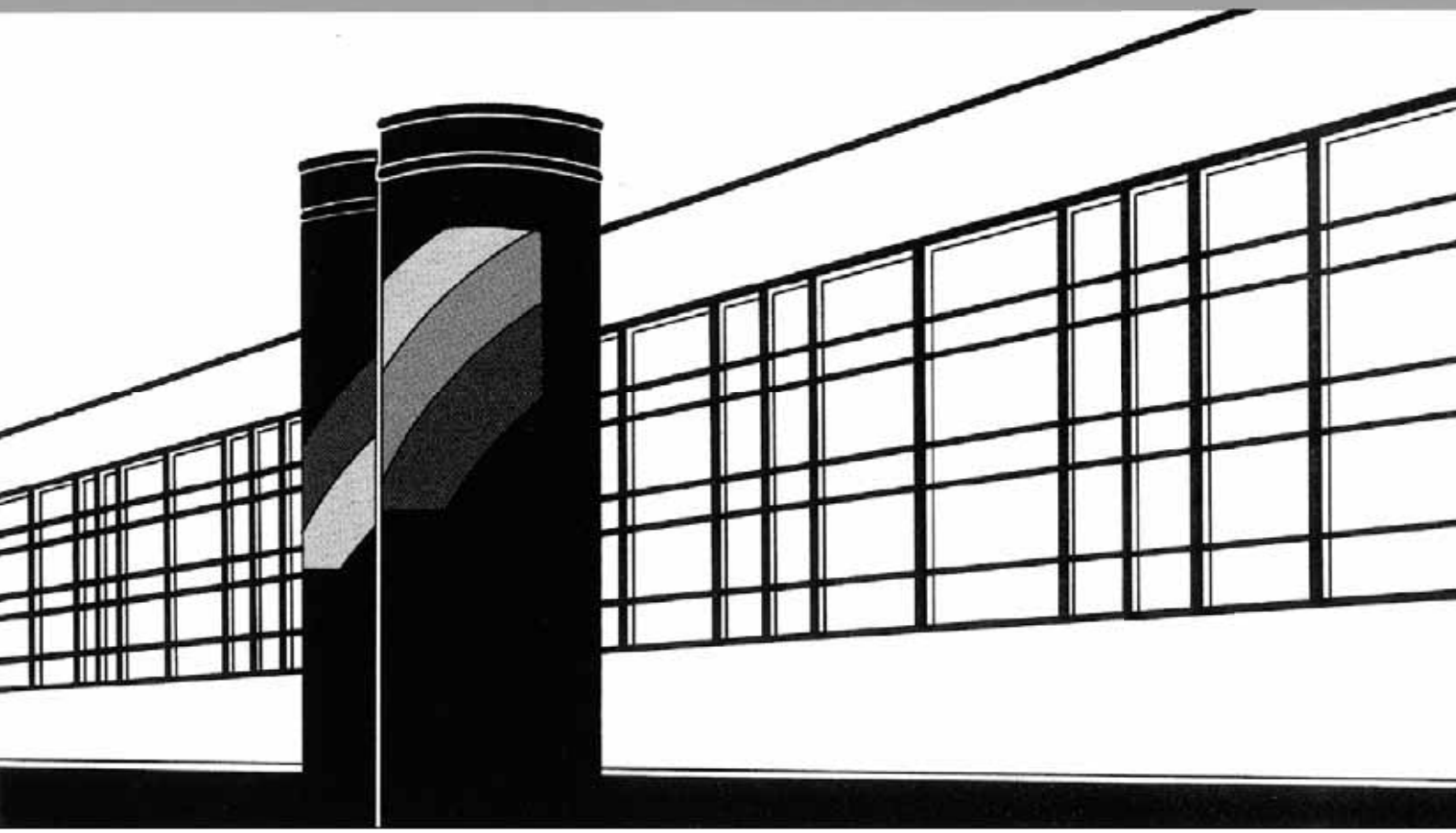
Universität Stuttgart

Institut für Wasser- und Umweltsystemmodellierung

Mitteilungen



Heft 224 Philipp Christoph Leube

Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application

Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application

Von der Fakultät Bau- und Umweltingenieurwissenschaften und dem
Stuttgart Research Centre for Simulation Technology
der Universität Stuttgart zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

Vorgelegt von

Philipp Christoph Leube

aus Pforzheim

Hauptberichter:	Jun.-Prof. Dr.-Ing. Wolfgang Nowak
Mitberichter:	Prof. Dr.-Ing. Rainer Helmig
Mitberichter:	Prof. Xavier Sanchez-Vila

Tag der mündlichen Prüfung: 24. Oktober 2013

Institut für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart
2013

Heft 224

Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application

von
Dr.-Ing.
Philipp Christoph Leube

Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart

D93 Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application

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

Leube, Philipp Christoph:
Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application. Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart. - Stuttgart: Institut für Wasser- und Umweltsystemmodellierung, 2013

(Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart: H. 224)
Zugl.: Stuttgart, Univ., Diss., 2013
ISBN 978-3-942036-28-3

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

Herausgegeben 2013 vom Eigenverlag des Instituts für Wasser- und Umwelt-systemmodellierung

Druck: Document Center S. Kästl, Ostfildern

Danksagung

Mein besonderer Dank gilt meinem Betreuer und Hauptberichter Wolfgang Nowak, der mich nicht nur in die wundersame Welt der stochastischen Hydrogeologie eingeführt hat, sondern auch ein allzeit konstruktiver und positivistischer Inspirator war und sich gleichzeitig für kein noch so unsinnig schönes Nebengesprächsthema zu schade war. Ebenso möchte ich mich bei Rainer Helmig und Xavier Sanchez-Vila für ihren Mitbericht bedanken. Letzterem besonders auch für seine Gastfreundschaft, mich gleich zweimal in seiner Arbeitsgruppe in Barcelona auf zu nehmen. Weiter möchte ich mich bei SimTech für die Finanzierung meiner Arbeit bedanken. Ich empfand es oft als Privileg, wissenschaftliches Lernen mit internationalen Austauschangeboten verbinden zu dürfen. Wenn man mehr Stunden am Tag im Büro verbringt als zu Hause, muss man sich zwangsweise mit seinen Kollegen arrangieren. Daher möchte ich meinem großen Kollegium danken, bei welchem dies ganz einfach funktioniert hat. Besonders angenehm war dabei die Büro-, Frei-, und Reisezeit mit Jonas und Andreas. Nicht zuletzt danke ich meinen Eltern, die vor mehr als drei Jahrzehnten der Grundstein für diese Arbeit gelegt haben. Teil einer Großfamilie zu sein, ist etwas wunderbares. Der letzter Dank gilt meinem nicht-technischen Umfeld Franziska, die mir wohl nicht nur durch ihre künstlerische Welt den Kopf verdreht hat.

Contents

List of Figures

List of Tables

Notation

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

Symbol	Definition	Dimension
Greek Letters:		
α_k	Arbitrary temporal characteristic of order k	[-]
α_L	Longitudinal dispersivity	[L]
α_T	Transversal dispersivity	[L]
β_{tot}	Capacity coefficient	[-]
γ	Slope parameter	
Γ_{Diri}	Dirichlet boundary	[-]
Γ_{Neu}	Neumann boundary	[-]
κ	Shape parameter of Matérn family	[-]
λ	Isotropical correlation length of quantity specified by subscript	[L]
μ	Mean value of quantity specified by subscript	[-]
ν	Slope parameter	[-]
ω	Moment of memory function	[-]
σ	Standard deviation of quantity specified by subscript	[-]
$\boldsymbol{\theta}$	Vector of uncertain structural parameters	[-]
$\boldsymbol{\xi}$	Vector of uncertain boundary parameters	[-]
ξ	Arbitrary base function	[T]
Latin Letters:		
b	Saturated thickness	[L]
c	Solute concentration of mobile domain	[M/L^3]
c_{im}	Solute concentration of immobile domain	[M/L^3]
D_e	Effective diffusion coefficient	[L^2/T]
D_m	Molecular diffusion coefficient	[L^2/T]
$\mathbf{D}$	Dispersion tensor	[L^2/T]
h	Water head	[L]
$\mathbf{k}$	Vector of uncertain conceptual parameters	[-]

k	Temporal moment order	[-]
K	Highest temporal moment order	[-]
m_k^*	k^{th} raw temporal moment	[T^k]
m_k	k^{th} normalized temporal moment	[T^k]
$m_{k,c}$	k^{th} centralized temporal moment	[T^k]
$m_{k,s}$	k^{th} standardized temporal moment	[-]
n_e	Porosity of mobile domain	[-]
n_{im}	Porosity of immobile domain	[-]
$\mathbf{n}$	Unit vector normal to interface	[-]
p	System pressure	[M/LT^2]
Q	Pumping term	[L^3/T]
r	General system response, e.g., pressure p	[-]
S	Storativity (2d)	[-]
s	Drawdown	[L]
T	Transmissivity (2d)	[L^2/T]
t	Time	[T]
$\mathbf{v}$	Seepage velocity	[L/T]
W	Source/ sink term	[L^3/T]

Operators and Other Symbols:

∂ ( )	Partial derivative
Δ	Differential operator
$\mathcal{D}$ ( )	Differential operator of arbitrary order applied to space
∇	Nabla operator
( ) $\cdot$ ( )	Scalar product
E [ ]	Expectation
Var [ ]	Variance
Cov [ ]	Covariance
L ( )	Laplace transform

Abbreviations:

ADE	Advection Dispersion Equation
BTC	Brake Through Curve
BF	Bootstrap Filter
cdf	Cummulative Density Function
DC	Drawdown Curve
FEM	Finite Element Method
FVM	Finite Volume Method
FPM	Fractured Porous Media
LC	Laplace coefficients

LT	Laplace transform
MC	Monte-Carlo (simulations)
MRMT	Multi-rate mass transfer
OD	Optimal Design
ORA	Optimal Resource Allocation
pdf	Probability Density Function
POD	Proper Orthogonal Decomposition
PreDIA	Pre-posterior Data Impact Assessor
PTRW	Particle Tracking Random Walk
RAL	Relative Averaged L^2-error
TM	Temporal Moments

Abstract

It is often stated that computational models are the only means to predict the future of hydro(geo)logical systems. Those cases that would inevitably benefit from a view into the future make reliable predictions often a challenging endeavor because they typically involve complex, coupled and dynamic systems addressing large-scale relationships. Examples include global warming [e.g., Schmidt et al., 2006], reactive transport on the catchment scale [e.g., Destouni et al., 2010], CO_2 sequestration [e.g., Kopp et al., 2009], radioactive waste disposal [e.g., Pollock, 1986; Olivella et al., 1994] or, very recently, hydraulic fracking [e.g., Myers, 2012].

Computational model complexity becomes even more drastic, when facing the ubiquitous fact that hydro(geo)logical systems can never be fully described because their parameters are incompletely known [Oreskes et al., 1994]. This leads to the need of quantifying the uncertainty and assessing the risk inherent in model predictions [Christakos, 1992; Oreskes et al., 1994; Rubin, 2003; Tartakovsky, 2007]. Incorporating field data into models helps to reduce the uncertainty of predictions [e.g., Kitanidis, 1995; Gómez-Hernández et al., 1997; Evensen, 2007; Franssen et al., 2009], but inevitably leads to so-called inverse problems [Tarantola, 1987] which imposes yet another computational burden.

For that reason, reducing computational complexity by adequate mathematical techniques has been the focus of many research efforts [Hooimeijer, 2001]. This allows the modeler or investigator to maintain the required prediction quality at a reasonably high level whilst controlling the computational demand, or, alternatively, to admit more conceptual complexity, finer resolutions or larger domains at the same computational costs, or to make brute force optimization tasks more feasible [Razavi et al., 2012].

The computational demand of complex hydro(geo)logical systems can be broken down into contributions from spatial, temporal and stochastic resolution, e.g., spatial grid resolution, time step size and number of repeated simulations dedicated to uncertainty quantification. Temporal model complexity is owed to the dynamic character of hydro(geo)logical systems which appears in time-dependent system response (prediction) curves. Examples include aquifer reactions due to recharge events, tidal pumping or changing river stages [e.g. Yeh et al., 2009], drawdown curves (DC) due to the excitation of the subsurface water level in pumping tests [e.g. Fetter, 2001], solute breakthrough curves (BTC) during the injection of water-borne tracers and contaminant spills [e.g. Fetter, 1999], or reactions of river discharge to precipitation in hydrological models [e.g. Nash and Sutcliffe, 1970].

In the last two decades, temporal model reduction has been subjected to many research efforts. The most important contributions have been based on integral transformations. This

line of research led among others to the so-called temporal moment (TM) generating equations [e.g., Harvey and Gorelick, 1995] paving the way for a variety of applications [Goode, 1996; Varni and Carrera, 1998; Cirpka and Kitanidis, 2000c; Cirpka and Nowak, 2004; Li et al., 2005; Zhu and Yeh, 2006; Luo et al., 2008; Pollock and Cirpka, 2008; Yin and Illman, 2009; Enzenhöfer et al., 2011; Molson and Frind, 2011, e.g.,]. The prevalent use of TM is not merely by chance, because they unite many advantages: Their conceptual simplicity, the possibility to jointly reduce models and observations allowing for a swift incorporation of data, and the non-intrusiveness when applying them to existing (commercial) codes.

In general, TM reduce the time-dependent governing equations to steady state and directly simulate the temporal characteristics of the system, if the equations are linear and coefficients are time-independent. This is achieved by an integral transform, projecting the dynamic system response (prediction) onto monomials in time before even solving the equations. In comparison to classical approaches of model reduction that involve orthogonal base functions, however, the base functions for TM are non-orthogonal. Also, most applications involving TM used only lower-degree TM without providing reasons for their choice.

This leads to a number of open research questions to be addressed in this thesis:

1. Does non-orthogonality impair the quality and efficiency of TM?
2. Can other temporal base functions reduce dynamic systems more efficiently than the monomials that lead to TM?
3. How can compression efficiency associated with temporal model reduction methods be quantified and how efficiently can information be compressed?
4. What is the value of temporal model reduction in competition with the computational demand of other reduced or discretized model dimensions, e.g., repetitive model runs through Monte-Carlo (MC) simulations?

The goal of this work is to analyze and assess existing techniques that reduce hydro(geo)logical models in time. This allows to answer the questions posed above and helps to better exploit the potential of temporal model reduction. To this end, I developed a fourfold approach described in the following. The individual steps are consecutive, following the mentality of typical top-down approaches.

Step (I) considers temporal model reduction from the most general perspective. To this end, I derived a formulation for temporal model reduction based on integral transformation with general base functions. This allows to classify base functions and to compare their reduction efficiency. Efficiency in this context means the ability of converting dynamic systems to simpler systems at reasonable computational times (see Chap. 5).

Step (II) assesses the most efficient model reduction techniques in terms of their compression efficiency. Here, efficiency refers to the absence of compression errors in data due to a loss-less compression. To this end, I developed a suitable method that can access the information content of data in a generic and rational way. This allows to measure and compare the performance of temporal model reduction techniques (see Chap. 6).

Step (III) judges the compression efficiency of temporal model reduction against the efficiency of other reduced or discretized model dimensions (e.g., spatial discretization or number of MC simulations). I developed a concept to combine errors from different model dimensions and assess their joint error in the light of available computational power. This allows to find the trade-off between different model dimensions and so optimally allocates computational resources (see Chap. 7).

Step (IV) employs TM in a new method for numerical upscaling of high-contrast fractured porous media. The method is based on flow-aligned blocks and uses multi-rate mass transfer (MRMT) models to parameterize unresolved sub-block heterogeneity. TM make the scale transition of parameters swift and simple (see Chap. 8).

Finally, I show the universality of my developed tools by applying them to other challenges of subsurface hydrogeology (see Chap. 9).

Based on the results of ***Steps (I) - (IV)*** I find the following conclusions most important. There is no way of temporal model reduction for dynamic systems based on arbitrary integral transforms with (non-)polynomial base functions that is better than the monomials leading to TM. This is due to the nature of TM allowing to mimic any temporal characteristic based on arbitrary polynomial base functions. Non-polynomial base functions generally lead to an inefficient or impossible evaluation of their temporal characteristics. The only remaining integral transform that even more efficiently reduces dynamic system responses is the Laplace transform (LT). However, the choice of orders (number of considered Laplace coefficients) remains unclear making LT impractical, e.g. in inverse problems.

The order of model reduction based on TM as opposed to the reduction or discretization of other model dimensions (e.g., number of repeated model runs trough MC simulations) should be carefully determined prior the model evaluation. This is because there always exists an optimal trade-off between these differently reduced or discretized model dimensions. The trade-off optimally exploits computational resources leading to the smallest joint error given a certain computational budget. Also, I found that TM help to improve the upscaling of high-contrast media leading to conceptual and computational advantages.

The tools I developed in this work can be employed in a variety of other applications. Optimal allocation of resources (from Chap. 7) can be applied to arbitrary combinations of different discretized or reduced model dimensions. The rational assessment of the explanatory power of data (from Chap. 6) is of utmost interest in optimal design of experiments. Here, my method outmatches linearized methods (such as Ensemble Kalman Filters) because linear methods fail to recognize relevant non-linear relations between potential measurement locations and the prediction goal, and hence oversample locations considered to be most informative from the limited viewpoint of linearized analysis.

Although TM have been used in different applications over the last two decades, their potential has, in my eyes, not been fully exploited. Based on my findings, I hope to encourage more studies to work with the concept of TM. Especially because the number of studies that employ TM with real data is vanishingly small, improved tests on existing data sets should be performed. Also, I hope to encourage those who limited their TM applications

to only lower order TM to consider a longer moment sequence. My study results specifically provide valuable advice for hydraulic tomography studies under transient conditions to use TM up to the fourth order. This might potentially alleviate the loss of accuracy used as argument against TM by certain authors.

Kurzfassung

Die Umweltforschung erfährt in den letzten Jahren einen drastischen Anstieg des Rechenbedarfs von Simulationsmodellen. Trotz der Vervielfachung von Rechenkapazitäten durch technischen Fortschritt stellen benötigte Rechenzeiten eine immerwährende Herausforderung dar. Diese Entwicklung wird vor allem dadurch unterstützt, dass die aktuell von Wissenschaftlern untersuchten Probleme immer komplexere Zusammenhänge auf immer größeren Simulationsgebieten abdecken, z. B. Globale Erwärmung [z. B. Schmidt et al., 2006], reaktiver Transport auf Einzugsgebietsgröße [z. B. Destouni et al., 2010], unterirdische CO_2 Speicherung [z. B. Kopp et al., 2009], atomare Endlagerung [z. B. Pollock, 1986; Olivella et al., 1994] oder seit jüngster Zeit auch hydraulische Frakturierung [z. B. Myers, 2012].

Während es vor zehn Jahren noch oft akzeptabel schien, Systeme im statischen Zustand zu beschreiben, sind heute dynamische Simulationen gefragt. Besonders die Tatsache, dass Umweltsysteme in ihren Beschaffenheiten niemals vollständig nachvollzogen werden können [Oreskes et al., 1994], macht die Berücksichtigung von Unsicherheiten in Modellvorhersagen und die damit verbundenen Risikobetrachtungen unausweichlich [Christakos, 1992; Oreskes et al., 1994; Rubin, 2003; Tartakovsky, 2007]. Dies führt zu einer weiteren Explosion von Rechenzeiten. Die Kalibrierung von Simulationen mit Messdaten ist dann die einzige Möglichkeit, die inhärenten Unsicherheiten in Modelvorhersagen zu reduzieren [Gómez-Hernández and Wen, 1998; Evensen, 2007; Franssen et al., 2009], erfordert aber die Invertierung des Modells, was ebenfalls einen weiteren Rechenaufwand mit sich bringt.

Oft ist dann der einzige Ausweg die sogenannte Modellreduktion, seit jeher ein Schwerpunkt vieler Forschungsbemühungen [Hooimeijer, 2001]. Modellreduktion reduziert Rechenzeiten, oder kann bei gleich bleibendem Rechenaufwand mehr Komplexität oder akkuratere räumlich/ zeitliche Auflösungen größerer Modellgebiete erlauben, oder macht aufwendige Optimierungsziele überhaupt erst möglich [Razavi et al., 2012]. Methoden zur Modellreduktion können generell eingeteilt werden in Methoden, die sich mit der Reduktion der räumlichen, zeitlichen oder stochastischen Dimensionalität von Modellen beschäftigen. Die zeitliche Dimensionalität von Modellen rührt z. B. von natürlichen Schwankungen im Wasserstand durch Regenereignisse oder Interaktionen mit Flüssen [Yeh et al., 2009] oder zeigt sich in hydraulischen Absenkkurven als Reaktion auf Pumptests [Fetter, 2001] oder Konzentrationsdurchbruchskurven die nach der Einleitung von Markierungsstoffen beobachtet werden können [Fetter, 1999].

In den letzten beiden Jahrzehnten stand die Modellreduktion der zeitlichen Dimensionalität im Fokus vieler Forschungsarbeiten. Die bedeutendsten Beiträge beruhten dabei auf Integraltransformationen, die zu den sog. zeitlichen Momenten (TM) [z. B. Harvey and Gorelick, 1995] führten und den Weg zu einer Vielzahl von Anwendungen ebneten [u. A., Goo-

de, 1996; Varni and Carrera, 1998; Cirpka and Kitanidis, 2000c; Cirpka and Nowak, 2004; Li et al., 2005; Zhu and Yeh, 2006; Luo et al., 2008; Pollock and Cirpka, 2008; Yin and Illman, 2009; Enzenhöfer et al., 2011; Molson and Frind, 2011]. TM vereinen viele Vorteile: Konzeptionelle Einfachheit, intuitive Aussagekraft, die Kompatibilität mit kommerziellen Programmen, sowie die Möglichkeit Modellreduktion und Datenkompression zu koppeln. Letzteres führt unter anderem auch zu einer schnellen Datenassimilierung und Modellkalibrierung.

Generell betrachtet, machen TM die zeitabhängigen Hauptgleichungen innerhalb eines Modells stationär, indem sie nicht das volle Zeitverhalten eines Systems simulieren, sondern lediglich eine Auswahl an ausgewählten Zeitcharakteristiken. Dies wird durch eine Integraltransformation erreicht, in welcher das Zeitverhalten auf bestimmte Ansatzfunktionen, sog. Monome, projiziert wird. Im Vergleich zu klassischen Modellreduktionen, die auf orthonormalen Ansatzfunktionen beruhen, setzen TM auf nicht-orthonormale Ansatzfunktionen. Auch verwenden die meisten Anwendungen nur TM niedrigen Grades, ohne diese Wahl auf konkrete Gründe zurück zu führen.

Dies führt mich zu einer Anzahl an offenen Fragen:

1. Beeinträchtigt nicht-Orthonormalität die Qualität und Effizienz von TM?
2. Gibt es zeitliche Ansatzfunktionen, die ein zeitabhängiges System effizienter reduzieren, als die Monome die zu TM führen?
3. Wie kann der Kompressionsgrad von Modellen gemessen werden und wie effizient können Daten komprimiert werden?
4. Was wiegt der Wert von zeitlicher Modellreduktion gegenüber anderen reduzierten oder diskretisierten Modelldimensionen, z. B. die Auflösung der stochastischen Dimension durch wiederholte Simulationen im Rahmen von Monte-Carlo (MC) Ansätzen.

Das Ziel dieser Arbeit ist es, Methoden zur zeitlichen Modellreduktion mit Hilfe dafür geeigneter Werkzeuge zu analysieren und zu bewerten. Dies ermöglicht es, die oben gestellten Fragen in Zusammenhang mit TM zu beantworten und das Potential zeitlicher Modellreduktion besser auszuschöpfen. Zu diesem Zweck habe ich den im Folgenden beschriebenen vierstufigen Ansatz entwickelt und angewandt. Der Ansatz ist an das klassische "Top-Down-Design" angelehnt.

Schritt (I) betrachtet zeitliche Modellreduktion so allgemein wie möglich. Dazu wurden alternative polynomiale und nicht-polynomiale Ansatzfunktionen klassifiziert und in ihrer Reduktionseffizienz verglichen. Reduktionseffizienz definiere ich hierbei als das Vermögen einer Ansatzfunktion ein Modell so zu konvertieren, dass dadurch eine signifikanter Rechenvorteil entsteht (siehe Kap. 5).

Schritt (II) fokusiert schließlich die Analyse und bewertet die effizientesten Techniken zu Modellreduktion hinsichtlich ihrer Kompressioneffizienz. Kompressioneffizienz ist dabei mit der Frage verbunden, wie effizient Informationen gespeichert werden können. Dazu wurde eine neue Methode namens PreDIA (pre-posterior data impact assessor) entwickelt und angewandt. PreDIA ist ein generisches Werkzeug, das in der

Lage ist, den Informationsgehalt von Daten und damit TM auf möglichst rationale Weise zu messen und ist damit linearisierte Methoden der Informationsverarbeitung überlegen (siehe Kap. 6).

Schritt (III) spiegelt die Ergebnisse aus den ***Schritten (I)-(II)*** im Lichte einer globalen und praktischen Betrachtung wider und vergleicht diese gegen andere reduzierte oder diskretisierte Modelldimensionen, wie z. B. die Anzahl an TM und die Anzahl an wiederholten Modelldurchläufen durch MC Simulationen. Dafür wurde eine weitere von mir entwickelte Methode zum optimalen Einsatz von Rechenleistungen (ORA) angewandt. ORA ermöglicht es, die vereinten Fehler aus verschiedenen Modelldimensionen kombiniert zu betrachten und minimiert den gemeinsamen Fehler für eine gegebene Rechenzeit. Das Fehlerminimum ist durch eine optimale Auflösung zweier Modelldimensionen gegeben (siehe Kap. 7).

Schritt (IV) adaptiert die Ergebnisse und wendet TM in einer neuen Methode zur numerischen Skalierung von hoch-kontrastierten (z. B. geklüftet-porösen) Medien an. Die Methode beruht auf einer Gitterstruktur die das Strömungsfeld nachbildet, sowie ein kinetisches Massentransfermodell (MRMT), um die anomalen Transportphänomene in geklüfteten Medien abzubilden. Durch den Einsatz von zeitlichen Momenten ist es möglich, die subskaligen Transportphänomene einzufangen und den Skalenübergang schnell und einfach zu gestalten (siehe Kap. 8).

In einem letzten Schritt unterstreiche ich den universellen Charakter, der von mir entwickelten Methoden und wende diese auf zwei weitere Fragestellungen an (siehe Kap. 9).

Basierend auf den Teilergebnissen aus allen vier Schritten kann zusammengefasst gesagt werden, dass es im Bereich zeitlicher Modellreduktionsmethoden die auf Integraltransformationen beruhen, keine effizienteren Ansatzfunktionen gibt, als die Monome, die zu zeitlichen Momenten führen. Dabei liegt es in der Natur der TM, dass diese alle Zeitcharakteristiken imitieren können, die auf polynomialen Ansatzfunktionen beruhen. Nicht-polynomiale Ansatzfunktionen führen generell zu einer ineffizienten oder gar unmöglichen Auswertung ihrer Zeitcharakteristiken. Die einzige Klasse von Ansatzfunktionen, die eine noch effizientere Berechnung ihrer Zeitcharakteristiken zulässt, ist die Laplace Transformation (LT). Jedoch bringt diese den Nachteil mit sich, dass die Ordnung (Anzahl an Zeitcharakteristiken), im Gegensatz zu TM, beliebig ist.

Der Grad an Modellreduktion, bestimmt durch die Anzahl von TM, sollte immer sorgfälltig gegen andere reduzierte oder diskretisierte Modelldimensionen (z. B. die Anzahl an wiederholten Modelldurchläufen durch sog. MC Simulationen) abgewägt werden. Hierbei lässt sich für jedes physikalische Rechenszenario eine individuelle und optimale Rechenaufteilung ermitteln, die das verfügbare Rechenaufgebot am Besten ausnutzt und den vereinten Fehler für ein gegebenes Rechenbudget minimiert. Der Einsatz von TM in einer neuen Methode zur numerischen Skalierung von geklüfteten Medien bringt konzeptionelle und rechnerische Vorteile.

Die von mir in dieser Arbeit entwickelten Werkzeuge lassen sich auf eine Vielzahl von weiteren Anwendungen übertragen. Der optimale Einsatz von Rechenleistungen (aus Kap. 7) ist dabei auf beliebige Kombinationen von Modelldimensionen anwendbar. Weiter spielt

die rationale Bemessung des Informationsgehaltes durch PreDIA (aus Kap. 6) bei der optimalen Versuchsplanung eine bedeutende Rolle. Hierbei zeigt PreDIA eindeutige Vorteile gegenüber linearisierten Methoden (z. B. Ensemble Kalman Filter), welche oft daran scheitern, dass sie den nicht-linearen Zusammenhang zwischen potentiellen Messwerten und dem Vorhersageziel nicht richtig erkennen können und daher Messorte mit ausschließlich linearem Informationsgehalt bevorzugen.

Obwohl TM in den letzten zwei Jahrzehnten durchaus eine breite Anwendung gefunden haben, ist ihr Potential in meinen Augen weitaus noch nicht voll ausgeschöpft. Basierend auf meinen Ergebnissen erhoffe ich mir eine noch breitere Anwendung von TM, besonders bei Arbeiten, welche mit realen TM arbeiten. Darüber hinaus hoffe ich, diejenigen zu ermutigen höhere TM zu benutzen, die ihre Auswahl bisher auf TM niedrigen Grades gestützt haben. Gerade im Bereich der hydraulischen Tomographie wäre die Verwendung von TM bis zum vierten Grad empfehlenswert und könnte den Verlust an Genauigkeit, welcher von gewissen Autoren als Argument gegen TM angeführt wurden, ausgleichen.

1. Introduction

1.1. Motivation and Goals

It is often stated that computational models are the only means to predict the future of hydro(geo)logical systems. Those cases that would inevitably benefit from a view into the future typically involve complex, coupled and dynamic systems addressing large-scales relationships and, hence, make reliable predictions a challenging endeavor. Examples include global warming [e.g., Schmidt et al., 2006], reactive transport on the catchment scale [e.g., Destouni et al., 2010], CO_2 sequestration [e.g., Kopp et al., 2009], radioactive waste disposal [e.g., Pollock, 1986; Olivella et al., 1994] or, very recently, hydraulic fracking [e.g., Myers, 2012].

Computational model complexity becomes even more drastic, when facing the ubiquitous fact that hydro(geo)logical systems can never be fully described because their parameters are incompletely known [Oreskes et al., 1994]. This leads to the need of quantifying the uncertainty and assessing the risk inherent in model predictions [Christakos, 1992; Oreskes et al., 1994; Rubin, 2003; Tartakovsky, 2007]. Incorporating field data into models helps to reduce the uncertainty of predictions [e.g., Kitanidis, 1995; Gómez-Hernández et al., 1997; Evensen, 2007; Franssen et al., 2009], but inevitably leads to so-called inverse problems [Tarantola, 1987] which imposes yet another computational burden.

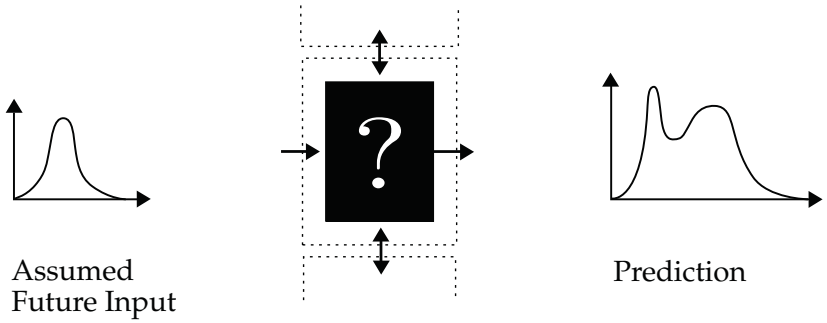


Figure 1.1.: Illustration of a complex, uncertain, dynamic system (black box) coupled with other systems (upper and lower dashed rectangles), dynamic input and prediction.

For that reason, reducing computational complexity by adequate mathematical techniques has been the focus of many research efforts [Hooimeijer, 2001]. This allows the modeler or investigator to maintain the required prediction quality at a reasonably high level whilst

controlling the computational demand, or, alternatively, to admit more conceptual complexity, finer resolution or larger domains at the same computational costs, or to make brute force optimization tasks more feasible [Razavi et al., 2012].

The computational demand of complex hydro(geo)logical systems can be broken down into contributions from spatial, temporal and stochastic resolution, e.g., spatial grid resolution, time step size and number of repeated simulations dedicated to uncertainty. Temporal model complexity is owed to the dynamic character of hydro(geo)logical systems which appears in time-dependent system response (prediction) curves (see Fig. 1.1 right side). Examples include aquifer reactions due to recharge events, tidal pumping or changing river stages [e.g. Yeh et al., 2009], drawdown curves (DC) due to the excitation of the subsurface water level in pumping tests [e.g. Fetter, 2001], solute breakthrough curves (BTC) during the injection of water-borne tracers and contaminant spills [e.g. Fetter, 1999], or reactions of river discharge to precipitation in hydrological models [e.g. Nash and Sutcliffe, 1970].

In the last two decades, temporal model reduction has been subject to many research efforts. The most important contributions have been based on integral transformations. This line of research led among others to the so-called temporal moment (TM) generating equations [e.g., Harvey and Gorelick, 1995] paving the way for a variety of applications [Goode, 1996; Cirpka and Kitanidis, 2000b; Li et al., 2005]. The prevalent use of TM is not merely by chance, because they unite many advantages: Their conceptual simplicity, the possibility to jointly reduce models and observations allowing for a swift incorporation of data, and the non-intrusiveness when applying them to existing (commercial) codes.

In general, TM reduce time-dependent governing equations to steady state and directly simulate the temporal characteristics of the system, if the equations are linear and coefficients are time-independent. This is achieved by an integral transform, projecting the dynamic system response (prediction) onto monomials in time before even solving the equations. In comparison to classical approaches of model reduction that involve orthogonal base functions, however, the base functions for TM are non-orthogonal. Also, most applications involving TM used only lower-degree TM without providing reasons for their choice.

This leads to a number of open research questions to be addressed in this thesis:

1. Does non-orthogonality impair the quality and efficiency of TM?
2. Can other temporal base functions more efficiently reduce dynamic systems than the monomials that lead to TM?
3. How can compression efficiency associated with temporal model reduction methods be quantified and how efficiently can information be compressed?
4. What is the value of temporal model reduction in competition with the computational demand of other discretized or reduced model dimensions, e.g., repetitive model runs through Monte-Carlo (MC) simulations?

The goal of this work is to analyze and assess existing techniques that reduce hydro(geo)logical models in time. This allows to answer the questions posed above and helps to better exploit the potential of temporal model reduction.

1.2. State of the Art

The most powerful contribution to the reduction of model complexity in time has been made by Harvey and Gorelick [1995]. Their approach reduces dynamic models to steady state models by employing a Laplacian transformation to the time dimension. After Taylor expansion of the Laplace coefficients (LC), this allows to directly simulate characteristics of the time-dependent response curves, the so called temporal moments (TM), with steady-state equations. Alternatively, TM can be derived by projecting the time-dependent governing equations onto a series of monomials t^k of order $k = 0 \ldots K$ [e.g., Cirpka and Kitanidis, 2000b]. The generating equations for TM are steady-state equivalents of the original governing equations, and so allow swift evaluation.

However, almost all applications of TM in the field of hydro(geo)logy (a detailed review on applications ca be found in Chap. 4.4) predominantly employed low-order moments. Reservations against using TM are, among other, the loss of information in inverse modeling based on an analysis only up to the first-order TM [Zhu and Yeh, 2005; Yin and Illman, 2009]. Also, all the studies I could find rarely provide reasons for the choice of order, and none of them assessed the information lost by not looking at higher orders.

Another way of removing the time dependence within the governing equation is to directly apply Laplace transform techniques without the transition to TM. They have been proven to be suitable in forward model reduction (including reconstruction of the full time-series from simulated LC) when considering sequences of more than 10 and up to 100 LC [e.g., Sudicky, 1989]. Here, the unresolved question is both how many, and which LC are required to properly represent the system.

Besides integral transformations, increasing attention has been drawn by snapshot-based model reduction methods [Vermeulen et al., 2004; McPhee and William, 2008]. Via Proper Orthogonal Decomposition (POD) into dominant spatial patterns [Papoulis, 1984], the model is reduced to some number of orthogonal base functions in physical space with time-dependent coefficients. Within other disciplines, this method is referred to as Principal Component Analysis (PCA) [Pearson, 1901], or Karhunen Loève Transform (KLT) [Loève, 1955]. I refer to these methods as spatial reduction methods since the model is, in its proper effects, reduced in physical space while the time-related model complexity remains untouched. The scope of this work, however, is strictly limited to temporal reduction methods. This strict focus is legitimate, because reduction methods in time can be evaluated independently of spatial methods. Reduction techniques in space and in time can be arbitrarily combined, because space and time are independent coordinates.

1.3. Approach

The general approach of my work is fourfold and described in the following in more detail. The individual steps are consecutive, following the mentality of typical top-down research designs. Starting from the very general perspective in ***Step (I)*** the analysis becomes more

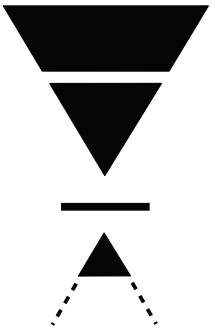


Figure 1.2.: General top-down approach illustrating the individual ***Steps (I)-(IV)*** to analyze and assess existing techniques that reduce hydro(geo)logical models in time.

focused in ***Step (II)*** and finally reflects the result in a global and practical light in ***Step (III)***. ***Step (IV)*** involves the application to a novel method.

Step (I) considers temporal model reduction from the most general perspective. To this end I derive a formulation for temporal model reduction based on integral transformation with general base functions. This allows to classify all possible base functions and to compare their reduction efficiency. Efficiency in this context means the ability of converting dynamic systems to simpler systems at reasonable computational times.

Step (II) assesses the most efficient model reduction techniques from ***Step (I)*** in terms of their compression efficiency. Here, efficiency refers to the absence of compression errors in data due to a loss-less compression. To this end, I develop a suitable method that can access the information content of data in a generic and rational way. This allows to objectively measure and compare the performance of temporal model reduction techniques.

Step (III) judges the compression efficiency of temporal model reduction against the efficiency of other reduced or discretized model dimensions (e.g., spatial discretization or number of MC simulations). I develop a concept to combine errors from different model dimensions and assess their joint error in the light of available computational power. This allows to find the trade-off between different reduced or discretized model dimensions and so optimally allocates computational resources.

Step (IV) employs TM in a new method for numerical upscaling of high-contrast fractured porous media.

Finally, I show the universality of my developed tools by applying them to other challenges of subsurface hydrogeology.

1.4. Structure of Work

The remainder of this work is organized as follows. Chap. 2 repeats the governing physical equations employed throughout the later analysis and introduces the basic features of statis-

tics and geostatistics. Chap. 3 introduces different physical scenarios to be employed repeatedly through the later work. Chap. 4 summarizes the concept of TM. Chap. 5 formulates a general approach for temporal model reduction based on integral transforms and discusses the reduction efficiency of alternative integral transforms (covering ***Step I***). Chap. 6 develops and applies a suitable method to asses the compression efficiency of temporal model reduction techniques (covering ***Step II***), and Chap. 7 features the compromise between TM and the discretization or reduction of other model dimensions (covering ***Step III***). Chap. 8 applies TM to fractured porous media (covering ***Step IV***) and Chap. 9 shows the generic nature of the methods developed in Chap. 6 and Chap. 7 by applying them to other tasks of subsurface hydrogeology.

The structure of the present work is partly reflected in a series of papers published by the author. Chap. 5 is based on the analysis in Leube et al. [2012b], whereas Chap. 6 falls back on the method developed in Leube et al. [2012a]. Chap. 7 is taken from the concept published in Leube et al. [2013]. The applications in Leube et al. [2012a] and Leube et al. [2013] both make up Chap. 9.

2. Governing Equations and Basic Methods

"State of the art" tools aiming to predict the complex behavior of hydro(geo)logical systems cope, among others, with two major challenges. There are (1) to describe as precisely as necessary the physical processes on all scales, and (2) to adequately fill the ubiquitous lack of knowledge on (hydro)geological properties. As my work touches both challenges, I dedicate this chapter to:

1. introduce the governing equations of groundwater flow and solute transport (Sec. 2.1), all of them repeatedly employed throughout the later derivations (Chap. 5) and applications (Chaps. 6-8)
2. refresh the geostatistical working hypothesis (Sec. 2.4) also used heavily throughout my work. For the sake of completeness and due to their familiarity to the temporal moments (TM) that will be introduced in Chap. 4, I briefly touch the fundamentals of descriptive statistics (Sec. 2.2) and theoretical distribution functions (Sec. 2.3) including their statistical moments. Also, I repeat Bayes Theorem in Sec. 2.5 and Bootstrap Filters in Sec. 2.6.

2.1. Governing Equations

In this section, I provide the governing equations of groundwater flow and solute transport, as well as the extension towards multi-rate mass-transfer (MRMT) [Haggerty and Gorelick, 1995].

2.1.1. Groundwater Flow

Two-dimensional flow in confined aquifers

Two-dimensional (depth-averaged) flow in confined aquifers can be described by the following partial differential equation [Bear, 1972]:

$$S\frac{\partial h}{\partial t} - \nabla \cdot (T\,\nabla h) = W - Q, \tag{2.1}$$

where S [-] is the specific storage coefficient, T [L^2/T] is the locally isotropic and depth-integrated transmissivity with $T = Kb$, where K [L/T] is the hydraulic conductivity and b [L] the saturated thickness. h [L] is the hydraulic head, W [L/T] is an internal volumetric

source/ sink term and Q [L^3/T] is the pumping rate, for example if the system is excited by a well. Generic initial and boundary conditions are

$$h = h_0 \text{ at } t = t_0 \tag{2.2}$$

$$h = h_{\text{Diri}} \text{ on } \Gamma_{\text{Diri}} \, \forall t \tag{2.3}$$

$$-\mathbf{n} \cdot (T \nabla h) = q_{\text{Neu}} \text{ on } \Gamma_{\text{Neu}} \, \forall t, \tag{2.4}$$

where h_0 [L] is the head prior to excitation, t [T] is time, t_0 [T] is the time at which the consideration starts (e.g., when pumping starts), Γ_{Diri} and Γ_{Neu} denote Dirichlet and Neumann boundaries with fixed-head Dirichlet head h_{Diri} [L] and constant specific Neumann fluxes q_{Neu} [L/T], respectively, and $\mathbf{n}$ is the unit vector normal to the boundaries pointing outward. No-flow boundaries are given by $q_{\text{Neu}} = 0$.

For constant-rate pumping $Q = \text{const.}$, time-independent source/ sink term W and $t \to \infty$, the system response $h(t \to \infty)$ reaches a state of equilibrium. This is also known as the steady-state for flow and is described by Eq. (2.5) [Bear, 1972]:

$$-\nabla \cdot (T \nabla h) = W - Q, \tag{2.5}$$

with generic boundary conditions

$$h = h_{\text{Diri}} \text{ on } \Gamma_{\text{Diri}} \tag{2.6}$$

$$-\mathbf{n} \cdot (T \nabla h) = q_{\text{Neu}} \text{ on } \Gamma_{\text{Neu}}. \tag{2.7}$$

For $Q = 0$ (no pumping), the steady-state $h(Q = 0)$ can be considered as initial head h_0 in Eq. (2.4) prior to pumping, and the drawdown s [L] is given by

$$s = h_0 - h. \tag{2.8}$$

For time-independent boundary conditions and source/sink term W, Eq. (2.1) can be subtracted from h_0 the corresponding steady-state arriving at the drawdown equation during pumping at any (possibly time-dependent pumping rate) $Q(t)$:

$$S \frac{\partial s}{\partial t} - \nabla \cdot (T \nabla s) = W - Q(t), \tag{2.9}$$

with initial and general boundary conditions

$$s = 0 \text{ at } t = t_0 \tag{2.10}$$

$$s = 0 \text{ on } \Gamma_{\text{Diri}} \, \forall t \tag{2.11}$$

$$-\mathbf{n} \cdot (T \nabla s) = q_{\text{Neu}} \text{ on } \Gamma_{\text{Neu}} \, \forall t. \tag{2.12}$$

Two-dimensional flow in unconfined aquifers

Opposed to the confined case, the saturated thickness changes under unconfined conditions and is equal to the head h. Thus, the transmissivity T becomes a product of K and h. Then, steady-state in Eq. (2.5) can be rewritten as [Fetter, 2001]:

$$-\frac{1}{2}\nabla \cdot (K\,\nabla h^2) = W - Q, \tag{2.13}$$

with boundary conditions

$$h = h_{\text{Diri}} \text{ on } \Gamma_{\text{Diri}} \tag{2.14}$$

$$-\mathbf{n} \cdot (K\,\nabla h^2) = q_{\text{Neu}} \text{ on } \Gamma_{\text{Neu}}. \tag{2.15}$$

2.1.2. Solute Transport

Transient transport in confined aquifers

Contaminant conservative transport is described by the well-known advection-dispersion equation [Fetter, 1999]

$$\frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{v}c - \mathbf{D}\nabla c) = 0, \tag{2.16}$$

with time t [T], concentration c [M/L^3], velocity field $\mathbf{v}$ [L/T], and dispersion tensor $\mathbf{D}$ [L^2/T] commonly described according to Scheidegger [1954]:

$$\mathbf{D} = \frac{\mathbf{v} \otimes \mathbf{v}}{|\mathbf{v}|}(\alpha_L - \alpha_T) + \mathbf{I}(|\mathbf{v}|\alpha_T + D_e) \tag{2.17}$$

where $\mathbf{v} \otimes \mathbf{v}$ is the tensor product of $\mathbf{v}$ with itself, $|\mathbf{v}|$ is the absolute value of $\mathbf{v}$, α_L [L] and α_T [L] are the local-scale longitudinal and transverse dispersivities, respectively, and D_e [L^2/T] is the effective diffusion coefficient with $D_e = D_m\, n_e$ where D_m [L^2/T] indicates the molecular diffusion coefficient and n_e [-] is the porosity.

Generic initial and boundary conditions are given by:

$$c = c_0 \text{ at } t = t_0 \tag{2.18}$$

$$c = c_{\text{Diri}} \text{ on } \Gamma_{\text{Diri}}\, \forall t \tag{2.19}$$

$$-\mathbf{n} \cdot (\mathbf{v}c - \mathbf{D}\nabla c) = J_{\text{Neu}} \text{ on } \Gamma_{\text{Neu}}\, \forall t, \tag{2.20}$$

where c_{Diri} [M/L^3] denotes a fixed-concentration at the Dirichlet boundaries Γ_{Diri} and J_{Neu} [M/TL^2] describes a constant specific mass flux over Neumann boundaries Γ_{Neu}. No-flux boundaries are given by $J_{\text{Neu}} = 0$.

Steady-state transport in confined aquifers

Under steady-state conditions, Eq. (2.16) simplifies to

$$\nabla \cdot (\mathbf{v}c - \mathbf{D}\nabla c) = 0, \tag{2.21}$$

with generic initial and boundary conditions:

$$c = c_{\text{Diri}} \text{ on } \Gamma_{\text{Diri}} \tag{2.22}$$

$$-\mathbf{n} \cdot (\mathbf{v}c - \mathbf{D}\nabla c) = J_{\text{Neu}} \text{ on } \Gamma_{\text{Neu}}. \tag{2.23}$$

2.1.3. Multi-Rate Mass Transfer

Complex mass transfer processes between slow and rapid zones below the scale described by the above equations overall result in complex non-Fickian behavior and cannot be described by Eq. (2.16) [Berkowitz, 2002]. Yet, this can be tackled with so-called dual-continuum models [Barenblatt et al., 1960; Warren and Root, 1963] by simply adding a transfer term to Eq. (2.16). The formulation of Carrera et al. [1998] arrives at:

$$\frac{\partial c}{\partial t} + \beta \frac{\partial c_{im}}{\partial t} = -\nabla \cdot (\mathbf{v}c - \mathbf{D}\nabla c), \tag{2.24}$$

where c is the concentration in the mobile domain, c_{im} [M/L^3] is the concentration in the immobile domain and $\beta = n_{im}/n_e$ [-] is the capacity coefficient describing the ratio of the porosities between the immobile and mobile domain [Harvey and Gorelick, 1995]. According to Carrera et al. [1998], the immobile concentration c_{im} can be written as function of the mobile concentration c:

$$c_{im} = \int_0^t \phi\left(t - t'\right) c \, \mathrm{d}\, t', \tag{2.25}$$

where $\phi(t)$ is a memory function describing the retention time in the immobile domain [Villermaux, 1987]. The formulation in Eq. (2.25) also allows to describe an entire spectrum if immobile domains, each one with its own capacity or rate coefficient.

Generic initial and boundary conditions are given by:

$$c = c_0 \text{ at } t = t_0 \tag{2.26}$$

$$c = c_{\text{Diri}} \text{ on } \Gamma_{\text{Diri}} \, \forall t \tag{2.27}$$

$$-\mathbf{n} \cdot (\mathbf{v}c - \mathbf{D}\nabla c) = J_{\text{Neu}} \text{ on } \Gamma_{\text{Neu}} \, \forall t, \tag{2.28}$$

where c_{Diri} denotes a fixed-concentration at the Dirichlet boundaries Γ_{Diri} and J_{Neu} describes constant specific mass fluxes over Neumann boundaries Γ_{Neu}. No-flux boundaries are given by $J_{\text{Neu}} = 0$.

2.2. Descriptive Statistics

Descriptive statistics is the discipline of quantitatively describing the main features of collected data. If only a single variable is considered and analyzed, e.g., all (hydro)geological data collected in a specific area with no respect to space, we speak of univariate statistics. If two variables are considered, e.g., comparing two subsets of collected data towards their dependency, we speak of bivariate statistics. The most prominent metrics in data analysis for both univariate and bivariate statistics shall be briefly described in the following.

Univariate Statistics

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

$$m = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{2.29}$$

The sample variance s^2 or in other words, the spreading of the data set about its mean m, is given by

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2 . \tag{2.30}$$

Taking the square root gives the standard deviation $s = \sqrt{s^2}$. The asymmetry of the spreading about the mean is described by the dimensionless skewness k_s

$$k_s = \left(\frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^3 \right) \Big/ s^3. \tag{2.31}$$

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

Bivariate Statistics

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

$$q = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - m \right) \left(x_i' - m' \right) . \tag{2.32}$$

From that the correlation coefficient ρ between the two sets x_i and x_i' is found by removing the individual standard deviations s and s' from q:

$$\rho = \frac{q}{s s'}. \tag{2.33}$$

2.3. Distribution Functions

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

$$F(u) = P(U \leq u), \tag{2.34}$$

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

$$\lim_{u \to -\infty} F(u) = 0 \qquad \text{and} \qquad \lim_{u \to \infty} F(u) = 1. \tag{2.35}$$

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

$$p(u) = \frac{\mathrm{d}\, F(u)}{\mathrm{d}\, u} \tag{2.36}$$

where $p(u)\ \mathrm{d}\, u$ is the probability that U lies within the infinitesimally small interval between u and $u + \mathrm{d}\, u$.

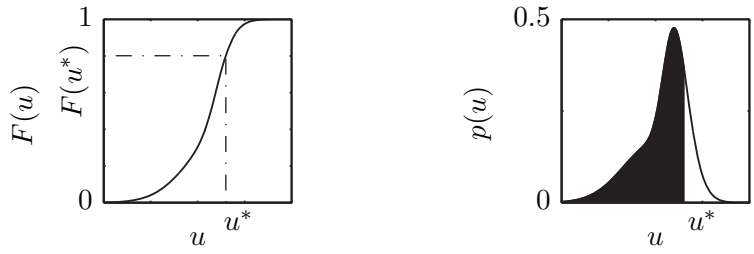


Figure 2.1.: Illustration of *cdf* (left) and *pdf* (right). The black area $\int_0^{u^*} p(u)\, \mathrm{d}\, u$ in right plot refers to the abscissa value $F(u^*)$ in the left plot.

For two jointly distributed variables u and u', the joint (bivariate) *pdf* is defined by $p(u, u')$. If both variables are independent their joint distribution simply factorizes to the product of their marginal distributions $p(u, u') = p(u)\, p(u')$. Integration over u or u' respectively, by $\int p(u, u')\, \mathrm{d}\, u' = p(u)$ and $\int p(u, u')\, \mathrm{d}\, u = p(u')$ yields the marginal distributions $p(u)$ and $p(u')$, respectively. From $p(u, u')$, the conditional *pdf*s $p(u|u')$ and $p(u'|u)$ can be defined. They represent the distribution of u given knowledge on u' and vice versa. Fig. 2.2 illustrates how the marginal, the conditional and the joint *pdf*s are linked to each other. Similar to the sample mean and variance described in Sec. 2.2, distribution functions can be characterized

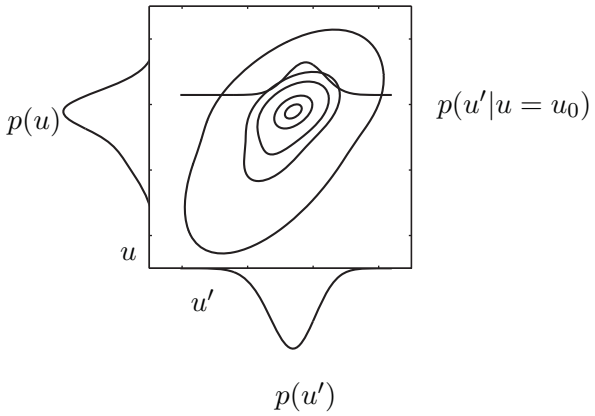


Figure 2.2.: Illustration of marginal *pdf*s $p(u)$ and $p(u')$, joint *pdf* $p(u, u')$ and exemplary conditional *pdf* $p(u'|u)$ for $u = u_0$.

by their means and variances. The mean μ of a random variable u is defined as expectation over the spectrum of possible values by a possibility-weighted integral:

$$\mathrm{E}\left[u\right] = \mu = \int_{-\infty}^{+\infty} u\, p(u)\ \mathrm{d}\, u, \tag{2.37}$$

The second central moment, or variance σ^2, is defined as

$$\mathrm{Var}\left[u\right] = \sigma^2 = \mathrm{E}\left[(u-\mu)^2\right] = \int_{-\infty}^{+\infty} (u\ -\mu)^2\, p(u)\ \mathrm{d}\, u. \tag{2.38}$$

The expression for $\mathrm{Var}\left[u\right]$ can be rewritten be means of expansion [Weiss, 2006]

$$\mathrm{Var}\left[u\right] = \mathrm{E}\left[u^2\right] - \mathrm{E}\left[u\right]^2 = \int_{-\infty}^{+\infty} u^2\, p(u)\ \mathrm{d}\, u - \left(\int_{-\infty}^{+\infty} u\, p(u)\ \mathrm{d}\, u\right)^2. \tag{2.39}$$

For conditional *pdf*s the condition mean $\mathrm{E}_{u|u'}\left[u\right]$ and conditional variance $\mathrm{Var}_{u|u'}\left[u\right]$ is defined in a similar fashion:

$$\mathrm{E}_{u|u'}\left[u\right] = \int_{-\infty}^{+\infty} u\, p(u|u')\ \mathrm{d}\, u, \tag{2.40}$$

and

$$\mathrm{Var}_{u|u'}\left[u\right] = \int_{-\infty}^{+\infty} \left(u\ -\mathrm{E}_{u|u'}\left[u\right]\right)^2 p(u|u')\ \mathrm{d}\, u = \mathrm{E}_{u|u'}\left[u^2\right] - \mathrm{E}_{u|u'}\left[u\right]^2. \tag{2.41}$$

Higher order moments, e.g., skewness, can be derived in a similar fashion but are not shown here. The derivation of statistical moments such as mean μ, variance σ^2 and higher order moments is similar to the definition of temporal moments as shown in Chap. 4.

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

$$\begin{aligned} \mathrm{Cov}\left[u, u'\right] &= \mathrm{E}\left[\left(u-\mu\right)\left(u'-\mu'\right)\right] && (2.42)\\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \left(u-\mu\right)\ \left(u'-\mu'\right)\ p(u,u')\ \mathrm{d}\, u\ \mathrm{d}\, u', && (2.43) \end{aligned}$$

with joint *pdf* $p(u, u')$. If u and u' are independent variables, their covariance $\mathrm{Cov}\,[u, u']$ equals zero.

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

- Gaussian Distribution

  A variable u influenced by a multitude of factors (natural and technical processes), such as measurement errors or marks on a test are often said to have a symmetric, and unimodal distribution. Meeting these prerequisites, the bell-shaped, normal or Gaussian distribution gained much popularity. This is also supported by the central limit theorem Weiss [2006]. Being conceptually and statistically attractive the Gaussian distribution depends on only two parameters, its mean μ and variance σ^2 and is defined as

  $$p(u) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(u-\mu)^2}{2\sigma^2}\right]. \tag{2.44}$$

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

- Log-normal Distribution

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

  $$p(u) = \frac{1}{u\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln u - \mu)^2}{2\sigma^2}\right]. \tag{2.45}$$

  The log-normal distribution is left-bounded on the interval $[0, +\infty]$. It is, hence, often applied to parameters or variables that are physically constrained to take positive values. Examples for physically restricted variables are concentrations, drawdown or precipitation. In geosciences, the most prominent parameter is the hydraulic conductivity, and it is typically assumed to be log-normally distributed [e.g., Gómez-Hernández and Wen, 1998]. An exemplary log-normal distribution is given in Fig. 2.3 (center).

- Exponential distribution

  Exponential distributions can be used to describe the lengths of the inter-arrival times in a Poisson process. The later is typically employed to define the distribution of fractures Priest and Hudson [1976]. The exponential distributions is defined only based on its mean value μ by

  $$p(u) = \frac{1}{\mu} \exp\left[\frac{-u}{\mu}\right]. \tag{2.46}$$

  The exponential distribution is supported on the interval $[0, +\infty]$. An exemplary exponential distribution is given in Fig. 2.3 (right).

- Multivariate Gaussian Distribution

  In most natural processes, a multitude of distributed parameters is involved. Typically, the values of these parameters depend on each other and need to be described including their mutual correlation by so-called multivariate distributions. For two Gaussian

distributed random variables u, u' acting jointly, their bivariate distribution is given by:

$$p(\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^2|\mathbf{\Sigma}}} \exp\left[-\frac{1}{2}(\mathbf{u}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{u}-\boldsymbol{\mu})\right], \tag{2.47}$$

with random variables $\mathbf{u} = \begin{pmatrix} u \\ u' \end{pmatrix}$, their mean $\boldsymbol{\mu} = \begin{pmatrix} \mu \\ \mu' \end{pmatrix}$ and covariance matrix $\mathbf{\Sigma} = \begin{pmatrix} \mathrm{Cov}[u,u] & \mathrm{Cov}[u,u'] \\ \mathrm{Cov}[u,u'] & \mathrm{Cov}[u',u'] \end{pmatrix}$.

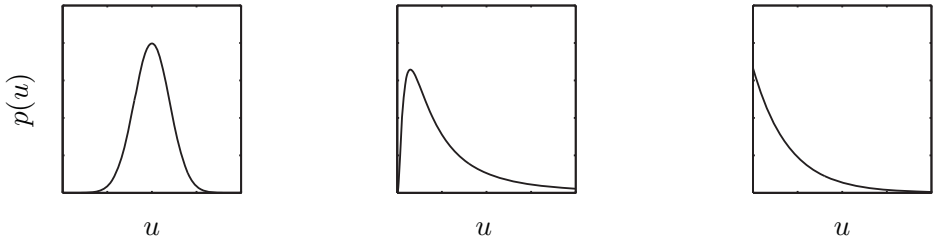


Figure 2.3.: Illustration of three exemplary *pdf*s featuring a normal (left), log-normal (center) and exponential (right) distribution.

2.4. Geostatistical Approach

In order to characterize and predict the spatial patterns of distributed variables in (hydro)geosystems, the spatial consideration of parameter variations and their mutual dependence becomes inevitable. This is where geostatistics drift from univariate statistics. In the most common geological models, two-point statistics come into play. Opposed to single-point statistics with no attention paid to the spatial arrangement of data, two-point statistics invite spatial dimensions into relevant statistical metrics. Thus, the statistics of spatial patterns can be observed. To this end, experimental *variograms* or more theoretical *covariance functions* can be employed. In this section, I briefly repeat how geostatistics can be employed in order to model and simulate spatial features. The statistical theories, however, consolidating the geostatistical hypothesis are not repeated here. For more information, I refer to Kitanidis [1997].

Spatial patterns of parameters $u(x)$ in locations x can be interpreted as random space functions (RSF). The most common descriptions of RSFs typically rely on the mean $\mu(x) = \mathrm{E}[u(x)]$, the variance $\sigma^2(x) = \mathrm{E}[u(x) - \mu(x))^2]$ and the covariance function $R(x,x') = \mathrm{E}[(u(x) - \mu(x))(u(x') - \mu(x))]$. With constant mean $\mu(x) = \mu$ and variance $\sigma^2(x) = \sigma^2$, and $R(x,x')$ depending only on the separation distance $h = ||x - x'||$ is classified as second-order stationary. Eqs. (2.48)-(2.49) give two exemplary theoretical covariance functions commonly employed in hydro(geo)logical models, the Gaussian and the exponential model. Both are

illustrated in Fig. 2.4. The Gaussian covariance model is defined as

$$R(h) = \sigma^2 \exp\left(-\frac{h^2}{L^2}\right), \tag{2.48}$$

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

$$R(h) = \sigma^2 \exp\left(-\frac{h}{L^2}\right). \tag{2.49}$$

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

To relax on assumptions associated with the covariance model, several recent studies, including Feyen [2003], Nowak et al. [2010] and Murakami et al. [2010], suggested to use the Matérn covariance function [Matérn, 1986]:

$$\begin{aligned} C(l) &= \frac{\sigma}{2^{\kappa-1}\Gamma(\kappa)}(2\sqrt{\kappa}l)^{\kappa} B_{\kappa}(2\sqrt{\kappa}l) \\ l &= \sqrt{\left(\frac{\Delta x_1}{\lambda_1}\right)^2 + \left(\frac{\Delta x_2}{\lambda_2}\right)^2}, \end{aligned} \tag{2.50}$$

with Gamma function $\Gamma(\cdot)$ and correlation length λ_1 and λ_2. $B_{\kappa}(\cdot)$ is the modified Bessel function of the third kind [Abramowitz and Stegun, 1972]. The additional shape parameter κ controls the shape of the covariance function, e.g.: $\kappa = 0.5$ is the exponential and $\kappa = \infty$ the Gaussian model. The benefits of the Matérn family have been discussed extensively by, e.g., Handcock and Stein [1993] and Diggle and Ribeiro Jr. [2002]. The relevance of the Matérn family within Bayesian geostatistics has recently been pointed out by Nowak et al. [2010].

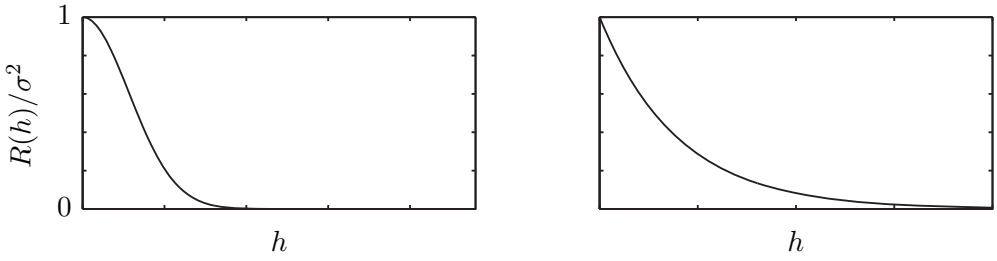


Figure 2.4.: Illustration of two exemplary covariance functions featuring a Gaussian (left), and an exponential (right) model.

Given only the mean μ, variance σ^2 and covariance function $R(h)$, spatially distributed parameter fields can be represented by multivariate Gaussian distributions (see Eq. 2.47) because these distribution parameter are sufficient to define the multivariate Gaussian distribution, and because as little as possible additional information is included in the so chosen

distribution shape. This assumption has received some criticism because the range of possible pattern types to describe is limited [Zinn and Harvey, 2003; Bárdossy and Li, 2008]. Nevertheless, its presence in hydro(geo)logical models is striking. Alternative representations resolving more complex spatial features include, among others, copulas [Bárdossy and Li, 2008; Haslauer et al., 2012].

Treating spatially distributed parameters as RSFs also allows to include uncertainty in their statistical descriptions, leading to Bayesian geostatistics. For more information, I refer to Diggle and Ribeiro Jr. [2007].

2.5. Bayes' Theorem

One of the most basic laws in statistics and geostatistics is the so-called Bayes' theorem [Papoulis, 1984]. Being a fundamental ingredient in parameter inference, Bayes' theorem describes how additional new information can reduce prior uncertainty.

To this end, different pdfs are considered. The prior pdf $p(u)$ refers to the initial uncertainty of some quantity u. The choice of $p(u)$ often results in flat (least subjective as possible) prior or sometimes improper priors Kass and Wasserman [1996]. The posterior pdf $p(u|y)$ defines the reduced state of uncertainty about u after having considered observations y distributed according to $p(y)$.

Both prior and posterior knowledge are related to each other via the so-called likelihood function $p(y|u)$ in the Bayes' theorem. The likelihood function describes the probability that given the prior pdf $p(u)$ how likely the observed data y are. The Bayes' theorem then states:

$$p(u|y) = \frac{p(y|u)p(u)}{p(y)}, \tag{2.51}$$

where $p(y)$ can be directly inferred from $p(y) \approx \sum p(y|u)$ [Robert and Casella, 2004].

2.6. Bootstrap Filter

Bootstrap filter (BF) Gordon et al. [1993] can be seen as a direct translation from Bayes' theorem to the sequential estimation of posterior distributions in a Monte-Carlo framework. BF generally comprise two repetitive steps (1) prediction and (2) update. Let us define model states u_k at time steps k and measurements y_k available at time steps k. Then their initial pdf at time step $k = 0$ is given by $p(u_0|y_0)$ with y_0 being a set of no measurements).

(1) In the prediction step, model states u_k at time k are calculated from previous states u_{k-1} at time $k-1$ via the model $f()$, e.g., an appropriate hydro(geo)logical model with $u_k = f(u_{k-1})$. From that the prior pdf $p(u_k|y_{1:k-1})$ can be calculated where $y_{1:k-1}$ represent available measurements up to time step $k-1$.

(2) At time step k new measurements y_k become available and may be used to update the prior $p(u_k|y_{1:k-1})$ via Bayes' theorem:

$$p(u_k, y_{1:k}) = \frac{p(u_k|y_k)p(u_k|y_{1:k-1})}{p(y_k|y_{1:k-1})}. \tag{2.52}$$

From that model states $p(u_{k+1}|y_{1:k})$ can be predicted during the next time step $k+1$.

In order to characterize the uncertainty about $p(u_k|y_{1:k})$, typically two metrics the mean and the variance are employed. Both metrics are calculated in the weighted sense. For n realizations of u_k, the weighted mean can be found by:

$$\mathrm{E}_{u_k|y_{1:k}}[u_k] \approx \frac{1}{v_1} \sum_{i=1}^{n} u_{ki}\, p(u_{ki}|y_{1:k}), \tag{2.53}$$

and the weighted variance

$$\mathrm{Var}_{u_k|y_{1:k}}[u_k] \approx \frac{v_1}{v_1^2 - v_2} \left[\sum_{i=1}^{n} u_{ki}^2\, p(u_{ki}|y_{1:k}) - \left(\sum_{i=1}^{n} u_{ki}\, p(u_{ki}|y_{1:k}) \right)^2 \right], \tag{2.54}$$

with $v_1 = \sum_{i=1}^{n} p(u_{ki}|y_{1:k})$ and $v_2 = \sum_{i=1}^{n} p(u_{ki}|y_{1:k})^2$. Both quantities are approximated in the weighted sample sense and, therefore, employ v_1 and v_2 [Weiss, 2006, p. 355]. The corresponding correction factor in Eq. (2.54) resembles the well-known factor $\frac{1}{n-1}$ for the non-weighted sample variance. This is an unbiased estimator of the population variance.

BF rely, as a matter of their nature, on MC simulations of random system responses. Thus, they have the desirable property that non-linear systems dependencies are entirely preserved during the updating step [Snyder et al., 2008] making them superior to linearized frameworks such as Ensemble Kalman Filters.

In the context of this thesis, BF and its derivative are used to update parameters, whereas BF, as described by Eq. (2.52), have been originally designed to sequentially update model states.

3. Scenarios

Throughout my thesis, I will employ five different scenarios of groundwater flow and solute transport. For that reason, I present them in the current chapter before continuing to develop my thesis in the following chapters. In the later analyses I only refer to the different scenarios and its individual cases by numbers.

3.1. Scenario (1): Simple Scenario based on Well Flow

Scenario (1) considers transient groundwater flow in a two-dimensional depth-averaged confined aquifer during a pumping test according to Eqs. (2.1)-(2.4). The well is located at $[x_w, y_w] = [50, 50]$ whereas the response is monitored at $[x_m, y_m] = [25, 25]$ (see Fig. 3.1 left). In order to reduce the impact of boundary effects, the numerical domain is chosen to be greater then the actual sample domain. Boundary conditions are Dirichlet-conditions at the western and eastern boundaries and Neumann no-flow conditions at the northern and southern boundaries. All relevant parameters are summarized in Table 3.1 (left column).

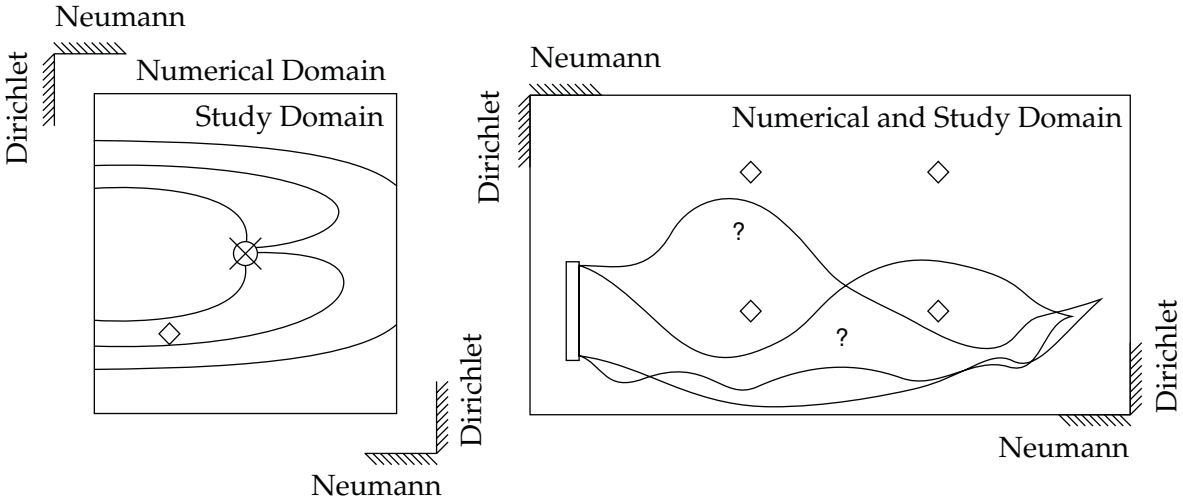


Figure 3.1.: Illustration of scenarios (1) (left) and (2) (right). The centered pumping well in scenario (1) is depicted by a crossed circle. Solid lines indicate the mean stationary flow field. The contaminant line source in scenario (2) is illustrated as rectangle, whereas the uncertainty about the plume evolution is shown by two exemplary shapes. Diamonds in both scenarios indicate monitoring locations.

			Scenario (1)	*Scenario (2)*
		Numerical Domain		
Domain size	$[L_1, L_2]$	[m]	[180, 180]	[300, 150]
Grid spacing	$[\Delta_1, \Delta_2]$	[m]	[2, 2]	[2, 1]
		Study Domain		
Domain size	$[L_1, L_2]$	[m]	[100, 100]	[300, 150]
Grid spacing	$[\Delta_1, \Delta_2]$	[m]	[2, 2]	[2, 1]
		Flow, Transport- and Geostatistical Parameters		
Mean ln T	μ_T	ln[m^2/s]	$-^*$	ln 10^{-5}
Variance ln T	σ_T^2	ln^2[(m^2/s)^2]	$-^*$	1
Integral scale ln T	$\lambda_{T,1,2}$	[m]	$-^*$	$[20, 20]$
Mean ln S	μ_S	$[-]$	$-^*$	$-$
Variance ln S	σ_S^2	$[-]$	$-^*$	$-$
Integral scale ln S	$\lambda_{S,1,2}$	$[-]$	$-^*$	$-$
Diffusion	D_m	[m^2/s]	$-$	1×10^{-9}
Dispersion	α_l, α_t	[m]	$-$	$0.1, 0.01$
Porosity	n_e	[-]	$-$	0.3
		Pumping Parameters		
Injection well	$[x_w, y_w]$	[m]	[50,50]	$-$
Monitoring well	$[x_m, y_m]$	[m]	[25,25]	$-$
Strength	Q	[m^3/s]	-0.05	$-$
Stress interval	Δt_s	[s]	72×10^2	$-$
Relaxation interval	Δt_r	[s]	72×10^3	$-$
Time spacing	Δt	[s]	120	$-$
		Source Parameters		
Source center	$[x_s, y_s]$	[m]	$-$	[20,50]
Source width	$[w_s]$	[m]	$-$	[70]
Fixed concentration	c_0	[kg/m^3]	$-$	1
Monitoring wells	$[x_m, y_m]$	[m]	$-$	$-^*$
Time spacing	Δt	[s]	$-$	120

Table 3.1.: Parameters for scenarios (1) and (2). Values indicated with * vary and are specified in Tab. 3.3.

3.2. Scenario (2): Simple Scenario based on Contaminant Transport

Scenario (2) considers transient solute transport in a two-dimensional depth-averaged confined aquifer according to Eqs. (2.16)-(2.20) in a stationary flow field according to Eqs. (2.5)-(2.7) (see Fig. 3.1, right). A contaminant line source is located at $x = 20$ m and ranges from $y = 15 - 85$ m. The system responses are monitored at four different locations. The underlying flow problem features Dirichlet-conditions at the western and eastern boundaries and Neumann no-flow conditions at the northern and southern boundaries. The transport problem has zero-flux boundaries at the western, northern and southern boundaries. All relevant parameters are summarized in Tab. 3.1 (right column).

3.3. Scenario (3): Complex Scenario based on Well Flow

Scenario (3) considers a two-dimensional depth-averaged unconfined aquifer with two pumping wells that influence the water budget of an ecologically sensitive wetland area (crossed box, see Fig. 3.2, left). The stationary flow problem involves Eqs. (2.5)-(2.7) and has Dirichlet-conditions at the western and eastern boundaries and Neuman zero-flux conditions at the northern and southern boundaries. All relevant parameters are summarized in Tab. 3.2 (left column).

3.4. Scenario (4): Complex Scenario based on Contaminant Transport

Scenario (4) considers a two-dimensional depth-averaged unconfined aquifer with a remediation (gray box) clean-up task as shown in Fig. 3.2 (right): Well 3 injects a reactant at concentration $c_0 = 1$ into a stationary groundwater flow. The injected reactant passes through the contaminated zone, and wells 1 and 2 extract the residual reactant after passing through the contamination. The stationary flow problem is described by Eqs. (2.5)-(2.7) featuring Dirichlet-conditions at the western and eastern boundaries and Neumann no-flow conditions at the northern. The stationary transport problem relies on Eqs. (2.21)-(2.23) with zero-flux boundaries at the western, northern and southern boundaries. All relevant parameters are summarized in Tab. 3.2 (right column).

3.5. Scenario (5): Scenario based on Fractured Porous Media

In scenario (5), I consider flow and transport in a fractured porous media (FPM) (see Fig. 3.3). The scenario features a two-dimensional depth-averaged confined aquifer with a contaminant line source extending over $1/3$ of the western boundary. The stationary flow problem described by Eqs. (2.5)-(2.7) features Dirichlet-conditions at the western and eastern

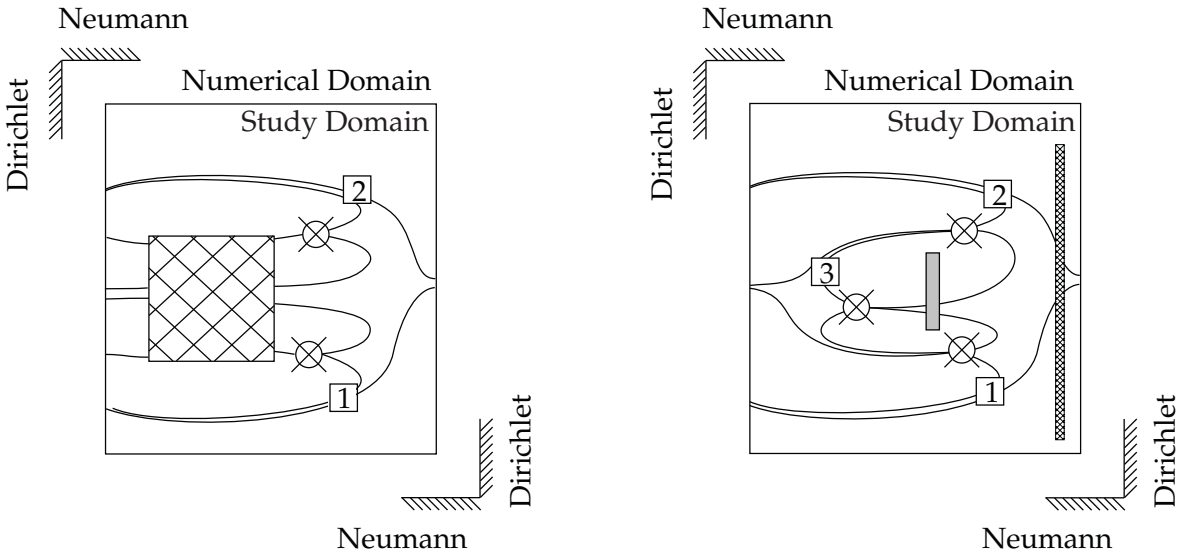


Figure 3.2.: Illustration of scenarios (3) (left) and (4) (right). Left: Wetland scenario featuring regional groundwater flow from left to right and two pumping wells (marked by crossed circles). The sensitive area is illustrated by a shaded box. Right: Scenario from remediation management with well (3) injecting a reactant passing through the contamination (gray box). The reactant is eventually being captured by the extraction wells (1) and (2). Lost reactant mass fluxes are assessed at the shaded control plane.

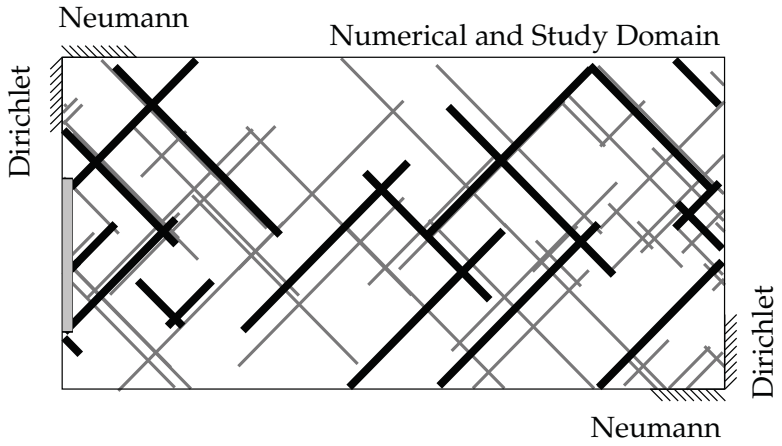


Figure 3.3.: Scenario (5) based on fractured porous media. The differently shaded lines indicate two fracture families with two different conductivities. Underlaid is the almost impermeable rock matrix. The contaminant line source extends over 1/3 of the western boundary.

boundaries and Neumann no-flow conditions at the northern and southern boundaries. The transport problem relies on Eqs. (2.16)-(2.20) and features zero-flux Neumann boundaries at

			Scenario (3)	*Scenario (4)*
		Numerical Domain		
Domain Size	$[L_1, L_2]$	[m]	[10 λ, 10 λ]	[12 λ, 12 λ]
Grid Spacing	$[\Delta x/\lambda]$	$[-]$	$-^*$	$-^*$
Sensitive Domain	$[L_1, L_2]$	[m]	[3 λ, 4 λ]	$[-]$
Control Plane	$[L_1, L_2]$	[m]	[3 λ, 4 λ]	[$-$, 12 λ]
Contaminated Area	$[L_1, L_2]$	[m]	$[-]$	[$-$, 1 λ]
		Flow, Transport- and Geostatistical Parameters		
Mean ln T	μ_T	ln[m^2/s]	ln 3×10^{-4}	ln 2.5×10^{-3}
Variance ln T	σ_T^2	ln[(m^2/s)^2]	3	3
Integral scale ln T	$\lambda_{T,1,2}$	[m]	$[64, 64]$	$[64, 64]$
Diffusion	D_m	[m^2/s]	$-$	1×10^{-9}
Dispersion	α_l, α_t	[m]	$-$	$0.1, 0.01$
Porosity	n_e	[-]	$-$	0.3
		Well Parameters		
Well 1	$[x_w, y_w]$	[m,m]	[7 λ,3 λ]	[8 λ,4 λ]
Well 2	$[x_w, y_w]$	[m,m]	[7 λ,7 λ]	[8 λ,8 λ]
Well 3	$[x_w, y_w]$	[m,m]	$-$	[4 λ,6 λ]
Strength 1	Q	[m^3/s]	0.4	0.005
Strength 2	Q	[m^3/s]	0.4	0.005
Strength 3	Q	[m^3/s]	$-$	-0.01

Table 3.2.: Parameters for scenarios (3) and (4). Values indicated with * vary and are specified in Sec. 9.2.1. All dimensions are specified as multiples of the integral scale λ for ln T

the western, northern and southern boundaries. All relevant parameters are summarized in Tabs. 8.1-8.2 of Chap. 8.

Throughout my later work, I vary some of the scenarios (1) - (5) regarding their geostatistical parameter assumptions, their prediction goals or the tools to evaluate target properties (e.g., in optimal design of experiments). For the sake of completeness, Tab. 3.3 summarizes all variations denoted be different cases.

3.6. Scenario Implementations

All scenarios (1) trough (4) feature transmissivity $T(x)$ as a discretized random space function represented by cell-wise values on a fine numerical grid. Following classical geostatistical ideas, I use $E[T]$ as a known constant, and assume that $\ln T' = \ln T - E[\ln T]$ is second-order stationary with isotropic Gaussian covariance function $C(h)$ that only depends on the

Scenario	*Case*	*Specific Variation*
(1)	(a)	Geostatistical Model (in Sec. 6.4 and Sec. 7.3.1)
	(b)	
	(c)	
	(d)	
(2)	(a)	Monitoring Location (in Sec. 6.5 and Sec. 7.3.2)
	(b)	
	(c)	
	(d)	
	(e)	Exploration of Information (in Sec. 9.1.3)
	(f)	
	(g)	Prediction Goal (in Sec. 9.1.3)
	(h)	
	(i)	Conceptual Model (in Sec. 9.1.3)
(3)	-	-
(4)	(a)	Error metric (in Sec. 9.2.2)
	(b)	
(5)	-	-

Table 3.3.: Case variations for all scenarios (1) - (5).

separation vector h [e.g., Kitanidis, 1997]. Realizations of $T(x)$ are generated with the same implementation of FFT-based methods [Newsam and Dietrich, 1994] as used in Nowak et al. [2008].

The flow problems of scenarios (1) and (3) are solved in parallel by using MODFLOW-2005 [Harbaugh, 2005] on 80 cores with 2.8 GHz. Solutions of the transport-related problems of scenarios (2) and (4) are based on Eulerian descriptions and rely on a finite element method (FEM) scheme already used in Nowak et al. [2008] and Nowak et al. [2010]. The flow problem of scenario (5) is solved by the Complex Systems Modeling Platform (CSMP) involving a FEM approximation [e.g., Geiger et al., 2010]. The transport solution of scenario (5) features different scales tackled by particle tracking random walk (PTRW) simulations [e.g., Salamon et al., 2006; Koch and Nowak, 2013], and a finite volume (FV) scheme [e.g., Cirpka et al., 1999a], respectively.

4. Temporal Moments

As listed in Sec. 1.2, there has been a great deal of work dedicated to the model reduction along the time dimension. The most striking one is the approach of Harvey and Gorelick [1995], establishing the so-called moment generating equations (MGE). As their work marks the starting point of my later analysis, I first refresh the idea of temporal moments (TM) in Sec. 4.1, discuss their physical meaning in Sec. 4.2, and then head to a brief discussion of MGE in Sec. 4.3 including applications Sec. 4.4.

4.1. Definition

Let $r(t)$ be a time-dependent response of a system to an external excitation starting at time $t_0 = 0$ and measured at some location x_0. Examples include drawdown curves (DC) due to the excitation of the subsurface water level in pumping tests [e.g. Fetter, 2001], reactions due to recharge events, tidal pumping or changing river stages [e.g. Yeh et al., 2009], solute breakthrough curves (BTC) during the injection of water-borne tracers and contaminant spills [e.g. Fetter, 1999], or reactions of river discharge to precipitation in hydrological models [e.g. Nash and Sutcliffe, 1970]. Such responses $r(t)$ may look like the solid black curve enveloping the gray shaded area in Fig. 4.1.

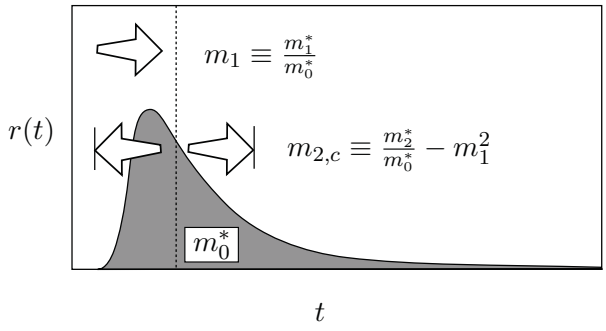


Figure 4.1.: Illustration of the first three TM, i.e., response strength m_0^*, mean response time m_1 and mean response duration $m_{2,c}$, and their relation to an exemplary response curve $r(t)$.

The k^{th} TM m_k^* of $r(t)$ is defined as

$$m_k^* = \int_0^\infty t^k r(t)\ \mathrm{d}\,t, \tag{4.1}$$

where t^k is a monomial of order k used as a base function. Then, the k^{th} raw moment m_k is calculated by normalization with m_0^*

$$m_k = \int_0^\infty t^k \frac{r(t)}{m_0^*} \, \mathrm{d}\, t = \frac{m_k^*}{m_0^*} \, [\mathrm{T}^k]. \tag{4.2}$$

The normalization by m_0^* makes $r(t)$ a function with density properties, i.e., $\int_{-\infty}^{\infty} r(t) \, \mathrm{d}\, t = 1$, similar to pdfs. Raw temporal moments m_k are then closely related to statistical moments, simply applied to time rather than to some random variable. Due to their familiarity with statistical moments, I refer the interested reader to Sec. 2.3 where distribution functions and associated statistical moments have been summarized.

Typically, higher order TM are centralized to $m_1 = 0$ and then standardized to $m_2 = 1$ using the binomial transform [e.g. Papoulis, 1984]

$$m_{k,c} = \sum_{j=0}^{k} \binom{n}{k} (-1)^{n-j} \frac{m_j}{m_0^*} \, m^{n-j} \, [\mathrm{T}^k], \tag{4.3}$$

and

$$m_{k,s}(x) = \frac{m_{k,c}}{m_{2,c}^{k/2}} \, [-], \tag{4.4}$$

with $m_{k,c}$ and $m_{k,s}$ being centralized and standardized TM, respectively.

This normalization, centralization and standardization is in analogy to image pattern recognition, where algebraic moment invariants are calculated in order to make image features invariant with respect to scale, translation and rotation [Prokop and Reeves, 1992]. TM can also be derived from the Laplacian transformation of $r(t)$ [Kubo, 1962; Harvey and Gorelick, 1995], as Taylor series coefficients in the spectral domain [Kendall and Stuart, 1977] (see Appendix A).

4.2. Physical Meaning

Based on the order k of the respective base function t^k, TM capture different individual features of the response curves $r(t)$. As summarized later in Sec. 4.4, most existing applications only consider lower-order TM. Fig. 4.1 illustrates the zeroth through second TM including the underlying response curve $r(t)$.

Zeroth TM The zeroth temporal moment m_0^* is a simple integral of the response $r(t)$ over time, and so measures the overall response strength. This is marked as the gray shaded area under the enveloping black solid curve in Fig. 4.1.

First TM The first raw (normalized) TM m_1 provides information on the time between excitation and bulk response, i.e., a characteristic response time of the system. This is marked by the vertical dashed line in Fig. 4.1.

Second TM The second normalized and centralized TM $m_{2,c}$ is the squared response deviation from the characteristic response time m_1. In other words, $m_{2,c}$ measures the characteristic response duration (horizontal double arrows).

Third TM The third normalized, centralized and standardized TM $m_{3,s}$ describes the asymmetry (skewness) of the response curve $r(t)$ and so characterizes the tailing, e.g., caused by kinetic sorption [e.g., Fetter, 1999].

Fourth TM The fourth normalized, centralized and standardized TM $m_{4,s}$ describes the peakedness (kurtosis) of the response curve $r(t)$.

In many situations, these features can be put into relation to the governing physical flow and transport processes and their parameters. In the following, two examples are provided:

First example based on drawdown (case 1a)

For a drawdown curve obtained from slug-like aquifer tests, m_0^* is related to the steady-state drawdown that would result from continuous pumping. m_1 is the characteristic relaxation time, also bearing some transient information that is needed to estimate the storativity [Li et al., 2005]. As a matter of physics, pressure waves propagate approximately radially sym-

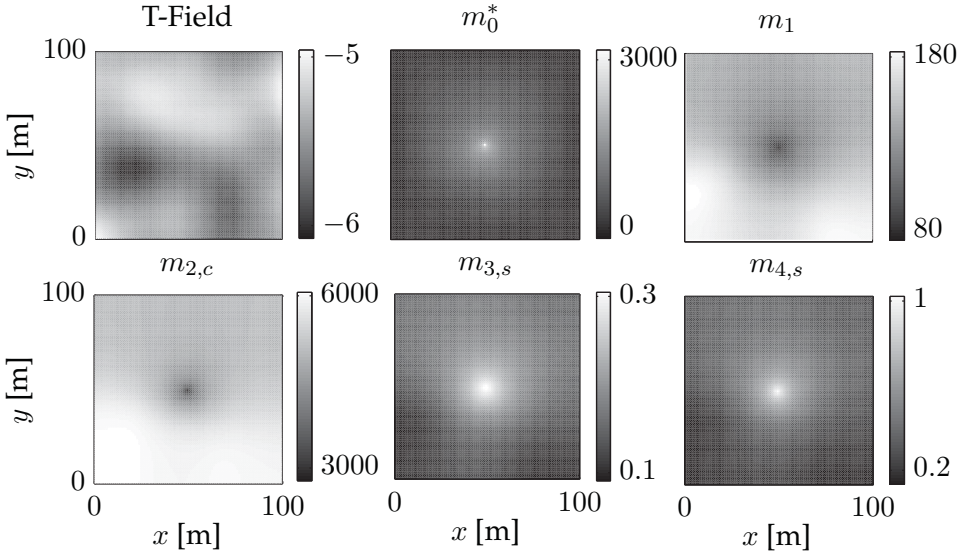


Figure 4.2.: Illustration of spatial TM distribution from m_0^* through $m_{4,s}$ based on scenario (1). The scenario features a centered pumping well.

metric (strictly in homogeneous isotropic media) from where the system has been excited. This is reflected in the spatial distribution of TM from drawdown curves revealing a strong spatial correlation as shown in Fig. 4.2 exemplary for the zeroth through fourth TM. The example features scenario (1) and the shown TM are calculated from the transient drawdown $r(t)$ due to extraction in the center of the domain. The transmissivity field has been generated based on $\mu_T = 10^{-4}$, $\sigma_T^2 = 1$, $\lambda_{T,1,2} = [20,20]$, and a Gaussian description of the spatial

correlation pattern. For more details on scenario (1), I refer to Sec. 3.1. Generally, the spatial pattern of TM is shaped by the heterogeneous nature of the underlying conductivity field, making TM from drawdown curves a promising measurement data type for estimating hydraulic conductivity fields [Li et al., 2005].

Second example based on solute transport (scenario 2)

Solute transport in porous or fractured-porous media is different to pressure-driven drawdown scenario, as information propagate from the point (volume) of solute injection pathwise to the point (area, volume) of observation. Hence, TM represent a path-integrated measure of transport characteristics. This makes TM from tracer experiments (e.g., measured tracer breakthrough curves) a useful data type to quantify aquifer properties different to the hydraulic conductivity, complementing the explanatory power of drawdown-based TM when jointly inverting conductivity fields [Cirpka and Kitanidis, 2000b; Nowak and Cirpka, 2006]. Fig. 4.3 shows an example for transport-based TM featuring scenario (2) for a pulse-like injection. The transmissivity field has been generated based on $\mu_T = 10^{-4}$, $\sigma_T^2 = 1$, $\lambda_{T,1,2}$ = [20,20], and an exponential description of the spatial correlation pattern. For a detailed description I refer to Sec. 3.2.

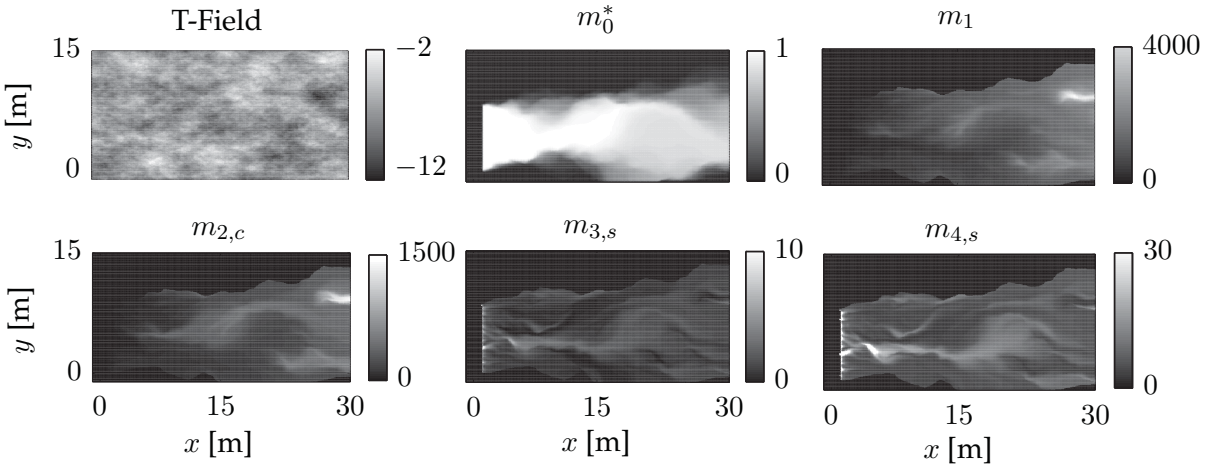


Figure 4.3.: Illustration of spatial TM distribution from m_0^* through $m_{4,s}$ based on scenario (2). The scenario features a contaminant line source ranging over 1/3 of the western boundary.

Here, the zeroth TM m_0^* is the total observed mass at a given point x_0. The first TM m_1 is related to the bulk arrival time along the travel path, reflecting the apparent average seepage velocity [Aris, 1956; Cirpka and Kitanidis, 2000a; Goode, 1996]. The second centralized TM $m_{2,c}$ describes solute spreading, a process attributed to diffusion, local dispersion, kinetic sorption [Valocchi, 1989], and macrodispersion [Gelhar and Axness, 1983; Cirpka and Kitanidis, 2000a]. Thus, $m_{2,c}$ may be used to define apparent dispersion coefficients [e.g., Cirpka and Kitanidis, 2000a]. Higher TM represent more complex information on structural

properties of porous media, e.g., caused by non-uniform grain-size distributions [Cunningham and Roberts, 1998]. Luo et al. [2008] derived a direct relation between TM of breakthrough curves (BTC) and TM of memory functions in multi-rate mass transfer processes (MRMT). A great advantage in this context is that TM remove the non-locality in time from MRMT equations, as will be explained in more detail in Chap. 8.

4.3. Moment Generating Equations

Besides their intuitive understanding, physical meaning and significance in theoretical analyses, TM have the advantage that they can be simulated at very low computational costs. Let us consider a generic linear dynamic and distributed system (e.g., described by a system of coupled PDEs or a single PDE such as Eq. (2.9) and Eq. (2.16)):

$$\frac{\partial r(t)}{\partial t} - \mathcal{D}(r(t)) = Q(t), \tag{4.5}$$

linear differential operator $\mathcal{D}(^*)$ of arbitrary order and forcing term $Q(t)$. For the sake of simplicity and without loss of generality, I set source/ sink terms W equal to zero and neglect system coefficients, e.g., the specific storage coefficient S_0.

Generic initial and boundary conditions are

$$r(t) = r_0 \text{ at } t = t_0 \tag{4.6}$$

$$r(t) = r_{Diri} \text{ on } \Gamma_{Diri} \forall t \tag{4.7}$$

$$-\mathbf{n} \cdot \nabla r(t) = J_{Neu} \text{ on } \Gamma_{Neu} \forall t, \tag{4.8}$$

where r_{Diri} and J_{Neu} are the corresponding values on the Dirichlet and Neumann conditions.

Applying Eq. (4.1) to Eqs. (4.5)-(4.8) reduces the transient PDE in Eq. (4.5) to a set of steady-state equations:

$$\mathcal{D}\left(m_k^*\right) = m_{k-1}^* + m_{Q,k}^* \tag{4.9}$$

$$\mathcal{D}\left(m_0^*\right) = m_{Q,k}^*, \tag{4.10}$$

with generic boundary conditions

$$m_k^* = m_{k,Diri}^* \text{ on } \Gamma_{Diri} \tag{4.11}$$

$$-\mathbf{n} \cdot \nabla m_k^* = m_{k,Neu}^* \text{ on } \Gamma_{Neu}, \tag{4.12}$$

where $m_{Q,k}^*$ is the k^{th} TM of the forcing term, now including the model forcing by the initial condition (if it is not equal to the steady-state at $t \to \infty$), and $m_{k,Diri}^*$ and $m_{k,Neu}^*$ are the k^{th} TM of the boundary values. The transition from Eq. (4.5) to Eqs. (4.9)-(4.12) requires integration by parts, which leads to the appearance of lower order TM as source/ sink terms on the right hand side.

Quite obviously, the considered TM of r_{Diri} and J_{Neu} have to be finite, which is satisfied if the forcing persists over a finite time. This is apparently given in most experiments with pulse-like system excitation such as slug tests. The detailed steps that lead to Eqs. (4.9)-(4.12) are provided for arbitrary base functions in Chap. 5. Auxiliary conditions required for integrating the time derivative are $t^k r(t) \rightarrow 0 \,\forall\, k$ and $t^k\, Q(t) \rightarrow 0 \,\forall\, k$ at $t \rightarrow \infty$, i.e., the system response has to asymptotically decay to zero faster than the highest power of t used in the analysis approaches infinity. Two more prerequisite are that Eq. (4.5) must be a linear (system) of PDEs or ODEs, and the coefficients must be time-invariant.

For drawdown from pumping tests in confined aquifers, which occurs as parabolic PDE, TM reduce Eqs. (4.8)-(4.10) to an elliptic PDE with formally time-independent boundary conditions. An important observation is that TM are now calculated recursively, where the previous TM of order $(k-1)$ serves as source term for the respective current TM of order k. This makes it impossible to directly access TM of higher orders. On the other hand, this type of recursive coupling is computationally very appealing as a fully coupled system of equations can be avoided.

Overall, this allows to simulate TM m_k^* at the computational costs of a few recursive steady-state simulations, avoiding the costly need for time marching schemes in transient simulations. Applications to specific problems existing in the literature can be found in Sec. 4.4.

4.4. Review of Temporal Moment Applications

As discussed in Sec. 4.2, TM capture the most significant aspects of a system response $r(t)$ such as strength, delay, duration, etc., and often have well-defined physical meanings. Thus, TM intuitively bear a high information density, dramatically reducing computational costs. This allows to cover many challenges in hydro(geo)logy such as prediction, uncertainty quantification, calibration/inversion, or probabilistic risk assessment. In the following, I review the applications of TM that can be found in the hydrogeological literature, with a specific focus on how many TM have been considered.

Harvey and Gorelick [1995] (three TM) where the first to provide moment-generating equations for transport with complex initial and boundary conditions and a variety of mass transfer models. Their moment-generating equations were of recursive-type to be solved numerically, and extended the earlier analytical work on TM studying chromatographic properties in chemical engineering by Kučěra [1965]; Kreft and Zuber [1978]; Villermaux [1981a,b], and later mass transfer in the subsurface by Valocchi [1985]; Goltz and Roberts [1987]; Valocchi [1990]; Sardin et al. [1991]; Cunningham and Roberts [1998]; Lawrence et al. [2002].

Representative numerical studies following the mentality of Harvey and Gorelick [1995] range from prediction of groundwater age or life expectancy by Goode [1996] (two TM), Molson and Frind [2011] (two TM), Varni and Carrera [1998] (two TM), and probabilistic assessment of well vulnerability zones [Enzenhöfer et al., 2011] (five TM) to solute travel time analysis [Cirpka and Nowak, 2004] (two TM). Cirpka and Kitanidis [2000b] (two TM), Li et al. [2005] (two TM), Zhu and Yeh [2006] (two TM), Pollock and Cirpka [2008] (two TM)

and Yin and Illman [2009] (two TM) applied TM to make calibration or stochastic inverse modeling more efficient. Cirpka and Kitanidis [2000a] (three TM) and Cirpka and Kitanidis [2000c] (three TM) deterministically related TM to transport characteristics. Luo et al. [2008] generalized the pioneer work of Harvey and Gorelick [1995] to mass transfer models with arbitrary formulations of the memory function.

When it comes to applications, I observed that almost all applications involve hypothetical data and scenarios to test and demonstrate their method, and almost no field applications exist. Among all work known to me, only Varni and Carrera [1998] (two TM), Cunningham and Roberts [1998] (four TM) and Nowak and Cirpka [2006] (three TM) compared their numerical simulations against field measurements of groundwater age, grain-size distributions and solute breakthrough curves, respectively. Nowak and Cirpka [2006] and Yin and Illman [2009] reduced measured real breakthrough curves and experimental drawdown curves, respectively, and then used TM in geostatistical inversion.

5. Reduction Efficiency via Alternative Integral Transforms

In the last chapter, we have seen that there is a great potential in reducing transient models, e.g., by means of temporal moments (TM). They, generally, reduce a time-dependent partial differential equation (PDE) into a recursive set of steady-state PDEs by projection onto a set of monomial base functions. Due to the simple coupling among higher and lower-order TM, a swift evaluation can be achieved. The only prerequisites are that (1) the partial differential equation is linear and (2) the coefficients must be independent of time and independent of the solution $r(t)$. The question is whether monomials are the only base functions that lead to this efficient recursive conversion of the governing equations. And, if no, are there other types of well-suited base functions and how efficiently do they reduce PDEs?

This leads me to ***Step (I)*** of my overall approach. In the following, I consider model reduction in time from the very most general perspective. First, I derive alternative integral transformations based on arbitrary base functions (Sec. 5.1). Second, I classify them and discuss their different properties towards reduction efficiency (Sec. 5.2). The material of the following chapter has been published in Leube et al. [2012b].

5.1. Alternative Base Functions

It can be anticipated from Eqs. (4.1)-(4.10), that choosing other base functions than the monomials t^k will lead to other, more general, temporal characteristics α_k than TM. The key question will be, whether their resulting generating equations are fully coupled, recursively coupled (like for TM) or independent. In order to analyze this issue, I replace the monomials t^k in Eqs. (4.1)-(4.12) with a set of yet unspecified base functions $\xi_k(t),\ k = 0 \dots K$, and repeat all steps analogously. This leads to a definition for arbitrary temporal characteristics:

$$\alpha_k = \int_0^\infty \xi_k(t) r(t)\ \mathrm{d}\, t, \tag{5.1}$$

with their corresponding generating equations

$$\underbrace{\int_0^\infty \xi_k(t)\,\frac{\partial r(t)}{\partial t}\,\mathrm{d}\,t}_{(1)} + \underbrace{\int_0^\infty \xi_k(t)\,\mathcal{D}\left(r(t)\right)\mathrm{d}\,t}_{(2)} = \underbrace{\int_0^\infty \xi_k(t)\,Q(t)\,\mathrm{d}\,t}_{(3)},\; k = 0 \ldots K. \tag{5.2}$$

The boundary conditions are similar to those defind for TM in Sec. 4.3

$$\int_0^\infty \xi_k(t)\,r(t)\;\mathrm{d}\,t \;=\; \int_0^\infty \xi_k(t)\,r_{Diri}(t)\;\mathrm{d}\,t \;\text{ on }\; \Gamma_{Diri}\,\forall\, t \tag{5.3}$$

$$\int_0^\infty \xi_k(t)\,\mathbf{n}\cdot\nabla\, r(t)\;\mathrm{d}\,t \;=\; \int_0^\infty \xi_k(t)\,\mathbf{n}\cdot\nabla\, r_{Neu}(t)\;\mathrm{d}\,t \;\text{ on }\; \Gamma_{Neu}\,\forall\, t. \tag{5.4}$$

Integrating terms (2) and (3) and the boundary conditions is trivial, since $\xi_k(t)$ and the time integral can be moved into the spatial differential operator. This leads to a differential expression for the new temporal characteristics α_k. Term (1) requires integration by parts and leads to terms (4) and (5) in the following equation:

$$\underbrace{\left[\xi_k(t)\,r(t)\right]_0^\infty}_{(4)} - \underbrace{\int_0^\infty \frac{\partial \xi_k(t)}{\partial t}\,r(t)\;\mathrm{d}\,t}_{(5)} + \underbrace{\mathcal{D}\left(\alpha_k\right)\mathrm{d}\,t}_{(2)} = \underbrace{\alpha_{Q,k}}_{(3)}, \tag{5.5}$$

where α_k are the temporal characteristics of order k that correspond to the TM in Eq. (4.1) when setting $\xi_k(t) = t^k$, and $\alpha_{Q,k}$ are the corresponding characteristics of the forcing function $Q(t)$. When the auxiliary conditions are changed accordingly, term (4) vanishes. The required conditions are $t^k r(t) \to 0\,\forall\, k$ and $t^k\, Q(t) \to 0\,\forall\, k$ at $t \to \infty$, i.e., the system response has to asymptotically decay to zero faster than the highest power of t used in the analysis approaches infinity. The associated boundary conditions become

$$\alpha_k \;=\; \hat{\alpha}_k \text{ on } \Gamma_{Diri} \tag{5.6}$$

$$\mathbf{n}\cdot\nabla\,\alpha_k \;=\; \hat{\alpha}_{Q,k} \text{ on } \Gamma_{Neu}. \tag{5.7}$$

Eq. (5.5) can be solved without reverting to a time-dependent solution of Eq. (4.8), if and only if the remaining term (5) can be expressed through a combination of characteristics with arbitrary orders k ranging from α_0 to α_k, $k = 0 \ldots K$ such that all time-related differential and integral operators disappear:

$$\underbrace{\int_0^\infty \frac{\partial \xi_k(t)}{\partial t}\,r(t)\;\mathrm{d}\,t}_{(5)} = c_{0k}\,\underbrace{\int_0^\infty \xi_0(t)\,r(t)\;\mathrm{d}\,t}_{\alpha_0} + \ldots + c_{Kk}\,\underbrace{\int_0^\infty \xi_K(t)\,r(t)\;\mathrm{d}\,t}_{\alpha_K} \tag{5.8}$$

where c_{Kk} are linear coefficients. Applying Eq. (5.1) and auxiliary conditions to both sides of Eq. (5.8) allows to replace the various integrals over $r(t)$ by the characteristics α_k of $r(t)$ and rewrite Eq. (5.8) as a system of ordinary differential equations (ODEs):

$$\frac{\partial \xi_k(t)}{\partial t} \;=\; c_{0k}\,\xi_0(t) + \ldots + c_{Kk}\,\xi_K(t). \tag{5.9}$$

This set of equations will allow to finally replace all remaining time-related operators in term (5) of Eq. (5.2), and leads to a coupling between the K replicants of Eq. (5.2) for all $k = 0 \ldots K$.

5.2. Coupling Matrices

I will now investigate specific coupling cases that can occur in Eq. (5.9). The goal is to find the set of base functions $\xi_k(t)$ that allows to most swiftly simulate temporal characteristics from Eq. (5.2). This way, I wish to find the approach that most efficiently reduces models in time. Putting Eq. (5.9) into matrix notation reveals different cases of coupling schemes between the replicates of Eq. (5.2) as illustrated in Fig. 5.1. Four specific cases are of particular relevance for further analysis and will be discussed in the following paragraphs.

5.2.1. Fully Populated Case

In the most general case (a), term (5) can only be expressed as a linear combination of all lower and higher order characteristics, leading to a fully populated coupling matrix. This will occur only if the base functions $\xi_k(t)$ are non-polynomial, e.g. rational, trigonometric, etc., such that none of their time derivatives vanish, or if they are polynomial approximations of arbitrary non-polynomial base functions truncated at order $K \geq k$.

For the final purpose of simulating temporal characteristics, this will lead to a fully coupled finite or even infinite system of equations in Eq. (5.2). This is unfeasible, because it will be much more expensive to solve than recursively coupled systems or decoupled equations (see the other cases). Also, it may exclude commercial software packages from being used if they do not allow solving coupled equations. Therefore, I can immediately remove case (a) from my further considerations.

5.2.2. Lower Order Case

Case (b) resembles the situation where term (5) can be expressed as a linear combination of characteristics of order only smaller than k, leading to a lower triangular coupling matrix. This can only occur if the base functions $\xi_k(t)$ are polynomials of order k (or polynomial approximations of, e.g., trigonometric, hyperbolic, square root, logarithmic, or any other arbitrary base functions, truncated at order k, sorted in ascending order). From Eq. (5.9) it can be seen, that, in case (b), the first line directly leads to $\xi_0(t) = const$, such that $\xi_1(t)$ must have first order in t, and so on.

Let us now consider an arbitrary polynomial base function expressed via linear combinations of monomials t^i

$$\xi_k(t) = \sum_{i=0}^{k} \psi_{ik}\, t^i, \tag{5.10}$$

with time-independent coefficients ψ_{ik}. When pursuing this approach, I get

$$\alpha_k = \int_0^\infty \xi_k(t)\, r(t)\ \mathrm{d}\, t = \int_0^\infty \sum_{i=0}^{k} \psi_{ik}\, t^i\, r(t)\ \mathrm{d}\, t. \tag{5.11}$$

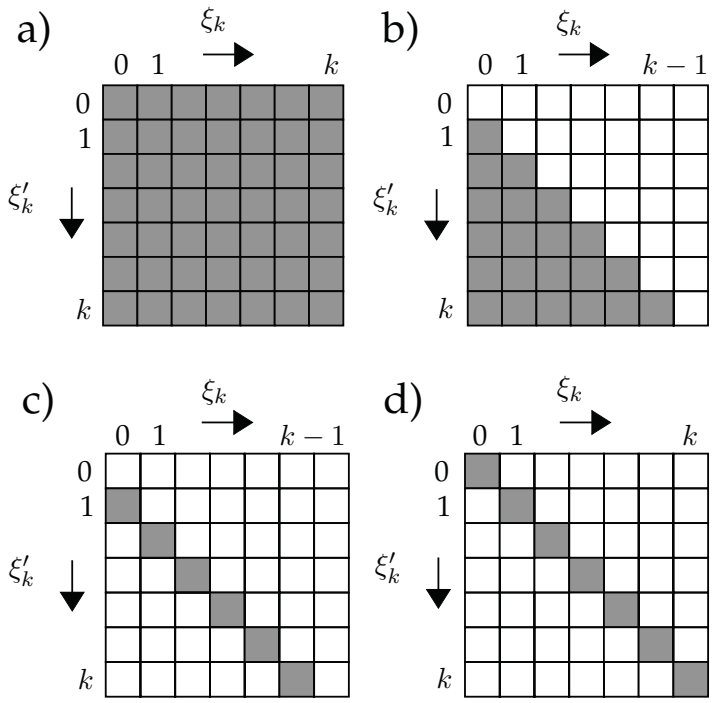


Figure 5.1.: Coupling matrices comprising four different cases: (a) full coupling, (b) the entire lower triangular matrix occupied for the general case, (c) only the secondary diagonal occupied in the Appell case, and (d) only the main diagonal occupied in the Laplace case.

Pulling the sum and ψ_{ik} outside the integral yields

$$\alpha_k = \sum_{i=0}^{k} \psi_{ik} \int_0^\infty t^i \, r(t) \, \mathrm{d}\, t, \tag{5.12}$$

which can be expressed as linear combination of TM m_i of order $i = 0 \ldots k$:

$$\alpha_k = \sum_{i=0}^{k} \psi_{ik} \, m_i^*. \tag{5.13}$$

From this, it follows that any temporal characteristic based on arbitrary polynomial base functions $\xi_k(t)$ obeying case (b) can be mimicked by TM through linear re-combination. Therefore, arbitrary polynomials of order $k = 0 \ldots K$ will capture the same temporal information, yet at slightly higher computational costs (due to the treatment of multiple source terms in Eq. 5.5).

5.2.3. Appell Case

Case (c) is a special case of (b) involving the so-called *Appell* sequences [Appell, 1880]. Appell sequences include Hermite polynomials, Bernoulli polynomials, Euler polynomials, and the monomials that lead to TM. They are in fact defined via an ODE system that is simpler than Eq. (5.9), occupying only the secondary diagonal of the coupling matrix. By the nature of this coupling, it is obvious that the recursive coupling is computationally the most efficient way to simulate temporal characteristics, together with the last case (d).

5.2.4. Laplace Diagonal Case

Case (d) considers the situation where the coupling term only occupies the main diagonal. Guaranteeing that $\frac{\partial \xi_k(t)}{\partial t}$ is proportional to $\xi_k(t)$ can be fulfilled if and only if $\xi_k(t) \propto e^{-kt}$, which directly leads to the Laplace transformation (LT). The relation of the LT to TM is recalled in Appendix A. In brief, the LT yields the spectrum of the system response, and TM are the Taylor expansion coefficients of the spectrum. As a consequence of diagonal coupling, the Laplace coefficients can be determined independently (uncoupled, non-recursively). This advantage is, however, bought at the fact that it is unclear which Laplace coefficients summarize the dynamic behavior in $r(t)$ best. As a direct consequence, applications employing the LT typically used between 10 and 40, sometimes even 100 Laplace coefficients to accurately restore the solution $r(t)$ [e.g., Li et al., 1992; Sudicky and McLaren, 1992]. I rate this case as fast but impractical since the choice of considered orders k remains unclear. A quantification of this statement for a specific system setup will be presented in Sec. 6.4.3.

5.2.5. Orthogonal Case

Characteristics α_k should summarize the dynamic behavior of $r(t)$ as good as possible, already with a small number. Only then, the sequence of considered characteristics can be truncated at low $k = 0 \dots K$, leading to a small set of replicates for Eq. (5.2) to be solved. In analogy to signal processing, I refer to this desired property as optimal compression. A prerequisite for optimal compression is, that temporal characteristics α_k in the order of the sequence $k = 0 \dots K$ add large and possibly non-redundant information units, sorted from most to least significant information units in descending order. Using the terminology of Fourier, Laplace, or more general integral transforms [e.g., Debnath and Bhatta, 2007], the spectrum $(\alpha_k)^2$ has to decay as fast as possible with increasing order k, by adequate choice of $\xi_k(t)$. The goal of non-redundancy can be achieved by guaranteeing orthogonality among the respective base functions $\xi_k(t)$.

Taking advantage of orthogonal base functions has already been done in fields different to my study. In image processing, Teague [1980] established orthogonal polynomials in order to derive moments invariant with respect to image translation. In the context of object reconstruction, Prokop and Reeves [1992] resumed that monomials are highly correlated and thus introduced orthogonalized moments in order to reduce the information redundancy among conventional moments. Furthermore, they concluded that orthogonal moments are more suitable in image reconstruction and may be used to determine the minimum number of moments required to adequately reconstruct and thus, uniquely characterize, a given image. In chromatography, Kučěra [1965] suggested to expand a time-dependent response in order to analytically solve the advection-dispersion equation including linear sorption kinetics. To this end, he suggested to use orthogonal Hermite polynomials.

I will now investigate whether any orthogonal base functions exist that allow to reduce Eq. (5.2). Generally, orthogonality between base functions $\xi_k(t)$ and $\xi_\ell(t)$ is defined as

$$\int_a^b \xi_k(t)\,\xi_\ell(t)\,w(t)dt = \begin{cases} 0, & \forall\, k \neq \ell \\ N_k, & \forall\, k = \ell \end{cases}, \tag{5.14}$$

with respect to the weighting function $w(t)$. N_k is the squared weighted L^2-norm and depends on the choice of the base function $\xi_k(t)$. In this context, the optimal choice of the base functions strongly depends on the associated weighting function $w(t)$ and its own moments [e.g., Abramowitz and Stegun, 1972; Oladyshkin et al., 2011] and on the integration interval $[a, b]$. For most dynamic and distributed systems of interest in hydro(geo)logical applications, I have $[a, b] = [0, \infty]$ and $w(t) = 1$. Finding an arbitrary orthogonal base function $\xi_k(t) = \sum_{i=0}^{k} \psi_{ik}\, t^i$ with time-independent coefficients ψ_{ik} meeting these constraints would lead to (according to Eq. (5.14)):

$$\int_0^\infty \xi_0^2\,\omega(t)\,\mathrm{d}\,t =: \int_0^\infty \psi_0^2\,\mathrm{d}\,t = \psi_0^2\,[t]_0^\infty = \psi_0^2\,\infty. \tag{5.15}$$

Under these conditions, it is impossible to define orthogonal base functions because there is no real-valued non-zero base function to fulfill Eq. (5.14). Orthogonal base functions can, since they do not exist for the class of problems I am interested in, be excluded from my considerations.

5.2.6. Cumulant Case

For the sake of completeness, I recall the case of cumulants (or semi-invariants). Similar to TM, cumulants characterize the nature of random variables, however, they do it a unique fashion [Gardiner, 1985]. Cumulants have the elegant property that they allow to reconstruct the dynamic response $r(t)$ using the so-called Edgeworth expansion which has highly advantageous convergence properties for nearly-Gaussian problems [e.g., Chatwin, 1970]. However, non-negativity of $r(t)$ is often a physical requirement, and the Edgeworth expansion can not guarantee non-negativity. Joint cumulants of several variables have the desirable property that if any of the involved random variables are independent, then their joint cumulants become zero (whereas the TMs would factorize).

As summarized in Appendix B cumulants can be expressed as a non-linear recombination of equal and lower order TM [Kubo, 1962]. Cumulants are also related to the Laplace transform (LT): Applying the natural logarithm to the moment generating equation (MGE) yields the spectrum of so called cumulants κ_k. Cumulants are not able to reduce Eq. (5.2) because applying the logarithm converts terms (1) and (2) to mixed integro-differential expressions and, hence, irreversibly changes the character of the parabolic PDE. For these reasons, I can also exclude cumulants for my purposes.

5.3. Summary and Conclusions

In this chapter, I performed ***Step (I)*** of my overall approach. I investigated base functions different to the monomials that lead to TM. By analyzing them towards their reduction efficiency, I found that they can be classified exhaustively by a limited set of cases. By comparing these cases to the monomials that lead to TM, I found the following conclusions most important:

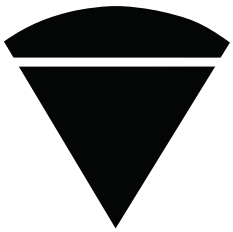
1. Any temporal characteristic based on arbitrary polynomial base functions $\xi_k(t)$ or on cumulants can be mimicked by TM through linear re-combination, and would not offer improved computational efficiency compared to TM.
2. Polynomial-based temporal characteristics in general contain the same information as TM, simply arranged in different linear combinations. They can, hence, not capture more information from the dynamic system.
3. As overall consequence, there is no way of temporal model reduction for dynamic systems based on arbitrary integral transforms with polynomial base functions that leads to more efficiently reduced systems than the monomials leading to TM.
4. The only remaining integral transform that reduces Eq. (4.5) to a non-coupled system of steady-state PDEs is the Laplace transform. In all applications of the Laplace transform involving forward problems that I could find, the number of characteristics necessary to capture the dynamic behavior was in the range of tens to hundreds. This is not satisfying for model reduction in comparison to the overall efficiency of TM, as I will illustrate in Chap. 6.

It should be mentioned that the derivations in Eqs. (5.1)-(5.9) hold for a quite generic form of PDEs. For example, this form includes parabolic PDEs (e.g., representing a dynamic confined groundwater model). The examples by Harvey and Gorelick [1995] or Luo et al. [2008] for advective-dispersive contaminant transport or MRMT, respectively, illustrate how this concept also holds for other parabolic or partial integro-differential equations. In fact, the results apply to any (system of) linear PDEs with the following properties:

- The spatial derivatives may have any arbitrary-order,
- There may be an arbitrary number of arbitrary-order time derivatives.
- For the integration by parts to work out, the coefficients must be independent of time and independent of the solution $r(t)$, and the dynamic model response has to decay to zero sufficiently fast.

In Chap. 6, I will proceed with ***Step (II)*** of my overall approach. To this end, I assess and compare the compression efficiency of the results from this chapter (i.e., the two based functions leading to the most swift evaluation of temporal characteristics, TM and LC).

6. Compression Efficiency via the Pre-Posterior Data Impact Assessor



In the last section I showed that there are only two ways of model reduction in time that lead to an efficiently reduced system, i.e., via the integral transformation based on monomial base functions that lead to temporal moments (TM) or via the Laplace transform providing the Laplace coefficients (LC). There remains, however, another fundamental question: How many TM or LC are necessary to achieve a sufficient degree of compression? Answering this question would also open the door to compare TM to LC. LC can be obtained from fully decoupled generating equations, as shown in Chap. 5. However, they conceal their most informative coefficients in the shadow of an arbitrary order. Comparing the resulting achievable compression efficiency is stated as ***Step (II)*** of my overall approach and carried out in the following chapter.

To this end, I develop a novel and versatile method called PreDIA (Preposterior Data Impact Assessor) to be introduced in Secs. 6.1-6.2. PreDIA analyzes, among others, the level of information carried by TM and LC. In Sec. 6.3 I employ PreDIA to scenarios where I consider reduced flow- and transport models, and discuss my observations in Secs. 6.4-6.5. Later in this thesis (Chap. 9), I will show the much more general application fields of PreDIA for optimal design of experiments. The material of the following chapter has been published in parts in Leube et al. [2012a] and Leube et al. [2012b].

6.1. General Approach

In order to answer the question on achievable compression efficiency, I propose first to rephrase the question and rather ask in a more intuitive context: What is the error between

a fully resolved dynamic model, e.g., represented by a time-series $r(t)$ and its appropriate reconstruction $\tilde{r}(t)$? The reconstruction would be based on a set of temporal characteristics α_k truncated at some order K with $k = 0 \dots K$. The answer is straightforward and can be expressed through the so-called L^2-norm:

$$L^2(K,t) = \left[r(t) - \tilde{r}(t)\right]^2 . \tag{6.1}$$

The individual choice of reconstruction techniques, e.g., maximum entropy [Jaynes, 1957], or Edgeworth series expansion [Kendall and Stuart, 1977] would, however, introduce an error of its own. To make my later analysis independent of the error in the specific reconstruction technique chosen, I propose to replace the L^2-norm by a statistically motivated norm like the Conditional Standard Deviation ($CStD$) (for similarity see Eq. 2.41):

$$CStD(K,t) = \left(Var_r\left[r(t)|\alpha_k\right]\right)^{1/2} \tag{6.2}$$

which is the conditional standard deviation of $r(t)$ given a set of temporal characteristics α_k with $k = 0 \dots K$ (e.g., TM or LC). The $CStD$ represents the motivation that TM (or LC) should at least be informative enough to identify response curves among a set of physically plausible random response curves $r(t)$. Such sets of plausible responses occur "naturally" in our context where a geostatistical model description is necessary in order to account for the ubiquitous lack of knowledge on model parameters (see Sec. 2.4). The $CStD$ is then build around the set of possible $r(t)$.

The crucial question is now, how to identify the conditional standard deviation. To emphasize this, I make a short excursion in order to point out the drawbacks of common approaches and make the reader more sensible to the novelty of my approach. The most commonly employed method to assess the conditional standard deviation of a data set u given just another data set u' is based on covariances. However, covariances restrict the analysis of the data sets to be linear or only weakly non-linear [e.g., Schweppe, 1973]. This behavior is illustrated in Fig. 6.1, showing a strictly linear case with proper fitting and a non-linear case revealing a total failure capturing the dependency with covariances. It should be mentioned that most cases occurring in nature are somewhere in between those two antipodes (strictly linear and strongly non-linear). There will be always a linear share of dependency properly captured and a non-linear share missed. The dependency between temporal characteristics α_k (e.g., TM or LC) and the underlying dynamic response $r(t)$ is not necessarily linear. Hence, a linear framework such as Ensemble Kalman Filters (EnKF) [e.g., Evensen, 2007; Nowak, 2009; Schöniger et al., 2012] or first-order second-moment methods (FOSM) [e.g., Kunstmann et al., 2002; Cirpka et al., 2004], might miss important features in a global assessment of TM or LC. To overcome this constraint, I propose a novel method being capable to properly capture non-linear dependencies (e.g., as occurring in Fig. 6.1 right).

The Pre-posterior Data Impact Assessor (PreDIA), as introduced in this chapter, is an extension of the Bootstrap Filters (BF) [Gordon et al., 1993] (see Sec. 2.6) towards information theory. BF rely on Monte Carlo (MC) simulations of random system responses and can, as a matter of nature, handle non-linear system dependencies [Snyder et al., 2008]. This opens the path to an assessment of the $CStD$ without loss of non-linear features. To this end, let us consider a sample of n_r potential response realizations $\mathbf{r}(t)$. Note, that $\mathbf{r}(t)$ is now written in

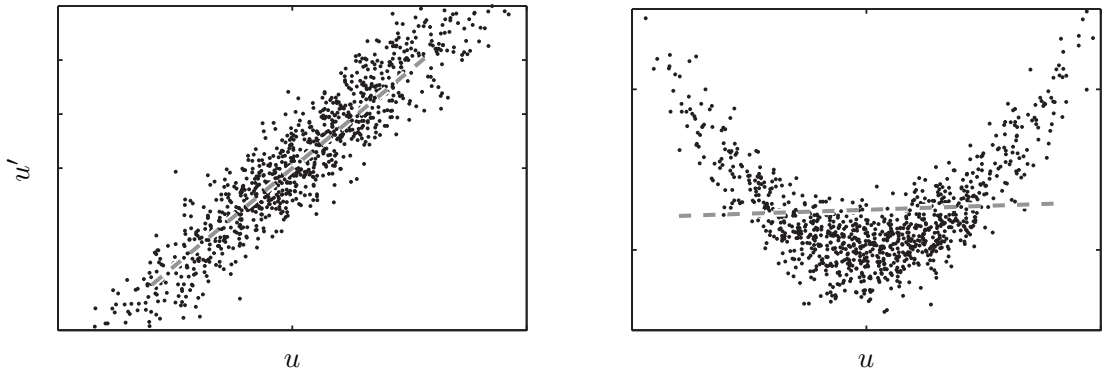


Figure 6.1.: Illustration of dependency structure between a linearly (left) and non-linearly (right) dependent set of data u and u' (black scattered cloud) and the performance of the covariance measure (gray dashed line) applied to the data sets in both cases.

bold as I consider a sample of n_r realizations. Then, for a given sample of n_α synthetically generated temporal characteristics $\boldsymbol{\alpha}_k = f(\mathbf{r}(t))$ with $k = 1 \ldots K$ and integral transformation $f()$ based on arbitrary base function ξ, the likelihood $L(\boldsymbol{\alpha}_k|\mathbf{r}(t))$ penalizing the discrepancy between any potential temporal characteristic α_k and the spectrum of overall possible characteristics $f(\mathbf{r}(t)$ can be calculated as

$$L(\boldsymbol{\alpha}_k|\mathbf{r}(t)) = pdf_{Normal}(f(\mathbf{r}(t)), \mathbf{R}_\varepsilon). \tag{6.3}$$

with Gaussian distribution around the mean $f(\mathbf{r}(t))$ and the $K \times K$ covariance matrix $\mathbf{R}_\varepsilon$. $\mathbf{R}_\varepsilon$ stems from the context of assimilating real data and refers to the physical measurement error [Evensen, 2007]. In our context, where $\boldsymbol{\alpha}_k$ are synthetically generated (and not taken from field observations), $\mathbf{R}_\varepsilon$ is better considered as a kernel bandwidth for kernel density estimations of pdfs [Silverman, 1986]. Given Bayes' theorem (see Sec. 2.51), the posterior distribution $p(\mathbf{r}(t)|\boldsymbol{\alpha}_k)$ can be determined according to:

$$p(\mathbf{r}(t)|\boldsymbol{\alpha}_k) \propto L(\boldsymbol{\alpha}_k|\mathbf{r}(t))\, p(\mathbf{r}(t)), \tag{6.4}$$

whereas $\mathbf{r}(t)$ represents the prior distribution. From that, $CStD(K,t)$ (see Eq. (2.41)) can be determined according to:

$$CStD(K,t) = \left\{ \mathrm{Var}_{\mathbf{r}(t)|\boldsymbol{\alpha}_k}[\mathbf{r}(t)] \right\}^{1/2}. \tag{6.5}$$

The final goal of PreDIA is to average $CStD(K,t)$ over the total spectrum of n_α possible $\boldsymbol{\alpha}_k$ values, arriving at the expectation (see Eq. (2.37)) over the Conditional Standard Deviation $R(K,t)$:

$$R(K,t) = \left\{ \mathrm{E}_{\boldsymbol{\alpha}_k} \mathrm{Var}_{\mathbf{r}(t)|\boldsymbol{\alpha}_k}[\mathbf{r}(t)] \right\}^{1/2}. \tag{6.6}$$

In the context of the analysis performed in this chapter, PreDIA is a specific case of Leube et al. [2012a]. A brief excursion to a different application as found in the original application is given in Chap. 9.

6.2. Implementation

The following section focuses on the implementation of PreDIA. For n_r realizations of $\mathbf{r}(t)$ independently drawn from $p(\mathbf{r}(t))$ and one (hypothetically given) α_k, a BF would evaluate an $n_r \times 1$ weight vector $\mathbf{w}$ according to Eq. (6.4) with $w_i = p(r_i(t)|\alpha_k)$. The weight vector $\mathbf{w}$ given a sequence of K TM is found be simply multiplying the individual likelihoods $L(\boldsymbol{\alpha}_K|\mathbf{r}(t)) = \prod_{k=0}^{K} L(\boldsymbol{\alpha}_k|\mathbf{r}(t))$. This is allowed because the individual likelihoods are assumed to be independent from each other (which is similar to the independence assumption among measurement errors [Evensen, 2007]). The optimal bandwidth for $\mathbf{R}_\varepsilon$ can be chosen according to simple "rule of thumbs" [Silverman, 1986] and is (in this analysis) kept constant for every time-step t of the response curves $r(t)$. Weighted averaging of n_r realizations $r_i(t)$, $i = 1...n_r$, yields $CStD(K,t)$ according to Eq. 2.39:

$$CStD(K,t) \approx \left\{ \frac{v_1}{v_1^2 - v_2} \left[\sum_{i=1}^{n_r} r_i(t)^2 \; w_i - \left(\sum_{i=1}^{n_r} r_i(t) \; w_i \right)^2 \right] \right\}^{1/2}, \tag{6.7}$$

with $v_1 = \sum_{i=1}^{n_r} w_i$ and $v_2 = \sum_{i=1}^{n} w_i^2$. Here, I approximate both quantities in the weighted sample sense and, therefore, employ v_1 and v_2 [Weiss, 2006, p. 355]. The corresponding correction factor in Eq. (6.7) resembles the well-known factor $\frac{1}{n-1}$ for the non-weighted sample variance. This makes Eq. (6.7) an unbiased estimator of the population variance.

Calculating the weight vector $\mathbf{w}$ for the set of n_α potential realizations of K sequences α_k yields an $n_r \times n_\alpha$ weight matrix $\mathbf{W}$

$$R(K,t) \approx \left\{ \frac{1}{n_\alpha} \sum_{j=1}^{n_\alpha} \frac{v_{1,j}}{v_{1,j}^2 - v_{2,j}} \left[\sum_{i=1}^{n_r} r_i(t)^2 \; W_{ij} - \left(\sum_{i=1}^{n_r} r_i(t) \; W_{ij} \right)^2 \right] \right\}^{1/2}. \tag{6.8}$$

Finally, $R(K,t)$ is normalized by its value for $k = 0$, i.e., by the uncertainty in absence of any TM (or LC). This yields the normalized compression error $R_n(K,t)$

$$R_n(K,t) = \frac{R(K,t)}{(\mathrm{Var}[\mathbf{r}(t)])^{1/2}}. \tag{6.9}$$

For time-integrated analysis, I first integrate $R(K,t)$ over time, normalize again, and define the total normalized compression error $R_t(K)$ by

$$R_t(K) = \left(\frac{\int_0^\infty R(K,t)^2 \; \mathrm{d}\, t}{\int_0^\infty \mathrm{Var}[\mathbf{r}(t)] \; \mathrm{d}\, t} \right)^{1/2}. \tag{6.10}$$

This nested scheme is illustrated in Fig. 6.2 illustrates the above introduced implementation. For the sake of simple illustration, I consider a function $y(t,a,b,c)$ described by three unknown parameters a through c:

$$y(t,a,b,c) = \exp\left(-a \; \log(t) - b \; \log(t)^2 - c \; \log(t)^3\right), \tag{6.11}$$

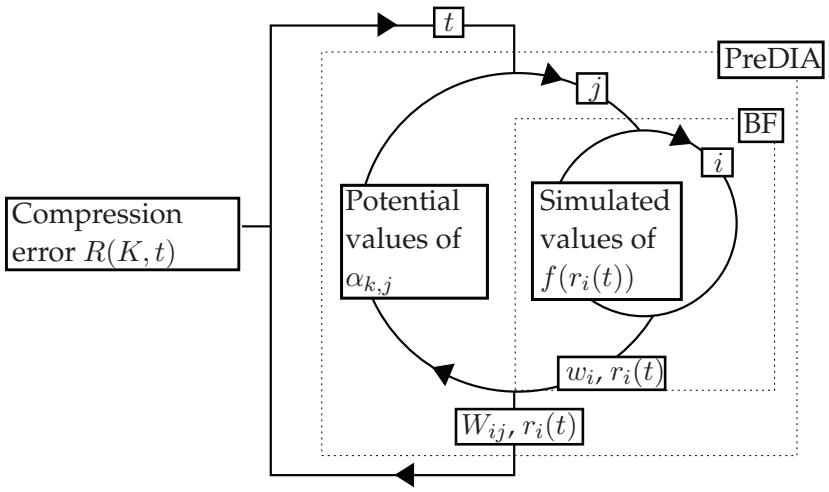


Figure 6.2.: Schematic illustration of PreDIA (Pre-posterior Data Impact Assessor) enveloping the BF (Bootstrap Filter). Both are nested together for the derivation of the compression error $R(K,t)$.

where parameters a through c are assumed to be log-normal distributed. The uncertainty in $y(t)$ due to the lack of knowledge on a through c is steadily decreased when learning about the parameters a through c, eventually decreasing to zero when knowing all 3 parameters. Storing only zero to three parameters instead of the entire curve resembles a data compression, and the resulting uncertainty in $y(t)$ integrated over time is the related compression error. When normalized by the compression error for zero parameters, this yields the normalized compression error $R_t(K)$ in Eq. (6.10).

6.3. Application

Two illustrative examples shall serve to investigate the compression efficiency of TM and LC. To this end, I employ scenario (1), the simple scenario from well flow (see Sec. 3.1) and scenario (2), the simple scenario from solute transport (see Sec. 3.2). For more details on the specific scenario parameters and boundary conditions I refer to Sec. 3.1 and 3.2. In order to investigate the sensitivity of the analysis results towards the choice of the underlying geostatistical scenario settings, I repeat the analysis for scenario (1) in several scenario variations where I vary the most relevant geostatistical parameters. This has been defined as cases (1a) - (1d). Varied parameters include σ_T^2, σ_S^2, $\lambda_{T,1,2}$ as summarized in Tab. 6.1. For the variation of the storage coefficient S, I follow the suggestion of Li et al. [2005], who reviewed the sparse literature on the variability of S and finally recommended to use a spatial constant with log-normal distribution and log-variance $\sigma_S^2 = 1$.

As for scenario (2), I vary the monitoring locations in cases (2a) - (2d). They are summarized in Tab. 6.2.

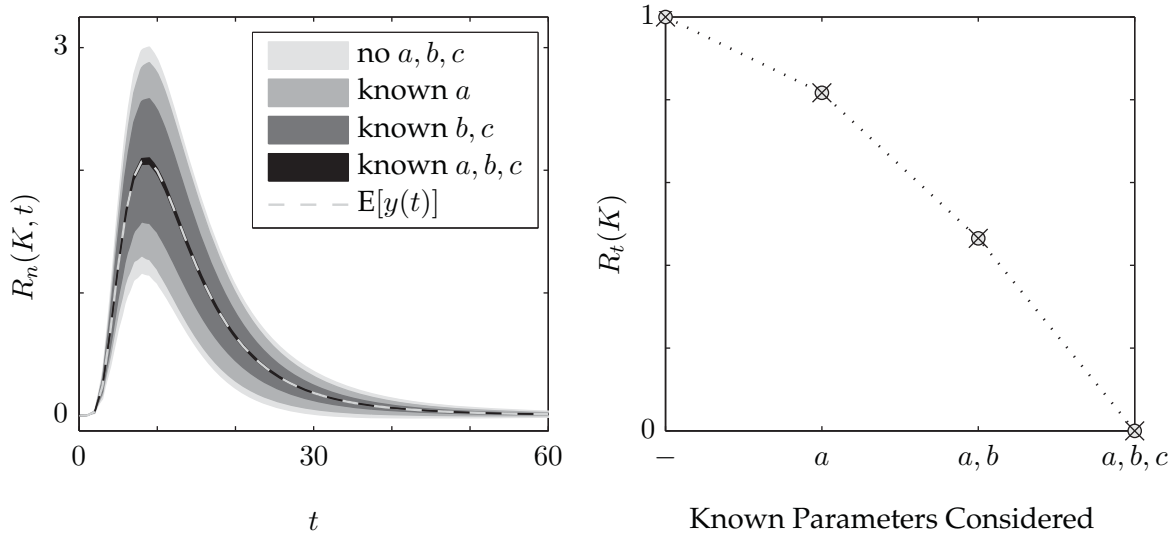


Figure 6.3.: Illustrative example on PreDIA. The unknown function y in Eq. (6.11) relying on parameters a through c is being analyzed regarding the remaining uncertainty when knowing only a limited number of its parameters. Left: Normalized compression error $R(K,t)$ ($\pm$ 2 standard deviations as shown as gray shaded areas) plotted around the expected response curve (dashed line). The lightest area indicates minimal knowledge. Knowledge of all three parameters eventually reduces the uncertainty down to zero. Right: The respective time-averaged error $R_t(K)$. Crossed circles indicate numerical values for $R_t(K)$.

Cases considered within scenario (1)

Case			(1a)	(1b)	(1c)	(1d)
Mean ln T	μ_T	ln[m^2/s]	10^{-3}	10^{-3}	10^{-3}	10^{-3}
Variance ln T	σ_T^2	ln^2[(m^2/s)^2]	0.8	2.5	0.8	0.8
Integral scale ln T	$\lambda_{T,1,2}$	[m]	[20,20]	[20,20]	[10,10]	[20,20]
Mean ln S	μ_S	ln[$-$]	10^{-3}	10^{-3}	10^{-3}	10^{-4}
Variance ln S	σ_S^2	ln^2[$-$]	-	-	-	1
Integral scale ln S	$\lambda_{S,1,2}$	[$-$]	-	-	-	[20,20]

Table 6.1.: Variation of geostatistical parameters in scenario (1).

To guarantee highly accurate sampling of $r(t)$ and its TM and LC respectively, I employ a MC ensemble consisting of 250k realizations for scenario (1) and 100k realizations for scenario (2), respectively. As a matter of the weighting-based importance resampling used in PreDIA, the accuracy of the analysis is degenerating with increasing TM sequence length K or, in general, with stronger conditioning on data [Leube et al., 2012a]. These limitations, also known as the "curse of dimensionality" or "filter degeneracy", have been the scope of

Cases considered within scenario (2)						
Case			(2a)	(2b)	(2c)	(2d)
Monitoring location	$[x_m, y_m]$	[m]	[140,50]	[260,90]	[80,120]	[140,120]

Table 6.2.: Variation of the monitoring location in scenario (2).

many studies in the past [e.g. Liu [2008]; Snyder et al. [2008]; Van Leeuwen [2009]]. I ensure that the study results are unaffected by this problem by assessing the associated MC-error of computing R, R_n and R_t by means of the non-parametric statistical Bootstrapping method [Efron, 1982]. This has, to the best of my knowledge, not been done before in the context of any reweighting-based MC analysis. For more details on the technical implementation of this Bootstrapping method I refer to Appendix C. For the solution of scenario (2), 100k Monte-Carlo realizations are generated in order to accurately resolve the spatial uncertainty.

6.4. Analysis based on Well Scenario (cases 1a - 1d)

In this section, I present and discuss the results based on scenario (1). In Sec. 6.4.1, I consider TM sequences of increasing higher order $0 \ldots K$ in order to investigate the cumulative compression efficiency towards loss-less compression. Sec. 6.4.2 considers individual TM of specific orders k to answer the question whether TM provide the most important information units first, and possibly in a strictly ordered fashion. Sec. 6.4.3 compares the results from TM against LC.

6.4.1. Cumulative Compression Efficiency

Fig. 6.4 (left) shows for case (1a) the expected response curve $\mathrm{E}[\mathbf{r}(t)]$. Here, $\mathbf{r}(t)$ represents the drawdown s at the monitoring location in absence of any TM data (dashed-dotted line), enveloped by its uncertainty $(\mathrm{Var}[\mathbf{r}(t)])^{1/2}$ (lightest gray shaded area). Using TM sequences of increasing highest order K then helps to know more and more details about the dynamic response as illustrated by the differently shaded areas $R(K,t)$ for various values of K. Obviously, a relatively high reduction of compression error is achieved by the first two TM in some time interval around the mean response time (zeroth TM only) and peak time (zeroth and first TM). The additional information when adding higher TM can hardly be seen in this type of visualization.

Fig. 6.4 (right) shows the time-averaged compression error $R_t(K)$ mapped against the highest TM order K for all four cases (1a) through (1d). The MC-error of $R_t(K)$, estimated by the Bootstrapping method, is visualized by the gray-shaded areas. These areas represent $\pm$ 2 standard deviations of assessing the $R_t(K)$-values. For all cases, $R_t(K)$ decreases strictly with increasing K, i.e., the longer the TM sequence considered. This is apparent, since

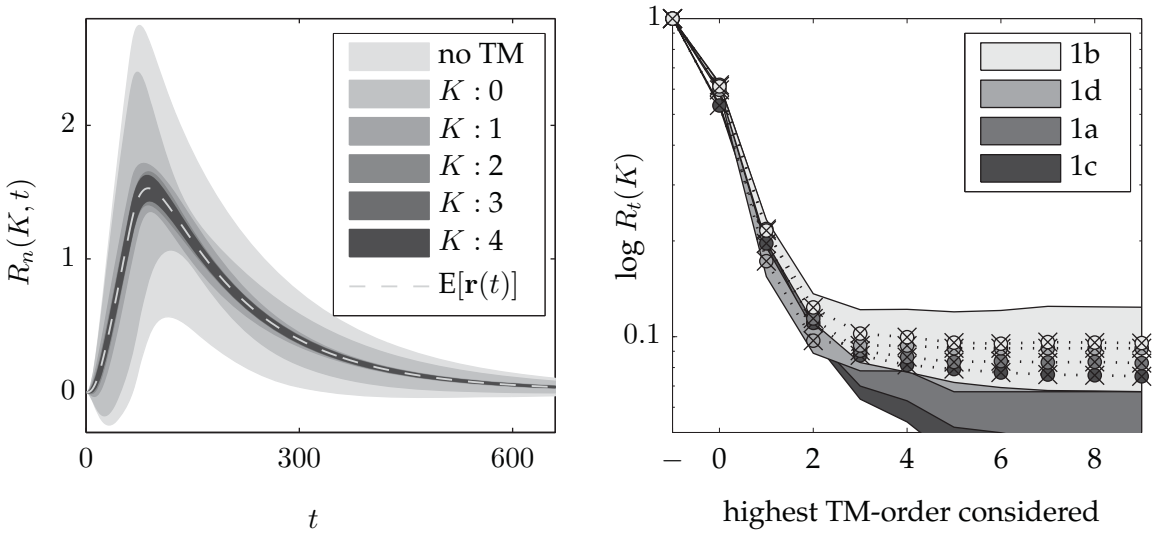


Figure 6.4.: Cumulative compression efficiency for TM featuring scenario (1). Left: Time-dependent compression errors $R(K,t)$ ($\pm$ 2 standard deviations as gray shaded areas) plotted around the expected response curve (dashed line) for TM sequences of increasing length K. Right: Their respective time integrals $R_t(K)$ for highest TM orders $0 \ldots 9$, and in absence of any TM data (marked by minus). Four cases (1a) - (1d) are compared. Crossed circles indicate numerical values for $R_t(K)$, shaded areas show $\pm$ 1 standard deviation of the MC-error in computing the $R_t(K)$ values.

longer TM sequences bear more information about the underlying time series. The first two TM (the zeroth and first) convey more than about 80 % of the information in all four cases, whereas the second and third TM contribute another 10 %. The remaining 10 % of information is distributed among an unquantifiable number of higher moments.

Comparing the different cases (1a) through (1d), I find the compression error generally identical with only slight differences (± 3 %). Although these differences appear to be small, they allow some meaningful insight into the driving physical processes: Case (1c) ranks comparatively best ($R_t(K)$ of (1c) is below that for 1a) in the sense that the overall information is concentrated best in the lower-order TM. This is because case (1c) causes less variability (smaller $\lambda_{T,1,2}$) associated with the possible dynamic shapes and features of drawdown curves compared to case 1a. With less variability in dynamic features, fewer units of information (a lower number of TM) suffice to infer the actual shape of the dynamic response.

The opposite behavior can be observed when analyzing cases (1b) and (1d). They introduce more variability compared to case (1a) (higher σ_T^2 for case 1b and uncertain S for case 1d), causing more variable dynamic features. Case (1d) produces even more drastic dynamic features through a much stiffer system with less diffuse behavior (small μ_S). Hence, both scenarios require additional information, i.e., more TM in order to achieve the same level of information.

All the above analyses may suffer to some extent from filter degeneracy making my results for longer TM sequences slightly less reliable. Based on my Bootstrapping-based error estimate, however, I found critical levels of filter degeneracy not to occur for a TM sequence length below 10. This is when most information (> 90 %) has already been captured, and so does not affect the conclusions I made above.

6.4.2. Individual Compression Efficiency

In the previous section, I analyzed the compression error of entire TM sequences of highest order $0 \dots K$. The final remaining question is: Is the order of $0 \dots k$ TM given by the recursive character of Eqs. (4.9)-(4.10) the one that provides the most informative TM at first? To this end, I analyze $R_t(k)$ individually for every TM of order k. The results are shown in Fig. 6.5 (left).

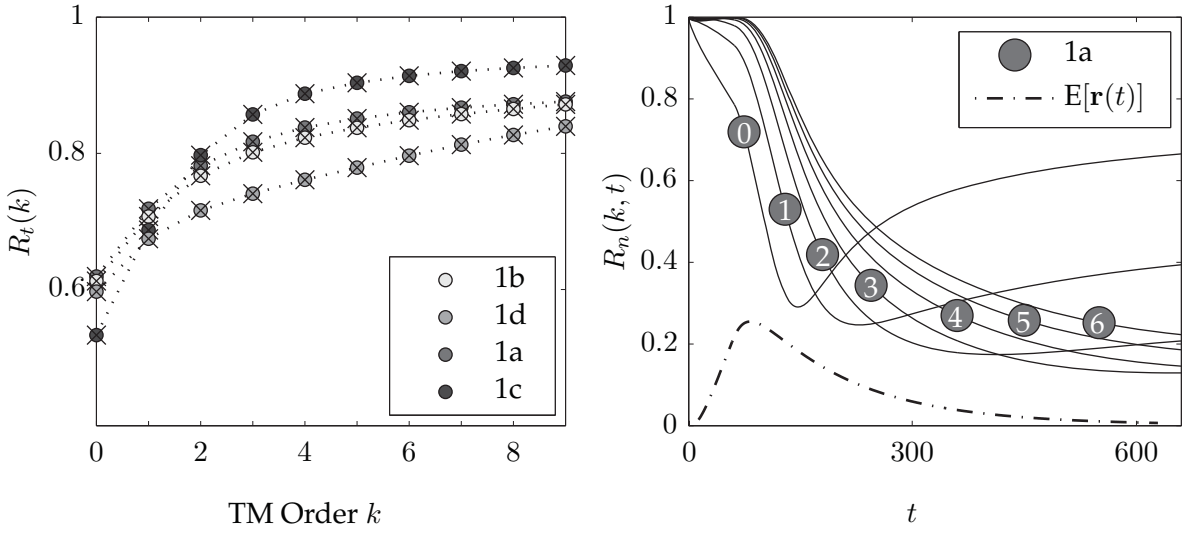


Figure 6.5.: Individual compression efficiency for TM in an application from groundwater flow (Scenario 1). Left: Analysis of the compression error $R_t(k)$ evaluated individually for every TM of order $k = 0 \dots 9$ for cases (1a) through (1d). Right: Temporal evolution of the compression error $R_n(k,t)$ for individual TM of order $k = 0 \dots 6$ in scenario 1a. The expected response curve $\mathrm{E}[\mathbf{r}(t)]$ is illustrated by the dashed-dotted curve.

Obviously, for all cases, lowest-order TM are again the more informative ones, and the compression error $R_t(k)$ of individual TM is steadily increasing with increasing order k and eventually climbs up to 90 %. The same behaviour can be observed for scenarios 1b through 1d.

For cases (1b) and (1d), higher-order TM convey more important information compared to cases (1a) and (1c), while they contribute much less in the cumulative analysis in Chap. 6.4.1. This seeming inconsistency is explained by the fact that TM are not statistically independent,

i.e., they are not orthogonal, convey partially redundant information and their information content is not simply additive. Apparently, the more variable cases (1b) and (1d) produce more redundancy among different TM.

Fig. 6.5 (right) shows the time dependence of $R_n(k,t)$ for TM of increasing order k on the example of case (1a). The curves indicate that each TM has specific time ranges in which it contributes most information. For higher-orders, the conveyed information is shifted to later times due to the increased leverage of higher order monomials at later times. Thus, higher-order TM capture later-time features of the response curve. This turns out important in slug-like pumping test analyses, where even the late-time features of drawdown recovery still contribute valuable information for estimating transmissivity [Oliver, 1993; Wu et al., 2005; Zhu and Yeh, 2006]. Also, the late-time behavior of solute breakthrough curves is important to identify non-Fickian transport phenomena, e.g., in multi-rate mass transfer models [Haggerty and Gorelick, 1995; Luo et al., 2008].

When working with TM of noisy time series measured in the field, in the context of inverse modeling, higher-order TM may be subject to large errors. Such errors have the potential to compromise their information content, requiring more TM to compensate for that loss. While this is not the scope of this chapter, it will be discussed further in the outlook.

6.4.3. Comparison against Laplace Case

As described in Chap. 5.2.4, LC are computationally attractive since they can be computed independently. However, it is a priori unknown which parts of the spectrum or precisely which LC will be most informative. This triggers the question if there exists a set of LC (and if yes, then which one) that is superior to TM in terms of information content.

I mimick the lack of knowledge on the optimal choice of LC sets by randomly sampling from a large spectrum of potential Laplace variables u, with $u = 10^{-5} \ldots 10^{5}$. I repeat this 500 times and measure the total normalized compression error $R_t(K)$ for different sequence lengths K in each repetition. As physical scenario I use scenario (1) with case (1a).

Fig. 6.6 shows the resulting total normalized compression error for the ensemble of 500 sets (gray lines), and the ensemble mean (dashed line). For comparison, I include the results for TM (dashed-dotted line) obtained in Sec. 6.4.1. I observe there is only a small fraction of LC sets performing slightly better than TM. The ensemble mean (expected Laplace performance), however, performs considerably worse than TM. Because the optimal set of LC is unknown in practical applications, the possibly better performance of LC can not be exploited.

Whether or not this disadvantage of LC could be outbalanced by the advantage of easier curve reconstruction for $K > 10/40$ [e.g., Li et al., 1992] will depend in the specific application context.

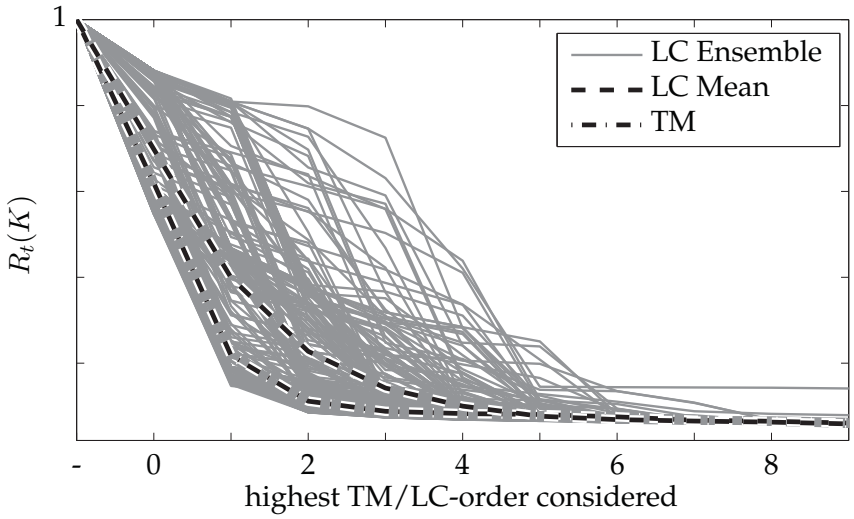


Figure 6.6.: Comparison between compression efficiency of LC and TM. Total normalized compression error $R_t(K)$ for sets of randomly drawn LC (grey lines), its mean (dashed line) and, for comparison the evolution of the respective TM case (dashed-dotted line).

6.5. Analysis based on Transport Scenario (cases 2a - 2d)

The next section considers the application of PreDIA to scenario (2) including cases (2a) - (2d). This section is concerned with the same analysis done in Sec. 6.4. Sec. 6.5.1 considers cumulative TM sequences of increasing order $0 \dots K$, whereas the individual analysis is not conducted for the transport scenario.

6.5.1. Cumulative Compression Efficiency

Similar to Fig. 6.4, Fig. 6.7 (left) shows the expected response curve $\mathrm{E}[\mathbf{r}(t)]$. Here $\mathbf{r}(t)$ is the normalized concentration c/c_0 for case (2b) in absence of any TM data (dashed-dotted line), enveloped by its uncertainty $(\mathrm{Var}[\mathbf{r}(t)])^{1/2}$ (lightest gray shaded area). Using TM sequences of increasing highest order K then helps to know more and more details about the dynamic response as illustrated by the decreasing shaded areas $R(K,t)$ for various values of K. Obviously, $(\mathrm{Var}[\mathbf{r}(t)])^{1/2}$ has a much smaller larger magnitude related to $\mathrm{E}[\mathbf{r}(t)]$ as compared to the drawdown in scenario (1). This means, the featured transport scenario is subject to more uncertainty. Also, the zeroth TM carries almost no information: The lightest gray shaded area is almost not distinguishable from the next darker. Only the first and next higher TM deliver considerable informations on $\mathbf{r}(t)$, reducing the gray-shaded areas more distinctly.

Fig. 6.7 (right) shows the time-averaged compression error $R_t(K)$ mapped against the highest TM order K for all four cases (2a) - (2d). The MC-error of $R_t(K)$, estimated by the Bootstrapping-based method, is again plotted around $R_t(K)$, however, the areas are not

visible with naked eyes in this case. This is due to the lower $(\mathrm{Var}[\mathbf{r}(t)])^{1/2}$ values leading to smaller MC-errors.

For all monitoring locations, $R_t(K)$ decreases strictly with increasing K, i.e., the longer the TM sequence considered, and reaches 10 % for sequences up to three and more TM. The remaining information is distributed among an unquantifiable number of higher moments. The interesting observation is that the cases featuring monitoring locations which are close to or aligned with the center line of the expected plume path, and hence are almost always hit by the plume (cases 2a and 2b) have a zeroth TM with no information content. This is obvious because all observable responses receive the identical mass (namely the initialized mass $c/c_0 = 1$). By contrast, the cases featuring monitoring locations flanking the expected plume path north (cases 2c and 2d) are seldom hit by the plume. The zeroth TM can take on values between zero and the initialized mass, and provides more distinguishable observations. This leads to an uncertainty reduction of almost 20 % (please note the logarithmic scale in Fig. 6.7, right).

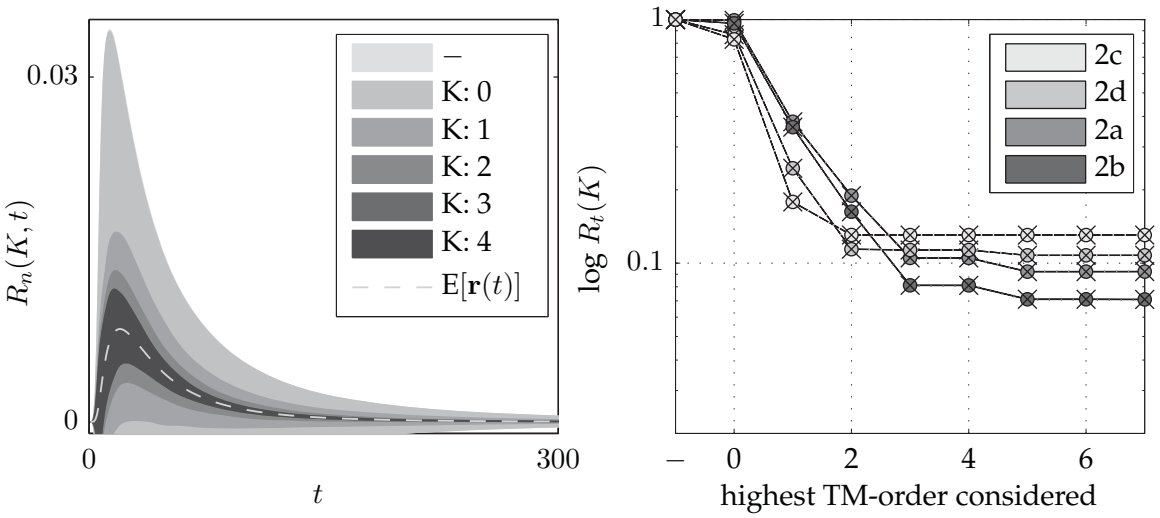


Figure 6.7.: Left: Time dependent compression errors $R(K,t)$ ($\pm$ 2 standard deviations as gray shaded areas) plotted around the expected response curve (dashed line) for TM sequences of increasing length K. Right: Their respective time integrals $R_t(K)$ for highest TM-orders $0 \ldots 7$, and in absence of any TM data (marked by minus). Four different cases (2a) - (2d) featuring different monitoring locations are compared. Crosses and circles indicate numerical values for $R_t(K)$, shaded areas (not visible by eyes in this case) show $\pm$ 1 standard deviation of the MC-error in computing the $R_t(K)$ values.

6.6. Summary and Conclusions

This chapter aimed at assessing the compression efficiency of temporal model reduction based on TM and LC defined as (***Step II***) of my thesis. To this end, I developed a novel and versatile method called PreDIA (Preposterior Data Impact Assessor). The method was then applied to two physical scenarios from subsurface flow and transport.

I found the following conclusions most important:

1. Based on an example from groundwater flow, the first two TM cover more than 80 % of the information required to identify a drawdown curve. Considering up to four TM captures 90 % or more of the overall information. The remaining 10 % of information are distributed among an unquantifiable number of higher moments. The lowest-order TM are always the most informative.

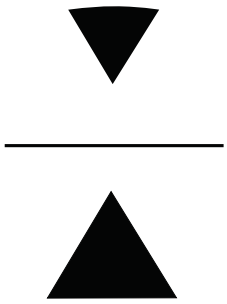
2. The distribution of information content over time differs among the TM orders. Late-time behavior can mostly be inferred from higher orders. The relevance of higher-order TM has to be judged in the light of any specific application task.

3. This is by far better than what I found for LC. One advantage of LC is that their equations are fully decoupled, such that arbitrary coefficients can be chosen in arbitrary order. This is, at the same time, their greatest disadvantage, because it is a priori unknown which ones are the most informative ones. Hence, it will be close to impossible to pick the optimal set of LC that could compete with TM.

4. For an example from contaminant transport, I arrive at almost the same conclusions than for groundwater flow. The interesting observation here is that the zeroth TM is only informative for monitoring locations flanking the plume (in the ensemble sense).

The next step of my overall approach (***Step III***) is to reflect these results against the discretization or reduction of other model dimensions, i.e., the number of repeated model runs through MC simulations. This step is performed in Chap. 7.

7. Reflection within Optimal Resource Allocation



In the last chapter I, showed how to rationally quantify the compression efficiency of temporal characteristics such as temporal moments (TM) or Laplace coefficients (LC). By comparing both TM and LC, I found that TM are superior to LC in terms of their compression efficiency. Also I found that, there is only a limited number of TM necessary to reduce the error in a truncated compression to an acceptable level. As the reduction of the temporal dimension is only one possibility to cope with increasing computational costs (see Sec. 1.1) it seems reasonable to ask how TM would compete against the discretization or reduction of other model dimensions (e.g., the numerical discretization of the physical domain or the number of repeating model runs through Monte Carlo (MC) simulation in order to resolve the parameter space).

This is defined as ***Step (III)*** of my overall approach and aims at more wisely allocating limited computational resources whilst keeping the joint error from combinations of discretization or model reductions at a global minimum. To this end, I present a new approach, denoted as optimal resources allocation (ORA), to be introduced in Sec. 7.1 and applied to a synthetic test scenario in Sec. 7.2. Parts of this chapter have been published in Leube et al. [2013].

7.1. Methodology

In the following, I treat the number of TM to be used as a resource allocation problem. The potential gain of accuracy (yield side) due to using more TM is assessed against the possible

gain in computational time by reduction or coarser discretizations of other model dimensions. The general idea is that there always exists an optimum ratio between these different model dimensions. The optimum can be found in the light of their associated computational costs (cost side). In order to find that optimal compromise, one has to properly quantify the yield and cost side via meaningful metrics. Sec. 7.1.1 shows how to generally quantify yields, whereas Sec. 7.1.2 addresses the computational costs. Sec. 7.1.3 then describes how to assemble these cost and error surfaces into an overall joint cost-to-error surface serving to finally determining the optimal trade-off between different combinations of discretized or reduced model dimensions.

7.1.1. Error Metrics

Let us suppose a hydro(geo)logical model output is defined by space varying in the $P \times T \times S$ domain, where P refers to the physical space, T is the time space and S represents the parameter or stochastic space. Then, errors associated with the numerical solution of SPDEs originate from either limited spatial resolution $\Delta\mathbf{x}$ of the physical space domain P, limited temporal resolution Δt of the time axis T or limited series of TM m_k with $k = 0 \dots K$, or limited statistical resolution $\Delta\xi$ of the parameter space S, respectively. In the most general formulation, the combined overall error in evaluating Ω, caused by its discretization in P, T and S, can be found by:

$$R_{Tot}(\Delta\mathbf{x}, \Delta t, \Delta\xi) = \frac{1}{V_a} \int_{\overline{a}} E_a(\Omega)\, \mathrm{d}\,\overline{a} \quad \text{with} \quad a \in P, T, S, \tag{7.1}$$

where E_a denotes some suitable, task-specific error norm (e.g., the L^2-norm) for the combined total error in the quantity of interest Ω and V_a is the total volume in space a to be integrated. E_a with $a \in P, T, S$, acts on either P, T or S, or on suitable combinations. It is enveloped by an outer, possibly multi-dimensional integral aggregating the yet remaining (complementary) dimensions $\overline{a}$ (such that $a \times \overline{a} = P \times S \times T$), ensuring R_{Tot} becomes a scalar quantity. Therefore, E_a can be employed to evaluate errors for arbitrary predictions, including point predictions, areal predictions, percentiles, mean values, variances or other statistics of predictions. Specific examples for illustration will be provided in Sec. 7.2.

While the L^1 or L^2-norms and variances are commonly used measures, most specific tasks will suggest the use of other norms. As an example, risk assessment problems may ask for minimal error probability in predicting that a critical risk percentile has been exceeded [e.g. de Barros and Rubin, 2008]. In general, one should measure the quality or confidence of the decision support offered by the model on the level of the overall modeling and simulation purpose [Saltelli et al., 2008]. In Chap. 9, I employ specific task-driven error norms and also show their applicability within the method.

The error as defined in Eq. (7.1) may simplify significantly, if independence of the individual error components within E_a from discretizations P, T, or S is assumed. For instance, when choosing L^2-type norms, independence leads to simple additivity of individual errors in each space P, S and T, respectively. The test cases in Chap. 9 will exemplify, however, that such assumptions have to be treated with care.

It should also be kept in mind that the notation of $\Delta\mathbf{x}$ for spatial discretization is not necessarily implying a fixed uniform grid spacing.

7.1.2. Budget Metrics

The available computational budget is typically limited. Thus, it is necessary to quantify the required budget of any proposed computational design. This can be done in different ways, either by measuring the simulation costs in units of elapsed time or by counting computational floating point operations (FLOPS). FLOPS are a direct measure and often used to specify asymptotic complexities (at the limit of fine discretizations) from theoretical properties and construction of specific algorithms. However, wall-clock run time may often be preferable because it accounts for inefficiencies of operation systems and programming languages, and always relies on the respective present computer architecture. Using wall-clock time in a given system will make the resource allocation specific to that system.

I define $B_{Tot}(\Delta\mathbf{x}, \Delta t, \Delta\xi)$ as the total required budget:

$$B_{Tot}(\Delta\mathbf{x}, \Delta t, \Delta\xi) = \int_S \Sigma(\Omega)\,\mathrm{d}\,S. \tag{7.2}$$

Here, Σ is some budget norm (e.g., FLOPS, wall-clock time) quantifying the budget required to evaluate one realization of the model at the given resolution $\Delta\mathbf{x}$, Δt. The integral over the probability space is required because the convergence of the space/ time solver may depend on the degree of value contrasts in the parameter vector, which may change between different realizations.

For the total budget B_{Tot}, independence assumptions are quite plausible. For example, the number of repeated simulations to resolve the event space S will often simply multiply with the average evaluation time per space/time solution, denoted as $\overline{\Sigma}$:

$$B_{Tot}(\Delta\mathbf{x}, \Delta t, \Delta\xi) = n_{MC} \cdot \overline{\Sigma}, \tag{7.3}$$

where n_{MC} is the budget multiplier to generate n_{MC} realizations Ω at the resolution $\Delta\mathbf{x}$ and Δt. If (asymptotic) computational complexity models for the spatial or temporal resolution are available, they can be fitted to experimental values of $\overline{\Sigma}$ or B_{Tot} at low $\Delta\mathbf{x}$ and Δt values. Sometimes, the budget may be linear in the number of time steps. However, linear relationships as in Eq. (7.3) are not necessarily true since later realizations or time steps might benefit from speedups from common grid-setup or from computing the solution to the first realization as initial guesses.

The proposed framework for optimal resource allocation can also support the choice between different numerical higher-level computing strategies. For example, if parallel computation is a viable option, there could be trade-offs involving running multiple small calculations in contrast to one large calculation influencing the balance between more MC sampling and finer grids. Such choices complicate the experimental designs required to search for the optimal resource application. Therefore, the examples provided later in this work will be restricted to a non-parallel strategy.

7.1.3. Methodology

The final goal of ORA is to assemble both the cost surface and the error surface into an overall joint cost-to-error surface in order to optimally allocate resources. Estimating $R_{Tot}(\Delta\mathbf{x}, \Delta t, \Delta\xi)$ as a function of $\Delta\mathbf{x}$, Δt and $\Delta\xi$, yields the error surface shown in Fig. 7.1 a). The same surface for time budget $B_{Tot}(\Delta\mathbf{x}, \Delta t, \Delta\xi)$ is shown in Fig. 7.1 b). The next step

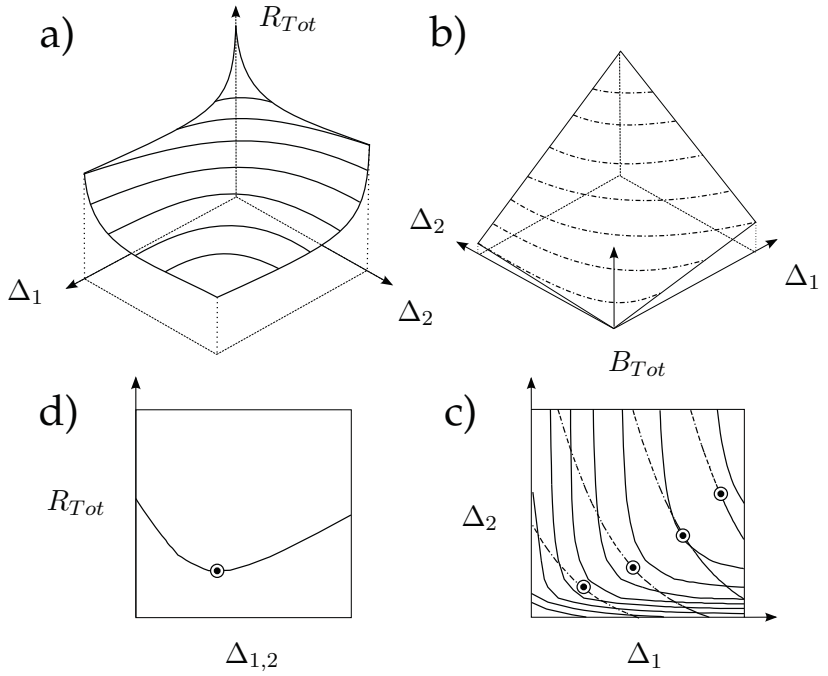


Figure 7.1.: Illustration of the ORA scheme. In clockwise direction: a) Joint error surface with iso-error-lines R_{Tot}, b) joint computational budget surface with iso-budget-lines B_{Tot}, c) assembled error-to-budget isoplot with optimal resource pairs (illustrated by black circles) and d) optimal resource pair Δ_1 and Δ_2 exemplary for one given budget. Optimal resource pairs are found by minimizing the joint error along the budget isoline that corresponds to a target budget.

is to derive the budget-to-error surface which is done by projecting $B_{Tot}(\Delta\mathbf{x}, \Delta t, \Delta\xi)$ onto $R_{Tot}(\Delta\mathbf{x}, \Delta t, \Delta\xi)$ as shown in Fig. 7.1 c).

The resulting surface reveals, for every iso-surface of given time budget B_{Tot}, an optimal ratio of $\Delta\mathbf{x} = \Delta\mathbf{x}_{opt}$ against $\Delta t = \Delta t_{opt}$ and $\Delta\xi = \Delta\xi_{opt}$, allowing the total error $R_{Tot,min}(\Delta\mathbf{x}_{opt}, \Delta\mathbf{t}_{opt}, \Delta\xi_{opt})$ to reach its minimum. The optima can be regarded as the minimal of the cost-slice along the contour for a given budget. I hypothesize that the error surface is convex, leading to convex functions along each budget isosurface, although the proof is beyond the scope of this thesis. The algorithm for constructing the budget-to-error surface involving all four prescribed steps is illustrated in Fig. 7.2.

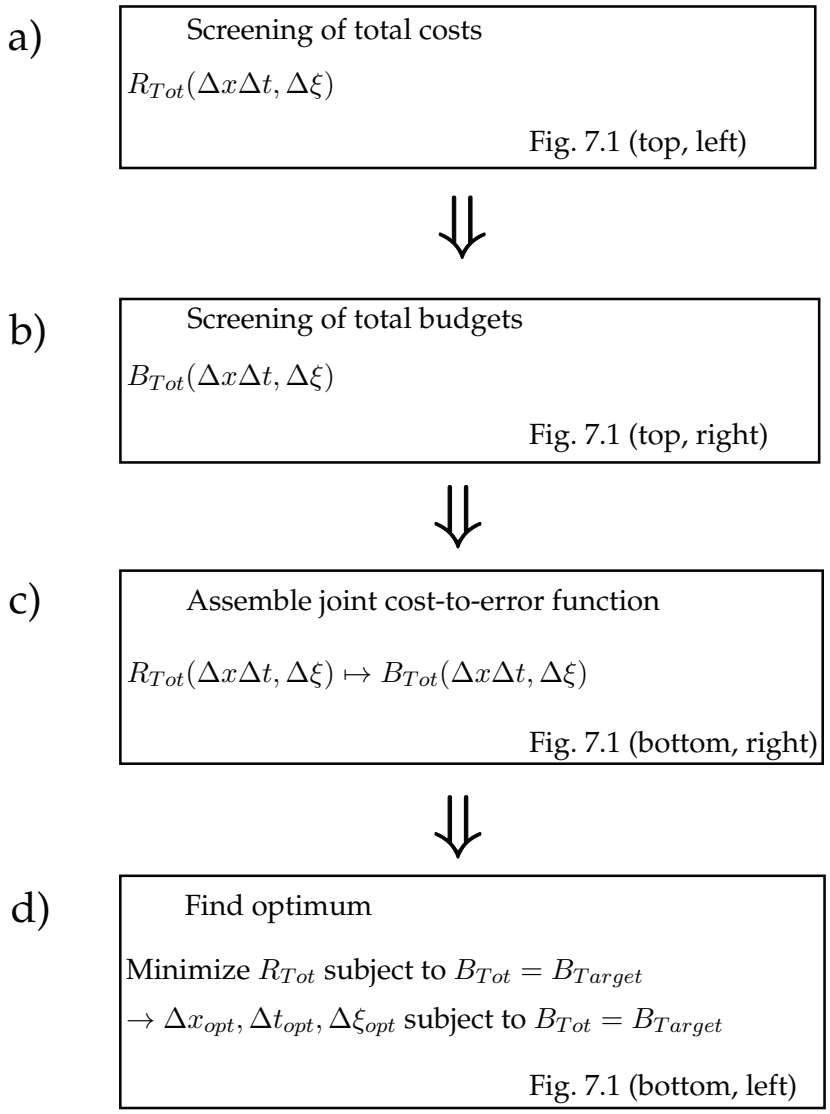


Figure 7.2.: Flowchart for assembling error and budget surfaces R_{Tot} and B_{Tot}, respectively and to find the optimal allocation for a given budget $B = B_{Target}$.

7.2. Application

In order to quantify and determine the optimal allocation of TM, I apply the methodology to the four cases (1a) - (1d) from scenario (1) and the four cases (2a) - (2d) from scenario (2). The time resolution parameter Δt will simply be replaced by the highest order K of TM considered.

The driving questions I wish to answer are:

- Does the hypothesized optimal trade-off between stochastical $\Delta \xi$ (number of repeated model runs through MC simulations) and temporal resolution Δt (highest order of

TM) actually exist?

- What key factors influence the position of the optimal trade-off?
- Are the different components of approximations errors independent anywhere in the relevant range around the optimum (tackled only in Chap. 9)?

The predictions $\Omega(\Delta t, \Delta\xi)$ for cases (1a) - (1d) are defined as the mean drawdown curve at the monitoring well $x_m = [25, 25]$. In order to investigate the sensitivity of ORA towards the choice of the underlying geostatistical scenario settings I vary again the most relevant geostatistical parameters throughout cases (1a) - (1d) similar to the application in Sec. 6.3. Varied parameters include σ_T^2, σ_S^2, $\lambda_{T,1,2}$ as summarized in Tab. 7.1. The predictions $\Omega(\Delta t, \Delta\xi)$ for

Cases considered within scenario (1)						
Case			(1a)	(1b)	(1c)	(1d)
Mean ln T	μ_T	ln[m^2/s]	10^{-3}	10^{-3}	10^{-3}	10^{-3}
Variance ln T	σ_T^2	ln^2[(m^2/s)^2]	0.8	2.5	0.8	0.8
Integral scale ln T	$\lambda_{T,1,2}$	[m]	[20,20]	[20,20]	[10,10]	[20,20]
Mean ln S	μ_S	ln[$-$]	10^{-3}	10^{-3}	10^{-3}	10^{-4}
Variance ln S	σ_S^2	ln^2[$-$]	-	-	-	1
Integral scale ln S	$\lambda_{S,1,2}$	[$-$]	-	-	-	[20,20]

Table 7.1.: Variation of geostatistical parameters in scenario (1).

cases (2a) - (2d) are defined as the mean breakthrough curves at different monitoring wells as defined in Tab. 7.2. To vary $\Delta\xi$ and Δt, respectively, $n_{MC} = 2 \times 10^5$ Monte Carlo realiza-

Cases considered within scenario (2)						
Case			(2a)	(2b)	(2c)	(2d)
Monitoring location	$[x_m, y_m]$	[m]	[140,50]	[260,90]	[80,120]	[140,120]

Table 7.2.: Variation of the monitoring location in scenario (2).

tions and different TM sequences ranging from $K = 0 \ldots 9$ were generated according to the transient and moment-generating equations in Sec. 2.1 and Sec. 4.3, respectively.

The error norm E_a accounts for both the temporal dimension P and the stochastic dimension S, respectively, in the sense of Eq. (7.1). Assuming independence of both errors significantly simplifies the joint error. For L^2-type norms, the joint error can be determine by simple addition of the individual errors:

$$R_{Tot}(K, n_{MC}) = E_P(K, n_{MC}) + E_S(K, n_{MC}). \tag{7.4}$$

For the temporal error $E_P(K, n_{MC})$, I employ the same error metric as introduced in Sec. 6.2 to assess the compression efficiency of TM:

$$E_P(K, n_{MC}) \approx$$
$$\int_0^\infty \frac{1}{n_{MC}} \sum_{j=1}^{n_{MC}} \frac{v_{1,j}}{v_{1,j}^2 - v_{2,j}} \left[\sum_{i=1}^{n_{MC}} \Omega_i(r(t), n_{MC})^2 \, W_{ij} - \left(\sum_{i=1}^{n_{MC}} \Omega_i(r(t), n_{MC}) \, W_{ij} \right)^2 \right] \mathrm{d}\, t. \tag{7.5}$$

The stochastic error $E_S(K, n_{MC})$ due to insufficient sampling by MC realizations is determined with the non-parametric statistical bootstrapping method [Efron, 1982]:

$$E_S(K, n_{MC}) = \frac{1}{B-1} \sum_{b=1}^{B} \left[\Omega^{(b)}(K, n_{MC}) - \overline{\Omega}(K, n_{MC}) \right]^2, \tag{7.6}$$

where $b = 1 \; ... \; B$ indicates the repetition over randomly drawn subsets with size $n_{MC}^* < 2 \times 10^5$ from the total number of potential realizations 2×10^5, and $\Omega^{(b)}$ is the respective prediction obtained by working with the $b-$th subset instead of n_{MC}. $\overline{\Omega}(K, n_{MC})$ denotes the average over the set of B realization subsets.

The average wall clock time required to generate one realization of $\Omega(K, n_{MC} = 1)$ at highest TM order, denoted as $\overline{\Sigma}$, is utilized as the computational budget norm for all cases. Then $B_{Tot}(\Delta \mathbf{x}, \Delta\xi)$ can be found by:

$$B_{Tot}(\Delta t, \Delta\xi) = n_{MC} \cdot \overline{\Sigma}, \tag{7.7}$$

where n_{MC} is the budget multiplier to generate n_{MC} realizations $\Omega(\Delta \mathbf{x}, \xi_k)$ at the resolution Δt.

Based on the algorithm described in Sec. 7.1 and Eqs. (7.4)-(7.7), I evaluate error-to-budget surfaces for all cases (1a) - (1d) and cases (2a) - (2d), and use these in order to find the optima. Results of the optimal resource patterns are presented in Sec. 7.3.

7.3. Results and Discussion

7.3.1. Simple Well Scenario (cases 1a - 1d)

The results for cases (1a) through (1d) are given in Fig. 7.3. Quite intuitively, their trade-offs move from low to high MC samples and from few to many TM. This behavior can generally observed for all involved cases (1a) trough (1d). The explanation is that temporal and statistical discretization should go hand in hand, because it does not make sense to statistically refine a poor dynamic description and vice versa. Surprisingly, when employing typical TM sequence lengths (K $\leq$ 4, see discussion in Sec. 4.4), all cases achieve this optima with less MC samples ($n_{MC} \leq 500$) then typically used in literature [e.g., Englert et al., 2006; Nowak et al., 2008].

By comparing the differences in the individual patterns, it can be observed that more variability (case 1b has $\sigma_T^2 = 2.5$ and $\lambda_{T,1,2} = [20, 20]$) tends to require larger Monte-Carlo

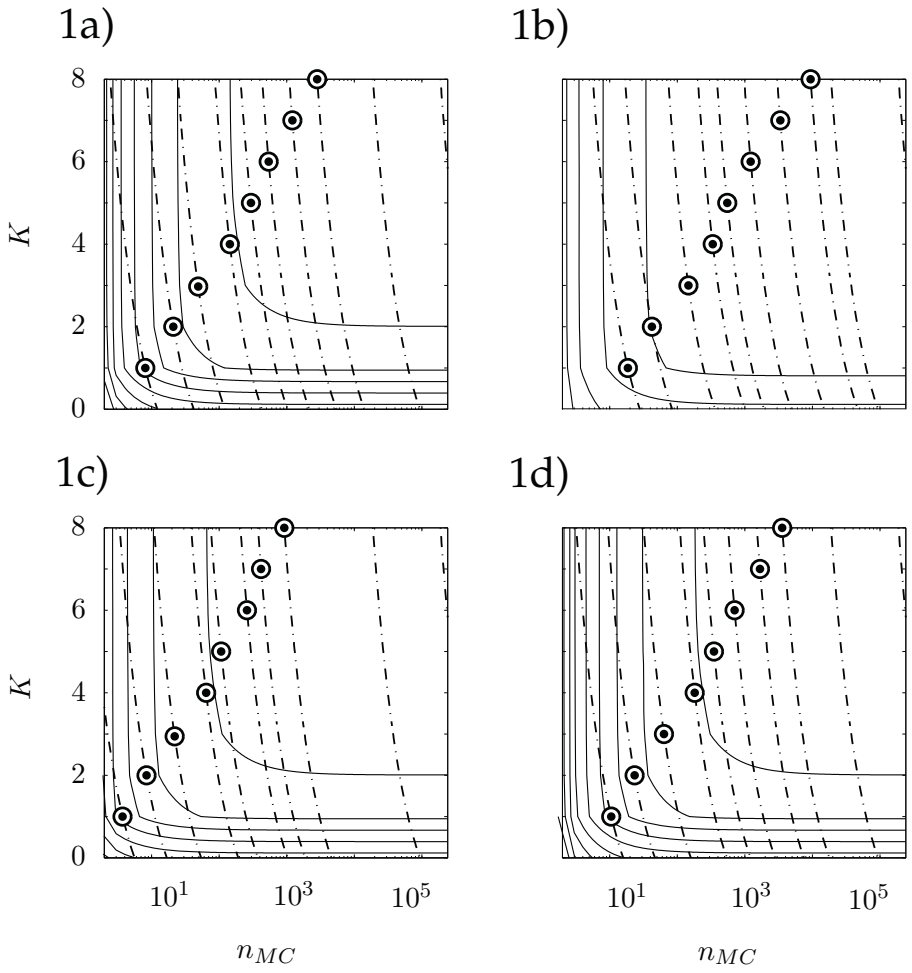


Figure 7.3.: Error-to-budget surfaces of the four pressure-based cases (1a) through (1d). Optima for given budgets are depicted by circles. Dashed-dotted contours mark the budget-surface, whereas the solid lines represent the error-surfaces.

samples, whereas less variability (case 1c has $\sigma_T^2 = 0.8$ and $\lambda_{T,1,2} = [10, 10]$) leads to less MC samples as compared to case (1a). Surprisingly, case (1d) is not in line with the aforementioned findings. Case (1d) reveals a pattern similar to case (1a) although it involves more variability due to the uncertainty in the storage coefficient S.

This surprising behavior can be explained by looking at the statistical variance and skewness of system responses as drivers for increased MC needs versus shape characteristics of individual response curves as drivers for increased time resolution needs. Fig. 7.4 (left) shows the time-dependent variance $R_t(K = 0)$ of $\Omega(K, n_{MC})$ at minimal reduction (zero TM) together with the time-averaged skewness values (see legend). For more details on the definition of $R_t(K)$, I refer to Eq. (6.8) in Sec. 6.2.

By comparing the magnitudes of $R_t(K = 0)$ for all four cases (Fig. 7.4, left), I find consistently increased magnitudes for the cases where more parameter uncertainty is involved (cases 1b and 1d) and a slightly decreased magnitude for case (1c) with less uncertainty. The

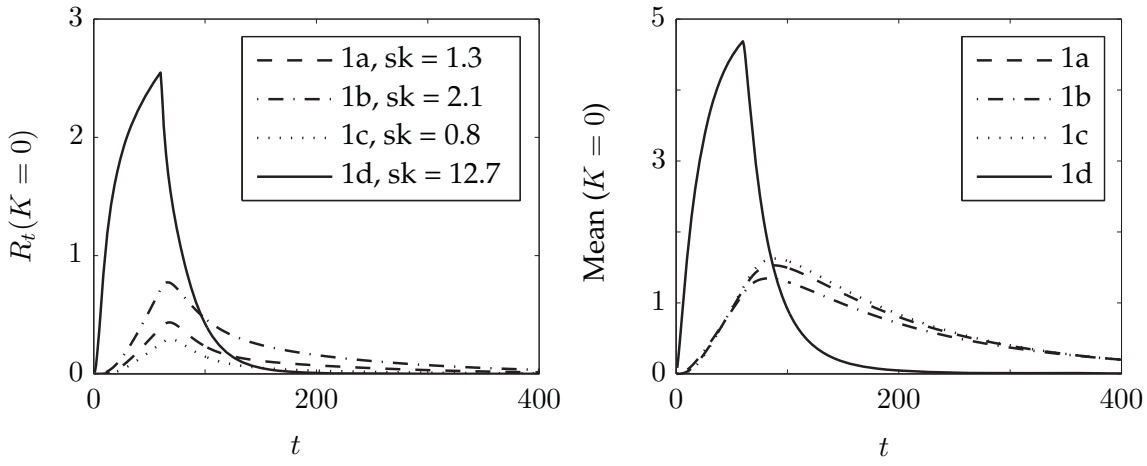


Figure 7.4.: Analysis of the variance (left) and mean (right) for all four cases (1a) through (1d).

resulting MC needs are made even more drastic by the individual skewness values depicted along with the legend. Larger variance and skewness values generally slow down the convergence of MC simulations and explain why the MC needs are largest in cases (1b) and (1d). However, by comparing the mean shapes of system responses in Fig. 7.4 (right), cases (1a) through (1c) show strong similarities, whereas case (1d) is distinctly different due to its strong peakedness. Here, the increased stiffness of the system (mean storage coefficient $\mu_S = 10^{-4}$ instead of 10^{-3}) comes into play, leading to more distinct shapes of system response curves. This shifts the TM needs to higher numbers as compared to smoother cases (1a) through (1c) because more distinct features require more TM to be captured (also see discussion in Sec. 6.4.1). This balances the increased TM and MC needs in case (1d), whereas the patterns in cases (1a) through (1c) tend to be controlled by their different MC needs.

7.3.2. Simple Transport Scenario (cases 2a - 2d)

The results for the four transport-based cases (2a) - (2d) are given in Fig. 7.5. Similar to cases (1a) - (1d) discussed in the previous section, I find the trade-offs moving from low to high MC samples and few to many TM. This behavior can generally observed for all involved cases (2a) trough (2d). Again, all cases achieve there optima with less MC samples ($n_{MC} \leq 500$) then typically used in literature [e.g., Englert et al., 2006; Nowak et al., 2008] when employing typical TM sequence lengths.

By comparing the differences in the individual patterns, two things can be observed: First, cases (2a) and (2b) reveal almost similar features, with test case (2b) tending only slightly towards more MC samples. Both monitoring locations involved in these cases are close to or aligned with the center line of the expected plume path (see Fig. 7.6, right). The monitoring location in case (2a) is closer to the contaminant release and thus observes stronger concentration contrasts over time whereas in case (2b) smoother concentration curves are received.

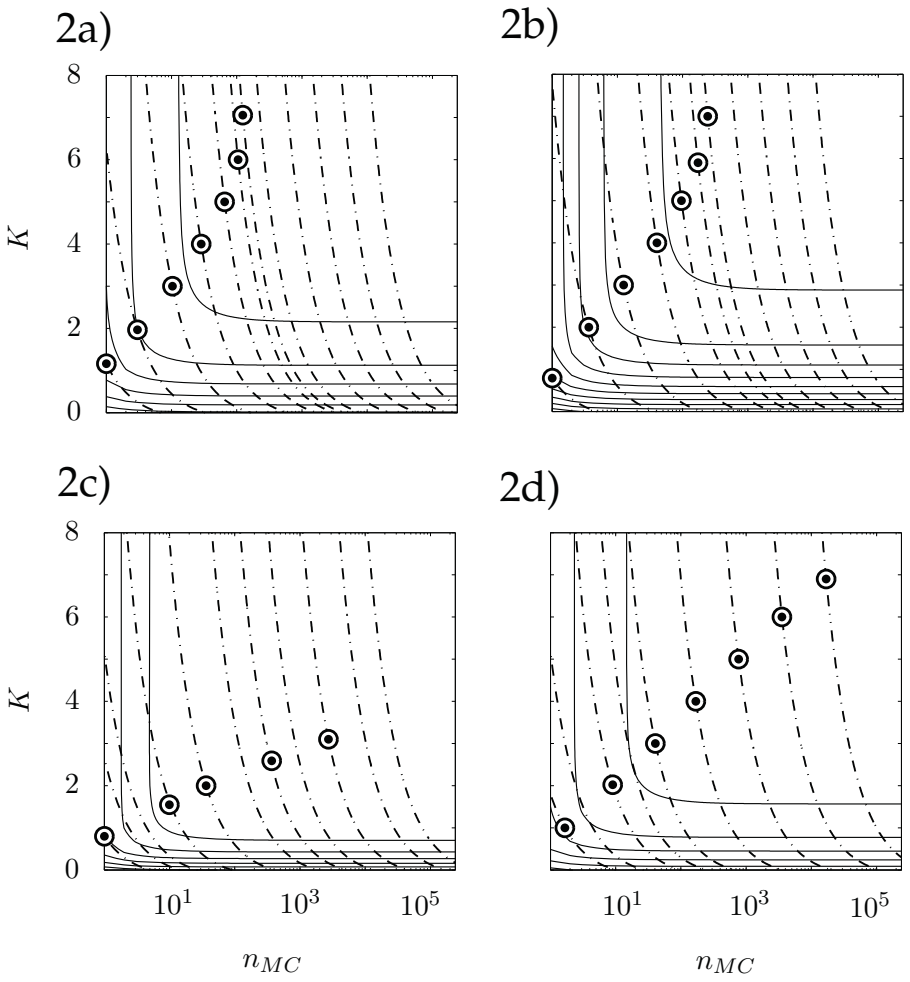


Figure 7.5.: Error-to-budget surfaces of the four transport-based cases (2a) through (2d). Optima for given budgets are depicted by circles. Dashed-dotted contours mark the budget-surface, whereas the solid lines represent the error-surfaces.

This can directly be explained by the fact that concentration gradients being generated by the system forcing and by heterogeneity have been dissipated only to a smaller extent before reaching the monitoring location of case (2a) as compared to case (2b). This triggers a larger need for time resolution for case (2a). This effect is counteracted by the time-dependent variance $R_t(K = 0)$ of $\Omega(K, n_{MC})$ at minimal reduction (zero TM), shown in Fig. 7.6 (left). Case (2a) has a much higher variance and mean skewness value as compared to case (2b). As overall result, this contrast in variance, however, is reflected only to a small extent in their trade-off patterns.

Second, trade-off patterns of cases (2c) and (2d) have a strong tendency towards more MC samples and strongly contrast with those of cases (2a) and (2b). This behavior cannot be explained by higher variance values as shown in Fig. 7.6 (left). However, when comparing their skewness values, a difference in the order of one magnitude can be observed. This slows down the convergence of MC simulations and increases the MC needs for cases (2c)

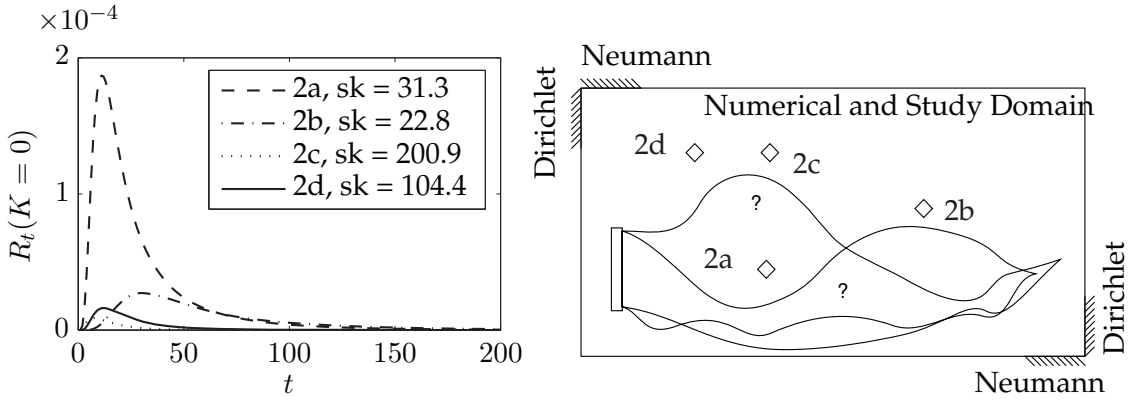


Figure 7.6.: Analysis of the variance for test cases (2a) through (2d) (left) and the spatial arrangement of the measurement locations (right).

and (2d). The extreme skewness values can be explained by the fact that the monitoring locations in both cases (2c) and (2d) flank the plume at a relatively strong lateral offset relative to the expected fringe location. This leads to very distinctive and skewed statistical distribution shapes (Bernoulli distribution, [e.g., Weiss, 2006]), revealing values equal to zero (location missed by plume) and only a few non-zero values (location hit by plume). The different needs for the zeroth TM as found in Sec. 6.5 are not reflected in the trade-off patterns of Fig. 7.5.

7.4. Summary and Conclusions

In this chapter, I showed the importance of reflecting the choice of TM-order against the discretization of other model dimensions, here featuring the number of MC realizations in order to resolve the parameter space. This was defined as ***Step (III)*** of my thesis. To this end, I developed a new method based on jointly considering the numerical error due to insufficiently reduced or discretized model dimensions (e.g., time, physical space, stochastical space) and their associated computational costs of different model dimensions. I obtained cost-to-error surfaces which served to find the optimal pair of computational resources given a certain computational budget. The method was applied to eight different cases from two scenarios from subsurface flow and transport. Each scenario involved four test cases.

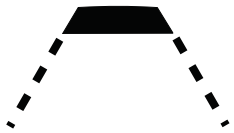
Based on my results, I highlight the following points:

- There always exists an optimal trade-off between the choice of discretizing different model dimensions, e.g. temporal and stochastic dimensions. The trade-off optimally exploits computational resources, leading to the smallest joint error given a certain computational budget.
- The choice of TM order heavily depends on the specific underlying physical problem (e.g., flow, or transport), the geostatistical model, the choice of error metric and the prediction task.

- When reflecting the choice of TM order against the number of MC runs, I found that the trade-off requires less MC runs as compared to typical numbers of MC runs found in literature.

In Sec. 9.2, the same methodology is applied to other model dimensions (spatial discretization and MC resolution), coupled with different and more complex scenarios, predictions and error metrics. The results found will lead to an extended discussion and conclusions towards more generality, hence emphasizing the meaning of optimal resources allocation. In the next chapter, I adapt TM to a novel method for upscaling of high-contrast fractured-porous media (***Step IV***).

8. Adaption to Numerical Upscaling of High Contrast Media



As shown in Sec. 4.2, temporal moments (TM) have useful properties and offer physically meaningful quantities when characterizing conservative transport in heterogeneous media. Examples include apparent velocities, apparent dispersion coefficients or apparent Péclet numbers from observed travel time distributions of solute components [Cirpka and Kitanidis, 2000a]. In other words, TM capture some of the systems behavior on the lower scale and express it on a larger scale defined by the support volume of observation by means of apparent values. This property can be very valuable for upscaling of transport simulations, especially when the underlying system shows large contrasts in material parameters that lead to complex system dynamics, such as in fractured porous media (FPM). In the following chapter, I present a new framework for upscaling high-contrast media. The idea involves TM at a certain step of the scale transition, making the whole procedure swift and accurate. I called this ***Step (IV)*** of my overall approach.

8.1. Background

High-contrast media (e.g., fractured porous media) pose one of the largest unresolved challenges for simulating large hydrogeological systems. Their presence is pervasive on earth, affecting the predictability of many current topics such as carbon sequestration and storage, hydraulic fracturing, groundwater exploitation, etc. Model-based simulations and predictions of such systems can be very helpful but are massively challenged by the complex nature of FPM: Low-permeable rock matrices are intersected by a network of disconnected or interconnected fractures, giving rise to a highly contrasted transport pattern with velocities typically ranging over several orders of magnitude throughout the domain [Berkowitz, 2002]. Dominant advection takes place in the high-permeable pathways, whereas the rock matrix contributes through much slower advection and solute diffusion into

almost-stagnant regions. Overall, this results in anomalous transport behavior with early peaks and late-time tails [Berkowitz, 2002]. Thus, the interaction between fractures and the matrix body are considered to be a key ingredient in the accurate prediction of solute transport in FPM.

Conceptually, there exist two ways to describe FPM: Discrete models or continuum models [Dietrich et al., 2005]. Discrete models allow for the most detailed description including features on the local (microscopic) level and, hence, lead to most accurate predictions of solute transport. However, applications are limited to small scales due to the enormous computational needs that would be required for larger problems, e.g., at the well catchment or reservoir scale. Also, detailed data on fracture geometries are hardly available, triggering issues of parameter and prediction uncertainties. The traditional alternative is to aggregate the system properties according to their role (weather they conduct or store mass) into two superposed continua, leading to so-called dual-continuum models [Barenblatt et al., 1960; Warren and Root, 1963]. The underlying idea is that flow is mainly occurring in the fractures whereas the rock matrix acts as storage term. In the two superimposed continua, steep gradients between solute concentrations causing a local disequilibrium have to be accounted for by so-called transfer functions [Berkowitz, 2002].

Dual-continuum models are frequently applied in the form of double-porosity (mobile-immobile) models, assuming that fluid flow occurs solely in the mobile continuum whereas the immobile zone is assumed to be impermeable [Gerke and van Genuchten, 1993; Simunek et al., 2003]. Transport into the immobile zone relies on diffusion only and can be approximated by first-order transfer terms [Huyakorn et al., 1983], higher-order transfer terms [Bibby, 1981; Dykhuizen, 1990; Zimmerman et al., 1993], multi-continuum models [Lichtner and Kang, 2007] and linear Boltzman transport equations [Benke and Painter, 2003; Painter and Cvetkovic, 2005]. Another more generalized class of transfer terms is considered by multi-rate mass transfer (MRMT) models [Villermaux, 1987; Brusseau et al., 1989; Valocchi, 1990; Sardin et al., 1991; Haggerty and Gorelick, 1995]. Instead of considering individual mass transfer models, MRMT models can simultaneously describe different types of mass transfer occurring on an entire spectrum of time scales, and hence cover a much greater spectrum of transfer features [Haggerty and Gorelick, 1995; Haggerty et al., 2000].

Already Haggerty et al. [2000] and McKenna et al. [2001] showed the applicability of MRTM to FPM and fitted simple mass transfer models to their observations. Hollenbeck et al. [1999] showed how to efficiently estimate mass transfer models from field experiments. In order to compensate for the discrepancy associated with transport prediction in transport upscaling, Willmann et al. [2008] suggested to fit tracer breakthrough curves via an appropriate choice of the memory function. MRMT models, however, make the governing equation non-local in time. This means that storage and re-mobilization does not depend solely on the concentration values in the mobile phase at the current time, but also on their history. This leads to so-called integro-differential equations and imposes its own challenge for the efficiency of numerical solution schemes.

Alternative non-local-in-time formulations have been introduced. For example Carrera et al. [1998] expressed MRMT as convolution of so-called memory functions and concentrations rather than a sum of predefined mass transfer models. Berkowitz and Scher [1995] mod-

eled the movements of solute particles as random walks in space and time, with space and time increments coupled. This method is known as continuous time random walk (CTRW). The different formulations have been compared and reviewed extensively by Carrera et al. [1998]; Dentz and Berkowitz [2003]; Silva et al. [2009]. In the context of FP, CTRW was introduced by Cortis and Birkholzer [2008] as an alternative to Eulerian approaches when modeling FPM. Due to the involved waiting time distribution, CTRW is also non-local in time, and the upscaling from discrete fracture simulations (i.e., inferring the waiting time distribution) is non-trivial [Geiger et al., 2010].

Compared to the great body of studies from the last almost four decades that reproduced observations by MRMT formulations, there are only few studies that obtain MRMT parameters from numerical upscaling of small-scale simulations. Fernàndez-Garcia et al. [2009] used MRMT formulations to compensate for the loss of information when upscaling transport simulations, whereas Li et al. [2011] extended their work to 3D. However, their work did not consider transport in FPM. Only Cortis and Birkholzer [2008] and later Geiger et al. [2010] upscaled numerical transport simulations in FPM using the CTRW formulation.

Most of the applications have done their upscaling globally, i.e., on the macro scale [e.g., Cortis and Birkholzer, 2008; Geiger et al., 2010]. This leads to a solution that depends on a single globally valid set of parameters describing the effective overall transport behavior within the entire domain. Lower-scale behavior and associated processes are not within the entire domain resolved but their effects on the domain scale are parameterized. Another mentality is the upscaling on the meso-scale or block-scale. This bears the advantage that transport behavior can also be described on the lower block-scale and so allows to resolve, e.g., plume location, plume shape, mixing, etc. However, block-scale upscaling is challenging by itself because blocks often require to represent the upscaled parameters by tensors in order to account for anisotropic effects on the block scale [e.g., Wen and Gómez-Hernández, 1996; Fernàndez-Garcia et al., 2009]. Another way out is to work with streamline-oriented grids [e.g., Cirpka et al., 1999b,c]. There, blocks are aligned with streamlines and isopotentials and, hence overcome to some extent the need of anisotropic descriptions on the block scale. Also, they are known to overcome the drawbacks of coarse Eulerian discretizations triggered by numerical dispersion and the related overestimation of mixing [Cirpka et al., 1999b,c]. Another advantage is their conceptual ease, as advection, longitudinal and transverse dispersion terms can clearly be separated by their direction relative to the blocks [Cirpka et al., 1999b].

8.2. Approach

In this framework, I consider transport in fractured-porous media (FPM) following a multi-scale mixed Lagrangian/Eulerian MRMT approach. Numerical tracer experiments are conducted, and all relevant physical small-scale features are account for via Lagrangian simulations (PTRW). For the macro-scale model, a Eulerian dual-porosity (mobile-immobile) approach is applied, accounting for matrix diffusion by MRMT with memory functions. Via the concept of TM, the large-scale MRMT equation is reduced. At the same time, this localizes the governing PDE in time, leading to a highly efficient large-scale model. The TM

approach allows to directly infer the memory function from the arrival time statistics of the local-scale results, yielding a conceptually smart approach for scale transition.

8.3. Methodology

The methodology involves five steps a) through e) as illustrated in Fig. 8.1. In the following, all steps shall be explained in more detail.

a) involves the fine grid generation including fully resolved complex features. From that, the local-scale pressure field can be calculated by solving the stationary 2d pressure equation in confined aquifers (see Eqs. (2.5)-(2.7)).

b) calculates the velocity field from Darcy's law which serves to calculate streamlines. One possible way, is via the so-called streamfunction $\Psi(x, y)$ [e.g., Bear, 1972]:

$$\Psi(x,y) = -\int_0^y v_x(x,y)dy = \int_0^x v_y(x,y)dx, \tag{8.1}$$

from which streamlines (values of $\Psi(x, y) = const$) can be picked. Alternatives include direct simulation of $\Psi(x, y)$ by streamfunction conductivities K_Ψ [e.g., Cirpka et al., 1999c] or trajectory tracking by simple advective particle tracking [e.g., Pollock, 1986].

Isopotentials (values of equal pressure) can be directly taken from the pressure field, sometimes also called isopotential function $\Phi(x, y)$. To guarantee orthogonality between isopotentials and streamlines, pseudopotential conductivities K_Φ can be determined [Matanga, 1988]. For steady-state flow, K_Φ satisfies the same type of equation as the head h, and Eq. (2.7) may therefore be solved by the same numerical methods. Both streamlines and isopotentials finally make up the block grid, featuring flow-aligned blocks.

c) isolates single blocks from the block grid on which small-scale transport simulations are run independently for all blocks. Here, a Lagrangian framework is implemented by particle tracking random walk (PTRW) following Eq. (8.2):

$$\mathbf{X}_p(t+\Delta t) = \mathbf{X}_p(t) + \mathbf{A}(\mathbf{X}_p,t)\Delta t + \mathbf{B}(\mathbf{X}_p,t)\cdot\boldsymbol{\xi}(t)\sqrt{\Delta t}, \tag{8.2}$$

where Δt is the time step, $\mathbf{X}_p(t)$ is the position of a particle at time t. $\mathbf{A}$ represents the advective movement with $\mathbf{A} = \mathbf{v}(\mathbf{X}_p, t) + \nabla \cdot \mathbf{D}(\mathbf{X}_p, t)$ and velocity $\mathbf{v}$, and $\boldsymbol{\xi}(t)$ is a vector of independent, normally distributed random variables with zero mean and unit variance. $\mathbf{B}$ is a displacement matrix that has to fulfill $\mathbf{B}\cdot\mathbf{B}^T = 2\mathbf{D}$ [Salamon et al., 2006]:

$$\mathbf{B} = \begin{pmatrix} \frac{v_x}{|\mathbf{v}|}\sqrt{2(\alpha_l|\mathbf{v}| + D_m)} & -\frac{v_y}{|\mathbf{v}|}\sqrt{2(\alpha_t|\mathbf{v}| + D_m)} \\ \frac{v_y}{|\mathbf{v}|}\sqrt{2(\alpha_l|\mathbf{v}| + D_m)} & \frac{v_x}{|\mathbf{v}|}\sqrt{2(\alpha_t|\mathbf{v}| + D_m)} \end{pmatrix}, \tag{8.3}$$

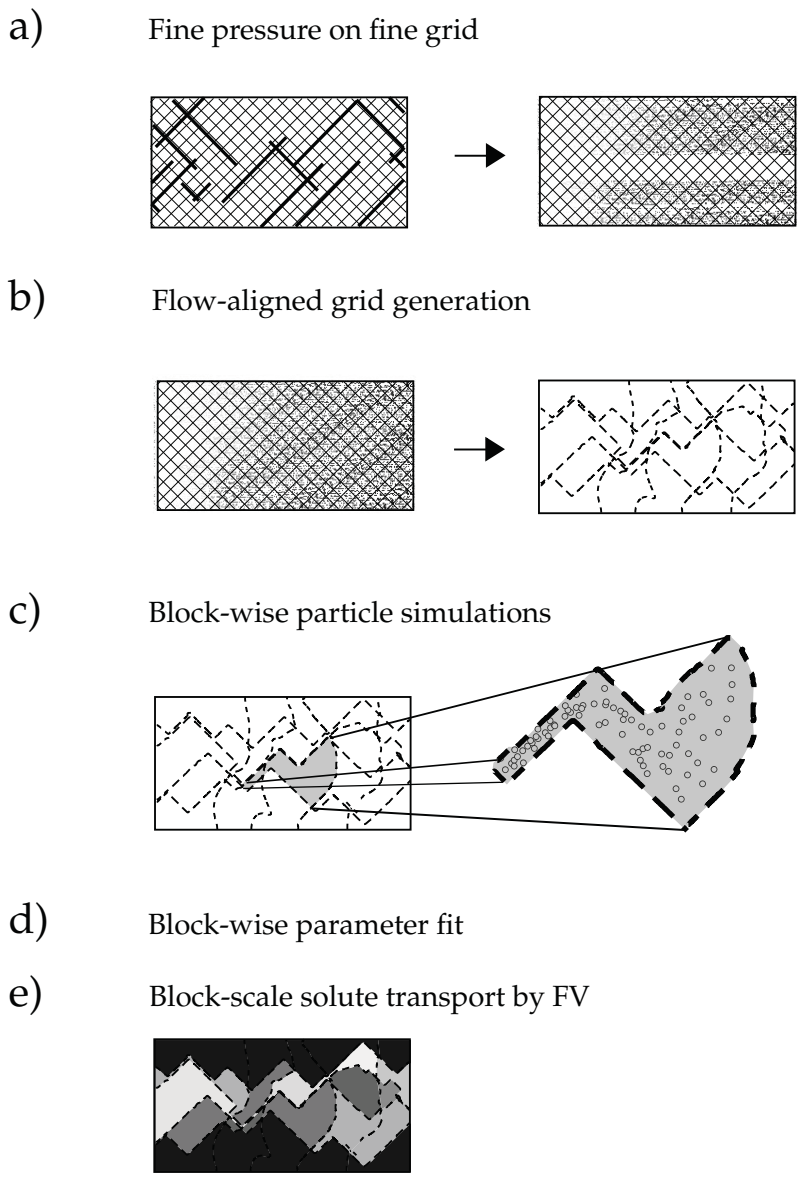


Figure 8.1.: Methodology illustrating the five steps for numerical upscaling of transport in fractured porous media.

with locale-scale longitudinal and transverse dispersivity coefficients $\alpha_{l,\mathrm{loc}}$ and $\alpha_{t,\mathrm{loc}}$, and molecular diffusion coefficient D_m. For more details on PTRW, I refer to [e.g., Salamon et al., 2006].

Particles are injected at the upstream boundary in a flux-weighted manner as an instantaneous injection $m^*_{0,in}$. TM m^*_0, m_1, $m_{2,c}$ are observed at the downstream and the second centralized transverse spatial moment $sm_{2,c}$ at the lower and upper neighbor-

ing outflow boundary. This procedure is illustrated by Fig. 8.2.

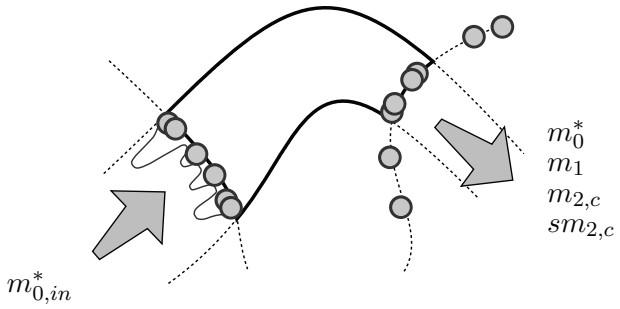


Figure 8.2.: Block-wise derivation of upscaling parameters by means of TM and particle tracking simulations. Particles indicated by red circles are injected at the inflow (left) according to their local velocity (flux-averaged). Averaged TM m_0^*, m_1, $m_{2,c}$ and the second centralized transverse spatial moment $sm_{2,c}$ capture the sub-block behavior (e.g., retardation, transverse spreading and effective dispersion, etc.). Solid lines indicate streamlines and iso-potentials.

d) determines the effective parameters of the upscaled model from the block-wise TM. To this end, a localized MRMT equation is used. The localization simplifies this task tremendously as TM can be directly matched with a 1d analytical solution of the localized MRTM equation:

$$\omega_0 = 1, \tag{8.4}$$

$$\beta = (m_1 - \frac{x}{v_{\text{eff}}})/(\omega_o \frac{x}{v_{\text{eff}}}), \tag{8.5}$$

$$\begin{aligned} \omega_1 = (m_2 - \frac{2x(D_e + \alpha_{l,\text{eff}}\, v_{\text{eff}})}{v_{\text{eff}}^3} - \frac{x^2}{v_{\text{eff}}^2}(1 + \beta\omega_0) \\ - \frac{x\beta(D_e + \alpha_{l,\text{eff}}\, v_{\text{eff}})\omega_0}{v_{\text{eff}}^3})/(\frac{2x\beta}{v_{\text{eff}}}), \end{aligned} \tag{8.6}$$

whereas v_{eff} and $D_e = D_m\, n_e$ are the absolute effective velocity and effective diffusion coefficient, respectively, with porosity n_e and x is the integrated path length along block centers. β is the capacity coefficient from Eq. (2.24), and ω_0 and ω_1 are the zeroth and first moment of the memory function [Luo et al., 2008]. All values β, ω_0 and ω_1 are spatially distributed, i.e. they differ from block to block. Absolute effective velocities v_{eff} are calculated from balancing the volume streamtube water fluxes, whereas the effective diffusion coefficient D_e is directly taken from the local scale. Note that Eqs. (8.4)-(8.6) serve only to make predictions up to second-order TM. This is, however, not a limitation of the methodology. For higher-order TM predictions the relevant equations for ω need to be derived.

It remains to be discussed how the effective dispersions coefficients $\alpha_{t,\text{eff}}$ and $\alpha_{l,\text{eff}}$ are determined. As the system of Eqs. (8.4)-(8.6) is underdetermined (the second TM has to determine both $\alpha_{l,\text{eff}}$ and ω_1) an independent assumption for $\alpha_{l,\text{eff}}$ is needed. For $\alpha_{l,\text{eff}}$, simply the local-scale value $\alpha_{l,\text{loc}}$ is taken. As for $\alpha_{t,\text{eff}}$ different definitions covering different scales are possible. As general suggestion, I opt to determine $\alpha_{t,\text{eff}}$ from the rate of increase in the second centralized transverse spatial moment. In the following, this value is referred to as $\alpha_{t,\text{mac}}$. The rate of increase is given as [e.g., Freyberg, 1986]

$$D_{mac} = \frac{1}{2}\frac{\Delta sm_{2,c}}{\Delta t}, \tag{8.7}$$

with macrodispersion coefficient D_{mac} and second centralized spatial transverse moment $sm_{2,c}$. By substituting D_{mac} with $D_e + \alpha_{l,\text{eff}}\, v_{\text{eff}}$ (for the 1d case) and Δt with $v_{\text{eff}}/\Delta x$, Eq. (8.7) can be solved for $\alpha_{t,\text{mac}}$:

$$\alpha_{t,\text{mac}} = \frac{1}{2}\frac{\Delta sm_{2,c}}{\Delta x} - \frac{D_e}{v_{\text{eff}}}, \tag{8.8}$$

$sm_{2,c}$ is determined from the spatial particle statistics as illustrated in Fig. 8.2. Please note that, in the later application, other definitions of $\alpha_{t,\text{eff}}$ on different scales are implemented and tested.

e) finally, employs effective parameters in a localized block-scale MRMT model on the coarse scale. The localized MRMT model is obtained from reducing Eqs. (2.24)-(2.25) by means of an integral transformation [Luo et al., 2008] to their moment generating equations:

$$v_{\text{eff}} \cdot \nabla m_k^* - \nabla \cdot (\mathbf{D}\nabla m_k^*) = k m_{k-1}^* + \beta k \sum_{l=0}^{k-1} \binom{k-1}{l} \omega_{k-1-l} m_l^* \tag{8.9}$$

with D_e according to Eq. (2.17) and boundary conditions

$$\mathbf{n} \cdot (v_{\text{eff}}\, m_k^* - \mathbf{D}\nabla m_k^*) = \frac{m_{Neu}}{Q}\mathbf{n} \cdot v_{\text{eff}}\delta_{k0} \;\text{ on }\; \Gamma_{Neu} \tag{8.10}$$

$$\mathbf{n} \cdot (\mathbf{D}\nabla m_k^*) = 0 \;\text{ on }\; \Gamma/\Gamma_{Neu}. \tag{8.11}$$

Here, Γ_{Neu} is the inflow boundary and Γ/Γ_{Neu} all other parts of the boundary. As matter of fact, the zeroth TM is free of sources and sinks, whereas higher-order TM are subject to sources and sinks controlled by the lower-order TM m_{k-1}^* and the memory function moments ω_{k-1-l}. This recursive coupling is known from the moment generating equations in Sec. 4.3 and has been extensively discussed in Chap. 5. Eqs. (8.9)-(8.11) are solved numerically by a Finite Volume (FV) approximation on the streamline-oriented grid aligned by the blocks, representing the solution on the block-scale.

Fig. 8.3 illustrates the different fluxes occurring in the FV implementation. Advective fluxes occur in the principal direction only. The dispersion tensor is a diagonal matrix. Hence, transverse dispersion occurs in transverse and longitudinal dispersion in principal direction only. These two effects not only simplify the solution of the advection-

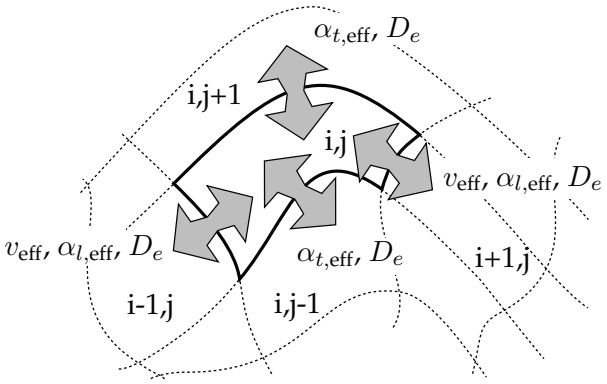


Figure 8.3.: FV scheme to evaluate the Block-scale solute transport, exemplary for one single block. Due to the streamline-oriented grid, the FV solution simplifies. Advective fluxes and longitudinal dispersion occur in the principal direction and transverse dispersion only in transverse direction only.

dispersion equation but also reduce numerical-dispersion. Eq. (8.12) gives the principal FV equation, exemplary for one block i, j:

$$\begin{aligned} & J_{\text{disp}}(i, j-1 \to i, j) + J_{\text{diff}}(i, j-1 \to i, j) \\ & + J_{\text{disp}}(i, j+1 \to i, j) + J_{\text{diff}}(i, j+1 \to i, j) \\ & + J_{\text{disp}}(i-1, j \to i, j) + J_{\text{diff}}(i-1, j \to i, j) \\ & + J_{\text{disp}}(i+1, j \to i, j) + J_{\text{diff}}(i+1, j \to i, j) \\ & + J_{\text{adv}}(i-1, j \to i, j) - J_{\text{adv}}(i+1, j \to i, j) \quad = \quad 0. \end{aligned} \tag{8.12}$$

8.4. Application

In order to show and discuss its applicability, I apply the new method to an illustrative example from solute transport in fractured porous media. The example is based on scenario (5) described in Sec. 3.5. Tabs. 8.1-8.2 summarize the relevant physical and geometrical parameters for the local and block-scale model.

Throughout the study, I vary the block resolution in order investigate its effect on the prediction accuracy. Overall, eight different block resolutions are considered, ranging from 304 to 9 blocks summarized in Tab. 8.3:

The goal of this study is to properly predict the spatially distributed mass m_0^* (zeroth TM), the arrival time m_1 (first TM), and the effective dispersion $m_{2,c}$ (second TM). A fine-scale PTRW simulation is conducted from which reference predictions for all three quantities are calculated. To make the scale of both the block-wise and the reference solution conforming for fair comparison, the reference solution is transfered to the block scale by block-wise TM evaluation. This allows to quantitatively compare the accuracy by scalar measures that

Fine Scale			
Domain size	$[L_1, L_2]$	[m]	[6, 3]
Grid spacing	$[n, n]$	[-]	[3675, 395]
FE Parameters			
Hydraulic Gradient	Δh	[-]	0.003
Fracture Transmissivity 1	T_F	[m/s]	1×10^{-3}
Fracture Transmissivity 2	T_F	[m/s]	1×10^{-5}
Matrix Transmissivity	T_M	[m/s]	1×10^{-9}
PTRW Parameters			
Porosity	Θ	[-]	0.1
Trans. Disp.	$\alpha_{t,\text{loc}}$	[m]	1×10^{-3}
Long. Disp.	$\alpha_{l,\text{loc}}$	[m]	5×10^{-3}
Diffusion	D_m	[m^2/s]	1×10^{-9}
Particle resolution	n_P	[-]	5000

Table 8.1.: Physical and geometrical parameters employed in the fine-scale model.

represent the scale of interest defined by the block scale. To this end, the relative volume-averaged global L^2-norm of mass discrepancies RAL_0 is calculated:

$$RAL_0 = \frac{\sum_{i=1}^{n} V_i (m_{0,i}^{*} - m_{0,i}^{*(Ref)})^2}{\sum_{i=1}^{n} V_i (m_{0,i}^{*(Ref)})^2}, \tag{8.13}$$

with total number of blocks n on the respective scale, block volumes V_i and block index i. As for the arrival time and the effective dispersion, I derive relative flux-averaged L^2-norms RAL_1 and RAL_2 at the outflow boundary according to Eqs. (8.14)-(8.15).

$$RAL_1 = \frac{\sum_{j=1}^{n_t} q_j (m_{1,j}^{*} - m_{1,j}^{*(Ref)})^2}{\sum_{j=1}^{n_t} q_j (m_{1,j}^{*(Ref)})^2}, \tag{8.14}$$

$$RAL_2 = \frac{\sum_{j=1}^{n_t} q_j (m_{2,j}^{*} - m_{2,j}^{*(Ref)})^2}{\sum_{j=1}^{n_t} q_j (m_{2,j}^{*(Ref)})^2}, \tag{8.15}$$

with total number of outflow-blocks n_t, respective flux q_j and block index j. Note, that mass m_0^* is assessed on the entire domain, whereas arrival time m_1 and effective dispersion $m_{2,c}$ are assessed only at the outflow boundary. This choice is aligned with typical prediction goals and quantities of theoretical interest found in literature.

Block-based RAL values are only one possible error measure. Alternatively, the error could have been determined on the local scale by mapping the block-wise values onto the reference scale, and then performing a finely resolved integration of squared differences over the entire domain.

	Block Scale		
Domain size	$[L_1, L_2]$	[m]	[6, 3]
Block resolution	$[n_S, n_T]$	[-]	*
	FV Parameters		
Trans. Disp.	$\alpha_{t,\text{eff}}$	[m]	*
Long. Disp.	$\alpha_{l,\text{eff}}$	[m]	5×10^{-3}
Diffusion	D_e	[m^2/s]	1×10^{-9}
Velocity	v_{eff}	[m/s]	tbd
-	β	[-]	tbd
-	ω_0	[-]	tbd
-	ω_1	[-]	tbd

Table 8.2.: Grid and simulation parameters used for the block-scale model. Parameters marked by * will be altered throughout the test case, whereas tbd means that the respective parameters need to be determined during the upscaling procedure.

		Resolution							
# Blocks	n	304	144	99	80	49	36	25	9
# Tubes	n_t	16	12	9	8	7	6	5	3
# Sections	n_s	19	12	11	10	7	6	5	3

Table 8.3.: Different resolutions on block-scale are considered throughout the study ranging from 304 to 9 blocks.

To test the choice of $\alpha_{t,\text{mac}}$ (denoted as "mac" in the following) as effective transverse dispersion coefficient other definitions covering other scales are implemented and tested. They include local-scale $\alpha_{t,\text{loc}}$ (denoted as "loc" in the following) and block-scale $\alpha_{t,\text{blo}}$ (defined as proportion of particles arriving at the upper or lower neighboring block outflow boundary multiplied by its specific discharge). The later is denoted as "blo" in the following. In the first step, however, only $\alpha_{t,\text{mac}}$ is considered and the block resolution is set to 144 blocks.

8.5. Results and Discussion

This section compares and discusses the results from the study described above. Sec. 8.5.1 shows the prediction of mass, arrival time and effective dispersion for the mac implementation while using 144 blocks. Sec. 8.5.2 varies the block resolution and analyzes its impact on the prediction accuracy, whereas Sec. 8.5.3 investigates the effect and the usefulness of other $\alpha_{t,\text{eff}}$ implementations, and Sec. 8.5.4 compares the MRMT-based predictions to a purely Fickian parameterization of dispersion. The later leads to the Fickian case based on the traditional advection-dispersion equation (ADE).

8.5.1. Prediction of Zeroth, First and Second Temporal Moment

Fig. 8.4 gives (in clock-wise direction) the fracture distribution, the pressure field, the velocity field and the coarse grid based on streamlines and isopotentials. Obviously, fast conduits connected to the left boundary take most of the water (indicated by high velocities greater than 1×10^{-6} m/s). The overall pressure pattern is shaped by the orientation of the few fast channels. Disconnected channel-networks impose sharp local pressure gradients when the gaps between fast conduits become small. The same can be observed for the coarse flow-aligned grid (based on streamlines and isopotentials). It follows the orientation of the few fast channels and reveals extremely thin and bent blocks.

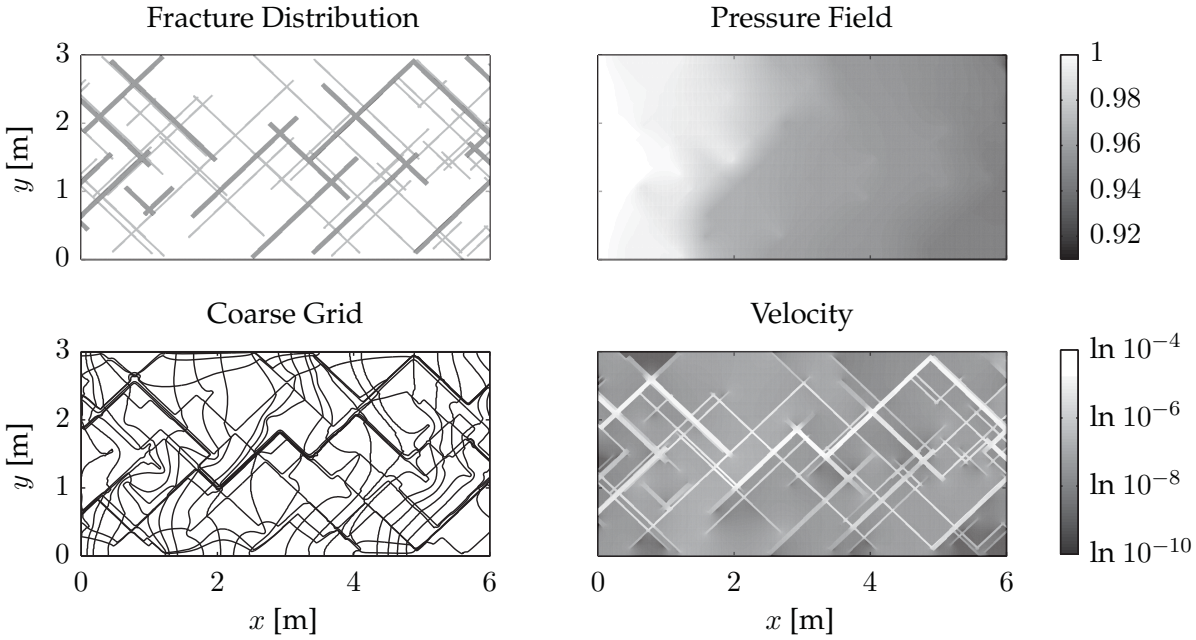


Figure 8.4.: In clock-wise direction: Synthetic fracture distribution (top left), pressure field (top right), velocity field (bottom right) and block grid based on flow-aligned blocks (bottom right).

Predicting mass distribution (zeroth TM)

The predictions of the spatial mass distribution for both the block-scale (left) and the reference case (right) are given in Fig. 8.5. Quantitatively, the shapes of both patterns are in good agreement. By calculating RAL_0, I find a deviation of merely 6 % from the reference solution. The overall good performance can be explained by two things. (1) The stream-line oriented grid significantly reduces numerical dispersion and location uncertainty. (2) Matching the second centralized spatial moment on the block scale measures the true spatial spreading at the local-scale and, hence, leads to a proper description of lateral mixing also on the block scale. Note that, in the case of predicting only the zeroth TM m_0^*, MRMT is obsolete as there

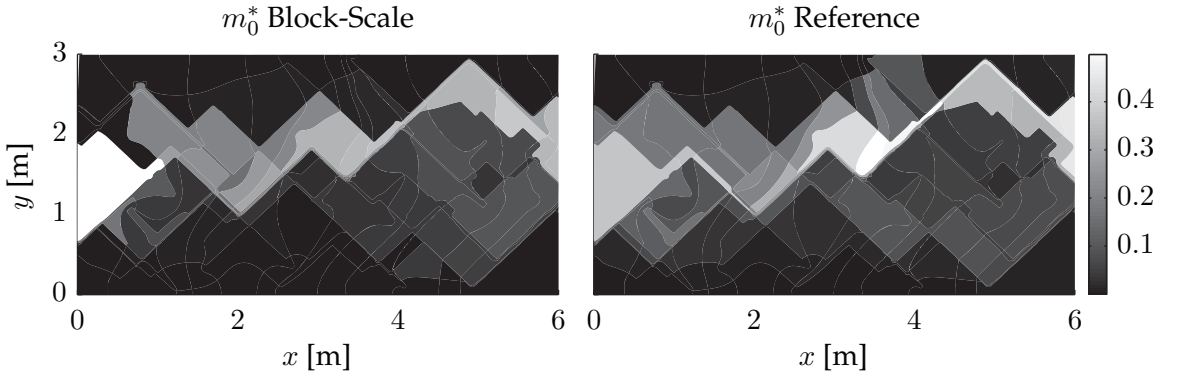


Figure 8.5.: Prediction of m_0^* for both the block-scale (left) and the reference (right) case.

is no dependency of MRMT parameters on the zeroth TM and because the governing equation for the zeroth TM contains no MRMT influence (see discussion at the end of Sec. 8.3).

Predicting arrival time (first TM)

The prediction of arrival times for both the block-scale (left) and the reference case (right) is given in Fig. 8.6. Again, the overall pattern is preserved. However, a visible mismatch of

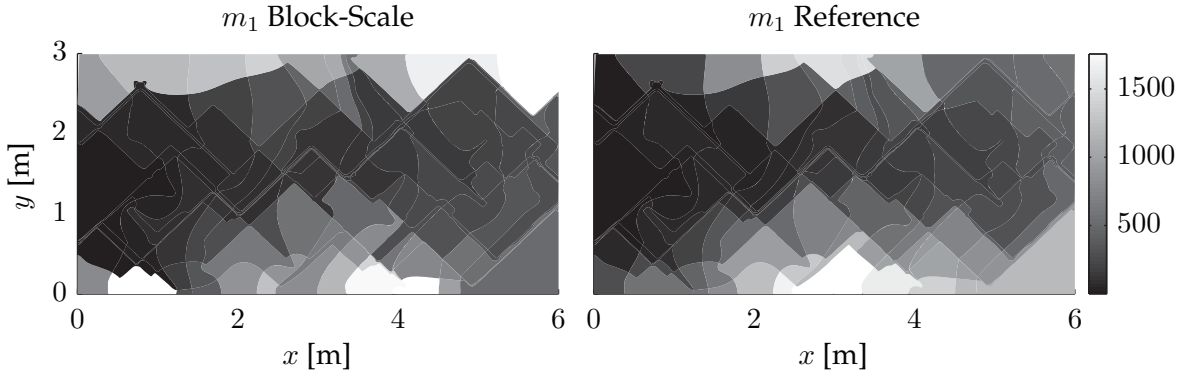


Figure 8.6.: Prediction of m_1 for both the block-scale (left) and the reference (right) case.

magnitudes at the northern and southern boundaries can be observed. Here, RAL_1 reveals a deviation of around 30 % from the reference solution. This worse performance as compared to the mass prediction can be explained by the fact that $\alpha_{t,\text{mac}}$ is now challenged by two competing requirements: (1) to quantify properly the degree of transverse spreading of mass and (2) to properly quantify the transverse mixing of different arrival times between neighboring streamlines. This makes the present choice of $\alpha_{t,\text{mac}}$ (designed for good transverse spreading without considering its effects on mixing) less appropriate.

Another issue is that the non-Fickian behavior of arrival times is compensated to some ex-

tent by introducing the β parameter from the MRMT model. However, β is determined by isolating single blocks whereas in the FV approximation the effects of β interact between neighboring streamlines. Since this is accounted for in the determination of β, errors might be invoked.

Predicting effective dispersion (second TM)

The prediction of effective dispersion for both the block-scale and the reference case is given in Fig. 8.7. The qualitative visual assessment reveals a good match with the overall disper-

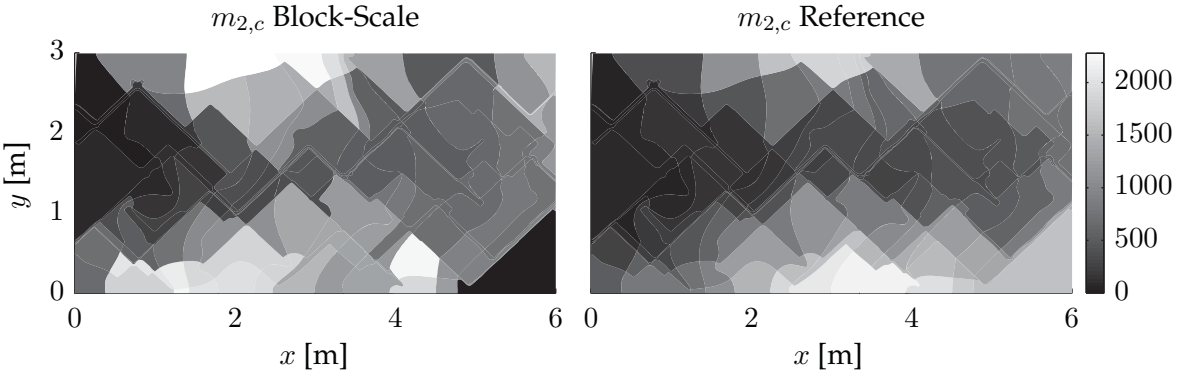


Figure 8.7.: Prediction of $m_{2,c}$ for both the block-scale (left) and the reference (right) case.

sion pattern, yet I find stronger deviations as compared to the arrival time case. By calculating RAL_2, a deviation of 40 % from the reference solution can be found. Again, $\alpha_{t,\text{mac}}$ has to account for both good transverse spreading and proper mixing. Also, the effects of β and, here additionally, ω_1 onto neighboring streamlines within the FV model (similar to the prediction of m_1) might invoke errors.

8.5.2. Comparing Different Block Resolutions

Fig. 8.8 (left) shows RAL-profiles for different block resolutions ranging from 9 to 304 blocks. I consider again the case of $\alpha_{t,\text{mac}}$. Starting with 304 blocks (right limit of Fig. 8.8), the block resolution is steadily reduced, finally arriving at a 9-block approximation (left limit of Fig. 8.8). For all three predictions m_0^*, m_1 and $m_{2,c}$ an increase of the error (from right to left) can be observed peaking somewhere between 1 and 50 blocks, and eventually dropping down to zero when approaching the left limit (single block domain).

To explain the above observed behavior I arrive at the following conclusion: Over the spectrum of block resolutions the error is driven by two things: (1) the amount of local-scale features to be properly transfered to the upper scale via MRMT parameters (β, ω_1) and (2) their interaction between neighboring streamlines in the upscaled model via block interfaces. RAL-values can be regarded as their combined product. Once either of both is zero,

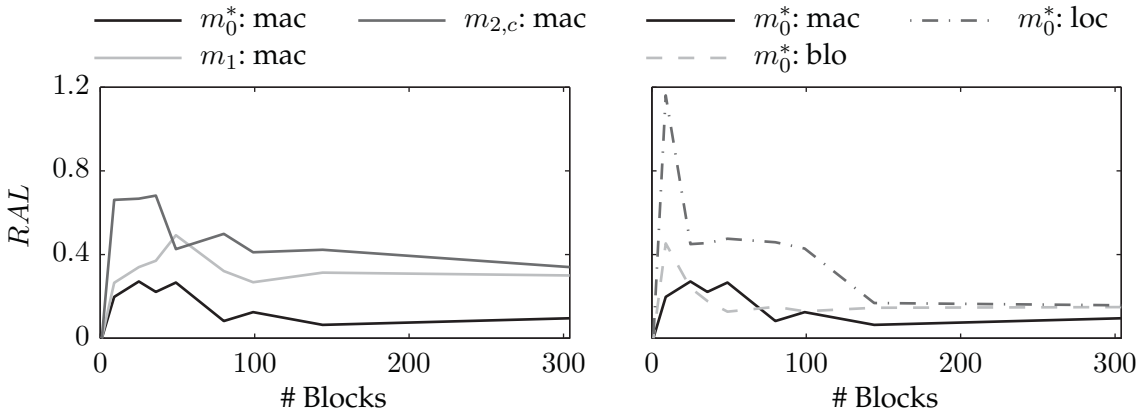


Figure 8.8.: Comparison of RAL-profiles for different temporal complexities m_0^*, m_1 and $m_{2,c}$ based on $\alpha_{t,\text{mac}}$ (left) and comparison of different $\alpha_{t,\text{eff}}$ definitions featuring m_0^* (right).

the product and hence RAL becomes zero. In other words: For high block numbers, the blocks itself take over role of resolving the anomalous transport behavior, making additional parametrization by MRMT unnecessary. With no remaining need for upscaling, the MRMT model becomes obsolete and the RAL-profiles converge to zero. For the opposite case with small block numbers, the number of block interfaces and the resolution of velocity contrasts decreases, and MRMT has to parameterize more and more sub-block contrasts. In case of a single-block domain, I arrive at zero interaction between blocks and their involved parameters. All sub-block features are captured by a single block which is defined to meet the prediction goals precisely at the scale of interest defined by the single (macroscopic) block. Thus, RAL become zero. In between these antipodes, the RAL-profiles reach their maximum, e.g. as observed in Fig. 8.8 (left).

By comparing the position of the error peaks in the number of used blocks, I find the RAL-profiles stretching towards higher resolution when increasing the order of temporal complexity (m_0^*, m_1 and $m_{2,c}$). Obviously, predicting higher order TM imposes more errors and requires higher resolutions to compensate for these errors. This is due to the fact that (1) higher-order TM balances on block-scale become increasingly erroneous, (2) more unconsidered effects in MRMT parameters β and ω_1 might invoke additional errors, and (3) upscaling-based errors propagate both spatially and throughout the recursive coupled TM orders.

8.5.3. Comparing Different Transversal Dispersion Coefficients

In the section, I analyze the sensitivity of RAL towards other definitions of $\alpha_{t,\text{eff}}$ (described in Sec. 8.3). To this end, I also employ local-scale $\alpha_{t,\text{loc}}$ and block-scale $\alpha_{t,\text{blo}}$. The RAL-profiles are shwon in Fig. 8.8 (right) and Fig. 8.9.

By comparing the results for the mass m_0^* (see Fig. 8.8, right), little difference can be iden-

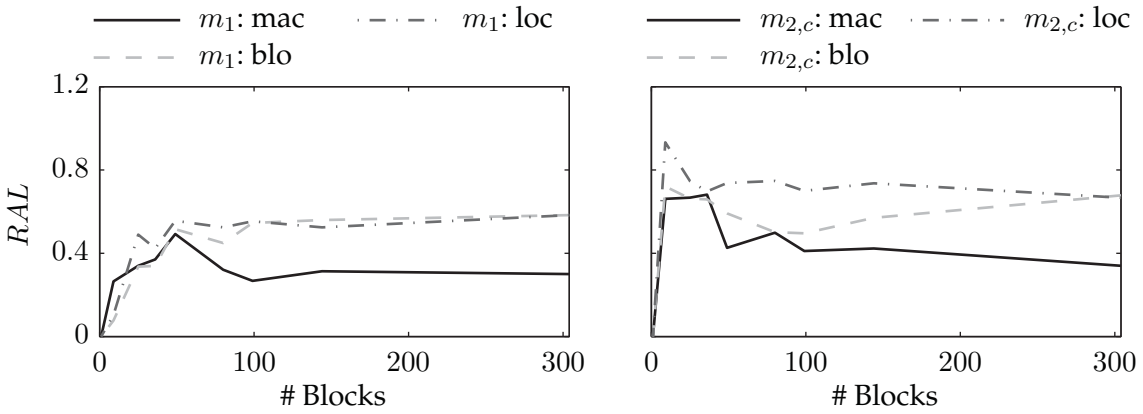


Figure 8.9.: Comparison of different $\alpha_{t,\text{eff}}$ definitions featuring m_1 (left) and $m_{2,c}$ (right).

tified for all definitions when considering more than around 150 blocks. For less than 150 blocks, however, local-scale $\alpha_{t,\text{loc}}$ (dashed line) starts deviating considerably. Here, $\alpha_{t,\text{loc}}$ becomes too small to guarantee sufficient lateral spreading. This is caused by the eliminated velocity contrasts within the homogenized blocks, and by the fact that the boosting effects of streamline focusing on transverse mixing [e.g., Werth et al., 2006] are partially lost. Only the components of streamline-focusing on block-scale are preserved. Consequently, the local-scale $\alpha_{t,\text{loc}}$ is too small to compensate for the lost components.

This can be explained by Fig. 8.10 (left) showing the median $\alpha_{t,\text{eff}}$ values for all three definitions $\alpha_{t,\text{loc}}$, $\alpha_{t,\text{blo}}$ and $\alpha_{t,\text{mac}}$, depending on the block resolution. Here $\alpha_{t,\text{loc}}$ is much smaller (1-2 orders of magnitude) than the other two definitions. $\alpha_{t,\text{mac}}$ yields the highest median value almost constant over the spectrum of resolutions. This is because $\alpha_{t,\text{mac}}$ represents by definition the proper value for transverse mixing including effects of streamline focusing on the reference scale. This type of effective block-wise transverse mixing coefficient has been shown to be an intrinsic medium property and scale-independent of travel time and, hence, also independent of the block volume used as sub-domain for measuring this effect.

The block-scale $\alpha_{t,\text{blo}}$ represents the idea of properly capturing the mass balance on the block scale and. This forces the block-scale mass fluxes on the FV-scheme and so allows for reasonable RAL-values when predicting m_0^*. Regarding the median $\alpha_{t,\text{blo}}$ in Fig. 8.10 (left), I find that $\alpha_{t,\text{blo}}$ increases with decreasing block resolution and ranks in between the magnitudes of $\alpha_{t,\text{loc}}$ and $\alpha_{t,\text{mac}}$.

As for arrival times m_1 (see Fig. 8.9, left), I find that the RAL-profiles for the different definitions of $\alpha_{t,\text{eff}}$ differ more strongly for higher resolutions and steadily converge for lower resolutions. Among all definitions, $\alpha_{t,\text{mac}}$ performs best over almost the entire spectrum of block resolutions. The local-scale $\alpha_{t,\text{loc}}$ underestimates transverse mixing of arrival times and, hence, cannot match the reference m_1. For less then around 50 blocks, however, the RAL profiles for all three $\alpha_{t,\text{eff}}$ definitions converge before dropping down to zero.

This is in line with the conclusions drawn in Sec. 8.5.2. For few blocks (here less than 50) the MRMT model (here β) parameterizes most of the sub-block contrasts and the interactions

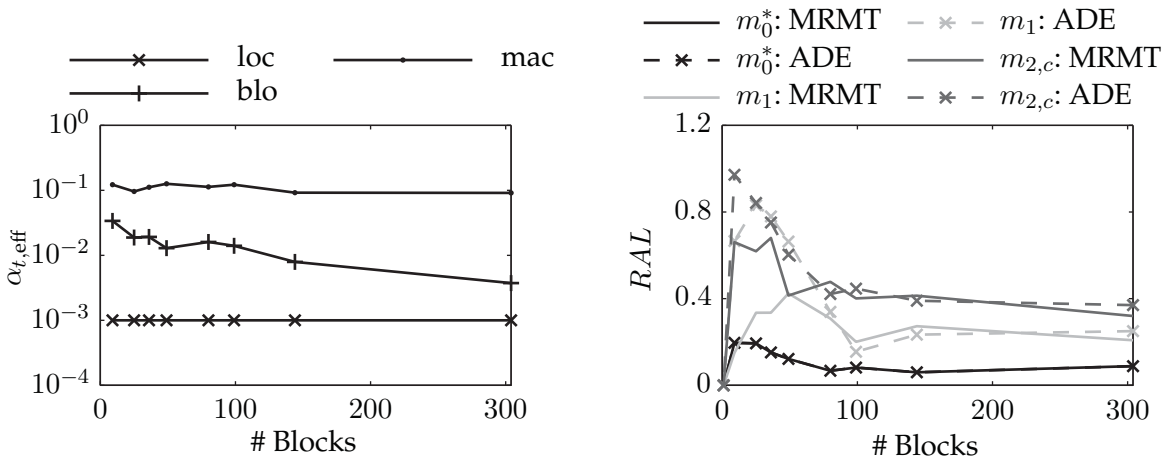


Figure 8.10.: Comparison of median $\alpha_{t,\text{eff}}$ values for different block resolutions (left) and the effect of a Fickian parameterization of dispersion (right).

between neighboring streamlines in the upscaled model via block interfaces becomes little. This makes RAL less sensitive to the choice of $\alpha_{t,\text{eff}}$.

Likewise for effective dispersion $m_{2,c}$ (see Fig. 8.9, right), the RAL-profiles for the different definitions of $\alpha_{t,\text{eff}}$ differ for higher resolutions and steadily converge for lower resolutions. Yet, this happens at greater RAL-values as compared to the m_1 case. Also, $\alpha_{t,\text{mac}}$ performs best over almost the entire spectrum of block resolutions. Here, the local-scale $\alpha_{t,\text{loc}}$ cannot match the reference $m_{2,c}$ because the underestimated transverse mixing of m_1 propagates through the prediction of $m_{2,c}$.

It should be mentioned that the median values of $\alpha_{t,\text{mac}}$ and $\alpha_{t,\text{blo}}$ might be an incomplete measure as they do not reflect the entire distribution. In my case, however, this choice is convenient because the local-scale $\alpha_{t,\text{loc}}$ is spatially constant.

8.5.4. Multi-Rate Mass Transfer Model vs. Advection Dispersion Equation

Fig. 8.10 (right) compares the RAL-profiles for all three predictions against the case where no MRMT model is considered to parameterize the lost sub-block features in upscaling. This means, in fact, to perform traditional transport upscaling with a purely Fickian parameterization of dispersion, and the claimed ADE. I choose $\alpha_{t,\text{mac}}$ for all MRMT-based and ADE-based predictions m_0^*, m_1 and $m_{2,c}$.

As a matter of fact, the case of predicting mass is insensitive to this change (the governing equation for m_0^* does not depend on any MRMT parameter even in the MRMT-based model). For both other cases m_1 and $m_{2,c}$, I observe that for resolutions higher than around 100 blocks, the ADE-based case has similar RAL-profiles as compared to the MRMT-based case. This is different for smaller resolutions. Here, RAL increases at a higher rate as compared to the MRMT-based case. This clearly indicates that the accuracy gain in upscaling by MRTM

takes full effect for resolutions smaller than around 100 blocks (in this study). This is because for more than 100 blocks most of the velocity contrasts are resolved by the blocks itself and there are no sub-block features to be parameterized by the MRMT model. Hence, there is no remaining need to use a MRMT-based model.

8.6. Summary and Conclusions

In this chapter, I introduced a new method for numerical upscaling of transport in high-contrast media. The idea is partially based on the useful properties of TM, and uses flow-aligned blocks for the upscaled model. The novelties include:

1. First-time application of flow-aligned blocks to FPM.
2. First-time application of TM to FPM.
3. First-time analysis of multi-scale arrival time statistics and effective dispersion in FPM via TM.
4. Efficient multi-scale modeling of FPM via MRMT models, combined with model reduction of both fine and large-scale models.
5. Efficient determination of memory function TM from fine-scale TM.

The method has been applied and tested in a scenario from contaminant solute transport in fractured-porous media. I found the following conclusions most relevant:

1. The prediction accuracy depends on the number of blocks representing the coarse-scale resolution and the complexity of the prediction goal (mass distribution of steady state, arrival time statistics or effective dispersion statistics). Generally, higher resolutions can yield better accuracy, but higher resolutions are also prone to errors because the nature of determining MRMT parameters does not account for their spatial effects within the block-scale model.
2. Different definitions of dispersion coefficients have been tested. Generally, the definition based on the centralized second transverse spatial moment yields best results. Local-scale and block-scale definition by mass balances underestimated transversal mixing. They lead to a worse performance.
3. If a certain block resolution is exceeded the domain contrasts are fully resolved by the block itself and the parameterization by MRMT becomes obsolete. This is similar to the traditional ADE-based solution.

After having carried the four major steps of my overall approach, I emphasize in Chap. 9 the universality of my developed methods. I show how the Pre-posterior Data Impact Assessor helps to significantly improve the optimal design of experiments (Sec. 9.1), and how optimal resource allocation can find the trade-off between other discretized or reduced model dimensions (Sec. 9.2).

9. Other Applications of the Developed Tools

In this work, I developed new tools in order to answer research questions associated with the analysis and assessment of model reduction techniques in time. These tools are, however, not limited to the applications shown in Chap. 6 and Chap. 7. Instead, they are generic tools that can be employed in a variety of different applications. In this chapter, I show how the Pre-posterior Data Impact Assessor (PreDIA) can be employed in optimal design of experiments (Sec. 9.1), and how Optimal Resource Allocation (ORA) can find the trade-off between the discretization of other model dimensions (Sec. 9.2).

9.1. Optimal Design of Experiments

PreDIA, as introduced in Chap. 6, has been described and employed as a method to identify response curves among a set of physically plausible random response curves $r(t)$ based on a set of temporal characteristics, e.g., TM or LC. In the expected sense, namely if the identification is repeated for all potential characteristics on all possible response curves $r(t)$, PreDIA can access the explanatory power of a temporal characteristic. In more general words, PreDIA finds the explanatory power (in the sense of statistical inferrence) of a some potential data y (temporal characteristics) towards a prediction z (response curve $r(t)$). This opens the door to optimal design of experiments.

9.1.1. Design Problem and Bayesian Analysis

In optimal design (OD), the problem statement typically seeks confidence associated with a certain hydro(geo)logical model prediction z, which shall be improved by collecting new and then conditioning on field data y. Yet, sampling and investigation campaigns are restricted by limited budgets, or by physical constraints, and therefore should be addressed in a rational and optimal way. This leads to the optimal design problem of finding the best sampling design or investigation strategy for the given problem at hand, i.e., the one that maximizes some kind of utility function ϕ (in Chap. 6 represented by the Conditional Standard Deviation $CStD$) under the given constraints. The impact or utility of a design is defined as its individual capability to reduce uncertainty associated with the prediction goal z, or to maximize some related measure of data utility [e.g., Federov and Hackl, 1997; Uciński, 2005; Müller, 2007].

The most important key ingredients to OD are adequate statistical or stochastic methodologies that properly transfer the uncertainty in model structure and parameters to model predictions z, while taking into account the impact of noisy measured and yet unmeasured (planned) data. Here, PreDIA comes into play, overcoming the limitations of (quasi-)linear statistical inference tasks commonly employed in OD tasks (see Fig. 6.1). Examples for linearizing methods include first-order second-moment methods (FOSM) [e.g., Kunstmann et al., 2002; Cirpka et al., 2004] or the ensemble Kalman filter (EnKF) [e.g., Evensen, 2007; Nowak, 2009; Schöniger et al., 2012].

The final step of OD is to find the best set of decision variables d (that specify, e.g., the number, locations, types, and experimental conditions for measurements which shall be acquired in the vector of measurement values $y(d)$) by maximizing the utility function ϕ:

$$d_{opt} = \arg\max_{d \in D}[\phi\{d\}], \tag{9.1}$$

where D is the space of admissible designs. This step can be managed by a broad spectrum of optimization schemes that do not scan the entire design space. For more details and background I refer to original publication by Leube et al. [2012b].

To make the consideration of uncertainty even more general, the current application embraces the concept of Bayesian geostatistics [Kitanidis, 1986]. This goes along with idea that the selection of a single geostatistical, structural or conceptual model is often unjustifiable. To reduce the subjectivity of a-priori assumptions, one may admit different model alternatives and weight them according to their a priori credibility. The modeling task is performed with all model alternatives, and posterior credibility values are assigned after comparison with available data. This procedure is called Bayesian model averaging (BMA) [e.g., Hoeting et al., 1999; Neuman, 2003]. When included into OD approaches it is called Bayesian (geostatistical) design [Nowak et al., 2010]

In the following application structural uncertainty is split into (1) structural parameters $\boldsymbol{\theta}$ related to potentially involved geostatistical models, (2) uncertainties of boundary/initial condition parameters $\boldsymbol{\xi}$ associated with each physical/conceptual model, and (3) uncertain conceptual model selections within $\mathbf{k}$ that formally switches between several available conceptual models, such as different structures of zonation or model forcing. Within PreDIA, averaging over the unknown meta-parameters $\boldsymbol{\theta}$, boundary/initial conditions $\boldsymbol{\xi}$ and model choice indicators $\mathbf{k}$ is done implicitly. For more details on continuous BMA, I refer to Leube et al. [2012b]. Tab. 9.1.1 summarizes all meta-parameters and their assigned distributions.

In the featured application case, the set of uncertain structural parameters $\boldsymbol{\theta}$ contains σ^2 accounting for the field variance, and λ_i being the correlation length scales in spatial directions x_i. To allow for more flexibility with the assumption on the covariance model, several recent studies suggested to use the Matérn family (see Sec. 2.4). I opt this mentality and implemented the Matérn family for the following scenarios.

Uncertain structural parameters $\boldsymbol{\theta}$			
Variance	σ_T^2	[-]	$\mathcal{N}(\mu = 2.0, \sigma = 0.3)$
Integral scale	λ	[m]	$\mathcal{N}(\mu = 15, \sigma = 2.0)$
Matérn Kappa	κ	[-]	$\mathcal{U}(a = 5, b = 36)$
Uncertain Boundary Parameters $\boldsymbol{\xi}$			
Deviation from center	ν	[°]	$\mathcal{N}(\mu = 0.0, \sigma = 10)$
Uncertain Conceptual Models $\boldsymbol{k}$			
Existence of hydraulic barrier	-	[-]	$\mathcal{B}(p = 0.3)$

Table 9.1.: Uncertain structural and boundary parameters and their assigned distributions.

9.1.2. Application

The features application is generally based on the setup of scenario (2) described in Sec. 3.2. This application also assumes a drinking water well or a similar sensitive location threatened by the upstream located source. This location is about seven expected-integral scales downstream of the contaminant source and about half an integral scale offset from the center line of the expected plume path. The goal of PreDIA is to find the sampling pattern which optimally reduces the uncertainty of predicting the long-term (steady state) contaminant concentration to be expected at the sensitive location.

The uncertain values for the Dirichlet flow boundary condition in $\boldsymbol{\xi}$ are determined by two uncertain parameters γ and ν which define the regional head gradient via its slope γ and orientation angle ν relative to the northern/southern boundaries.

Different to the scenario description in Sec. 3.2, this application also features an uncertain conceptual model choice manifested in $\boldsymbol{k}$. $\boldsymbol{k}$ encodes a possibly present hydraulic barrier south of the prediction target due to uncertainty in geological medium boundaries. For the sake of scenario variation, it is assumed that local hydrogeologists are uncertain about the extent of a narrow zone filled with a different geological facies which might be present in that area. For simplicity, this is implement as a rectangle ($x = 180$ m, width = 10 m, length = 75 m) with a different mean value for log-conductivity of T = ln 10^{-7}. The prior probability of this alternative model is set to 30 %. Please note that the possibly present barrier is only considered in one case.

Concentrations c are considered to be not available as measurement data, because the spill just happened and the plume has not evolved yet. Instead, only head and transmissivity data shall be optimally collected in order to maximize the reduction of uncertainty in prediction z. I define data on transmissivity T and hydraulic head h to have measurement errors $\sigma_{r,T}$ and $\sigma_{r,h}$, respectively, to be measurable at the point scale, e.g., by disturbed core-samples and by small monitoring wells. For instructive reasons, transmissivities T are not sampled at the same locations as hydraulic head h by default, since this will help to better exhibit and discuss the underlying physics associated with the respective choice of location

and data type. Locations where T is informative may not be informative for h measurements, because different physical flow and transport-related phenomena may co-ordinate the individual data types to different informative locations. However, my framework could easily handle constraints such that T and h measurement locations can be forced to coincide.

A large sample size of 50,000 realizations has been chosen to ensure that my discussion of the method and resulting designs is not compromised by statistical noise. I use a greedy search followed by a sequential exchange algorithm [Christakos, 1992] in order to optimize the design, and the utility of each design candidate is evaluated with PreDIA.

In order to quantify the filter degeneracy of PreDIA (see Sec. 6.3), I use an extension of the Effective Sample Size (ESS) introduced by [Liu, 2008]. To this end, I average the ESS over the spectrum of potential measurement values $y(d)$ arriving at the Averaged Effective Sample Size (AESS). To guarantee proper preposterior statistics, the AESS is monitored carefully during the optimization procedure.

In the following, I consider five different cases (2f) - (2i) based on scenario (2), each one following a different research objective. They will serve to show that PreDIA can (1) include arbitrary prediction goals regardless of their non-linearity and (2) that it can include arbitrary task-driven formulations. Also, I address the consideration of additional conceptual model uncertainty, i.e., via incorporating a hydraulic barrier. The resulting cases are

Cases (2e) and (2f) Minimum-variance prediction of a contaminant concentration c at the sensitive location. To emphasize the difference to conventional linear methods, I compare the results of my method to results from an Ensemble Kalman Filter (EnKF) [e.g. Herrera and Pinder, 2005; Evensen, 2007]. Therefore, I run a first scenario variation with PreDIA (case 2e) and compare the results to a sampling pattern obtained from an EnKF (case 2f).

Cases (2g) and (2h) Maximum-confidence prediction of whether a critical concentration threshold will be exceeded. This is equivalent to predicting an indicator quantity $\mathbf{z} = I(c > c_{crit})$, with $E[I] = P(c > c_{crit})$. Since the indicator is a discrete variable that depends very non-linearly on model parameters, it does not meet the requirements under which EnKFs can be used for comparison. Instead, two threshold values are considered with PreDIA: $c_{crit} = P_{15}$ (case 2g) and $c_{crit} = P_{75}$ (case 2h), where P_{15} and P_{75} are the c-values below which 15 % and 75 % of the c-values may be found, respectively.

Case (2i) Consideration of a possibly present hydraulic barrier and minimum-variance prediction of a contaminant concentration c at the sensitive location.

9.1.3. Results and Discussion

In this section, the sampling patterns resulting from the synthetic test case and its variations defined in the previous section are presented and discussed.

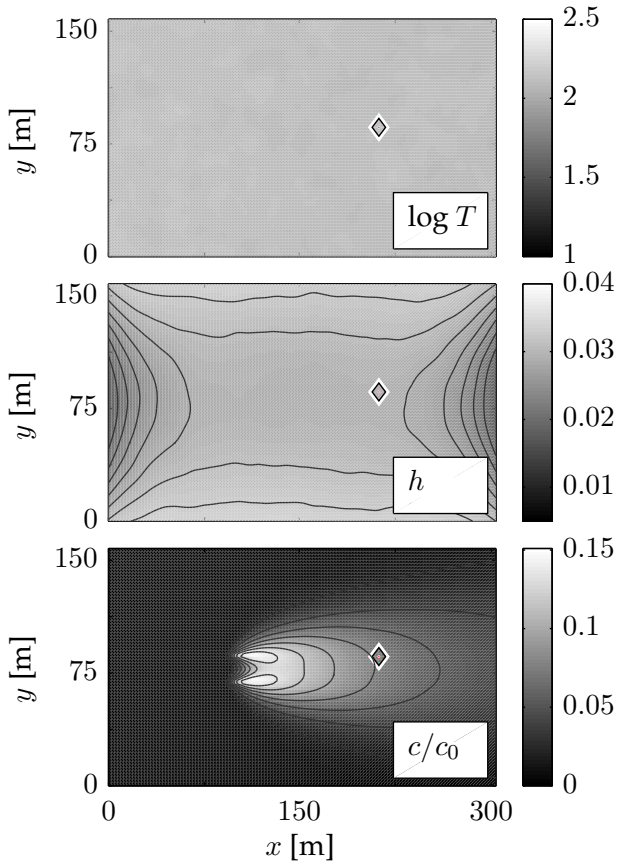


Figure 9.1.: Prior uncertainties (variance) associated with transmissivity (top), hydraulic head (center) and concentration (bottom) based on the uncertain structural and boundary parameters listed in Tab. 9.1.1.

Sampling Pattern Optimized for Predicting Concentration (Case 2e)

Case (2e) features optimal sampling for minimum-variance prediction of concentrations at the sensitive location. The resulting sampling pattern, obtained with PreDIA, is shown in Fig. 9.2 (left). Fig. 9.1 shows the respective variances of T, h and c prior to investigation. In Fig. 9.2 (left), I also included the expected conditional variance of transmissivity (top), hydraulic head (center) and predicted concentration (bottom). The basic characteristics of the design pattern mostly coincide with the results found in Nowak et al. [2010] who considered a similar scenario. However, there are important differences since they used an EnKF and I employ PreDIA. With regard to the sampling pattern, I find two predominant groups: (1) measurements gathering around the source and (2) measurements flanking the expected migration path of the plume. Near-source measurements are exclusively occurring as transmissivity measurements. They are highly informative since they provide information about the volumetric flow rate through the source area. The flow rate through the source, in turn, is

a dominant factor that dictates the total contaminant mass flux, the expected width and the dispersion characteristics of the plume further downstream [de Barros and Nowak, 2010].

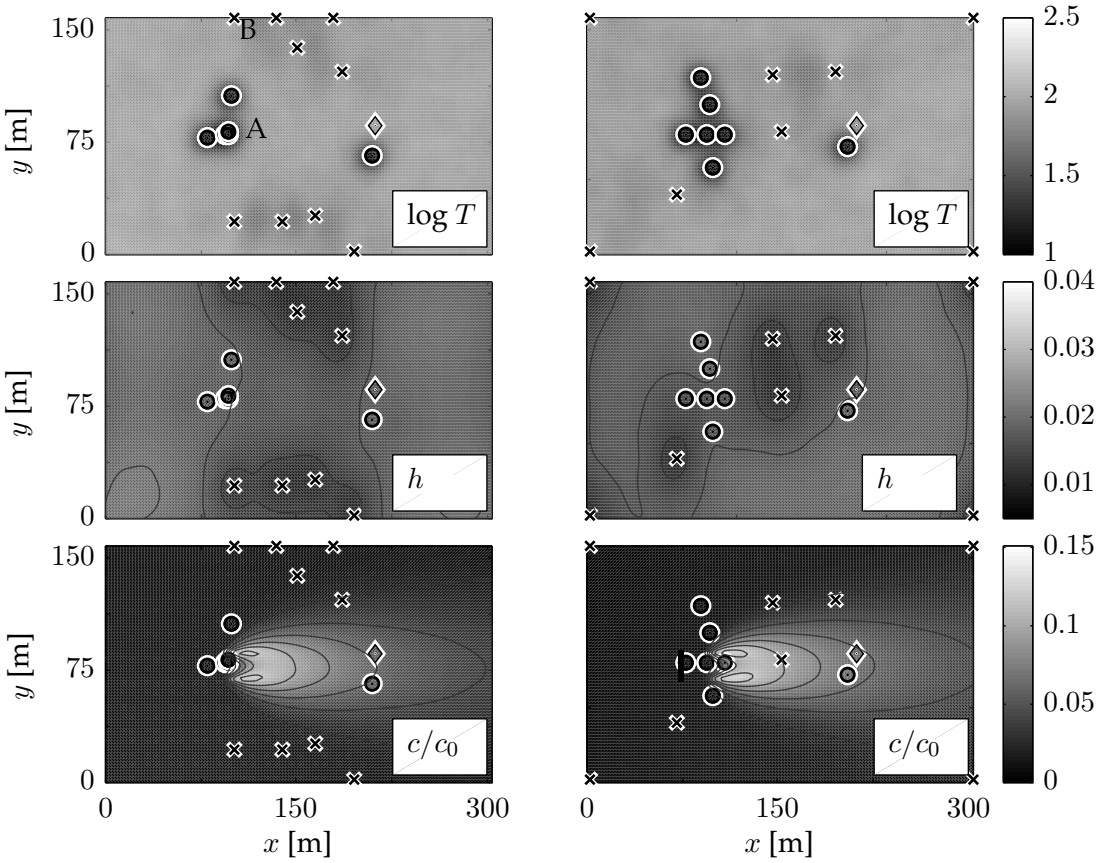


Figure 9.2.: PreDIA-based (left, case 2e) and EnKF-based (right, case 2f) sampling pattern optimized for minimum prediction variance of concentration at the sensitive location. Head measurements (crosses), transmissivity measurements (circles), source (box) and target (diamond). Maps in the background are expected pre-posterior variances for transmissivity (top), hydraulic head (center) and concentration (bottom).

The measurements flanking the plume are head measurements which capture both the large-scale drift of the plume (due to the uncertain regional head gradient) and the meso-scale meandering of the plume (caused by heterogeneity).

In principle, the prediction task leads to information needs that manifest themselves most in those regions where the statistical dependency between the measurable quantities (transmissivity or hydraulic head) and the prediction goal is highest, while avoiding mutually too close measurements that would merely convey to redundant information. Fig. 9.3 shows the statistical dependencies between observable quantities at potential measurement locations and the prediction target for a near-source transmissivity measurement location (A,

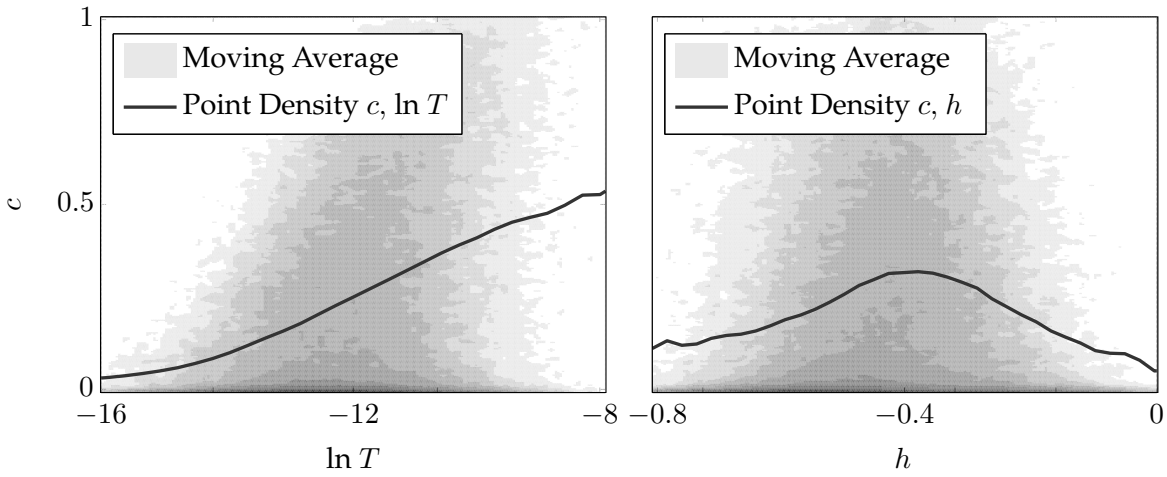


Figure 9.3.: Scatter density plots depicting the relation between the sample of predicted concentrations and the sample of transmissivity values at a near-source location A (left) and hydraulic head values at a near-boundary location B (right). The solid line illustrates the relation via moving average.

left figure pane) and a near-boundary head measurement location (B, left figure pane). The statistical dependencies are obtained by plotting the sample of possible measurement values against the sample of predicted concentrations. I additionally illustrate the non-linear dependency in the scatter plot by a moving average line.

Obviously, T at the near-source location (A) has a mostly linear relation to the predicted concentration. The higher the transmissivity at the source, the higher is the source discharge and the broader is the plume on average after leaving the source. Therefore, the plume is far more likely to maintain high concentrations even over long travel distances, and is more likely to hit the target [de Barros and Nowak, 2010].

Opposed to that, h at the near-boundary location (B) exhibits a non-linear dependency to the prediction goal. Extreme angles of the regional flow gradient divert the plume away from the target location, for both positive and negative values of the angle. By contrast, regional flow in the straight uniform direction drives the plume, most likely, through the target. The resulting dependency between hydraulic heads close to the boundary and the predicted concentration has an almost purely quadratic behavior, and shows almost no correlation in a linear sense, i.e. has almost zero covariance.

Fig. 9.4 (left) illustrates how the individual transmissivity or hydraulic head measurements added during the greedy part of the optimization reduce the variance of the prediction goal and related physical quantities. The latter include the total solute mass flux through the source, the angle of the boundary condition (causing a large-scale drift), the width of the plume at the target (lateral spreading) and the lateral position of the plume's centroid (also affected by meso-scale meandering caused by heterogeneity).

I can clearly see that transmissivity measurements located closely to the source greatly re-

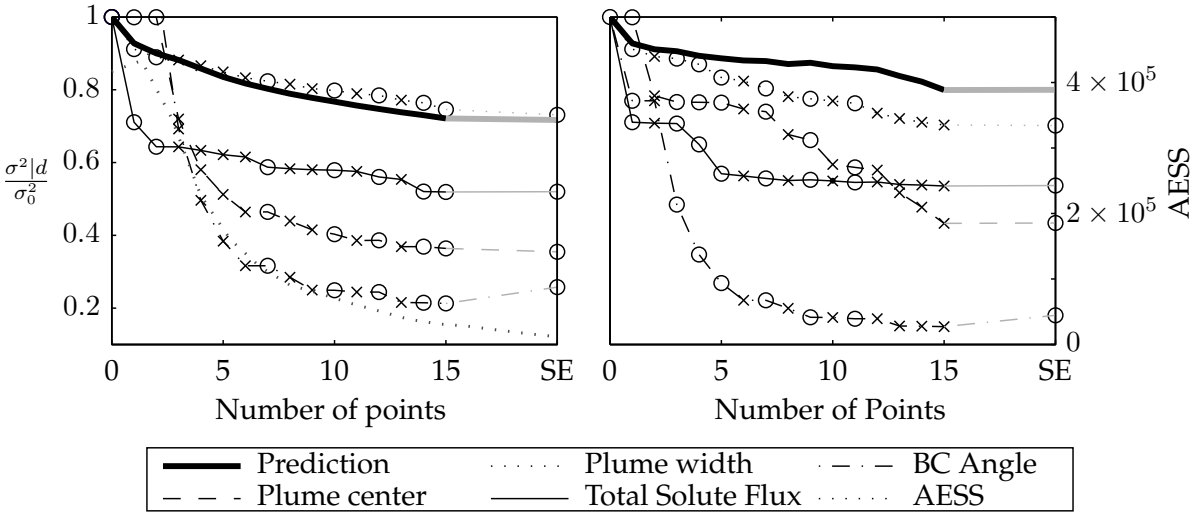


Figure 9.4.: Expected variance reduction for PreDIA (left, case 2e) and EnKF (right, case 2f) during greedy-based placement of samples for different auxiliary quantities. The sequential exchange (SE) phase is not shown in detail but only indicated by the gray lines. Hydraulic head measurements are denoted by cross marks and transmissivity measurements by circle marks. The right axis quantifies, for the PreDIA-based optimization, the respective Averaged Effective Sample Size (AESS).

duce the prediction uncertainty of the total solute flux (also see Fig. 9.4) for this case, while the head measurements along the flanks are almost not informative to the total solute flux. Instead, the uncertainty of the boundary condition (regional flow direction) is greatly reduced by the head measurements, whereas the transmissivity measurements around the source contribute almost no related information (also see Fig. 9.4). Likewise, the position of the plume center is revealed almost solely by head measurements. For the plume width at the prediction target, I find a sensitivity to both head and transmissivity measurements, where the first two transmissivity measurements at the source are clearly the most valuable ones.

Comparison to EnKF (Case 2f)

The sampling pattern provided by the Ensemble Kalman Filter (EnKF) relies on exactly the same geostatistical and boundary parameters used in case (2e), and hence uses the very same MC set of possible sample data. For technical insights in the EnKF formalism, please be referred to Herrera and Pinder [2005] or Evensen [2007]. The resulting pattern is shown in Fig. 9.2 (right column). The underlaid maps of expected conditional variance are evaluated by PreDIA, because the maps provided by the EnKF are inaccurate and would not be comparable to those shown in the left part of Fig. 9.2.

Compared to the PreDIA-based sampling pattern (case 2e), I find again the group of trans-

missivity samples in the source area. However, the number of measurements in this group is much larger. The next fundamental difference to the PreDIA-based sampling pattern is that the group of head measurements at the northern and southern domain boundary is smaller in favor of head measurements in the corners of the design domain. Apparently, the relevance of the variable boundary conditions that induce large-scale drift of the plume is also recognized, but judged differently by the EnKF analysis scheme.

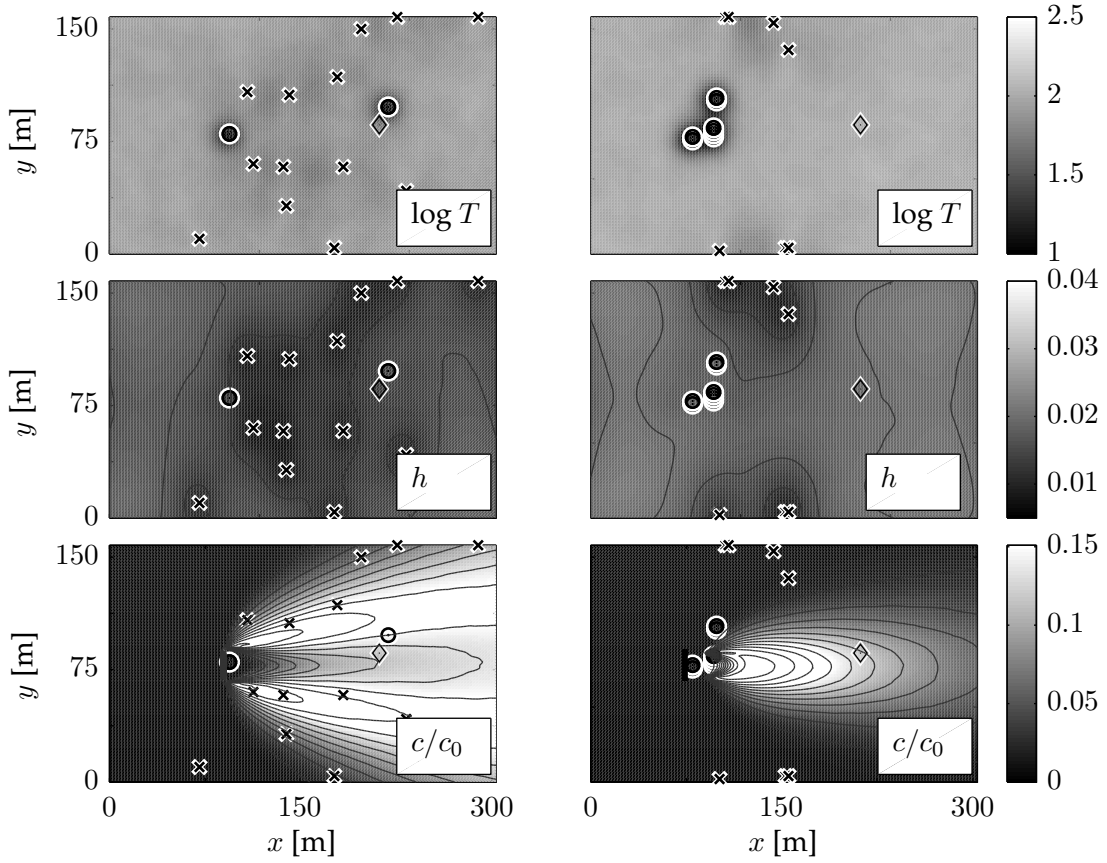


Figure 9.5.: PreDIA-based sampling pattern optimized for predicting the exceedance of a low c_{crit} (left, case 2g) and high c_{crit} (right, case 2h). Head measurements (crosses), transmissivity measurements (circles), source (box) and target (diamond). Maps in the background are preposterior variances for transmissivity (top), hydraulic head (center) and indicator variable (bottom). A selected near-source location is marked by A, whereas a near-boundary location is marked by B.

The EnKF assesses statistical dependencies only via covariances, which are a measure for linear dependence only. It is unable to capture even-order (e.g., quadratic) dependencies such as between head measurements near the northern and southern boundary and the prediction goal (see Fig. 9.3). Therefore, it simply ignores these head measurement locations as potential sources of valuable information. Hence, crucial information about the meso-scale

meandering of the plume is neglected. However, four measurement locations were placed at the corners of the allowable design locations. Apparently, their non-linear dependency exhibits a sufficiently large linear component due to the slight asymmetry of my setup.

Overall, this leads to a significantly worse performance in reducing the uncertainty associated with the plume center, even though the EnKF captures the uncertain boundary condition reasonably well. This can be seen by comparing the expected conditional variance within Fig. 9.4 (left and right). With a higher relative emphasis on the mostly linear source transmissivity information, the plume width and total solute flux are determined comparably well. Still, the overall prediction quality of concentration c is reduced by ignoring and misinterpreting non-linear information, such that PreDIA clearly outmatches the EnKF. In my setup, PreDIA achieves 25 $\%$ more uncertainty reduction with the same number of sampling positions than the EnKF.

In more general terms, EnKFs and all linear(ized) methods can only measure correlation, which is a very incomplete access to statistical dependence. For example, zero correlation between a zero-mean variable and its square does not imply at all that a squared value is independent of its square root. Hence, the limitations of linear(ized) methods illustrated in my specific example generalize to all non-linear applications.

Sampling Patterns Optimized for Predicting Exceedance Probability (Cases 2g and 2h)

In this test case, I desire maximum-confidence prediction whether a critical concentration value (e.g. imposed by a regulatory threshold) will be exceeded or not. The PreDIA-based sampling patterns for cases (2g) and (2h) are shown in Fig. 9.5, again obtained from the same MC sample.

Case (2g) ($c_{crit} = P_{15}$) exhibits a sampling pattern which is mainly based on head measurements at near-boundary and towards-target locations. Transmissivity measurements exploring the source region are practically absent. For predicting low threshold values, it is only important, and therefore sufficient, to know that the plume misses the sensitive location. This information is obtained by head measurements flanking the plume, which can reveal transverse gradients that could divert the plume from hitting the sensitive location.

Case (2h) ($c_{crit} = P_{85}$) shows an inverted behavior, where the source is sampled repeatedly using six transmissivity samples that are hardly distinguishable in Fig. 9.5. Two additional transmissivity samples north of the source support the near-source samples by addressing the contrast in transmissivity between the source and its surroundings. Instead, head measurements closely flanking the plume are disregarded. This is a direct consequence of the different information needs between case (2g) and (2h). For high threshold values, it is necessary to know whether the plume preserves its initial peak concentration over large travel distances up to the sensitive location. Highly conductive sources favor this behavior, and can be identified by increasing the source sampling density. In addition, highly conductive sources statistically imply an increased downstream plume width. With the plume sufficiently wide, the chances of bypassing the sensitive location by meso-scale meandering

decrease and only a globally rotated mean flow direction can prevent the plume from hitting the sensitive location. That is the reason why (1) transverse gradients and the related head measurements are not closely flanking the plume, and (2) there are more remote head samples at the northern and southern boundaries that help to infer the global flow direction without being disturbed by heterogeneity-induced smaller-scale head fluctuations.

In order to emphasize the task-specific character of the individual design patterns towards their respective prediction goal, I applied each design pattern to the prediction goals of all other test cases. This yields the performance indices summarized in Tab. 9.2.

Case	(2e)	(2f)	(2g)	(2h
(2e)	**100.00** %	75.14 %	79.10 %	95.99 %
(2g)	81.41 %	76.03 %	**100.00** %	69.01 %
(2h)	90.43 %	38.79 %	27.54 %	**100.00** %

Table 9.2.: Performance indices for every sampling design when applying on different prediction goals.

The performance indices show that the PreDIA-based design pattern (2e) clearly outmatches the EnKF (2f) for all three prediction goals. The EnKF-based design pattern is even surpassed in its own objective by the PreDIA-based sampling patterns designed for cases (2g) and (2h). The worst performance was found for the pattern of case (2g) (low threshold) when applied to the objective of case (2h) (high threshold). This can be explained by the fact that these two patterns lay their focus on opposed features in their respective design objectives, i.e. on meso-scale meandering versus source conductivity. The opposite case (applying the pattern of case 2g to case 2h) performs better. Obviously, in my specific examples, many source conductivity measurements are more generic all-purpose information than head measurements populating the boundaries.

Sampling Patterns Accounting for Conceptual Model Uncertainty (Case 2i)

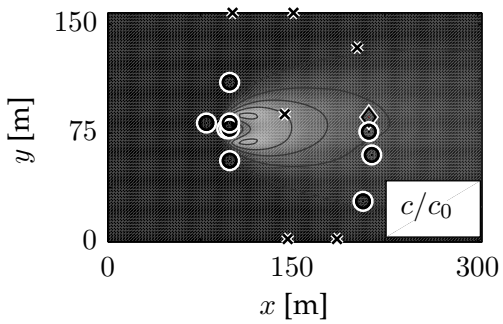


Figure 9.6.: Sampling pattern (case 2i) when considering conceptual model uncertainty exemplary represented by a hydraulic barrier.

The optimized sampling pattern for case (2i) is shown in Fig. 9.6. Opposed to the previous cases, case (2i) also considers conceptual model uncertainty, represented by a possibly present hydraulic barrier. If present, the barrier causes a flow regime which forces the plume to swerve northwards and so increases the chance that the plume hits the sensitive location. The strong dependence of the predicted concentration on the presence of the hydraulic barrier requires an adequate model choice. Therefore, the sampling pattern reacts to this additional uncertainty. Compared to case 1a, three transmissivity measurements are placed in the area of the possibly present barrier, whereas most other design features are preserved.

Although I did not use model choice as objective function for the design (the importance of model choice is only implicit via its role in my chosen prediction goal), the reliability of correct model choice is improved by the adapted sampling pattern provided by PreDIA. This effect can be illustrated best by computing the preposterior weights of the two different hypothesized models: Among all possible data sets generated with the barrier, the model with barrier obtains (on average over all those data sets) a weight of of 98%. Among all possible data sets generated without the barrier, the model without the barrier receives an average weight of 50%. Weighting both preposterior cases by their prior probabilities to occur (i.e. 70% and 30% respectively) yields an expected reliability of 85% to choose the correct model. This is a significantly increased reliability compared to the prior stage, where the reliability lies at 58%.

PreDIA also allows to perform a full BMA analysis including measures like pre-posterior inter-model and intra-model variances, because all statistics are available. However, I omit this analysis here for the sake of brevity. As for the computational costs and convergence issues, the AESS drops in case (2i) from 500 (cases 2e - 2h) to about 200. This is owed to the increased variability and uncertainty in hydraulic conductivity introduced by the possibly present hydraulic barrier.

9.1.4. Summary and Conclusions

In this chapter, I applied PreDIA to an optimal design problem taken from contaminant hydrogeology, where I illustrated its applicability to different sources of uncertainty, various prediction tasks and task-driven objective functions. Within a groundwater quality example, I considered non-co-located hydraulic head and transmissivity measurements. In order to show the limitations of linearized methods, I compared the optimal design patterns obtained via PreDIA to those from an EnKF.

When applying PreDIA in a optimal desing framework, I found the following conclusions most important:

1. PreDIA outmatches linearized methods (such as EnKFs) because linear methods fail to recognize relevant non-linear relations between potential measurement locations and the prediction goal, and hence oversample locations considered to be most informative from the limited viewpoint of linearized analysis.

2. PreDIA can handle arbitrary task-driven formulations of optimal design. I demonstrate this in a scenario variation that involves predicting the exceedance of a regula-

tory threshold value, which is important for risk management [e.g., de Barros et al., 2009]. The sampling pattern for the task-driven prediction strongly depends on the level of the threshold value, because different information needs are triggered by the underlying flow and transport physics. Neither this difference nor such classes of task-driven formulations could be handled by linearized methods.

3. The number of MC realizations needed by PreDIA for convergence rises with the number of planned sampling points and their measurement accuracy. This is inherited from BFs in general. The Averaged Effective Sample Size (AESS) serves as a sound measure to monitor statistical convergence. However, the problem of filter degeneracy is still a challenge when planning extensive sampling campaigns. An extension of PreDIA towards more efficient stochastic methods would help to further increase the affordable sampling size. Here, linear methods are superior as they benefit from fast analytical solutions.
4. Bayesian model averaging is implicit in PreDIA at no additional conceptual costs, and allows to reduce the subjectivity of prior assumptions on, e.g., geostatistical parameters, boundary parameters or physical/conceptual model alternatives. Introducing more variability to models might increase the computational costs or might lead to a decrease in the AESS. Incorrect prior assumptions could negatively affect the quality of the resulting optimal designs.
5. My specific illustrative example showed that the uncertain direction of a regional groundwater flow has a significant impact on the uncertainty of predicting contaminations, and should hence not be neglected. This additional uncertainty can be quickly reduced by hydraulic head measurements at large distances.
6. In my specific case, the optimal design predominantly addressed uncertainty in head boundary conditions and contaminant source hydraulics, rather than structural uncertainty in the geostatistical model. This will change according to the relative importance of individual sources of uncertainty, and the availability of data types that are adequate to address these individual uncertainties.

9.2. Optimal Resource Allocation in other Model Dimensions

ORA as introduced in Chap. 7 has been introduced as versatile tool to explore the potential of optimally allocating limited computational resources. In an application the optimum between the number of TM and the number of repeated MC simulation has been determined and discussed. In the following section, I will the extend the range of application towards any other model dimension, two more complex scenarios and a more ambitious error measure.

9.2.1. Application

The principal question is again: Can a hypothesized optimal trade-off be found for between the reduction or discretization of two different model dimensions, given a certain computa-

tional budget? In this application, I choose the number of repeated MC simulations denoted by $\Delta\xi$ and the spatial discretization denoted by $\Delta\mathbf{x}$ as the model dimensions to be investigated. Overall, I consider two scenarios, the complex well scenario featuring a wetland management (scenario 3) and the complex transport scenario featuring a remediation campaign (scenario 4) (see Chap. 3). Considering two different physical setups allows to investigate to some extent the influence of system physics and optimal trade-offs. In order to evaluate the sensitivity with respect to a goal-oriented choice the error metric E_a, scenario (4) features two different predictions and overall modeling goals denoted by case (4a) and case (4b). In the following, all three cases including their predictions $\Omega(\Delta\mathbf{x}, \Delta\xi)$ are described in more detail.

Case (3) The prediction $\Omega(\Delta\mathbf{x}, \Delta\xi)$ is the spatial and statistical 95-th percentile s_{95} of all drawdown values s in the sensitive area. This is the drawdown value s which is exceeded in only 5% of the wetland area, in only 5 % of all realizations. Due to the steady-state condition and the selection of a single value from the spatial domain, the prediction is independent of space and time. Thus, only spatial and statistical discretization is required, and no spatial or temporal aggregation is performed in the error metric E_a Eq. 7.1.

Case (4a) The prediction $\Omega_{4a}(\Delta\mathbf{x}, \Delta\xi)$ is the mean total mass flux of reactants passing the control plane. The mean total mass flux is defined as the integral of the concentration $c(\mathbf{x}, \xi_k)$ [M/L^3] times the velocity $v(\mathbf{x}, \xi_k)$ [L/T] over the plane, averaged over the ensemble of MC realizations $k = 1 \ldots n_{MC}$.

Case (4b) Motivated from decision theory, the prediction $\Omega_{4b}(\Delta\mathbf{x}, \Delta\xi)$ is defined as the model supported decision to conduct ($\Omega_{4b} = 1$) or abandon ($\Omega_{4b} = 0$) the remediation endeavor that is involved in scenario 4. The remediation is conducted if a predefined critical reactant loss (lost mass flux = 0.018 kg/s) is deceeded in 90 % of all MC realizations. This is a hypothesis-driven decision analysis, similar to the concept featured by Nowak et al. [2012] for optimal design of site exploration.

In all three cases featured here, the performance of ORA is quantified by an error norm E_a that accounts for both the spatial dimension P and the stochastic dimension S, respectively, in the sense of Eq. (7.1). For the spatial dimension, the most common, straightforward and reliable technique to assess numerical errors are systematic grid-convergence studies [e.g., Roache, 1997]. Such studies compare the numerical solution against a finer reference resolution $\Delta\mathbf{x}_{Ref}$ and Δt_{Ref} in both space and time. For cases (3) and (4a), I use the L^2-norm, equivalent to the mean-squared-error, for error quantification. However, any other quantity of interest could be used as well, as will be done for case (4b). When using the L^2-norm, the metric E_P for spatial discretization error can be found by

$$E_P(\Delta x, \Delta\xi_{ref}) = \left[\Omega(\Delta\mathbf{x}, \Delta\xi_{ref}) - \Omega(\Delta\mathbf{x}_{Ref}, \Delta\xi_{ref})\right]^2 . \tag{9.2}$$

Cases (3) and (4a) rely on error variances as metric for statistical resolution. Since the spatial and the statistical errors are not necessarily independent, I evaluate the total error jointly by

enveloping Eq. (9.2) with the non-parametric statistical bootstrapping method [Efron, 1982]:

$$R_{Tot}(\Delta x, \Delta\xi) = \frac{1}{B-1}\sum_{b=1}^{B}\left[\Omega^{(b)}(\Delta x, \Delta\xi) - \Omega(\Delta x_{ref}, \Delta\xi_{ref})\right]^2, \tag{9.3}$$

where $b = 1 \; ... \; B$ indicates the repetition over randomly drawn subsets with size $n^*_{MC} < n_{MC,ref}$ (according to $\Delta\xi$) from the total number of potential realizations $n_{MC,ref}$, and $\Omega^{(b)}$ is the respective prediction obtained by working with the $b-$th subset instead of $\Delta\xi_{ref}$.

Case (4b) considers a binary outcome with events 0 or 1, representing the decision to conduct (1) or abandon (0) the remediation plan. As total error metric, I assess the probability of providing the wrong decision (compared to the decision resulting from the reference resolutions) due to insufficient spatial or statistical resolution.

The average wall clock time required to generate one realization of $\Omega(\Delta\mathbf{x}, \xi_k)$ at resolution $\Delta\mathbf{x}$, denoted as $\overline{\Sigma}$ is utilized as the computational budget norm for all scenarios. Thus, the required total budget $B_{Tot}(\Delta\mathbf{x}, \Delta\xi)$ can be found by:

$$B_{Tot}(\Delta\mathbf{x}, \Delta\xi) = n_{MC} \cdot \overline{\Sigma}, \tag{9.4}$$

where n_{MC} is the budget multiplier to generate n_{MC} realizations $\Omega(\Delta\mathbf{x}, \xi_k)$ at the resolution $\Delta\mathbf{x}$.

To generate the data required for Eqs. (9.2)-(9.4), I evaluate $\Omega_k(\Delta\mathbf{x}, \Delta\xi_k)$ for different discretizations $\Delta\mathbf{x}$ and $\Delta\xi$ with $\Delta\mathbf{x} = 1/32\,\lambda \; ... \; 1\,\lambda$ and $\Delta\xi = 1 \times 10^1 \; ... \; 1 \times 10^4$ MC realizations. Notice that both examples (wetland and remediation management) consider steady-state conditions, therefore only spatial discretization matters and resolutions of time is not required. I define $\Delta\mathbf{x} = 1/32\,\lambda$ to be the reference resolution $\Delta\mathbf{x}_{Ref}$, with a total number of grid points $\approx 1 \times 10^5$ for the wetland example and $\approx 1.5 \times 10^5$ for the remediation example. The randomly generated reference fields $T(\mathbf{x}, \xi_k)$ in each MC realizations are used to derive the coarser parameter fields by simply homogenizing out the respective sub-grid cells (by taking the geometric mean of 2×2 finer cells in each coarsening step). The reference predictions $\Omega(\Delta\mathbf{x}_{Ref}, \Delta\xi_k)$ serve to determine the numerical error norm $E_P(\Delta\mathbf{x})$. This is a simple way of implementation and appears suitable for this study, but is not meant to be a general suggestion for future applications. For alternative and more practical implementations, I refer to Sec. 9.2.3. I define $\Delta\mathbf{x} = 1\,\lambda$ as the lowest grid resolution. Even lower resolutions can directly be excluded in my case, because they would loose too much of the parameter variability by homogenization. Also, I do avoid compensating for sub-grid variability by sub-grid dispersion coefficients as done by, e.g., Rubin et al. [1999], Bellin et al. [2004] and de Barros and Rubin [2011].

Based on the algorithm described in Sec. 7.1 and Eqs. (9.2)-(9.4), I evaluate error-to-budget surfaces for all three scenarios and use these in order to find the optima for allocating computational resources between the spatial and stochastical discretization. Results of both error-to-cost surfaces and optimal resource patterns are presented in Sec. 9.2.2.

9.2.2. Results and Analyis

Fig. 9.7 displays examples of slices through the cost-to-error surface for different combinations of Δx and $\Delta\xi$ for case (3). Both errors decrease with finer spatial resolutions and with larger number of MC runs as expected. The errors eventually converge to the remaining spatial error and MC error, respectively as shown in Fig. 9.7 (left) and Fig. 9.7 (right). Fig. 9.7 (left) also reveals that the total error curves over varying MC resolution n_{MC} for given spatial discretizations are not parallel to each other. This indicates that possible assumptions on the independence of errors (leading to simple additivity for L^2-type norms, where one would have $E_{Tot} = E_P + E_S$) do not hold for my test case. Even for high spatial resolution ($\Delta x < 1/16\,\lambda$), the curves are still far from being parallel. The same behavior can be observed in the error curves over Δx for fixed n_{MC}, see Fig. 9.7 (right).

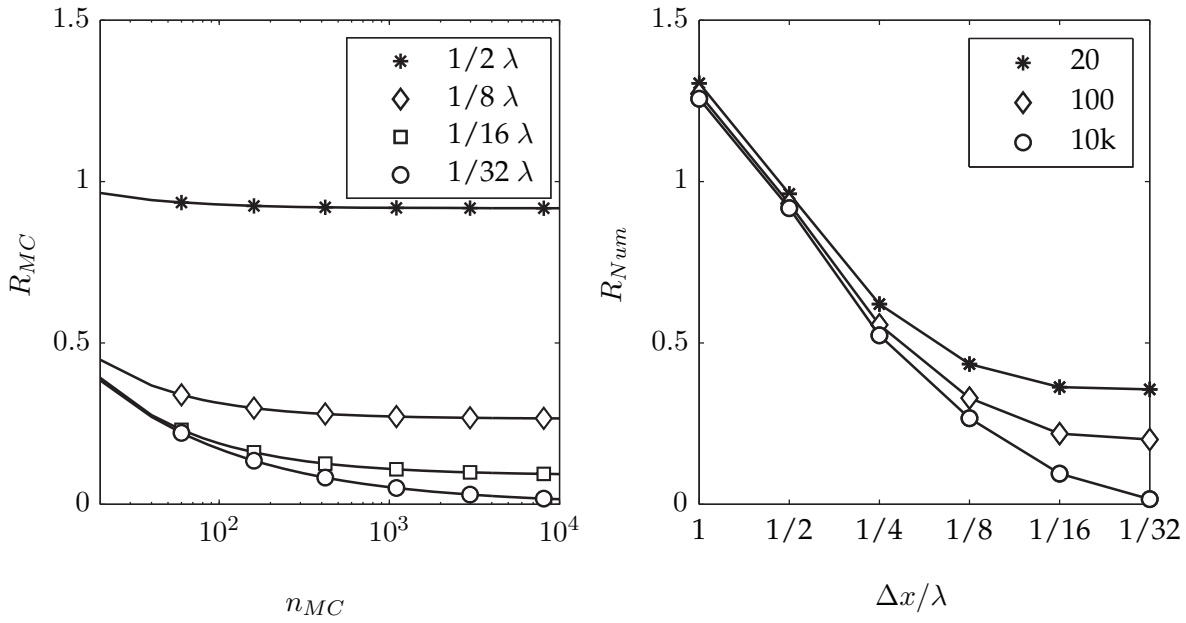


Figure 9.7.: Evolution of statistical error (left) and spatial error (right) for given statistical discretization (number of MC realizations n_{MC}) and spatial discretization Δx, respectively.

The resulting optimal allocations for cases (3), (4a) and (4b) are shown in Fig. 9.8. The optimal trade-offs depend on the available budget and move from low to high spatial and stochastical discretizations for the first two scenarios. By comparing case (3) (Fig. 9.8, left) and case (4a) (Fig. 9.8, center), the case (3) tends to achieve its optima with fewer MC realizations but with finer grids, compared to the case (4a). Both scenarios use different numerical schemes for different types of differential equations, which leads to different numerical convergence [Cainelli et al., 2012]. Still, the difference is surprising, because the solutions of a Laplace-type (elliptic) equation (groundwater flow equation in case 3) are much smoother than the numerical solution of the advective-dispersive equation at large Péclet numbers (in case 4a). This makes the spatial discretization error in hydraulic heads less sensitive to

increases in the grid size and to related issues of numerical dispersion. Additionally, the velocity field from the groundwater flow equation is by far more sensitive to coarse resolutions than the pressure solution alone, and enters only the remediation example, not the wetland case. In the optimal resource allocation of these specific two scenarios, however, I observe the opposite behavior: The transport-based case (4a) shows a tendency towards finer MC-resolution, whereas the pressure-based case (3) prefers finer spatial resolution. One possible

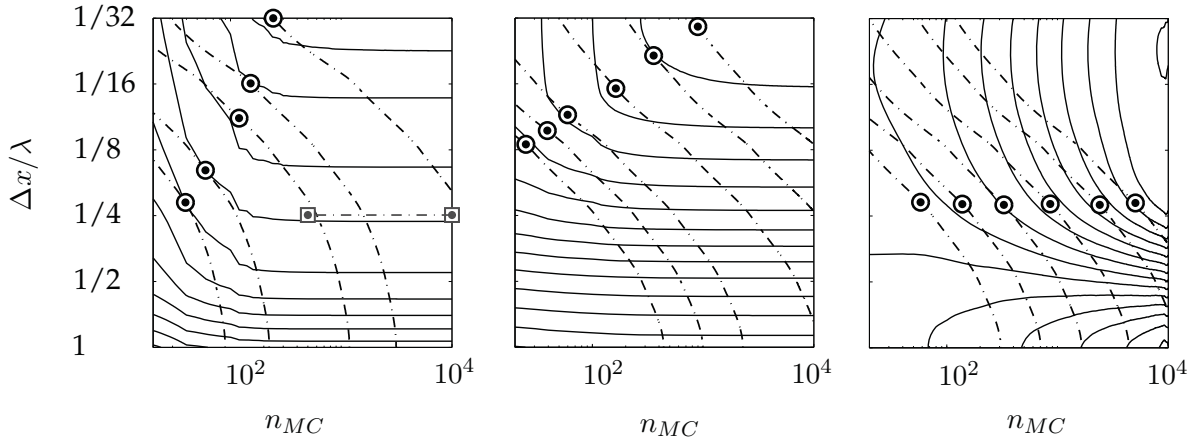


Figure 9.8.: Error-to-budget surfaces for cases (3) (left), (4a) (center) and (4b) (right). Optima for given budgets are depicted by circles. Dashed-dotted contours mark the budget-surface, whereas the solid lines represent the error-surfaces. Gray boxes interconnected with gray dashed-dotted lines mark exemplary pairs suggested by literature to be compared at their performance in Sec. 9.2.2.

explanation is that concentrations and mass flux estimates tend not to be Gaussian distributed. In general, for this class of predictions, the pdf are asymmetric [Bellin and Tonina, 2007; Schwede et al., 2008; Dentz and Tartakovsky, 2010; Cirpka et al., 2012]. Analyzing the histograms of the featured model predictions for both case (3) and (4a), I find a two times larger skewness in the remediation scenario (see Fig. 9.9). Larger skewness generally slows down the convergence of MC simulations and explains why the total error in the remediation scenario is controlled by the MC error. This, in turn, shifts the optimal allocation towards more MC realizations in the remediation scenario. Please note that, when choosing alternative numerical strategies with different convergence properties, the optimal allocation of resources may be different and would lead to different conclusions.

When comparing case (4b) in Fig. 9.8 (right) against cases (3) and (4a), I observe significantly different patterns. This is driven by the error-surface which is distinctly different in case (4b) compared to that from case (4a) (remember that cases 4a and 4b share the exactly same physics and budget surface). Here, the errors depend on how much the individual 90-th percentiles of lost mass flux for each $\Delta\mathbf{x}$ differ from the critical lost mass. MC resolution is preferred over spatial discretization, which remains fixed at $\Delta\mathbf{x} = 1/4$. This is explained by Fig. 9.10, revealing a consistent downward drift of the 90-th percentiles from low to high spatial resolutions. The most plausible reason for this effect is the increased numerical dis-

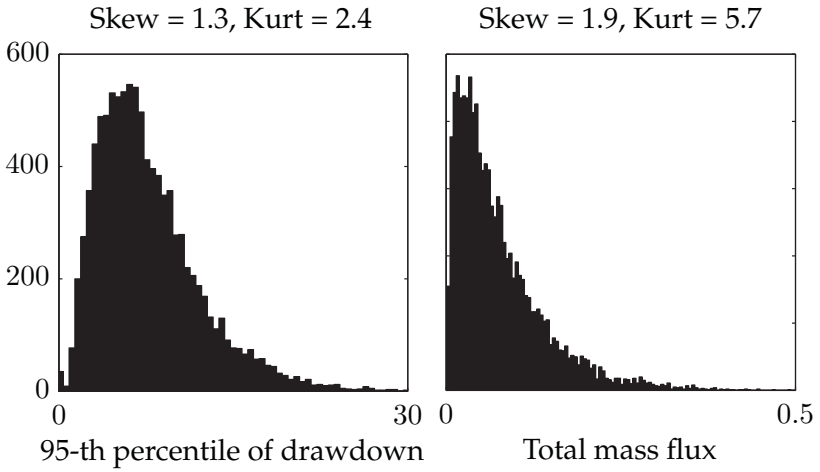


Figure 9.9.: Histograms analysis for case (3) (left) and case (4a) (right). Higher order statistical moments (skewness and kurtosis) are reported.

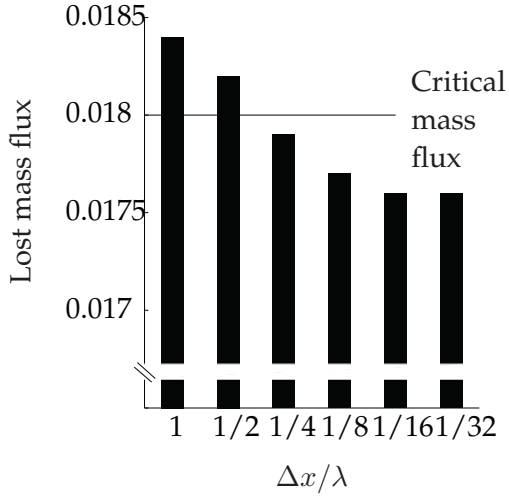


Figure 9.10.: Percentile analysis featuring the 90-th percentile for case (4b) indicating a downward drift from low to high resolutions. Obviously, the 90-th percentile of $\Delta\mathbf{x} = 1/4$ is the lowest resolution falling below the critical lost mass flux m = 0.018 kg/s.

persion at lower discretizations, driving mass away from the pumping wells. Discretizations of $\Delta\mathbf{x} = 1/4$ or finer, however, are sufficiently accurate to predict that the 90-th percentile is lower than its critical value. This makes resolutions finer than $\Delta\mathbf{x} = 1/4$ unnecessary in the light of ORA, although grid convergence for the percentile seems to be not achieved before $\Delta\mathbf{x} = 1/16$.

The apparent lesson behind these three scenarios is that, depending on the model prediction and the chosen uncertainty and error measures, the overall joint error can be controlled by different sources of error with a complex interplay and trade-offs that depend on many

influence factors. Hence, the ORA can differ substantially from individual case to individual case. Also, these examples demonstrate clearly that ORA should be viewed within a goal oriented approach. This establishes a strong similarity between optimal design of computer experiments and the goal-oriented character of optimal design for field experiments [de Barros and Rubin, 2008; de Barros et al., 2009; Nowak et al., 2010; de Barros et al., 2012].

9.2.3. Discussions and Practical Application

The results above showed that the trade-off patterns within ORA depend on the system, numerical schemes, and modeling objectives under consideration. This includes the type of physical problem, the general setup and boundary conditions, discretization types, the geostatistical assumptions, the physical type of predictions Ω and the character of the involved error measure E_{Tot}. A general suggestion on optimal resource allocation can, hence, not be given. Instead, the optimal trade-off needs to be determined carefully prior to each individual investigation endeavor.

For the sake of further discussions, I quantify the computational budget which could be saved due to optimal allocation compared to typical discretization choices found in the literature. To this end, I calculated the corresponding budgets and errors for literature values in the wetland problem (case 3). For the spatial discretization, some studies from the stochastic hydrogeological literature suggest to choose $\Delta\mathbf{x}$ according to $1/(1+\sigma)\,\lambda$ [Ababou et al., 1989], or $\Delta\mathbf{x} = 1/4\,\lambda$ [Bellin et al., 1992] in order to sufficiently resolve spatial patterns of geostatistical parameter fields. For n_{MC} found in the literature, numbers range between $500 \sim 10000$, or sometimes even higher [e.g., Englert et al., 2006; Nowak et al., 2008], independent of the complexity of the problem. Thus, I computed the joint errors for two exemplary pairs ($\Delta\mathbf{x} = 1/4\,\lambda$, $n_{MC} = 500$) and ($\Delta\mathbf{x} = 1/4\,\lambda$, $n_{MC} = 10000$), and estimated the saved budget by comparison to optimally allocated resources with the same errors (also see Fig. 9.8, left). I found that, for case (3), 81 % to 99 % of the computational time could be saved. These numbers are merely rough guesses. Still, they indicate that there is a high potential in saving computational resources or a corresponding potential for reducing the overall errors.

While the cited references mentioned above had good reasons for suggesting $\Delta\mathbf{x} = 1/4\,\lambda$ or $\Delta\mathbf{x} = 1/(1+\sigma)\,\lambda$ in their featured problem setups, I find from my analysis that this choice is good only together with small MC ensembles in my specific applications. This is a very strong evidence that the computational optimal resource allocation analysis has to be evaluated for each individual application, and guidelines for specific task should not be generalized without great care, if at all.

A further situation-specific influence is the degree of system uncertainty. The synthetic test cases I performed were both characterized by a log-conductivity variance of σ^2 = 3. With increasing variance, i.e. more parameter uncertainty, I expect the optimal trade-off to put less emphasis on spatial discretization as a first guess. This becomes obvious from the opposite limiting case: in absence of uncertainty (i.e., in a deterministic model), the spatial resolution will receive all of the computational budget, as MC runs becomes unnecessary. However, the increasing need for larger MC ensembles is in general not necessarily monotonic with

increasing heterogeneity. The reason is that the numerical error of spatial discretization also increases with contrasting parameter values in neighboring cells, and these contrasts also increase with the log-conductivity variance [Cainelli et al., 2012].

Also, I point out that possible assumptions of independent errors should not be taken lightly. The cited literature suggests that a spatial discretization in the order of $\Delta\mathbf{x} = 1/4\,\lambda$ is sufficient to accurately capture spatial statistics. This statement would imply independent (and hence additive) L^2-type errors at $\Delta x = 1/4\lambda$ and above. However, I found even for higher spatial discretizations ($\Delta\mathbf{x} \leq 1/32\,\lambda$) that the MC errors and the numerical errors are still dependent on each other. I conclude that, only for immensely fine spatial discretizations (beyond $\mathbf{x} = 1/32\lambda$), the assumption of independent errors may become legitimate in my cases.

The resulting methodology of computing all potential pairs of $\Delta\mathbf{x}$ and n_{MC} in Eq. (7.1) is more expensive than the final optimized computing tasks. Thus, my current procedure may be suitable for fundamental investigations of problem classes, but not for specific practical applications. As possible solutions for practical applications, I briefly outline three alternative approaches which could help bring forth these ideas into practice.

- **Approach 1** is based on the idea of optimal search algorithms. As both the budget and error surfaces are smooth and monotonically decreasing, it is straightforward to employ a simple iterative search scheme to minimize the error under the constraint of fixed budget. This scheme avoids evaluating the surface at every location.
- **Approach 2** has its roots in the extrapolation of errors and budgets. For example, the MC-error is monotonically decreasing down to zero, and a simple parametric function could be fitted and used for extrapolation. This would require to evaluate only small MC-ensembles ($n_{MC} < 200$). In fine discretization regions, the numerical error surface may be approximated well by well-known a priori error scaling laws, such as $R_{Num} \approx 1/\Delta\mathbf{x}$ for first-order accurate schemes. Again, fitting and extrapolation may be possible.
- **Approach 3** is a possible modification of approaches 1 and 2. Their most expensive individual step is to evaluate the reference simulation. Instead of using a reference, it could be possible to work with error gradients, where each finer resolution serves as reference for each coarser resolution. This would require to re-assess the error metrics from Sec. 7.1.1.

All three concepts will deliver challenges for future research.

Another important issue to be discussed in this context is conceptual model uncertainty. In my work, I assumed that the conceptual model and its parameterization are fixed. In many applications, however, there is uncertainty about the adequate choice of model equations, boundary conditions or parameterizations. Nevertheless, there are several methods in the community that could take into account these issues. For example, structural uncertainty in geostatistics can be tackled using Bayesian geostatistics [Kitanidis, 1986; Nowak et al., 2010; Leube et al., 2012a]. Also, anomalous transport models such as continuous-time random walk [e.g. Berkowitz et al., 2006] or multi-rate mass transfer [e.g. Carrera et al., 1998] might

be perceived as competing models to the advection dispersion equation in Bayesian model averaging frameworks [e.g., Hoeting et al., 1999; Neuman, 2003]. Within my framework, these alternatives could be included within the MC ensemble by using different conceptual models, different numerical schemes or different boundary conditions in each realization. Alternatively, one might seek the respective optimal allocation under each of the competing models before they enter the model averaging procedure. While my optimal allocation would serve to evaluate each of the competing models best, the conceptual dimension of uncertainty cannot be resolved unless by collecting additional data. The optimal sampling of environmental systems for minimal uncertainty is an entirely different issue and specific techniques to reduce conceptual uncertainty have recently been proposed [e.g., Leube et al., 2012a; Neuman et al., 2012; Nowak et al., 2012; Wöhling et al., 2013].

As final note, I would like to add that ambiguity may not only arise from conceptual uncertainty in model formulations, but also from ambiguity or multi-objective formulations of modeling goals and the related error measures. This may be approached via multi-objective optimization techniques applied to the same underlying problem formulation, and will reveal the trade-offs between designing computations for one or another task-specific and competing objective.

9.2.4. Summary and Conclusions

In this section, I showed the application of ORA to a different combination of model dimensions than in Chap. 7, looking at spatial versus statistical resolution instead of statistical resolution versus number of TM for approximating the time coordinate. Also the current application used two more complex scenarios and a more ambitious error measure.

Based on my results, I highlight the following points:

- The optimal allocation of computational resources is strongly sensitive to the physical scenario, the prediction and the employed types of error metrics.
- For fine spatial resolution, I found a smaller number of MC runs when compared to those found in the literature.
- Allocating computational budgets wisely reveals, that for my specific case (3a), 81 % to 99 % of the computational time could be saved.

Despite the fact that my focus was on hydrogeological applications, the ideas put forth here could be used for a variety of environmental problems where computation also plays a strong role. Examples are pollution in open water bodies [e.g., Fischer et al., 1979] or in the atmosphere [e.g., Seinfeld and Pandis, 1997], energy fluxes on fully coupled systems [e.g., Maxwell et al., 2007] and nitrogen exchange in the hyporheic zone [e.g., Marzadri et al., 2011].

Overall, the conclusions from my study are highly similar to those from goal-oriented frameworks for optimal data acquisition campaigns in environmental multi-component systems [de Barros and Rubin, 2008]. In addition, other issues (not investigated in this paper) can

influence the optimal allocation of computational resources. For example, de Barros et al. [2009] showed within a risk-based context that the ratios between relevant scaling metrics in contaminant transport affects the need for finer numerical meshes, characterization efforts and MC runs.

On a last note, I remind that my results and conclusions are based on the chosen illustrations, and that ORA should be conducted for each specific case. The challenge and relevance of distributing computational resources wisely increases when dealing with optimal design of experiments in large-scale hydro-systems [e.g. Reed et al., 2007] and when the underlying physical and statistical models are uncertain [e.g. Neuman, 2003; Nowak et al., 2010].

10. Summary, Conclusions and Outlook

Model reduction techniques are essential tools to control the overburdening costs of complex models. One branch of such techniques is the reduction of the time dimension. Major contributions to solve this task have been based on integral transformation. They have the elegant property that by choosing suitable base functions, e.g., the monomials that lead to the so-called temporal moments (TM), the dynamic model can be simulated via steady-state equations. TM allow to maintain the required accuracy of hydro(geo)logical applications (e.g., forward predictions, model calibration or parameter estimation) at a reasonably high level whilst controlling the computational demand, or, alternatively, to admit more conceptual complexity, finer resolutions or larger domains at the same computational costs, or to make brute force optimization tasks more feasible.

In comparison to classical approaches of model reduction that involve orthogonal base functions, however, the base functions that lead to TM are non-orthogonal. Also, most applications involving TM used only lower-degree TM without providing reasons for their choice. This led to a number of open research questions:

1. Does non-orthogonality impair the quality and efficiency of TM?
2. Can other temporal base functions more efficiently reduce dynamic systems than the monomials that lead to TM?
3. How can compression efficiency associated with temporal model reduction methods be quantified and how efficiently can information be compressed?
4. What is the value of temporal model reduction in competition with the computational demand of other discretized or reduced model dimensions, e.g., repetitive model runs through Monte-Carlo (MC) simulations?

In this work, I successfully developed tools to analyze and assess existing techniques that reduce hydro(geo)logical models in time, and answered the questions posed above. To this end, I pursuit an approach starting from a classical top-down perspective considering temporal model reduction from the most general point of view first (***Step I***). This has been done by classifying alternative polynomial and non-polynomial base functions and comparing their reduction efficiency (Chap. 5). Efficiency in this context means the ability of converting dynamic systems to simpler systems at reasonable computational times.

The analysis, then, became more focused and the most efficient model reduction techniques were assessed in terms of their compression efficiency (***Step II***). Here, efficiency refers to the absence of compression errors in data due to a loss-less compression. To this end, I developed and applied a new method denoted as pre-posterior data impact assessor (PreDIA) and measured the information content of TM in a rational and generic way (Chap. 6).

Next, the results from steps ***Steps (I)-(II)*** were reflected in a global and practical light against discretized or reduced model dimensions other than the temporal one (***Step III***). For this purpose, I developed and applied a method for optimal resource allocation (ORA). The method allows to combine errors from different reduced or discretized model dimension (e.g., temporal, spatial, stochastical) and minimizes the combined error for a given computational budget (Chap. 7).

The results were then adapted and TM employed in a new method for numerical upscaling of high-contrast fractured porous media (***Step IV***). The method is based on flow-aligned blocks and uses multi-rate mass transfer (MRMT) models to parameterize unresolved sub-block heterogeneity. TM are used to efficiently capture sub-block features in dynamic model responses and so make the scale transition of parameters swift and simple (Chap. 8).

Finally, I proofed the universality of my developed tools by applying them to other tasks of subsurface hydrogeology (Chap. 9). The most important conclusions from ***Step (I)-(IV)*** identified in Chaps. 5-8 and Chap. 9 are summarized in the next five paragraphs, followed by a brief outlook.

10.1. Summary of Conclusions

Reduction Efficiency (via Alternative Integral Transforms) - *Step (I)*

1. Polynomial or non-polynomial base functions for integral transforms can be classified exhaustively by a limited set of cases. Each case has different properties regarding its reduction efficiency. The most efficient case comprises the so-called Appell sequences. They lead to the most simple recursive coupling in the reduced steady-state model equation. The Appell sequences contain the monomial base functions that lead to TM.
2. Due to the linear combination of monomials inherent in all polynomials, any temporal characteristic based on arbitrary polynomial base functions (or on cumulants) can be mimicked by TM through (non-)linear re-combination. Thus, more general polynomials would, by no means, offer improved computational efficiency as compared to TM. Also, they contain the same information as TM, simply arranged in different (non-)linear combinations. They can, hence, not capture more information from the dynamic system.
3. The only remaining integral transform that reduces the dynamic linear system of partial differential equations (PDE) to a non-coupled system of steady-state PDEs is the Laplace transform (LT). However, the choice of orders (number of considered Laplace coefficients LC) remains unclear, making LC impractical, e.g. in inverse problems.
4. There exist no set of orthogonal base functions meeting the requirements of typical hydro(geo)logical applications.
5. In total, I can conclude that TM and LC are the only temporal characteristics that can be obtained from an efficiently reduced dynamic model. It remains to investigate and compare their compression efficiency.

The results apply to any (system of) linear PDEs with the following properties:

- The spatial derivatives may have any arbitrary order,
- There may be an arbitrary number of arbitrary-order time derivatives.
- For the integration by parts to work out, the coefficients must be independent of time and independent of the dynamic model response, and the dynamic model response has to decay to zero sufficiently fast.

Compression Efficiency (via the Pre-posterior Data Impact Assessor) - *Step (II)*

1. Based on an example from groundwater flow, I found that the first two TM cover more than 80 % of the information required to characterize dynamic system response curves. Considering up to four TM captures 90 % or more of the overall information. The remaining 10 % of information are distributed among an unquantifiable number of higher moments. The lowest-order TM are always the most informative. I found similar results for an example from solute transport.
2. The distribution of information content over time differs among the TM orders. Late-time behavior can mostly be inferred from higher orders. The relevance of higher-order TM has to be judged in the light of any specific application task.
3. This is by far better than what I found for LC. One advantage of LC is that their equations are fully decoupled, such that arbitrary coefficients can be chosen in arbitrary order. This turns, however, into their greatest disadvantage, because it is a priori unknown which ones are the most informative ones. Hence, it will be close to impossible to pick the optimal set of LC that could compete with TM.
4. In total, I can conclude that there is no better way for physically based model reduction in time, than by the monomials leading to TM. This holds at least for any (system of) PDEs with the properties listed above.

Reflection (within Optimal Resource Allocation) - *Step (III)*

1. There always exists an optimal trade-off between arbitrary combinations of different discretized or reduced model dimensions, e.g., the temporal and stochastical discretization. The trade-off optimally exploits computational resources, leading to the smallest joint error given a certain computational budget.
2. When reflecting the choice of TM order against the number of MC runs, I found that their optimum heavily depends on the specific underlying physical problem (e.g., flow, or transport), the geostatistical model and the prediction task under which the optimization has been carried out.
3. Also, the optimal trade-off requires less MC runs as compared to typical numbers of MC runs found in literature, at least for my specific example.

4. In total, I can conclude that the choice of TM should always be weighted carefully against other discretized or reduced model dimensions prior to the modeling endeavor.

Adaption (to Numerical Upscaling in Fractured Porous Media) - *Step (IV)*

1. Applying TM in a new method for numerical upscaling of fractured porous media (FPM) allows for a swift scale transformation of parameters. Also, it is a first-time application of TM to FPM. Overall, the method provides an efficient tool for multi-scale modeling of FPM based on flow-aligned blocks and MRMT models.
2. The prediction accuracy of the coarse simulation depends on the coarse-scale resolution and the predicted temporal complexity (number of TM). Generally, higher resolutions can yield better accuracy, but higher resolutions are also prone to errors because the nature of determining MRMT parameters does not account for their spatial effects within the block-scale model.
3. Different definitions of dispersion coefficients have been tested. Generally, the definition based on the centralized second transverse spatial moment yields best results. Local-scale and block-scale definition by mass balances underestimate transversal mixing. They lead to a worse performance.
4. If a certain block resolution is exceeded the domain contrasts are fully resolved by the block itself and the parameterization by MRMT becomes obsolete. This is similar to the traditional ADE-based solution.
5. In total, I can conclude that TM help to improve the upscaling of highly complex systems.

Other Applications of the Developed Tools I (to Optimal Design of Experiments)

1. In the field of optimal design of experiments, the application of the Pre-posterior Data Impact Assessor (PreDIA) clearly outmatches linearized methods (such as ensemble Kalman filters) because linear methods fail to recognize relevant non-linear relations between potential measurement locations and the prediction goal, and hence oversample locations considered to be most informative from the limited viewpoint of linearized analysis.
2. PreDIA can handle arbitrary task-driven formulations of optimal design. I demonstrate this in a scenario variation that involves predicting the exceedance of a regulatory threshold value, which is important for risk management. The resulting sampling pattern strongly depends on the level of the threshold value, because different information needs are triggered by the underlying flow and transport physics. Neither this difference nor such classes of task-driven formulations could be handled by linearized methods.

3. Bayesian model averaging is implicit in PreDIA at no additional conceptual costs, and allows to reduce the subjectivity of prior assumptions on, e.g. geostatistical parameters, boundary parameters or physical/conceptual model alternatives (like hydraulic barriers).
4. In total, I can conclude that fully non-linear and non-parametric inference engines for optimal design may be computationally more expensive than existing (quasi-)linear approaches, but these computational costs pay off in the form of better experimental designs and larger freedom in relaxing assumptions and in pursuing task-driven optimization goals.

Other Applications of the Developed Tools II (to Optimal Resource Allocation in other Model Dimensions)

1. The application spectrum of ORA has been successfully translated to other model dimensions. By considering the spatial (grid resolution) and stochastical dimension (number of repeated model runs through MC simulations), a clear optimum can be found revealing interesting insights into the physical scenarios.
2. The optimal allocation of computational resources is strongly sensitive to physical scenario, the type of prediction and the employed error metric.
3. Allocating computational budgets wisely leads, for my specific examples, to a speedup in the order of 80 - 99 %.
4. In total, it can be concluded that ORA is a generic tool that helps to optimally exploit computational resources in arbitrary combinations of different model discretizations or reductions.

10.2. Overall Conclusions

Although TM have been used in different applications over the last two decades, their potential has, in my eyes, not been fully exploited. Based on my findings, I hope to encourage more studies to work with the concept of TM. Especially because the number of studies found in the literature that employ TM with real data is small, more improved tests on existing data sets should be performed as proof of concept for practical applications in real world scenarios. Also, I hope to encourage those who limited their TM applications to only lower-order TM to consider a longer moment sequence. My study results specifically provide valuable advice for hydraulic tomography studies under transient conditions to use TM up to the fourth order. This might potentially alleviate the loss of accuracy used as argument against TM by certain authors.

10.3. Outlook

Following the conclusions from above, a few things remain to be said. The analysis and conclusions made in this work were mainly based on forward consideration of model reduction in time. The context of inverse modeling has shortly been touched upon in the discussion of LC, but was not subject to an independent quantitative analysis. The striking advantage here is that TM offer a joint model reduction and data compression, leading to a simple model calibration by measured data. As a matter of fact, measured data such as pressure time series observed in pumping tests, are prone to noise. This imposes its own error to the temporal characteristics and makes their determination inaccurate. Consequently, the compression efficiency of TM reduces. For a quantitative assessment, however, measurement errors have to be determined properly which has not been done in this thesis. This would allow to assess the performance of TM versus alternatives as truncated TM or LC in the inverse problem and should be tested in future research. I expect similar results for inverting tracer data based on TM. This hypothesis is supported in parts by the study of Nowak and Cirpka [2006] who showed that including the second TM of tracer breakthrough curves for geostatistical inversion leads to better results.

A way to tackle the problem of inaccurate TM due to noisy time series is the early truncation of time series leading to truncated TM [Jawitz, 2004; Luo et al., 2006]. This helps to control the integral under the noisy time-series by cutting off late-time data that would lead to enormous integrals. Rather than sharply cutting off data time series like in truncated TM, time series could be smoothly truncated by a weighting function inside the transformation. For example, I found that weighting by e^{-t} leads to an accurate integration and efficiently decreases the influence of measurement errors on higher-order TM in the late-time parts of the integration integral. Also, the dynamic model can still be reduced to steady-state with such a steady-state weighting. However, the model reduction introduces a new type of linear PDE slightly different to the well known Laplace-type steady-state equations. The efficiency should, hence, be carefully assessed and also the potential explored in the light of information loss due the smoothing-based truncation versus the otherwise imminent information loss due to noise. Also, reduction and compression efficiency should be compared against those of the truncated TM of Jawitz [2004]; Luo et al. [2006].

Another issue is the reconstruction of time series from their TM. This is of interest in the forward problem and has been addressed for quite some time. Kučěra [1965] came up with the idea of calculating orthonormal characteristics based on Hermite polynomials obtained from linear recombinations of TM. This allows to reconstruct the original dynamic response by means of expansion from TM [Kučěra, 1965]. However, reconstructed time series from polynomial expansion techniques tend to oscillate and can produce non-physical values. This is especially undesired if late-time features such as tailing are of interest. A familiar idea based on expansions is the so-called Edgeworth expansion involving cumulants [e.g., Chatwin, 1970]. Unfortunately, this series expansion requires TM to be close to the TM of Gaussian curves. Otherwise, the series can again generate negative values in the tailing. If the model reduction is based on the Laplace transform, distribution curves can be derived by the inverse Laplace transform. This reveals reasonable accuracy, however, at the costs of 10 to 40 Laplace coefficients [e.g., Li et al., 1992; Sudicky and McLaren, 1992].

A completely different idea is based on the maximum entropy (ME) method [e.g., Jaynes, 1957; Harvey and Gorelick, 1995]. Its greatest advantage is that ME can guarantee physically reasonable curves. However, I found that ME lacks performance and accuracy when using more than three TM. This is because the method has to solve a high-dimensional non-linear optimization problem. Furthermore, I observed that the basic parametric functional shape proposed by the principle of ME leads to undesired artifacts in the shape of reconstructed curves. This is, in parts, due to oscillating polynomials within an exponential function. In my eyes, this promising method could be strongly improved at two fronts: (1) The computational demands of the involved optimization could be boosted by using specifically tailored and adaptive Gauss-Hermite integration rules for calculating the expensive integrals appearing within the optimization scheme. (2) In order to avoid the undesired artifacts, the method could be extended to the principle of minimum relative entropy [e.g., Woodbury and Ulrych, 1993], which allows for a broader class of parametric shapes during reconstruction, and allows to include prior knowledge, desired from physical principles, about the expected shape of the time series. Such improved versions of curve reconstruction could further help TM-based model-reduction to penetrate deeper into practice.

Bibliography

R. Ababou, D. McLaughlin, L. W. Gelhar, and A. F. B. Tompson. Numerical simulation of three-dimensional saturated flow in randomly heterogeneous porous media. *Transport in Porous Media*, 4(6):549–565, 1989. doi: 10.1007/BF00223627.

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, 10th edition, 1972.

P. Appell. Sur une classe de polynômes. *Annales Scientifiques de l'École Normale Supérieure (2)*, 9:119–144, 1880.

R. Aris. On the dispersion of linear kinematic waves. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 235:67–78, 1956. doi: 10.1098/rspa.1958.0082.

A. Bárdossy and J. Li. Geostatistical interpolation using copulas. *Water Resources Research.*, 44(W07412), 2008. doi: 10.1029/2007WR006115.

G. Barenblatt, I. Zheltov, and I. Kochina. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks. *Journal of Applied Mathematics and Mechanics*, 24(5): 1286–1303, 1960.

J. Bear. *Dynamics of Fluids in Porous Media*. Dover Publications, New York, USA, 1st edition, 1972.

A. Bellin and D. Tonina. Probability density function of non-reactive solute concentration in heterogeneous porous formations. *Journal of Contaminant Hydrology*, 94(1):109–125, 2007. doi: 10.1016/j.jconhyd.2007.05.005.

A. Bellin, P. Salandin, and A. Rinaldo. Simulation of dispersion in heterogeneous porous formations: Statistics, first-order theories, convergence of computations. *Water Resources Research*, 28(9):2211–2227, 1992. doi: 10.1029/92WR00578.

A. Bellin, A. E. Lawrence, and Y. Rubin. Models of sub-grid variability in numerical simulations of solute transport in heterogeneous porous formations: Three-dimensional flow and effect of pore-scale dispersion. *Stochastic Environmental Research and Risk Assessment*, 18(1):31–38, 2004. doi: 10.1007/s00477-003-0164-2.

R. Benke and S. Painter. Modeling conservative tracer transport in fracture networks with a hybrid approach based on the boltzmann transport equation. *Water Resources Research*, 39 (11):1324, 2003. doi: 10.1029/2003WR001966.

B. Berkowitz. Characterizing flow and transport in fractured geological media: A review. *Advances in Water Resources*, 25(8):861–884, 2002. doi: 10.1016/S0309-1708(02)00042-8.

B. Berkowitz and H. Scher. On characterization of anomalous dispersion in porous and fractured media. *Water Resources Research*, 31(6):1461–1466, 1995. doi: 10.1029/95WR00483.

B. Berkowitz, A. Cortis, M. Dentz, and H. Scher. Modeling non-Fickian transport in geological formations as a continuous time random walk. *Reviews of Geophysics*, 44(2), 2006. doi: 10.1029/2005RG000178.

R. Bibby. Mass transport of solutes in dual-porosity media. *Water Resources Research*, 17(4): 1075–1081, 1981. doi: 10.1029/WR017i004p01075.

M. Brusseau, R. Jessup, and P. Rao. Modeling the transport of solutes influenced by multiprocess nonequilibrium. *Water Resources Research*, 25(9):1971–1988, 1989. doi: 10.1029/WR025i009p01971.

O. Cainelli, A. Bellin, and M. Putti. On the accuracy of classic numerical schemes for modeling flow in saturated heterogeneous formations. *Advances in Water Resources*, 47:43–55, 2012. doi: 10.1016/j.advwatres.2012.06.016.

J. Carrera, X. Sánchez-Vila, I. Benet, A. Medina, G. Galarza, and J. Guimerà. On matrix diffusion: formulations, solution methods and qualitative effects. *Hydrogeology Journal*, 6 (1):178–190, 1998. doi: 10.1007/s100400050143.

P. Chatwin. The approach to normality of the concentration distribution of a solute in a solvent flowing along a straight pipe. *Journal of Fluid Mecha*, 43(2):321–352, 1970. doi: 10.1017/S0022112070002409.

G. Christakos. *Random Field Models in Earth Sciences*. Dover Publications, New York, 2nd edition, 1992.

O. A. Cirpka and P. K. Kitanidis. Characterization of mixing and dilution in heterogeneous aquifers by means of local temporal moments. *Water Resources Research.*, 36(5):1221–1236, 2000a. doi: 10.1029/1999WR900354.

O. A. Cirpka and P. K. Kitanidis. Sensitivity of temporal moments calculated by the adjoint-state method and joint inversing of head and tracer data. *Advances in Water Resources*, 24 (1):89–103, 2000b. doi: 10.1016/S0309-1708(00)00007-5.

O. A. Cirpka and P. K. Kitanidis. An advective-dispersive stream tube approach for the transfer of conservative-tracer data to reactive transport. *Water Resources Research*, 36(5): 1209–1220, 2000c. doi: 10.1029/1999WR900355.

O. A. Cirpka and W. Nowak. First-order variance of travel time in nonstationary formations. *Water Resources Research.*, 40(3), 2004. doi: 10.1029/2003WR002851. W03507.

O. A. Cirpka, E. O. Frind, and R. Helmig. Streamline-oriented grid generation for transport modelling in two-dimensional domains including wells. *Advances in Water Resources*, 22 (7):697–710, 1999a. doi: 10.1016/S0309-1708(98)00050-5.

O. A. Cirpka, E. O. Frind, and R. Helmig. Numerical simulation of biodegradation controlled by transverse mixing. *Journal of Contaminant Hydrology*, 40(2):159–182, 1999b. doi: 10.1016/S0169-7722(99)00044-3.

O. A. Cirpka, E. O. Frind, and R. Helmig. Streamline-oriented grid generation for transport modelling in two-dimensional domains including wells. *Advances in Water Resources*, 22 (7):697–710, 1999c. doi: 10.1016/S0309-1708(98)00050-5.

O. A. Cirpka, C. M. Bürger, W. Nowak, and M. Finkel. Uncertainty and data worth analysis for the hydraulic design of funnel-and-gate systems in heterogeneous aquifers. *Water Resources Research.*, 40(11), 2004. doi: 10.1029/2004WR003352. W11502.

O. A. Cirpka, M. Rolle, G. Chiogna, F. P. J. de Barros, and W. Nowak. Stochastic evaluation of mixing-controlled steady-state plume lengths in two-dimensional heterogeneous domains. *Journal of Contaminant Hydrology*, 138–139:22–39, 2012. doi: 10.1016/j.jconhyd.2012.05.007.

A. Cortis and J. Birkholzer. Continuous time random walk analysis of solute transport in fractured porous media. *Water Resources Research*, 44, 2008. doi: 10.1029/2007WR006596. W06414.

J. A. Cunningham and P. V. Roberts. Use of temporal moments to investigate the effects of nonuniform grain-size distribution on the transport of sorbing solutes. *Water Resources Research.*, 34(6):1415–1425, 1998. doi: 10.1029/98WR00702.

F. P. J. de Barros and W. Nowak. On the link between contaminant source release conditions and plume prediction uncertainty. *Journal of Contaminant Hydrology*, 116:24–34, 2010. doi: 10.1016/j.jconhyd.2010.05.004.

F. P. J. de Barros and Y. Rubin. A risk-driven approach for subsurface site characterization. *Water Resources Research*, 44(1):W01414, 2008. doi: 10.1029/2007WR006081.

F. P. J. de Barros and Y. Rubin. Modelling of block-scale macrodispersion as a random function. *Journal of Fluid Mechanics*, 676:514 545, 2011. doi: 10.1017/jfm.2011.65.

F. P. J. de Barros, Y. Rubin, and R. M. Maxwell. The concept of comparative information yield curves and its application to risk-based site characterization. *Water Resources Research.*, 45 (6), 2009. doi: doi:10.1029/2008WR007324. W06401.

F. P. J. de Barros, S. Ezzedine, and Y. Rubin. Impact of hydrogeological data on measures of uncertainty, site characterization and environmental performance metrics. *Advances in Water Resources*, 36:51–63, 2012. doi: 10.1016/j.advwatres.2011.05.004.

L. Debnath and D. Bhatta. *Integral transforms and their applications*. Chapman and Hall/CRC, US, 2nd edition, 2007.

M. Dentz and B. Berkowitz. Transport behavior of a passive solute in continuous time random walks and multirate mass transfer. *Water Resources Research*, 39(5):1111, 2003. doi: 10.1029/2001WR001163.

M. Dentz and D. M. Tartakovsky. Probability density functions for passive scalars dispersed in random velocity fields. *Geophysical Research Letters*, 37:L24406, 2010. doi: 10.1029/2010GL045748.

G. Destouni, K. Persson, C. Prieto, and J. Jarsjö. General quantification of catchment-scale nutrient and pollutant transport through the subsurface to surface and coastal waters. *Environmental Science & Technology*, 44(6):2048–2055, 2010. doi: 10.1021/es902338y.

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

P. J. Diggle and P. J. Ribeiro Jr. Bayesian inference in Gaussian model-based geostatistics. *Geographical and Environmental Modelling*, 6(2):129–146, 2002. doi: 10.1080/1361593022000029467.

P. J. Diggle and P. J. Ribeiro Jr. *Model-based geostatistics*. Springer series in statistics. Springer, New York, 1st edition, 2007.

R. Dykhuizen. A new coupling term for dual-porosity models. *Water Resources Research*, 26 (2):351–356, 1990. doi: 10.1029/WR026i002p00351.

B. Efron. *The Jackknife, the Bootstrap and Other Resampling Plans*, volume 1. Society for Industrial Mathematics, Philadelphia, 1st edition, 1982.

A. Englert, J. Vanderborght, and H. Vereecken. Prediction of velocity statistics in three-dimensional multi-Gaussian hydraulic conductivity fields. *Water Resources Research*, 42: 15, 2006. doi: 10.1029/2005WR004014. W03418.

R. Enzenhöfer, W. Nowak, and R. Helmig. Probabilistic Exposure Risk Assessment with Advective-Dispersive Well Vulnerability Criteria. *Advances in Water Resources*, 2011. doi: 10.1016/j.advwatres.2011.04.018. in press.

G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer, Heidelberg, 2nd edition, 2007.

V. V. Federov and P. Hackl. *Model-Oriented Design of Experiments*. Springer, New York, 1st edition, 1997.

D. Fernàndez-Garcia, G. Llerar-Meza, and J. J. Gómez-Hernández. Upscaling transport with mass transfer models: Mean behavior and propagation of uncertainty. *Water Resources Research*, 45(10), 2009. doi: 10.1029/2009WR007764.

C. W. Fetter. *Contaminant Hydrogeology*. Prentice Hall, Inc., Upper Saddle River, NJ, 2nd edition, 1999.

C. W. Fetter. *Applied Hydrogeology*. Prentice Hall, Inc., Upper Saddle River, NJ, 4th edition, 2001.

L. Feyen. A Bayesian approach to stochastic capture zone delineation incorporating tracer arrival times, conductivity measurements, and hydraulic head observations. *Water Resources Research.*, 39(5), 2003. doi: 10.1029/2002WR001544.

H. Fischer, E. List, R. Koh, J. Imberger, and N. Brooks. *Mixing in inland and coastal waters*. Academic Press, Ca, US, 1st edition, 1979.

H. J. H. Franssen, A. Alcolea, M. Riva, M. Bakr, N. van der Wiel, F. Stauffer, and A. Guadagnini. A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. *Advances in Water Resources*, 32:851–872, 2009. doi: doi:10.1016/j.advwatres.2009.02.011.

D. L. Freyberg. A natural gradient experiment on solute transport in a sand aquifer: 2. spatial moments and the advection and dispersion of nonreactive tracers. *Water Resources Research*, 22(13):2031–2046, 1986. doi: 10.1029/WR022i013p02031.

C. W. Gardiner. *Handbook of stochastic methods*. 1st. Springer, New York, US, 1985.

S. Geiger, A. Cortis, and J. T. Birkholzer. Upscaling solute transport in naturally fractured porous media with the continuous time random walk method. *Water Resources Research*, 46(12), 2010. doi: 10.1029/2010WR009133. W12530.

L. W. Gelhar and C. L. Axness. Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resources Research*, 19(1):161–180, 1983. doi: 10.1029/WR019i001p00161.

H. Gerke and M. van Genuchten. A dual-porosity model for simulating the preferential movement of water and solutes in structured porous media. *Water Resources Research*, 29 (2):305–319, 1993. doi: 10.1029/92WR02339.

M. N. Goltz and P. V. Roberts. Using the method of moments to analyze three-dimensional diffusion-limited solute transport from temporal and spatial perspectives. *Water Resources Research*, 23(8):1575–1585, 1987. doi: 10.1029/WR023i008p01575.

J. J. Gómez-Hernández and X.-H. Wen. To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. *Advances in Water Resources*, 21(1):47–61, 1998. doi: 10.1016/S0309-1708(96)00031-0.

J. J. Gómez-Hernández, A. Sahuquillo, and J. E. Capilla. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data - 1. Theory. *Journal of Hydrology*, 203(1-4):162–174, 1997. doi: 10.1016/S0022-1694(98)00138-3.

D. J. Goode. Direct simulation of groundwater age. *Water Resources Research.*, 32(2):289–296, 1996. ISSN 0043-1397. doi: 10.1029/95WR03401.

N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings-F*, 140(2):107–113, 1993. doi: 10.1049/ip-f-2.1993.0015.

R. Haggerty and S. M. Gorelick. Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity. *Water Resources Research*, 31(10): 2383–2400, 1995. doi: 10.1029/95WR01583.

R. Haggerty, S. A. McKenna, and L. C. Meigs. On the late-time behavior of tracer test breakthrough curves. *Water Resources Research*, 36(12):3467–3479, 2000. doi: 10.1029/2000WR900214.

M. S. Handcock and M. L. Stein. A Bayesian analysis of kriging. *American Statistical Association and American Society for Quality*, 35(4):403–410, 1993.

A. W. Harbaugh. *MODFLOW-2005, The US Geological Survey Modular Ground-water Model–the Ground-water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16*. US Geological Survey, 2005.

C. F. Harvey and S. M. Gorelick. Temporal moment-generating equations: Modeling transport and mass transfer in heterogeneous aquifers. *Water Resources Research.*, 31:1895–1911, 1995. doi: 10.1029/95WR01231.

C. P. Haslauer, P. Guthke, A. Bárdossy, and E. A. Sudicky. Effects of non-gaussian copula-based hydraulic conductivity fields on macrodispersion. *Water Resources Research*, 48:18, 2012. doi: 10.1029/2011WR011425. W07507.

G. S. Herrera and G. F. Pinder. Space-time optimization of groundwater quality sampling networks. *Water Resources Research.*, 41, 2005. doi: 10.1029/2004WR003626. W12407.

J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky. Bayesian model averaging: A tutorial. *Statistical Science*, 14(4):382–417, 1999.

K. J. Hollenbeck, C. F. Harvey, R. Haggerty, and C. J. Werth. A method for estimating distributions of mass transfer rate coefficients with application to purging and batch experiments. *Journal of Contaminant Hydrology*, 37(3-4):367–388, 1999. doi: 10.1016/S0169-7722(98)00165-X.

M. A. Hooimeijer. *Reduction of Complex Computational Models*. PhD thesis, Delft University of Technology, 2001.

P. S. Huyakorn, B. H. Lester, and J. W. Mercer. An efficient finite element technique for modeling transport in fractured porous media: 1. single species transport. *Water Resources Research*, 19(3):841–854, 1983. doi: 10.1029/WR019i003p00841.

J. W. Jawitz. Moments of truncated continuous univariate distributions. *Advances in Water Resources*, 27(3):269–281, 2004. doi: 10.1016/j.advwatres.2003.12.002.

E. T. Jaynes. Information theory and statistical mechanics. *The Physical Review*, 106(4):620–630, 1957. doi: 10.1103/PhysRev.106.620.

R. E. Kass and L. Wasserman. The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91:1343–1370, 1996.

M. G. Kendall and A. Stuart. *The advanced theory of statistics*. Griffin, London, 4th edition, 1977.

P. K. Kitanidis. Parameter uncertainty in estimation of spatial functions: Bayesian analysis. *Water Resources Research.*, 22(4):499–507, 1986. doi: 10.1029/WR022i004p00499.

P. K. Kitanidis. Quasi-linear geostatistical theory for inversing. *Water Resources Research.*, 31 (10):2411–2419, 1995. doi: 10.1029/95WR01945.

P. K. Kitanidis. *Introduction to geostatistics: Applications to hydrogeology*. Cambridge University Press, New York, 1st edition, 1997.

J. Koch and W. Nowak. A concept to implement Dirichlet boundary conditions in PTRW simulations. *Water Resources Research*, 2013. under review.

A. Kopp, H. Class, and R. Helmig. Investigations on CO2 storage capacity in saline aquifers–Part 2: Estimation of storage capacity coefficients. *International Journal of Greenhouse Gas Control*, 3(3):277–287, 2009. doi: 10.1016/j.ijggc.2008.10.001.

A. Kreft and A. Zuber. On the physical meaning of the dispersion equation and its solutions for different initial and boundary conditions. *Chemical Engineering Science*, 33(11):1471–1480, 1978. doi: 10.1016/0009-2509(78)85196-3.

R. Kubo. Generalized cumulant expansion method. *Journal of the Physical Society of Japan*, 17: 1100, 1962. doi: 10.1143/JPSJ.17.1100.

E. Kucěra. Contribution to the theory of chromatography: Linear non-equilibrium elution chromatography. *Journal of Chromatography A*, 19:237–248, 1965. doi: 10.1016/S0021-9673(01)99457-9.

H. Kunstmann, W. Kinzelbach, and T. Siegfried. Conditional first-order second-moment method and its application to the quantification of uncertainty in groundwater modeling. *Water Resources Research.*, 38(4)(1035), 2002. doi: 10.1029/2000WR000022.

A. E. Lawrence, X. Sanchez-Vila, and Y. Rubin. Conditional moments of the breakthrough curves of kinetically sorbing solute in heterogeneous porous media using multirate mass transfer models for sorption and desorption. *Water Resources Research*, 38(11):30, 2002. doi: doi:10.1029/2001WR001006.

P. C. Leube, A. Geiges, and W. Nowak. Bayesian assessment of the expected data impact on prediction confidence in optimal design. *Water Resources Research.*, 48(2), 2012a. doi: 10.1029/2010WR010137. W02501.

P. C. Leube, W. Nowak, and G. Schneider. Temporal Moments revisited: Why there is no better way for physically-based model reduction in time. *Water Resources Research*, 48(11): 14, 2012b. doi: 10.1029/2012WR011973. W11527.

P. C. Leube, W. Nowak, F. P. J. de Barros, and R. Rajagopal. Towards optimal allocation of computer resources: trade-offs between uncertainty quantification, discretization and model reduction. *Environmental Modelling & Software*, 50:97–107, 2013. doi: 10.1016/j.envsoft.2013.08.008.

L. Li, H. Zhou, and J. Gómez-Hernández. Transport upscaling using multi-rate mass transfer in three-dimensional highly heterogeneous porous media. *Advances in Water Resources*, 34 (4):478–489, 2011. doi: 10.1016/j.advwatres.2011.01.001.

S. G. Li, F. Ruan, and D. McLaughlin. A space-time accurate method for solving solute transport problems. *Water Resources Research*, 28(9):2297–2306, 1992. doi: 10.1029/92WR01009.

W. Li, W. Nowak, and O. A. Cirpka. Geostatistical inverse modeling of transient pumping tests using temporal moments of drawdown. *Water Resources Research.*, 41(8), 2005. doi: 10.1029/2004WR003874. W08403.

P. C. Lichtner and Q. Kang. Upscaling pore-scale reactive transport equations using a multiscale continuum formulation. *Water Resources Research*, 43(12), 2007. doi: 10.1029/2006WR005664. W12S15.

J. S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer, New York, 2008.

M. Loève. *Probability Theory*. Van Nostrand, Princeton, New Jersey, 1st edition, 1955.

J. Luo, O. A. Cirpka, and P. K. Kitanidis. Temporal-moment matching for truncated breakthrough curves for step or step-pulse injection. *Advances in Water Resources*, 29(9):1306–1313, 2006. doi: 10.1016/j.advwatres.2005.10.005.

J. Luo, O. A. Cirpka, M. Dentz, and J. Carrera. Temporal moments for transport with mass transfer described by an arbitrary memory function in heterogeneous media. *Water Resources Research*, 44(1), 2008. doi: 10.1029/2007WR006262. W01502.

A. Marzadri, D. Tonina, and A. Bellin. A semianalytical three-dimensional process-based model for hyporheic nitrogen dynamics in gravel bed rivers. *Water Resources Research*, 47 (11), 2011. doi: 10.1029/2011WR010583. W11518.

G. B. Matanga. Pseudopotential functions in construction of flow nets for contaminant transport modeling. *Water Resources Research*, 24(4):553–560, 1988. doi: 10.1029/WR024i004p00553.

B. Matérn. *Spatial Variation*. Springer, Berlin, Germany, 2nd edition, 1986.

R. M. Maxwell, F. K. Chow, and S. J. Kollet. The groundwater-land-surface-atmosphere connection: Soil moisture effects on the atmospheric boundary layer in fully-coupled simulations. *Advances in Water Resources*, 30(12):2447–2466, 2007. doi: 10.1016/j.advwatres.2007.05.018.

S. A. McKenna, L. C. Meigs, and R. Haggerty. Tracer tests in a fractured dolomite: 3. double-porosity, multiple-rate mass transfer processes in convergent flow tracer tests. *Water Resources Research*, 37(5):1143–1154, 2001. doi: 10.1029/2000WR900333.

J. McPhee and W. G. Y. William. Groundwater management using model reduction via empirical orthogonal functions. *Journal of Water Resources Planning and Management*, 134: 161, 2008. doi: 10.1061/(ASCE)0733-9496(2008)134:2(161).

J. W. Molson and E. O. Frind. On the use of mean groundwater age, life expectancy and capture probability for defining aquifer vulnerability and time-of-travel zones for source water protection. *Journal of Contaminant Hydrology*, 127:76–87, 2011. doi: 10.1016/j.jconhyd.2011.06.001.

W. G. Müller. *Collecting Spatial Data*. Springer, Berlin, Germany, 3rd edition, 2007.

H. Murakami, X. Chen, M. S. Hahn, Y. Liu, M. L. Rockhold, V. R. Vermeul, J. M. Zachara, and Y. Rubin. Bayesian approach for three-dimensional aquifer characterization at the Hanford 300 area. *Hydrology and Earth System Sciences*, 7(2):2017–2052, 2010.

T. Myers. Potential Contaminant Pathways from Hydraulically Fractured Shale to Aquifers. *Ground Water*, pages 1–11, 2012. doi: 10.1111/j.1745-6584.2012.00933.x.

J. E. Nash and J. V. Sutcliffe. River flow forecasting through conceptual models part 1: A discussion of principles. *Journal of Hydrology*, 10(3):282–290, 1970. doi: 10.1016/0022-1694(70)90255-6.

S. P. Neuman. Maximum likelihood Bayesian averaging of uncertain model predictions. *Stochastic Environmental Research and Risk Assessment*, 17(5):291–305, 2003. doi: 10.1007/s00477-003-0151-7.

S. P. Neuman, L. Xue, M. Ye, and D. Lu. Bayesian analysis of data-worth considering model and parameter uncertainties. *Advances in Water Resources*, 36:75–85, 2012. doi: 10.1016/j.advwatres.2011.02.007.

G. N. Newsam and C. R. Dietrich. Bounds on the size of nonnegative definite circulant embeddings of positive definite toeplitz matrices. *Information Theory, IEEE Transactions on*, 40(4):1218–1220, 1994. doi: 10.1109/18.335952.

W. Nowak. Best unbiased ensemble linearization and the quasi-linear Kalman ensemble generator. *Water Resources Research.*, 45(4), 2009. doi: 10.1029/2008WR007328. W04431.

W. Nowak and O. A. Cirpka. Geostatistical inference of hydraulic conductivity and dispersivities from hydraulic heads and tracer data. *Water Resources Research*, 42(8), 2006. doi: 10.1029/2005WR004832. W08416.

W. Nowak, R. Schwede, O. A. Cirpka, and I. Neuweiler. Probability density functions of hydraulic head and velocity in three-dimensional heterogeneous porous media. *Water Resources Research*, 44, 2008. doi: 10.1029/2007WR006383. W08452.

W. Nowak, F. P. J. de Barros, and Y. Rubin. Bayesian geostatistical design - task-driven optimal site investigation when the geostatistical model is uncertain. *Water Resources Research.*, 46, 2010. doi: 10.1029/2009WR008312. W03535.

W. Nowak, Y. Rubin, and F. P. Barros. A hypothesis-driven approach to optimize field campaigns. *Water Resources Research*, 48(6), 2012. doi: 10.1029/2011WR011016.

S. Oladyshkin, H. Class, R. Helmig, and W. Nowak. A concept for data-driven uncertainty quantification and its application to carbon dioxide storage in geological formations. *Advances in Water Resources*, 34(11):15081518, 2011. doi: 10.1016/j.advwatres.2011.08.005.

S. Olivella, J. Carrera, A. Gens, and E. Alonso. Nonisothermal multiphase flow of brine and gas through saline media. *Transport in Porous Media*, 15(3):271–293, 1994. doi: 10.1007/BF00613282.

D. S. Oliver. The influence of nonuniform transmissivity and storativity on drawdown. *Water Resources Research.*, 29(1):169–178, 1993. doi: 10.1029/92WR02061.

N. Oreskes, K. Shrader-Frechette, and K. Belitz. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263(5147):641–646, 1994. doi: 10.1126/science.263.5147.641.

S. Painter and V. Cvetkovic. Upscaling discrete fracture network simulations: An alternative to continuum transport models. *Water Resources Research*, 41(2), 2005. doi: 10.1029/2004WR003682. W02002.

A. Papoulis. *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 2nd edition, 1984.

K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6):559–572, 1901.

D. Pollock and O. A. Cirpka. Temporal moments in geoelectrical monitoring of salt tracer experiments. *Water Resources Research.*, 44(12), 2008. doi: 10.1029/2008WR007014. W12416.

D. W. Pollock. Simulation of fluid flow and energy transport processes associated with high-level radioactive waste disposal in unsaturated alluvium. *Water Resources Research*, 22(5): 765–775, 1986. doi: 10.1029/WR022i005p00765.

S. D. Priest and J. A. Hudson. Discontinuity spacings in rock. 13(5):135–148, 1976. doi: 10.1016/0148-9062(76)90818-4.

R. J. Prokop and A. P. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processing*, 54(5):438–460, 1992. doi: 10.1016/1049-9652(92)90027-U.

S. Razavi, B. A. Tolson, and D. H. Burn. Review of surrogate modeling in water resources. *Water Resources Research*, 48:32, 2012. doi: 10.1029/2011WR011527. W07401.

P. Reed, J. B. Kollat, and V. K. Devireddy. Using interactive archives in evolutionary multi-objective optimization: A case study for long-term groundwater monitoring design. *Environmental Modelling & Software*, 22(5):683–692, 2007. doi: 10.1016/j.envsoft.2005.12.021.

P. J. Roache. Quantification of uncertainty in computational fluid dynamics. *Annual Review of Fluid Mechanics*, 29(1):123–160, 1997. doi: 10.1146/annurev.fluid.29.1.123.

C. Robert and G. Casella. *Monte Carlo statistical methods*. Springer Verlag, New York, 2nd edition, 2004.

Y. Rubin. *Applied stochastic hydrogeology*. Oxford University Press, USA, New York, 1st edition, 2003.

Y. Rubin, A. Sun, R. Maxwell, and A. Bellin. The concept of block-effective macrodispersivity and a unified approach for grid-scale-and plume-scale-dependent transport. *Journal of Fluid Mechanics*, 395:161–180, 1999. doi: 10.1017/S0022112099005868.

P. Salamon, D. Fernàndez-Garcia, and J. J. Gómez-Hernández. A review and numerical assessment of the random walk particle tracking method. *Journal of Contaminant Hydrology*, 87(3):277–305, 2006. doi: 10.1016/j.jconhyd.2006.05.005.

A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global sensitivity analysis: the primer*. Wiley-Interscience, Cornwall, GB, 1st edition, 2008.

M. Sardin, D. Schweich, F. Leij, and M. T. Genuchten. Modeling the nonequilibrium transport of linearly interacting solutes in porous media: A review. *Water Resources Research*, 27 (9):2287–2307, 1991. doi: 10.1029/91WR01034.

A. E. Scheidegger. Statistical hydrodynamics in porous media. *Journal of Applied Physics*, 25(8):994–1001, 1954. doi: 10.1063/1.1721815.

G. A. Schmidt, R. Ruedy, J. E. Hansen, I. Aleinov, N. Bell, M. Bauer, S. Bauer, B. Cairns, V. Canuto, Y. Cheng, et al. Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data. *Journal of Climate*, 19(2):153–192, 2006. doi: 10.1175/JCLI3612.1.

A. Schöniger, W. Nowak, and H.-J. Franssen. Parameter estimation by Ensemble Kalman Filters with transformed data: Approach and application to hydraulic tomography. *Water Resources Research*, 48(4), 2012. doi: 10.1029/2011WR010462. W04502.

R. L. Schwede, O. A. Cirpka, W. Nowak, and I. Neuweiler. Impact of sampling volume on the probability density function of steady state concentration. *Water Resources Research*, 44 (12), 2008. doi: 10.1029/2007WR006668. W12433.

F. C. Schweppe. *Uncertain Dynamic Systems*. Prentice-Hall, Englewood Cliffs, NJ, 1st edition, 1973.

J. H. Seinfeld and S. N. Pandis. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. Wiley-Interscience, NY, US, 1st edition, 1997.

O. Silva, J. Carrera, M. Dentz, S. Kumar, A. Alcolea, and M. Willmann. A general real-time formulation for multi-rate mass transfer problems. *Hydrology and Earth System Sciences*, 13 (8):1399–1411, 2009. doi: 10.5194/hess-13-1399-2009.

B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, London, New York, 1st edition, 1986.

J. Simunek, N. J. Jarvis, M. T. Van Genuchten, and A. Gärdenäs. Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone. *Journal of Hydrology*, 272(1):14–35, 2003. doi: 10.1016/S0022-1694(02)00252-4.

C. Snyder, T. Bengtsson, P. Bickel, and J. Anderson. Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, 136(12):4629–4640, 2008. doi: 10.1175/2008MWR2529.1.

E. A. Sudicky. The Laplace transform Galerkin technique: A time-continuous finite element theory and application to mass transport in groundwater. *Water Resources Research.*, 25(8): 1833–1846, 1989. doi: 10.1029/WR025i008p01833.

E. A. Sudicky and R. G. McLaren. The Laplace transform Galerkin technique for large-scale simulation of mass transport in discretely fractured porous formations. *Water Resources Research*, 28(2):499–514, 1992. doi: 10.1029/91WR02560.

A. Tarantola. *Inverse problem theory: Methods for data fitting and model parameter estimation*. Elsevier Science, Amsterdam, Netherland, 1st edition, 1987.

D. M. Tartakovsky. Probabilistic risk analysis in subsurface hydrology. *Geophysical Research Letters*, 34(5), 2007. doi: 10.1029/2007GL029245.

M. R. Teague. Image analysis via the general theory of moments. *Journal of the Optical Society of America*, 70(8):920–930, 1980. doi: 10.1364/JOSA.70.000920.

D. Uciński. *Optimal Measurement Methods for Distributed Parameters System Identification*. CRC Press, Florida, USA, 1st edition, 2005.

A. J. Valocchi. Validity of the local equilibrium assumption for modeling sorbing solute transport through homogeneous soils. *Water Resources Research*, 21(6):808–820, 1985. doi: 10.1029/WR021i006p00808.

A. J. Valocchi. Spatial moment analysis of the transport of kinetically adsorbing solutes through stratified aquifers. *Water Resources Research*, 25(2):273–279, 1989. doi: 10.1029/WR025i002p00273.

A. J. Valocchi. Use of temporal moment analysis to study reactive solute transport in aggregated porous media. *Geoderma*, 46(1-3):233–247, 1990. doi: 10.1016/0016-7061(90)90017-4.

N. G. Van Kampen. *Stochastic processes in physics and chemistry*. North Holland, Amsterdam, NL, 3rd edition, 2007.

P. J. Van Leeuwen. Particle filtering in geophysical systems. *Monthly Weather Review*, 137 (12):4089–4114, 2009. doi: 10.1175/2009MWR2835.1.

M. Varni and J. Carrera. Simulation of groundwater age distributions. *Water Resources Research.*, 34(12):3271–3281, 1998. doi: 10.1029/98WR02536.

P. T. M. Vermeulen, A. W. Heemink, and C. B. M. Te Stroet. Reduced models for linear groundwater flow models using empirical orthogonal functions. *Advances in Water Resources*, 27(1):57–69, 2004. doi: 10.1016/j.advwatres.2003.09.008.

J. Villermaux. The chromatographic reactor. In A. . E. Rodrigues and D. . Tondeur, editors, *Percolation Processes: Theory and Applications*, number 33 in E, pages 539–588. Springer, 1981a.

J. Villermaux. Theory of linear chromatography. In A. . E. Rodrigues and D. . Tondeur, editors, *Percolation Processes: Theory and Applications*, number 33 in E, pages 83–140. Springer, 1981b.

J. Villermaux. Chemical engineering approach to dynamic modelling of linear chromatography : A flexible method for representing complex phenomena from simple concepts. *Journal of Chromatography A*, 406:11–26, 1987. doi: 10.1016/S0021-9673(00)94014-7.

J. E. Warren and P. J. Root. The behavior of naturally fractured reservoirs. *Old SPE Journal*, 3(3):245–255, 1963. doi: 10.2118/426-PA.

N. A. Weiss. *A Course in Probability*. Addison Wesley Longman, USA, 1st edition, 2006.

X.-H. Wen and J. J. Gómez-Hernández. Upscaling hydraulic conductivities in heterogeneous media: An overview. *Journal of Hydrology*, 183(1):ix–xxxii, 1996. doi: 10.1016/S0022-1694(96)80030-8.

C. J. Werth, O. A. Cirpka, and P. Grathwohl. Enhanced mixing and reaction through flow focusing in heterogeneous porous media. *Water Resources Research*, 42(12), 2006. doi: 10.1029/2005WR004511.

M. Willmann, J. Carrera, and X. Sánchez-Vila. Transport upscaling in heterogeneous aquifers: What physical parameters control memory functions? *Water Resources Research*, 44(12), 2008. doi: 10.1029/2007WR006531. W12437.

T. Wöhling, A. Geiges, W. Nowak, S. Gayler, P. Högy, and H. Wizemann. Towards Optimizing Experiments for Maximum-confidence Model Selection between Different Soil-plant Models. *Procedia Environmental Sciences*, 19:514523, 2013. doi: 10.1016/j.proenv.2013.06.058.

A. D. Woodbury and T. J. Ulrych. Minimum relative entropy: Forward probabilistic modeling. *Water Resources Research.*, 29(8):2847–2860, 1993. doi: 10.1029/93WR00923.

C. M. Wu, T. C. J. Yeh, J. Zhu, T. H. Lee, N. S. Hsu, C. H. Chen, and A. F. Sancho. Traditional analysis of aquifer tests: Comparing apples to oranges. *Water Resources Research*, 41(9), 2005. W09402.

T. C. J. Yeh, J. Xiang, R. M. Suribhatla, K. C. Hsu, C. H. Lee, and J. C. Wen. River stage tomography: A new approach for characterizing groundwater basins. *Water Resources Research*, 45(5), 2009. doi: 10.1029/2008WR007233. W05409.

D. Yin and W. A. Illman. Hydraulic tomography using temporal moments of drawdown recovery data: A laboratory sandbox study. *Water Resources Research.*, 45(1), 2009. doi: 10.1029/2007WR006623. W01502.

J. Zhu and T. C. J. Yeh. Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resources Research*, 41(7), 2005. doi: 10.1029/2004WR003790. W07028.

J. Zhu and T. C. J. Yeh. Analysis of hydraulic tomography using temporal moments of drawdown recovery data. *Water Resources Research*, 42, 2006. doi: 10.1029/2005WR004309. W02403.

R. W. Zimmerman, G. Chen, T. Hadgu, and G. S. Bodvarsson. A numerical dual-porosity model with semianalytical treatment of fracture/matrix flow. *Water Resources Research*, 29 (7):2127–2137, 1993. doi: 10.1029/93WR00749.

B. Zinn and C. F. Harvey. When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields. *Water Resources Research.*, 39(3)(1051), 2003. doi: 10.1029/2001WR001146.

A. Moment Generating Equation

The Laplace transform (LT) of a response curve $r(t)$ is expressed as:

$$L(r(t))_u = \int_0^\infty \mathrm{e}^{-ut}\, r(t)\, \mathrm{d}\, t, \tag{A.1}$$

where u is the Laplace variable. Inserting the Taylor expansion of e^{ut} about $t = 0$

$$\mathrm{e}^{ut} = \sum_{n=0}^{\infty} \frac{(ut)^k}{k!} = 1 + ut + \frac{u^2t^2}{2!} + ... + \frac{u^kt^k}{k!} \tag{A.2}$$

into Eq. (A.1) decomposes the spectral domain. Temporal moments (TM) $m_k^* = \int_0^\infty t^k r(t)\ \mathrm{d}\, t$ now occur as coefficients of the Tayler series and yield the so-called moment generating equation [Van Kampen, 2007]:

$$L(r(t)) = m_0^* + m_1^* u + \frac{m_2^* u^2}{2!} + ... + \frac{m_k^* u^k}{k!}. \tag{A.3}$$

Taking the derivative of $L(r(t))$ with respect to u, yields the k^{th} TM

$$\lim_{u \to 0} \left[\frac{\partial^k L(r(t))}{\partial u^k} \right] = m_k^*. \tag{A.4}$$

B. Cumulant Functions and their Relation to Temporal Moments

Applying the first natural logarithm to both sides of Eq. (A.2) and then performing just another Taylor expansion (for the logarithm) about $t = 0$ yields the cumulants κ_k and the associated cumulant function [Kubo, 1962]:

$$\log\, L(r(t)) = \kappa_1 u + \frac{\kappa_2 u^2}{2!} + ... + \frac{\kappa_k u^k}{k!}. \tag{B.1}$$

Cumulants can be expressed as a non-linear re-combination of equal- and lower order normalized TM [Kubo, 1962]:

$$\begin{aligned}
\kappa_1 &= m_1 \\
\kappa_2 &= m_2 - m_1^2 \\
\kappa_3 &= m_3 - 3\, m_2\, m_1 + 2\, m_1^3 \\
\kappa_4 &= m_4 - 4\, m_3\, m_1 - 3\, m_2^2 + 12\, m_2\, m_1^2 - 6\, m_1^4 \\
&\vdots \\
\kappa_k &= m_k - \sum_{j-1}^{k-1} \binom{k-1}{j-1} \kappa_j\, m_{k-j}
\end{aligned}$$

C. Pre-Posterior Data Impact Assessor Error

In order to assess the so-called Monte Carlo (MC)-error associated with the computed measures R, R_n and R_t in Chap. 6.2, I developed and implemented a swift procedure based on the well-known Bootstrapping method [Efron, 1982]. The general idea of the Bootstrapping method is strikingly simple. In order to arrive at the sampling distribution of the full ensemble (e.g., a set of ensemble mean values), it attempts to mimic the repetitive redrawing of the full ensemble by redrawing subsets of the (one) full ensemble.

The same mentality can be applied to the Pre-posterior Data Impact Assessor (PreDIA). Here, the most costly operation is the derivation of the $n \times m$ weight matrix W_{ij} with $i = 1 \dots n$ and $j = 1 \dots m$. Calculating W_{ij} once, allows to redraw B random subsets b from W_{ij} each of it with $i = 1 \dots n^b$ and $n^b \leq n$, and so opens the door to repetitive evaluations of PreDIA on B subsets.

$$CStD(K,t)^{(b)} \approx \sum_{i=1}^{n^b} r_i(t)^2 \; W_{ij} - \left(\sum_{i=1}^{n^b} r_i(t) \; W_{ij} \right)^2 . \tag{C.1}$$

If n^b is relatively large compared to n (which is typical for Bootstrapping) the scheme can further be sped up by calculating the terms $\sum_{i=1}^{n} r_i(t)^2 \; W_{ij}$ and $\sum_{i=1}^{n} r_i(t) \; W_{ij}$ for the full ensemble with $i = 1 \dots n$ once and deduct the respective quantities for the B so-called complementary subsets with $i = 1 \dots n^c$ and $n^c + n^b = n$ in order to arrive at $CStD(K,t)^{(b)}$. This, however, requires some reweighting and compensation for the squared terms in Eq. (C.1):

$$\begin{aligned} CStD(K,t)^{(b)} \approx & \left(\sum_{i=1}^{n} r_i(t)^2 \; W_{ij} - \sum_{i=1}^{n^c} r_i(t)^2 \; W_{ij} \right) v_j \\ & - \left(\sum_{i=1}^{n} r_i(t) \; W_{ij} \right)^2 v_j^2 - \left(\sum_{i=1}^{n^c} r_i(t) \; W_{ij} \right)^2 v_j^2 \\ & + 2 \left(\sum_{i=1}^{n} r_i(t) \; W_{ij} \right) \left(\sum_{i=1}^{n^c} r_i(t) \; W_{ij} \right) v_j^2, \end{aligned} \tag{C.2}$$

with reweighting term $v_j = \sum_{i=1}^{n} W_{ij} / (\sum_{i=1}^{n} W_{ij} - \sum_{i=1}^{n^c} W_{ij})$. Taking the variance about the spectrum of subsets B in $CStD(K,t)^{(b)}$ yields the MC-error associated with the ensemble of n potential realizations $\mathbf{r}(t)$ to be finally averaged about the set of m potential realizations of $\mathbf{m}_k^*$ (similar to Eq. (6.8):

$$Var_B[R(K,t)] \approx \left\{ \frac{1}{m} \sum_{j=1}^{m} \frac{v_{1,j}}{v_{1,j}^2 - v_{2,j}} \frac{1}{n^b} \sum_{1=b}^{B} \left(CStD(K,t)^{(b)} - \frac{1}{n^b} \sum_{1=b}^{B} CStD(K,t)^{(b)} \right)^2 \right\}^{1/2} \tag{C.3}$$

Institut für Wasser- und Umweltsystemmodellierung
Universität Stuttgart

Pfaffenwaldring 61
70569 Stuttgart (Vaihingen)
Telefon (0711) 685 - 64717/64749/64752/64679
Telefax (0711) 685 - 67020 o. 64746 o. 64681
E-Mail: iws@iws.uni-stuttgart.de
http://www.iws.uni-stuttgart.de

Verzeichnis der Mitteilungshefte

5 Plate, Erich: *Beitrag zur Bestimmung der Windgeschwindigkeitsverteilung in der durch eine Wand gestörten bodennahen Luftschicht*,
und
Röhnisch, Arthur; Marotz, Günter: *Neue Baustoffe und Bauausführungen für den Schutz der Böschungen und der Sohle von Kanälen, Flüssen und Häfen; Gestehungskosten und jeweilige Vorteile*, sowie Unny, T.E.: *Schwingungsuntersuchungen am Kegelstrahlschieber*, 1967

6 Seiler, Erich: *Die Ermittlung des Anlagenwertes der bundeseigenen Binnenschiffahrtsstraßen und Talsperren und des Anteils der Binnenschiffahrt an diesem Wert*, 1967

7 *Sonderheft anläßlich des 65. Geburtstages von Prof. Arthur Röhnisch mit Beiträgen von* Benk, Dieter; Breitling, J.; Gurr, Siegfried; Haberhauer, Robert; Honekamp, Hermann; Kuz, Klaus Dieter; Marotz, Günter; Mayer-Vorfelder, Hans-Jörg; Miller, Rudolf; Plate, Erich J.; Radomski, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1967

8 Jumikis, Alfred: *Beitrag zur experimentellen Untersuchung des Wassernachschubs in einem gefrierenden Boden und die Beurteilung der Ergebnisse*, 1968

9 Marotz, Günter: *Technische Grundlagen einer Wasserspeicherung im natürlichen Untergrund*, 1968

10 Radomski, Helge: *Untersuchungen über den Einfluß der Querschnittsform wellenförmiger Spundwände auf die statischen und rammtechnischen Eigenschaften*, 1968

11 Schwarz, Helmut: *Die Grenztragfähigkeit des Baugrundes bei Einwirkung vertikal gezogener Ankerplatten als zweidimensionales Bruchproblem*, 1969

12 Erbel, Klaus: *Ein Beitrag zur Untersuchung der Metamorphose von Mittelgebirgsschneedecken unter besonderer Berücksichtigung eines Verfahrens zur Bestimmung der thermischen Schneequalität*, 1969

13 Westhaus, Karl-Heinz: *Der Strukturwandel in der Binnenschiffahrt und sein Einfluß auf den Ausbau der Binnenschiffskanäle*, 1969

14 Mayer-Vorfelder, Hans-Jörg: *Ein Beitrag zur Berechnung des Erdwiderstandes unter Ansatz der logarithmischen Spirale als Gleitflächenfunktion*, 1970

15 Schulz, Manfred: *Berechnung des räumlichen Erddruckes auf die Wandung kreiszylindrischer Körper*, 1970

16 Mobasseri, Manoutschehr: *Die Rippenstützmauer. Konstruktion und Grenzen ihrer Standsicherheit*, 1970

17 Benk, Dieter: *Ein Beitrag zum Betrieb und zur Bemessung von Hochwasserrückhaltebecken*, 1970

18 Gàl, Attila: *Bestimmung der mitschwingenden Wassermasse bei überströmten Fischbauchklappen mit kreiszylindrischem Staublech*, 1971, vergriffen

19 Kuz, Klaus Dieter: *Ein Beitrag zur Frage des Einsetzens von Kavitationserscheinungen in einer Düsenströmung bei Berücksichtigung der im Wasser gelösten Gase*, 1971, vergriffen

20 Schaak, Hartmut: *Verteilleitungen von Wasserkraftanlagen*, 1971

21 *Sonderheft zur Eröffnung der neuen Versuchsanstalt des Instituts für Wasserbau der Universität Stuttgart mit Beiträgen von* Brombach, Hansjörg; Dirksen, Wolfram; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Holthoff, Franz-Josef; Kuz, Klaus Dieter; Marotz, Günter; Minor, Hans-Erwin; Petrikat, Kurt; Röhnisch, Arthur; Rueff, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1972

22 Wang, Chung-su: *Ein Beitrag zur Berechnung der Schwingungen an Kegelstrahlschiebern*, 1972

23 Mayer-Vorfelder, Hans-Jörg: *Erdwiderstandsbeiwerte nach dem Ohde-Variationsverfahren*, 1972

24 Minor, Hans-Erwin: *Beitrag zur Bestimmung der Schwingungsanfachungsfunktionen überströmter Stauklappen*, 1972, vergriffen

25 Brombach, Hansjörg: *Untersuchung strömungsmechanischer Elemente (Fluidik) und die Möglichkeit der Anwendung von Wirbelkammerelementen im Wasserbau*, 1972, vergriffen

26 Wildenhahn, Eberhard: *Beitrag zur Berechnung von Horizontalfilterbrunnen*, 1972

27 Steinlein, Helmut: *Die Eliminierung der Schwebstoffe aus Flußwasser zum Zweck der unterirdischen Wasserspeicherung, gezeigt am Beispiel der Iller*, 1972

28 Holthoff, Franz Josef: *Die Überwindung großer Hubhöhen in der Binnenschiffahrt durch Schwimmerhebewerke*, 1973

29 Röder, Karl: *Einwirkungen aus Baugrundbewegungen auf trog- und kastenförmige Konstruktionen des Wasser- und Tunnelbaues*, 1973

30 Kretschmer, Heinz: *Die Bemessung von Bogenstaumauern in Abhängigkeit von der Talform*, 1973

31 Honekamp, Hermann: *Beitrag zur Berechnung der Montage von Unterwasserpipelines*, 1973

32 Giesecke, Jürgen: *Die Wirbelkammertriode als neuartiges Steuerorgan im Wasserbau*, und Brombach, Hansjörg: *Entwicklung, Bauformen, Wirkungsweise und Steuereigenschaften von Wirbelkammerverstärkern*, 1974

33 Rueff, Helge: *Untersuchung der schwingungserregenden Kräfte an zwei hintereinander angeordneten Tiefschützen unter besonderer Berücksichtigung von Kavitation*, 1974

34 Röhnisch, Arthur: *Einpreßversuche mit Zementmörtel für Spannbeton - Vergleich der Ergebnisse von Modellversuchen mit Ausführungen in Hüllwellrohren*, 1975

35 *Sonderheft anläßlich des 65. Geburtstages von Prof. Dr.-Ing. Kurt Petrikat mit Beiträgen von*: Brombach, Hansjörg; Erbel, Klaus; Flinspach, Dieter; Fischer jr., Richard; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Haberhauer, Robert; Hafner Edzard; Hausenblas, Bernhard; Horlacher, Hans-Burkhard; Hutarew, Andreas; Knoll, Manfred; Krummet, Ralph; Marotz, Günter; Merkle, Theodor; Miller, Christoph; Minor, Hans-Erwin; Neumayer, Hans; Rao, Syamala; Rath, Paul; Rueff, Helge; Ruppert, Jürgen; Schwarz, Wolfgang; Topal-Gökceli, Mehmet; Vollmer, Ernst; Wang, Chung-su; Weber, Hans-Georg; 1975

36 Berger, Jochum: *Beitrag zur Berechnung des Spannungszustandes in rotationssymmetrisch belasteten Kugelschalen veränderlicher Wandstärke unter Gas- und Flüssigkeitsdruck durch Integration schwach singulärer Differentialgleichungen*, 1975

37 Dirksen, Wolfram: *Berechnung instationärer Abflußvorgänge in gestauten Gerinnen mittels Differenzenverfahren und die Anwendung auf Hochwasserrückhaltebecken*, 1976

38 Horlacher, Hans-Burkhard: *Berechnung instationärer Temperatur- und Wärmespannungsfelder in langen mehrschichtigen Hohlzylindern*, 1976

39 Hafner, Edzard: *Untersuchung der hydrodynamischen Kräfte auf Baukörper im Tiefwasserbereich des Meeres*, 1977, ISBN 3-921694-39-6

40 Ruppert, Jürgen: *Über den Axialwirbelkammerverstärker für den Einsatz im Wasserbau*, 1977, ISBN 3-921694-40-X

41 Hutarew, Andreas: *Beitrag zur Beeinflußbarkeit des Sauerstoffgehalts in Fließgewässern an Abstürzen und Wehren*, 1977, ISBN 3-921694-41-8, vergriffen

42 Miller, Christoph: *Ein Beitrag zur Bestimmung der schwingungserregenden Kräfte an unterströmten Wehren*, 1977, ISBN 3-921694-42-6

43 Schwarz, Wolfgang: *Druckstoßberechnung unter Berücksichtigung der Radial- und Längsverschiebungen der Rohrwandung*, 1978, ISBN 3-921694-43-4

44 Kinzelbach, Wolfgang: *Numerische Untersuchungen über den optimalen Einsatz variabler Kühlsysteme einer Kraftwerkskette am Beispiel Oberrhein*, 1978, ISBN 3-921694-44-2

45 Barczewski, Baldur: *Neue Meßmethoden für Wasser-Luftgemische und deren Anwendung auf zweiphasige Auftriebsstrahlen*, 1979, ISBN 3-921694-45-0

46 Neumayer, Hans: *Untersuchung der Strömungsvorgänge in radialen Wirbelkammerverstärkern*, 1979, ISBN 3-921694-46-9

47 Elalfy, Youssef-Elhassan: *Untersuchung der Strömungsvorgänge in Wirbelkammerdioden und -drosseln*, 1979, ISBN 3-921694-47-7

48 Brombach, Hansjörg: *Automatisierung der Bewirtschaftung von Wasserspeichern*, 1981, ISBN 3-921694-48-5

49 Geldner, Peter: *Deterministische und stochastische Methoden zur Bestimmung der Selbstdichtung von Gewässern*, 1981, ISBN 3-921694-49-3, vergriffen

50 Mehlhorn, Hans: *Temperaturveränderungen im Grundwasser durch Brauchwassereinleitungen*, 1982, ISBN 3-921694-50-7, vergriffen

51 Hafner, Edzard: *Rohrleitungen und Behälter im Meer*, 1983, ISBN 3-921694-51-5

52 Rinnert, Bernd: *Hydrodynamische Dispersion in porösen Medien: Einfluß von Dichteunterschieden auf die Vertikalvermischung in horizontaler Strömung*, 1983, ISBN 3-921694-52-3, vergriffen

53 Lindner, Wulf: *Steuerung von Grundwasserentnahmen unter Einhaltung ökologischer Kriterien*, 1983, ISBN 3-921694-53-1, vergriffen

54 Herr, Michael; Herzer, Jörg; Kinzelbach, Wolfgang; Kobus, Helmut; Rinnert, Bernd: *Methoden zur rechnerischen Erfassung und hydraulischen Sanierung von Grundwasserkontaminationen*, 1983, ISBN 3-921694-54-X

55 Schmitt, Paul: *Wege zur Automatisierung der Niederschlagsermittlung*, 1984, ISBN 3-921694-55-8, vergriffen

56 Müller, Peter: *Transport und selektive Sedimentation von Schwebstoffen bei gestautem Abfluß*, 1985, ISBN 3-921694-56-6

57 El-Qawasmeh, Fuad: *Möglichkeiten und Grenzen der Tropfbewässerung unter besonderer Berücksichtigung der Verstopfungsanfälligkeit der Tropfelemente*, 1985, ISBN 3-921694-57-4, vergriffen

58 Kirchenbaur, Klaus: *Mikroprozessorgesteuerte Erfassung instationärer Druckfelder am Beispiel seegangsbelasteter Baukörper*, 1985, ISBN 3-921694-58-2

59 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1984/85 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart), 1985, ISBN 3-921694-59-0, vergriffen

60 Spitz, Karlheinz: *Dispersion in porösen Medien: Einfluß von Inhomogenitäten und Dichteunterschieden*, 1985, ISBN 3-921694-60-4, vergriffen

61 Kobus, Helmut: *An Introduction to Air-Water Flows in Hydraulics*, 1985, ISBN 3-921694-61-2

62 Kaleris, Vassilios: *Erfassung des Austausches von Oberflächen- und Grundwasser in horizontalebenen Grundwassermodellen*, 1986, ISBN 3-921694-62-0

63 Herr, Michael: *Grundlagen der hydraulischen Sanierung verunreinigter Porengrundwasserleiter*, 1987, ISBN 3-921694-63-9

64 Marx, Walter: *Berechnung von Temperatur und Spannung in Massenbeton infolge Hydratation*, 1987, ISBN 3-921694-64-7

65 Koschitzky, Hans-Peter: *Dimensionierungskonzept für Sohlbelüfter in Schußrinnen zur Vermeidung von Kavitationsschäden*, 1987, ISBN 3-921694-65-5

66 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1986/87 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart) 1987, ISBN 3-921694-66-3

67 Söll, Thomas: *Berechnungsverfahren zur Abschätzung anthropogener Temperaturanomalien im Grundwasser*, 1988, ISBN 3-921694-67-1

68 Dittrich, Andreas; Westrich, Bernd: *Bodenseeufererosion, Bestandsaufnahme und Bewertung*, 1988, ISBN 3-921694-68-X, vergriffen

69 Huwe, Bernd; van der Ploeg, Rienk R.: *Modelle zur Simulation des Stickstoffhaushaltes von Standorten mit unterschiedlicher landwirtschaftlicher Nutzung*, 1988, ISBN 3-921694-69-8, vergriffen

70 Stephan, Karl: *Integration elliptischer Funktionen*, 1988, ISBN 3-921694-70-1

71 Kobus, Helmut; Zilliox, Lothaire (Hrsg.): *Nitratbelastung des Grundwassers, Auswirkungen der Landwirtschaft auf die Grundwasser- und Rohwasserbeschaffenheit und Maßnahmen zum Schutz des Grundwassers*. Vorträge des deutsch-französischen Kolloquiums am 6. Oktober 1988, Universitäten Stuttgart und Louis Pasteur Strasbourg (Vorträge in deutsch oder französisch, Kurzfassungen zweisprachig), 1988, ISBN 3-921694-71-X

72 Soyeaux, Renald: *Unterströmung von Stauanlagen auf klüftigem Untergrund unter Berücksichtigung laminarer und turbulenter Fließzustände*,1991, ISBN 3-921694-72-8

73 Kohane, Roberto: *Berechnungsmethoden für Hochwasserabfluß in Fließgewässern mit überströmten Vorländern*, 1991, ISBN 3-921694-73-6

74 Hassinger, Reinhard: *Beitrag zur Hydraulik und Bemessung von Blocksteinrampen in flexibler Bauweise*, 1991, ISBN 3-921694-74-4, vergriffen

75 Schäfer, Gerhard: *Einfluß von Schichtenstrukturen und lokalen Einlagerungen auf die Längsdispersion in Porengrundwasserleitern*, 1991, ISBN 3-921694-75-2

76 Giesecke, Jürgen: *Vorträge, Wasserwirtschaft in stark besiedelten Regionen; Umweltforschung mit Schwerpunkt Wasserwirtschaft*, 1991, ISBN 3-921694-76-0

77 Huwe, Bernd: *Deterministische und stochastische Ansätze zur Modellierung des Stickstoffhaushalts landwirtschaftlich genutzter Flächen auf unterschiedlichem Skalenniveau*, 1992, ISBN 3-921694-77-9, vergriffen

78 Rommel, Michael: *Verwendung von Kluftdaten zur realitätsnahen Generierung von Kluftnetzen mit anschließender laminar-turbulenter Strömungsberechnung*, 1993, ISBN 3-92 1694-78-7

79 Marschall, Paul: *Die Ermittlung lokaler Stofffrachten im Grundwasser mit Hilfe von Einbohrloch-Meßverfahren*, 1993, ISBN 3-921694-79-5, vergriffen

80 Ptak, Thomas: *Stofftransport in heterogenen Porenaquiferen: Felduntersuchungen und stochastische Modellierung*, 1993, ISBN 3-921694-80-9, vergriffen

81 Haakh, Frieder: *Transientes Strömungsverhalten in Wirbelkammern,* 1993, ISBN 3-921694-81-7

82 Kobus, Helmut; Cirpka, Olaf; Barczewski, Baldur; Koschitzky, Hans-Peter: *Versucheinrichtung zur Grundwasser und Altlastensanierung VEGAS, Konzeption und Programmrahmen*, 1993, ISBN 3-921694-82-5

83 Zang, Weidong: *Optimaler Echtzeit-Betrieb eines Speichers mit aktueller Abflußregenerierung*, 1994, ISBN 3-921694-83-3, vergriffen

84 Franke, Hans-Jörg: *Stochastische Modellierung eines flächenhaften Stoffeintrages und Transports in Grundwasser am Beispiel der Pflanzenschutzmittelproblematik*, 1995, ISBN 3-921694-84-1

85 Lang, Ulrich: *Simulation regionaler Strömungs- und Transportvorgänge in Karstaquiferen mit Hilfe des Doppelkontinuum-Ansatzes: Methodenentwicklung und Parameteridentifikation*, 1995, ISBN 3-921694-85-X, vergriffen

86 Helmig, Rainer: *Einführung in die Numerischen Methoden der Hydromechanik*, 1996, ISBN 3-921694-86-8, vergriffen

87 Cirpka, Olaf: *CONTRACT: A Numerical Tool for Contaminant Transport and Chemical Transformations - Theory and Program Documentation* -, 1996, ISBN 3-921694-87-6

88 Haberlandt, Uwe: *Stochastische Synthese und Regionalisierung des Niederschlages für Schmutzfrachtberechnungen*, 1996, ISBN 3-921694-88-4

89 Croisé, Jean: *Extraktion von flüchtigen Chemikalien aus natürlichen Lockergesteinen mittels erzwungener Luftströmung*, 1996, ISBN 3-921694-89-2, vergriffen

90 Jorde, Klaus: *Ökologisch begründete, dynamische Mindestwasserregelungen bei Ausleitungskraftwerken*, 1997, ISBN 3-921694-90-6, vergriffen

91 Helmig, Rainer: *Gekoppelte Strömungs- und Transportprozesse im Untergrund - Ein Beitrag zur Hydrosystemmodellierung-*, 1998, ISBN 3-921694-91-4, vergriffen

92 Emmert, Martin: *Numerische Modellierung nichtisothermer Gas-Wasser Systeme in porösen Medien,* 1997, ISBN 3-921694-92-2

93 Kern, Ulrich: *Transport von Schweb- und Schadstoffen in staugeregelten Fließgewässern am Beispiel des Neckars*, 1997, ISBN 3-921694-93-0, vergriffen

94 Förster, Georg: *Druckstoßdämpfung durch große Luftblasen in Hochpunkten von Rohrleitungen 1997*, ISBN 3-921694-94-9

95 Cirpka, Olaf: *Numerische Methoden zur Simulation des reaktiven Mehrkomponententransports im Grundwasser*, 1997, ISBN 3-921694-95-7, vergriffen

96 Färber, Arne: *Wärmetransport in der ungesättigten Bodenzone: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1997, ISBN 3-921694-96-5

97 Betz, Christoph: *Wasserdampfdestillation von Schadstoffen im porösen Medium: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1998, ISBN 3-921694-97-3

98 Xu, Yichun: *Numerical Modeling of Suspended Sediment Transport in Rivers*, 1998, ISBN 3-921694-98-1, vergriffen

99 Wüst, Wolfgang: *Geochemische Untersuchungen zur Sanierung CKW-kontaminierter Aquifere mit Fe(0)-Reaktionswänden,* 2000, ISBN 3-933761-02-2

100 Sheta, Hussam: *Simulation von Mehrphasenvorgängen in porösen Medien unter Einbeziehung von Hysterese-Effekten*, 2000, ISBN 3-933761-03-4

101 Ayros, Edwin: *Regionalisierung extremer Abflüsse auf der Grundlage statistischer Verfahren*, 2000, ISBN 3-933761-04-2, vergriffen

102 Huber, Ralf: *Compositional Multiphase Flow and Transport in Heterogeneous Porous Media*, 2000, ISBN 3-933761-05-0

103 Braun, Christopherus: *Ein Upscaling-Verfahren für Mehrphasenströmungen in porösen Medien*, 2000, ISBN 3-933761-06-9

104 Hofmann, Bernd: *Entwicklung eines rechnergestützten Managementsystems zur Beurteilung von Grundwasserschadensfällen*, 2000, ISBN 3-933761-07-7

105 Class, Holger: *Theorie und numerische Modellierung nichtisothermer Mehrphasenprozesse in NAPL-kontaminierten porösen Medien,* 2001, ISBN 3-933761-08-5

106 Schmidt, Reinhard: *Wasserdampf- und Heißluftinjektion zur thermischen Sanierung kontaminierter Standorte*, 2001, ISBN 3-933761-09-3

107 Josef, Reinhold:, *Schadstoffextraktion mit hydraulischen Sanierungsverfahren unter Anwendung von grenzflächenaktiven Stoffen*, 2001, ISBN 3-933761-10-7

108 Schneider, Matthias: *Habitat- und Abflussmodellierung für Fließgewässer mit unscharfen Berechnungsansätzen*, 2001, ISBN 3-933761-11-5

109 Rathgeb, Andreas: *Hydrodynamische Bemessungsgrundlagen für Lockerdeckwerke an überströmbaren Erddämmen*, 2001, ISBN 3-933761-12-3

110 Lang, Stefan: *Parallele numerische Simulation instätionärer Probleme mit adaptiven Methoden auf unstrukturierten Gittern*, 2001, ISBN 3-933761-13-1

111 Appt, Jochen; Stumpp Simone: *Die Bodensee-Messkampagne 2001, IWS/CWR Lake Constance Measurement Program 2001*, 2002, ISBN 3-933761-14-X

112 Heimerl, Stephan: *Systematische Beurteilung von Wasserkraftprojekten,* 2002, ISBN 3-933761-15-8, vergriffen

113 Iqbal, Amin: *On the Management and Salinity Control of Drip Irrigation,* 2002, ISBN 3-933761-16-6

114 Silberhorn-Hemminger, Annette: *Modellierung von Kluftaquifersystemen: Geostatistische Analyse und deterministisch-stochastische Kluftgenerierung,* 2002, ISBN 3-933761-17-4

115 Winkler, Angela: *Prozesse des Wärme- und Stofftransports bei der In-situ-Sanierung mit festen Wärmequellen,* 2003, ISBN 3-933761-18-2

116 Marx, Walter: *Wasserkraft, Bewässerung, Umwelt - Planungs- und Bewertungsschwerpunkte der Wasserbewirtschaftung,* 2003, ISBN 3-933761-19-0

117 Hinkelmann, Reinhard: *Efficient Numerical Methods and Information-Processing Techniques in Environment Water,* 2003, ISBN 3-933761-20-4

118 Samaniego-Eguiguren, Luis Eduardo: *Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments*, 2003, ISBN 3-933761-21-2

119 Neunhäuserer, Lina: *Diskretisierungsansätze zur Modellierung von Strömungs- und Transportprozessen in geklüftet-porösen Medien*, 2003, ISBN 3-933761-22-0

120 Paul, Maren: *Simulation of Two-Phase Flow in Heterogeneous Poros Media with Adaptive Methods*, 2003, ISBN 3-933761-23-9

121 Ehret, Uwe: *Rainfall and Flood Nowcasting in Small Catchments using Weather Radar*, 2003, ISBN 3-933761-24-7

122 Haag, Ingo: *Der Sauerstoffhaushalt staugeregelter Flüsse am Beispiel des Neckars - Analysen, Experimente, Simulationen -,* 2003, ISBN 3-933761-25-5

123 Appt, Jochen: *Analysis of Basin-Scale Internal Waves in Upper Lake Constance,* 2003, ISBN 3-933761-26-3

124 Hrsg.: Schrenk, Volker; Batereau, Katrin; Barczewski, Baldur; Weber, Karolin und Koschitzky, Hans-Peter: *Symposium Ressource Fläche und VEGAS - Statuskolloquium 2003, 30. September und 1. Oktober 2003,* 2003, ISBN 3-933761-27-1

125 Omar Khalil Ouda: *Optimisation of Agricultural Water Use: A Decision Support System for the Gaza Strip,* 2003, ISBN 3-933761-28-0

126 Batereau, Katrin: *Sensorbasierte Bodenluftmessung zur Vor-Ort-Erkundung von Schadensherden im Untergrund*, 2004, ISBN 3-933761-29-8

127 Witt, Oliver: *Erosionsstabilität von Gewässersedimenten mit Auswirkung auf den Stofftransport bei Hochwasser am Beispiel ausgewählter Stauhaltungen des Oberrheins*, 2004, ISBN 3-933761-30-1

128 Jakobs, Hartmut: *Simulation nicht-isothermer Gas-Wasser-Prozesse in komplexen Kluft-Matrix-Systemen*, 2004, ISBN 3-933761-31-X

129 Li, Chen-Chien: *Deterministisch-stochastisches Berechnungskonzept zur Beurteilung der Auswirkungen erosiver Hochwasserereignisse in Flussstauhaltungen,* 2004, ISBN 3-933761-32-8

130 Reichenberger, Volker; Helmig, Rainer; Jakobs, Hartmut; Bastian, Peter; Niessner, Jennifer: *Complex Gas-Water Processes in Discrete Fracture-Matrix Systems: Upscaling, Mass-Conservative Discretization and Efficient Multilevel Solution,* 2004, ISBN 3-933761-33-6

131 Hrsg.: Barczewski, Baldur; Koschitzky, Hans-Peter; Weber, Karolin; Wege, Ralf: *VEGAS - Statuskolloquium 2004,* Tagungsband zur Veranstaltung am 05. Oktober 2004 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2004, ISBN 3-933761-34-4

132 Asie, Kemal Jabir: *Finite Volume Models for Multiphase Multicomponent Flow through Porous Media*. 2005, ISBN 3-933761-35-2

133 Jacoub, George: *Development of a 2-D Numerical Module for Particulate Contaminant Transport in Flood Retention Reservoirs and Impounded Rivers,* 2004, ISBN 3-933761-36-0

134 Nowak, Wolfgang: *Geostatistical Methods for the Identification of Flow and Transport Parameters in the Subsurface,* 2005, ISBN 3-933761-37-9

135 Süß, Mia: *Analysis of the influence of structures and boundaries on flow and transport processes in fractured porous media*, 2005, ISBN 3-933761-38-7

136 Jose, Surabhin Chackiath: *Experimental Investigations on Longitudinal Dispersive Mixing in Heterogeneous Aquifers,* 2005, ISBN: 3-933761-39-5

137 Filiz, Fulya: *Linking Large-Scale Meteorological Conditions to Floods in Mesoscale Catchments,* 2005, ISBN 3-933761-40-9

138 Qin, Minghao: *Wirklichkeitsnahe und recheneffiziente Ermittlung von Temperatur und Spannungen bei großen RCC-Staumauern*, 2005, ISBN 3-933761-41-7

139 Kobayashi, Kenichiro: *Optimization Methods for Multiphase Systems in the Subsurface - Application to Methane Migration in Coal Mining Areas*, 2005, ISBN 3-933761-42-5

140 Rahman, Md. Arifur: *Experimental Investigations on Transverse Dispersive Mixing in Heterogeneous Porous Media,* 2005, ISBN 3-933761-43-3

141 Schrenk, Volker: *Ökobilanzen zur Bewertung von Altlastensanierungsmaßnahmen,* 2005, ISBN 3-933761-44-1

142 Hundecha, Hirpa Yeshewatesfa: *Regionalization of Parameters of a Conceptual Rainfall-Runoff Model*, 2005, ISBN: 3-933761-45-X

143 Wege, Ralf: *Untersuchungs- und Überwachungsmethoden für die Beurteilung natürlicher Selbstreinigungsprozesse im Grundwasser,* 2005, ISBN 3-933761-46-8

144 Breiting, Thomas: *Techniken und Methoden der Hydroinformatik - Modellierung von komplexen Hydrosystemen im Untergrund*, 2006, 3-933761-47-6

145 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Müller, Martin: *Ressource Untergrund: 10 Jahre VEGAS: Forschung und Technologieentwicklung zum Schutz von Grundwasser und Boden,* Tagungsband zur Veranstaltung am 28. und 29. September 2005 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2005, ISBN 3-933761-48-4

146 Rojanschi, Vlad: *Abflusskonzentration in mesoskaligen Einzugsgebieten unter Berücksichtigung des Sickerraumes,* 2006, ISBN 3-933761-49-2

147 Winkler, Nina Simone: *Optimierung der Steuerung von Hochwasserrückhaltebecken-systemen,* 2006, ISBN 3-933761-50-6

148 Wolf, Jens: *Räumlich differenzierte Modellierung der Grundwasserströmung alluvialer Aquifere für mesoskalige Einzugsgebiete*, 2006, ISBN: 3-933761-51-4

149 Kohler, Beate: *Externe Effekte der Laufwasserkraftnutzung*, 2006, ISBN 3-933761-52-2

150 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS-Statuskolloquium 2006,* Tagungsband zur Veranstaltung am 28. September 2006 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2006, ISBN 3-933761-53-0

151 Niessner, Jennifer: *Multi-Scale Modeling of Multi-Phase - Multi-Component Processes in Heterogeneous Porous Media*, 2006, ISBN 3-933761-54-9

152 Fischer, Markus: *Beanspruchung eingeerdeter Rohrleitungen infolge Austrocknung bindiger Böden,* 2006, ISBN 3-933761-55-7

153 Schneck, Alexander: *Optimierung der Grundwasserbewirtschaftung unter Berücksichtigung der Belange der Wasserversorgung, der Landwirtschaft und des Naturschutzes*, 2006, ISBN 3-933761-56-5

154 Das, Tapash: *The Impact of Spatial Variability of Precipitation on the Predictive Uncertainty of Hydrological Models,* 2006, ISBN 3-933761-57-3

155 Bielinski, Andreas: *Numerical Simulation of CO_2 sequestration in geological formations*, 2007, ISBN 3-933761-58-1

156 Mödinger, Jens: *Entwicklung eines Bewertungs- und Entscheidungsunterstützungssystems für eine nachhaltige regionale Grundwasserbewirtschaftung*, 2006, ISBN 3-933761-60-3

157 Manthey, Sabine: *Two-phase flow processes with dynamic effects in porous media - parameter estimation and simulation,* 2007, ISBN 3-933761-61-1

158 Pozos Estrada, Oscar: *Investigation on the Effects of Entrained Air in Pipelines,* 2007, ISBN 3-933761-62-X

159 Ochs, Steffen Oliver: *Steam injection into saturated porous media – process analysis including experimental and numerical investigations*, 2007, ISBN 3-933761-63-8

160 Marx, Andreas: *Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage, 2007,* ISBN 3-933761-64-6

161 Hartmann, Gabriele Maria: *Investigation of Evapotranspiration Concepts in Hydrological Modelling for Climate Change Impact Assessment*, 2007, ISBN 3-933761-65-4

162 Kebede Gurmessa, Tesfaye: *Numerical Investigation on Flow and Transport Characteristics to Improve Long-Term Simulation of Reservoir Sedimentation*, 2007, ISBN 3-933761-66-2

163 Trifković, Aleksandar: *Multi-objective and Risk-based Modelling Methodology for Planning, Design and Operation of Water Supply Systems*, 2007, ISBN 3-933761-67-0

164 Götzinger, Jens: *Distributed Conceptual Hydrological Modelling - Simulation of Climate, Land Use Change Impact and Uncertainty Analysis*, 2007, ISBN 3-933761-68-9

165 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS – Kolloquium 2007,* Tagungsband zur Veranstaltung am 26. September 2007 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2007, ISBN 3-933761-69-7

166 Freeman, Beau: *Modernization Criteria Assessment for Water Resources Planning; Klamath Irrigation Project, U.S.*, 2008, ISBN 3-933761-70-0

167 Dreher, Thomas: *Selektive Sedimentation von Feinstschwebstoffen in Wechselwirkung mit wandnahen turbulenten Strömungsbedingungen*, 2008, ISBN 3-933761-71-9

168 Yang, Wei: *Discrete-Continuous Downscaling Model for Generating Daily Precipitation Time Series*, 2008, ISBN 3-933761-72-7

169 Kopecki, Ianina: *Calculational Approach to FST-Hemispheres for Multiparametrical Benthos Habitat Modelling,* 2008*,* ISBN 3-933761-73-5

170 Brommundt, Jürgen: *Stochastische Generierung räumlich zusammenhängender Niederschlagszeitreihen*, 2008, ISBN 3-933761-74-3

171 Papafotiou, *Alexandros: Numerical Investigations of the Role of Hysteresis in Heterogeneous Two-Phase Flow Systems*, 2008, ISBN 3-933761-75-1

172 He, Yi: *Application of a Non-Parametric Classification Scheme to Catchment Hydrology*, 2008, ISBN 978-3-933761-76-7

173 Wagner, Sven: *Water Balance in a Poorly Gauged Basin in West Africa Using Atmospheric Modelling and Remote Sensing Information*, 2008, ISBN 978-3-933761-77-4

174 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias; Schrenk, Volker: *VEGAS-Kolloquium 2008 Ressource Fläche III*, Tagungsband zur Veranstaltung am 01. Oktober 2008 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2008, ISBN 978-3-933761-78-1

175 Patil, Sachin: *Regionalization of an Event Based Nash Cascade Model for Flood Predictions in Ungauged Basins*, 2008, ISBN 978-3-933761-79-8

176 Assteerawatt, Anongnart: *Flow and Transport Modelling of Fractured Aquifers based on a Geostatistical Approach,* 2008, ISBN 978-3-933761-80-4

177 Karnahl, Joachim Alexander: *2D numerische Modellierung von multifraktionalem Schwebstoff- und Schadstofftransport in Flüssen*, 2008, ISBN 978-3-933761-81-1

178 Hiester, Uwe: *Technologieentwicklung zur In-situ-Sanierung der ungesättigten Bodenzone mit festen Wärmequellen,* 2009, ISBN 978-3-933761-82-8

179 Laux, Patrick: *Statistical Modeling of Precipitation for Agricultural Planning in the Volta Basin of West Africa,* 2009, ISBN 978-3-933761-83-5

180 Ehsan, Saqib: *Evaluation of Life Safety Risks Related to Severe Flooding,* 2009, ISBN 978-3-933761-84-2

181 Prohaska, Sandra: *Development and Application of a 1D Multi-Strip Fine Sediment Transport Model for Regulated Rivers,* 2009, ISBN 978-3-933761-85-9

182 Kopp, Andreas: *Evaluation of CO_2 Injection Processes in Geological Formations for Site Screening,* 2009, ISBN 978-3-933761-86-6

183 Ebigbo, Anozie: *Modelling of biofilm growth and its influence on CO_2 and water (two-phase) flow in porous media,* 2009, ISBN 978-3-933761-87-3

184 Freiboth, Sandra: *A phenomenological model for the numerical simulation of multiphase multicomponent processes considering structural alterations of porous media*, 2009, ISBN 978-3-933761-88-0

185 Zöllner, Frank: *Implementierung und Anwendung netzfreier Methoden im Konstruktiven Wasserbau und in der Hydromechanik*, 2009, ISBN 978-3-933761-89-7

186 Vasin, Milos: *Influence of the soil structure and property contrast on flow and transport in the unsaturated zone*, 2010, ISBN 978-3-933761-90-3

187 Li, Jing: *Application of Copulas as a New Geostatistical Tool*, 2010, ISBN 978-3-933761-91-0

188 AghaKouchak, Amir: *Simulation of Remotely Sensed Rainfall Fields Using Copulas*, 2010, ISBN 978-3-933761-92-7

189 Thapa, Pawan Kumar: *Physically-based spatially distributed rainfall runoff modelling for soil erosion estimation*, 2010, ISBN 978-3-933761-93-4

190 Wurms, Sven: *Numerische Modellierung der Sedimentationsprozesse in Retentionsanlagen zur Steuerung von Stoffströmen bei extremen Hochwasserabflussereignissen*, 2011, ISBN 978-3-933761-94-1

191 Merkel, Uwe: *Unsicherheitsanalyse hydraulischer Einwirkungen auf Hochwasserschutzdeiche und Steigerung der Leistungsfähigkeit durch adaptive Strömungsmodellierung*, 2011, ISBN 978-3-933761-95-8

192 Fritz, Jochen: *A Decoupled Model for Compositional Non-Isothermal Multiphase Flow in Porous Media and Multiphysics Approaches for Two-Phase Flow*, 2010, ISBN 978-3-933761-96-5

193 Weber, Karolin (Hrsg.): *12. Treffen junger WissenschaftlerInnen an Wasserbauinstituten*, 2010, ISBN 978-3-933761-97-2

194 Bliefernicht, Jan-Geert: *Probability Forecasts of Daily Areal Precipitation for Small River Basins,* 2011, ISBN 978-3-933761-98-9

195 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2010 In-situ-Sanierung - Stand und Entwicklung Nano und ISCO -,* Tagungsband zur Veranstaltung am 07. Oktober 2010 an der Universität Stuttgart, Campus Stuttgart-Vaihingen*,* 2010*,* ISBN 978-3-933761-99-6

196 Gafurov, Abror: *Water Balance Modeling Using Remote Sensing Information - Focus on Central Asia,* 2010, ISBN 978-3-942036-00-9

197 Mackenberg, Sylvia: *Die Quellstärke in der Sickerwasserprognose: Möglichkeiten und Grenzen von Labor- und Freilanduntersuchungen,* 2010, ISBN 978-3-942036-01-6

198 Singh, Shailesh Kumar: *Robust Parameter Estimation in Gauged and Ungauged Basins,* 2010, ISBN 978-3-942036-02-3

199 Doğan, Mehmet Onur: *Coupling of porous media flow with pipe flow,* 2011, ISBN 978-3-942036-03-0

200 Liu, Min: *Study of Topographic Effects on Hydrological Patterns and the Implication on Hydrological Modeling and Data Interpolation,* 2011, ISBN 978-3-942036-04-7

201 Geleta, Habtamu Itefa: *Watershed Sediment Yield Modeling for Data Scarce Areas,* 2011, ISBN 978-3-942036-05-4

202 Franke, Jörg: *Einfluss der Überwachung auf die Versagenswahrscheinlichkeit von Staustufen,* 2011, ISBN 978-3-942036-06-1

203 Bakimchandra, Oinam: *Integrated Fuzzy-GIS approach for assessing regional soil erosion risks,* 2011, ISBN 978-3-942036-07-8

204 Alam, Muhammad Mahboob: *Statistical Downscaling of Extremes of Precipitation in Mesoscale Catchments from Different RCMs and Their Effects on Local Hydrology,* 2011, ISBN 978-3-942036-08-5

205 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2011 Flache Geothermie - Perspektiven und Risiken,* Tagungsband zur Veranstaltung am 06. Oktober 2011 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2011, ISBN 978-3-933761-09-2

206 Haslauer, Claus: *Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas with a Focus on Solute Transport Behaviour,* 2011, ISBN 978-3-942036-10-8

207 Dung, Nguyen Viet: *Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta,* 2011, ISBN 978-3-942036-11-5

208 Hung, Nguyen Nghia: *Sediment dynamics in the floodplain of the Mekong Delta, Vietnam,* 2011, ISBN 978-3-942036-12-2

209 Kuhlmann, Anna: *Influence of soil structure and root water uptake on flow in the unsaturated zone,* 2012, ISBN 978-3-942036-13-9

210 Tuhtan, Jeffrey Andrew: *Including the Second Law Inequality in Aquatic Ecodynamics: A Modeling Approach for Alpine Rivers Impacted by Hydropeaking,* 2012, ISBN 978-3-942036-14-6

211 Tolossa, Habtamu: *Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach,* 2012, ISBN 978-3-942036-15-3

212 Tatomir, Alexandru-Bodgan: *From Discrete to Continuum Concepts of Flow in Fractured Porous Media,* 2012, ISBN 978-3-942036-16-0

213 Erbertseder, Karin: *A Multi-Scale Model for Describing Cancer-Therapeutic Transport in the Human Lung,* 2012, ISBN 978-3-942036-17-7

214 Noack, Markus: *Modelling Approach for Interstitial Sediment Dynamics and Reproduction of Gravel Spawning Fish,* 2012, ISBN 978-3-942036-18-4

215 De Boer, Cjestmir Volkert: *Transport of Nano Sized Zero Valent Iron Colloids during Injection into the Subsurface*, 2012, ISBN 978-3-942036-19-1

216 Pfaff, Thomas: *Processing and Analysis of Weather Radar Data for Use in Hydrology*, 2013, ISBN 978-3-942036-20-7

217 Lebrenz, Hans-Henning: *Addressing the Input Uncertainty for Hydrological Modeling by a New Geostatistical Method*, 2013, ISBN 978-3-942036-21-4

218 Darcis, Melanie Yvonne: *Coupling Models of Different Complexity for the Simulation of CO_2 Storage in Deep Saline Aquifers*, 2013, ISBN 978-3-942036-22-1

219 Beck, Ferdinand: *Generation of Spatially Correlated Synthetic Rainfall Time Series in High Temporal Resolution - A Data Driven Approach*, 2013, ISBN 978-3-942036-23-8

220 Guthke, Philipp: *Non-multi-Gaussian spatial structures: Process-driven natural genesis, manifestation, modeling approaches, and influences on dependent processes*, 2013, ISBN 978-3-942036-24-5

221 Walter, Lena: *Uncertainty studies and risk assessment for CO_2 storage in geological formations*, 2013, ISBN 978-3-942036-25-2

222 Wolff, Markus: *Multi-scale modeling of two-phase flow in porous media including capillary pressure effects,* 2013, ISBN 978-3-942036-26-9

223 Mosthaf, Klaus Roland: *Modeling and analysis of coupled porous-medium and free flow with application to evaporation processes,* 2013, ISBN 978-3-942036-27-6

224 Leube, Philipp Christoph: *Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application,* 2013, ISBN 978-3-942036-28-3

Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: www.iws.uni-stuttgart.de zur Verfügung.