$$

where ε_{ij} is the j^{th} residual of the contaminant i . Since only one type of measurements was used in each step, set specific weights were not specified in Eq.4.15, assuming that all measurements contribute equally to the objective function. The calibration was performed with the global optimization algorithm SCE-UA [43].

4.3. Comparison between the geochemical and simplified model on the short term column data

The calibrated parameters are reported in Tab.4.1. As can be seen minor differences were found between the degradation rates determined with the STS and with the exponential deactivation model. Only for trans-DCE the degradation rates differ substantially, with the geochemical model providing an higher degradation rate. The RRN_i parameters reached their upper bounds for trans-DCE, cis-DCE and VC, fixed at 10% of the initial reactive surface value divided by the effective velocity. This can be seen in Fig.4.2(c) and Fig.4.2(d), where the reactive surface area at 604 PV is constant at $99.4 \text{ m}^2 \text{ L}^{-1}$. Higher values of the residual reactivity for trans-DCE and cis-DCE were estimated when the upper bound was increased, resulting in a underestimation of the contaminant concentration in the validation step.

In Fig.4.3 the calibrated PCE and VC profiles for the two models are shown. As can be seen, the simulated profiles are similar between the two models, as confirmed by the similar RMSE obtained in calibration (Tab.4.2). The main difference between the two models is how the reactive surface depletion is modeled. In the STS model the maximum depletion always occurs at the inlet (Fig.4.2(b)) whereas in the geochemical model occurs at 3 cm from the inlet, where the simulated pH rises enough to cause the precipitation of carbonate minerals (Fig.4.2(a)). The differences in the simulation of the deactivation mechanism are reflected in the simulated contaminant profiles. For example, PCE profiles at 160 and 209 PV at the column inlet are more flattened for the STS model compared to the geochemical model (Fig.4.3(a) and Fig.4.2(b)).

In the validation step the STS is not able to reproduce the measurements as well as the geochemical model (exponential deactivation without iron dissolution). The PCE profiles computed by the STS model (Fig.4.4(b)) are overestimated due to the low reactive surface area at 325 and 605 PV (Fig.4.2(b)). The reactive surface area estimated by the geochemical model at 604 PV is around $80 \text{ m}^2 \text{ L}^{-1}$ whereas for the STS model the PCE reactive surface area is $3.4 \text{ m}^2 \text{ L}^{-1}$. The worse fit obtained in validation for the STS model is also confirmed by the larger RMSE (Tab.4.2). In Tab.4.2 the log likelihoods for the two geochemical models (exponential and linear deactivation without iron dissolution) and for the STS model are reported. The log likelihood value was

4.3 Comparison between the geochemical and simplified model on the short term column data

estimated assuming a Gaussian distribution of the model residuals:

$$\log(l(\phi, \Sigma|Y)) \propto -\frac{1}{2}\log|\Sigma| - \frac{1}{2}\varepsilon^T \Sigma^{-1} \varepsilon \quad (4.16)$$

where Σ is the covariance matrix of the residual vector ε (here assumed a column vector containing all residuals ε_{ij}). The covariance matrix Σ was assumed diagonal (no residual correlation) and the variances to be equal within each measurements set i (which is equivalent to the fixed weight approach used in chapter 3). The residual variance of each set was estimated from the residuals calculated at the optimal parameter vector.

Table 4.1.: Calibrated parameters using the STS model and the geochemical model

Step	Parameter	STS Model	Geochemical model (exp. deactivation)
1	$\log k_{PCE-PRB}$ (mol m ² s ⁻¹)	-12.95	-13.01
	$\log DPN$ (s ² m ⁻¹)	2.05	-
	$\log TZN$ (s)	4.98	-
	$\log RRN_{PCE}$ (s m ⁻¹)	2.80	-
2	$\log k_{TCE-PRB}$ (mol m ² s ⁻¹)	-12.41	-12.47
	$\log RRN_{TCE}$ (s m ⁻¹)	4.29	
3	$\log k_{trans-DCE-PRB}$ (mol m ² s ⁻¹)	-13.85	-13.66
	$\log RRN_{trans-DCE}$ (s m ⁻¹)	4.29	
4	$\log k_{cis-DCE-PRB}$ (mol m ² s ⁻¹)	-12.17	-12.22
	$\log RRN_{cis-DCE}$ (s m ⁻¹)	4.29	
5	$\log k_{VC-PRB}$ (mol m ² s ⁻¹)	-13.30	-13.31
	$\log RRN_{VC}$ (s m ⁻¹)	4.29	

As can be seen from the table the higher log likelihood values were obtained for the geochemical model with the exponential deactivation formula, both in calibration and validation, although the STS model provides the lowest RMSE in calibration. In validation, the linear geochemical model has the lowest likelihood value and the highest RMSE error, indicating that even if the mineral precipitation reactions are included in the model, its performance is worse than that of the STS model. Therefore, the appropriate description of the reactive surface deactivation is crucial to obtain good results both in

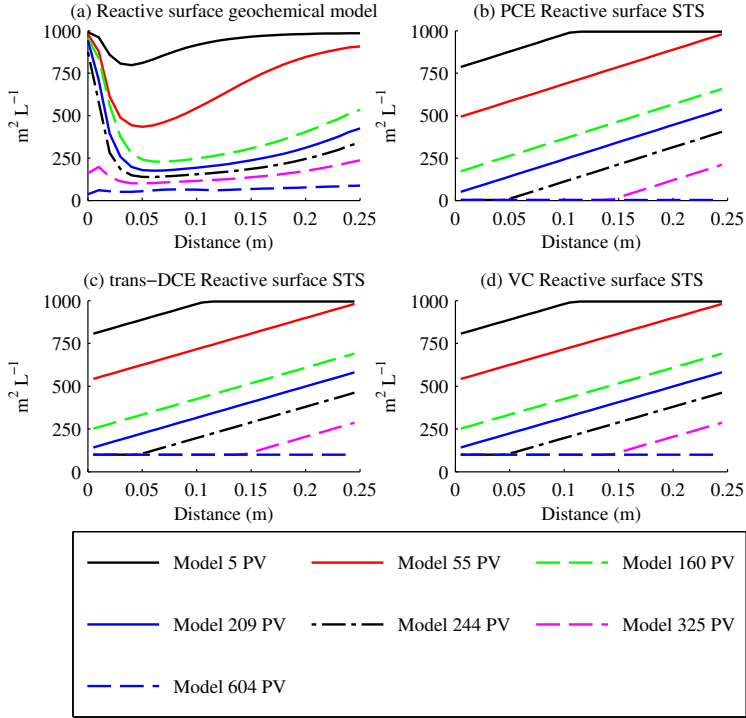


Figure 4.2.: Reactive surface area estimated by the geochemical model (a) and by the STS model ((b), (c) and (d)).

calibration and validation. From the table it can clearly be concluded that the exponential deactivation model with two deactivation parameters (α_{CC} and α_{IC}) is more likely given the data (note that the reported values are in logarithmic scale).

4.4. Long term predictions

Long term extrapolations were performed by lowering the effective groundwater velocity and recalculating the DP , TZ and RR_i parameters from the calibrated DPN , TZN and RRN_i . The calculation assumes that the front velocity scales linearly with the effective groundwater velocity (Eq.4.2-Eq.4.4).

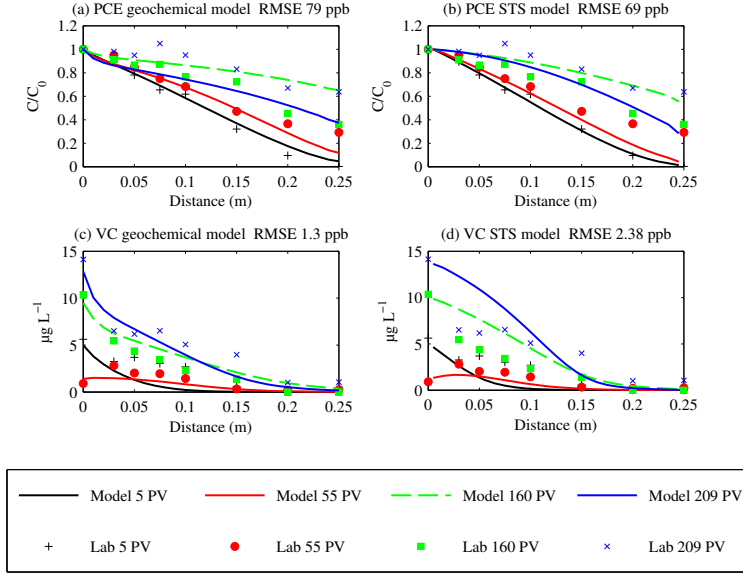


Figure 4.3.: Measured and simulated profiles for the geochemical ((a) and (c)) and the STS model ((b) and (d)) in the calibration step.

In Fig. 4.5 the PCE breakthrough curves estimated by the three models (STS and two geochemical models without iron dissolution) at different groundwater velocities are shown. From the figure it can be seen that for the geochemical model with exponential deactivation (dashed lines) the breakthrough time does not scale linearly with the effective groundwater velocity: the breakthrough time at 2 m y^{-1} is more than 2.5 times larger than that obtained with an effective groundwater velocity of 5 m y^{-1} . If the assumption that the front velocity scales linearly with the velocity holds also for the geochemical model, the breakthrough time for the red dashed curve should be equal to 68 years.

The relative difference between the breakthrough times provided by the STS and the exponential deactivation model increases as the effective groundwater velocity decrease (compared solid and dashed lines in Fig. 4.5). Such difference is related to the mineral precipitation rates and can be observed in Fig. 4.6, where the total reactive surface in function of time is shown. For the exponential deactivation model forced with 2 m y^{-1} effective groundwater velocity, the aragonite precipitation rate decreases about 30% in 12 years, causing a lower

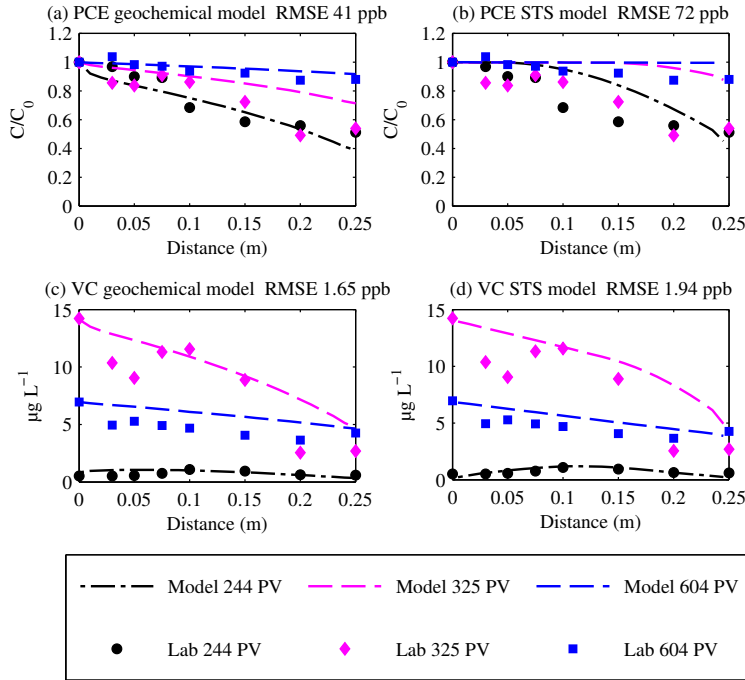


Figure 4.4.: Measured and simulated profiles for the geochemical ((a) and (c)) and the STS model ((b) and (d)) in the validation step.

deactivation and increasing the breakthrough time (Fig. 4.6(c)). Using the same model forced with a groundwater flow velocity of 10 m y^{-1} the aragonite precipitation rate decreased only of 8% in 2.5 years, even when the reactive surface begins to be substantially deactivated. In this case the aragonite precipitation rate is less sensitive to deactivation because the high carbonate load ensures an almost constant precipitation rate. As a results, the rate of reactive surface area deactivation is also constant and the STS and the exponential geochemical model provide similar estimations of the PRB longevity (Fig. 4.6(a)).

Assuming that the exponential geochemical model resembles the real deactivation mechanism more closely compared to the other models, it can be concluded that the STS can be used to extrapolate the PRB longevity for effective groundwater velocities larger than 10 m y^{-1} . However, this conclusion might be site-specific and depended on the carbonate concentration of

Table 4.2.: Log likelihood and RMSE values for the geochemical models and the STS model.

	Calibration		
	Exponential	Linear	STS
Log likelihood	-389.46	-390.55	-430.24
RMSE (ppb)	48.21	50.94	42.96
	Validation		
	Exponential	Linear	STS
Log likelihood	-304.33	-363.23	-327.49
RMSE (ppb)	35.73	61.65	46.40

the influent. The linear deactivation model tends to underestimate the barrier longevity because the reactive surface area is completely deactivated by the minerals precipitated after few years of operation.

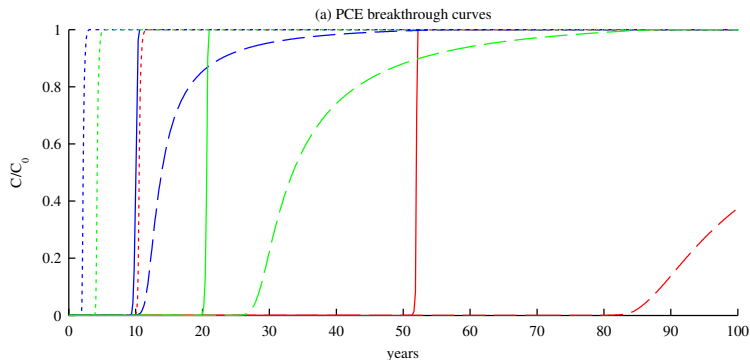


Figure 4.5.: PCE breakthrough curves estimated by the STS model (solid lines) and by the geochemical models (dashed lines correspond to the exponential deactivation model and dotted lines to the linear deactivation model without iron dissolution). Blue curves correspond to a groundwater velocity of 10 m y⁻¹, green curves to 5 m y⁻¹ and red curves to 2 m y⁻¹.

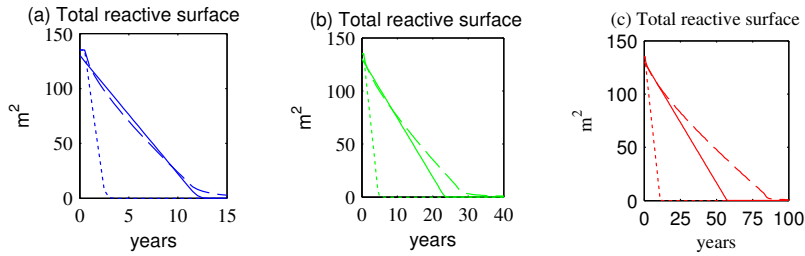


Figure 4.6.: Total reactive surface available for contaminant degradation estimated by the STS model (solid line), by the geochemical model with exponential deactivation (dashed line) and by the geochemical model with linear deactivation (dotted lines). Figure(a) corresponds to a groundwater velocity of 10 m y^{-1} , Figure(b) to a groundwater velocity of 5 m y^{-1} and Figure(c) to a groundwater velocity of 2 m y^{-1} .

4.5. Conclusions

In this chapter a simplified iron deactivation model (STS model) was described and used for the simulation of the column experiment presented in chapter 3. The model was proposed as an alternative to more complex geochemical models that explicitly account for complex mineral precipitation reactions.

The STS model was unable to reproduce the contaminant measurements in the validation step, due to an excessive deactivation of the reactive surface area at the end of the experiment. Moreover, the residual reactivities RR_i of the STS model could not be precisely estimated from the contaminant profiles used in calibration. Instead, the geochemical model with exponential deactivation provided good results both in calibration and validation, indicating that explicitly accounting for mineral precipitation reactions improves the simulation of the short term column experiment.

In order to extrapolate the barrier longevity under low flow velocity (field conditions) the front parameters estimated from the column measurements were scaled by the effective groundwater velocity. This procedure did not take into account the non-linearities present in the true geochemical system. In particular, at low flow rates the aragonite precipitation rate might slow down substantially, due to the depletion of the reactive surface area. In this case the STS model always provided a conservative estimation of the PRB longevity. The main advantages of the STS model over the geochemical models is that it

4.5 Conclusions

does not require inorganic concentration measurements for the estimation of its parameters and the shorter computation time.

5. Inference of reactive transport model parameters using a Bayesian multivariate approach¹

Abstract: Parameter estimation of subsurface transport models from multi-species data requires the definition of an objective function that includes different types of measurements. Common approaches are weighted least squares (WLS), where weights are specified a priori for each measurement, and weighted least squares with weight estimation (WLS(we)) where weights are estimated from the data together with the parameters. In this study, we formulate the parameter estimation task as a multivariate Bayesian inference problem. The WLS and WLS(we) methods are special cases in this framework, corresponding to specific prior assumptions about the residual covariance matrix. The Bayesian perspective allows for generalizations to cases where residual correlation is important and for efficient inference by analytically integrating out the (co)variances (or weights) from the joint posterior. Specifically, the WLS and WLS(we) methods are compared to a multivariate (MV) approach that accounts for residual correlations without the need for explicit weight estimation. When applied to inference of reactive transport model parameters from column-scale data on dissolved species concentrations, the following results were obtained: (1) accounting for residual correlation between species provides more accurate parameter estimation for high residual correlation levels whereas its influence for predictive uncertainty is negligible, (2) integrating out the (co)variances leads to an efficient estimation of the full joint posterior with a reduced computational effort compared to the WLS(we) method, and (3) in the presence of model structural errors, none of the methods is able to identify the correct parameter values.

¹Based on: Carniato, L., Schoups, G. and van de Giesen, N. (2014). Inference of reactive transport model parameters using a Bayesian multivariate approach. *Water Resources Research*, 50(8): 6406-6427.

5.1. Introduction

Reactive transport models are increasingly used to simulate complex chemical reactions in the subsurface. These models are typically characterized by flow, transport and kinetic parameters which need to be estimated through inverse modelling. During the last 30 years, inverse modelling methodologies have been extensively applied in the context of groundwater flow and contaminant transport [83, 167]. A recent review of the algorithms and methodologies used to infer flow and transport parameters from lab and field measurements is given by Vrugt et al. [157]. Inverse modelling of reactive transport is difficult because it is often ill-posed and different models might reproduce the same set of measurements. Hill and Tiedeman [58] report common issues encountered in the calibration of groundwater and transport models, suggesting that unreasonable parameter estimates and large error variances indicate the presence of model error. Matott and J. [108] developed a method where parameters and rate models are calibrated simultaneously from measurements, effectively demonstrating that different model formulations result in a similar fit to the data. Thus, one major question in reactive transport model applications is how deficiency in model structures can be identified and corrected, allowing the modeller to formulate models capable of providing reliable predictions. Gupta et al. [48] review how model error (or model structural adequacy) is addressed in different scientific communities (groundwater, unsaturated zone, terrestrial hydrometeorology, and surface water communities), including subsurface reactive transport modelling. In their work, it is acknowledged that identification of reaction networks from data is a new and challenging problem, requiring new methods to identify model inadequacy.

From a Bayesian point of view, parameters are estimated from measurements by their posterior density function (pdf) [44, 154, 113, 158]. Wagner [158] showed how groundwater heads and concentration measurements can be integrated in a unique function to estimate flow and transport parameters. In the approach of Wagner, the integration of different measurement groups in a unique objective function requires the estimation of the observation error variances, in order to normalize the contribution of different sets. A similar approach was adopted in Barlebo et al. [6], where error variances were estimated iteratively during parameter estimation. Medina and Carrera [113] and Fienen et al. [44] extended the approach of Wagner by including prior knowledge of the parameter values in the objective function. Vrugt et al. [154] considered the case where the distribution of the model residuals is not normal, providing a general form of the posterior parameter density function.

An important issue in the aforementioned studies is the specification of the covariance matrix of the model residuals. Different assumptions about the structure of the covariance matrix lead to different estimation methodologies. In generalized least squares (GLS, Draper and Smith [42]), the residuals (observed values minus simulated counterparts) are assumed to be normally distributed with zero mean and a known covariance matrix. The weighted least squares approach (WLS) further assumes that the covariance matrix is diagonal with known variances (with weights inversely proportional to the variances), whereas the ordinary least squares approach (OLS) also assumes all variances to be equal (homoscedastic). Several applications of these methods are reported in the literature. Hill et al. [57] used a WLS method to estimate hydraulic parameters of a synthetic aquifer using heads and flow data. Simunek et al. [140] and Arora et al. [2] used a WLS method to estimate the model parameters of a soil column experiment. Barth and Hill [7] used a WLS method to estimate model parameters for a virus transport model. Bell et al. [9] applied an OLS method to determine the sorption and diffusion parameters from a laboratory cell experiment. They used a data transformation to ensure that the assumptions of the OLS method were satisfied (uncorrelated and homoscedastic model residuals).

For the Maximum A Posteriori (MAP) estimation of the model parameters with the GLS, WLS and OLS methods, efficient gradient based algorithms have been used, such as the Levenberg-Marquardt algorithm (e.g. UCODE, Poeter et al. [123] and PEST, Doherty [38]). However, optimal parameter values found by these algorithms may represent a local minimum. To prevent the estimation of sub-optimal parameters, global optimization algorithms can be used instead. The application of heuristic algorithms to the minimization of a WLS and OLS objective function has been less extensively reported. Mertens et al. [114] and Shawn Matott and Rabideau [138] compared several algorithms for the minimization of WLS and OLS objective functions to infer the parameters of a complex biochemical model and a pesticide lysimeter test, respectively. Vrugt et al. [157] applied a multi-algorithm method to infer the parameters of a three-dimensional soil vapour extraction model, relying heavily on parallel computing.

Bayesian approaches aim to integrate the full joint parameter posterior, usually using efficient sampling techniques (such as Markov Chain Monte Carlo or MCMC sampling [156]) and allowing an accurate quantification of parameter and prediction uncertainty. The cost of integrating the joint posterior is much higher compared to identifying only its mode (MAP estimation) and the application of Bayesian approaches in reactive transport is still limited to few

studies [169, 130]

In most of the applications listed above, the WLS method was applied assuming that the total residual variance could be estimated a priori (for example, from an estimate of measurement and model error). In reactive transport models, residuals may be dominated by unknown model structural errors, for example due to the incomplete knowledge of the chemical reaction network. In this case, residual variances (or weights) can be estimated together with model parameters using an iterative two-stage procedure, referred to here as WLS with weight estimation, or WLS(we). In the first stage, the model parameters are estimated keeping the residual variances fixed, whereas in the second stage the maximum likelihood method is used to estimate the variances, keeping the model parameters fixed at values obtained in the first stage. The method is described in Sun [143] and applied in Dai and Samper [32, 31].

Finally, residuals may also be correlated (in space, in time, or between chemical species), as explicitly acknowledged in the GLS approach, where a full covariance matrix is specified a priori. However, to the extent that such correlations arise from a priori unknown model errors, it may be necessary to also estimate residual correlations from the data [134, 88]. For this purpose, typically simple correlation structures, such as first-order autoregressive models, are postulated a priori and their parameters estimated from the data.

In this study, we formulate the reactive transport model parameter estimation task as a multivariate Bayesian inference problem, thereby unifying and generalizing existing approaches. Specifically, the WLS and WLS(we) methods are compared to a multivariate (MV) approach to parameter inference that accounts for general (linear) residual correlations without the need for simultaneous estimation of the covariance matrix. The study is organized as follows. Section 5.2 presents the Bayesian framework for parameter inference. Section 5.3 introduces the case studies and reactive transport model used here as an application, section 5.4 compares the inference results obtained for the different methods for synthetic and real data and section 5.5 provides conclusions.

5.2. Bayesian framework

A reactive transport model generally contains a number of parameters, whose values are not known a priori, but which can be estimated from data, for example from measured dissolved species concentrations. Here a dataset consisting of n independent sampling events is considered, where for each sampling event one measures concentrations of m chemical species at a particular location and

time. The data can be assembled into a matrix n times m matrix Y where each row corresponds to a sampling event of m measurements. Data matrix Y can be subtracted to a corresponding matrix of simulated concentrations, $\eta(\phi)$, to yield a matrix of residuals ε ,

$$\varepsilon = F(Y) - F(\eta(\phi, Z)) = \begin{pmatrix} \varepsilon_{11} & \cdots & \varepsilon_{1j} & \cdots & \varepsilon_{1m} \\ \vdots & & & & \vdots \\ \varepsilon_{i1} & \cdots & \varepsilon_{ij} & \cdots & \varepsilon_{im} \\ \vdots & & & & \vdots \\ \varepsilon_{n1} & \cdots & \varepsilon_{nj} & \cdots & \varepsilon_{nm} \end{pmatrix} \quad (5.1)$$

where ϕ is the vector of model parameters, and F is a matrix of monotonic transformations (e.g. log transforms) applied to both measured and simulated concentrations, such that for each sampling event $i = 1, \dots, n$, data vectors $\varepsilon_{(i)}$ (corresponding to rows in ε) are (approximately) independent and identically distributed according to an m -variate zero-mean Gaussian $N_m(0, \Sigma)$ with m times m covariance matrix Σ . This Gaussian accounts for random fluctuations of the (transformed) residuals around zero due to a combination of measurement errors (e.g. sampling and analytical errors), model input errors (e.g. misspecified boundary conditions), and model structural errors (e.g. when an important reaction is not accounted for in the model). Hence, in this formulation errors from various sources are lumped together into an overall residual term, as opposed to explicitly separating out model and observation errors [130, 75]. Given these assumptions, the likelihood function $l(\phi, \Sigma | Y)$ is obtained as the product of n m -variate Gaussian [15],

$$l(\phi, \Sigma | Y) \equiv p(\varepsilon | \phi, \Sigma) = (2\pi)^{-\frac{mn}{2}} |\Sigma|^{-\frac{n}{2}} \exp \left(-\frac{1}{2} \sum_{l=1}^n \varepsilon_{(l)}^T \Sigma^{-1} \varepsilon_{(l)} \right) \quad (5.2)$$

where $|\Sigma|$ is the determinant of the covariance matrix. For example, in the case study reported below, we consider measurements of four different chemical species ($m = 4$), namely calcium, iron, total inorganic carbon or TIC, and pH, such that the Σ matrix becomes,

$$\Sigma = \begin{pmatrix} \sigma_{Ca^{2+}}^2 & cov_{Ca^{2+}Fe^{2+}} & cov_{Ca^{2+}TIC} & cov_{Ca^{2+}pH} \\ cov_{Fe^{2+}Ca^{2+}} & \sigma_{Fe^{2+}}^2 & cov_{Fe^{2+}Ca^{2+}} & cov_{Fe^{2+}pH} \\ cov_{TICCa^{2+}} & cov_{TICFe^{2+}} & \sigma_{TIC}^2 & cov_{TICpH} \\ cov_{pHCa^{2+}} & cov_{pHFe^{2+}} & cov_{pHTIC} & \sigma_{pH}^2 \end{pmatrix} \quad (5.3)$$

where σ_j^2 are the variances of the different species, and cov_{ij} are the covariances between species. Hence, whereas residuals are assumed independent between sampling events (different locations and times), residual correlation between species is taken into account. For our application, F consists of log-transforms, except for pH, which already is a log-transformed concentration. It will be useful to rewrite the likelihood in Eq.5.2 as follows:

$$l(\phi, \Sigma | Y) \propto |\Sigma|^{-\frac{n}{2}} \exp \left(-\frac{1}{2} tr \Sigma^{-1} S(\phi) \right) \quad (5.4)$$

where tr stands for trace or sum of diagonal elements of the $S(\phi)$ matrix defined as,

$$S_{ij}(\phi) = \sum_{u=1}^n \varepsilon_{ui} \varepsilon_{uj} \quad i, j = 1, \dots, m \quad (5.5)$$

The likelihood function is augmented with prior probability density functions (pdf) for ϕ and Σ , to yield the posterior pdf $p(\phi, \Sigma | Y)$, which forms the basis for estimating ϕ and Σ from the data,

$$p(\phi, \Sigma | Y) \propto l(\phi, \Sigma | Y) p(\phi) p(\Sigma) \quad (5.6)$$

where ϕ and Σ are assumed independent a priori. Following Box and Tiao [15], we assume a uniform prior for ϕ , i.e. $p(\phi) \propto 1$, such that the joint posterior becomes,

$$p(\phi, \Sigma | Y) \propto |\Sigma|^{-\frac{n}{2}} \exp \left(-\frac{1}{2} tr \Sigma^{-1} S(\phi) \right) p(\Sigma) \quad (5.7)$$

Various cases can now be distinguished depending on which prior assumptions are made about the covariance matrix Σ :

Σ is fixed a priori and assumed diagonal (WLS): In this case, the off-diagonal elements of Σ are zero, and the diagonal elements (variances) are specified a priori, e.g. based on known measurements errors (as in Tab.5.1 for the case

study described in the following section) or estimates of model errors (Hill and Tiedeman [58], pages 300-301). As such, the posterior reduces to,

$$p(\phi, \Sigma | Y) \propto \exp\left(-\frac{1}{2}tr\Sigma^{-1}S(\phi)\right) = \exp\left(-\frac{1}{2}\sum_{j=1}^m \frac{1}{\sigma_j^2}S_{jj}(\phi)\right) \quad (5.8)$$

with set-specific weights equal to σ_j^{-2} . This posterior has been widely applied in inverse modelling studies [57, 140, 7, 2]. Maximizing Eq.5.8 is equivalent to minimize its negative logarithm, which leads to the WLS objective function [21, 55] and to the solution of the WLS method [143]:

$$OBF_{WLS} = \sum_{j=1}^m \frac{1}{\sigma_j^2} S_{jj}(\phi) \quad (5.9)$$

Σ is unknown a priori and assumed diagonal (WLS(*we*)): Often it is not possible to specify Σ a priori, for example when the residuals are affected by a priori unknown model input and/or model structural errors. In this case, Σ needs to be estimated from the data [101]. A common approach is to assume that the off-diagonal elements of Σ are zero (no residual correlation between species), and to then estimate the diagonal elements of Σ (variances) from the data. If one takes a uniform prior, i.e. $p(\Sigma) \propto 1$, the posterior becomes,

$$p(\phi, \Sigma | Y) \propto |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2}tr\Sigma^{-1}S(\phi)\right) = \left(\prod_{j=1}^m \sigma_j^{-n}\right) \exp\left(-\frac{1}{2}\sum_{j=1}^m \frac{1}{\sigma_j^2}S_{jj}(\phi)\right) \quad (5.10)$$

where variances (one for each measurement set j) and model parameters are inferred jointly from the data. In principle the variances should be estimated for each residual ε_{ij} , because errors can be different at each time and location. However, this leads to an underdetermined inverse problem, where the number of parameters (model plus error parameters) exceeds the number of measurements. Therefore assumptions about the statistical properties of the

errors must be introduced [134]. In this study is assumed that the residual variance σ_j^2 to have the following form:

$$\sigma_j^2 = \sigma_{MEj}^2 + \sigma_{MOj}^2 \quad j = 1, \dots, m \quad (5.11)$$

where σ_{MEj}^2 is the measurement error variance (assumed to be known) and σ_{MOj}^2 is the remaining residual variance due to unknown model input and/or model structural errors (which typically derives from model application and was estimated). In our approach σ_j^2 is assumed to be constant within the set j (note that Eq.5.11 refers to the variance of transformed residuals). Maximizing the posterior Eq.5.10 leads to the solution of the WLS method with weight estimation or WLS(we), where weights are given by σ_j^{-2} for each specie j . Hence, weights are estimated for each measurement species j , unlike robust regression methodologies which aim to detect and adjust the weights of specific outliers [60].

An arguably better prior (than uniform) for Σ that is scale-invariant is the noninformative (Jeffreys) prior suggested by Box and Tiao [15], namely $p(\Sigma) \propto \prod_{j=1}^m \sigma_j^{-n}$. Note that this prior prefers small residual variances σ_j^2 , as opposed to the uniform prior used in Eq.5.10. It results in the following posterior,

$$p(\phi, \Sigma \mid Y) \propto \left(\prod_{j=1}^m \sigma_j^{-n-2} \right) \exp \left(-\frac{1}{2} \sum_{j=1}^m \frac{1}{\sigma_j^2} S_{jj}(\phi) \right) = \prod_{j=1}^m \left[\chi_n^{-2} \left(\frac{\sigma_j^2}{S_{jj}(\phi)} \right) S_{jj}(\phi)^{-\frac{n}{2}-1} \right] \quad (5.12)$$

where, given ϕ , each term $\frac{\sigma_j^2}{S_{jj}(\phi)}$ has an inverse-chi-square distribution with n degrees of freedom. It is useful to analytically integrate out the variances (and thus Σ), leading to the following marginal posterior for the model parameters [15],

$$p(\phi \mid Y) \propto \prod_{j=1}^m [S_{jj}(\phi)]^{-\frac{n}{2}} \quad (5.13)$$

An advantage of the marginalization step is that joint estimation of model parameters ϕ and the m variances (or weights) can be avoided. Instead, the model parameters can be inferred separately using their marginal posterior (for example using MCMC), following by conditional sampling of the variances. In other words, writing the joint posterior of ϕ and Σ given Y as:

$$p(\phi, \Sigma | Y) = p(\phi | Y)p(\Sigma | \phi, Y)$$

we first sample ϕ from the marginal posterior $p(\phi | Y)$ using MCMC, and then sample each variance σ_j^2 of the covariance matrix Σ from $p(\Sigma | \phi, Y)$, which amounts to drawing a number from a χ_n^{-2} distribution and multiplying it by $S_{jj}(\phi)$. Hence, no additional runs with the reactive transport model are required to estimate the full posterior once samples from the marginal posterior $p(\phi | Y)$ are available.

Σ is unknown a priori and allowed to be non-diagonal (MV): This is the most general case, which accounts for residual correlation between species (although still assuming independence between sampling events). Following Box and Tiao [15], we adopt a noninformative (Jeffreys) prior for Σ , i.e. $p(\Sigma) \propto |\Sigma|^{-\frac{1}{2}(m+1)}$, yielding the posterior (for $n \geq m$),

$$p(\phi, \Sigma | Y) \propto |\Sigma|^{-\frac{1}{2}(n+m+1)} \exp\left(-\frac{1}{2}\text{tr}\Sigma^{-1}S(\phi)\right) = W_m^{-1}(\Sigma; S(\phi), n)Z(\phi) \quad (5.14)$$

where $W_m^{-1}(\Sigma; S(\phi), n)$ is an inverse-Wishart distribution for Σ (given ϕ) with scale matrix $S(\phi)$, n degrees of freedom, and normalizing constant $Z(\phi) = c |S(\phi)|^{-\frac{n}{2}}$, with c a constant independent of ϕ . Integrating out Σ from the joint posterior then yields the following marginal posterior for ϕ [15],

$$p(\phi | Y) \propto |S(\phi)|^{-\frac{n}{2}} \quad (5.15)$$

with the conditional posterior of Σ , $p(\Sigma | \phi, Y)$, equal to $W_m^{-1}(\Sigma; S(\phi), n)$ and matrix $S(\phi)$ defined in Eq.5.5. As in the diagonal case, an advantage of the marginalization step is that joint estimation of model parameters ϕ and all $(1/2)m(m+1)$ distinct elements of Σ can be avoided. Similar to the diagonal case, the full posterior is obtained by MCMC sampling ϕ from $p(\phi | Y)$, followed by sampling a covariance matrix Σ from $W_m^{-1}(\Sigma; S(\phi_{MAP}), n)$.

The mode of $W_m^{-1}(\Sigma; S(\phi_{MAP}), n)$ provides a direct estimation of Σ for $\phi = \phi_{MAP}$:

$$\Sigma_{MAP} \approx \frac{S(\phi_{MAP})}{n + m + 1} \quad (5.16)$$

an important constraint in arriving at equation Eq.5.15 is that the number of measurement events n is equal to or greater than the number of different groups m . Hence, assigning each data point to its own group is not allowed, since then $n = 1 < m$. As indicated above, here we group residuals by species, but, for our case studies, alternatively grouping residuals by location (depth) or time, to account for correlation in space or time, would also comply with the constraint $n \geq m$. Finally, note that in deriving Eq.5.15, no assumptions were made, beyond grouping the residuals, about the (linear) correlation structure. This is in contrast to using an a priori fixed correlation structure, such as a first-order autoregressive model [134, 88].

For the case studies reported below we use the DREAM-ZS (DiffeRential Evolution Adaptive Metropolis) algorithm developed by Vrugt et al. [155] for MCMC sampling from the parameter posteriors, with the total number of MCMC samples equal to the number of parameters multiplied by 2000. This value was generally sufficient to achieve convergence of the MCMC algorithm.

5.3. Application to reactive transport model

5.3.1. Model modifications

The WLS, WLS(we) and MV methods were used to infer the parameters of the column-scale reactive transport model described in chapter 3. Compared to chapter 3 some changes have been applied.

First, the reactive transport code PHAST was substituted with MIN3P [110]. Similarly to PHAST, MIN3P solves saturated groundwater flow and reaction chemistry using the partial equilibrium approach [141, 98]. However, MIN3P adjusts the time increments adaptively depending on the stiffness of the geochemical integration problem and uses a globally implicit approach to couple transport and reactions, which suppresses the numerical diffusion introduced by the operator splitting. This allows increasing the time stepping with small loss of accuracy when an upstream spatial weighting scheme is used.

These technical details are important in that they greatly reduce the CPU time required for one forward simulation, compared to the PHAST implementation and allow the inference of the full joint posterior with MCMC sampling. Since in our system the major geochemical changes occur at the beginning of the simulation, with only gradual changes at later times, time steps in MIN3P can be increased without a substantial loss of accuracy. In our runs, MIN3P was up to 200 times more efficient than PHAST for the simulation of the first 209 PV. Some preliminary simulations were performed to compare the simulation obtained with an accurate numerical scheme and discretization (central spatial weighting with grid Péclet number equal to 1.5 and grid Courant number equal to 0.5) with those obtained with the upstream spatial weighting used in the following case studies. Small differences were found between the two simulations, indicating that the parameter values inferred by the different estimation methods can safely be considered independent of the space and time discretization.

Secondly, only the inorganic chemistry was modeled since the aim of the study is to compare different methodologies to infer model parameters and not to provide predictions of the barrier contaminant remediation efficiency. The inorganic part was assumed not to be affected by low contaminant concentrations.

Finally, it was also assumed that iron can be oxidized by carbonate anions [109, 129] according to the following rate:

$$r_{IrCTIC} = -k_{IrC,TIC}S_r \frac{[CO_3^{2-}]}{k_{h,CO_3^{2-}} + [CO_3^{2-}]} \quad (5.17)$$

where $k_{IrC,TIC}$ is the iron corrosion by carbonate rate coefficient per unit of iron reactive surface area ($\text{mol m}^{-2} \text{s}^{-1}$) and $k_{h,CO_3^{2-}}$ is the half saturation constant reported in the MIN3P database ($10^{-8} \text{ mol L}^{-1}$). From preliminary simulations it was found that the addition of corrosion by carbonate improved the fit to the iron measurements. The reason can be deduced from Tab. A.5, where anaerobic iron corrosion by water produces 2 moles of OH^- for each mole of zero valent iron oxidized iron whereas anaerobic iron corrosion by TIC consumes 2.5 moles H^+ . Therefore, to increase the pH required for mineral precipitation less iron needs to be corroded when iron corrosion by carbonate is included in the model, reducing the overestimation of the iron measurements (Fig. 3.6).

In this study only the exponential deactivation model without the dissolution

term (Eq.3.6) was used. As indicated in chapter 3, the dissolution term does not have an appreciable influence on the simulations of the short term column experiment and the exponential deactivation model could fit the data slightly better than the linear deactivation model.

Based on a preliminary sensitivity analysis, seven parameter were considered to affect the inorganic simulation the most: the rate coefficient for iron corrosion by water ($k_{IrC,W}$, Eq.3.2), the rate coefficient for iron corrosion by carbonate ($k_{IrC,TIC}$, Eq.5.17), the precipitation rate coefficients for aragonite and iron hydroxide carbonate ($k_{eff,CC}$ and $k_{eff,IC}$ in Eq.3.3), the deactivation constants for aragonite and iron hydroxy carbonate (α_{CC} and α_{IC} in equation Eq.3.6) and the solubility of iron hydroxy carbonate (K_{CC}^m in Eq.3.3). The solubility of iron hydroxy carbonate was assumed adjustable because the logarithmic value of this parameter has been reported to vary significantly from -1.56 [92] to 0.075. Flow and transport initial and boundary conditions for the model were maintained as described in sec. 3.2.4.

5.3.2. Case studies

The methods presented in sec. 5.2 were first tested on two synthetic datasets generated from the model described above using the following procedure. First, model simulations of each set were generated at 0.03, 0.075, 0.15 and 0.25 m from the column inlet at 5, 55, 110, 160 and 209 pore volumes (PV) using the correct parameter values, yielding a 20 by 4 model simulation matrix $\eta(\phi)$ (20 measurements events for each set). Then the model simulations were log transformed and corrupted with measurement errors generated from a multivariate Gaussian normal distribution $N_m(0, \Sigma)$ with known Σ (20 independent samples were extracted from $N_m(0, \Sigma)$, yielding a 20 by 4 matrix of residuals ε , which were added to the model simulation matrix $\eta(\phi)$ to obtain the measurement matrix Y). The diagonal elements of Σ were set equal to the estimated σ_{MEj} reported in Tab. 5.1, fifth column. For the first dataset (low correlation) the off diagonal elements of Σ were calculated such that the off diagonal elements of the relative correlation matrix were all equal to 0.1. For the second synthetic dataset (high correlation), the off-diagonal elements of the correlation matrix Σ were calculated such that the off diagonal elements of the correlation matrix were all equal to 0.9. Due to the small sample size (20 draws) the Σ_{res} matrix estimated from the residuals [34] was different from the Σ used to generate the dataset. The maximum absolute off-diagonal value of Σ_{res} was 0.44 for the low correlation case, whereas the minimum absolute off-diagonal value of Σ_{res} was 0.82 for the high correlation case.

In the first two cases (synthetic case with low and high residual correlation), we used the WLS, WLS(we) and MV methods to infer the rate coefficients of iron carbonate precipitation, iron corrosion by water, aragonite precipitation and iron corrosion by carbonate ($k_{eff,IC}$, $k_{IrC,W}$, $k_{eff,IC}$ and $k_{IrC,TIC}$) from the two synthetic datasets. In these cases, the model includes all the reactions used for the generation of the synthetic data and the only source of error is the measurement error. In the third case (synthetic case with model error), the precipitation of iron carbonate was excluded from the model and the iron corrosion by water, aragonite precipitation and iron corrosion by carbonate ($k_{IrC,W}$, $k_{eff,IC}$ and $k_{IrC,TIC}$) were calibrated using the low correlation dataset. This experiment was performed to assess the performance of the three methods in a case where model error is introduced.

In the fourth case, the column measurements were used to infer all sensitive parameters. In addition to the parameters inferred in the synthetic cases, the deactivation constants for iron carbonate and calcium carbonate (α_{IC} and α_{CC} in Eq.3.6) and the equilibrium constant for iron hydroxy carbonate (K_{CC}^m in Eq.3.3) were included in the parameter inference. Therefore, seven parameters were estimated using the measurements at 0.03, 0.075, 0.15 and 0.25 m from the column inlet at 5, 55, 110, 160 and 209 pore volumes (PV), similarly to the synthetic cases. The measurements available at 244, 325 and 604 PV were used as a validation dataset.

5.3.3. Assessing model structural error

When residuals are not only affected by measurement errors, but also by model structural errors (and/or model input errors), estimated residual variances σ_j become larger than corresponding measurement error variances σ_{MEj} , and the excess variance σ_{MOj} can be used as a measure of model error. The WLS(we) method is the only one that explicit estimate the variances σ_{MOj} (Eq.5.11). For the other methods σ_{MOj} were estimated subtracting σ_{MEj} (assumed to be known a priori) from σ_j . For the WLS method the total residual variance σ_j was estimated from the residuals $\varepsilon_{i,j}$ computed at the MAP parameter estimation. For the MV method the total residual variances at the MAP parameter estimations were obtained from the diagonal elements of the Σ_{MAP} matrix (Eq.5.16). Another measure of model error is the standard error (SE),

$$SE = \sqrt{\frac{OBF_{WLS}}{mn - p}} \quad (5.18)$$

where $OBFWLS$ is defined in Eq.5.9 (with $\sigma_j = \sigma_{MEj}$) and p is the number of model parameters. As reported by Hill and Tiedeman [58], SE should be equal to one in the absence of model error and when the weighting scheme reflects data accuracy. In this study the SE was used to indicate the presence of model errors and the estimated σ_{MOj} to indicate the species that were affected the most by it. The SE was also used to measure goodness of fit.

5.4. Results

Results are organized into three subsections. First, we investigate the effect of residual correlation on parameter and predictive uncertainty by means of synthetically generated datasets (first and second case study from sec. 5.3.2). Second, we evaluate performance of the various methods in the presence of model structural errors by means of both a synthetically generated dataset and a real dataset (third and fourth case study from sec. 5.3.2). Finally, computational efficiency of the different methods is compared.

5.4.1. Effect of residual correlation

In Fig. 5.1(a-d) the profiles simulated with the correct parameters and the synthetic data corrupted with low correlated noise level are shown. The profiles obtained with the MAP parameter set identified by each method are hardly distinguishable from the correct simulations and the different results are not shown (the corresponding RMSE values were almost identical to those indicated in Fig. 5.1). From the Figure it can be seen that larger residuals were generated for calcium and TIC near the influent and lower ones for iron and pH, which agrees with the specified $\sigma_{MEij,untr}$ reported in Tab. 5.1 (larger for calcium and TIC and proportional to the magnitude of the measurement). No clear correlation is visually evident between residuals for different species.

In Fig. 5.1(e-h) the simulated profiles with the correct parameter values and high correlated residuals are shown (as for the low correlations case, the simulations with the MAP parameters were very similar for all methods). From the Figure it is evident that a positive residual correlation exists. An example is the sampling event at 55 PV at 0.075 m from the column inlet, where the calcium, iron and TIC measurements are lower than the correct concentration profiles. In real applications, systematic over or underestimation might be caused by model errors and/or misspecified boundary conditions, for example

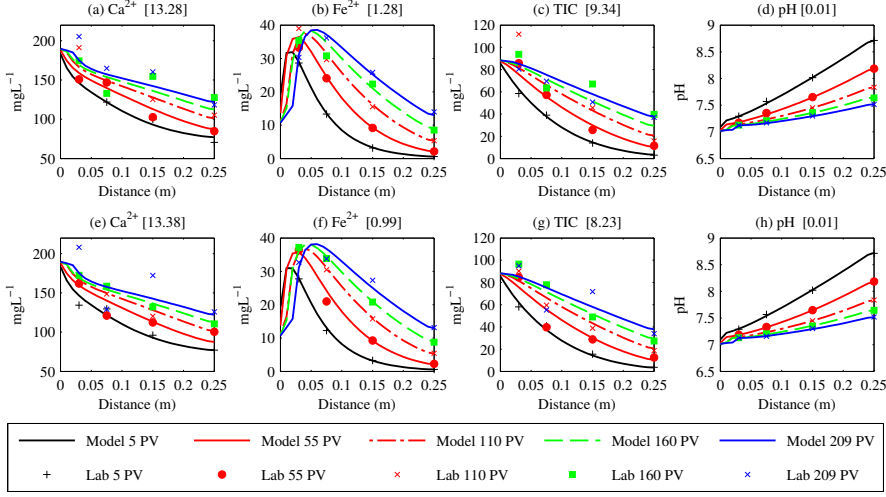


Figure 5.1.: Species concentrations simulated with the correct parameter values (solid lines) and synthetic measurements (symbols) based on artificial residuals with either low (top row) or high correlation (bottom row). In square brackets the RMSEs for the correct parameter values are reported.

due to unaccounted variations of inflow rates and in field applications due to well construction problems.

Normal probability plots [58, 42, 24] for weighted residuals are used to check the Gaussian assumption underlying all methods studied here; weighted residuals are considered to be normally distributed if they plot on a straight line. Weights for each measurement set were specified inversely proportional to the residual standard deviations computed as reported in Tab. 5.1 (columns 5 to 7). For the MV method graphical checks of multivariate normality can also be performed using multivariate normality plots based on the Mahalanobis distance which accounts for residual correlation [104]. However, these plots are not directly comparable to normality plots of the WLS and WLS(we) methods, because the multivariate normality plots display a point for each measurement event (consisting of measured concentration for each of the 4 species), resulting in a plot with only 20 points instead of a normality plot with 80 points. Normality plots can be used to check normality of the weighted residuals also for the MV method, because if the prior assumption of the multivariate residual distribution $N_m(0, \Sigma)$ holds, the marginal distribution of the residuals belong-

ing to the set j must be normal with zero mean and standard deviation σ_j . Substantial deviations of the weighted residuals from the normality line certainly indicate that the multivariate normality assumption is not satisfied. In addition to normality plots, Lilliefors tests [100] at 5% significance level were also used to test for normality. The results in Fig. 5.2 show that the residual distributions can be considered normal for all methods, for both high and low correlation levels.

At low correlation, residuals are more aligned to the normality line compared to the high correlation case, with the WLS method having mean and standard deviation closest to 0 and 1, respectively (Fig. 5.2(a)). The adjustment of the weights in the WLS(we) method improved the alignment to the normality line for the calcium residual (209 PV, 0.03 m) located at the positive tail of the WLS residual distribution (Fig. 5.2(b)). A similar effect can also be seen for the MV method, where the residual variances (i.e. the weights) were integrated out from the joint posterior (Fig. 5.2(c)). In this case the standard deviation of the normalized residual distribution was slightly larger than one. Similar results were also obtained for the high correlation case.

In Fig. 5.3(a-d) and (j-m) the posterior cumulative distribution functions (CDFs) of the model parameters for the low and high correlation cases are shown. As can be seen, the methods provide similar results at low correlation level, with the lowest L_2 norm of the bias vector (defined as the difference between the correct and the MAP parameter vector) obtained for the MV method. Also for high correlation levels the MV method provides the best MAP parameter estimation and lower parameter uncertainties, as can be noted from the steep green curves close to the correct values. Such result is due to the fulfillment of the statistical assumptions about the residual correlation, which were considered only in the MV method. An additional experiment with an increased measurement error variance (50% increase) was performed, with the MV method providing the best estimates of the correct parameter values with narrower confidence intervals. In this test, the true value of $\log_{\text{keff},IC}$ was included in the 5-95% confidence interval only for the MV method. From these results it is clear that the MV method is superior to the WLS and WLS(we) methods in identifying the correct parameters in the case of highly correlated residuals and in absence of model errors.

The posterior CDFs of the estimated variances are shown in Fig. 5.3(e-h) and Fig. 5.3(n-q). As can be seen, the WLS method uses fixed values of the residual variances, based on the prior knowledge of the measurement errors (Tab. 5.1), which might differ from the variances computed from the correct residuals (in these synthetic cases such difference is attributable to the small sample size).

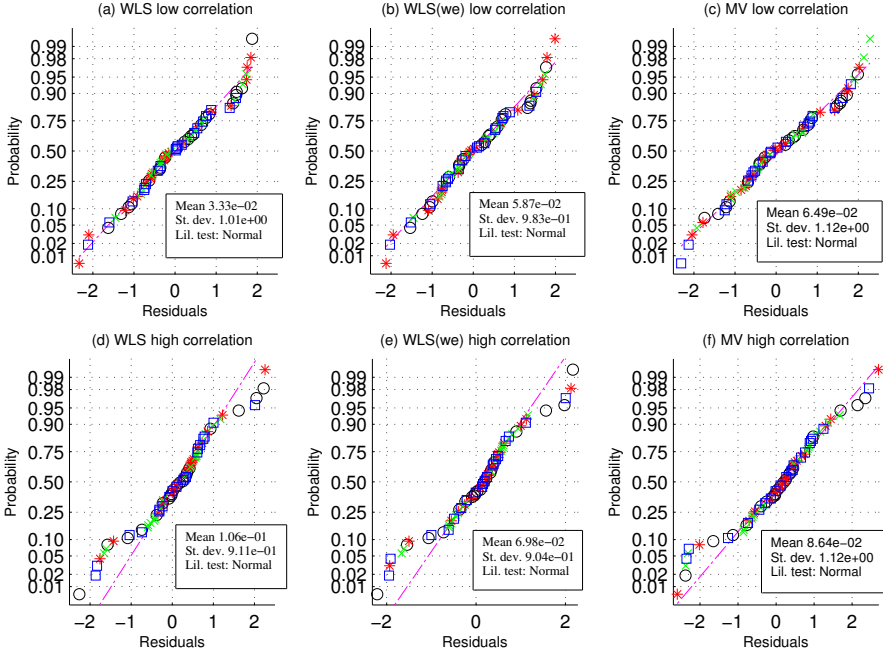


Figure 5.2.: Normal probability plots of the weighted residuals for the synthetic case without model error: plots (a) to (c) show the residual distributions for the low correlation case, the plots (d) to (f) for the high correlation case. Black circles represent calcium residuals, red asterisks iron residuals, green crosses TIC residuals and blue squares pH residuals. Lilliefors tests at 5% significance level indicate normal residual distributions for all cases.

Such result is also confirmed in Tab. B.1(b) and Tab. B.1(b) of the appendix, where positive σ_{MOj} were estimated for the WLS method despite the absence of model errors. For the WLS(we) method the residual variance was estimated by adding a positive σ_{MOj}^2 to the measurement error variance (Eq.5.11). For the cases where the sample variance was lower than the measurement error variance σ_{MEj} , adding a positive value did not improve the estimations (as can be seen in Fig. 5.3(g) for the low correlation case and Fig. 5.3(o-q) for the high correlation case). Making fewer assumptions on the residual variances as in the MV method has the advantage of providing more reliable estimations, with the correct variances always included in the 5-95% confidence intervals.

The MV method was also able to provide a reliable estimation of the covariance elements of Σ_{res} matrix, as shown in Fig. 5.3(i) and (r) for the $cov_{Ca2+Fe2+}$ element and in Fig. B.1 of the appendix for the other elements. As can be seen from Fig. 5.3, $cov_{Ca2+Fe2+}$ is zero for the WLS(we) and WLS methods, as assumed a priori.

Fig. 5.4 shows the predictive QQ plot and predictive uncertainty bounds for all measurements used in calibration. When the empirical cumulative distribution function of the sampled values at each observation agrees with the uniform distribution (correspondence with the 1:1 line), the hypothesis made in the calibration framework are consistent with the data [144]. For the low correlation case (Fig. 5.4(a)) all methods approach the 1:1 line. In this case, the amount of measurements included in the 95% prediction is closest to the correct value for the WLS method. For the high correlation case (Fig. 5.4(b)) the different methods also perform similarly, with the MV method marginally closer to the 1:1 line. In both cases, the MV method slightly overestimates predictive uncertainty, as can be seen in Fig. 5.4(c) and Fig. 5.4(d), where slightly larger predictive bounds were obtained for calcium and iron. Such slight overestimation of predictive uncertainty could also be guessed from Fig. 5.3, where wider posteriors for the variances were obtained for the MV method. Overall, Fig. 5.4 shows that accounting for correlation has a negligible effect on predictive uncertainty.

The SE measure (Fig. 5.4(a) and Fig. 5.4(b)) correctly indicates the absence of model errors, since all values are close to one both at high and low correlation levels. The σ_{MOj} values for each measurement set are reported in Tab. B.1 and Tab. B.2 of the appendix; these values are close to zero, indicating no model error, as expected, but not exactly zero due to the small sample size.

5.4.2. Effect of model structural error

5.4.2.1. Synthetic case with model error

In this case, the methods were used to infer parameters related to iron corrosion by water, aragonite precipitation and iron corrosion by carbonate ($k_{IrC,W}$, $k_{eff,CC}$, and $k_{IrC,TIC}$), while excluding precipitation of iron hydroxy carbonate from the model. In Fig. 5.5, the simulated profiles provided by the different methods are shown.

As can be seen, iron and TIC concentrations were largely overestimated, whereas a relatively good fit was obtained for calcium. More carbonate (TIC) remains in solution because carbonate precipitation in the form of iron hy-

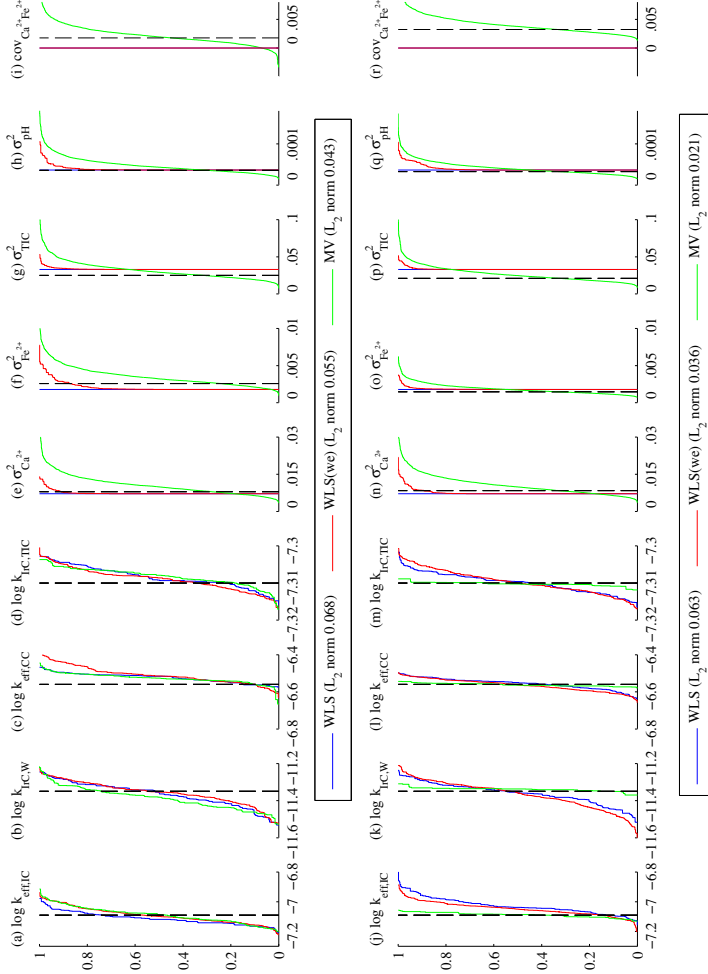


Figure 5.3.: Posterior cumulative distribution functions for model parameters and residual variances for the low (figures (a) to (i)) and high correlation (figures (j) to (r)) cases. The black dashed lines represent the correct parameter values and the variances computed from the correct residuals. In the legend boxes the L_2 norms of the bias parameter vectors are reported.

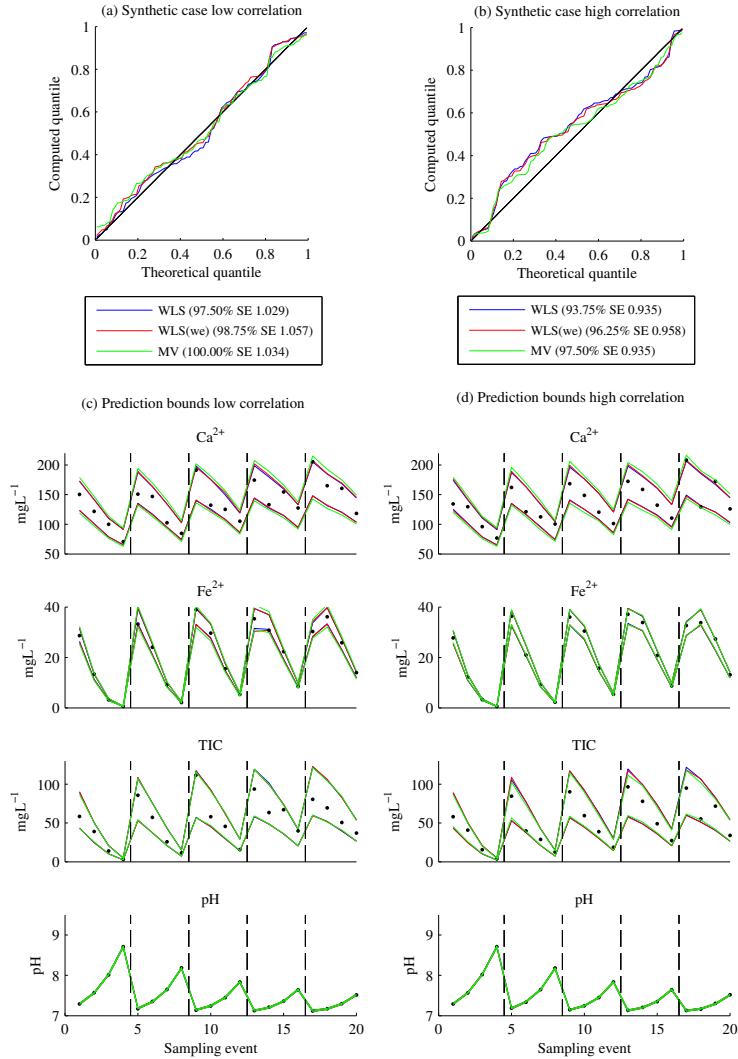


Figure 5.4.: Predictive QQ plot and uncertainty bounds for the synthetic case with low residual correlation (figures (a) and (c)) and with high residual correlation (figures (b) and (d)). In the legend boxes the percentage of measurements contained in the 95% prediction bounds and the standard error values are reported. Dashed lines delimit the pore volumes (5, 55, 110, 160 and 209 PV). Blue, red and green lines correspond to the WLS, WLS(we) and MV methods respectively.

Table 5.1.: Experimental errors for each of the measurement sets

Dataset	Measure- ment error ^b	Standard deviation $\sigma_{MEij,untr}$ ^c	Transfor- mation F	Standard deviation σ_{MEj}^2 , WLS ^d	Standard deviation $\sigma_j = \sqrt{\sigma_{MEj}^2 + \sigma_{MOj}^2}$, WLS(we)	Covariance matrix Σ , MV
Calcium	14%	$\frac{0.14C}{1.65}$	Log	$\frac{0.14}{1.65}$	$\left[\left(\frac{0.14}{1.65} \right)^2 + \sigma_{MO_{Ca^{2+}}}^2 \right]^{0.5}$	Σ_{MAP}
Iron	7%	$\frac{0.07C}{1.65}$	Log	$\frac{0.07}{1.65}$	$\left[\left(\frac{0.07}{1.65} \right)^2 + \sigma_{MO_{Fe^{2+}}}^2 \right]^{0.5}$	Σ_{MAP}
TIC ^a	30%	$\frac{0.3C}{1.65}$	Log	$\frac{0.03}{1.65}$	$\left[\left(\frac{0.30}{1.65} \right)^2 + \sigma_{MO_{TIC}}^2 \right]^{0.5}$	Σ_{MAP}
pH	0.01 pH units	$\frac{0.01}{1.65}$	None	$\frac{0.01}{1.65}$	$\left[\left(\frac{0.14}{1.65} \right)^2 + \sigma_{MO_{pH}}^2 \right]^{0.5}$	Σ_{MAP}

^a total inorganic carbon.^b laboratory information.^c standard deviation was calculated assuming a normal distribution of the residuals [58].
^d weights for the WLS and WLS(we) were calculated as the inverse of the standard deviations. Note that for the WLS case, σ_{MEj} is assumed equal to the coefficient of variation cv_j of the residuals before the transformation ($\sigma_{MEij,untr} = cv_j Y_{ij}$). This assumption is acceptable for small σ_{MEj} values.

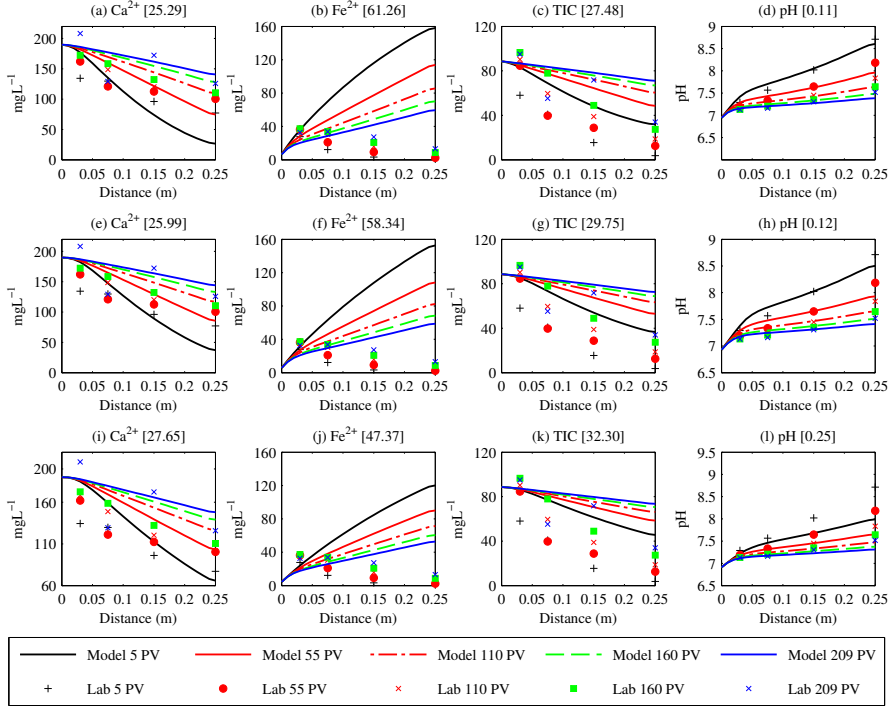


Figure 5.5.: Model fits using WLS (top row), WLS(we) (central row) and MV method (bottom row) for the synthetic case with model error. In square brackets the RMSEs are provided.

droxy carbonate was excluded from the model. As a result, the ion activity product (IAP) of aragonite is higher and parameter $k_{eff,CC}$ is underestimated by more than one order of magnitude (see Tab. B.3 in combination with Eq.3.3 which shows that the precipitation rate of calcium carbonate is proportional to $k_{eff,CC}$ and IAP). Moreover, the $k_{IrC,W}$ parameter almost reaches its lower bound ($\log k_{IrC,W} = -12$), which already suggests the presence of model error. In this case, none of the methods was able to retrieve the correct parameters.

The residual distributions at the MAP parameter estimates (Fig. 5.6(a-c)) deviate from normality for all methods, with heavy negative and positive tails (corresponding to an overestimation of the dissolved iron and the underestimation of pH). The WLS method, which uses a fixed weight approach, deviates

the most from normality, whereas the WLS(we) and MV methods give better results, but with the Lilliefors test suggesting non-normality for all methods. Therefore, the statistical assumptions made in the calibration phase were not met in this synthetic case with model error.

The posterior parameter CDFs provided by the different estimation methods are shown in Fig. 5.7(a-c). The WLS method provides the sharpest CDFs, underestimating parameter uncertainty. Such underestimation can be attributed to the assumption that the residual variance only consists of measurement errors, whereas in this test model error is also an important component of the residual variance. The WLS(we) and MV methods provide different CDFs with the correct value of the $\log k_{IRC,W}$ parameter included in the 5-95% confidence interval. The correct value of $k_{eff,CC}$ was included in the 5-95% confidence interval only for the MV method (Fig. 5.7(b)), and none of the methods were able to identify the correct value $k_{IRC,TIC}$ value, indicating that the proposed methods can not correct the large errors induced by omitting iron hydroxy carbonate precipitation from the model. This is also confirmed by the large L_2 norms of the bias vector reported in the top legend of Fig. 5.7.

Different parameter estimates between the WLS(we) and MV methods can mostly be attributed to the different assumptions made about residual correlation. This hypothesis was confirmed by performing an additional experiment in which only the diagonal elements of the $S(\phi)$ matrix were considered a priori (referred to as MV(diagonal) in Fig. 5.7), corresponding to the assumption of no residual correlation and resulting in the Eq. 5.13 posterior. Neglecting residual correlation resulted in parameter CDFs that were indeed very similar to those obtained with the WLS(we) method (Fig. 5.7(a-c)). Small differences between the WLS(we) and MV(diagonal) parameter posteriors may be due to different assumptions about the prior of the covariance matrix (uniform in the WLS(we) case and Jeffreys prior in the MV methods). The MV(diagonal) method consistently finds σ_j^2 values that are smaller than the WLS(we) method (Fig. 5.7(d-g)), which may be related to the Jeffreys prior preferring small variances. Note that using the Jeffreys prior in the MV method does not always improve the estimation of the variances when model error is present, as can be seen in Fig. 5.7(d), where the MV method underestimates the correct calcium residual variance. Moreover, the introduction of model error did not allow the correct inference of the covariances (Fig. 5.7(h)).

The WLS(we) and MV methods adjust the residual variances to decrease the weights of extreme outliers (mostly iron measurements) and increase the value of the respective posterior densities (Tab. B.3(a) of the appendix). The WLS method underestimates the correct residual variances, which are fixed at the

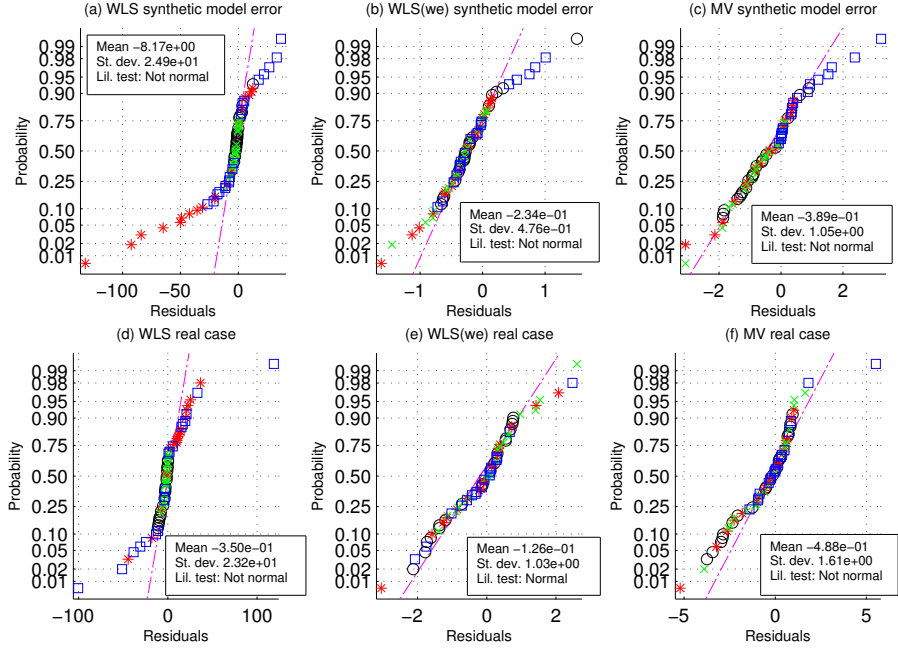


Figure 5.6.: Normal probability plots of the weighted residuals for the synthetic case with model error and the real case: plots (a) to (c) show the residual distributions for the synthetic case with model error, and plots (d) to (f) for the real case. Black circles represent calcium residuals, red asterisks iron residuals, green crosses TIC residuals and blue squares pH residuals.

measurement error variance (Fig. 5.7(d-g)). The weight adjustment resulted in a lower overestimation of the iron measurements from the WLS to the WLS(we) and MV (Fig. 5.5(b,f and j)), as opposed to the simulation with the correct parameter values where a severe overestimation of iron and pH was observed (with dissolved iron concentrations up to 180 mg L^{-1} and pH values up to 9.2). Therefore, estimating the weights from the data in presence of model bias did not allow the correct simulation of the concentration profiles.

The predictive quantile-quantile (QQ) plot shown in Fig. 5.8(a) indicates that the WLS method underestimates prediction uncertainty, with most of the measurements outside the prediction bounds (most of the computed quantiles close to 0 and 1). Consistent with the large estimated residual variances (Fig. 5.8),

the WLS(we) method provides the largest prediction bounds with all measurement contained in the prediction bounds and a typical S-shape of the QQ plot, indicating a severe overestimation of predictive uncertainty. For the MV methods (MV and MV(diagonal)), the fraction of measurements contained in the 95% prediction bounds approaches the theoretical value, with the computed quantiles closer to the 1:1 line. However, a reliable quantification of predictive uncertainty in the synthetic case with model error also requires the statistical description of the introduced model error, which was not included in our methods. For example, Reichert and Schuwirth [130] indicate that neglecting model bias in the description of the model output can lead to an incorrect estimation of predictive uncertainty. In our case, large uncertainty bounds were predicted for dissolved iron, with values up to 300000 mg L⁻¹ for the WLS(we) method and 8000 mg L⁻¹ for the MV method. Obviously, these values are not realistic and are caused by the presence of large model errors.

The presence of model error is correctly indicated by the large estimated *SEs* (Fig. 5.8(a)). The estimated σ_{MOj} variances are reported in Tab.B.3(b) of the appendix. Compared to cases without model error, larger positive values were obtained for all measurement sets and methods, with the largest values obtained for iron and TIC. Such information could be used as a guide towards improving the model structure, in this case indicating the need of including the omitted iron carbonate precipitation reaction which consumes dissolved iron and TIC.

5.4.2.2. Real case

Finally, we also test and compare the different methods on a real dataset. In this case, the measurements of a real column experiment were used to infer seven model parameters, as described in sec.5.3.2. In Fig. 5.9 the model fits for the different MAP parameter estimates are shown together with the corresponding RMSE values. The WLS method fits the pH measurements accurately because large weights were assigned to pH residuals and the pH measurements were sensitive to mineral precipitation and iron corrosion rates. In the WLS(we) method, the adjustment of the variances resulted in a redistribution of the weights with lower RMSE for calcium, iron and TIC measurements and a higher RMSE for pH. Similar results were obtained for the MV method. The simulated concentrations shown in Fig. 5.9 can be explained by analyzing the MAP parameter estimates reported in Tab.B.4(a) of the appendix. The calcium carbonate precipitation rate coefficient ($k_{eff,CC}$) was estimated about two orders of magnitude higher in the WLS(we) and MV methods compared

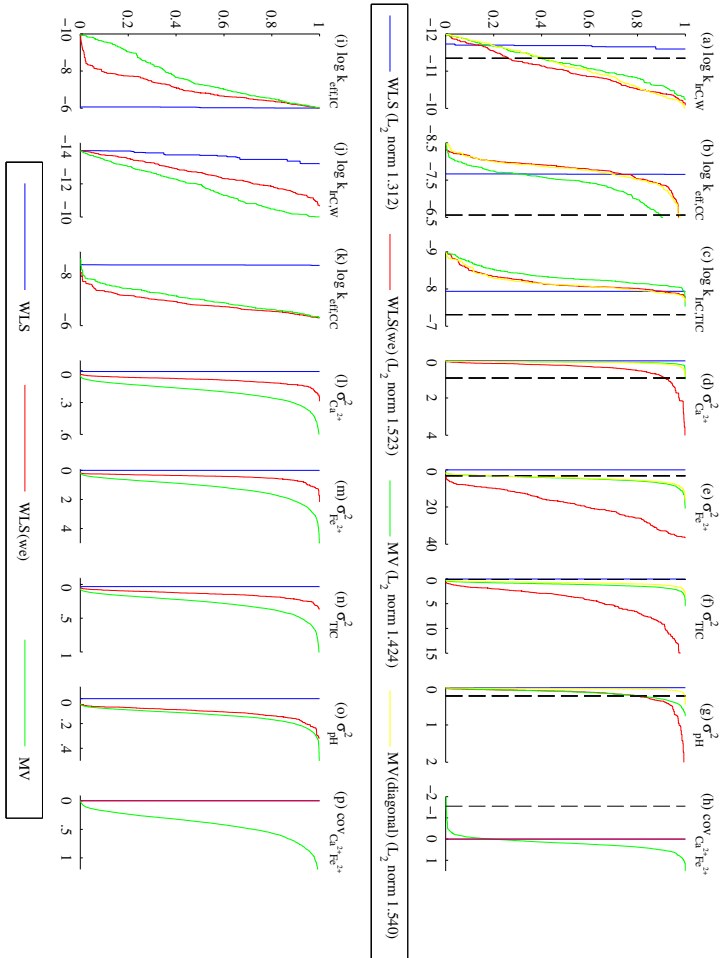


Figure 5.7.: Posterior cumulative distribution functions for model parameters and residual variances for the synthetic case with model error (figures (a) to (h)) and for the real case (figures (i) to (p)). The black dashed lines represent the correct parameter values and the variances estimated from the correct residuals. In the top legend box the L_2 norms of the bias parameter vectors are reported.

to the WLS method. This leads to more precipitation of calcium carbonate (aragonite), thereby removing carbonate from solution, reducing precipitation of iron hydroxy carbonate, and thus leaving more iron in solution.

Normality plots of the weighted residuals are shown in Fig. 5.6(d-f). As in the synthetic case with model error, the WLS residual distribution deviates the most from normality, with heavy tails consisting of pH residuals (at 5 PV and 0.15 m and 0.25 m from the influent). The normality assumption holds only for the WLS(we) method. This does not necessarily imply the absence of model error, which might still be present as indicated by a high *SE* value. Overall, the tails of the residual distribution in the real case deviate less from normality than in the synthetic case with model error, indicating a better simulation of measured concentrations (as it can be confirmed by comparing Fig. 5.9 to Fig. 5.5).

Posterior CDFs for $k_{eff,IC}$, $k_{IC,W}$, $k_{eff,CC}$ parameters are shown in Fig. 5.7(i-k). Parameter uncertainty increases considerably from the WLS method to the other methods, due to residuals affected by other errors than just measurement errors. Accounting for residual correlation in the MV method further increases parameter uncertainty (as indicated by the confidence intervals reported in Tab.B.4(a) of the appendix) and in this case also increases the estimated residual variances (Fig. 5.7(l-o)).

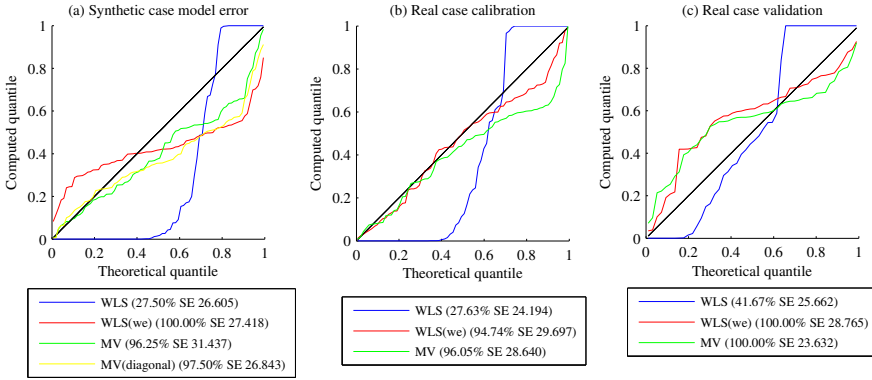


Figure 5.8.: Predictive QQ plots for the synthetic case with model error (a) and for the real case in calibration (b) and validation (c). In the legend boxes the percentage of measurements contained in the 95% prediction bounds and the standard error values are reported.

As for the synthetic case with model error, the WLS method underestimates prediction uncertainty in calibration (Fig. 5.8(b)). For the WLS(we) and MV methods, larger 95% prediction bounds were estimated in calibration, which contained almost the correct amount of measurements with a closer correspondence to the 1:1 line. Accuracy of the estimated prediction bounds crucially depends on the presence of model bias, since none of the proposed error models explicitly account for it. The large *SEs* estimated in calibration suggests the presence of model errors also in the real case.

Compared to the calibration step, lower RMSEs were generally obtained in validation (Fig. 5.10), probably because at early stages the geochemical conditions in the column were not fully stabilized (see for example the TIC measurements in Figure 10 at 5 PV at 0.03 m and 0.05 m). Moreover, similar concentration profiles were simulated by the WLS(we) and MV methods, with the MV method providing the best fit for iron and pH measurements and slightly worse fit for calcium and TIC. The better fit obtained in validation is also confirmed by the lower *SEs* for the WLS(we) and MV methods (Fig. 5.8(c)), with the MV method providing the lowest value. For the WLS method, the underestimation of predictive uncertainty is less severe compared to calibration whereas the WLS(we) and MV methods overestimate uncertainty. The σ_{MO_j} standard deviations reported in Tab. B.4(b) of the appendix indicate that iron residuals are the most affected by model errors, suggesting that model improvements should be sought in the description of iron corrosion and precipitation processes.

In all studied cases the MV method was applied considering correlation between different species, since a strong interplay between different ions was expected (for example, a pH increases promote mineral precipitation and causes a decrease of dissolved TIC and calcium). However, correlations between measurements at different depths or at different pore volumes can also be evaluated. Using the residuals at the MAP parameter estimates of the MV method (Fig. 5.9(i-l)) the maximum off-diagonal correlation coefficient between different species was 0.86 (iron and TIC couple), whereas when the residuals were grouped by depth the maximum off-diagonal correlation coefficient was 0.84, for the measurements taken at 0.05 and 0.10 m from the column influent. When the residuals were grouped by pore volumes the maximum significant correlation coefficient was 0.95 (between pore volume 110 and 160). Based on these results, the MV method with residuals grouped by pore volumes could also be considered and the obtained results compared with those obtained in this study.

5.4 Results

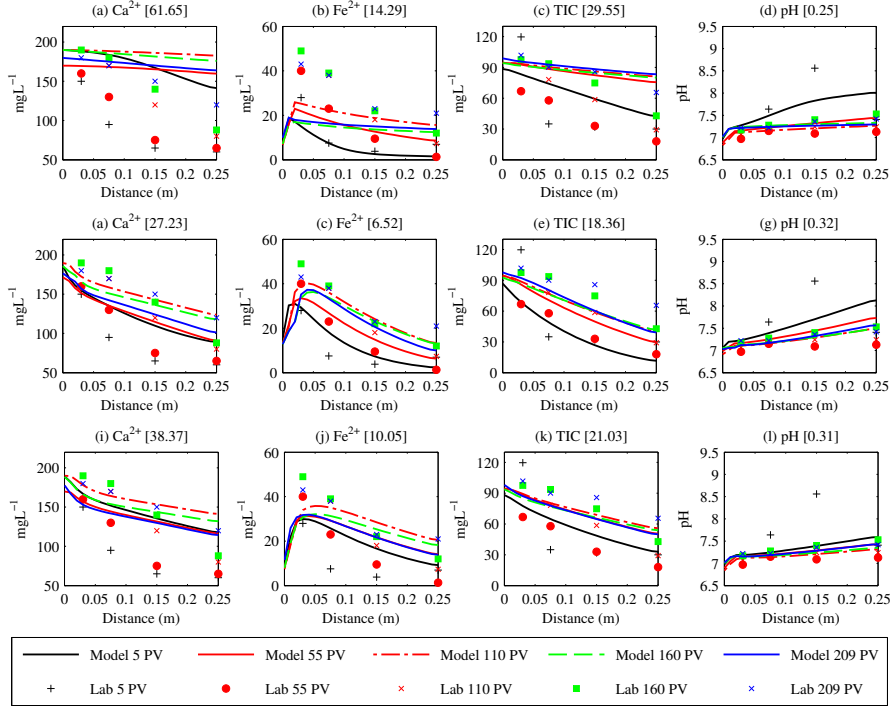


Figure 5.9.: Calibration results for the real case: WLS (top row), WLS(we) (central row) and MV method (bottom row). In square brackets the RMSEs are provided.

5.4.3. Computational efficiency

In Tab. 5.2, the CPU hours required to achieve convergence of DREAM-ZS (with convergence indicated by the Gelman-Rubin R statistic below 1.2) are reported together with the number of model runs. Except for the synthetic case with high correlation, the MV method outperforms the WLS(we) method, with a speed up factor ranging from 1.2 (for the synthetic case with low correlation) to 3.8 (for the synthetic case with model error), thanks to analytical marginalization of the covariance matrix before running the MCMC algorithms. However, for the synthetic case with high correlation, chain convergence was reached faster in the WLS(we) method. In a hypothetical case where all elements of the covariance matrix are explicitly estimated from the

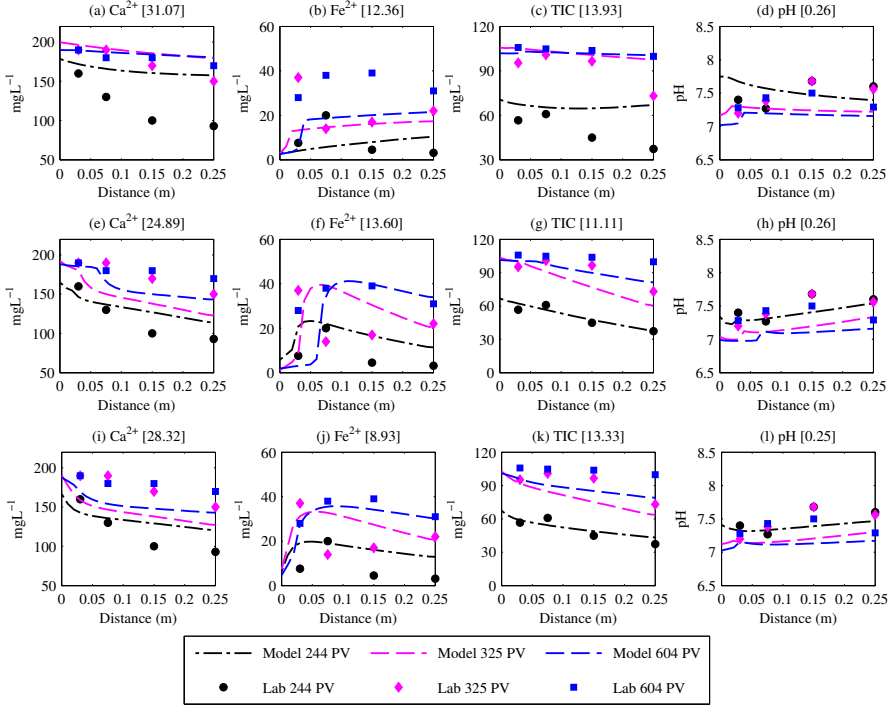


Figure 5.10.: Validation results for the real case: WLS (top), WLS(we) (centre) and MV method (bottom). In square brackets the RMSEs are provided.

data (yielding a total of 14 parameters in our synthetic case), it is likely that the WLS(we) method will become more computational demanding than the MV method. For the WLS method the convergence was found to be slow and reached only at the specified maximum number of model runs (except for the synthetic case with high correlation).

5.5. Conclusions

A multivariate Bayesian methodology was presented for estimating reactive transport model parameters from multi-species data. It was shown that existing approaches, specifically the weighted least-squares method (WLS) and the weighted least-squares method with weight estimation (WLS(we)), fit within

Table 5.2.: CPU hours required to achieve convergence of DREAM-ZS (on a single core of an Intel Core 2 Duo CPU E8400 at 3.00 GHz). In brackets the number of model runs is reported.

Method	Synthetic case low correlation	Synthetic case high correlation	Synthetic case with model error	Real case
WLS	32.1 (8000)	19.8 (4788)	21.9 (6000)	154.2 (14000)
WLS(we)	18.1 (4788)	24.7 (6384)	16.5 (4893)	111.0 (12078)
MV	15.0 (3591)	36.5 (7182)	4.3 (1200)	55.7 (6291)

this framework and are obtained by making specific prior assumptions about the residual covariance matrix. The Bayesian perspective allows for generalizations to cases where residual correlation is important and for more efficient parameter estimation by analytically integrating out the (co)variances (or weights) from the posterior. One such method, referred to here as the MV method (for MultiVariate), incorporates these two benefits. For the cases studied here, analytically integrating out the variances from the posterior leads to estimation of parameter and predictive uncertainty that is more efficient, in terms of CPU time, compared to the alternative of jointly estimating model parameters and the full covariance matrix.

Using numerical experiments, the MV method was compared to WLS and WLS(we) to evaluate the effect of residual correlation (between species) on parameter and predictive uncertainty. It was found that residual correlation does not have a strong effect on predictive uncertainty, and that parameter estimates and their posteriors were only affected at relatively high levels of residual correlation between species (0.9).

The results also show that all methods can be considered robust only for zero-mean residual distributions, as they all rely on this a priori assumption. In the presence of model structural errors, for example due to an unaccounted for or improperly described reaction, strongly biased residual distributions and parameters values were obtained for all methods, as confirmed by large misfits between simulated and observed data and by estimated values for the parameters that reach prior bounds. The complete reaction network is rarely known in real systems and model structural errors are likely to be present. In such cases, more effort should be directed towards the identification of missing processes, or one should try to capture model misfits by introducing statistical bias in the likelihood function [130]. Close scrutiny of the model residuals

should then be used as a guide to model improvement, including calculating residual variances not attributable to measurement errors to get an idea of where the largest model errors occur.

A limitation of the multivariate (MV) method used here is that it cannot be implemented using a full covariance matrix with a dimension equal to the number of all measurements. Instead, residuals must be grouped such that the number of groups (m) does not exceed the number of measurements per group (n). In our cases, residuals were grouped by species, and thus only correlations between species were accounted for. Other groupings, corresponding to different prior assumptions about the correlation structure, are possible as well, e.g. grouping by space and/or time, as long as m is less or equal to n .

6. Site model ¹

Abstract: In this study a numerical groundwater reactive transport model of a shallow groundwater aquifer contaminated with volatile organic compounds is developed. In addition to advective-dispersive transport, the model includes contaminant release from source areas, natural attenuation, abiotic degradation by a permeable reactive barrier at the site, and dilution by infiltrating rain. Aquifer heterogeneity is parameterized using pilot points for hydraulic conductivity, specific yield and groundwater recharge. A methodology is developed and applied to estimate the large number of parameters from the limited data at the field site (groundwater levels, groundwater concentrations of multiple chemical species, point-scale measurements of soil hydraulic conductivity, and lab-scale derived information on chemical and biochemical reactions). The proposed methodology relies on pilot point parameterization of hydraulic parameters and groundwater recharge, a regularization procedure to reconcile the large number of spatially distributed model parameters with the limited field data, a step-wise approach for integrating the different data sets into the model, and high performance computing. The methodology was proven to be effective in reproducing multiple contaminant plumes and in reducing the prior parameter uncertainty of hydraulic conductivity and groundwater recharge. Our results further indicate that contaminant transport predictions are strongly affected by the choice of the groundwater recharge model and flow parameters should be identified using both head and concentration measurements.

6.1. Introduction

Subsurface reactive transport models are increasingly applied to real contaminant sites. However, major challenges are present when applying such models to reproduce the observed concentration measurements and attempt to use

¹Based on: Carniato L., Schoups G., van De Giesen N. and Seuntjens P., Bastiaens L., and Sapion H. (2014). Highly parameterized inversion of groundwater reactive transport for a complex field site (under review in *Journal of Contaminant Hydrology*).

them as predictive tools. One of these challenges is the correct conceptualization and description of the subsurface processes using data from lab experiment or field measurements [48, 108]. Usually a limited number of field data are available to conceptualize complex transport models. Moreover, when reactions are also modelled the computational times increase dramatically. In practical terms, a simulation of solute transport over a period of years to decades using a fine discretization for a large region may require hours or days to complete on modern personal computers [82]. For these reasons and to avoid complexity not justified by the available data, simplifying assumptions are often made about the distribution of aquifer parameters, groundwater flow and transport dynamics, degradation rates and parameter inference procedures [56].

In numerous studies homogeneous aquifer properties (within zones, layers or over the entire model domain) are used [151, 74, 30]. Disadvantages of simple parameterization schemes are that the real response of the physical system might not be captured, introducing model structural noise [46] and hampering the ability to understand the natural system through model calibration [61]. Only a limited number of field-scale reactive transport studies directly inferred heterogeneity from available data [44, 147], whereas abundant literature is present for flow only problems [52, 89, 168]. Tonkin and Doherty [148] applied the pilot point method [35] to infer spatially variable flow and transport parameters (hydraulic conductivity, recharge and porosity) at the Hampton Bays Site, New York. In their approach a large number of model parameters were defined and the computational burden limited through the definition of “super parameters”. The method was used to reproduce the observed MTBE (Methyl tert-butyl ether) concentrations at the site, effectively demonstrating that a better reproduction of the observations was obtained when aquifer heterogeneity is accounted for. Kowalsky et al. [86] studied the effect of different parameterizations of the hydraulic conductivity field (e.g. pilot point number and locations) in a synthetic tracer experiment, indicating the need for real applications where various data types are used in parameter inference. Fienen et al. [44] used different data types (head, oxygen and tritium isotopic measurements) with a geostatistical method [59, 77] to infer the hydraulic conductivities of an isthmus comprised in two lakes in the Trout Lake watershed, northern Wisconsin, United States. They concluded that the flexible parameterization offered by the geostatistical method and the simultaneous use of multiple data sources improves the flow path delineation in the isthmus zone. In all cases, hundreds to thousands model parameters were used, combined with parameter regularization [19, 167] to include prior parameter

knowledge. Pilot point and geostatistical methods allow the optimal level of complexity to be inferred directly from the data, provided that the model is correctly conceptualized and the correct assumption about prior parameter knowledge are made.

Besides heterogeneity in hydraulic conductivity, spatial and temporal variation of groundwater recharge is important in modelling contaminant transport in shallow unconfined aquifers, due to the possible dilution effect by infiltrating rainwater. For example, Kowalsky et al. [87] incorporated geochemical and time-lapse resistivity data in a hydrogeochemical model of the Oak Ridge Integrated Field Research Challenge (IFRC) site in Tennessee, United States. In their model it was assumed that groundwater recharge determines nitrate concentrations in the unconfined aquifer, in particular in the shallow part. Their study also points to the need of spatially extensive dataset (concentration but also geophysical measurements) to monitor recharge related concentration variations. Şengör and Ünlü [136] developed a groundwater transport model to determine the extension of acrylonitrile contamination at a spill site in Turkey. In their long term simulations (2001 to 2011) the size of the plume in the high permeability zone shrank significantly due to the dilution by groundwater recharge. In these types of problems, flow and transport in the unsaturated zone should ideally also be modelled, further increasing the computational demand [166]. A simpler approach is to assume groundwater recharge equal to the average net balance at the water table [30] or a spatially variable fraction of this balance [52, 148]. Assefa and Woodbury [3] estimated groundwater recharge in the Okanagan Basin basin (Canada) calibrating a HYDRUS-1D model [139] on soil moisture data and extrapolating the results over the entire basin (245 km²) based on soil characteristics. Their results show a remarkable spatial variability of the average recharge value, varying from 12 to 170 mm y⁻¹. On a larger scale (3600 km²), Hayley et al. [52] estimate groundwater recharge and horizontal hydraulic conductivity from head and flow data, also obtaining variable groundwater recharge rates.

Another important issue concerns the calibration methodology. In the majority of the reactive transport field studies flow parameters are calibrated first against head and flow data to provide the flow field to be used in reactive transport calculations [30, 125, 47]. Afterwards, the remaining reactive transport parameters are calibrated on concentration data, using independent estimations (e.g. lab experiments or field observations) and refining the initial guesses by manual calibration. The limitation of this approach is that the information about flow parameters contained in the concentration measurements is not used. Many synthetic and lab-scale studies demonstrate that

concentration data are more informative of the local subsurface heterogeneity than groundwater head data [124, 158].

In this study, we explicitly tackle the challenges discussed above, by means of an application of existing methodologies to a complex real-world contaminated site. Our study goes beyond existing applications which typically are restricted to single reactive species or synthetic aquifers. Applications of highly parameterized inversion of multi-component reactive transport models are not widely reported, and our study illustrates both the benefits and limitations of such an approach for real-world applications. Specifically, we show that parameter estimation in shallow heterogeneous contaminated aquifers strongly influenced by recharge dynamics is possible by a combination of regularization, high-performance computing, and a step-wise approach of parameter refinement using laboratory and field data. The methodology is evaluated in terms of reproducing the observations, realism of the estimated parameters and reduction of the prior parameter uncertainty.

6.2. Groundwater flow model

A conceptual picture of relevant flow processes at the site is shown in Fig. 6.1. Groundwater flow in confined aquifers is described with the following equation:

$$\frac{\partial}{\partial x} \left(K_{xx} \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_{yy} \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_{zz} \frac{\partial h}{\partial z} \right) + W = S_s \frac{\partial h}{\partial t} \quad (6.1)$$

where K_{xx} , K_{yy} and K_{zz} are the saturated hydraulic conductivity values in x , y , and z directions (m d^{-1}); h is the total hydraulic head (m); W is the source/sink term representing recharge and evaporation (d^{-1}) and S_s is the specific storage (m^{-1}) for confined aquifers. In unconfined aquifers Eq. 6.1 is integrated over the vertical direction and the specific storage is replaced by the aquifer specific yield S_y (-). Groundwater flow was simulated using MODFLOW-2000 [50], within the MODFLOW domain indicated in Fig. 6.3. The model is oriented along the main flow direction and is composed of 2 layers, 133 rows and 125 columns. The bottom layer coincides with the screened interval of the piezometers. The shallow layer was modelled as unconfined (convertible), while the bottom one was modelled as confined, since observed groundwater levels never drop below the top of the bottom layer. In this case the specific storage coefficient was substituted with the specific yield value

divided by the bottom layer thickness. This approach produced almost identical results of treating both layers as convertible but avoided the drying and re-wetting for some cells (which might cause the simulation failure).

Cauchy type boundaries (general heads) were used at the north (upstream) and south (downstream) sides of the model domain, while no-flow boundary conditions were specified for the west and east boundaries. Three drains were included in the pasture area of the model, as observed in the field (Fig. 6.3). No flow is assumed to occur below the bottom layer, due to presence of the low permeable clay (Boom clay formation, Fig. 2.4). Special attention was paid to modeling aquifer heterogeneity and spatial variations in recharge entering the aquifer in the pasture area, as discussed in the following subsection.

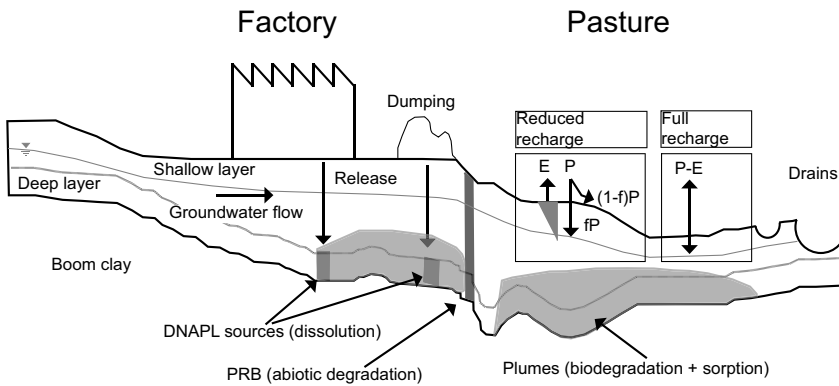


Figure 6.1.: Conceptual model of the site.

In Tab. 6.1 the time discretizations used in the inversion tests (chapter 7) are reported. For discretizations 1, 2 and 4 a steady state stress period is firstly simulated, followed by weekly stress periods. The choice of each time discretization depends on the simulation type: when contaminant transport was simulated starting from the 2003 interpolated concentration fields, discretization 2 was used. In the case that the simulation aims to reconstruct the contamination history from the first spillage events up to 2012, discretization 1 was used. Discretization 3 was used for the geochemical simulation with the barrier installed and discretization 4 for the geochemical simulation without the barrier.

Table 6.1.: Time discretizations used in the model (SS = steady state, WT = weekly transient, MT = monthly transient).

Discretiza- tion	Period	Duration (d)	Stress periods	Stress period types
1	1957-2012	20237	649	0 - 15701 SS, 15701 - 20237 WT
2	2003-2012	3302	471	0-14 SS, 14 - 3302 WT
3	2005-2035	10687	618	0- 2429 WT, 2429 - 10687 MT
4	1957-2035	28488	919	0 - 15701 SS, 15701 - 20230 WT, 20230 - 28488 MT

6.2.1. Aquifer properties heterogeneity

Spatial variation in hydraulic conductivity, specific yield and recharge was modeled using the pilot point method [22, 128]. In the pilot point method the unknown parameters are estimated for a set of points collocated within the domain. The corresponding spatially distributed parameter field was obtained kriging the pilot points values over the entire model domain. As shown in Fig. 6.2(a), specific yield pilot points were placed on a regular grid (22.5 m apart near the barrier and further apart in areas with fewer groundwater and concentration measurements), resulting in 127 pilot points. Infiltration fraction pilot points were placed only in the permeable pasture area where measurements were collected (Fig. 6.2(b)), resulting in 43 infiltration fraction pilot points. Hydraulic conductivity parameterization was similar to the one used for specific yield (Fig. 6.2(c)), with a local refinement along the main plume direction (Fig. 6.2(d)), yielding 178 pilot points. Spatial correlation of the heterogeneous aquifer properties were assumed to be described by an exponential variogram:

$$\gamma(h) = b(1 - \exp(-ah)) \quad (6.2)$$

where d is spatial distance, b is the sill and a is the range. A range of 150 m was adopted for specific yield and hydraulic conductivity fields and 105 m for infiltration fraction pilot points, corresponding to the maximum pilot point to grid cell distance [40]. The variogram range was kept fixed in the inversion tests as the inversion results are expected to be more influenced by locations and number of pilot points used in the inversion [86].

The sill for log-transformed specific yield was estimated as 0.137, which corresponds to one fourth of the variation range squared, assuming values for

specific yield from 0.01 to 0.3 [69] and a log normal distribution of the specific yields values. Similarly, the sill for log-transformed infiltration fractions was estimated at 0.063 (assuming a variation range for recharge from 0.1 to 1).

The sill for log-transformed horizontal hydraulic conductivity was estimated as 0.254, based on the variance of values for log-transformed hydraulic conductivity determined from slug tests. Finally, in the absence of direct measurements, vertical hydraulic conductivity was assumed to be 10 times smaller than the corresponding horizontal hydraulic conductivity [146]. Furthermore, vertical variation in aquifer properties was neglected, and thus the same values were adopted for both model layers. This latter assumption is supported by the similar estimation of the hydraulic conductivity from the slug test data at piezometers 9 and 10 (Fig. 2.4), which are screened in the deep and shallow layer respectively. The choice of the pilot point locations was not based on sensitivity approaches that optimize pilot point placement on the basis of measurement locations [71], but many pilot point were used in the barrier and plume area and afterwards singular values decomposition was used to further regularize the problem.

6.2.2. Recharge

The pasture area of the shallow unconfined aquifer exhibits dynamic variations in groundwater levels caused by the complex interplay between precipitation, infiltration, runoff, evaporation, and recharge. To quantify recharge and/or evaporation from the, mostly shallow, water tables at the site, two different models were considered (Fig. 6.1).

In the first model, referred to as the “full recharge model”, net “recharge” (positive or negative) to/from the water table was estimated as the difference between weekly precipitation and potential evaporation data from a nearby meteorological station (Westdorpe, Netherlands). In this model, runoff is assumed to occur when the water table reaches the land surface implemented here with the MODFLOW drain package by specifying virtual drains at the land surface with a large drain conductance value. The full recharge model might be justified by the shallow depth of the water table (in most circumstances measured to be no lower than 1.5 meters from the ground surface), but might introduce too much water in the aquifer which can be drained out from the system only by high hydraulic conductivity values.

In the second recharge model, referred to as the “reduced recharge model”, infiltration is first estimated as a (temporally constant, but spatially variable) fraction of precipitation, with the remaining water assumed to be lost as runoff.

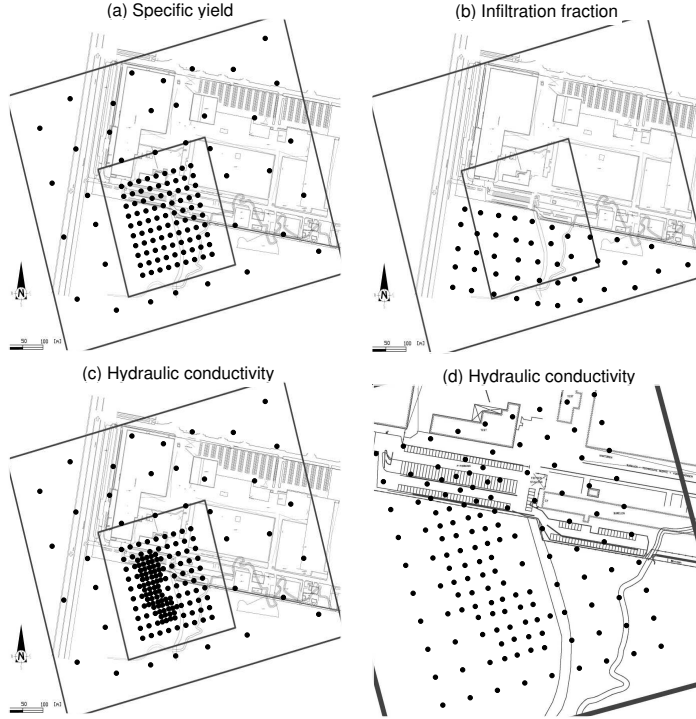


Figure 6.2.: Pilot points for (a) specific yield, (b) infiltration fraction, and (c) hydraulic conductivity. In Figure (d) the pilot point refinement along the main plume direction is shown.

Infiltrated water is then added as recharge to the water table, from where it can be lost again by evaporation. Hence, recharge and evaporation are treated separately, and are implemented using the MODFLOW recharge and evapotranspiration packages, respectively. The rate of evaporation from the groundwater table is assumed to vary linearly as a function of water table depth, with a maximum equal to potential evaporation when the water table is at the land surface, and zero when the water table is more than 2 meters below land surface (this extinction depth was estimated from an empirical relation and the soil properties reported in the study of Shah et al. [137]). The fraction of infiltrating precipitation was assumed to be spatially variable and parameterized with pilot points (Fig. 6.2(b)), using a similar approach as for

hydraulic conductivity and specific yield. Temporal variation in infiltrating fractions was not taken into account.

6.3. Reactive transport model

Three-dimensional advective-dispersive reactive transport was simulated using the RT3D code [27] for an aerial extent contained within the MODFLOW domain (Fig. 6.3). Transport was simulated with the standard advective-dispersive-reaction equation:

$$\frac{\partial C_k}{\partial t} = \frac{\partial}{\partial x_i} \left(D_{ij} \frac{\partial C_k}{\partial x_j} \right) - \frac{\partial}{\partial x_i} \left(\frac{q_i}{por} C_k \right) + \frac{W}{por} C_{Wk} + r_k \quad k = 1, \dots, m_c \quad (6.3)$$

$$\frac{d\tilde{C}_k}{dt} = \tilde{r}_k \quad k = 1, \dots, (n_c - m_c) \quad (6.4)$$

where i and j represent the spatial direction (x , y or z) and indexes repetition implies summation, n_c is the total number of specie, m_c is the total number of aqueous-phase (mobile) species (thus, m_c minus n_c is the total number of immobile species), C_k is the aqueous-phase concentration of the k^{th} species (mol L^{-1}), $\tilde{C}_k$ is the solid-phase concentration of the k^{th} chemical specie (mol mol^{-1} , contaminant mass per unit mass of porous media), D_{ij} is the hydrodynamic dispersion coefficient ($\text{m}^2 \text{d}^{-1}$), q_i is the Darcy flow in the i direction (m d^{-1}), por is the soil porosity, W is the volumetric flux of water per unit volume of aquifer representing sources and sinks (for example groundwater recharge, d^{-1}), C_{Wk} is the concentration of source/sink (mol L^{-1}), r_k represents the rate of all reactions that occur in the aqueous-phase ($\text{mol L}^{-1} \text{d}^{-1}$) and $\tilde{r}_k$ represents the rate of all reactions that occur in the soil-phase ($\text{mol mol}^{-1} \text{d}^{-1}$). The Darcy flow q_i is provided to RT3D through an external flow-link file saved by MODFLOW, therefore assuming the transport processes not affecting the flow field. The reaction term r_k which deplete or produce the contaminant k , can be deduced from Fig. 6.4:

$$r_k = r_{k-SOU} - r_{k-DEG} + \sum_i^p r_{i-DEG} - r_{k-PRB} \quad (6.5)$$

where r_{k-SOU} is the release rate of the compound k from the sources (which depends on the source location), r_{k-DEG} is the biodegradation rate of compound k , p is the number of compounds that can degrade to k and r_{k-PRB} is the degradation rate of k into the reactive barrier. An example for TCE is reported below:

$$r_{TCE} = r_{TCE-SOU,i} - r_{TCE-DEG} + r_{PCE-DEG} - r_{TCE-PRB} \quad (6.6)$$

since all concentrations were expressed in mol L⁻¹ and biodegradation was assumed to occur sequentially (Fig. 6.4), stoichiometric yield values were not necessary. Contaminant degradation by zero-valent iron is assumed to mainly occur by elimination and much less by hydrogenolysis, with no formation of intermediates. RT3D uses a operator splitting numerical strategy, which allows the specification of general formulas for r_k and $\tilde{r}_k$. The advection term of Eq.6.3 is solved with an explicit in time total variation diminishing scheme (TVD), which has stability constraints that leads to long simulation times. Only using the TVD scheme in the field scale geochemical model (PHT3D, sec.7.7), the column contaminant profiles were correctly reproduced (chapter 3). Other advection schemes (such as upstream in space finite differences) that could reduce the computation time by taking larger time steps, showed instabilities or introduced too much numerical dispersion.

6.3.1. Contaminant dissolution from the source areas

Parent contaminants (PCE, TCE and TCA) are known to form dense non aqueous phase liquid phases (DNAPL), which tend to migrate downwards due to the gravity force and slowly dilute in groundwater. It is assumed that pure phase contaminants were released on the ground surface and have migrated and accumulated in the bottom layer (Fig. 6.1). The pure phase DNAPL pools in the bottom layer dilute in groundwater forming the observed contaminant plumes. Hence, the change in time of DNAPL phase mass for each parent compound is described by the following expression:

$$\tilde{r}_{k-SOU,i} = \min[-k_{la}(C_{k,i}^* - C_k), 0] + k_{k-SOU,i} \quad (6.7)$$

where k_{la} is a mass transfer coefficient, $C_{k,i}^*$ represents the equilibrium aqueous concentration in groundwater for parent contaminant k (TCA, PCE or TCE)

at the specific source location i (1, 2 or 3, as shown in Fig. 2.2), C_k is parent contaminant concentration in groundwater and $k_{k-SOU,i}$ is the release rate of the parent contaminant k during the release period at source location i (Tab. 2.1). The equilibrium contaminant solubility of each parent contaminant compound was described by the Raoult's law [135]:

$$C_k^* = C_{k-SOL} \gamma_k \frac{\tilde{C}_k}{\sum_{i=1}^{N_p} \tilde{C}_{k,i}} \quad (6.8)$$

where C_{k-SOL} is the solubility of compound k in groundwater, γ_k is the activity coefficient (assumed to be unity) and N_p is the number of parent contaminants in the specific source area i (based on Tab. 2.1 sources 1 and 3 were assumed to contain PCE and TCE, whereas source 2 only TCA). Contaminant solubility for PCE, TCE and TCA were $8.37\text{e-}3 \text{ mol L}^{-1}$, $1.21\text{e-}3 \text{ mol L}^{-1}$ and $9.74\text{e-}3 \text{ mol L}^{-1}$ [119]. For the mobile domain, the release from the sources equals only to the first part of Eq. 6.7, but with opposite sign:

$$r_{k-SOU,i} = \max [k_{la}(C_{k,i}^* - C_k), 0] \quad (6.9)$$

since no quantitative information was available about contaminant disposal and release rate through time, $k_{k-SOU,i}$ for each parent compound and source area i were estimated from groundwater concentration data, together with the mass transfer coefficient k_{la} . Eq. 6.7 provided more realistic simulations of the contaminant plumes compared to assign constant concentration boundary at the sources and allows the simulation of DNAPL dissolution in the long term scenarios.

6.3.2. Contaminant sorption to aquifer material

Sorption was modelled as a reversible equilibrium process assuming a linear isotherm:

$$\tilde{C}_k = K_{d-k} C_k \quad (6.10)$$

the distribution coefficients K_{d-k} (L kg^{-1}) were calculated as:

$$K_{d-k} = K_{oc-k} f_{oc} \quad (6.11)$$

where K_{oc-k} is the contaminant partition coefficient (L kg^{-1} , [160]) and f_{oc} the carbon content measured in soil samples from the site. Measurements of the organic carbon content report an average f_{oc} of 0.3% for the deep layer (piezometer 5) and 1.5% for the shallow layer (piezometer 26, Fig. 2.4). The retardation factor R_k under equilibrium conditions can be calculated as:

$$R_k = 1 + \frac{\rho_b}{por} K_{d-k} \quad (6.12)$$

Here, ρ_b is the dry bulk density, set equal to 1628 kg L^{-1} (determined from angled core samples) and por is the effective porosity, which was estimated from groundwater concentration data. A first estimation of the retardation factors can be obtained assuming por equal to 0.1 (a plausible value for sandy clay soils [69]). The resulting retardation factors are reported in Tab. 6.2 and should be interpreted as initial estimations.

Table 6.2.: Partition coefficients and calculated retardation rates.

Component	K_{oc}^a (L kg^{-1})	$R_{shallow}$ (-)	R_{deep} (-)
PCE	269.2	14.15	66.73
TCE	18.2	1.89	5.44
cis-DCE	24	2.17	6.86
VC	1.06	1.06	1.28
TCA	14.76	14.76	69.82
DCA	3.81	3.81	15.05
Total contamination	50 ^b	3.44	13.21

^a from RCBA database [160]. ^b weighted average of the individual K_{oc} coefficients on contaminant concentrations.

6.3.3. Abiotic contaminant degradation within the reactive barrier

As groundwater flows through the permeable reactive barrier, dissolved contaminants come in contact with the reactive material (zero-valent iron), and are converted to non-toxic substances (such as ethene) via chemical (abiotic) reduction. Individual kinetic rates of abiotic contaminant degradation in the

barrier are described a mixed-order rate model (Eq.3.1), with parameters determined from the column experiment. As shown in Fig. 6.3, the barrier is not oriented perpendicularly to the RT3D model domain: in the model it is represented as a series of reactive connected 2.5 m by 2.5 m grid cells, resulting in a stair-type representation of the 0.3 m wide barrier.

This coarse representation was necessary to keep the forward model runs below 40 minutes and make feasible the inference of spatially distributed parameters using concentration data. This assumption is also justified by the small influence of the barrier installation on the measured concentrations. Indeed only 9 piezometers showed an abrupt decline of the concentrations after the barrier installation, indicating that the groundwater treated by the PRB did not traveled far after October 2005.

To account for the coarse barrier discretization, the degradation rates determined from the column experiment k_{k-PRB} were divided by 8.3, the ratio between the block size (2.5 m) and the effective barrier thickness (0.3 m). Whereas in principle mineral precipitation within the barrier may decrease its reactivity over time, this was not accounted for in the current model, an assumption supported by reactivity tests performed in 2012 on reactive zero-valent iron material extracted from the in-situ barrier after 7 years of operation.

After parameter inference, geochemical scenarios were simulated, also accounting for mineral precipitation and its effect on iron reactivity. Compared to the reactive transport domain used for the inversions, the geochemical domain is rotated 24.9 degrees clockwise (green square in Fig. 6.3), to be able to represent the barrier correctly, with rows of the PHT3D domain parallel to the barrier. In this case the 0.3 cm thick barrier was discretized using 0.06 m wide blocks, employing directly the geochemical model calibrated in chapter 3.

6.3.4. Treatment of contaminant biodegradation and substrate consumption

Following Johnson and Truex [70], the degradation of source contaminants (PCE, TCE and TCA) was modelled as a sequence of irreversible kinetic degradation reactions (Fig. 6.4). The network does not include trans-DCE, 1,1,2-TCA and 1,2-DCA contaminants, as these were measured at low concentrations. The biodegradation reactions were assumed to occur under anaerobic conditions with production of intermediate contaminants (TCE, cis-DCE, VC and DCA). The dashed arrow indicate dehydrochlorination, where one chloride atom and one proton are eliminated. The reaction is abiotic but studies indicate that the reaction can be enhanced/catalyzed by bacteria [111] and/or

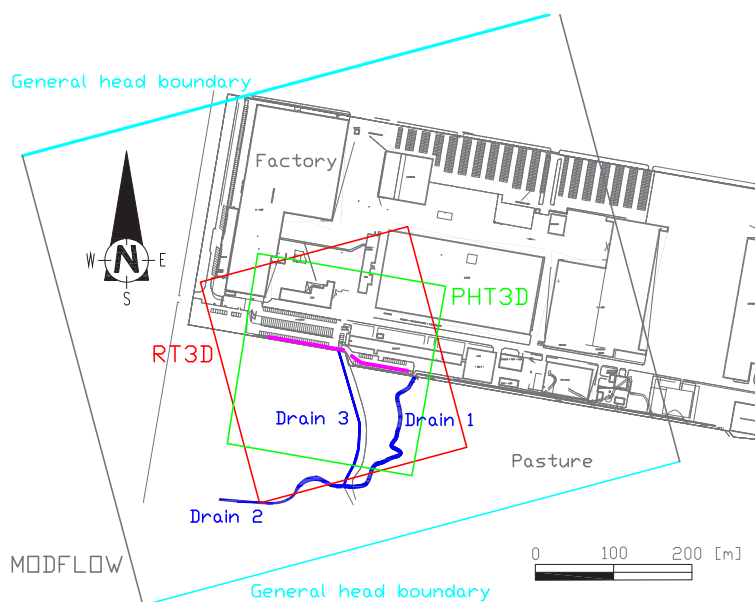


Figure 6.3.: Model domains and flow boundary conditions. In the external flow domain (MODFLOW), general head boundary conditions were used at the north and south boundaries, while no flow condition were used at the east and west boundaries. The three drain conditions are indicated by the blue lines. The RT3D domain is indicated by the red square within the flow domain, oriented along the main flow direction. Geochemical simulations (PHT3D) were performed in the green square. The PRB is indicated by the magenta line.

minerals (e.g., clay [23]). Usually dehydrochlorination proceeds at slow rates, as indicated in Tab. 6.3.

In Fig. 6.4 the solid arrows indicate reductive dechlorination reactions, where the contaminants are used as electron acceptor by microorganism. Electron donor is usually hydrogen or acetate, produced by organic matter fermentation [53, 37]. As such, dechlorination both affects and depends on the availability of fermentable substrate, with low degradation rates and substrate levels in the interior of the contamination plume, and high degradation rates and substrate levels at the fringes of the plume, where the plume is in contact with pristine aquifer material (sec. 2.2). However, dechlorinators have to compete

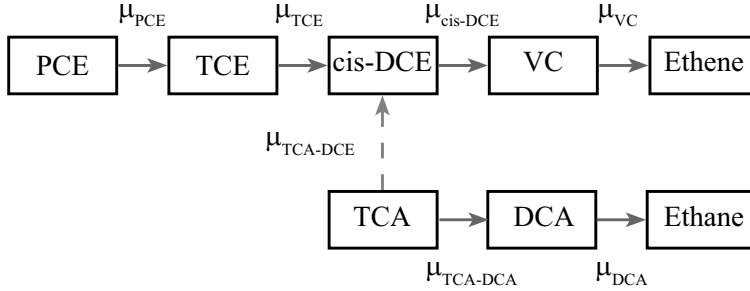


Figure 6.4.: Biodegradation network. Solid arrows indicate reductive dechlorination (biotic) and the dashed arrow dehydrochlorination (abiotic).

Table 6.3.: First order degradation coefficients for the degradation network reported in Fig. 6.4 [142].

Rate constant	Literature range (d ⁻¹)
μ_{PCE}	0-0.08
μ_{TCE}	0-0.023
$\mu_{cis-DCE}$	0.001-0.006
μ_{VC}	0-0.007
$\mu_{TCA-DCE}^a$	0.00076-0.0017
$\mu_{TCA-DCA}$	0-0.125
$\mu_{DCA-Ethe}$	0-0.011

^a based on the studies of Jeffers [67], Mabey and Mill [103] and Vogel and McCarty [153].

for the reducing equivalents against micro-organisms that employ other terminal electron accepting processes (Tab. 6.4). The amount of substrate required to complete the dechlorination of one mole of contaminant depends on composition of organic matter, which is difficult to estimate and can vary widely from site to site. The hypothetical molecule $C_5H_7O_2N$ can be considered representative of the organic carbon present in real sites (assuming that organic carbon is mainly composed by biomass). One mole of this molecule can yield 20 electron equivalents when it is completely oxidized to CO_2 (or 10 H_2 equivalents):

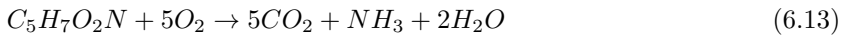


Table 6.4.: Biological processes using hydrogen

Biological process	Reaction rate
Dechlorination	$PCE + H_2 \rightarrow TCE + H^+ + Cl^-$
	$TCE + H_2 \rightarrow cis - DCE + H^+ + Cl^-$
	$cis - DCE + H_2 \rightarrow VC + H^+ + Cl^-$
	$VC + H_2 \rightarrow Ethe + H^+ + Cl^-$
Competing TEAP's	$0.25CO_2 + H_2 \rightarrow 0.25CH_4 + 0.5H_2O$
	$SO_4^{2-} + 4H_2 \rightarrow H_2S + 2OH^- + 2H_2O$
	$2FeOOH + H_2 \rightarrow 2Fe^{2+} + 4OH^-$
Substrate fermentation	Butyrate + $2H_2O \rightarrow 2Acetate + H^+ + 2H_2$
	Lactate + $2H_2O \rightarrow Acetate + HCO_3^- + H^+ + 2H_2$

The sparse field data indicate that only 1% of this potential can be actually exploited. And the studies about the dechlorination reaction in mixed microbial communities indicate that of this actual amount only 1 to 10% can be channeled to dechlorinators (due to competitive processes with other , TEAPs). Therefore, between 100 to 1000 mol of $C_5H_7O_2N$ are required to dechlorinate one mole of contaminant to the next reaction step in the network (e.g. from PCE to TCE). Bauer et al. [8] measured an electron donating capacity from 0.07 to 1.5 meeq g C^{-1} for dissolved organic matter (DOM) in a peat bog. Assuming again $C_5H_7O_2N$ as representative of DOM, 22.22 to 476.19 moles of DOM are necessary to remove one chloride atom from one mole of chlorinated contaminant.

Two alternative biodegradation models were tested on the batch experiment performed with aquifer material and polluted groundwater from piezometer 23 (sec. 2.2), before including one of these in the field scale model. The corresponding conceptual models are shown in Fig. C.1 of the appendix. In both models the assumptions are:

- The competition for hydrogen among different bacteria communities is not explicitly modeled. Regarding methanogenic bacteria, this might be a good assumption because low concentrations of methane were measured in the batch headspaces.
- Inhibitory effects due to sulphites and low pH's are not modeled [85].
- pH buffer due to carbonate anions is not modeled.
- The biomass grow is not considered, in accordance with previous studies [132, 28].
- Desorption from the soil matrix is assumed to be negligible.

- The reactions occurs under anaerobic conditions.

In the first model hydrogen production by substrate fermentation was explicitly described [105, 26, 131]:

$$r_{SUB-DEG} = \mu_{SUB} \frac{SUB}{SUB + K_{SUB}} I_{H_2} \text{ where } I_{H_2} = \exp\left(-\frac{H_2}{H_2^{scale}}\right) \quad (6.14)$$

$$r_{H_2} = f_{H_2} r_{SUB-DEG} - f_{DEC} \sum_{i=1}^n r_{i-DEG} \quad (6.15)$$

$$r_{PCE-DEG} = \mu_{PCE} \frac{PCE}{PCE + K_{PCE} \left(1 + \frac{TCE}{K_{TCE}}\right)} \frac{H_2 - H_2^*}{K_{H_2} + (H_2 - H_2^*)} \quad (6.16)$$

$$r_{TCE-DEG} = \mu_{TCE} \frac{TCE}{TCE + K_{TCE} \left(1 + \frac{PCE}{K_{PCE}}\right)} \frac{H_2 - H_2^*}{K_{H_2} + (H_2 - H_2^*)} \quad (6.17)$$

$$r_{cis-DCE-DEG} = \mu_{cis-DCE} \frac{cis-DCE}{cis-DCE + K_{cis-DCE} \left(1 + \frac{VC}{K_{VC}}\right)} \frac{H_2 - H_2^*}{K_{H_2} + (H_2 - H_2^*)} \quad (6.18)$$

$$r_{VC-DEG} = \mu_{VC} \frac{VC}{VC + K_{VC} \left(1 + \frac{cis-DCE}{K_{cis-DCE}}\right)} \frac{H_2 - H_2^*}{K_{H_2} + (H_2 - H_2^*)} \quad (6.19)$$

where μ_{SUB} is the substrate degradation rate of fermenting bacteria ($\text{mol L}^{-1} \text{d}^{-1}$), SUB is the substrate concentration ($\text{mol}_{SUB} \text{L}^{-1}$), I_{H_2} is the hydrogen inhibition factor, H_2 is the hydrogen concentration (mol L^{-1}), H_2^{scale} is the inhibitory aqueous hydrogen concentration, f_{H_2} is the stoichiometric coefficient of for hydrogen (-), f_{DEC} is the stoichiometric coefficient for dechlorination reactions (-), μ and K are the maximum degradation rates and half saturation constant for each contaminant, H_2^* is an hydrogen threshold concentration for

dechlorinators and K_{H_2} is the half velocity coefficient for hydrogen use by dechlorinators. The model described by Eq.6.14-Eq.6.19 was implemented in a RT3D batch module and the five maximum degradation coefficients μ were inferred from the measurements. Other model parameters were kept fixed, as reported in Tab.C.1. Since three separate experiments were performed (A, B and C), the inversion was repeated three times, one for each experiment, obtaining three μ values for each contaminant.

The simulation of each batch was divided in two parts, one from day 1 to 167 (first groundwater addition) and the other from day 174 to 384 (second groundwater addition, see Fig.2.5). The initial hydrogen and substrate concentrations specified at the beginning of the second part were assumed equal to the values simulated at the end of the first part. The sum of the squared residuals was minimized using SCE-UA global optimization algorithm.

In the second model a direct dependence of the degradation rates on the substrate concentration was assumed, without explicitly describing the substrate fermentation (Fig. C.1 of the appendix):

$$r_{SUB-DEG} = f_{SUB/VOC} \sum_{i=1}^{nc} r_{i-DEG} \quad (6.20)$$

$$r_{PCE-DEG} = SUB \mu_{PCE} \frac{PCE}{PCE + K_{PCE} \left(1 + \frac{TCE}{K_{TCE}}\right)} \quad (6.21)$$

$$r_{TCE-DEG} = SUB \mu_{TCE} \frac{TCE}{TCE + K_{TCE} \left(1 + \frac{PCE}{K_{PCE}}\right)} \quad (6.22)$$

$$r_{DCE-DEG} = SUB \mu_{cis-DCE} \frac{cis-DCE}{cis-DCE + K_{TCE} \left(1 + \frac{VC}{K_{VC}}\right)} \quad (6.23)$$

$$r_{VC-DEG} = SUB \mu_{VC} \frac{VC}{VC + K_{VC} \left(1 + \frac{cis-DCE}{K_{cis-DCE}}\right)} \quad (6.24)$$

Where $f_{SUB/VOG}$ is a stoichiometric coefficient expressing the amount of substrate required for the degradation of one mole of contaminant. This second model assumes that the fermentative step is not limiting dechlorination. In this case the maximum grow rates μ are scaled by the substrate content, with unit $\text{mol mol}_{SUB}^{-1} \text{ d}^{-1}$.

6.3.4.1. Results of the biodegradation models for the batch experiments

In Tab. 6.5 the maximum degradation rates estimated for the first model with a stoichiometric coefficient $f_{H_2} = 10$ are reported. The degradation coefficients are within the range reported in literature, except for the low value obtained for μ_{PCE} . The best fit was obtained for test B. As can be seen from Fig. 6.5, the second part of the experiment was better reproduced, possibly due to a lag phase for the degrading bacteria.

A second calibration experiment using $f_{H_2} = 0.1$ was also performed. In this case the μ_{SUB} rate was higher, compensating for the lower conversion of substrate to hydrogen, whereas the other rate coefficients were similar to the previous experiment. This result indicates that f_{H_2} and μ_{SUB} can not uniquely be estimated from contaminant concentration measurements. Additional hydrogen and substrate concentration measurements could support the estimation of f_{H_2} and μ_{SUB} coefficients required for modelling substrate fermentation with hydrogen production.

Since only contaminant concentration measurements are available, the use of the second simpler model might be justified. In the last experiment only the μ coefficients of the second model were calibrated, keeping the stoichiometric $f_{SUB/VOG}$ fixed at 600 (as mentioned above this stoichiometric coefficient might be variable between 100 and 1000). The $f_{SUB/VOG}$ stoichiometric coefficient was found to be almost insensitive to the measured concentrations because the organic substrate was not substantially consumed in the simulations (with a $f_{SUB/VOG}$ of 600 only 22.7% of the initial substrate content was consumed). In the long term scenarios, where several decades are simulated, the substrate consumption might be more significant. In the next chapter, the stoichiometric coefficient $f_{SUB/VOG}$ was directly inferred from the field measurements.

As can be seen from Fig. 6.5 the fits provided by the first and second model for batch B are similar, with $\mu_k SUB$ coefficients of the second model lower than the μ_k coefficients of the first model (due to the removal of the hydrogen term in equations Eq.6.20-Eq.6.24). The average log likelihood of the second model (Eq.4.16) is -1.55e6, similar to that obtained with the second model with f_{H_2}

Table 6.5.: Calibrated parameters for the first model with $f_{H_2} = 10$

Parameter	504 A	504 B	504 C	Aver- age value
μ_{PCE} (mol L ⁻¹ d ⁻¹)	1.84e-8	4.11e-8	1.64e-8	2.53e-8
μ_{TCE} (mol L ⁻¹ d ⁻¹)	4.27e-6	1.75e-5	3.68e-6	8.49e-6
μ_{DCE} (mol L ⁻¹ d ⁻¹)	8.43e-7	3.92e-6	8.71e-7	1.88e-6
μ_{VC} (mol L ⁻¹ d ⁻¹)	2.73e-7	3.03e-7	2.62e-7	2.79e-7
μ_{SUB} (mol L ⁻¹ d ⁻¹)	6.51e-3	1.55e-7	8.28e-4	2.44e-3
RMSE (μg L ⁻¹)	629.27	355.10	609.16	531.18

equal to 10 (-1.36e6, with all measurements weights equal to unity). Thus, the second model provides a description of natural attenuation process comparable to the first model, without introducing the substrate parameters f_{H_2} and μ_{SUB} , which were found correlated. In the field application the biodegradation of contaminants was described with the second simpler model.

TCA and DCA were also measured in the field, although not in the batch experiments. The degradation of TCA to DCA and of DCA to ethene was considered biologically mediated and depended on the amount of organic substrate in the soil. The reactions consume organic substrate and the resulting rates add to Eq.6.20. For simplicity TCA and DCA degradation rates were modeled with a first order kinetic rate:

$$r_{k-DEG} = \mu_k SUB C_k \quad (6.25)$$

Table 6.6.: Calibrated parameters for the second model with $f_{SUB/VOC} = 600$

Parameter	504 A	504 B	504 C	Aver- age value	Aver- age μ
$\mu_{PCE}SUB$ (mol L ⁻¹ d ⁻¹)	1.64e-8	1.53e-8	1.38e-8	1.52e-8	7.30e-8
$\mu_{TCE}SUB$ (mol L ⁻¹ d ⁻¹)	3.56e-6	5.41e-6	2.80e-6	3.92e-6	1.89e-5
$\mu_{DCE}SUB$ (mol L ⁻¹ d ⁻¹)	7.23e-7	1.25e-6	7.30e-7	9.02e-7	4.34e-6
$\mu_{VC}SUB$ (mol L ⁻¹ d ⁻¹)	2.47e-7	1.17e-7	2.30e-7	1.98e-7	9.55e-7
RMSE (μg L ⁻¹)	678.84	378.65	650.99	596.49	

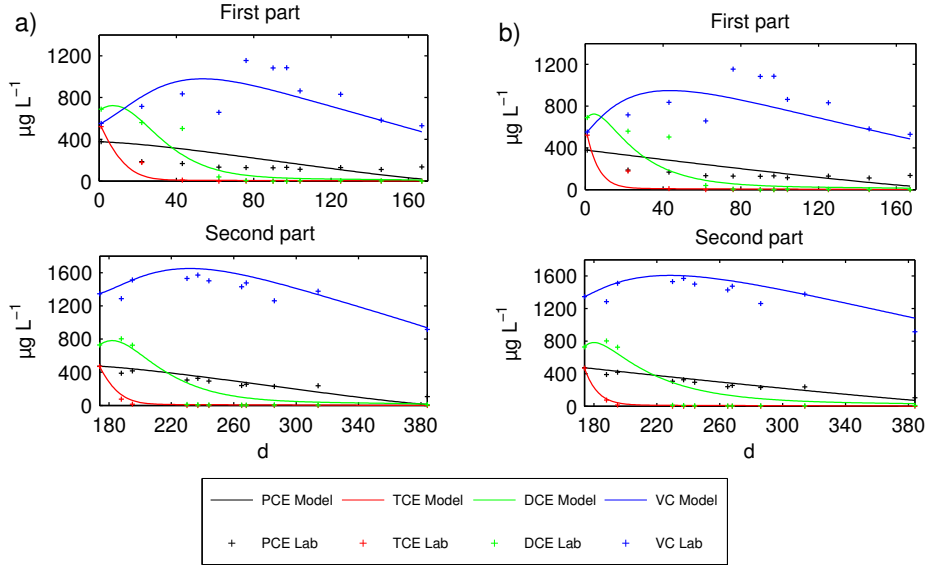


Figure 6.5.: Simulated (solid lines) and measured (crosses) concentrations: (a) with the first model and $f_{H_2} = 10$ and (b) with the second model and $f_{SUB/VOC} = 600$. Top figures show the first part of the experiment (1 to 167 days), bottom figures the second part (174 to 384 days).

The initial degradation rates $\mu_{TCA-DCA}$ and μ_{DCA} were derived from table Tab.6.3, dividing the maximum first order rate by the estimated substrate content in the deep layer. The initial substrate concentration was estimated from the carbon content (sec.6.3.2). Assuming all carbon is available as substrate, a total porosity of 0.4 (estimated from angled core samples), a bulk density of 1628 kg L^{-1} and $\text{C}_5\text{H}_7\text{O}_2\text{N}$ as representative molecule of the substrate, the initial substrate concentration was estimated at $1.02 \text{ mol}_{SUB} \text{ L}^{-1}$ in the shallow layer and $0.204 \text{ mol}_{SUB} \text{ L}^{-1}$ in the deep layer.

The degradation of TCA to DCE was assumed abiotic and not depended on the organic matter content. Therefore, the calculated reaction rate was not summed to Eq.6.20. Again, a first order kinetic rate was used:

$$r_{(TCA-DCE)-DEG} = \mu_{TCA-DCE} TCA \quad (6.26)$$

In this case $\mu_{TCA-DCE}$ was assumed equal to the maximum value reported in Tab. 6.3.

7. Inverse modelling at the site and prediction of the PRB efficiency¹

In this chapter the parameters of the site model (chapter 6) are inferred using PEST and high performance computing resources (Tab.D.1). In the first part the the inversion theory on which PEST is based and the procedure used to infer the model parameters are described. In the second part the results of the inversion experiments are presented and discussed. Finally, the long term PRB efficiency is assessed using the calibrated parameters and the geochemical model developed in chapter 3.

7.1. Non linear inversion theory

In this section a brief explanation of the inversion theory used to infer the parameter values at the site is presented. At the initial parameter guess ϕ_0 the vector of simulated field observations is $\eta(\phi_0, Z)$ (thereafter $\eta(\phi_0, Z) = Y_0$). By using the Taylor's rule, the vector of the model simulation at the parameter value ϕ can be approximated as:

$$\eta(\phi, Z) \approx Y_0 + X(\phi - \phi_0) = Y_0 + X\Delta_\phi \quad (7.1)$$

where X is the Jacobian matrix of the model η at the parameter value ϕ_0 . The Jacobian matrix contains n rows (one for each observation), the p element of each row is the derivative of a particular observation respect the parameter p .

¹Based on: Carniato L., Schoups G., van De Giesen N. and Seuntjens P., Bastiaens L., and Sapion H. (2014). Highly parameterized inversion of groundwater reactive transport for a complex field site (under review in *Journal of Contaminant Hydrology*).

The sum of the squared weighted errors between $\eta(\phi, Z)$ and Y_0 can thus be defined as:

$$\Phi_m = (\eta(\phi, Z) - Y_0 - X\Delta_\phi)^T Q (\eta(\phi, Z) - Y_0 - X\Delta_\phi) \quad (7.2)$$

where Q is the weighting matrix (which diagonal contains the squares of the observation weights). Supposing that the “true” parameter vector is ϕ such that $\eta(\phi, Z) = Y$, the aim of the gradient based estimation (on which PEST is based) is to find the optimal parameter upgrade Δ_ϕ that minimizes Eq.7.2, defined as measurement objective function. A necessary condition is that the first derivative of Eq.7.2 is equal to zero. Such condition is satisfied when the parameter upgrade vector is:

$$\Delta_\phi = (X^T Q X)^{-1} X^T Q (Y - Y_0) \quad (7.3)$$

since the linear approximation is valid only around ϕ_0 , the parameter vector must be updated iteratively until the measurement objective function reaches a small value below which the parameters are considered estimated:

$$\phi_{IT} = \phi_{IT-1} + \Delta_{\phi IT} \quad (7.4)$$

where IT denotes the iteration counter. One problem in the calculation of the parameter upgrade vector is that the inversion of the $X^T Q X$ matrix can be difficult if it is rank deficient (or nearly singular). To prevent this situation, $X^T Q X$ matrix can be supplemented with a diagonal term:

$$\Delta_\phi = (X^T Q X + \lambda I)^{-1} X^T Q (Y - Y_0) \quad (7.5)$$

here I denotes the identity matrix and λ the Marquardt lambda parameter [106, 93]. The λ parameter is adjusted during the estimation process, usually starting with a high value (which corresponds to the “steepest descend” method) and ending with a low value. High value of λ ensure that the $(X^T Q X + \lambda I)$ matrix can be inverted.

Another problem that can be encountered is that the optimal parameter vector corresponds to a local minimum of the objective function. Global optimization methods, such as SCE-UA (chapter 4) or Monte Carlo methods can be

used instead (chapter 5). However, these methods require more model runs, especially when the number of estimated model parameters increases. Use of Eq.7.4 and Eq.7.5 is very efficient and it is often employed when the number of parameters is high and the model execution time is long (such as in reactive transport models). Moreover, an extra regularizing term can be added that restrict the parameter search on feasible areas of the parameter space. The regularization term acts on the parameter values and can assume the following form:

$$\Phi_r = (Z\phi - j)^T R(Z\phi - j) \quad (7.6)$$

where Z is a p by p matrix which contains the regularization relationships between parameters, j the “observed” values of these relations (assuming that these relations are linear and can be expressed by a matrix operation) and R is a relative weight matrix assigned to regularization observations j . For example a row of the Z matrix could be zero except for two elements equal to 1 and -1 and the relative element of j equal to 0. In this way it is assumed that two particular parameters (corresponding to the column indexes of the Z matrix) will be forced to be equal (homogeneous condition). Other constraints can be implemented in the Z matrix, such as a preferred value condition. The resulting global objective function is then reformulated as:

$$\Phi = \Phi_m + \beta^2 \Phi_r \quad (7.7)$$

where β^2 is the regularization weight factor. The parameter upgrade vector including the β^2 parameter and the regularization observations is:

$$\Delta_p = (X^T Q X + \lambda I + \beta^2 Z^T R Z)^{-1} (X^T Q (Y - Y_0) + \beta^2 Z^T R (j - j_0)) \quad (7.8)$$

where j_0 equals to $Z\phi_0$. Selection of an appropriate β^2 value is crucial. An high value of β^2 ignore the fitting of the observation in favor of fitting the regularization constraints j . A small value of β^2 will produce parameter estimates that strongly differs from the regularization constraints, which often contains our assumptions about reasonable parameter values. To minimize Eq.7.7 BeoPEST [62] was used, a particular version of PEST software package [39] which allows an efficient parallelization of the model runs. The value of β^2 is iteratively determined by PEST from the user specified target measurement

(PHIMLIM, Φ_l^m), which express the value below which the model is deemed calibrated.

After the calculation of the Jacobian matrix, PEST decreases or increases the current β^2 (based on the current value of the measurement objective function), estimating the corresponding parameter upgrade from Eq.7.8 and the measurement objective function from Eq.7.2 (assuming linearity). The optimal weight factor for the current iteration is found when the measurement objective function equals to Φ_l^m . After the optimal weight factor is determined, PEST tests different values of the Marquardt parameter, calculating the corresponding parameter upgrades and choosing the parameter vector with the largest reduction of the global objective function (using the full model).

Despite the addition of the λI matrix in Eq.7.8, strong rank deficient $X^T Q X$ matrices can be calculated when many parameters are involved in the estimation process. Besides the Marquardt parameter and regularization, truncated singular value decomposition can be applied on the first part of Eq.7.8 :

$$(X^T Q X + \lambda I + \beta^2 Z^T R Z) = V E V^T \quad (7.9)$$

where V is the matrix of the eigenvectors of $(X^T Q X + \lambda I + \beta^2 Z^T R Z)$ and E is the diagonal matrix of eigenvalues. If $(X^T Q X + \lambda I + \beta^2 Z^T R Z)$ is positive semidefinite, its eigenvalues are real, and its eigenvectors are orthogonal ($V^T = V^{-1}$), V can be written as:

$$V = [V_1 V_2] \quad (7.10)$$

where V_1 contains eigenvectors corresponding to the largest eigenvalues and V_2 contains the remaining eigenvectors, including those having zero eigenvalues. In the inversions experiments of this chapter, V_1 matrix contains the eigenvectors having a ratio eigenvalue/maximum eigenvalue larger than 1e-6. To exclude the estimation of insensitive parameter only V_1 is considered (operation is known as “truncation”). The parameter upgrade vector can be redefined after truncation as:

$$\Delta_\phi = V_1 E_1^{-1} V_1^T (X^T Q (Y - Y_0) + \beta^2 Z^T S (j - j_0)) \quad (7.11)$$

where E_1 corresponds to the diagonal matrix of the largest eigenvalues. The use of singular value decomposition is also beneficial for correlated parameters [58]. The use of all suites of regularization techniques provided in PEST

enables the estimation of many parameters, as those that generates from the pilot point technique.

Our approach differs from the one used by Tonkin and Doherty [148], where $X^T Q X$ is decomposed around a pre-calibrated parameter values and super parameters are defined only at the beginning of the optimization. In their approach it is assumed that the super parameters constitutes a linear combination of estimable parameter set, despite the Jacobian matrix X changes as the inversion progresses (model non linearity). Therefore there might be chances that the super parameters defined at the beginning of the inversion process do not span the estimable parameter space at the end of the inversion, resulting in an unsatisfactory model calibration. To avoid this problem the full Jacobian matrix and the subsequent decomposition and truncation can be performed at each iteration, as was done in the our inversion experiments. However, the computational effort compared to the Tonkin and Doherty [148] approach is greatly increased.

7.1.1. Parameter and prediction uncertainty

Parameter uncertainty after regularized inversion was assessed using the PRE-DUNC7 utility of the PEST software package, where is assumed that a linear relation between the measurements Y and the optimal parameter set holds:

$$Y = X\phi + \varepsilon \quad (7.12)$$

where ε is the vector of residuals (due to measurement and model errors). Under this assumption the following equation holds:

$$\begin{bmatrix} Y \\ \phi \end{bmatrix} = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} \varepsilon \\ \phi \end{bmatrix} \quad (7.13)$$

Using the standard matrix relationships for covariance propagation, the joint covariance of measurements and model parameters equals to:

$$cov \left(\begin{bmatrix} Y \\ \phi \end{bmatrix} \right) = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} I & 0 \\ X^T & I \end{bmatrix} = \begin{bmatrix} X C X + \Sigma & X C \\ C X^T & C \end{bmatrix} \quad (7.14)$$

where Σ is the covariance matrix of the measurement errors and C is the prior covariance matrix of the model parameters, expressing the pre-calibration parameter uncertainty. When parameters are conditioned to the measurements Y the pre-calibration parameter uncertainty decreases. The posterior covariance of model parameters can be expressed as [33]:

$$\tilde{C} = C - CX^T[XX^T + \Sigma]^{-1}XC \quad (7.15)$$

with a measurement error covariance:

$$\Sigma = \sigma_r^2 Q \quad (7.16)$$

where Q is the weighting matrix (weights assumed inversely proportional to the error standard deviations with cross terms equal to zero) and σ_r^2 is a reference variance (calculated at the end of the optimization as $\sigma_r^2 = \Phi_m/(n-p)$, where n is the number of observations and p the number of parameters).

To estimate the standard deviation of the predictions PREDUNC6 utility was used. Also in this utility it is assumed that linearity exist between the prediction s and model parameters :

$$s = y\phi \quad (7.17)$$

where y is the sensitivity of model prediction s to parameters ϕ (corresponding to a row of the Jacobian matrix). The post calibration variance of the prediction s can be obtained from Eq.7.15 with basic matrix manipulations:

$$\sigma_s^2 = y\tilde{C}y^T = yCy^T - yCX^T[XX^T + \Sigma]^{-1}XCy^T \quad (7.18)$$

The prior covariance C is reconstructed from prior knowledge of parameter uncertainty as described in sec.7.6.1.1. It should be noted that assuming linearity is a reasonable approximation for small parameter variations. For large parameter variations (large posterior parameter standard deviations) Monte Carlo analysis should be used to characterize the parameter and prediction uncertainty [158]. However, this approach was unfeasible in this study, due to the large number of model parameters and the long computation times required to solve the reactive transport equations. An overview of the inversion tests and prediction uncertainty estimations is reported in Tab.D.1.

7.2. Inversion procedure

Several model parameters, listed in Tab.7.1, were not directly measured in situ and were estimated from available groundwater levels and concentration data. Three different optimization problems were set up, each with its own objective function, data, and set of parameters to be estimated:

1. Optimization of flow parameters

$$\Phi = \Phi_{m,1} + \beta^2 \Phi_r \quad (7.19)$$

2. Optimization of reactive transport parameters

$$\Phi = \sum_{j=2}^7 \Phi_{m,j} \quad (7.20)$$

3. Joint optimization of flow and reactive transport parameters

$$\Phi = \sum_{j=1}^7 \Phi_{m,j} + \beta^2 \Phi_r \quad (7.21)$$

where $\Phi_{m,j}$ is the measurement objective functions for the set j , j is the measurement set index ($j = 1$ groundwater levels, $j = 2$ PCE, $j = 3$ TCE, $j = 4$ cis-DCE, $j = 5$ VC, $j = 6$ TCA, $j = 7$ TCA). For hydraulic conductivity a preferred value of 0.19 m d^{-1} was specified as regularization condition, equal to the mean hydraulic conductivity value estimated from the slug in and slug out tests (Fig.7.5). The preferred value for specific yield was specified at 0.1 [69]. A smoothness regularization constrain was imposed for the recharge pilot points, with squared weights inversely proportional to the values of the variogram model.

In the flow optimization (step 1), 351 flow parameters (Tab.7.1) were calibrated on 420 transient groundwater head levels collect in 2011, whereas the measurements before 2011 were used as validation dataset. In the optimization of reactive transport parameters (step 2), 18 parameters were calibrated on 2240 concentration data (corresponding to degradation, source and transport parameters of Tab.7.1), keeping the flow field fixed as estimated from step 1. In the joint optimization (step 3), flow (178 values of hydraulic conductivities and the 3 drain conductances) and transport parameters were estimated from head and concentration data (420 head and 2240 concentration measurements). Therefore, two optimal parameter vectors (369 parameters)

were compared. The first obtained after step 2, with flow parameters inferred from head measurements and transport and chemical parameters inferred from concentration measurements. The second obtained after step 3, refining the estimations of hydraulic conductivities and drain conductances obtained at step 1 with concentration data and determining reactive transport parameters from head and concentration measurements.

7.3. Optimization of the flow parameters

As described in sec. 7.1, the regularization weight is determined by β^2 , which is calculated at the starting of each Levenberg-Marquardt iteration to ensure that Φ_m rises no higher than a user specified target objective function Φ_l^m . Therefore, Φ_l^m has a strong influence on the estimated parameter fields, and should not be specified below the expected measurement and model errors. For the optimization of the flow parameters, Φ_l^m was set to 9.45, which corresponds to a RMSE of 0.15 m. Such value was deemed reasonable and accounts for the precision of the measurements and neglected daily groundwater fluctuations. This assumption was supported by detailed information provided by pressure loggers, where the hydraulic pressure was measured every 15 minutes for almost 1 year. In Fig. 7.1 the raw data and an example of a model simulation are compared. As can be seen the weekly model outputs differ from the pressure logger time series, in particular during strong precipitation events (for example around 30/11/11). The model was discretized with weekly stress period to keep the model runtime reasonable for a pilot point calibration (around 4 minutes for a flow simulation). Therefore, the simulated heads represent average weekly values and discrepancies with the punctual measurements must be allowed to obtain a reasonable parameter estimation [45].

7.3.1. Full recharge model results

As can be seen from Fig. 7.2(a), the estimated hydraulic conductivity assumes the preferred value over large portions of the model domain (corresponding to a value of -0.712 in the log scale). High values were calculated in the pasture area, with a peak of 15.309 m d⁻¹ close to piezometer 16. Fig. 7.5 shows the estimated hydraulic conductivities from slug in and slug out tests. As can be seen, the estimation of hydraulic conductivities removing groundwater (slug out) are on average four times lower than those adding groundwater (slug in). The results indicates that these tests can provide a raw estimation of the local

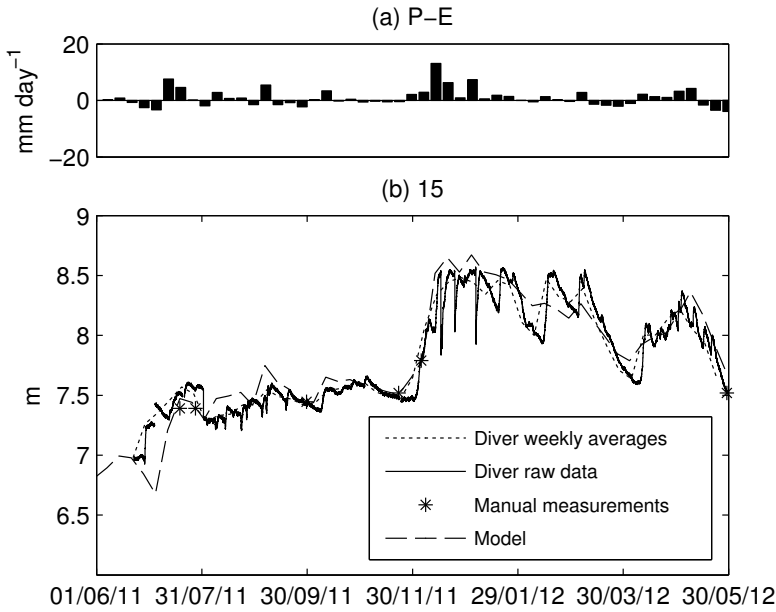


Figure 7.1.: (a) Weekly precipitation minus evaporation time series and (b) pressure logger data and model results.

hydraulic conductivity value[17]. The estimated values with the full recharge model are closer to the higher values of the slug in estimations, because higher hydraulic conductivities allow more water to be drained out from the system.

In Fig. 7.2(b) shows the estimated specific yield field. Specific yield ranges from 0.009 to 0.966, with extreme values close to piezometer 5. These high values are outside the upper value suggested for clay soils and might indicate the presence of model errors [58]. Large variations of the influent fluxes in the full recharge model cause large fluctuations of the water levels in the pasture area and in the close factory area, which are dumped by high specific yield values. In the areas where no groundwater levels measurements were taken, the specific yield was close to the preferred value (0.1) imposed by the regularization constraints.

In Fig. 7.2(c) and (d) the residual histograms (observed minus simulated) for the calibration and validation data sets are shown. In calibration the residual histogram is unbiased, with a mean value of -0.004 m, while in validation is

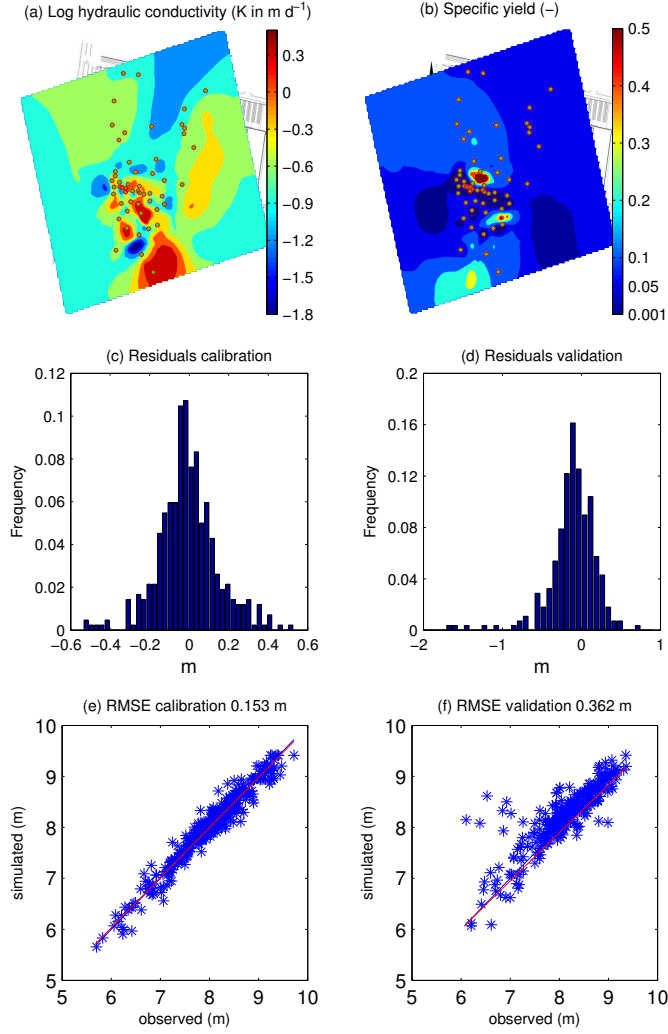


Figure 7.2.: Results with the full recharge model and $\Phi_l^m = 9.45$. (a) Horizontal hydraulic conductivity field, (b) specific yield field (c), residual histograms for (d) calibration and (d) validation and cross plots for (e) calibration and (f) validation.

negatively biased (-0.118 m). Fig. 7.2(e) and Fig. 7.2(f) show the calibration and validation cross plots. Here, the RMSE in the calibration step is 0.153 m, close to the average error imposed with a Φ_l^m value of 9.45 (0.15 m), while in validation the RMSE is considerably increased (0.362 m). The validation dataset includes some outliers, taken close to the barrier before its installation, which were up to 2 meters lower than post barrier installation measurements and could have been wrongly measured (due for example to the misspecification of the piezometer elevation). This problem was not present in calibration dataset, because the position and elevation of all monitored piezometers was surveyed before starting the 2011 monitoring campaign. The calibration cross plot shows a good match of the simulation against the observations.

In Fig. 7.3 the transient simulated groundwater heads and the measurements used in calibration and validation are shown for four piezometers along the plume direction. As can be seen, the amplitude of groundwater level fluctuations is higher for the piezometers in the pasture area (piezometers 10, 21 and 16) compared to the piezometer in the paved area (piezometer 3). Visually a good fit was obtained both in calibration and validation. For piezometer 21 the water table approaches to the land surface, as confirmed by the water ponds observed after strong precipitation events.

7.3.2. Reduced recharge model results

The groundwater recharge at the water table is determinant for the estimation of the hydraulic conductivities. Lower hydraulic conductivities are expected at lower net recharge values (272 mm year⁻¹ for the full recharge model compared to 42 mm year⁻¹ for the reduced one). This result is confirmed in Fig. 7.4, where lower hydraulic conductivities were obtained in the pasture area compared to the full recharge model. In this case the maximum value in the pasture area is 1.04 m d⁻¹ (close to piezometer 16) and 3.41 d⁻¹ at the north boundary, where groundwater flow is poorly characterized (few head measurements available). From Fig. 7.5 it is evident that the reduced recharge estimations in the pasture area agree more favorably with the slug-in estimations whereas in the factory area the model estimations are comprised between the slug-in and slug-out values.

Despite the reduction of the infiltration rate at the water table discrepancy between the hydraulic conductivities estimated with pilot points and slug data still persist. A similar result was found by Kowalsky et al. [86] and was attributed to a possibly incorrect analysis of the slug data, such as the specification of the influence radii, or the different averaging of the hydraulic conductiv-

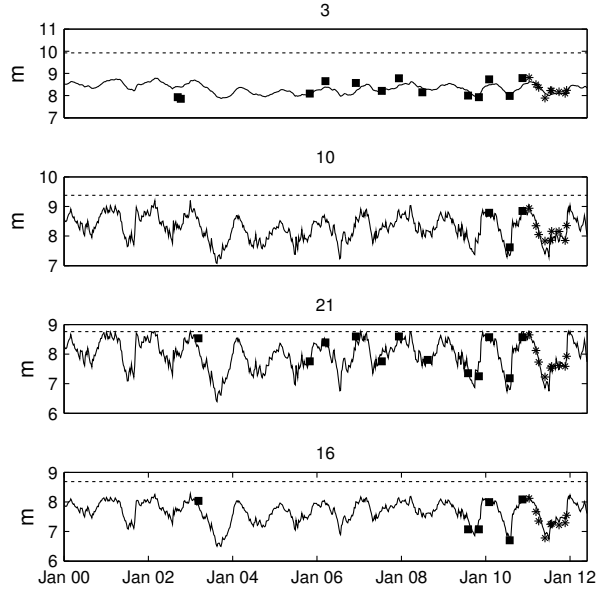


Figure 7.3.: Transient groundwater heads in the factory and pasture areas using the full recharge model. The 2011 measurements were used for calibration (asterisks) whereas the groundwater levels collected before were used for validation (squares).

ities in the slug test estimations and in the pilot point estimations. Moreover, the aquifer tests provide an estimation of the hydraulic conductivity at a scale which is typically smaller than the model discretization [6]. Therefore, some discrepancy between the hydraulic conductivities estimated through inverse modeling and the slug tests can be considered acceptable.

In Fig. 7.4(b) the estimated specific yield field is shown. Compared to the full recharge case, no unreasonable high values were obtained close to the piezometer 5, with a minimum specific yield value of 0.011 in the pasture area and a maximum value of 0.164 in the factory area, which are more plausible for a sandy clay soil [69]. The estimated recharge field (Fig. 7.14) was homogenous, with a maximum and minimum value of 0.354 and 0.328 respectively.

The residual distribution in calibration and validation step (Fig. 7.4(c,d)) can be considered unbiased (-0.014 m and -0.01 m respectively), in contrast to

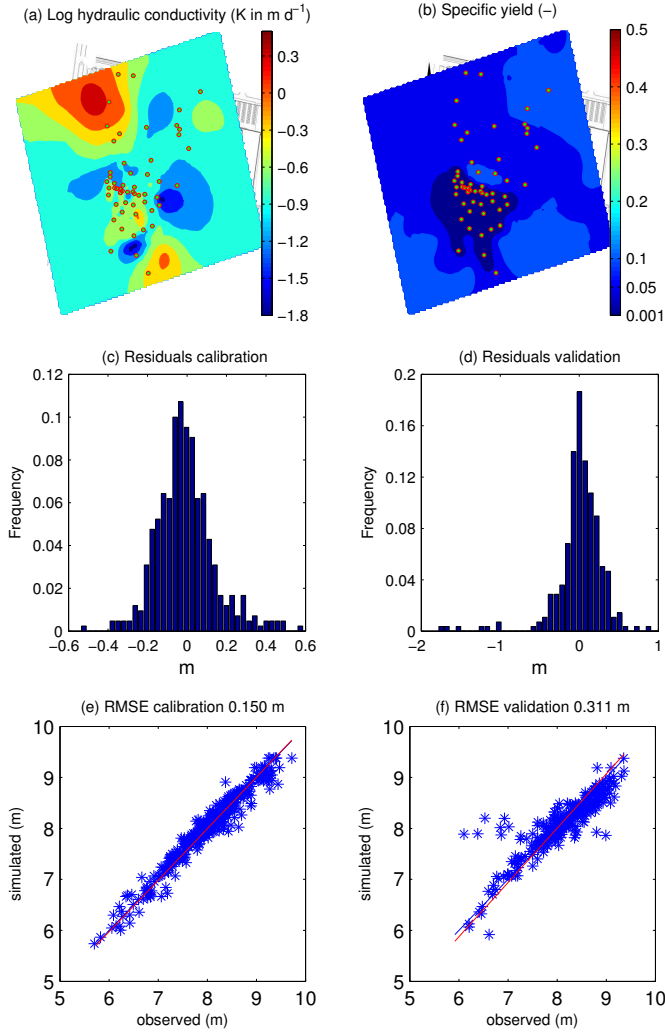


Figure 7.4.: Results with the reduced recharge model and $\Phi_l^m = 9.45$. (a) Horizontal hydraulic conductivity field, (b) specific yield field (c), residual histograms for (d) calibration and (d) validation and cross plots for (e) calibration and (f) validation.

the full recharge model where an appreciable negative bias was estimated in validation (-0.118 m). The results achieved in calibration are comparable to the full recharge case, since the same Φ_l^m was imposed (Fig. 7.4(e)). However, better results were obtained in the validation step, with a reduction of the RMSE of 0.051 m (Fig. 7.4(f)), indicating that slightly better predictions of the groundwater levels are provided by the reduced recharge model.

Since the two recharge models provide a similar fit to the groundwater levels, even with different hydraulic conductivity fields and net recharge values at the groundwater table, the choice of the appropriate recharge model should also consider the information content of the concentration measurements. In Fig. 7.10 the calibrated drain conductances are reported (white bars), together with an estimation of the 95% confidence interval (based on the estimation of the posterior parameter covariance matrix, Eq.7.15, thus accounting for prior estimation of parameter uncertainty). Compared to the initial estimations (black bars), the drain conductances were increased.

7.3.3. Effect of low regularization constrains

The model parameters of the reduced recharge flow model were recalibrated imposing a Φ_m^l value equal to 1.05 (corresponding to a root mean square error of 0.05 m). The algorithm was not able to fit the head values to the

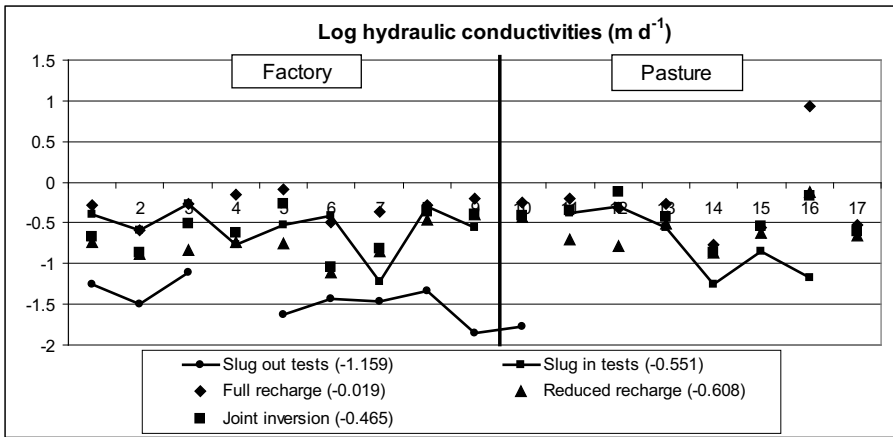


Figure 7.5.: Estimated hydraulic conductivities from slug tests and from inverse modelling. In brackets the average values are reported.

specified target and the minimum measurement objective function was 3.633, corresponding to a RMSE of 0.082 m. In calibration the average value of the residuals was $-4.60\text{e-}3$ m, which can be considered an unbiased estimation. Therefore, imposing a closer fit decreases the negative residual bias previously calculated. In validation the average value of the residuals was $9.34\text{e-}4$ m (Fig. 7.6(d)).

Compared to the previous experiment, the RMSE in calibration was reduced of 0.068 m, producing an highly heterogeneous hydraulic conductivity field (Fig. 7.6(a)) and unrealistic specific yield values in the factory area (with values up to 1, Fig. 7.6(b)). The RMSE in validation was only marginally reduced. Hydraulic conductivity has a considerable effect on the spreading of the contaminants and should also be conditioned by concentration measurements.

7.4. Conservative transport simulations

7.4.1. Description of the experiments

Before simulating multicomponent reactive transport some preliminary experiments considering only the total contaminant concentration were performed (TVOC, total concentration of volatile organic compounds). In these experiments neither retardation nor biodegradation were included.

The aim was testing the effect of different recharge models and specifications of the initial concentration fields on the simulated concentrations. The barrier is supposed to be installed on 1 October 2005 and able to degrade the total contamination accordingly to Eq.3.1. Effective porosity and longitudinal dispersivity were fixed at 0.1 [112] and 5 meters respectively. The model simulations were extracted at the 2009-2012 sampling events and averaged. The mean observed and simulated values were interpolated using inverse-distance-squared interpolation. In this manner the maps of the simulation results can directly be compared with the maps of the average measured values.

For the first two experiments (2(a-b) in Tab.D.1) the simulations start in June 2003 from the average total concentration field before the barrier installation (Fig. 7.7(a)), with measurements mostly collected between April 2002 and April 2003. At each source a residual DNAPL of 1000 kg was assumed to be left from the release period (1957-1996).

For the next two experiments (2(c-d)) the simulations start in January 1957 (discretization 3) from a pristine condition. In these experiments a contaminant release rate of $1.5\text{e-}3$ kg $\text{m}^{-2}\text{d}^{-1}$ was assumed ($k_{k-SOU,i}$ in Eq.6.7), yielding

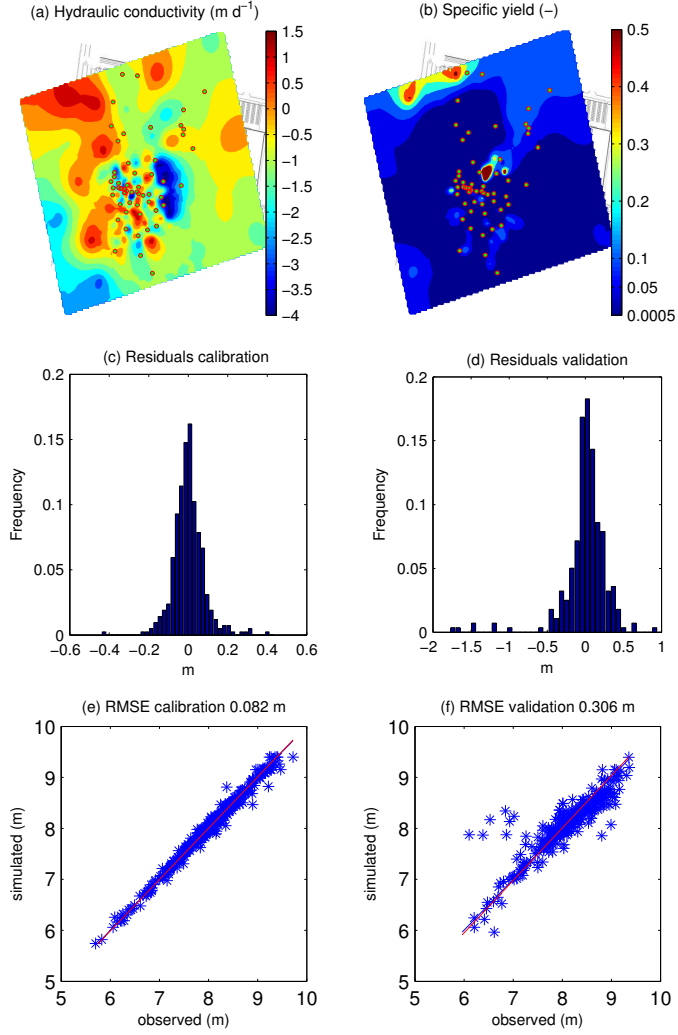


Figure 7.6.: Results with the reduced recharge model and $\Phi_l^m = 1.05$. (a) Horizontal hydraulic conductivity field, (b) specific yield field (c), residual histograms for (d) calibration and (d) validation and cross plots for (e) calibration and (f) validation.

a total released DNAPL mass of 1000 kg. The source parameters were deliberately high to simulate the maximum achievable contamination.

7.4.2. Results

In Fig. 7.7(c) the results using the full recharge model and the average concentration field after the barrier installation are shown. By comparing the simulation results with the post installation measurements (Fig. 7.7(b)) it is evident that the contaminant concentrations in the pasture area are underestimated. This underestimation can be explained by an excessive infiltration rate and by a poor specification of the initial concentration field. Indeed, the piezometer in the pasture area with the highest concentration (piezometer 15) was installed only after the barrier and was not included in the initial concentration field.

Using the reduced recharge model (Fig. 7.7(d)) the plume is less diluted, as can be seen from the wider plume extension and the higher contaminant concentrations at source 3. However, this did not fix the underestimation of the concentrations at piezometer 15. A more realistic simulation should start from 1957 with a pristine aquifer, thereby allowing the plume to fully develop in the pasture (also in the area that was not monitored before the barrier installation).

In Fig. 7.7(e) the simulation using the full recharge model and starting the contaminant spillage in 1957 in a pristine aquifer is shown. Comparing this figure with Fig. 7.7(b), it is evident that the contamination in the pasture area as well as the residual DNAPL at the source 3 is diluted by rainwater. Even when a higher contaminant mass was released at the sources (10000 kg at each source) and a higher mass transfer was used ($1e-3 \text{ d}^{-1}$), the contaminant concentrations in the pasture area were underestimated and in the factory area largely overestimated. This result indicates that the underestimation of contamination in the pasture with the full recharge model is not due to underestimation of the released mass but to excessive dilution by infiltrating rainwater.

In Fig. 7.7(f) the results starting from 1957 and using the reduced recharge model are shown. Compared to the full recharge case, the dilution effect in the pasture area and at the source 3 is reduced, although the concentrations at piezometers 15 and 23 were still underestimated. The contaminant plume originating from source 3 is clearly divided after the barrier installation, as confirmed by the measured concentrations at piezometers 27 and 9, where an abrupt decline of pollutant concentrations was measured after barrier instal-

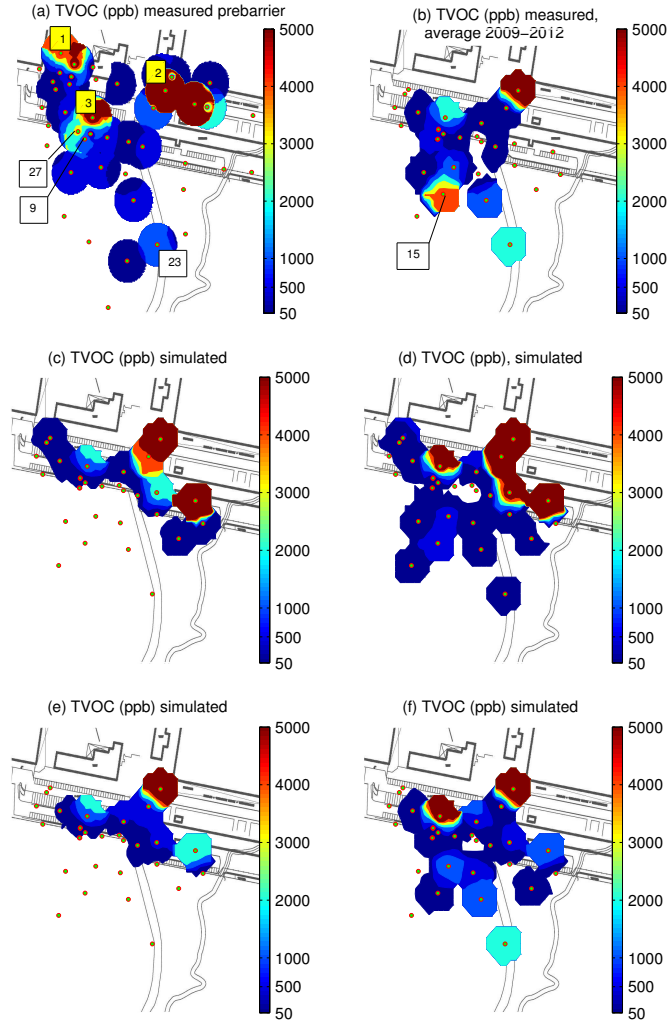


Figure 7.7.: (a) Interpolated TVOC concentrations before the barrier installation (average 2000–2003 period) and (b) after the barrier installation (2005–2012 period), (c) simulated post barrier concentrations with the full recharge model and (d) with the reduced recharge model (simulations starting from 2003), (e) simulated post barrier concentrations with the full recharge model and (f) with the reduced recharge model (simulations starting from 1957). In figure (a) and (b) the source locations are indicated by yellow text boxes and the piezometer numbers by white text boxes.

lation. As for the Fig. 7.7(e) the plume originating from source 2 traveled too far towards the right part of the barrier, without being diluted by rainwater or biodegraded. The overestimation of contaminant concentrations at these locations might be corrected by introducing sorption and biodegradation and by conditioning the hydraulic conductivity (i.e. the flow direction) on concentration data.

With regard to contaminant dilution in the pasture area, an alternative to reducing recharge would be to introduce different hydraulic conductivity values for the top and bottom layer. In particular, a higher conductivity of the first layer would allow more rainwater to be drained in the first layer, whereas a lower conductivity of the bottom layer would impede rainwater to percolate to the bottom layer, where most contamination is present. We tested this by repeating the flow optimization, using a preferred hydraulic conductivity value for the bottom layer equal to the average value of the slug tests (since most of the slug tests were performed for piezometers screened in the deep layer) and a ten times higher hydraulic conductivity value for the top layer. Such a conceptualization is supported by bore logs, which indicate a gradual passage from sandy clay (top layer) to clay sediments (bottom layer). Enforcing hydraulic conductivity differences between the two layers resulted in a slightly worse fit to measured heads (0.156 m in calibration and 0.347 m in validation), but partially avoided contaminant dilution. However, concentrations in the source areas were overestimated due to the lower hydraulic conductivity (and effective velocities) in the deep layer. In the conservative transport tests, the large overestimation at the source areas can only be reduced by lowering the source parameters, bringing back the situation where concentrations in the pasture area are underestimated. Moreover, the hydraulic conductivities in the top layer appeared too high, as indicated by the difference between the slug-in estimate in the shallow piezometer 10 (0.15 m d^{-1}) and the corresponding estimation obtain through inverse modelling (1.763 m d^{-1}).

From these conservative experiments it can be concluded that:

- The reduced recharge model is more appropriate as it avoids the strong dilution effect simulated with the full recharge model, resulting in severe underestimation of contaminant spreading, even without accounting for sorption or biodegradation. High contaminant concentrations were observed 220 meters downstream from the factory area (piezometer 23), which could not be reproduced using the full recharge model and unrealistically large source parameters.
- Starting the contaminant transport simulations from 1957 with a pristine aquifer gives better results than starting from 2003 because representa-

tive fields of the pre-installation concentrations in the pasture area are not available and due to the slow groundwater velocity, which makes the simulation of short periods strongly dependent from the choice of the initial concentration field.

7.5. Optimization of reactive transport parameters

7.5.1. Description of the experiments

Using the calibrated flow field ($\Phi_m^l = 9.45$, reduced recharge model) the model was used to simulate the contaminant concentrations from January 1957 to June 2012. The model parameters are listed in Tab. 7.1.

The initial amount of biodegradable substrate *SUB* was calculated from the total organic carbon content, as described in sec. 6.3.4.1. The substrate consumption by biodegradation was accounted for by setting the $f_{S/VOC}$ coefficient equal to 600. The organic carbon fraction f_{oc} was kept fixed to the initial estimates (sec. 6.3.2) and the consumption of organic carbon by biodegradation was neglected in the calculation of the retardation coefficient.

The initial values of the model parameters are reported in the third column of Tab. 7.1. As indicated in the table, most of the model parameters were derived from literature values, except for the source parameters and four degradation parameters that were manually adjusted to obtain realistic initial estimation of the concentration fields. The release rates $k_{k-SOU,i}$ at each source (for which no information is available from the historical records) were adjusted from the value used in the conservative transport simulations ($1.5e-3 \text{ kg m}^{-2}\text{d}^{-1}$), which produce an overestimation of the concentrations at the sources.

Four numerical experiments assuming known flow field were performed. The first experiment uses the degradation coefficients determined from batch experiments (Tab. 6.6). In the second experiment the degradation coefficients μ_{PCE} , μ_{TCE} , $\mu_{cis-DCE}$ and μ_{VC} were manually adjusted to reproduce the average post barrier concentration maps. The parameter vector obtained from manual calibration represent the starting condition for the third experiment, where an inversion with equal weights within each measurement set was performed (similarly to the approach used in chapter 3). In the last experiment the weights were assigned inversely proportional to the measured concentrations, only accounting for the reported analytical error.

7.5.2. Simulations with degradation rates determined from batch experiments

Using the degradation coefficients determined from the batch experiment the TCE, cis-DCE and VC plumes were strongly underestimated in the pasture area and the substrate content significantly depleted (around 97% of the initial substrate was consumed in the contaminated areas, with a 30 fold decrease of the initial degradation rates). This indicates that the μ_{PCE} , μ_{TCE} , $\mu_{cis-DCE}$ and μ_{VC} coefficients reported in Tab. 6.6 were too high when applied directly to the field model.

7.5.3. Manual tuning of the degradation rates

Since the TCE, cis-DCE and VC contamination was not reproduced using the degradation coefficients determined with batch experiments, the μ_{PCE} , μ_{TCE} , $\mu_{cis-DCE}$ and μ_{VC} coefficients were manually tuned to reproduce the contamination. A limitation of this exercise is that the flow field was kept fixed, as conditioned from head measurements and adjusting the chemical parameters can not fix any error in the flow field.

The resulting parameter values reported in Tab. 7.1 (second column) were lower than those determined using batch data. Fig. 7.8 shows the maps of the average concentrations for each species, calculated in the same way as the TVOC concentration maps shown in Fig. 7.7 and considering all sampling events (the top row depicts the maps of average observed concentrations and the central row the maps of average simulated concentrations).

From the figures it is evident that PCE and TCE concentrations were underestimated. For cis-DCE the high concentration measurements at the piezometer 15 were not reproduced, because the piezometer is located at the plume fringe. The same problem occurs for the cis-DCE plume originating from source 2, which is directed southwards whereas the measurements suggest a more eastward orientation (Fig. 7.8(c)). As a result, the highest concentrations were simulated in the area between piezometers 18 and 6 where no measurements were collected. The elevated VC concentrations at piezometer 23 (Fig. 7.8(d)) were not simulated. As for the cis-DCE plume, the VC plume originating from source 2 is directed southwards with the concentration peak in a area with no measurements. Moreover, the simulated peak was lower than the maximum measured VC concentration, suggesting that the $\mu_{TCA-DCE}$ rate coefficient must be increased.

Cross plots for cis-DCE, VC and DCA are shown in Fig. 7.9. As can be seen

from the figures extreme outliers are present for these contaminants, corresponding to the the measurements taken at piezometers 18 and 19 during the first monitoring campaign in 2003. The composite sensitivity of these outliers was estimated at the manually calibrated parameters. Observation composite sensitivity was calculated with [39]:

$$s_j = (Q(XX^T))_{j,j}^{0.5}/n \quad (7.22)$$

where j represent a particular observation and n the number of observations. Weights (square roots of the diagonal elements of the Q matrix) were calculated as described in the next subsection. Outliers with high composite sensitivities were eliminated from the calibration dataset of the next experiments, since they might prevent fitting other measurements. For cis-DCE about 1.07 % of the total composite sensitivity was retained by the extreme outlier depicted in Fig. 7.9(a) (piezometer 18 on 12/07/2002), for VC the three outliers between 3000 and 4000 ppbs (Fig. 7.9(b)) had a smaller influence on the inversion (the three measurements sum up to 0.17 % of the total composite sensitivity). For DCA only the measurements collected at piezometer 18 are sensitive and the simulated value around 48000 ppb in Fig. 7.9(c) retains the 0.03% of the total composite sensitivity. Based on this analysis, only the extreme cis-DCE outlier was excluded from the calibration dataset. In the next experiments it was found that the exclusion of this single outlier improved the simulation, allowing a better reproduction of the cis-DCE concentrations in the pasture area.

7.5.4. Inversion with equal weights

After determining a plausible initial parameter vector via manual calibration and removing the cis-DCE outlier, the parameters listed in Tab. 7.1 (degradation, source and transport parameters) were automatically estimated, specifying equal observation weights within each measurement set. The weights were calculated to ensure that the initial contribution of each set to the measurement objective function Φ_m to be equal to the number of measurements. The reason of this choice can be derived from the negative log likelihood of the posterior density function [143]:

$$-\log(p(\phi|Y)) \propto \sum_{j=1}^m \sum_{i=1}^{n_j} \left[\log \sigma_{ij} + \frac{\varepsilon_{ij}^2}{2\sigma_{ij}^2} \right] \quad (7.23)$$

Table 7.1.: Estimated model parameters with initial guesses

Parameters	Manual adjust- ment	Notes
Flow model		
Hydraulic conductivity (m d^{-1})	0.194	Average value from all slug tests
Specific yield (-)	0.1	Plausible value for sandy clay soil [69]
Precipitation fraction	1	Initially, it is assumed that all precipitation infiltrates
Drain 1, 2 and 3 conductances (m d^{-1})	70, 10, 10	Drain widths are estimated as 7, 7 and 1 m , drain beds 0.1 m thick with hydraulic conductivities equal to 1 m d^{-1}
Degradation parameters ^a		
μ_{PCE} (mol $\text{mol}_{\text{SUB}}^{-1} \text{d}^{-1}$)	1.80e-8	Adjusted from the batch experiment estimate (7.3e-8)
μ_{TCE} (mol $\text{mol}_{\text{SUB}}^{-1} \text{d}^{-1}$)	1.00e-7	Adjusted from the batch experiment estimate (1.89e-5)
$\mu_{cis-DCE}$ (mol $\text{mol}_{\text{SUB}}^{-1} \text{d}^{-1}$)	3.00e-9	Adjusted from the batch experiment estimate (4.34e-6)
μ_{VC} (mol $\text{mol}_{\text{SUB}}^{-1} \text{d}^{-1}$)	1.71e-9	Adjusted from the batch experiment estimate (9.55e-7)
$\mu_{TCA-DCA}$ (mol $\text{mol}_{\text{SUB}}^{-1} \text{d}^{-1}$)	6.14e-1	Derived from Suarez and Rifai [142]
$\mu_{TCA-DCE}$ (d^{-1})	1.70e-3	Maximum rate coefficient reported in Suarez and Rifai [142]
μ_{DCA} (mol $\text{mol}_{\text{SUB}}^{-1} \text{d}^{-1}$)	5.41e-3	Derived from Suarez and Rifai [142]
$f_{S/TVOC}$ (-)	600	See sec. 6.3.4
Source parameters		
k_{ta} (d^{-1})	1e-4	Higher mass transfer coefficients caused solver failure (fast mass transfer)
$k_{PCE-SOU,1}$ ($\text{kg d}^{-1} \text{m}^{-2}$)	1.66e-5	Manual calibration (11.05) ^b
$k_{TCE-SOU,1}$ ($\text{kg d}^{-1} \text{m}^{-2}$)	3.94e-5	Manual calibration (26.26)
$k_{TCA-SOU,2}$ ($\text{kg d}^{-1} \text{m}^{-2}$)	5.34e-4	Manual calibration (355.48)
$k_{PCE-SOU,3}$ ($\text{kg d}^{-1} \text{m}^{-2}$)	1.65e-6	Manual calibration (1.10)
$k_{TCE-SOU,3}$ ($\text{kg d}^{-1} \text{m}^{-2}$)	1.31e-6	Manual calibration (0.88)
Transport parameters		
DSP (m) ^c	5	Manual calibration
TRPT (-)	0.1	Manual calibration
TRPV (-)	0.01	Manual calibration
porosity (-)	0.1	Plausible value for sandy clay soils [112]
Total number of parameters	369	

^a degradation parameters were estimated only for piezometer 23, for the other piezometers biodegradation was negligible. Suarez and Rifai [142] reported that degradation rates in lab conditions are higher than field conditions, with difference between 1 to 2 orders of magnitude.
^b in brackets the total released mass at each source is reported (in kg). ^c DSP indicates the longitudinal dispersivity, TRPT the horizontal transverse to longitudinal dispersivity ratio and TRPV the vertical transverse to longitudinal dispersivity ratio.

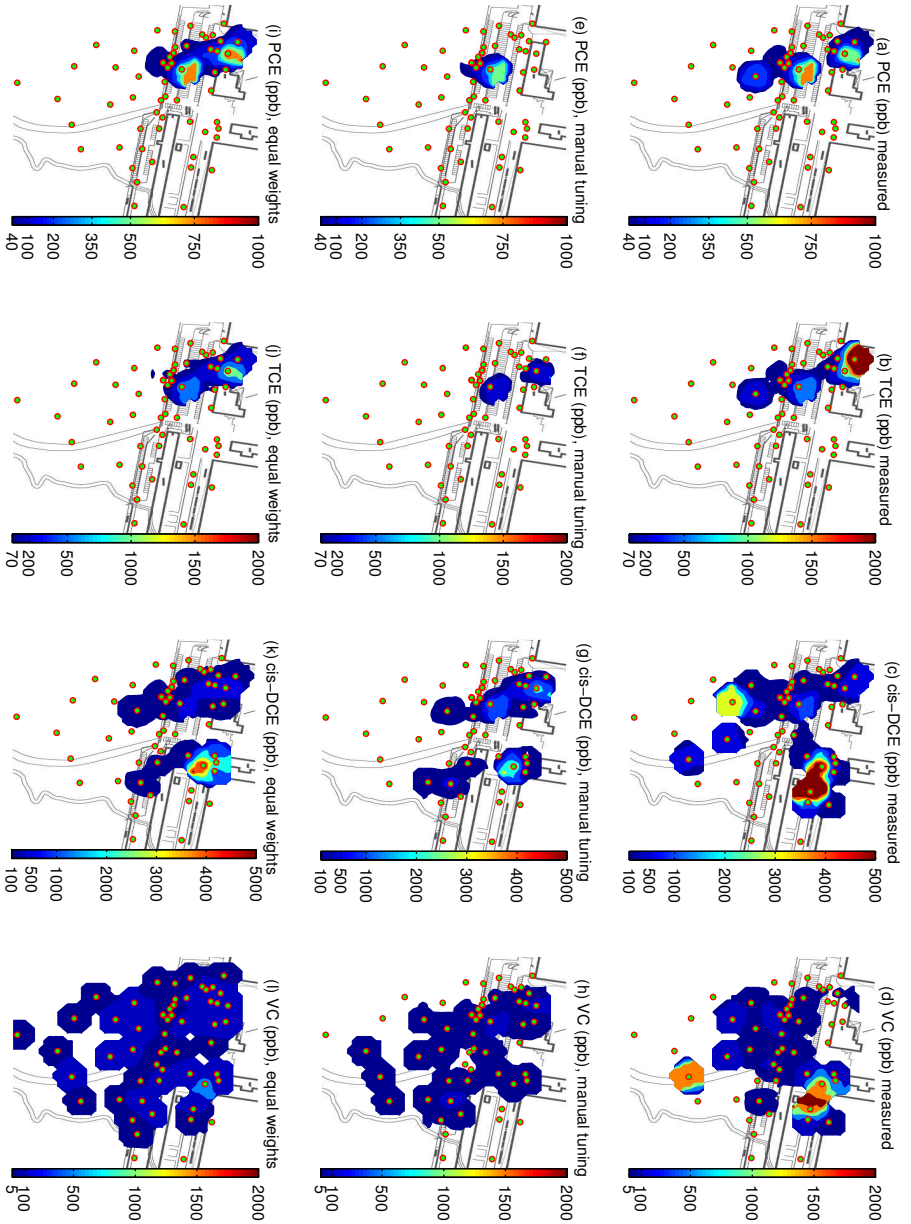


Figure 7.8: Top row (figures (a-d)): average measured concentrations. Central row (figures (e-h)): average simulated concentrations with manually tuned parameters. Bottom row (figures (i-l)): average simulated concentrations after inversion with equal observation weights.

where m is the number of measurement sets and n_j is the number of measurement in the set j . In Eq.7.23 it is assumed that the residuals are normally distributed, not correlated and uniform prior density functions of model parameters ϕ and variances σ_{ij}^2 . By deriving Eq.7.23 respect the variance σ_{ij}^2 and equating it to zero (necessary condition for the maximization of $-\log(p(\phi|Y))$) it can be shown that the variance σ_{ij}^2 must be equal to the squared residual ε_{ij}^2 . If all variances are assumed equal within each measurement set, the variance of the set j can be estimated as:

$$\sigma_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} \varepsilon_{ij}^2 \quad (7.24)$$

assuming the observation weights of the set j inversely proportional to the standard deviation ($w_j = 1/\sigma_j$), the contribution of each measurement set at the starting of the optimization is equal to the number of measurements n_j contained in that set. This procedure can be used when the residual variances are significantly affected by unknown model errors/model inputs. However, the residual mean is assumed zero, which might not hold in presence of strong model errors.

The estimated degradation, source and transport parameters are shown in Fig.7.10. As can be seen, most of the estimated degradation parameters (white bars) are lower than the manually tuned parameters (black bars), which were already lowered from the batch estimates (Tab.7.1) for the simulation of a realistic contaminant concentrations. This result confirms the findings of

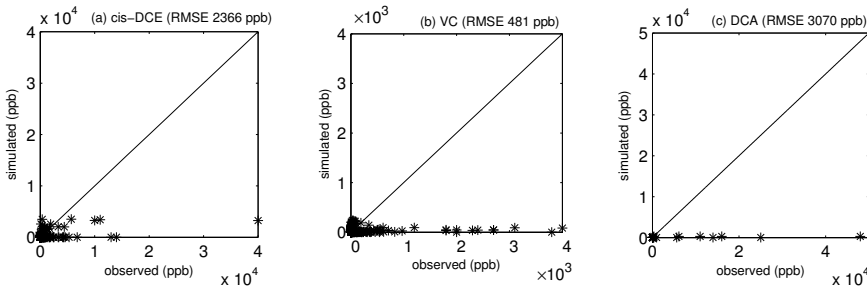


Figure 7.9.: Cross plots for cis-DCE, VC and DCA after manual tuning of the biodegradation rates (fixed flow field).

Suarez and Rifai [142], where first-order dechlorination rate coefficients determined from laboratory studies were almost two orders of magnitude higher than those determined from in-situ studies. Extremely low values were obtained for μ_{PCE} and μ_{VC} ($2.59\text{e-}10$ and $8.79\text{e-}13$ mol mol_{SUB}⁻¹ d⁻¹). Only the $\mu_{TCA-DCE}$ coefficient was increased, to allow for simulation of the high cis-DCE measurements at source 2 (Fig. 7.8(c)), whereas μ_{DCA} was strongly decreased to correct the initial DCA underestimation (result not shown).

The mass transfer coefficient k_{la} reached its upper bound, fixed at $1\text{e-}4$ d⁻¹. The largest $k_{k-SOU,i}$ release rate was estimated at source 2, with a released TCA mass of 10140 kg, whereas for the other sources the estimated released mass was below 40 kg. The lowest estimation of the total released mass (not accounting for biodegradation and dilution processes) was obtained integrating the total contaminant concentration over the entire aquifer volume. The results of this integration yield a contaminant mass of 22 kg, far below the one obtained integrating the initial release rates over the spillage period.

A low value of the $f_{S/VOC}$ parameter was estimated (5.06), indicating a small consumption of the organic substrate by dechlorinators and constant degradation rates coefficients μ in the biodegradation model, as assumed in most field scale models [30, 28]. On the other hand, this contradicts the correspondence found in the batch experiments between the low organic content at piezometer 9, 10, 21 and 16 ($< 0.3\%$) and the observed low biodegradation rates, suggesting the need for more investigations of the organic carbon consumption mechanism at the site.

Large values of the longitudinal dispersivities (56.11 m), transverse and vertical dispersivity ratios were also estimated, yielding a wide spread of the VC plume (Fig. 7.11(c)). From the figure it is evident that the longitudinal and transversal dispersivities must be reduced. Moreover, the estimated effective porosity (0.57) was above the total porosity estimated from core samples (0.4). These unrealistic values point to the need for an estimation procedure which uses the additional information contained in the concentration measurements to constrain the flow parameters affecting solute transport (such as hydraulic conductivity).

In Fig. 7.8, maps of average observed concentrations (top row) are compared with maps of average simulated concentrations (bottom row). Simulated PCE concentrations at source 1 were overestimated and TCE concentrations underestimated. Only a few PCE and TCE measurements were taken close to source 1 (three for PCE and three for TCE before the barrier installation) and their influence on the estimation was small, amounting to 8.6% of the total PCE and TCE composite sensitivities. Close to source 3, more measurements

were taken, with a larger contribution to PCE and TCE composite sensitivities (about 64.11%). Since the measurements at source 3 are more sensitive and matching the TCE concentrations at source 1 would cause an overestimation of TCE concentrations at source 3, the inversion algorithm fits TCE and PCE measurements at source 3 more closely.

The most mobile contaminants (having lower K_{d-k} distribution coefficient) are cis-DCE and VC, as shown by the larger dimensions of the plumes (Fig. 7.8(k-l)). The cis-DCE plume originating from sources 1 and 3 is correctly oriented although the concentration at piezometer 15 in the pasture is underestimated. Most of the cis-DCE composite sensitivity was retained by the high concentration measurements close to source 2 (30.62%) and source 3 (14.18%). It was difficult to match these measurements simultaneously to those collected in the pasture area (amounting to 8.98% of the total cis-DCE composite sensitivity).

As in the manual tuning case (central row) the cis-DCE and VC plumes originating from source 2 are oriented towards the second barrier segment (Fig. 7.8(k-l)), whereas the measurements suggest a more south-east orientation, with high cis-DCE concentrations at piezometer 19. Also in this case the high VC concentrations at piezometer 23 were not reproduced. A correct simulation of the VC pollution at piezometer 23 is important because VC is the most mobile and toxic chlorinated compound. From these results it is clear that a separate calibration of flow and reactive transport parameters does not provide a satisfactory simulation of the cis-DCE and VC plumes.

7.5.5. Inversion with weights proportional to the inverse of the concentrations

In this last experiment with fixed flow field, the observation weights were specified inversely proportional to the measurements ($w_{ij} = 1/(0.1C_{ij})$). This approach can lead to strongly biased parameters because it assumes the residual variances to be determined solely by the analytical measurement error. Unaccounted for measurement error might also be due to contaminant volatilization during the field sampling.

Using these observation weights, the concentrations of all contaminants were underestimated because high weights were assigned to low concentration values. A clean aquifer was simulated in the pasture area, due to the extremely low mass transfer rate k_{la} ($1.15\text{e-}6 \text{ d}^{-1}$). Ideally the weights assigned to each measurement should be estimated from the data when model error is expected, or from preliminary estimates of the total residual variance as reported above.

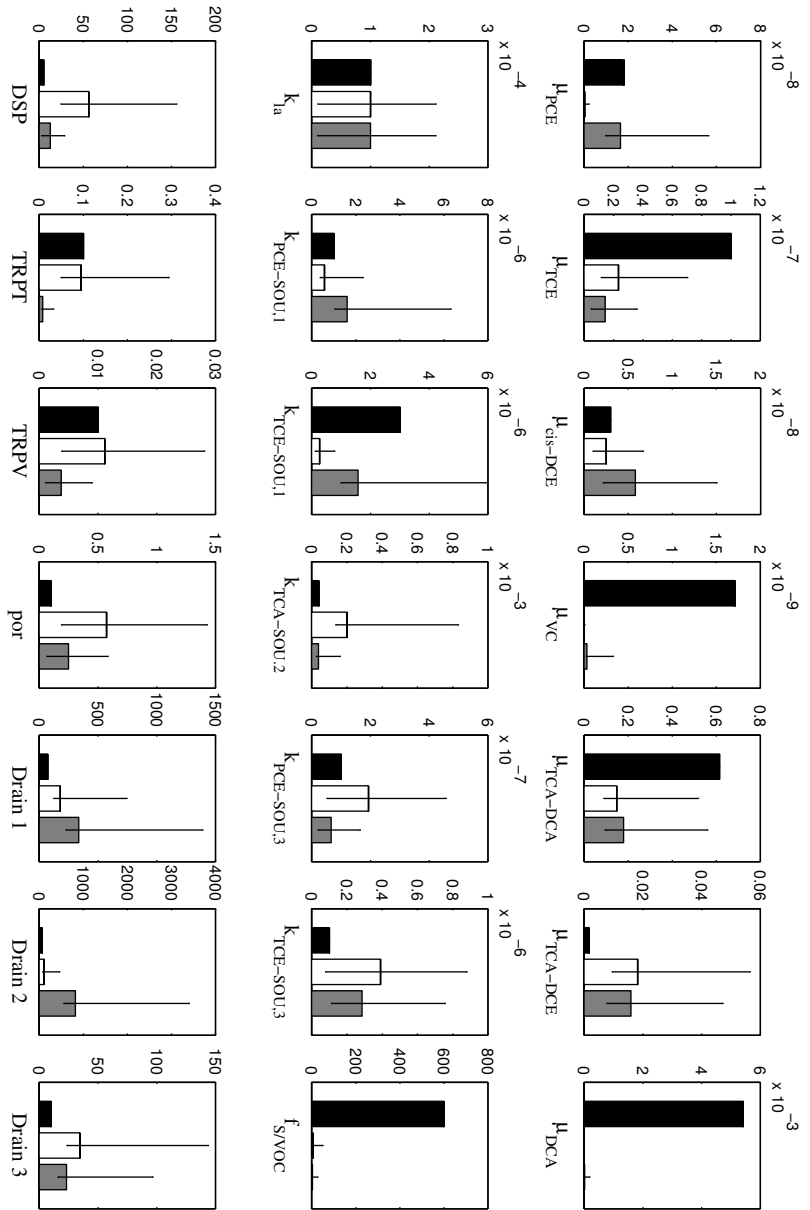


Figure 7.10: Initial parameter values (black bars), after separate optimization of flow and transport parameters (white bars) and after joint optimization (gray bars). Vertical lines indicate 95% confidence intervals, estimated from the posterior covariance matrix Eq.7.15 (2 standard deviations were added and subtracted). Parameter units are reported in Tab. 7.1.

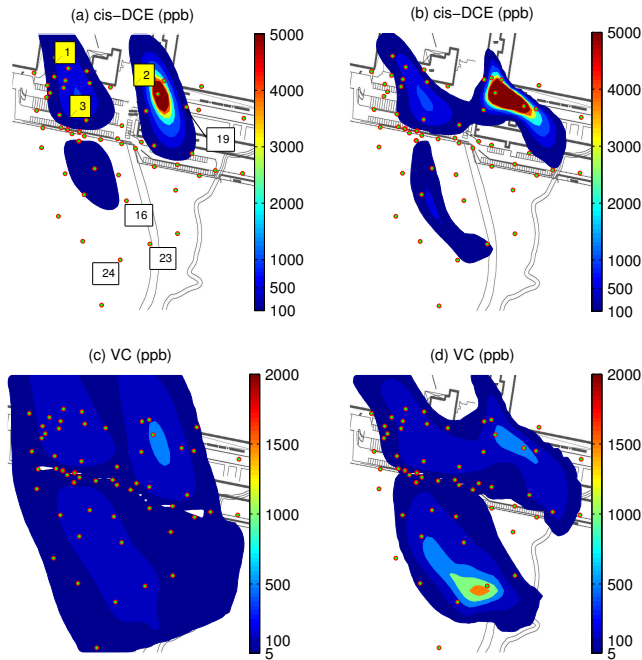


Figure 7.11.: Contaminant plumes on 1 January 2012 (a) cis-DCE with separate estimation of flow and transport parameters and (b) with joint optimization. (c) VC with separate estimation of flow and transport parameters and (d) with joint optimization.

7.6. Joint optimization of flow and reactive transport parameters

It is well recognized that concentrations measurements contain information about the flow field [149, 158]. In this experiment all available concentration measurements and the 2011 groundwater measurements were used to infer flow and reactive transport parameters jointly (experiments 4 in Tab.D.1). When dealing with different measurement types in presence of model errors, the observation weights should be determined together with the model parameters. For the field scale application it was not possible to apply a Monte Carlo method as in chapter 5, due to the large number of model parameters and the

long model run time. In this case an iterative procedure was used to adjust the observation weights honoring the regularization constrains.

The procedure is shown in Fig. 7.12. In the preliminary part of the procedure (top box), the total residual variances σ_j^2 are estimated with Eq.7.24, using the raw residuals ε_{ij} obtained running the model with the manually tuned parameters. The square of the initial observation weight $w_{0,j}^2$ is assumed equal to the inverse of the variance σ_j^2 . The target objective function for each set j ($\Phi_{m,j}^l$) is calculated from the user specified target $RMSE_{t,j}$ (as for the flow inversions, such target is subjective and depends on the expected fit). The values $GFACT_j$ and $OLDGFACT_j$ are set to 1, the first representing the ratio of the current squared weight to the initial weights ($w_{IT,j}^2/w_{0,j}^2$) and the second a variable to save the current $GFACT_j$ value.

The optimization routine is than started (bottom box). At the first step the measurement objective function $\Phi_{m,j,IT}$ with the initial weights $w_{0,j}^2$ is calculated, using the weights and the model results of the previous iteration $IT - 1$. At the second step the $GFACT_j$ ratio is adjusted such that is increased if $\Phi_{m,j,IT}$ is greater than the user specified target $\Phi_{m,j}^l$ and decreased otherwise. At step 3 the $GFACT_j$ are scaled such that after the adjustment of the weights the sum of $\Phi_{m,j}^l GFACT_j$ equals to Φ_m^l specified at the starting of the optimization routine. At step 4 the ratio $DTEMP_j = w_{IT,j}^2/w_{IT-1,j}^2$ is calculated and used in the step 5 to calculate the new weights w_{IT}^2 . In the same step the current $GFACT_j$ is saved in $OLDGFACT_j$. At step 6 the model is run $p + 1$ times to fill the Jacobian matrix, with p equal to the number of parameters. At step 7 the the optimal weight factor β^2 is estimated, testing different parameters upgrades and assuming model linearity (as explained in sec. 7.1). At step 8 the global objective function Φ_{IT} is formed and different λ parameters are tested to minimize the value of Φ_{IT} . At the end of the lambda search the minimized Φ_{IT} is saved and a new iteration started. The best parameter set is the one that provides the lowest Φ_{IT} among the different iterations. In PEST the internal reweighing is such that $GFACT_j$ is never higher or lower by a factor more than 128.

Moreover, the increase/decrease per iteration is never greater than 2, used to ensure a stable regression (large changes in the weighting can cause large changes in the objective function surface and poor convergence). The reweighing approach in non linear parameter estimation was used in several studies (Barlebo et al. [6], Carrera and Neuman [20], Barth and Hill [7] among the others), mostly with two measurement sets (head and concentration data) and different reweighing strategies (for example scaling the concentration weights with the ratio head/concentration measurement objective function).

```

1. Estimate  $\varepsilon_{i,j}$  with  $i = 1, \dots, n_j$  and  $j = 1, \dots, m$  ( $n_j$  is the number observations in the set  $j$  and  $m$  is the number of sets)

For  $j = 1 : m$ 

2.  $w_{0,j}^2 = 1/(\sum_i \varepsilon_{i,j}^2 / n_j)$ ,  $\Phi_{m,j}^l = n_j w_{0,j}^2 \text{RMSE}_{i,j}^2$ ,  $\Phi_m^l = \sum_j \Phi_{m,j}^l$ ,  $\Phi_{m,j,0} = w_{0,j}^2 \sum_i \varepsilon_{i,j}^2$ 
   GFACTj=1, OLDGFACTj=1
end

For IT = 1 : MIT (MIT is the maximum number of iterations)

   For  $j = 1 : m$ 

   1.  $\Phi_{m,j,IT} = w_{IT,j}^2 \sum_i \varepsilon_{i,j}^2 / \text{GFACT}_j$ 
   2.  $\text{GFACT}_j = \text{GFACT}_j \Phi_{m,j,IT} / \Phi_{m,j}^l$ 
   3.  $\text{GFACT}_j = \text{GFACT}_j (\Phi_m^l / \sum_j (\Phi_{m,j,IT}^l \text{GFACT}_j))$ 
   4.  $\text{DTEMP}_j = \text{GFACT}_j / \text{OLDGFACT}_j$ ,  $\text{OLDGFACT}_j = \text{GFACT}_j$ 
   5.  $w_{IT,j}^2 = w_{IT-1,j}^2 \text{DTEMP}_j$ 

   end

6. Fill jacobian matrix, run the model  $p + 1$  times
7. Calculate regularization weight factor  $\beta^2$ 
8.  $\Phi_{m,IT} = \sum_j w_{IT,j}^2 \sum_i \varepsilon_{i,j}^2$ ,  $\Phi_{i,IT} = \sum_j w_{IT,j}^2 \varepsilon_{i,j}^2$ ,  $\Phi_{IT} = \Phi_{m,IT} + \beta^2 \Phi_{i,IT}$ , minimize  $\Phi_{IT}$ 
9. save  $\Phi_{IT}$ ,  $\text{RMSE}_{IT,j}$ 

end

The best solution is  $\min(\Phi_{IT})$ 

```

Figure 7.12.: Iterative procedure to adjust the observation weights honoring the regularization constrains.

As in the flow inversions the individual targets had a strong influence on the final parameter estimations. Low $\Phi_{m,j}^l$ (low $\text{RMSE}_{t,j}$) make the inverse problem unstable (hydraulic conductivity fields with high variances and extreme ranges) because the Φ_r term in Eq.7.21 becomes underweighted. To determine the maximum achievable fit, a first test with low $\Phi_{m,i}^l$ targets was performed ($\text{RMSE}_{t,j}$ at 200 ppb for each contaminant, Tab.7.2). The joint optimization was then repeated with higher $\Phi_{m,i}^l$ as deduced from the first test. For groundwater levels the target $\Phi_{m,j}^l$ was specified at 420, which is equivalent to keep a RMSE equal to 0.15 m (as obtained from the flow optimization).

In the joint optimization the degradation, source and transport parameters re-

Table 7.2.: RMSEs and target $\Phi_{m,i}^l$ for each inversion test and measurement set. RMSEs for concentrations are expressed in ppb and for head measurements in meters. $\Phi_{m,i}^l$ values are dimensionless.

	PCE	TCE	cis-DCE	VC	TCA	DCA	Head
Manual tuning RMSEs	182	351	1415	481	433	3068	0.150
Inversion with fixed flow field RMSEs	155	310	1366	454	114	2672	0.150
Low regularization test							
$\Phi_{m,i}^l$	448	121	8	26	80	2	420
Achieved Φ_m	323	267	352	311	29	331	347
Target RMSEs	200	200	200	200	200	200	0.150
Achieved RMSEs	170	296	1370	440	120	2890	0.137
High regularization test							
$\Phi_{m,i}^l$	448	273	188	160	90	159	420
Achieved Φ_m	388	317	183	201	26	160	525
Target RMSEs	200	300	1000	300	200	2000	0.150
Achieved RMSEs	186	323	989	354	115	2008	0.168

ported in Tab. 7.1 were estimated with the recharge and hydraulic conductivity parameters. From preliminary tests it was found that the composite sensitivities of the specific yield pilot points were on the average one order of magnitude lower than the sensitivities of hydraulic conductivity pilot points and 300 times lower than the highest parameter composite sensitivity. Such low composite sensitivities usually belongs to parameters difficult to estimate for the given dataset and weighting scheme [?]. Therefore, the specific yield field was kept fixed during the joint optimization process. The starting biodegradation, source and transport parameter values were those estimated by manual calibration (sec. 7.5.3).

7.6.1. Joint optimization results

Tab. 7.2 reports the achieved RMSEs for the tests with low and high regularization constrains. As can be seen, only for PCE, TCE and groundwater levels the achieved RMSEs in the low regularization test are lower than those obtained for the high regularization test. Moreover, the resulting hydraulic conductivity field was strongly heterogeneous with extreme variations over short distances and a peak value of 17.48 m d^{-1} , which was deemed not realistic.

For the high regularization test the RMSEs for cis-DCE, VC and TCA were strongly reduced, mainly due to the improved orientation of the contaminant

plume originating from source 2. Since a more realistic hydraulic conductivity field and lower RMSEs for the most mobile contaminants were obtained in the high regularization test, the optimized parameter provided by this test were considered as the best values and the high regularization test is thereafter simply referred to as joint optimization.

In the joint optimization the average VC concentrations at piezometer 23 were better reproduced compared to the other optimization tests (Fig. 7.13(l)). High VC concentrations in this area were also caused by a negative recharge flux, as larger evaporative fluxes are simulated in areas with shallow groundwater table. This was verified using the ZONEBUDGET program [49]. The amount of rainwater infiltrating in the area close to piezometer 23 during the period 1957-2000 (first steady state period) amounts to 160.22 mm y^{-1} whereas simulated evaporation is 278.32 mm y^{-1} , yielding a negative imbalance of $-118.11 \text{ mm y}^{-1}$. This indicates that the downslope portion of the pasture functions as a groundwater discharge area, a finding that should be confirmed with in-situ infiltration measurements.

The degradation parameters obtained in the joint optimization were similar to those obtained for the fixed flow field inversion (compare white and gray bars in the top row of Fig. 7.10), except for an increased μ_{PCE} coefficient, which caused an underestimation of PCE concentrations at source 1 and 3 (Fig. 7.13(i)) and an increased RMSE for PCE (Tab. 7.2). The estimated released mass at source 3 decreased, with a total estimated DNAPL mass of 8 kg. This low estimated released mass might indicate an accidental spillage not related to a particular production activity, as suggested by historical information on the production areas. A compressor test area is reported at source 1 and confirmed by higher release coefficients, with a total released DNAPL mass of 79 kg. The release rate at source 2 was reduced compared to the inversion with fixed flow field (from 10140 kg to 860 kg), with a narrower confidence interval. In the inversion with fixed flow field, $k_{TCA-SOU,2}$ was increased to match the high concentrations measured at piezometer 19, despite the southward orientation of the cis-DCE plume (Fig. 7.11(a)).

The joint optimization provided lower dispersion parameters compared to the separate flow and transport optimizations, due to a better representation of the hydraulic conductivity heterogeneity (Fig. 7.10). Similar results were also found by Tonkin and Doherty [148] and it agrees with a lowering of mechanical dispersion as more heterogeneity is explicitly accounted for in the model [82, 18]. A more realistic value of the effective porosity (0.25) was also estimated in the joint optimization. The effect of the reduced dispersion coefficients (both longitudinal and transverse) can be seen in Fig. 7.11, where the VC plume in

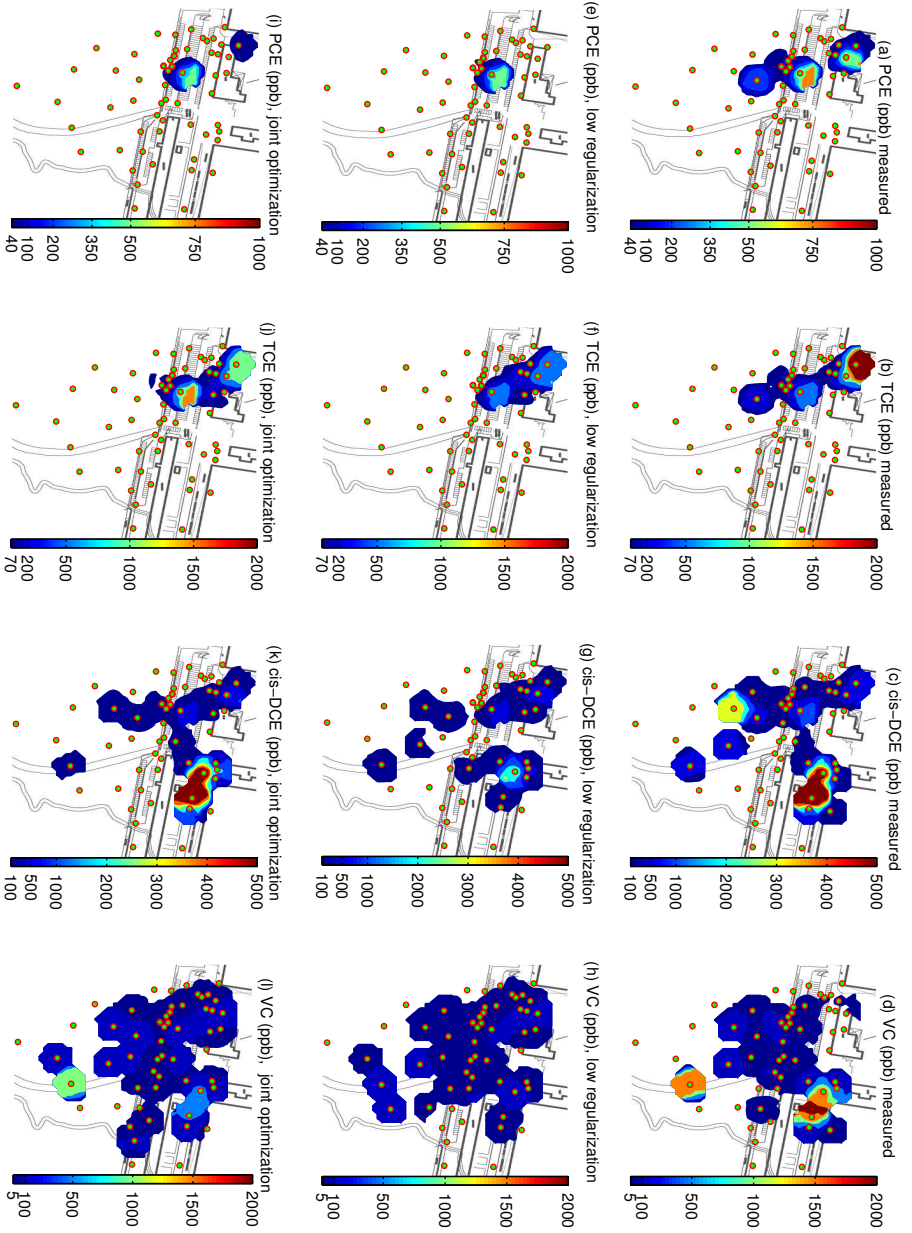


Figure 7.13: First row (figures a-d) average measured concentrations, central row (figures e-h) average simulated concentrations for the low regularization test, and bottom row (figures i-l) average simulated concentration for the high regularization test (“joint optimization”).

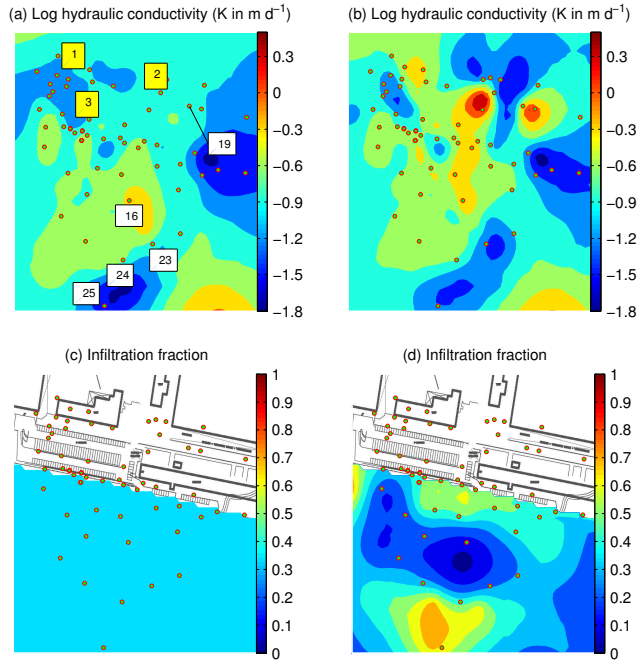


Figure 7.14.: (a) Hydraulic conductivity estimated using groundwater heads only and (b) from the joint optimization. (c) Infiltration fraction estimated using head data and (d) from the joint optimization.

the joint optimization case (Fig. 7.11(d)) is less spread out compared to the case with the separate estimation of flow and transport parameters (Fig. 7.11(c)).

The hydraulic conductivity fields in the polluted area at the start and at the end of the joint inversion are shown in Fig. 7.14(a-b). As can be seen from the figures, the low hydraulic conductivities formations identified from head data at piezometer 25 and on the right part of the barrier were retained in the joint inversion, with point-scale hydraulic conductivity estimates similar to those obtained from slug-in tests (Fig. 7.5). Most of the differences between the two fields are located in the source areas, with maximum hydraulic conductivity values close to source 2, to correct the plume direction towards piezometer 19 and the central part of the barrier (Fig. 7.11(c)). These estimates should be supported by additional field measurements because they might also compen-

sate for an incorrect localization of the second source.

Fig. 7.14(c) shows estimated infiltration rates from head only measurements and Fig. 7.14(d) in the joint optimization. As can be seen, in the inversion with head only data the homogeneous regularization constraint is respected (with a value of 0.36), whereas in the joint inversion a heterogeneous field was estimated, with the lowest value between piezometer 16 and 24 (0.062), the latter at the fringe of the VC plume (Fig. 7.11(d)). This low value indicates that almost no water infiltrates and seems to represent an inversion artifact rather than a realistic infiltration rate at the site, because larger recharge rates are expected for deep water tables, as opposed to the obtained results, where a lower infiltration fraction was estimated in zones of deeper groundwater table (about 1 m). Infiltration measurements are required to corroborate the estimations of the recharge parameters.

7.6.1.1. Posterior parameter and prediction uncertainty

The prior covariance matrix C required for the estimation of the posterior parameter covariance $\tilde{C}$ was reconstructed by assembling different sub matrices. For hydraulic conductivity, specific yield and recharge the covariance matrices were derived from the corresponding variograms. The other model parameters were assumed to be uncorrelated and the variances (corresponding to the diagonal elements of C) estimated from plausible variation intervals and assuming the prior distribution of model parameters to be log normal. The biodegradation rate coefficients were assumed to vary within two orders of magnitudes whereas the drain conductances, porosity and dispersivities values were assumed to vary within one order of magnitude. Dividing the variation ranges by 4 and squaring the obtained values, the estimated prior variance of the degradation coefficients and physical parameters were 0.25 and 0.0625 respectively.

Two model parameters posterior covariances were estimated with Eq. 7.15, one at the optimal parameter vector obtained after the separate optimization of flow and transport parameters and another one at the optimal parameter vector obtained after the joint optimization. The posterior parameter standard deviations were then compared to assess the reduction of prior parameter uncertainty when assimilating head and concentrations jointly against the separate assimilation of head and concentrations.

Fig. 7.15 shows that prior standard deviations for hydraulic conductivity (0.5) and infiltration fraction (0.251) were reduced over larger areas when head and concentration measurements were used, as opposed to using only head

measurements.

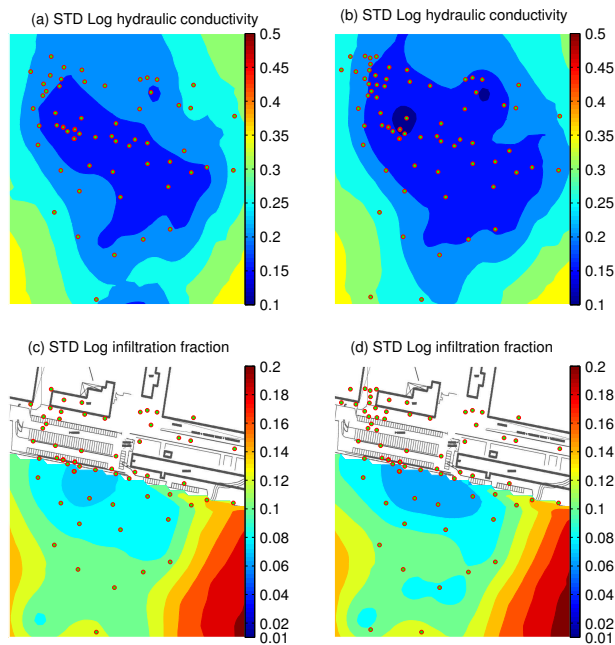


Figure 7.15.: (a) Standard deviation of log hydraulic conductivity with separate estimation of flow and transport parameters and (b) with joint optimization. (c) Standard deviation of log infiltration fraction with separate estimation of flow and transport parameters and (d) with joint optimization.

The maximum reduction of the prior standard deviation for hydraulic conductivity was 72.8 % for separate optimization of flow and transport parameters and 73% for the joint optimization. Corresponding percentages for the infiltration fraction are 71.6% and 74.7%. Interestingly, when using only head data the reduction of the prior hydraulic conductivity standard deviation is localized in the pasture area, where elevated fluctuations of the groundwater levels were observed (thus containing more information about the hydraulic parameters). When head and concentrations are used jointly, the reduction of the prior hydraulic conductivity standard deviation also occurs in the factory area, clearly delineating source 2 and 3.

Model predictions at 2035 were simulated only using the parameters inferred

with the joint optimization. Fig. 7.16 and Fig. 7.17 show the concentration time series for some selected piezometers along the main plume direction. In these predictions the PRB is fully reactive, as supported from in-situ core samples.

For cis-DCE (Fig. 7.16) the temporal variation in measured concentrations was not reproduced and the best simulations that could be obtained were concentration time series passing through the measurements (solid lines in piezometers 3, 20, 21 and 23). This could be due to the use of temporally averaged (weekly) values for precipitation and evaporation. However, significant correlations patterns between the measured concentrations and the net precipitation minus evaporation time series at the different piezometer locations were not found from preliminary statistical analysis. Part of the observed variation might also be due to unaccounted for measurement noise (for example due to contaminant volatilization during field sampling). Finally, model errors are likely to be important as well, e.g. due to numerous assumptions made in the advective-dispersive models [82] and lack of information about the released contaminant mass. For piezometer 16 the concentrations were consistently underestimated, this piezometer being located close to the cis-DCE plume fringe (Fig. 7.11(b)). As mentioned before, most of the cis-DCE sensitivity belongs to the measurements collected at source 2 and a good cis-DCE fit at source 2 and in the pasture area was difficult to achieve. The dashed lines in Fig. 7.16 and Fig. 7.17 represent the prediction uncertainty due to parameter and residual uncertainty (adding and subtracting to the simulations two standard deviations, calculated from Eq. 7.18). As can be seen the amplitude of the linear prediction bound depends on the piezometer location. For piezometer 20, which is located 6 cm downgradient from the barrier, prediction uncertainty is negligible after the removal of contaminants in the barrier. For piezometer 21, prediction uncertainty decreases as the contamination super pass the piezometer (the sensitivity of the model prediction y in Eq. 7.17 becomes small).

The VC time series are shown in Fig. 7.17. In this case the model bias in the factory area (piezometer 3) is larger compared to cis-DCE. As can be seen from Fig. 7.13(d), measured VC concentrations in the factory area at source 1 and 3 were lower than the corresponding simulated concentrations shown in Fig. 7.13(l). Using release rates $k_{k-SOU,i}$ that vary through time, with more contaminant release at the start of the simulation (1957) and less release close to 1996, might correct the overestimation in the factory area but also introduce additional parameters difficult to determine from the calibration dataset. Good results were obtained for piezometer 23, with the average observed VC

concentrations more closely reproduced. Simulated VC concentrations indicate that in the year 2035 VC concentrations might be as high as the ones measured during 2002-2011, due to the small effect of contaminant dilution and degradation. As for cis-DCE, the measured concentrations do not show a clear seasonal pattern.

Summarizing, with the joint optimization the plume configuration was greatly improved and the average TCE and cis-DCE patterns reproduced. For VC, which is the most mobile specie, a large transverse spreading and elevated concentrations at piezometer 23 were simulated, caused by a small but constant net evaporative flux during the first steady state stress period.

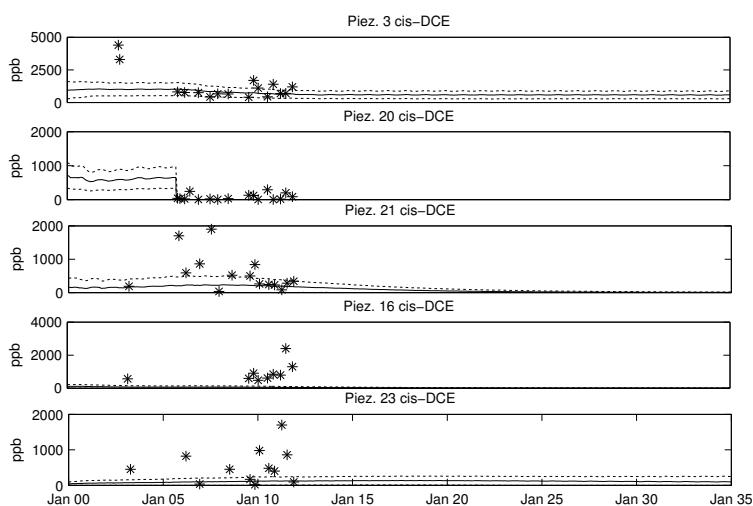


Figure 7.16.: cis-DCE concentrations along transect defined by piezometers 3-23. The solid lines correspond to the model simulation with the optimal parameter values, dashed lines defines the 95% uncertainty bounds estimated using Eq.7.18 and asterisks concentration measurements.

7.7. Geochemical scenarios

Geochemical scenarios until 2035 were simulated in the local geochemical model (Fig. 6.3) using the flow and transport parameters estimated from the joint optimization. The geochemical model developed for the column experi-

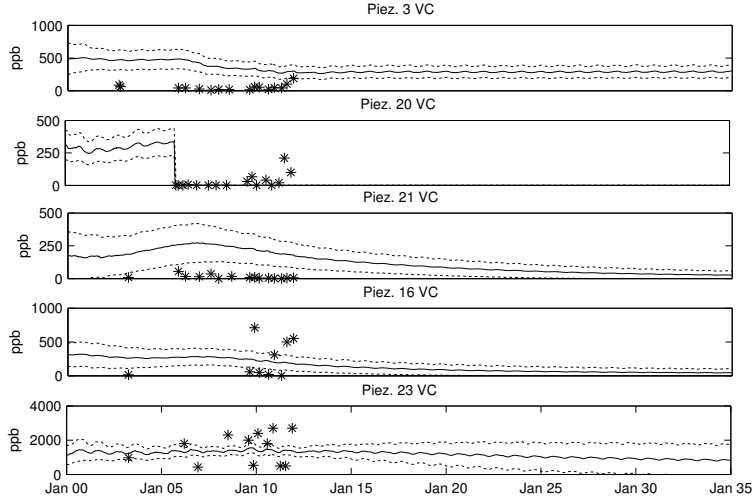


Figure 7.17.: VC concentrations along transect defined by piezometers 3-23. The solid lines correspond to the model simulations with the optimal parameter values, dashed lines define the 95% uncertainty bounds estimated using Eq.7.18 and asterisks concentration measurements.

ment (chapter 3) was integrated with the biodegradation model. Beside contaminants and substrate, the geochemical model includes Ca^{2+} , Cl^- , Na^+ , Fe^{2+} , CO_3^{2-} , H^+ and $\text{H}_{2(\text{aq})}$. The geochemical simulation was performed with PHT3D [127]. Similarly to PHAST used in chapter 3 for the simulation of the column experiment, PHT3D employs PHREEQC to solve kinetic and equilibrium reactions, providing identical simulations of the laboratory column experiment. PHT3D was chosen for the field scale simulation over PHAST because it can use the flow field computed by MODFLOW.

In PHT3D the computational mesh was refined to correctly include the barrier, increasing considerably the number of time steps. Moreover, PHREEQC requires more CPU time to integrate the chemical reactions compared to the RT3D simulations, due to: (1) increased number of chemical species included in the model, (2) numerical solution of coupled kinetic and equilibrium reactions, and (3) inclusion of the iron corrosion reaction that has a relatively large rate compared to the contaminant degradation reactions (stiffness). For these reasons, a single geochemical simulation of the PRB required about 35 days on a single CPU core (Tab. D.1) and could be performed only once. The goals

of the geochemical scenarios are:

1. Assess the long term impact of the barrier installation on the contaminant plume: this is accomplished by comparing scenarios with and without PRB.
2. Assess the long term barrier performance: this is accomplished estimating the reactivity decline over time.

In these geochemical scenarios the main assumptions are:

1. Declining in reactivity is only caused by carbonate minerals, as assumed in chapter 3.
2. The porosity of the barrier is constant (no porosity updating). This assumption provides the lowest reactivity estimate (larger residence times in the PRB with larger accumulation of precipitates).
3. Hydrogen produced by iron corrosion is not used by dechlorinations and is assumed to volatilize once its solubility limit is reached. Such assumption provides the largest contaminant concentrations in the pasture area.

The geochemical scenario without PRB starts from 01-01-1957 assuming a clean aquifer and an homogeneous concentration field for inorganic ions, as derived from the concentrations measured in the piezometers upstream the barrier and reported in Tab. 7.3 (inorganic concentrations were available only after the barrier installation). The groundwater levels calculated by MODFLOW were interpolated along the edges of the PHT3D model (Fig. 6.3). Along these edges groundwater can exit from the model domain or enter with a specified concentration. In the latter case the influent concentration was set equal to the average aquifer concentration (Tab. 7.3). In Tab. 7.3 the composition of the infiltrating rainwater is also reported.

The heterogeneous aquifer properties fields (hydraulic conductivity, specific yields and recharge fractions) were interpolated in the PHT3D model grid from the pilot point values obtained from the joint optimization. The rotation procedure was developed to include a “fine” barrier discretization and simulate the groundwater flow and transport similarly to the unrotated domain used for the inverse modelling.

As can be seen in Fig. 7.18(a) and (b), the groundwater heads RMSEs before and after the rotation procedure for the piezometers within the PHT3D model domain are almost identical, indicating similar flow fields in the rotated and unrotated domains. The simulated TCE plumes (Fig. 7.18(c, d)) and VC plumes (Fig. 7.18(e, f)) on 1 October 2005 (barrier installation date) are also similar for the two domains. The small differences are caused by the grid rotation, since the reaction modules used in RT3D and PHT3D provided identical

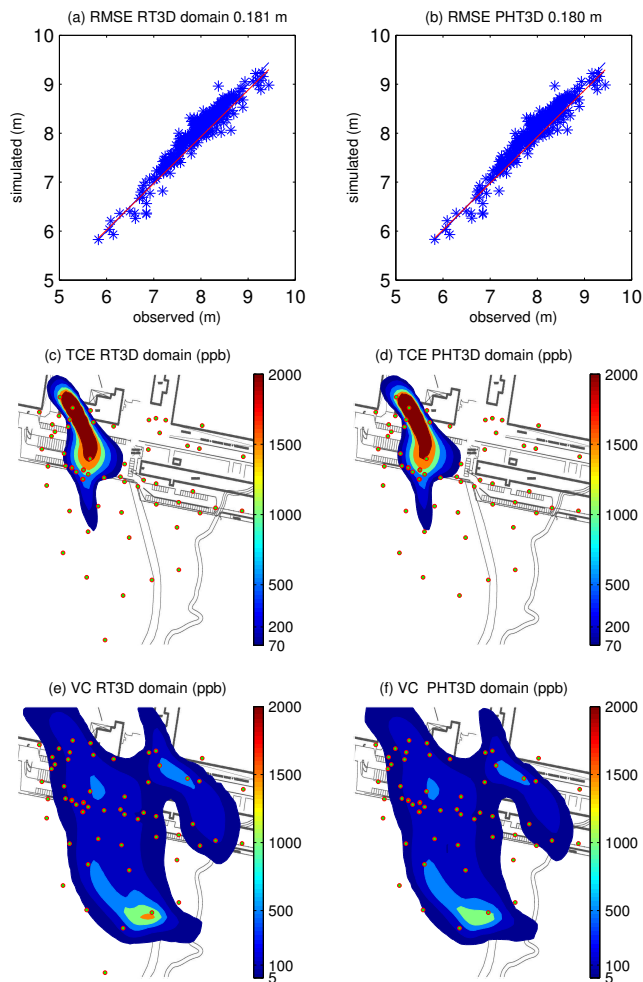


Figure 7.18.: Results obtained for the RT3D and PHT3D models. (a) Head cross plot for the RT3D model (piezometers within the PHT3D model domain), and (b) for the PHT3D model, (c) TCE plume on 1 October 2005 for the RT3D model and (d) for the PHT3D model, (e) VC plume for the RT3D model and (f) for the PHT3D model.

Table 7.3.: Aquifer and rainwater [13] inorganic compositions.

Concentration (mol L ⁻¹)	Aquifer	Rainwater
TIC (mol L ⁻¹)	6.30e-3	0
Ca ²⁺ (mol L ⁻¹)	4.88e-3	1.00e-5
Fe ²⁺ (mol L ⁻¹)	6.02e-5	0
Cl ⁻ (mol L ⁻¹)	1.25e-2	4.92e-5
Na ⁺ (mol L ⁻¹)	7.09e-3	2.60e-5
pH	6.73	5.50
pe	6.13e-1	4

results when tested in batch (results not shown). However, these differences can be considered acceptable if compared to the inability of the model to reproduce the observed concentration fluctuations (Fig. 7.17).

The geochemical scenario with the PRB starts from 1 October 2005, using the concentration fields provided by the geochemical scenario without PRB (TCE and VC concentration fields are shown in Fig. 7.18). In addition to including the deactivation processes, this scenario also differs from the scenario without PRB in the hydraulic conductivity and porosity assigned to the barrier zone, which were increased at 10.2 m d⁻¹ and 0.417 respectively (column data). The iron medium is assumed to deactivate accordingly to the exponential model without iron dissolution (Eq.3.6). The exponential model was chosen over the linear model because the column measurements were fitted slightly better, in particular in validation (chapter 3 and chapter 4). The iron corrosion rate was lowered of 100 times because using the corrosion rate estimated from the column experiment (1.32e-10 mol m² s⁻¹) resulted in a large overestimation of the pH measurements in the piezometers close to the barrier (piezometers 26 and 27). After the adjustment of the corrosion rate, the pH residuals were more normally distributed. The corrosion rate adopted for the field simulation (1.32e-12 mol m² s⁻¹) is close to the values used by Li et al. [96] for the simulation of an hypothetical field installation (2.31e-12 mol m² s⁻¹) and by Mayer et al. [109] for the simulation of the PRB at the Coast Guard Support Center near Elizabeth City, North Carolina (5.01e-12 mol m² s⁻¹). The effect of iron dissolution on iron deactivation was found to be negligible at the corrosion rate adopted for the field simulation.

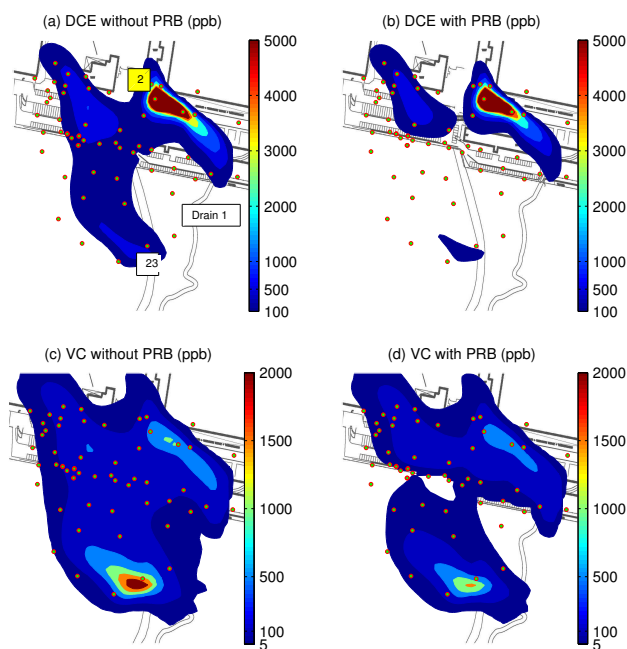


Figure 7.19.: Simulated cis-DCE and VC contaminant concentrations on 1 January 2035. Figures (a) and (c) show the contamination without the PRB and figures (b) and (d) with the PRB installed.

7.7.1. Effect of the barrier installation on contamination

In Fig. 7.19 the cis-DCE and VC plumes on 1 January 2035 without the barrier installed are compared with those with the barrier installed. As can be seen, the PRB effectively reduces the cis-DCE and VC contamination in the pasture area compared to the case without the barrier and is able to stop the migration of TCE pollution originating from source 1, where high pollution levels are predicted to persist up to 2035 (result not shown). However, the VC contamination that reached piezometer 23 and the pasture area before its installation (Fig. 7.18(f)) is persistent due to the low biodegradation rates and the scarce dilution by rain, as was already concluded from Fig. 7.17. To achieve the VC remediation target (contaminant concentrations below 5 ppb) additional actions might be required in the pasture area (for example

enhancing biodegradation with injection of fermentable substrate).

Moreover, the cis-DCE and VC contamination originating from source 2 reached drain 1, located on the right part of the second barrier segment (Fig. 7.19(b,d)). In this location the barrier is not installed and contaminants might further pollute the pasture area in the future. The reliability of this prediction needs to be improved by collecting additional measurements nearby source 2 to better characterize the direction of the plume originating from this source.

7.7.2. Effect of the barrier installation on inorganic concentrations

In Fig. 7.20 the simulated concentration maps on 1 January 2035 are shown. As expected, the pH increase (Fig. 7.20(a)) occurs close to the barrier area, with larger values in the low hydraulic conductivities zones (central part of the first barrier segment and right part of the second barrier segment, see Fig. 7.14(b)), corresponding to low effective velocities and higher groundwater residence times within the barrier. pH values range between 6.6 (background value) and 10.32 (within the PRB), similarly to the simulations of Li et al. [96].

Hydrogen concentrations close to the saturation limit (0.8 mg L^{-1}) were simulated downstream the barrier (Fig. 7.20(b)) because hydrogen was assumed to be not consumed by biodegrading bacteria and the dilution effect by rainwater to be small. As described in sec. 2.2, hydrogen enhances contaminant degradation and might further reduce the VC plume in the pasture area. Moreover, including hydrogen consumption by biodegrading bacterias would provide more realistic predictions of the hydrogen concentrations. Interestingly, hydrogen did not migrate in the central part of the first barrier segment, where low hydraulic conductivities were estimated (Fig. 7.14(b)). An insitu quantification of hydrogen consumption rate by dechlorinating bacteria might be useful to improve the biodegradation model (for example allowing the use the first biodegradation model proposed in sec. 6.3.4, where biodegradation depends on hydrogen concentration). Such predictions will also be affected by the effective insitu corrosion rate, groundwater recharge and the estimated hydraulic conductivity field.

Calcium consumption within the barrier due to mineral precipitation can not be seen in Fig. 7.20(c), although a small concentration gradient is present in the simulations. High calcium concentrations were simulated at the discharge area close to piezometer 23, similarly to the VC plume shown in Fig. 7.19(d).

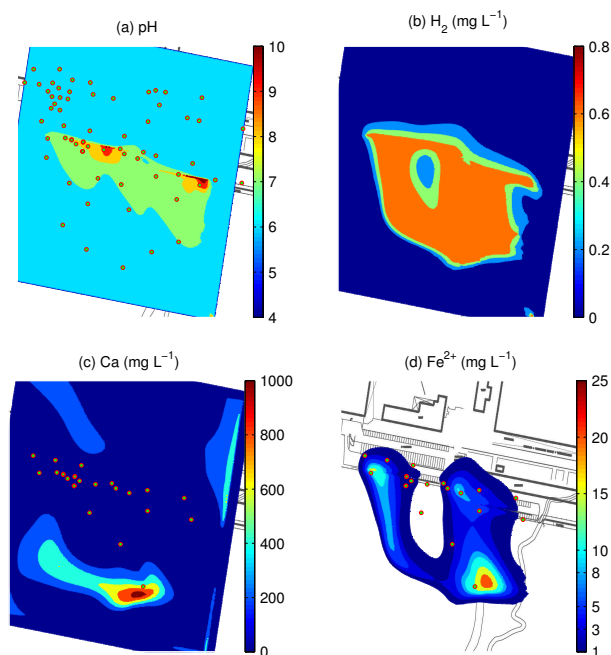


Figure 7.20.: Simulated inorganic concentrations on 1 January 2035 with the PRB installed: (a) pH, (b) dissolved H_2 , (c) dissolved calcium and (d) dissolved iron.

Measured calcium concentrations at piezometer 23 were far below the simulated concentrations (consisting of two measurements around 230 mg L^{-1}), indicating that the high concentration area shown in Fig. 7.20(c) might be caused by an excessive evaporative flux.

Part of the dissolved iron (Fig. 7.20(d)) produced by the corrosion reaction precipitates back as amorphous iron hydroxide and iron hydroxy carbonate (for each mole of corroded iron, 0.04 moles of iron hydroxide and 0.43 moles of iron hydroxy carbonate precipitate within the barrier) and the other part travels downstream in the pasture area. In the central part of the barrier the precipitation of iron carbonate is higher and the dissolved iron concentrations are lower.

7.7.3. Long term PRB efficiency

Fig. 7.21 shows the reactivity decline over time. As can be seen, the average reactivity loss amounts to 38% whereas the maximum to 46%. The maximum loss of reactive surface was simulated in the central and right part of the barrier, in correspondence to low hydraulic conductivities and high pH values (Fig. 7.20(a)). As reported by Li et al. [96], minerals accumulate where the influx rates do not exceed the precipitation rates, in this case in the low hydraulic conductivity zones.

Despite the loss of reactivity, the barrier is able to degrade the contaminants originating from the factory area, as can be seen in Fig. 7.19(d). The VC contamination in the central part was produced before its installation and could not be diluted by rainwater or biodegraded.

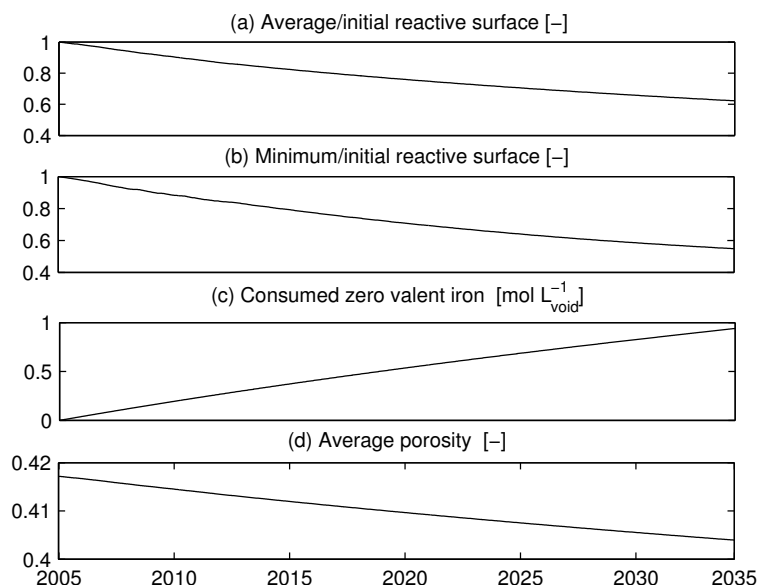


Figure 7.21.: Predicted long term PRB efficiency over time: (a) ratio average/initial reactive surface (b) ratio minimum/initial reactive surface (c) consumed iron and (d) porosity.

Only 5.5% of the initial iron content is predicted to be corroded by water, a results strongly depended on the value adopted for the iron corrosion rate

(e.g. with a 10 times higher corrosion rate, 50% of the initial iron mass is corroded). The maximum expected porosity decline due to mineral precipitation (Fig. 7.21(e)) was small (3.2 % of the initial value), justifying the approach used in the geochemical simulation, where a constant porosity was assumed.

These results agree qualitatively with the core sample analysis. In these analysis the reactivity of the PRB core samples towards TCE was compared to the reactivity of the fresh mixture used at the site, indicating no substantial reactivity decline after 7 years of operation. The EMPA/XRD analysis indicates that a negligible amount of aragonite/calcite covers the iron grains, visually much smaller than that shown in Fig. 3.2 for the column experiment. This information supports the simulation results, where the average aragonite mineral fraction within the PRB amounts to 0.26 %.

7.8. Conclusions

In this work a multicomponent reactive transport model of a permeable reactive barrier site was developed and calibrated to reproduce the present and future contamination. The model includes the processes considered important for the reproduction of the observed pollution, including different recharge patterns in the factory and in the pasture area, contaminant biodegradation and abiotic degradation in the reactive barrier. Two alternative groundwater recharge models were tested (full and reduced recharge) and hydraulic conductivity, specific yield and infiltrating precipitation fractions were parameterized with pilot points, allowing the mapping of the expected parameter variability in the model domain. Flow parameters were estimated from head measurements, followed by inference of reactive transport parameters from concentration measurements. In an additional inversion test, flow and reactive transport parameters were inferred jointly from head and concentration measurements. Parameter and prediction uncertainties were also estimated, assuming known prior parameter covariances and linearizing the model at the optimum parameters.

Groundwater recharge in this shallow groundwater system plays a crucial role in contaminant migration. Our results indicate that the two recharge models fit the head data similarly well, with different estimates for spatially variable hydraulic conductivity and specific yield. However, when using the full recharge model, contamination in the pasture area was strongly diluted, even when neglecting sorption and biodegradation, and assuming elevated source release parameters. The reduced recharge model avoids this problem by generating

less contaminant dilution and yielding estimates for hydraulic conductivity and specific yield that more closely correspond to measured values. Inferred infiltration rates are fairly low however, and in-situ infiltration measurements are required to verify the validity of this approach. Introducing layered heterogeneity in the model, as an alternative to the reduced recharge model for reducing contaminant dilution in the bottom model layer, did not result in satisfying results, and thus the reduced recharge model was retained.

The use of pilot point parameterization introduced a large number of model parameters. To stabilize the inversion, mathematical (singular value decomposition [4]) and explicit regularization [118, 145] were imposed to the parameter fields. The weight of the regularization term in the objective function was adjusted iteratively to achieve specified fits to the individual chemical compounds. This fit was expressed by target objective functions which have a strong influence on the final parameter estimation, and thus need to be carefully chosen. Such problem was also recognized by Fienen et al. [44], where it was stated that changes in regularization constraints (i.e. specific fitting targets) can result in very different parameter fields. In our application, the target objective function values were determined from preliminary model runs. As these values reflect expected misfits between model and data, they account for both measurement and model errors. In our case the model was able to reproduce spatial patterns of time-averaged concentrations, but did not adequately mimic temporal variation of measured concentration at individual piezometers.

The advantage of introducing a flexible parameterization of flow parameters (in particular hydraulic conductivity) is that measured concentrations can be better reproduced. The best matches between measured and simulated cis-DCE and VC plumes were obtained when head and concentration data were used jointly to estimate model parameters. Concentration data contained information about hydraulic conductivity and infiltrating rainwater fractions also in the impervious model area, as demonstrated by the larger reduction of the prior parameters standard deviations compared to the separate estimation of flow and transport parameters. Moreover, a more realistic estimation of the effective porosity and a reduction of the dispersion coefficients were obtained in the joint inversion, in agreement with previous modelling studies [148]. However, some features of the infiltrating fraction and hydraulic conductivity fields might represent an inversion artifact. For example, large variations of hydraulic conductivity were estimated close to the second source and might compensate for an incorrect localization of this source.

Two geochemical scenarios at 2035 were simulated using the optimal parameter

set estimated from the joint optimization and the geochemical model developed from the column experiment (chapter 3). The corrosion rate used in the column simulations produced an overestimation of pH field measurements and was reduced to a value similar to those used in other simulations of field scale PRBs. In the long term geochemical scenario, a negligible reactivity and porosity decline was simulated and the PRB was effective in reducing the contamination in pasture area. However, VC contamination is predicted to persist in the pasture due to the small effect of contaminant dilution and degradation. Therefore, the reliability of the predictions crucially depends on the accurate simulation of the recharge and contaminant biodegradation processes. The hydrogen produced by iron corrosion was assumed not to enhance biodegradation, providing a worst case pollution scenario. The positive effect of hydrogen on contaminant degradation was identified in the lab experiments but its effect should be quantified under field conditions.

To provide more reliable predictions, the current description of flow and transport processes at the site should be improved using additional information, as indicated from the results of this first modelling effort. These include: (1) direct infiltration measurements in the pasture area, (2) additional concentration measurements between source 2 and the second barrier segment, (3) quantification of the measurement error beyond the analytical uncertainty (such as the quantification of contaminant volatilization during the field sampling), and (4) biodegradation tests in additional piezometers (for example at the second source) to confirm the current description of the biodegradation process. From the results obtained in this study a correct description of the water balance is also crucial to predict long-term contamination at the site. In our specific case, reactive transport modelling improved the conceptual understanding of flow and contaminant migration at the site, showing its utility and limitations.

8. Conclusions and future research

This thesis studied and addressed three challenges encountered in developing and applying groundwater reactive transport models, namely: (1) accounting for conceptual model uncertainty, (2) incorporating data from multiple sources into the model, and (3) incorporating spatial heterogeneity in the model in the presence of limited field data. These challenges were tackled in the context of groundwater contamination and remediation at a field site, for which lab and field scale data were used to develop groundwater reactive transport models.

Challenge 1: model conceptual uncertainty Model conceptual uncertainty was encountered both at the lab-scale (chapter 3 and chapter 4) and at the field-scale (chapter 2 and chapter 7). At the lab-scale, significant prior uncertainty existed on the exact mechanism of iron deactivation caused by mineral precipitation in iron-based permeable reactive barriers (PRB). This was accounted for by including in the analysis four different deactivation models (chapter 3) and further developing an alternative simplified deactivation model (chapter 4). At the field-scale, model conceptual uncertainty related to the amount of recharge reaching the aquifer, which was complicated by the shallowness and dynamic nature of the groundwater system. Two conceptual models for recharge were proposed.

Challenge 2: multiple data sources Groundwater reactive transport models need to incorporate data from multiple sources, e.g. measurements of multiple chemical species. A common approach is to optimize parameters using an overall measure of fit, such as the weighted sum of squared residuals (differences between observed and simulated values), where each data set is given a weight based on prior knowledge of the measurement and model errors. In chapter 5, a multivariate Bayesian approach was developed that formally integrates multi-species data, while accounting for correlations. The approach improves upon existing methods by (i) accounting for correlations, and (ii) avoiding an explicit estimation of weights, making it more computationally efficient.

Challenge 3: spatial heterogeneity In chapter 2 and chapter 7 a three-dimensional groundwater reactive transport model was developed, that accounts for spatial heterogeneity in hydraulic parameters and recharge, and for relevant flow and reactive transport processes at the field site. A methodology was developed and applied to estimate the large number of parameters in this model from the limited data at the field site (groundwater levels, groundwater concentrations of multiple chemical species, point-scale measurements of soil hydraulic conductivity, and lab-scale derived information on chemical and biochemical reactions). The proposed methodology relies on (i) pilot points to quantify spatial heterogeneity of hydraulic parameters and recharge, (ii) a regularization procedure to reconcile the large number of spatially distributed model parameters with the limited field data, (iii) a step-wise approach for integrating the different data sets into the model, and (iv) high performance computing. Finally, the field-scale model was used to predict future contaminant migration and remediation efficiency at the site.

The following two sections list the major conclusions and provide suggestions for future research.

8.1. Conclusions

1. The use of column measurements alone did not allow a unique identification of the iron deactivation mechanism (linear and exponential model, iron hydroxide reactive or not) in the PRB, resulting in significant uncertainty of the predicted lifetime of the PRB. In situ angled core samples of the PRB together with reactivity tests were necessary to restrict the choice of the appropriate deactivation model used in the field-scale geochemical simulations.
2. In this context the use of a simpler deactivation model might be justified. In the simplified deactivation model proposed in chapter 4, the speed of the deactivation front relative to the effective groundwater velocity was derived from the observed decline in contaminant degradation, without explicitly modeling iron corrosion and mineral precipitation reactions. This model does not take into account the slowdown of the precipitation rates as the barrier ages and as such provides a conservative estimate of barrier longevity compared to the exponential deactivation model without iron dissolution.
3. The multivariate method was advantageous to infer model parameters from multiple data types compared to existing weighted least squares

methods (WLS and WLS(we)) because it avoids a priori specification of observation weights and allows considering the full residual correlation matrix at a reduced computational cost. It was found that residual correlation does not have a strong effect on predictive uncertainty, and that parameter estimates and their posteriors were only affected at relatively high levels of residual correlation between chemical species (0.9).

4. At the field site alternative descriptions of the groundwater recharge process provided similar simulations of the groundwater levels but different simulations of the contaminant plumes, with different estimates of the hydraulic conductivity fields. The use of groundwater level and concentration data for inference of flow and transport parameters (joint optimization) provided more realistic estimates of the transport parameters (porosity and dispersivities) and improved considerably the fit to the concentration measurements.
5. In the joint inference, areas with low infiltration values were estimated in the pasture area, mainly to prevent contaminant dilution by infiltrating rainwater. Large variations of the hydraulic conductivity field were estimated between the second contaminant source and the second barrier segment, in a poorly characterized area. These estimates should be supported by additional field measurements because they could represent inversion artifacts rather than realistic parameter estimates.
6. Degradation parameters derived from the batch experiments and the iron corrosion rate estimated from the column data were not directly applicable to the field scale model because they provided an underestimation of the contaminant plume and an overestimation of the simulated pH values, respectively. Such differences might be due to the different resolutions and conditions of the lab and field scale models.
7. After adjusting the iron corrosion rate to reproduce pH observations at the field site, the maximum decline in barrier reactivity amounted to 46% by 2035. Despite this reactivity decline, the PRB is predicted to remain fully functional up to 2035.
8. Vinyl Chloride (VC) contamination in the pasture is predicted to persist even after 2035, due to the low biodegradation rates and the small amount of infiltrating rainwater (no contaminant dilution).

8.2. Future research

Below, some interesting directions of future research are listed:

1. Further research on iron deactivation mechanisms is necessary to more accurately predict the long-term reactivity of PRBs. Decrease of iron reactivity was attributed to carbonate minerals coatings only, despite complexes deriving from silica dissolution were observed in the angled core sample of the site PRB (EMPA images), which might hinder the access of contaminants to the iron surface [81]. Dissolved organic matter present in groundwater may also adsorb onto iron hydroxide surfaces reducing the degradation potential of the iron medium [78]. As new data and process knowledge become available, these can be incorporated into the models, resulting in a subsequent decrease in prediction uncertainty.
2. The estimation methods proposed in chapter 5 can be considered robust only for zero-mean residual distributions. In presence of model structural errors, for example due to an unaccounted reaction, strongly biased parameter estimates were obtained. To obtain reliable parameter and prediction uncertainty estimates in presence of model errors, a likelihood function which includes model bias should be included, as proposed in Reichert and Schuwirth [130].
3. Additional data collection at the site is suggested to improve model predictions. In particular:
 - Infiltration rates in the pasture area could be measured to estimate average (in time) infiltration fractions. These estimates could be compared with the results obtained through inverse modelling.
 - Additional concentration measurements (and eventually slug tests) between the second source and the second barrier segment might be useful. These additional measurements can provide valuable information for the estimation of hydraulic conductivities, degradation and source parameters in this poorly characterized area.
 - A precise quantification of the measurement error beyond the analytical uncertainty (such as the quantification of contaminant volatilization during the field sampling) would help to separate concentration fluctuations attributable to measurement error from those attributable to other physical processes (such as groundwater flow variations).
4. Model parameters of the field-scale model were inferred using a gradient based algorithm and parameter and predictive uncertainty assessed assuming linearity at the optimal parameter values, due to the large number of model parameters and the long simulation time. Ongoing research aims to use reduced models to accelerate Bayesian inference of highly parameterized CPU-demanding models. Laloy et al. [89] approxi-

mated the flow response of a real aquifer system using Polynomial Chaos Expansion, achieving a 2x to 5x speed-up for the exploration of a high-dimensional parameter posterior. A similar approach could also be used for the field-scale reactive transport model developed in this thesis. In this case the challenge concerns the construction of an accurate reduced model, which requires an initial phase where several model runs are performed, potentially more than those required for a gradient based MAP parameter inference. However, full parameter inference provides accurate parameter and prediction uncertainty estimates and might also give more accurate estimates of the joint parameter posterior mode (gradient based methods get easily trapped in local minima).

5. The depth of the groundwater table was in most circumstances measured no lower than 1.5 meters from the ground surface. In this case the dynamics of the unsaturated zone have a strong effect on the model results (for example determining the amount of infiltrating rainwater). The dynamics of the unsaturated zone could be included in the model, using a full three-dimensional saturated-unsaturated reactive transport model or approximating the unsaturated zone as a one-dimensional column and estimating the infiltrating recharge at the bottom of the column (thus not modelling contaminant transport in the unsaturated zone but only water flow). When one of these approaches is used, the CPU time required for a forward run increases and the inference of spatially variable aquifer properties will also be more challenging.

A. Appendix to chapter 3

Mass transfer limitations

At the macroscopic scale, the model explicitly accounts for the interplay between mass transfer (advection, dispersion, diffusion) in the bulk solution and reaction rates: as flow velocity decreases, simulated residence time increases, affecting all simulated degradation and geochemical reactions (including pH, iron corrosion, mineral precipitation, and iron deactivation).

At the microscopic scale, mass transfer from bulk solution to the iron grain surface is by molecular diffusion only; since flow velocities are small there is no turbulent transfer. An order of magnitude comparison between molecular diffusion and surface reaction rates indicates that reaction rates are smaller (more than 100 times) and hence control the degradation process irrespective of flow velocities in the bulk solution. An example is here reported:

$$J_D = -D \frac{\partial C}{\partial r} = 10e-9 \left(\frac{m^2}{s} \right) \frac{1.01e-5 \left(\frac{mol}{m^3} \right)}{1e-4(m)} = 1.01e-10 \left(\frac{mol}{m^2s} \right) \quad (A.1)$$

In this example a concentration gradient between the bulk solution and the iron grain surface of 1 g L⁻¹ is assumed (1.01 mol m⁻³ for cis-DCE). It is also assumed that the mass transfer layer is as big as the average particle size (conservative estimate). The resulting molecular diffusion flux is 165 higher than the highest degradation rate (6.13e-13 mol m⁻² s⁻¹ for cis-DCE).

Table A.1.: Influent flow rates and concentrations for the column experiment (BDL: below detection limit).

Pore volume	0 to	31 to	61 to	94 to	139 to	187 to	230 to	284 to	404 to	538 to
	30	60	93	138	186	229	283	403	537	604
Flow rate (L s ⁻¹)	2.35e-6	3.04e-6	3.02e-6	3.05e-6	2.77e-6	1.97e-6	2.51e-6	2.31e-6	2.47e-6	3.04e-6
TCE (mol L ⁻¹)	3.61e-6	2.50e-6	3.85e-6	3.10e-6	3.65e-6	2.36e-6	1.50e-6	2.65e-6	1.54e-6	2.71e-6
PCE (mol L ⁻¹)	4.38e-6	2.52e-6	4.05e-6	2.18e-6	4.65e-6	2.54e-6	1.40e-6	2.55e-6	1.99e-6	3.25e-6
trans-DCE (mol L ⁻¹)	8.28e-8	5.61e-8	8.01e-8	7.82e-8	7.62e-8	5.77e-8	3.56e-8	6.25e-8	3.74e-8	6.10e-8
cis-DCE (mol L ⁻¹)	4.01e-6	3.09e-6	3.52e-6	4.30e-6	4.23e-6	3.36e-6	2.33e-6	4.20e-6	1.99e-6	3.51e-6
VC (mol L ⁻¹)	8.99e-8	1.45e-8	2.38e-7	1.56e-7	1.65e-7	2.26e-7	BDL	2.27e-7	1.61e-8	1.11e-7
pH (pH)	6.85	6.80	6.80	6.75	6.90	6.88	7.59	7.12	7.08	7.01
Eh (mV)	54	132	132	45	361	58	312	169	280	193
Ca ²⁺ (mol L ⁻¹)	4.74e-3	4.24e-3	4.24e-3	4.74e-3	4.74e-3	4.49e-3	4.49e-3	4.99e-3	5.24e-3	4.74e-3
Mg ²⁺ (mol L ⁻¹)	9.87e-4	1.07e-3	1.07e-3	1.03e-3	1.03e-3	1.03e-3	1.03e-3	1.03e-3	1.03e-3	1.03e-3
Fe ²⁺ (mol L ⁻¹)	1.43e-6	1.34e-5	1.34e-5	2.86e-6	2.33e-6	5.19e-6	1.11e-5	5.37e-6	8.42e-5	3.40e-5
Na ⁺ (mol L ⁻¹)	4.35e-3	4.78e-3	4.78e-3	4.57e-3	4.57e-3	4.57e-3	4.57e-3	4.57e-3	4.57e-3	4.57e-3
TiC (mol L ⁻¹)	7.41e-3	7.91e-3	7.91e-3	7.91e-3	7.88e-3	8.28e-3	5.97e-3	8.83e-3	8.49e-3	8.49e-3
SO ₄ ²⁻ (mol L ⁻¹)	1.46e-3	1.35e-3	1.35e-3	1.35e-3	1.25e-3	1.25e-3	1.25e-3	1.35e-3	1.56e-3	1.35e-3
Cl ⁻ (mol L ⁻¹)	5.36e-3	6.21e-3	6.21e-3	5.78e-3	5.78e-3	5.78e-3	5.78e-3	5.78e-3	5.78e-3	5.78e-3
Mn ²⁺ (mol L ⁻¹)	1.27e-6	1.27e-6	1.27e-6	1.27e-6	1.27e-6	1.27e-6	1.27e-6	1.27e-6	1.27e-6	1.27e-6

Table A.2.: Overview of the methods used to determine pH, inorganic ions and contaminants concentrations (adapted from Van Nooten et al. [150]).

Parameter	Method
Contaminants	GC-MS ^a
pH	PH535 (WTW)
Eh	PHM62 electrode (Radiometer Copenhagen) (Ag/AgCl reference electrode)
Ca ²⁺ , Fe ²⁺ , Na ⁺	Inductively coupled plasma / atomic emission spectrometry (ICP-AES)
SO ₄ ²⁻ , Mn ²⁺ , Cl ⁻	Ionchromatography

^a concentrations were determined via headspace analyses with a Thermo Finnigan Trace GC-MS, using a DB5-ms column (60 m length; 0.25 mm internal diameter) and helium as carrier gas at a constant flow rate of 0.9 mL min⁻¹.

Table A.3.: Minerals and liquid-gas equilibriums and equilibrium constants.

Mineral phase	Mineral/phase name	$\log K_i^m$
$CaCO_3 \rightleftharpoons Ca^{2+} + CO_3^{2-}$	Aragonite	-8.336 ^a
$Fe_2(OH)_2CO_3 + 2H^+ \rightleftharpoons 2Fe^{2+} + CO_3^{2-} + 2H_2O$	Iron hydroxy carbonate	1.56 ^b
$Fe(OH)_{2(am)} \rightleftharpoons Fe^{2+} + 2H_2O - 2H^+$	Amorphous iron hydroxide	13.905 ^c
$FeOOH + 3H^+ \rightleftharpoons Fe^{3+} + 2H_2O$	Goethite	-1.0 ^a
$Fe_3O_4 + 8H^+ \rightleftharpoons 2Fe^{3+} + Fe^{2+} + 4H_2O$	Magnetite	3.737 ^d
$FeS_{(am)} + H^+ \rightleftharpoons Fe^{2+} + HS^-$	Amorphous iron sulfide	-4.648 ^a
$H_{2(g)} \rightleftharpoons H_{2(aq)}$	Hydrogen	-3.150 ^a
$CO_{2(g)} \rightleftharpoons CO_{2(aq)}$	Carbon	-1.468 ^a

^a from PHREEQC database [120]. ^b value estimated by Lee et al. [92].

^c from EQ/6 database [163]. ^d from WATEQ4F database [5].

Table A.4.: Aqueous speciation reactions and equilibrium constants.

Equilibrium reaction	Log K
$H_2O \rightleftharpoons OH^- + H^+$	-14
$2H_2O \rightleftharpoons O_2 + 4H^+ + 4e^-$	-86.08
$2H^+ + 2e^- \rightleftharpoons H_2$	-3.15
$CO_3^{2-} + H^+ \rightleftharpoons HCO_3^-$	10.329
$CO_3^{2-} + 2H^+ \rightleftharpoons CO_2 + H_2O$	16.681
$SO_4^{2-} + H^+ \rightleftharpoons HSO_4^-$	1.988
$Ca^{2+} + H_2O \rightleftharpoons CaOH^+ + H^+$	-12.780
$Ca^{2+} + CO_3^{2-} \rightleftharpoons CaCO_3$	3.224
$Ca^{2+} + CO_3^{2-} + H^+ \rightleftharpoons CaHCO_3^+$	11.435
$Ca^{2+} + SO_4^{2-} \rightleftharpoons CaSO_4$	2.300
$Ca^{2+} + HSO_4^- \rightleftharpoons CaHSO_4^+$	1.080
$Na^+ + H_2O \rightleftharpoons NaOH + H^+$	-14.180
$Na^+ + CO_3^{2-} \rightleftharpoons NaCO_3^-$	1.270
$Na^+ + HCO_3^- \rightleftharpoons NaHCO_3$	-0.250
$Na^+ + SO_4^{2-} \rightleftharpoons NaSO_4^-$	0.700
$Fe^{2+} + H_2O \rightleftharpoons FeOH^+ + H^+$	-9.500
$Fe^{2+} + 2H_2O \rightleftharpoons Fe(OH)_2 + 2H^+$	-20.570
$Fe^{2+} + 3H_2O \rightleftharpoons Fe(OH)_3 + 3H^+$	-31.000
$Fe^{2+} + Cl^- \rightleftharpoons FeCl^+$	0.140
$Fe^{2+} + CO_3^{2-} \rightleftharpoons FeCO_3$	4.380
$Fe^{2+} + HCO_3^- \rightleftharpoons FeHCO_3^-$	2.000
$Fe^{2+} + SO_4^{2-} \rightleftharpoons FeSO_4$	2.250
$Fe^{2+} + HSO_4^- \rightleftharpoons FeHSO_4^+$	1.080
$Fe^{2+} \rightleftharpoons Fe^{3+} + e^-$	-13.020
$Fe^{3+} + H_2O \rightleftharpoons FeOH^{2+} + H^+$	-2.190
$Fe^{3+} + 2H_2O \rightleftharpoons Fe(OH)^{2+} + 2H^+$	-5.670
$Fe^{3+} + 3H_2O \rightleftharpoons Fe(OH)_3 + 3H^+$	-12.560
$Fe^{3+} + 4H_2O \rightleftharpoons Fe(OH)_4^- + 4H^+$	-21.60
$2Fe^{3+} + 2H_2O \rightleftharpoons Fe_2(OH)_2^{4+} + 2H^+$	-2.950
$3Fe^{3+} + 4H_2O \rightleftharpoons Fe_3(OH)_4^{5+} + 4H^+$	-6.300
$Fe^{3+} + Cl^- \rightleftharpoons FeCl^{2+}$	1.480
$Fe^{3+} + 2Cl^- \rightleftharpoons FeCl_2^+$	2.130
$Fe^{3+} + 3Cl^- \rightleftharpoons FeCl_3$	1.130
$Fe^{3+} + SO_4^{2-} \rightleftharpoons FeSO_4^+$	4.040

Equilibrium reaction	Log K
$\text{Fe}^{3+} + 2\text{SO}_4^{2-} \rightleftharpoons \text{Fe}(\text{SO}_4)_2^-$	5.380

Table A.5.: VOCs degradation reaction network and iron corrosion reaction stoichiometry.

Reaction	Stoichiometry
PCE degradation	$\text{C}_2\text{Cl}_4 + 3.61\text{Fe}^0 + 3.61\text{H}^+ \rightarrow 3.61\text{Fe}^{2+} + 0.13\text{C}_2\text{HCl}_3 + 0.87\text{C}_2\text{H}_4 + 3.61\text{Cl}^-$
TCE degradation	$\text{C}_2\text{HCl}_3 + 2.94\text{Fe}^0 + 2.94\text{H}^+ \rightarrow 2.94\text{Fe}^{2+} + 0.03\text{C}_2\text{H}_2\text{Cl}_2 + 0.97\text{C}_2\text{H}_4 + 2.94\text{Cl}^-$
cis-DCE degradation	$\text{C}_2\text{H}_2\text{Cl}_2 + 1.94\text{Fe}^0 + 1.94\text{H}^+ \rightarrow 1.94\text{Fe}^{2+} + 0.06\text{C}_2\text{H}_3\text{Cl} + 0.94\text{C}_2\text{H}_4 + 1.94\text{Cl}^-$
trans-DCE degradation	$\text{C}_2\text{H}_2\text{Cl}_2 + 1.99\text{Fe}^0 + 1.99\text{H}^+ \rightarrow 1.99\text{Fe}^{2+} + 0.01\text{C}_2\text{H}_3\text{Cl} + 0.99\text{C}_2\text{H}_4 + 1.99\text{Cl}^-$
VC degradation	$\text{C}_2\text{H}_3\text{Cl} + \text{Fe}^0 + \text{H}^+ \rightarrow \text{Fe}^{2+} + \text{C}_2\text{H}_4 + \text{Cl}^-$
Anaerobic iron corrosion by water	$\text{Fe}^0 + 2\text{H}_2\text{O} \rightarrow \text{Fe}^{2+} + 2\text{OH}^- + \text{H}_2$
Anaerobic iron corrosion by carbonate	$\text{Fe}^0 + \frac{1}{4}\text{CO}_3^{2-} + \frac{5}{2}\text{H}^+ \rightarrow \text{Fe}^{2+} + \frac{1}{4}\text{CH}_4(\text{aq}) + \frac{3}{4}\text{H}_2\text{O}$
Aerobic iron corrosion	$\text{Fe}^0 + \frac{1}{4}\text{O}_{2(\text{aq})} + 3\text{H}^+ \rightarrow \text{Fe}^{3+} + \frac{1}{2}\text{H}_2\text{O} + \text{H}_2$
Sulfate biodegradation	$\text{SO}_4^{2-} + 4\text{H}_{2(\text{aq})} + \text{H}^+ \rightarrow \text{HS}^- + 4\text{H}_2\text{O}$

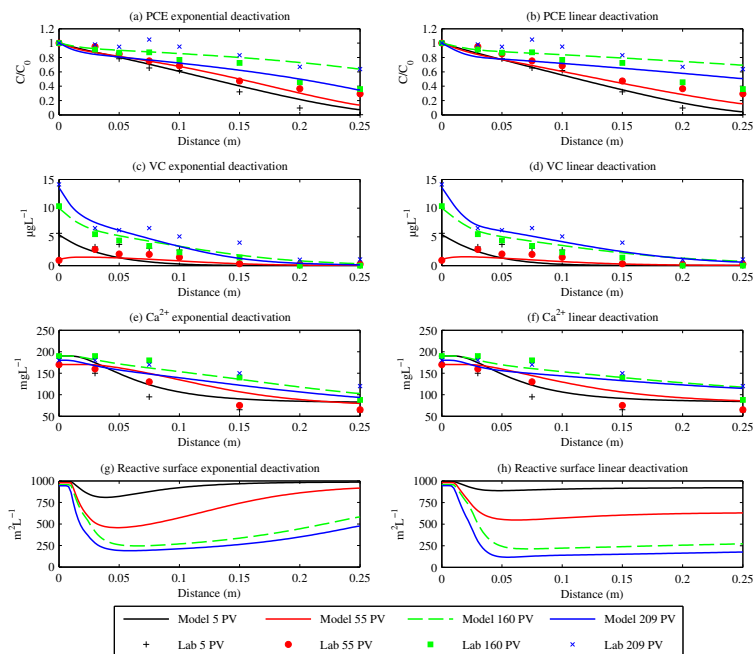


Figure A.1.: Measured and simulated profiles for the calibration step including the dissolution term. Note the reactive surface declining (figures g and h) at the column inlet due to iron dissolution.

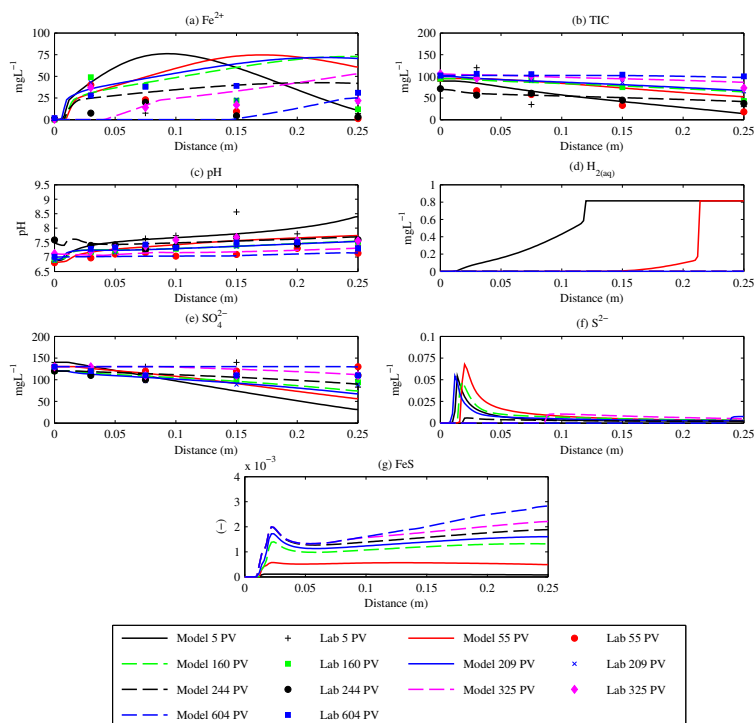


Figure A.2.: Measured and simulated profiles with the exponential version of the model, including sulphate biodegradation and amorphous iron sulphite precipitation. (a) Fe^{2+} , (b) TIC, (c) pH, (d) H_2 , (e) SO_4^{2-} , (f) S^{2-} and (g) iron sulphite mineral fraction.

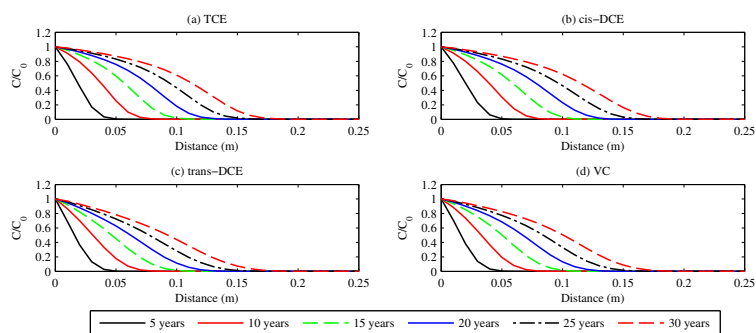


Figure A.3.: Simulated VOCs concentrations predicted by the calibrated model forced with an effective groundwater velocity of 0.01 m d^{-1} . In this picture the reactive surface deactivation was described by the exponential relationship of Jeen et al. [65] without any modification.

B. Appendix to chapter 5

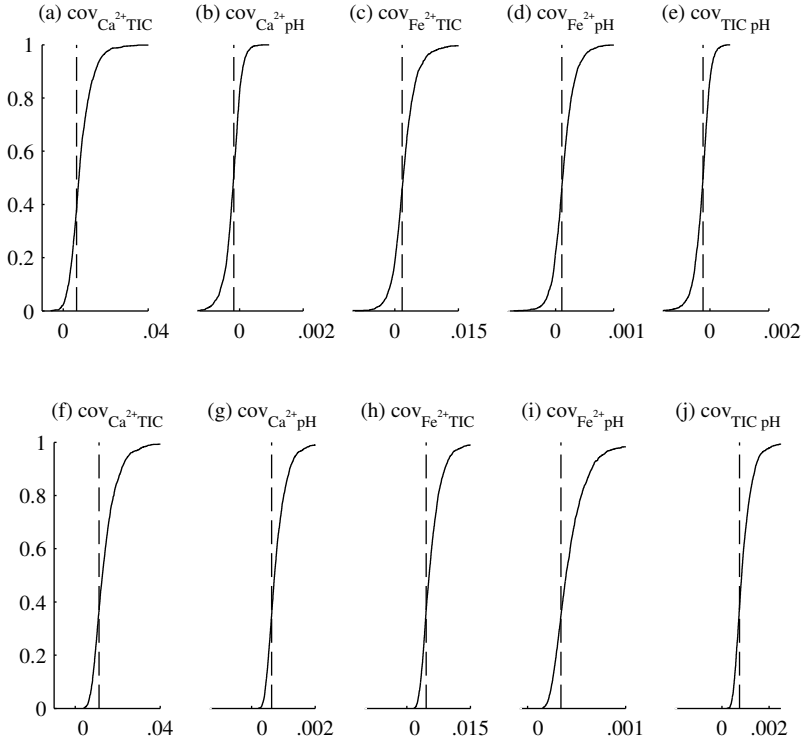


Figure B.1.: Cumulative distribution functions of the off-diagonal Σ_{res} matrix elements inferred by the MV method: plots (a) to (e) show the covariances for the low correlation case and plots (f) to (j) the covariances for the high correlation case. The black dashed lines correspond to the values computed from the correct residuals.

Table B.1.: Parameter inference results for the synthetic case with low correlation

a)	Parameters				L_2 bias vector ^b	SE^c	Log posterior density ^d
	$\log k_{eff,IC}$	$\log k_{IrC,W}$	$\log k_{eff,CC}$	$\log k_{IrC,TIC}$			
Correct values	-7.09	-11.35	-6.56	-7.31	0	1.043	133.93 WLS 135.12 WLS(we) MV
WLS	-7.11 (-7.17 to -7.05) ^a -7.08	-11.39 (-7.49 to -11.27) -11.36	-6.51 (-6.57 to -6.49) -6.51	-7.307 (-7.314 to -7.304) -7.307	0.068	1.029	134.97
WLS(we)	(-7.17 to -6.99) -7.08	(-11.48 to -11.27) -11.38	(-6.58 to -6.43) -6.53	(-7.314 to -7.304) -7.308	0.055	1.057	135.25
MV	(-7.18 to -6.99)	(-11.47 to -11.29)	(-6.57 to -6.49)	(-7.313 to -7.305)	0.043	1.034	135.98
b)	Estimated σ_{MOj}						
	Ca ²⁺	Fe ²⁺	TIC	pH			
WLS	2.84e-2	2.76e-2	0 ^e	0			
WLS(we)	3.10e-2 (8.86e-6 to 4.45e-2)	8.97e-3 (8.97e-6 to 4.45e-2)	5.93e-3 (9.38e-6 to 5.80e-2)	1.27e-5 (1.08e-5 to 3.90e-3)			
MV	0	1.37e-2	0	0			

^a In brackets the 5-95 parameter quantiles are reported. ^b The bias vector is defined as the difference between the correct parameter vector and the estimated one. ^c SE was calculated using equation 5.18. ^d Computed accordingly to the theoretical section (equations 5.8, 5.10 and 5.15). For the WLS(we) method, the density function with the correct parameter set was calculated using a σ_{MO_j} equal to 10^{-5} (negligible additional standard deviation). ^e Reported values of 0 indicate that the estimated total residual variance was slightly lower than the measurement error variance, providing a negative σ_{MO_j} (which is not realistic and was set to 0).

Table B.2.: Parameter inference results for the synthetic case with high correlation

a)	Parameters					
	$\log k_{eff,IC}$	$\log k_{I-C,W}$	$\log k_{eff,CC}$	$\log k_{I-C,TIC}$	L_2 bias vector	SE
Correct values	-7.09	-11.35	-6.56	-7.31	0	0.942
WLS	-7.03 (-7.10 to -6.91)	-11.36 (-11.47 to -11.27)	-6.54 (-6.61 to -6.52)	-7.308 (-7.314 to -7.306)	0.063	0.935
WLS(we)	-7.06 (-7.12 to -6.97)	-11.33 (-11.50 to -11.26)	-6.57 (-6.62 to -6.51)	-7.311 (-7.315 to -7.304)	0.036	0.958
MV	-7.07 (-7.12 to -7.07)	-11.35 (-11.37 to -11.33)	-6.55 (-6.57 to -6.54)	-7.310 (-7.311 to -7.309)	0.021	0.935
b)	Estimated σ_{MOj}					
	Ca^{2+}	Fe^{2+}	TIC	pH		
WLS	3.61e-2	0	0	0		
WLS(we)	1.87e-2 (8.64e-6 to 5.51e-2)	2.23e-5 (1.11e-5 to 1.63e-2)	2.45e-5 (1.18e-5 to 5.48e-2)	4.65e-5 (8.10e-6 to 4.85e-3)		
MV	0	0	0	0		

Table B.3.: Parameter inference results for the synthetic case with model error

a)	$\log k_{TC,W}$ $\log k_{eff,CC}$ $\log k_{TC,TIC}$			L_2 bias vector	SE	Log posterior density
Correct values	-11.35	-6.56	-7.31	0	53.337	-1.09e5 WLS -1.09e5 WLS(we)
WLS	-11.70	-7.66	-7.94	1.312	26.605	42.4 MV -2.72e4
	(-11.70 to -11.60)	(-7.67 to -7.65)	(-7.94 to -7.94)			
	-10.97	-7.82	-8.07	1.523	27.418	-27.37
WLS(we)	(-11.83 to -10.35)	(-8.21 to -7.05)	(-8.77 to -7.89)			
MV	-11.84	-7.61	-8.14	1.424	31.437	-0.25
	(-11.91 to -10.48)	(-8.00 to -6.21)	(-8.86 to - 8.06)			
b)	Estimated σ_{MOj}			pH		
	Ca^{2+}	Fe^{2+}	TIC			
WLS	0.287	1.673	0.544	0.110		
WLS(we)	0.474	3.441	1.568	0.249		
	(0.269 to 1.171)	(2.801 to 5.890)	(1.163 to 3.464)	(0.191 to 0.768)		
MV	0.169	1.807	0.738	0.118		

Table B.4.: Parameter inference results for the real case.

a)	Parameters					
	$\log k_{eff,IC}$	$\log k_{I-C,W}$	$\log k_{eff,CC}$	$\log K_{IC}^n$	$\log k_{I-C,TIC}$	$\log \alpha_{IC}$
WLS	-6.02 (-6.06 to -6.01)(a)	-13.58 (-14.00 to -13.23)	-8.45 (-8.45 to -8.44)	-1.25 (-1.27 to -1.24)	-7.85 (-7.85 to -7.84)	2.22 (2.21 to 2.23)
WLS(we)	-6.67 (-8.29 to -6.10)	-12.61 (-13.88 to -11.20)	-6.60 (-7.58 to -6.39)	-1.60 (-2.00 to 0.96)	-7.50 (-7.62 to -7.45)	0.72 (0.66 to 2.07)
MV	-6.87 (-9.81 to -6.13)	-11.91 (-13.71 to -10.08)	-6.85 (-7.87 to -6.44)	-1.54 (-1.86 to 0.04)	-7.81 (-9.24 to -7.63)	1.81 (0.91 to 2.73)
b)	Estimated σ_{MOj}					
	Ca^{2+}	Fe^{2+}	TIC	pH		
WLS	0.340	0.749	0.408	0.258		
WLS(we)	0.231 (0.207 to 0.375)	0.519 (0.495 to 1)	0.315 (0.205 to 0.467)	0.345 (0.249 to 0.457)		
MV	0.316	0.715	0.318	0.275		

C. Appendix to chapter 6

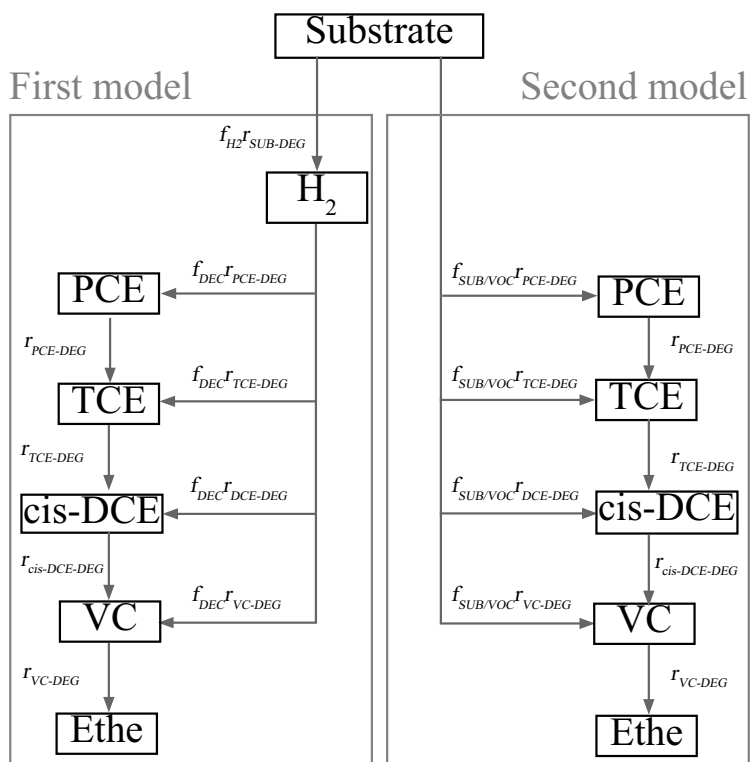


Figure C.1.: Biodegradation models.

Table C.1.: Initial parameter values of the first biodegradation model.

Parameter	Units	Initial guess	Estimated	Literature range
μ_{PCE}	mol L ⁻¹ d ⁻¹	1e-5	Yes	4.05e-7 ^a to 2.07e-1 ^b
μ_{TCE}	mol L ⁻¹ d ⁻¹	1e-5	Yes	5.49e-6 ^a to 2.16e-1 ^b
$\mu_{cis-DCE}$	mol L ⁻¹ d ⁻¹	1e-5	Yes	7.2e-7 ^a to 6.35e-2 ^b
μ_{VC}	mol L ⁻¹ d ⁻¹	1e-5	Yes	7.2e-7 ^a to 6.35e-2 ^b
μ_{SUB}	mol L ⁻¹ d ⁻¹	1e-5	Yes	1.32e-4 ^c
K_{PCE}	mol L ⁻¹	1.1e-7 ^d	No	1.10e-7 ^d to 8.8e-7 ^b
K_{TCE}	mol L ⁻¹	1.4e-6	No	5e-8 ^e to 1.58e-6 ^c
K_{DCE}	mol L ⁻¹	3.3e-6	No	2.16e-6 ^c to 3.3e-6 ^d
K_{VC}	mol L ⁻¹	2.6e-6	No	2.6e-6 ^d to 1.76e-3 ^c
K_{SUB}	mol L ⁻¹	1.7e-5	No	1.7e-5 ^c
K_{H_2}	mol L ⁻¹	2e-8	No	2.0e-8 ^c to 1e-7 ^a
H_2^*	mol L ⁻¹	2e-9	No	1.5e-9 ^b to 2e-9 ^c
H_2^{scale}	mol L ⁻¹	8e-8	No	8e-9 ^c
f_{H_2}	-	10 ^f	No	3-14
f_{DEC}	-	1.41	No	1.41 ^a

^a Christ and Abriola [25]. ^b Cupples et al. [29].

^c Kouznetsova et al. [85]. ^d half saturation constants values were taken from Haston and McCarty [51]. ^e Lee et al. [91]. ^f assuming $C_5H_7O_2N$ as substrate molecule. Manoli et al. [105] report a conversion factor ranging from 3 (lactate) to 14 (linoleate), without accounting for bacteria competition (such as methanogenic and sulphate reducer).

D. Appendix to chapter 7

Table D.1.: Inversions and simulations for the site model

Experiment	Discr.	Target (Tab. 6.1)	Parameters	Algorithm settings and initial conditions	CPU hours
1a Transient flow, full recharge	1	420 transient ground- water levels (2011)	178 hydraulic conductivities, 127 specific yield and 3 drain conductances	$\Phi_m^l = 9.45$. Initial parameter values of Tab. 7.1	111.31 (29 CPUs)
1b Transient flow, reduced recharge	1	420 transient ground- water levels (2011)	178 hydraulic conductivities, 127 specific yield and 3 drain conductances, 43 infiltration fractions pilot points	$\Phi_m^l = 9.45$. Initial parameter values of Tab. 7.1.	1201.57 (29 CPUs)
1c Transient flow, reduced recharge, low regu- larization	1	420 transient ground- water levels (2011)	178 hydraulic conductivities, 127 specific yield and 3 drain conductances, 43 infiltration fractions pilot points	$\Phi_m^l = 1.05$. Initial parameter values of Tab. 7.1.	833.75 (29 CPUs)

Experiment	Discr.	Target (Tab. 6.1)	Parameters	Algorithm settings and initial conditions	CPU hours
2a Con- servative transport (total contami- nation), full recharge	2	Average 2009-2012 total concen- tration map	Forward run	Flow parameters from 1a. Initial concentration field equal to the average prebarrier concentration map. At each source 1000 kg of DNAPL are present at the starting of the simulation. Mass transfer and transport parameters are reported in Tab. 7.1	12.30 (1 CPU)
2b Con- servative transport (total contami- nation), reduced recharge	2	Average 2009-2012 total concen- tration map	Forward run	Flow parameters from 1b. Initial concentration, source and transport parameters as in 2a	0.06 (1 CPU)

Experiment	Discr.	Target (Tab. 6.1)	Parameters	Algorithm settings and initial conditions	CPU hours
2c Con- servative transport (total contami- nation), full recharge	1	Average 2009-2012 total concen- tration map	Forward run	Flow parameters from 1a. Pristine aquifer. At each source the release rate was specified at $1.5\text{e-}4 \text{ mol L}^{-1} \text{ d}^{-1}$ (corresponding to 1000 kg of DNAPL released during the 1957-1996 period). Other source and transport parameters are reported in Tab. 7.1	1.42 (1 CPU)
2d Con- servative transport (total contami- nation), reduced recharge	1	Average 2009-2012 total concen- tration map	Forward run	Flow parameters from 1b. Pristine aquifer. Degradation, source and transport parameters as in step 2c	0.23 (1 CPU)

Experiment	Discr.	Target (Tab. 6.1)	Parameters	Algorithm settings and initial conditions	CPU hours
3a Multi- compo- nent reactive trans- port, batch ex- periment degrada- tion rates	1	Average concen- tration maps using all measure- ments	Forward run	Flow parameters from step 1b. Degradation parameters from PCE to VC fixed as determined from batch experiments (Tab. 7.1, third column), the other degradation, source and transport parameters are reported in the same table	1.35 (1 CPU)
3b Multi- compo- nent reactive trans- port, manual calibra- tion	1	Average concen- tration maps using all measure- ments	Several forward runs	Flow parameters from 1b. μ_{PCE} , μ_{TCE} , $\mu_{cis-DCE}$ and μ_{VC} degradation parameters were adjusted whereas the other degradation, source and transport parameters were kept fixed as reported in Tab. 7.1	16.90 (1 CPU)

Experiment	Discr.	Target (Tab. 6.1)	Parameters	Algorithm settings and initial conditions	CPU hours
3c Multi- compo- nent reactive trans- port, equal weights	1	2240 concen- trations measure- ments (all data)	Degradation, source and transport parameters (Tab. 7.1, 18 parameters)	Flow parameters from 1b. Initial degradation, source and transport parameters from 3b. No regularization, equal initial contribution of each measurement set to the initial Φ_m	332.80 (18 CPUs)
3d Multi- compo- nent reactive trans- port, inverse weights	1	2240 concen- trations measure- ments	Degradation, source and transport parameters (Tab. 7.1, 18 parameters)	Flow parameters from step 1b. Initial degradation, source and transport parameters from step 3b. No regularization, weights inversely proportional to the concentrations	648.80 (18 CPUs)
4a Multi- compo- nent reactive transport	1	420 transient ground- water levels (2011), 2240 concen- trations	178 hydraulic conductivities (refined), 3 drain conductances, 43 recharge pilot points, 18 degradation parameters	Initial flow field from step 1b. Initial degradation, source and transport parameters from 3b. Iterative weighting scheme with target RMSE of 200 ppb for each measurement set	2827.07 (28 CPUs)

Experiment	Discr.	Target (Tab. 6.1)	Parameters	Algorithm settings and initial conditions	CPU hours
4b Multi- compo- nent reactive transport	1	420 transient ground- water levels (2011), 2240 con- taminant concen- trations	178 hydraulic conductivities (refined), 3 drain conductance, 43 recharge pilot points for recharge, 17 degradation parameters	Initial flow field from step 1b. Initial degradation, source and transport parameters from step 3b. Iterative weighting scheme with target $\Phi_{m,i}^l$, with target RMSEs reported in Tab. 7.2	1876.20 (36 CPUs)
5a Multi- compo- nent reactive transport	4	Predic- tion	None	None, calculating the Jacobian at the best parameter set (tests 1b and 3c) for parameter uncertainty	298.83 (10 CPUs)
5b Multi- compo- nent reactive transport	4	Predic- tion	None	None, calculating the jacobian at the best parameter set (step 4b) for parameter and prediction uncertainty	550.50 (10 CPUs)
6a Geo- chemical scenario without PRB	4	Predic- tion	None	None, long term scenario	173.20 (1 CPU)
6b Geo- chemical scenario with PRB	3	Predic- tion	None	None, long term scenario	834.53 (1 CPU)
Total amount of CPU hours					9710.93

Bibliography

- [1] Arnold, W. A. and Roberts, A. L. (2000). Pathways and kinetics of chlorinated ethylene and chlorinated acetylene reaction with Fe0 particles. *Environmental Science & Technology*, 34(9):1794–1805.
- [2] Arora, B., Mohanty, B. P., and McGuire, J. T. (2011). Inverse estimation of parameters for multidomain flow models in soil columns with different macropore densities. *Water Resources Research*, 47(4):W04512.
- [3] Assefa, K. A. and Woodbury, A. D. (2013). Transient, spatially varied groundwater recharge modeling. *Water Resources Research*, 49(8):4593–4606.
- [4] Aster, R. C., Borchers, B., and Thurber, C. H. (2005). *Parameter estimation and inverse problems*. Elsevier Academic Press.
- [5] Ball, J. W. and Nordstrom, D. K. (1991). Users manual for WATEQ4F, with revised thermodynamic database and test cases for calculating speciation of major, trace and redox elements in natural waters. open file report. *US Geol. Surv.*, pages 91–183.
- [6] Barlebo, H. C., Hill, M. C., and Rosbjerg, D. (2004). Investigating the macrodispersion experiment (MADE) site in Columbus, Mississippi, using a three-dimensional inverse flow and transport model. *Water Resources Research*, 40(4):W04211.
- [7] Barth, G. and Hill, M. C. (2005). Numerical methods for improving sensitivity analysis and parameter estimation of virus transport simulated using sorptive–reactive processes. *Journal of Contaminant Hydrology*, 76(3–4):251–277.
- [8] Bauer, M., Heitmann, T., Macalady, D. L., and Blodau, C. (2006). Electron transfer capacities and reaction kinetics of peat dissolved organic matter. *Environmental Science & Technology*, 41(1):139–145.
- [9] Bell, L. S. J., Binning, P. J., Kuczera, G., and Kau, P. M. H. (2002). Rigorous uncertainty assessment in contaminant transport inverse modelling: a case study of fluoride diffusion through clay liners. *Journal of Contaminant Hydrology*, 57(1–2):1–20.

-
- [10] Bi, E., Devlin, J. F., Huang, B., and Firdous, R. (2010). Transport and kinetic studies to characterize reactive and nonreactive sites on granular iron. *Environmental Science & Technology*, 44(14):5564–5569.
- [11] Blowes, D. W., Ptacek, C. J., Benner, S. G., McRae, C. W. T., Bennett, T. A., and Puls, R. W. (2000). Treatment of inorganic contaminants using permeable reactive barriers. *Journal of Contaminant Hydrology*, 45(1–2):123–137.
- [12] Bonin, P. M. L., Odziemkowski, M. S., Reardon, E. J., and Gillham, R. W. (2000). In situ identification of carbonate-containing green rust on iron electrodes in solutions simulating groundwater. *Journal of Solution Chemistry*, 29(10):1061–1074.
- [13] Boschloo, D. J. and Stolk, A. P. (1999). *National network for analysis of rainwater chemistry (in Dutch)*. Number RIVM Rep. 723101049/054.
- [14] Bouwer, H. and Rice, R. C. (1976). A slug test for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells. *Water Resources Research*, 12(3):423–428.
- [15] Box, G. E. P. and Tiao, G. C. (1992). *Bayesian Inference in Statistical Analysis*, 588 pp. John Wiley, New York.
- [16] Bredehoeft, J. (2005). The conceptualization model problem-surprise. *Hydrogeology Journal*, 13(1):37–46.
- [17] Campbell, M. D., Starrett, M. S., Fowler, J. D., and Klein, J. (1990). Slug test and hydraulic conductivity. *Ground Water Manage*, 4:85–99.
- [18] Carrera, J. (1993). An overview of uncertainties in modelling groundwater solute transport. *Journal of Contaminant Hydrology*, 13:23–48.
- [19] Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L. J. (2005). Inverse problem in hydrogeology. *Hydrogeology Journal*, 13(1):206–222.
- [20] Carrera, J. and Neuman, S. P. (1986). Estimation of aquifer parameters under transient and steady state conditions: 1. maximum likelihood method incorporating prior information. *Water Resources Research*, 22(2):199–210.
- [21] Carroll, R. J. and Ruppert, D. (1982). A comparison between maximum likelihood and generalized least squares in a heteroscedastic linear model. *Journal of The American Statistical Association*, 77(380):878–882.
- [22] Certes, C. and de Marsily, G. (1991). Application of the pilot point method to the identification of aquifer transmissivities. *Advances in Water Resources*, 14(5):284–300.
- [23] Cervini-Silva, J., Kostka, J. E., Larson, R. A., Stucki, J. W., and Wu, J. (2003). Dehydrochlorination of 1,1,1-trichloroethane and pentachloroethane

- by microbially reduced ferruginous smectite. *Environmental Toxicology and Chemistry*, 22(5):1046–1050.
- [24] Chambers, J., Cleveland, W., Kleiner, B., and Tukey, P. (1983). *Graphical Methods for Data Analysis*. Wadsworth International Group, Belmont, California.
- [25] Christ, J. A. and Abriola, L. M. (2007). Modeling metabolic reductive dechlorination in dense non-aqueous phase liquid source-zones. *Advances in Water Resources*, 30(6–7):1547–1561.
- [26] Clapp, L., Semmens, M., Novak, P., and Hozalski, R. (2004). Model for in situ perchloroethene dechlorination via membrane-delivered hydrogen. *Journal of Environmental Engineering*, 130(11):1367–1381.
- [27] Clement, T. (1997). RT3D: A modular computer code for simulating reactive multi-species transport in 3-Dimensional groundwater aquifers. *Pacific Northwest National Laboratory*.
- [28] Clement, T. P., Johnson, C. D., Sun, Y., Klecka, G. M., and Bartlett, C. (2000). Natural attenuation of chlorinated ethene compounds: model development and field-scale application at the Dover site. *Journal of Contaminant Hydrology*, 42(2–4):113–140.
- [29] Cupples, A. M., Spormann, A. M., and McCarty, P. L. (2003). Growth of a dehalococcoides-like microorganism on vinyl chloride and cis-dichloroethene as electron acceptors as determined by competitive PCr. *Applied and Environmental Microbiology*, 69(2):953–959.
- [30] D’Affonseca, F. M., Prommer, H., Finkel, M., Blum, P., and Grathwohl, P. (2011). Modeling the long-term and transient evolution of biogeochemical and isotopic signatures in coal tar-contaminated aquifers. *Water Resources Research*, 47(5):W05518.
- [31] Dai, Z. and Samper, J. (2004). Inverse problem of multicomponent reactive chemical transport in porous media: Formulation and applications. *Water Resources Research*, 40(7):W07407.
- [32] Dai, Z. and Samper, J. (2006). Inverse modeling of water flow and multicomponent reactive transport in coastal aquifer systems. *Journal of Hydrology*, 327(3):447–461.
- [33] Dausman, A. M., Doherty, J., Langevin, C. D., and Sukop, M. C. (2010). Quantifying data worth toward reducing predictive uncertainty. *Ground Water*, 48(5):729–740.
- [34] Davis, J. C. and Sampson, R. J. (2002). *Statistics and data analysis in geology*, volume 3. Wiley New York.

- [35] De Marsily, G (1984). Spatial variability of properties in porous media: a stochastic approach. In *Fundamentals of Transport Phenomena in Porous Media*, pages 719–769. Springer.
- [36] Devlin, J. F. and Allin, K. O. (2005). Major anion effects on the kinetics and reactivity of granular iron in glass-encased magnet batch reactor experiments. *Environmental Science & Technology*, 39(6):1868–1874.
- [37] Distefano, T. D., Gossett, J. M., and Zinder, S. H. (1992). Hydrogen as an electron-donor for dechlorination of tetrachloroethene by an anaerobic mixed culture. *Applied and Environmental Microbiology*, 58(11):3622–3629.
- [38] Doherty, J. (2003). Ground water model calibration using pilot points and regularization. *Ground Water*, 41(2):170–177.
- [39] Doherty, J. (2007). *PEST: Model Independent Parameter Estimation*. Watermark Numerical Computing.
- [40] Doherty, J., Fienen, M., and Hunt, R. (2010). Approaches to highly parameterized inversion: Pilot-point theory, guidelines, and research directions. Sci. Invest. Rep 5168, US Geol. Surv.
- [41] Doherty, J. and Welter, D. (2010). A short exploration of structural noise. *Water Resources Research*, 46(5):W05525.
- [42] Draper, N. R. and Smith, H. (1998). *Applied regression analysis*. John Wiley and Sons, New York.
- [43] Duan, Q., Sorooshian, S., and Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4):1015–1031.
- [44] Fienen, M., Hunt, R., Krabbenhoft, D., and Clemo, T. (2009 (b)). Obtaining parsimonious hydraulic conductivity fields using head and transport observations: A Bayesian geostatistical parameter estimation approach. *Water Resources Research*, 45(8):W08405.
- [45] Fienen, M. N., Muffels, C. T., and Hunt, R. J. (2009 (a)). On constraining pilot point calibration with regularization in PEST. *Ground Water*, 47(6):835–844.
- [46] Gallagher, M. R. and Doherty, J. (2007). Parameter interdependence and uncertainty induced by lumping in a hydrologic model. *Water Resources Research*, 43(5):W05421.
- [47] Gorelick, S. M., Evans, B., and Remson, I. (1983). Identifying sources of groundwater pollution: An optimization approach. *Water Resources Research*, 19(3):779–790.

- [48] Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., and Ye, M. (2012). Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, 48:W08301.
- [49] Harbaugh, A. W. (1990). *A computer program for calculating subregional water budgets using results from the US Geological Survey modular three-dimensional finite-difference ground-water flow model*. US Geol. Surv.
- [50] Harbaugh, A. W., Banta, E. R., Hill, M. C., and McDonald, M. G. (2000). *MODFLOW-2000, the US Geological Survey modular ground-water model: User guide to modularization concepts and the ground-water flow process*. US Geol. Surv.
- [51] Haston, Z. C. and McCarty, P. L. (1998). Chlorinated ethene half-velocity coefficients (ks) for reductive dehalogenation. *Environmental Science & Technology*, 33(2):223–226.
- [52] Hayley, K., Schumacher, J., MacMillan, G., and Boutin, L. (2014). Highly parameterized model calibration with cloud computing: an example of regional flow model calibration in northeast Alberta, Canada. *Hydrogeology Journal*, 22(3):729–737.
- [53] He, J. Z., Sung, Y., Dollhopf, M. E., Fathepure, B. Z., Tiedje, J. M., and Löffler, F. E. (2002). Acetate versus hydrogen as direct electron donors to stimulate the microbial reductive dechlorination process at chloroethene-contaminated sites. *Environmental Science & Technology*, 36(18):3945–3952.
- [54] Henderson, A. D. and Demond, A. H. (2007). Long-term performance of zero-valent iron permeable reactive barriers: A critical review. *Environmental Engineering Science*, 24(4): 401-423.
- [55] Hill, M. C. (1998). *Methods and guidelines for effective model calibration*. US Geol. Surv., Denver, Colorado, USA.
- [56] Hill, M. C. (2006). The practical use of simplicity in developing ground water models. *Ground Water*, 44(6):775–781.
- [57] Hill, M. C., Cooley, R. L., and Pollock, D. W. (1998). A controlled experiment in ground water flow model calibration. *Ground Water*, 36(3):520–535.
- [58] Hill, M. C. and Tiedeman, C. R. (2007). *Effective groundwater model calibration: with analysis of data, sensitivities, predictions, and uncertainty*. John Wiley and Sons.
- [59] Hoeksema, R. J. and Kitanidis, P. K. (1984). An application of the geo-statistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resources Research*, 20(7):1003–1020.

-
- [60] Huber, P. J. and Ronchetti, E. M. (2009). *Robust statistics*. Hoboken. John Wiley, N. J.
- [61] Hunt, R. J., Doherty, J., and Tonkin, M. J. (2007). Are models too simple? arguments for increased parameterization. *Ground Water*, 45(3):254–262.
- [62] Hunt, R. J., Luchette, J., Schreuder, W. A., Rumbaugh, J. O., Doherty, J., Tonkin, M. J., and Rumbaugh, D. B. (2010). Using a cloud to replenish parched groundwater modeling efforts. *Ground Water*, 48(3):360–365.
- [63] Jeen, S.-W. (2005). *Effects of Mineral Precipitation on Long-Term Performance of Granular Iron Permeable Reactive Barriers: Column Experiments and Numerical Simulation*. PhD thesis.
- [64] Jeen, S.-W., Gillham, R. W., and Blowes, D. W. (2006). Effects of carbonate precipitates on long-term performance of granular iron for reductive dechlorination of TCE. *Environmental Science & Technology*, 40(20):6432–6437.
- [65] Jeen, S.-W., Jambor, J. L., Blowes, D. W., and Gillham, R. W. (2007a). Precipitates on granular iron in solutions containing calcium carbonate with trichloroethene and hexavalent chromium. *Environmental Science & Technology*, 41(6):1989–1994.
- [66] Jeen, S.-W., Mayer, K. U., Gillham, R. W., and Blowes, D. W. (2007b). Reactive transport modeling of trichloroethene treatment with declining reactivity of iron. *Environmental Science & Technology*, 41(4):1432–1438.
- [67] Jeffers, P. M., Ward, L. M., Woytowitch, L. M., and Wolfe, N. L. (1989). Homogeneous hydrolysis rate constants for selected chlorinated methanes, ethanes, ethenes, and propanes. *Environmental Science & Technology*, 23(8):965–969.
- [68] Jin suk, O., Jeen, S.-W., Gillham, R. W., and Gui, L. (2009). Effects of initial iron corrosion rate on long-term performance of iron permeable reactive barriers: Column experiments and numerical simulation. *Journal of Contaminant Hydrology*, 103(3-4):145–156.
- [69] Johnson, A. I. (1963). *Specific yield: compilation of specific yield for various materials*. US Government Printing Office.
- [70] Johnson, C. D. and Truex, M. J. (2006). *RT3D reaction modules for natural and enhanced attenuation of Chloroethanes, Chloroethenes, Chloromethanes, and daughter products*. Pacific Northwest National Laboratory.
- [71] Jung, Y., Ranjithan, R. S., and Mahinthakumar, G. (2011). Subsurface characterization using a D-optimality based pilot point method. *Journal of Hydroinformatics*, 13(4):775–793.

- [72] Kamolpornwijit, W., Liang, L., Moline, G. R., Hart, T., and West, O. R. (2004). Identification and quantification of mineral precipitation in Fe0 filings from a column study. *Environmental Science & Technology*, 38(21):5757–5765.
- [73] Kamolpornwijit, W., Liang, L., West, O. R., Moline, G. R., and Sullivan, A. B. (2003). Preferential flow path development and its influence on long-term PRB performance: column study. *Journal of Contaminant Hydrology*, 66(3):161–178.
- [74] Karlsen, R., Smits, F., Stuyfzand, P., Olsthoorn, T., and van Breukelen, B. (2012). A post audit and inverse modeling in reactive transport: 50 years of artificial recharge in the Amsterdam water supply dunes. *Journal of Hydrology*, 454-455(0):7 – 25.
- [75] Kennedy, M. C. and O’Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464.
- [76] Kipp, K. L. (1997). Guide to the revised heat and solute transport simulator: HST3D version 2. Water resources investigation report. *US Geol. Surv.*, 97-4157(4157).
- [77] Kitanidis, P. K. and Vomvoris, E. G. (1983). A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resources Research*, 19(3):677–690.
- [78] Klausen, J., Vikesland, P. J., Kohn, T., Burris, D. R., Ball, W. P., and Roberts, A. L. (2003). Longevity of granular iron in groundwater treatment processes: solution composition effects on reduction of organohalides and nitroaromatic compounds. *Environmental Science & Technology*, 37(6):1208–1218.
- [79] Köber, R., Schlicker, O., Ebert, M., and Dahmke, A. (2002). Degradation of chlorinated ethylenes by Fe0: inhibition processes and mineral precipitation. *Environmental Geology*, 41(6):644–652.
- [80] Köber, R., Welter, E., Ebert, M., and Dahmke, A. (2005). Removal of arsenic from groundwater by zerovalent iron and the role of sulfide. *Environmental Science & Technology*, 39(20):8038–8044.
- [81] Kohn, T. and Lynn Roberts, A. (2006). The effect of silica on the degradation of organohalides in granular iron columns. *Journal of Contaminant Hydrology*, 83(1-2):70–88.
- [82] Konikow, L. F. (2011). The secret to successful solute-transport modeling. *Ground Water*, 49(2):144–159.

-
- [83] Kool, J. B., Parker, J. C., and van Genuchten, M. T. (1987). Parameter estimation for unsaturated flow and transport models –A review. *Journal of Hydrology*, 91(3):255–293.
- [84] Kouznetsova, I., Bayer, P., Ebert, M., and Finkel, M. (2007). Modelling the long-term performance of zero-valent iron using a spatio-temporal approach for iron aging. *Journal of Contaminant Hydrology*, 90(1):58–80.
- [85] Kouznetsova, I., Mao, X., Robinson, C., Barry, D. A., Gerhard, J. I., and McCarty, P. L. (2010). Biological reduction of chlorinated solvents: Batch-scale geochemical modeling. *Advances in Water Resources*, 33(9):969–986.
- [86] Kowalsky, M. B., Finsterle, S., Williams, K. H., Murray, C., Commer, M., Newcomer, D., Englert, A., Steefel, C. I., and Hubbard, S. S. (2012). On parameterization of the inverse problem for estimating aquifer properties using tracer data. *Water Resources Research*, 48(6):W06535.
- [87] Kowalsky, M. B., Gasperikova, E., Finsterle, S., Watson, D., Baker, G., and Hubbard, S. S. (2011). Coupled modeling of hydrogeochemical and electrical resistivity data for exploring the impact of recharge on subsurface contamination. *Water Resources Research*, 47(2):W02509.
- [88] Kuczera, G. (1983). Improved parameter inference in catchment models: 1. evaluating parameter uncertainty. *Water Resources Research*, 19(5):1151–1162.
- [89] Laloy, E., Rogiers, B., Vrugt, J. A., Mallants, D., and Jacques, D. (2013). Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion. *Water Resources Research*, 49(5):2664–2682.
- [90] Lasaga, A. C. (1998). *Kinetic Theory in the Earth Sciences*. Princeton University Press, Princeton, New Jersey.
- [91] Lee, I.-S., Bae, J.-H., Yang, Y., and McCarty, P. L. (2004). Simulated and experimental evaluation of factors affecting the rate and extent of reductive dehalogenation of chloroethenes with glucose. *Journal of Contaminant Hydrology*, 74(1–4):313–331.
- [92] Lee, T. R. and Wilkin, R. T. (2010). Iron hydroxy carbonate formation in zerovalent iron permeable reactive barriers: Characterization and evaluation of phase stability. *Journal of Contaminant Hydrology*, 116(1–4):47–57.
- [93] Levenberg, K. (1944). A method for the solution of certain problems in least squares. *Quarterly of applied mathematics*, 2:164–168.
- [94] Li, L. and Benson, C. H. (2010). Evaluation of five strategies to limit the impact of fouling in permeable reactive barriers. *Journal of Hazardous Materials*, 181(1):170–180.

- [95] Li, L., Benson, C. H., and Lawson, E. M. (2005). Impact of mineral fouling on hydraulic behavior of permeable reactive barriers. *Ground Water*, 43(4):582–596.
- [96] Li, L., Benson, C. H., and Lawson, E. M. (2006). Modeling porosity reductions caused by mineral fouling in continuous-wall permeable reactive barriers. *Journal of Contaminant Hydrology*, 83(1):89–121.
- [97] Liang, L., Moline, G. R., Kamolpornwijit, W., and West, O. R. (2005). Influence of hydrogeochemical processes on zero-valent iron reactive barrier performance: A field investigation. *Journal of Contaminant Hydrology*, 78(4):291–312.
- [98] Lichtner, P. C. (1985). Continuum model for simultaneous chemical reactions and mass transport in hydrothermal systems. *Geochimica et Cosmochimica Acta*, 49(3):779–800.
- [99] Lichtner, P. C. (1996). Continuum formulation of multicomponent-multiphase reactive transport. *Reviews in Mineralogy and Geochemistry*, 34(1):1–81.
- [100] Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318):399–402.
- [101] Lu, D., Ye, M., and Hill, M. C. (2012). Analysis of regression confidence intervals and Bayesian credible intervals for uncertainty quantification. *Water Resources Research*, 48(9):W09521.
- [102] Ma, R., Zheng, C., Liu, C., Greskowiak, J., Prommer, H., and Zachara, J. M. (2014). Assessment of controlling processes for field-scale uranium reactive transport under highly transient flow conditions. *Water Resources Research*, 50(2):WR013835.
- [103] Mabey, W. and Mill, T. (1978). Critical review of hydrolysis of organic compounds in water under environmental conditions. *Journal of Physical and Chemical Reference Data*, 7(2):383–415.
- [104] Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Sciences (Calcutta)*, 2:49–55.
- [105] Manoli, G., Chambon, J. C., Bjerg, P. L., Scheutz, C., Binning, P. J., and Broholm, M. M. (2012). A remediation performance model for enhanced metabolic reductive dechlorination of chloroethenes in fractured clay till. *Journal of Contaminant Hydrology*, 131(1):64–78.
- [106] Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society of Industrial and Applied Mathematics*, 11(2):431–441.

-
- [107] Matheson, L. J. and Tratnyek, P. G. (1994). Reductive dehalogenation of chlorinated methanes by iron metal. *Environmental Science & Technology*, 28(12):2045–2053.
- [108] Matott, L. S. and J., R. A. (2008). Calibration of complex subsurface reaction models using a surrogate-model approach. *Advances in Water Resources*, 31(12):697–1707.
- [109] Mayer, K. U., Blowes, D. W., and Frind, E. O. (2001). Reactive transport modeling of an in situ reactive barrier for the treatment of hexavalent chromium and trichloroethylene in groundwater. *Water Resources Research*, 37(12):3091–3103.
- [110] Mayer, K. U., Frind, E. O., and Blowes, D. W. (2002). Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions. *Water Resources Research*, 38(9):1174.
- [111] Maymó-Gatell, X., Anguish, T., and Zinder, S. H. (1999). Reductive dechlorination of chlorinated ethenes and 1,2-dichloroethane by "Dehalococcoides ethenogenes" 195. *Applied and Environmental Microbiology*, 65(7):3108–3113.
- [112] McWhorter, D. B. and Sunada, D. K. (1977). *Ground water hydrology and hydraulics*. Water Resources Publication, Fort Collins, Colorado, USA.
- [113] Medina, A. and Carrera, J. (1996). Coupled estimation of flow and solute transport parameters. *Water Resources Research*, 32(10):3063–3076.
- [114] Mertens, J., Madsen, H., Kristensen, M., Jacques, D., and Feyen, J. (2005). Sensitivity of soil parameters in unsaturated zone modelling and the relation between effective, laboratory and in situ estimates. *Hydrological Processes*, 19(8):1611–1633.
- [115] Morkin, M., Devlin, J. F., Barker, J. F., and Butler, B. J. (2000). In situ sequential treatment of a mixed contaminant plume. *Journal of Contaminant Hydrology*, 45(3):283–302.
- [116] Morrison, S. (2003). Performance evaluation of a permeable reactive barrier using reaction products as tracers. *Environmental Science & Technology*, 37(10):2302–2309.
- [117] Noubactep, C. (2009). An analysis of the evolution of reactive species in Fe⁰/H₂O systems. *Journal of Hazardous Materials*, 168(2):1626–1631.
- [118] Ory, J. and Pratt, R. (1995). Are our parameter estimators biased? the significance of finite-difference regularization operators. *Inverse Problems*, 11(2):397.

- [119] Pankow, J. F. and Cherry, J. A. (1996). *Dense chlorinated solvents and other DNAPLs in groundwater: History, behavior, and remediation*. Waterloo Press, Portland, Oregon.
- [120] Parkhurst, D. and Appelo, C. (1999). User's guide to PHREEQC (Version 2): a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. *US Geological Survey Water-Resources Investigations Report*, 312.
- [121] Parkhurst, D., Kipp, K., and Charlton, S. (2010). PHAST Version 2 a program for simulating groundwater flow, solute transport, and multicomponent geochemical reactions. *US Geological Survey techniques and methods*, 6:A35.
- [122] Phillips, D. H., Watson, D. B., Roh, Y., and Gu, B. (2003). Mineralogical characteristics and transformations during long-term operation of a zerovalent iron reactive barrier. *Journal of Environmental Quality*, 32(6):2033–2045.
- [123] Poeter, E., Hill, M., Banta, E., Mehl, S., and Christensen, S. (2005). UCODE_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation. *US geological survey techniques and methods*, 6:A11.
- [124] Pollock, D. and Cirpka, O. A. (2012). Fully coupled hydrogeophysical inversion of a laboratory salt tracer experiment monitored by electrical resistivity tomography. *Water Resources Research*, 48(1):W01505.
- [125] Prommer, H., Anneser, B., Rolle, M., Einsiedl, F., and Griebler, C. (2009). Biogeochemical and isotopic gradients in a BTEX/PAH contaminant plume: Model-based interpretation of a high-resolution field data set. *Environmental Science & Technology*, 43(21):8206–8212.
- [126] Prommer, H., Aziz, L. H., Bolaño, N., Taubald, H., and Schüth, C. (2008). Modelling of geochemical and isotopic changes in a column experiment for degradation of tce by zero-valent iron. *Journal of Contaminant Hydrology*, 97(1-2):13–26.
- [127] Prommer, H., Barry, D., and Zheng, C. (2003). MODFLOW/MT3DMS-based reactive multicomponent transport modeling. *Ground Water*, 41(2):247–257.
- [128] RamaRao, B. S., LaVenue, A. M., De Marsily, G., and Marietta, M. G. (1995). Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 1. theory and computational experiments. *Water Resources Research*, 31(3):475–493.

-
- [129] Reardon, E. J. (1995). Anaerobic corrosion of granular iron: Measurement and interpretation of hydrogen evolution rates. *Environmental Science & Technology*, 29(12):2936–2945.
- [130] Reichert, P. and Schuwirth, N. (2012). Linking statistical bias description to multiobjective model calibration. *Water Resources Research*, 48(9):WR011391.
- [131] Rittmann, B. E. and McCarty, P. L. (2001). *Environmental biotechnology*. McGraw-Hill, New York.
- [132] Rotter, B. E., Barry, D. A., Gerhard, J. I., and Small, J. S. (2011). Modeling the effectiveness of U(VI) biomineralization in dual-porosity porous media. *Journal of Hydrology*, 402(1):14–24.
- [133] Sarr, D. (2001). Zero-valent-iron permeable reactive barriers - how long will they last? *Remediation*, 11(2):1–18.
- [134] Schoups, G. and Vrugt, J. A. (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, 46(10):W10531.
- [135] Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M. (2005). *Environmental organic chemistry*. John Wiley and Sons.
- [136] Şengör, S. S. and Ünlü, K. (2013). Modeling contaminant transport and remediation at an acrylonitrile spill site in turkey. *Journal of Contaminant Hydrology*, 150:77–92.
- [137] Shah, N., Nachabe, M., and Ross, M. (2007). Extinction depth and evapotranspiration from ground water under selected land covers. *Ground Water*, 45(3):329–338.
- [138] Shawn Matott, L. and Rabideau, A. J. (2008). Calibration of subsurface batch and reactive-transport models involving complex biogeochemical processes. *Advances in Water Resources*, 31(2):269–286.
- [139] Simunek, J., Van Genuchten, M. T., and Sejna, M. (2005). The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. *University of California-Riverside Research Reports*, 3:1–240.
- [140] Simunek, J., Wendroth, O., and van Genuchten, M. T. (1999). Estimating unsaturated soil hydraulic properties from laboratory tension disc infiltrometer experiments. *Water Resources Research*, 35(10):2965–2979.
- [141] Steefel, C. and Lasaga, A. C. (1994). A coupled model for transport of multiple chemical species and kinetic precipitation/dissolution reactions

- with application to reactive flow in single phase hydrothermal systems. *American Journal of Science*, 294:529.
- [142] Suarez, M. P. and Rifai, H. S. (1999). Biodegradation rates for fuel hydrocarbons and chlorinated solvents in groundwater. *Bioremediation Journal*, 3(4):337–362.
- [143] Sun, N.-Z. (1994). *Inverse Problems in groundwater modeling*. Kluwer Academic.
- [144] Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S. W., and Srikanthan, S. (2009). Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. *Water Resources Research*, 45(12):WR006825.
- [145] Tikhonov, A. N. and Arsenin, V. Y. (1977). *Solutions of ill-posed problems*. Winston, Washington, DC.
- [146] Todd, D. K. (1959). *Groundwater hydrology*. John Wiley and Sons.
- [147] Tonkin, M. and Doherty, J. (2009). Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. *Water Resources Research*, 45(12):W00B10.
- [148] Tonkin, M. J. and Doherty, J. (2005). A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resources Research*, 41(10):W10412.
- [149] Valstar, J. R., McLaughlin, D. B., te Stroet, C. B. M., and van Geer, F. C. (2004). A representer-based inverse method for groundwater flow and transport applications. *Water Resources Research*, 40(5):W05116.
- [150] Van Nooten, T., Springael, D., and Bastiaens, L. (2008). Positive impact of microorganisms on the performance of laboratory-scale permeable reactive iron barriers. *Environmental Science & Technology*, 42(5):1680–1686.
- [151] Vandenbohede, A., Wallis, I., Houtte, E., and Ranst, E. (2013). Hydrogeochemical transport modeling of the infiltration of tertiary treated wastewater in a dune area, Belgium. *Hydrogeology Journal*, 21(6):1307–1321.
- [152] Vikesland, P. J., Klausen, J., Zimmermann, H., Lynn Roberts, A., and Ball, W. P. (2003). Longevity of granular iron in groundwater treatment processes: changes in solute transport properties over time. *Journal of Contaminant Hydrology*, 64(1):3–33.
- [153] Vogel, T. M. and McCarty, P. L. (1987). Abiotic and biotic transformations of 1,1,1-trichloroethane under methanogenic conditions. *Environmental Science & Technology*, 21(12):1208–1213.

-
- [154] Vrugt, J. A., Bouten, W., Gupta, H. V., and Hopmans, J. W. (2003). Toward improved identifiability of soil hydraulic parameters: On the selection of a suitable parametric model. *Vadose Zone Journal*, 2(1):98–113.
- [155] Vrugt, J. A., Braak, C. J. F. t., Diks, C. G. H., Robinson, B. A., Hyman, J. M., and Higdon, D. (2009a). Accelerating Markov chain Monte Carlo simulation by differential evolution with self adaptive randomized subspace sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, 10(3):273–290.
- [156] Vrugt, J. A., Robinson, B. A., and Hyman, J. M. (2009b). Self-adaptive multimethod search for global optimization in real-parameter spaces. *IEEE Transactions on Evolutionary Computation*, 13(2):243–259.
- [157] Vrugt, J. A., Stauffer, P. H., Wöhling, T., Robinson, B. A., and Vesselinov, V. V. (2008). Inverse modeling of subsurface flow and transport properties: A review with new developments. *Vadose Zone Journal*, 7(2):843–864.
- [158] Wagner, B. J. (1992). Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modelling. *Journal of Hydrology*, 135(1-4):275–303.
- [159] Wang, J. and Farrell, J. (2003). Investigating the role of atomic hydrogen on chloroethene reactions with iron using tafel analysis and electrochemical impedance spectroscopy. *Environmental Science & Technology*, 37:3891–3896.
- [160] Wiedemeier, T. H., Rifai, H. S., Newell, C. J., and Wilson, J. T. (2007). Appendix B: RCBA chemical database, natural attenuation of fuels and chlorinated solvents in the subsurface. *Natural Attenuation of Fuels and Chlorinated Solvents in the Subsurface*, pages 593–603.
- [161] Wilkin, R. T. and Puls, R. W. (2003). Capstone report on the application, monitoring, and performance of permeable reactive barriers for ground-water remediation: Volume 1: Performance evaluations at two sites. Technical report, DTIC document.
- [162] Wilkin R.T., Puls R.W., S. G. (2003). Long-term performance of permeable reactive barriers using zero-valent iron: Geochemical and microbiological effects. *Ground Water*, 41(4):493–503.
- [163] Wolery, T. J. and Daveler, S. A. (1992). *EQ6, a computer program for reaction path modeling of aqueous geochemical systems: theoretical manual, user's guide and related documentation (Version 7.0)*. Lawrence Livermore Laboratory, University of California.
- [164] Wüst, W. F., Köber, R., Schlicker, O., and Dahmke, A. (1999). Combined zero-and first-order kinetic model of the degradation of TCE and

- cis-DCE with commercial iron. *Environmental Science & Technology*, 33(23):4304–4309.
- [165] Yabusaki, S., Cantrell, K., Sass, B., and Steefel, C. (2001). Multicomponent reactive transport in an in situ zero-valent iron cell. *Environmental Science & Technology*, 35(7):1493–1503.
- [166] Yabusaki, S. B., Fang, Y., Williams, K. H., Murray, C. J., Ward, A. L., Dayvault, R. D., Waichler, S. R., Newcomer, D. R., Spane, F. A., and Long, P. E. (2011). Variably saturated flow and multicomponent biogeochemical reactive transport modeling of a uranium bioremediation field experiment. *Journal of Contaminant Hydrology*, 126(3):271 – 290.
- [167] Yeh, W. W.-G. (1986). Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resources Research*, 22(2):95–108.
- [168] Yoon, H., Hart, D. B., and McKenna, S. A. (2013). Parameter estimation and predictive uncertainty in stochastic inverse modeling of groundwater flow: Comparing null-space Monte Carlo and multiple starting point methods. *Water Resources Research*, 49(1):536–553.
- [169] Zhang, G., Lu, D., Ye, M., Gunzburger, M., and Webster, C. (2013). An adaptive sparse-grid high-order stochastic collocation method for Bayesian inference in groundwater reactive transport modeling. *Water Resources Research*, 49(10):6871–6892.
- [170] Zhang, Y. and Gillham, R. W. (2005). Effects of gas generation and precipitates on performance of fe0 PRBs. *Ground Water*, 43(1):113–121.

Acknowledgments

Five years and a half ago I left my small village in the Veneto countryside to start a PhD in Delft. At the beginning I thought that getting a PhD was only a matter of knowledge and predisposition for scientific reasoning. It has taken me some time to realize that this is only a part of the whole story. Another part requires good communication with colleagues and that kind of persistence that comes from the encouragement of people who love you.

First of all I want to thank Gerrit, for reviewing my research and for providing useful suggestions. Gerrit, thank you for listening to my complex reasoning for such a long time. I am sure that the freedom you gave me ensured that I learned something useful from my mistakes. Thanks to Nick van de Giesen for providing me with the opportunity to do a PhD, thanks to my colleagues for the useful discussions and thanks to Hanneke, Betty and Luz for being friendly. In particular, I want to thank Koen and Ruud for sharing the office and moments of the PhD life with me.

Thanks to Alberto Barausse, Luca Palmeri and Yuri Artioli for encouraging me at the beginning of my PhD. Thanks to Milica Velimirovic, Leen Bastiaens, Queenie Simons for explaining patiently how your lab experiments were performed. Thanks to Hans Sapion for answering my questions and sharing some practical knowledge on contaminated sites management.

Thanks to Ronnie Schwede for introducing me to the high performance computing word, Olaf Cirpka for shedding some light on my research and the other colleagues of Tübingen University for the short but intense period I spent in Germany. I am grateful to my mother for giving me the opportunity to study engineering and to my sister and my brother.

Finally I want to thank the person who encouraged me the most, my girlfriend Marianna. Thanks Marianna for moving to the Netherlands to stay with me, I hope I am giving you the same support and love you are giving me every day.

About the author

Luca Carniato was born on 15 January 1982 in Treviso, Italy. He obtained his high school diploma in quantity surveying and construction science in September 2001. Afterward, he studied at the University of Padova, obtaining his bachelor degree in Environmental Engineering in July 2003 and his master degree in May 2007. In his master thesis, Luca estimated nitrogen and phosphorus fluxes in the Adriatic Sea using a steady state model based on salinity balances and a hydrodynamic model. This work was performed under the supervision of Luca Palmeri at the Environmental Systems Analysis Lab (LASA).

The collaboration with LASA continued after the master, where Luca worked as consultant for the National Institute of Oceanography and Experimental Geophysics of Trieste and the Regional Environmental Agency of Padova. During this period, Luca developed quantitative models for the prediction of the riverine nutrient load into the Adriatic sea and analyzed data from a real-time monitoring network of the Venice lagoon.

In June 2009 Luca joined the water management section at TU Delft as a PhD researcher. His research was founded by a large scale European project on remediation technologies (AQUAREHAB).

Before finalizing his PhD, Luca started working for Petrotechnical Data Systems in Rijswijk, Netherlands. In his spare time, Luca enjoys playing guitar and traveling.

Publications

Journal papers

Carniato L., Schoups G., Seuntjens P., Van Nooten T., Simons Q., and Bastiaens L. (2012). Predicting longevity of iron permeable reactive barriers using multiple iron deactivation models, *Journal of Contaminant Hydrology*, 142–143(0), 93–108.

Carniato, L., Schoups, G. and van de Giesen, N. (2014). Inference of reactive transport model parameters using a Bayesian multivariate approach. *Water Resources Research*, 50(8): 6406–6427.

Carniato L., Schoups G., van De Giesen N. and Seuntjens P., Bastiaens L., and Sapon H. (2014). Highly parameterized inversion of groundwater reactive transport for a complex field site (under review in *Journal of Contaminant Hydrology*).

Velimirovic, M., Carniato, L., Simons, Q., Schoups, G., Seuntjens, P., Bastiaens, L. (2014). Corrosion rate estimations of microscale zerovalent iron particles via direct hydrogen production measurements. *Journal of Hazardous Materials*, 270(0): 18–26, 2014.

Book chapters

Copetti, D., Carniato, L., Crise, A., Guyennon, N., Palmeri, L., Pisacane, G., Tartari, G. (2013). Impacts of climate change on water quality. In *Regional Assessment of Climate Change in the Mediterranean*, pp. 307–332, Springer, Netherlands.

Conference proceedings

Carniato L., Schoups G., Bastiaens L., Seuntjens P. Simplified models to predict long term efficiency of Fe⁰ permeable reactive barriers, Second European

Symposium on Water Technologies & Management, 20-21 November 2013, Leuven, Belgium.

Carniato L., Schoups G., Sapion H., Simons Q., Seuntjens P. and Bastiaens L. Predictive assessment of groundwater remediation efficiency for a permeable reactive barrier site, First European Symposium on Remediation Technologies and their Integration in Water Management, 25-26 September 2012, Barcelona, Spain.

Carniato L., Schoups G., Seuntjens P. and Bastiaens L. Long term efficiency assessment of permeable reactive barrier using lab-scale column tests, AQUARE-HAB Second open user meeting, 18 January 2011, Copenhagen, Denmark.

Carniato L., Schoups G., Seuntjens P. and Bastiaens L. Calibration of a subsurface reactive transport model to assess long-term efficiency of a Permeable Reactive Barrier, Geophysical Research Abstracts, EGU General Assembly 2011, Vol. 13, EGU2011-4051.

Carniato L., Schoups G., Seuntjens P. and Bastiaens L. Geochemical modelling of Fe⁰ permeable reactive barriers. Proceedings of the 4th International symposium on permeable reactive barriers and reactive zones, July 6-8 2010, Antwerp, Belgium.

Carniato L., Schoups G., Seuntjens P. and Bastiaens L. Modeling Fe⁰ permeable reactive barriers for groundwater remediation, Geophysical Research Abstracts. EGU General Assembly 2010, Vol. 12, EGU2010-6009-1.