

# Mathematical Analysis of Natural and Artificial Tracers for the Quantification of River-Groundwater Exchange

## Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät  
der Eberhard Karls Universität Tübingen  
zur Erlangung des Grades eines  
Doktors der Naturwissenschaften  
(Dr.rer.nat)

vorgelegt von

Zijie Liao

aus Tianjin, Volksrepublik China

Tübingen  
2014

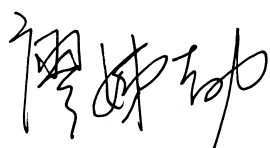
**Tag der mündlichen Qualifikation: 14.07.2014**

**Dekan: Prof. Dr. Wolfgang Rosenstiel**

**1. Berichterstatter: Prof. Dr.-Ing Olaf A. Cirpka**  
**2. Berichterstatter: PD Dr. Jan H. Fleckenstein**

## Declaration of Authorship

I herewith declare that I myself wrote the PhD thesis. I made use of only the sources and auxiliary means in word or context mentioned as such, and which I listed up completely and marked as such. I declare that I adhered to the guidelines of securing good scientific practical experience of the University of Tübingen (conclusion of the Academic Senate of 25 May 2000). I hereby affirm in lieu of an oath (Eidesstattliche Versicherung) that all these declarations are true and that I did not conceal anything. I am well aware that false declarations made under oath will be punished by a prison sentence up to three years or by a monetary penalty.



Zijie Liao

Den Haag, 09.April.2014

## Acknowledgements

I would like to take this opportunity to express my great thanks to my supervisor Prof. Dr.-Ing. Olaf A. Cirpka for his instructions and help during my study. He gave me the chance to study hydrogeology; he spent time to help me with many courses and my research whenever I need him; he corrected my mistakes when I was on the wrong track of the study and inspired me with new ideas when I was frustrated with unsuccessful work. Without his great guidance and successive support, I would not have finished this work at all. His dedication to work and his rigorous attitude towards research encourages and influences me all the time. I am really lucky and appreciate to get the chance to work with him.

I would like to thank specially my committee members: PD Dr. Jan H. Fleckenstein, Prof. Dr. Peter Dietrich, and Prof. Dr. Peter Grathwohl, for their constructive comments and thoughtful suggestions in improving this work.

Besides, I would like to thank my colleges in hydrogeology work group for their support and help when I need them. Many thanks to Dr. Dennis Lemke, for his great cooperation in data collecting and analyzing. We have worked closely together, exchanging ideas and discussing about the models and results. I would like to thank also all the colleges in WESS, especially Dr. Karsten Osenbrück, for his remarks and advice on my papers and thesis.

Last but not the least, I am very thankful to meet my husband Dr. Ronnie L. Schwede during my study and made the biggest step in my life. He is always supportive and helpful in many ways. Many thanks also goes to my parents, my parents-in-law and my sister, who support me unconditionally all the time.

There are many more names that I would like to mention, but it is not easy, because I have received so many helps, suggestions, supports, and encouragements that I cannot express my thankfulness enough. I am just very grateful for everyone who I spent the time with in the past five years in Tübingen University.

## Abstract

Evaluating the exchange between surface water and groundwater is of vital importance in understanding the ecological status of a river system and the groundwater quality in the associated riparian zone. Surface water, which is recognized as rich in oxygen and organic carbon, infiltrates into the adjacent aquifer, facilitating biogeochemical reactions, filtering hazardous particles, and degrading dissolved contaminants. These activities primarily determine the self-cleaning capacity of the river and the quality of extracted groundwater in adjacent pumping wells. A potential saving in exerting extra treatment of the extracted water is possible, when the travel times between the river and an extraction well is long enough to ensure sufficient filtration. Likewise the efficiency of self-cleaning depends on the intensity of hyporheic exchange. Therefore, both bank-filtration and hyporheic exchange have been intensively investigated in the past decades.

A well-known proxy for estimating the degree of filtration is the travel time of the infiltrating water within the subsurface. The longer the infiltrating water experiences the subsurface conditions, the less extra treatment is needed. Due to the natural morphology of the river channel and the inhomogeneous properties of the sediments, the travel time of the infiltrating river can be seen as a continuous random variable, ranging between zero and infinity. While the mean travel time may be used as the representative residence time in the subsurface, the non-linear dependence of biogeochemical reactions in the subsurface on travel time call for estimating the full statistical distribution of the travel time as a key point in characterizing the exchange process. The main goal of this work lies in estimating the travel-time distribution in bank-filtration systems and upon hyporheic exchange in rivers.

A classical way to obtain a travel-time distribution is by performing a tracer test. Both natural and artificial tracers are used in estimating the residence time of river water in the subsurface. In this work, continuous time series of a natural tracer, i.e. electrical conductivity, is used in the analysis of bank-filtration, whereas an artificial, reactive, and fluorescent tracer is considered in investigating the hyporheic exchange in the river channel. The reactive tracer resazurin (Raz) has been used for quantifying the metabolic activity in the stream system in recent years. It is irreversibly reduced to resorufin (Rru) under mildly reducing conditions. In this work, I mathematically analyze tests in which Raz together with a conservative tracer fluorescein (Fluo) are introduced into the river and observed at a downstream location. The differences among the breakthrough curves (BTCs) of the different tracers are used to infer the fraction of the river water undergoing hyporheic exchange and other properties of the hyporheic zone.

Many models have been used and developed to obtain a good estimation of the travel time distribution and the solute transport in the subsurface from continuous tracer signals. In the analysis of bank-filtration cross-correlation and autoregressive moving average (ARMA) models, among others, have been used under stationary conditions, whereas windowed cross-correlation, autoregressive integrated moving average (ARIMA), and the dynamic time warping (DTW) methods have been developed for the non-stationary case. An advanced method using nonparametric deconvolution for the stationary case was proposed by *Cirpka et al.* (2007) to infer the travel time distribution without predefining the shape of the distribution, which may reveal unconventional patterns of the hyporheic travel-time distribution of the infiltrating water

such as broad peaks or heavy tails; however, a corresponding non-stationary method has not yet been developed. In the hyporheic exchange models, different parametric functions are utilized for the travel-time distribution, but the discussion about which one is more appropriate than others is still ongoing, depending on the field sites and experimental conditions.

In this work, the two main goals are: 1) to estimate the river-to-well travel-time distribution of a bank-filtration system under non-stationary conditions; 2) to estimate the travel-time distribution of the river water in the hyporheic zone and its application. To achieve goal 1, I have developed a nonparametric deconvolution method with a sliding window to deal with the non-stationarity of the bank-filtration system. This method is able to capture dynamic properties of the system, and the nonparametric deconvolution facilitates the detection of unconventional shapes of the travel-time distribution, which have not yet been discussed in the literature. For goal 2, I have derived a methodology of embedding the nonparametric deconvolution into a transport model, and tested it with a synthetic tracer test. The model has been extended to account for a river undergoing hyporheic exchange with two-site sorption and three linear reactions in the hyporheic zone and has applied to a field data set.

The newly developed methods contribute to an advanced understanding of river-groundwater exchange. They provide an insight of solute transport in infiltrating river water and extend the available methodologies to model surface water-groundwater interactions.

## Zusammenfassung

Die Quantifizierung des Austauschs zwischen Oberflächen- und Grundwasser ist von essentieller Wichtigkeit, um den ökologischen Zustand eines Flusssystemes und die Grundwasserqualität in der Nähe des Flusses zu verstehen. Oberflächenwasser enthalten typischerweise viel Sauerstoff und organischen Kohlenstoff und infiltrieren in den angrenzenden Grundwasserleiter. Dadurch werden biogeochemische Reaktionen im Sediment ermöglicht, umweltschädliche Partikel ausgefiltert und gelöste Schadstoffe abgebaut. Diese Mechanismen bestimmen maßgeblich das Selbstreinigungsvermögen des Flusses und die Grundwasserqualität in ufernahen Grundwasserentnahmehäusern. Bei ausreichend langer Fließstrecke zwischen dem Fluss und dem Entnahmehaus ist eine ausreichende Filtration gewährleistet. Damit kann eine teure zusätzliche Wasseraufbereitung eingespart werden. Auf ähnliche Weise hängt die Selbstreinigungskapazität des Flusses von der Intensität des hyporheischen Austauschs ab. Unter anderem aus diesen Gründen wurden in den letzten Jahrzehnten sowohl die Uferfiltration als auch der hyporheische Austausch intensiv erforscht.

Eine bekannte Ersatzgröße um die Filtration abzuschätzen, ist die Verweilzeit des infiltrierten Wassers im Untergrund. Je länger das Wasser im Untergrund verweilt und die oben genannten Prozesse wirken, desto weniger zusätzliche Wasseraufbereitung ist nötig. Die unregelmäßige Morphologie des Flussbetts und die inhomogenen Eigenschaften des Sediments führen dazu, dass die Verweilzeit am besten als Zufallsvariable betrachtet wird, die Werte zwischen null und unendlich annehmen kann. Während die mittlere Verweilzeit möglicherweise als repräsentative Verweildauer des Wassers im Untergrund verwendet werden kann, erfordert die Simulation von nicht-linearen biogeochemischen Reaktionen im Untergrund die komplette statistische Verteilung der Verweilzeit. Dies ist ein wichtiger Punkt für das wissenschaftliche Verständnis der Austauschprozesse. Das Hauptziel der vorliegenden Arbeit besteht darin, die Verteilungsfunktion der Verweilzeit in der Uferfiltration und beim hyporheischen Austausch in Flusssystemen abzuschätzen.

Der klassische Ansatz, um die Verweilzeitverteilung im Untergrund zu bestimmen, ist der Tracerversuch. Hierzu können sowohl natürliche als auch künstliche Tracer ausgewertet werden. In der vorliegenden Arbeit wurden kontinuierliche Zeitreihen von natürlichen Tracern, nämlich der elektrischen Leitfähigkeit, zur Bestimmung der Verweilzeit bei der Uferfiltration untersucht. Dagegen wurde die Verweilzeit in der hyporheischen Zone von Flüssen mithilfe von künstlichen, reaktiven Tracern bestimmt. Der reaktive Fluoreszenzfarbstoff, Resazurin (Raz) wurde in den letzten Jahren für die Bestimmung der metabolischen Aktivität in Flusssystemen benutzt. Resazurin wird unter leicht reduzierten Bedingungen irreversibel zu dem ebenfalls fluoreszierenden Farbstoff Resorufin (Rru) umgewandelt. In der vorliegenden Arbeit habe ich Tracerversuche analysiert, bei denen gleichzeitig Raz und der konservative Markierstoff Fluorescein (Fluo) in den Fluss eingeleitet wurden und dann flussabwärts zusammen mit dem entstandenen Rru ihre Konzentration gemessen wurde. Die Unterschiede in den Durchbruchkurven der verschiedenen Tracer wurden benutzt, um die Stärke des hyporheischen Austausches, die Verweilzeitverteilung in der hyporheischen Zone sowie Sorptions- und Reaktionseigenschaften in der hyporheischen Zone zu bestimmen.

Es bestehen verschiedene Methoden, um aus kontinuierlichen Zeitreihen von Tracern die

Verteilungsfunktion der Verweilzeit abzuschätzen. Um die Verweilzeiten zwischen Flüssen und Grundwasserbeobachtungsstellen unter stationären Bedingungen zu ermitteln, wurden Methoden benutzt, die auf der Kreuzkorrelation und dem ARMA Ansatz (Autoregressive-Moving Average) basieren. Für den nichtstationären Fall wurden stückweise Kreuzkorrelations-Methoden, ARIMA-Modelle (Autoregressive Integrated Moving Average) und DTW-Methoden (dynamic time warping) angewendet. *Cirpka et al.* (2007) entwickelte eine formfreie Methode zur Bestimmung von stationären Fließzeitverteilungen. Diese Methode beruht auf einer nicht-parametrischen Dekonvolution und kann ungewöhnliche Fließverteilungen detektieren, die zum Beispiel weitständige Extremalwertbereiche aufweisen oder endlastig sind. Eine entsprechende Methode für den nicht-stationären Fall wurde bis jetzt noch nicht entwickelt. Für die Fließzeitverteilungen in der hyporheischen Zone wurden verschiedene parametrische Funktionen benutzt. Je nachdem an welchem Standort und unter welchen Bedingungen die Tests durchgeführt wurden, haben sich unterschiedliche Parametrisierungen bewährt, wobei ein klares Muster schwer zu erkennen ist.

Die vorliegende Arbeit verfolgt zwei Hauptziele: 1. Die Bestimmung der Fließzeitverteilung zwischen einem Fluss und einer Grundwasserbeobachtungsstelle in Uferfiltrationssystemen unter nichtstationären Bedingungen, 2. Die Bestimmung der Verweilzeitverteilung von Flusswasser in der hyporheischen Zone sowie die Anwendung der entwickelten Methoden. Um das erste Ziel zu erreichen, habe ich die nicht parametrische Dekonvolution mit einem gleitenden-Intervall Ansatz (sliding window) kombiniert. Mit dieser Methode können sowohl die unkonventionellen Formen der Verteilungsfunktion durch die Verwendung der nichtparametrischen Dekonvolution als auch die dynamischen Eigenschaften des Systems anhand des gleitenden Intervalls detektiert werden. Um das zweite Ziel zu erreichen, habe ich eine Methode entwickelt, die die nicht parametrische Dekonvolution auf die Bestimmung hyporheischen Verweilzeiten überträgt. Diese neue Methode habe ich dahingehend erweitert, dass ich mehrere Sorptions- und Umwandlungsprozesse der reaktiven Tracer in der hyporheischen Zone mitberücksichtigen konnte und diese auf Felddaten angewendet habe.

Die von mir neu entwickelten und benutzten Methoden tragen zu einem tieferen Verständnis des Austauschs zwischen Flüssen und dem Grundwasser bei. Sie geben einen tieferen Einblick in den Transport von Stoffen, die mit dem infiltrierenden Wasser in den Untergrund transportiert werden und liefern Wissenschaftlern, die sich mit dem Austausch von Oberflächen- und Grundwasser beschäftigen, eine neue Methodik zur Datenanalyse.



# Contents

<b>List of Figures</b>	<b>XI</b>
<b>List of Tables</b>	<b>XVII</b>
<b>1 General Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Review . . . . .	3
1.2.1 Tracer test . . . . .	3
1.2.2 Models . . . . .	3
1.3 Objectives . . . . .	5
1.4 Theory . . . . .	5
1.4.1 Nonparametric deconvolution ( <i>Cirpka et al., 2007</i> ) . . . . .	5
1.4.2 Nonparametric deconvolution under non-stationary condition (Chapter 2)	7
1.4.3 In-channel transport with hyporheic exchange (Chapter 3: <i>Liao and Cirpka (2011)</i> ) . . . . .	8
1.4.4 Governing equation for the transport of the Raz-Rru system in the hyporheic zone . . . . .	10
1.4.5 In-channel transport with hyporheic exchange including two-site sorption (Chapter 4: <i>Liao et al. (2013)</i> ) . . . . .	11
1.5 Outline . . . . .	12
<b>2 Non-Stationary Nonparametric Inference of River-to-Groundwater Travel-Time Distributions</b>	<b>15</b>
2.1 Introduction . . . . .	15
2.2 Method . . . . .	18
2.2.1 Overview of the Method . . . . .	18

2.2.2	Data Detrending . . . . .	18
2.2.3	Principles of Auto-Correlated Deconvolution with a Sliding Window . . . . .	21
2.2.4	Choice of the Window Function . . . . .	22
2.2.5	Discrete Parameter-Estimation Procedure . . . . .	22
2.2.6	Goodness of Fit . . . . .	25
2.3	Application to Field Data . . . . .	25
2.3.1	Field Site and Data . . . . .	25
2.3.2	Detrending of the Data . . . . .	26
2.3.3	Results of Deconvolution . . . . .	26
2.3.4	Estimated Transfer Functions . . . . .	27
2.3.5	Model Performance . . . . .	30
2.3.6	Relationship to Base Flow . . . . .	33
2.4	Conclusions . . . . .	35
2.A	Detailed information of the model . . . . .	39
2.A.1	Regularization of the Transfer Functions $g(\tau)$ . . . . .	39
2.A.2	Selection Matrix $\mathbf{H}$ and Rules of Activating and Deactivating Lagrange Multipliers . . . . .	40
2.B	Peak-Finding Method . . . . .	40
<b>3</b>	<b>Shape-Free Inference of Hyporheic Travel-Time Distributions from Synthetic Conservative and 'Smart' Tracer Tests in Streams</b>	<b>43</b>
3.1	Introduction . . . . .	44
3.2	Governing Equations . . . . .	46
3.2.1	One-Dimensional Advective-Dispersive Transport in Streams . . . . .	46
3.2.2	Solution in the Laplace Domain . . . . .	48
3.3	Estimating parameters . . . . .	49
3.3.1	Shape-Free Inference . . . . .	49
3.3.2	Initial Guesses . . . . .	53
3.4	Application to a Hypothetical Test Case . . . . .	54
3.4.1	Model Setup . . . . .	55
3.4.2	Results of Inversion . . . . .	55
3.5	Discussion and Conclusions . . . . .	58
3.A	Equation comparison . . . . .	61

3.B	Derivation of the Objective Function $W(\mathbf{p})$ . . . . .	62
3.C	Sensitivity Analysis . . . . .	64
3.D	Temporal Moment Analysis . . . . .	66
<b>4</b>	<b>Modeling and Inverting Reactive Stream Tracers Undergoing Two-Site Sorption and Decay in the Hyporheic Zone</b> . . . . .	<b>69</b>
4.1	Introduction . . . . .	70
4.2	Model Formulation . . . . .	72
4.2.1	General Setup Considered . . . . .	72
4.2.2	Governing Equations . . . . .	72
4.2.3	Hyporheic Transfer Functions for Conservative and Reactive Tracers . . . . .	74
4.2.4	Joint Estimation of the Shape-Free Hyporheic Travel-Time Distribution and Reactive Parameters . . . . .	78
4.3	Application to Field Data . . . . .	79
4.3.1	Experimental Setup . . . . .	79
4.3.2	Results of Data Analysis . . . . .	80
4.4	Discussion and Conclusions . . . . .	86
4.A	Solution in the Laplace Domain . . . . .	89
4.B	Transport equations . . . . .	90
4.C	Solution of Governing Equation . . . . .	91
4.D	Discretization . . . . .	92
<b>5</b>	<b>Consistent Simulation of Pulse-Like Conservative and Reactive Stream-Tracer Experiments on Multiple Scales</b> . . . . .	<b>95</b>
5.1	Introduction . . . . .	96
5.2	Experiments . . . . .	96
5.2.1	Column experiments . . . . .	96
5.2.2	Gravel-bar experiments . . . . .	98
5.3	Theory . . . . .	98
5.4	Results . . . . .	99
5.4.1	Column experiment . . . . .	99
5.4.2	Gravel-bar experiment . . . . .	100
5.4.3	Comparison of the parameters . . . . .	100
5.5	Conclusions . . . . .	102

<b>6 Conclusions and Outlook</b>	<b>103</b>
<b>Bibliography</b>	<b>107</b>

# List of Figures

1.1	A simple schematic plot of the river water-groundwater interaction in a riparian region. The red dashed lines indicate the hyporheic exchange, whereas the light blue dashed lines show the bank-filtration process. . . . .	2
2.1	Overview of the method at time step $t_c$ . Within each iteration step, the current guess of the transfer function $g(\tau, t_c)$ is convoluted with the detrended input signal. The simulated and measured output signals are windowed, weighted by the weighting function, and compared to each other to obtain part 2 of the objective function. As first contribution to the objective function, the method penalized differences between the current transfer function $g(\tau, t_c)$ and that of the previous time step $g(\tau, t_c - \Delta t)$ . The transfer function $g(\tau, t_c)$ is then updated based on the minimization of the sum of the objective functions. . . . .	19
2.2	Original EC data and the baselines for well 1, 2 and 3. Black: EC data in the river; blue: EC data in the wells. . . . .	20
2.3	Weighting function according to Eq. (2.7). $t_c$ : center point, $w$ : width of the window. The rising and falling limbs are determined by the Gaussian cumulative distribution function. . . . .	23
2.4	Detrended data from November, 2003 to September, 2005. Black: detrended EC signal in the river; blue: detrended EC signal in the wells. . . . .	27
2.5	Estimated stationary transfer functions from the river to wells 1, 2, and 3, respectively, using stationary method. Log-normal distributions with the same mean and variance are used as the initial guess for the first window in the non-stationary method. . . . .	28

2.6	Simulated EC time series of well 1, 2, and 3 using the stationary and non-stationary methods, respectively. Gray line: measured EC signal in the wells; Red line: simulation of the well EC time series using the stationary method; Blue line: simulation using the non-stationary method. . . . .	30
2.7	Transfer function obtained by the non-stationary method for well 1. Truncation travel time: 15 days. . . . .	31
2.8	Transfer function obtained by the non-stationary method for well 2. Truncation travel time: 6 days. Note that the scale of travel time $\tau$ for well 2 is different from the other two wells (see Figures 2.7 and 2.9). . . . .	32
2.9	Transfer function obtained by the non-stationary method for well 3. Truncation travel time: 15 days. . . . .	33
2.10	Estimated parameter changing over time with uncertainties. Black lines are for well 1; red lines are for well 2; and green lines are for well 3. The gray bands are the parameter uncertainties. Water level and base flow is plotted for analysis purpose. . . . .	34
2.11	Recovery rate of the EC signal in the wells. Horizontal lines are the mean values. Black line: well 1 (mean of 41%); red line: well 2 (mean of 78%); green line: well 3 (mean of 47%). . . . .	35
2.12	Peak height comparison between the simulation and the output signals of wells 1, 2, and 3 for the stationary and non-stationary methods. On the left: the top two figures are for well 1; the middle two figures are for well 2; the bottom two figures are for well 3. On the right: corresponding box plot of the data used for plotting on the left. . . . .	37
2.13	Peak time comparison for stationary and non-stationary method between the simulation and the output signals for well 1, 2 and 3. On the left: the top two figures are for well 1; the middle two figures are for well 2; the bottom two figures are for well 3. On the right: corresponding box plot of the data used for plotting on the left. . . . .	38
2.14	Peak height and time differences are found by the peak-matching algorithm for the simulation and the observation data, using the method of <i>Dugge</i> (2012). . .	41

3.1	Illustration of the transient-storage model concept. Advective-dispersive transport within the stream is considered one-dimensional. Transport through the hyporheic zone is assumed to be along non-interacting streamlines, in which the longitudinal spatial coordinate is replaced by travel time. . . . .	46
3.2	Flow chart of the general approach. The shape-free inference method requires an initial guess provided by parametric inference. . . . .	50
3.3	True hyporheic travel-time distribution $g(\tau)$ in the virtual test case (solid line), generated by summing an exponential function (dotted line) and a log-normal distribution (dashed line). . . . .	57
3.4	Virtual true breakthrough curves of conservative and reactive tracers. Solid lines: breakthrough curves calculated from forward model; dotted lines: noise-added breakthrough curves used as virtual measurements in the inversion. . . .	58
3.5	Fitted breakthrough curves obtained by the joint shape-free inference method. Solid lines: fitted breakthrough curves of conservative and reactive tracers; dotted lines: measurements of tracer concentration used for the inversion. . . . .	60
3.6	Estimated hyporheic travel-time distributions $g(\tau)$ . Solid line: mean value of $g(\tau)$ from the set of conditional realizations; gray band: range of 16% – 84% probability of the conditional statistical distributions; dotted line: virtual true hyporheic travel-time distribution; dashed line: exponential function obtained in the parametric inference step (initial guess of shape-free inference). . . . .	61
4.1	General setup of the model. The conservative tracer and reactive parent compound are injected as a joint pulse into the stream. No reaction takes place in the stream for all tracers, while in the hyporheic zone the reactive parent compound undergoes first-order transformation to the reaction product and both reactive tracers are subject to two-site sorption, i.e., equilibrium and kinetic sorption. the latter is shown as two different small boxes for each cell within the hyporheic zone. The strength of hyporheic exchange is quantified by $q_{he}$ as discharge of exchanged water per water volume in the stream. Transport within the hyporheic zone is described as advection in travel time rather space. The hyporheic travel-time distribution $g(\tau)$ determines the age distribution of the water returning into the stream. Breakthrough curves of all three tracers are measured downstream.	73

- 4.2 Principle of calculating transfer functions for different tracers from the hyporheic travel-time distribution  $g(\tau)$ . A: length profiles of reactive-tracer concentrations with respect to travel time  $\tau^*$  for different times  $\tau$ , indicated by changing colors (lighter colors mark later times). Solid lines: reactive parent compound; dotted lines: reactive daughter compound; black dashed line: hyporheic travel-time distribution  $g(\tau^*)$ . B: corresponding transfer functions, reflecting the sampling of concentrations in A according to  $g(\tau^*)$ . . . . . 76
- 4.3 Independent fits of breakthrough curves for each monitoring station (MS). Left column: measured and fitted breakthrough curves obtained at three monitoring stations (MS) which are located 830m, 1075m and 1210m downstream of the injection point. Green lines: Fluo (conservative tracer), blue lines: Raz (reactive parent compound), red lines: Rru (reaction product). Horizontal lines: quantification limits of the measurement device for each tracer. Right column: estimated transfer functions for all tracers. . . . . 81
- 4.4 Joint fits of all breakthrough curves and all tracers; measured and fitted breakthrough curves at the three monitoring stations; single set of hyporheic transfer functions obtained from jointly fitting the data of all three monitoring stations (MS). Green lines: Fluo (conservative tracer), blue lines: Raz (reactive parent compound), red lines: Rru (reaction product). Horizontal lines: quantification limits of the measurement device for each tracer. . . . . 82
- 4.5 Uncertainty of the transfer functions obtained by the joint fit of all breakthrough curves at all monitoring stations. The gray band is 95% of the uncertainty for the transfer functions based on 1120 conditional realizations. . . . . 83
- 5.1 Conservative tracer (Fluo) and reactive tracer (Raz) are used in the tracer tests, where the tracer concentration of these tracers and the reaction product (Rru) are measured at the observation points. . . . . 97
- 5.2 Schematic plot for the column experiments. Two fluorometers were placed at the inlet and the outlet of the column. Sieved sediments with diameter  $< 4mm$  were filled in to the column. The experiment was done at room temperature and the pH value was kept about 7. . . . . 97



5.3 Schematic plot of the gravel-bar experiments. One fluorometer is placed in front of the gravel bar in the river, and the other one is measuring the signal in a piezometer within the gravel-bar. Fluo and Raz solution is introduced as a pulse into the river, and their concentrations including Rru are measured by the fluorometers. . . . . 98

5.4 (a) Simulation of column experiment. Top: Fluo; middle: Raz; bottom: Rru. Dotted lines are input signals; black solid lines are measured output signals; and colored lines are the simulation results of the model. (b) Transfer function for the column experiment. Green: transfer function of Fluo; blue: transfer function of Raz; red: transfer function of Rru; cyan: transfer function from Raz to Rru. . . 100

5.5 (a) Simulation of gravel-bar experiment. Top: Fluo; middle: Raz; bottom: Rru. Dotted lines are input signals; black solid lines are output signals; and colored lines are the simulation results of the model. (b) Transfer function of the gravel-bar experiment. Green: transfer function of Fluo; blue: transfer function of Raz; red: transfer function of Rru; cyan: transfer function from Raz to Rru. . . . . 101



# List of Tables

1.1	The overview of the chapters in this work. Chapter 3, Chapter 4 and Chapter 5 address the problem of the hyporheic exchange under stationary condition, whereas Chapter 2 deals with the bank-filtration system under non-stationary condition. . . . .	12
2.1	Parameters for deconvolution of this work compared to those of <i>Cirpka et al.</i> (2007). Truncation time: length of the estimated transfer function; $\theta$ : smoothness parameter of the transfer function within each window; . . . . .	29
2.2	Comparison of estimated parameter characteristics between non-stationary deconvolution of raw data (this work) and stationary deconvolution of detrended data ( <i>Cirpka et al.</i> , 2007). . . . .	36
2.3	Model performance of the non-stationary deconvolution simulation. RMSE: root mean square error; NS: Nash-Sutcliffe model efficiency coefficient; stat.: stationary method; non-stat.: non-stationary method. . . . .	36
2.4	Correlation coefficients of river base flow and estimated parameters. Note that these values are calculated based on the normalized data. . . . .	39
3.1	Parameters of the virtual test case. . . . .	56
3.2	Comparison between true and estimated parameter values without hyporheic travel-time distribution. "Single est.": best nonnegative shape-free estimate without conditional realizations; "cond. real.": results of conditional realizations. . . . .	59

4.1	Estimated parameters from breakthrough curve of Fluo, Raz, and Rru at three difference monitoring stations (MS) and joint fit of all three stations. The uncertainty of the estimated parameters is given as coefficients of variation in per cent. The quality of the fits is quantified by the root mean square error (RMSE) in the reproduction of the measured breakthrough curves by the model. . . . .	84
5.1	Recovery rate and mean travel time for column and gravel-bar experiment. . . . .	101
5.2	Estimated parameters from breakthrough curves of Fluo, Raz, and Rru from the column and gravel-bar experiments. $R_1[-]$ : retardation factor for Raz; $R_2[-]$ : retardation factor for Rru; $K_1[-]$ : distribution coefficient for Raz; $K_2[-]$ : distribution coefficient for Rru; $k_1[1/s]$ : sorption mass transfer coefficient for Raz; $k_2[1/s]$ : sorption mass transfer coefficient for Rru; $\lambda_1[1/s]$ : decay coefficient of Raz to unknown compound; $\lambda_{12}[1/s]$ : decay coefficient of Raz to Rru; $\lambda_2[1/s]$ : decay coefficient of Rru to unknown compound. . . . .	102

# Chapter 1

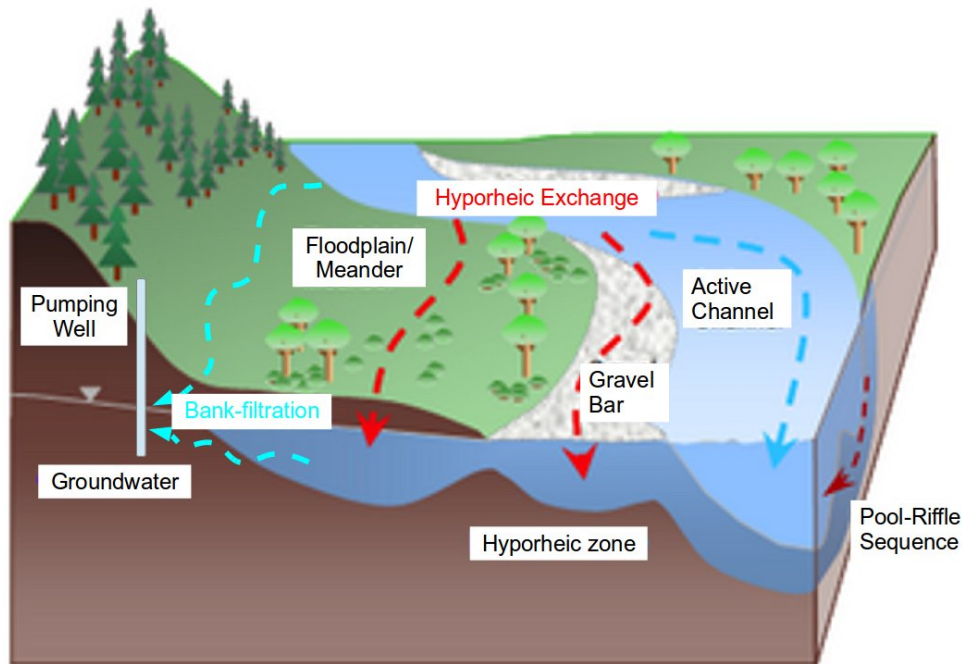
## General Introduction

### 1.1 Motivation

River water and groundwater, among others, are the most important water sources for land use. These two water bodies are successively interchanging, which has become a hot spot in hydrogeological research over the past decades (*Sontheimer, 1980; Schubert, 2002; Bencala and Walters, 1983; Boulton et al., 1998; Sophocleous, 2002*, among others). The typical river water, which is normally rich in oxygen, nutrients and dissolved organic carbon, infiltrates into the subsurface and facilitate the physical and chemical activities or reactions in the microbial communities. This is of vital importance of the water quality, self-cleaning capacity of the river, and the water supply of drinking water. The degree of the exchange between the river water and ground water is influenced by many factors, for example, the distance between the river water and the extraction point of the groundwater, or the characteristic of the aquifer among others (*Gooseff et al., 2003; Cardenas et al., 2004; Briggs et al., 2009; Bianchin et al., 2010; Camporese et al., 2014*).

In this work, river water is discussed intensively about how it interacts with groundwater in the adjacent aquifer. Two different types of exchanges are investigated, namely bank-filtration and the hyporheic exchange. Figure 1.1 shows a schematic plot for the river system and its surroundings. Bank-filtration is shown as the light blue dashed lines and hyporheic exchange is shown as the red dashed lines. Both processes are driven by hydraulic head differences and they are the two main means of the river water-groundwater interaction.

A bank-filtration system constitutes of a groundwater well, or well gallery, in the vicinity of a river, accompanied with a water-purification facility if necessary. The extracted groundwater mainly originates from river-water infiltrate. The hyporheic zone, the exact definition of which is still under discussion (*Hester et al., 2013; Zhou and Endreny, 2013*), is normally recognized as the transition zone between the river and the aquifer, where the river water infiltrates into the subsurface and ex-filtrates back into the river channel again. Both bank-filtration and hyporheic exchange require that the fresh water infiltrates into the subsurface, the difference is that the exfiltration point appears at the extraction well in the case of bank-filtration and in the river channel in the case of hyporheic exchange. In many cases bank-filtration plays a vital roll in the water quality of an extraction well used for drinking purpose (*Sontheimer, 1980; Kuehn and*



Modified from <http://www.wess.info/typo3/index.php?id=60>

Figure 1.1: A simple schematic plot of the river water-groundwater interaction in a riparian region. The red dashed lines indicate the hyporheic exchange, whereas the light blue dashed lines show the bank-filtration process.

*Mueller, 2000; Ray et al., 2002; Schubert, 2002; Weiss et al., 2005b*), and hyporheic exchange determines the self-cleaning capacity of the river water (*Boulton et al., 1998; Sophocleous, 2002; Bencala and Walters, 1983; Findlay, 1995; Brunke and Gonser, 1997; Hester and Gooseff, 2010; Robertson and Wood, 2010; Krause et al., 2011*), for example when contamination events take place in the river.

How much reactions and activities occur in the subsurface depend highly on the time that the fresh water remains in the subsurface area. The longer the water stays therein, the more it experiences mineralization, and the more efficient is the filtration of the pathogenic particles and the degradation of the organic contaminant (*Weiss et al., 2005b*). The water body passes through the aquifer or the hyporheic zone, and reaches the extraction well or returns back into the river channel due to the hydraulic head difference following Darcy's law, including the dispersion and adsorption processes in the transport. Some solutes travel the shortest way and ex-filtrate early, whereas other solutes might be trapped in a small pore or absorbed onto the surface of the sediment, so that they reach the extraction point much later. Therefore the travel time is best described as a distribution, rather than a single value. The time that the water needs to travel through the subsurface is a key descriptor of the exchange process and is also the main target of this work.

## 1.2 Review

### 1.2.1 Tracer test

A traditional way of estimating the travel time is by tracer tests, where natural or artificial tracers are observed at an observation point in the river or in the aquifer. Tracer tests have been widely used in hydrogeological research (*Davis et al.*, 1980). Different types of tracers are suitable for different time scales, spatial scales, field sites, convenience, cost performances and so forth. For bank-filtration, the travel-time distribution of the infiltrating river water can be either calculated by injecting an artificial tracer into the river and measure the concentration time series, denoted breakthrough curve, at the extraction well, or by deconvoluting natural tracer time series in the river channel and the observation well (*Sheets et al.*, 2002; *Luo et al.*, 2006; *Maloszewski and Zuber*, 1993). Both methods may be seen as a direct measurement method, because the breakthrough curve in the observation point is already the probability density function of the travel times when a point injection of a conservative tracer is conducted. However, the travel time within the hyporheic zone is comparably harder to obtain, because it is practically impossible to collect all hyporheic water returning back into the river. Some studies have focused on the storage time of the water in the whole system, which include the hyporheic zone and the surface water body before it reaches the observation point (*Haggerty et al.*, 2002), but this travel time is not restricted to the hyporheic zone, therefore it is difficult to investigate how long the fresh water has remained exclusively in the hyporheic zone. *Haggerty et al.* (2008) proposed to use resazurin (Raz) as an artificial tracer in river tracer tests. It is a fluorescent dye which can be detected at low concentration and is irreversibly reduced to another fluorescent dye, resorufin (Rru). It is believed that this transformation predominantly takes place in the hyporheic zone where microbiological activity prevails (*González-Pinzón et al.*, 2012). Applying this reactive tracer is useful for two purposes: (1) quantifying the overall metabolic activity of the river, predominantly its hyporheic zone, and (2) distinguishing between in-stream dispersion and longitudinal spreading of solutes caused by hyporheic exchange.

### 1.2.2 Models

Many models investigating the residence time of fresh water in the subsurface have been proposed in the literature (a review can be found by *Krause et al.* (2014)). Most of the models are falling into four categories: a) transport models in the river assuming that the river channel is a one-dimensional or quasi one-dimensional system (ADE, with one or more storage zone(s)) (e.g. *Bencala and Walters*, 1983; *Harvey et al.*, 1996; *Haggerty et al.*, 2002; *Wörman et al.*, 2002; *Marion et al.*, 2008; *Deng and Jung*, 2009; *Gooseff et al.*, 2005a; *Choi et al.*, 2000); b) transport both in the river and the subsurface by considering the morphology of the river bed and the path ways of solute transport (e.g. *Revelli et al.*, 2008; *Boano et al.*, 2010a; *Cardenas*, 2007, 2008a); c) localized investigation of the seepage velocity and transport (e.g. *Sawyer et al.*, 2011; *Hatch et al.*, 2006; *Vogt et al.*, 2010); d) adapting methods from time series analysis onto the continuously recorded data and decipher the residence time (e.g. *Stichler et al.*, 1986; *Hunt et al.*, 2005; *Sheets et al.*, 2002; *Schmidt et al.*, 2012).

The residence time distribution of the solute transport in the river bed has been intensively discussed for different field sites and experiments. The most classical one is an exponential function, which is inherited from the transient storage model (*Bencala and Walters, 1983; Wörman et al., 2002; Lemke et al., 2013a*). The assumption is that the mixing in the storage zone is immediate and complete. A two-storage zone model has been proposed because in-stream storage and hyporheic exchange pose the same result in terms of tracking water particles, but water in the in-stream storage is exposed to sunlight and open air implying a completely different environment than the hyporheic zone (*Briggs et al., 2009; Choi et al., 2000; Gooseff et al., 2004*). Other functions used to parameterize the travel-time distribution are based on specific characteristics of the field sites (*Gooseff et al., 2005b*), such as the truncated power-law distribution (*Dentz et al., 2004*), the Gamma distribution (*Tetzlaff et al., 2014; Luo et al., 2006*), and the log-normal distribution (*Wörman et al., 2002*). More general models (*Haggerty and Gorelick, 1995; Varni and Carrera, 1998; Marion et al., 2008*, among others) have extended the travel-time distribution from one specific parametric function to a family of exponential functions. The final shape of the distribution is estimated by the measurements themselves. This to some extent solves the question of which function is the most suitable one for a field site without trying out all possible candidates. However, due to the heterogeneity of the aquifer or the hyporheic zone, different flow paths might be involved, and therefore unconventional features of the travel-time distribution, for example long tails or multiple peaks, may be needed in describing the residence time distribution.

Experiments on bank-filtration utilize both natural and artificial tracers (*Leibundgut et al., 2009; Lin et al., 2003; Darling et al., 2010*, among others). In principle, solute transport in groundwater can be described by the advection-dispersion equation, but detailed characterization of the aquifer is required to obtain suitable spatial parameter distributions (*Malaguerra et al., 2013*). Many methods consider only the input and output tracer signals, and treat the aquifer as a black box (*Sheets et al., 2002*). The comparison of these signals provides an insight in the mean travel time and the damping of the signals, for example in the cross-correlation method for estimating the time lags. However, the distribution of the residence time is important in understanding the characteristics of the system. *Cirpka et al. (2007)* proposed a non-parametric deconvolution method to accommodate this situation, where no prior shape of the travel-time distribution is assumed.

In many cases, stationarity of the system is assumed during the experimental period. This simplification may be valid when the experimental time is too short that the changes of the flow field site is not possible to detect. In other cases, when dynamic changes are rather intense or the experimental period is relatively long, it is not appropriate to ignore non-stationarity. To address this problem, some stationary methods have been modified to a non-stationary version and applied in order to understand more about the dynamic system. An intuitive way to extend a stationary into non-stationary method is to use the sliding window technique (*Boker et al., 2002; Schmidt et al., 2012; Keogh et al., 2001*), where the long experimental time is sliced into short time periods for which stationarity is assumed. The window slides along the experimental time and resulting in a set of parameters changing with time. This has been successfully implemented in the cross-correlation method (*Boker et al., 2002; Schmidt et al., 2012; Dürrenmatt et al., 2013*), where the most likely time lag, which has the highest correlation coefficient, is calculated



for each window and the dynamic change of the system is manifested in time-varying time lags. Another method, the dynamic time warping (DTW), has also become popular in searching the similarities between input and output signals (*Schmidt et al.*, 2012). It is similar to the windowed cross-correlation method, but other than searching the most likely time lag of a single window, DTW yields the changes of the time lags for each time point. However, methods that determine how the shape of the travel-time distribution changes over time are missing in recent literatures.

## 1.3 Objectives

Based on the literature review above, this work focuses on improving our understanding about the interactions between river water and groundwater in its adjacent aquifer, filling the methodological gap of estimating the shape of the transfer function for the bank-filtration systems under non-stationary conditions and evaluating the travel-time distribution of infiltrating water in the hyporheic zone. In general, the following questions are addressed and discussed in this work:

- Question 1: How can the approach of *Cirpka et al.* (2007) for estimating the river-to-groundwater travel-time distribution be extended from stationary to non-stationary conditions, and how well can dynamic changes of the system be characterized by analyzing natural tracer time series using the extended approach?
- Question 2: What is a suitable methodology of estimating the travel-time distribution of river water in the hyporheic zone? Is it possible to determine unconventional shapes of the hyporheic travel-time distributions?
- Question 3: Can this method be applied to field measurements? How can possible reaction and adsorption processes of tracers in the hyporheic zone be addressed in this framework?
- Question 4: Is this method ready for upscaling? How does the model perform for tracer tests at different scales?

## 1.4 Theory

To answer the above questions, a series of models have been developed and used for different cases. Because the notations used in the following chapters vary, I summarize the models and governing equations of these chapters using a consistent notation, thus showing the connections among them.

### 1.4.1 Nonparametric deconvolution (*Cirpka et al.*, 2007)

The kernel of this work is the nonparametric deconvolution proposed by *Cirpka et al.* (2007) assume input tracer concentration signal  $c_{in}(t)[ML^{-3}]$  and output signal  $c_{out}(t)[ML^{-3}]$  are collected in the river and the well, respectively. The travel-time distribution (for a conservative tracer) can be calculated by inverting the convolution integral:

$$c_{out}(t) = \int_0^{\infty} g(\tau)c_{in}(t-\tau)d\tau \quad (1.1)$$

in which the travel-time distribution  $g(\tau)[T^{-1}]$  indicates the probability density that the water solute reaches the output at time  $\tau$ . By saying travel-time distribution, we mean the distribution of conservative tracers travel through the system. Because  $g(\tau)$  is a probability density function, the integral of the travel-time distribution is 1 by definition. For non-conservative tracers, we borrow the term transfer function from signal analysis, which expresses only how the system responds to a pulse input including all possible processes like reaction, decay and adsorption. The transfer function does not need to integrate to unity. Therefore, the travel-time distribution is the special case of a transfer function for a conservative tracer.

The discrete form of Eq. 1.1 is:

$$\mathbf{y} = \mathbf{X}\mathbf{g} \quad (1.2)$$

in which  $\mathbf{y}$  is the discrete form of the output signal  $c_{out}$ , and  $\mathbf{x}$  denotes the discrete form of input signal  $c_{in}$ , and

$$\mathbf{X}_{ij} = \Delta t x_{i-j+1} \quad (1.3)$$

where  $\mathbf{X}$  is a Toeplitz matrix constructed from the discrete input signal  $\mathbf{x}$ . To circumvent over-fitting of the fluctuating signals, the travel-time distribution  $\mathbf{g}[T^{-1}]$  is assumed to be a second-order intrinsic random time variable with a linear semivariogram:

$$\gamma_g(h) = E\left[\frac{1}{2}(g(\tau+h) - g(\tau))^2\right] = \theta|h| \quad (1.4)$$

in which  $\theta[T^{-3}]$  is a smoothness parameter, determining how strongly the transfer function is expected to fluctuate with travel time  $\tau[T]$ ;  $h[T]$  is the distance between two comparable elements in  $g(\tau)$ , which is usually an integer multiple of the travel time increment  $\Delta t$ . As additional constraint, we need to consider that the travel-time distribution is a strictly non-negative function.

The objective function is then:

$$W(\mathbf{g}|\mathbf{x}, \mathbf{y}) = \underbrace{(\mathbf{y} - \mathbf{X}\mathbf{g})^T \mathbf{C}^{-1} (\mathbf{y} - \mathbf{X}\mathbf{g})}_{\text{residual}} - \underbrace{\mathbf{g}^T \mathbf{\Gamma}_{gg}^{-1} \mathbf{g}}_{\text{smoothness of } g} + \underbrace{\mathbf{v}^T \mathbf{H}\mathbf{g}}_{\text{non-negativity}} \quad (1.5)$$

in which  $\mathbf{C}$  is the covariance matrix of the measurement error,  $\mathbf{\Gamma}_{gg}$  is the discrete semivariogram matrix, defined by  $\Gamma_{gg}(i, j) = \gamma_g(|t_i - t_j|)$ ,  $\mathbf{v}$  is the Lagrange multiplier where the non-negativity constraint is applied, matrix  $\mathbf{H}$  is an selected matrix, defined by:

$$\mathbf{H} = \begin{cases} 1, & \text{if } g_j \text{ is affected by the } i\text{-th constraint} \\ 0, & \text{otherwise} \end{cases} \quad (1.6)$$

The scheme is iterative because setting the constraints by introducing Lagrange multipliers depends on the estimated transfer function  $g(\tau)$  itself. The rules for activating and deactivating the constraints are as follows:

$$\begin{aligned}
g_j \leq 0 &\rightarrow \text{add constraint for element } j \\
v_i \leq 0 &\rightarrow \text{keep constraint } i \\
v_i > 0 &\rightarrow \text{remove constraint } i
\end{aligned} \tag{1.7}$$

The objective function is minimized by solving the following system of linear equations via taking derivatives with respect to each parameters and setting it to zero:

$$\begin{bmatrix} \mathbf{X}^\top \mathbf{C}^{-1} \mathbf{X} - \Gamma_{gg}^{-1} & \mathbf{H}^\top \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{C}^{-1} \mathbf{y} \\ \mathbf{0} \end{bmatrix} \tag{1.8}$$

This method is used onto the input-output signal when analyzing the data from bank-filtration applications. The smoothness parameter  $\theta$  of travel-time distribution  $g(\tau)$  is estimated by generating conditional realizations and applying the expectation-maximization method (McLachlan and Krishnan, 1997). More details can be found in Cirpka et al. (2007) and Liao and Cirpka (2011) (Chapter 3).

### 1.4.2 Nonparametric deconvolution under non-stationary condition (Chapter 2)

To extend the nonparametric deconvolution method of Cirpka et al. (2007) to non-stationary conditions, Eq. 1.1 is applied only to a small windowed time series. Without detrending the data, we have the following governing equation at  $t_c$  as the forward operator:

$$c_{out}(t, t_c) = \int_0^\infty g(\tau, t_c) c_{in}(t - \tau) d\tau \tag{1.9}$$

in which  $t_c [T]$  is the center time of the selected window;  $c_{out}(t, t_c) [ML^{-3}]$  is the simulation based on the transfer function at time  $t_c$ . The discrete form becomes:

$$\mathbf{y}_{t_c} = \mathbf{X} \mathbf{g}_{t_c} \tag{1.10}$$

in which the matrix  $\mathbf{X}$  is defined the same as in Eq. 1.3;  $\mathbf{y}_{t_c}$  is the discrete form of the simulation corresponding to time  $t_c$ . Therefore, the objective function has the following form:

$$\begin{aligned}
W(\mathbf{g}_{t_c} | \mathbf{x}, \mathbf{y}, \mathbf{g}_{t_c - \Delta t}) = & \underbrace{[\phi \circ (\mathbf{y} - \mathbf{X} \mathbf{g}_{t_c})]^\top \mathbf{C}^{-1} [\phi \circ (\mathbf{y} - \mathbf{X} \mathbf{g}_{t_c})]}_{\text{residual}} - \underbrace{\mathbf{g}_{t_c}^\top \Gamma_{gg}^{-1} \mathbf{g}_{t_c}}_{\text{smoothness of } g_{t_c}} + \underbrace{\mathbf{v}^\top \mathbf{H} \mathbf{g}_{t_c}}_{\text{non-negativity}} \\
& + \underbrace{\lambda (1 - \mathbf{1}^\top \mathbf{g}_{t_c})}_{\text{integral constraint}} + \underbrace{(\mathbf{g}_{t_c} - \mathbf{g}_{t_c - \Delta t})^\top \mathbf{C}_{gg}^{-1} (\mathbf{g}_{t_c} - \mathbf{g}_{t_c - \Delta t})}_{\text{smoothness among time steps}}
\end{aligned} \tag{1.11}$$

in which the first three terms are very similar to the terms used in Eq. 1.5, but in the residual term, instead of checking the overall residual of the simulation and the measurement, only the residual in a certain window is checked, which is defined by the weighting function  $\phi$ ; the fourth term comes from the constraint that the recovery rate must not exceed unity; the last term stems from the non-stationarity, where no large fluctuations are allowed compared to the previous step. This

objective function is minimized following the same procedure and the system of linear equations reads as:

$$\begin{bmatrix} (\Phi \circ \mathbf{X})^\top \mathbf{C}^{-1} (\Phi \circ \mathbf{X}) - \Gamma_{gg}^{-1} + \mathbf{C}_{gg}^{-1} & \mathbf{H}^\top & \mathbf{1} \\ \mathbf{H} & \mathbf{0} & \mathbf{0} \\ -\mathbf{1}^\top & \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{g}_{t_c} \\ \mathbf{v} \\ \lambda \end{bmatrix} = \begin{bmatrix} (\Phi \circ \mathbf{X})^\top \mathbf{C}^{-1} (\phi \circ \mathbf{y}) + \mathbf{C}_{gg}^{-1} \mathbf{g}_{t_c - \Delta t} \\ \mathbf{0} \\ -1 \end{bmatrix} \quad (1.12)$$

in which the matrix  $\Phi$  is a blown-up matrix of the weighting function  $\phi$ . This equation is solved in each window to obtain a changing transfer function  $g(\tau, t)$  along time.

### 1.4.3 In-channel transport with hyporheic exchange (Chapter 3: *Liao and Cirpka (2011)*)

Unlike in the bank-filtration, the model for solute transport in a river undergoing hyporheic exchange is a combination of the advection-dispersion equation (ADE) with a convolution integral expressing transient storage in the hyporheic zone:

$$\frac{\partial c}{\partial t} + v \frac{\partial c}{\partial x} - D \frac{\partial^2 c}{\partial x^2} = q_{he} \left( \int_0^t g(\tau) c(t - \tau) d\tau - c(t) \right) \quad (1.13)$$

subject to the initial and boundary conditions:

$$\begin{aligned} c(t = 0, x) &= 0, \\ \left( vc - D \frac{\partial c}{\partial x} \right) \Big|_{x=0} &= \frac{m}{A} \delta(t), \\ \lim_{x \rightarrow \infty} c(t, x) &= 0, \quad \forall t \end{aligned} \quad (1.14)$$

in which  $c [ML^{-3}]$  is the tracer concentration in the river channel;  $v [LT^{-1}]$  is the in-stream velocity;  $D [L^2 T^{-1}]$  is the in-stream dispersion coefficient;  $q_{he} [T^{-1}]$  is the hyporheic exchange rate coefficient, that is, the volume of exchanged water per time unit and volume of channel water;  $t [T]$  is the time after the tracer has been introduced into the system;  $x [L]$  is the distance from the injection point;  $A [L^2]$  is the cross-sectional area of the river at the injection point;  $m [M]$  is the injected mass of the tracer. The transfer function  $g(\tau) [T^{-1}]$  denotes the distribution of time that the water remains within the hyporheic zone before returning to the stream. Adopting the method of *Cirpka et al. (2007)*, the transfer function has no predefined shape and therefore regularization is needed. This equation is solved analytically in the Laplace domain, with numerical back transformation into the time domain. The solution in the Laplace domain has the form:

$$\tilde{c}(s, x) = \tilde{c}_0(s) \exp(-\alpha(s)x), \quad (1.15)$$

with the parameters:

$$\begin{aligned}\alpha(s) &= \frac{-v + \sqrt{v^2 + 4D(s + q_{he} - q_{he}\tilde{g}(s))}}{2D} \\ \tilde{c}_0(s) &= \frac{m}{A} \cdot \frac{1}{v + \alpha(s)D}\end{aligned}\quad (1.16)$$

in which the Laplace transformation is defined as:

$$\tilde{f}(s) = \int_0^\infty e^{-st} f(t) dt. \quad (1.17)$$

The transfer function  $g(\tau)$  can be either a travel-time distribution for a conservative tracer, which has an integral of 1, or a transfer function of a non-conservative tracer, where the zeroth moment may be smaller than one. In Chapter 3 (*Liao and Cirpka, 2011*) only a first-order decay is considered for the non-conservative tracer, but in the following a more complicated model with linear chemical transformation and two-site sorption is considered and accounted for in the derivation of transfer function. In the model, not only the transfer function  $g(\tau)$  needs to be estimated, but also the other parameters:  $v$ ,  $D$ , and  $q_{he}$ .

I define the parameter vector  $\mathbf{p} = [\mathbf{g}^\top, v, D, q_{he}]^\top$  (in the real application, I used  $\ln(v)$ ,  $\ln(D)$ , and  $\ln(q_{he})$  instead to avoid getting non-positive values), and the set of measured tracer concentrations in the stream as  $\mathbf{c}_{meas}$  for the discrete form of the breakthrough curve  $c(t)_{meas}$ ; then the corresponding objective function becomes:

$$W(\mathbf{p}|\mathbf{c}_{meas}) = \underbrace{(\mathbf{c}_{meas} - \mathbf{c}(\mathbf{p}))^\top \mathbf{C}^{-1} (\mathbf{c}_{meas} - \mathbf{c}(\mathbf{p}))}_{\text{residual}} - \underbrace{\mathbf{g}^\top \Gamma_{gg}^{-1} \mathbf{g}}_{\text{smoothness of } g} + \underbrace{\mathbf{v}^\top \mathbf{H} \mathbf{g}}_{\text{non-negativity}} \quad (1.18)$$

This form looks similar to Eq. 1.5, but the forward model  $\mathbf{c}(\mathbf{p})$  is nonlinear with respect to the parameters, which requires an iterative estimation method, such as the Gauss-Newton method. The system of linear equations:

$$\begin{bmatrix} \mathbf{J}_g^\top \mathbf{C}^{-1} \mathbf{J}_g - \Gamma_{gg}^{-1} & \mathbf{J}_g^\top \mathbf{C}^{-1} \mathbf{J}_{vDq} & \mathbf{H}^\top \\ \mathbf{J}_{vDq}^\top \mathbf{C}^{-1} \mathbf{J}_g & \mathbf{J}_{vDq}^\top \mathbf{C}^{-1} \mathbf{J}_{vDq} & \mathbf{0} \\ \mathbf{H} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ \mathbf{p}_{vDq} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_g^\top \mathbf{C}^{-1} \mathbf{c}_* \\ \mathbf{J}_{vDq}^\top \mathbf{C}^{-1} \mathbf{c}_* \\ \mathbf{0} \end{bmatrix} \quad (1.19)$$

in which  $\mathbf{c}_* := \mathbf{c}_{meas} - \mathbf{c}(\mathbf{p}_{last}) + \mathbf{J}\mathbf{p}_{last}$ ;  $\mathbf{J}_g$  and  $\mathbf{J}_{vDq}$  are the Jacobian matrices for  $\mathbf{g}$  and  $[v, D, q_{he}]^\top$ , respectively, that is, the matrices of partial derivatives of all simulated measurements with respect to all parameters. Because the parameter set  $\mathbf{p}_{last}$  is used in the calculation of  $\mathbf{c}_*$ , the iterative procedure is repeated until the parameter set converges. The usage of the joint fit of both conservative and reactive tracer breakthrough curves provides the possibility to estimate the transfer function  $g(\tau)$  and the other parameters at the same time.

### 1.4.4 Governing equation for the transport of the Raz-Rru system in the hyporheic zone

Laboratory studies have indicated that considering linear decay as only reactive process of the reactive tracer is not enough, especially when using the Raz-Rru reactive-tracer system. In order to understand a little more about the transport and reaction of the tracers undergoing hyporheic exchange, a more complicated model is proposed. Before involving decay and two-site sorption process into the transport model, I focus on one tube in the hyporheic zone, where the infiltrating river water ex-filtrates back to the river channel after remaining in the subsurface for a distribution of travel time. The corresponding transfer functions  $g_{Raz}$  and  $g_{Rru}$  are correlated with the travel time distribution  $g_{Fluo}$  and other parameters via following transport equation in the hyporheic zone:

$$\begin{aligned}
 R_{Raz} \frac{\partial c_{Raz}^{hz}}{\partial \tau} + K_{kin,Raz} \frac{\partial c_{Raz}^*}{\partial \tau} + \frac{\partial c_{Raz}^{hz}}{\partial \tau^*} &= -(\lambda_{Raz} + \lambda_{Raz-Rru}) c_{Raz}^{hz} \\
 \frac{\partial c_{Raz}^*}{\partial \tau} &= k_{mt,Raz} (c_{Raz}^{hz} - c_{Raz}^*) \\
 c_{Raz}^{hz}(\tau, 0) &= b_{Raz} \delta(\tau)
 \end{aligned} \tag{1.20}$$

and the governing equations for the reactive product Rru are:

$$\begin{aligned}
 R_{Rru} \frac{\partial c_{Rru}^{hz}}{\partial \tau} + K_{kin,Rru} \frac{\partial c_{Rru}^*}{\partial \tau} + \frac{\partial c_{Rru}^{hz}}{\partial \tau^*} &= \lambda_{Raz-Rru} c_{Raz}^{hz} - \lambda_{Rru} c_{Rru}^{hz} \\
 \frac{\partial c_{Rru}^*}{\partial \tau} &= k_{mt,Rru} (c_{Rru}^{hz} - c_{Rru}^*) \\
 c_{Rru}^{hz}(\tau, 0) &= b_{Rru} \delta(\tau)
 \end{aligned} \tag{1.21}$$

in which  $c_i^{hz}$ ,  $i = Raz$  or  $Rru$  [ $ML^{-3}$ ] are the tracer concentration of Raz and Rru in the hyporheic zone;  $c_i^*$ ,  $i = Raz$  or  $Rru$  [ $ML^{-3}$ ] are the tracer concentration at the kinetically sorbing sites;  $b_i$ ,  $i = Raz$  or  $Rru$  are the mass of the tracer introduced into the system;  $\tau$  [ $T$ ] is the travel time and  $\tau^*$  [ $T$ ] is the travel time of the conservative tracer. Solving these equations results in the concentration of Raz:  $c_{Raz}^{hz}$ , Rru transformed from Raz:  $c_{Raz,Rru}^{hz}$ , and Rru originates from existing Rru:  $c_{Rru,Rru}^{hz}$ , therefore the transfer functions are a weighted integral of these concentrations with the travel-time distribution  $g_{Fluo}$ :

$$\begin{aligned}
 g_{Raz}(\tau) &= \int_0^\infty g_{Fluo}(\tau^*) c_{Raz}^{hz}(\tau, \tau^*) d\tau^* \\
 g_{Raz,Rru}(\tau) &= \int_0^\infty g_{Fluo}(\tau^*) c_{Raz,Rru}^{hz}(\tau, \tau^*) d\tau^* \\
 g_{Rru}(\tau) &= \int_0^\infty g_{Fluo}(\tau^*) c_{Rru,Rru}^{hz}(\tau, \tau^*) d\tau^*
 \end{aligned} \tag{1.22}$$

which are also solved analytically in the Laplace domain and transformed back into the time domain numerically.

### 1.4.5 In-channel transport with hyporheic exchange including two-site sorption (Chapter 4: *Liao et al. (2013)*)

With the derivation from Section 1.4.4, the governing equations for tracer fluorescein (Fluo) and resazurin (Raz) remain the same as Eq. 1.13:

$$\frac{\partial c_i}{\partial t} + v \frac{\partial c_i}{\partial x} - D \frac{\partial^2 c_i}{\partial x^2} = q_{he} \left( \int_0^t g_i(\tau) c_i(t - \tau) d\tau - c_i \right), \quad i = \text{Fluo or Raz} \quad (1.23)$$

where the governing equation for tracer esorufin (Rru) reads as:

$$\frac{\partial c_{Rru}}{\partial t} + v \frac{\partial c_{Rru}}{\partial x} - D \frac{\partial^2 c_{Rru}}{\partial x^2} = q_{he} \left( \int_0^t (g_{Raz-Rru}(\tau) c_{Raz}(t - \tau) + g_{Rru}(\tau) c_{Rru}(t - \tau)) d\tau - c_{Rru}(t) \right) \quad (1.24)$$

in which the source term in the convolution is composed of two parts: one originates from the existing Rru in the system, and the other is newly transformed from Raz through undergoing the hyporheic exchange. These equations are again solved analytically in the Laplace domain, followed with a numerical back transformation into the time domain. In Eq. 1.23 and 1.24, the parameters to estimate are: the travel-time distribution, i.e. the transfer function of Fluo  $g_{Fluo}[T^{-1}]$ ; the in-channel characteristic parameters  $v[LT^{-1}]$ ,  $D[L^2T^{-1}]$  and  $q_{he}[T^{-1}]$ ; the sorption coefficients  $R_{Raz}[-]$ ,  $R_{Rru}[-]$  from the equilibrium sorption and  $K_{kin,Raz}[-]$ ,  $K_{kin,Rru}[-]$ ,  $k_{mt,Raz}[T^{-1}]$ ,  $k_{mt,Rru}[T^{-1}]$  from the kinetic sorption; the decay coefficients  $\lambda_{Raz}[T^{-1}]$ ,  $\lambda_{Raz-Rru}[T^{-1}]$ , and  $\lambda_{Rru}[T^{-1}]$ . A joint fit of all three tracers provides the chance to estimate all parameters all together. The details of how the transfer function  $g_{Raz}[T^{-1}]$  and  $g_{Rru}[T^{-1}]$  is correlated to the travel-time distribution  $g_{Fluo}$  through all the parameters can be found in Chapter 4 (*Liao et al., 2013*). The objective function remains similar as Eq. 1.18:

$$W(\mathbf{p}|\mathbf{c}_{meas}) = \underbrace{(\mathbf{c}_{meas} - \mathbf{c}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{c}_{meas} - \mathbf{c}(\mathbf{p}))}_{\text{residual}} - \underbrace{\mathbf{g}^T \Gamma_{gg}^{-1} \mathbf{g}}_{\text{smoothness of } g} + \underbrace{\mathbf{v}^T \mathbf{H} \mathbf{g}}_{\text{non-negativity}} \quad (1.25)$$

and the system of linear equations are:

$$\begin{bmatrix} \mathbf{J}^T \mathbf{C}^{-1} \mathbf{J} - \Gamma_{pp}^{-1} & \mathbf{H}^T \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{J}^T \mathbf{C}^{-1} (\mathbf{c}_{meas} + \mathbf{J} \mathbf{p}_{last} - \mathbf{c}(\mathbf{p}_{last})) \\ \mathbf{0} \end{bmatrix} \quad (1.26)$$

in which  $\Gamma_{pp}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Gamma_{gg}^{-1} \end{bmatrix}$ , and  $\mathbf{J}$  is the Jacobian matrix for all the parameters. The difference between this model and the one in previous section is that there are more parameters to be estimated and joint fit of three tracers together. The difficulty lies in the estimation of the transfer functions  $g_{Raz}$  and  $g_{Rru}$ , which appears to be a weighted integration of the travel-time distribution  $g_{Fluo}$  and the concentration change from a pulse injection into the hyporheic zone, which is calculated by assuming a pure advection transport in the experimental time and the travel time.

In general, I have listed governing equations, objective functions and its corresponding system of linear equations from each chapter/work. The existing method from *Cirpka et al. (2007)* appears as a deconvolution. Under non-stationary conditions, this deconvolution is applied in a small time span over which stationary conditions are assumed. In the river tracer test model, the convolution integral becomes the source term embedded into the advection-dispersion equation of in-stream transport. This model is then extended to account for more parameters characterizing decay and two-site sorption of the reactive tracers in the hyporheic zone.

## 1.5 Outline

The main objective of this work is to estimate the travel-time distributions of the river water in the subsurface and some other parameters characterizing solute transport in streams. The outline of this work is shown in Tab. 1.1. Two situations are addressed here: bank-filtration and hyporheic exchange. For both situations, stationary and non-stationary conditions are discussed. The basic of this work is the paper of *Cirpka et al. (2007)*, in which a non-parametric deconvolution method is proposed to estimate the travel-time distribution of bank-filtration systems under stationary conditions. I have extended it to the non-stationary case (Chapter 2). On the other hand, the solute transport in the river channel undergoing hyporheic exchange is proposed under stationary conditions (Chapter 3), and extended to a more complicated model to include decay and sorption of the reactive tracers in Chapter 4. Chapter 5 is an overall comparison of the model performance from different scales.

	stationary	non-stationary
Bank-filtration	<i>Cirpka et al. (2007)</i>	Chapter 2
Hyporheic exchange	Chapter 3; Chapter 4; Chapter 5	future work

Table 1.1: The overview of the chapters in this work. Chapter 3, Chapter 4 and Chapter 5 address the problem of the hyporheic exchange under stationary condition, whereas Chapter 2 deals with the bank-filtration system under non-stationary condition.

An outline of this work is as following:

**Chapter 1:** A general introduction of the problem at hand: bank-filtration and hyporheic exchange existing methods and models in the literature. A summary of the equations used in this work are listed as well.

**Chapter 2:** This chapter deals with estimating river-to-groundwater travel-time distribution in non-stationary bank-filtration. The available non-parametric deconvolution method is extended to non-stationary conditions by applying the sliding window technique. The non-stationary method provides better agreement between the simulation and the measurements in both overall comparison and in peak event time point. This work is submitted to *Journal of Hydrology*.



**Chapter 3:** From this chapter on, an extensive analysis of hyporheic exchange in river channels is conducted. This chapter starts with a synthetic tracer test, where the non-parametric deconvolution method is adopted to the estimation of the non-local source term in the advection-dispersion equation. The joint fit using conservative tracer and reactive tracer provides the opportunity to distinguish between in-stream dispersion and hyporheic exchange. This work is published in *Water Resources Research* (Liao and Cirpka, 2011).

**Chapter 4:** An application of the proposed model to a real field data is shown in this chapter. Furthermore, I extend the analysis of hyporheic processes to include several linear transformation processes and two-site sorption of the reactive tracers in the hyporheic zone. A joint fit of all three tracers facilitates the estimation of the parameters and a shape-free travel-time distribution. This work is published in *Water Resources Research* (Liao et al., 2013).

**Chapter 5:** The applicability of the proposed model is tested in this chapter on different scales: a column experiment at the scale of 30 cm, a field gravel-bar experiment at the scale of 2m, and a stream tracer test covering 1210m. The results from all three scales show quite good fit of the model and give reasonable parameter values. This work has been presented as an oral presentation at the Fall Meeting of the American Geophysical Union (AGU), 2012.

**Chapter 6:** This chapter summarizes the achievements from the above works, and discusses limitations and prospects.



## Chapter 2

# Non-Stationary Nonparametric Inference of River-to-Groundwater Travel-Time Distributions

Z. Liao, K. Osenbrück, O.A. Cirpka: Non-stationary nonparametric inference of river-to-groundwater travel-time distributions. *J. Hydrol.* (submitted), 2014.

### Abstract

The travel-time distribution between rivers and groundwater observation points and the mixing of freshly infiltrated river water with groundwater of other origin is of high relevance in riverbank filtration. These characteristics usually are inferred from the analysis of natural-tracer time series, typically relying on a stationary input-output relationship. However, non-stationarity is a significant feature of the riparian zone causing time-varying river-to-groundwater transfer functions. We present a non-stationary extension of nonparametric deconvolution by performing stationary deconvolution with windowed time series, enforcing smoothness of the determined transfer function in time and travel time. The nonparametric approach facilitates the identification of unconventional features in travel-time distributions, such as broad peaks, and the sliding-window approach is an easy way to accommodate the method to dynamic changes of the system under consideration. By this, we obtain time-varying signal-recovery rates and travel-time distributions, from which we derive the mean travel time and the spread of the distribution as function of time. We apply our method to electric-conductivity data collected at River Thur, Switzerland, and adjacent piezometers. The non-stationary approach reproduces the groundwater observations significantly better than the stationary one, both in terms of overall metrics and in matching individual peaks. We compare characteristics of the transient transfer function to base flow which indicates shorter travel times at higher river stages.

### 2.1 Introduction

Bank filtration, in which river water infiltrates into the river banks and is extracted at some distance, plays a major role for water supply in many countries (e.g., *Sontheimer*, 1980; *Kuehn and Mueller*,

2000; Ray *et al.*, 2002; Schubert, 2002). The water passage through the banks filters particles such as pathogenic bacteria (e.g., Weiss *et al.*, 2005a) and facilitates natural cycling of nutrients as well as removal of biodegradable contaminants (e.g., Doussan *et al.*, 1997; Hoppe-Jones *et al.*, 2010). Thus, bank filtration may partially replace technical water-purification systems. Its effectiveness highly depends on the travel time from the river to the well, which is often used as a proxy for the assessment of infiltrating river-water quality. In the regulations of several countries, inner protection zones for groundwater wells are defined by isochrones, e.g., in Germany by 50 days of travel time (DVGW, 1995) and in Switzerland by 10 days (BUWAL, 2004). Even though these protection zones regulate land use, the agencies recommend that infiltrating river water also exceeds the required travel time before being extracted. Thus, determining the travel times from rivers to observation and pumping wells is of vital importance in the management of river-fed groundwater resources.

The traditional technique of determining travel times is by tracer tests. In artificial-tracer experiments an easily detectable, conservative, and harmless compound, that is not yet present in the system, is injected into the stream and measured at all observation wells of interest (Leibundgut *et al.*, 2009). These tests directly provide the volumetric fraction of river water in the well and the travel-time distribution. At larger rivers, however, very large tracer masses need to be injected, the largest fraction of which is not undergoing infiltration. Also, artificial-tracer tests yield only information about river-groundwater exchange at the time of the experiment.

As an alternative, natural tracers bearing a time signal have been analyzed. Besides concentrations of dissolved gases, such as radon (e.g., Hoehn and von Gunten, 1989), noble gases (e.g., Stute *et al.*, 1997; Massmann *et al.*, 2008), and chlorofluorocarbons (e.g., Beyerle *et al.*, 1999; Darling *et al.*, 2010), time series of naturally fluctuating properties, such as water isotopes (e.g., Stichler *et al.*, 1986; Hunt *et al.*, 2005), temperature (e.g., Anderson, 2005) and specific electric conductivity (e.g., Cox *et al.*, 2007; Cirpka *et al.*, 2007) can be used to infer river-to-groundwater travel times. Specific advantages of using time series of physical properties are their low cost of detection and their continuous quality, which can be recorded over years. This becomes particularly important if high-frequency time series are required due to non-stationary flow or boundary conditions, as is the case in most river systems. In the present study, we will focus on non-stationary travel time distributions and therefore restrict the analysis to the interpretation of continuously measured fluctuating specific electrical conductivity.

The most profound way of analyzing tracer time series in a bank-filtration system would be by calibrating a mechanistic, spatially explicit coupled river-groundwater flow-and-transport model (e.g., Malaguerri *et al.*, 2013). This is often not possible, because the exact river bathymetry, distribution of hydraulic subsurface parameters, and boundary conditions are not known. Under such conditions, the target of tracer-data analysis is restricted to identifying a linear input-output relationship between the measured times series in the river (input) and the observation or pumping well (output).

Many approaches of linear input-output relationships can be casted as convolution models, in which the input signal is convoluted with the transfer function to obtain the output signal. If average values of the signals differ, an additional systematic offset may be needed. The transfer function is also denoted impulse-response function or Green's function. Its integral is the signal-recovery rate, and the normalized transfer function (integrating to unity) is the travel-time distribution.

An easy approach of identifying a single characteristic transfer time between the two time series is by cross-correlation, in which the time shift with the highest correlation coefficient is interpreted as the effective travel time (e.g., Sheets *et al.*, 2002). In order to account for smoothing of the input signal by the transfer process, the input signal may be filtered before cross-correlation, and the optimal combination of filter width and time shift is sought for (e.g., Vogt *et al.*, 2009). Together with the linear-regression coefficients obtained for the optimal combination of filter width and time shift, the results of this analysis

may be interpreted as parametric travel-time distribution amended by a recovery rate and a systematic offset in the time series.

A simple and parsimonious approach of inferring a non-negative transfer function is by assuming a non-negative distribution, such as the log-normal, gamma, or inverse Gaussian distribution, which is fully described by a few shape parameters, and fitting these parameters (e.g., Luo *et al.*, 2006; Maloszewski and Zuber, 1993). Fitting one-dimensional transport models with constant coefficients would also fall into this category (for a comprehensive overview of 1-D transport solutions see Toride *et al.*, 1993), but the one-dimensional interpretation of data resulting from complex three-dimensional transport may lead to a misguided assignment of physical properties to the system under investigation.

Even when parametric travel-time distributions without direct relationship to a specific one-dimensional transport problem are chosen as transfer functions, the inferred solutions are restricted to predefined shapes. Multi-modal distributions, or distributions allowing broad peaks and long tails are mostly not tested. However, the complexity of riverbed morphology and sediments could facilitate such transfer functions. To overcome this difficulty, our workgroup has developed non-parametric methods of estimating non-negative, smooth transfer functions and applied them to bank-filtration problems (Cirpka *et al.*, 2007; Vogt *et al.*, 2010), stream-to-stream tracer tests (Payn *et al.*, 2008), and the identification of hyporheic travel-time distributions (Liao and Cirpka, 2011; Liao *et al.*, 2013). Recently, McCallum *et al.* (2014) suggested a similar method in which the smoothness regularization of Cirpka *et al.* (2007) has been replaced by applying the singular-value-decomposition based pseudo-inverse in the solution of the resulting close-to-singular system of linear equations.

All approaches discussed above are based on the assumption of stationarity. That is, the response of the system to a unit-pulse input signal is assumed to be always identical. This is in contrast to riparian systems being dynamic. As a consequence, the transfer functions inferred by stationary methods reflect a weighted time average of the true dynamic system behavior at best.

The problem of non-stationarity occurs in various applications of time-series analysis. Many methods of handling non-stationarity are extensions of stationary approaches. A way of circumventing non-stationarity is by applying a sliding window to the time series, resulting in smaller sections of the data, which are then analyzed by a suitable stationary method. The overall scheme is non-stationary because the stationary analysis is repeated for each window, so that the inferred parameters can change with the sliding of the window. Small window sizes allow for stronger non-stationarity, and applying overlapping windows leads to smoother transitions of the inferred parameters. Boker *et al.* (2002) combined cross-correlation and the sliding-window method to obtain the dynamic properties of a system from associated time series.

In a similar way, Schmidt *et al.* (2012) proposed an extension of the dynamic-time-warping method (Dürrenmatt *et al.*, 2013; Berndt and Clifford, 1994), where variations in the time lag between time series of electrical conductivity were derived from an optimal (warping) path within the correlation matrix.

All these dynamic methods have improved the understanding of river-groundwater systems in terms of the mean travel time, but dynamic non-parametric deconvolution, which considers both the potential unconventional shape of the transfer function and the non-stationary property of the river-to-groundwater system, has not yet been conducted.

In the present work, we propose a non-stationary extension of the nonparametric deconvolution method of Cirpka *et al.* (2007) to determine travel-time distributions between rivers and observation wells. Non-stationarity is achieved by applying the sliding-window technique with local stationarity for the windowed input-to-output relationship. We choose short overlapping windows and penalize rapid changes in the determined transfer function, yielding a continuously and smoothly changing travel-time distribution. Like in the approach of Cirpka *et al.* (2007), we do not predefine the functional shape of the

travel-time distribution but require smoothness for regularization. To the best of our knowledge, this is the first time that dynamic, nonparametric travel-time distributions are determined from continuous tracer time series.

We apply our method to a more than one-year long time series of electric conductivity (EC) collected at River Thur, Switzerland, and adjacent observation wells. This field site has been intensively explored in the past years (*Cirpka et al.*, 2007; *Coscia et al.*, 2011; *Schneider et al.*, 2011, among others). From the time-dependent transfer functions estimated by the method, we infer the recovery rate, the mean travel time, and the spread of the travel-time distribution from the river to the piezometers as function of time and analyze their potential relationships with river stage.

## 2.2 Method

### 2.2.1 Overview of the Method

Before performing any deconvolution, we detrend the data by subtracting the baseline of each time series, thus accounting for systematic differences between the river and observation well time series caused by measurement bias or a physical/chemical process. The approach used for that is described in Section 2.2.2.

Figure 2.1 schematically illustrates the method at a central time from a window  $t_c$ . The transfer function  $g(\tau, t_c)$ , which depends on the time  $t_c$  and the travel time  $\tau$ , is estimated based on an iterative procedure, which balances misfit of the simulated output function (objective function 2) and the difference of the estimated current transfer function  $g(\tau, t_c)$  to that of the previous time  $g(\tau, t_c - \Delta t)$  (objective function 1), where  $\Delta t$  is the time increment of window step, which is equal to time increment in the measurements. More symbols and definitions can be found in Section 2.2.3. A weighting function is applied to the windowed simulation and output signal, so that differences at the edge of the window are diminished. Upon convergence, the same procedure is applied to the next step  $t_c + \Delta t$  until the shifted window reaches the end of the time series. To obtain an initial guess of the transfer function for the first time step and a reference to compare to, we apply the stationary method of *Cirpka et al.* (2007) to the entire detrended time series. Then a log-normal distribution with the same mean and variance is chosen as the initial guess, to avoid problems that may arise from the non-negativity constraint used in non-parametric deconvolution.

Because the initial guess for the first window is obtained from the stationary method, we would have to account for a burn-in period until the estimated transfer functions become independent of the initial guess. To partially overcome this problem, we first slide the window forward and then backward, making the transfer function of the first window depend on the following transfer functions rather than the stationary solution.

### 2.2.2 Data Detrending

The time series of river and observation well are detrended by finding the baselines of the original EC data and taking the differences of the measurements and the baselines. A moving average filter is first applied to the original signals to remove outliers. The local maxima are then defined by the peak-finding method of *Dugge* (2012), explained in 2.B, combined with manual selection. At last, the chosen maxima are smoothly interpolated by a cubic, monotonicity-preserving spline method. Detrending is finished by

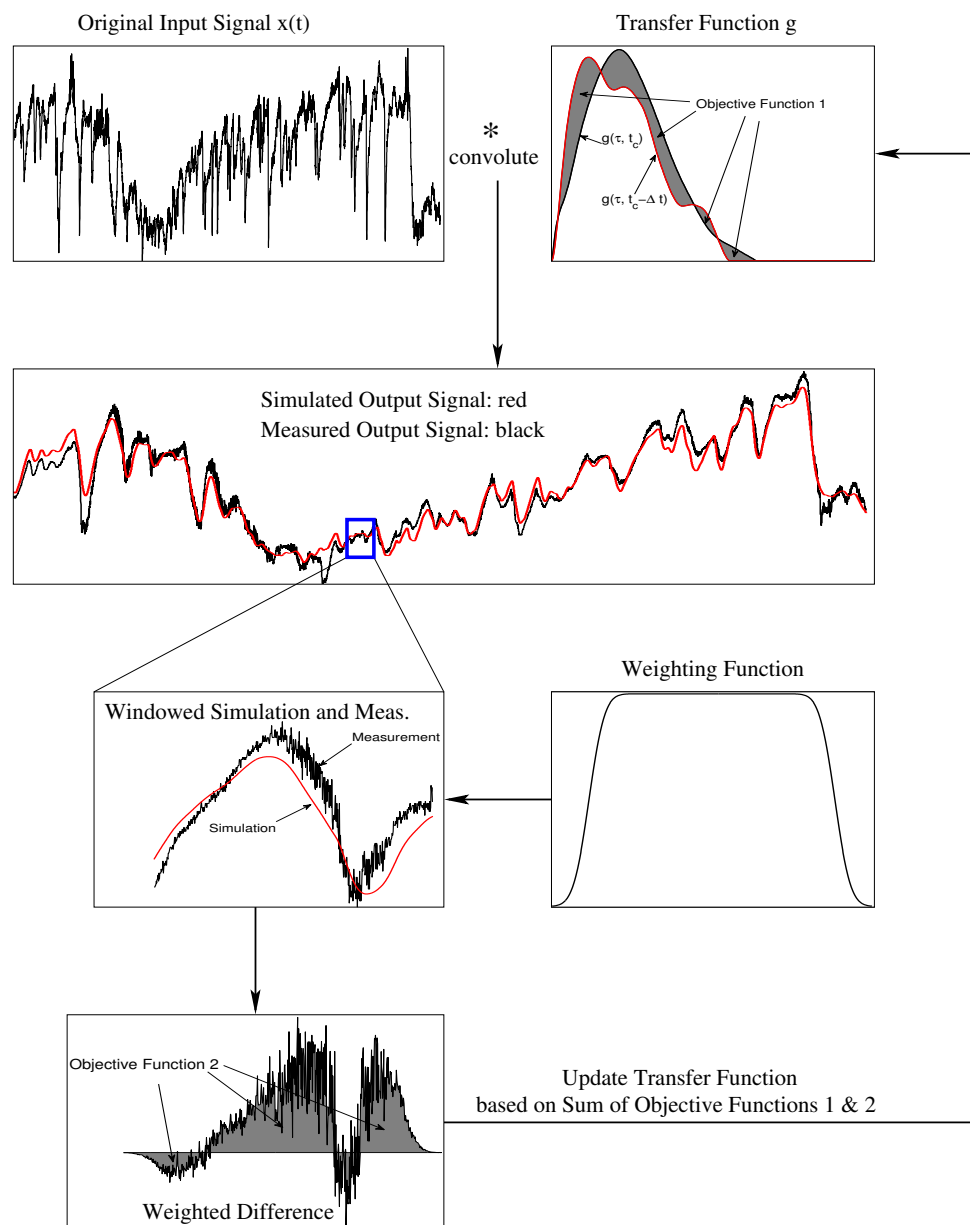


Figure 2.1: Overview of the method at time step  $t_c$ . Within each iteration step, the current guess of the transfer function  $g(\tau, t_c)$  is convoluted with the detrended input signal. The simulated and measured output signals are windowed, weighted by the weighting function, and compared to each other to obtain part 2 of the objective function. As first contribution to the objective function, the method penalized differences between the current transfer function  $g(\tau, t_c)$  and that of the previous time step  $g(\tau, t_c - \Delta t)$ . The transfer function  $g(\tau, t_c)$  is then updated based on the minimization of the sum of the objective functions.

subtracting the baseline from the original EC data, and the detrended time series are subsequently used

as the input and output signals in the model.

Figure 2.2 shows the original EC data and the inferred baselines. The maximum line is chosen as the baseline because rainfall events normally cause a drop in the measured EC time series. These event-based drops are considered the signals to be analyzed. From the baselines, we notice a period of low EC-values in the river around May in both 2004 and 2005. The observations in the well show similar patterns, but not to the same extent as those in the river. We conjecture that snow melt in the upper catchment of the river causes this drop (see also *Cirpka et al., 2007*). During the other time periods, the baselines of the input signals are a little lower than the baseline of the output signals for well 1 and 3, whereas the baseline of well 2 and the river are basically identical outside of the presumed snow-melt period. Differences in the baselines may be attributed to different mixing ratios between fresh river infiltrate and old groundwater, but may as well be caused by different intensities of water-rock interactions along the flow lines leading to the individual well screens, located at different distance to the river and depth.

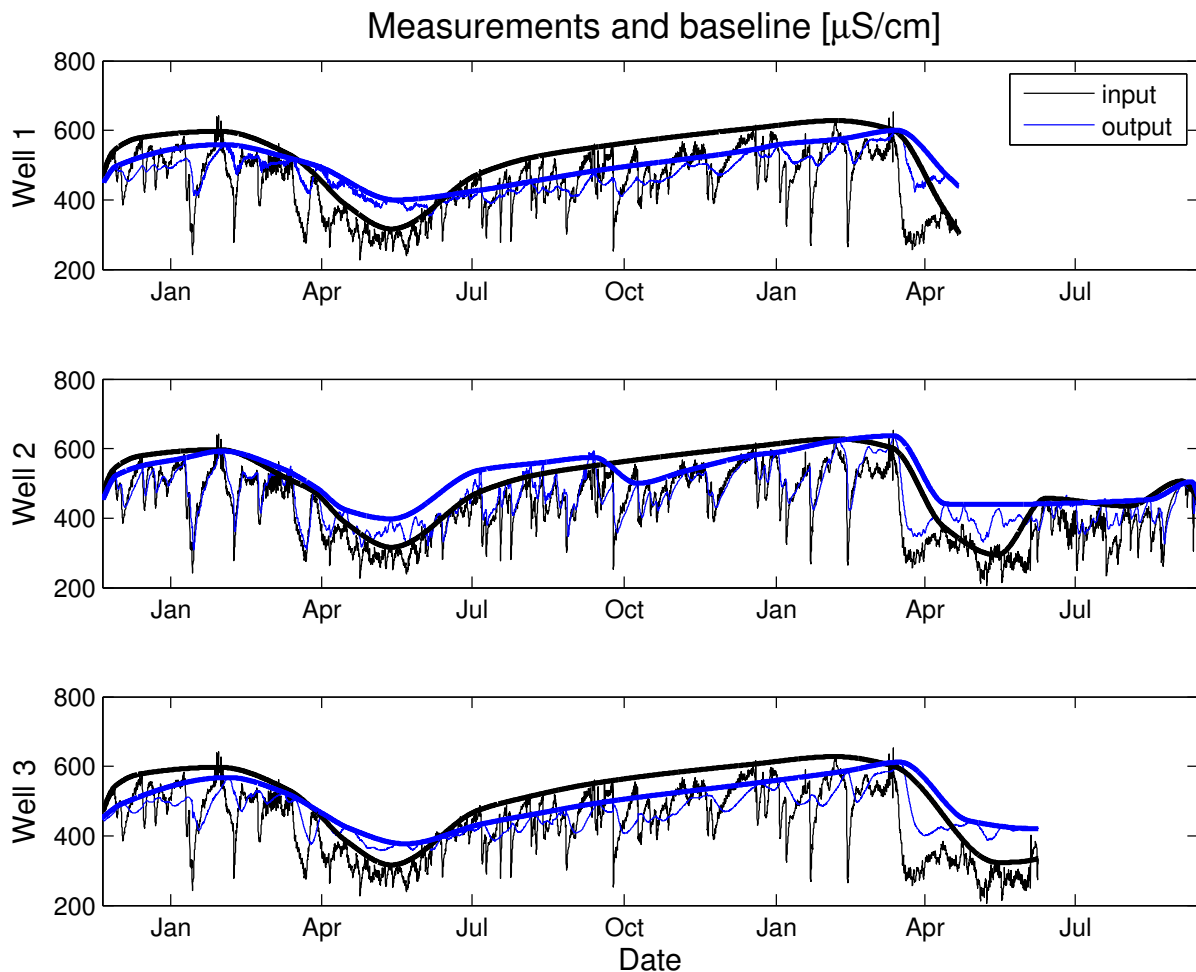


Figure 2.2: Original EC data and the baselines for well 1, 2 and 3. Black: EC data in the river; blue: EC data in the wells.



### 2.2.3 Principles of Auto-Correlated Deconvolution with a Sliding Window

The in- and output-signal time series  $x(t)$  and  $y(t)$ , respectively, of a linear, stationary system are related to each other by the convolution integral:

$$y(t) = \int_0^{\infty} g(\tau)x(t-\tau)d\tau, \quad (2.1)$$

in which  $t [T]$  is time,  $g(\tau) [T^{-1}]$  denotes the stationary transfer function, or impulse-response function of the system, and  $\tau [T]$  is the time offset.

Under non-stationary conditions, the transfer function  $g(\tau)$  varies with time. We make the transfer function  $g(\tau, t_c)$  dependent on a central time  $t_c [T]$ . In the following, we denote the modeled output signal  $f(t)$ , and the base lines for input and output signals  $x_{BL}(t)$  and  $y_{BL}(t)$  respectively. Depending on  $t_c$ , we now get different estimates of  $f(t)$  because  $g(\tau, t_c)$  changes with  $t_c$ :

$$f(t, t_c) = \int_0^{\infty} g(\tau, t_c)(x(t-\tau) - x_{BL}(t-\tau))d\tau + y_{BL}(t). \quad (2.2)$$

The purpose of deconvolution is to infer the non-stationary transfer function  $g(\tau, t_c)$  from the in- and output time series  $x(t)$  and  $y(t)$ . Towards this end, we study only fractions of the in- and output signals, and determine  $g(\tau)$  for these fractions. This is implemented by weighting the residual  $\varepsilon(t, t_c)$  between the modeled and measured output signals with a window-function  $\phi(t, t_c) [-]$ , centered about the central time  $t_c$  and dropping to zero for times significantly different to  $t_c$ :

$$\varepsilon(t, t_c) = \phi(t, t_c)(f(t, t_c) - y(t)), \quad (2.3)$$

which is minimized for each discrete value of  $t_c$  in an extended least-square sense with two smoothness constraints for  $g(\tau, t_c)$ : one in  $\tau$ , and the other in  $t_c$ . This leads to the following objective function:

$$W(g(\tau, t_c)|g(\tau, t_c - \Delta t_c)) = \frac{1}{\Delta t} \int_0^{\infty} \frac{\varepsilon^2(t, t_c)}{\sigma_y^2} dt + \frac{1}{2\theta} \int_0^{\infty} \left( \frac{dg(\tau, t_c)}{d\tau} \right)^2 d\tau + a \int_0^{\infty} (g(\tau, t_c) - g(\tau, t_c - \Delta t_c))^2 d\tau, \quad (2.4)$$

subject to the non-negativity constraint:

$$g(\tau, t_c) \geq 0 \quad \forall \tau, \quad (2.5)$$

and a second constraint enforcing that the integral of the transfer function does not exceed unity:

$$\int_0^{\infty} g(\tau, t_c) d\tau \leq 1 \quad \forall t_c \quad (2.6)$$

in which the first term of Eq. (2.4) is the windowed likelihood term, weighted by the squared measurement error  $\sigma_y$  (units identical to those of the signals) of the output signal  $y(t)$ ; the second term penalizes derivatives of the transfer function using the smoothness parameter  $\theta [T^{-3}]$ ; whereas the third term penalizes differences between the determined transfer function  $g(\tau, t_c)$  at the given central time  $t_c$  to the last estimates at central time  $t_c - \Delta t_c$ , where  $a$  is a smoothness parameter regarding the non-stationarity. The

two constraints prevent unphysical behavior, namely negative concentrations and a tracer recovery larger than unity.

The smoothness constraint of  $g(\tau, t_c)$  in  $\tau$  has already been discussed by *Cirpka et al.* (2007) for stationary nonparametric deconvolution. The latter authors posed the constraint as geostatistical regularization with linear semi-variogram function. Without this constraint, strong fluctuations of  $g(\tau, t_c)$  in  $\tau$  are caused by overfitting. Enforcing that  $g(\tau, t_c)$  does not vary too dramatically with time  $t_c$  has been inspired by an identical constraint within dynamic harmonic regression (*Young et al.*, 1999), therein applied to time-dependent sine- and cosine base functions. This constraint is particularly necessary when the input data exhibit periods with hardly any signal to be transferred. Over these periods, our approach makes the transfer function keeping to the last meaningful function rather than being modified to map spurious noise in the signals. Finally, considering windowed time series has been done in windowed cross-correlation (*Boker et al.*, 2002), among other techniques of time-series analysis. An inherent difficulty of the sliding window method, however, lies in choosing the appropriate window shape, size, and increment.

Similar to dynamic harmonic regression (*Young et al.*, 1999), we minimize the difficulties of a burn-in period in the smoothness of  $g(\tau, t_c)$  with respect to  $t_c$  by first marching in the positive  $t_c$ -direction, conditioning the estimated parameters on the previous estimate, followed by a simulation, in which we march in the negative  $t_c$ -direction, now conditioning on  $g(\tau, t_c + \Delta t_c)$ . Section 2.2.5 contains more details of the discrete inference procedure.

## 2.2.4 Choice of the Window Function

The weighting function of the sliding window should have the following properties: (a) the starting and ending points must have a weight of 0; (b) the maximum weight should be 1; (c) the rising and falling limbs should be smooth to prevent high-frequency artifacts introduced by windowing. These requirements are met by the following function used in our study:

$$\phi(t, t_c, w, s) = \Phi_{cdf}(t, t_c - w/2, s) - \Phi_{cdf}(t, t_c + w/2, s), \quad (2.7)$$

in which  $\Phi_{cdf}(x, \mu, s)$  is the cumulative distribution function of a Gaussian distribution:

$$\Phi_{cdf}(x, \mu, s) = \frac{1}{2} \left[ 1 + \operatorname{erf} \left( \frac{x - \mu}{\sqrt{2}s^2} \right) \right], \quad (2.8)$$

$w [T]$  is the width of the window, and  $s [T]$  is the standard deviation of the Gaussian distribution, determining how smoothly  $\phi(t, t_c, w, s)$  navigates between zero and one. An example of  $\phi(t, t_c, w, s)$  is shown in Fig. 2.3.

By choosing a smooth window function, a strong signal gradually influences the determined transfer function when it moves into the window; this prevents too rapid changes. The width  $w$  of the window should be as small as possible to obtain the full dynamics of the system, but must be bigger than the length of the transfer function. The smoothness parameter  $s$  must be considerably smaller than  $w$ .

A suitable alternative to our window function would be the Tukey windows, which extends a rectangular window by half cosine-functions leading to a smooth transition from zero to one, and the other way round, over a finite distance (*Tukey*, 1967).

## 2.2.5 Discrete Parameter-Estimation Procedure

Upon discretization, the detrended input and output signals become vectors  $\mathbf{x}$  and  $\mathbf{y}$  of size  $n_x \times 1$  and  $n_y \times 1$ , respectively. In the following, we assume that the data are collected at regular time intervals

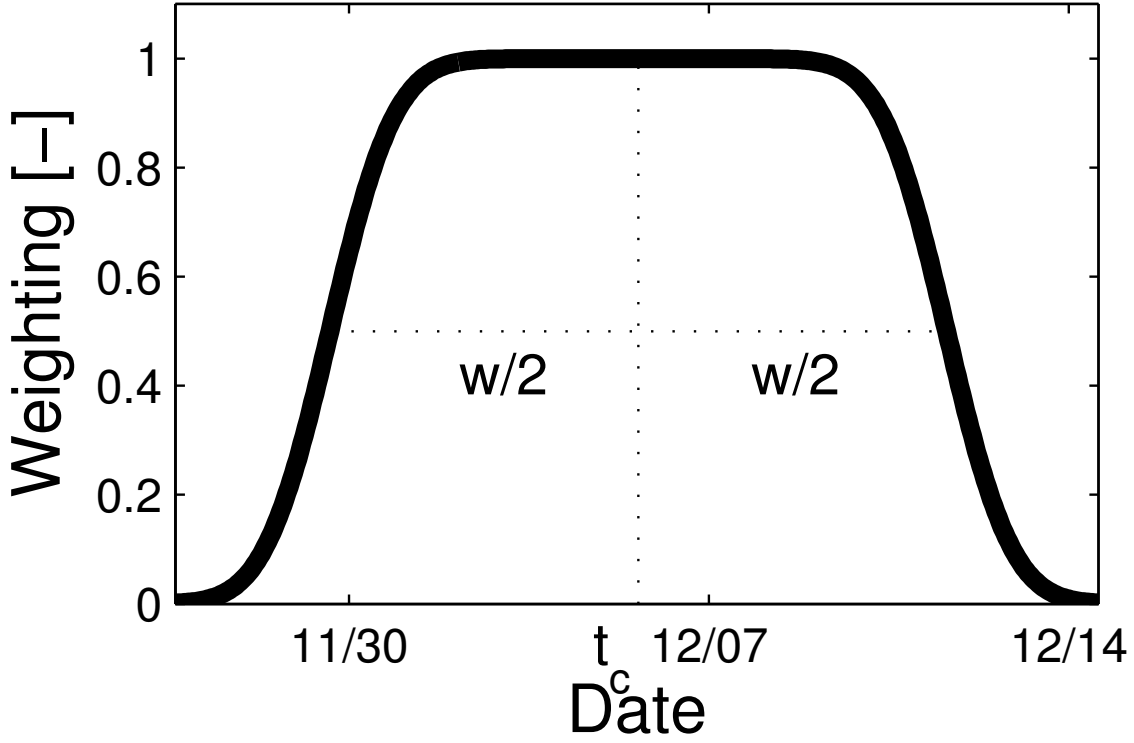


Figure 2.3: Weighting function according to Eq. (2.7).  $t_c$ : center point,  $w$ : width of the window. The rising and falling limbs are determined by the Gaussian cumulative distribution function.

$\Delta t$ , which are identical for the in- and output signals. Then, the transfer function at central time  $t_c$  is also a vector  $\mathbf{g}(t_c)$  of discrete values spaced by  $\Delta\tau = \Delta t$ . We truncate this vector after  $n_g$  elements, when the coefficients are expected to be negligible. Then, the discrete (detrended) model becomes:

$$\mathbf{f}(\mathbf{g}) = \mathbf{X}\mathbf{g}, \quad (2.9)$$

in which  $\mathbf{X}$  is a  $n_y \times n_g$  Toeplitz matrix, generated from  $\mathbf{x}$  by  $X_{ij} = \Delta\tau x_{i-j+1}$  (e.g., *Cirpka et al.*, 2007).

The objective function of Eq. (2.4), subject to the inequality constraints of Eq. (2.5) and Eq. (2.6), now reads in discrete form for each window:

$$W(\mathbf{g}_t | \mathbf{g}_{t-1}, \mathbf{y}, \mathbf{x}) = \underbrace{\frac{[\phi_t \circ (\mathbf{f}(\mathbf{g}_t) - \mathbf{y})]^\top [\phi_t \circ (\mathbf{f}(\mathbf{g}_t) - \mathbf{y})]}{\sigma_y^2}}_{\text{residual}} + \underbrace{\mathbf{g}_t^\top \mathbf{G}_{gg} \mathbf{g}_t}_{\text{smoothness of } \mathbf{g}_t} \quad (2.10)$$

$$+ \underbrace{\lambda_1^\top \mathbf{H} \mathbf{g}_t}_{\text{non-negativity}} + \underbrace{\lambda_2 (1 - \mathbf{1}^\top \mathbf{g}_t)}_{\text{integration}} + \underbrace{(\mathbf{g}_t - \mathbf{g}_{t-1})^\top \mathbf{C}_{gg}^{-1} (\mathbf{g}_t - \mathbf{g}_{t-1})}_{\text{smoothness constraint among time steps}},$$

in which  $\mathbf{g}_t$  is the  $n_g \times 1$  discretized travel-time distribution. The subscript  $t$  marks the current window, whereas  $t-1$  denotes the last estimated window, which can be either the window at an earlier time step in

case of forward-window-sliding or at a later time step in backward-window-sliding.  $\mathbf{f}(\mathbf{g}_t)$  is the discrete model according to Eq. (2.9).  $\phi_t = [\phi_1, \phi_2, \dots, \phi_{n_y}]^\top$  is the  $n_y \times 1$  vector of discrete weights corresponding to window  $t$ ; and  $\mathbf{a} \circ \mathbf{b}$  denotes elementwise multiplication of the vectors  $\mathbf{a}$  and  $\mathbf{b}$ , also known as Hadamard or Schur product.

The  $n_g \times n_g$  matrix  $\mathbf{G}_{gg}$  is tridiagonal with elements  $\frac{\Delta\tau}{\theta}[-\frac{1}{2}, 1, -\frac{1}{2}]$  on the three center diagonals. In the geostatistical interpretation of the smoothness constraint,  $\mathbf{G}_{gg}$  is the negative inverse semivariogram matrix  $-\Gamma_{gg}^{-1}$  (Cirpka *et al.*, 2007).

$\lambda_1 = [\lambda_1, \lambda_2, \dots, \lambda_{n_L}]^\top$  is the  $n_L \times 1$  vector of Lagrange multipliers resulting from the non-negativity constraint, in which  $n_L$  is the number of constraints set;  $\mathbf{H}$  is the corresponding  $n_L \times n_g$  selection matrix. As explained by Cirpka *et al.* (2007), a particular element of  $\mathbf{g}_t$  is forced to be zero when the unconstrained optimization yields a negative value, and the sign of the Lagrange multiplier indicates whether a constraint is really needed. Constraints are activated and deactivated until they do not change any more from one iteration to the next.  $\lambda_2$  is a scalar, which is only activated when the integral of  $g(\tau)$  is larger than 1, where the integral is forced to be unity.  $\mathbf{1}$  denotes an  $n_g \times 1$  vector of unit elements, so that  $\mathbf{1}^\top \mathbf{g}_t$  is the sum of all elements of  $\mathbf{g}_t$ .

$\mathbf{C}_{gg}^{-1} = \text{diag} \left( \frac{\Delta t}{s_{g_1}^2}, \dots, \frac{\Delta t}{s_{g_{n_g}}^2} \right)$  is the  $n_g \times n_g$  smoothness-constraint matrix of the parameters between two neighboring windows, where  $s_{g_i}^2$  is the covariance of the  $i$ -th discrete parameter relating the present and previous estimate.

Eq. (2.10) is the discrete version of Eq. (2.4). Because the model of Eq. (2.9) is linear with respect to the parameters, the objective function is a quadratic form of the parameters. The only non-linearity stems from the inequality constraints. For given constraints, the objective function is minimized by setting the derivatives with respect to the parameters to zero, resulting in the following system of linear equations:

$$\begin{bmatrix} \mathbf{W} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{g}_t \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \frac{(\Phi_t \circ \mathbf{X})^\top (\phi_t \circ \mathbf{y})}{\sigma_y^2} + \mathbf{C}_{gg}^{-1} \mathbf{g}_{t-1} \\ \mathbf{0} \\ -\frac{1}{2} \end{bmatrix}, \quad (2.11)$$

with

$$\mathbf{W} = \begin{bmatrix} \frac{(\Phi_t \circ \mathbf{X})^\top (\Phi_t \circ \mathbf{X})}{\sigma_y^2} + \mathbf{G}_{gg} + \mathbf{C}_{gg}^{-1} \end{bmatrix}, \quad (2.12)$$

$$\mathbf{B} = \frac{1}{2} \begin{bmatrix} \mathbf{H} \\ -\mathbf{1}^\top \end{bmatrix}, \quad (2.13)$$

where  $\Phi_t = [\phi_t, \phi_t, \dots, \phi_t]$  is a blown-up  $n_y \times n_g$  matrix of the weighting function  $\phi_t$ . For each window, we alternately solve this system of linear equations and update the constraints regarding non-negativity and the integral of  $g(\tau, t_c)$  until convergence is reached. The results are used as the prior for the estimate of the next window.

The parameter uncertainty is expressed by the following posterior covariance matrix, derived by linear uncertainty propagation of Eq. (2.11):

$$\begin{aligned} \mathbf{C}_{gg}^t &= \sigma_y^2 \frac{\partial \mathbf{g}_t}{\partial \mathbf{y}^\top} \left( \frac{\partial \mathbf{g}_t}{\partial \mathbf{y}^\top} \right)^\top + \frac{\partial \mathbf{g}_t}{\partial \mathbf{g}_{t-1}^\top} \mathbf{C}_{gg} \left( \frac{\partial \mathbf{g}_t}{\partial \mathbf{g}_{t-1}^\top} \right)^\top \\ &= \sigma_y^2 \mathbf{A}_{11} \mathbf{R} \mathbf{R}^\top \mathbf{A}_{11}^\top + \mathbf{A}_{11} \mathbf{C}_{gg}^{-1} \mathbf{A}_{11}^\top, \end{aligned} \quad (2.14)$$

in which

$$\mathbf{A}_{11} = \mathbf{W}^{-1} - \mathbf{W}^{-1} \mathbf{B}^{\top} (\mathbf{B} \mathbf{W}^{-1} \mathbf{B}^{\top})^{-1} \mathbf{B} \mathbf{W}^{-1}, \quad (2.15)$$

$$\mathbf{R} = \frac{1}{\sigma_y^2} [ (\Phi_t \circ \Phi_t \circ \mathbf{X})^{\top} ]. \quad (2.16)$$

## 2.2.6 Goodness of Fit

For the comparison between observed and simulated output signals, we choose the prediction at the central time of each window,  $t = t_c$ , as the only prediction of that particular window. As metrics for judging the quality of the simulations, we use the root mean square error (RMSE), the Nash-Sutcliffe coefficient (NS) of model efficiency (*Nash and Sutcliffe, 1970*), the error in peak heights, and the time shift of peaks.

The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{obs}^i - y_{model}^i)^2}, \quad (2.17)$$

in which  $N$  is the total number of measurements, whereas  $y_{obs}$  and  $y_{model}$  are the observed and simulated output signals, respectively. The RMSE has the same unit as the observation data and quantifies the mean discrepancy between the model and the measurements.

The Nash-Sutcliffe coefficient of model efficiency (NS) is defined as:

$$NS = 1 - \frac{\sum_{i=1}^N (y_{obs}^i - y_{model}^i)^2}{\sum_{i=1}^N (y_{obs}^i - \bar{y}_{obs})^2}, \quad (2.18)$$

in which  $\bar{y}_{obs}$  is the mean of the observations. NS-values close to one indicate an accurate model, whereas NS close to zero means that the model is not better than using the mean of the observations.

We also match peaks of the observed and simulated output time series by identifying local minima within certain windows (different from the sliding window used in the parameter inference) (*Dugge, 2012*, see also Appendix B), see Figure 2.14. We choose minima as peaks in our application, because the time series used (electric conductivity) exhibits rapid reductions during hydrological events. The peaks of the observed and simulated time series are compared in their height and timing. These metrics quantify to which extent the obvious signals in the data are matched.

## 2.3 Application to Field Data

### 2.3.1 Field Site and Data

We apply our non-stationary deconvolution method to electrical-conductivity (EC) time series collected at River Thur, Switzerland, and three observation wells adjacent to the river at the Widen site, close to Frauenfeld, canton Thurgau. The time series contain hourly measured data from November 20th, 2003 until at least April 21st, 2005. A detailed description of the field site and the data is given by *Cirpka et al. (2007)*. In our application, the detrended EC data from the river is used as the input signal, whereas the detrended data from wells 1, 2, 3 are used as output signals to estimate the travel-time distributions from the river to three different well locations. Wells 1 and 2 are adjacent to each other at the edge of the

overbank. Both wells are screened over two meters, well 1 at the bottom of the aquifer (approximately 12-10m), and well 2 at the top (approximately 4-2m). Well 3 is fully screened (approximately 12-2m) and located about 50 meters away from the main channel.

### 2.3.2 Detrending of the Data

Figure 2.4 shows the detrended data, which are calculated by subtracting the baseline from the original signals. Black lines, which show more spikes and rapid changes, are the detrended input signals, whereas blue lines mark the detrended output signals in the wells. Upon detrending, we have eliminated seasonal or other trends in the time series that presumably are not related to the EC signals caused by hydrological events. Note that *Cirpka et al. (2007)* detrended the same data by fitting sinusoidal signals with frequencies of  $1/a$ ,  $2/a$ ,  $3/a$ , and  $4/a$  to the original time series. The latter approach, however, implicitly assumes that the baselines are identical in each calendar year.

The detrended data of well 2 and the river are quite similar, which indicates a quick response of the observed signal to the river signals. There are also time periods, such as in June 2004, during which the calculated output signal is even larger than the input signal, which is an artifact requiring the integral constraint to be activated. For wells 1 and 3, the output signals are smoother than the input signals, and the values are smaller. For well 3, a small time delay in the output signal can visually be observed, which we attribute to this well being located at larger distance to the river channel than the other two wells.

### 2.3.3 Results of Deconvolution

For all detrended data sets, we start the analysis with stationary deconvolution of the entire data series using the method of *Cirpka et al. (2007)*. This yields the initial guess for non-stationary deconvolution. The estimated transfer functions are plotted in Figure 2.5, in which the black lines mark the estimated transfer functions, whereas the gray bands are the 16% and 84% quantiles of the parameter uncertainties. The truncation time of the transfer functions are 15 days for well 1, 6 days for well 2, and 15 days for well 3, while *Cirpka et al. (2007)* used 14 days, 7 days, and 20 days, respectively (see also Table 2.1). In the present study, the smoothness parameter  $\theta$  is chosen identical to that of *Cirpka et al. (2007)* for simplicity, because the estimation of this comes at high computational costs and the model performance indicates that the chosen value is reasonable also in this work.

The most important difference between the present study and that of *Cirpka et al. (2007)* lies in the non-stationary analysis of the data. However, we have also changed the approach of detrending the data from subtracting a time-periodic trend to subtracting a non-periodic baseline. This may partially explain, why we have selected slightly different parameters for deconvolution in the present study than *Cirpka et al. (2007)*. Nonetheless, the general shape (including the tailing) and the mean travel time of the stationary transfer functions estimated in this study are similar to those obtained by *Cirpka et al. (2007)*.

A log-normal distribution function with the same mean and variance as the transfer function from the stationary deconvolution is taken as initial state  $\mathbf{g}_0$  of the first window,  $t = 1$ , in the non-stationary deconvolution method. From there on, the last parameter set  $\mathbf{g}_{t-1}$ , is needed in the determination of  $\mathbf{g}_t$  of the current window  $t$ . The step size from one window to the next is identical to the step size  $\Delta t$  of the data. The window size is chosen as 20 days, 15 days, and 20 days for wells 1, 2, and 3, respectively. The first few windows are influenced by the choice of the initial guess  $\mathbf{g}_0$ . In order to diminish the influence of  $\mathbf{g}_0$ , we slide first the window forward in time, until we reach the end of the data set, and then slide the window backward, conditioning on  $\mathbf{g}_{t+1}$ , thus overwriting all estimated parameter sets.

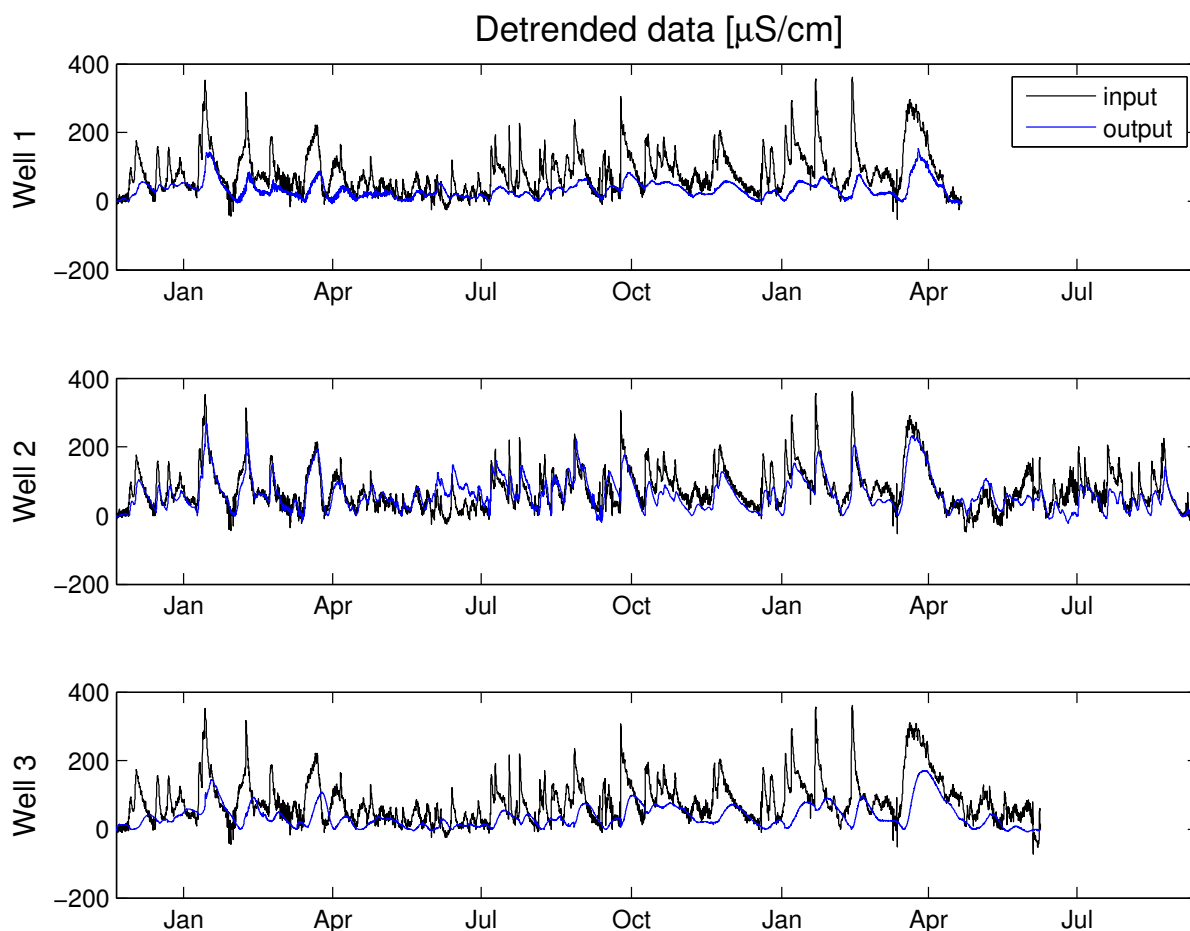


Figure 2.4: Detrended data from November, 2003 to September, 2005. Black: detrended EC signal in the river; blue: detrended EC signal in the wells.

Figure 2.6 shows the observed output data together with the simulated outputs using the transfer functions of the stationary and non-stationary methods for wells 1, 2, and 3. The gray lines mark the output EC-signals in the wells; the red lines show the simulated output, where the detrended data are first applied with the stationary method to obtain the simulated signal without trend, then add back the baseline for the comparison in the figure. The blue lines are the simulated outputs from the non-stationary method. The blue lines are closer to the gray lines than the red lines, indicating a better agreement between the non-stationary model output and the output observations. The associated model performance will be discussed in Section 2.3.5.

### 2.3.4 Estimated Transfer Functions

The estimated transfer functions for wells 1, 2, and 3 are plotted in Figures 2.7-2.9. The two horizontal axes denote the date and the travel time  $\tau$  of the signal between the river and the well, respectively,

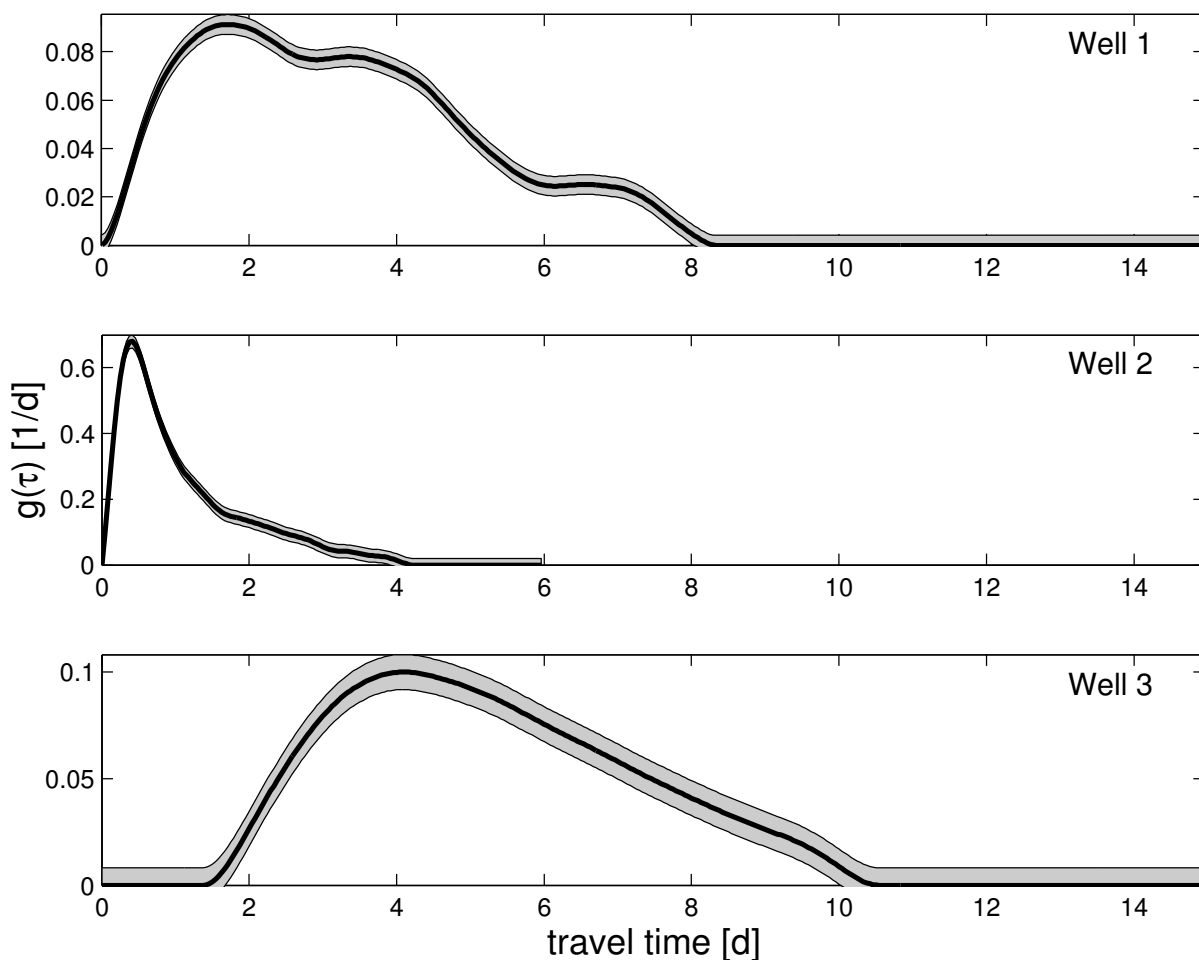


Figure 2.5: Estimated stationary transfer functions from the river to wells 1, 2, and 3, respectively, using stationary method. Log-normal distributions with the same mean and variance are used as the initial guess for the first window in the non-stationary method.

whereas the vertical axis is the value of the transfer function  $g(\tau, t_c)$ . The general shapes of the transfer function are similar to those obtained by the stationary method, but the peak arrival time, width of the peak, and the tailing behavior change with time. In spring 2004, the transfer functions of all three wells show relative high peak values and fast responses to the signals in the river, whereas towards the end of 2004, the transfer functions have lower peak values. These seasonal patterns coincide with low EC baseline values in spring 2004 and high EC values at the end of 2004 within the observation data.

Figure 2.10 shows the offset, mean travel time, and spread of the estimated transfer functions as function of time. The offset shown in the top left panel is the difference between the baselines of the original input and output EC signals. The offset of the three wells show uniformly high values in May 2004 and May 2005, which were times of high base flow. The base flow is determined by similar technique as acquiring the trend of the EC data, where the high water stage is believed to be a result of snow melt. The latter is visualized in the top right panel of Figure 2.10 showing the water stage in the river over the experimental period in blue and its base-flow component in black. While floods occurred at all times of



Table 2.1: Parameters for deconvolution of this work compared to those of *Cirpka et al. (2007)*. Truncation time: length of the estimated transfer function;  $\theta$ : smoothness parameter of the transfer function within each window;

		This work	<i>Cirpka et al. (2007)</i>
Truncation time	Well 1	15d	14d
	Well 2	6d	7d
	Well 3	15d	20d
$\theta [h^{-3}]$	Well 1	$2.3 \times 10^{-8}$	
	Well 2	$3.4 \times 10^{-6}$	
	Well 3	$9.4 \times 10^{-9}$	

the year, the base flow exhibited a clear seasonal pattern with high water stages during the snow-melt periods in April and May and lowest water levels in December.

The remaining panels to the left of Figure 2.10 show the mean travel times from the river to each well as a function of time. They are calculated as the first moments of the travel-time distributions, which are the transfer functions normalized by their zeroth moment. The remaining panels on the right of Figure 2.10 show the spread of the travel-time distributions, which is the square-root of the normalized second central moment. The corresponding uncertainties for each parameter were calculated using Eq. 2.14 and are shown as gray bands ( $\pm$  denotes one standard deviation of estimation). The mean travel times of the three wells are fluctuating during the study period, but in April 2004 all mean travel times are quite small. Well 3 (green line) shows the clearest seasonal trend in the mean travel time with a distinct maximum in November-December 2004. These patterns can be related to that of the base flow: During spring, the water level is higher, facilitating potentially short travel paths of the infiltrating water to the observation wells, whereas in winter the water level reaches its minimum, and the infiltrating water needs more time to reach the wells. The spread of the travel-time distributions also shows fluctuations, which partially follow the patterns of the mean travel time, but appears more difficult to interpret.

Figure 2.11 shows the signal recovery rates, that is, the zeroth moment of the transfer functions. Horizontal lines mark the mean values. Well 2 has the highest mean recovery rate of 0.78, whereas well 1 has the lowest with a value of 0.41. This is so even though the distance of these two wells to the river is identical, but the screening of well 2 is in the top part of the aquifer, whereas well 1 is screened at depth. The recovery rate may potentially be explained as volumetric fraction of the river-borne water in the mixture with water carrying no river signal in the observation well. In this framework, the detected difference in the recovery rates of wells 1 and 2 could be an indication that fresh river infiltrate can be found predominantly at the top of the aquifer, whereas older water of different origin may be found at the bottom. As already discussed by *Cirpka et al. (2007)*, the interpretation of the signal recovery as a mixing ratio may be incomplete because electric conductivity is not a conservative property of water. Mineralization of the water leads to an increase in EC, but may also affect the recovery of EC signals.

Table 2.2 lists average characteristics obtained by non-stationary deconvolution of the detrended EC signals (this study) in comparison to the same characteristics obtained by stationary deconvolution of detrended data (*Cirpka et al., 2007*). The recovery rates determined by the non-stationary methods are basically smaller than those from the stationary method, the mean travel times differ slightly, whereas the spread is systematically larger in the stationary approach. The latter is consistent with the stationary

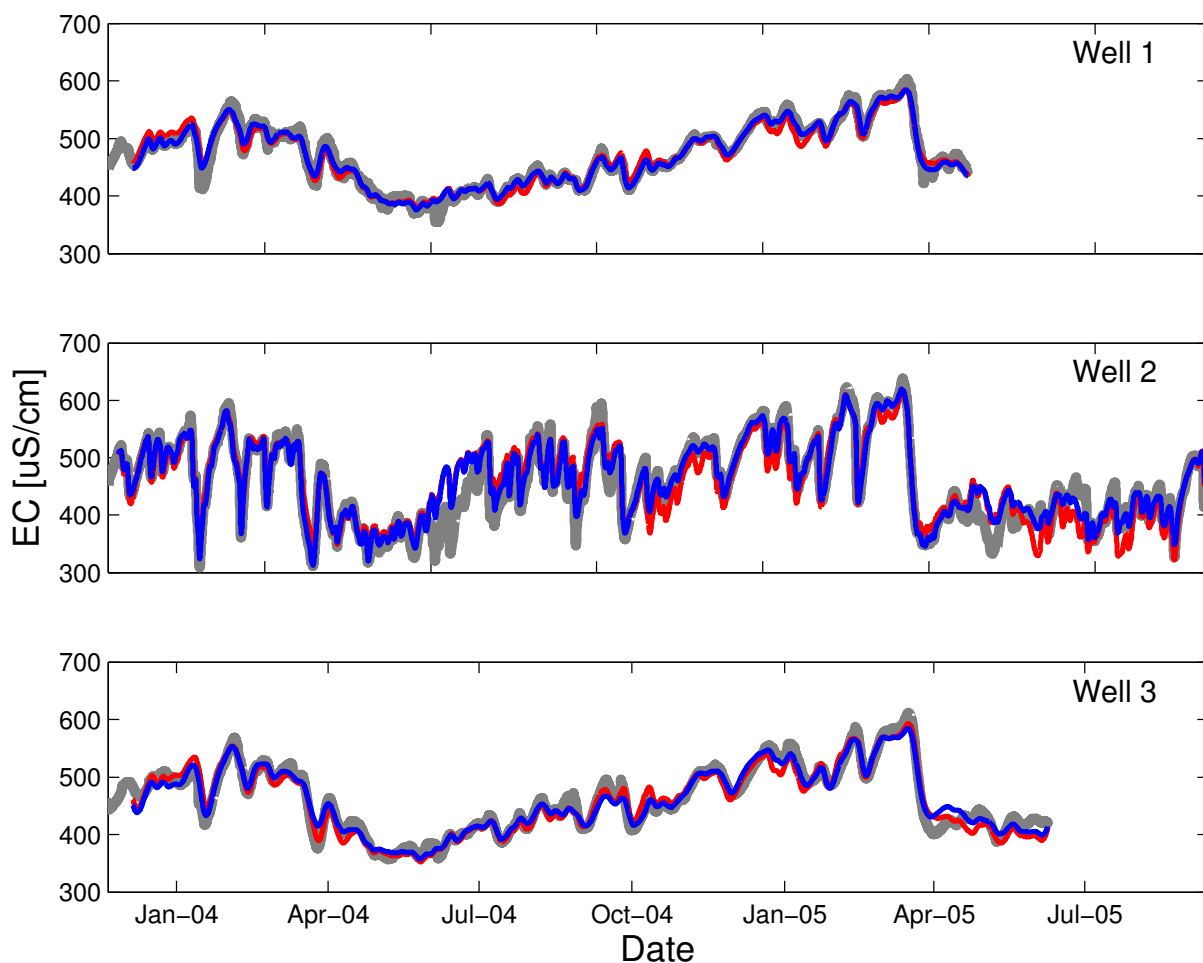


Figure 2.6: Simulated EC time series of well 1, 2, and 3 using the stationary and non-stationary methods, respectively. Gray line: measured EC signal in the wells; Red line: simulation of the well EC time series using the stationary method; Blue line: simulation using the non-stationary method.

travel-time distribution representing a time average.

### 2.3.5 Model Performance

Table 2.3 lists the root mean square error (RMSE) and the Nash-Sutcliffe coefficient of model efficiency (NS) of the simulated output signals in comparison to the observations. The abbreviation 'stat.' stands for the stationary method, and 'non-stat.' for the non-stationary one. The RMSE-values are 11, 26 and 14  $\mu\text{S}/\text{cm}$  for wells 1, 2, and 3, respectively, when applying the stationary method. These values are reduced to 7, 19, and 12  $\mu\text{S}/\text{cm}$  when using the non-stationary method. The improved agreement between models and observations is confirmed by the Nash-Sutcliffe coefficients, which are 0.91, 0.84 and 0.87 for wells 1, 2, and 3, when using the non-stationary method, and have values of 0.80, 0.71, and

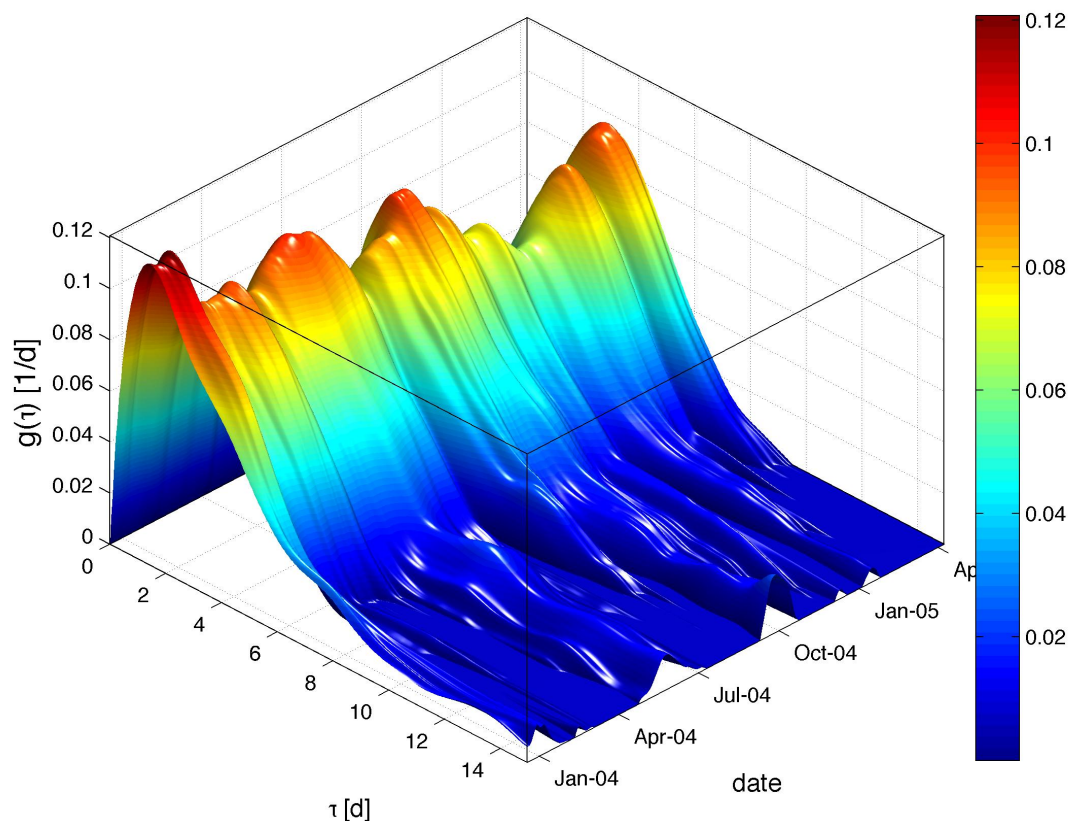


Figure 2.7: Transfer function obtained by the non-stationary method for well 1. Truncation travel time: 15 days.

0.82 when applying the stationary method.

Figures 2.12&2.13 show differences of individual peak heights and times between the model output and the observations. The black circles in Figure 2.12 are differences of individual peak heights for the stationary and non-stationary methods. The horizontal dashed lines mark the range of  $\pm$  one standard deviation about the mean value. Towards the right, box-and-whisker plots visualize the statistics of this metrics for the two deconvolution methods. The range of the peak-height errors obtained by the non-stationary method is apparently smaller than that obtained by the stationary method. Table 2.3 contains the norm of these errors. By changing to the non-stationary method, the error norms have improved by 23%, 33%, and 9% of the values obtained by the stationary method for wells 1, 2, and 3, respectively.

Figure 2.13 shows the same type of comparison as Figure 2.12, but now for the errors in the peak times. As listed in Table 2.3, the corresponding error norms have improved by 15%, 49%, and 45% of the values obtained by the stationary method when changing to the non-stationary approach. The improvement is the largest for well 2, both in peak height and peak time. This particular well responds the quickest to EC fluctuations in the river, with the best preservation of the EC signals and is also influenced the most by the water stage and hydrological events. Here, the signals and the input-output

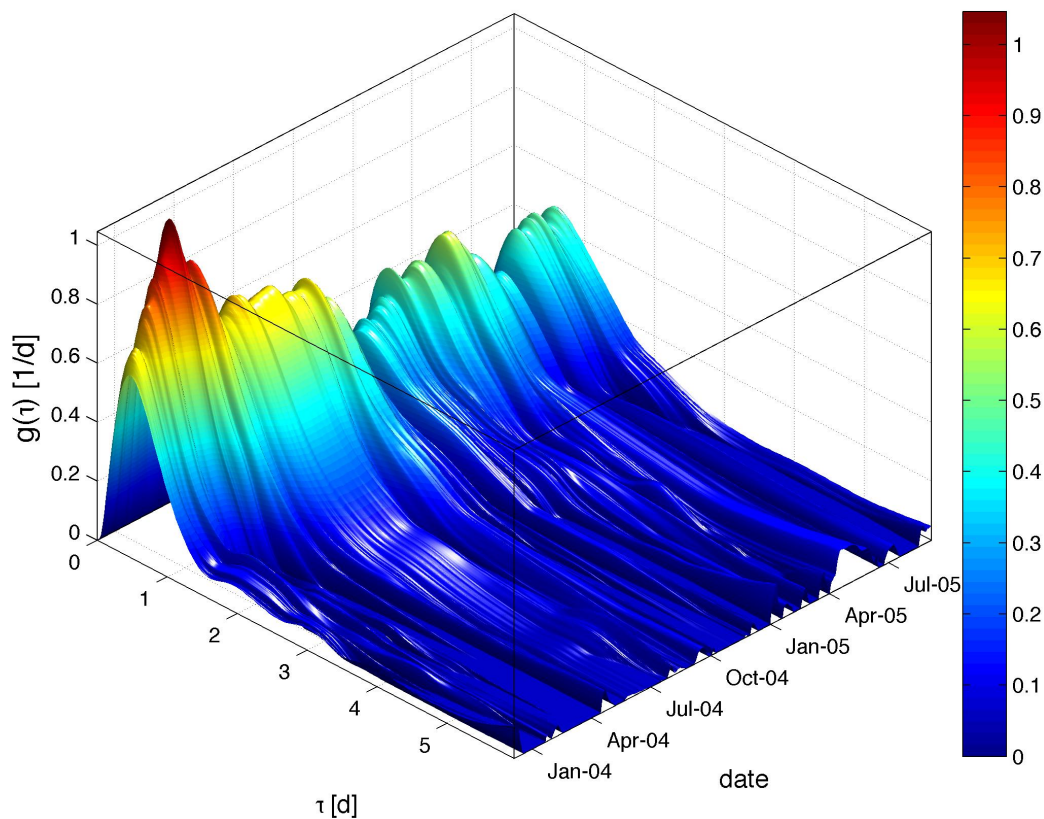


Figure 2.8: Transfer function obtained by the non-stationary method for well 2. Truncation travel time: 6 days. Note that the scale of travel time  $\tau$  for well 2 is different from the other two wells (see Figures 2.7 and 2.9).

relationship between the river and well data are the clearest, and implementing the non-stationary method yields the highest gain. For wells 1 and 3 the time delay between input and output signal as well as the offset are bigger, indicating less clear signals and also less sensitivity towards rapid fluctuations of the flow field.

In general, the non-stationary method shows a better model performance than the stationary method, both with respect to overall metrics and in the agreement of single events. At the particular field site studied here, the stationary method also works fine in terms of the mean travel time and the spread of the travel-time distributions, which we attribute to the quick hydraulic response of the groundwater levels to river-stage fluctuations. That is, at this particular site the groundwater flow patterns do not change so much with the river stage, so that even the alpine hydrological regime of River Thur does not cause overly dramatic non-stationarity in the river-to-groundwater transfer functions.

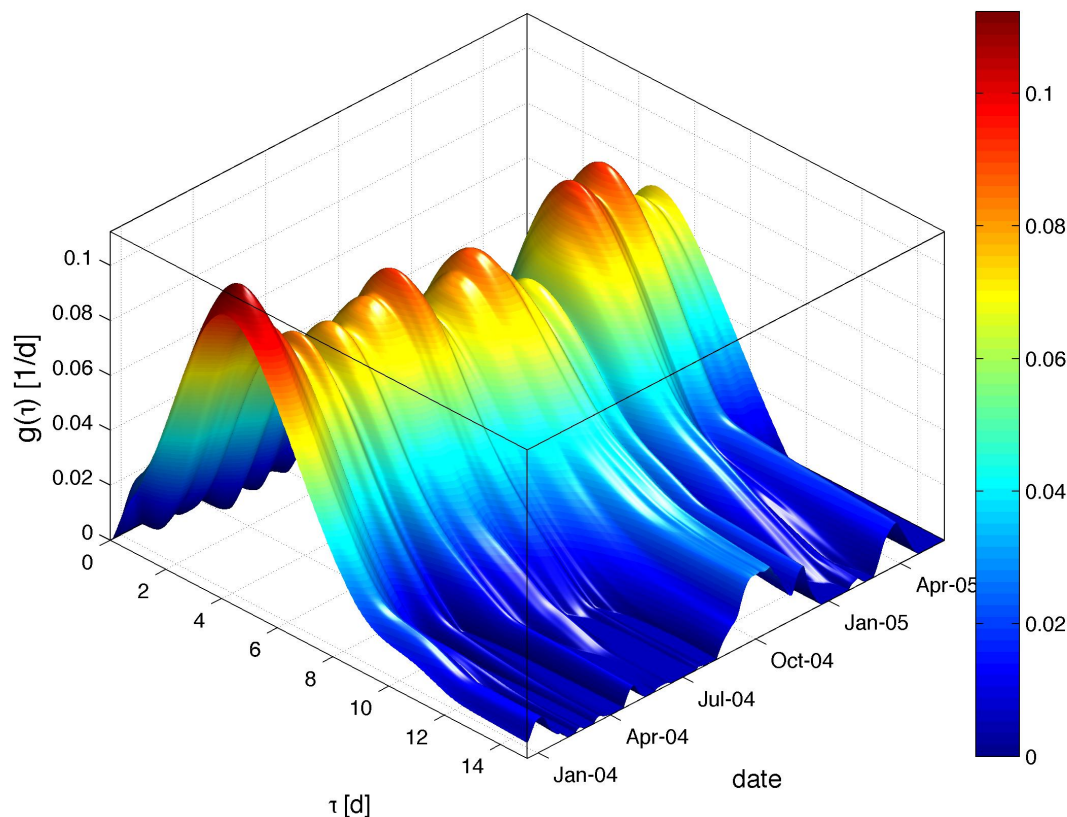


Figure 2.9: Transfer function obtained by the non-stationary method for well 3. Truncation travel time: 15 days.

### 2.3.6 Relationship to Base Flow

Though there are no drastic changes in flow patterns, the temporal variation of the river-to-groundwater transfer functions plotted in Figures 2.7-2.9 reflect changes of travel paths and velocities over time. Since such changes must also be reflected in hydraulic heads, we analyze whether the metrics of the transfer functions, plotted in Figure 2.10, are related to river stage. The actual river stage is very dynamic during hydrological events. The duration of these events is often shorter than the mean travel time. We thus correlate the signal recovery, the mean travel time, and the spread of the travel-time distribution to the base-flow component. Table 2.4 lists the corresponding correlation coefficients. The correlation coefficients are small or moderate (less than 0.7), but general trends can be identified for the different wells. The correlations between parameters are statistically significant at 5% level.

The base flow and the recovery rate are positively correlated for well 1 and 2, which might be an indicator that at high water levels more river water infiltrates at the location close to the river channel, and the EC signals are less diluted by mixing of river water with groundwater of other origin. The negative correlation coefficient of well 3 might be a sign that the response in EC to an increase in base flow in this well is different to that in wells 1 and 2. A potential explanation may be that travel paths change with

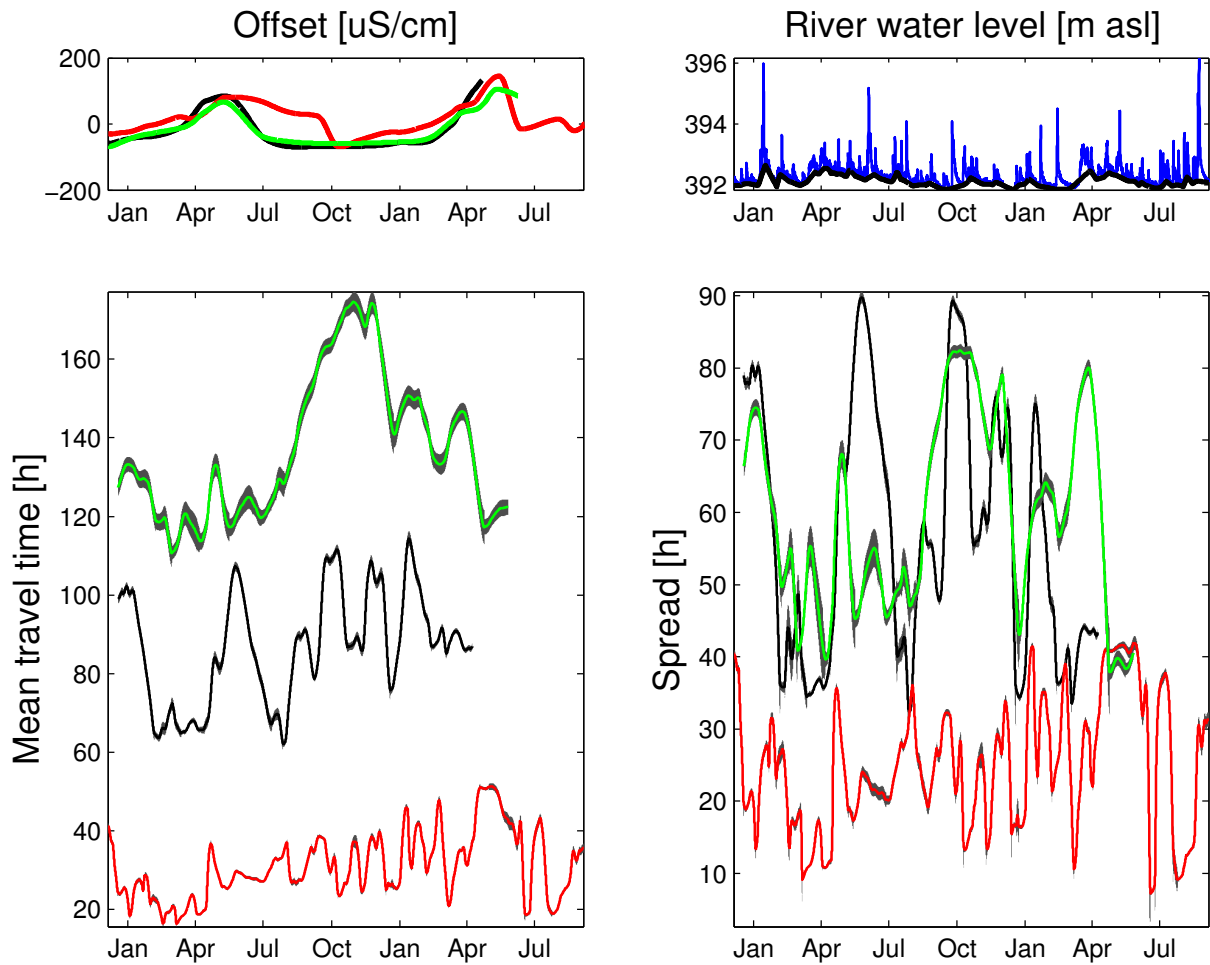


Figure 2.10: Estimated parameter changing over time with uncertainties. Black lines are for well 1; red lines are for well 2; and green lines are for well 3. The gray bands are the parameter uncertainties. Water level and base flow is plotted for analysis purpose.

river stage which also shifts the boundary surface between river water and old groundwater, presumably in a non-trivial way. The correlation coefficients between the base flow and the mean travel time are all negative. That is, all signals tend to propagate more quickly towards the wells when the base river stage is higher, which might be caused by higher velocities (due to higher hydraulic gradients) or the activation of shorter flow paths. It may be noteworthy that the correlation for well 2 is not as high as for the other two wells, because it is closely located to the main river channel and responds so fast, that the water stage might not be a factor for the travel time of the infiltrating water. The relationship between the offset and base flow is clearly positive, which means that during high-water-stage periods, the difference of the input and output base line increases. The correlation mainly originates from the EC baseline in the river channel decreasing at times of high base flow, whereas the baselines in the piezometers are less affected. We assume that the origin of river water at high base flow is somewhat different than at low base flow,

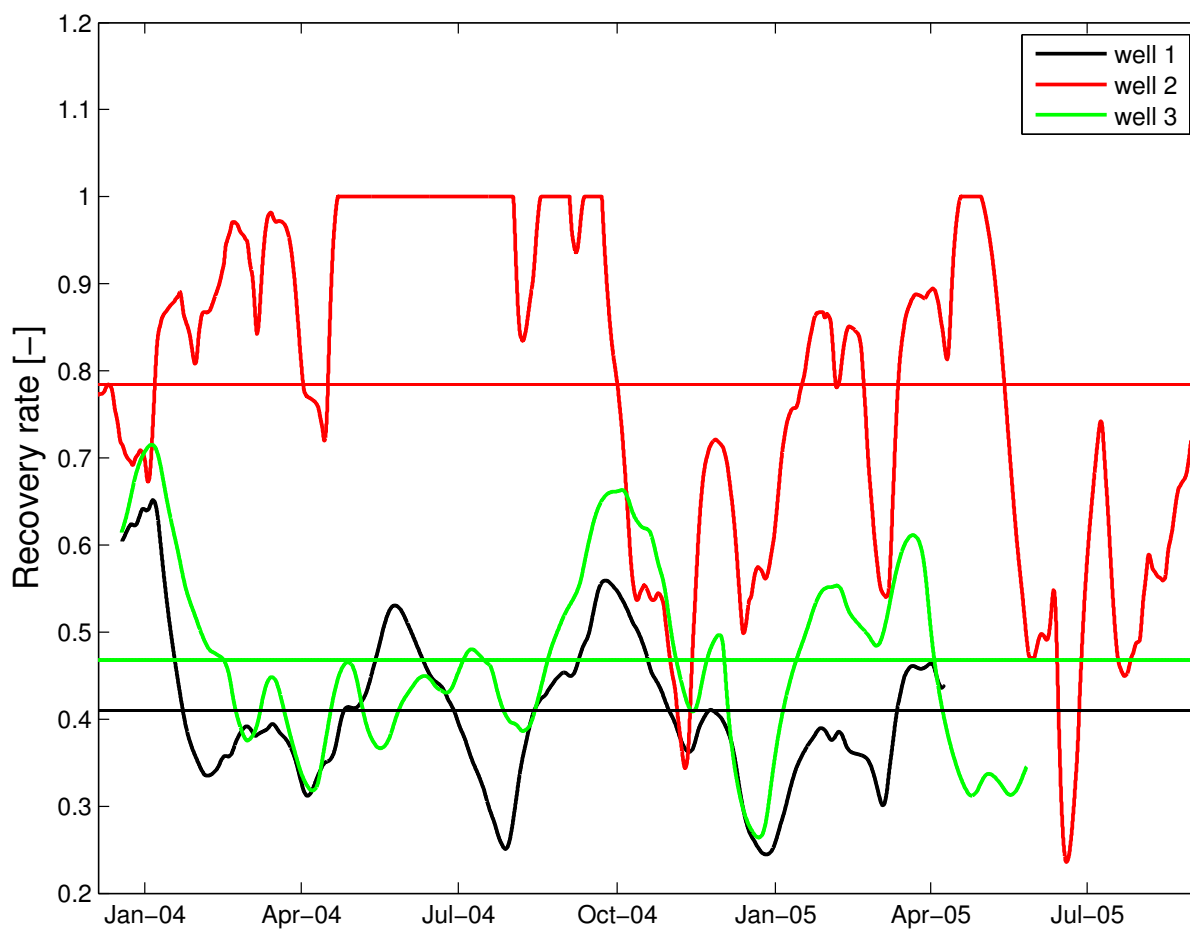


Figure 2.11: Recovery rate of the EC signal in the wells. Horizontal lines are the mean values. Black line: well 1 (mean of 41%); red line: well 2 (mean of 78%); green line: well 3 (mean of 47%).

in particular the fraction of fast runoff components, including snow-melt water, should be higher. This water may not have yet reached chemical equilibrium with the sediments, particularly for carbonate. At low base flow, a substantial fraction of river water originates from groundwater in the upper catchment and is thus closer to equilibrium with sediments than the other runoff components. At the groundwater observation wells, by contrast, the river water has already lost parts of its aggressiveness, so that the seasonality is less pronounced.

## 2.4 Conclusions

In the present study, we have combined the nonparametric deconvolution method of *Cirpka et al.* (2007) with the sliding-window approach to obtain a non-stationary, nonparametric deconvolution method for the determination of dynamic river-to-groundwater travel-time distributions based on natural-tracer

Table 2.2: Comparison of estimated parameter characteristics between non-stationary deconvolution of raw data (this work) and stationary deconvolution of detrended data (*Cirpka et al.*, 2007).

		This work	<i>Cirpka et al.</i> (2007)
(average) recovery rate	Well 1	0.41	0.51
	Well 2	0.78	0.94
	Well 3	0.47	0.60
(average) mean travel time [d]	Well 1	3.52	3.75
	Well 2	1.40	1.46
	Well 3	5.66	6.08
(average) spread [d]	Well 1	2.36	3.67
	Well 2	1.11	1.54
	Well 3	2.42	3.92

Table 2.3: Model performance of the non-stationary deconvolution simulation. RMSE: root mean square error; NS: Nash-Sutcliffe model efficiency coefficient; stat.: stationary method; non-stat.: non-stationary method.

	Well 1	Well 2	Well 3
RMSE stat. [ $\mu\text{S}/\text{cm}$ ]	11	26	14
RMSE non-stat. [ $\mu\text{S}/\text{cm}$ ]	7	19	12
NS coeff. stat. [-]	0.80	0.71	0.82
NS coeff. non-stat. [-]	0.91	0.84	0.8
Peak height stat. [ $\mu\text{S}/\text{cm}$ ]	16	35	13
Peak height non-stat. [ $\mu\text{S}/\text{cm}$ ]	12	23	12
Peak time stat.	14h	22h	1d13h
Peak time non-stat.	12h	11h	21h

time series. The sliding-window technique is easy to implement, but it is necessary to enforce autocorrelation of the parameters in time, because the data do not contain signals sufficiently strong for meaningful analysis at all times. The implementation of auto-correlated time-varying parameters follows similar concepts applied in dynamic harmonic regression (*Young et al.*, 1999). As already highlighted by *Cirpka et al.* (2007), the nonparametric deconvolution has the advantage that no prior assumptions regarding the shape for the travel-time distribution is made.

Because rivers are dynamic systems, the exchange between rivers and groundwater is also dynamic. Stationary time-series approaches applied to natural-tracer data cannot reveal how the transfer of river-water properties to observation wells changes with time. The latter requires a non-stationary approach such as the one presented in this study. The application to electric-conductivity time series collected at River Thur and in adjacent piezometers shows that the non-stationary method outperforms the stationary



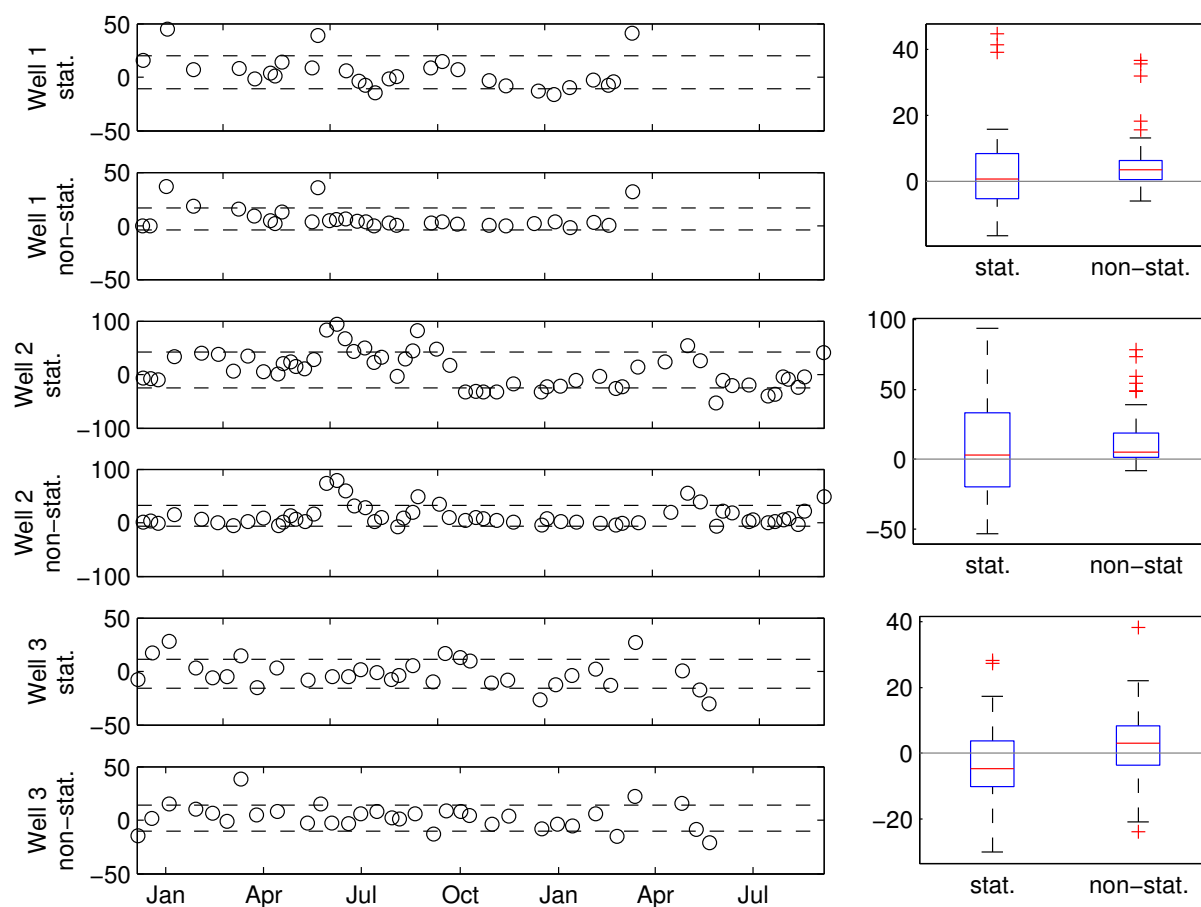


Figure 2.12: Peak height comparison between the simulation and the output signals of wells 1, 2, and 3 for the stationary and non-stationary methods. On the left: the top two figures are for well 1; the middle two figures are for well 2; the bottom two figures are for well 3. On the right: corresponding box plot of the data used for plotting on the left.

one both in overall metrics of output-signal reproduction and in matching individual events. At the specific field site studied here, the stationary results are not too bad, especially for well 1 and 3, where the signal is more averaged and damped. However, other rivers show much more pronounced dynamics of the river-groundwater interactions, and our method might be a promising candidate for analyzing natural-tracer times series collected at these sites.

In our application, we have related metrics of the determined transfer functions to base flow. At high base river stage, the travel times between River Thur and the observation wells are apparently shorter, and the river signal is better recovered within the piezometer which is the closest to the river channel. Based on conceptually simple input-output relationships, however, it is not possible to decide whether the main cause is enhanced infiltration with higher flow velocities or the activation of new flow paths.

High-quality data sets are paramount for the application of the method. Raw data often contain spikes, jumps, or trends. Accounting for a time-varying offset, continuously changing base lines pose

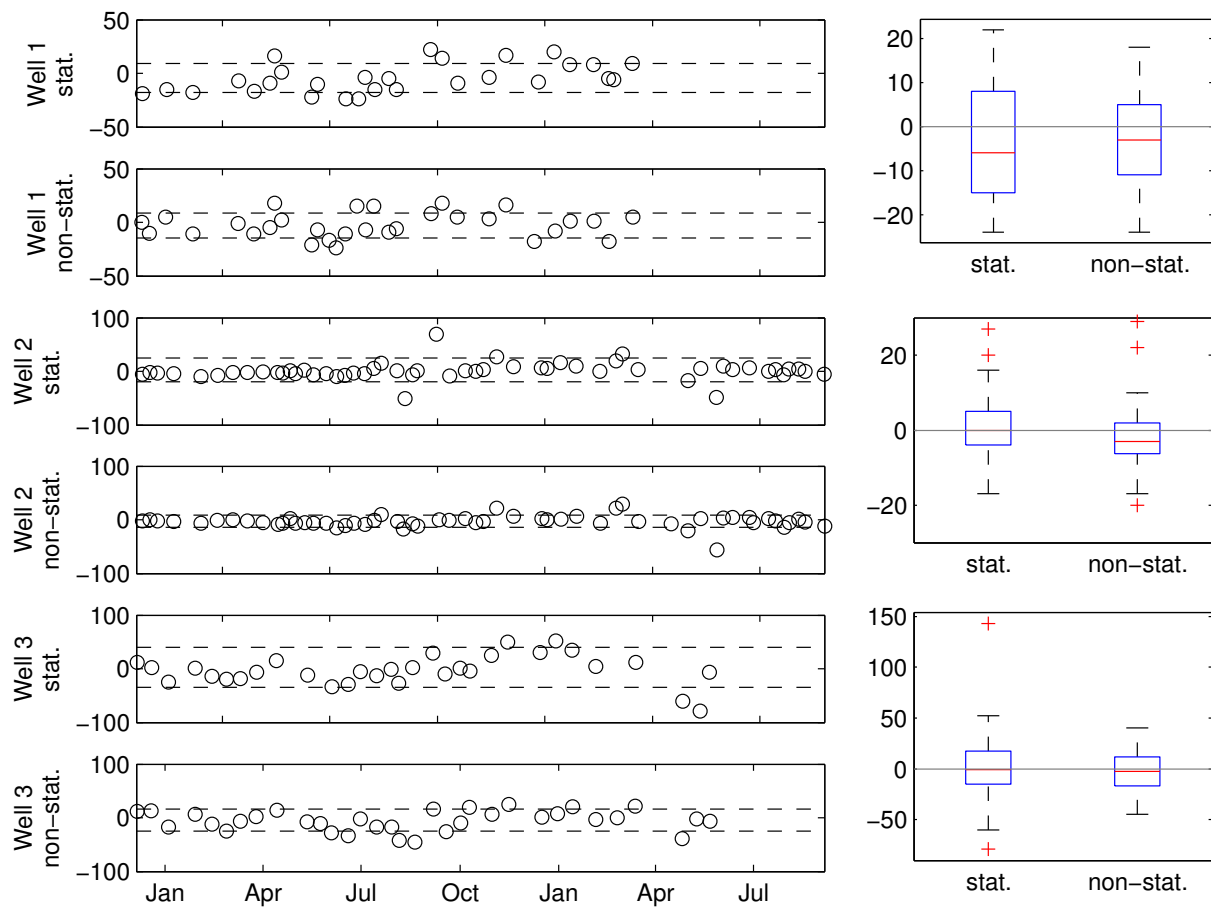


Figure 2.13: Peak time comparison for stationary and non-stationary method between the simulation and the output signals for well 1, 2 and 3. On the left: the top two figures are for well 1; the middle two figures are for well 2; the bottom two figures are for well 3. On the right: corresponding box plot of the data used for plotting on the left.

no problems in the identification of transfer functions, whereas jumps in the data would erroneously be interpreted as significant signals. Thus, cleaning the data to be deconvoluted is absolutely necessary. We have set up the approach to deal with continuous time series of identical temporal resolution. Extensions to irregularly spaced data are possible but destroy the convenient property of  $\mathbf{X}$  being a Toeplitz matrix. Also, the time series must be considerably longer than the estimated transfer function.

When applying the approach to electric-conductivity time series, detrending is a critical step in data processing, because systematic offsets between EC values in rivers and groundwater observation wells may change with time. If we had taken a uniform mean as trend, we would have obtained signal recovery rates that need to be constrained to unity at spring times with large base flow. The separation between trends and signals, however, remains somewhat subjective. In particular tails of the estimated transfer functions are affected by choosing the baselines. We assume that a more conservative property than EC, such as the concentration of dissolved chloride, would require less baseline correction, provided that the

Table 2.4: Correlation coefficients of river base flow and estimated parameters. Note that these values are calculated based on the normalized data.

		Recovery	Mean travel time	Offset
Base flow	Well 1	0.02	-0.34	0.66
	Well 2	0.40	-0.06	0.38
	Well 3	-0.24	-0.50	0.59

associated probes are not affected by sensor drifts.

A key difficulty of our approach is the high non-uniqueness of the determined transfer functions, requiring careful regularization by enforcing smoothness in time and travel time. While *Cirpka et al.* (2007) outlined a computationally expensive approach to estimate the smoothness parameter  $\theta$  from the data within the stationary nonparametric method, we have not yet developed similar techniques for the two smoothness parameters needed in the non-stationary approach. Our present recommendation is to use the optimal  $\theta$ -value from the stationary method and test various values for smoothness within time. Similarly, the window size has to be chosen manually. A large window size stabilizes the estimates but rapid changes in the transport behavior gets missed. Conversely, small window sizes facilitate the detection of rapid changes but may be prone to matching noise within the data. More research will be needed to develop more objective criteria for the smoothness parameters and the window size.

## Acknowledgements

This work was supported by a grant from the Ministry of Science, Research and Arts of Baden-Württemberg (AZ Zu 33-721.3-2), the Helmholtz Center for Environmental Research - UFZ, Leipzig, and the German Academic Exchange Service (DAAD).

## 2.A Details of *Cirpka et al.* (2007) Applied in the Present Work

### 2.A.1 Regularization of the Transfer Functions $g(\tau)$

Smoothness constraint is added onto the travel time distribution  $g(\tau)$  within each sliding window, where a second-order intrinsic model with a linear semivariogram is applied:

$$E\left[\frac{1}{2}(g(\tau+h) - g(\tau))^2\right] = \theta|h| \quad (2.19)$$

in which  $\theta[T^{-3}]$  is the smoothness parameter of  $g(\tau)$ . The discretized value of  $g(\tau)$  only depends on the travel time difference between the two, where small value of  $\theta$  ensures small variation in the transfer function  $g(\tau)$ .

## 2.A.2 Selection Matrix $\mathbf{H}$ and Rules of Activating and Deactivating Lagrange Multipliers

In Eq.(2.10), the selection matrix  $\mathbf{H}$  is used in the non-negativity term by the Lagrange multipliers. This matrix is defined for the element of  $g(\tau)$ , which needs to be constrained because it shows a negative value in the estimation process:

$$\mathbf{H}_{ij} = \begin{cases} 1, & \text{if } g_j \text{ is activated by the } i\text{-th constraint} \\ 0, & \text{otherwise} \end{cases} \quad (2.20)$$

The rules for Lagrange multipliers are:

$$\begin{aligned} g_i < 0, & \rightarrow \text{add constraint for element } j \\ \lambda_i \leq 0, & \rightarrow \text{keep constraint } i \\ \lambda_i > 0, & \rightarrow \text{remove constraint } i \end{aligned} \quad (2.21)$$

## 2.B Peak-Finding Method of *Dugge* (2012)

We have applied the method of *Dugge* (2012) to evaluate the model performance in terms of peak event agreement. The steps of finding and matching the peaks are as following:

1. Devide the time series into small sections with designated window size.
2. Find the points with maximum and minimum values within each window.
3. Delete those points where the differences between maximum and minimum values are smaller than a given threshold to rule out small peaks.
4. Find closest local maxima in two time series, when the time difference is smaller than a given threshold.
5. Remove the selected points preventing reuse.

A schematic figure for the peak matching is shown in Figure 2.14.

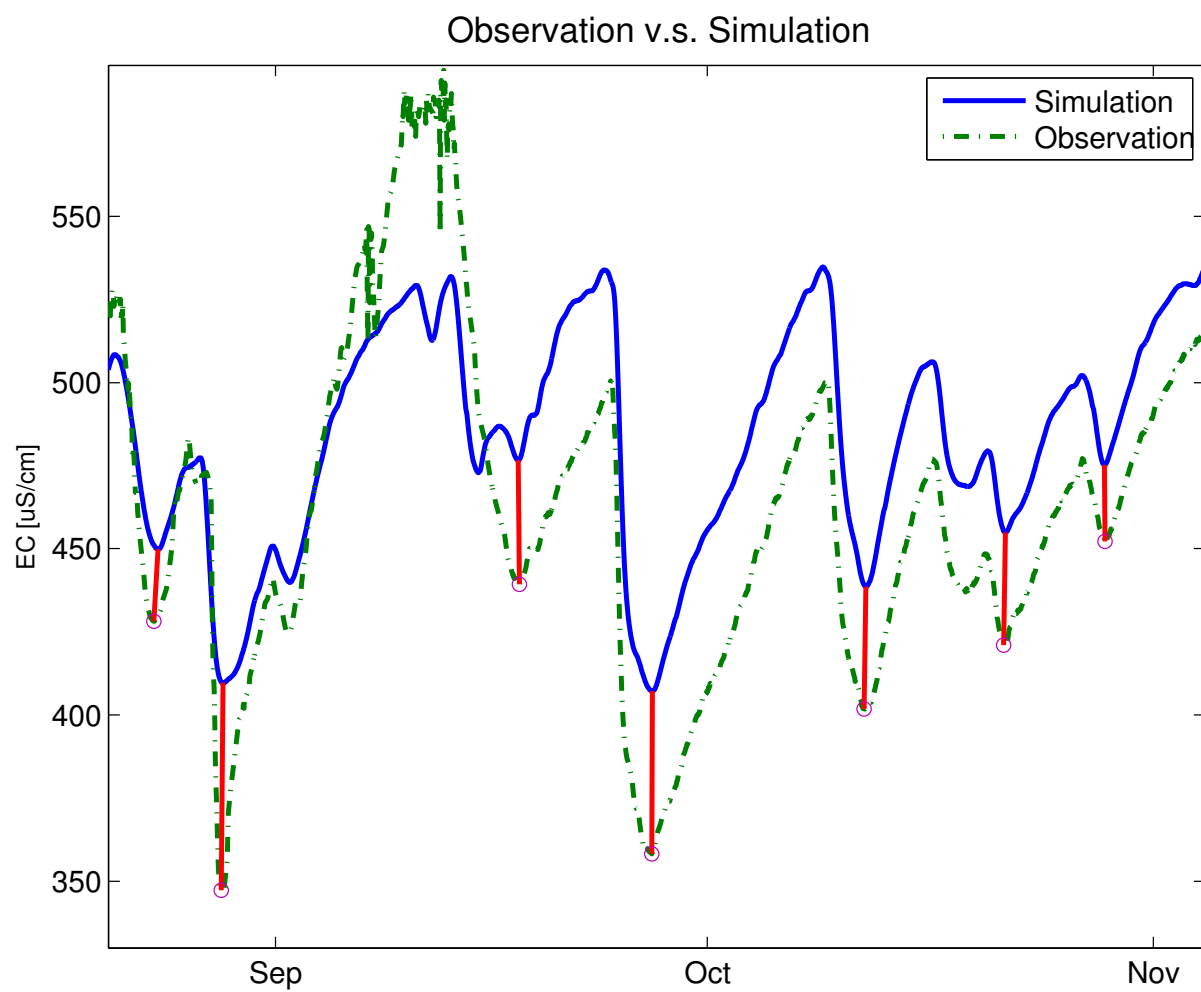


Figure 2.14: Peak height and time differences are found by the peak-matching algorithm for the simulation and the observation data, using the method of *Dugge (2012)*.



## Chapter 3

# Shape-Free Inference of Hyporheic Travel-Time Distributions from Synthetic Conservative and 'Smart' Tracer Tests in Streams

Liao, Z., and O. A. Cirpka (2011), Shape-free inference of hyporheic traveltime distributions from synthetic conservative and 'smart' tracer tests in streams, *Water Resour. Res.*, 47, W07510, doi:10.1029/2010WR009927

### Abstract

The hyporheic zone has been identified important for river ecology, natural biogeochemical turnover, filtration of particles, degradation of dissolved pollutants and thus for the self-cleaning capacity of streams and for groundwater quality. Good estimation of the travel-time distribution in the hyporheic zone is required to achieve a better understanding of transport in the river system. The transient storage model has been accepted as an appropriate tool for reach-scale transport in rivers undergoing hyporheic exchange, but the choice of the best parametric function for the hyporheic travel-time distribution has remained unclear. We present an approach of obtaining hyporheic travel-time distributions from synchronous conservative and 'smart' tracer experiments that does not rely on a particular functional form of the hyporheic travel-time distribution, but treats the latter as a continuous function. Nonnegativity of the hyporheic travel-time distribution is enforced by the method of Lagrange multipliers. A smoothness parameter, needed for regularization, and uncertainty bounds are obtained by the expectation-maximization method relying on conditional realizations. The shape-free inference provides the opportunity of capturing unconventional shapes, e.g. multiple peaks, in the estimation. We test the approach by applying it to a virtual test case with a bimodal hyporheic travel-time distribution, which is recaptured in the inversion of noisy data.

### 3.1 Introduction

The transition zone between rivers and groundwater, denoted as hyporheic zone, has been identified important for river ecology, natural biogeochemical turnover, filtration of particles, degradation of dissolved pollutants and thus for the self-cleaning capacity of streams and for groundwater quality (*Boulton et al.*, 1998; *Sophocleous*, 2002). Hyporheic exchange is the intermediate passage of river water through the river bed and/or alluvial aquifer. In our definition, it includes the water flux through ripples and dunes, which is predominantly a vertical exchange process, as well across river banks through meanders, which is predominantly a horizontal process. It is driven by hydraulic head differences along the river bed, and controlled by bathymetry, pressure distributions within the stream, and composition of bed material (*Gooseff et al.*, 2003; *Cardenas et al.*, 2004; *Briggs et al.*, 2009).

Hyporheic exchange brings river water, which may be rich in organic carbon and oxygen, in contact with biofilms on the grain surfaces of the bed material, facilitating enhanced biogeochemical turnover. As a simplifying model, it may be assumed that the turnover of river-water components in the hyporheic zone can be related to the time over which the water remains in the sediments. Due to the irregular shape of the river bed and the heterogeneity of the adjacent aquifer, such travel times should be described as statistical distributions rather than single values. The key objective of this study is to present methods for estimating the strength of hyporheic exchange and the associated travel-time distribution. These parameters are valuable in the estimation of effective biogeochemical turnover in the hyporheic zone and therefore crucial for water resources management.

In order to model the effects of hyporheic exchange on effective transport in streams, transient-storage models (TSM) have become well accepted (*Bencala and Walters*, 1983; *Harvey et al.*, 1996). These models are mostly based on the advection-dispersion equation for transport within the stream, coupled to transient terms accounting for storage in the hyporheic zone. The simplest transient-storage model is based on a single immobile storage undergoing first-order exchange with the stream water, which implies an exponential function of the travel-time distribution. Other studies have indicated that such a dual-domain model may not always be appropriate (*Gooseff et al.*, 2005b). The extension to multiple immobile domains, each undergoing first-order mass exchange with the mobile domain, is denoted multi-rate mass transfer model (MRMT) (*Haggerty and Gorelick*, 1995). The key in the application of the MRMT model, is to identify the number of immobile domains, their sizes, and the corresponding rate coefficients. Generalized methods based on travel-time distributions, where any kind of function can be used to fit the observed data, are potentially superior to the single transient storage model. At certain field sites, power-law models (e.g., *Haggerty et al.*, 2002) and the log-normal distribution (*Wörman et al.*, 2002) were proved as appropriate estimate of the residence time distribution according to the experimental data. The conceptually similar continuous-time random walk (CTRW) approaches (*Boano et al.*, 2007) describe the effect of hyporheic exchange on river transport by waiting times associated to each storage process. Recently, the Solute Transport in Rivers (STIR) model (*Marion et al.*, 2008) and the Variable Residence Time (VART) model (*Deng and Jung*, 2009) were proposed to extend the previous parametric models.

Even though substantial effort has been spent on relating the functional shape of the hyporheic travel-time distribution to morphological features of the river, we doubt that it is always possible to decide a priori which parametric model describes hyporheic exchange at a particular site the best. This is so because often multiple bedforms, such as both pool-ripple sequences and river meanders, synchronously cause hyporheic exchange. Also, the sediments of the hyporheic zone and the associated alluvial aquifer are typically heterogeneous, exhibiting dominant features, such as palaeo channels, or not. While the direct evaluation of hyporheic travel times is possible in coupled numerical models (e.g. *Cardenas et al.*,



2004), this is not possible in field experiments. In a stream-tracer experiment, the solute may undergo hyporheic exchange several times, and it is also transported within the stream. Thus, the distribution of residence times for a single passage through the hyporheic zone, which is the target of interest here, is hidden in a tracer signal caused by multiple, interacting processes. Separating the effects of hyporheic exchange on stream transport from the other processes, and quantifying hyporheic travel times, is particularly important when data obtained by tracer tests are to be transferred to reactive transport, in which biogeochemical reactions within the hyporheic zone are coupled to physical transport processes. In such a context, confusing hyporheic exchange with in-stream mixing processes, or working with erroneous hyporheic travel-time distributions, may lead to wrong predictions of overall reactive transport.

Given a hyporheic travel-time distribution, the simulation of effective tracer transport in the river can best be achieved by solution in the Laplace domain (*Haggerty and Gorelick, 1995; Haggerty et al., 2002; Wörman, 2000; Wörman et al., 2002; Marion et al., 2008*). All approaches suggested so far require a parametric model for the hyporheic travel-time distribution, and parameter estimation implies identifying a few parameters (e.g., the effective velocity in the river, the in-stream dispersion coefficient, an exchange coefficient, and two to three parameters characterizing the hyporheic travel-time distribution). Even with this limited number of parameters, ambiguity poses a problem.

Various experimental methods exist to obtain the travel-time distribution between an infiltrating river and an observation point in groundwater, including artificial tracer tests (e.g. *Davis et al., 1980; Lin et al., 2003*), the analysis of dissolved gases (e.g. *Alvarado et al., 2005*), stable isotopes (e.g. *Massmann et al., 2008*), and time-series analysis of naturally fluctuating physical properties such as temperature and electric conductivity (*Anderson, 2005; Cirpka et al., 2007; Constantz, 2008; Constantz et al., 2003; Hoehn and Cirpka, 2006; Keery et al., 2007*). Conservative tracer experiments in streams have been widely used for the investigation of exchange between surface freshwater and groundwater, in which the breakthrough curve (BTC) measured downstream is used to infer parameters of advection, dispersion within the stream, and hyporheic exchange. However, the spreading of the breakthrough curve is caused by the combination of many distinct processes, such as in-stream dispersion, exchange with in-stream dead zones, and hyporheic exchange, so that the distinction between different spreading mechanisms might be unclear by using a conservative tracer alone. Recently, *Haggerty et al. (2008)* have introduced resazurin as a 'smart' tracer that selectively undergoes decay in the hyporheic zone. These authors hope to distinguish in-stream dispersion and hyporheic exchange by analyzing the difference in breakthrough curves of a conservative tracer and a reactive tracer, resazurin. A rigorous quantitative analysis, however, is still missing. In particular, techniques of identifying hyporheic travel-time distributions without predefining the type of function have not been developed so far.

In the context of analyzing stream-to-well travel times, *Cirpka et al. (2007)* have developed a non-parametric deconvolution method allowing the determination of the travel-time distribution between the river and a groundwater observation point from fluctuations of electric conductivity without relying on a predefined shape of the distribution function. The key contribution of the present study is to transfer this technique to the identification of hyporheic travel-time distributions from tracer tests within streams.

In this study, we simulate an artificial tracer experiment in the stream by using both conservative and reactive tracers, where breakthrough curves measured downstream are jointly inverted to estimate stream-transport parameters and the travel-time distribution within the hyporheic zone. Transport within the stream reach is considered to follow the 1-D advection-dispersion equation, and a fraction of the river water is assumed to infiltrate into the hyporheic zone and return back to the same point in the river over a distribution of travel times (Fig. 3.1). In reality, the assumption of returning back at exactly the same point where the water infiltrates into the hyporheic zone is not appropriate, but the river reach is considered to have a length on the order of kilometers, which makes the small spatial offset by a few meters between

infiltration and exfiltration points negligible. We also assume the travel-time distribution to be uniform over time and reach length. This implies similar hydraulic properties of the hyporheic zone over the river reach.

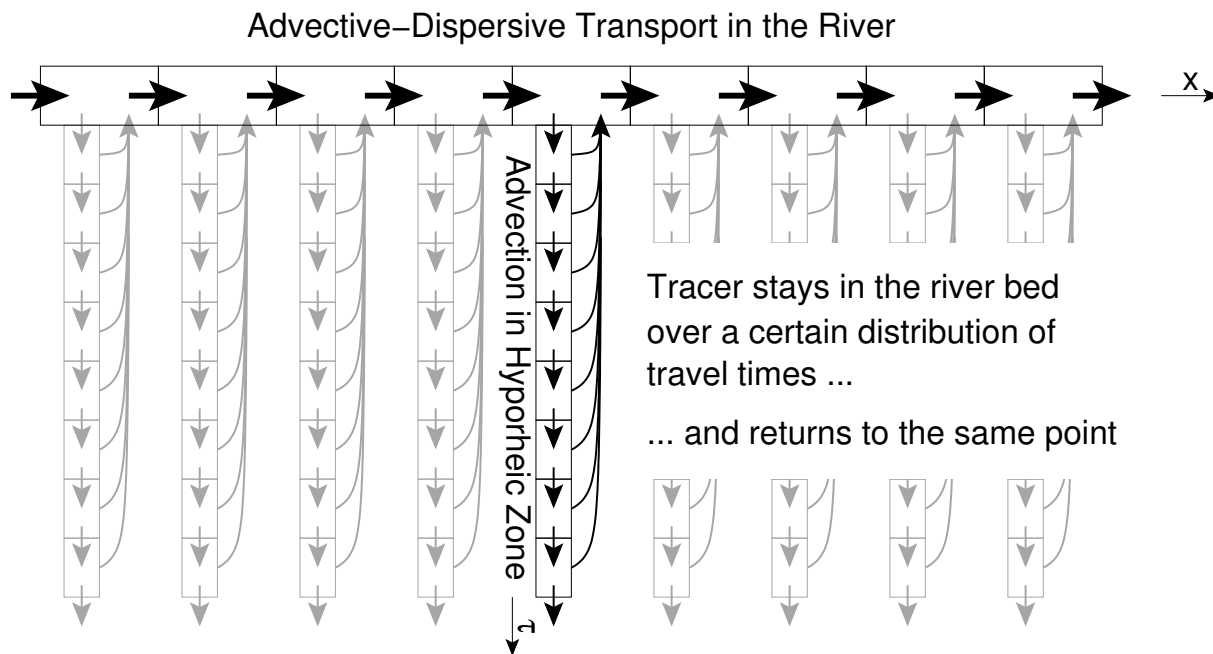


Figure 3.1: Illustration of the transient-storage model concept. Advective-dispersive transport within the stream is considered one-dimensional. Transport through the hyporheic zone is assumed to be along non-interacting streamlines, in which the longitudinal spatial coordinate is replaced by travel time.

The main objective of this study is to develop an inverse method for the estimation of hyporheic travel-time distributions from conservative and reactive tracer experiments in streams, in which the approach of *Cirpka et al.* (2007) is extended to be combined with nonlocal-in-time methods to describe solute transport in rivers undergoing hyporheic exchange (*Haggerty et al.*, 2002). In particular, we want to derive a shape-free method to determine travel-time distributions from tests with tracers that selectively react in the hyporheic zone (*Haggerty et al.*, 2008).

## 3.2 Governing Equations

### 3.2.1 One-Dimensional Advective-Dispersive Transport in Streams

The effect of the hyporheic zone on solute transport within a stream can be expressed by adding the change of storage within the hyporheic zone to the mass balance equation of the solute. *Haggerty et al.* (2002) expressed the change of hyporheic storage as the convolution of a hyporheic memory function  $g^*(t)$  [ $T^{-1}$ ] with the time derivative of the solute concentration  $c(t)$  [ $ML^{-3}$ ]. considering this term in the conventional advection-dispersion equation (ADE) for one-dimensional transport in a river yields:

$$\frac{\partial c}{\partial t} + \beta \int_0^t \frac{\partial c(t-\tau)}{\partial t} g^*(\tau) d\tau + v \frac{\partial c}{\partial x} - D \frac{\partial^2 c}{\partial x^2} = \frac{\partial c_{inj}}{\partial t} \quad (3.1)$$

in which  $t$  [T] is the time since the experiment has started;  $\tau$  [T] is a lag time;  $\beta$  [-] is the ratio of the storage capacities in the hyporheic zone over that in the channel;  $v$  [ $LT^{-1}$ ] is the mean velocity in the stream;  $D$  [ $L^2T^{-1}$ ] is the in-stream dispersion coefficient;  $c_{inj}$  [ $ML^{-3}$ ] is the concentration of the tracer injected into the channel, assumed to be a pulse injection throughout the paper, i.e.,  $c_{inj} = m/Q\delta(t)$ , in which  $m$  [M] is the injected mass,  $Q$  [ $L^3T^{-1}$ ] is the stream discharge, and  $\delta(t)$  [ $T^{-1}$ ] is the Dirac delta function applied to time. The memory function  $g^*(t)$  can be considered as the probability density that a tracer molecule entering the hyporheic zone at time zero will still be in the hyporheic zone at time  $t$ . In Eq. (3.1), we have made the implicit assumption that the only water gained by the river originates from the river itself. For simplicity, we will keep this assumption in the following.

Instead of using the memory function  $g^*(t)$ , we introduce the travel time distribution  $g(\tau)$  [ $T^{-1}$ ] in our model, which describes the probability density of a tracer particle entering the hyporheic zone at time zero to leave it after the time period  $\tau$ . We also account for first-order decay within the hyporheic zone. Under these assumptions, our model formulation reads as:

$$\frac{\partial c}{\partial t} + v \frac{\partial c}{\partial x} - D \frac{\partial^2 c}{\partial x^2} = q_{he} \left( \int_0^t g(\tau) c(t-\tau) \exp(-\lambda\tau) d\tau - c(t) \right) \quad (3.2)$$

subject to the initial and boundary conditions:

$$c(t=0, x) = 0 \quad (3.3)$$

$$\left( vc - D \frac{\partial c}{\partial x} \right) \Big|_{x=0} = \frac{m}{A} \delta(t) \quad (3.4)$$

$$\lim_{x \rightarrow \infty} c(t, x) = 0 \quad \forall t \quad (3.5)$$

in which  $q_{he}$  [ $T^{-1}$ ] is the exchange-rate coefficient of river-water undergoing hyporheic exchange, which is the amount of the river water, entering the hyporheic zone and returning back into the river, per volume of river water;  $\lambda$  [ $T^{-1}$ ] is the first-order decay coefficient of the reactive tracer in the hyporheic zone (for a conservative tracer,  $\lambda = 0$ ); and  $A$  [ $L^2$ ] is the cross-sectional area of the river at the injection point.

The formulation of Eq. (3.2) resembles the general approach of (Wörman *et al.*, 2002) who considered transport along an ensemble of hyporheic flowpaths. Eq. (3.2) is identical to the STIR model, if Fickian in-stream transport is assumed (Eq. (28) of Marion *et al.*, 2008). In contrast to Eq. (3.1), we do not need to take the derivative of the hyporheic travel-time distribution. For the shape-free inference method discussed below,  $g(\tau)$  is constrained to be non-negative for all values of  $\tau$ . Such a constraint can easily be implemented by the method of Lagrange multipliers (see below). In Eq. (3.1), by contrast,  $g^*(\tau)$  must be a monotonically decreasing function, which is less straight-forward to be enforced. The model formulation of Eq. (3.1) requires the ratio  $\beta$  of the storage capacities in the hyporheic zone over that in the stream, which needs to be estimated from tracer data in practical applications. Our formulation in Eq. (3.2) requires the strength of hyporheic exchange  $q_{he}$ , which is equally difficult to obtain. Both models require a probability density function of times, namely the age distribution  $g^*(\tau)$  within the hyporheic zone in case Eq. (3.1), and the age distribution  $g(\tau)$  of the solutes returning to the stream in case of Eq. (3.2). The models in Eqs. (3.1 & 3.2) are mathematically equivalent when the following identities hold (details can be found in Appendix 3.A):

$$q_{he} = \beta g^*(0) \quad (3.6)$$

$$g(\tau) = -\frac{1}{g^*(0)} \frac{\partial g^*(\tau)}{\partial \tau} \quad (3.7)$$

### 3.2.2 Solution in the Laplace Domain

The solution of the forward model in the time domain is tedious, because it involves a convolution in time. It is significantly expedited by changing into the Laplace domain. The Laplace transformation of a time-dependent variable  $f(t)$  is defined as:

$$\tilde{f}(s) := \mathcal{L}\{f(t)\} = \int_0^{\infty} e^{-st} f(t) dt \quad (3.8)$$

in which  $s [T^{-1}]$  is the complex Laplace coordinate. Applying the Laplace transformation of Eq. (3.2) in time yields:

$$(s + q_{he}(1 - \tilde{g}(s + \lambda))) \tilde{c}(s, x) + v \frac{d\tilde{c}(s, x)}{dx} - D \frac{d^2\tilde{c}(s, x)}{dx^2} = 0 \quad (3.9)$$

$$\left( v\tilde{c}(s, x) - D \frac{d\tilde{c}(s, x)}{dx} \right) \Big|_{x=0} = \frac{m}{A} \quad (3.10)$$

$$\lim_{x \rightarrow \infty} \tilde{c}(s, x) = 0 \quad (3.11)$$

in which the Laplace transformation has converted the time derivative  $\partial c / \partial t$  to the product  $s\tilde{c}(s)$  and the convolution integral  $\int_0^t g(\tau)c(t - \tau)d\tau$  into a multiplication of two Laplace-transformed functions  $\tilde{g}(s)\tilde{c}(s)$ . Eqs. (3.9-3.11) must be met for all values of the Laplace coordinate  $s$ .

In the following, we assume that the in-stream velocity  $v$  and the dispersion coefficient  $D$  are uniform over the entire length of the reach. Then, Eq. (3.9) is a second-order homogeneous linear ordinary differential equation with constant coefficients having the following solution:

$$\tilde{c}(s, x) = \tilde{c}_0(s) \exp(-\alpha x) \quad (3.12)$$

in which complex coefficient  $\alpha$  and  $\tilde{c}_0(s)$  are defined by:

$$\alpha(s) = \frac{2(s + q_{he}(1 - \tilde{g}(s + \lambda)))}{\sqrt{v^2 + 4D(s + q_{he}(1 - \tilde{g}(s + \lambda)))} + v} \quad (3.13)$$

$$\tilde{c}_0(s) = \frac{m}{A} \frac{1}{v + \alpha D} \quad (3.14)$$

which includes the conservative case ( $\lambda = 0$ ).

In our simulations, we use the analytical expression in the Laplace domain, Eqs. (3.13 & 3.14), and perform numerical back-transformation into the time domain using the approach of *De Hoog et al.* (1982). If  $v$  or  $D$  varied in space, we could solve Eqs. (3.9-3.11) by numerical simulation. If the coefficient varied in time, the solution strategy by Laplace transformation is no more feasible.

### 3.3 Inference of Hyporheic Travel-Time Distribution $g(\tau)$ and other Transport Parameters

Our model, Eq.(3.2), contains four variables to be estimated. Besides the in-stream velocity  $v$  and the in-stream dispersion coefficient  $D$ , our main target quantities are the exchange flux  $q_{he}$  and the non-negative travel-time distribution  $g(\tau)$  in the hyporheic zone. Rather than choosing a parametric model for  $g(\tau)$  and estimating the parameters of this model, we want to infer  $g(\tau)$  as a continuous function. As will be shown below, we require a structural parameter  $\theta$ , which determines the smoothness of the travel-time distribution, which can also be estimated from the data within the overall inference scheme.

In this section, we first explain the shape-free inference method. This involves discretizing the hyporheic travel-time distribution  $g(\tau)$ , formulating the regularization, constraining the results to be non-negative, including the data from both conservative and reactive tracer tests, and finally estimating the smoothness parameter  $\theta$ . The shape-free method requires an initial guess of all parameters, including  $g(\tau)$ . We do this by assuming a parametric model of  $g(\tau)$  and estimating the corresponding parameters. The procedure followed for that is presented at the end of this section. Fig. 3.2 shows a schematic representation of the overall approach in the order of the computational steps involved.

#### 3.3.1 Shape-Free Inference

##### Discretization of Hyporheic Travel-Time Distribution

The travel-time distribution  $g(\tau)$  in the hyporheic zone is a probability density function, describing the likelihood that the river water returning from the hyporheic zone has stayed therein over the duration  $\tau$ . In general  $g(\tau)$  could have any functional form, as long as  $g(\tau) \geq 0$  for all values of  $\tau$ . Rather than using a parametric function with a few shape parameters, we now express the continuous distribution  $g(\tau)$  as a piecewise linear function, spanned by a large set of discrete values  $g_k$  of  $g(\tau)$  at times  $\tau = k\Delta\tau$ :

$$g(\tau) = \frac{k\Delta\tau - \tau}{\Delta\tau} g_{k-1} + \frac{\tau - (k-1)\Delta\tau}{\Delta\tau} g_k \text{ for } (k-1)\Delta\tau \leq \tau \leq k\Delta\tau \quad (3.15)$$

in which  $k$  is an integer, ranging from 1 to  $n_\tau$ . Here,  $n_\tau$  is the total number of discretize points describing  $g(\tau)$ . The discrete representation of the travel-time distribution  $g(\tau)$  in the Laplace domain is:

$$\tilde{g}(s) = \sum_{k=1}^{n_\tau} \frac{1}{s^2 \Delta\tau} (\exp(-s(k-1)\Delta\tau) (g_k - g_{k-1} (1 - s\Delta\tau)) - \exp(-sk\Delta\tau) (g_k (1 + s\Delta\tau) - g_{k-1})) \quad (3.16)$$

##### Implementation of Shape-Free Inference

The problem of estimating  $g(\tau)$  is conceptually similar to inferring a transfer function in linear input-output relations. In the following, we adopt the non-parametric deconvolution approach of *Cirpka et al.* (2007) to the problem at hand.

The vector  $\mathbf{g}$  of discretize  $g(\tau)$  values is typically chosen large (a few hundred values) to allow the inference procedure to resolve features of  $g(\tau)$  on multiple scales. However, a fine discretization of  $g(\tau)$  makes the estimation unstable. For regularization, we assume that  $g(\tau)$  is a second-order intrinsic time variable with a linear semivariogram:

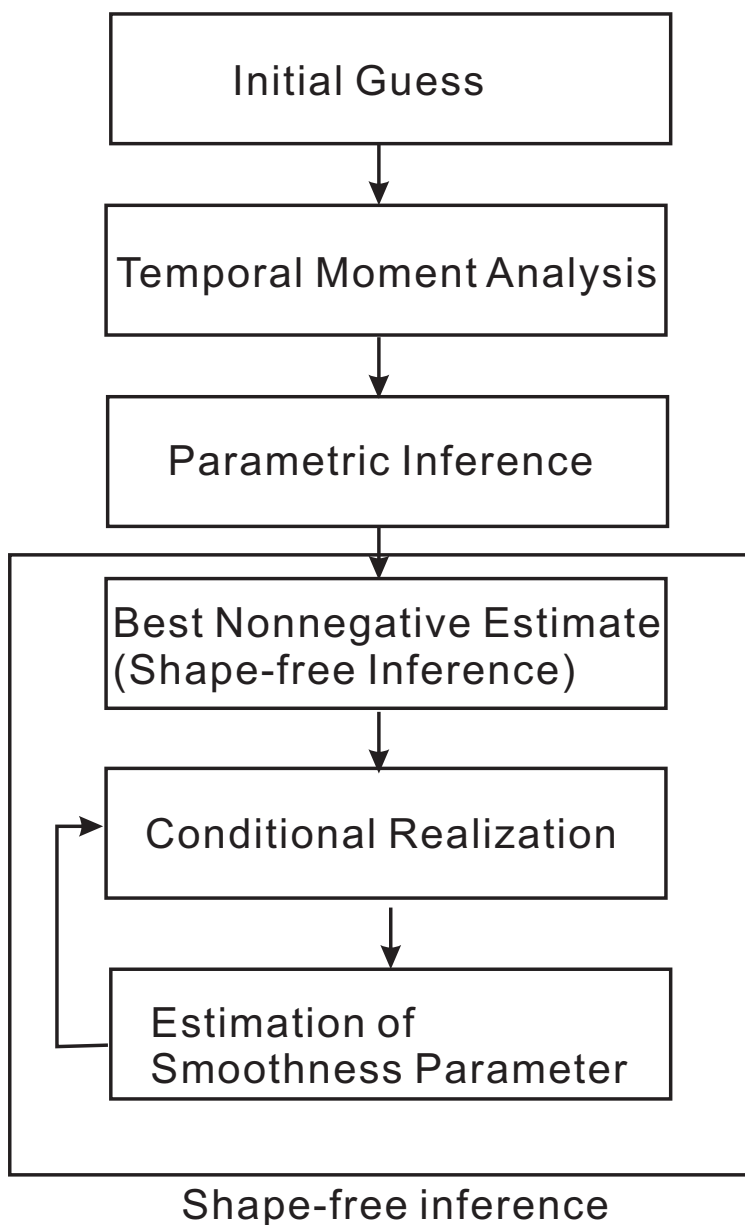


Figure 3.2: Flow chart of the general approach. The shape-free inference method requires an initial guess provided by parametric inference.

$$\gamma_g(h) = E \left[ \frac{1}{2} (g(\tau+h) - g(\tau))^2 \right] = \theta |h| \quad (3.17)$$

which needs to be considered only for values of  $h$  that are integer multiples of  $\Delta\tau$ . Regularization is commonly used in the inference of continuous parameter functions. It is actually needed as soon as the inverse problem becomes underdetermined. Assuming that the function to be determined is second-order intrinsic with a linear variogram is mathematical identical to first-order Tikhonov regularization, in which a term is introduced in the objective function penalizing strong gradients of the estimated function, here

$g(\tau)$  (Kitanidis, 1997).

Jointly with  $\mathbf{g}$ , we also estimate the parameters  $v$ ,  $D$ , and  $q_{he}$ . We consider the logarithms of  $v$ ,  $D$  and  $q_{he}$  to ensure non-negativity of the original variables. The nonnegativity of  $g(\tau)$  is not enforced by the logarithmic transformation of  $g(\tau)$ , because the true value of  $g(\tau)$  might be zero. Lagrange multipliers are employed for this purpose later.

Collecting all parameters to be estimated in a single  $n_p \times 1$  vector  $\mathbf{p}$  may be expressed as:

$$\mathbf{p} = \begin{bmatrix} \mathbf{g} \\ \mathbf{p}_{vDq} \end{bmatrix}_{n_p \times 1} \quad (3.18)$$

in which  $n_p (= n_\tau + 3)$  represents the total number of estimated parameters, and  $\mathbf{p}_{vDq} = [\ln(v), \ln(D), \ln(q_{he})]^T$ .

Applying Bayes theorem, assuming multi-Gaussian prior probability density functions, linearizing the functional dependence  $\mathbf{c}(\mathbf{p})$  about the current estimate of  $\mathbf{p}$ , and seeking for the most likely parameter combination  $\mathbf{p}$ , we obtain as objective function  $W(\mathbf{p})$  the negative logarithm of the posterior log-likelihood function. This objective is iteratively minimized in the inference procedure:

$$W(\mathbf{p}) = \frac{1}{2}(\mathbf{c}_* - \mathbf{J}\mathbf{p})^T \mathbf{C}^{-1}(\mathbf{c}_* - \mathbf{J}\mathbf{p}) - \frac{1}{2}\mathbf{g}^T \mathbf{\Gamma}_{gg}^{-1} \mathbf{g} + \mathbf{v}^T \mathbf{H}\mathbf{g} \quad (3.19)$$

in which  $c_*$ , stemming from successive linearization, is defined in Eq. (3.46);  $\mathbf{J}$  is the Jacobian matrix given in Eq. (3.45);  $\mathbf{C}$  is the covariance matrix of the measured concentrations, expressing both measurement and model errors;  $\mathbf{\Gamma}_{gg}$  is the  $n_g \times n_g$  matrix of discrete semivariogram values obtained for all pairs of elements in  $\mathbf{g}$ ;  $\mathbf{v}$  is the vector of Lagrange multipliers;  $\mathbf{H}$  is a  $n_v \times n_g$  selection matrix with  $n_v$  being the number of active constraints. The objective function  $W(\mathbf{p})$  consists of three terms. The first term quantifies the discrepancy between model outcome and measurements; the second term penalizes strong fluctuations of  $g(\tau)$ ; and the last term stems from the non-negativity constraint of  $g(\tau)$ . Details of the derivation can be found in Appendix 3.B. The objective function  $W(\mathbf{p})$  is minimized by setting the derivative with respect to  $\mathbf{p}$  and  $\mathbf{v}$  to zero, resulting in the following system of equations:

$$\begin{bmatrix} \mathbf{J}_g^T \mathbf{C}^{-1} \mathbf{J}_g - \mathbf{\Gamma}_{gg}^{-1} & \mathbf{J}_g^T \mathbf{C}^{-1} \mathbf{J}_{vDq} & \mathbf{H}^T \\ \mathbf{J}_{vDq}^T \mathbf{C}^{-1} \mathbf{J}_g & \mathbf{J}_{vDq}^T \mathbf{C}^{-1} \mathbf{J}_{vDq} & \mathbf{0} \\ \mathbf{H} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ \mathbf{p}_{vDq} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_g^T \mathbf{C}^{-1} \mathbf{c}_* \\ \mathbf{J}_{vDq}^T \mathbf{C}^{-1} \mathbf{c}_* \\ \mathbf{0} \end{bmatrix} \quad (3.20)$$

Activation and deactivation of the constraints depend on the behavior of the solution. The following rules apply:

$$\begin{aligned} g_j < 0 &\rightarrow \text{add constraint for element } j \\ v_i \leq 0 &\rightarrow \text{keep constraint } i \\ v_i > 0 &\rightarrow \text{remove constraint } i \end{aligned} \quad (3.21)$$

The implementation requires nested iterations: the outer iteration is the Gauss-Newton method, in which the forward operator is linearized about the current estimate, Eq. (3.44). Each Gauss-Newton step requires evaluation of the Jacobian  $\mathbf{J}$ . For a given Jacobian  $\mathbf{J}$ , we apply an inner iteration in order to identify the correct set of active constraints. The inner loop consists of alternating between solving the

system of linear equations, Eq. (3.20) and updating the constraints according to Eq. (3.21). The Gauss-Newton method is stabilized by a line search, that is, a Gauss-Newton step determined by the inner loop may be reduced if the objective function does not decrease from one outer iteration step to the next.

### Joint Inversion of Conservative and Reactive Tracer Data

A key difficulty in the joint estimation of the in-stream transport coefficients  $\nu$ ,  $D$  and the hyporheic parameters  $q_{he}$ ,  $g(\tau)$  is the non-uniqueness of the problem. Increasing  $q_{he}$  or putting larger probability mass onto higher values of  $\tau$  leads to later simulated breakthrough and more tailing. However, the same can be achieved by decreasing  $\nu$  and increasing  $D$ . To separate between in-stream mixing and mixing by hyporheic exchange, it is thus important to include data in the inversion that is particularly sensitive to hyporheic exchange. This is achieved by considering, in addition to a conservative tracer, a reactive tracer that undergoes degradation selectively in the hyporheic zone (Haggerty *et al.*, 2009).

Merging the concentration vectors for the conservative and reactive tracers, the variables in the objective function, Eq. (3.19), become  $(n_{con} + n_{rea}) \times 1$  vectors:

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}_{con} \\ \mathbf{c}_{rea} \end{bmatrix} \quad \text{and} \quad \mathbf{c}_* = \begin{bmatrix} \mathbf{c}_{*con} \\ \mathbf{c}_{*rea} \end{bmatrix} \quad (3.22)$$

in which  $n_{con}$  and  $n_{rea}$  are the numbers of conservative-tracer and reactive-tracer measurements, respectively. The covariance matrix is a  $(n_{con} + n_{rea}) \times (n_{con} + n_{rea})$  square matrix:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{con} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{rea} \end{bmatrix} \quad (3.23)$$

in which we have assumed that the measurements of the two tracers are uncorrelated. Finally, the sensitivity matrices  $\mathbf{J}_g$  and  $\mathbf{J}_{\nu Dq}$  are expressed by:

$$\mathbf{J}_g = \begin{bmatrix} \mathbf{J}_{g,con} \\ \mathbf{J}_{g,rea} \end{bmatrix} \quad \text{and} \quad \mathbf{J}_{\nu Dq} = \begin{bmatrix} \mathbf{J}_{\nu Dq,con} \\ \mathbf{J}_{\nu Dq,rea} \end{bmatrix} \quad (3.24)$$

All other aspects of the inference remain the same.

### Estimation of the Smoothness Parameter $\theta$

In the regularization step, we assume that the travel-time distribution  $g(\tau)$  is autocorrelated, expressed by a linear semivariogram with slope  $\theta$ , Eq. (3.17), which enforces the smoothness of the function. The smaller  $\theta$  is, the smoother is the function. In order to calculate the objective function, Eq.(3.19), this smoothness parameter  $\theta$  must be specified. In the current study, we estimate the smoothness parameter  $\theta$  by applying the expectation-maximization method (McLachlan and Krishnan, 1997). This method includes two steps: In the first step, we generate a set of conditional realizations with a given value of  $\theta$ , whereas in the second, we evaluate the value of  $\theta$  which best describes the smoothness of the conditional realizations of the hyporheic travel-time distribution.

The generation of conditional realizations is similar to the estimation of the nonnegative best estimate according to Eq. (3.20). We follow the procedure of the method of smallest possible modification (e.g., Journal and Huijbregts, 1978). Instead of estimating a single hyporheic travel-time distribution



$g(\tau)$ , we consider multiple realizations, each of them expressed as the sum of an unconditional realization  $\mathbf{g}'_u$ , exhibiting fluctuations in the travel-time distribution on all scales, and a smooth corrector term  $\mathbf{g}_c$  which is statistically isomorphic to  $\mathbf{g}$ :

$$\mathbf{g} = \mathbf{g}'_u + \mathbf{g}_c \quad (3.25)$$

Going through the same derivation steps as those leading to Eq. (3.20), the corresponding linear system of equations becomes:

$$\begin{bmatrix} \mathbf{J}_g^T \mathbf{C}^{-1} \mathbf{J}_g - \mathbf{\Gamma}_{gg}^{-1} & \mathbf{J}_g^T \mathbf{C}^{-1} \mathbf{J}_{vDq} & \mathbf{H}^T \\ \mathbf{J}_{vDq}^T \mathbf{C}^{-1} \mathbf{J}_g & \mathbf{J}_{vDq}^T \mathbf{C}^{-1} \mathbf{J}_{vDq} & \mathbf{0} \\ \mathbf{H} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{g}_c \\ \mathbf{p}_{vDq} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_g^T \mathbf{C}^{-1} (\mathbf{c}_* - \mathbf{J}_g \mathbf{g}'_u) \\ \mathbf{J}_{vDq}^T \mathbf{C}^{-1} (\mathbf{c}_* - \mathbf{J}_g \mathbf{g}'_u) \\ -\mathbf{H} \mathbf{g}'_u \end{bmatrix} \quad (3.26)$$

in which all symbols appearing before remain identical, and the rules of adjusting the constraints is the same as in calculating the best nonnegative estimate, see Eq. (3.21). The unconditional realization  $\mathbf{g}'_u$  is a random vector  $\mathbf{L}\mathbf{u}$  with the covariance properties of the system, where  $\mathbf{L}$  is the Cholesky decomposition of the negative inverse semivariogram matrix,  $-\mathbf{\Gamma}_{gg}$ , and  $\mathbf{u}$  is a vector of uncorrelated numbers drawn from a standard normal distribution.

Based on the set of conditional realizations, the smoothness parameter  $\theta$  can be updated in the maximization step of the expectation-maximization method. Because we consider a linear semivariogram, we only need to consider the pairwise squared difference of the element  $g(\tau)$  (Michalak and Kitanidis, 2003; Cirpka et al., 2007). The probability density function of  $\mathbf{g}$  given  $\theta$  can be expressed as:

$$p(\mathbf{g}|\theta) = \frac{1}{\sqrt{(2\pi)^{n_g} \det(-\mathbf{\Gamma}_{gg})}} \exp\left(-\frac{1}{2} \mathbf{g}^T \mathbf{\Gamma}_{gg}^{-1} \mathbf{g}\right) \quad (3.27)$$

and the updated smoothness parameter  $\theta$  is the one maximizing:

$$\varphi(\theta) = \sum_{j=1}^{n_r} \ln p(\mathbf{g}_j|\theta) \quad (3.28)$$

in which  $n_r$  is the number of realizations and  $n_g$  is the number of discretization points of the hyporheic travel-time distribution. After updating  $\theta$ , a new set of realizations is generated using the current value of  $\theta$ . This procedure is repeated until  $\theta$  has converged. As convergence criterion, we require that the relative difference between two solutions of  $\theta$  is less than 0.5%.

### 3.3.2 Initial Guesses

The process outlined above is iterative in nature. It requires initial guesses of all parameters, including the discretized hyporheic travel-time distribution. As initial guess for the shape-free inference of  $g(\tau)$ , we fit a parametric model to the tracer data. In the following steps of the estimation, this parameterization will be relaxed. As the simplest parametric model for  $g(\tau)$ , we use the exponential distribution, which is fully determined by a single parameter  $k$  [ $T^{-1}$ ]:

$$g(\tau) = k \exp(k\tau) \quad (3.29)$$

In the parametric inference, we estimate the parameters  $v$ ,  $D$ ,  $q_{he}$ , and  $k$  by fitting simulated tracer breakthrough curves to measured ones using an iterative procedure. This requires an initial guess of each parameter.

Of all parameters, the in-stream velocity  $v$  may be estimated most accurately (*Wörman and Wachniew, 2007*). We may do this by dividing the distance between the injection and observation points by the peak arrival time. Once  $v$  is given, we can use the semi-empirical equation of the asymptotic longitudinal dispersion coefficient  $D$  in a river by *Fischer (1967, 1979)* to obtain an initial guess of  $D$ :

$$D \approx 0.011 \frac{v^2 w^2}{v_* h} \quad (3.30)$$

in which  $w$  and  $h$  [L] are the width and depth of the river, respectively, and  $v_* \approx \sqrt{I_0 g h}$  [ $LT^{-1}$ ] is the bottom shear-stress velocity, here approximated by assuming normal depth (i.e., friction slope equals bottom slope),  $I_0$  [-] is the bottom slope, and  $g = 9.81$  [ $m/s^2$ ] is the acceleration constant due to gravity.

To obtain initial guesses of the hyporheic exchange rate  $q_{he}$  and characteristic parameter  $k$  of the hyporheic travel-time distribution for given values of  $v$  and  $D$ , we match temporal moments of the observed breakthrough curve (for details see Appendix 3.D):

$$k = \frac{2 \left( \frac{\mu_1 v}{\mu_0 x} - 1 \right) x}{\frac{\mu_{2c}}{\mu_0} v - \frac{2\mu_1^2 D}{\mu_0^2 x}} \quad (3.31)$$

$$q_{he} = \frac{2 \left( \frac{\mu_1 v}{\mu_0 x} - 1 \right)^2 x}{\frac{\mu_{2c}}{\mu_0} v - \frac{2\mu_1^2 D}{\mu_0^2 x}} \quad (3.32)$$

This set of parameters is used as initial guess for the parametric inference procedure, in which the sum of weighted squared differences between model output and measured concentration values is minimized. As optimization method, we use the Melder-Nead-simplex method implemented in the matlab function `fminsearch` (*Lagarias et al., 1999*). The results of the parametric estimation are thereafter used as initial guesses for the shape-free inference step.

### 3.4 Application to a Hypothetical Test Case

In this section, we apply the proposed shape-free inference method to a virtual tracer test experiment in a stream. Virtual breakthrough curves of a conservative and a reactive tracer are computed by solving Eq.(3.2) for given velocity  $v$ , dispersion coefficient  $D$ , decay coefficient  $\lambda$ , hyporheic-exchange coefficient  $q_{he}$ , and a hyporheic travel-time distribution  $g(\tau)$ , and adding noise to the data. Subsequently, the true parameters are forgotten, and we estimate all parameters, including the hyporheic travel-time distribution  $g(\tau)$ , by joint inversion of the noisy conservative and reactive tracer BTCs using the shape-free method explained in Section 3.3.1.

### 3.4.1 Model Setup

We consider a river undergoing hyporheic exchange with in-stream velocity  $v \approx 0.44m/s$  and in-stream dispersion  $D \approx 1.67m^2/s$  which are assumed to be uniform over the river reach. These parameters are consistent with a river having an approximately rectangular profile with width  $w = 5m$ , depth  $h = 0.5m$ , bottom slope  $I_0 = 0.1\%$ , and Manning's roughness coefficient  $n = 0.04$ , assuming normal depth. We simulate a pulse-like injection experiment with the introduction of  $1kg$  tracer and compute the tracer concentration  $10km$  downstream of the injection point. In this example, all parameters, including the hyporheic travel-time distribution  $g(\tau)$ , do neither change over the length of the river reach nor over the duration of the experiment. The values of all parameters are listed in Table 3.1.

The virtual hyporheic travel-time distribution  $g(\tau)$  is constructed by summing an exponential and a log-normal distribution, leading to a secondary peak (shown in Fig. 3.3). Such a hyporheic travel-time distribution could result from advective-dispersive transport through preferential flow paths in the hyporheic zone. Double peaked breakthrough curves are very well known in transport of compounds undergoing inter-phase mass transfer (e.g., *Quindoz and Valocchi, 1993*), which is mathematically identical to the transient-storage model. This phenomenon already occurs with a standard single immobile domain undergoing first-order mass transfer with the mobile domain. That is, double peaks in the breakthrough curve do not require a secondary peak in the hyporheic travel-time distribution. Late-time peak in travel-time distribution can be found in *Cardenas et al. (2008)*, *Cardenas (2008a)* and *Cardenas (2007)*. The given type of hyporheic travel-time distribution cannot be approximated well by a single parametric distribution commonly used in hyporheic-transport studies, such as the exponential and power-law functions. We truncate the hyporheic travel-time distribution at  $\tau = 22h$ , and discretize it with time increments of  $10min$ . The simulated breakthrough curves of the conservative and reactive tracers are discretized by  $1min$  increments.

Fig. 3.4 shows the simulated, noise-free breakthrough curves of the conservative and reactive tracers at the downstream observation point. In total, we simulate about 30 hours. Both breakthrough curves exhibit bimodal behavior. The first, rather narrow peak represents transport within the stream practically not undergoing any hyporheic exchange. Because the reaction is restricted to the hyporheic zone, the first peak is almost identical for the conservative and reactive tracers. The second, long tailed peak represents transport of solutes that have undergone substantial hyporheic exchange. As a consequence, this peak is much more pronounced for the conservative tracer than for the reactive one. As mentioned above, the difference of the two breakthrough curves helps distinguishing between in-stream mixing and longitudinal spreading due to hyporheic exchange.

In the following inversion, we use noisy breakthrough curves by adding noise to both curves applying the following rule:

$$c_{m,i} = c_i + (\epsilon_{rel} \times c_i + \epsilon_{abs} \times \max(c)) \times z_i \quad (3.33)$$

in which  $\epsilon_{rel} [-]$  is the relative error,  $\epsilon_{abs} \times \max(c) [ML^{-3}]$  is the absolute error, here expressed as a fraction of the maximum concentration, and  $z_i$  is a random variable drawn from a standard normal distribution. Fig. 3.4 shows a comparison between the noise-free (solid line) and noisy (dotted line) breakthrough curve of the conservative tracer. Only the latter is used as virtual measurement in the example application.

### 3.4.2 Results of Inversion

As shown in Fig. 3.5, we obtain a fairly good fit for both the conservative and reactive tracer tracers BTCs. Again, we hardly see any differences for the first peaks between conservative and reactive tracers,

Table 3.1: Parameters of the virtual test case.

Physical Parameters		
$v$	0.44 [m/s]	in-stream velocity
$D$	1.67 [m <sup>2</sup> /s]	in-stream dispersion
$m$	1 [kg]	injected mass of the tracer compound
$x$	1 × 10 <sup>4</sup> [m]	length of river reach
$\lambda$	5 × 10 <sup>-5</sup> [1/s]	first-order decay coefficient in hyporheic zone
$q_{he}$	1 × 10 <sup>-4</sup> [1/s]	hyporheic exchange rate coefficient
Parameters to Construct the Hyporheic Travel-Time Distribution $g(\tau)$		
$\mu_{exp}$	5 × 10 <sup>3</sup> [s]	first moment of exponential contribution
$\mu_{LN}$	2 × 10 <sup>4</sup> [s]	mean value of the lognormal contribution
$\sigma$	7.4 × 10 <sup>3</sup> [s]	variance of the lognormal contribution
Numerical Parameters		
$t$	5 - 31.2 [h]	range of simulation time
$\Delta t$	1 [min]	discretization in time
$\tau$	0 - 21 [h]	range of hyporheic travel time
$\Delta \tau$	10 [min]	discretization in travel time
$\epsilon_{abs}$	0.01 [-]	absolute error/maximum concentration
$\epsilon_{rel}$	0.05 [-]	relative error

whereas the long tails are well reproduced in the model fit. The noise-free true BTCs in Fig. 3.4 and the fitted BTCs in Fig. 3.5 hardly differ, and the model passes the Chi-square test quite well ( $P = 0.29$ ). The estimated velocity, dispersion coefficient, and strength of hyporheic exchange are compared to the true values in Table 3.2, which also includes the uncertainty of the estimates obtained from the set of conditional realizations. The deviations between the true parameter values and the mean estimates are all within the obtained uncertainty bounds. The coefficients of variation for these three parameters are:  $CV_v = 0.0042$ ,  $CV_D = 0.2263$ , and  $CV_{q_{he}} = 0.0649$  respectively. This result indicates that the in-stream velocity can be estimated the most accurate among these three parameters, whereas the dispersion coefficient remains the most uncertain. This is to some extent consistent with the understanding of our model. Different combinations of in-stream dispersion coefficient and hyporheic exchange rate might lead to a similar shape of the BTC measured downstream. The joint inversion of conservative and reactive tracers has improved the process differentiation, but some ambiguity remains. The correlation coefficients of estimation between the hyporheic exchange rate  $q_{he}$  and the other two parameters are  $r_{v,q_{he}} = 0.94$  and  $r_{D,q_{he}} = -0.42$ , respectively, indicating a high positive relationship between the uncertainty of the in-stream velocity  $v$  and the hyporheic exchange rate  $q_{he}$ . However, the estimation variance of the hyporheic exchange rate  $q_{he}$  is so small that the correlation of uncertainty does not pose any parameter identification problem. The negative relationship between in-stream dispersion and hyporheic exchange rate may again be explained by different combinations of these two parameters providing similar BTCs. If the in-stream dispersion  $D$  is estimated higher, the hyporheic exchange rate  $q_{he}$  is estimated lower, and vice versa, because the hyporheic exchange contributes to the spreading of the BTC like in-stream dispersion.

Fig. 3.6 shows the estimated hyporheic travel-time distribution. The smooth dotted line is the artificial true travel-time distribution used for generating the virtual measurement data. The dashed line represents the estimated travel-time distribution using an exponential function. We see that it estimates

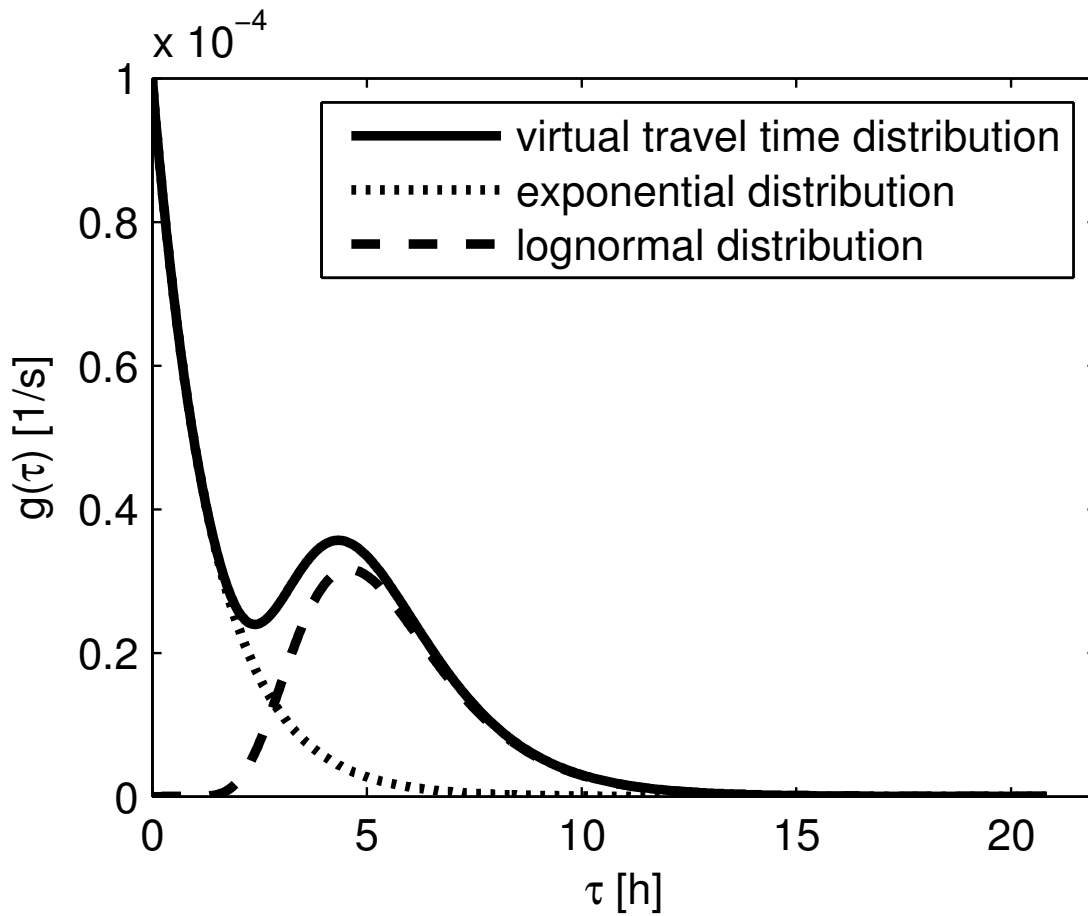


Figure 3.3: True hyporheic travel-time distribution  $g(\tau)$  in the virtual test case (solid line), generated by summing an exponential function (dotted line) and a log-normal distribution (dashed line).

the early return times well, but fails to capture the main feature of the true travel-time distribution, namely the secondary peak. The solid line is the estimated hyporheic travel-time distribution using the shape-free method, which clearly presents the feature of the secondary peak. After approximately 1 hour travel time, the true travel-time distribution lies within the 16%-84% probability band of the conditional realizations. This probability band has been chosen because it represents the mean value  $\pm$  one standard deviation in case of a Gaussian distribution. The probability of very small travel times is overestimated. It may be worth noting that hyporheic exchange with small travel times is particularly difficult to separate from in-stream dispersion. Thus, the slight underestimation of the dispersion coefficient (see Table 3.2) may have been compensated by the overestimation of early-time hyporheic exchange. The long tails of the BTCs are not affected by this mismatch.

In the estimation of the smoothness parameter  $\theta$ , we use an initial guess of  $1 \times 10^{-14} [1/s^{-3}]$ . After 22 iterations,  $\theta$  converges to  $7.6 \times 10^{-14} [1/s^{-3}]$  which describes the smoothness of the set of solutions best. The true travel-time distribution  $g(\tau)$  is considerably smoother (see Fig. 3.6), indicating that noisy data cause somewhat noisy estimated hyporheic travel-time distributions. However, most of the true travel-time distribution lies within the band of 16%-84% probability computed by the conditional real-

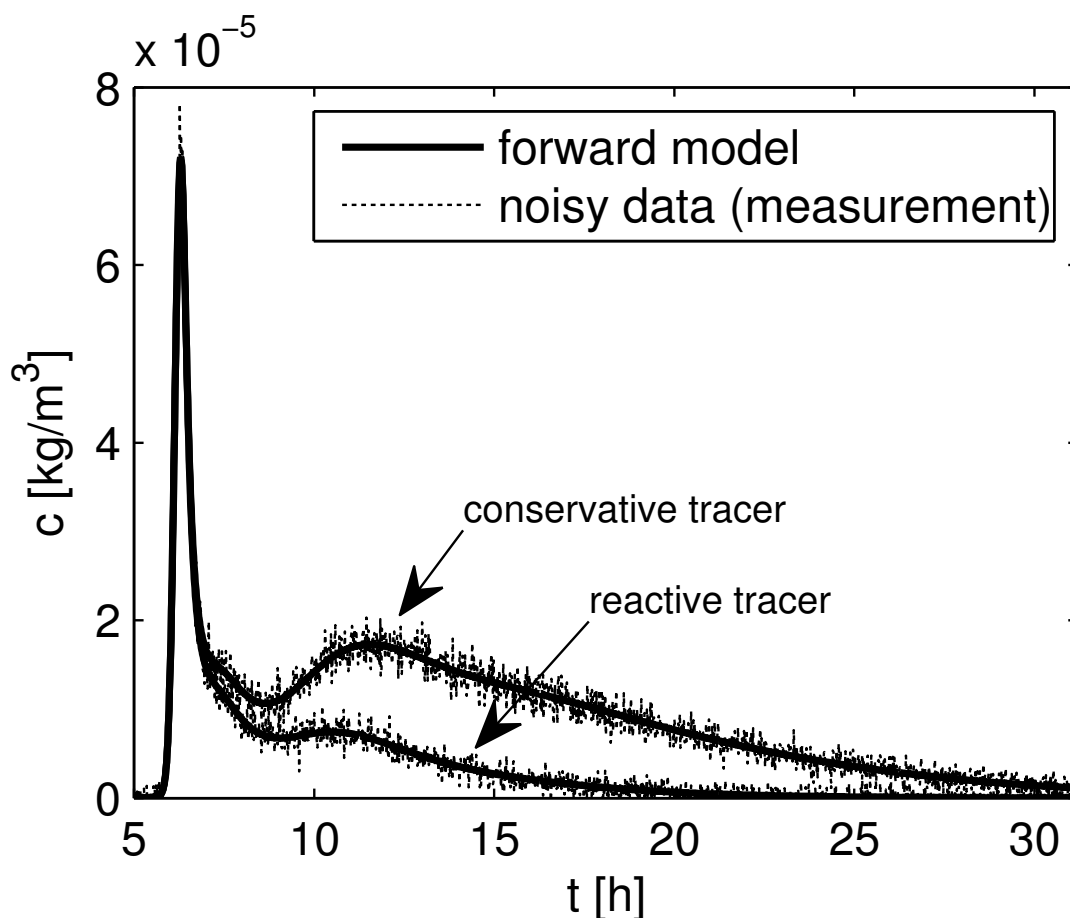


Figure 3.4: Virtual true breakthrough curves of conservative and reactive tracers. Solid lines: breakthrough curves calculated from forward model; dotted lines: noise-added breakthrough curves used as virtual measurements in the inversion.

izations.

The computational costs of the estimation procedure highly depend on the number of discretization points of the travel-time distribution and the BTCs, the initial guess, and the convergence criterion. The generation of plausible realizations is implemented as parallel Matlab code, and takes about 58 hours on a quad core processor, including the estimation of the smoothness parameter  $\theta$ . As indicated above, the generation of conditional realizations without estimation of  $\theta$  would be about 20 times quicker.

### 3.5 Discussion and Conclusions

The transient storage model is a well accepted framework for the description of reach-scale solute transport in streams undergoing substantial hyporheic exchange. We have presented a method to infer in-stream transport parameters together with the hyporheic travel-time distribution from combined conservative and 'smart' tracer tests without relying on a particular functional form of the hyporheic travel-time distribution. Such an approach is advantageous because it overcomes debates of the best para-

Table 3.2: Comparison between true and estimated parameter values without hyporheic travel-time distribution. "Single est.": best nonnegative shape-free estimate without conditional realizations; "cond. real.": results of conditional realizations.

	True	Estimated	
		single est.	cond. real.
$v$ [m/s]	0.441	0.443	$0.444 \pm 0.002$
$D$ [ $m^2/s$ ]	1.67	1.64	$1.58 \pm 0.37$
$q_{he} \times 10^{-4}$ [1/s]	1	1.05	$1.12 \pm 0.073$
$\theta$ [ $1/s^{-3}$ ]	-	-	$7.6 \times 10^{-4}$

metric model and facilitates identifying unconventional features such as broad or multiple peaks of the distribution.

The approach presented in this study has been transferred from a previously published method of identifying linear transfer functions from fluctuating input and output signals (*Cirpka et al., 2007; Payn et al., 2008*). The current application is more challenging, because the measured data depend on the distribution to be estimated in a nonlinear way, and the dependence is less direct. The stream tracer test does not directly provide the time over which the tracer has remained in the hyporheic zone, but an integral signal of tracers that have undergone hyporheic exchange multiple times. Therefore, one should not expect obtaining the same accuracy in inversion as in straight linear deconvolution.

Because in-stream mixing processes and hyporheic exchange have similar effects on the overall breakthrough curve of a conservative compound measured further downstream, it is important to include data sets with different sensitivity towards hyporheic exchange. For this purpose, *Haggerty et al. (2008)* have introduced resazurin as 'smart' tracer selectively undergoing decay during the hyporheic passage. The current contribution makes use of this concept and presents a method of joint quantitative analysis of a conservative tracer and resazurin. We assess this approach to be promising, even though some ambiguity between in-stream dispersion and rapid hyporheic exchange remains.

Our approach is based on a series of assumptions, the most important one being stationarity of all parameters in space and time. If there was clear evidence of space-dependent in-stream transport parameters, we could account for this by replacing the analytical expression of solute transport in the Laplace domain by numerical simulations. Identifying different hyporheic travel-time distributions along the length of the river, by contrast, would be conceptually and numerically demanding. Most likely, we would not recommend attempting such an inversion using a shape-free inference method. Even for parametric inference, one would have to doubt that data sets of sufficient quality exist to calibrate such models.

The computational effort of the inverse approach is larger than in the estimation of shape parameters when a parametric model of  $g(\tau)$  is used. The highest effort comes with the estimation of the smoothness parameter  $\theta$ . The latter may be replaced by a manual choice of  $\theta$  based on visual inspection of the estimated hyporheic travel-time distributions. If the estimated distribution  $g(\tau)$  resembles a particular functional shape, it is permissible to continue with that functional form. In tests, not shown, we could demonstrate that our approach recovers commonly used distributions. The decisive advantage is that we do not need to guess the correct functional shape, we determine it from the data themselves.

Testing our method by application to a virtual tracer test, we are fairly optimistic with respect to the

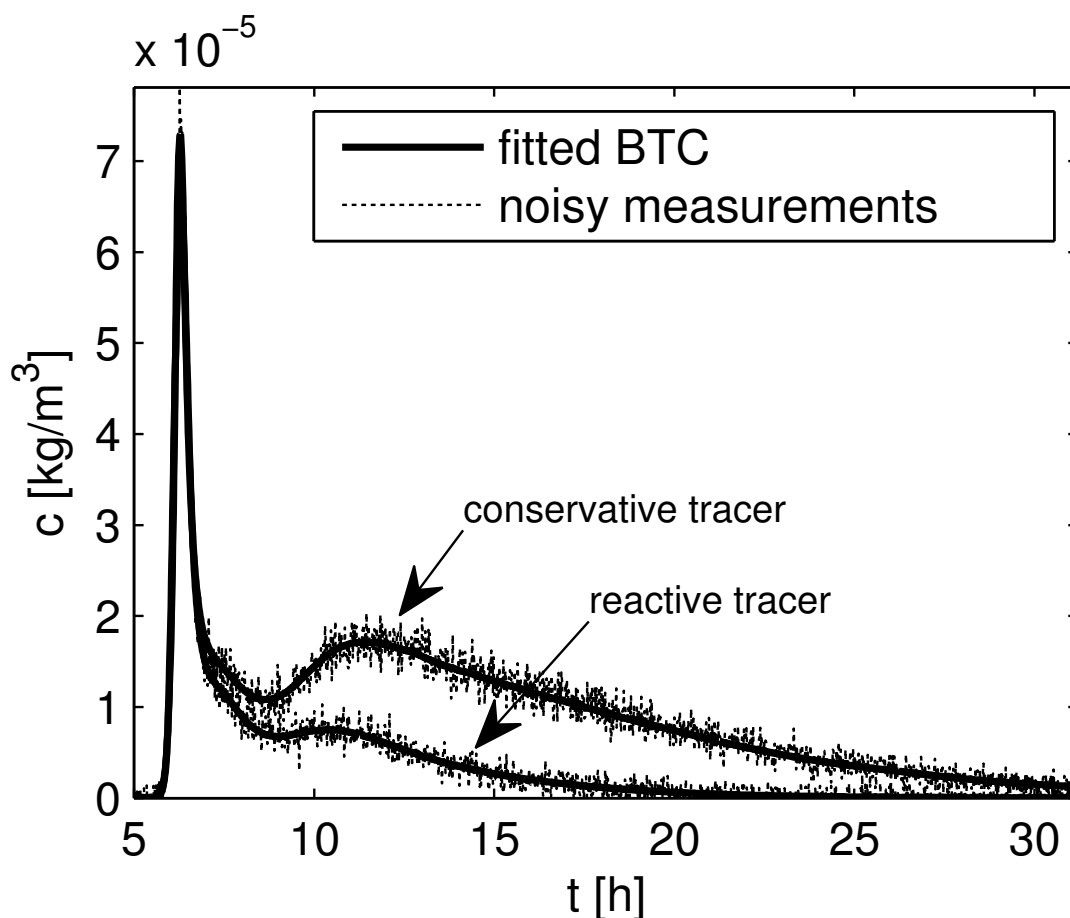


Figure 3.5: Fitted breakthrough curves obtained by the joint shape-free inference method. Solid lines: fitted breakthrough curves of conservative and reactive tracers; dotted lines: measurements of tracer concentration used for the inversion.

amount of data included in the measurements. Specifically, we have assumed that time series, as obtained by continuous sensor measurements, exist. While such measurements are in principle possible, the field studies known to us are based on sampling (*Haggerty et al.*, 2008, 2009). It should be clear that a full distribution cannot be obtained from a few samples. However, it might be possible to combine sensor-based conservative-tracer measurements (e.g., of fluorescein) with a limited number of reactive-tracer samples. Our approach could be adopted to such type of data.

## Acknowledgments

We are grateful to Michael Gooseff, Timothy Ginn, and an anonymous reviewer for their constructive comments. The German Academic Exchange Program DAAD supported the first author by the Ph.D scholarship “Mathematical Analysis of Natural and Artificial Tracers for the Quantification of River-Groundwater Exchange”



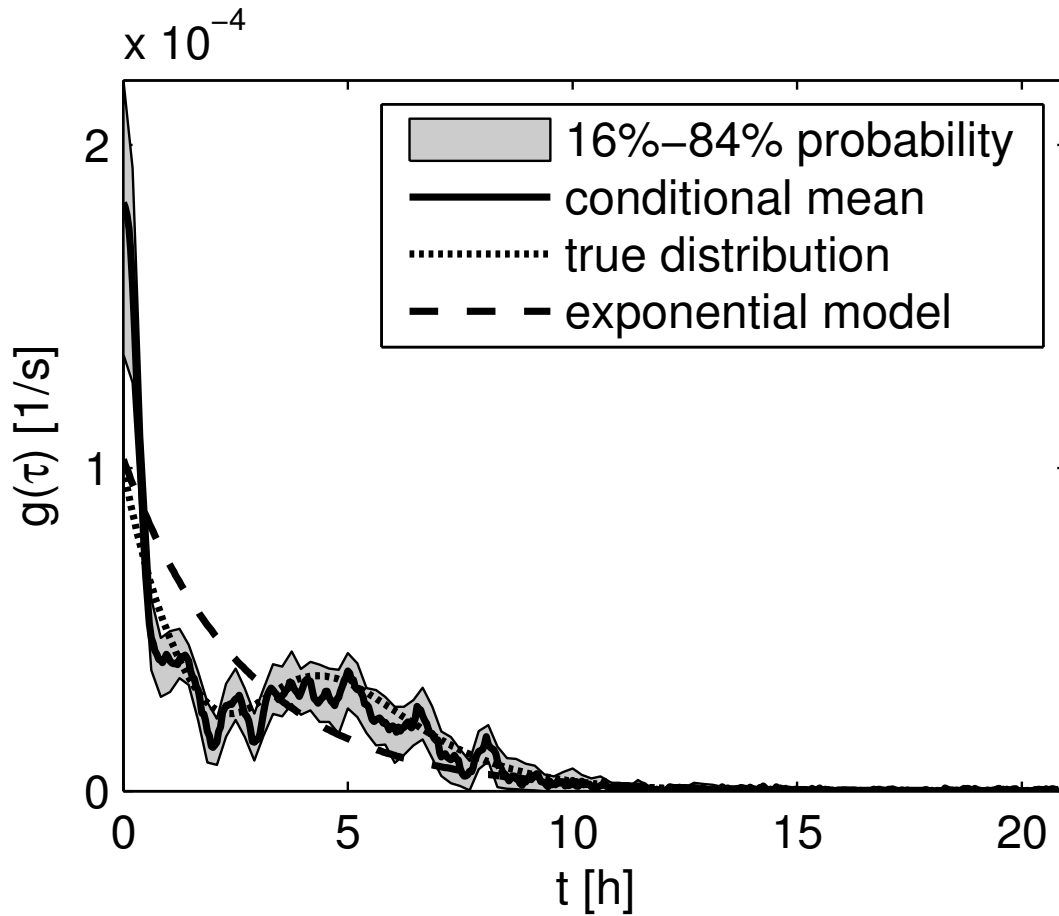


Figure 3.6: Estimated hyporheic travel-time distributions  $g(\tau)$ . Solid line: mean value of  $g(\tau)$  from the set of conditional realizations; gray band: range of 16% – 84% probability of the conditional statistical distributions; dotted line: virtual true hyporheic travel-time distribution; dashed line: exponential function obtained in the parametric inference step (initial guess of shape-free inference).

### 3.A Equivalence between Eq. (3.1) and Eq. (3.2)

In Eq. (3.1), the injected tracer concentration  $c_{inj}$  is considered as a pulse injection, which is included in the boundary condition in our model, Eq. (3.4). For the purpose of showing the equivalence between the two equations, we consider a conservative tracer, that is,  $\lambda = 0$  in Eq. (3.2). The advection-dispersion operator is identical in both equations, so that the only difference lies in the term related to the hyporheic exchange. In Eq. (3.1), the hyporheic exchange is described as:

$$\begin{aligned}
-\beta \int_0^t \frac{\partial c(t-\tau)}{\partial t} g^*(\tau) d\tau &= \beta g^*(\tau) c(t-\tau) \Big|_0^t - \beta \int_0^t c(t-\tau) \frac{\partial g^*(\tau)}{\partial \tau} d\tau \\
&= \beta g^*(t) c(0) - \beta g^*(0) c(t) - \beta \int_0^t c(t-\tau) \frac{\partial g^*(\tau)}{\partial \tau} d\tau \\
&= -\beta g^*(0) c(t) - \beta \int_0^t c(t-\tau) \frac{\partial g^*(\tau)}{\partial \tau} d\tau
\end{aligned} \tag{3.34}$$

which we compare to the corresponding expression in Eq. (3.2):

$$q_{he} \left( \int_0^t g(\tau) c(t-\tau) \exp(-\lambda \tau) d\tau - c(t) \right) \tag{3.35}$$

which is identical to the term of Eq. (3.1) if the following identities hold:

$$q_{he} = \beta g^*(0) \tag{3.36}$$

$$g(\tau) = -\frac{1}{g^*(0)} \frac{\partial g^*(\tau)}{\partial \tau} \tag{3.37}$$

### 3.B Derivation of the Objective Function $W(\mathbf{p})$

For a given measured breakthrough curve  $c_{m,i} = c(t_i, x_{meas})$  at the measurement location  $x_{meas}$ , the conditional probability density function  $p(\mathbf{g}, \ln(v), \ln(D), \ln(q_{he}) | \mathbf{c}_m)$ , given the measurement  $\mathbf{c}_m$ , is given by Bayes' theorem:

$$p(\mathbf{p} | \mathbf{c}_m, \theta) = \frac{p(\mathbf{c}_m | \mathbf{p}) p(\mathbf{p} | \theta)}{p(\mathbf{c}_m)} \tag{3.38}$$

in which  $p(\mathbf{c}_m | \mathbf{p})$  is the likelihood of measurements  $\mathbf{c}_m$  for given parameters  $\mathbf{p}$ ,  $p(\mathbf{p} | \theta)$  is the prior distribution of the parameters for given slope  $\theta$  of the semivariogram, and  $p(\mathbf{c}_m)$  is independent of the parameters to be estimated.

We assume that the likelihood and the prior probability density functions are multi-Gaussian:

$$p(\mathbf{c}_m | \mathbf{p}) \propto \exp \left( -\frac{1}{2} (\mathbf{c}_m - \mathbf{c}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{c}_m - \mathbf{c}(\mathbf{p})) \right) \tag{3.39}$$

$$p(\mathbf{p} | \theta) \propto \exp \left( \frac{1}{2} \mathbf{p}^T \Gamma_{pp}^{-1} \mathbf{p} \right) \tag{3.40}$$

in which  $\mathbf{c}_m$  is the vector of measured concentrations, whereas  $\mathbf{c}(\mathbf{p})$  is the corresponding model outcome for given parameters  $\mathbf{p}$ ,  $\mathbf{C}$  is the covariance matrix expressing both measurement and model errors,  $\Gamma_{pp}$  is the negative generalized prior covariance matrix of all parameters; its inverse is a partitioned matrix with the following blocks:

$$\Gamma_{pp}^{-1} = \begin{bmatrix} \Gamma_{gg}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}_{n_p \times n_p} \tag{3.41}$$

$\Gamma_{gg}$  is the  $n_g \times n_g$  matrix of discrete semivariogram values obtained for all pairs of elements in  $\mathbf{g}$ , with  $\Gamma_{gg}(i, j) = \theta|i - j|\Delta\tau$ . Note that:

$$\mathbf{p}^T \Gamma_{pp}^{-1} \mathbf{p} = \mathbf{g}^T \Gamma_{gg}^{-1} \mathbf{g} \quad (3.42)$$

so that the log-conditional posterior probability density function  $L(\mathbf{p}|\mathbf{c}, \theta) = \log(p(\mathbf{p}|\mathbf{c}, \theta))$  of the parameters for given measurements is:

$$L(\mathbf{p}|\mathbf{c}_m, \theta) = -\frac{1}{2}(\mathbf{c}_m - \mathbf{c}(\mathbf{p}))^T \mathbf{C}^{-1}(\mathbf{c}_m - \mathbf{c}(\mathbf{p})) + \frac{1}{2}\mathbf{g}^T \Gamma_{gg}^{-1} \mathbf{g} + const. \quad (3.43)$$

which we maximize in order to infer the most likely values of  $\mathbf{p}$ .

Our forward operator is the nonlocal transport equation (3.2). The relationships between concentration  $c(t)$  and all parameters  $g(\tau)$ ,  $v$ ,  $D$ , and  $q_{he}$  are nonlinear, since they appear as products in the governing equation. This implies that the mode of  $L(\mathbf{p}|\mathbf{c}, \theta)$  cannot be obtained by solving a single system of linear equations. Thus, we apply successive linearization about the last estimate  $\mathbf{p}_k$  (with  $k$  being the iteration index):

$$\mathbf{c}(\mathbf{p}) \approx \mathbf{c}(\mathbf{p}_k) + \mathbf{J}(\mathbf{p} - \mathbf{p}_k) \quad (3.44)$$

in which  $\mathbf{J}$  is the Jacobian, or sensitivity matrix, defined by

$$\mathbf{J} = \left[ \frac{\partial \mathbf{c}}{\partial \mathbf{g}}, \frac{\partial \mathbf{c}}{\partial \ln(v)}, \frac{\partial \mathbf{c}}{\partial \ln(D)}, \frac{\partial \mathbf{c}}{\partial \ln(q_{he})} \right] = [\mathbf{J}_g, \mathbf{J}_{vDq}] \quad (3.45)$$

with  $\mathbf{J}_g := \frac{\partial \mathbf{c}}{\partial \mathbf{g}}, \mathbf{J}_{vDq} := \left[ \frac{\partial \mathbf{c}}{\partial \ln(v)}, \frac{\partial \mathbf{c}}{\partial \ln(D)}, \frac{\partial \mathbf{c}}{\partial \ln(q_{he})} \right]$

The numerical method of obtaining the Jacobian  $\mathbf{J}$  is described in detail in Appendix 3.C. We also denote:

$$\mathbf{c}_* = \mathbf{c}_m - \mathbf{c}(\mathbf{p}_k) + \mathbf{J}\mathbf{p}_k \quad (3.46)$$

for further use.

Substituting Eqs. (3.44 & 3.46) into the log-likelihood function Eq. (3.43), we arrive at:

$$L(\mathbf{p}|\mathbf{c}, \theta) \approx -\frac{1}{2}(\mathbf{c}_* - \mathbf{J}\mathbf{p})^T \mathbf{C}^{-1}(\mathbf{c}_* - \mathbf{J}\mathbf{p}) + \frac{1}{2}\mathbf{g}^T \Gamma_{gg}^{-1} \mathbf{g} + const. \quad (3.47)$$

The solution of  $\mathbf{p}$  maximizing Eq. (3.47) may contain negative entries of the discrete travel-time distribution  $\mathbf{g}$ , which contradicts that  $g(\tau)$  is a probability density function:

$$\mathbf{g} \geq \mathbf{0} \quad (3.48)$$

We enforce non-negativity of  $\mathbf{g}$  by the method of Lagrange multipliers (e.g., *Vogel*, 2002). Because Eq. (3.48) is an inequality constraint, we need to apply an iterative procedure, in which equality constraints of the type  $g_i = 0$  are set wherever necessary. For the entire vector  $\mathbf{g}$  this may be expressed by:

$$\mathbf{H}\mathbf{g} = \mathbf{0} \quad (3.49)$$

in which  $\mathbf{H}$  is a  $n_v \times n_g$  selection matrix with  $n_v$  being the number of active constraints.  $H_{i,j}$  is unity if  $i$ -th constraint enforces the  $j$ -th element of  $\mathbf{g}$  to be zero.

In the method of Lagrange multipliers, we multiply each constraint  $i$  with a multiplier  $v_i$  and add it to the objective function, which is the negative of the linearized log-likelihood given in Eq. (3.47). That is, our constrained objective function below, which needs to be minimized:

$$W = \frac{1}{2}(\mathbf{c}_* - \mathbf{Jp})^T \mathbf{C}^{-1}(\mathbf{c}_* - \mathbf{Jp}) - \frac{1}{2}\mathbf{g}^T \Gamma_{gg}^{-1} \mathbf{g} + \mathbf{v}^T \mathbf{Hg} \quad (3.50)$$

### 3.C Sensitivity Analysis

In order to compute sensitivities of concentrations with respect to all transport parameters, we perform a small-perturbation analysis of Eq. (3.2). We separate each variable into its mean value, about which the sensitivity is derived, and the fluctuations about the mean:

$$c = \bar{c} + c', \quad v = \bar{v} + v', \quad D = \bar{D} + D', \quad q_{he} = \bar{q}_{he} + q'_{he}, \quad g = \bar{g} + g' \quad (3.51)$$

in which symbols with an overbar are the mean values and those with the prime are the fluctuations. Substituting Eq. (3.51) into Eqs. (3.2-3.5), dropping products of perturbations, and separating the unperturbed and perturbed parts results in:

$$\frac{\partial \bar{c}}{\partial t} + \bar{v} \frac{\partial \bar{c}}{\partial x} - \bar{D} \frac{\partial^2 \bar{c}}{\partial x^2} - \bar{q}_{he} \left( \int_0^\infty \bar{g}(\tau) \bar{c}(t - \tau) \exp(-\lambda \tau) d\tau - \bar{c}(t) \right) = 0 \quad (3.52)$$

$$\bar{c}(t = 0, x) = 0 \quad (3.53)$$

$$\left( \bar{v} \bar{c} - \bar{D} \frac{\partial \bar{c}}{\partial x} \right) \Big|_{x=0} = \frac{m}{A} \delta(t) \quad (3.54)$$

$$\lim_{x \rightarrow \infty} \bar{c} = 0 \quad \forall t \quad (3.55)$$

$$\frac{\partial c'}{\partial t} + \bar{v} \frac{\partial c'}{\partial x} - \bar{D} \frac{\partial^2 c'}{\partial x^2} - \bar{q}_{he} \left( \int_0^\infty \bar{g}(\tau) c'(t - \tau) \exp(-\lambda \tau) d\tau + c'(t) \right) = A \quad (3.56)$$

$$c'(t = 0, x) = 0 \quad (3.57)$$

$$\left( \bar{v} c' - \bar{D} \frac{\partial c'}{\partial x} \right) \Big|_{x=0} = B \quad (3.58)$$

$$\lim_{x \rightarrow \infty} c' = 0 \quad \forall t \quad (3.59)$$

with:

$$A = \begin{cases} -v' \frac{\partial \bar{c}}{\partial x} & \text{if perturbed in } v \\ D' \frac{\partial^2 \bar{c}}{\partial x^2} & \text{if perturbed in } D \\ q'_{he} \left( \int_0^\infty \bar{g}(\tau) \bar{c}(t-\tau) \exp(-\lambda \tau) d\tau - \bar{c}(t) \right) & \text{if perturbed in } q_{he} \\ \bar{q}_{he} \int_0^\infty g'(\tau) \bar{c}(t-\tau) \exp(-\lambda \tau) d\tau & \text{if perturbed in } g(\tau) \end{cases} \quad (3.60)$$

$$B = \begin{cases} -v' \bar{c}(x=0) & \text{if perturbed in } v \\ D' \frac{\partial \bar{c}(x=0)}{\partial x} & \text{if perturbed in } D \\ 0 & \text{if perturbed in } q_{he} \text{ or } g(\tau) \end{cases} \quad (3.61)$$

Eqs. (3.52-3.55) are identical to Eqs. (3.2-3.5), except for the overbar, and the result in the Laplace domain is given in Eqs. (3.12-3.14). The Laplace transforms of Eqs. (3.56-3.59) are:

$$(s + \bar{q}_{he}(1 - \bar{g}(s + \lambda))) \bar{c}' + \bar{v} \frac{d\bar{c}'}{dx} - \bar{D} \frac{d^2 \bar{c}'}{dx^2} = \bar{A} \quad (3.62)$$

$$\left( \bar{v} \bar{c}' - \bar{D} \frac{d\bar{c}'}{dx} \right) \Big|_{x=0} = \bar{B} \quad (3.63)$$

$$\lim_{x \rightarrow \infty} \bar{c}' = 0 \quad (3.64)$$

with

$$\bar{A} = \begin{cases} -v' \frac{d\bar{c}}{dx} & \text{if perturbed in } v \\ D' \frac{d^2 \bar{c}}{dx^2} & \text{if perturbed in } D \\ q'_{he} (\bar{g}(s + \lambda) - 1) \bar{c} & \text{if perturbed in } q_{he} \\ \bar{q}_{he} \bar{g}'(s + \lambda) \bar{c} & \text{if perturbed in } g(\tau) \end{cases} \quad (3.65)$$

$$\bar{B} = \begin{cases} -v' \bar{c}(x=0) & \text{if perturbed in } v \\ D' \frac{d\bar{c}(x=0)}{dx} & \text{if perturbed in } D \\ 0 & \text{if perturbed in } q_{he} \text{ and } g(\tau) \end{cases} \quad (3.66)$$

The solution of these equations has the general form:

$$\bar{c}' = (kx + a) \exp(-\alpha x) \quad (3.67)$$

with coefficients  $k$  and  $a$  of:

$$k = \begin{cases} \frac{\alpha v' \tilde{c}_0}{\bar{v} + 2\alpha \bar{D}} & \text{if perturbed in } v \\ \frac{\alpha^2 D' \tilde{c}_0}{\bar{v} + 2\alpha \bar{D}} & \text{if perturbed in } D \\ \frac{q'_{he} (\tilde{g}(s + \lambda) - 1) \tilde{c}_0}{\bar{v} + 2\alpha \bar{D}} & \text{if perturbed in } q_{he} \\ \frac{\bar{q}_{he} \tilde{g}(s + \lambda) \tilde{c}_0}{\bar{v} + 2\alpha \bar{D}} & \text{if perturbed in } g(\tau) \end{cases} \quad (3.68)$$

$$a = \begin{cases} \frac{k\bar{D} - v' \tilde{c}_0}{\bar{v} + \alpha \bar{D}} & \text{if perturbed in } v \\ \frac{k\bar{D} - D' \alpha \tilde{c}_0}{\bar{v} + \alpha \bar{D}} & \text{if perturbed in } D \\ \frac{k\bar{D}}{\bar{v} + \alpha \bar{D}} & \text{if perturbed in } q_{he} \text{ and } g(\tau) \end{cases} \quad (3.69)$$

### 3.D Temporal Moment Analysis

The  $i$ th temporal moment of a breakthrough curve obtained at location  $x$  is defined as:

$$\mu_i(x) = \int_0^\infty t^i c(x, t) dt \quad (3.70)$$

in which  $c(x, t)$  is the tracer concentration measured in the stream. The  $i$ th central moment is defined as:

$$\mu_{ic}(x) = \int_0^\infty \left( t - \frac{\mu_1}{\mu_0} \right)^i c(x, t) dt \quad (3.71)$$

We derived the recursive expression for  $k$ th temporal moment from Eq.(3.2):

$$-k\mu_{k-1} + v \frac{\partial \mu_k}{\partial x} - D \frac{\partial^2 \mu_k}{\partial x^2} = q_{he} \sum_{l=0}^k \binom{k}{l} \mu_l \omega_{k-l} - q_{he} \mu_k \quad (3.72)$$

in which

$$\omega_k = \int_0^\infty t^k g(t) dt \quad (3.73)$$

We only need the first two (central) temporal moments, and the following expressions are used to get the initial guess for  $q_{he}$  and  $k$  in Eq.(3.31 & 3.32):

$$\mu_1 = (q_{he}\omega_1 + 1)\mu_0 \frac{x}{v} \quad (3.74)$$

$$\mu_{2c} = \left( \frac{2D}{\mu_0} \left( \frac{\partial \mu_1}{\partial x} \right)^2 + q_{he}\mu_0\omega_2 \right) \frac{x}{v} \quad (3.75)$$





## Chapter 4

# Modeling and Inverting Reactive Stream Tracers Undergoing Two-Site Sorption and Decay in the Hyporheic Zone

Liao, Z., D. Lemke, K. Osenbrück, and O. A. Cirpka (2013), Modeling and inverting reactive stream tracers undergoing two-site sorption and decay in the hyporheic zone, *Water Resour. Res.*, 49, 34063422, 10.1002/wrcr.20276.

### Abstract

Performing stream-tracer experiments is an accepted technique to assess transport characteristics of streams undergoing hyporheic exchange. Recently, combining conservative and reactive tracers, in which the latter presumably undergoes degradation exclusively within the hyporheic zone, has been suggested to study in-stream transport, hyporheic exchange, and the metabolic activity of the hyporheic zone. The combined quantitative analysis to adequately describe such tests, however, has been missing. In this paper, we present mathematical methods to jointly analyze breakthrough curves of a conservative tracer (fluorescein), a linearly degrading tracer (resazurin), and its daughter compound (resorufin), which are synchronously introduced into the stream as pulses. In-stream transport is described by the one-dimensional advection-dispersion equation, amended with a convolution term to account for transient storage within the hyporheic zone over a distribution of travel times, transformation of the reactive tracer in the hyporheic zone, and two-site sorption of the parent and daughter compounds therein. We use a shape-free approach of describing the hyporheic travel-time distribution, overcoming the difficulty of identifying the best functional parameterization for transient storage. We discuss how this model can be fitted to the breakthrough curves of all three compounds and demonstrate the method by an application to a tracer test in the third-order stream Goldersbach in Southern Germany. The entire river water passes once through the hyporheic zone over a travel distance of about 200 m with mean hyporheic residence times ranging between 16 and 23 minutes. We also observed a secondary peak in the transfer functions at about 1 hour indicating a second hyporheic flow path. We could jointly fit the breakthrough curves of all compounds in three monitoring stations, and evaluated the parameter uncertainty of the individual and joint fits by a method based on conditional realizations of the hyporheic travel-time distribution. The approach gives insight into in-stream transport, hyporheic exchange, metabolic activity, and river-bed sorption of the

stream under investigation.

## 4.1 Introduction

The transition zone between groundwater and surface water bodies, denoted hyporheic zone, is believed to be the key for evaluating the ecological status of the stream with respect to habitat formation, filtration of hazardous particles, biogeochemical turnover, and degradation of organic pollutants (e.g., *Bencala and Walters, 1983; Findlay, 1995; Brunke and Gonser, 1997; Boulton et al., 1998; Sophocleous, 2002; Hester and Gooseff, 2010; Robertson and Wood, 2010; Krause et al., 2011*). Hyporheic exchange can take place on multiple scales, such as ripples and dunes of the river bed, riffle-pool sequences, and river meanders, leading to hyporheic travel paths of multiple lengths and with multiple travel times.

In the past years, detailed modeling studies have shown that the extent of hyporheic flow cells, the strength of hyporheic exchange, and the distribution of hyporheic travel times can be predicted based on an exact morphological and hydraulic description of bedforms and bathymetry (*Gooseff et al., 2006; Cardenas and Wilson, 2007; Cardenas et al., 2008; Feng and Michaelides, 2009; Boano et al., 2010b*), river logs (*Sawyer et al., 2011*), fluvial islands (*Shope et al., 2012*), and meanders (*Cardenas, 2008a; Revelli et al., 2008; Boano et al., 2010a*, among others), to name only the most important features. The most common model used in hyporheic stream transport is the transient-storage model (*Bencala and Walters, 1983*), in which the hyporheic zone is conceptualized as an immobile water body exchanging solutes with the mobile stream. Solute transport within the stream is commonly described by the one-dimensional advection-dispersion equation. In the original version of the transient-storage model, a well mixed immobile water zone is assumed, and mass transfer between mobile and immobile water is proportional to the concentration difference between the two domains. This model has been extended by accounting for multiple immobile zones, each of which having a different size and undergoing first-order mass transfer with the stream according to a different mass-transfer coefficient, in so-called multi-rate mass-transfer models (*Haggerty and Gorelick, 1995; Carrera et al., 1998*) and in multiple-storage-zone models (*Marion et al., 2008*). Rather than discrete immobile zones, also distributions of mass-transfer rates may be used, in which the distributions are commonly described by functions with a few parameters (*Haggerty et al., 2000*). As another variant, *Deng and Jung (2009)* suggested to make the mass-transfer rate of the single-transient-storage model time dependent. Conceptually similar are certain variants of the continuous-time random-walk method to describe stream transport including a waiting-time distribution for solute particles in the immobile zone (*Boano et al., 2007*). All of these extensions have been developed to improve the description of tailing behavior observed in stream-tracer tests (*Drummond et al., 2012; Gooseff et al., 2003*).

Most of the transient-storage models listed above consider the size of the storage zone and the residence-time distribution within that zone, deriving the transport equation from mass conservation of the lumped system of the stream and the hyporheic zone. In the present study, we follow an approach similar to that of *Wörman et al. (2002)* that does not require the size of the transient storage zone and depends on a volumetric exchange rate between the stream and the hyporheic zone as well as on the age distribution of the returning hyporheic water instead. These two concepts are equivalent and can be transferred into each other (e.g. *Liao and Cirpka, 2011*). We denote the age distribution of the returning hyporheic as hyporheic travel-time distribution  $g(\tau)$ , with the hyporheic travel time  $\tau$ , and the discharge of exchanged water per volume of stream water as hyporheic exchange rate coefficient  $q_{he}$ .

Several parametric models have been suggested for  $g(\tau)$ . *Wörman et al. (2002)* listed: (1) the exponential distribution, which is equivalent to the single linear-storage model, (2) the Dirac delta distribution,

reflecting advective transport in a single hyporheic flow path, (3) the log-normal distribution, resembling advective-dispersive transport in a single hyporheic flow path, and the uniform distribution. Power-law functions have also frequently been applied (*Haggerty et al.*, 2002; *Cardenas et al.*, 2008; *Cardenas*, 2008b). Despite efforts to relate hyporheic travel-time distributions to geometric and other properties of single features (*Sawyer et al.*, 2011; *Cardenas et al.*, 2008, among others), it remains unclear which parameterization is the best in real applications affected by many features. To overcome this problem, *Liao and Cirpka* (2011) proposed a shape-free method to estimate the hyporheic travel-time distribution by jointly analyzing breakthrough curves of conservative and reactive tracers obtained from pulse-like tracer tests.

While stream-tracer tests with conservative tracers have been used in stream characterization for several decades, *Haggerty et al.* (2008) only recently proposed to use resazurin (Raz) as "smart" or reactive tracer that undergoes irreversible transformation to resorufin (Rru) under mildly reducing conditions as typically present in the microbiologically active hyporheic zone (see also *Haggerty et al.*, 2009). In previous field experiments, Raz was continuously injected into the stream. This has the advantage that transient effects on concentrations are excluded and metabolic activity can be quantified more accurately from the Raz:Rru ratio. Conversely, it prevents using the difference between the conservative- and reactive-tracer breakthrough curves for a better quantification of the strength of hyporheic exchange and the hyporheic travel-time distribution as suggested by *Liao and Cirpka* (2011). Experiments with joint pulse-like injection of the conservative tracer fluorescein (Fluo) and Raz into a stream are discussed in a companion paper (*Lemke et al.*, 2013a) together with a simplified analysis of the observed breakthrough curves. A much closer look into the mathematical analysis of combined tracer breakthrough curves will be provided in the present paper.

Our previous attempts to model transport of Raz in streams with a metabolically active hyporheic zone had two major shortcomings. *Liao and Cirpka* (2011) assumed that the reactive tracers do not sorb, whereas *Lemke et al.* (2013a) assumed linear sorption at equilibrium. Already *Haggerty et al.* (2008) reported non-equilibrium sorption in column experiments which is in agreement with our own column studies that we could simulate well by a two-site sorption model (data not shown here). For the rather short hyporheic residence reported by *Lemke et al.* (2013a), it should be clear that kinetic sorption can have an effect on the observed breakthrough curves. The general impact of sorption within the hyporheic zone on stream transport has already been discussed by *Jonsson et al.* (2004). A second shortcoming of both *Liao and Cirpka* (2011) and *Lemke et al.* (2013a) is that these studies do not consider the breakthrough curves of the degradation product Rru even though the latter can be measured more easily than Raz by fluorometric methods.

The main objective of the present paper is to extend the shape-free method proposed by *Liao and Cirpka* (2011) for the determination of hyporheic travel-time distributions from joint conservative and reactive stream tracer tests with pulse-like injection. The extensions include (1) the consideration of two-site sorption of the reactive compounds within the hyporheic zone to be compliant with findings from column tests, and (2) the analysis of the breakthrough curve of the reactive daughter compound together with the parent compound and the conservative tracer. The focus of the present study is on the derivation of the predictive models, because the inversion methodology does not significantly differ from that of *Liao and Cirpka* (2011). Finally, we apply the method to the field data presented by *Lemke et al.* (2013a). The proposed method yields: (1) parameters of in-stream transport, namely the in-stream velocity and dispersion coefficient, (2) the characterization of hyporheic exchange by the hyporheic exchange rate and the hyporheic travel-time distribution, (3) equilibrium and kinetic sorption parameters of the hyporheic zone, and (4) a measure of metabolic activity within the hyporheic zone in the form of a first-order rate coefficient of Raz-to-Rru transformation.

## 4.2 Model Formulation

### 4.2.1 General Setup Considered

The main objective of this work is to estimate transport parameters of streams from reactive tracer experiments: the velocity and in-stream dispersion coefficient, the rate of hyporheic exchange, and the distribution of times over which water parcels stay within the hyporheic zone. While breakthrough curves of tracers in observation wells directly yield the travel-time distribution of the tracer in the system by deconvoluting the input and output signals (e.g. *Cirpka et al., 2007*), tracer breakthrough curves in streams depend only indirectly on the travel-time distribution in the hyporheic zone, because no measurements are taken in the hyporheic zone or in the river bed itself. This leads to a coupled system of in-stream transport plus a convolution term that quantifies how the fraction of tracer, which has been transported into the hyporheic zone by infiltration, returns back into the stream.

The setup of the stream-tracer application considered is shown in Figure 4.1. The conservative tracer (Flou) and the reactive tracer (Raz) are jointly introduced by pulse injections into the stream. Within the hyporheic zone, the reactive tracer partially undergoes reduction to its reaction product (Rru). The breakthrough curves of all three tracers are measured far downstream, in order to quantify hyporheic processes on the scale of a river reach, at which the effects of individual features are averaged.

Within the model, solute transport in the stream is considered to be advective-dispersive, where a fraction of river water infiltrates into the hyporheic zone and comes back at the same location over a distribution of travel times. Both in-stream dispersion and hyporheic exchange lead to spreading of the conservative-tracer breakthrough curve, making the distinction between these two processes difficult when only conservative tracer tests are analyzed. The reactive parent compound, however, is converted to the daughter compound upon exposure to hyporheic-zone conditions. *Liao and Cirpka (2011)* claimed that jointly analyzing conservative and reactive tracer breakthrough curves improves our ability to distinguish between spreading by in-stream dispersion and hyporheic exchange. However, in this study we also consider that the reactive tracers may undergo two-site sorption within the hyporheic zone. As a consequence, the reactive tracer signals are affected by the metabolic activity and sorption properties of the stream bed. Parameters describing these processes need to be inferred together with the physical transport properties in the analysis of the breakthrough curves.

### 4.2.2 Governing Equations

The one-dimensional advection-dispersion equation (ADE) is commonly used to describe solute transport in streams. As proposed by *Liao and Cirpka (2011)*, the effect of hyporheic exchange on conservative solute transport is quantified by an exchange rate and a hyporheic travel-time distribution  $g(\tau)$ . The governing equation for a conservative compound reads as:

$$\frac{\partial c}{\partial t} + v \frac{\partial c}{\partial x} - D \frac{\partial^2 c}{\partial x^2} = q_{he} \left( \int_0^t g(\tau) c(t - \tau) d\tau - c(t) \right) \quad (4.1)$$

in which  $c [ML^{-3}]$  is the tracer concentration in the river;  $t [T]$  is the time since injection of the tracer into the stream;  $v [LT^{-1}]$  is the flow velocity of the stream;  $D [L^2T^{-1}]$  is the in-stream dispersion coefficient;  $q_{he} [T^{-1}]$  is the hyporheic exchange rate coefficient;  $g(\tau) [T^{-1}]$  is the distribution over which water parcels stay within the hyporheic zone, and  $\tau [T]$  is the travel time within the hyporheic zone. Eq. (4.1) is a nonlocal-in-time, linear partial differential equation of second-order. The corresponding initial and boundary conditions are for a point injection into the flow and a semi-infinite domain:

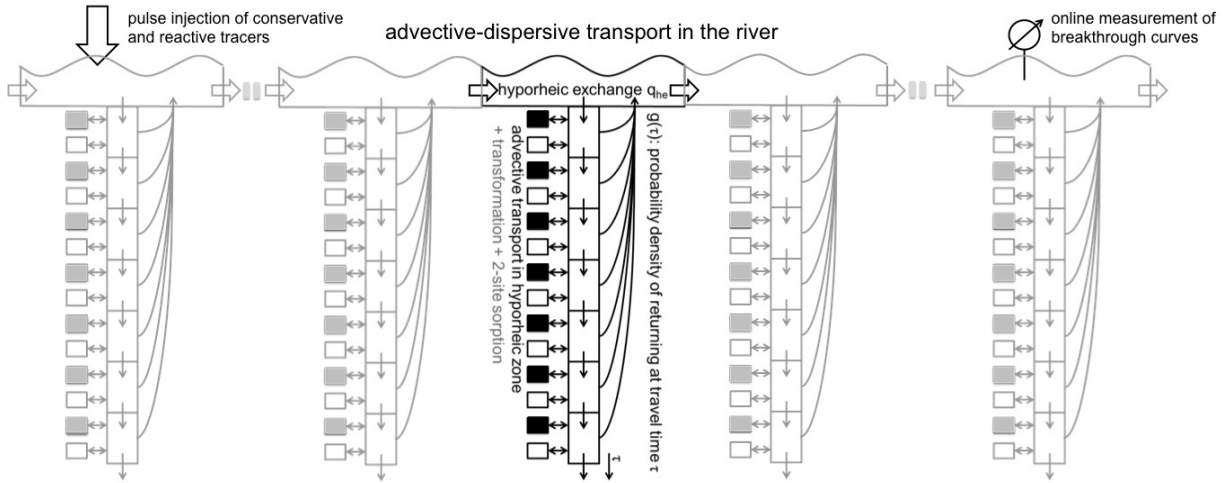


Figure 4.1: General setup of the model. The conservative tracer and reactive parent compound are injected as a joint pulse into the stream. No reaction takes place in the stream for all tracers, while in the hyporheic zone the reactive parent compound undergoes first-order transformation to the reaction product and both reactive tracers are subject to two-site sorption, i.e., equilibrium and kinetic sorption. the latter is shown as two different small boxes for each cell within the hyporheic zone. The strength of hyporheic exchange is quantified by  $q_{he}$  as discharge of exchanged water per water volume in the stream. Transport within the hyporheic zone is described as advection in travel time rather space. The hyporheic travel-time distribution  $g(\tau)$  determines the age distribution of the water returning into the stream. Breakthrough curves of all three tracers are measured downstream.

$$c(t, x) = 0, \quad \forall x, t = 0 \quad (4.2)$$

$$vc - D \frac{\partial c}{\partial x} = \frac{m}{A} \delta(t), \quad x = 0 \quad (4.3)$$

$$c(t, x) = 0, \quad x \rightarrow \infty, \forall t \quad (4.4)$$

in which  $m [M]$  is the total mass of the tracer injected in the river;  $A [L^2]$  is the cross-sectional area of the river channel; and  $\delta(\cdot)$  is the Dirac delta function with units inverse to those of the argument.

We have introduced  $g(\tau)$  as a probability density function of times over which water parcels stay within the hyporheic zone. By adopting the concept of signal processing and other fields, it may also be considered as a more generalized transfer function for any type of linear transport. A transfer function describes the response in the system output to a unit pulse input, in which the hyporheic zone is the system of interest in the current application. In this sense,  $g(\tau)$  quantifies the concentration of a compound measured in the water returning from the hyporheic zone to the stream at time  $\tau$  when the compound was introduced into the hyporheic zone as a unit pulse at time 0. The generalized hyporheic transfer function can differ from the travel-time distribution due to linear processes such as linear sorption, both kinetic and at local equilibrium, and first-order transformations. The convolution integral of Eq. (4.1), however, would not be valid if nonlinear reactions or nonlinear mass-transfer processes had to be considered (Dennery and Krzywicki, 1967, chapter 7).

Since three tracers are actually measured and analyzed by the model, we need three governing equations, where those for the conservative tracer and the reactive parent compound look identical even though the corresponding hyporheic transfer functions differ, whereas the governing equation of the reaction product differs because it depends on the concentration of the parent compound:

$$\begin{aligned}\frac{\partial c_i}{\partial t} + v \frac{\partial c_i}{\partial x} - D \frac{\partial^2 c_i}{\partial x^2} &= q_{he} \left( \int_0^t g_i(\tau) c_i(t - \tau) d\tau - c_i \right), \quad i = 0, 1 \\ \frac{\partial c_2}{\partial t} + v \frac{\partial c_2}{\partial x} - D \frac{\partial^2 c_2}{\partial x^2} &= q_{he} \left( \int_0^t (g_{12}(\tau) c_1(t - \tau) + g_2(\tau) c_2(t - \tau)) d\tau - c_2 \right)\end{aligned}\quad (4.5)$$

subject to:

$$\begin{aligned}vc_i - D \frac{\partial c_i}{\partial x} \Big|_{x=0} &= \frac{m_i}{A} \delta(t), \quad i = 0, 1 \\ vc_2 - D \frac{\partial c_2}{\partial x} \Big|_{x=0} &= 0\end{aligned}\quad (4.6)$$

in which the subscript 0 is for the conservative tracer (Flou), 1 for the reactive tracer (Raz), and 2 for the product (Rru). The governing equation for the reaction product contains two terms in the convolution integral, because one part of the tracer originates from the transformation of the reactive tracer within the hyporheic zone, whereas the other part has already been introduced into the hyporheic zone by river-water infiltration. Therefore, we use  $g_0(\tau)$  to denote the hyporheic travel time distribution, which is the hyporheic transfer function of the conservative tracer,  $g_1(\tau)$  describes the hyporheic transfer function of the reactive parent compound that is not converted in the hyporheic zone,  $g_2(\tau)$  denotes the hyporheic transfer function of the reaction product, already entering the hyporheic zone as reaction product, whereas  $g_{12}(\tau)$  quantifies the response of the daughter compound concentration in the returning water caused by a unit pulse of the parent compound in the water entering the hyporheic zone. Note that the three transfer functions include the effects of all processes within the hyporheic zone, namely physical transport, the linear transformation and two-site sorption of the reactive tracers. In Appendix 4.A, we derive analytical expressions for tracer concentrations  $\tilde{c}_i(s, x)$ ,  $i = 0, 1, 2$  in the Laplace domain, requiring the Laplace transformed hyporheic transfer function  $\tilde{g}_i(s)$ ,  $i = 0, 1, 2, 12$  with  $s [T^{-1}]$  being the complex Laplace coordinate (see Eq. (4.21)).

### 4.2.3 Hyporheic Transfer Functions for Conservative and Reactive Tracers

The hyporheic transfer functions for the tracers are denoted by  $g_i(\tau)$ ,  $i = 0, 1, 2, 12$ , where  $\tau$  is the travel time in the hyporheic zone starting from the time point when the water enters the river bed. All reactive processes of the solutes in the hyporheic zone, namely microbial transformations and sorption, modify  $g_i(\tau)$ . Thus while the conservative hyporheic transfer function  $g_0(\tau)$  is a travel time distribution reflecting the geometry and conductivity of the hyporheic zone and the pressure distribution at the stream bed, the reactive hyporheic transfer functions  $g_1(\tau)$  of the parent compound Raz,  $g_2(\tau)$  of the daughter compound Rru, and the cross-compound transfer function  $g_{12}(\tau)$  differ in shape. In the fitting procedure, we want to estimate  $g_0(\tau)$  and the reactive parameters modifying  $g_0(\tau)$  to  $g_1(\tau)$ ,  $g_2(\tau)$ , and  $g_{12}(\tau)$ . In this paper, rather than assuming a specific functional shape of  $g_0(\tau)$ , such as the exponential or log-normal distributions, we identify shape-free functions, which may exhibit non-classical features such as broad peaks, multiple peaks, or long tails.

In the following derivations, it is advantageous to express transport within the hyporheic zone as function of travel time rather than spatial coordinates, thus avoiding the explicit consideration of flow paths, velocities, and dispersion coefficients within the hyporheic zone. The latter approach is valid in linear transport with uniform reaction coefficients, in which the order of mixing water parcels of different age and linear reactions does not influence the model outcome.

### General Relationship between Hyporheic Travel-Time Distributions and Hyporheic Transfer Functions

The general approach for calculating the transfer function of a tracer in the system for a given travel-time distribution  $g(\tau)$  of the returning water is expressed by:

$$g_i(\tau) = \int_0^{\infty} g(\tau^*) c_i^{hz}(\tau^*, \tau) d\tau^*, \quad i = 0, 1, 2, 12 \quad (4.7)$$

in which  $g_i(\tau)$  is the corresponding hyporheic transfer function with indices discussed above;  $g(\tau^*)$  is the hyporheic travel-time distribution, indicating how the pulse of an ideal tracer, which does not undergo transformation or inter-phase mass-transfer processes, returns back into the stream, in which  $\tau^*$  is the travel time for the ideal tracer (substituting the spatial coordinate);  $c_i^{hz}$  is the tracer concentration of compound  $i$  in the hyporheic zone at time  $\tau$  and travel time  $\tau^*$ .

The concept of computing  $g_i(\tau)$  may be best understood by considering a pulse injection of compound  $i$  into the hyporheic zone at time  $\tau = 0$ , as illustrated in Figure 4.2. As discussed below, we consider advection in travel-time coordinates, two-site sorption, and linear decay of the reactive parent compound. The solid lines of Figure 4.2A show the corresponding length profiles of the parent compound  $c_i^{hz}$  in the travel-time domain (coordinate  $\tau^*$ ) for different times  $\tau$  as solid lines. With increasing time, the color changes from black via red to yellow. If advection was the only transport process, the length profile would be a moving Dirac pulse. At early times, such a pulse can be observed. At later times, however, kinetic sorption leads to a hump following the pulse; and at very late times the profile becomes Gaussian with respect to  $\tau^*$ . The dotted lines in Figure 4.2A show the corresponding length profiles of the reaction product with matching colors indicating identical time points. With increasing time, the reaction product starts to dominate over the parent compound, but it vanishes again due to further degradation of the reaction product. As shown in Section 4.2.3, we account for an unspecified degradation of the parent compound, the transformation from parent compound to the reaction product and a further degradation of the reaction product, requiring three decay coefficients in our model.

The black dashed lines in Figure 4.2A and B mark the hyporheic travel-time distribution  $g(\tau^*)$  of the water returning into the stream. As expressed in Eq. (4.7), the concentration of the compounds returning into the stream at any time  $\tau$  can be computed by considering the full profile  $c_i^{hz}(\tau^*)$  at that time, weighting it with the probability density  $g(\tau^*)$  that the returning water has the age  $\tau^*$ , and integrating over all possible values of  $\tau^*$ . The resulting responses upon a pulse injection (that is, the transfer functions) are illustrated in Figure 4.2B. For an ideal tracer, the transfer function  $g_0(\tau)$  equals the hyporheic travel-time distribution  $g(\tau)$ . The transfer function of the reactive parent compound  $g_1(\tau)$  rapidly drops at early times due to mass-transfer from the aqueous to the sorbing phase within the hyporheic zone. This mass is later released, leading to tailing of the transfer function  $g_1(\tau)$  in comparison to the hyporheic travel-time distribution  $g(\tau)$ . The transfer function  $g_{12}(\tau)$  of the reaction product due to a pulse injection of the parent compound indicates to which extent the released parent compound is transformed to the daughter compound. In the profiles shown in Figure 4.2A the concentration of the daughter compound significantly exceeds the concentration of the parent compound at late times. However, this happens in a

range of travel-time values  $\tau^*$  that is hardly sampled by  $g(\tau^*)$ , so that the corresponding cross-compound transfer function  $g_{12}(\tau)$  even decreases at the decisive late times. The hyporheic transfer function  $g_2(\tau)$  of the daughter compound in response to an injection already as daughter compounds, not shown in Figure 4.2B, is similar to that of  $g_1(\tau)$  because the same process types, namely linear decay and two-site sorption, are considered.

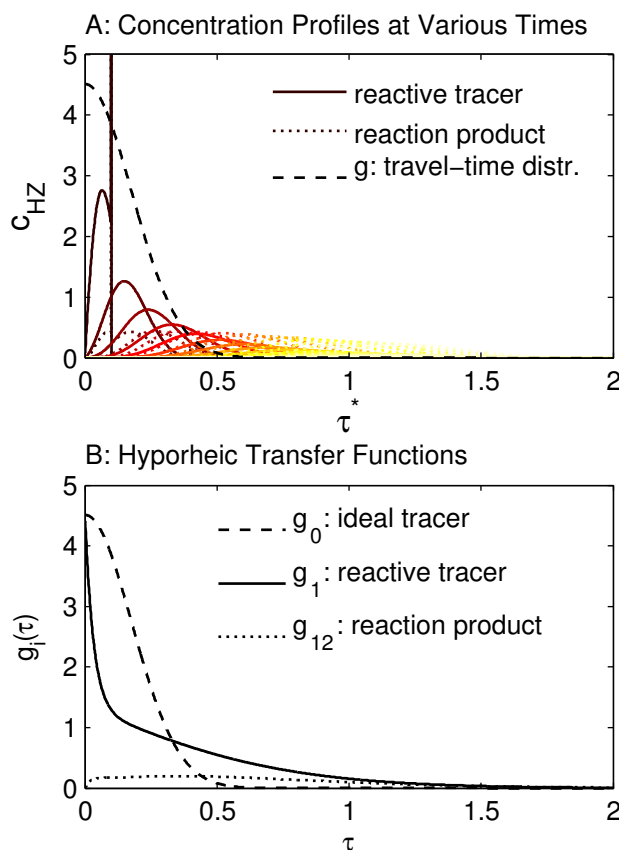


Figure 4.2: Principle of calculating transfer functions for different tracers from the hyporheic travel-time distribution  $g(\tau)$ . A: length profiles of reactive-tracer concentrations with respect to travel time  $\tau^*$  for different times  $\tau$ , indicated by changing colors (lighter colors mark later times). Solid lines: reactive parent compound; dotted lines: reactive daughter compound; black dashed line: hyporheic travel-time distribution  $g(\tau^*)$ . B: corresponding transfer functions, reflecting the sampling of concentrations in A according to  $g(\tau^*)$ .

### Considering Decay and Two-Site Sorption

The simplest case for the evaluation of the hyporheic transfer function is that of the conservative tracer, where the tracer concentration within the hyporheic zone for a pulse injection remains a pulse traveling with  $d\tau^*/d\tau = 1$ , resulting in:

$$g_0(\tau) = \int_0^\infty g(\tau^*)\delta(\tau - \tau^*)d\tau^* = g(\tau), \quad \forall \tau \quad (4.8)$$



in which the delta function  $\delta(\tau - \tau^*)$  represents the conservative-tracer concentration within the hyporheic zone for a pulse injection.

For the Raz-Rru system, the expressions of  $c_i^{hz}$ ,  $i = 1, 2$  are somewhat more complicated. Here, we assume the following processes within the hyporheic zone: (1) linear decay of Raz to Rru, quantified by a first-order rate coefficient  $\lambda_{12}$ ; (2) linear decay of Raz to unidentified reaction products, quantified by  $\lambda_1$ ; (3) linear decay of Rru to unidentified reaction products, quantified by  $\lambda_2$ ; (4) linear equilibrium sorption of both Raz and Rru, quantified by corresponding retardation factors  $R_1$  and  $R_2$ ; (5) linear kinetic sorption of both Raz and Rru, quantified by dimensionless distribution coefficients  $K_{kin,1}$  and  $K_{kin,2}$  and rate coefficients of mass transfer  $k_{mt,1}$  and  $k_{mt,2}$ , respectively. In general, hyporheic transport is expressed in travel-time rather than spatial coordinates.

This model differs slightly from that proposed by *Haggerty et al.* (2009), who considered only equilibrium sorption, even though the same authors have applied multi-rate mass-transfer models for sorption elsewhere (*Haggerty and Gorelick*, 1995). Since hyporheic travel times can range from seconds to days, it is not reasonable to assume that sorption is always at equilibrium. Also, column experiments, not shown here, indicate the need to account at least for two sorption sites. A second difference to *Haggerty et al.* (2009) is that in our model the transport in the travel-time domain is restricted to advection. Effects of dispersion are included in the travel-time distribution  $g(\tau^*)$  appearing in Eq. (4.7): More dispersive transport would lead to a wider distribution  $g(\tau^*)$ .

The governing equation for the transport of the parent compound (Raz) within the hyporheic zone is:

$$\begin{aligned} R_1 \frac{\partial c_1^{hz}}{\partial \tau} + K_{kin,1} \frac{\partial c_1^*}{\partial \tau} + \frac{\partial c_1^{hz}}{\partial \tau^*} &= -(\lambda_1 + \lambda_{12})c_1^{hz} \\ \frac{\partial c_1^*}{\partial \tau} &= k_{mt,1}(c_1^{hz} - c_1^*) \\ c_1^{hz}(\tau, 0) &= b_1 \delta(\tau) \end{aligned} \quad (4.9)$$

and that for the daughter compound (Rru):

$$\begin{aligned} R_2 \frac{\partial c_2^{hz}}{\partial \tau} + K_{kin,2} \frac{\partial c_2^*}{\partial \tau} + \frac{\partial c_2^{hz}}{\partial \tau^*} &= \lambda_{12}c_1^{hz} - \lambda_2c_2^{hz} \\ \frac{\partial c_2^*}{\partial \tau} &= k_{mt,2}(c_2^{hz} - c_2^*) \\ c_2^{hz}(\tau, 0) &= b_2 \delta(\tau) \end{aligned} \quad (4.10)$$

in which  $c_i^{hz}[ML^{-3}]$ ,  $i = 1, 2$  are the concentrations of the reactive parent and daughter compounds Raz and Rru in the hyporheic zone, respectively;  $c_i^*[ML^{-3}]$ ,  $i = 1, 2$  are concentrations at the kinetically sorbing sites, and  $K_{kin,i}$ ,  $i = 1, 2$  are dimensionless distribution coefficients. The coefficients  $b_1$  and  $b_2$  represent how much mass of the parent and daughter compounds are introduced into the hyporheic zone at time  $\tau = 0$ . For the evaluation of transfer functions,  $b_{1,2}$  is dimensionless and can either be zero or one. Appendix 4.B contains a derivation of the equations, and Appendix 4.C the analytical solutions in the Laplace domain.

Given the tracer concentrations in the hyporheic zone, the hyporheic transfer function for the reactive parent and daughter compounds can be calculated according to Eq. (4.7) as follows:

$$\begin{aligned}
g_1(\tau) &= \int_0^\infty g(\tau^*) c_1^{hz}(\tau, \tau^*) d\tau^* \\
g_{12}(\tau) &= \int_0^\infty g(\tau^*) c_{12}^{hz}(\tau, \tau^*) d\tau^* \\
g_2(\tau) &= \int_0^\infty g(\tau^*) c_{22}^{hz}(\tau, \tau^*) d\tau^*
\end{aligned} \tag{4.11}$$

in which  $c_1^{hz}(\tau, \tau^*)$  is the concentration of Raz caused by injection of Raz in the system for  $b_1 = 1$  in Eq. (4.9);  $c_{12}^{hz}(\tau, \tau^*)$  is the concentration of Rru caused by transformation from Raz at conservative travel time  $\tau^*$  and time  $\tau$ , stemming from the input of  $c_1$  (that is, for  $b_1 = 1$  and  $b_2 = 0$  in Eq. (4.9)), and  $c_{22}^{hz}(\tau, \tau^*)$  is the concentration of Rru caused by injection of the same compound into the domain, stemming from the input of  $c_2$  (that is, for  $b_1 = 0$  and  $b_2 = 1$  in Eq. (4.9)).

The weighting average with respect to  $\tau^*$  expressed in Eq. (4.11) also holds for the Laplace-transformed hyporheic transfer functions. Here the hyporheic concentrations  $c_i^{hz}(\tau, \tau^*)$  are replaced by their Laplace transforms  $\tilde{c}_i^{hz}(s, \tau^*)$  given in Appendix 4.C:

$$\begin{aligned}
\tilde{g}_0(s) &= \int_0^\infty g_0(\tau^*) \exp(-s\tau^*) d\tau^* \\
\tilde{g}_1(s) &= \int_0^\infty g_0(\tau^*) \exp(-\beta_1 \tau^*) d\tau^* \\
\tilde{g}_{12}(s) &= \int_0^\infty g_0(\tau^*) \frac{\lambda_{12}}{\beta_1 - \beta_2} (\exp(-\beta_2 \tau^*) - \exp(-\beta_1 \tau^*)) d\tau^* \\
\tilde{g}_2(s) &= \int_0^\infty g_0(\tau^*) \exp(-\beta_2 \tau^*) d\tau^*
\end{aligned} \tag{4.12}$$

in which we have made use of  $g_0(\tau^*) = g(\tau^*)$  according to Eq. (4.8). The  $\beta_{1,2}$  coefficients are derived in Appendix 4.C. The integrals appearing in Eq. (4.12) are comparably easy to compute for a simple parametric hyporheic travel-time distribution  $g_0(\tau)$ , such as the exponential function. In the shape-free approach,  $g_0(\tau)$  is discretized as a piecewise linear function. The corresponding analytical expressions of the integrals in Eq. (4.12) are listed in Appendix 4.D.

#### 4.2.4 Joint Estimation of the Shape-Free Hyporheic Travel-Time Distribution and Reactive Parameters

The parameter-estimation scheme is essentially identical to the method of *Cirpka et al.* (2007), even though more data are included and more parameters are estimated. In the shape-free approach, the hyporheic travel-time distribution  $g_0(\tau)$  is estimated at a set of discrete travel times  $\tau_i$  with regular travel-time increment  $\Delta\tau$ , chosen to be identical to the time increment  $\Delta t$  of the tracer measurements. The largest travel time  $\tau$  must not be larger than the duration of the experiment. Therefore, the total number of parameters is in the same order as the number of measurement (hundreds to thousands), requiring regularization. Adopting the method of *Cirpka et al.* (2007), we assume that the hyporheic travel-time distribution  $g_0(\tau)$  is stochastic, auto-correlated travel-time variable described by a second-order intrinsic model with a linear semivariogram:

$$\gamma_{g_0}(h) = E \left[ \frac{1}{2} (g_0(\tau + h) - g_0(\tau))^2 \right] = \theta |h| \tag{4.13}$$

in which  $h$  is the travel-time difference between two values in the vector  $g_0(\tau)$ . The slope  $\theta [T^{-3}]$  of the semivariogram quantifies the smoothness of  $g_0(\tau)$ , where small values of  $\theta$  enforce particularly smooth behavior. Although more complex semivariogram models could be used here, we have chosen the linear

model because it has only one structural parameter, and the lack of information about  $g_0(\tau)$  results in the simplest form of the semivariogram model. The chosen approach is identical to first-order Tikhonov regularization, in which solutions of  $g_0(\tau)$  with large derivatives  $\partial g_0/\partial \tau$  are penalized (Kitanidis, 1997). With the regularization we avoid reproducing noise in the data when estimating  $g_0(\tau)$ . The model is rather sensitive to the value of  $\theta$ , which therefore needs to be chosen carefully in the estimation procedure. Because  $g_0(\tau)$  is a probability density function, we must ensure that no element  $g_0(\tau_i)$  is negative.

Besides  $g_0(\tau)$ , we estimate the following parameters:  $v$ ,  $D$ ,  $q_{he}$ ,  $\lambda_i$ , ( $i = 1, 2, 12$ ),  $R_i$ , ( $i = 1, 2$ ),  $k_{mt,i}$ , ( $i = 1, 2$ ) and  $K_{kin,i}$  ( $i = 1, 2$ ), which are used to characterize stream transport and to compute the hyporheic transfer functions  $g_1(\tau)$ ,  $g_2(\tau)$  and  $g_{12}(\tau)$ . By applying Bayes' theorem and assuming multi-Gaussian prior probability density functions, we obtain the objective function:

$$W(\mathbf{p}) = (\mathbf{c}(\mathbf{p}) - \mathbf{c}_m)^T \mathbf{C}^{-1} (\mathbf{c}(\mathbf{p}) - \mathbf{c}_m) + \mathbf{g}_0^T \mathbf{\Gamma}_{gg}^{-1} \mathbf{g}_0 + \boldsymbol{\nu}^T \mathbf{H} \mathbf{g}_0 \quad (4.14)$$

in which  $\mathbf{c}_m$  is the vector of measured tracer concentrations in the stream;  $\mathbf{c}(\mathbf{p})$  is the simulated tracer concentration based on the set of parameters  $\mathbf{p} := (\mathbf{p}'_{para}, \mathbf{g}'(\tau))'$ , where we denote  $\mathbf{p}_{para} := (\log(v), \log(D), \log(q_{he}), \log(\lambda), \log(R-1), \log(k_{mt}), \log(K_{kin}))'$  for simplicity;  $\mathbf{C}$  is the covariance matrix expressing uncertainty of the measurements;  $\mathbf{\Gamma}_{gg}$  is a  $n_\tau \times n_\tau$  matrix, describing the discrete semivariogram values for all pairs of elements in  $g_0(\tau)$ ;  $\boldsymbol{\nu}$  is the set of Lagrange multipliers for the nonnegativity constraints on  $g_0(\tau)$ ;  $\mathbf{H}$  is a  $n_\nu \times n_\tau$  selection matrix, which changes in every iteration step according to the number of constraints  $n_\nu$  and the position at which the nonnegativity constraints are activated. The model simulation is linearized about the last estimate  $\mathbf{p}_{last}$ :  $\mathbf{c}(\mathbf{p}) \approx \mathbf{J}(\mathbf{p} - \mathbf{p}_{last})$ , where  $\mathbf{J}$  is the Jacobian matrix. Minimization of the objective function is achieved by setting the derivatives with respect to  $\mathbf{p}$  and  $\boldsymbol{\nu}$  to zero, resulting in the following system of linear equations:

$$\begin{bmatrix} \mathbf{J}^T \mathbf{C}^{-1} \mathbf{J} - \mathbf{\Gamma}_{pp}^{-1} & \mathbf{H}^T \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \boldsymbol{\nu} \end{bmatrix} = \begin{bmatrix} \mathbf{J}^T \mathbf{C}^{-1} (\mathbf{c}_m + \mathbf{J} \mathbf{p}_{last} - \mathbf{c}(\mathbf{p}_{last})) \\ \mathbf{0} \end{bmatrix} \quad (4.15)$$

in which  $\mathbf{\Gamma}_{pp}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}_{gg}^{-1} \end{bmatrix}$ . This system of equations is solved iteratively with an inner iteration to identify the correct set of Lagrange multipliers and an outer iteration for the successive linearization about the last estimate. The rules to activate/deactivate the nonnegativity constraints can be found in Cirpka et al. (2007). The scheme is stabilized by a line search in the Gauss-Newton direction.

## 4.3 Application to Field Data

### 4.3.1 Experimental Setup

A tracer test with the conservative and reactive fluorescent tracers Fluo and Raz was performed at River Goldersbach in southern Germany in July 2012. In the following, we give a brief overview of the field experiment. Details on the field site, the experimental setup, and the separation and correction of the tracer signals are given by Lemke et al. (2013a). River Goldersbach is part of the Neckar catchment within the Rhine basin. The section of River Goldersbach used for the tracer is located about 30 km south of Stuttgart. The river flows along a highly meandering course through a narrow floodplain with a morphology characterized by riffle-pool sequences with abundant point bars. The river bed material is mainly composed of coarse sand and gravels.

A premixed solution of  $0.87 \times 10^{-2}$  moles of Fluo and  $1.13 \times 10^{-1}$  moles of Raz was injected into the river as a pulse, distributed uniformly over the river cross section. The injected tracers as well as the

reaction product Rru were measured every 10 s using on-line field fluorometers (Lemke *et al.*, 2013b). Three monitoring stations were chosen to place the fluorometers at distances of 830 m, 1075 m, and 1210 m downstream the injection point, respectively. The experiment was started after sunset and completed before sunrise to avoid photodegradation of the tracers (Haggerty *et al.*, 2009). During the entire tracer test, the pH of the river was fairly stable around a value of 8. At this pH, the conservative tracer Fluo can be treated as an ideal tracer (Smith and Pretorius, 2002). The discharge during the experimental period was also stable, which is around 60 s/L.

Measured breakthrough curves, are shown in the left column of Figure 4.3. The injected pulses were spread upon transport in the stream, and reduction of Raz lead to the formation of Rru, even though this remained a small fraction. The recovery of the conservative tracer Fluo was high ( $> 94\%$ ) for all three monitoring stations. The recovered number of moles of Raz and Rru at each monitoring station did not add up to the number of moles of Raz injected, which were attributed to the decay of Raz and Rru to unknown (non-fluorescent) compounds Lemke *et al.* (2013a).

### 4.3.2 Results of Data Analysis

We analyze the breakthrough curves obtained at the three monitoring stations and estimate transport as well as reactive parameters along with the hyporheic travel-time distribution  $g(\tau)$  using the shape-free approach described in Section 4.2.2. The same data set is analyzed by Lemke *et al.* (2013a) who use an exponential function as the hyporheic travel-time distribution, consider only equilibrium sorption within the hyporheic zone, and do not analyze the information contained in the breakthrough curves of the reaction product (Rru).

The shape-free method is applied to jointly fit the breakthrough curves of the conservative tracer (Fluo), reactive parent compound (Raz), and reactive product (Rru), first at each individual monitoring station (Figure 4.3), followed by a joint fit of the breakthrough curves for all compounds at all monitoring stations (Figure 4.4). The left column of Figure 4.3 shows the individually fitted, semi-logarithmic breakthrough curves for the time range between 1 and 12 hours after tracer injection. The measured molar concentrations are shown as dots, whereas the colored lines are the simulated breakthrough curves. The three solid horizontal lines in the plots indicate the quantification limits of the fluorometers. Only measured concentrations above this line are considered reliable (Lemke *et al.*, 2013b) and are used for parameter estimation. From top to bottom, the quantification limits are  $1 \mu\text{g}/\ell$  for Raz,  $0.5 \mu\text{g}/\ell$  for Rru, and  $0.1 \mu\text{g}/\ell$  for Fluo. Because much more molar mass of Raz than that of Fluo was injected, the peak concentrations of Fluo are the smallest, but the relative accuracy of the measurements is the highest. This can be explained by the high quantum yield of Fluo in comparison to Rru, and even more so Raz, which is also indicated by the different quantification limits for the three compounds. From the breakthrough curves, we can see that the concentration of Fluo reaches its peak the earliest, Raz is somewhat retarded, which can be explained by sorption within the hyporheic zone, and Rru peaks the latest. The latter is expected since Rru needs time to form.

The different tailing behavior of the three compounds is captured very well by the fit within the quantification limit. While Fluo shows an exponential tailing behavior for about two hours, the concentration drops more slowly at late times. Such a behavior cannot be captured with a model that uses an exponential hyporheic travel-time distribution, like the one used by Lemke *et al.* (2013a). The shape-free approach, by contrast, leads to near-to-perfect agreement in the tailing of the conservative tracer even at very late times. This is important because the simulated breakthrough curves of the reactive compounds directly depend on the hyporheic travel-time distribution. If we presented a good model fit of the reactive tracers together with a bad one of the conservative tracer, bias in the estimated reactive tracers was needed

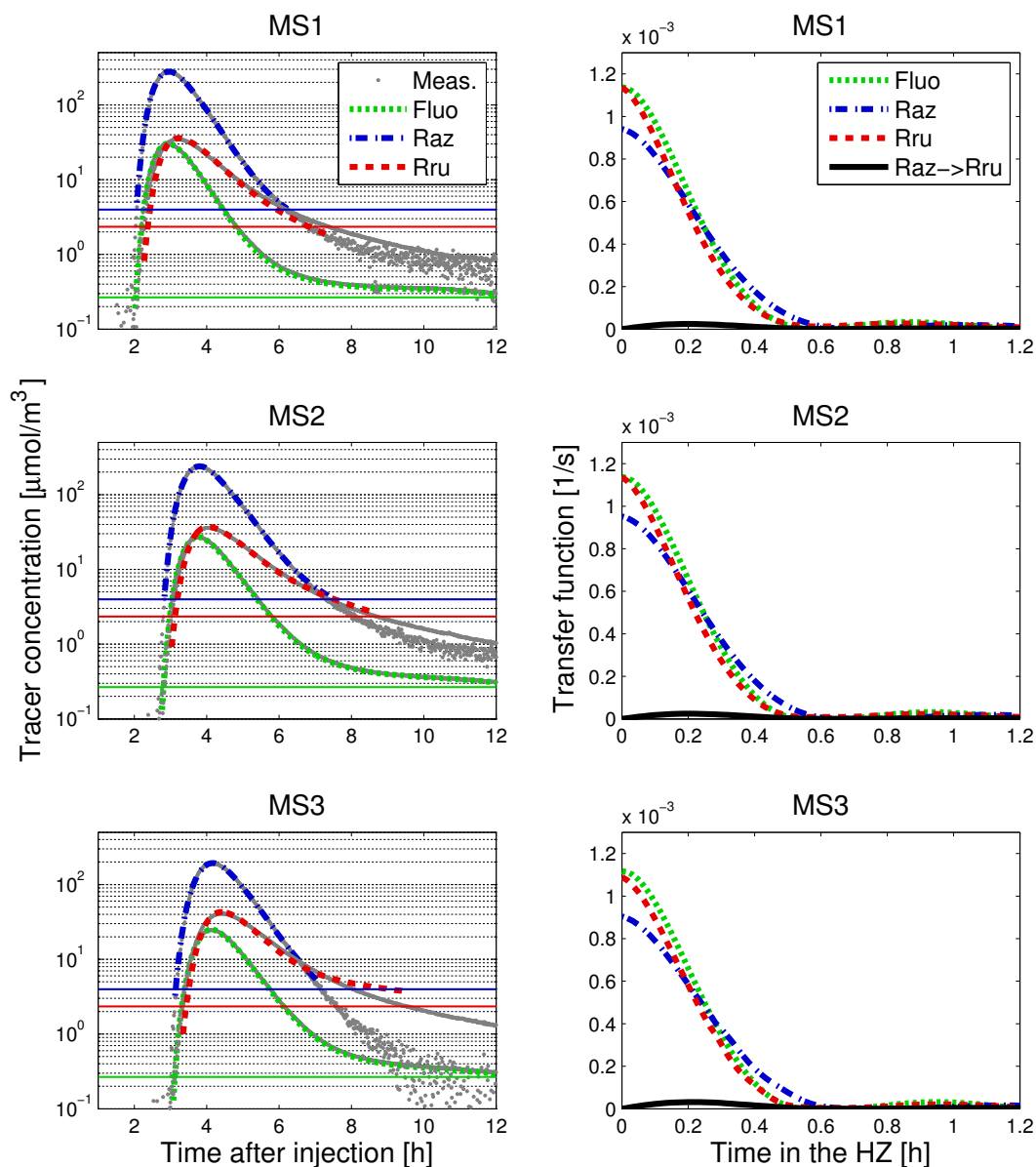


Figure 4.3: Independent fits of breakthrough curves for each monitoring station (MS). Left column: measured and fitted breakthrough curves obtained at three monitoring stations (MS) which are located 830m, 1075m and 1210m downstream of the injection point. Green lines: Fluo (conservative tracer), blue lines: Raz (reactive parent compound), red lines: Rru (reaction product). Horizontal lines: quantification limits of the measurement device for each tracer. Right column: estimated transfer functions for all tracers.

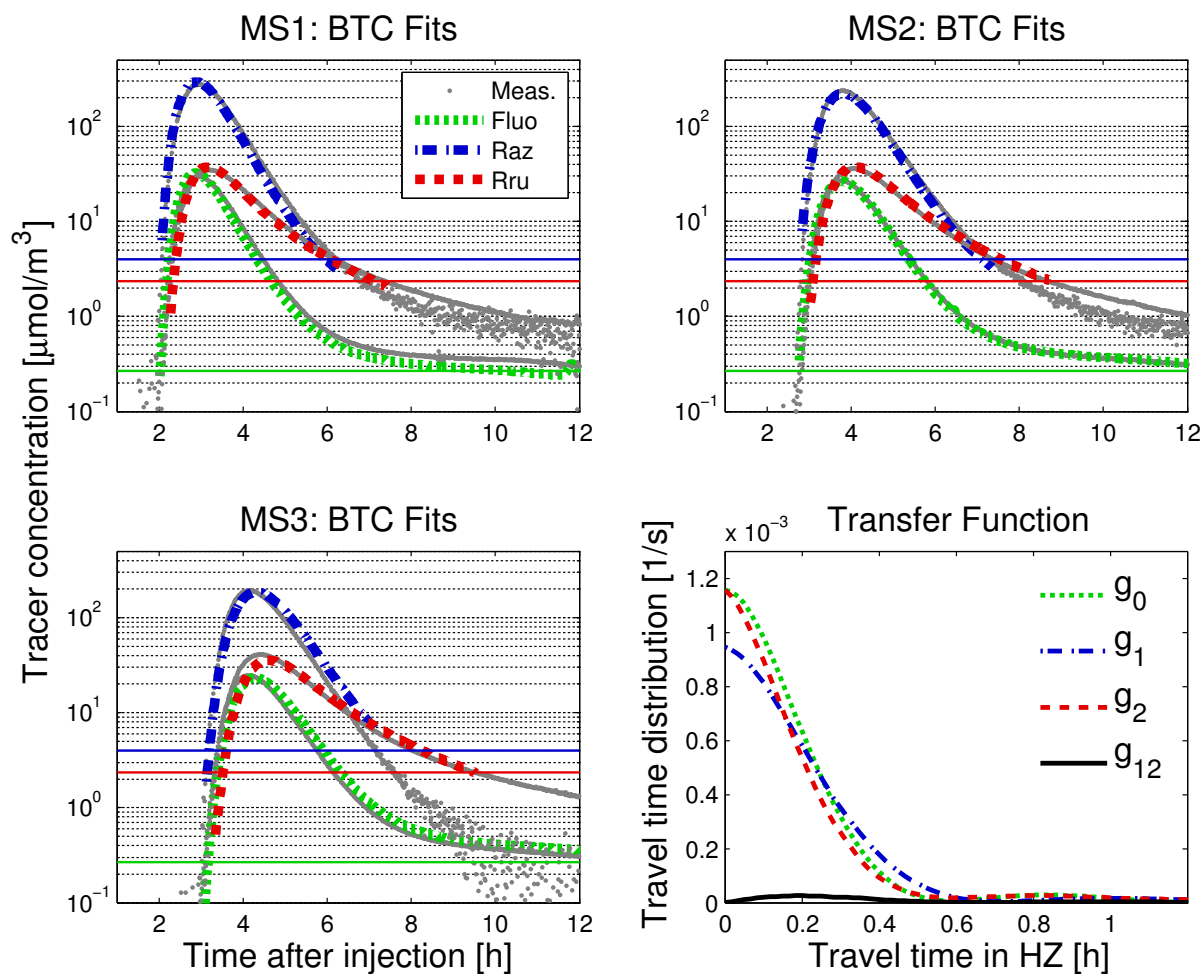


Figure 4.4: Joint fits of all breakthrough curves and all tracers; measured and fitted fitted breakthrough curves at the three monitoring stations; single set of hyperheic transfer functions obtained from jointly fitting the data of all three monitoring stations (MS). Green lines: Fluo (conservative tracer), blue lines: Raz (reactive parent compound), red lines: Rru (reaction product). Horizontal lines: quantification limits of the measurement device for each tracer.

to cancel the error introduced by the erroneous hyperheic travel-time distribution, so that all parameters were wrong.

The reactive parent compound Raz shows the steepest drop in the tail, which can be attributed to decay contributing to a loss of mass. The transformation of Raz to Rru makes the concentration of the latter drop rather slowly. The model fits agree quite well with the observations, capturing most of the tailing behavior. A little offset of the fitted Rru concentrations in comparison to the measurements can be observed at very late times and is amplified by the semi-logarithmic plot, but the main pattern of the tail is still very well captured. A closer inspection of the simulated Rru breakthrough curves at all three monitoring stations reveals that the simulated concentration peaks appear a little bit later than the measured ones, and tailing is slightly overestimated. *Bencala and Walters (1983)* reported that sorption onto the immobile sediment in the stream becomes a major control of solute transport in low-flow periods,

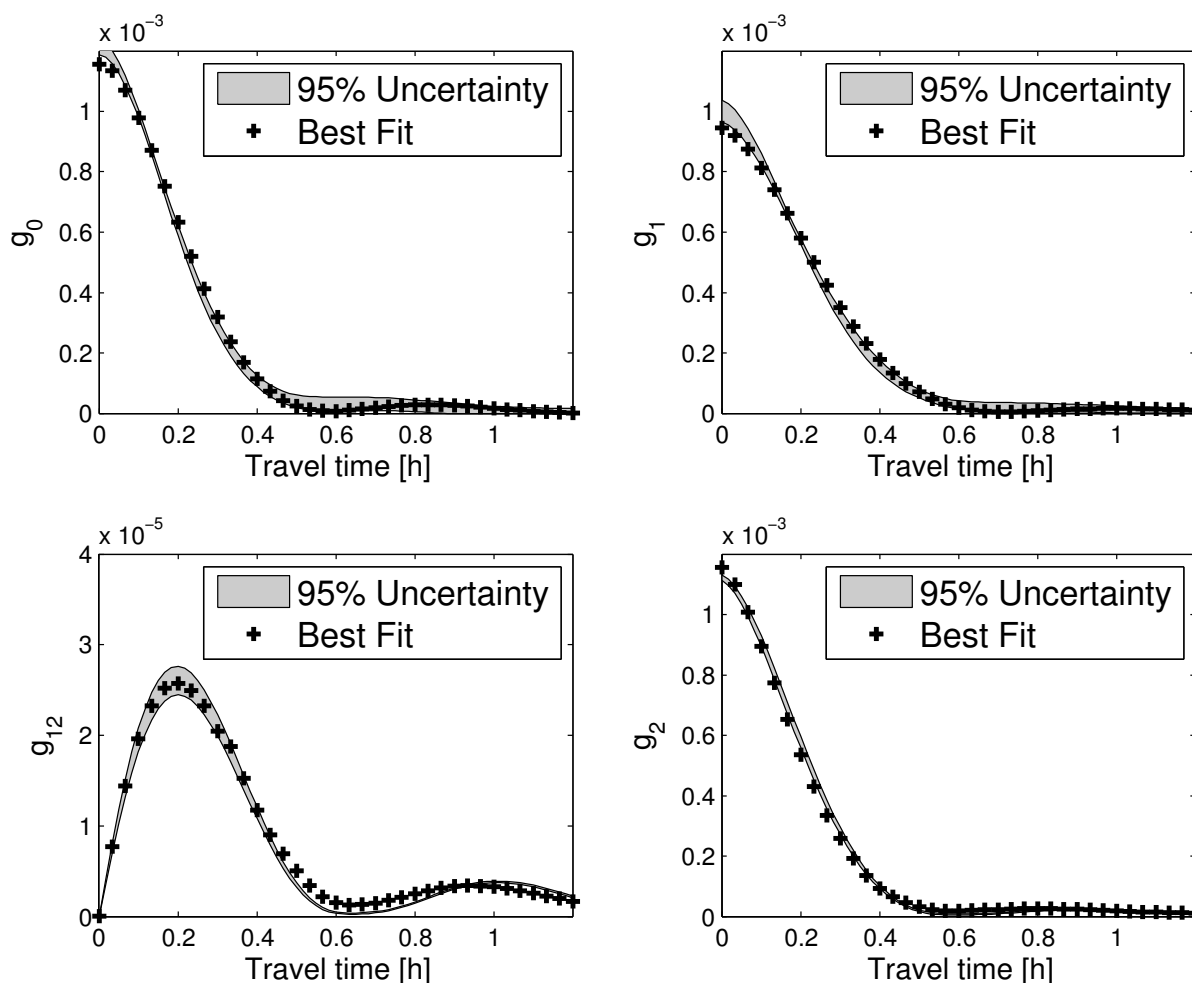


Figure 4.5: Uncertainty of the transfer functions obtained by the joint fit of all breakthrough curves at all monitoring stations. The gray band is 95% of the uncertainty for the transfer functions based on 1120 conditional realizations.

and we have observed similar findings in the fitting of other tracer tests with small in-stream velocity. We conjecture that the slow transport in the stream facilitates non-negligible contact of Raz to metabolically active biofilms at the river bottom rather than within the hyporheic zone, leading to earlier transformation of Raz to Rru than in the model, in which the transformation is strictly restricted to the hyporheic zone where transport is much slower than within the stream. Altogether, however, we assess the fits of the breakthrough curves obtained by the shape-free method to be very good, including the tailing behavior. This is also corroborated by the values of the root mean square error (RMSE) of the model fit listed in Table 4.1.

In addition, we perform a joint fit using the breakthrough curves of all compounds at all three monitoring stations to estimate a single set of parameters suitable for all stations. The results are shown in Figure 4.4. All breakthrough curves are fitted rather well, and the transfer functions have similar patterns as shown in the individual fits of Figure 4.3. In the joint fit, the mean residence time of Fluo in the hyporheic zone is 16 min, of Raz 14 min, and of Rru 23 min. As expected, the joint fit results in higher

root mean square error compared to the individual fits (see Table 4.1), but are considered still acceptable. *Bencala and Walters* (1983) have shown similar results. The good fit for several stations is promising for the application of our method to larger river reaches, provided that no major changes in river morphology and flow regime occur over the length of the reach under investigation.

Table 4.1: Estimated parameters from breakthrough curve of Fluo, Raz, and Rru at three difference monitoring stations (MS) and joint fit of all three stations. The uncertainty of the estimated parameters is given as coefficients of variation in per cent. The quality of the fits is quantified by the root mean square error (RMSE) in the reproduction of the measured breakthrough curves by the model.

Parameter	MS1: 830m	MS2: 1075m	MS3: 1210m	Joint fit
$v[m/s]$	$0.095 \pm 0.64\%$	$0.097 \pm 0.22\%$	$0.098 \pm 0.18\%$	$0.095 \pm 0.61\%$
$D[m^2/s]$	$0.198 \pm 7.7\%$	$0.118 \pm 3.2\%$	$0.145 \pm 2.8\%$	$0.185 \pm 13.4\%$
$q_{he}[1/s](\times 10^{-4})$	$5.10 \pm 2.6\%$	$5.05 \pm 0.7\%$	$4.55 \pm 0.6\%$	$4.75 \pm 2.8\%$
$\lambda_{Raz}[1/s](\times 10^{-5})$	$3.87 \pm 13.0\%$	$2.50 \pm 18.1\%$	$4.43 \pm 19.1\%$	$3.73 \pm 23.6\%$
$\lambda_{Raz \rightarrow Rru}[1/s](\times 10^{-5})$	$6.51 \pm 7.4\%$	$6.31 \pm 5.4\%$	$8.42 \pm 7.1\%$	$7.07 \pm 4.0\%$
$\lambda_{Rru}[1/s](\times 10^{-5})$	$9.29 \pm 37.5\%$	$7.81 \pm 26.7\%$	$1.74 \pm 55.1\%$	$7.70 \pm 20.4\%$
$R_{Raz}[-]$	$1.21 \pm 0.7\%$	$1.19 \pm 1.2\%$	$1.23 \pm 1.2\%$	$1.22 \pm 1.7\%$
$R_{Rru}[-]$	$1.00 \pm 0.6\%$	$1.00 \pm 0.8\%$	$1.03 \pm 4.6\%$	$1.00 \pm 0.5\%$
$K_{kin,Raz}[-]$	$0.132 \pm 9.4\%$	$0.296 \pm 2.3\%$	$0.515 \pm 3.0\%$	$0.70 \pm 71.0\%$
$K_{kin,Rru}[-]$	$0.419 \pm 21.0\%$	$0.716 \pm 18.1\%$	$2.589 \pm 29.5\%$	$0.66 \pm 24.0\%$
$k_{mt,Raz}[1/s](\times 10^{-5})$	$7.75 \pm 15.7\%$	$4.77 \pm 2.8\%$	$3.15 \pm 2.7\%$	$2.29 \pm 76.0\%$
$k_{mt,Rru}[1/s](\times 10^{-4})$	$3.31 \pm 24.6\%$	$1.95 \pm 18.1\%$	$0.661 \pm 36.1\%$	$2.74 \pm 13.0\%$
Quality of the Fits				
RMSE Fluo [ $\mu\text{mol}/\text{m}^3$ ]	$0.20 \pm 11.4\%$	$0.22 \pm 20.6\%$	$0.21 \pm 19.2\%$	$0.97 \pm 1.70\%$
RMSE Raz [ $\mu\text{mol}/\text{m}^3$ ]	$1.80 \pm 19.0\%$	$2.43 \pm 25.2\%$	$2.08 \pm 23.0\%$	$13.60 \pm 0.81\%$
RMSE Rru [ $\mu\text{mol}/\text{m}^3$ ]	$1.50 \pm 9.0\%$	$1.06 \pm 9.4\%$	$1.50 \pm 11.8\%$	$2.68 \pm 3.80\%$

For each distance and the joint fit, we estimate one set of parameters by jointly fitting the breakthrough curves of Fluo, Raz, and Rru. Table 4.1 lists the estimated transport and reaction parameters. The estimation uncertainties of the parameters are expressed as coefficients of variation in per cent. They are calculated by generating 1000 conditional realizations of the hyporheic travel-time distribution  $g(\tau)$  by the method outlined by *Liao and Cirpka* (2011). Each conditional realization of  $g(\tau)$  yields a different set of the other transport and reaction parameters.

The estimated in-stream velocity  $v$  of approximately  $0.096[m/s]$  is very consistent across the three monitoring stations, and the value obtain by the joint fit lies in the range of individual fits. The estimated in-stream dispersion coefficients  $D \approx 0.12[m^2/s]$  show a higher variability among the monitoring stations. Here, the joint fit has the highest uncertainty. The same holds for the hyporheic exchange rate coefficient  $q_{he} \approx 5 \times 10^{-4}[1/s]$ . This is so because solute spreading caused by hyporheic flow paths with small hyporheic travel times can hardly be distinguished from spreading due to Fickian in-stream dispersion (see also *Liao and Cirpka*, 2011). The average distance over which the stream water passes



on average once through the hyporheic zone is about 195 m. The transformation from Raz to an undetected compound  $\lambda_1 \approx 3.7 \times 10^{-5} [1/s]$  appears to be slower than the transformation rate from Raz to Rru  $\lambda_{12} \approx 7.1 \times 10^{-5} [1/s]$ , which is consistent with the findings of *Haggerty et al. (2009)*. Regarding the sorption of the reactive compounds within the hyporheic zone, Raz sorption appears to be partially at local equilibrium ( $R_{Raz} > 1$ ) and partially kinetically controlled ( $K_{kin,Raz} > 0$ ), whereas practically all sorption of Rru seems to be kinetically controlled according to the model fit ( $R_{Rru} \approx 1$ ). The estimated first-order mass-transfer coefficients  $k_{mt,i}$  are in the range of  $0.08 - 0.28 [1/h]$  for Raz and  $0.24 - 1.2 [1/h]$  for Rru, respectively. Overall, the estimated parameters describing chemical transformations and sorption within the hyporheic zone vary more strongly among the monitoring stations than the stream parameters and the hyporheic exchange rate coefficient. Nonetheless, the uncertainties of the joint-fit parameters are not necessarily larger than those of the individual fits. This is so, because the joint-fit parameters express a compromise solution, trying to meet the data of all measurement stations simultaneously. The range of the compromise solution is fairly small for several parameters, even though the individual fits differ to a larger extent. Joint-fit parameter values deviating too much from the compromise solution would be penalized by a severe misfit in breakthrough curves of one or two monitoring stations. The deviations between the measured and simulated breakthrough curves obtained by both the individual and joint fits that cannot be explained by parameter variability represent the conceptual model uncertainty, which is higher in the joint fit than in the individual fits, as expressed by the RMSE-values listed in Table 4.1.

The reactive parameters modify the transfer functions of Raz and Rru in comparison to the hyporheic travel-time distribution. The right column of Figure 4.3 shows the estimated hyporheic transfer functions for all three compounds, including the cross-compound transfer function  $g_{12}(\tau)$ , for all three monitoring stations. The smoothness parameter  $\theta$  can theoretically be inferred from the data using the expectation-maximization method requiring the repeated generation of multiple conditional realizations (*Cirpka et al., 2007; Liao and Cirpka, 2011*). However, due to the associated high computational effort, we have manually chosen  $\theta = 1 \times 10^{-12} [1/s^3]$  by inspecting the model fit and the shape of the estimated hyporheic travel-time distribution, which should not fluctuate too much within very small travel-time increments.

The green dotted lines in the right column of Figure 4.3 and the right bottom plot of Figure 4.4 are the estimated hyporheic transfer functions  $g_0(\tau)$  for the conservative tracer Fluo, which are identical to the underlying hyporheic travel-time distributions  $g(\tau)$ . The initial guess for  $g(\tau)$  is an exponential function, which is equivalent to using the standard transient-storage model with a single linearly reacting storage zone. This initial guess is substantially modified in the inversion procedure. While the exponential function drops rapidly at small values of  $\tau$  and more slowly at large ones, all estimated hyporheic travel-time distributions  $g(\tau)$  estimated by the shape-free approach show first the behavior of an approximately truncated Gaussian function, which starts off comparably flat and drops off more rapidly at larger values of  $\tau$ , followed by a clear secondary peak at  $\approx 1 [h]$ . This additional peak is needed to achieve the tailing of the observed breakthrough curves. It has independently been identified in the analysis of the breakthrough curves at all monitoring stations, and we observe the same behavior of the transfer functions in the joint fit shown in Figure 4.4. Our interpretation is that stream transport is affected by at least two dominant hyporheic exchange mechanisms: the exchange by micro bedforms in which the stream water stays only more or less 16 minutes on average, and a second feature with travel times of about one hour. Even though the ratio of two residence times is smaller than 5, which was proposed by *Choi et al. (2000)* as characteristic value for the separation of different regimes, we see a significant time deviation in the transfer functions. With a pulse injection and an overall duration of the experiment of 12 hours, we cannot exclude that larger structures leading to considerably longer hyporheic travel-times exist in the stream system. These would remain unnoticed due to quantification limits (*Ward et al., 2010*). Comparing the

transfer function obtained by the individual fits of the monitoring stations and by the joint fit, we observe that the joint fit transfer function exhibits no more a clear disconnection between the two peaks.

The blue dash-dotted lines in the right column of Figure 4.3 and the right bottom plot of Figure 4.4 are the estimated hyporheic transfer functions  $g_1(\tau)$  of the reactive parent compound Raz. The most obvious difference to  $g(\tau)$  is that the truncated-Gaussian like hump at small values of  $\tau$  is lower and smaller for Raz than for Fluo. This is caused by the substantial equilibrium sorption of Raz estimated by the inversion procedure, which is also corroborated by the shift of the secondary peak. Kinetic sorption and decay have only a minor influence on the hyporheic transfer functions of Raz. The latter also becomes clear when considering the cross-compound transfer function  $g_{12}(\tau)$ , quantifying the response of the Rru concentration in the hyporheic water returning to the stream due to a unit pulse of Raz entering the hyporheic zone. The cross-compound transfer function  $g_{12}(\tau)$  is plotted as black solid lines in the right column of Figure 4.3 and the right bottom plot of Figure 4.4. More Rru is generated with increasing distance in stream, which is consistent with our understanding that along the river reach, Raz infiltrates more often into the hyporheic zone, where Rru is formed. On the time-scales that the stream water stays within the hyporheic zone, only a small fraction of Raz is transformed to Rru. This behavior is in agreement to the finding that over the entire reach less than 20% of Raz has been transformed to Rru (Lemke *et al.*, 2013a).

The red dashed line in the right column of Figure 4.3 and the right bottom plot of Figure 4.4 are the estimated hyporheic transfer functions  $g_2(\tau)$  of the reaction product Rru, that is, the concentration of Rru in the returning water due to a unit pulse of Rru introduced into the hyporheic zone. Because the model fits estimate negligible equilibrium sorption of Rru, and kinetic sorption has estimated characteristic time scales of 0.8-4.2 hours when most of the hyporheic water has already returned back into the stream, the hyporheic transfer functions  $g_2(\tau)$  are very similar to the underlying hyporheic travel-time distributions  $g(\tau)$  at small values of  $\tau$ . Only at larger travel-time values, some modification in the shape becomes visible.

Based on 1000 conditional realizations, Figure 4.5 shows the uncertainty of all transfer functions. The single best fit assuming a perfectly smooth hyporheic travel-time distribution lies mainly in the 95% uncertainty band of the conditional realizations, but the uncertainty itself is rather narrow for all parameters. As discussed above, the small parameter uncertainty represents the narrow parameter range for a compromise solution trying to meet all breakthrough curves as well as possible, rather than a perfect match of simulated and measured breakthrough curves.

## 4.4 Discussion and Conclusions

We have presented a model analyzing stream-tracer tests combining a conservative tracer, Fluo, and a linearly decaying tracer, Raz, that are jointly introduced into the stream as a single pulse, in which we also consider the concentration of the reaction product, Rru. The hyporheic travel-time distributions estimated by the shape-free approach do not resemble a simple parametric function. The independent fits for the individual monitoring stations agreed in a rather broad truncated Gaussian-like distribution of hyporheic travel times at small values, with a mean residence time of 16 min, amended by a secondary peak at about 1 hour. Revealing such features is a major advantage of the shape-free approach, even though the exact details of the estimated hyporheic travel-time distributions should be handled with care. In contrast to parametric models, the shape-free method is able to reproduce the tails of the observed concentration breakthrough curves. This is only possible when accounting for relatively longer (or slower) hyporheic travel paths. Of course, we cannot make any statements about the significance of possible hyporheic

travel paths with travel times beyond the duration the experiment. As discussed by *Harvey et al.* (1996), only the fast component of hyporheic exchange can be detected by performing stream-tracer experiments. The fits of the breakthrough curves are excellent, even though small mismatches are still unavoidable.

Fits using all three tracers for each monitoring station or even across all stations is superior to only using the conservative-tracer data. Even though not shown in this paper, we have compared the results from only fitting the Fluo data and jointly fitting the breakthrough curves of all compounds. While the in-stream velocity is estimated very consistently among different monitoring stations and between fits of Fluo alone or all tracers, the in-stream dispersion coefficient tends to be overestimated and the hyporheic exchange rate coefficient underestimated, when we use only the Fluo breakthrough curves compared to the results from fitting the breakthrough curves of all compounds. This is so because both parameters explain spreading of the conservative-tracer breakthrough curves. By considering the reactive tracers in the model analysis, the distinction between in-stream mixing and spreading due to hyporheic exchange has improved, because the transformation of Raz to Rru takes place only in the hyporheic zone.

Accounting for the reaction product is an important consistency check for both the experimental data and the model. Fitting only the conservative tracer and the reactive parent compound Raz, as done by *Liao and Cirpka* (2011) and *Lemke et al.* (2013a), yields a total loss coefficient of Raz, but it is not possible to quantify the fraction that is transformed to Rru. The latter information is important because the Raz-to-Rru transformation has been related to aerobic respiration of the stream bed (*González-Pinzón et al.*, 2012; *Argerich et al.*, 2011), whereas an unspecified loss of Raz may be caused by processes that are completely independent of metabolic activity. The mismatch between the total mass of generated Rru and the remaining Raz compared to the total mass of injected Raz may indicate long travel times in the hyporheic zone that cannot be detected within the experimental time, or unknown transformation processes.

A key motivation for developing the approach was to better distinguish in-stream mixing processes from spreading due to hyporheic exchange. We cannot exclude that some mismatch still remains. The estimated in-stream dispersion coefficient has the largest value at the first monitoring station, while the hyporheic exchange coefficient decreases with travel distance. This could be an aliasing effect, because with increasing time the concentration length profiles in systems with transient storage increasingly resemble Gaussian profiles, characteristic for Fickian dispersion. The inverse model is thus allowed to trade missing hyporheic exchange for an overestimation of in-stream dispersion if only the conservative tracer is considered. The latter should be penalized by misfits of the reactive breakthrough curves because an underestimation of hyporheic exchange causes a decrease in overall simulated transformation from Raz to Rru. In the current implementation, however, such a misfit can be balanced by increasing the estimated rate coefficients, which can actually be observed in Table 4.1. Thus, the joint fit of breakthrough curves from several monitoring stations provides more reliable estimates of the parameters. We see that the joint-fit values of in-stream velocity  $v$ , in-stream dispersion coefficient  $D$  and the hyporheic exchange-rate coefficient  $q_{he}$  all lie in the range of the individual fits. While the root mean square error of the joint fit is higher than in the individual fit, we consider the joint model fit as well acceptable (see also *Bencala and Walters*, 1983).

Our model accounts for two-site sorption of the reactive compounds within the hyporheic zone. This degree of complexity was necessary in the analysis of column tests and also when fitting other stream-tracer tests not shown in this paper. In the current application, the Raz sorption appears to be partially in local equilibrium whereas Rru sorption is almost entirely controlled by kinetics. We cannot fully exclude that unidentified processes, such as a contribution of Raz-to-Rru transformation within the stream rather than within the hyporheic zone has introduced bias in the estimated sorption parameters. Neglecting sorption of the reactive tracers altogether, as done by *Liao and Cirpka* (2011), however, is clearly not

valid for the Raz/Rru-system.

In recent years, several studies have shown that transient storage occurs not only in the hyporheic zone but also within the stream itself (e.g., *Harvey et al.*, 2005; *Briggs et al.*, 2009, 2010; *Gooseff et al.*, 2011). Separating the effects of individual storage zones on stream-tracer breakthrough curves is difficult without additional information. While our stream gives no visual indication (large pools, dead zones, etc.) of intensive in-stream storage, we cannot guarantee that all transient storage identified by our method is located within the hyporheic zone. This implies that all identified parameters are effective values, describing the net behavior of the stream over the length of the reach. We don't assess this as over-critical because this effective description is what is targeted for in reach-scale solute transport anyway. The model indicates net reach-scale solute retention, metabolic activity, and sorptivity, maybe not only located in the hyporheic zone but most likely transferrable to other compounds (oxygen, sorbing contaminants) of interest.

We are fully aware that hyporheic exchange is a complex three-dimensional process, depending on multiple properties of the stream, its bed, and the adjacent aquifer. Rather than resolving all influencing factors, we describe transport as a "one-and-a-half" dimensional process with one-dimensional advective-dispersive transport in the stream coupled to one-dimensional, formally purely advective transport in the hyporheic zone. The effects of the complex flow geometry on solute transport in the real system are captured by the hyporheic travel-time distribution. This approach is valid only when the reach length under consideration is much larger than the dominant flow paths of hyporheic exchange. For simplicity, we also assume stationarity of the hyporheic travel-time distribution. That is, the same distribution is considered along the entire reach length simulated. This requires that the general geomorphological and other characteristics of the stream, determining hyporheic exchange processes, do not significantly change over the distance under consideration. In its current implementation, the model does not consider gaining or losing of river water in the experiment reach. Losing river water cannot be detected by tracer tests with a single injection point, whereas gaining conditions would lead to dilution of all tracers. Towards detecting gross losses and gains, *Payn et al.* (2009) performed a series of conservative-tracer/dilution-gauging tests along a river reach, and *Covino et al.* (2011) on the scale of entire river networks. Both losing and gaining lead to discharge and thus velocity values varying over the length of the reach. These processes can be included if sufficient knowledge about the location of major gains and losses exist, detected, e.g., by differential dilution gauging. Most likely, the analytical solution of the transport equation in the Laplace domain had to be replaced by numerical simulation, whether in the time or Laplace domain.

In summary, the method and model presented in this paper fits the breakthrough curves from tracer experiment in River Goldersbach very well. The shape-free method allows revealing unconventional patterns of the hyporheic travel-time distribution, yielding better fit of the breakthrough curves than obtained by parametric hyporheic travel-time distributions (*Lemke et al.*, 2013a). Jointly fitting high-resolution breakthrough curves of Fluo, Raz, and Rru gives insights into in-stream transport, hyporheic exchange, metabolic activity and sorption properties of the stream bed.

## Acknowledgments

We are thankful to the reviewers Fulvio Boano, Aaron Packman and Roy Haggerty for their helpful comments and constructive suggestions. This work was supported by a grant from the Ministry of Science, Research and Arts of Baden-Württemberg (AZ Zu 33-721.3-2), the Helmholtz Center for Environmental Research - UFZ, Leipzig, and the German Academic Exchange Service (DAAD).

## 4.A Solution in the Laplace Domain

The initial value problem Eq. (4.5) with the boundary conditions Eq. (4.6) is solved in the Laplace domain, where the convolution term becomes a multiplication, such that the partial differential equation is converted into a second-order homogeneous (for Fluo and Raz) or inhomogeneous (for Rru) linear ordinary differential equation:

$$\begin{aligned} (s + q_{he} - q_{he}\tilde{g}_0)\tilde{c}_0(s, x) + v\frac{d\tilde{c}_0(s, x)}{dx} - D\frac{d^2\tilde{c}_0(s, x)}{dx^2} &= 0 \\ (s + q_{he} - q_{he}\tilde{g}_1)\tilde{c}_1(s, x) + v\frac{d\tilde{c}_1(s, x)}{dx} - D\frac{d^2\tilde{c}_1(s, x)}{dx^2} &= 0 \\ (s + q_{he} - q_{he}\tilde{g}_2)\tilde{c}_2(s, x) - q_{he}\tilde{g}_{12}\tilde{c}_1 + v\frac{d\tilde{c}_2(s, x)}{dx} - D\frac{d^2\tilde{c}_2(s, x)}{dx^2} &= 0 \end{aligned} \quad (4.16)$$

in which  $s$  is the complex Laplace variable and the tilde marks the corresponding Laplace transform in time. Applying the Laplace transformation on the boundary conditions, we obtain:

$$\begin{aligned} v\tilde{c}_0(s, x) - D\frac{d\tilde{c}_0(s, x)}{dx}\Big|_{x=0} &= \frac{m_0}{A} \\ v\tilde{c}_1(s, x) - D\frac{d\tilde{c}_1(s, x)}{dx}\Big|_{x=0} &= \frac{m_1}{A} \\ v\tilde{c}_2(s, x) - D\frac{d\tilde{c}_2(s, x)}{dx}\Big|_{x=0} &= 0 \end{aligned} \quad (4.17)$$

For simplicity we denote:

$$\begin{aligned} f_0 &= s + q_{he} - q_{he}\tilde{g}_0 \\ f_1 &= s + q_{he} - q_{he}\tilde{g}_1 \\ f_2 &= s + q_{he} - q_{he}\tilde{g}_2 \\ f_{12} &= -q_{he}\tilde{g}_{12} \end{aligned} \quad (4.18)$$

Therefore, the governing equations in the Laplace domain become:

$$\begin{aligned} f_0\tilde{c}_0(s, x) + v\frac{d\tilde{c}_0(s, x)}{dx} - D\frac{d^2\tilde{c}_0(s, x)}{dx^2} &= 0 \\ f_1\tilde{c}_1(s, x) + v\frac{d\tilde{c}_1(s, x)}{dx} - D\frac{d^2\tilde{c}_1(s, x)}{dx^2} &= 0 \\ f_2\tilde{c}_2(s, x) + f_{12}\tilde{c}_1 + v\frac{d\tilde{c}_2(s, x)}{dx} - D\frac{d^2\tilde{c}_2(s, x)}{dx^2} &= 0 \end{aligned} \quad (4.19)$$

subject to:

$$\begin{aligned} v\tilde{c}_i(s, x) - D\frac{d\tilde{c}_i(s, x)}{dx}\Big|_{x=0} &= \frac{m_i}{A}, i = 0, 1, 2 \\ \lim_{x \rightarrow \infty} \tilde{c}_i(s, x) &= 0, \forall s \end{aligned} \quad (4.20)$$

The solution in the Laplace domain reads as:

$$\begin{aligned}
\tilde{c}_0(s,x) &= \tilde{c}_0(s,0) \exp(-\alpha_0 x), & \alpha_0 &= \frac{-v + \sqrt{v^2 + 4Df_0}}{2D}, & \tilde{c}_0(s,0) &= \frac{m_0}{A} \frac{1}{v + \alpha_0 D} \\
\tilde{c}_1(s,x) &= \tilde{c}_1(s,0) \exp(-\alpha_1 x), & \alpha_1 &= \frac{-v + \sqrt{v^2 + 4Df_1}}{2D}, & \tilde{c}_1(s,0) &= \frac{m_1}{A} \frac{1}{v + \alpha_1 D} \\
\tilde{c}_2(s,x) &= a_2 \exp(-\alpha_2 x) + a_{12} \exp(-\alpha_1 x), \\
\alpha_2 &= \frac{-v + \sqrt{v^2 + 4Df_2}}{2D}, & a_{12} &= \frac{f_{12}}{f_1 - f_2} \tilde{c}_1(s,0) & a_2 &= -\frac{f_{12}}{f_1 - f_2} \frac{m_1}{A} \frac{1}{v + \alpha_2 D}
\end{aligned} \tag{4.21}$$

which is back-transformed into the time domain by the numerical method of *De Hoog et al.* (1982).

## 4.B Transport Equations for the Hyporheic Zone with Two-Site Sorption

We use the one-dimensional advection equation to describe solute transport in the hyporheic zone, replacing the spatial domain by the travel-time coordinate  $\tau^*$ . Considering first-order decay of the reactive tracer in the hyporheic zone, we arrive at:

$$n \frac{\partial c_{hz}}{\partial t} + (1-n)\rho_s \frac{\partial s}{\partial t} + n \frac{\partial c_{hz}}{\partial \tau^*} = -\lambda n c_{hz} \tag{4.22}$$

subject to a zero initial condition and a pulse-like boundary condition at  $\tau^* = 0$ :

$$c(t=0) = 0 \quad \forall \tau^*; \quad c_{\tau^*=0}(t) = \delta(t) \tag{4.23}$$

in which  $c_{hz} [ML^{-3}]$  is the tracer concentration in the aqueous phase of the hyporheic zone;  $s [MM^{-1}]$  is the mass of sorbing tracer per mass of solids;  $n [-]$  is the effective porosity in the hyporheic zone;  $\rho_s [ML^{-3}]$  is the mass density of the solids;  $\lambda [T^{-1}]$  is the first-order decay coefficient;  $t [T]$  is real time, whereas  $\tau^* [T]$  is travel time within the hyporheic zone, which is a substitute for space. We distinguish the sorbing-phase concentration in a contribution at equilibrium sites and another one at kinetic sites:

$$s = s_{eq} + s_{kin} \tag{4.24}$$

assuming linear behavior of both equilibrium and kinetic sorption:

$$s_{eq} = K_d^{eq} c_{hz} \tag{4.25}$$

$$\frac{\partial s_{kin}}{\partial t} = k_{mt} (K_d^{kin} c_{hz} - s_{kin}) \tag{4.26}$$

in which  $K_d^{eq} [L^3 M^{-1}]$  is the distribution coefficient between the equilibrium sorption sites and water,  $K_d^{kin} [L^3 M^{-1}]$  is the distribution coefficient between the kinetic sorption sites and water, and  $k_{mt} [T^{-1}]$  is the rate coefficient of kinetic mass-transfer.

Substituting Eqs. (4.25 & 4.26) into Eq. (4.22), we arrive at:

$$\begin{aligned}
& n \frac{\partial c_{hz}}{\partial t} + (1-n) \rho_s K_d^{eq} \frac{\partial c_{hz}}{\partial t} + (1-n) \rho_s \frac{\partial s_{kin}}{\partial t} + n \frac{\partial c_{hz}}{\partial \tau^*} = -\lambda n c_{hz} \\
\Rightarrow & \left(1 + \frac{1-n}{n} \rho_s K_d^{eq}\right) \frac{\partial c_{hz}}{\partial t} + \frac{1-n}{n} \rho_s \frac{\partial s_{kin}}{\partial t} + \frac{\partial c_{hz}}{\partial \tau^*} = -\lambda c_{hz} \\
\Rightarrow & R \frac{\partial c_{hz}}{\partial t} + \frac{\partial c_{hz}}{\partial \tau^*} = \frac{1-n}{n} \rho_s k_{mt} (s_{kin} - K_d^{kin} c_{hz}) - \lambda c_{hz}
\end{aligned} \tag{4.27}$$

in which we denote:

$$R = 1 + \frac{1-n}{n} \rho_s K_d^{eq} \tag{4.28}$$

as the retardation factor due to equilibrium sorption.

We now introduce the aqueous saturation concentration  $c_i$  [ $ML^{-3}$ ] with respect to the concentration at the kinetic sorption sites  $s_{kin}$ :

$$c_i = \frac{s_{kin}}{K_d^{kin}} \tag{4.29}$$

yielding the following system of equations for Eqs. (4.27 & 4.26):

$$\begin{aligned}
R \frac{\partial c_{hz}}{\partial t} + \frac{\partial c_{hz}}{\partial \tau^*} &= K_{kin} k_{mt} (c_i - c_{hz}) - \lambda c_{hz} \\
\frac{\partial c_i}{\partial t} &= k_{mt} (c_{hz} - c_i)
\end{aligned} \tag{4.30}$$

in which

$$K_{kin} = \frac{1-n}{n} \rho_s K_d^{kin}. \tag{4.31}$$

Eq. (4.30) describes how the tracer concentration in the hyporheic zone changes over time in both the aqueous and sorbing phases. As seen in the governing equation Eq.(4.1), these values are not used directly in the model, but needed for calculating  $g_i(\tau)$  in Eq. (4.11).

## 4.C Solution of Governing Equation with Two-Site Sorption in the Laplace Domain

In the Laplace domain, Eqs. (4.9 & 4.10) become:

$$(Rs + \lambda_1 + \lambda_{12}) \tilde{c}_1^{hz} + K_{kin} s \tilde{c}_1^* + \frac{d\tilde{c}_1^{hz}}{d\tau^*} = 0, \quad (s + k_{mt}) \tilde{c}_1^* = k_{mt} \tilde{c}_1^{hz}, \quad \tilde{c}_1^{hz} = b_1 \tag{4.32}$$

$$(Rs + \lambda_2) \tilde{c}_2^{hz} + K_{kin} s \tilde{c}_2^* + \frac{d\tilde{c}_2^{hz}}{d\tau^*} = \lambda_{12} \tilde{c}_1^{hz}, \quad (s + k_{mt}) \tilde{c}_2^* = k_{mt} \tilde{c}_2^{hz}, \quad \tilde{c}_2^{hz} = b_2$$

$$\begin{aligned}
\Rightarrow & \beta_1 \tilde{c}_1^{hz} + \frac{d\tilde{c}_1^{hz}}{d\tau^*} = 0, \quad \Rightarrow \tilde{c}_1^{hz} = b_1 \exp(-\beta_1 \tau^*) \\
& \beta_2 \tilde{c}_2^{hz} - \lambda_{12} \tilde{c}_1^{hz} + \frac{d\tilde{c}_2^{hz}}{d\tau^*} = 0, \quad \Rightarrow \tilde{c}_2^{hz} = d_2 \exp(-\beta_2 \tau^*) + d_{12} \exp(-\beta_1 \tau^*)
\end{aligned} \tag{4.33}$$

where

$$\beta_1 = Rs + \lambda_1 + \lambda_{12} + K_{kin}s \frac{k_{mt}}{s + k_{mt}} \quad (4.34)$$

$$\beta_2 = Rs + \lambda_2 + K_{kin}s \frac{k_{mt}}{s + k_{mt}} \quad (4.35)$$

$$d_2 = b_2 + \frac{\lambda_{12}}{\beta_1 - \beta_2} b_1 \quad (4.36)$$

$$d_{12} = -\frac{\lambda_{12}}{\beta_1 - \beta_2} b_1 \quad (4.37)$$

## 4.D Discretization of the Hyporheic Transfer Functions $g_i(\tau)$

Using the same technique as *Liao and Cirpka* (2011), the hyporheic travel-time distribution  $g(\tau^*)$  is not forced to follow a specific function characterized by few parameters. Instead,  $g(\tau^*)$  is discretized into  $n_\tau$  sections with identical increment  $\Delta\tau^*$ , and we apply piece-wise linear interpolation of discrete values  $g(\tau_i^*)$ :

$$g(\tau^*) = \sum_{i=1}^{n_\tau} \xi_i + \eta_i \tau^* \quad (4.38)$$

in which  $\xi_i = g(\tau_{i-1}^*) - \eta_i \tau_{i-1}^*$  is the intersection of the linear function, and  $\eta_i = \frac{g(\tau_i^*) - g(\tau_{i-1}^*)}{\Delta\tau^*}$  is the slope.

The hyporheic transfer functions  $g_0(\tau)$ ,  $g_1(\tau)$ ,  $g_2(\tau)$  and  $g_{12}(\tau)$  depend on the hyporheic travel-time distribution  $g(\tau^*)$  via Eq. (4.11). Because  $g(\tau^*)$  is defined as piecewise linear function, the Laplace transforms  $\tilde{g}_i(s)$ ,  $i = 0, 1, 2, 12$  must also be computed piecewise. Substituting Eq. (4.38) into Eq. (4.12) and performing the integration, yields:

$$\begin{aligned} \tilde{g}_0(s) &= \sum_{i=1}^{n_\tau} \int_{\tau_{i-1}^*}^{\tau_i^*} (\xi_i + \eta_i \tau^*) \exp(-s\tau^*) d\tau^* \\ &= \sum_{i=1}^{n_\tau} \left[ \left( -\frac{\xi_i}{s} - \frac{\eta_i}{s} \tau_i^* - \frac{\eta_i}{s^2} \right) \exp(-s\tau_i^*) + \left( \frac{\xi_i}{s} + \frac{\eta_i}{s} \tau_{i-1}^* + \frac{\eta_i}{s^2} \right) \exp(-s\tau_{i-1}^*) \right] \end{aligned} \quad (4.39)$$

$$\begin{aligned} \tilde{g}_1(s) &= \sum_{i=1}^{n_\tau} \int_{\tau_{i-1}^*}^{\tau_i^*} (\xi_i + \eta_i \tau^*) \exp(-\beta_1 \tau^*) d\tau^* \\ &= \sum_{i=1}^{n_\tau} \left[ \left( -\frac{\xi_i}{\beta_1} - \frac{\eta_i}{\beta_1} \tau_i^* - \frac{\eta_i}{\beta_1^2} \right) \exp(-\beta_1 \tau_i^*) + \left( \frac{\xi_i}{\beta_1} + \frac{\eta_i}{\beta_1} \tau_{i-1}^* + \frac{\eta_i}{\beta_1^2} \right) \exp(-\beta_1 \tau_{i-1}^*) \right] \end{aligned} \quad (4.40)$$



$$\begin{aligned}
\tilde{g}_{12}(s) &= \frac{\lambda_{12}}{\beta_1 - \beta_2} \sum_{i=1}^{n_\tau} \int_{\tau_{i-1}^*}^{\tau_i^*} (\xi_i + \eta_i \tau^*) (\exp(-\beta_2 \tau^*) - \exp(-\beta_1 \tau^*)) d\tau^* \\
&= \frac{\lambda_{12}}{\beta_1 - \beta_2} \sum_{i=1}^{n_\tau} \left[ \left( -\frac{\xi_i}{\beta_2} - \frac{\eta_i}{\beta_2} \tau_i^* - \frac{\eta_i}{\beta_2^2} \right) \exp(-\beta_2 \tau_i^*) + \left( \frac{\xi_i}{\beta_2} + \frac{\eta_i}{\beta_2} \tau_{i-1}^* + \frac{\eta_i}{\beta_2^2} \right) \exp(-\beta_2 \tau_{i-1}^*) \right. \\
&\quad \left. - \left( -\frac{\xi_i}{\beta_1} - \frac{\eta_i}{\beta_1} \tau_i^* - \frac{\eta_i}{\beta_1^2} \right) \exp(-\beta_1 \tau_i^*) - \left( \frac{\xi_i}{\beta_1} + \frac{\eta_i}{\beta_1} \tau_{i-1}^* + \frac{\eta_i}{\beta_1^2} \right) \exp(-\beta_1 \tau_{i-1}^*) \right] \quad (4.41)
\end{aligned}$$

$$\begin{aligned}
\tilde{g}_2(s) &= \sum_{i=1}^{n_\tau} \int_{\tau_{i-1}^*}^{\tau_i^*} (\xi_i + \eta_i \tau^*) \exp(-\beta_2 \tau^*) d\tau^* \\
&= \sum_{i=1}^{n_\tau} \left[ \left( -\frac{\xi_i}{\beta_2} - \frac{\eta_i}{\beta_2} \tau_i^* - \frac{\eta_i}{\beta_2^2} \right) \exp(-\beta_2 \tau_i^*) + \left( \frac{\xi_i}{\beta_2} + \frac{\eta_i}{\beta_2} \tau_{i-1}^* + \frac{\eta_i}{\beta_2^2} \right) \exp(-\beta_2 \tau_{i-1}^*) \right] \quad (4.42)
\end{aligned}$$



## Chapter 5

# Consistent Simulation of Pulse-Like Conservative and Reactive Stream-Tracer Experiments on Multiple Scales

Z. Liao, M. Gritsch, J. Knapp, D. Lemke, O.A. Cirpka: Consistent simulation of pulse-like conservative and reactive stream-tracer experiments on multiple scales, abstract H14B-07, presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 3-7 Dec, 2012.

Liao, Z.; Gritsch, M.; Lemke, D.; Ostenbrück, K.; Cirpka, O.A.: Data analysis for tracer experiments using shape-free deconvolution method on multiple scales, NovCare, Leipzig, 13-16 May, 2013

### Abstract

The redox-sensitive compound resazurin (Raz) has recently been introduced as reactive tracer for streams undergoing hyporheic exchange. Resazurin is converted to resorufin (Rru) in viable cells so that the Raz-Rru system acts as a probe for the metabolic activity of a stream, which is believed to be concentrated in the hyporheic zone. The reactive tracers, however, can also be sorbed. Advective-dispersive transport in the stream, potentially non-conventional transient storage in the hyporheic zone, sorption at the riverbed material, and transformation of the compounds leads to a coupled system that needs to be considered when analyzing pulse-like tracer experiments with the Raz-Rru system and a conservative tracer.

We present a consistent model formulation of the reactive-transport system and apply it in the interpretation of experiments on multiple scales. (1) Column experiments with a length of about 30cm indicate that Raz and Rru undergo two-site sorption and chemical transformation during the passage through riverbed material. (2) We have analyzed breakthrough curves of the conservative and reactive tracers in piezometers within gravel bars during stream-tracer experiments. The conservative-tracer breakthrough curves in the stream and the adjacent gravel bar were deconvoluted to obtain the stream-to-piezometer travel-time distribution. Fitting the reactive tracer breakthrough curves, while accounting for the conservative travel-time distribution and the input signal in the stream, confirmed the two-site-sorption plus decay model of the Raz-Rru system in the hyporheic zone. (3) In Chapter 4 I have analyzed solute transport on the scale of entire stream reaches. There, I have jointly analyzed the conservative and reactive tracer breakthrough curves using a shape-free approach for the hyporheic travel-time distribution, two-site sorption of Raz and Rru in the hyporheic zone, and the Raz-to-Rru transformation. By this, we

could identify the stream velocity, in-stream dispersion coefficient, the hyporheic exchange rate, the time distribution over which the water stays within the hyporheic zone, the transformation coefficient, and sorption parameters of the riverbed. Consistent data sets for multiple distances along the stream could be identified.

## 5.1 Introduction

Hyporheic exchange plays an important role in river water quality. The infiltration of river water into the subsurface facilitates microbial activity, filtration of particles including pathogens, and degradation of dissolved pollutants. The travel time of the river water in the river bed determines the strength and completeness of these activities and therefore the hyporheic residence time is of great interest in understanding the river water-groundwater exchange system.

Performing tracer tests is a well-known method in investigating the transport of solutes in rivers subject to short stays in adjacent aquifers. Here an easy-to-detect compound is introduced into the system and observed at point further downstream. 'Smart' tracers, which are irreversibly reduced in the hyporheic zone, have recently become popular in stream tracer tests (*Haggerty et al.*, 2008, 2009), because the difference of the breakthrough curves (BTCs) between conservative and reactive tracers provides insight into the loss of the reactive tracer upon hyporheic exchange.

In Chapter 3 and Chapter 4, I have proposed a shape-free deconvolution method for analyzing reach-scale tracer tests using the conservative tracer Fluorescein (Fluo) and the reactive tracer Resazurin (Raz). The reaction product Resorufin (Rru) is also included in the analysis in Chapter 4 to provide more information of the reactive-tracer behavior. The model performs fairly well in fitting the BTCs for all three tracers at different observation stations and estimates the shape-free hyporheic travel-time distribution and some other parameters. So far, this model has been demonstrated to work in the analysis of data from reach-scale experiments, therefore the scaling characteristics of the model still needs to be investigated. This chapter focuses on the application of the model proposed in Chapter 4 to different scales of experiments: a) a column experiment in the lab (around 30cm); b) an experiment at a gravel bar in the stream channel (around 0.5-2m). The river scale tracer test is demonstrated in Chapter 4.

## 5.2 Experiments

We have performed three tracer tests at different scales. The sediments of the column experiment are from a sand bar in River Schaich, where the gravel-bar experiment was carried out. The reach-scale tracer test is done in River Goldersbach, near Tübingen, Germany (see also Chapter 4).

In the following analysis, we inject the conservative tracer Fluo and the reactive tracer Raz into the system, and the observations include the concentration of Fluo, Raz and its reaction product Rru. Within the system, decay and two-site sorption is considered for Raz and Rru, as shown in Chapter 4 (see Figure. 5.1).

### 5.2.1 Column experiments

The column experiments were carried out in the lab. A schematic plot of the experiment can be seen in Figure 5.2. The column, which is about 30cm long, was filled with the sediments from a sand bar at River Schaich. Two fluorimeters were installed at the inlet and outlet of the column, measuring

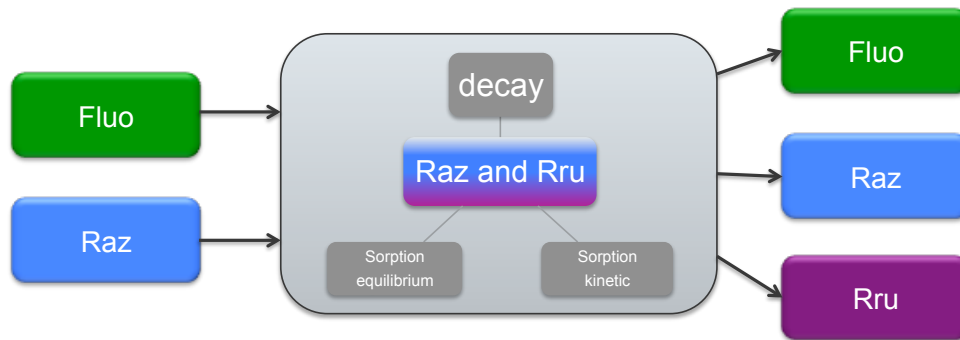


Figure 5.1: Conservative tracer (Fluo) and reactive tracer (Raz) are used in the tracer tests, where the tracer concentration of these tracers and the reaction product (Rru) are measured at the observation points.

the concentration of Fluo, Raz, and Rru. The tracer solution was pumped through the column, which is placed in a constant temperature room at 20°C. The pH value was kept about 7, the flow rate was 20ml/min, and the sediments were sieved to a particle diameter smaller than 4mm. The column was first flushed with water, then the tracer solution flowed through it for about 10L in 7 hours, after which the flow was switched back to tracer-free water to obtain the tailing behavior in the breakthrough curves.

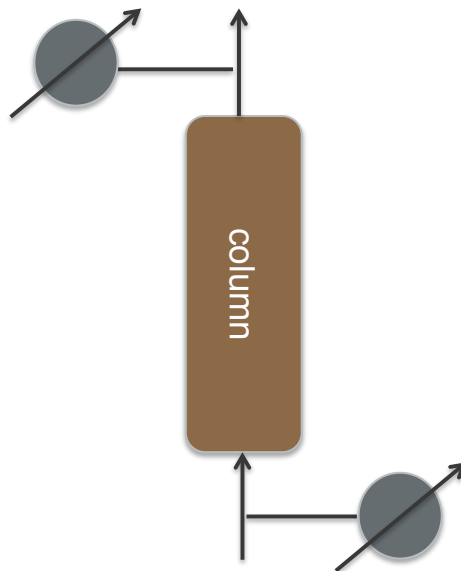


Figure 5.2: Schematic plot for the column experiments. Two fluorometers were placed at the inlet and the outlet of the column. Sieved sediments with diameter  $< 4\text{mm}$  were filled in to the column. The experiment was done at room temperature and the pH value was kept about 7.

## 5.2.2 Gravel-bar experiments

The gravel-bar experiment was performed at River Schaich, where a mixture of Fluo and Raz was injected as a pulse into the river, and two fluorimeters were used to measure the concentrations of Fluo, Raz, and the reaction product Rru right in front of the gravel bar in the river and within a piezometer in the gravel bar. A schematic plot can be found in Figure 5.3. The gravel bar in the field resembled the column in the lab, but the input signals were much smoother and the incoming water already contained the reaction product Rru because the stream water had already undergone hyporheic exchange before reaching the spot with the gravel-bar under investigation. The temperature was about 18°C and the pH values approximately 8.

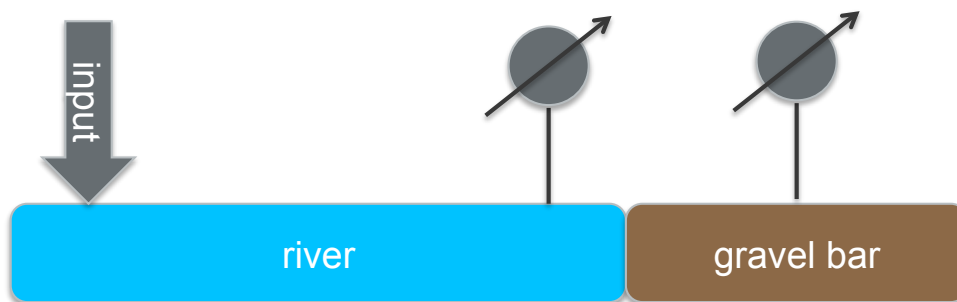


Figure 5.3: Schematic plot of the gravel-bar experiments. One fluorimeter is placed in front of the gravel bar in the river, and the other one is measuring the signal in a piezometer within the gravel-bar. Fluo and Raz solution is introduced as a pulse into the river, and their concentrations including Rru are measured by the fluorimeters.

## 5.3 Theory

From Section 5.2, it is clear that the column experiments in the lab and the gravel-bar experiments share the same governing equations, namely an input-output type of model as proposed in *Cirpka et al.* (2007) for the conservative tracer, which needs to be extended for the reactive compounds. The model for the in-stream tracer test is consists of the one-dimensional advection-dispersion equation (ADE) in the river channel combined with a convolution reflecting hyporheic exchange, as shown in Chapter 3 and 4. The governing equations for the conservative compound Fluo and the reactive tracer Raz are formally similar, whereas the model for Raz requires a cross-tracer term in the master equation. An overview of the governing equations of column and gravel-bar experiment can be found in the following:

$$\left\{ \begin{array}{l} \text{Fluo and Raz: } c_{i,out} = \int_0^t g_i(\tau) c_{i,in}(t - \tau) d\tau \\ \text{Rru: } c_{i,out}(t) = \int_0^t \left( g_{12}(\tau) c_{1,in}(t - \tau) + g_2(\tau) c_{2,in}(t - \tau) \right) d\tau \end{array} \right. \quad (5.1)$$

$$\left. \begin{array}{l} \text{Rru: } c_{i,out}(t) = \int_0^t \left( g_{12}(\tau) c_{1,in}(t - \tau) + g_2(\tau) c_{2,in}(t - \tau) \right) d\tau \end{array} \right. \quad (5.2)$$

Eq. 5.1 and Eq. 5.2 are the governing equations for all three tracers at two different scales. In the equations, I use the index 0 for Fluo, 1 for Raz, and 2 for Rru in the subscript.  $c_i[ML^{-3}]$  is the tracer concentration;  $t[T]$  is the time since start of injection; and  $g_i[T^{-1}]$  is the transfer functions for each tracer.

The basic governing equation is the Eq. 5.1, where the governing equations for Fluo and Raz are simply the convolution of transfer function  $g_i(\tau)$  with the input signal  $c_i(t)$ . The transfer function  $g_0$  for Fluo is identical to the travel-time distribution, which has an integral of 1, whereas the transfer function  $g_1$  for Raz may be integrated to a number smaller than unity because Raz may be transformed within the porous medium. The transfer function itself includes all possible linear processes in the subsurface, e.g. decay, adsorption, dispersion etc.. For Rru, the convolution of Eq. 5.2 is composed of two terms. The first term  $g_{12} * c_{1,in}$  comes from the transformation of Raz to Rru within the hyporheic zone, whereas the second term  $g_2 * c_{2,in}$  is related to Rru that already enters the system.

Eq. 5.1 and Eq. 5.2 are convolutions and can be solved as matrix-vector products after discretization. The key point is to obtain the transfer functions  $g_i$ . *Lemke et al.* (2013c) stated that adsorption of Raz and Rru onto the matrix of the porous medium should be considered in the model. Even though kinetic sorption is dominant, including both equilibrium and kinetic sorption together better describes the system. This is also why we have extended the existing model of Chapter 3 to a more complicated model accounting for several transformations and two-site sorption of Raz and Rru in Chapter 4.

As explained in Chapter 4, the transfer functions can be expressed as:

$$g_i(\tau) = \int_0^{\infty} g_0(\tau^*) c_i^{hz}(\tau^*, \tau) d\tau^*, \quad i = 0, 1, 2, 12 \quad (5.3)$$

in which the function  $g_0(\tau)$  is the travel-time distribution obtained from the conservative tracer, and  $c_i^{hz}[ML^{-3}]$  is the tracer concentration at time  $\tau$  and travel time  $\tau^*$  originating from a hypothetical pulse injection, which follows the transport equations Eqs.(4.9) and (4.10). The details of solving these transfer functions can be found in Section 4.2.3.

## 5.4 Results

### 5.4.1 Column experiment

A detailed description of the column experiments can be found in *Gritsch* (2013) and *Lemke et al.* (2013a). Figure 5.4a shows the result of a column experiment using the sediments from River Schaich. From top to bottom, there are figures for Fluo, Raz, and Rru, respectively. The black dotted lines are the input signals at the inlet of the column and the black lines are the measured output signals at the outlet. The duration of the tracer injection was about 7 hours. The colored lines are the simulation results of the models, all showing very good agreement with the measurements at the outlet of the column. In the top figure, representing Fluo, the plateau value of the output signal is identical to that of the input signal because Fluo behaves as a conservative tracer, whereas in the figure for Raz, we notice a decrease in the plateau concentration in the outlet, which is caused by the transformation to Rru and an unknown transformation product. The bottom figure, related to Rru, shows a clear increase in the plateau concentration in the output in comparison to the input.

The corresponding transfer functions are shown in Figure 5.4b. The green line is the transfer function of Fluo, which is also the travel-time distribution because Fluo is a conservative tracer. The shape looks like a Gaussian distribution, which indicates that a classical transport model most likely would also be appropriate for this data. The long tails in  $g_1$  and  $g_2$  come from kinetic sorption, where some solutes are released at later times from the sorbing sites. The cyan line is the transfer function of Raz transforms into Rru. The recovery rate, which is the zeroth moment of the transfer functions, and the mean travel time for all three tracers are listed in Table 5.1.

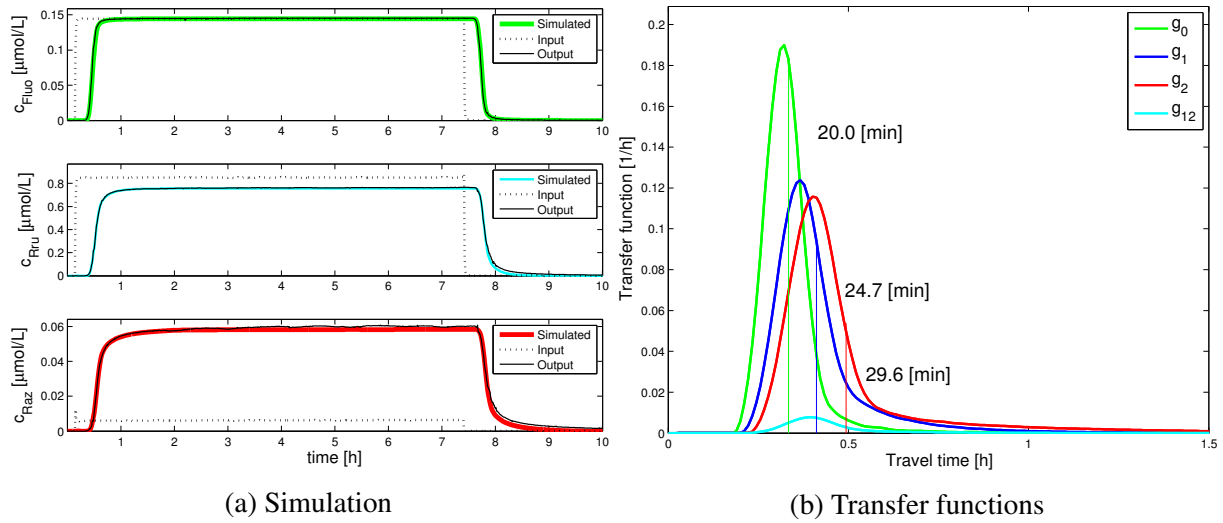


Figure 5.4: (a) Simulation of column experiment. Top: Fluo; middle: Raz; bottom: Rru. Dotted lines are input signals; black solid lines are measured output signals; and colored lines are the simulation results of the model. (b) Transfer function for the column experiment. Green: transfer function of Fluo; blue: transfer function of Raz; red: transfer function of Rru; cyan: transfer function from Raz to Rru.

## 5.4.2 Gravel-bar experiment

The experiment of the gravel bar is quite similar to the column experiment, only the column is replaced by the gravel bar in the field. Figure 5.5a shows the measurements and the corresponding simulation results of the three tracers. Again, the dotted lines are the signals in the river close to the gravel bar; the black solid lines are the measurements in the piezometer; and the colored lines are the model results. The ratio of observed mass in the piezometer to the input signal in the river for Fluo and Raz are close to 1, whereas the Rru has a slight increase to 1.2. This is caused by the transformation from Raz to Rru during the passage through the gravel bar. We have checked the zeroth moment of the transfer functions for the three tracers (see Table 5.1). The recovery of Fluo is constraint to 1, and the recovery rates for Raz and Rru are 0.88 and 0.92, respectively.

Figure 5.5b shows the estimated transfer functions. By applying the shape-free deconvolution method, we are able to obtain unconventional shapes of the transfer functions, for example the secondary peak showing in this figure. The green line is the travel-time distribution  $g_0$  of Fluo. I notice a small secondary peak appearing after 1 hour travel time. This peak might be a sign of different flow path through the gravel bar. The transfer function  $g_1$  for Raz and  $g_2$  for Rru have long and broad tails, which comes from kinetic sorption. More gravel-bar experiments have been reported by *Gritsch* (2013).

## 5.4.3 Comparison of the parameters

The main purpose of this chapter is to collect tracer tests performed at different scales and compare the estimated parameters. I only focus on the result from column and gravel-bar experiment, because the sediments are from the same site at River Schaich (see Table 5.2), whereas the stream tracer test is carried out in River Goldersbach. In general, the parameters estimated from different experiments provide similar



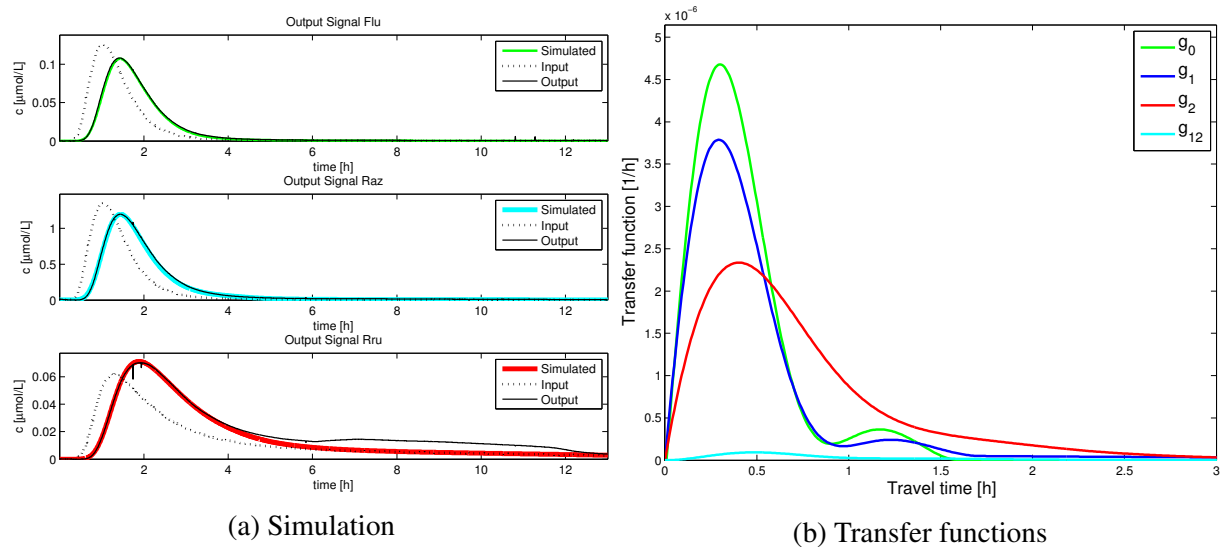


Figure 5.5: (a) Simulation of gravel-bar experiment. Top: Fluo; middle: Raz; bottom: Rru. Dotted lines are input signals; black solid lines are output signals; and colored lines are the simulation results of the model. (b) Transfer function of the gravel-bar experiment. Green: transfer function of Fluo; blue: transfer function of Raz; red: transfer function of Rru; cyan: transfer function from Raz to Rru.

Table 5.1: Recovery rate and mean travel time for column and gravel-bar experiment.

		Column	Gravel bar
Recovery rate [-]	Fluo	1	1
	Raz	0.89	0.88
	Rru	0.94	0.92
Mean travel time [min]	Fluo	20.0	20.6
	Raz	24.7	29.2
	Rru	29.6	44.4

values. From the distribution coefficient  $K_1$  and  $K_2$ , we notice that less tracer is sorbed at the kinetic sites in the column experiment than in the gravel-bar experiment. A potential reason could be that the sediment in the column was sieved and better sorted, and also the temperature was somewhat higher than in the field. The mass transfer coefficient  $k_1$  for Raz was clearly higher in the column experiment than in the gravel-bar experiment. The transformation of Raz to both an unknown compound and Rru was also faster for the column experiment according to the  $\lambda_1$  and  $\lambda_{12}$  values (compared to gravel-bar experiment).

Table 5.2: Estimated parameters from breakthrough curves of Fluo, Raz, and Rru from the column and gravel-bar experiments.  $R_1[-]$ : retardation factor for Raz;  $R_2[-]$ : retardation factor for Rru;  $K_1[-]$ : distribution coefficient for Raz;  $K_2[-]$ : distribution coefficient for Rru;  $k_1[1/s]$ : sorption mass transfer coefficient for Raz;  $k_2[1/s]$ : sorption mass transfer coefficient for Rru;  $\lambda_1[1/s]$ : decay coefficient of Raz to unknown compound;  $\lambda_{12}[1/s]$ : decay coefficient of Raz to Rru;  $\lambda_2[1/s]$ : decay coefficient of Rru to unknown compound.

Parameter	Column experiment	Gravel-bar experiment
$R_1[-]$	1.18	1.07
$R_2[-]$	1.31	1.47
$K_1[-]$	0.10	0.90
$K_2[-]$	0.25	0.46
$k_1[1/s]$	$1.9 \times 10^{-3}$	$9.2 \times 10^{-5}$
$k_2[1/s]$	$6.8 \times 10^{-4}$	$9.3 \times 10^{-4}$
$\lambda_1[1/s]$	$4.5 \times 10^{-5}$	$2.9 \times 10^{-5}$
$\lambda_{12}[1/s]$	$5.9 \times 10^{-5}$	$2.7 \times 10^{-5}$
$\lambda_2[1/s]$	$4.4 \times 10^{-5}$	$5.6 \times 10^{-5}$

## 5.5 Conclusions

In this chapter, I have presented tracer tests at different scales, from lab to the field, using conservative and reactive tracers, implementing the shape-free deconvolution method, and including two-site sorption in the model. The shape-free deconvolution reveals unconventional features of the transfer functions, and the simulations of the model agree very well with the observations. The results show a great success in applying this model to completely different experimental scales in terms of fitting the measurements. Even though the data from the same field site is missing for the stream tracer experiment, the application in River Goldersbach is shown in Chapter 4. All together, successfully using the same mathematical expressions for transport processes at different scales gives hope that upscaling of hyporheic reactive transport is possible. I observed that the sorption and decay processes are stronger and faster in the lab experiment than in the field, which may be affected by temperature effects and the pretreatment of the sediments in the lab experiments. Of course a parameter comparison of tracer tests from the same field site would have been better, but the gravel-bar experiments in River Goldersbach failed, and the breakthrough curves of the stream-tracer experiments in River Schaich showed deficiencies as well.

# Chapter 6

## Conclusions and Outlook

In Chapter 1, I have posed four research questions addressed in Chapters 2 to 5, thus contributing to better understanding river-groundwater exchange when analyzing natural tracer time series and artificial tracer tests.

- Question 1: How can the approach of *Cirpka et al. (2007)* for estimating the river-to-groundwater travel-time distribution be extended from stationary to non-stationary conditions, and how well can dynamic changes of the system be characterized by analyzing natural tracer time series using the extended approach?

Chapter 2 demonstrated a potential way to extend the shape-free deconvolution method from stationary to non-stationary conditions by combining it with a sliding window technique, where the long time series is windowed into small sections for which local stationary conditions are assumed. By sliding the window along time, a set of parameters is estimated, which requires reasonable smoothness constraints to prevent unrealistic variations caused by noise in the data. At the field site at River Thur, Switzerland, the stationary method provides a decent estimation of river-to-groundwater travel times, but the non-stationary model improved the model fitting, depicted changes of the system revealing the influence of the river-water stage on the groundwater observation wells.

- Question 2: What is a suitable methodology of estimating the travel-time distribution of river water in the hyporheic zone? Is it possible to determine unconventional shapes of the hyporheic travel-time distributions?

Chapter 3 proposed a methodology to estimate the transfer function of solutes entering the hyporheic zone and returning back to the stream, separating the spread in the observed BTC caused by in-stream dispersion and from that originating from hyporheic exchange. This method is tested by a synthetic tracer experiments with a conservative tracer and a single reactive tracer, considering first-order decay of the reactive tracer in the hyporheic zone. The method is able to reveal unconventional shapes of the hyporheic travel-time distribution because it is nonparametric.

- Question 3: Can this method be applied to field measurements? How can possible reaction and adsorption processes of tracers in the hyporheic zone be addressed in this framework?

Chapter 4 showed how the model is applied to a field data set. Substantiate reasons support that two-site sorption should be considered in the model, and the joint analysis of a conservative tracer,

a reactive tracer and its reaction product requires a more complicated formulation that accounts for three first-order reactions and compound-specific sorption parameters. The method facilitates jointly analyzing the breakthrough curves of three tracers at multiple observation points. The model shows excellent agreement to the observations, including the tailing behavior.

- Question 4: Is this method ready for upscaling? How does the model perform for tracer tests at different scales?

Chapter 5 demonstrated that the proposed method not only works very well in the analysis of stream tracer tests at different observation distances, but also provides decent results for column experiments and gravel-bar experiments. This corroborates that the method is suitable for upscaling and has a potential to be used at multiple scales.

In general, this work has extended an existing non-parametric deconvolution method to non-stationary conditions, adopted it to the stream-tracer tests, proposed a method to jointly use tracer data of a conservative tracer, a reactive tracer compound, and its degradation product, including sorption and multiple decay processes in the hyporheic zone, and tested it with experiments at different scales. The series of models proposed here are applicable for bank-filtration and stream tracer tests, providing shape-free travel-time distributions in all applications. This method may contribute to understanding surface water-ground water interactions by shading light into transient storage.

Despite first successful application of the methods to field data, a series of difficulties and open questions remain:

- Robustness of the estimation is one of the biggest issues of this method. The method, as presented, relies on gradient based minimization, which requires the absence of local minima in the objective function and a good initial guess. *Lemke et al. (2013a)* and *Lemke et al. (2013c)* applied a Markov-Chain Monte Carlo method to find the global minimum of parameters in a parametric version of the estimation problem, which improves the reliability and stability of parameter estimation. However, due to the large number of parameters in the shape-free deconvolution method, a Markov-Chain Monte Carlo approach appears unfeasible. Therefore, a good and reasonable initial guess appears paramount. A better way of getting the initial guess or a method to find the global minimum need to be investigated.
- Another problem arises from over-parameterization. The shape-free deconvolution method overcomes the problem of choosing 'the right' parametric function, making it possible to reveal unconventional shapes of the transfer functions. However, this comes at the price that the transfer function is discretized into a large set of values in the travel time. For regularization we need a smoothness constraint, the value of which also needs to be estimated. With this large number of parameters we are able to fit broad peaks and long tails in a decent way, but whether the estimation of so many parameters is worthwhile, compared to the improvement in the simulation, has not yet been discussed, and rigorous methods on the data worth are difficult to derive.
- Computational costs are also relevant. In the previous point, I mentioned the smoothness constraint, the value of which is estimated by the Expectation-Maximization method (*McLachlan and Krishnan, 1997*). This method requires many sets of conditional realizations and the convergence of the method is slow. Besides that, the sensitivity calculation also consumes a lot of computer time, depending on the number of parameters and measurements. Thus, approaches to speed up the overall method would be very useful.

- Last but not the least, I have not developed a non-stationary method for the estimation of hyporheic travel times (see the missing item in Table 1.1). For this purpose, performing a pulse-injection experiment might not be suitable anymore, because of the short duration of such an experiment. A natural tracer that is indicative of hyporheic exchange will be needed, but it's not yet clear which tracer this could be.



# Bibliography

- Alvarado, J. A. C., R. Purtschert, K. Hinsby, L. Troldborg, M. Hofer, R. Kipfer, W. Aeschbach-Hertig, and H. Arno-Synal (2005),  $^{36}\text{Cl}$  in Modern Groundwater Dated by a Multi-Tracer Approach ( $^3\text{H}/^3\text{He}$ ,  $\text{SF}_6$ , CFC-12 and  $^{85}\text{Kr}$ ): A Case Study in Quaternary Sand Aquifers in the Odense Pilot River Basin, Denmark, *Appl. Geochem.*, 20(3), 599–609, doi:10.1016/j.apgeochem.2004.09.018.
- Anderson, M. P. (2005), Heat as a ground water tracer, *Ground Water*, 43(6), 951–968.
- Argerich, A., R. Haggerty, E. Martí, F. Sabater, and J. Zarnetske (2011), Quantification of metabolically active transient storage (mats) in two reaches with contrasting transient storage and ecosystem respiration, *J. Geophys. Res.*, 116(G3), G03,034, doi:10.1029/2010JG001379.
- Bencala, K. E., and R. A. Walters (1983), Simulation of solute transport in a mountain pool-and-riffle stream: a transient storage model, *Water Resour. Res.*, 19(3), 718–724, doi:10.1029/WR019i003p00718.
- Berndt, D. J., and J. Clifford (1994), Using dynamic time warping to find patterns in time series., in *KDD workshop*, vol. 10, pp. 359–370, Seattle, WA.
- Beyerle, U., W. Aeschbach-Hertig, M. Hofer, D. Imboden, H. Baur, and R. Kipfer (1999), Infiltration of river water to a shallow aquifer investigated with  $\text{H-3}/\text{He-3}$ , noble gases and CFCs, *J. Hydrol.*, 220(3-4), 169–185, doi:10.1016/S0022-1694(99)00069-4.
- Bianchin, M., L. Smith, and R. Beckie (2010), Quantifying hyporheic exchange in a tidal river using temperature time series, *Water Resources Research*, 46(7).
- Boano, F., A. Packman, A. Cortis, R. Revelli, and L. Ridolfi (2007), A Continuous Time Random Walk approach to the stream transport of solutes, *Water Resour. Res.*, 43(10), 10,425, doi:10.1029/2007WR006062.
- Boano, F., A. Demaria, R. Revelli, and L. Ridolfi (2010a), Biogeochemical zonation due to intrameander hyporheic flow, *Water Resour. Res.*, 46(2), W02,511, doi:10.1029/2008WR007583.
- Boano, F., R. Revelli, and L. Ridolfi (2010b), Effect of streamflow stochasticity on bedform-driven hyporheic exchange, *Adv. Water Resour.*, 33(11), 1367–1374, doi:10.1016/j.advwatres.2010.03.005.
- Boker, S. M., J. L. Rotondo, M. Xu, and K. King (2002), Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series., *Psychological Methods*, 7(3), 338.

- Boulton, A. J., S. Findlay, P. Marmonier, E. H. Stanley, and H. M. Valett (1998), The functional significance of the hyporheic zone in streams and rivers, *Annual Review of Ecology and Systematics*, 29(1), 59–81.
- Briggs, M. A., M. N. Gooseff, C. D. Arp, and M. A. Baker (2009), A method for estimating surface transient storage parameters for streams with concurrent hyporheic storage, *Water Resour. Res.*, 45(2), W00D27, doi:10.1029/2008WR006959.
- Briggs, M. A., M. N. Gooseff, B. J. Peterson, K. Morkeski, W. M. Wollheim, and C. S. Hopkinson (2010), Surface and hyporheic transient storage dynamics throughout a coastal stream network, *Water Resour. Res.*, 46(6), W06,516, doi:10.1029/2009WR008222.
- Brunke, M., and T. Gonser (1997), The ecological significance of exchange processes between rivers and groundwater, *Freshwater Biol.*, 37(null), 1–33, doi:10.1046/j.1365-2427.1997.00143.x.
- BUWAL (2004), *Wegleitung Grundwasserschutz*, Vollzug Umwelt, Bundesamt für Umwelt, Wald und Landschaft, Bern, Switzerland.
- Camporese, M., D. Penna, M. Borga, and C. Paniconi (2014), A field and modeling study of nonlinear storage-discharge dynamics for an alpine headwater catchment, *Water Resources Research*.
- Cardenas, M. B. (2007), Potential contribution of topography-driven regional groundwater flow to fractal stream chemistry: Residence time distribution analysis of Tóth flow, *Geophys. Res. Lett.*, 34(5), L05,403.
- Cardenas, M. B. (2008a), The effect of river bend morphology on flow and timescales of surface water-groundwater exchange across pointbars, *J. Hydrol.*, 362(1-2), 134–141, doi:10.1016/j.jhydrol.2008.08.018.
- Cardenas, M. B. (2008b), Surface water-groundwater interface geomorphology leads to scaling of residence times, *Geophys. Res. Lett.*, 35(8), L08,402, doi:10.1029/2008GL033753.
- Cardenas, M. B., and J. L. Wilson (2007), Exchange across a sediment-water interface with ambient groundwater discharge, *J. Hydrol.*, 346(null), 69–80, doi:10.1016/j.jhydrol.2007.08.019.
- Cardenas, M. B., J. L. Wilson, and V. A. Zlotnik (2004), Impact of Heterogeneity, Bed Forms, and Stream Curvature on Subchannel Hyporheic Exchange, *Water Resour. Res.*, 40(8), W08,307.
- Cardenas, M. B., J. L. Wilson, and R. Haggerty (2008), Residence time of bedform-driven hyporheic exchange, *Advances Water Res.*, 31(10), 1382–1386, doi:10.1016/j.advwatres.2008.07.006.
- Carrera, J., X. Sanchez-Vila, I. Benet, A. Medina, G. Galarza, and J. Guimera (1998), On matrix diffusion: formulations, solution methods and qualitative effects, *Hydrogeol. J.*, 6(1), 178–190, doi:10.1007/s100400050143.
- Choi, J., J. W. Harvey, and M. H. Conklin (2000), Characterizing multiple timescales of stream and storage zone interaction that affect solute fate and transport in streams, *Water Resources Research*, 36(6), 1511–1518.



- Cirpka, O. A., M. N. Fienen, M. Hofer, E. Hoehn, A. Tessarini, R. Kipfer, and P. K. Kitanidis (2007), Analyzing bank filtration by deconvoluting time series of electric conductivity, *Ground Water*, 45(3), 318–328, doi:10.1111/j.1745-6584.2006.00293.x.
- Constantz, J. (2008), Heat as a tracer to determine streambed water exchanges, *Water Resour. Res.*, 44(12), W00D10.
- Constantz, J., M. H. Cox, and G. W. Su (2003), Comparison of heat and bromide as ground water tracers near streams, *Ground Water*, 41, 647–656.
- Coscia, I., S. A. Greenhalgh, N. Linde, J. Doetsch, L. Marescot, T. Guenther, T. Vogt, and A. G. Green (2011), 3-D crosshole ERT for aquifer characterization and monitoring of infiltrating river water, *Geophysics*, 76(2), G49–G59, doi:10.1190/1.3553003.
- Covino, T., B. McGlynn, and J. Mallard (2011), Stream-groundwater exchange and hydrologic turnover at the network scale, *Water Resources Research*, 47(12), W12,521, doi:10.1029/2011WR010942.
- Cox, M., G. Su, and J. Constantz (2007), Heat, chloride, and specific conductance as ground water tracers near streams, *Groundwater*, 45(2), 187–195.
- Darling, W., D. Gooddy, J. Riches, and I. Wallis (2010), Using environmental tracers to assess the extent of river-groundwater interaction in a quarried area of the english chalk, *Appl. Geochem.*, 25(7), 923–932.
- Davis, S. N., G. M. Thompson, H. W. Bentley, and G. Stiles (1980), Ground-water tracers - a short review, *Ground Water*, 18(1), 14–23.
- De Hoog, F. R., J. H. Knight, and A. N. Stokes (1982), An improved method for numerical inversion of Laplace transforms, *SIAM J. Sci. Stat. Comput.*, 3, 357–366, doi:10.1137/0903022.
- Deng, Z., and H. Jung (2009), Variable residence time-based model for solute transport in streams, *Water Resour. Res.*, 45(3), W03,415, doi:10.1029/2008WR007000.
- Dennery, P., and A. Krzywicki (1967), *Mathematics for Physicists*, Harper's Physics Series, Harper and Row, New York.
- Dentz, M., A. Cortis, H. Scher, and B. Berkowitz (2004), Time behavior of solute transport in heterogeneous media: Transition from anomalous to normal transport, *Advances in Water Resources*, 27(2), 155–173.
- Doussan, C., G. Poitevin, E. Ledoux, and M. Detay (1997), River bank filtration: Modelling of the changes in water chemistry with emphasis on nitrogen species, *J. Contam. Hydrol.*, 25(1-2), 129–156, doi:10.1016/S0169-7722(96)00024-1.
- Drummond, J. D., T. P. Covino, A. F. Aubeneau, D. Leong, S. Patil, R. Schumer, and A. I. Packman (2012), Effects of solute breakthrough curve tail truncation on residence time estimates: A synthesis of solute tracer injection studies, *J. Geophys. Res.*, 117, G00N08, doi:10.1029/2012JG002019.
- Dugge, J. (2012), Case Study of Distributed Continuous Hydrological Modelling Forced with Different Precipitation Products, Master's thesis, Swiss Federal Institute of Technology Zurich.

- Dürrenmatt, D., D. Del Giudice, and J. Rieckermann (2013), Dynamic time warping improves sewer flow monitoring, *Water research*.
- DVGW (1995), Richtlinie für Trinkwasserschutzgebiete, 1. Teil: Schutzgebiete für Grundwasser, *Richtlinie W 101*, Deutscher Verein des Gas- und Wasserfachs, Bonn, Germany.
- Feng, Z.-G., and E. E. Michaelides (2009), Secondary flow within a river bed and contaminant transport, *Environ. Fluid Mech.*, 9(6), 617–634, doi:10.1007/s10652-009-9132-9.
- Findlay, S. (1995), Importance of surface-subsurface exchange in stream ecosystems: The hyporheic zone, *Limnol. Oceanog.*, pp. 159–164.
- Fischer, H. B. (1967), The mechanics of dispersion in natural streams, *J. Hydraul. Div. ASCE*, 93(6), 187–216.
- Fischer, H. B. (1979), *Mixing in inland and coastal waters*, Academic Press.
- González-Pinzón, R., R. Haggerty, and D. D. Myrold (2012), Measuring aerobic respiration in stream ecosystems using the resazurin-resorufin system, *J. Geophys. Res.*, 117(null), G00N06, doi: 10.1029/2012JG001965.
- Gooseff, M., K. Bencala, D. Scott, R. Runkel, and D. McKnight (2005a), Sensitivity analysis of conservative and reactive stream transient storage models applied to field data from multiple-reach experiments, *Advances Water Res.*, 28(5), 479–492.
- Gooseff, M. N., S. M. Wondzell, R. Haggerty, and J. Anderson (2003), Comparing transient storage modeling and residence time distribution (RTD) analysis in geomorphically varied reaches in the Lookout Creek basin, Oregon, USA, *Advances Water Resour.*, 26(9), 925–937, doi:10.1016/S0309-1708(03)00105-2.
- Gooseff, M. N., D. M. McKnight, R. L. Runkel, and J. H. Duff (2004), Denitrification and hydrologic transient storage in a glacial meltwater stream, mcmurdo dry valleys, antarctica, *Limnology and Oceanography*, 49(5), 1884–1895.
- Gooseff, M. N., J. LaNier, R. Haggerty, and K. Kokkeler (2005b), Determining in-channel (dead zone) transient storage by comparing solute transport in a bedrock channel-alluvial channel sequence, Oregon, *Water Resour. Res.*, 41(6), W06,014.
- Gooseff, M. N., J. K. Anderson, S. M. Wondzell, J. LaNier, and R. Haggerty (2006), A modeling study of hyporheic exchange pattern and the sequence, size, and spacing of stream bedforms in mountain stream networks, oregon, usa, *Hydrol. Process.*, 20(11), 2443–2457, doi:10.1002/hyp.5790.
- Gooseff, M. N., D. A. Benson, M. A. Briggs, M. Weaver, W. Wollheim, B. Peterson, and C. S. Hopkinson (2011), Residence time distributions in surface transient storage zones in streams: Estimation via signal deconvolution, *Water Resour. Res.*, 47(5), W05,509, doi:10.1029/2010WR009959.
- Gritsch, M. (2013), The resazurin smart tracer system in gravelbars and columns, Master's thesis, University of Tübingen.

- Haggerty, R., and S. M. Gorelick (1995), Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity, *Water Resour. Res.*, 31(10), 2383–2400, doi:10.1029/95WR10583.
- Haggerty, R., S. A. McKenna, and L. C. Meigs (2000), On the late-time behavior of tracer test breakthrough curves, *Water Resour. Res.*, 36(12), 3467–3479, doi:10.1029/2000WR900214.
- Haggerty, R., S. M. Wondzell, and M. A. Johnson (2002), Power-law residence time distribution in the hyporheic zone of a 2nd-order mountain stream, *Geophys. Res. Lett.*, 29(13), 1640, doi:10.1029/2002GL014743.
- Haggerty, R., A. Argerich, and E. Martí (2008), Development of a smart tracer for the assessment of microbiological activity and sediment-water interaction in natural waters: The resazurin-resorufin system, *Water Resour. Res.*, 44, W00D01, doi:10.1029/2007WR006670.
- Haggerty, R., E. Martí, A. Argerich, D. von Schiller, and N. Grimm (2009), Resazurin as a "smart" tracer for quantifying metabolically active transient storage in stream ecosystems, *J. Geophys. Res. - Biogeosci.*, 114, G03,014, doi:10.1029/2008JG000942.
- Harvey, J., B. Wagner, and K. Bencala (1996), Evaluating the reliability of the stream tracer approach to characterize stream-subsurface water exchange, *Water Resour. Res.*, 32(8), 2441–2451, doi:10.1029/96WR01268.
- Harvey, J., J. Saiers, and J. Newlin (2005), Solute transport and storage mechanisms in wetlands of the everglades, south florida, *Water Resour. Res.*, 41(5), W05,009, doi:10.1029/2004WR003507.
- Hatch, C. E., A. T. Fisher, J. S. Revenaugh, J. Constantz, and C. Ruehl (2006), Quantifying surface water-groundwater interactions using time series analysis of streambed thermal records: Method development, *Water Resources Research*, 42(10).
- Hester, E. T., and M. N. Gooseff (2010), Moving beyond the banks: Hyporheic restoration is fundamental to restoring ecological services and functions of streams, *Environ. Sci. Tech.*, 44(5), 1521–1525.
- Hester, E. T., K. I. Young, and M. A. Widdowson (2013), Mixing of surface and groundwater induced by riverbed dunes: Implications for hyporheic zone definitions and pollutant reactions, *Water Resources Research*, 49(9), 5221–5237.
- Hoehn, E., and O. Cirpka (2006), Assessing hyporheic zone dynamics in two alluvial flood plains of the southern alps using water temperature and tracers, *Hydrol. Earth Syst. Sci.*, 10(4), 553–563.
- Hoehn, E., and H. von Gunten (1989), Radon in groundwater - a tool to assess infiltration from surface waters to aquifers, *Water Resour. Res.*, 25(8), 1795–1803, doi:10.1029/WR025i008p01795.
- Hoppe-Jones, C., G. Oldham, and J. E. Drewes (2010), Attenuation of total organic carbon and unregulated trace organic chemicals in us riverbank filtration systems, *Water Res.*, 44(15), 4643–4659, doi:10.1016/j.watres.2010.06.022.
- Hunt, R., T. Coplen, N. Haas, D. Saad, and M. Borchardt (2005), Investigating surface water-well interaction using stable isotope ratios of water, *J. Hydrol.*, 302, 154–172.

- Jonsson, K., H. Johansson, and A. Wörman (2004), Sorption behavior and long-term retention of reactive solutes in the hyporheic zone of streams, *J. Environ. Eng.*, 130, 573, doi:10.1061/(ASCE)0733-9372(2004)130:5(573).
- Journel, A. J., and C. J. Huijbregts (1978), *Mining Geostatistics*, Academic Press, London.
- Keery, J., A. Binley, N. Crook, and J. Smith (2007), Temporal and spatial variability of groundwater-surface water fluxes: Development and application of an analytical method using temperature time series, *J. Hydrol.*, 336(1-2), 1–16.
- Keogh, E., S. Chu, D. Hart, and M. Pazzani (2001), An online algorithm for segmenting time series, in *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pp. 289–296, IEEE.
- Kitanidis, P. K. (1997), The minimum structure solution to the inverse problem, *Water Resour. Res.*, 33(10), 2263–2272, doi:10.1029/97WR01619.
- Krause, S., D. Hannah, J. Fleckenstein, C. Heppell, D. Kaeser, R. Pickup, G. Pinay, A. Robertson, and P. Wood (2011), Inter-disciplinary perspectives on processes in the hyporheic zone, *Ecohydrology*, 4(4), 481–499.
- Krause, S., F. Boano, M. O. Cuthbert, J. H. Fleckenstein, and J. Lewandowski (2014), Understanding process dynamics at aquifer-surface water interfaces: An introduction to the special section on new modelling approaches and novel experimental technologies, *Water Resources Research*.
- Kuehn, W., and U. Mueller (2000), Riverbank filtration - an overview, *J. Am. Water Works Ass.*, 92(12), 60–69.
- Lagarias, J., J. Reeds, M. Wright, and P. Wright (1999), Convergence properties of the nelder-mead simplex method in low dimensions, *SIAM Journal on Optimization*, 9(1), 112–147.
- Leibundgut, C., P. Maloszewski, and C. Külls (2009), Tracers in hydrology.
- Lemke, D., Z. Liao, K. Osenbrück, and O. A. Cirpka (2013a), Concurrent conservative and reactive tracer tests in a stream undergoing hyporheic exchange, *Water Resour. Res.*, (submitted).
- Lemke, D., P.-A. Schnegg, M. Schwientek, K. Osenbrück, and O. A. Cirpka (2013b), On-line fluorometry of multiple reactive and conservative tracers in streams, *Environ. Earth Sci.*, (in press), doi: 10.1007/s12665-013-2305-3.
- Lemke, D., R. González-Pinzón, Z. Liao, T. Wöhling, K. Osenbrück, R. Haggerty, and O. Cirpka (2013c), Sorption and transformation of the reactive tracers resazurin and resorufin in natural river sediments., *Hydrology & Earth System Sciences Discussions*, 10(10).
- Liao, Z., and O. A. Cirpka (2011), Shape-free inference of hyporheic travel-time distributions from synthetic conservative and "smart" tracer tests in streams, *Water Resour. Res.*, 47(7), W07,510, doi: 10.1029/2010WR009927.
- Liao, Z., D. Lemke, K. Osenbrück, and O. A. Cirpka (2013), Modeling and inverting reactive stream tracers undergoing two-site sorption and decay in the hyporheic zone, *Water Resources Research*.

- Lin, A., J. Debroux, J. Cunningham, and M. Reinhard (2003), Comparison of rhodamine WT and bromide in the determination of hydraulic characteristics of constructed wetlands, *Ecol. Eng.*, 20(1), 75–88.
- Luo, J., O. A. Cirpka, M. N. Fienen, W.-m. Wu, T. L. Mehlhorn, J. Carley, P. M. Jardine, C. S. Criddle, and P. K. Kitanidis (2006), A parametric transfer function methodology for analyzing reactive transport in nonuniform flow, *Journal of contaminant hydrology*, 83(1), 27–41.
- Malaguerra, F., H. Albrechtsen, and P. Binning (2013), Assessment of the contamination of drinking water supply wells by pesticides from surface water resources using a finite element reactive transport model and global sensitivity analysis techniques, *J. Hydrol.*, 476, 321–331.
- Maloszewski, P., and A. Zuber (1993), Principles and practice of calibration and validation of mathematical models for the interpretation of environmental isotopes in aquifers, *Adv. Water Resour.*, 16(3), 173–190.
- Marion, A., M. Zaramella, and A. Bottacin-Busolin (2008), Solute transport in rivers with multiple storage zones: The STIR model, *Water Resour. Res.*, 44(10), W10,406, doi:10.1029/2008WR007037.
- Massmann, G., J. Sültenfuß, U. Dünnebier, A. Knappe, T. Taute, and A. Pekdeger (2008), Investigation of groundwater residence times during bank filtration in Berlin: A multi-tracer approach, *Hydrol. Proc.*, 22(6), 788–801.
- McCallum, J. L., N. B. Engdahl, T. R. Ginn, P. Cook, et al. (2014), Nonparametric estimation of groundwater residence time distributions: What can environmental tracer data tell us about groundwater residence time?, *Water Resources Research*, 50(3), 2022–2038.
- McLachlan, G., and T. Krishnan (1997), *The EM algorithm and extensions*, Wiley, New York.
- Michalak, A., and P. Kitanidis (2003), A method for enforcing parameter nonnegativity in bayesian inverse problems with an application to contaminant source identification, *Water Resour. Res.*, 39(2), 1033.
- Nash, J., and J. Sutcliffe (1970), River flow forecasting through conceptual models part ia discussion of principles, *Journal of hydrology*, 10(3), 282–290.
- Payn, R. A., M. N. Gooseff, D. A. Benson, O. A. Cirpka, J. P. Zarnetske, W. B. Bowden, J. McNamara, and J. H. Bradford (2008), Comparison of Instantaneous and Constant-Rate Stream Tracer Experiments Through Non-Parametric Analysis of Residence Time Distributions, *Water Resour. Res.*, 44(6), W06,404.
- Payn, R. A., M. N. Gooseff, B. L. McGlynn, K. E. Bencala, and S. M. Wondzell (2009), Channel water balance and exchange with subsurface flow along a mountain headwater stream in montana, united states, *Water RESour. Res.*, 45, W11,427, doi:10.1029/2008WR007644.
- Quindoz, H., and A. J. Valocchi (1993), Stochastic Analysis of the Transport of Kinetically Sorbing Solutes in Aquifers with Randomly Heterogeneous Hydraulic Conductivity, *Water Resour. Res.*, 29(9), 3227–3240.
- Ray, C., T. Grischek, J. Schubert, J. Wang, and T. Speth (2002), A perspective of riverbank filtration, *J. Am. Water Works Ass.*, 94(4), 149–160.

- Revelli, R., F. Boano, C. Camporeale, and L. Ridolfi (2008), Intra-meander hyporheic flow in alluvial rivers, *Water Resour. Res.*, 44(12), W12,428, doi:10.1029/2008WR007081.
- Robertson, A. L., and P. J. Wood (2010), Ecology of the hyporheic zone: origins, current knowledge and future directions, *Fundam. Appl. Limnol., Arch. Hydrobiol.*, 176(4), 279–289, doi:10.1127/1863-9135/2010/0176-0279.
- Sawyer, A. H., M. B. Cardenas, and J. Buttles (2011), Hyporheic exchange due to channel-spanning logs, *Water Resour. Res.*, 47(08), doi:10.1029/2011WR010484.
- Schmidt, C., A. Musolff, N. Trauth, M. Vieweg, and J. Fleckenstein (2012), Transient analysis of fluctuations of electrical conductivity as tracer in the streambed, *Hydrology and Earth System Sciences Discussions*, 9(5), 6345–6365.
- Schneider, P., T. Vogt, M. Schirmer, J. Doetsch, N. Linde, N. Pasquale, P. Perona, and O. A. Cirpka (2011), Towards improved instrumentation for assessing river-groundwater interactions in a restored river corridor, *Hydrol. Earth Sys. Sci.*, 15(8), 2531–2549, doi:10.5194/hess-15-2531-2011.
- Schubert, J. (2002), Hydraulic aspects of riverbank filtration - field studies, *J. Hydrol.*, 266(3-4), 145–161, doi:10.1016/S0022-1694(02)00159-2.
- Sheets, R., R. Darner, and B. Whitteberry (2002), Lag times of bank filtration at a well field, Cincinnati, Ohio, USA, *J. Hydrol.*, 266(3-4), 162–174.
- Shope, C. L., J. E. Constantz, C. A. Cooper, D. M. Reeves, G. Pohl, and W. A. McKay (2012), Influence of a large fluvial island, streambed, and stream bank on surface water-groundwater fluxes and water table dynamics, *Water Resour. Res.*, 48(6), W06,512, doi:10.1029/2011WR011564.
- Smith, S. A., and W. A. Pretorius (2002), The conservative behaviour of fluorescein, *Water S. Africa Pretoria*, 28(4), 403–406.
- Sontheimer, H. (1980), Experience with riverbank filtration along the Rhine River, *J. Am. Water Works Ass.*, 72(7), 386–390.
- Sophocleous, M. (2002), Interactions between groundwater and surface water: the state of the science, *Hydrogeol. J.*, 10(1), 52–67, doi:10.1007/s10040-001-0170-8.
- Stichler, W., P. Maloszewski, and H. Moser (1986), Modelling of river water infiltration using oxygen-18 data, *J. Hydrol.*, 83, 355–365.
- Stute, M., J. Deák, K. Révész, J. Böhlke, E. Deseö, R. Weppernig, and P. Schlosser (1997), Tritium/<sup>3</sup>He dating of river infiltration: An example from the danube in the szigetköz area, hungary, *Groundwater*, 35(5), 905–911.
- Tetzlaff, D., C. Birkel, J. Dick, J. Geris, and C. Soulsby (2014), Storage dynamics in hydrogeological units control hillslope connectivity, runoff generation, and the evolution of catchment transit time distributions, *Water Resources Research*, 50(2), 969–985.

- Toride, N., F. J. Leij, and M. T. Genuchten (1993), A comprehensive set of analytical solutions for nonequilibrium solute transport with first-order decay and zero-order production, *Water Resour. Res.*, 29(7), 2167–2182.
- Tukey, J. W. (1967), An introduction to the calculations of numerical spectrum analysis, *Spectral Analysis of Time Series*, pp. 25–46.
- Varni, M., and J. Carrera (1998), Simulation of groundwater age distributions, *Water Resour. Res.*, 34(12), 3271–3281.
- Vogel, C. (2002), *Computational methods for inverse problems, volume 23 of Frontiers in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.
- Vogt, T., E. Hoehn, P. Schneider, and O. Cirpka (2009), Untersuchung der Flusswasserinfiltration in voralpinen Schottern mittels Zeitreihenanalyse, *Grundwasser*, 14(3), 179–194.
- Vogt, T., E. Hoehn, P. Schneider, A. Freund, M. Schirmer, and O. A. Cirpka (2010), Fluctuations of electrical conductivity as a natural tracer for bank filtration in a losing stream, *Advances in water resources*, 33(11), 1296–1308.
- Ward, A. S., M. N. Gooseff, and K. Singha (2010), Characterizing hyporheic transport processes - interpretation of electrical geophysical data in coupled stream-hyporheic zone systems during solute tracer studies, *Advance Water Resor.*, 33(11), 1320–1330, doi:10.1016/j.advwatres.2010.05.008.
- Weiss, W., E. Bouwer, R. Aboytes, M. LeChevallier, C. O’Melia, B. Le, and K. Schwab (2005a), River filtration for control of microorganisms results from field monitoring, *Water Res.*, 39(10), 1990–2001, doi:10.1016/j.watres.2005.03.018.
- Weiss, W. J., E. J. Bouwer, R. Aboytes, M. W. LeChevallier, C. R. O’Melia, B. T. Le, and K. J. Schwab (2005b), Riverbank filtration for control of microorganisms: Results from field monitoring, *Water research*, 39(10), 1990–2001.
- Wörman, A. (2000), Comparison of models for transient storage of solutes in small streams, *Water Resour. Res.*, 36(2), 455–468.
- Wörman, A., and P. Wachniew (2007), Reach scale and evaluation methods as limitations for transient storage properties in streams and rivers, *Water Resour. Res.*, 43(10), 10,405.
- Wörman, A., A. Packman, H. Johansson, and K. Jonsson (2002), Effect of flow-induced exchange in hyporheic zones on longitudinal transport of solutes in streams and rivers, *Water Resour. Res.*, 38(1), 1001, doi:10.1029/2001WR000769.
- Young, P. C., D. J. Pedregal, and W. Tych (1999), Dynamic harmonic regression, *Journal of forecasting*, 18(6), 369–394.
- Zhou, T., and T. A. Endreny (2013), Reshaping of the hyporheic zone beneath river restoration structures: Flume and hydrodynamic experiments, *Water Resources Research*, 49(8), 5009–5020.