INAUGURAL – DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich–Mathematischen Gesamtfakultät

der

Ruprecht – Karls – Universität

Heidelberg

vorgelegt von

Diplom-Mathematikerin Angelika Dieses

aus Neustadt/Weinstraße

Tag der mündlichen Prüfung: 15. Dezember 2000

Numerical Methods for Optimization Problems in Water Flow and Reactive Solute Transport Processes of Xenobiotics in Soils

Gutachter: **Prof. Dr. Hans Georg Bock**

Prof. Dr. Otto Richter

Acknowledgements

First and foremost, I wish to thank my advisors Prof. Dr. Hans Georg Bock, Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, and Prof. Dr. Otto Richter, Institute of Geoecology, Technical University of Braunschweig, for their interest in planning and supervising this interdisciplinary project and for their continuous support during the last three years. In addition, I am greatly indebted to Dr. Johannes Schlöder, IWR, for many inspiring and fruitful discussions and for the excellent supervision he gave me all the time.

I wish to thank my colleagues in the group of Prof. Dr. Hans Georg Bock and Dr. Johannes Schlöder for their friendship and their assistance. Without the advanced software and the enormous knowledge base available in the group, the development of the tools ECOFIT and ECOPLAN would not have been possible in such a short period of time. In particular, I am indebted to Dr. Irene Bauer and Stefan Körkel for providing the software packages DAESOL and VPLAN and for various discussions about integration methods and optimal experimental design. I also thank my colleague and roommate Oliver Großhans for always having an open ear for my questions and for his friendly help with diverse soft- and hardware problems. In this context, I also want to mention the excellent environment and atmosphere of the IWR for doing interdisciplinary research.

I also would like to thank Prof. Dr. Otto Richter and all his co-workers, in particular Dr. Dagmar Söndgerath, Dr. Ralf Seppelt and Michael Flake, for their warm welcome during my four months stay at their institute and for the introduction to the secrets of soil science and environmental fate modeling of xenobiotics.

Special thanks go to Dr. Anna Schreieck, Dr. Marite de Mehr and Prof. Dr. Josef Kallrath, BASF AG Ludwigshafen, and Dr. Bernhard Gottesbüren, BASF Agricultural Center Limburgerhof, for their interest in this project and for the financial support. The financial support granted by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 359, “Reaktive Strömungen, Diffusion und Transport” is also gratefully acknowledged.

I owe many thanks to Karin Aden and Andreas Horn, Institute of Geoecology, Technical University of Braunschweig, for the good cooperation and for their great effort to supply me with laboriously collected experimental data. I am grateful to the BASF Agricultural Center in Limburgerhof, the Staatliche Lehr- und Forschungsanstalt (SLFA) in Neustadt/Weinstr., and the Novartis Crop Protection AG in Basel for giving me access to interesting studies and experimental data.

And last, but certainly far from least, I thank my parents for their love and their continuous support during all the years of study, and my fiancé Stephan for his love and his encouraging support while doing this PhD project.

Contents

List of Symbols

	Lowercase Latin Characters	
c	concentration in the liquid phase	$[ML^{-3}]$
c_{ref}	fixed reference concentration for Freundlich isotherm	$[ML^{-3}]$
c_T	total concentration	$[ML^{-3}]$
f	correction factor for the mass balance	$[1]$
g	concentration in the gaseous phase	$[ML^{-3}]$
$\bar{g}$	acceleration of gravity	$[9.81\ ms^{-2}]$
k_c	degradation rate in the liquid phase	$[T^{-1}]$
k_s	degradation rate in the solid phase	$[T^{-1}]$
k_{eff}	apparent rate constant	$[T^{-1}]$
n, m	fitting parameters	$[1]$
q	volumetric water flux density	$[LT^{-1}]$
s	concentration in the solid phase	$[MM^{-1}]$
t	time	$[T]$
v	pore water velocity	$[LT^{-1}]$
$\bar{w}$	conversion factor $\bar{w} = \rho_w \bar{g}$	$[ML^{-3}ms^{-2}]$
z	vertical coordinate (positive downward)	$[L]$
	Uppercase Latin Characters	
$C(\psi_m)$	specific soil water capacity in ψ_m	$[M^{-1}LT^2]$
$\bar{C}(\theta)$	inverse of specific soil water capacity in θ	$[ML^{-1}T^{-2}]$
D_0	coefficient of molecular diffusion in pure water	$[L^2T^{-1}]$
D_{air}	coefficient of molecular diffusion in the air	$[L^2T^{-1}]$
D_d	coefficient of molecular diffusion in the liquid phase	$[L^2T^{-1}]$
D_h	hydrodynamic dispersion term	$[L^2T^{-1}]$
D_m	coefficient of dispersion	$[L^2T^{-1}]$
D_s	coefficient of molecular diffusion in the gaseous phase	$[L^2T^{-1}]$
$\bar{D}(\theta)$	soil water diffusivity	$[L^2T^{-1}]$
J_c	mass flux due to convection	$[ML^{-2}T^{-1}]$
J_d	mass flux in the liquid phase due to diffusion	$[ML^{-2}T^{-1}]$
J_g	mass flux in the gaseous phase	$[ML^{-2}T^{-1}]$
J_m	mass flux in the liquid phase due to dispersion	$[ML^{-2}T^{-1}]$

J_T	total mass flux	$[ML^{-2}T^{-1}]$
$K(\psi_m)$	hydraulic conductivity in ψ_m	$[M^{-1}L^3T]$
$\bar{K}(\theta)$	hydraulic conductivity in θ	$[M^{-1}L^3T]$
K_d	equilibrium constant of linear isotherm	$[M^{-1}L^3]$
K_f	sorption coefficient of Freundlich isotherm	$[M^{-N}L^{3N}]$
$\bar{K}_f$	sorption coefficient of scaled Freundlich isotherm	$[M^{-1}L^3]$
K_l	sorption coefficient of Langmuir isotherm	$[M^{-1}L^3]$
K_s	saturated hydraulic conductivity	$[M^{-1}L^3T]$
K_M	Michaelis constant	$[ML^{-3}]$
N	Freundlich exponent	$[1]$
Q	source/sink term in the convection-dispersion equation	$[ML^{-3}T^{-1}]$
S	source/sink term in the Richards equation in ψ_m	$[T^{-1}]$
$\bar{S}$	source/sink term in the Richards equation in θ	$[T^{-1}]$
V_{max}	maximal reaction velocity (Michaelis-Menten kinetics)	$[ML^{-3}T^{-1}]$

Lowercase Greek Characters

α	fitting parameter	$[M^{-1}LT^2]$
α_L	dispersion length	$[L]$
$\bar{\alpha}$	adsorption-desorption rate	$[T^{-1}]$
β	finite adsorption capacity	$[MM^{-1}]$
ϵ	volumetric air content	$[L^3L^{-3}]$
λ	pore size distribution index	$[1]$
ϕ	porosity of the soil	$[L^3L^{-3}]$
ψ_b	bubbling pressure	$[ML^{-1}T^{-2}]$
ψ_g	gravitational potential	$[ML^{-1}T^{-2}]$
ψ_h	soil water potential	$[ML^{-1}T^{-2}]$
ψ_m	matric potential	$[ML^{-1}T^{-2}]$
ψ_s	osmotic potential	$[ML^{-1}T^{-2}]$
ψ_{tp}	tensiometer potential	$[ML^{-1}T^{-2}]$
ρ	bulk density	$[ML^{-3}]$
ρ_w	mass density of water	$[ML^{-3}]$
θ	volumetric water content	$[L^3L^{-3}]$
θ_r	residual water content	$[L^3L^{-3}]$
θ_s	saturated water content	$[L^3L^{-3}]$
$\tau(\theta)$	tortuosity factor	$[1]$

Uppercase Greek Characters

Θ	normalized water content	$[1]$

Introduction

Xenobiotics such as pesticides, herbicides or fungicides, play a key role in an efficient and economic agriculture. Their critical impact on the environment, in particular on the groundwater, has long been recognized. In recent years, care about environmental effects and human safety has become of major public interest. Authorities have responded by introducing stricter regulations and imposing stringent test procedures before releasing a substance for use in fields (Plimmer, 1999 [101]).

On the European level, the Registration Directive 91/414/EEC, concerning the placing of plant protection products (PPP) on the EU market, came into force in 1993 (Boesten et al., 1999 [24]). Uniform principles for the registration process, e.g. the type and scope of experimental studies, were defined. In particular, it is required that in addition to the experiments carried out for the assessment of the fate and behavior of plant protection products in the environment, estimates of predicted environmental concentrations in soil, water, and air based on mathematical modeling have to provided (Kloskowski et al., 1999 [73]). Several simulation tools, e.g. PELMO (Klein, 1995 [72]; Jene, 1998 [69]), PEARL (Tiktak et al., 1999 [125]), PRZM (Carsel et al., 1998 [29]) or MACRO (Jarvis and Larsson, 1998 [68]), have become available and are widely used by research institutes as well as companies and regulatory agencies. Developments in recent years have shown that due to its enormous potential to predict the substance behavior for different soils and climatic conditions, mathematical modeling and simulation will be of increasing importance for the registration process.

In order to solve the simulation problem, i.e. the forward problem, for predicting system states, e.g. concentrations, the values of all parameters used in the mathematical model have to be available. Usually, the unknown parameters in the model, such as sorption coefficients or half-lifes, are determined from batch experiments or incubation studies using sieved soils under controlled temperature and moisture conditions in the laboratory. However, the validity of the extrapolation of laboratory data to transport and degradation processes in undisturbed soil cores or in the field have been questioned.

An alternative and very promising way to obtain parameters suitable for predicting the environmental fate of xenobiotics is inverse modeling. The idea of inverse modeling is to estimate the unknown parameters directly from lysimeter or field measurement data. This is done by mathematical optimization, where an objective function containing weighted deviation between the computed and the observed data is minimized.

Inverse modeling has been common practice for saturated flow problems in ground-water hydrology for years (Yeh, 1986 [142]). Its application to the unsaturated zone is relatively new. It has thus far been mainly limited to the water transport equation to infer hydraulic parameters of the soil. For overviews see (Kool et al., 1987 [81]) and (Hopmans and Šimůnek, 1999 [61]). There are only few papers considering parameter estimation for reactive solute transport under unsaturated conditions described by coupled water and solute transport equations (see e.g. Mishra and Parker, 1989 [92]; Medina and Carrera, 1996 [90]; Abbaspour et al., 1997 [1]; Šimůnek and van Genuchten, 1999 [116]; Dieses et al., 1999 [41]; Dieses et al., 1999 [43]). In the context of registration studies, the idea has been pursued to compute unknown parameters by coupling simulation tools with commonly available nonlinear regression software, e.g. PEST (Doherty, 1994 [44]). However, the results thus obtained are mainly unsatisfying.

Up to now, the registration of new substances on the basis of parameters derived by inverse modeling has been granted only in some specific cases. However, the role of inverse modeling in the registration of plant protection products is under discussion. Members of different national authorities of European countries, e.g. Germany, Denmark, United Kingdom, The Netherlands, Belgium and France, gave statements on the current status of the use of inverse modeling in national pesticide registration at the recent "Workshop on Inverse Modeling" at the Research Center Jülich in May 2000. An expert of the German regulatory agency, the Biologische Bundesanstalt für Land- und Forstwirtschaft (BBA), for example, stated that "up to now inverse modeling is a procedure that is not very well tested and accepted" and that "more information on the method is needed until it can be accepted." In summary, there was no consensus as such. Even though the benefits were granted, the broad majority of the representatives expressed a rather cautious attitude towards the results obtained so far. Most adopted a policy of wait-and-see asking for more information and expertise in order to judge the method's validity. In particular, more sophisticated mathematical methods were required in order to overcome the deficiencies encountered by combining simulation models with model-independent nonlinear regression tools. The INRA, France, for example, claimed that "more insight into the techniques of inversion (feasibility, mathematical pitfalls, robustness of the method,...) is needed."

However, in order to establish inverse modeling in the registration process, in addition to reliable tools for parameter estimation, optimized experimental designs for column, lysimeter and field experiments are necessary. Methods are needed to determine experimental conditions, such as irrigation or application schemes, and sampling designs leading to measurement data that is suitable for parameter estimation. The choice of the experimental design is of particular interest, as the parameter estimation problems investigated so far have frequently been shown to be ill-posed. Even though this problem of ill-posedness, mainly caused by parameters that are practically insensitive to the measurement data, has been encountered in many studies (e.g. Hornung, 1983 [63]; Kool et al., 1985 [80]; Toorman et al., 1992 [126]; van Dam et al., 1992 [127]), only a few approaches for optimizing the experimental designs have been reported. In some few cases the problem to derive better experimental conditions and sampling schemes has been interpreted and formulated as a mathematical optimization problem. Due to the lack

of suitable optimization procedures, the problem has often been solved by an enumerative grid search through the design space (e.g. Knopmann and Voss, 1987 [74]; 1989 [76]). Sun (1994 [123]) even stated that "in practice to solve such a general problem is too difficult" and "that up to date, only few hypothetical examples and simplified cases have been reported".

Hence, in addition to the scientific interest there is an increasing demand by industry and also by national authorities for sophisticated and reliable methods for both parameter estimation and optimal experimental design in transport and degradation processes of xenobiotics in soils.

This is the background setting for this work which aims at providing methods that simplify investigations and enable a better understanding of the observed processes. In order to support the mathematical modeling, the tools ECOFIT (Dieses et al., 1999 [41]; 1999 [43]) for parameter estimation and ECOPLAN (Dieses et al., 2000 [42]) for optimal experimental design in water and reactive solute transport processes in soils have been developed. In this thesis new contributions are made to the following fields:

- Modeling of parameter estimation problems and optimal experimental design problems
- Numerical methods for parameter estimation in water flow and reactive solute transport processes of xenobiotics in soils
- Numerical methods for optimal experimental design in water flow and reactive solute transport processes of xenobiotics in soils
- Application of the developed methods to various real-life problems

In the following we give a summary containing the new contributions made to these topics. We have organized this thesis in essentially independent chapters because we assume that not all readers are interested in all chapters to the same extent. Those readers who are mainly interested in the modeling aspects and in the application of the tools ECOFIT and ECOPLAN to column, lysimeter and field studies are referred to the Chapters 1 and 5. For those readers who are interested in the mathematical part and want to get a deeper understanding of the multiple shooting method and the Generalized Gauss-Newton method, a review is provided in Section 2.1 and 2.2 based on (Bock, 1981 [19]; 1983 [20]; 1987 [21]; Schlöder and Bock, 1983 [109]; Schlöder, 1988 [108]).

Modeling for Parameter Estimation and Optimal Experimental Design (Chapter 1)

When studying the environmental fate of xenobiotics, the vadoze zone, i.e. the first 30 to 100 centimeters of the soil where the relevant processes such as degradation take place, is of particular interest. As this zone is generally not saturated with water, the transport of dissolved substances is greatly affected by the flow of water. Thus, unless studying column experiments in the laboratory where steady-state conditions can be met, the transport of both water and solute have to be considered simultaneously. The mathematical modeling of water and solute

transport processes in the unsaturated zone leads to instationary partial differential equations (PDEs) coupled with nonlinear ordinary differential (ODEs) or differential algebraic equations (DAEs).

For the sake of comparability and in order to enable the use of the developed tools for registration studies, the choice of model equations considered in this work is oriented at the equations implemented in the simulation models used so far for registration. In addition to these models, presented in Section 1.1 and 1.2, we investigate more complex models, not yet available in the commonly used simulation tools, that include e.g. nonlinear sorption described by Langmuir isotherms or nonlinear degradation according to Michaelis-Menten kinetics. A very difficult problem in the modeling of column, lysimeter or field experiments is the appropriate description of initial and boundary conditions. In Section 1.3, we summarize frequently encountered conditions in practice and work out suitable modeling approaches for initial, upper and lower boundary conditions.

Section 1.4 is devoted to the parameter estimation problem in environmental fate experiments. In Section 1.4.1, we present a formulation for the inverse problem that covers a wide range of typical column and lysimeter experiments. In Section 1.4.2, the parameter estimation methods used in common practice are reviewed and analyzed. In contrast to the frequently reported approach where the simulation and the optimization problems are treated separately, we pursue a different approach where the parameter estimation problem is interpreted as a weighted least-squares problem constrained by a set of PDEs and ODEs (Bock, 1981 [19]).

The problem of how to optimize the experimental design for column and lysimeter experiments is addressed in Section 1.5. In Section 1.5.1, based on our approach for the parameter estimation problem, the optimal experimental design problem is derived and formulated as an optimal control problem. Hereby, we build on the approach worked out by Körkel et al. (1999 [83]) and Bauer et al. (1999 [9]; 2000 [10]) for systems described by ODEs and DAEs. We distinguish between several types of optimization variables. Experimental conditions such as initial conditions of the soil column and irrigation/application schemes are described by time-independent control variables and control functions, respectively. In order to optimize the sampling scheme, binary weights for each possible measurement point are introduced. The few approaches reported in literature for optimizing the design are reviewed in Section 1.5.2.

The set-up of a meaningful model for such complex systems as the environmental fate of xenobiotics in soils is, however, generally an iterative process which requires the support of suitable tools. For this purpose, we have developed the tools ECOFIT and ECOPLAN for parameter estimation and optimal experimental design which are presented in the following.

ECOFIT: **An efficient method for parameter estimation (Chapter 2 and 3)**

While enormous progress has been made in the development of powerful methods for the simulation of multi-component multi-phase flow in two and three dimensions, the study of the inverse problem has up to now been mainly limited to smaller test problems (Sun, 1994 [123]). This might be due to the fact that the adequate treatment of parameter estimation problems as

they arise from column or mini-lysimeter studies under unsaturated flow conditions is a very demanding and complex task requiring methods that

- can treat high nonlinearities,
- enable sufficiently fine spatial grids in order to handle high spatial activity within the soil cores,
- enable the fast solution of the resulting large scale systems,
- guarantee the efficient generation of sufficiently accurate derivatives,
- incorporate prior information,
- are based on reliable termination criteria.

The development of the parameter estimation tool ECOFIT according to these requirements has only been possible within this period of time since we could built upon the sophisticated methods and the knowledge available in the research group of Bock and Schlöder. Essentially, the developed approach rests on the following two pillars:

- a *reduced* Generalized Gauss-Newton method (Chapter 2)
- very efficient strategies for the generation of the required derivatives (Chapter 3).

In Chapter 2, the focus is upon the efficient solution of large scale parameter estimation problems. In the first section, we outline how the parameter estimation problem constrained by PDEs and ODEs is transformed by discretization into a large scale nonlinear constrained least-squares problem. Finite differences are employed for spatial discretization. Discretization in time is done by multiple shooting. A short review of the Generalized Gauss-Newton method (Bock 1981 [19]; 1983 [20]; 1987 [21]) and its convergence results is given in Section 2.2. The need for a specially tailored approach for large scale problems is motivated.

In Section 2.3, we present an approach based on the reduced Generalized Gauss-Newton method (Schlöder, 1988 [108]) that is capable to handle these large scale problems arising from the discretization of PDEs. We exploit the fact that the initial conditions for the states are fixed and thus the systems have only few degrees of freedom, namely equal to the number of unknown parameters. Using directional derivatives for setting up the linearized problems, we end up with essentially the same computational effort and storage requirements as for the single shooting method while maintaining the advantages of multiple shooting.

In Chapter 3, efficient strategies are presented for the computation of derivatives in discretized PDE-systems. The analysis of the reduced Generalized Gauss-Newton method shows that by far the majority of computational effort is spent in the computation of derivatives. Thus, in order to further speed up the code, we develop specially tailored, highly efficient methods that exploit structures on several levels.

In ECOFIT, the state-of-the-art integrator DAESOL (Bauer, 1999 [11]; 2000 [8]), a multi-step method code with a variable step size and order control based on Backward Differentiation Formulae (BDF), is used. So far, DAESOL has been the only tool that provides the solution of the forward problem as well as the computation of both first and second order derivatives within the framework of *Internal Numerical Differentiation* (Bock, 1981 [19]). The required directional derivatives are computed by the solution of the corresponding variational differential equations.

In the first section, the performance of a standard version of DAESOL to handle also large scale problems as they arise from discretized PDEs is investigated for both the finite difference mode (FD) and the automatic differentiation mode (AD). In the latter mode the derivatives of the right hand side f with respect to the states y and the parameters p are provided by the automatic differentiation tool ADIFOR (Bischof et al., 1992 [15], 1994 [16], 1998 [17]). Analysis of the performance reveals that the computational effort for parameter estimation is mainly dominated by the frequent computation of f_y whose computational complexity increases quadratically with the number of spatial nodes $\mathcal{O}(n_y^2)$.

In Section 3.2, we first outline an approach that circumvents the computation of f_y in each BDF step by using a modified Newton method (Dieses et al., 1999 [41]; Bock et al., 1995 [23]). The idea is to reduce the number of f_y-computations by substituting them by directional derivatives of type $f_y w$.

Secondly, in Section 3.3, we present an approach that removes the complexity order $\mathcal{O}(n_y^2)$ in order to further speed up the code. We exploit the fact that due to the use of fixed spatial grids the sparsity pattern induced by the spatial discretization of the PDEs remains unchanged in the course the reduced Generalized Gauss-Newton method. By identifying structural orthogonal columns of f_y (Curtis et al., 1974 [34]), f_y is computed via a compressed matrix requiring only $l \ll n_y$ instead of n_y directional derivatives. Thus, the computation of f_y is only dependent on the order of the spatial discretization routine and the number of PDEs and is independent of the number of spatial nodes. By this strategy, we manage to reduce the computation complexity of f_y to the same complexity order $\mathcal{O}(n_y)$ as the evaluation of the right hand side f.

Combing both strategies, the modified Newton method and the compressed approach for the computation of f_y, finally a speed up by a factor of 40 is gained for a parameter estimation problem with 2 PDEs and 961 spatial nodes. The CPU time is reduced from originally 2.5 days to less than 1.5 hours.

ECOPLAN: **An approach for optimizing experimental conditions and sampling schemes (Chapter 4)**

Due to the numerous problems encountered by estimating unknown parameters from water and solute transport processes, many researchers have claimed the need for methods to optimize experimental designs.

In order to address this problem, it has become popular among the soil science community within recent years to derive improved designs by analyzing two-dimensional response surfaces

or by calculating sensitivity coefficients from hypothetical data. However, in both cases this is done on a tedious and cumbersome trial-and-error basis, rather than embedding the problem into the framework of optimization. There are only few studies that consider this problem in the context of optimal experimental design theory using statistical design criteria. The basis for this is the variance-covariance matrix, which describes the statistical quality of the parameter estimates. The objective of optimal experimental design is to minimize a function Φ of the variance-covariance matrix of the underlying parameter estimation problem. Often one of the classical designs, D-, A-, or E-optimality are used, which aim at minimizing the determinant, the trace, and the largest eigenvalue of the variance-covariance matrix, respectively. So far this approach has only been used to derive optimal sampling schemes, i.e. optimal allocations of measurement points in time and space.

However, considering in particular column or mini-lysimeter studies, there are by far more possibilities to influence the experiments in order to obtain good measurement data with respect to parameter estimation. In this work, for the first time an approach is presented that enables to optimize both

- the experimental conditions, e.g. initial soil profiles, irrigation and application schemes, and
- the sampling design

with respect to parameter estimation in transport and degradation processes of xenobiotics in soils. The approach presented follows the concepts worked out within the BMBF project "Optimale Versuchsplanung für nichtlineare Prozesse" (FKZ: 03 D 0043, principal investigators: Bock, Schlöder). Within this project the optimal experimental tool VPLAN for parameter estimation problems constrained by ODEs and DAEs was developed (Bauer et al., 1999 [9]; Körkel et al., 1999 [83]; Bauer et al., 2000 [10]). On the basis of VPLAN and the parameter estimation tool ECOFIT, the new tool ECOPLAN that is suitable for optimal experimental design in water and reactive solute transport processes has been developed.

In Section 4.1, the optimal experimental design problem is formulated as an optimal control problem. Hereby, in addition to time-independent control variables and time-dependent control functions, binary weights for feasible measurement points in time and space are considered. For the numerical solution, as described in Section 4.2, a direct approach is employed where time-dependent control functions and state constraints are discretized on a suitable grid. This results in a finite dimensional, nonlinear constrained optimization problem that is solved by a structured SQP-method. In Section 4.3, we discuss practical requirements for optimal experimental designs and possible extensions, such as sequential designs or designs for model discrimination.

Application of ECOFIT and ECOPLAN to column, lysimeter and field studies (Chapter 5)

As this is an interdisciplinary work, the application of the developed tools ECOFIT and ECOPLAN to column, lysimeter and field experiments is of particular interest. Different types of examples are presented, some of them using hypothetical data, some of them building on data that

is provided by experiments. In the first two examples, in order to enable a controlled scenario, measurement data is generated by solving the forward problem for a predefined true parameter set followed by adding pseudo-normally distributed noise. In the first example, Section 5.1, the performance of ECOFIT is investigated for identifying water and solute transport parameters from different noisy data sets starting from poor initial guesses. The second example, Section 5.2, is devoted to the treatment of parameter estimation problems in layered soils.

In Section 5.3, we also study a field experiment even though the main focus of this work is upon inverse modeling in column and mini-lysimeter experiments. Here, ECOFIT is used for estimating the van Genuchten parameters n, α and K_s from field data. The experimental part of this study was carried out by Aden (1999 [3]; [2]) at a BASF test site in the upper Rhine valley. Time domain reflectrometry (TDR) was employed to monitor the volumetric water contents in several depths. For this type of experiments, an adequate model is developed.

In two cases ECOFIT is applied to mini-lysimeter studies as they are performed for registration purposes. In the first case, Section 5.4, the transport and sorption behavior of three European soils are determined. The experiments were carried out by the Staatliche Lehr- und Forschungsanstalt (SLFA), Neustadt/Weinstr. (Fent, 1999 [50]). After the application of the non-reactive tracer bromide and of a ^{14}C-labeled test substance X, the undisturbed soil cores were irrigated by a constant daily rate and leachate volumes were sampled. A model is worked out for this typical class of outflow experiments and the unknown transport and sorption parameters are estimated.

In a second study, Section 5.5, we investigate the environmental fate of the grass herbicide S-Metolachlor and its two main metabolites by means of ECOFIT. The data used for inverse modeling was obtained by mini-lysimeter experiments performed by Horn (1999 [62]) at a Novartis test site in Switzerland. In this study the mini-lysimeters were exposed to normal climatic conditions. Again, the unknown parameters, such as the linear sorption coefficients and the degradation rates, are determined based on leachate data. This parameter estimation problem, however, suffers from the problem of ill-posedness due to the insufficient information provided by the available data.

In the last section, the potential and the features of ECOPLAN are demonstrated on a hypothetical column outflow experiment where both water and solute transport parameters are estimated simultaneously. In a first scenario, the impact of optimized experimental (boundary) conditions, such as irrigation scheme and substance concentration in the irrigation water, is studied for a given sampling scheme. The objective of a second scenario is the simultaneous optimization of both the sampling scheme and the experimental conditions. Comparing intuitive designs with the experimental design optimized by ECOPLAN reveals the huge potential of optimal experimental design. With the same number of observations parameter variances can be drastically reduced. Moreover, questions about the influence of neglecting or adding certain types of data on the quality of parameter estimates can be easily answered. We can, for example, decide a priori, whether profile concentrations obtained by slicing soil columns at the end of an experiment can essentially improve estimation results or not. Several examples are presented and discussed in detail.

All computational results presented in this work were obtained on a workstation SUN ULTRA SPARC 10 (300MHz) running Solaris version 7.

Chapter 1

Mathematical Problems in the Environmental Fate Modeling of Xenobiotics

Registration of xenobiotics requires, besides the ecotoxicological risk assessment, studies about the environmental fate of the employed substances. On the one hand, regulatory authorities are interested in parameter values such as sorption coefficients, e.g. K_d-values, or degradation rates. On the other hand, more recently simulation results have to be submitted as well. A European consensus about a simulation model has not been obtained yet. In most Western European countries the one-dimensional models PELMO (Klein, 1995 [72]; Jene, 1998 [69]), PEARL (Tiktak et al., 1999 [125]), PRZM (Carsel et al., 1998 [29]) or MACRO (Jarvis and Larsson, 1998 [68]) are used.

As the tools developed in this work should also be applicable to registration studies, the choice of the model equations considered is based on the simulation models used in current practice by industry and regulatory authorities.

In the first two sections the relevant model equations applied for registration calculations, i.e. the Richards equation for the water transport, the convection-dispersion equation for the solute transport, linear sorption and degradation, are presented. In addition, more recent model approaches for nonlinear processes such as Langmuir sorption or Michaelis-Menten kinetics are discussed. These more complex models are generally not available in commonly used simulation tools. Up till now, nonlinear approaches of this type are of minor interest, as the risk assessment of xenobiotics required by authorities is mainly constrained to linear approaches.

In Section 1.3 special aspects associated with the modeling of column, lysimeter and field experiments, in particular the adequate description of upper and lower boundary conditions are addressed. The following section is devoted to the formulation of the parameter estimation problem as it typically arises in the environmental fate studies. An overview of parameter estimation tools used in current practice is given. Frequently encountered pitfalls and shortcomings are summarized and the need for more sophisticated methods is motivated. In Section 1.5 an optimal control problem is derived for the determination of optimal experimental conditions and

sampling schemes for column and mini-lysimeter studies with respect to parameter estimation. Strategies employed in practice to treat the design problem are reviewed.

1.1 Transient Flow of Water in the Unsaturated Soil

Under natural conditions, as they arise in lysimeter or field experiments, water flow through soil is in general instationary, i.e. it varies in time and space. In the following the Richards equation for water transport in the unsaturated zone is derived. Frequently used parameterization of the soil water characteristic and the unsaturated hydraulic conductivity are presented which are assumed to hold for soils with uni-modal pore size distributions.

However, in the case of macroscopic structures like fractures or macropores more complicated models, e.g. two domain approaches considering macropore flow (Jarvis, 1991 [67]) or capillary flow (e.g. Diekkrüger, 1992 [36]; Richter et al., 1996 [104]) should be employed. Even though these approaches are not discussed in the following, they can, in principle, be treated by the methods developed in this work.

1.1.1 Richards Equation

The transport equation for water in a variably saturated rigid porous medium can be derived by the mass conservation equation, which in the one-dimensional case is of the form

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} + S, \tag{1.1}$$

where θ $[L^3L^{-3}]$ is the volumetric water content, q $[LT^{-1}]$ is the volumetric water flux density, S $[T^{-1}]$ is a source/sink term, z $[L]$ is the vertical coordinate (positive downward), and t $[T]$ is time.

We assume Darcy's law

$$q = -\bar{K}\frac{\partial \psi_h}{\partial z} \tag{1.2}$$

to hold, which relates the water flux q with a hydraulic gradient by the hydraulic conductivity $\bar{K}$ $[M^{-1}L^3T]$. However, in contrast to the case of saturated conditions, the hydraulic conductivity is a function, $\bar{K}(\theta)$ $[M^{-1}L^3T]$, that not only depends on the geometry of the pore space and the physical properties of the water phase but also on the geometry of the water phase.

The soil water potential ψ_h $[ML^{-1}T^{-2}]$ is composed of several partial potentials, i.e. the gravitational potential ψ_g $[ML^{-1}T^{-2}]$, the osmotic potential ψ_s $[ML^{-1}T^{-2}]$ and the tensiometer potential ψ_{tp} $[ML^{-1}T^{-2}]$:

$$\psi_h = \psi_g + \psi_s + \psi_{tp}. \tag{1.3}$$

The gravitational potential

$$\psi_g = -\rho_w \bar{g}(z - z_0) \tag{1.4}$$

describes the energy that is necessary to move water from a reference level z_0 to a depth z, where ρ_w $[ML^{-3}]$ is the mass density of water and $\bar{g}$ is the acceleration of gravity $[9.81\ ms^{-2}]$. In the following we neglect the osmotic potential ψ_s and assume the air pressure to be constant. Considering only one component of the tensiometer pressure ψ_{tp}, namely the matric potential ψ_m $[ML^{-1}T^{-2}]$, (1.3) reduces to

$$\psi_h = \psi_g + \psi_m = \psi_m - \bar{w}(z - z_0), \tag{1.5}$$

with $\bar{w} = \rho_w \bar{g}$ $[ML^{-3}ms^{-2}]$. The matric potential ψ_m is the energy that is required to move water into the porous medium.

Combining the mass conservation equation (1.1) with the Darcy law (1.2) and putting the reference level $z_0 = 0$ we get the Richards equation

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\big[\bar{K}(\theta)\frac{\partial}{\partial z}\big(\psi_m - \bar{w}z\big)\big] + S. \tag{1.6}$$

Treating θ as a function of ψ_m, as outlined in detail in the next section, we can apply the chain rule to $\partial\theta/\partial t$ and rewrite $\bar{K}$ as a function of ψ_m, $K(\psi_m)$ $[M^{-1}L^3T]$. Thus, we end up with the Richards equation in the matric potential form

$$C(\psi_m)\frac{\partial \psi_m}{\partial t} = \frac{\partial}{\partial z}\big[K(\psi_m)\frac{\partial}{\partial z}\big(\psi_m - \bar{w}z\big)\big] + S, \tag{1.7}$$

where $C(\psi_m) = d\theta/d\psi_m$ $[M^{-1}LT^2]$ is the specific soil water capacity. Note, that (1.7) is valid for both the saturated and the unsaturated case.

On the other hand, we can also interpret ψ_m as a function of θ. By applying the chain rule to $\partial\psi_m/\partial z$, we obtain the Richards equation in the water content form

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\big[\bar{D}(\theta)\frac{\partial \theta}{\partial z} - \bar{w}\bar{K}(\theta)\big] + \bar{S}, \tag{1.8}$$

where

$$\bar{D}(\theta) := \bar{K}(\theta)\bar{C}(\theta), \qquad \bar{C}(\theta) = \frac{d\psi_m}{d\theta} \tag{1.9}$$

is the soil water diffusivity $[L^2T^{-1}]$. $\bar{S} := \bar{S}(\theta)$ $[T^{-1}]$ denotes the source/sink term in θ in contrast to $S := S(\psi_m)$. Due to the fact that $\bar{C}(\theta) \to \infty$ $[ML^{-1}T^{-2}]$ as the soil becomes saturated, (1.8) is only defined for the unsaturated case. Note, that (1.8) can be formally interpreted as a convection-dispersion equation, where formally $V(\theta) = \bar{w}(d\bar{K}/d\theta)$ denotes the convection velocity and $\bar{D}(\theta)$ the dispersion coefficient (Roth, 1996 [106]).

1.1.2 Parameterization of Hydraulic Functions

Several parameterizations of the unsaturated soil hydraulic properties, i.e. the soil water characteristic and the unsaturated hydraulic conductivity are known. In order to enable a compact formulation we introduce the normalized water content [1]

$$\Theta = \frac{\theta - \theta_r}{\theta_s - \theta_r}, \tag{1.10}$$

where θ_s $[L^3L^{-3}]$ is the saturated volumetric water content and θ_r $[L^3L^{-3}]$ is the residual volumetric water content.

One of the first models was given by Brook and Corey (1964 [27]) and Burdine (1953 [28]):

$$\Theta(\psi_m) = \begin{cases} \left(\frac{\psi_b}{\psi_m}\right)^{\lambda} & \text{for} \quad \psi_m \leq \psi_b \\ 1 & \text{for} \quad \psi_m > \psi_b \end{cases} \tag{1.11}$$

$$K(\psi_m) = K_s\Theta^{3+2/\lambda}, \tag{1.12}$$

where ψ_b $[ML^{-1}T^{-2}]$ denotes the bubbling pressure, λ $[1]$ the pore size distribution index, and K_s $[M^{-1}L^3T]$ the saturated hydraulic conductivity.

Basing on (1.11) van Genuchten (1980 [129]) developed a class of functions that is continuously differentiable and applicable to a broader range of soil types:

$$\Theta(\psi_m) = \begin{cases} \left[1 + (\alpha|\psi_m|)^n\right]^{-m} & \text{for} \quad \psi_m \leq 0 \\ 1 & \text{for} \quad \psi_m > 0, \end{cases} \tag{1.13}$$

where m $[1]$, n $[1]$ and α $[M^{-1}LT^2]$ are positive fitting parameters. In connection with the statistical pore size distribution model of Mualem (1976 [94]) for the unsaturated hydraulic conductivity

$$K(\psi_m) = K_s\frac{(1 - (\alpha|\psi_m|)^{n-1}(1 + (\alpha|\psi_m|)^n)^{-m})^2}{(1 + (\alpha|\psi_m|)^n)^{\frac{m}{2}}} \tag{1.14}$$

frequently $m = 1 - 1/n$ is used.

Conversely, both functions can be also expressed as functions of θ:

$$|\psi_m(\theta)| = \alpha^{-1}\left(\Theta^{-\frac{1}{m}} - 1\right)^{\frac{1}{n}} \qquad \text{for} \quad \psi_m \leq 0 \tag{1.15}$$

$$\bar{K}(\theta) = K_s\Theta^{\frac{1}{2}}\left[1 - \left(1 - \Theta^{\frac{1}{m}}\right)^m\right]^2. \tag{1.16}$$

Remark 1.1 (Hysteresis)
The parameterization for the soil water characteristic mentioned above does not consider hysteresis. For models that include hysteresis see for example (Kool et Parker, 1987 [78]). It should be noted that the hydraulic conductivity in ψ_m generally shows a strong hysteresis while hysteretic effects for the formulation in θ are often negligible.

1.2 Solute Transport through the Subsurface with Transient Flow of Water

The processes associated with the environmental fate of xenobiotics in soils are very complex and by far not completely understood. Their adequate modeling, in particular for heterogeneous structures, is subject of current research.

In this section, we derive the convection-dispersion equation which is assumed to hold for the description of the transport of dissolved substances on the column and mini-lysimeter scale. Under unsaturated conditions the convection-dispersion equation has to be solved together with the Richards equation. Several approaches for the modeling of linear and nonlinear sorption and degradation processes are presented.

1.2.1 Convection-Dispersion Equation

The transport of dissolved chemicals, e.g. xenobiotics, in a porous medium such as soil is driven by convection, molecular diffusion and dispersion. Considering also mass conservation for these processes the convection-dispersion equation for solute transport in the unsaturated zone can be derived.

Starting from the general continuity equation mass conservation of the total concentration c_T $[ML^{-3}]$ in the one-dimensional case requires

$$\frac{\partial c_T}{\partial t} = -\frac{\partial J_T}{\partial z} + Q, \tag{1.17}$$

where J_T $[ML^{-2}T^{-1}]$ denotes the total mass flux and Q $[ML^{-3}T^{-1}]$ is a source or sink term representing the creation or disappearance of substance.

The total concentration c_T can be decomposed into the respective concentrations occurring in the liquid, solid and gaseous phase:

$$c_T = \theta c + \rho s + \epsilon g, \tag{1.18}$$

where c $[ML^{-3}]$ is the concentration in the liquid phase, s $[MM^{-1}]$ is the concentration in the solid phase, ρ $[ML^{-3}]$ is the bulk density, g $[ML^{-3}]$ is the concentration in the gaseous phase, and ϵ $[L^3L^{-3}]$ is the volumetric air content.

For the total mass flux J_T several components can be distinguished, namely molecular diffusion in the liquid and gaseous phase, dispersion, and convective transport along with the water flow. Each of these terms will be discussed in the following and model approaches will be given.

Molecular diffusion in the liquid phase

Molecular diffusion of dissolved substances in the liquid phase is caused by Brownian motion and leads to mixing due to concentration gradients. It is described by Fick's law

$$J_d = -D_d(\theta)\frac{\partial c}{\partial z}, \tag{1.19}$$

where D_d $[L^2T^{-1}]$ is called the coefficient of molecular diffusion. However, D_d is not a constant but a function of θ and parameterizes the geometry of the water phase:

$$D_d(\theta) = D_0\tau(\theta), \tag{1.20}$$

where D_0 $[L^2T^{-1}]$ is the coefficient of molecular diffusion in pure water, and $\tau(\theta)$ $[1]$ is the tortuosity factor. This empirical tortuosity factor $\tau(\theta)$ is a function of the volumetric water content θ and takes into account longer flow paths due to the space geometry.

Often used approximations are based on the Millington-Quirk models for gaseous diffusion (Millington and Quirk, 1961 [91]), such as

$$\tau(\theta) = \frac{\theta^{10/3}}{\phi^2} \tag{1.21}$$

or as proposed by Jin and Jury (1996 [70])

$$\tau(\theta) = \frac{\theta^2}{\phi^{2/3}}. \tag{1.22}$$

Here ϕ $[L^3L^{-3}]$ denotes the porosity of the soil. Kemper and van Schaik (1966 [71]) suggested the empirical parameterization

$$\tau(\theta) = a \exp(b\theta), \tag{1.23}$$

where a $[1]$ and b $[1]$ are positive constants.

Molecular diffusion in the gaseous phase

According to Fick's law the mass flux of substance in the gaseous phase J_g $[ML^{-2}T^{-1}]$ due to molecular diffusion is given by

$$J_g = -D_s \frac{\partial g}{\partial z}, \tag{1.24}$$

where D_s $[L^2T^{-1}]$ is the coefficient of molecular diffusion in the gaseous phase. Employing the approach of Millington and Quirk (1961 [91]) we get

$$D_s = \frac{\epsilon^{10/3}}{\phi^2} D_{air}, \tag{1.25}$$

with the coefficient of molecular diffusion in the air D_{air} $[L^2T^{-1}]$.

Dispersion

Superimposed on the diffusive transport is the dispersive transport. Mechanical dispersion arises from inhomogeneity of the pore space and leads to mixing caused by random water movement:

$$J_m = -\theta D_m \frac{\partial c}{\partial z}, \tag{1.26}$$

where D_m $[L^2T^{-1}]$ is the coefficient of dispersion in the liquid phase. For the one-dimensional case it is often assumed that D_m increases linearly with the pore water velocity $v = |q|/\theta$ $[LT^{-1}]$

$$D_m(q) = \alpha_L \frac{|q|}{\theta}. \tag{1.27}$$

The dispersion length α_L $[L]$ is determined by the geometry of the transport volume.

Remark 1.2
For the saturated case, the transport volume equals the pore space and can thus be treated as a constant. Under unsaturated conditions, however, the geometry of the water phase varies. Thus, α_L relies on the volumetric water content θ which varies in time and space.

The molecular diffusion in the liquid phase (1.19) and the dispersion term (1.26) are both proportional to the gradient $\partial c/\partial z$ and thus are often combined to the hydrodynamic dispersion term

$$D_h(\theta, q) = \frac{D_d(\theta)}{\theta} + D_m(q). \tag{1.28}$$

Convection

The convective flux J_c $[ML^{-2}T^{-1}]$ translates the dissolved substance along with the water flux q

$$J_c = qc. \tag{1.29}$$

As convection is often the dominant component in the transport the correct description of water transport is a prerequisite for a meaningful modeling of the solute transport.

Collecting (1.28), (1.24) and (1.29) the total solute flux is given by

$$J_T = -\theta D_h \frac{\partial c}{\partial z} - D_s \frac{\partial g}{\partial z} + qc. \tag{1.30}$$

By combining now the continuity equation (1.17) with the total solute flux (1.30) we obtain the convection-dispersion equation

$$\frac{\partial c_T}{\partial t} = \frac{\partial}{\partial t}(\theta c + \rho s + \epsilon g) = \frac{\partial}{\partial z}\big(\theta D_h \frac{\partial c}{\partial z} + D_s \frac{\partial g}{\partial z} - qc\big) + Q, \tag{1.31}$$

which describes the solute transport in the unsaturated zone.

1.2.2 Modeling of Sorption Processes in Soils

Sorption processes with the soil matrix generally influence both the transport and the degradation of substances. In the following the impact of sorption on the transport behavior is studied. How sorption affects degradation is discussed in detail in the next section.

Depending on the time scale we distinguish equilibrium and non-equilibrium sorption. Considering equilibrium sorption we assume the sorption process to be very fast compared to the transport. In this case the retardation factor

$$R = 1 + \frac{\rho}{\theta}\frac{ds}{dc} \tag{1.32}$$

is a measure for the delay of the transport caused by sorption. For non-equilibrium bindings, however, the velocity of the sorption process is of the same order of magnitude as the transport velocity or the decay rate.

In the following the classical one-binding-site models for linear and kinetic sorption are presented for equilibrium and non-equilibrium conditions. For models taking several binding sites into account the reader is referred to e.g. Richter et al. (1996 [104]), Nörtersheuser (1993 [96]), Beulke (1998 [14]).

Linear sorption

The non-equilibrium kinetics of sorption for a first-order reaction can formulated as

$$\frac{d}{dt}(\rho s) = \rho\bar{\alpha}(K_d c - s), \tag{1.33}$$

where $K_d\,[M^{-1}L^3]$ is the equilibrium constant for linear adsorption and $\bar{\alpha}\,[T^{-1}]$ is the adsorption-desorption rate. In the equilibrium state, where the time derivative ds/dt is zero, (1.33) reduces to

$$s = K_d c. \tag{1.34}$$

The respective retardation factor is then given by

$$R = 1 + \frac{\rho}{\theta}K_d. \tag{1.35}$$

Langmuir isotherms

For nonlinear sorption processes that are characterized by finite binding capacity the Langmuir isotherm

$$\frac{d}{dt}(\rho s) = \rho\bar{\alpha}\big(\frac{K_l\beta c}{1 + K_l c} - s\big) \tag{1.36}$$

may be used, where $K_l\ [M^{-1}L^3]$ denotes the sorption coefficient of the Langmuir isotherm and $\beta\ [MM^{-1}]$ the finite adsorption capacity. Under equilibrium conditions (1.36) becomes

$$s = \frac{K_l \beta c}{1 + K_l c} \tag{1.37}$$

and we obtain the retardation factor

$$R = 1 + \frac{\rho}{\theta} \frac{K_l \beta}{(1 + K_l c)^2}. \tag{1.38}$$

For small concentrations $c \ll K_l^{-1}$ (1.37) converges to the linear isotherm (1.34) with $K_d = K_l \beta$. At high concentrations $c \gg K_l^{-1}$ the absorbed concentration s is bounded by the finite absorption capacity β.

Freundlich isotherms

Nonlinear sorption that is not bounded by a finite number of binding sites is traditionally modeled by

$$\frac{d}{dt}(\rho s) = \rho \bar{\alpha}(K_f c^N - s), \tag{1.39}$$

where $K_f\ [M^{-N}L^{3N}]$ is the sorption coefficient of the Freundlich isotherm and $N\ [1]$ is the Freundlich exponent. Considering the equilibrium case for (1.39)

$$s = K_f c^N, \tag{1.40}$$

we come up with a retardation factor of the form

$$R = 1 + \frac{\rho}{\theta} N K_f c^{N-1}. \tag{1.41}$$

However, it should be noted that for $N < 1$ and $c = 0$ the retardation factor (1.41) is not defined.

Recently, several formulations have been given (see e.g. Roth 1996 [106]; Tiktak et al., 1999 [125]), which avoid a unit of the Freundlich coefficient K_f that depends on the Freundlich exponent N. The idea is to introduce a fixed reference concentration $c_{ref}\ [ML^{-3}]$. Applying this strategy we can rewrite (1.39) as

$$\frac{\partial}{\partial t}(\rho s) = \rho \bar{\alpha} \big(\bar{K}_f c_{ref} \big(\frac{c}{c_{ref}} \big)^N - s \big), \tag{1.42}$$

where the Freundlich coefficient $\bar{K}_f$ is now of the dimension $[M^{-1}L^3]$.

1.2.3 Model Approaches for Degradation of Xenobiotics in Soils

As the gaseous phase is of no importance for the application problems treated in this work, we concentrate in the following on the liquid and the solid phases. For the sake of presentation the operator

$$L(\theta, q)c := \frac{\partial}{\partial z}\big(\theta D_h \frac{\partial c}{\partial z} - qc\big). \tag{1.43}$$

is introduced.

Linear degradation models

The simplest approach which considers degradation in both the liquid and the solid phase is given by a first-order decay model

$$\frac{\partial}{\partial t}\big(\theta c + \rho s\big) = L(\theta, q)c - k_c\theta c - k_s\rho s, \tag{1.44}$$

where k_c $[T^{-1}]$ and k_s $[T^{-1}]$ denote the degradation rates in the liquid and the solid phase, respectively.

During degradation in soil xenobiotics are often transformed into several metabolites, which undergo further transformation processes. Let us consider a system with a parent c_1 and one metabolite (child) c_2. Assuming that degradation only occurs in the liquid phase we can write

$$\frac{\partial}{\partial t}(\theta c_1) = L_1(\theta, q)c_1 - f k_{c1}\theta c_1 \tag{1.45}$$

$$\frac{\partial}{\partial t}(\theta c_2) = L_2(\theta, q)c_2 + f k_{c1}\theta c_1 - k_{c2}\theta c_2, \tag{1.46}$$

where f $[1]$ is a correction factor for the mass balance of the transformation process.

Nonlinear degradation models

As mentioned in the previous section sorption may not only affect the transport in the soil but also the degradation. This is worked out in the following for equilibrium conditions.

Even in the case where the degradation rate k_c is independent of the concentration c

$$R\theta\frac{\partial c}{\partial t} = L(\theta, q)c - k_c\theta c, \tag{1.47}$$

nonlinear sorption isotherms, i.e. Langmuir or Freundlich isotherms, lead to nonlinear degradation behavior

$$\theta\frac{\partial c}{\partial t} = \frac{1}{R}L(\theta, q)c - k_{eff}\theta c, \tag{1.48}$$

where the apparent rate constant $k_{eff}\ [T^{-1}]$ is a function of the concentration c :

$$k_{eff} = \frac{k_c}{R} = \frac{k_c}{1 + \frac{\rho}{\theta}\frac{ds}{dc}} = \begin{cases} k_c\big(1 + \frac{\rho}{\theta}\frac{K_l\beta}{(1+K_l c)^2}\big)^{-1} & \text{Langmuir} \\ k_c\big(1 + \frac{\rho}{\theta}NK_f c^{N-1}\big)^{-1} & \text{Freundlich} \end{cases} \tag{1.49}$$

Thus, in contrast to linear sorption, nonlinear sorption induces nonlinear degradation.

In addition to nonlinear sorption nonlinearities in kinetics are often due to biological processes such as degradation by microorganisms. When microorganisms are involved in the degradation process good results have been obtained by the use of an enzymatical catalytic reaction model (see e.g. Richter et al., 1992 [105]; Richter et al., 1996 [104]). Thus, capacity limited degradation processes may be described by Michaelis-Menten kinetics

$$\frac{\partial}{\partial t}(\theta c + \rho s) = L(\theta, q)c - \theta\frac{V_{max}c}{c + K_M}, \tag{1.50}$$

where $V_{max}\ [ML^{-3}T^{-1}]$ represents the maximal reaction velocity and $K_M\ [ML^{-3}]$ the Michaelis constant. For $c \ll K_M$ the Michaelis-Menten term approximately obeys linear degradation with $V_{max}c/K_M \sim k_c c$. However, for $c \gg K_M$ the concentration c drops out and we end up with a reaction of zero order.

The influence of environmental parameters on degradation

So far the degradation rate has been assumed to be a constant. In the soil, however, kinetic processes are influenced by various environmental parameters, e.g. the organic carbon content or the pH-value of the soil. Degradation rates, in particular, are dependent on temperature and humidity. A comparatively simple approach to model temperature dependence is the Arrhenius law as it is used in chemical reaction kinetics. However, the activity of microorganisms does not increase above a certain temperature. This effect may be modeled by O'Neills' function (O'Neill, 1972 [97]) which is based on the assumption that an optimal temperature exists and that above and below this temperature the degradation rate decreases.

Similarly, the influence of humidity, i.e. the soil water content, on degradation can be formulated. In addition to the most commonly used model of Walker and Allen (1984 [137]), which assumes that an increase in the water content leads to an increase in the degradation, Nörtersheuser (1993 [96]) and Richter et al. (1996 [104]) propose a model that takes decreasing degradation near saturation into account.

The influence of temperature and humidity, for example, plays an important role in the degradation behavior of herbicides as several studies have shown (see e.g. Nörtersheuser, 1993 [96]; von Götz, 1997 [134]; Beulke, 1998 [14]; Aden et al., 1999 [3]). If required the models mentioned above can easily be added in ECOFIT and ECOPLAN.

1.3 Mathematical Modeling for Column, Lysimeter and Field Experiments

The meaningful description of upper and lower boundary conditions is a very difficult problem in the modeling of column, lysimeter and field experiments. Often the practitioner carrying out the experiments can only give a descriptive characterization of what happens at the top and the bottom of the soil core considered. In order to close this gap between a descriptive and a proper mathematical formulation, a lot of modeling work has to be done before coming up with reliable boundary conditions that well approximate experimental reality.

In this section, we first define different types of boundary conditions and discuss their physical meaning and their suitability for column, lysimeter and field experiments. Even though the main focus of this work is on column and mini-lysimeter studies which are carried out under controllable conditions a short overview of the additional difficulties that one encounters when modeling lysimeter or even field experiments is given.

1.3.1 Modeling of Initial and Boundary Conditions

In order to set up the initial boundary value problem for the PDEs

$$C(\psi_m)\frac{\partial \psi_m}{\partial t} = \frac{\partial}{\partial z}\Big[K(\psi_m)\frac{\partial}{\partial z}\big(\psi_m - \bar{w}z\big)\Big] + S \tag{1.51}$$

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\Big[\bar{D}(\theta)\frac{\partial \theta}{\partial z} - \bar{w}\bar{K}(\theta)\Big] + \bar{S} \tag{1.52}$$

$$\frac{\partial}{\partial t}(\theta c + \rho s) = \frac{\partial}{\partial z}\Big[\theta D_h \frac{\partial c}{\partial z} - qc\Big] + Q \tag{1.53}$$

both initial and boundary conditions have to be specified.

For the one-dimensional domain $\Omega = [z_0, z_e]$ under consideration initial values for the state variables y ($y = \psi_m, \theta, c$) at the time $t = 0$ are required:

$$y(t = 0, z) = y_0(z) \qquad z \in \Omega. \tag{1.54}$$

Moreover, initial values s_0 have to be provided in case that the solid phase s is described by a non-equilibrium isotherm, i.e. by an ODE.

Definition 1.1
In general, we can distinguish the following types of boundary conditions (see e.g. Zauderer, 1989 [144]; Schwarz, 1997 [110]) given by

$$a(t)\frac{\partial y(t,z)}{\partial \vec{n}} + b(t)y(t,z)\Big|_{\partial\Omega} = g(t), \tag{1.55}$$

where $a(t)$, $b(t)$ and $g(t)$ are given functions on the boundary $\partial\Omega$, $\vec{n}$ is the outward normal vector. In particular, notice that a and b are independent of the state variable y. If $a = 0$ and $b \neq 0$, (1.55) is referred to as a first-type or Dirichlet boundary condition. For $a \neq 0$ and $b = 0$ (1.55) is called a second-type or Neumann boundary condition. A linear combination of both conditions, i.e. $a \neq 0$ and $b \neq 0$, is named third-type or Cauchy boundary condition.

Water transport

In the following the different types of boundary conditions are discussed for the Richards equation in the matric potential form (1.51). The boundary conditions for the description in the water content form (1.52) can be derived analogously.

Dirichlet condition: prescribes the matric potential on the boundary $\partial\Omega$:

$$\psi_m(t,z)\Big|_{\partial\Omega} = f_1(t). \tag{1.56}$$

Neumann condition: prescribes a gradient normal to the boundary $\partial\Omega$:

$$\frac{\partial\psi_m(t,z)}{\partial\vec{n}}\Big|_{\partial\Omega} = f_2(t). \tag{1.57}$$

A special case of (1.57) is an impervious boundary where the gradient equals zero.

McCord (1991 [89]) points out that only the specified matric potential gradient (1.57) is a true second-type boundary condition. However, it has become standard in hydrologic literature to also refer to the *specified flux condition*

$$-\big[K(\psi_m)\frac{\partial}{\partial z}\big(\psi_m - \bar{w}z\big)\big]\Big|_{\partial\Omega} = \bar{f}_2(t) \tag{1.58}$$

as Neumann boundary condition.

Cauchy condition: prescribes a linear combination of the matric potential and the gradient:

$$a(t)\frac{\partial\psi_m(t,z)}{\vec{n}} + b(t)\psi_m(t,z)\Big|_{\partial\Omega} = f_3(t). \tag{1.59}$$

Upper boundary

For soil column experiments in the laboratory, where the boundary conditions can be controlled, conditions are used that can in general be described by Dirichlet conditions or specified flux conditions.

Considering experiments under natural climatic conditions, i.e. lysimeter or field studies, the situation is more complicated because the net infiltration flux through the soil surface has to be determined from the precipitation rate (plus eventually the irrigation rate), the interception rate due to crop canopy and the evapotranspiration rate. In most simulation tools for field water movement, e.g. SWATRER (Dierckx et al., 1986 [39]), AMBETI (Braden, 1995 [26]), SIMULAT (Diekkrüger et al., 1995 [37]), or PEARL (Tiktak, 1999[125]), the calculation of the evapotranspiration is based on the Penman-Monteith equation (Penman, 1949 [100]; Monteith, 1965 [93]).

As long as no saturation occurs at the upper boundary the specified flux condition (1.58) is used. For a saturated upper boundary infiltration is limited by the saturated hydraulic conductivity K_s. In this case Diekkrüger (1992 [36]) proposes to prescribe ψ_m, i.e. to switch to a Dirichlet boundary condition, in order to calculate the real infiltration and thus, based on the difference of rainfall and infiltration rates, the surface water run off.

Lower boundary

In soil column experiments frequently the matric potential at the bottom of the column is controlled which corresponds to a Dirichlet condition. In order to describe column outflow experiments also an infinite lower boundary condition, i.e. $\partial\psi_m(t,\infty)/\partial z = 0$ or $\psi_m(t,\infty) = 0$, may be assumed.

Investigating field studies in general three situations are distinguished (see e.g. Diels, 1994 [38]; Tiktak et al., 1999 [125]):

1. The groundwater level $\bar{L}$, where $\psi_m(t,\bar{L}) = 0$, is specified as a function time.

2. The matric potential ψ_m at the lower boundary is known as a function of time (Dirichlet condition).

3. The flux through the lower boundary is given as a function of time (specified flux condition).

   (a) Zero flux at the lower boundary, i.e. $\bar{f}_2(t) = 0$.

   (b) Free drainage condition at the lower boundary, i.e. $\partial\psi_m/\partial\vec{n} = 0$.
   Here it is assumed that the flux through the lower boundary equals the hydraulic conductivity $K(\psi_m)$ at the lower boundary. This implies that the matric potential ψ_m at the bottom of the column is constant with respect to depth and that the water flow is only driven by gravity. In order to satisfy this condition the water table must be sufficiently deep. Moreover, we have to guarantee by the choice of the depth of the lower boundary that the infiltrated water never reaches this lower boundary.

A good model for the lower boundary of a lysimeter is still a challenging task because none of the previous conditions can be applied. Due to the drainage or outlet system installed at the bottom of the lysimeter outflow normally only occurs when the bottom is saturated. To describe this situation often a mixture of Dirichlet and specified flux condition is used.

Solute Transport

For the convection-dispersion equation (1.72) the following three types of boundary conditions may be formulated:

Dirichlet condition: prescribes the concentration on the boundary $\partial\Omega$

$$c(t,z)\Big|_{\partial\Omega} = g_1(t) \tag{1.60}$$

or the convective transport over the boundary $\partial\Omega$

$$q(t,z)c(t,z)\Big|_{\partial\Omega} = \bar{g}_1(t). \tag{1.61}$$

Neumann condition: prescribes the concentration gradient on the boundary

$$\frac{\partial c(t,z)}{\partial \vec{n}}\Big|_{\partial\Omega} = g_2(t) \tag{1.62}$$

or the hydrodynamic dispersive flux over the boundary

$$-\theta(t,z)D_h(\theta,q)\frac{\partial c(t,z)}{\partial \vec{n}}\Big|_{\partial\Omega} = \bar{g}_2(t). \tag{1.63}$$

In practice, the specification of a Neumann condition is only possible for impervious boundaries, i.e $\partial c/\partial z = 0$.

Cauchy condition: prescribes the total flux over the boundary

$$J_T(t,z)\big|_{\partial\Omega} = -\theta(t,z)D_h(\theta,q)\frac{\partial c(t,z)}{\partial \vec{n}} + q(t,z)c(t,z)\Big|_{\partial\Omega} = g_3(t). \tag{1.64}$$

Upper boundary

For the boundary at the soil surface often a Cauchy-type boundary condition is chosen (Richter et al., 1996 [104]):

$$J_T(t,0) = -\theta(t,0)D_h(\theta(t,0),q(t,0))\frac{\partial c(t,0)}{\partial z} + q(t,0)c(t,0) = \begin{cases} q(t,0)c_0(t) & q(t,0) > 0 \\ 0 & q(t,0) \leq 0, \end{cases} \tag{1.65}$$

where $c_0(t)$ denotes the substance concentration in the irrigation water. Here it is assumed that the substance is applied to the soil surface together with the irrigation water, i.e. $q(t,0) > 0$. When irrigation is stopped, i.e. $q(t,0) = 0$, or when evaporation becomes an issue, i.e. $q(t,0) < 0$, no substance enters the soil system.

Alternatively, substance can be incorporated within the first centimeters of the soil column. This is modeled by the initial condition

$$c(t=0,z) = \begin{cases} c_0(z) & z \leq z_0 \\ 0 & z > z_0 \end{cases} \tag{1.66}$$

combined with a zero flux conditions $J_T(t,0) = 0$ at the upper boundary.

Lower boundary
At the lower boundary the hydrodynamic dispersion term $D_h(\theta, q)$ may be set to zero, i.e. only the convective transport along with the flow of water is then taken into account. Frequently, also an infinite lower boundary condition, i.e. $\partial c(t,\infty)/\partial z = 0$ or $c(t,\infty) = 0$, is chosen. The latter is usually used for the modeling of column outflow experiments where at the real lower boundary of the soil column flux concentrations are computed from resident concentrations (Parker and van Genuchten, 1984 [99], 1984 [131]).

1.3.2 Modeling of Outdoor Conditions

When field applications are considered this causes a host of new problems for modeling. Besides the models discussed for the transport of water, additional models are required, e.g. for the lateral movement of water due to sloped soil surface or for drainage systems which become active for high ground water levels. Moreover, the uptake of water by plant roots may essentially affect the transport of water.

In the case of crop cover also the description of the upper boundary for the transport of xenobiotics becomes more complicated. Depending on the application method, e.g. spraying, a part of the substance may be intercepted by the crop canopy or may drift to adjacent fields. Another part may dissipate on the soil surface due to volatilization before entering the soil system. Another source of loss is the substance uptake by plant roots. As mentioned in the previous section evapotranspiration plays an important role, which itself is dominated by plant growth. In contrast to column or mini-lysimeter experiments the temperature can not be assumed to be constant for field experiments. Being aware of the influence of soil temperature on degradation and transport this effect can not be neglected in the field scale. This implies that the heat conduction equation, taking into account the properties of the soil, has to be solved simultaneously with the water and solute transport equation in order to describe the temperature field adequately.

For most of these difficulties arising in field or even lysimeter studies models have been developed. Nevertheless, these complex processes, in particular in combination with heterogeneous soil structures, remain only partially understood and are by far not validated.

In principle, however, if models for more complex processes are available, they can be incorporated with relatively little effort into ECOFIT and ECOPLAN.

1.4 The Problem of Parameter Estimation in Environmental Fate Studies

Due to the lack of reliable and validated models for (large) lysimeter and field studies we concentrate in this work mainly on the environmental fate of xenobiotics on the column and mini-lysimeter scale which we assume to be governed by the Richards equation, the convection-dispersion equation, and the sorption and degradation processes discussed in the previous sections.

Simulation tools for both water and solute transport in the variably saturated soil are widely used. Enormous progress has been made in the development of powerful methods for the simulation of multi-component multi-phase flow in two or three dimensions. However, in order to meaningfully simulate the fate of substances reliable parameters are necessary. In contrast to the developments for simulation methods, the study of the inverse problem is up to now mainly limited to smaller test problems.

So far it is still common practice to derive the parameters by manually changing their values and running simulations until a satisfying agreement between simulated and observed curves is obtained. However, this trial-and-error approach may be very time consuming and tedious, particularly if more than two parameters have to be fitted simultaneously.

In contrast to this manual optimization, we will present in Chapter 2 an approach that supports the identification of unknown parameters on the basis of measurement data in an automated way by mathematical optimization methods.

In this section we first give a formulation of an inverse problem that covers a wide range of typical column and mini-lysimeter experiments. In our approach the parameter estimation problem is interpreted as a weighted least-squares problem constrained by a set of PDEs and ODEs. In addition, an overview of the tools used in current practice for the solution of the parameter estimation problem is presented and frequently encountered pitfalls are outlined. The need for the new solution methods that are developed in this thesis is motivated.

1.4.1 Formulation of the Inverse Problem

For the sake of clarity and due to the fact that the substances investigated in Chapter 5 are non-volatile, we restrict ourselves to the liquid and solid phase. However, the formulation given in the following can be easily extended to include the gaseous phase.

In the inverse modeling context we are facing the situation that measurement data η_{kij} for one or more species k ($k = \psi_m, \theta, c_l$) or functions of them is given, e.g. for the soil matric potential ψ_m, the water content θ or for the concentrations c_l of the different substances ($l = 1, \ldots, n$), for example parents and their metabolites. The data is recorded in depths z_j ($j = 1, \ldots, m_2$) at time points t_i ($i = 1, \ldots, m_1$). Depending on the experimental set-up state variables can be measured directly, e.g. water contents by TDR sensors and matric potentials by tensiometers, or indirectly by devices that are described by observation functions, e.g. time-averaged integrals of concentrations.

Point measurements of substance concentrations during the experiment are hardly accessible without disturbing the flow regime or even destroying the soil core. In general two types of measurements are available for the substance concentration. First of all, in column and mini-lysimeter experiments the outflow or also called the leachate, i.e. the volume of water with the dissolved substance that leaves the soil bottom, is collected and analyzed at certain time intervals. Secondly, by slicing the soil column respectively the mini-lysimeter at the end of the experiment concentration data for certain depths can be obtained. Depending on the substance, this option may be very tedious and expensive in particular if bound residues are expected.

In the practical studies investigated here, we assume that measurements can be described by observation functions b_k of the model responses $k(t,z)$ plus an error that follows a normal distribution with zero expectation $\epsilon_{kij} \sim \mathcal{N}(0, \sigma_{kij}^2)$:

$$\eta_{kij} = b_k\big(t_i, k(t_i, z_j), p, \bar{q}, u\big) + \epsilon_{kij}. \tag{1.67}$$

Here, $p \in \mathbb{R}^{n_p}$ denotes the parameters, $\bar{q} \in \mathbb{R}^{N_1}$ time-independent control variables and $u : [t_0, t_e] \to \mathbb{R}^{N_2}$ time-dependent control functions. In the context of parameter estimation, however, $\bar{q}$ and u are fixed.

The task of parameter estimation is to determine an unknown parameter vector p, e.g. van Genuchten parameters for the water transport, degradation rates or K_d-values for the solute transport, and solutions $k(t,z)$ that fit the data best:

$$\min \sum_{k=\psi_m,\theta,c_l} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} w_{kij} \frac{\big(\eta_{kij} - b_k(t_i, k(t_i, z_j), p, \bar{q}, u)\big)^2}{\sigma_{kij}^2} \tag{1.68}$$

such that

$$\begin{aligned}
C(\psi_m; p, \bar{q}, u)\frac{\partial \psi_m}{\partial t} &= \frac{\partial}{\partial z}\big[K(\psi_m; p, \bar{q}, u)\frac{\partial}{\partial z}\big(\psi_m - \bar{w}z\big)\big] + S(\psi_m; p, \bar{q}, u) && (1.69)\\
\frac{\partial \theta}{\partial t} &= \frac{\partial}{\partial z}\big[\bar{D}(\theta; p, \bar{q}, u)\frac{\partial \theta}{\partial z} - \bar{w}\bar{K}(\theta; p, \bar{q}, u)\big] + \bar{S}(\theta; p, \bar{q}, u) && (1.70)\\
\frac{\partial}{\partial t}\big(\theta c_l + \rho s_l\big) &= \frac{\partial}{\partial z}\big(\theta D_{h_l}(\theta; p, \bar{q}, u)\frac{\partial c_l}{\partial z} - q(p, \bar{q}, u)c_l\big) && (1.71)\\
&\quad + Q_l(\theta, c_1, \dots, c_n, s_1, \dots, s_n; p, \bar{q}, u) && (1.72)
\end{aligned}$$

$$+ \quad \text{ODEs for } s_l \ (l = 1, \dots, n) \tag{1.73}$$

$$+ \quad \text{initial and boundary conditions,} \tag{1.74}$$

are satisfied.

The term $Q_l(\theta, c_1, \dots, c_n, s_1, \dots, s_n; p)$ (1.72) summarizes all sources and sinks, i.e. all processes creating or consuming substances, e.g. first-order or Michaelis-Menten degradation terms $[ML^{-3}T^{-1}]$. The ODEs for s_l $(l = 1, \dots, n)$ describe interactions between liquid and solid phase concentrations due to the sorption processes described in Section 1.2.2.

The least-squares functional (1.68) has to be minimized under the infinite constraint that the PDEs for the water and the solute transport (1.69)-(1.73) including their initial and boundary conditions (1.74) are satisfied. The weights $w_{kij} \in \{0, 1\}$ describe a selection of measurements points out of a set of feasible ones. As it is the case for the control variables $\bar{q}$ and the control functions $u(t)$, they are fixed for parameter estimation. In the framework of optimal experimental design, however, these variables are free, i.e. they become optimization variables, as we will outline in Section 1.5. Here, the individual terms of the least-squares functional are weighted by the standard deviations σ_{kij} of the measurements η_{kij}.

To keep the formulation as general as possible both the ψ_m- and the θ-version of the Richards equation are presented. The version used depends on the experimental set-up and the data available. Note that the use of observation functions enables the simultaneous treatment of profile and also leachate outflow data.

1.4.2 Solution Methods in Current Practice

Inverse modeling has been used in groundwater hydrology for saturated flow problems for years (see e.g. Yeh, 1986 [142]). But due to the nonlinearities inherent to the equations describing unsaturated flow and transport processes, its application to the unsaturated zone is relatively new. Up to now, parameter estimation methods have been mainly used to determine soil hydraulic parameters in the water transport equation.

Estimation of unsaturated soil hydraulic functions started with studies of Zachmann et al. (1981 [143]), Dane and Hruska (1983 [35]) and Hornung (1983 [63]). Reviews can be found in Kool et al. (1987 [81]) and Hopmans and Šimůnek (1999 [61]). In the mid eighties inverse modeling was mainly used with one-step outflow experiments (see e.g. Kool et al., 1985 [80]; Parker et al., 1985 [98]; Kool and Parker, 1988 [79]; Toorman et al., 1992 [126]; van Dam et al., 1992 [127]) and was later expanded to multi-step outflow experiments (see e.g. van Dam et al., 1994 [128]; Durner et al., 1996 [45]). Several applications of parameter estimation to field experiments basing on e.g. tensiometer and TDR measurements (see e.g. Arning, 1994 [4]), tension disc infiltrometers experiments (see e.g. Šimůnek and van Genuchten, 1996 [114], 1997 [115]; Šimůnek et al., 1998 [112], 1999 [121]) extraction methods (see e.g. Inoue et al., [66]) or cone permeameter methods (see e.g. Gribb et al., 1996 [57]; Šimůnek et al., 1999 [113]) are also reported.

In contrast, parameter estimation for transient solute transport in the unsaturated zone described by coupled water and solute transport equations or even the simultaneous identification of soil hydraulic, solute transport and reaction parameters is barely reported (Mishra and Parker, 1989 [92]; Medina and Carrera, 1996 [90]; Abbaspour et al., 1997 [1]; Šimůnek and van Genuchten, 1999 [116]; Dieses et al., 1999 [41]; Dieses et al., 1999 [43]).

Several tools for inverse modeling in groundwater flow, e.g. MODFLOWP (Hill, 1998 [60]) and coupled water flow and solute transport, e.g. HYDRUS-1D (Šimůnek et al., 1998 [119]), HYDRUS-2D (Šimůnek et al., 1996 [118]), are available. It has become standard among hydrologists and soil scientists to estimate unknown parameters in these equations by coupling simulation tools, e.g. SWMS_2D (Šimůnek et al., 1994 [117]), PEARL (Tiktak et al.,

1999 [125]), SIMULAT (Diekkrüger et al., 1995 [37]), ESHPIM (Zurmühl and Durner, 1998 [146]) etc., with nonlinear optimization software packages, in particular implementations of the Levenberg-Marquardt (Marquardt, 1963 [88]) or the Gauss-Newton algorithm. Often used packages are e.g. PEST (Doherty, 1994 [44]), UCODE (Poeter and Hill, 1998 [102]), BMDP (Bard, 1974 [7]), MATLAB (Grace, 1992 [56]). Parameters are determined by repeated calls of the optimizer to the numerical simulator. Simulation, i.e. the solution of the forward problem, and optimization are treated as two separate procedures communicating only by a suitable interface.

At first glance this approach seems convenient and straightforward. However, frequently unsatisfactory results are obtained. Often estimation results are highly dependent on the initial guesses used. In particular, difficulties have been encountered with re-estimating the correct parameter values in nonlinear problems on the basis of undisturbed data when starting from only slightly perturbed parameter values. Moreover, these approaches are not able to account for problem-inherent structures which results in long computing times, even for small problems, i.e. coarse spatial grids. For this reason, spatial discretization with less than 50 spatial nodes are often used ignoring the spatial discretization error.

In Chapter 2 and 3, we will present solution methods that overcome these limitations (Dieses et al., 1999 [41]; 1999 [43]). As the formulation of the inverse problem (1.68)-(1.74) indicates, we pursue a different approach where the parameter estimation problem is interpreted as a least-squares problem that is constrained by coupled PDEs and ODEs. Both the simulation and the optimization problem are solved simultaneously. These methods provide the basis for the newly developed tool ECOFIT (Dieses et al., 1999 [41]; 1999 [43]) suitable for parameter estimation in nonlinear transport and degradation processes of xenobiotics in soils.

1.5 The Problem of Optimal Experimental Conditions in Environmental Fate Studies

Many authors who investigated parameter estimation for transport processes in the unsaturated zone encountered problems with ill-posedness of the inverse problems (see e.g. Hornung, 1983 [63]; Kool et al., 1985 [80]; Toorman et al., 1992 [126]; van Dam et al., 1992 [127]). They often described the situation that parameters were practically insensitive to observed data. Thus, it was concluded that data should be collected at points in time and space which showed a high sensitivity to the parameters under consideration.

One approach to determine points that deliver data that is most sensitive to parameters is optimal experimental design. Since the mid eighties, Yeh (1986 [142]), Kool et al. (1987 [81]), Kool and Parker (1988 [79]) and others claimed the need for optimal experimental design such as optimal allocation of sampling points or optimal input boundary conditions.

However, in practice experiments are often performed intuitively or with the objective that experimental conditions, e.g. irrigation schemes, resemble as much as possible outdoor conditions. Sampling schemes generally follow simple rules, e.g. measurements are carried out once per day or once a week. However, measurement data thus obtained might not be necessarily

suitable for parameter estimation. The contrary is frequently the case where inverse modeling on the basis of such data leads to unsatisfactory results or even to (nearly) singular parameter estimation problems, i.e. estimation of all required parameters is not possible. In the latter case, often regularization strategies are applied where e.g. in the simplest case some of the parameter values are fixed and only a subset of parameters is determined. Even if parameter estimation is possible, the resulting variances are often large and thus the estimates unreliable.

In contrast to this intuitive approach, the objective of optimal experimental design is to optimize experimental conditions and sampling schemes such that parameter estimation based on this data results in the best possible reliability of the estimates.

In this section, the optimal experimental design problem for the parameter estimation problem set up in the previous section is derived and formulated as an optimal control problem. The approach pursued is orientated by the one worked out in (Bauer et al., 1999 [9]; Körkel et al., [83]; Bauer et al., 2000 [10]). We discuss possible options to control respectively influence the system in the context of column and mini-lysimeter studies. A first insight into the mathematical treatment of the different optimization variables, i.e. control variables, control functions and weighting factors for measurement points, is given. In Section 1.5.2, we review the state-of-the-art in experimental design with respect to parameter estimation in water flow and solute transport processes. It will become obvious that due to the lack of suitable solution methods for these highly demanding problems, up to now hardly any practical optimal experimental design problems have been addressed. The development of new, sophisticated methods is indispensable if optimal experimental design should be established also in this area.

1.5.1 Formulation as Optimal Control Problem

In column and mini-lysimeter studies experimental conditions that can be optimized are, for example, the initial and the boundary conditions. In particular at the top of the soil column, i.e. at the upper boundary, the irrigation scheme and the concentrations of the added substance can generally be controlled and thus be optimized. The sampling scheme, e.g. where and when measurements are carried out, depends on the experimental set-up. Out of a set of feasible measurements, e.g. leachate or point measurements, the optimal sampling scheme has to be chosen.

In order to formulate the optimal control problem we consider the discretization of (1.68)-(1.74) in time and space. As it will be discussed in Section 2.1.1 and 2.1.2, this results in a finite dimensional, nonlinear constrained least-squares problem of the form

$$\min_{s,p} \|r_1(s,p)\|_2^2 \tag{1.75}$$

$$r_c(s,p) \;=\; 0. \tag{1.76}$$

Here, s is a finite dimensional vector parameterizing the solutions of the PDEs and ODEs. r_1 denotes the vector of residuals and r_c the equality conditions.

Sophisticated parameter estimation tools usually enable a statistical analysis of the solution, e.g. correlations, variances or confidence intervals for the parameter estimates. The variance-

covariance matrices C are generally derived from the linearized vector of residuals and constraints. The smaller the variances, i.e. the entries in the diagonal of C, the more reliable are the parameter estimates. Thus, a reasonable goal to improve estimation results is to identify sampling schemes and experimental conditions that are most likely to yield parameter estimates with low variances.

In the optimal experimental design problem different types of control/optimization variables arising from the description of the experimental conditions and the sampling scheme can be distinguished. In the specification of experimental conditions, control variables $\bar{q}_s$ $(s = 1, \ldots, N_1)$, such as initial concentrations, or time dependent control functions $u_l(t)$ $(l = 1, \ldots, N_2)$, such as water infiltration rates, can occur. For optimization purposes every control function $u_l(t)$ is parameterized, e.g. by piecewise constant or piecewise linear controls. For all control variables $\bar{q}$ and control functions $u(t)$ defining experimental conditions feasible ranges, in particular upper and lower bounds, are chosen according to the experimental possibilities.

Secondly, in order to derive optimal sampling schemes the total number of measurements, L_{max}, and a set of L_{fea} feasible measurement points $\Xi = \{\xi_r, r = 1, \ldots, L_{fea}\}$ have to be specified. With every possible measurement point ξ_r $(r = 1, \ldots, L_{fea})$ a weight w_r being 0 or 1 is associated. The aim is to determine an optimal sampling scheme

$$w = (w_1, \ldots, w_{L_{fea}})^T, \quad w_r \in \{0, 1\} \tag{1.77}$$

with $\sum_{r=1}^{L_{fea}} w_r = L_{max}$.

It is important to note that different assumptions are made for parameter estimation and optimal experimental design. In the context of parameter estimation, the given measurement data is obtained from a given design, i.e. all controls and weights are known and fixed. On the basis of this data parameters are optimized.

This is in contrast to optimal experimental design where control variables, control functions and weights are obtained by optimization for a fixed set of parameters p. Thus, here the variance-covariance matrix $C := C(s, p, \bar{q}, u, w)$ is considered as a function of the controls $\bar{q}$ and $u(t)$ and the weights w.

The variance-covariance matrix C is used for the description of the statistical quality of the estimated parameters. The objective of optimal experimental design is to minimize a function Φ of the variance-covariance matrix C of the underlying parameter estimation problem. Often, one of the classical design criteria, D-, A-, or E-optimality, which are outlined in Section 4.1.2, is used. The most common criterion, the D-optimality, for example, aims at the maximization of the determinant of the information matrix which is equivalent to the minimization of the determinant of the variance-covariance matrix (see e.g. Silvey, 1980 [111]).

The optimal experimental design problem can now be formulated as a nonlinear state-constrained optimal control problem:

$$\min_{s,\bar{q},u,w} \Phi(C(s, p, \bar{q}, u, w)) \tag{1.78}$$

$$r_c(s, p, \bar{q}, u) = 0. \tag{1.79}$$

Additionally, equality and inequality conditions on controls, weights and states can be formulated and are summarized as

$$e(s, p, \bar{q}, u, w) \left\{ \begin{array}{c} = \\ \geq \end{array} \right\}. \tag{1.80}$$

Note, that (1.78)-(1.80) describe a mathematically highly intricate problem. It is necessary to treat an objective function on the variance-covariance matrix which is implicitly defined by the Jacobian of the underlying parameter estimation problem.

1.5.2 Solution Methods in Current Practice

So far measurement points and experimental conditions that deliver data being highly sensitive to parameters have been mainly determined by either the analysis of response surfaces and/or the study of parameter sensitivities (see e.g. Toorman et al., 1992 [126]; van Dam et al., 1992 [127]; Šimůnek and van Genuchten, 1996 [114], 1997 [115]; Weiss and Smith, 1998 [139]; Inoue et al., 1998 [66]; Šimůnek et al, 1998 [120]).

The first approach uses two-dimensional response surfaces of objective functions. These are computed by evaluating the objective function for two parameters with varying values at a time while keeping the other parameters fixed. Even though this approach takes into account nonlinearities it is very tedious since the objective function has to be evaluated typically several hundred times for every parameter combination to provide sufficient data for contour plots. Moreover, even taking the computed cross-sections together, this gives only a rough idea about the complete response space spanned by all parameters (Šimůnek and van Genuchten, 1996 [114]).

In the second approach, sensitivity coefficients are calculated a priori from hypothetical data. According to the rule that high sensitivities correspond to well defined minima and small parameter uncertainties, measurement points in time and space are determined. Again this approach is very cumbersome.

Other authors studied the impact of different boundary conditions on parameter estimation results. Durner et al. (1999 [46]), for example, evaluated one-step, multi-step and continuous boundary conditions for inflow/outflow experiments and investigated their suitability to estimation of soil hydraulic properties on the basis of sensitivity studies.

Knopman and Voss (1987 [74]) were the first who considered experimental design problems for transport processes in soil in the framework of optimization problems by using statistical design criteria. According to the idea of optimal experimental design they aimed to determine sampling schemes that deliver observations being most sensitive with respect to unknown parameters according to some prescribed optimality criteria on the variance-covariance matrix. D-optimal sampling designs, i.e. optimal allocations of measurement points in time and space, for an one-dimensional solute transport model with two unknown parameters under steady-state conditions were found by Knopman and Voss (1987 [74]) by an enumerative grid search through a discretized design space. This concept was further extended to multi-objective sampling design which includes model discrimination, variance reduction in parameter estimation

and sampling cost (Knopman and Voss, 1989 [76]). Again, due to lack of suitable optimization procedures, this problem could only be solved by enumerative grid search. Knopman and Voss (1989 [76]) proposed that "to solve the optimization problem, integer programming could be used to identify optimal designs for both single and multi-objective problems". But they supposed that "even with a relatively simple set of constraints and linear objective functions, the multi-objective integer programming problem may be difficult to solve".

Some authors chose the minimization of experimental cost as objective function while the reliability of identified parameters was used as constraints (e.g. Nishikawa and Yeh, 1989 [95]; Hsu and Yeh, 1989 [64]). For a coupled two-dimensional water and reactive solute transport model Wagner (1995 [136]) presented a sampling strategy that minimizes model prediction uncertainty subject to a constraint on a given data collection budget. For the solution of the integer programming problem a branch-and-cut and a genetic algorithm were used.

In order to relax the uniqueness requirement of inverse solutions Sun and Yeh (1990 [124]) defined the concept of extended identifiability and showed their relation to experimental design.

Sun (1994 [123]) gave a general formulation of an optimal experimental design problem in ground-water modeling that contained two types of decision variables. He distinguished decision variables associated with system excitation, e.g. irrigation rates, time periods of irrigation, concentration in the irrigation water, and those associated with system observation, e.g. locations and time points of measurements. Some of the variables are described by continuous functions such as irrigation schemes. Other variables, e.g. measurement points, take discrete values or were even binary. Sun (1994 [123]) stated that "in practice to solve such a general problem is too difficult" and "that up to date, only few hypothetical examples and simplified cases have been reported".

To our knowledge, so far no study has addressed the problem of optimizing both sampling schemes and experimental conditions, e.g. boundary conditions such as irrigation schemes, simultaneously for coupled water and reactive solute transport in the unsaturated zone on the basis of optimal experimental design theory. In Chapter 4 an approach is presented that allows to solve this highly demanding class of problems. This approach has been implemented in the tool ECOPLAN suitable for optimal experimental design in nonlinear transport and degradation processes of xenobiotics in soils (Dieses et al., 2000 [42]).

Chapter 2

Fast Numerical Methods for the Solution of Large Scale Parameter Estimation Problems

In this chapter we present a very efficient *reduced* Generalized Gauss-Newton method for the solution of parameter estimation problems constrained by PDEs and ODEs. So far, the application of the Generalized Gauss-Newton method developed by Bock (1981 [19], 1983 [20], 1987 [21]) has been mainly limited to parameter estimation problems constrained by ODEs and DAEs. In addition, modifications of the Generalized Gauss-Newton method have been developed that are suitable e.g. for large scale inverse problems in ODEs/DAEs (Schlöder, 1988 [108]), or for multi-experiment problems in ODEs/DAEs (Schlöder and Bock, 1983 [109]; Schlöder, 1988 [108]; von Schwerin, 1998 [135]). Zieße et al. (1996 [145]) applied the Generalized Gauss-Newton method to systems constrained by PDEs. In order to handle the huge amount of computational work arising from the discretization of PDEs a parallel approach was used (Gallitzendörfer, 1997 [53]).

Our approach, implemented in the newly developed tool ECOFIT (Dieses et al., 1999 [41]; 1999 [43]) for parameter estimation in multi-species transport reaction systems, rests on two pillars:

- a *reduced* Generalized Gauss-Newton method and
- a very efficient strategy for the generation of the required derivatives.

Content of the current chapter is the derivation of the *reduced* Generalized Gauss-Newton method. In Chapter 3, we will present a very efficient approach for the computation of the derivatives where in general most of the CPU time is spent.

This chapter is structured as follows: In the first section we discuss approaches for the discretization in space and time. In Section 2.2 the Generalized Gauss-Newton method (Bock, 1981 [19]; 1983 [20]; 1987 [21]) is reviewed. It will be shown that the problem sizes induced by the PDEs require a specially tailored approach. In Section 2.3 we present a special form

of the *reduced* approach of Schlöder (1988 [108]) which enables us to handle these large scale problems by exploiting given structures on several levels. How the corresponding variance-covariance matrices of the solutions are computed is outlined in Section 2.4.

2.1 Problem Discretization

2.1.1 Spatial Discretization

In order to solve the least-squares problem (1.68)-(1.74) numerically the PDEs (1.69)-(1.72) and the ODEs (1.73) describing infinite constraints have to be transformed into finite ones. We employ the method of lines in its vertical version (see e.g. Großmann and Ross, 1992 [58]) to ensure modularity and flexibility of the developed tools ECOFIT and ECOPLAN. The idea of the method of lines, which has become a common approach for solving one-dimensional time-dependent PDEs, is to reduce the initial boundary problem to a system of ODEs in time by discretizing in space. In a second step, the resulting, in general stiff initial value problem is solved by an appropriate integration method. The convenience and power of the method of lines lies in the fact that highly developed solvers for stiff ODEs and DAEs have become available.

The spatial discretization may have various forms, e.g. finite difference, finite element or finite volume schemes. For the regular geometry considered in this work we choose finite differences. In our approach the public domain spatial differentiation routines of order four DSS004 and DSS020 by Schiesser (1991 [107]) are applied. For parabolic-type components in the PDEs, i.e. the diffusion-dispersion term, the routine DSS004 based on central differences schemes, is used. Approximations of hyperbolic-type components, i.e. the convective term, may cause numerical problems. Good results have been obtained with the five point biased upwind finite difference scheme of Carver and Hinds (1978 [30]) implemented in the routine DSS020 (Schiesser, 1991 [107]). Details about these routines and studies about their performance can be found in Appendix A.

Depending on the spatial activity exhibited by the system sufficiently fine spatial grids are required to guarantee a certain degree of accuracy. Basically there are two ways to handle this problem. First of all, methods may be used which attempt to adjust the spatial grid automatically. Over the past decade the interest in adaptive spatial discretization methods for the solution of the forward problem in the framework of the method of lines has rapidly increased (see e.g. Verwer et al., 1989 [132]; Furzeland et al., 1990 [52], Berzins and Ware, 1996 [13]; Li et al., 1998 [85]; Berzins et al., 1998 [12]). Adaptive methods are very attractive with regard to simulation and are thus widely used. Their application in the context of optimization, however, is more difficult and subject of current research. The development of reliable adaptive spatial discretization methods suitable for parameter estimation using the Generalized Gauss-Newton method is a topic on its own and out of the scope of this work.

Here the primary focus is on applications and on the reliable solution of real-life problems. We pursue in the following an approach that is based on a fixed spatial grid. However, in order to handle high spatial activity a sufficiently high spatial resolution for the whole domain, i.e. several hundred up to thousand spatial nodes, may be necessary leading to a large scale

optimization problem. Since the parameter estimation problem as such does not change during the optimization process problem-inherent structures can be exploited on several levels. This enables us to design a very efficient parameter estimation code, even for systems consisting of several thousand ODEs.

Discretizing the parameter estimation problem (1.68)-(1.74) on a fixed spatial grid results in a least-squares problem constrained by ODEs of the form

$$\min ||\tilde{F}(y(t,p))||_2^2 = \min \sum_{i,j} \sigma_{ij}^{-2} \big(\tilde{\eta}_{ij} - y_i(t_j,p)\big)^2 \tag{2.1}$$

$$\dot{y}(t) = f(t,y,p) \qquad y(t_0) = y_0 \tag{2.2}$$

with $f : [t_0,t_e] \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_y}$ and $n_y = N \times (n_{PDE} + n_{ODE})$, where N, n_{PDE} and n_{ODE} denote the number of spatial nodes, PDEs (1.69)-(1.72) and ODEs (1.73), respectively.

2.1.2 The Multiple Shooting Method for Parameterization in Time

As outlined in Section 1.4.2 it is most widespread to solve inverse problems in PDEs/ODEs by combining tools for the solution of the forward problem with iterative optimization routines. In general, an initial value problem approach is used.

In contrast to this, we pursue the approach of Bock (1981 [19]; 1983 [20]) who interpreted the parameter estimation problems of type (2.1)-(2.2) as a constrained multipoint boundary value problem. For its solution a boundary value problem approach using multiple shooting was developed. The idea behind this method is not to integrate the ODEs (2.2) on the complete interval $[t_0,t_e]$ but to divide it into m subintervals $[\tau_j,\tau_{j+1}]$ $(j = 0,\ldots,m-1)$:

$$t_0 = \tau_0 < \tau_1 < \ldots < \tau_m = t_e. \tag{2.3}$$

Additional variables s_k are introduced at each multiple shooting node. On the basis of the initial values s_k initial value problems are formulated on each of the m subintervals

$$\dot{y}(t) = f(t,y,p) \qquad y(\tau_k) = s_k, \qquad k = 0,\ldots,m-1. \tag{2.4}$$

Thus, an only piecewise continuous solution $y(t;s_k,p)$ is obtained. Continuity of the final solution is enforced by requiring the additionally defined matching conditions

$$h_k(s_k,s_{k+1},p) = y(\tau_{k+1};s_k,p) - s_{k+1}, \qquad k = 0,\ldots,m-1 \tag{2.5}$$

to be satisfied.

It should be noted that the multiple shooting method guarantees the existence of a start solution on the complete interval by choosing adequately small subintervals and reasonable initial guesses. In contrast, the single shooting method may lead to singularities in case of bad initial guesses for parameters. Thus, integration until the end of the time interval may be impossible. This situation is often reported for real-life problems where due to poor first initial guesses the solver fails even though the parameter estimation problem as such is well-posed for

the true parameters. These pitfalls of the single shooting method may also be encountered, for example, for the Richards equation which is not defined when ψ_m respectively $C(\psi_m)$ becomes zero.

Example 2.1.1
We consider the following system describing water transport for a typical column outflow experiment. For the soil hydraulic functions the van Genuchten-Mualem parameterization (1.13)-(1.14) is used.

$$C(\psi_m)\frac{\partial \psi_m}{\partial t} = \frac{\partial}{\partial z}\big[K(\psi_m)\frac{\partial}{\partial z}\big(\psi_m - \bar{w}z\big)\big] \tag{2.6}$$

with

$$\begin{aligned} K(\psi_m) &= K_s \frac{(1-(\alpha|\psi_m|)^{n-1}(1+(\alpha|\psi_m|)^n)^{1/n-1})^2}{(1+(\alpha|\psi_m|)^n)^{\frac{1-1/n}{2}}} \\ C(\psi_m) &= \alpha(n-1)(\theta_s-\theta_r)(\alpha|\psi_m|)^{n-1}(1+(\alpha|\psi_m|)^n)^{1/n-2} \end{aligned} \tag{2.7}$$

- *Initial condition:* $\psi_m(0,z) = -670, \qquad z \geq 0$
- *Upper boundary:*

$$q(t,0) = -K(\psi_m(t,0))\frac{\partial}{\partial z}\big(\psi_m(t,0) - \bar{w}z\big) = \begin{cases} 0.2 & t \leq 6.0 \\ 0.0 & t > 6.0 \end{cases}$$

- *Lower boundary:* $\partial\psi_m(t,20)/\partial z = 0, \qquad t \geq 0$

Suppose $n = 1.2\ [1]$, $\alpha = 0.0102\ [hPa^{-1}]$, $K_s = 10.0\ [cm\,h^{-1}]$, $\theta_s = 0.4\ [1]$ *and* $\theta_r = 0.05\ [1]$ *(Seppelt, private communication; Vink et al., 1994 [133]) are the true van Genuchten parameters. For these values the forward problem can be solved without any problems. Measurement data is generated for several depth and several points in time by simulation and - for test purposes - no error was added. The aim is to estimate the parameters* n, α *and* K_s. *However, starting the parameter estimation for the values* θ_s *and* θ_r *as given above and with the only slightly perturbed initial guesses* $n = 1.3$, $\alpha = 0.0107$ *and the correct value* $K_s = 10$ *the integration and thus the parameter estimation fails. Due to the initial guesses,* ψ_m *goes to zero at the upper boundary and the problem becomes undefined at* $t = 5.9$. *Using multiple shooting, i.e. by restarting the integration on each subinterval, the trajectory is prevented to drift too far away from the true trajectory, i.e.* ψ_m *is prevented from becoming 0. Employing 20 multiple shooting nodes, even for initial guesses* $n = 1.4$, $\alpha = 0.015$ *and* $K_s = 20$ *convergence to the true parameter values is reached within 10 iterations.*

Another important advantage of multiple shooting over single shooting is that multiple shooting allows for the use of prior information. In particular, measurement data is not only used in the objective function (2.1) but it may also be exploited for selecting the initial guesses

for the initial values s_k. By this means, it is easily possible to incorporate available information about the solution. Moreover, multiple shooting reduces the nonlinearity of the problem.

Parameterizing the semi-discretized parameter estimation problem (2.2)-(2.1) in time results in a finite dimensional, nonlinear constrained structured least-squares problem in the augmented vector $x = (s_0, \dots, s_m, p)^T \in \mathbb{R}^n$ with $n = (m+1)n_y + n_p$

$$\min_x ||r_1(s_0, \dots, s_m, p)||_2^2 \tag{2.8}$$

$$r_2(s_0, p) = 0 \tag{2.9}$$

$$\bar{r}_3(s_0, \dots, s_m, p) = 0. \tag{2.10}$$

Here, $r_1 \in C^3(D, \mathbb{R}^{n_1})$, $D \subset \mathbb{R}^n$, denotes the n_1 least-squares conditions and $r_2 \in C^3(D, \mathbb{R}^{n_2})$ the n_2 initial conditions. $\bar{r}_3 = (r_3, h_0, \dots, h_{m-1})^T \in C^3(D, \mathbb{R}^{\bar{n}_3})$ describes other possible equality conditions $r_3 \in C^3(D, \mathbb{R}^{n_3})$ including the m matching conditions h_i, altogether $\bar{n}_3 = mn_y + n_3$.

Note that these parameter estimation problems arising from the discretization of PDEs are large scale. Considering, for example, a problem constrained by 3 PDEs and employing 400 spatial nodes and 10 multiple shooting nodes, we end up with an optimization problem in 12000 variables. Commonly available nonlinear least-squares solvers are not capable of coping with such problem sizes.

However, in the parameter estimation problems arising from column and mini-lysimeter studies the initial values s_0 are generally known, i.e. r_2 is independent of $s_1, \dots, s_m$. This induces a special structure of the Jacobian which is exploited by the methods derived in Section 2.3 in order to reduce the computational effort for the solution of the problem.

2.2 The Generalized Gauss-Newton Method

For the solution of nonlinear constrained least-squares problems of the type just discussed, Bock (1981 [19]) proposed a generalization of the Gauss-Newton method which originally was only applicable to unconstrained least-squares problems. In the following the Generalized Gauss-Newton method (Bock, 1981 [19], 1983 [20], 1987 [21]) is presented which has proven to be stable and efficient for a series of real-life parameter estimation problems constrained by ODEs and DAEs. A big advantage of the Generalized Gauss-Newton approach is that it is not only superior to "lower order methods", e.g. manual optimization, but also to "higher order methods", e.g. SQP methods. It shows good linear local convergence behavior, i.e. the convergence rate improves as better data and model fit. In contrast to SQP methods which require second order information for quasi-linear convergence behavior, only first order information has to be provided.

First, the main steps of the Generalized Gauss-Newton algorithm, which is the core routine of the parameter estimation tool PARFIT (Bock, 1981 [19], 1983 [20], 1987 [21]), are outlined. After a summary of the underlying theory including optimality criteria and convergence results, an estimation of the computational effort and the storage requirements is given. It will become

obvious that for the large scale systems the standard implementation is not suitable and that more efficient strategies are required.

2.2.1 Algorithm

The most important steps of the algorithm applied to the nonlinear constrained least-squares problem (2.8)-(2.10) are:

1. Start with an initial guess x_0.

2. Improve the solutions iteratively by

$$x_{k+1} = x_k + \lambda_k \Delta x_k,$$

where Δx_k solves the constrained linear problem

$$\min_{\Delta x} ||r_1(x_k) + J_1(x_k)\Delta x_k||_2^2 \tag{2.11}$$

$$r_2(x_k) + J_2(x_k)\Delta x_k = 0 \tag{2.12}$$

$$\bar{r}_3(x_k) + \bar{J}_3(x_k)\Delta x_k = 0 \tag{2.13}$$

with the Jacobians $J_i(x) := \partial r_i(x)/\partial x$ $(i = 1, 2)$ and $\bar{J}_3(x) := \partial \bar{r}_3(x)/\partial x$. $\lambda_k \in (0, 1]$ denotes a relaxation factor of a globalization strategy.

The corresponding Jacobian $J(x) = \partial r/\partial x$ is large scale and shows the typical block structure due to the multiple shooting discretization:

$$J = \begin{pmatrix} J_1 \\ J_2 \\ \bar{J}_3 \end{pmatrix} = \begin{pmatrix} D_1^0 & D_1^1 & \dots & \dots & D_1^m & D_1^p \\ D_2^0 & 0 & \dots & \dots & 0 & D_2^p \\ D_3^0 & D_3^1 & \dots & \dots & D_3^m & D_3^p \\ G_0 & -I & & & & G_0^p \\ & G_1 & -I & & 0 & \vdots \\ & & \ddots & \ddots & & \vdots \\ & 0 & & G_{m-1} & -I & G_{m-1}^p \end{pmatrix} \qquad r = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ h_0 \\ \vdots \\ h_{m-1} \end{pmatrix} \tag{2.14}$$

$$D_i^j = \partial r_i(s_0, \dots, s_m, p)/\partial s_j, \quad D_i^p = \partial r_i(s_0, \dots, s_m, p)/\partial p,$$
$$G_i = \partial y(\tau_{i+1}; s_i, p)/\partial s_i, \qquad G_i^p = \partial y(\tau_{i+1}; s_i, p)/\partial p,$$

with

$$\begin{aligned} D_1^j &\in \mathbb{R}^{n_1 \times n_y}, & D_1^p &\in \mathbb{R}^{n_1 \times n_p}, \\ D_2^j &\in \mathbb{R}^{n_2 \times n_y}, & D_2^p &\in \mathbb{R}^{n_2 \times n_p}, \\ D_3^j &\in \mathbb{R}^{n_3 \times n_y}, & D_3^p &\in \mathbb{R}^{n_3 \times n_p}, \\ G_i,\, I &\in \mathbb{R}^{n_y \times n_y}, & G_i^p &\in \mathbb{R}^{n_y \times n_p} \end{aligned}$$

Here the block columns of J describe the derivatives with respect to the variables s_i and p. The block rows with the D_i matrices denote the derivatives of the functionals r_i with respect to the initial values s_i, where the last column contains the derivatives with respect to the parameters p. The block rows with the G_i matrices correspond to the derivatives of the continuity conditions h_i.

Note, that for the particular situation considered here, where the initial conditions $r_2(s_0, p)$ only depend on s_0 and p, a special structure in the second block row is induced. This can be exploited as we will see in Section 2.3.

2.2.2 Optimality Criteria

For the ease of presentation, we summarize all equality conditions in $r_c = (r_2^T, \bar{r}_3^T)^T \in C^3(D, \mathbb{R}^{n_c})$, $n_c = n_2 + \bar{n}_3$, and rewrite the nonlinear problem (2.8)-(2.10) as

$$\min_x ||r_1(x)||_2^2 \tag{2.15}$$

$$r_c(x) \quad = \quad 0. \tag{2.16}$$

Analogously, the linear problem (2.11)-(2.13) is of the form

$$\min_{\Delta x} ||r_1(x_k) + J_1(x_k)\Delta x_k||_2^2 \tag{2.17}$$

$$r_c(x_k) + J_c(x_k)\Delta x_k \quad = \quad 0. \tag{2.18}$$

Definition 2.1 (Constraint Qualification [CQ])
The Constraint Qualification [CQ] is satisfied if

$$Rg(J_c(x)) = n_c. \tag{2.19}$$

A point $x \in D \subset \mathbb{R}^n$ is said to be regular if x satisfies the [CQ].

Definition 2.2 (Positive Definiteness [PD])

$$Rg \begin{pmatrix} J_1(x) \\ J_c(x) \end{pmatrix} = n \qquad \forall x \in D \tag{2.20}$$

Then the following equivalence holds:

$$[PD] \Longleftrightarrow w^T J_1(x)^T J_1(x) w > 0 \quad \forall w \in Ker\, J_c / \{0\} \tag{2.21}$$

Definition 2.3 (Lagrangian Function)
The Lagrangian function for the nonlinear problem (2.15)-(2.16) is defined as

$$L(x, \lambda) = \frac{1}{2}||r_1(x)||_2^2 - \lambda^T r_c(x), \tag{2.22}$$

where $\lambda \in \mathbb{R}^{n_c}$ is called Lagrange multiplier.

Theorem 2.4 (Necessary Conditions)
Let x^* be regular and a solution of (2.15)-(2.16). Then x^* is feasible, i.e. $r_c(x^*) = 0$, and the first order necessary condition holds: There exists a multiplier vector λ^* that uniquely solves the stationary condition

$$\frac{\partial}{\partial x} L(x^*, \lambda^*) = r_1(x^*)^T J_1(x^*) - \lambda^{*T} J_c(x^*) = 0. \tag{2.23}$$

Additionally, the second order necessary condition is satisfied: For all directions

$$\forall w \neq 0 \qquad J_c(x^*)w = 0$$

the Hessian matrix of L is positive semi-definite

$$w^T H(x^*, \lambda^*) w \geq 0, \qquad H(x^*, \lambda^*) := \frac{\partial^2}{\partial x^2} L(x^*, \lambda^*). \tag{2.24}$$

Definition 2.5 (Karush-Kuhn-Tucker Point (KKT Point))
A vector (x^*, λ^*) which is feasible and satisfies the stationary condition (2.23) is called a Karush-Kuhn-Tucker point (KKT point).

Lemma 2.6
If $[CQ]$ and $[PD]$ are satisfied, then the following equivalence holds

(x^*, λ^*) *is a KKT point of (2.15)-(2.16)* $\iff$ $(0, \lambda^*)$ *is a KKT point of (2.17)-(2.18),*

i.e. a KKT point of the nonlinear problem (2.15)-(2.16) is a fixed point of the Gauss-Newton iteration.

A great advantage of the Generalized Gauss-Newton method is that a generalized inverse J^+ can be defined which solves the linearized problem (2.17)-(2.18). This enables a uniform theoretical treatment of nonlinear equation systems, unconstrained and constrained least-squares problems as we will see in the next section.

Theorem 2.7 (Existence of a Generalized Inverse)
If $J = (J_1^T, J_c^T)^T$ in the linear constrained least-squares problem (2.17)-(2.18) satisfies the conditions $[CQ]$ and $[PD]$, then it follows:

1. *For any $r = (r_1^T, r_c^T) \in \mathbb{R}^{n_1+n_c}$ exists exactly one Karush-Kuhn-Tucker point $(\Delta x, \lambda)$ for (2.17)-(2.18), and Δx is a strict minimum.*

2. *There exists a linear mapping $J^+ : \mathbb{R}^{n_1+n_c} \to \mathbb{R}^n$, such that*

   $$\Delta x = -J^+ r \tag{2.25}$$

   solves (2.17)-(2.18) for any $r \in \mathbb{R}^{n_1+n_c}$ with

   $$J^+ = \begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} J_1^T J_1 & J_c^T \\ J_c & 0 \end{pmatrix}^{-1} \begin{pmatrix} J_1^T & 0 \\ 0 & I \end{pmatrix}. \tag{2.26}$$

3. *The solution operator J^+ is a generalized inverse and satisfies the defining condition*

   $$J^+ = J^+ J J^+. \tag{2.27}$$

2.2.3 Convergence Results

The *Local Convergence Theorem* of Bock (1987 [21]) presented in the following gives conditions for the local convergence of the full step Generalized Gauss-Newton method ($\lambda_k = 1$) and quantifies it. This theorem also holds for the unconstrained case, where J^+ becomes the Moore-Penrose pseudo-inverse, and for nonlinear equation systems $F(x) = 0$, where J^+ reduces to the normal inverse J^{-1}. In the latter case, convergence can also be shown for approximations of J^{-1}.

Theorem 2.8 (Local Convergence Theorem)
Let $J(x)^+$ be the generalized inverse of the Jacobian $J(x) = (J_1(x)^T, J_c(x)^T)^T$ with respect to the function $r = (r_1^T, r_c^T)^T \in C^1(D, \mathbb{R}^{n_1+n_c})$ in a nonlinear constrained least-squares problem. The Jacobian J respectively its generalized inverse J^+ satisfy the following Lipschitz conditions:

$$||J(y)^+\big(J(x+t(y-x)) - J(x)\big)(y-x)|| \leq \omega t||y-x||^2, \qquad \omega < \infty \tag{2.28}$$

$$||\big(J(z)^+ - J(x)^+\big)R(x)|| \leq \kappa(x)||z-x|| \leq \kappa||z-x|| \tag{2.29}$$

$$R(x) := r(x) - J(x)J(x)^+r(x) \qquad (\text{residuum}) \qquad \kappa < 1$$

$\forall\, t \in [0,1],\ \forall\, x, y, z \in D$ *and* $x - y = J(x)^+r(x)$.

Then for all $x_0 \in D$ with

$$\delta_0 := \frac{\alpha_0\omega}{2} + \kappa < 1, \qquad \alpha_j := ||J(x_j)^+r(x_j)|| \tag{2.30}$$

$$D_0 := K\big(x_0, \frac{\alpha_0}{1-\delta_0}\big) \subset D \tag{2.31}$$

follows:

1. *The iteration*

$$x_{j+1} = x_j + \Delta x_j, \qquad \Delta x_j = -J(x_j)^+r(x_j) \tag{2.32}$$

   is well-defined and remains in D_0.

2. *$x_j \to x^* \in D_0$ $(j \to \infty)$ with $J(x^*)^+r(x^*) = 0$.*

3. *The convergence is linear with*

$$||\Delta x_{j+1}|| \leq \big(\frac{\alpha_j\omega}{2} + \kappa\big)||\Delta x_j|| =: \delta_j||\Delta x_j||.$$

4. For the jth iteration the following a priori estimation holds

$$||x_j - x^*|| \leq \delta_0^j \frac{\alpha_0}{1-\delta_0}.$$

Remark 2.1
The Lipschitz constant ω characterizes the nonlinearity of the model. Its inverse limits the region where the linearization is valid. Therefore ω is of great importance for the selection of the relaxation factor λ_k of the damped Generalized Gauss-Newton method (Bock, 1981 [19]; 1987 [21]).

The Lipschitz constant of J^+, κ, measures the compatibility of the data with the model. For $\kappa < 1$, the fixed point x^* is not only a stationary point but a strict local minimum. In contrast to this, for $\kappa \geq 1$, there may exist perturbations of the measurement data in same order of magnitude as the measurements, such that the fixed point is stationary but not a minimum (Bock, 1987 [21]).

Considering real-life parameter estimation problems which are highly complex and nonlinear, in general no initial guesses x_0 can be provided that are sufficiently close to the solution to satisfy the requirements of the *Local Convergence Theorem.* Thus, it may happen that the start increment points out of the region where the linear approximation is valid.

In order to extend the convergence domain, the iteration (2.32) is relaxed to

$$x_{k+1} = x_k + \lambda_k \Delta x_k \qquad \lambda_k \in (0, 1] \tag{2.33}$$

in each iteration. The relaxation factor λ_k is chosen according to special globalization strategies (Bock, 1987 [21]; Bock et al., 2000 [22]) in order to come closer to the solution.

2.2.4 Estimation of the Computational Effort

According to the Generalized Gauss-Newton algorithm in Section 2.2.1, the solution of the nonlinear constrained least-squares problem (2.8)-(2.10) is obtained iteratively by solving linear least-squares problems of the form (2.11)-(2.13). The computation of the augmented vector of variables

$$\Delta x = (\Delta s_0, \ldots, \Delta s_m, \Delta p) = -J^+ r \tag{2.34}$$

is done by recursive decomposition methods whereby systems of dimension $n_y + n_p$ are generated. In the following, the main steps for the solution of the linearized subproblems are summarized (Bock 1981 [19], 1987 [21]).

In order to solve the linear subproblem (2.8)-(2.10) of dimension $(n_y(m+1) + n_p)$, we first reduce the large but sparse and structured system to a smaller one by condensing. The idea is not to directly solve the original system but to iteratively exploit structures by employing suitable transformations. Using, e.g. a block Gauss elimination, which can be interpreted as a

transformation of the original system to a triangular form, we end up with a condensed problem of dimension $(n_y + n_p)$ depending only on the initial values Δs_0 and the parameters Δp:

$$\min_{\Delta s_0, \Delta p} ||E_1 \Delta s_0 + P_1 \Delta p + u_1||_2^2 \tag{2.35}$$

$$E_2 \Delta s_0 + P_2 \Delta p + u_2 = 0 \tag{2.36}$$

$$E_3 \Delta s_0 + P_3 \Delta p + u_3 = 0 \tag{2.37}$$

Hereby, the matrices E_l and P_l and the vectors u_l $(l = 1, 3)$ can be computed according to the following scheme (Bock, 1987 [21]; Schlöder, 1988 [108]):

$$E_l = D_l^0 + \sum_{i=1}^{m} D_l^i \prod_{j=1}^{i} G_{i-j}, \qquad l = 1, 3 \tag{2.38}$$

$$E_2 = D_2^0 \tag{2.39}$$

$$P_l = D_l^p + \sum_{i=1}^{m} D_l^i \Big(\sum_{j=0}^{i-1} \Big(\prod_{k=1}^{i-1-j} G_{i-k} \Big) G_j^p \Big), \qquad l = 1, 3 \tag{2.40}$$

$$P_2 = D_2^p \tag{2.41}$$

$$u_l = r_l + \sum_{i=1}^{m} D_l^i \Big(\sum_{j=0}^{i-1} \Big(\prod_{k=1}^{i-1-j} G_{i-k} \Big) h_j \Big), \qquad l = 1, 3 \tag{2.42}$$

$$u_2 = r_2 \tag{2.43}$$

For the sake of clarity we introduce $\prod_{j=l}^{n} M_j := M_l \cdot M_{l+1} \ldots M_n$ for $l \leq n$ and $\prod_{j=l}^{n} M_j := I$ for $l > n$.

In order to compute Δs_0 and Δp the condensed system (2.35)-(2.37) is reduced to a triangular form using equivalence transformations:

$$\begin{pmatrix} E_2 & P_2 \\ E_3 & P_3 \\ E_1 & P_1 \end{pmatrix} = \begin{pmatrix} \bar{J}_c \\ \bar{J}_1 \end{pmatrix} = \begin{pmatrix} T & 0 \\ L & Q \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \\ 0 & 0 \end{pmatrix}. \tag{2.44}$$

In a first step, $\bar{J}_c$, i.e. the part associated with equality conditions, is reduced to an upper triangular form by a suitable transformation T, e.g. by Gauss elimination or orthogonal transformation. By the use of the matrix L the variables corresponding to the triangular matrix R_{11} are eliminated from the least-squares part. In a second step, the remaining least-squares part is triangularized by an orthogonal transformation Q, e.g. Householder transformation, and the increments Δs_0 and Δp are computed.

After the solution of the condensed system (2.35)-(2.37) the other increments $\Delta s_1, \ldots, \Delta s_m$ are computed by the forward recursion

$$\Delta s_{j+1} = G_j \Delta s_j + G_j^p \Delta_p + h_j, \qquad j = 0, \ldots, m-1. \tag{2.45}$$

In the standard implementation of the Generalized Gauss-Newton method, PARFIT (Bock, 1987 [21]), the right hand side r and all derivative matrices D_i^j and G_i in the Jacobian are computed in each iteration. Basically, in order to derive the condensed system (2.35)-(2.37) explicitly all matrix and vector operations described in (2.38)-(2.43) have to be carried out. In PARFIT, however, given structures in the D_i matrices are exploited.

In summary, in each Generalized Gauss-Newton step the following work has to be done: In order to evaluate the right hand side r the ODE system arising from semi-discretization (2.2) has to be integrated. Additionally, to set up the Jacobian the solution of $n_y + n_p$ variational differential equations respectively varied trajectories are required. As $n_p \ll n_y$ we can state that the computational effort is of order $\mathcal{O}(n_y)$, i.e. it is linear in the number of spatial nodes. Also the storage requirements for the Jacobian are determined by the spatial grid. A single storage of the Jacobian requires

$$(n_1 + n_3) \times (n_y(m+1) + n_p) + n_y \times (n_y + n_p) \times m + n_2 \times (n_y + n_p) = \mathcal{O}(n_y^2) \quad (2.46)$$

memory cells.

A simple estimation illustrates that for the degradation and transport processes studied here the standard approach leads to an unacceptable computational effort and to enormous storage requirements even for moderately fine spatial grids.

Example 2.2.1
Let us consider a least-squares problem constrained by two PDEs (e.g. water and solute transport). 6 parameters are unknown. For a spatial discretization with 400 nodes we obtain a system of 800 ODEs. A multiple shooting parameterization with 10 multiple shooting nodes gives an optimization problem in 8006 variables. Suppose that the parameters have to be estimated on the basis of 40 measurements. Not considering given structures in the D_i matrices, according to (2.46) the corresponding Jacobian then would have 7445040 entries. Assuming that 8 bytes are required per entry, already 56 MB would be necessary just to store the Jacobian once.

Hence, to treat these large scale problems in thousand of variables efficiently non-standard techniques are necessary. One way to cope with the huge bulk of computational work required is by parallelizing. This was done by Zieße et al. (1996 [145]) employing the parallel version PARFIT/MP (Gallitzendörfer, 1997 [53]). In this approach the evaluation of the right hand side r and the computation of the D_i^j and G_i matrices are performed completely in parallel on each subinterval. Moreover, on the linear algebra level parallelism is used.

We want to pursue a different approach which exploits given structures to reduce the dimension of the optimization problem. This reduced approach is outlined in the next section.

2.3 The Reduced Approach

Parameter estimation problems in PDEs result in large scale optimization problems for which very efficient solution methods are required. The approach presented in the following is a

special case of a reduced Generalized Gauss-Newton method which was developed and implemented in the code FIXFIT by Schlöder (1988 [108]). The idea of Schlöder's reduced approach is to exploit equality conditions to reduce the computational effort for the solution of the problem.

In the newly developed parameter estimation tool for multi-species transport reaction systems, ECOFIT (Dieses et., 1999 [41]; 1999 [43]), we exploit the fact that the initial conditions for all states are fixed. Thus, the systems have only few degrees of freedom, namely as many as unknown parameters are present. In contrast to the standard procedure discussed in the previous section, directional derivatives are therefore used for setting up the linear subproblems.

2.3.1 An Efficient Condensing Algorithm

For the sake of presentation we assume in the following that either the s_i directly enter r_l or that separability for $r_l = \sum_{i=0}^{m} r_l^i(s_i, p)$ is given which implies separability for $D_l^p = dr_l/dp = \sum_{i=0}^{m} r_l^i(s_i,p)/dp =: \sum_{i=0}^{m} {}^iD_l^p$. Then we can add the parameters p to each multiple shooting variable s_i and consider a problem in $\bar{x}_i = (s_i^T, p_i^T)^T$, $p_i = p_{i+1}$, instead. This formally leads, compared to the formulation in (2.14), to a slightly modified Jacobian $\bar{J}$ and right hand side $\bar{r}$:

$$\bar{J} = \begin{pmatrix} \bar{D}_1^0 & \bar{D}_1^1 & \dots & \dots & \bar{D}_1^m \\ \bar{D}_2^0 & 0 & \dots & \dots & 0 \\ \bar{D}_3^0 & \bar{D}_3^1 & \dots & \dots & \bar{D}_3^m \\ \bar{G}_0 & -\bar{I} & & & \\ & \bar{G}_1 & -\bar{I} & & 0 \\ & & \ddots & \ddots & \\ & 0 & & \bar{G}_{m-1} & -\bar{I} \end{pmatrix} \qquad \bar{r} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \bar{h}_0 \\ \vdots \\ \bar{h}_{m-1} \end{pmatrix} \tag{2.47}$$

with

$$\bar{D}_i^j := \begin{pmatrix} D_i^j & {}^jD_i^p \end{pmatrix} \in \mathbb{R}^{n_i \times (n_y + n_p)} \tag{2.48}$$

$$\bar{G}_i := \begin{pmatrix} G_i & G_i^p \\ 0 & \hat{I} \end{pmatrix} \in \mathbb{R}^{(n_y+n_p)\times(n_y+n_p)} \tag{2.49}$$

$$\bar{h}_i : = \begin{pmatrix} y(\tau_{i+1}; s_i, p_i) - s_{i+1} \\ p_i - p_{i+1} \end{pmatrix} \tag{2.50}$$

$$\bar{I} \in \mathbb{R}^{(n_y+n_p)\times(n_y+n_p)} \qquad \hat{I} \in \mathbb{R}^{n_p \times n_p}. \tag{2.51}$$

Applying block Gauss elimination to (2.47) results in the same condensed system as in Section 2.2.4

$$\min_{\Delta s_0, \Delta p} ||E_1 \Delta s_0 + P_1 \Delta p + u_1||_2^2 \tag{2.52}$$

$$E_2 \Delta s_0 + P_2 \Delta p + u_2 = 0 \tag{2.53}$$

$$E_3 \Delta s_0 + P_3 \Delta p + u_3 = 0. \tag{2.54}$$

This time, however, slightly different formulas for P_l and u_l are obtained

$$E_l = D_l^0 + \sum_{i=1}^{m} D_l^i \prod_{j=1}^{i} G_{i-j} \qquad l = 1,3 \tag{2.55}$$

$$E_2 = D_2^0 \tag{2.56}$$

$$P_l = {}^0D_l^p + \sum_{i=1}^{m} \Big\{ {}^iD_l^p + D_l^i \big(\sum_{j=0}^{i-1} \big(\prod_{k=1}^{i-1-j} G_{i-k} \big) G_j^p \big) \Big\} \qquad l = 1,3 \tag{2.57}$$

$$P_2 = D_2^p = {}^0D_2^p \tag{2.58}$$

$$u_l = r_l^0 + \sum_{i=1}^{m} \Big\{ r_l^i + D_l^i \big(\sum_{j=0}^{i-1} \big(\prod_{k=1}^{i-1-j} G_{i-k} \big) h_j \big) \Big\} \qquad l = 1,3 \tag{2.59}$$

$$u_2 = r_2 = r_2^0, \tag{2.60}$$

with $\prod_{j=l}^{k} M_j := M_l \cdot M_{l+1} \ldots M_k$ for $l \leq k$ and $\prod_{j=l}^{k} M_j := I$ for $l > k$. Decomposing D_l^p resp. r_l on each subinterval into its separable components ${}^iD_l^p$ resp. r_l^i is a crucial prerequisite for the use of directional derivatives as outlined in the following.

Assuming that D_2^0 has full rank - which is the case in many practical applications - we can formally eliminate the variables fixed by the initial conditions

$$\Delta s_0 = -(D_2^0)^{-1} [{}^0D_2^p \Delta p + r_2^0]. \tag{2.61}$$

This results in a reduced condensed system with only few degrees of freedom, i.e. number of parameters n_p.

$$\min_{\Delta p} ||\tilde{P}_1 \Delta p + \tilde{u}_1||_2^2 \tag{2.62}$$

$$\tilde{P}_3 \Delta p + \tilde{u}_3 = 0, \tag{2.63}$$

where

$$\tilde{P}_l = P_l + E_l \big[-(D_2^0)^{-1} \cdot {}^0D_2^p \big] \qquad l = 1,3 \tag{2.64}$$

$$= D_l^0 [-(D_2^0)^{-1} \cdot {}^0D_2^p] + {}^0D_l^p + \tag{2.65}$$

$$\sum_{i=1}^{m} \Big\{ {}^iD_l^p + D_l^i \big[(\prod_{j=1}^{i} G_{i-j}) [-(D_2^0)^{-1} \cdot {}^0D_2^p] + \sum_{j=0}^{i-1} \big(\prod_{k=1}^{i-1-j} G_{i-k} \big) G_j^p \big] \Big\} \tag{2.66}$$

$$\tilde{u}_l = u_l + E_l \big[-(D_2^0)^{-1} r_2^0 \big] \tag{2.67}$$

$$= D_l^0 [-(D_2^0)^{-1} r_2^0] + r_l^0 + \tag{2.68}$$

$$\sum_{i=1}^{m} \Big\{ r_l^i + D_l^i \big[(\prod_{j=1}^{i} G_{i-j}) [-(D_2^0)^{-1} r_2^0] + \sum_{j=0}^{i-1} \big(\prod_{k=1}^{i-1-j} G_{i-k} \big) h_j \big] \Big\}. \tag{2.69}$$

2.3.2 Efficient Evaluation by Directional Derivatives

The explicit evaluation of $\tilde{P}_l$ and $\tilde{u}_l$ according to (2.64)-(2.69) is very expensive if - as it is the case in the standard procedure - all expressions D_l^i, G_i etc. are independently evaluated and if all matrix and vector products are explicitly computed. Basing on the work of Schlöder (1988 [108]), the explicit computation and storage of D_l^i, G_i etc. is avoided in the approach used here by a successive evaluation of directional derivatives. Defining recursively for $k = 0, \ldots, m-1$

$$
\begin{aligned}
{}^{-1}G_p &:= -(D_2^0)^{-1} \cdot {}^0D_2^p \qquad & {}^kG_p &:= G_k \cdot {}^{k-1}G_p + G_k^p && (2.70)\\
{}^{-1}G_r &:= -(D_2^0)^{-1} r_2^0 \qquad & {}^kG_r &:= G_k \cdot {}^{k-1}G_r + h_k && (2.71)
\end{aligned}
$$

the forward recursion for $i = 1, \ldots, m$ and $l = 1, 3$

$$
\begin{aligned}
\hat{P}_l^0 &:= D_l^0 \cdot {}^{-1}G_p + {}^0D_l^p \qquad & \hat{P}_l^i &:= \hat{P}_l^{i-1} + D_l^i \cdot {}^{i-1}G_p + {}^iD_l^p && (2.72)\\
\hat{u}_l^0 &:= D_l^0 \cdot {}^{-1}G_r + r_l^0 \qquad & \hat{u}_l^i &:= \hat{u}_l^{i-1} + D_l^i \cdot {}^{i-1}G_r + r_l^i && (2.73)
\end{aligned}
$$

results in the required

$$
\tilde{P}_l = \hat{P}_l^m \qquad \tilde{u}_l = \tilde{u}_l^m. \tag{2.74}
$$

If the initial values s_0 are given explicitly, $D_2^0 = I$. For problems, however, where neither the s_i directly enter r_l nor the separability assumption made above holds, it is necessary to store all the directions ${}^iG_p$ and ${}^iG_r$ until a complete forward integration is performed to obtain the required information for r_l and thus for the recursion (2.72)-(2.73).

The great advantage of the reduced approach is that for the recursion presented above the computation of only $n_p + 1$ directional derivatives is needed, namely n_p for ${}^kG_p$ resp. $D_l^i \cdot {}^{i-1}G_p$ and an additional one for ${}^kG_r$ resp. $D_l^i \cdot {}^{i-1}G_r$. How these directional derivatives are obtained is outlined in the following for ${}^kG_p$.

Instead of formulating the j th column $({}^kG_p)_j$ as a product of G_k with the vector $({}^{k-1}G_p)_j$ plus a vector $(G_k^p)_j$, we interpret it as a directional derivative:

$$
({}^kG_p)_j = \begin{pmatrix} G_k & G_k^p \end{pmatrix} \begin{pmatrix} {}^{k-1}G_p \\ \hat{I} \end{pmatrix}_j = \frac{d}{d\epsilon} y\Big(\tau_{k+1}, \bar{x}_k + \epsilon \begin{pmatrix} {}^{k-1}G_p \\ \hat{I} \end{pmatrix}_j\Big)\Big|_{\epsilon=0}, \tag{2.75}
$$

with $\bar{x}_k = (s_k^T, p^T)^T$.

These directional derivatives may be approximated by the difference quotient between the varied and the nominal trajectory as it is done in FIXFIT (Schlöder, 1988 [108]):

$$
({}^kG_p)_j = \frac{y(\tau_{k+1}, \bar{x}_k + \epsilon W_j^k) - y(\tau_{k+1}, \bar{x}_k)}{\epsilon} + \mathcal{O}(\epsilon). \tag{2.76}
$$

where $W^k := \begin{pmatrix} {}^{k-1}G_p \\ \hat{I} \end{pmatrix}$.

In our approach we compute $({}^{k}G_p)_j$ by solving a so called variational differential equation (VDE):

$$\begin{pmatrix} \dot{G}_k & \dot{G}_k^p \end{pmatrix} W_j^k = \begin{pmatrix} f_y & f_p \end{pmatrix} \bar{G}_k W_j^k \qquad\qquad \bar{G}_k(\tau_k) W_j^k = W_j^k, \tag{2.77}$$

with $\bar{G}_k = \begin{pmatrix} G_k & G_k^p \\ 0 & \hat{I} \end{pmatrix} \in \mathbb{R}^{(n_y+n_p)\times(n_y+n_p)}$. Formally, this VDE (2.77) is solved together with

$$\dot{y} = f(t, y, p) \qquad\qquad y(\tau_k) = s_k. \tag{2.78}$$

An advantage of the VDEs (2.77) over the varied trajectories (2.76) is, that if the directional derivatives of type

$$\begin{pmatrix} f_y & f_p \end{pmatrix} \bar{d}, \qquad\qquad \bar{d} \in \mathbb{R}^{n_y+n_p} \tag{2.79}$$

can be obtained analytically, it is possible to compute $({}^{k}G_p)_j$ without the approximation error $\mathcal{O}(\epsilon)$. An efficient way to solve (2.77) is discussed in Chapter 3. The same strategy is applied to the other matrices ${}^{k}G_r$, $D_l^i \cdot {}^{i-1}G_p$ and $D_l^i \cdot {}^{i-1}G_r$.

Once the reduced condensed system (2.62)-(2.63) is set up, equivalence transformations are applied to the Jacobian $\tilde{J} := (\tilde{P}_3^T, \tilde{P}_1^T)^T$ and the right hand side $\tilde{u} := (\tilde{u}_3, \tilde{u}_1)$ reducing the system to a triangular form

$$\begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta p_1 \\ \Delta p_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \begin{matrix} \} \, n_3 \\ \\ \} \, n_1 \end{matrix} \tag{2.80}$$

with $\Delta p = (\Delta p_1^T, \Delta p_2^T)^T$. The solution is given by $\Delta p_2 = R_{22}^{-1} v_2$ and $\Delta p_1 = R_{11}^{-1}(v_1 - R_{12} v_2)$. Finally, in order to solve the original linear subproblem (2.11)-(2.13), we compute the missing unknown increments $\Delta s_1, \ldots, \Delta s_m$, which were eliminated by condensing, with low computational effort by forward recursion:

$$\Delta s_j = {}^{j-1}G_p \Delta p + {}^{j-1}G_r, \quad j = 1, \ldots, m. \tag{2.81}$$

In summary, we can state the following main results for the reduced approach: For the solution of the linear subproblem (2.11)-(2.13) now only n_p+1 instead of the n_y+n_p directional derivatives as required in the standard approach are necessary. Thus, the number of derivatives, which generally cause the main bulk of computational work, is independent of the spatial grid. Due to the fact that only n_p+1 directional derivatives are needed, we end up with essentially the same computational effort as for the single shooting method while maintaining the advantages of multiple shooting.

2.4 Sensitivity Analysis of the Solution

In addition to the solution of the parameter estimation problem, a statistical analysis of the solution is of great practical importance. The estimated parameters are only meaningful if an estimate of their reliability, i.e. variance-covariance matrices or confidence intervals, can be provided. In the current section, we outlined how variance-covariance matrices and confidence intervals (Bock, 1987 [21]) can be derived in the framework of the Generalized Gauss-Newton method. This provides the basis for the solution of optimal experimental design problems which aim at the minimization of a function on the variance-covariance matrix.

Let J be the Jacobian of the complete system (2.14) in the solution point x^* and J^+ the respective generalized inverse. Under the assumption about the measurement error made in (1.67), with the vector of residuals r_1, the solution $\Delta x = -J^+ r$ is also a random variable. An approximation of the corresponding variance-covariance matrix is given by

$$C(x) := \mathbb{E}(\Delta x \Delta x^T) = \mathbb{E}(J^+ r r^T J^{+T}) = J^+ \mathbb{E}(r r^T) J^{+T} = \beta^2 J^+ \Lambda J^{+T}, \tag{2.82}$$

where

$$\mathbb{E}(r r^T) = \beta^2 \begin{pmatrix} I_{n_1} & 0 \\ 0 & 0 \end{pmatrix} = \beta^2 \Lambda. \tag{2.83}$$

In the framework of the Generalized Gauss-Newton method an approximation of the variance-covariance matrix for the parameters p is easily accessible. Using equivalence transformations as described in (2.44) resp. (2.80) for the solution of the reduced condensed system (2.62)-(2.63), we can write the Jacobian $\tilde{J} = (\tilde{P}_3^T, \tilde{P}_1^T)^T$ as

$$\tilde{J} = \tilde{T} R = \begin{pmatrix} T & 0 \\ L & Q \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \\ 0 & 0 \end{pmatrix}. \tag{2.84}$$

The generalized inverse of $\tilde{J}$ can be formulated as

$$\tilde{J}^+ = R^+ \tilde{T}^{-1}, \tag{2.85}$$

where

$$R^+ = \begin{pmatrix} R_{11}^{-1} & -R_{11}^{-1} R_{12} R_{22}^{-1} & 0 \\ 0 & R_{22}^{-1} & 0 \end{pmatrix}. \tag{2.86}$$

Due to its special form, the variance-covariance matrix can then be easily computed by

$$C(p) = \beta^2 \tilde{J}^+ \tilde{\Lambda} \tilde{J}^{+T} = \beta^2 R^+ \tilde{T}^{-1} \tilde{\Lambda} \tilde{T}^{-T} R^{+T} = \beta^2 R^+ \tilde{\Lambda} R^+ = \beta^2 \bar{C}, \tag{2.87}$$

with $\tilde{\Lambda} = \begin{pmatrix} 0 & 0 \\ 0 & I_{n_1} \end{pmatrix}$. The standard deviations of the parameters p are given by

$$\sigma(p_i) = \beta \bar{C}_{ii}^{\frac{1}{2}}, \qquad i = 1, \ldots, n_p. \tag{2.88}$$

In a similar way, the standard deviations for the variables $s_0, \ldots, s_m$ can be obtained (Schlöder, 1988 [108]). These are of particular interest when the quality of the solution of states has to be assessed for which no data is available.

For the case that the common factor β^2 is unknown, an independent estimate can be obtained by

$$b^2 := \frac{||r_1(s_0, \ldots, s_m, p)||_2^2}{l_2}, \tag{2.89}$$

where $l_2 := n_1 - l_1$ and $l_1 := (n_y + n_p) - (n_2 + n_3)$. Approximations of the corresponding confidence intervals θ_i can be computed by

$$\theta_i = b\big(\bar{C}_{ii} l_1 F_{l_1,l_2;1-\alpha}\big)^{\frac{1}{2}} \qquad i = 1, \ldots, n_p, \tag{2.90}$$

where $F_{l_1,l_2;1-\alpha}$ denotes the $(1-\alpha)$-quantile of the F_{l_1,l_2}-distribution.

Chapter 3

Efficient Generation of Derivatives in Discretized PDE-Systems

The primary focus of this chapter is on efficient strategies for the computation of accurate derivatives. In the framework of the reduced Generalized Gauss-Newton method by far the largest part of computational work is caused by the computation of the directional derivatives required to generate the reduced condensed system (2.62)-(2.63). Even though the explicit computation of the derivative matrices D_i^j, G_i, etc. is avoided by the reduced approach, there remain $n_p + 1$ directional derivatives to be computed in each iteration. Thus, in addition to solving the stiff initial value problem (2.2), sufficiently accurate directional derivatives have to be provided to ensure the convergence to the correct solution.

In ECOFIT, for the solution of the stiff initial value problems, the integrator DAESOL (Bauer, 1999 [11]; 2000 [8]), suitable for ODEs and DAEs of index 1, is used. DAESOL is a multi-step method code with a variable step size and order control based on Backward Differentiation Formulae (BDF). So far, DAESOL has been the only tool that provides the solution of the forward problem as well as the computation of both first and second order derivatives within the framework of *Internal Numerical Differentiation* (Bock, 1981 [19]). However, in order to handle also large scale parameter estimation problems arising from PDEs efficiently, further developments are necessary.

In Section 3.1, we first overview commonly used, state-of-the-art methods for the computation of the different types of derivatives required. It will be shown that even when applying these highly sophisticated methods to our class of problems high CPU times are needed, unless only small sized problems with coarse spatial grids are considered. Analyzing the performance reveals that the computational effort for parameter estimation is mainly dominated by the frequent and expensive computation of f_y, i.e. the derivative of the right hand side f with respect to the states y.

Two strategies are developed to remove this bottleneck. In Section 3.2, we first present a modified Newton method which permits a significant reduction of the number of f_y-computations by substituting them by directional derivatives of type $f_y w$. In a second step, Section 3.3, a specially tailored strategy is derived to efficiently generate the remaining f_y. A significant

speed up is obtained by combining both strategies resulting in a very fast parameter estimation code for large scale problems involving several thousand ODEs.

3.1 Computation of Derivatives Based on IND

In this section, we study the different types of derivatives required in the course of the reduced Generalized Gauss-Newton method and discuss methods for their generation. Basically, we can distinguish two classes of derivatives:

- Derivatives resp. directional derivatives of the right hand side f with respect to the states y and to the parameters p, such as f_y, $f_y w$ and f_p.
- Sensitivity matrices: Directional derivatives of the solution $y(t; s_i, p)$ with respect to s_i and p , such as ${}^kG_p$, ${}^kG_r$, $D_l^i \cdot {}^{i-1}G_p$ and $D_l^i \cdot {}^{i-1}G_r$.

In the first section, the BDF discretization scheme for the solution of the forward problem is reviewed and state-of-the-art methods for the computation of f_y required by the BDF method are discussed. In Section 3.1.2, after summarizing the pitfalls frequently encountered in the computation of sensitivity matrices, we outline their computation in the context of Internal Numerical Differentiation (Bock,1981 [19]). Finally, in Section 3.1.3, the performance of the state-of-the-art integrator DAESOL applied to a representative soil column experiment is investigated.

3.1.1 Derivatives Required by the BDF Method

The evaluation of the least-squares conditions r_1 and the equality conditions r_2 and r_3 requires the solution of the forward problem (2.2). Using a BDF method, the derivatives of the right hand side f of the initial value problem (2.2) with respect to the states y have to be provided in each iteration.

In the $(n+1)$st step the BDF discretization scheme is of the form

$$F(y_{n+1}) := \sum_{i=0}^{k} \alpha_i y_{n+1-i} + h f(t_{n+1}, y_{n+1}, p) = 0. \tag{3.1}$$

This nonlinear equation system in y_{n+1} is solved by a modified Newton method. The evaluation of the corresponding Jacobian

$$J(y_{n+1}) := \alpha_0 I + h f_y(t_{n+1}, y_{n+1}, p) \tag{3.2}$$

requires the expensive computation of f_y. In DAESOL, a monitoring strategy is used which keeps the Jacobian frozen as long as possible to avoid the expensive evaluation and decomposition of $J(y_{n+1})$ in each BDF step.

In small-sized problems, the computation of f_y is of minor importance. Considering large scale systems, however, the efficient generation of f_y plays the key role in the development of

a fast parameter estimation code as we will outline in Section 3.2 and 3.3. In the following, we overview commonly used methods for computing f_y.

Essentially, we can distinguish four approaches frequently applied in common practice to generate derivatives of an explicitly given function $u(x)$, such as the right hand side f: *hand-coding, symbolic differentiation, finite differences and automatic differentiation.* A straightforward way to compute the derivative of a function $u(x)$ with respect to x is by *hand-coding*. This produces accurate and generally efficient code. However, hand-coding may be a tedious and very error-prone process, especially if highly complex functions are involved. A major practical drawback is that whenever the original function is changed, the derivative-code has to be modified too. Another commonly used option to obtain derivatives is by employing a *symbolic differential* tool such as Mathematica (Wolfram, 1997 [140]) or Maple (Gander and Hrebicek, 1993 [54]). Despite their powerful capabilities for manipulating algebraic expressions, they are ill-suited for handling routines with branches, loops, or calls to subroutines. Due to its combinatorial character, symbolic differentiation requires enormous resources, even for small sized problems. The most popular approach to approximate derivatives is to use *finite differences*. The derivative of $u(x)$ with respect to the ith component at a point x_0 can be approximated, e.g. by first-order accurate forward differences

$$\frac{\partial u(x)}{\partial x_i}\Big|_{x=x_0} = \frac{u(x_0 + \eta e_i) - u(x_0)}{\eta} + \mathcal{O}(\eta), \tag{3.3}$$

where e_i denotes the ith unit vector and η a perturbation factor. Critical in this context is the choice of a reasonable perturbation factor η. On the one hand, a small η is needed to minimize the truncation error, on the other hand the subtraction of nearly identical function values may lead to numerical cancellation.

In ECOFIT, we generate the required derivatives of the right hand side f with respect to states y and parameters p by *automatic differentiation*, which will be shown to be very advantageous in our case. Automatic differentiation is essentially based on the fact that each function can be interpreted as a sequence of elementary operations such as addition, multiplication, or elementary functions as for example sine or cosine. By successively applying the chain rule to the function made up of elementary functions, e.g.

$$\frac{\partial}{\partial x} u(g(x))\Big|_{x=x_0} = \Big(\frac{\partial}{\partial s} u(s)\Big|_{s=g(x_0)}\Big)\Big(\frac{\partial}{\partial x} g(x)\Big|_{x=x_0}\Big), \tag{3.4}$$

the derivatives computed are correct up to machine precision. In general, two ways of propagating derivatives in automatic differentiation can be distinguished. In the forward mode, derivatives of intermediate values are computed with respect to the input variables. As the runtime and the storage requirements of the forward mode are roughly linear in the number of input variables, it is very efficient for computing derivatives of a large number of output variables with respect to few input variables. In the reverse mode, which is in a certain sense complementary to the forward mode, derivatives of the final result with respect to the intermediate values are generated. Here, the runtime is linear in the number of output variables. The storage requirements, however, are more difficult to access as all intermediate values have to be stored. The

reverse mode is particularly suited to derive sensitivities of a small number of output variables with respect to a large number of input variables.

In ECOFIT, we use the automatic differentiation tool ADIFOR (Bischof et al., 1992 [15], 1994 [16], 1998 [17]), which is based on a hybrid forward/reverse mode approach. As we will see in the course of this chapter, in our case it is very advantageous that ADIFOR produces a product $f_y \times S$, instead of generating f_y as such. Hereby, S denotes the so called seed matrix which has to be initialized by the user. If the full Jacobian f_y is needed, the seed matrix S is set to the identity matrix I. Directional derivatives $f_y w$ are obtained by setting $S = w$.

3.1.2 Generation of Directional Derivatives Using IND

By far the most CPU time is spent on the computation of the sensitivity matrices ${}^kG_p$, ${}^kG_r$, $D_l^i \cdot {}^{i-1}G_p$, and $D_l^i \cdot {}^{i-1}G_r$, required to set up the linear subproblem (2.62)-(2.63). In each of the iterations directional derivatives of the general type

$$G(t;\nu)w = \frac{\partial y(t;\nu)}{\partial \nu} w = \frac{d}{d\eta} y(t;\nu+\eta w)\Big|_{\eta=0} \qquad \nu = (y_0, p) \tag{3.5}$$

with

$$\dot{y}(t) = f(t,y,p) \qquad y(t_0) = y_0 \tag{3.6}$$

are needed. These can be obtained e.g. by finite differences of the form

$$G(t;\nu)w = \frac{y(t;\nu+\eta w) - y(t;\nu)}{\eta} + \mathcal{O}(\eta). \tag{3.7}$$

In contrast to the finite difference scheme used in the previous section for an explicitly given function $u(x)$, the situation becomes more complicated when finite differences are applied to the output $y(t;\nu)$ of an integrator. When using a sophisticated integrator with an automatic step size and order control, the solution $y(t;\nu)$ is generally a discontinuous function of the input, i.e. the initial values y_0 and the parameters p. Applying finite differences to such an integrator output generally leads to poor derivatives, unless both the nominal trajectory $y(t;\nu)$ and the varied trajectory $y(t;\nu+\eta w)$ are computed with an extremely high accuracy. This approach is also referred to as *External Numerical Differentiation (END)*. As a rule of thumb, to gain a certain degree of accuracy in the derivatives the square of this accuracy has to be provided for the solution of $y(t;\nu)$ leading to an enormous computational effort.

This fact, however, may cause problems when using so called model-independent, derivative-free parameter estimation codes such as PEST (Doherty, 1994 [44]). There, parameters are estimated by repeated calls of the optimization routine to the simulation routine for the solution of the forward problem. Typically, these optimization routines treat the simulator as a black box. Consequently, unless integrators with a fixed step size and order are used, the derivatives are generated according to the concept of External Numerical Differentiation, employing finite difference schemes of type (3.7). This implies that high integration accuracies are necessary for

the computation of $y(t;\nu+\eta w)$ and $y(t;\nu)$ in order to avoid poor derivatives and thus wrong parameter estimates. In addition, these optimization routines are generally extremely inefficient since possible saving potentials in the linear algebra due to similar structures in the nominal and in the perturbed solution are not exploited.

A remedy is provided by computing the derivatives according to the concept of *Internal Numerical Differentiation (IND)* which was introduced by Bock (1981 [19]). The basic idea of this approach is to compute the derivatives of the discretization scheme as such, rather than the derivatives of the solution trajectory $y(t;\nu)$. Regarding the finite difference scheme (3.7), IND implies the computation of the varied trajectory $y(t;\nu+\eta w)$ with the same step size and order as the nominal trajectory $y(t;\nu)$. As linear algebra components of the nominal trajectory can be reused for the varied trajectory, the computational effort can be drastically reduced. Bock (1981 [19]) has shown this approach to be stable especially for low integration accuracies. In contrast to END, in order to calculate a required degree of accuracy for the derivatives, only solutions of this accuracy have to be provided.

A realization of this strategy, i.e. finite difference schemes using IND, was applied by Schlöder (1988 [108]) for the computation of the sensitivity matrices required by the reduced approach.

In our context, it is advantageous to compute the required derivatives of the general form (3.5) by solving - within the framework of IND - a variational differential equation (VDE). Due to the fact that we can obtain directional derivatives of type $f_y w$ analytically, i.e. by automatic differentiation, this approach enables us to get rid of the accuracy limit $\mathcal{O}(\eta)$ in (3.7) which is inherent to finite difference schemes.

In the following, we outline the concept of IND for the solution of the VDE within the BDF discretization scheme. Differentiating the initial value problem (3.6) with respect to $\nu = (y_0, p)$, we obtain for the sensitivity matrix $G(t;\nu) = \partial y(t;\nu)/\partial\nu$ the corresponding variational differential equation

$$\dot{G}(t;\nu) = f_y(t,y,p)G(t;\nu) + f_\nu(t,y,p) \qquad G(t_0;\nu) = \frac{\partial y_0}{\partial \nu} \tag{3.8}$$

$$\dot{y}(t) = f(t,y,p) \qquad y(t_0) = y_0, \tag{3.9}$$

which has to be solved along with the nominal trajectory (3.9). The differentiation of the BDF discretization scheme (3.1) itself with respect to ν results in

$$\sum_{i=0}^{k} \alpha_i y_{\nu,n+1-i} + h f_y(t_{n+1}, y_{n+1}, p) y_{\nu,n+1} + h f_\nu(t_{n+1}, y_{n+1}, p) + \frac{\partial h}{\partial \nu} f(t_{n+1}, y_{n+1}, p) = 0. \tag{3.10}$$

When the VDE (3.8) is computed - according to the concept of IND - with the same step size and order as the nominal trajectory (3.9), the error term $\partial h/\partial\nu$ is zero. Then the solution for the VDE is the exact derivative of the nominal trajectory approximation.

In every BDF step the linear system

$$\left[\alpha_0 I + h f_y(t_{n+1}, y_{n+1}, p)\right] y_{\nu,n+1} = -h f_\nu(t_{n+1}, y_{n+1}, p) - c \tag{3.11}$$

has to be solved for $y_{\nu,n+1}$, where $c = \sum_{i=1}^{k} \alpha_i y_{\nu,n+1-i}$. This direct strategy, i.e. the explicit solution of the linear system (3.11), is typically used in the literature for sensitivity analysis.

3.1.3 Performance of State-of-the-Art Methods

In the following example we investigate the performance of this standard technique in the integrator DAESOL for the two options available by means of a set of two PDEs describing a typical column experiment. In the first mode, the derivatives f_y and f_p are computed by finite differences (FD-mode). In the second mode, the derivative information generated by automatic differentiation (AD-mode) is used.

Example 3.1.1
As a reference example throughout this chapter we consider a soil column of 20 cm length described by the Richards equation and the convection-dispersion equation containing a nonlinear Michaelis-Menten degradation term:

Water transport

$$C(\psi_m)\frac{\partial \psi_m}{\partial t} = \frac{\partial}{\partial z}\left[K(\psi_m)\frac{\partial}{\partial z}\left(\psi_m - \bar{w}z\right)\right] \tag{3.12}$$

with

$$\begin{aligned} K(\psi_m) &= K_s \frac{(1-(\alpha|\psi_m|)^{n-1}(1+(\alpha|\psi_m|)^n)^{1/n-1})^2}{(1+(\alpha|\psi_m|)^n)^{\frac{1-1/n}{2}}} \\ C(\psi_m) &= \alpha(n-1)(\theta_s-\theta_r)(\alpha|\psi_m|)^{n-1}(1+(\alpha|\psi_m|)^n)^{1/n-2} \end{aligned} \tag{3.13}$$

- *Initial condition:* $\psi_m(0,z) = -670, \qquad z \geq 0$
- *Upper boundary:*

$$q(t,0) = -K(\psi_m(t,0))\frac{\partial}{\partial z}\left(\psi_m(t,0) - \bar{w}z\right) = \begin{cases} 0.2 & t \leq 6.0 \\ 0.0 & t > 6.0 \end{cases}$$

- *Lower boundary:* $\partial\psi_m(t,20)/\partial z = 0, \qquad t \geq 0$

Solute transport

$$\frac{\partial}{\partial t}(\theta c) = \frac{\partial}{\partial z}\left(\theta D_h(\theta)\frac{\partial c}{\partial z} - qc\right) - \theta\frac{V_{max}c}{c+K_M} \tag{3.14}$$

with

$$D_h(\theta) = \frac{0.0046\exp(b\theta)}{\theta} + D_m \tag{3.15}$$

- *Initial condition:* $c(0,z) = 0, \qquad z \geq 0$
- *Upper boundary:*

$$-\theta(t,0)D_h(\theta(t,0),q(t,0))\frac{\partial c(t,0)}{\partial z} + q(t,0)c(t,0) = \begin{cases} 10.0\,q(t,0) & t \leq 6.0 \\ 0 & t > 6.0 \end{cases}$$

- *Lower boundary:* $\partial c(t,20)/\partial z = 0, \qquad t \geq 0$

The van Genuchten parameters characterizing the hydraulic properties are fixed to $n = 1.2$ $[1]$, $\alpha = 0.0102$ $[hPa^{-1}]$, $\theta_r = 0.05$ $[1]$, $\theta_s = 0.4$ $[1]$ *and* $K_s = 10.0$ $[cm\,h^{-1}]$ *(Seppelt, private communication; Vink et al., 1994 [133]). In the solute transport equation, the exponent* b *in the diffusion term, the dispersion coefficient* D_m *and the parameters* V_{max} *and* K_M *in the Michaelis-Menten term are unknown. Measurement data for 4 depths (*$z = 2, 7, 12, 17\,[cm]$*) and 11 points in time (*$t = 1, 2, \ldots, 11\,[h]$*) is generated by solving the PDEs (3.12) and (3.14) for the initial and boundary conditions presented using the "true" parameters as given in Table 3.1. No noise is added. Starting from the initial guesses for the parameters in Table 3.1, the*

	True Parameters	Initial Guesses
b	10.0	17.0
D_m	5.0	2.0
V_{max}	0.2	0.1
K_M	1.0	0.5

Table 3.1: True Parameter Values and Initial Guesses

parameter estimation problem is solved according to the standard approach by directly solving the linear equation system (3.11) in each BDF step. Two different modes for the generation of f_y *and* f_p *are used:*

- *Finite difference schemes* (FD)
- *Automatic differentiation* (AD) *using the tool* ADIFOR 2.0 (Revision D)

In Table 3.2, the CPU times, the number of f*-calls and the number of iterations required for parameter estimation in the* FD *mode are summarized for increasingly finer spatial grids. The huge number of* f*-calls is due to the frequent evaluation of the right hand side* f *in order to approximate* f_y *and* f_p *via finite differences. In the* AD *mode, subroutines for the computation of* f_y *and* f_p *generated by* ADIFOR *are provided. In addition to Table 3.2 , Table 3.3 shows for the* AD *mode the number of calls to the subroutines* f_y *and* f_p*. Comparing both modes, as done in Figure 3.1, where CPU times are plotted against the number of spatial nodes, illustrates that the* AD *mode is roughly a factor 2 faster than the* FD *mode. Moreover, parameter estimation in the* AD *mode requires on average fewer iterations than in the* FD *mode, which is likely due to the higher accuracy of derivatives. However, Figure 3.1 also reveals that neither mode is*

# Nodes	# ODEs	CPU	# f-Calls	# Iter.
41	82	6 min 23 sec	565708	6
81	161	28 min 47 sec	1340518	8
161	321	1 h 27 min	2076748	5
321	641	6 h 04 min	4368027	5
641	1282	24 h 51 min	8958814	5
961	1922	57 h 08 min	13701043	5

Table 3.2: Standard Technique Using Finite Differences (FD) for f_y and f_p

# Nodes	CPU	# f-Calls	# f_y-Calls	# f_p-Calls	# Iter.
41	3 min 6 sec	37926	5557	5293	5
81	15 min 13 sec	49073	7236	6897	7
161	45 min 11 sec	46012	6274	5994	5
321	3 h 19 min	50033	6754	6454	5
641	13 h 54 min	52708	6979	6625	5
961	33 h 55 min	56850	7484	7046	5

Table 3.3: Standard Technique Using Automatic Differentiation (AD) for f_y and f_p

capable to cope with the large scale systems leading to unacceptable CPU times of several days. Unsatisfying results are obtained if more than 600 ODEs are involved. The computational effort to evaluate f_y is in both cases of complexity order $\mathcal{O}(n_y^2)$ as exhibited in Figure 3.2. Figure 3.3, showing the quotient between the CPU-time required for computing f_y and the overall CPU time, demonstrates that almost the complete CPU time is spent in the computation of f_y.

We can conclude that the derivatives generated by ADIFOR outperform the finite difference approximations in terms of efficiency, accuracy and convergence rates of the reduced Generalized Gauss-Newton method. However, in summary we have to state that even when using a highly sophisticated, state-of-the-art integrator such as DAESOL - even in combination with ADIFOR - the computing times obtained for the large scale parameter estimation problems studied here are by far not acceptable. Hence, for the treatment of real-life problems the use of faster strategies is indispensable.

The analysis in Example 3.1.1 shows that the computational effort is dominated by the computation of f_y in every BDF step. The CPU time for the computation of f_y increases quadratically with the number of spatial nodes. Consequently, strategies to speed up parameter estimation, in particular for fine spatial grids, must aim at removing this bottleneck. Basically, this can be approached in two ways by avoiding the computation of f_y in every BDF step and/or by drastically reducing the computing times for f_y, i.e. by getting rid of the complexity order $\mathcal{O}(n_y^2)$. In the next two sections approaches for both strategies are presented.

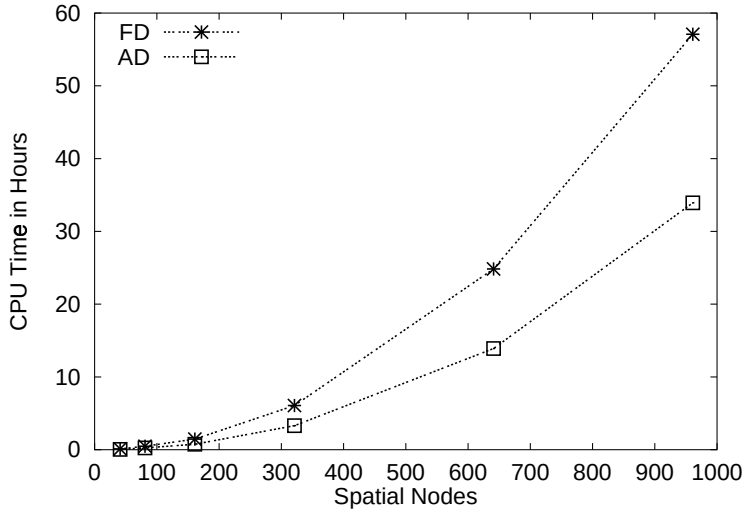

Figure 3.1: Overall CPU times for parameter estimation in the FD and the AD mode.

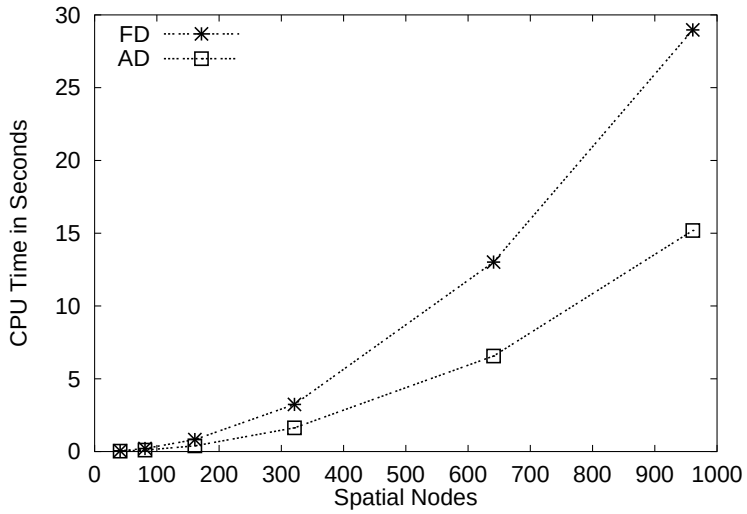

Figure 3.2: CPU times for the computation of f_y in the FD and the AD mode.

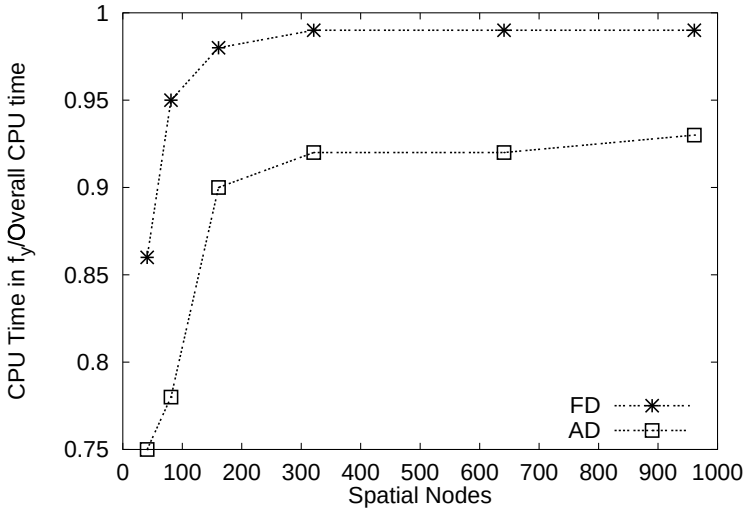

Figure 3.3: Fraction of the overall CPU time spent for the computation of f_y in the FD and the AD mode.

3.2 Solution of VDEs Using a Modified Newton Method

In this section a new approach is outlined that circumvents the computation of f_y in each BDF step by using a modified Newton method (Dieses et al., 1999 [41]) instead of a direct approach for the solution of the linear system (3.11). This approach is also a realization of the principles of Internal Numerical Differentiation (see Bock et al., 1995 [23]; Bauer, 2000 [8]).

3.2.1 Description of the Approach

In our approach we make use of the main result of Chapter 2: In the *reduced* Generalized Gauss-Newton method the VDE (2.77) has to be solved only for a few directions, namely only $n_p + 1$. In this case it is more efficient to solve the linear system (3.11) by means of a modified Newton method. This avoids the expensive computation of f_y in every BDF step:

$$F(y_{\nu,n+1}) := \big[\alpha_0 I + h f_y(t_{n+1}, y_{n+1}, p)\big] y_{\nu,n+1} + h f_\nu(t_{n+1}, y_{n+1}, p) + c \stackrel{!}{=} 0 \tag{3.16}$$

For the mth modified Newton iteration we get

$$\begin{aligned} \tilde{J}\Delta y^{(m)}_{\nu,n+1} &= -F(y^{(m)}_{\nu,n+1}) \\ y^{(m+1)}_{\nu,n+1} &= y^{(m)}_{\nu,n+1} + \Delta y^{(m)}_{\nu,n+1}. \end{aligned} \tag{3.17}$$

Here $\tilde{J}$ denotes an approximation for $J = \partial F(y_{\nu,n+1})/y_{\nu,n+1}$. Applying the monitoring strategy originally used in DAESOL not only for the computation of the nominal trajectory but also for the solution of (3.17), $\tilde{J}$ is kept frozen as long as possible to keep the computational effort as low as possible. Since $\tilde{J}$ is already needed in the integration of the nominal trajectory, no additional

# Nodes	CPU	# f-calls	# f_y-calls	# f_p-calls	# $f_y w$-calls	# Iter.
41	1 min 42 sec	62547	1029	5974	25173	6
81	4 min 01 sec	64245	952	5801	25419	5
161	12 min 00 sec	70792	998	6215	27778	5
321	39 min 31 sec	75869	941	6507	29487	5
641	2 h 30 min	81158	960	6874	31415	5
961	5 h 38 min	83991	1010	7047	32845	5

Table 3.4: Strategy OPT-I: Using a Modified Newton Method with $f_y w$

computation and decomposition of $\tilde{J}$ is required. Hence the expensive computation of f_y is avoided. Instead, to set up $F(y_{\nu,n+1})$, only the directional derivatives $f_y(t_{n+1}, y_{n+1}, p)y_{\nu,n+1}$ have to be provided. As a result of the reduced approach altogether only $n_p + 1$ directional derivatives are needed. These can be generated efficiently and with high accuracy via ADIFOR which enables a direct evaluation of each directional derivative at the cost of 3 to 5 function evaluations. In summary, this approach leads to an enormous reduction of the overall CPU time as the results in the example discussed in the next section will demonstrate.

3.2.2 Performance Study

In this section, we investigate the performance of the reduced Generalized Gauss-Newton method using the modified Newton approach for the solution of the VDEs. The results are compared to the ones obtained in Section 3.1.3 employing DAESOL in the FD and the AD mode.

Example 3.2.1 (Continuation of Example 3.1.1)
The strategy of employing a modified Newton method, as described in Section 3.2.1, instead of a direct method for the solution of the linear system (3.11) is applied to the reference example, Example 3.1.1, outlined in the previous section. For this mode, in the following referred to as OPT-I*, in addition to the CPU times, the number of f-calls etc., the number of directional derivatives is listed in Table 3.4. Comparing these results with the ones obtained in the* AD *mode, Table 3.3, shows that the number of the expensive computations of f_y can be reduced by approximately a factor of 6. As Figure 3.4 illustrates, this results in significant savings of the overall CPU times. The speed up can reach factors of 5 resp. 10 compared to the* AD *resp.* FD *mode. Although the number of f_y-computations is drastically reduced, still more than the half of the total computational work is spent for the computation of f_y (Figure 3.5).*

First of all, we can state that the new strategy performs extremely well for this class of problems requiring only $n_p + 1$ directional derivatives. The reduction of the computational effort gained is considerable. However, Example 3.2.1 also reveals that in order to further speed up the code it is necessary to reduce the CPU times for the computation of f_y as such. In the next section we will present an approach that removes the complexity order $\mathcal{O}(n_y^2)$ for the computation of f_y.

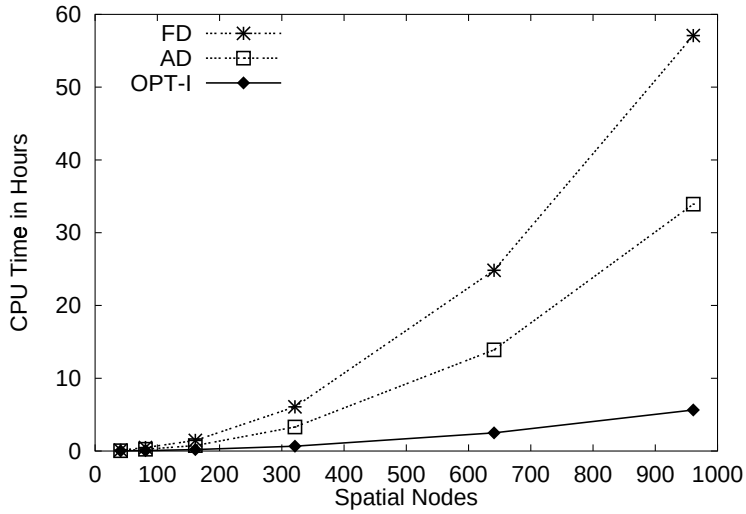

Figure 3.4: Overall CPU times for parameter estimation in the FD, AD and OPT-I mode.

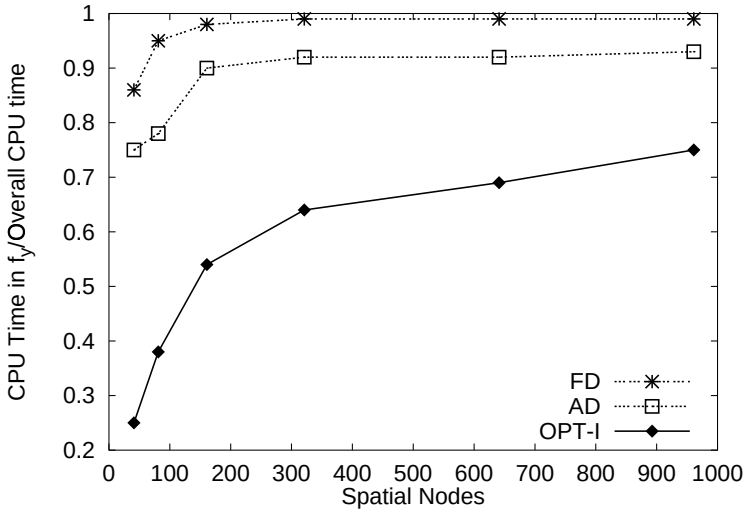

Figure 3.5: Fraction of the overall CPU time spent for the computation of f_y in the FD, AD and OPT-I mode.

3.3 Specially Tailored Methods for the Computation of Jacobians

3.3.1 Description of the Approach

In this section we outline a strategy for the computation of f_y which is of the same complexity order as the evaluation of the right hand side f itself, namely $\mathcal{O}(n_y)$. We exploit the fact that, as a result of the use of fixed spatial grids, the sparsity pattern remains unchanged in the course of the reduced Generalized Gauss-Newton method.

Even though ADIFOR outperforms finite difference approximations, it does not automatically take into account the special sparsity structure of f_y induced by the spatial discretization of the PDEs. Due to the fact that ADIFOR is mainly based on the forward mode (Bischof et al., 1996 [18]), it is in its standard version, as used in the previous sections, rather expensive for computing sparse and structured Jacobians J. In order to generate the full $n_y \times n_y$ Jacobian $J = f_y \in \mathbb{R}^{n_y \times n_y}$ the corresponding seed matrix is the identity matrix $I \in \mathbb{R}^{n_y \times n_y}$ consisting of n_y columns, i.e. the so called leading dimension lp is n_y. Thus, roughly $lp = n_y$ operations for every assignment statement in the original function f are required. Consequently, the computation of f_y requires n_y times as many operations as the computation of f. As the latter one is of complexity order $\mathcal{O}(n_y)$, the generation of f_y is of order $\mathcal{O}(n_y^2)$ as exemplified in Figure 3.2.

However, it is well known that the number of function evaluations required to compute an approximation to the Jacobian by divided differences can be much less than n_y if J is sparse. One option to exploit the sparsity structure of a Jacobian J is by computing a compressed Jacobian V (Curtis et al., 1974 [34]). The key idea of this approach is to identify structural orthogonal columns of J. The columns of J are partitioned into groups such that columns of the same group do not have nonzeros in the same row position. Suppose J has been partitioned into l groups, each group consisting of structurally orthogonal columns, J can be derived by computing l directional derivatives $J\bar{s}_i$ $(i = 1, \ldots, l)$. Hereby, a vector $\bar{s}_i$ is associated with each of the l groups, whose jth component $\bar{s}_{ij} = 1$ if the jth column of J is in this group, and $\bar{s}_{ij} = 0$ otherwise. Thus, we can substitute the computation of n_y directional derivatives Je_i $(i = 1, \ldots, n_y)$, e_i ith unit vector, in the standard approach by computing only $l \ll n_y$ directional derivatives $J\bar{s}_i$ $(i = 1, \ldots, l)$. The resulting compressed Jacobian V, a $n_y \times l$ matrix, has to be extracted to obtain the required Jacobian J. Because of the structural orthogonality properties this extraction is unique.

The crucial point in this approach is the identification of the structural orthogonal columns. Curtis et al. (1974 [34]) proposed an algorithm, also referred to as CPR algorithm, where successively groups are formed by scanning the original Jacobian column by column. If the actual column under consideration has not been included in one of the previous groups, it is a potential candidate for the current group. For the case that it does not have a nonzero in the same row position as one of the other columns already in the group, this column is added to the group, otherwise it is skipped. Coleman and Moré (1983 [33]) interpreted this partitioning problem as a graph coloring problem. By employing methods from this field, Coleman et al. (1984 [32], 1999 [31]) were able to improve the original CPR algorithm.

Averick et al. (1994 [6]) applied these ideas to automatic differentiation. If the structure of the Jacobian is known a priori, the seed matrix has to be initialized according to the identified partitioning in structurally orthogonal columns.

For the class of problems treated here, the structure of the Jacobian f_y is induced by the spatial discretization schemes for the PDEs. Employing the spatial discretization routines DSS004 and DSS020 (Schiesser, 1991 [107]) of order 4, a bandwidth of 5 is obtained for first order spatial derivatives in f. Accordingly, the second order spatial derivatives lead to a bandwidth of 9. Considering the Jacobian f_y for n_{PDE} PDEs and N spatial nodes, i.e. $n_y = n_{PDE} \times N$, at most $9 \times n_{PDE} \times n_y$ out of $n_y \times n_y$ entries in f_y are nonzero. For example, for two PDEs and 321, 641 and 961 spatial nodes only 5.6 %, 2.8 % and 1.9 % of the entries, respectively, are nonzeros. Thus, the resulting Jacobians are sparse even for moderately sized problems.

Analyzing the structure of f_y induced by the spatial discretization routines used in this work, it can be easily seen that the Jacobian can be grouped into $9 \times n_{PDE}$ sets of structurally orthogonal columns, independent of the number of spatial nodes. As Figure 3.6 illustrates for a small example with 2 PDEs and 21 spatial nodes, 18 structural orthogonal groups can be formed where the columns 1, 10 and 19 are in the first group, the columns 2, 11 and 20 in the second group, the columns 22, 31 and 40 in the 10th group and so on. The structure of the corresponding seed matrix is given in Figure 3.7 together with the resulting compressed Jacobian.

The main advantage of this approach is that the leading dimension lp is no longer equal to the number of variables $n_y = n_{PDE} \times N$ but is independent of the spatial discretization. Now the leading dimension lp is determined only by the number of PDEs n_{PDE}, which is for the class of problems studied here typically less than 10, such that lp is only a small multiple of the bandwidth induced by the second order derivatives in space, i.e. $lp = 9 \times n_{PDE}$. As a result, the computation of f_y is now of the same complexity order $\mathcal{O}(n_y)$ as the evaluation of the right hand side f.

Using this compressed approach to derive f_y within the reduced Generalized Gauss-Newton method enables considerable savings in the overall CPU times for parameter estimation. This will be exemplified in the next section by means of the reference problem used so far.

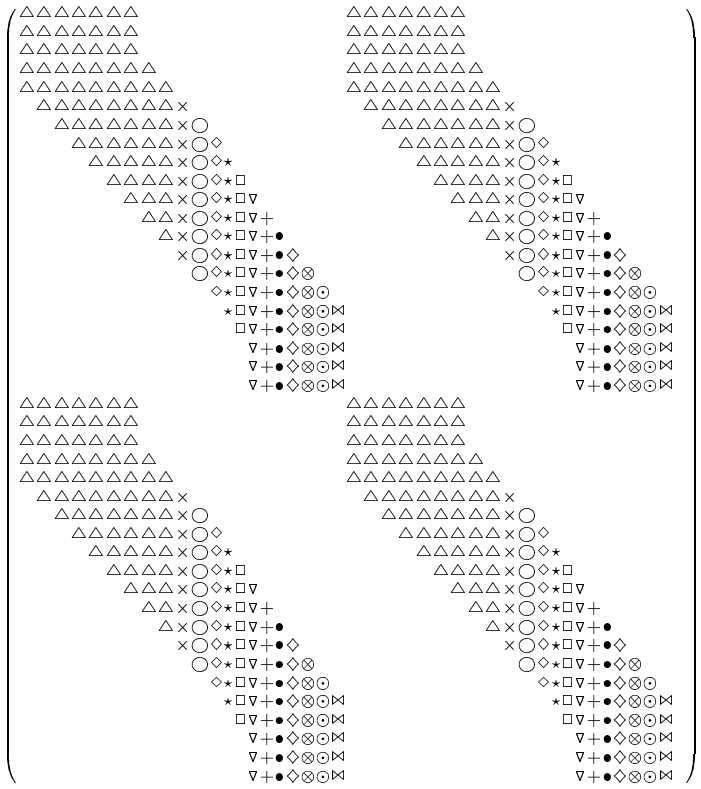

Figure 3.6: Sparsity pattern of the Jacobian for 2 PDEs using 21 spatial nodes.

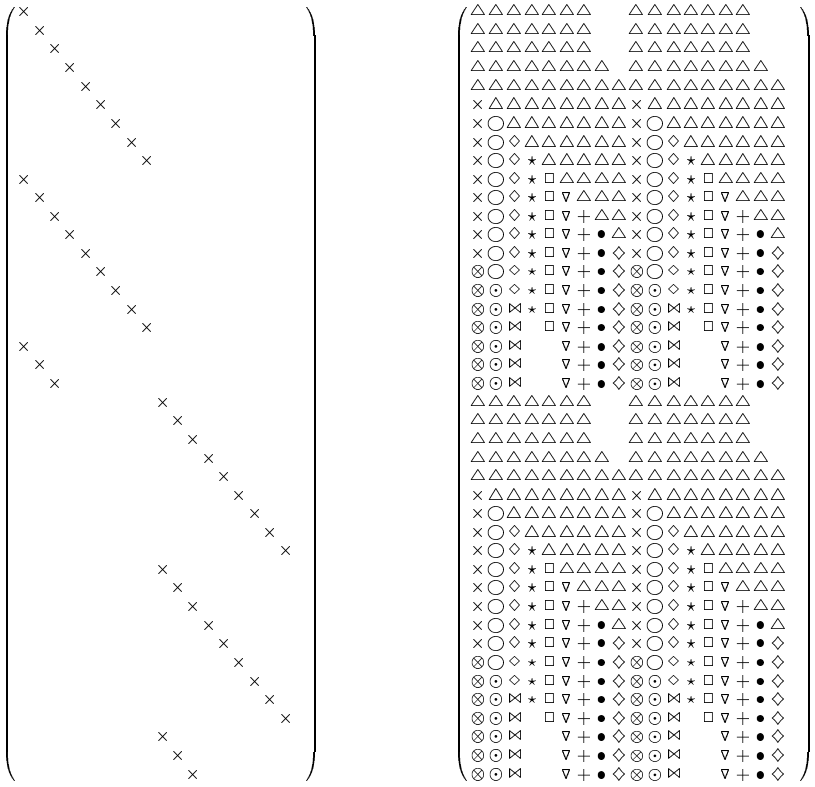

Figure 3.7: Seed matrix and compressed Jacobian for 2 PDEs using 21 spatial nodes.

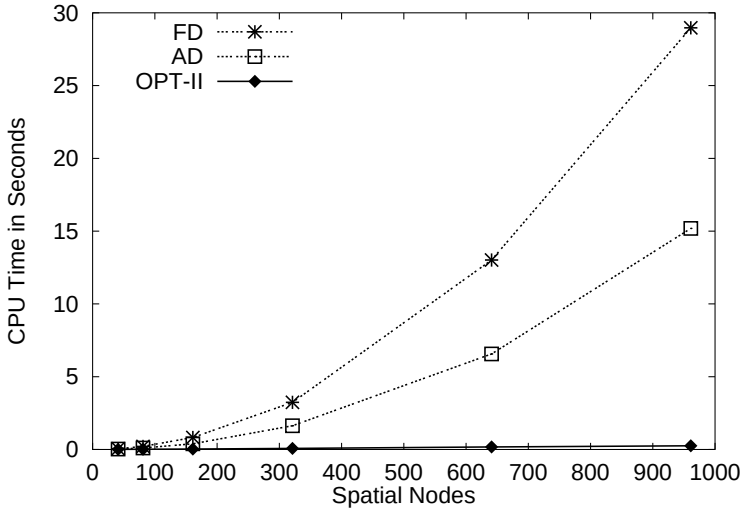

Figure 3.8: CPU times for the computation of f_y using the FD, the standard AD and the matrix compression mode (OPT-II).

3.3.2 Performance Study

In this section, we investigate the performance of the reduced Generalized Gauss-Newton method using the specially tailored methods for the computation of f_y as just derived.

Example 3.3.1 (Continuation of Example 3.1.1 and 3.2.1)
The efficient compressing strategy to derive f_y is applied to the reference problem discussed in Example 3.1.1 and 3.2.1. First of all, we investigate the complexity behavior for the computation of f_y in dependence of the spatial grid. As shown in Figure 3.8, using matrix compressing the computational effort for f_y increases only linearly with the number of spatial nodes, instead of quadratically as before. This confirms the theoretical results that both the computation of f and f_y are of the same complexity order $\mathcal{O}(n_y)$.

Before combining this efficient strategy with the modified Newton method developed in Section 3.2, we first study the performance of matrix compression in the context of the originally used direct approach for the solution of the linear problem (3.11). This combination is referred to in the following as OPT-II*. The corresponding results for* OPT-II *are given in Table 3.5. Figure 3.9 compares the performance of* OPT-II *with the optimized mode* OPT-I *based on the standard computation of f_y in the* AD *mode and the modified Newton method. We can state that for this example it is more efficient to speed up the computation of f_y as such without reducing the number of required f_y-computations than to avoid a part of the f_y-computations in the original, expensive mode. However, these results might be reversed for different examples as other studies have shown.*

The best results are finally obtained by combining the matrix compressing strategy for f_y with the modified Newton method of Section 3.2. As shown in Table 3.6, even for large scale problems with 961 spatial nodes, i.e. 1922 ODEs, less than 1.5 hours are required for parameter

# Nodes	CPU	# f-calls	# f_y-calls	# f_p-calls	# Iter.
41	1 min 24 sec	37926	5557	5293	5
81	3 min 53 sec	49073	7236	6897	7
161	8 min 03 sec	46012	6274	5994	5
321	27 min 54 sec	50033	6754	6454	5
641	1 h 52 min	52708	6979	6625	5
961	4 h 21 min	56850	7484	7046	5

Table 3.5: Strategy OPT-II: Compressed Computation of f_y and Direct Solution of the Linear System 3.11

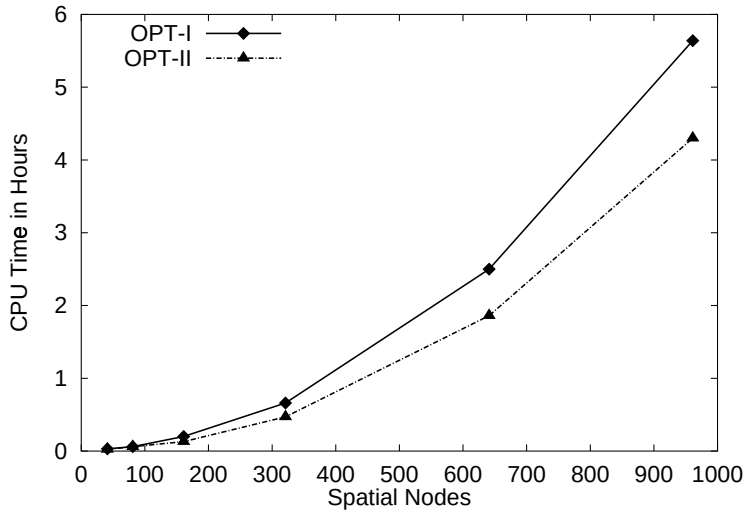

Figure 3.9: Overall CPU times for parameter estimation in the OPT-I and OPT-II mode.

# Nodes	CPU	# f-calls	# f_y-calls	# f_p-calls	# $f_y w$-calls	# Iter.
41	1 min 23 sec	62547	1029	5974	25173	6
81	2 min 41 sec	64245	925	5801	25419	5
161	5 min 59 sec	70792	998	6215	27778	5
321	14 min 41 sec	75869	941	6507	29487	5
641	42 min 31 sec	81158	960	6874	31415	5
961	1 h 25 min	83991	1010	7047	32845	5

Table 3.6: Strategy OPT-III: Compressed Computation of f_y and Modified Newton Method for the Solution of the Linear System 3.11

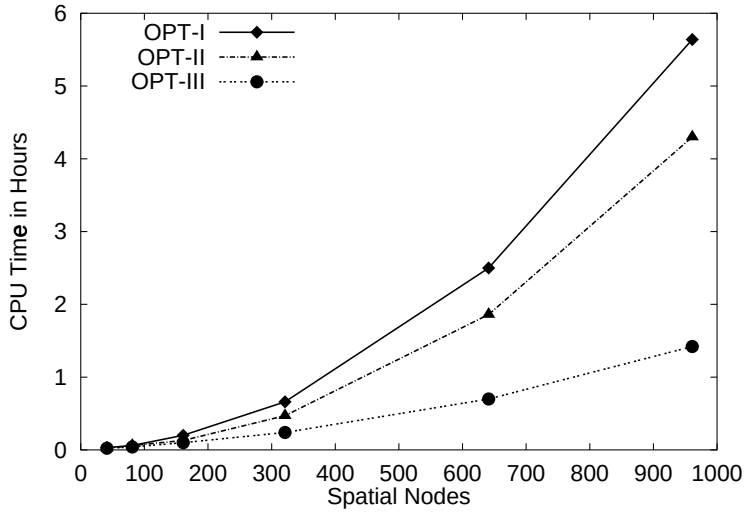

Figure 3.10: Overall CPU times for parameter estimation in the OPT-I, OPT-II and OPT-III mode.

estimation instead of the approximately 2.5 days needed in the finite differences mode (FD) *described in Section 3.1, Table 3.2. This gives a speed up by a factor of 40. Compared to the modes* OPT-I *and* OPT-II *additionally an acceleration by a factor of 4 respectively 3 is yielded. Figure 3.10 summarizes the performance results of the three modes* OPT-I*,* OPT-II *and* OPT-III*.*

These results exhibit that for the large scale parameter estimation problem studied here neither of the two strategies will provide the speed up gained in the end on its own. The first strategy OPT-I, which aims at substituting a part of the expensive f_y-computations by directional derivatives, suffers from the fact that still more than one half of the overall CPU time is spent for the computation of f_y due to its computational complexity of order $\mathcal{O}(n_y^2)$ (Figure 3.11).

The second strategy OPT-II, even though enabling a very fast computation of f_y, is still dominated by the huge amount of f_y-calls. Hence, still 20 $\%$ of the total computational effort is due to the evaluation of f_y (Figure 3.11). Only by combining both strategy do we manage to reduce the fraction spent for the computation to less than 5 $\%$ as visualized in Figure 3.11.

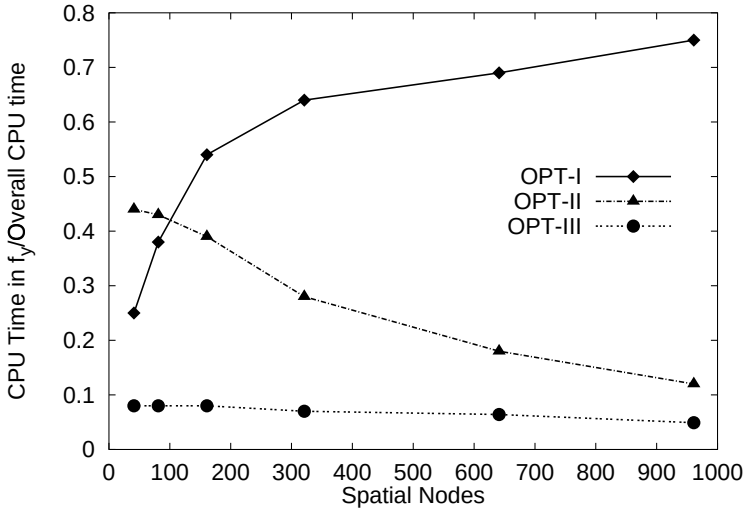

Figure 3.11: Fraction of the overall CPU time spent for the computation of f_y in the OPT-I, OPT-II and OPT-III mode.

In summary, we can state that by means of synergy effects of the two outlined strategies, the modified Newton method and the compressed computation of f_y, a very efficient approach for the solution of VDEs has been developed. Using this specially tailored approach in the framework of the reduced Generalized Gauss-Newton method, we end up with a fast method for parameter estimation in large scale problems arising from discretized PDEs.

Chapter 4

Optimization of Experimental Conditions and Sampling Design

Relatively few studies about optimal experimental design for parameter estimation problems constrained by differential equations are reported. Lohmann (1992 [87], 1993 [86]), for example, developed methods for optimal experimental design in chemical reaction systems. His focus was on the determination of optimal sampling designs. Hilf (1996 [59]) investigated optimal experimental design problems for the calibration of robots which were described by DAEs. The optimization of both sampling designs and controls were considered. In contrast to the problems treated here, the underlying parameter estimation problems were unconstrained.

Based on the results of Lohmann and Hilf, recently an approach for optimal experimental design in parameter estimation problems constrained by ODEs and DAEs has been developed in the framework of a BMBF project [1]. These strategies have been implemented in the software tool VPLAN (Bauer et al., 1999 [9]; Körkel et al., 1999 [83]; Bauer et al., 2000 [10]) and are used for optimal experimental design in chemical reaction systems (Körkel, [82]).

In this work we extend this approach to parameter estimation problems constrained by PDEs. It is used for the determination of optimal sampling schemes and experimental conditions of column experiments. On the basis of VPLAN and ECOFIT (Dieses et al., 1999 [41]; 1999 [43]) the new tool ECOPLAN (Dieses et al., 2000 [42]) has been developed. In particular, ECOPLAN is suitable for optimal experimental design in water flow and reactive solute transport processes.

The main focus of this chapter is on the formulation of an optimal experimental design problem where the underlying parameter estimation problem is constrained by PDEs and ODEs as discussed in the previous chapters and on the presentation of state-of-the-art solution methods. Hereby, the way of presentation is orientated by Bauer et al. (2000 [10]). In the first section, the optimal experimental design problem is derived in the framework of the Generalized Gauss-Newton method. We discuss the different types of optimization variables and objective functions and formulate the design problem as an optimal control problem. In Section 4.2, numerical

[1] Verbundvorhaben “Optimale Versuchsplanung für nichtlineare Prozesse” sponsored by BMBF (FKZ: 03 D 0043)

solution methods are outlined. Using a direct approach, the time-dependent control functions and state constraints are discretized on a suitable grid. The resulting finite dimensional, nonlinear constrained optimization problem is finally solved by a structured SQP-method. Section 4.3 is devoted to the discussion of practical requirements on optimal experimental designs and on possible extensions, such as sequential designs or designs for model discrimination.

4.1 Formulation of the Optimization Problem

4.1.1 Variance-Covariance Matrix

The formulation of the optimal experimental design problem as derived in the following is based on an approximation of the variance-covariance matrix. As this matrix characterizes the statistical quality of the parameter estimates, a reasonable approach for optimal experimental design is to determine sampling schemes and experimental conditions that yield estimates with low variances.

In the framework of the approach outlined in Chapter 2, the discretization of the original parameter estimation problem (1.68)-(1.74) in time and space results in a finite dimensional nonlinear constrained least-squares problem of the form

$$\min_{s,p} ||r_1(s,p)||_2^2 \tag{4.1}$$

$$r_c(s,p) \;=\; 0. \tag{4.2}$$

Hereby, r_1 denotes the least-squares conditions and r_c the equality conditions, i.e. the equality conditions arising from the parameterization of the PDEs and ODEs including their initial and boundary conditions and potentially further equality conditions. In every iteration of the Generalized Gauss-Newton algorithm a problem of the form

$$\min_{\Delta s,\Delta p} ||r_1(\hat{s},\hat{p}) + J_1(\hat{s},\hat{p})(\Delta s,\Delta p)||_2^2 \tag{4.3}$$

$$r_c(\hat{s},\hat{p}) + J_c(\hat{s},\hat{p})(\Delta s,\Delta p) \;=\; 0 \tag{4.4}$$

is solved, where J_1 and J_c are the Jacobians of the vector of residuals r_1 and of the constraints r_c, respectively. With the generalized inverse of the Jacobian $J^T = (J_1\ J_c)^T$

$$J(\hat{s},\hat{p})^+ = \begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} J_1^T J_1 & J_c^T \\ J_c & 0 \end{pmatrix}^{-1} \begin{pmatrix} J_1^T & 0 \\ 0 & I \end{pmatrix}, \tag{4.5}$$

which is the solution operator for the constrained linear problem (4.3)-(4.4), the variance-covariance matrix for the parameter estimates (Bock, 1987 [21]) can be written as

$$C(\hat{s},\hat{p}) = J(\hat{s},\hat{p})^+ \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} J(\hat{s},\hat{p})^{+T}. \tag{4.6}$$

Thus, the variance-covariance matrix is easily and in comparison inexpensively accessible within the Generalized Gauss-Newton method. This provides the basis for optimal experimental design which aims at minimizing a function on the variance-covariance matrix.

4.1.2 Optimization Variables and Objective Functions

In order to formulate the optimal experimental design problem we have to specify optimization variables and an objective function.

In the column and mini-lysimeter studies considered here we have several possibilities to control respectively influence the experimental conditions and the sampling design. In the following we distinguish three types of control/optimization variables:

- **Control variables** $\bar{q}$ are time-independent optimization variables, such as initial conditions for the soil column, i.e. the initial matric potential $\psi_m(t = 0, z)$ and/or the initial substance concentrations $c(t = 0, z)$. For optimization, feasible ranges for $\bar{q}$, in particular upper and lower bounds, have to be defined.

- **Control functions** $u(t)$ are in contrast to control variables $\bar{q}$ time-dependent. In our case, possible control functions enter upper and lower boundary conditions. For example, we can control the water input flux $q(t, 0)$ or the substance input concentration $c_0(t, 0)$ at the top of the soil column. At the lower boundary, new technologies allow to vary the matric potential $\psi_m(t, L)$ as a function of time. Similar to the control variables, upper and lower bounds for $u(t)$ have to be specified.

- **Weights** w are used to describe the sampling design, i.e. where and when which measurements are carried out. A set of L_{fea} feasible measurements

$$\Xi = \{\xi_r = (t_i, z_j, m_k) |\, t_i \in T, z_j \in Z, m_k \in M, r = 1, \ldots, L_{fea}\} \tag{4.7}$$

  has to be defined. Here, T and Z denote the temporal and spatial domains, respectively. M represents the different types of feasible measurements, i.e. which species can be measured by means of which type of measurement device, e.g. point or outflow measurements. For every measurement point ξ_r a guess of the accuracy of the measurement, i.e. the standard deviation, $\bar{\sigma}_r$ has to be provided. We associate with every measurement point ξ_r $(r = 1, \ldots, L_{fea})$ a weight w_r being 0 or 1. Out of the set of the L_{fea} feasible measurements Ξ a prescribed number of L_{max} measurements has to be chosen satisfying for $w = (w_1, \ldots, w_{L_{fea}})^T$ with $w_r \in \{0, 1\}$

$$\sum_{r=1}^{L_{fea}} w_r = L_{max}. \tag{4.8}$$

As mentioned in Section 1.5.1, in the context of optimal experimental design the parameters p are fixed, i.e. an optimal experimental design is determined for a fixed set of parameters p. The corresponding variance-covariance matrix C is now a function of the optimization variables/functions $\bar{q}$, $u(t)$ and w

$$C := C(s, p, \bar{q}, u, w). \tag{4.9}$$

In order to set up the experimental design problem a suitable objective function, i.e. a design criterion, has to be chosen. Frequently, one of the classical objective functions Φ on the variance-covariance matrix (see e.g. Fedorov, 1972 [47] , Atkinson and Donev, 1992 [5], Pukelsheim, 1993 [103]) is used:

- **A-Criterion**: minimizes the average variances of the estimated parameters

$$\Phi_1(C) = \frac{1}{n_p} trace(C). \tag{4.10}$$

- **E-Criterion**: minimizes the largest eigenvalues of the variance-covariance matrix

$$\Phi_\infty(C) = \lambda_{max}(C). \tag{4.11}$$

- **D-Criterion**: minimizes in the case of an unconstrained parameter estimation problem the determinant of the variance-covariance matrix

$$\Phi_0(C) = (det\, C)^{\frac{1}{n_p}}. \tag{4.12}$$

  Due to the fact that the variance-covariance matrix is singular for constrained parameter estimation problems, a restriction to a regular submatrix of C is necessary

$$\Phi_0(C) = (det\, K^T C K)^{\frac{1}{n_K}}, \tag{4.13}$$

  where K is a full rank, $n_K \times n_p$ matrix such that $K^T C K$ is regular.

If the main focus is on standard deviations respectively confidence intervals, a min-max-criterion (Bock 1987, [21]; Lohmann et al., 1992 [87]) may be employed which minimizes the maximal standard deviation.

4.1.3 The Optimal Control Problem

The optimal experimental design problem can now be formulated as a nonlinear state-constrained optimal control problem:

$$\min_{s,\bar{q},u,w} \Phi(C(J(s,p,\bar{q},u,w))) \tag{4.14}$$

$$r_c(s,p,\bar{q},u) = 0. \tag{4.15}$$

Additionally, equality and inequality conditions on the controls $\bar{q}$ and $u(t)$ and the weights w can be defined and are summarized as

$$c(\bar{q},u,w) \left\{ \begin{array}{c} = \\ \geq \end{array} \right\}. \tag{4.16}$$

In particular, (4.16) contains the constraints on the weights $w = (w_1, \ldots, w_{L_{fea}})^T$

$$\sum_{r=1}^{L_{fea}} w_r = L_{max}, \qquad w_r \in \{0, 1\}. \tag{4.17}$$

Moreover, state constraints of the form

$$d(y, p, \bar{q}, u) \geq 0 \tag{4.18}$$

can be specified.

Considering column and mini-lysimeter studies, constraints on the control variables $\bar{q}$ are e.g. upper and lower bounds for feasible initial values for the matric potential $\psi_m(t = 0, z)$ and/or for the substance concentration $c_0(t = 0, z)$. Analogously, we can define, for example, upper and lower bounds for feasible water input fluxes $q(t, z = 0)$ and/or for feasible substance input concentrations $c_0(t, z = 0)$ at the upper boundary by formulating the constraints on the corresponding control functions. Using state constraints of type (4.18), we can avoid e.g. the Richards equation to become undefined by requiring $\psi_m(t, z = 0) > \varepsilon > 0$ at the upper boundary.

In summary, we end up with a highly intricate optimal control problem constrained by infinite dimensional equality and inequality conditions. In particular, we have to treat an objective function on the variance-covariance matrix which is implicitly defined by the Jacobian of the underlying parameter estimation problem. Thus, the objective function already contains first-order derivatives which implies that for optimization employing e.g. SQP-methods sufficiently accurate second-order derivatives have to be provided.

4.2 Solution of the Optimal Control Problem

4.2.1 Direct Approach

The optimal control problem (4.14)-(4.18) is treated by means of a direct approach. The idea is to transform the infinite dimensional problem (4.14)-(4.18) into a finite dimensional nonlinear constrained optimization problem. To this end the following steps are carried out:

- **Parameterization of control functions**
  The time-dependent control functions u ($u : [t_0, t_e] \rightarrow \mathbb{R}^{N_2}$) are parameterized on a suitable grid, e.g. by piecewise constant or linear functions. For every control function u_l ($l = 1, \ldots, N_2$) M control variables $\bar{q}_k$ ($k = N_1 + (l-1)M + 1, \ldots, N_1 + lM$) are introduced. Together with the N_1 time-independent control variables $\bar{q} \in \mathbb{R}^{N_1}$, an augmented vector of control variables is defined

$$\hat{q} := (\bar{q}_1, \ldots, \bar{q}_{N_1}, \bar{q}_{N_1+1}, \ldots, \bar{q}_{N_1+MN_2})^T. \tag{4.19}$$

- **Discretization of state constraints**
  State constraints are also discretized on a suitable grid and substituted by corresponding interior-point conditions.

- **Relaxation of integrality constraints**
  In order to enable the use of a Successive Quadratic Programming (SQP) method we use instead of the binary weights $w \in \{0,1\}^{L_{fea}}$ a relaxed formulation

$$w \in [0,1]^{L_{fea}}. \tag{4.20}$$

  Thus, also fractional weights, e.g. 0.75, may be identified by optimization. By applying a suitable rounding heuristic an integer solution can be obtained which typically hardly increases the optimality criterion value. However, for the numerical results presented in Chapter 5 no rounding strategies are necessary because the solution satisfies already the integrality constraints.

4.2.2 Numerical Solution of the Discretized Problem

SQP-method

Using a direct approach as described in the previous section results in a finite dimensional nonlinear constrained optimization problem. Rearranging equality and inequality constraints we can write

$$\min_v F_1(v) \tag{4.21}$$

$$F_2(v) = 0 \tag{4.22}$$

$$F_3(v) \geq 0, \tag{4.23}$$

where $v = (s, \hat{q}, w)$.

As in the Generalized Gauss-Newton method (2.11)-(2.13), Section 2.2.1, the solution is iteratively improved by

$$v_{k+1} = v_k + \hat{\lambda}_k \Delta v_k, \qquad \hat{\lambda}_k \in (0,1]. \tag{4.24}$$

This time, however, the iterate Δv_k solves a quadratic subproblem of the form

$$\min_{\Delta v} \frac{1}{2} \Delta v^T H_k \Delta v + \hat{J}_1(v_k)\Delta v \tag{4.25}$$

$$F_2(v_k) + \hat{J}_2(v_k)\Delta v = 0 \tag{4.26}$$

$$F_3(v_k) + \hat{J}_3(v_k)\Delta v \geq 0, \tag{4.27}$$

with $\hat{J}_i(v_k) = \nabla F_i(v_k)^T$. Here, H_k denotes an approximation of the Hessian of the Lagrangian function of the nonlinear problem (4.21)-(4.23)

$$L(v, \lambda, \mu) = F_1(v) - \lambda^T F_2(v) - \mu^T F_3(v), \tag{4.28}$$

where λ and μ are the corresponding Lagrange multipliers.

For the solution of the discretized optimization problem (4.21)-(4.23) the SQP-method SNOPT developed by Gill et al. (1998 [55]) is used.

Generation of Derivatives

In order to solve the finite dimensional nonlinear constrained optimization problem, the evaluation of F_2 and F_3 and the computation of the Jacobians $\hat{J}_i$ $(i = 1, 2, 3)$ are required in each iteration of the SQP-method. In addition, an approximation of the Hessian H_k has to be provided. As in the Generalized Gauss-Newton method, the main load of computational work is again due to the generation of derivatives.

Applying the chain rule we obtain for the Jacobians

$$\frac{dF_1}{dv} = \frac{\partial \Phi}{\partial C}\frac{\partial C}{\partial J}\frac{dJ}{dv} \tag{4.29}$$

$$\frac{dF_i}{dv} = \frac{\partial F_i}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial F_i}{\partial v}, \qquad i = 2, 3. \tag{4.30}$$

Of particular interest is the computation of (4.29) because the objective function Φ is implicitly defined by the Jacobian J of the underlying parameter estimation problem. For more detail about the computation of the individual terms see (Bauer et al., 2000 [10]; Bauer, 2000 [8]).

In the course of the computation of (4.29) and (4.29) most of the CPU time is spent for computing derivatives of the solution $y(t; s, p, \hat{q})$ with respect to s, p and $\hat{q}$. In addition to the first-order derivatives

$$\frac{\partial y(t; s, p, \hat{q})}{\partial s}, \quad \frac{\partial y(t; s, p, \hat{q})}{\partial p} \quad \text{and} \quad \frac{\partial y(t; s, p, \hat{q})}{\partial \hat{q}}, \tag{4.31}$$

which can be efficiently derived by the methods outlined in Chapter 3, also the mixed, second-order derivatives

$$\frac{\partial^2 y(t; s, p, \hat{q})}{\partial \hat{q} \partial s} \quad \text{and} \quad \frac{\partial^2 y(t; s, p, \hat{q})}{\partial \hat{q} \partial p} \tag{4.32}$$

are needed. Similar as for the first-order derivatives, the second-order derivatives are generated by again solving variational differential equations (Bauer, 2000 [8]). These strategies are implemented in DAESOL. Thus, DAESOL is able to provide all the information required to set up the quadratic subproblem (4.25)-(4.27).

An approximation for the Hessian H_k is obtained by update formulas based on derivatives of the Lagrangian function (4.28).

4.3 Practical Requirements on Experimental Designs

In contrast to parameter estimation, where the parameters are free variables, we have so far assumed the parameters for optimal experimental design to be fixed while optimizing the controls and the weights. This may seem a paradox in so far as the identification of the (unknown) parameters is the principal target of optimizing the design. As this dependency on the parameters is inherent to the use of one of the classical design criteria for nonlinear problems, several strategies have been developed to circumvent these difficulties.

First of all, sequential procedures have been suggested (see e.g. Box and Lucas (1959) [25]). The idea is that one starts with an initial parameter guess, identifies an optimal experimental design according to one of the design criteria, performs the experiments, determines an update of the parameter values by parameter estimation incorporating the new data, and repeats this loop until a prescribed convergence criterion is reached. Nishikawa and Yeh (1989 [95]), for example, applied such a strategy for determining a D-optimal pumping test design in a groundwater system. The optimal experimental design tool VPLAN by Körkel et al. (1999 [83]) has been developed in such a way that it supports this sequential approach. They use special strategies to reduce the computational effort by freezing parts of the underlying Jacobian.

Secondly, parameter-robust design methods are employed that are intended to cover a reasonable range of parameter values (see e.g. Walter and Prozato, 1987 [138]). Determining a design that may reduce parameter variances for a broad range of parameter values is more likely to minimize the risk of missing important sensitivity information than if only one parameter set was used which may turn out to be wrong.

In Section 5.6, we will investigate the robustness of a computed A-optimal design for parameter estimation in a column outflow experiment. It will be shown that even though the A-criterion value increases for shifting the parameters within reasonable ranges, these optimized designs by far outperform commonly used straightforward designs.

However, as least as important as optimizing experimental designs to increase the reliability of parameter estimates is optimal experimental design for model discrimination. So far, optimal experimental design for parameter estimation has been derived under the assumption that the underlying "true" model is known. In practice, this is frequently not the case. Often several models seem to be possible, and the question arises which of these competing models should be chosen. Model discrimination aims at answering this type of questions. In general, first a reliable model should be selected before identifying optimal experimental designs with respect to parameter estimation. As the latter one depends on the chosen model, wrong model assumptions lead to optimized design which are, however, not optimal for estimating the true parameters.

Recently, more and more scientists dealing with column, (mini-)lysimeter or field experiments have claimed the need for reliable methods respectively tools for model discrimination.

In the literature, several strategies mainly based on a sequential approach are reported (see e.g., Hunter and Reiner, 1965 [65]; Fedorov and Malyutov, 1972 [49]; Fedorov, 1975 [48]). These approaches aim at identifying sequentially the measurement points where the competing models under consideration differ most. The idea is that measurement points of greatest difference in the model predictions contribute the most information to the discriminatory power of the sampling design. Knopman and Voss (1988 [75], 1989 [76]) and Knopman et al. (1991 [77]), for example, studied sampling designs for model discrimination in one-dimensional analytical models for transient solute transport in porous media.

However, methods for model discrimination in systems described by differential equations are barely reported. We have developed an approach for model discrimination in dynamic systems described by ODEs and DAEs, which enables to optimize in addition to the sampling design some or all initial values of the dynamic system (Dieses, 1997 [40]). As this approach

can in principle be extended to the problem class considered here, we sketch in the following the main steps of the algorithm. For the sake of presentation we assume all initial values to be optimized.

1. Set $k = 0$. Given are two models M_1 and M_2 described by ODEs/DAEs, initial values $y_1^0 = y_2^0$, initial guesses for the parameters p_1^0 and p_2^0 and measurement data corresponding to a start design Ω^0.

2. Set $k := k + 1$. Estimate the new parameters p_1^k and p_2^k on the basis of the measurement data corresponding to the design Ω^{k-1}, the initial guesses p_1^{k-1} and p_2^{k-1} and the initial values $y_1^{k-1} = y_2^{k-1}$.

3. Solve for the models M_1 and M_2 updated by the new parameters p_1^k and p_2^k, respectively, a suitable optimization problem. Determine, for example, a measurement point t^k and initial values $y_1^k = y_2^k$ within a defined range, such that the difference for the model outputs of the measured species is maximized.

4. Run a new experiment for the identified initial values $y_1^k = y_2^k$ and carry out the measurement at $t = t^k$.

5. Update the sampling design Ω^k and go to step 2.

In general, after a few iterations the underlying sampling pattern becomes obvious. The procedure is stopped when a suitable determination criterion, e.g. based on significance tests, is reached. For the solution of the parameter estimation problem constrained by ODEs/DAEs, step 2, the Generalized Gauss-Newton method PARFIT (Bock, 1981 [19]; 1983 [20]; 1987 [21]) is used. The optimization problem described in step 3 is solved by means of the direct multiple shooting method MUSCOD-II (Leineweber, 1999 [84]).

Chapter 5

Numerical Results

In this chapter the tools ECOFIT and ECOPLAN developed for parameter estimation and optimal experimental design are applied to several column, mini-lysimeter and field experiments. The objective is to determine hydraulic parameters as well as environmental fate parameters.

In the first section, the main focus is upon investigating the performance of ECOFIT for identifying van Genuchten parameters and solute transport parameters from noisy data starting from poor initial guesses (Dieses et al., 1999 [41]). In order to have a controlled scenario, measurement data is generated by solving the forward problem for a predefined true parameter set followed by adding pseudo-normally distributed noise.

The second section is devoted to the treatment of parameter estimation problems in layered soils. Considering a homogeneous sand layer on top of a homogeneous clay layer, the aim is to determine the hydraulic parameters for both layers in one experiment. Here, measurement data is generated using the van Genuchten parameters of a Dutch soil.

In Section 5.3 we describe how ECOFIT is used for estimating the van Genuchten parameters n, α and K_s from a field experiment (Dieses et al., 1999 [43]). The experimental part of this study was carried out by Aden (1999 [3]; 2000 [2]) at a BASF test site in the upper Rhine valley. Time domain reflectrometry (TDR) was employed to monitor the volumetric water contents in several depths. For these types of experiments, an adequate model is developed.

In Section 5.4 ECOFIT is applied to a mini-lysimeter study in order to determine the transport and sorption behavior of three European soils. The experiments were carried out by the Staatliche Lehr- und Forschungsanstalt (SLFA), Neustadt/Weinstr. (Fent, 1999 [50]). After the application of the non-reactive tracer bromide and of a $^{14}\mathrm{C}$-labeled test substance X, the undisturbed soil cores are irrigated by a constant daily rate and leachate volumes are sampled. A model is worked out for this typical class of outflow experiments and the unknown parameters are estimated.

In the following section we demonstrate how the environmental fate of the grass herbicide S-Metolachlor and its two main metabolites is investigated by means of ECOFIT. The data used for parameter estimation is obtained by mini-lysimeter experiments performed by Horn (1999 [62]) at a Novartis test site in Switzerland. In contrast to the mini-lysimeter study in Section 5.4, the mini-lysimeters are here exposed to normal climatic conditions. Again, the un-

known parameters, e.g. the linear sorption coefficients or the degradation rates, are determined based on leachate data. This parameter estimation problem, however, suffers from the problem of ill-posedness due to the insufficient information provided by the data available.

In the last section, the performance of the optimal design tool ECOPLAN is studied which aims at avoiding the ill-posedness problem by optimizing the sampling designs and the irrigation/application schemes in order to obtain data containing sufficient information for parameter estimation. In particular, ECOPLAN is used to design typical column outflow experiments for identifying both water and solute transport parameters with high accuracy in one experiment by means of leachate data (Dieses et al., 2000 [42]).

5.1 Column Experiment with Nonlinear Michaelis-Menten Kinetics

In the following a hypothetical soil column experiment is studied to investigate the performance of ECOFIT for estimating parameters from noisy data (Dieses et al., 1999 [41]). Generating the measurement data by solving the forward problem for the true parameter values and disturbing it, enables us to verify the results.

The aim of this column experiment is to determine the hydraulic parameters for the water transport and the environmental fate parameters for the solute transport in an unsaturated soil. In accordance with general practice, we first estimate the van Genuchten parameters n, α and K_s on the basis of point measurement data of the matric potential ψ_m and the water content θ. In a second step, the hydraulic parameters are fixed and the solute transport parameters are estimated by means of substance concentration data.

5.1.1 Estimation of Parameters in the Water Transport Equation

First, we consider the flow of water in an unsaturated soil column of 20 cm length which we model by the Richards equation in the ψ_m- respectively θ-form in order to estimate the hydraulic parameters in the van Genuchten-Mualem model:

Richards equation in ψ_m $[hPa]$

$$C(\psi_m)\frac{\partial \psi_m}{\partial t} = \frac{\partial}{\partial z}\big[K(\psi_m)\frac{\partial}{\partial z}\big(\psi_m - \bar{w}z\big)\big] \tag{5.1}$$

with

$$K(\psi_m) = K_s\frac{(1-(\alpha|\psi_m|)^{n-1}(1+(\alpha|\psi_m|)^n)^m)^2}{(1+(\alpha|\psi_m|)^n)^{m/2}}, \qquad m = 1-1/n \tag{5.2}$$

$$C(\psi_m) = \alpha(n-1)(\theta_s-\theta_r)(\alpha|\psi_m|)^{n-1}(1+(\alpha|\psi_m|)^n)^{1/n-2} \tag{5.3}$$

- Initial condition: $\psi_m(0,z) = -970, \qquad z \geq 0$

- Upper boundary:

$$q(t,0) = -K(\psi_m(t,0))\frac{\partial}{\partial z}\big(\psi_m(t,0) - \bar{w}z\big) = \begin{cases} 0.2 & t \le 6.0 \\ 0.0 & t > 6.0 \end{cases} \tag{5.4}$$

- Lower boundary: $\partial\psi_m(t,20)/\partial z = 0, \qquad t \ge 0.$

Richards equation in θ [1]

$$\frac{\partial\theta}{\partial t} = \frac{\partial}{\partial z}\big[\bar{D}(\theta)\frac{\partial\theta}{\partial z} - \bar{w}\bar{K}(\theta)\big] \tag{5.5}$$

with

$$\bar{K}(\theta) = K_s\Theta^{\frac{1}{2}}\big[1 - \big(1 - \Theta^{\frac{1}{m}}\big)^m\big]^2, \qquad \Theta = \frac{\theta - \theta_r}{\theta_s - \theta_r} \tag{5.6}$$

$$\bar{C}(\theta) = \frac{1}{\alpha n m}\big(\Theta^{-\frac{1}{m}} - 1\big)^{-m}\Theta^{-\frac{1}{m}}(\theta - \theta_r)^{-1}, \qquad m = 1 - \frac{1}{n} \tag{5.7}$$

$$\bar{D}(\theta) = \bar{K}(\theta)\bar{C}(\theta) \tag{5.8}$$

- Initial condition: $\theta(0,z) = 0.27, \qquad z \ge 0$

- Upper boundary:

$$q(t,0) = -\big[\bar{D}(\theta(t,0))\frac{\partial\theta(t,0)}{\partial z} - \bar{w}\bar{K}(\theta(t,0))\big] = \begin{cases} 0.2 & t \le 6.0 \\ 0.0 & t > 6.0 \end{cases} \tag{5.9}$$

- Lower boundary: $\partial\theta(t,20)/\partial z = 0, \qquad t \ge 0.$

The column experiment lasts 12 hours $[h]$. During the first 6 hours water is infiltrated with a constant rate $q(t,0) = 0.2\ [cm\,h^{-1}]$. Then the infiltration is stopped. This is modeled by a specified flux condition at the upper boundary. The lower boundary is described by a free drainage condition, i.e. the gradient of ψ_m respectively θ is zero.

Measurement data for ψ_m and θ for 4 depths ($z = 2, 7, 12, 17\,[cm]$) and 11 points in time ($t = 1, 2, \ldots, 11\,[h]$) is generated by solving (5.1) and (5.5) for the true parameter values $n = 1.2$ [1], $\alpha = 0.0102\ [hPa^{-1}]$, $K_s = 10.0\ [cm\,h^{-1}]$, $\theta_r = 0.05$ [1] and $\theta_s = 0.4$ [1] (Seppelt, private communication; Vink et al., 1994 [133]). This "error-free" data is disturbed as follows:

$$\eta_i = y_i(1 + \beta\epsilon) \qquad \epsilon_i \sim \mathcal{N}(0,1), \tag{5.10}$$

where β describes the factor of perturbation, e.g. $\beta = 5\%$. The solution surfaces for the true parameter values illustrating the effect of the irrigation scheme are given in Figure 5.1 and 5.2.

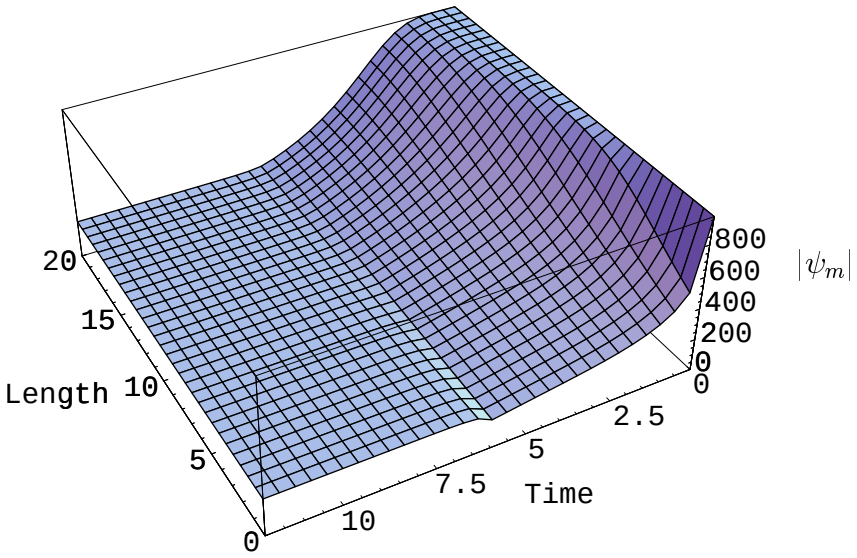

Figure 5.1: Solution surface for the Richards equation (5.1) in ψ_m $[hPa]$.

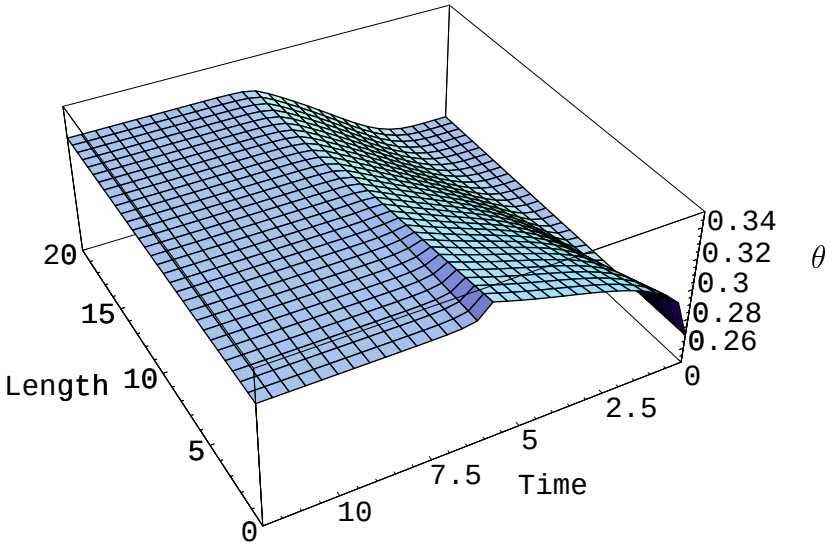

Figure 5.2: Solution surface for the Richards equation (5.5) in θ $[1]$.

	n	α	K_s	# Iter.
True Values	1.2	0.0102	10.0	
Initial Guesses	2.0	0.08	100.0	
$\beta = 1\%$	1.1993 ± 0.0011	0.01025 ± 0.00035	10.33 ± 0.88	12
$\beta = 5\%$	1.1964 ± 0.0055	0.01088 ± 0.00188	12.82 ± 5.40	16
$\beta = 10\%$	1.1918 ± 0.0108	0.01277 ± 0.00411	20.53 ± 17.23	16

Table 5.1: Parameter Estimates for n, α and K_s and 95% Confidence Intervals

	$\beta = 1\%$		$\beta = 5\%$		$\beta = 10\%$	
	α	K_s	α	K_s	α	K_s
n	-0.56	-0.68	-0.51	-0.65	-0.37	-0.54
α	-	0.97	-	0.96	-	0.97

Table 5.2: Correlation Matrices for $\beta = 1\%$, 5% and 10%

In the following we study different scenarios with increasing complexity. First, we investigate the identifiability of n , α and K_s while keeping θ_r and θ_s fixed. This is a common scenario in practice as θ_r and θ_s are often obtained separately. Secondly, we study whether it is possible to estimate from the same data also θ_r in addition to n, α and K_s. Finally, we try to determine all five parameters simultaneously. This latter estimation problem is known to be very hard as the parameters are generally highly correlated.

In the first scenario, n, α and K_s are estimated on the basis of 1%, 5% and 10% pseudo-normally disturbed data. As Table 5.1 shows, even though poor initial guesses are used all parameters can be identified and lie within the 95% confidence intervals. The corresponding correlation matrices, which do not differ much for $\beta = 1\%$, 5% and 10%, are summarized in Table 5.2.

The estimation of θ_r in addition to n, α and K_s, however, is only possible for data with $\beta = 1\%$ and improved parameter guesses (Table 5.3). As the correlation matrix, Table 5.4, reveals, the four parameters are highly correlated.

	n	α	K_s	θ_r
True Values	1.2	0.0102	10.0	0.05
Initial Guesses	1.5	0.02	20.0	0.07
$\beta = 1\%$	1.1977 ± 0.0333	0.01027 ± 0.00055	10.46 ± 2.75	0.048 ± 0.041

Table 5.3: Parameter Estimates for n, α, K_s and θ_r and 95% Confidence Intervals

	α	K_s	θ_r
n	-0.86	-0.96	0.99
α	-	0.95	-0.86
K_s	-	-	-0.97

Table 5.4: Correlation Matrix for n, α, K_s and θ_r

If we now try to determine all five parameters n, α, K_s, θ_r and θ_s, the parameter estimation problem becomes nearly singular, even for error-free data. The data does not contain enough information to identify the five van Genuchten parameters simultaneously. This situation is often encountered in practice, in particular for the identification of the hydraulic parameters.

Considering, however, an experiment with a slightly different irrigation scheme, all five parameters can be estimated. In contrast to the irrigation scheme described by the upper boundary conditions (5.4) and (5.9) we now use an alternating irrigation scheme of the form

$$\begin{aligned} q(t,0) &= -K(\psi_m(t,0))\frac{\partial}{\partial z}\big(\psi_m(t,0) - \bar{w}z\big) \\ &= -\big[\bar{D}(\theta(t,0))\frac{\partial\theta(t,0)}{\partial z} - \bar{w}\bar{K}(\theta(t,0))\big] = \begin{cases} 0.2 & 0.0 < t \leq 3.0 \\ 0.0 & 3.0 < t \leq 6.0 \\ 0.2 & 6.0 < t \leq 9.0 \\ 0.0 & 9.0 < t \leq 12.0 \end{cases} . \end{aligned} \tag{5.11}$$

This causes more spatial and temporal activity of the solution surface, as Figure 5.3 and 5.4 show, and leads to measurement data which exhibits a higher sensitivity to the unknown parameters. Now the parameter estimation problem as such is well-posed and it is possible to identify all parameters from undisturbed data starting from the initial guesses $n = 1.3$, $\alpha = 0.01$, $K_s = 11.0$, $\theta_r = 0.07$ and $\theta_s = 0.5$.

Here, the modification of the upper boundary was chosen intuitively based on the insight gained by studying similar problems. Considering real life problems, however, designs are required that guarantee well-posed parameter estimation problems even for noisy data. Thus, the choice of good irrigation schemes should be embedded in the framework of optimal experimental design which aims at optimizing the designs according to some prescribed design criteria. In Section 5.6. we will outline how an optimal experimental design can be determined by ECOPLAN that enables the simultaneous estimation of both water and solute transport parameters in only one outflow experiment.

5.1.2 Estimation of Parameters in the Solute Transport Equation

After the hydraulic parameters have been identified, the aim is now to determine the Michaelis-Menten parameters V_{max} $[ML^{-3}d^{-1}]$ and K_M $[ML^{-3}]$ and a parameterization of the diffusion-dispersion term, b $[1]$ and D_m $[cm^2d^{-1}]$. The transport of the dissolved substance, e.g. a pesti-

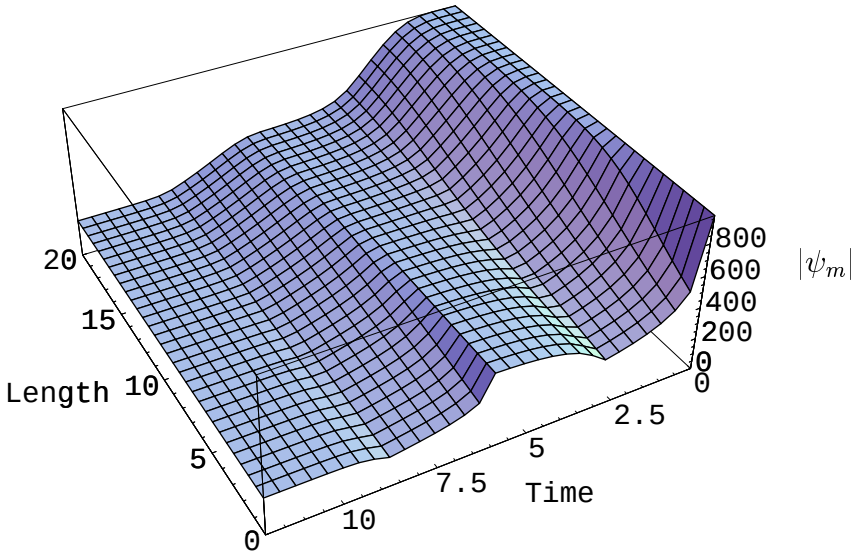

Figure 5.3: Solution surface for the Richards equation (5.1) in ψ_m $[hPa^{-1}]$ with the modified upper boundary condition (5.11).

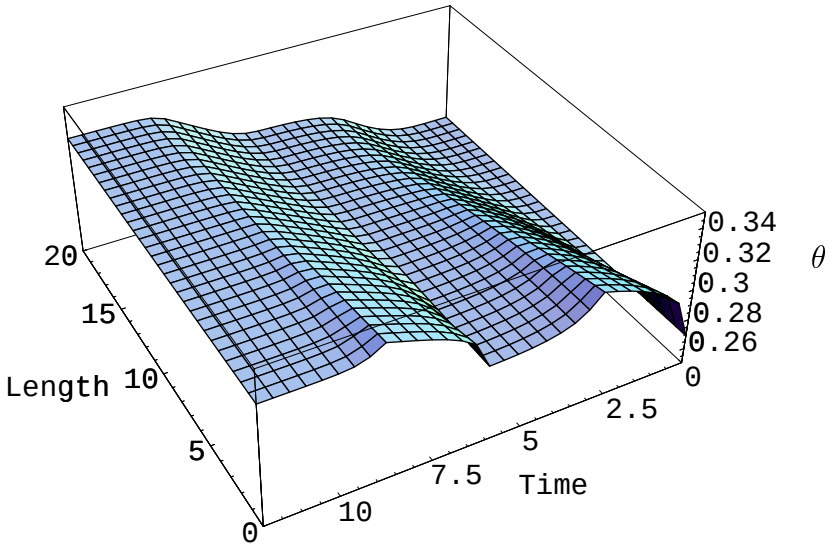

Figure 5.4: Solution surface for the Richards equation (5.5) in θ $[1]$ with the modified upper boundary condition (5.11).

cide, in the unsaturated soil is modeled by two coupled PDEs, namely the Richards equation in the ψ_m-form (5.1) as discussed in the previous section and the following convection-dispersion equation:

Solute transport

$$\frac{\partial}{\partial t}(\theta c) = \frac{\partial}{\partial z}\big(\theta D_h(\theta)\frac{\partial c}{\partial z} - qc\big) - \theta\frac{V_{max}c}{c+K_M} \tag{5.12}$$

with

$$D_h(\theta) = \frac{0.0046\exp(b\theta)}{\theta} + D_m \tag{5.13}$$

- Initial condition: $c(0,z) = 0, \qquad z \geq 0$
- Upper boundary:

$$-\theta(t,0)D_h(\theta(t,0),q(t,0))\frac{\partial c(t,0)}{\partial z} + q(t,0)c(t,0) = \begin{cases} c_0\, q(t,0) & t \leq 6.0 \\ 0 & t > 6.0 \end{cases}$$

- Lower boundary: $\partial c(t,20)/\partial z = 0, \qquad t \geq 0.$

In contrast to the initial condition used in Section 5.6.1, now a wetter column is employed with $\psi_m(t,0) = -670\ [hPa]$ for $z > 0$. The van Genuchten parameters are fixed to the true values $n = 1.2\ [1]$, $\alpha = 0.0102\ [hPa^{-1}]$, $K_s = 10.0\ [cm\,h^{-1}]$, $\theta_r = 0.05\ [1]$ and $\theta_s = 0.4\ [1]$. Both equations (5.1) and (5.12) are coupled by the water flux $q(t,z)$ and the water content $\theta(t,z)$ which is related to the matric potential $\psi_m(t,z)$ by the water retention curve

$$\theta(\psi_m) = (\theta_s - \theta_r)\big(1 + (\alpha|\psi_m|)^n\big)^{-m} + \theta_r, \qquad m = 1 - \frac{1}{n}. \tag{5.14}$$

During the first half of the experiment with $T_{end} = 12\ [h]$ the dissolved substance $c_0 = 10.0\ [ML^{-3}]$ is infiltrated together with the water, then the infiltration is stopped. Figure 5.5 illustrates the fate of the substance which is degraded within the first 10 cm such that no substance reaches the end of the column.

1%, 2%, 3%, 5% and 10% pseudo-normally disturbed data for the substance concentration c is generated for 4 depths ($z = 2, 7, 12, 17\ [cm]$) and 11 points in time ($t = 1, 2, \ldots, 11\ [h]$) by means of the procedure used for the ψ_m- and θ-data. Starting from the initial guesses given in Table 5.5, the parameters can be estimated in all cases and lie within the 95% confidence intervals. As the corresponding correlation matrices are similar, only the one for $\beta = 1\%$ is shown here representatively in Table 5.6.

In summary, we can state that ECOFIT enables to identify even highly correlated parameters from noisy data. Even for poor initial guesses, convergence to the true parameters is reached within few iterations.

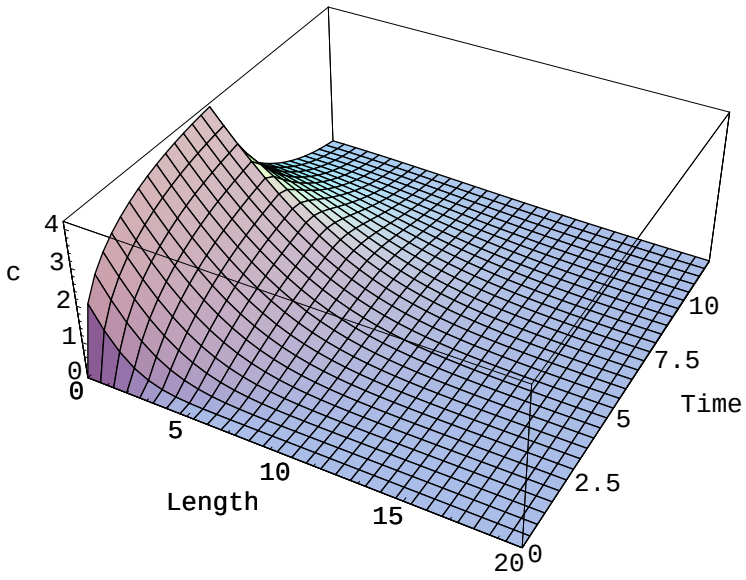

Figure 5.5: Solution surface of the convection-dispersion equation (5.12) for the substance concentration $c\,[ML^{-3}]$.

	V_{max}	K_M	b	D_m	# Iter.
True Values	0.2	1.0	10.0	5.0	
Initial Guesses	0.1	0.5	17.0	2.0	
$\beta = 1\%$	0.205 ± 0.011	1.045 ± 0.117	10.22 ± 2.51	4.95 ± 0.31	5
$\beta = 2\%$	0.209 ± 0.024	1.095 ± 0.240	10.41 ± 4.41	4.91 ± 0.55	5
$\beta = 3\%$	0.214 ± 0.036	1.139 ± 0.371	10.64 ± 6.24	4.86 ± 0.84	5
$\beta = 5\%$	0.228 ± 0.068	1.276 ± 0.711	10.87 ± 9.20	4.77 ± 1.38	5
$\beta = 10\%$	0.270 ± 0.178	1.703 ± 1.897	11.50 ± 13.95	4.56 ± 2.53	5

Table 5.5: Parameter Estimates for V_{max}, K_M, b and D_m and 95% Confidence Intervals

	K_M	b	D_M
V_{max}	0.99	-0.54	0.32
K_M	-	-0.48	0.24
b	-	-	-0.97

Table 5.6: Correlation Matrix for V_{max}, K_M, b and D_M for 1% Disturbed Data

5.2 Estimation of Hydraulic Parameters in a Layered Soil

In the column experiment investigated in the previous section, we assumed the soil to be homogeneous. In general, this assumption only holds for repacked soil columns. Considering, however, lysimeter studies where undisturbed soil cores are used, or even field experiments, the soils might exhibit horizontally parallel layers.

In the following we study a two-layer soil column of 40 cm length composed of a homogeneous sand layer on top of a homogeneous clay layer, both of 20 cm length. The aim is to estimate the hydraulic parameters n, α and K_s of both layers in one experiment based on ψ_m- and θ-measurement data. As in Section 5.1.1, the water transport in the unsaturated soil is modeled by the Richards equation in ψ_m (5.1) and θ (5.5). In addition to the initial and boundary conditions, a transition condition at the layer interface has to be formulated. This is discussed in the following for ψ_m $[hPa]$. The corresponding condition for θ $[1]$ can be derived analogously.

- Initial condition: $\psi_m(0,z) = -970 + 20 \cdot z, \qquad z \geq 0$
- Upper boundary:

$$q(t,0) = -K(\psi_m(t,0))\frac{\partial}{\partial z}\big(\psi_m(t,0) - \bar{w}z\big) = \begin{cases} 0.2 & 0.0 < t \leq 6.0 \\ 0.0 & 6.0 < t \leq 12.0 \\ 0.2 & 12.0 < t \leq 18.0 \\ 0.0 & 18.0 < t \leq 24.0 \end{cases} \tag{5.15}$$

- Lower boundary: $\partial\psi_m(t,40)/\partial z = 0, \qquad t \geq 0.$
- **Condition at the layer interface**
  As the water flux $q(t,20)$ is the same in both layers, at the layer interface the following transition condition must hold:

$$\begin{aligned} q(t,20) &= -K_{sand}(\psi_m(t,20))\big[\frac{\partial}{\partial z}\big(\psi_m(t,20) - \bar{w}z\big)\big]_{sand} \\ &= -K_{clay}(\psi_m(t,20))\big[\frac{\partial}{\partial z}\big(\psi_m(t,20) - \bar{w}z\big)\big]_{clay}. \end{aligned} \tag{5.16}$$

As true values for the parameter estimation problem the van Genuchten parameters in Table 5.7 are used. They are based on the Dutch texture classes and characterize a loam-poor fine sand and a light sandy clay (Wösten et al., 1994 [141]).

Figure 5.6 and 5.7 illustrate the solution surfaces for the Richards equation in ψ_m $[hPa]$ and θ $[1]$ using the van Genuchten parameters given in Table 5.7. For each layer measurement data for ψ_m and θ for 3 depths ($z = 0.5, 7.5, 15$; $z = 22.5, 25.5, 32.5$ $[cm]$) and 24 points in time ($t = 1, 2, \ldots, 24$) are generated employing the true values in Table 5.7. The results in Table 5.8 demonstrate that the parameters n^1, α^1 and K_s^1 in the first layer, where the solution surface exhibits a high spatial and temporal activity, are well determined by the data and thus can be

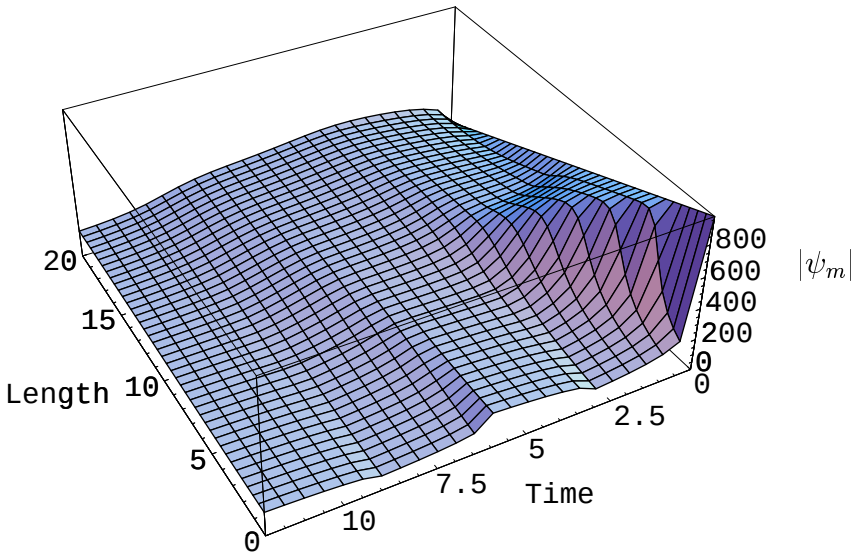

Figure 5.6: Two layers: Solution surface for the Richards equation in ψ_m $[hPa]$.

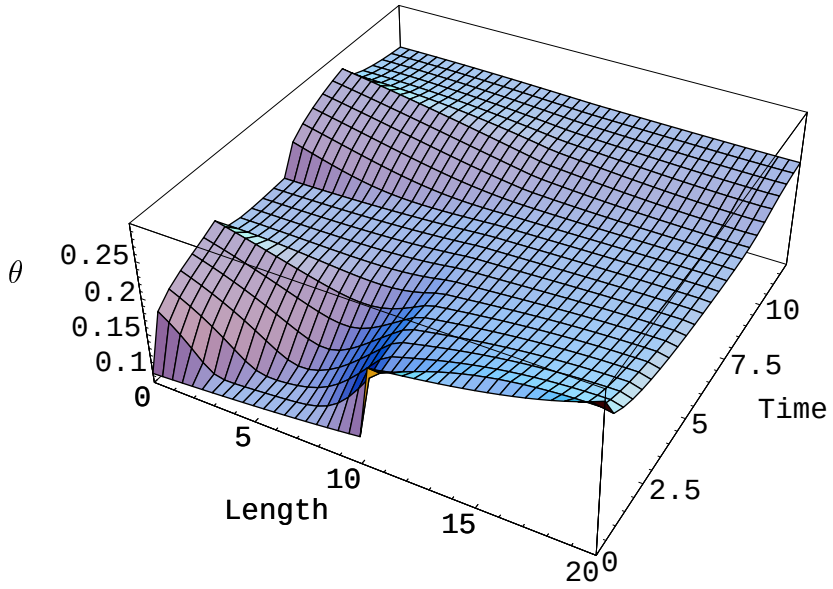

Figure 5.7: Two Layers: Solution surface for the Richards equation in θ $[1]$.

	$n\ [1]$	$\alpha\ [hPa^{-1}]$	$K_s\ [cm\,d^{-1}]$	$\theta_r\ [1]$	$\theta_s\ [1]$
Sand	1.507	0.0249	17.46	0.43	0.01
Clay	1.250	0.0194	14.07	0.40	0.0

Table 5.7: Van Genuchten Parameters for a Loam-Poor Fine Sand and a Light Sandy Clay (Wösten et al., 1994 [141])

estimated with high accuracy. The estimation of the parameters n^2, α^2 and K_s^2 in the second layer, however, is more difficult since these parameters are insufficiently determined by the data. As Table 5.9 illustrates, they are much higher correlated than the parameters n^1, α^1 and K_s^1 in the sand layer. This example demonstrates that, in order to identify parameters in deeper layers, better experimental designs are required to guarantee identifiability of all parameters in practice.

	True Values	Initial Guesses	$\beta = 1\%$	$\beta = 5\%$
n^1	1.507	2.0	1.5189 ± 0.0321	1.5731 ± 0.1732
α^1	0.0249	0.03	0.02488 ± 0.00062	0.02461 ± 0.00379
K_s^1	17.46	20.0	17.83 ± 1.32	19.13 ± 6.85
n^2	1.250	2.0	1.2465 ± 0.0256	1.2452 ± 0.1064
α^2	0.0194	0.03	0.01690 ± 0.01663	0.01261 ± 0.05592
K_s^2	14.07	20.0	9.79 ± 22.86	4.24 ± 43.29

Table 5.8: Estimates for the Parameters n^1, α^1 and K_s^1 in the Sand Layer and for the Parameter n^2, α^2 and K_s^2 in the Clay Layer Including 95% Confidence Intervals

	α^1	K_s^1	n^2	α^2	K_s^2
n^1	-0.62	0.48	-0.51	-0.34	-0.34
α^1	-	0.88	0.40	-0.65	-0.65
K_s^1	-	-	-0.19	0.78	-0.77
n^2	-	-	-	0.79	0.80
K_s^2	-	-	-	-	0.99

Table 5.9: Correlation Matrix for n^1, α^1, K_s^1, n^2, α^2 and K_s^2

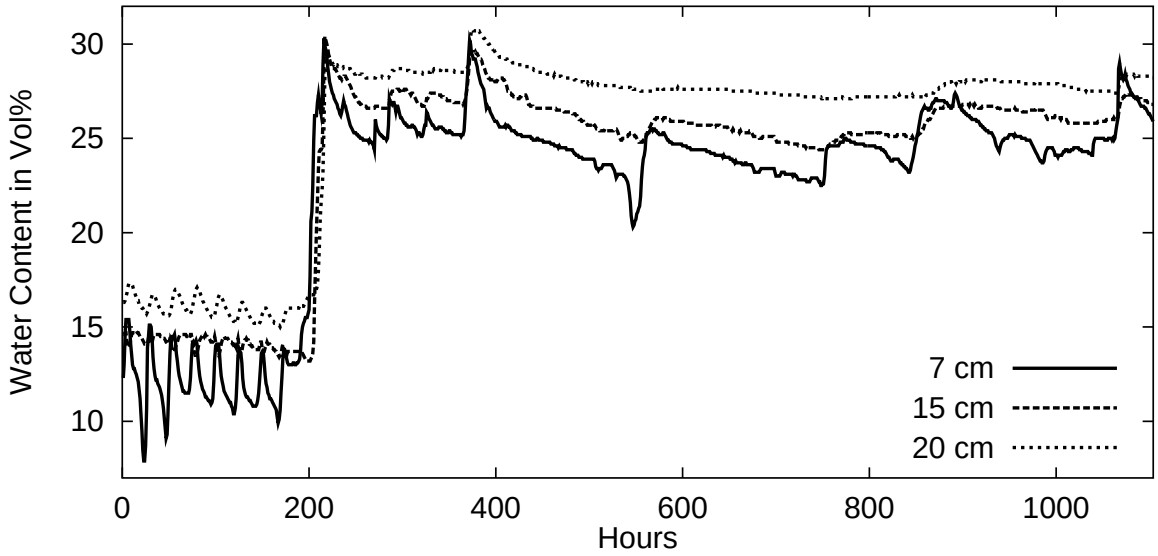

Figure 5.8: Hourly water content data in $7\ cm$, $15\ cm$ and $20\ cm$ depth.

5.3 Estimation of Van Genuchten Parameters in a Field Experiment

In the following application ECOFIT is used to determine the water transport parameters from a field experiment (Dieses et al., 1999 [43]). The experimental data used here was obtained in the context of an environmental fate study of the herbicide Metazachlor, which was carried out by Aden (1999 [3]; 2000 [2]) in her PhD project at the BASF AG, Agricultural Center Limburgerhof, and the Institute of Geoecology, Technical University of Braunschweig. As a prerequisite for a reliable modeling of Metazachlor under outdoor conditions, the correct representation of the water movement is necessary. The purpose is to identify the van Genuchten parameters n, α and K_s from water content data. The residual and the saturated water content, θ_r and θ_s, are determined by other methods and are kept fixed during the estimation process ($\theta_r = 0.03\ [1]$, $\theta_s = 0.334\ [1]$).

The sampling is conducted on an agricultural site without crop cover in the upper Rhine valley. The top zone of the loamy sand soil, which is of particular interest when studying the dissipation behavior of xenobiotics, consists of 10 % clay, 13 % silt and 77 % sand and has an organic carbon content of 0.6 %. Time-domain reflectrometry (TDR) is used to monitor the volumetric water content θ in the field. The TDR probes are inserted at $7\ cm$, $15\ cm$ and $20\ cm$ depth. Water content data is recorded on an hourly basis. Figure 5.8 shows the monitoring curves during a time period of 46 days (28 October - 13 December 1997). As Figure 5.8 reveals, the TDR probes exhibit strong daily variations which particularly affect the measurements of the most upper TDR probe. To smooth these fluctuations the hourly data is averaged to obtain daily ones.

Due to the loose texture of the dry soil within the first $10\ cm$, the actual position of the upper TDR probe which is originally fixed in $7\ cm$ depth is highly uncertain. In addition, the heavy rainfall, which starts after approximately 200 hours (3 November 1997), leads to a compression of the upper soil horizon which again shifts the position of the upper TDR probe. Therefore, the determination of its actual depth should also be part of the optimization process.

5.3.1 Modeling of TDR-Measurements

The water transport is modeled by the Richards equation in θ [1]:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\big[\bar{D}(\theta)\frac{\partial \theta}{\partial z} - \bar{w}\bar{K}(\theta)\big] \tag{5.17}$$

with

$$\bar{K}(\theta) = K_s \Theta^{\frac{1}{2}}\big[1 - \big(1 - \Theta^{\frac{1}{m}}\big)^m\big]^2, \qquad \Theta = \frac{\theta - \theta_r}{\theta_s - \theta_r} \tag{5.18}$$

$$\bar{C}(\theta) = \frac{1}{\alpha n m}\big(\Theta^{-\frac{1}{m}} - 1\big)^{-m}\Theta^{-\frac{1}{m}}(\theta - \theta_r)^{-1}, \qquad m = 1 - \frac{1}{n} \tag{5.19}$$

$$\bar{D}(\theta) = \bar{K}(\theta)\bar{C}(\theta). \tag{5.20}$$

In contrast to the column experiments discussed so far, here the initial and upper boundary conditions can not be controlled. They are defined by the climatic conditions at the agricultural site, i.e. by the history of the soil and the precipitation rates. This makes a suitable modeling much more difficult. A straightforward way to describe the upper boundary is to formulate a specified flux condition for the precipitation rates. Unfortunately, due to strong local climatic differences during the experimental period between the agricultural site and the weather station where the precipitation data is recorded, this data can not be used for defining the upper boundary condition.

A remedy is provided by pursuing the idea to formulate Dirichlet conditions at the upper and the lower boundary employing the measurement data in $7\ cm$ and $20\ cm$ depth, respectively. In this approach we assume the measurement data used for defining the boundary conditions to be free of measurement errors. In general, TDR performs well for determining volumetric water contents and delivers data which is of high accuracy relatively to each other. The absolute measurement error lies in the range of 1 Vol$\%$. Thus, taking into account the other uncertainties, such as the position of the TDR probes or probable inhomogeneities of the soil, it is justifiable to neglect this error.

In order to describe the initial profile on the first day, i.e. the initial conditions, a linear interpolation of the corresponding water content measurements in $7\ cm$, $15\ cm$ and $20\ cm$ depth are used.

5.3.2 Estimation Results

The aim is now to estimate the hydraulic parameters n, α and K_s by fitting the TDR data in $15\ cm$ depth. In addition, due to uncertainty about the correct depth of the upper TDR probe,

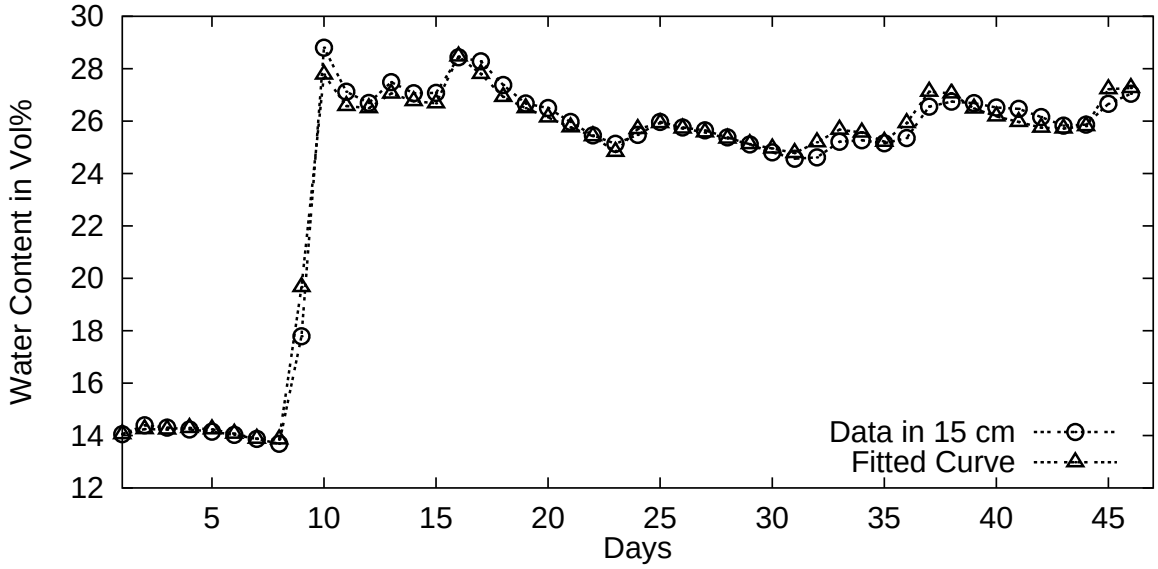

Figure 5.9: Fitted and measured curves in $15\ cm$ depth.

	$n\ [1]$	$\alpha\ [hPa^{-1}]$	$K_s\ [cm\,d^{-1}]$
Initial Guesses	1.5	0.05	35.0
Estimates	1.262 ± 0.036	0.0324 ± 0.0024	20.92 ± 1.68

Table 5.10: Parameter Estimates for n, α and K_s and 95% Confidence Intervals

the determination of its actual depth is also included in the optimization process.

As Table 5.10 shows, all parameters can be identified. The corresponding 95% confidence intervals and the correlation matrix are given in Table 5.10 and Table 5.11, respectively.

For the position of the upper TDR probe a depth of $12\ cm$ is found. A good fit is gained as the comparison of the fitted and measured curves in Figure 5.9 illustrates.

In summary, we can state that very encouraging results for the field situations are obtained by ECOFIT in spite of numerous uncertainties in the water transport model (e.g. inhomogeneity of the soil, cracks etc.) and in the measurement data (e.g. depths of the TDR probes, temperature dependence of TDR).

	α	K_s
n	0.14	-0.61
α	-	-0.94

Table 5.11: Correlation Matrix for n, α and K_s and 95% Confidence Intervals

5.4 Mini-Lysimeter Study: Determination of Transport and Sorption Parameters for Bromide and a Substance X

The purpose of this study is to determine the transport and sorption behavior of a substance X in three European soils. The corresponding mini-lysimeter experiments were carried out at the Staatliche Lehr- und Forschungsanstalt (SLFA) Neustadt/Weinstr. (Fent, 1999 [50]). ECOFIT is used for the estimation of the transport and sorption parameters.

The experiments are based on undisturbed soil columns which are treated with the conservative tracer bromide and a radio-labeled test substance X. The leachates are collected at the base of the column and the concentrations are analyzed every 12 hours. The unknown parameters, e.g. the diffusion-dispersion coefficient or sorption parameters, are estimated by fitting the breakthrough curve generated by the leachate concentrations.

For this study three test sites without crop cover are chosen that are specified by the soil characteristics given in Table 5.12. From each test site several undisturbed soil cores are sam-

	Clay $[\%]$	Silt $[\%]$	Sand $[\%]$	C_{org} $[\%]$
Soil A	4.1	21.8	74.1	0.6
Soil B	6.0	7.0	87.0	1.3
Soil C	16.0	22.0	62.0	1.2

Table 5.12: Soil Characteristic Data for the Test Soils A, B and C

pled by forcing steel tubes of $21.1\,cm$ diameter and $30\,cm$ length into the ground by means of a hammer. When the steel tube has reached $30\,cm$ depth, the soil around the tube is removed and the soil at the lower end of the tube is cut off with a knife.

Before the outflow experiment is started, the soil columns are equilibrated for 7 days by applying the same irrigation rate of $0.34\ cm\,d^{-1}$ as later in the experiment.

For the application solutions are used containing $0.5\,g$ bromide respectively $0.737\,mg$ of the test substance. The replicate columns of each soil are applied at the same time by means of a mask with 20 holes. The mask is placed on the top of the column. Through each hole $200\ \mu l$ of the bromide solution are applied to the soil surface by use of a pipettor. Then the mask is rotated about 45 degrees and the dissolved test solution is added in a similar way. Immediately after application, the soil columns are irrigated with $0.2\ cm$ water within 72 minutes. During the experiment an average irrigation rate of $0.34\ cm\,d^{-1}$ is applied.

5.4.1 Parameter Estimation for Bromide Outflow Data

In a first step the solute transport behavior in the three soils is investigated employing the conservative tracer bromide which shows neither sorption nor degradation.

Figures 5.10, 5.13 and 5.16 illustrate the analyzed concentrations of bromide in the measured leachates as a function of time. Even though the replicates partially differ considerably, all columns exhibit the typical breakthrough behavior.

Modeling

As the experiment is carried out under steady-state conditions both the water flux q and the water content θ are constant, and it is sufficient to only consider the convection-dispersion equation of the form

$$\frac{\partial c_r}{\partial t} = D\frac{\partial^2 c_r}{\partial z^2} - v\frac{\partial c_r}{\partial z}. \tag{5.21}$$

Here, $c_r\ [g\,l^{-1}]$ is the (resident) concentration, $D\ [cm^2 d^{-1}]$ is the diffusion-dispersion coefficient and $v = q/\theta\ [cm\,d^{-1}]$ is the pore water velocity.

Investigating the initial and boundary conditions, we have found that it is advantageous to assume the bromide to be incorporated within the first $0.4\ cm$ of the soil column, rather than modeling the instantaneous pulse of the 20 holes as such. This leads to the following initial condition:

$$c_r(0,z) = \begin{cases} c_0 & z \le 0.4\,cm \\ 0 & z > 0.4\,cm, \end{cases} \tag{5.22}$$

where $c_0\ [g\,cm^{-3}]$ denotes the soil concentration of the incorporated bromide. As no bromide is added with the irrigated water during the experiment, a zero flux condition is formulated at the upper boundary:

$$\Big(v c_r(t,z) - D\frac{\partial c_r(t,z)}{\partial z}\Big)\Big|_{z=0} = 0 \qquad t > 0. \tag{5.23}$$

In the examples discussed so far, we have used a Neumann condition of the form $\partial c/\partial z = 0$ at the lower boundary assuming the dispersive flux to be zero and considering only the convective transport along with the water flow. Here, according to the approach of van Genuchten and Parker (1984 [131], 1984 [99]) for the modeling of outflow experiments, formally an infinite lower boundary is defined

$$\frac{\partial c_r(t,z)}{\partial z}\Big|_{z=\infty} = 0, \tag{5.24}$$

which in our case is substituted by

$$\frac{\partial c_r(t,z)}{\partial z}\Big|_{z=L} = 0, \tag{5.25}$$

with a sufficiently large L, e.g. $L = 60\,cm$. The idea behind choosing such a type of boundary is to compute at the actual lower boundary of the soil column, i.e. $L_e = 30\,cm$, so called

flux-averaged concentrations $c^{flux}(t,z)$. They are frequently used for the description of the concentrations passing the boundary. The flux-averaged concentrations $c^{flux}(t,z)$ can be derived from the normal, so called resident concentrations $c_r(t,z)$ via

$$c^{flux}(t,z) = c_r(t,z) - \frac{D}{v}\frac{\partial c_r(t,z)}{\partial z}. \tag{5.26}$$

As the leachate concentrations are not analyzed continuously, but only every 12 hours, the measured concentrations are averaged concentrations which we model by

$$M^i(t_i, L_e) = \frac{1}{t_i - t_{i-1}} \int_{t_i-1}^{t_i} c^{flux}(\tau, L_e) d\tau \qquad i = 1, \cdots, n. \tag{5.27}$$

In ECOFIT, the quantities $M^i(t_i, L_e)$ can be easily computed by introducing one additional ODE.

Estimation Results

Parameters to be estimated are the pore water velocity v, the diffusion-dispersion coefficient D and the initial soil concentration c_0 of the incorporated tracer. In principle, we can derive c_0 by the fraction of the amount of bromide used and the volume of soil water V_w in the consider soil volume V_s

$$c_0 = \frac{0.5g}{V_w}. \tag{5.28}$$

Hereby, the soil volume V_s is given by

$$V_s = surface\ of\ the\ soil\ column \times incorporation\ depth = 139.8\, cm^3. \tag{5.29}$$

In order to compute $V_w = \theta V_s$, the actual water content θ is required, which is not available a priori. A posteriori, however, we can approximate the water content by the estimate of the pore water velocity via $\theta = q/v$ with $q = 0.34\, cm\, d^{-1}$. This independent value for c_0 can finally be used to validate the estimates.

For all soil columns the parameters can be estimated. The estimates including their 95% confidence intervals and the corresponding correlation matrices are summarized in the Tables 5.13 and 5.14. As the Figures 5.11-5.18 show, good fits are obtained in all cases.

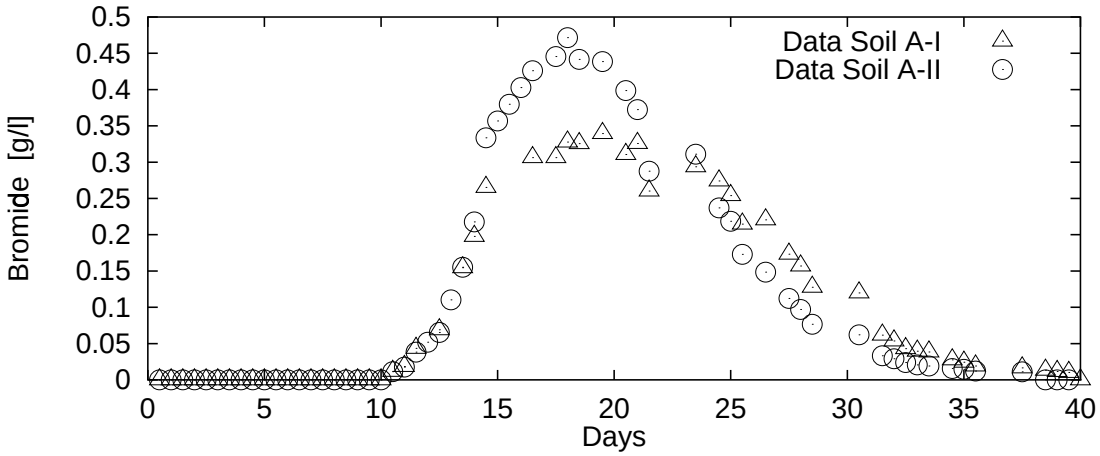

Figure 5.10: Measured bromide concentrations for the columns A-I and A-II.

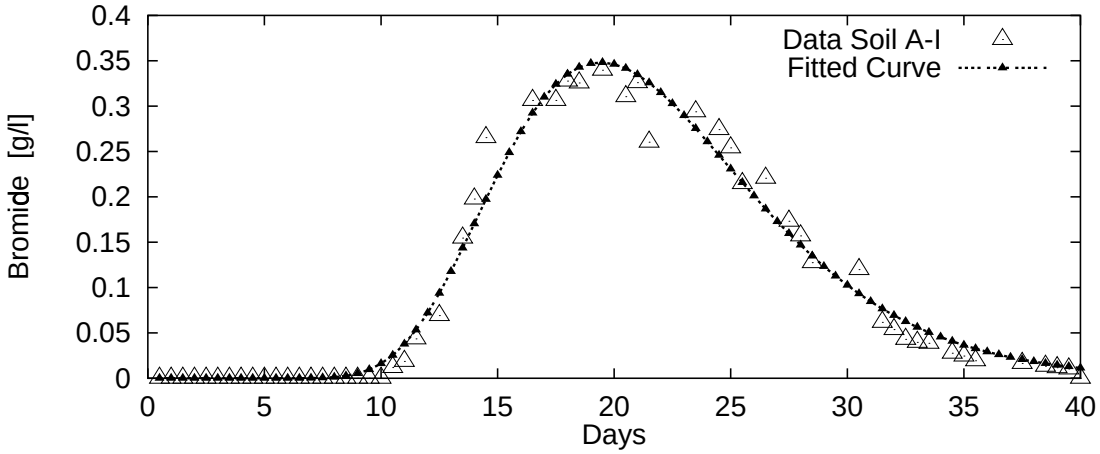

Figure 5.11: Fitted and measured bromide concentrations for column A-I.

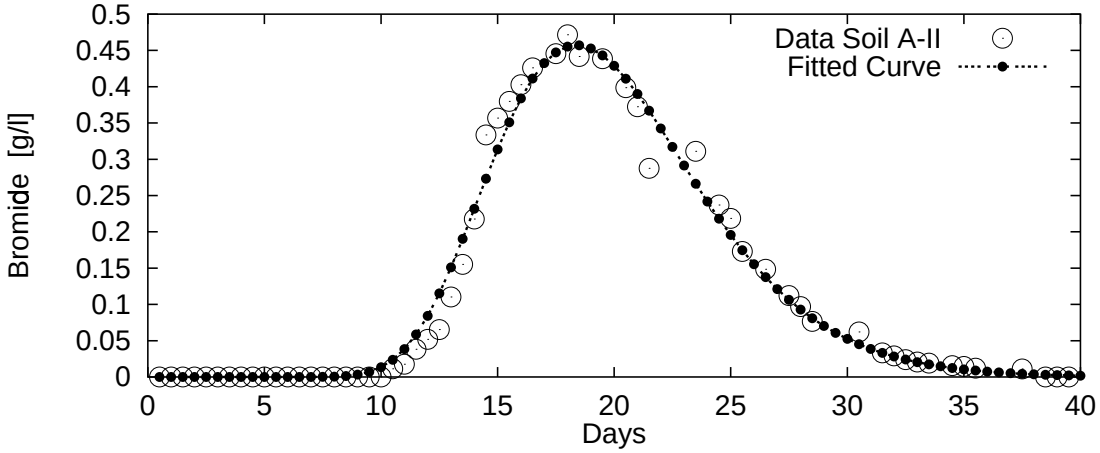

Figure 5.12: Fitted and measured bromide concentrations for column A-II.

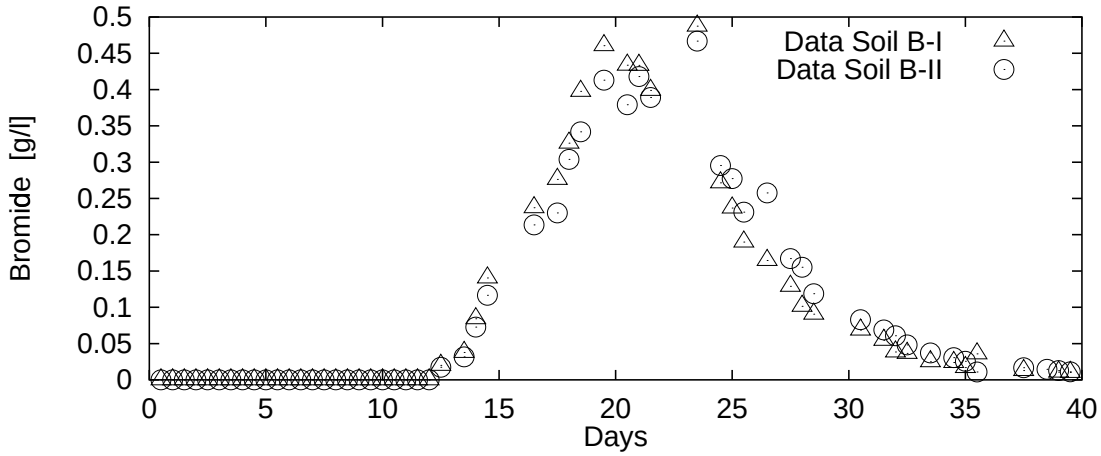

Figure 5.13: Measured bromide concentrations for the columns B-I and B-II.

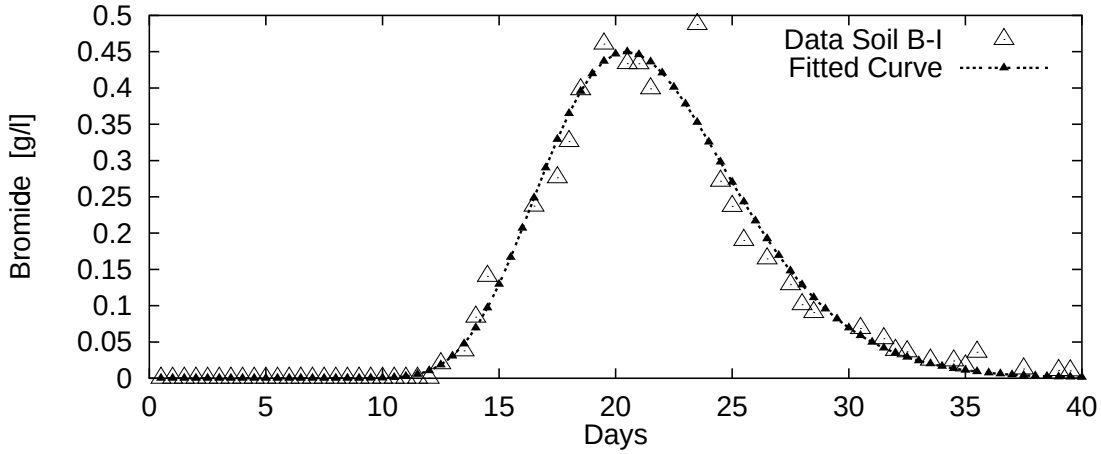

Figure 5.14: Fitted and measured bromide concentrations for column B-I.

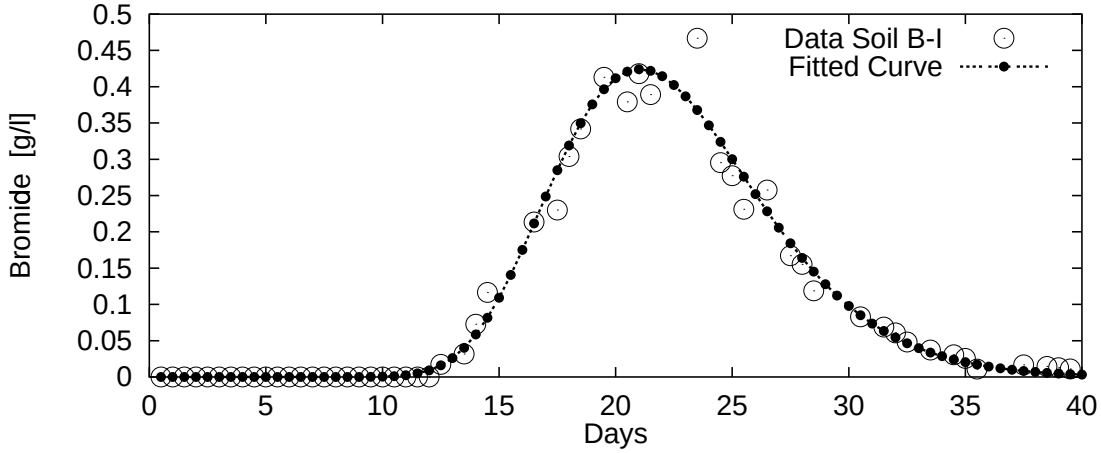

Figure 5.15: Fitted and measured bromide concentrations for column B-II.

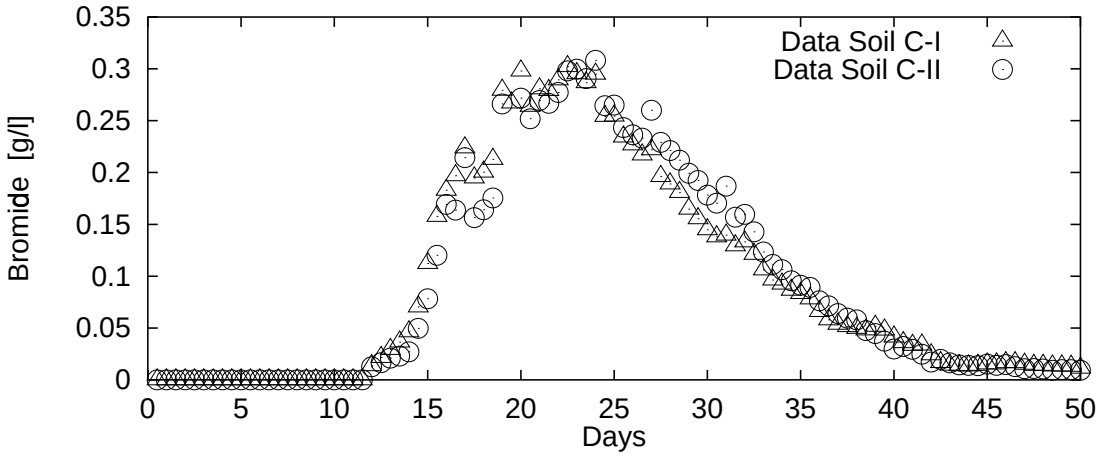

Figure 5.16: Measured bromide concentrations for the columns C-I and C-II.

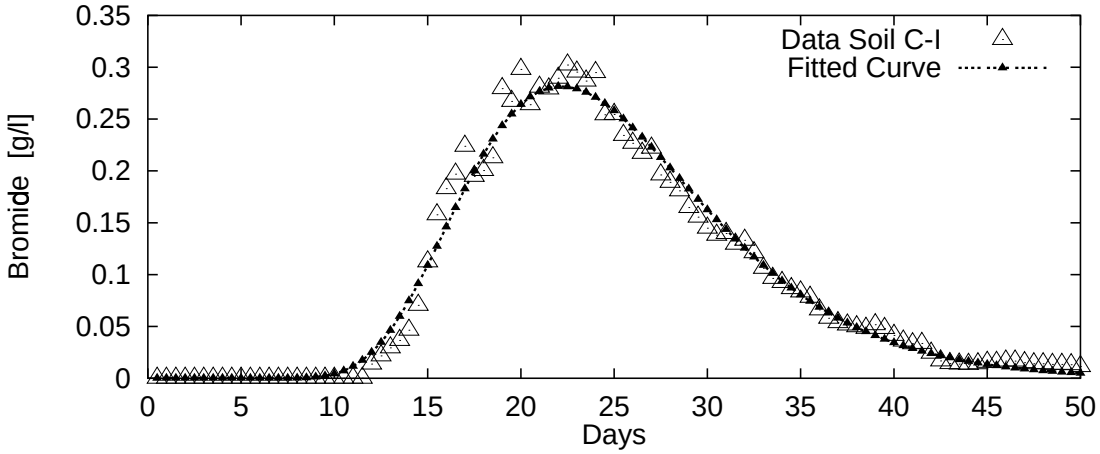

Figure 5.17: Fitted and measured bromide concentrations for column C-I.

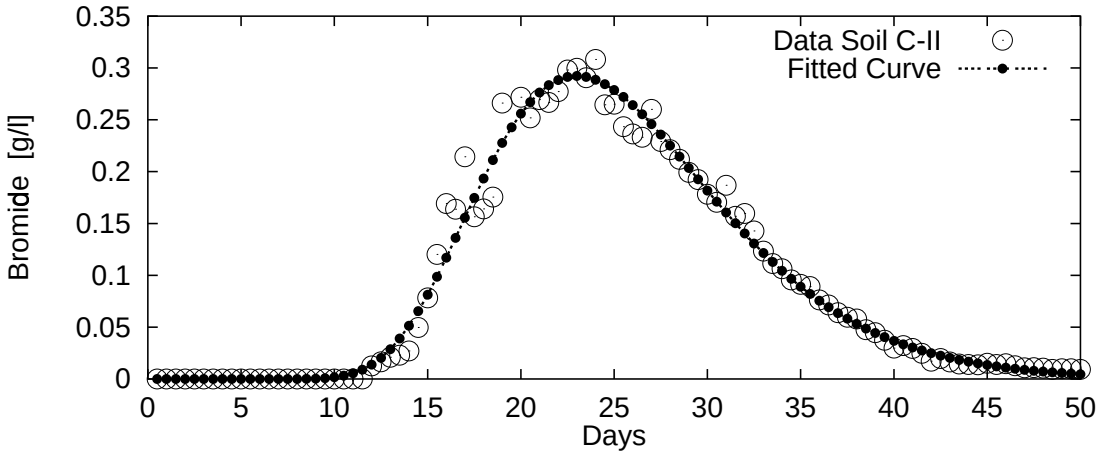

Figure 5.18: Fitted and measured bromide concentrations for column C-II.

	$v\ [cm\,d^{-1}]$	$D\ [cm^2 d^{-1}]$	$c_0\ [g\,l^{-1}]$
Soil A-I	1.387 ± 0.011	1.646 ± 0.095	20.98 ± 0.11
Soil A-II	1.502 ± 0.007	1.273 ± 0.061	23.66 ± 0.10
Soil B-I	1.39 ± 0.006	0.84 ± 0.048	20.25 ± 0.10
Soil B-II	1.35 ± 0.004	0.88 ± 0.031	19.86 ± 0.07
Soil C-I	1.20 ± 0.006	1.54 ± 0.063	17.47 ± 0.07
Soil C-II	1.18 ± 0.003	1.33 ± 0.027	17.30 ± 0.03

Table 5.13: Estimates of v, D and c_0 and 95 $\%$ Confidence Intervals

	v-D	v-c_0	D-c_0
Soil A-I	-0.77	-0.18	0.55
Soil A-II	-0.62	0.45	0.47
Soil B-I	-0.60	0.24	0.56
Soil B-II	-0.65	-0.21	0.61
Soil C-I	-0.61	0.74	0.02
Soil C-II	-0.59	0.71	0.19

Table 5.14: Correlation Matrix

Now it remains to check whether the estimated value c_0 and the one computed by (5.28) are of the same order of magnitude. For Soil A-I, for example, we obtain an approximation of the actual water content of $\theta = q/v = 0.245\ [1]$ which results according to (5.28) and (5.29) in an initial concentration of $c_0^{calc} = 14.5\ [g\,l^{-1}]$. Considering all the uncertainties involved in the experiment, e.g. the application and the irrigation, and in the modeling, the match of both values, 14.5 and 20.98, is very satisfying and validates the values found by optimization.

5.4.2 Parameter Estimation for Substance Outflow Data

Studying reactive substances, the sorption behavior is generally of particular interest. In the following, two modeling approaches are discussed for the sorption behavior of the test substance X, namely a linear sorption approach and a nonlinear one described by a Langmuir isotherm. In both cases equilibrium conditions are assumed.

Linear Sorption

In a first step we study the performance of the linear sorption model, which is the standard model required by authorities for registration. Similar to the non-reactive tracer, the test substance X is modeled by a convection-dispersion equation

$$R\frac{\partial c_r}{\partial t} = D\frac{\partial^2 c_r}{\partial z^2} - v\frac{\partial c_r}{\partial z}. \tag{5.30}$$

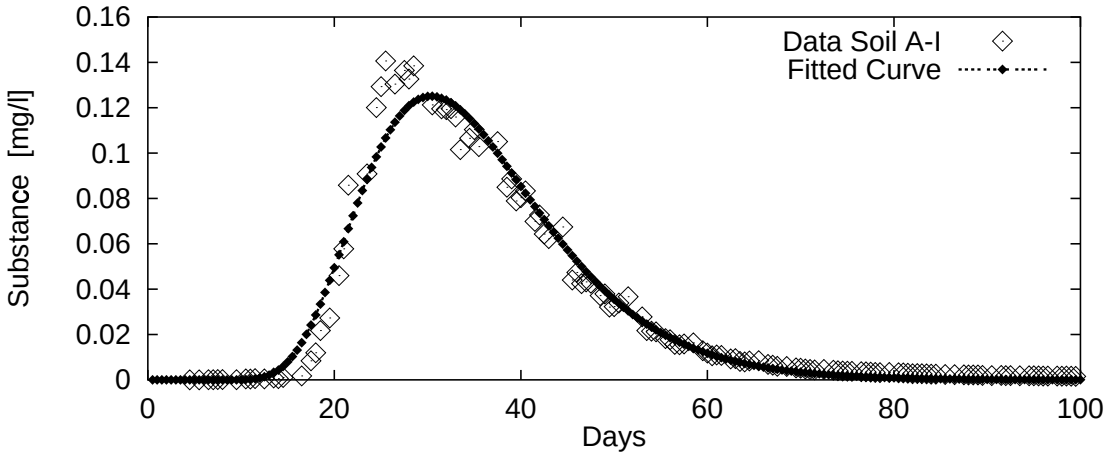

Figure 5.19: Linear sorption approach: fitted and measured concentrations of the test substance X in column A-I.

This time, however, a retardation factor is considered which is for the linear sorption of the form

$$R = 1 + \frac{\rho}{\theta} K_d. \tag{5.31}$$

The initial and boundary conditions are identical to the ones used for the tracer bromide (5.22), (5.23) and (5.25). Here, $v = 1.387\ [cm\, d^{-1}]$ is fixed to the value calculated in the previous section. ρ is put to 1.53 $[g\, cm^{-3}]$. Parameters to be estimated are the linear sorption coefficient K_d, the diffusion-dispersion coefficient D and the initial value c_0.

Even though all parameters can be identified, Table 5.15 and 5.16, the fitting results in Figure 5.19 reveal that the linear sorption model can neither account for the observed rising flank nor for the amplitude. Thus, we can state that this linear approach is not suitable to describe the observed process properly. For this reason, we investigate in the following whether a nonlinear

	$K_d\ [cm^3 g^{-1}]$	$D\ [cm^2 d^{-1}]$	$c_0\ [mg\, l^{-1}]$
Soil A-I	0.162 ± 0.002	2.086 ± 0.102	8.26 ± 0.04

Table 5.15: Estimates of K_d, D and c_0 and 95 $\%$ Confidence Intervals

	K_d-D	K_d-c_0	D-c_0
Soil A-I	0.67	-0.84	-0.44

Table 5.16: Correlation Matrix

model delivers better results. Due to the fact that the Freundlich approach is not defined for $c_r(t, z) = 0$ and a Freundlich exponent $N < 1$, we choose the Langmuir approach which does not suffer from this limitation.

Nonlinear Langmuir Sorption

The model equation for the Langmuir sorption is the same as before (5.30), only the retardation factor R is substituted by

$$R = 1 + \frac{\rho}{\theta} \frac{K_l \beta}{(1 + K_l c_r)^2.} \tag{5.32}$$

In this approach, in addition to D and c_0 the Langmuir parameters K_l and β have to be identified. The estimation results are presented in Table 5.17 and 5.18.

	$D\ [cm^2 d^{-1}]$	$c_0\ [mg\, l^{-1}]$	$K_l\ [l\, mg^{-1}]$	$\beta\ [1]$
Soil A-I	1.582 ± 0.065	12.99 ± 0.05	2.155 ± 0.235	0.128 ± 0.010

Table 5.17: Estimates of D, c_0, K_l and β and 95 % Confidence Intervals

	D	c_0	K_l
β	0.92	-0.96	0.99
D	-	-0.78	0.90
c_0	-	-	-0.95

Table 5.18: Correlation Matrix

As Figure 5.20 shows, employing the nonlinear Langmuir approach, a by far better fit in comparison to the linear approach is achieved. The sum of squared residuals is reduced by approximately a factor of 3. Up till now, however, only the linear approach and the nonlinear Freundlich approach are available in the simulation tools used for registration studies. As these results exemplify, there are good reasons to also consider the nonlinear Langmuir approach when investigating the sorption behavior of new substances.

5.5 Mini-Lysimeter Study: Environmental Fate of S-Metolachlor and Its Main Metabolites

The aim of the following parameter identification problem is to determine the degradation, sorption and dispersion coefficients for the grass herbicide S-Metolachlor (2-chloro-6'-ethyl-N-(2-methoxy 1-methylethyl)-O-aceto toluidide) and its two major metabolites on the basis of mini-lysimeter leachate data. The transformation paths of interest are shown in Figure 5.21.

The experimental part of this study was conducted by Horn in his diploma thesis at the Institute of Geoecology, Technical University of Braunschweig and at the Novartis Crop Protection AG, Basel. Here the discussion is restricted to the experimental aspects required for modeling and parameter estimation. For more details about the experiments we refer to Horn's diploma thesis (1999 [62]).

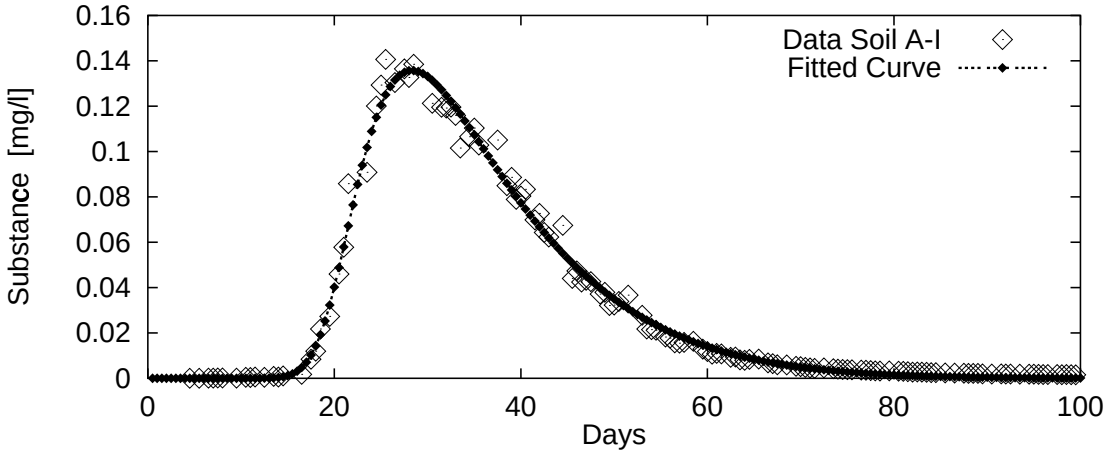

Figure 5.20: Nonlinear Langmuir approach: fitted and measured concentrations of test substance X in column A-I.

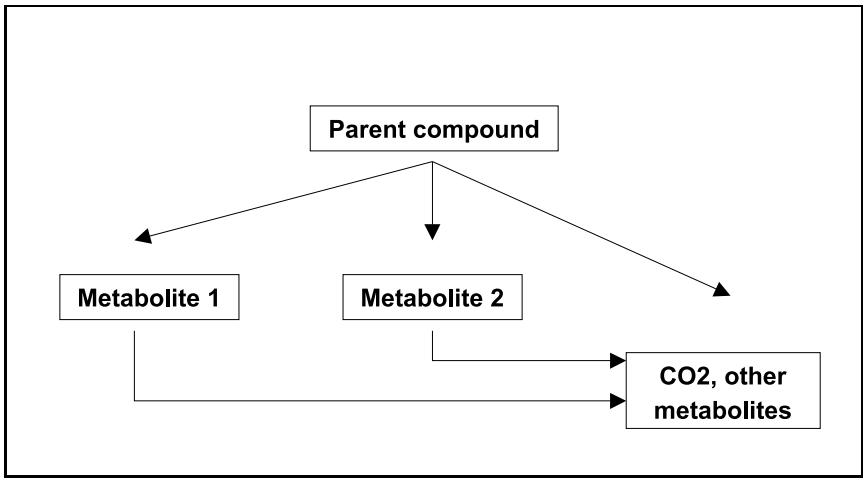

Figure 5.21: Scheme of transformation of S-Metolachlor in soil[1].

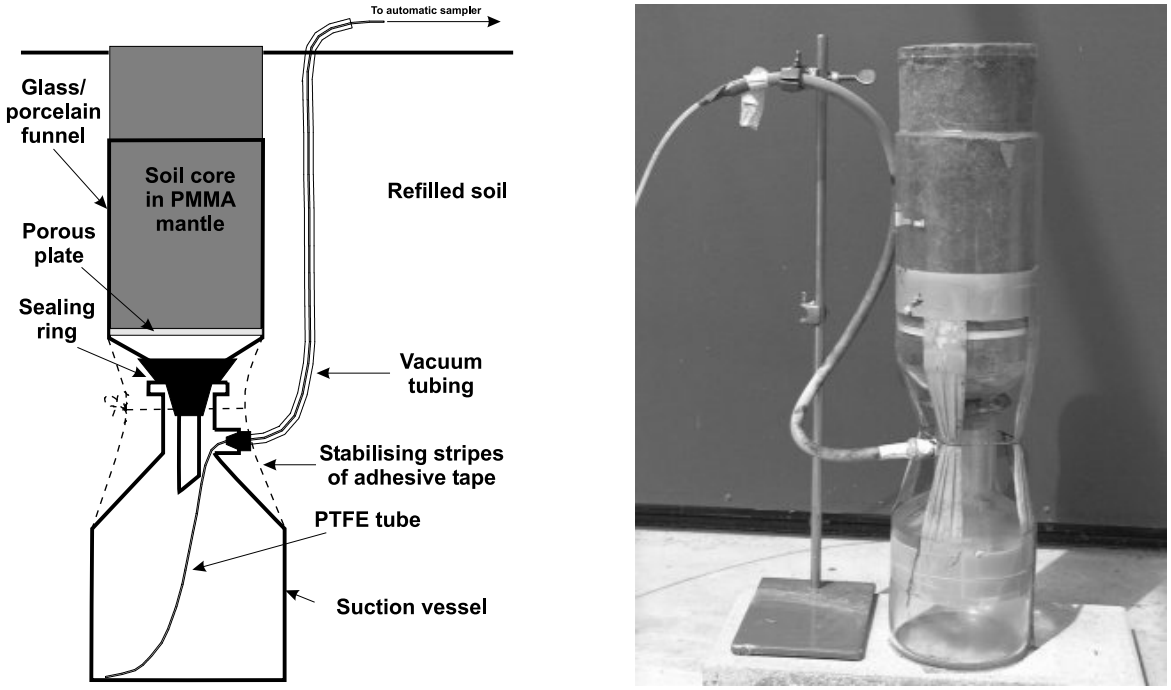

Figure 5.22: Scheme and photo of the glass mini-lysimeter[1].

5.5.1 Experimental Set-Up

For the study soil samples of a sandy loam are taken from an agricultural site in Switzerland. The soil chosen consists of 12% clay, 21% silt and 67% sand and has an organic carbon content of 1.2%. Glass mini-lysimeters of $0.3\ m$ length and $0.145\ m$ diameter, as shown in Figure 5.22, are installed on a lysimeter test site in Stein, Switzerland.

After a preconditioning phase of four weeks, soybean seeds are put into the soil column. Two days after sowing (17 May 1999) a solution of the ^{14}C-labeled parent compound S-Metolachlor is dripped on the bare soil with a syringe. During the first 24 hours the mini-lysimeters are protected from precipitation in order to enable an optimal efficiency of the substance. Then the mini-lysimeters are exposed to normal climatic conditions for 97 days (18 May to 23 August 1999). Meteorological data, e.g. air temperature, irrigation and precipitation volumes, are daily recorded at the testing facility and are summarized in Figure 5.23. The leachates are collected in intervals of 14 days and are analyzed for their radioactivity amount. The amount of S-Metolachlor and its metabolites are quantified by thin layer chromatography (TLC).

In addition, a mini-lysimeter with a nominal suction of $-90\ hPa$ at the lower boundary is installed to study the water transport. TDR is used to monitor the volumetric water contents. TDR probes are inserted in $5\ cm$, $15\ cm$, and $25\ cm$ depth and data is recorded hourly during the experiment.

[1]Figures were provided by Horn

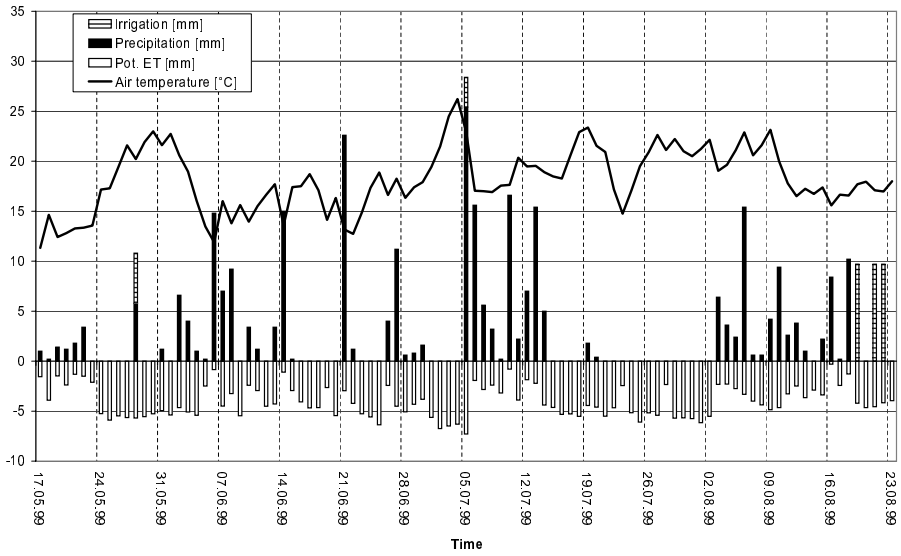

Figure 5.23: Meteorological data from the testing $\mathrm{facility}$[1].

5.5.2 Modeling

Water Transport

The water transport is modeled by the Richards equation in θ (5.17)-(5.20) as described for the field experiment in Section 5.3. Again the idea is to employ the TDR measurements at $5\ cm$ and $25\ cm$ depth as upper and lower boundary conditions and to estimate the unknown van Genuchten parameters n, α and K_s by fitting the TDR data in $15\ cm$ depth. This time, however, this approach fails. In contrast to the TDR data in Section 5.3, Figure 5.8, the mini-lysimeter studied here shows a higher water content at $15\ cm$ than at $5\ cm$ and $25\ cm$. There are several possibilities to explain this unusual behavior. It may be caused by the suction installed at the lower boundary, but also by the influence of preferential flow processes or by inhomogeneities of the soil core. Studies show that in this case it is not possible to reproduce the observed behavior by the Richards equation.

As a reliable description of the flow of water is necessary to meaningfully simulate the transport and degradation behavior of S-Metolachlor and its main metabolites, the experimenters propose to approximate the required quantities, i.e. the volumetric water content $\theta(t, z)$ and the water flux density $q(t, z)$ as follows. First of all, it is assumed that the volumetric water contents θ as well as the water flux densities q are constant across the soil profile. Weekly averaged volumetric water contents are obtained from TDR data. Water flux densities are also adjusted in weekly intervals. Their averaged values are calculated as the mean of average drainage flux densities and average precipitation/irrigation flux densities within the time intervals considered.

Solute Transport

After preliminary studies the following model consisting of three PDEs is set up to describe the environmental fate of the parent compound S-Metolachlor, P, and its main metabolites, M_1 and M_2, for the transformation paths given in Figure 5.21:

$$R_P \frac{\partial c_P}{\partial t} = \frac{D_s}{\theta}\frac{\partial^2 c_P}{\partial z^2} - \frac{q}{\theta}\frac{\partial c_P}{\partial z} - k_1 R_P c_P - k_2 R_P c_P - k_r R_P c_P \tag{5.33}$$

$$R_{M_1} \frac{\partial c_{M_1}}{\partial t} = \frac{D_s}{\theta}\frac{\partial^2 c_{M_1}}{\partial z^2} - \frac{q}{\theta}\frac{\partial c_{M_1}}{\partial z} + f_1 k_1 R_P c_P - k_{el1} R_{M_1} c_{M_1} \tag{5.34}$$

$$R_{M_2} \frac{\partial c_{M_2}}{\partial t} = \frac{D_s}{\theta}\frac{\partial^2 c_{M_2}}{\partial z^2} - \frac{q}{\theta}\frac{\partial c_{M_2}}{\partial z} + f_2 k_2 R_P c_P - k_{el2} R_{M_2} c_{M_2}, \tag{5.35}$$

with retardation factors describing linear equilibrium sorption

$$R_l = 1 + \frac{\rho}{\theta} K_{d,l}, \qquad l = P, M_1, M_2. \tag{5.36}$$

In all cases the degradation process is considered to be first order linear and to take place in both the dissolved and the adsorbed phase. Assuming that dispersion affects the transport of all three species to the same extent, only one apparent dispersion coefficient is defined

$$D_s = \lambda |q| + a \exp(b\theta) D_w. \tag{5.37}$$

For the sake of clarity, we summarize all variables introduced above and explain them in the following table:

c_P, c_{M_1}, c_{M_2}	Resident solute concentration of the parent compound, of metabolite 1, and of metabolite 2 $[g\,m^{-3}]$
$K_{d,P}$, K_{d,M_1}, K_{D,M_2}	Linear sorption coefficient of the parent compound, of metabolite 1, and of metabolite 2 $[m^3 g^{-1}]$
k_1, k_2, k_r	Degradation rate of the parent compound to metabolite 1, to metabolite 2 and to other metabolites and CO_2 $[d^{-1}]$
k_{el1}, k_{el2}	Degradation rate of metabolite 1 and metabolite 2 to other metabolites, and CO_2 $[d^{-1}]$
f_1, f_2, f_3	Correction factors for the mass balance of the transformation processes $[1]$
θ	Volumetric water content $[1]$
q	Volumetric water flux density $[m\,d^{-1}]$
$v = q/\theta$	Pore water velocity $[m\,d^{-1}]$
ρ	Bulk density $[g\,m^{-3}]$
D_s	Apparent dispersion coefficient $[m^2\,d^{-1}]$
λ	Dispersion length $[m]$
D_w	Molecular dispersion coefficient in water ($4.3 \cdot 10^{-5}$ $[m^2 d^{-1}]$)
a	Soil parameter, for sandy loam (0.005 $[1]$)
b	Soil parameter, for sandy loam $[1]$

As in the mini-lysimeter study discussed in Section 5.4, we assume the substance, in this case the parent, to be incorporated in the upper $5\ cm$ of the soil defining the following initial condition

$$c_P(0,z) = \begin{cases} c_0 & z \leq 0.05 \\ 0 & z > 0.05 \end{cases} \qquad z > 0. \tag{5.38}$$

At the beginning of the experiment the mini-lysimeter is assumed to be free of metabolite 1 and metabolite 2

$$c_{M_1}(0,z) = c_{M_2}(0,z) = 0.0 \qquad z > 0. \tag{5.39}$$

Afterwards, no more substance is added. Therefore, zero flux conditions are formulated at the upper boundary

$$\left(vc_l(t,z) - \frac{D_s}{\theta}\frac{\partial c_l(t,z)}{\partial z}\right)\Big|_{z=0} = 0 \qquad l = P, M_1, M_2 \qquad t > 0. \tag{5.40}$$

The lower boundary of the mini-lysimeter is modeled by means of the semi-infinite approach of van Genuchten and Parker (1984 [131], 1984 [99]). For the computation a column of $L = 1.0\,m$ length with a Neumann condition at the lower boundary is used

$$\frac{\partial c_l(t,z)}{\partial z}\Big|_{z=L} = 0 \qquad l = P, M_1, M_2. \tag{5.41}$$

At the actual lower boundary of the mini-lysimeter, $L_e = 0.3\,m$, flux-averaged concentrations for the three species are obtained via

$$c_l^{flux}(t,z) = c_l(t,z) - \frac{D_s}{q}\frac{\partial c_l(t,z)}{\partial z}, \qquad l = P, M_1, M_2. \tag{5.42}$$

The leachate concentrations M_l^i for the parent P and the metabolites M_1 and M_2, sampled in a time interval $[t_{i-1}, t_i]$, are modeled by

$$M_l^i(t_i, L_e) = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} c_l^{flux}(\tau, L_e) d\tau, \qquad i = 1, \cdots, n \qquad l = P, M_1, M_2. \tag{5.43}$$

In order to simulate the transport and degradation behavior of the three species, values for all parameters used in the model (5.33)-(5.43) are required. Some of the parameters can be derived from literature and from (laboratory) experiments, while the others have to be determined by parameter estimation.

The linear sorption coefficient for the parent S-Metolachlor, which is known to be a rather immobile substance, is calculated from the organic carbon partition coefficient $K_{OC,P} = 2.46 \cdot 10^{-4}\,m^3 g^{-1}$ and the organic carbon content of the soil $C_{org} = 1.2\ [\%]$ by

$$K_{d,P} = K_{OC,P}\frac{C_{org}\,[\%]}{100} = 2.95 \cdot 10^{-6} \quad [m^3 g^{-1}]. \tag{5.44}$$

The mass balance correction factors of the transformation process are derived by

$$f_l = \frac{M_{mol,product}}{M_{mol,educt}} \qquad l = 1, 2, 3, \tag{5.45}$$

where $M_{mol,product}$ and $M_{mol,educt}$ denote the molecular weight of the product and the molecular weight of the educt, respectively. For the substances considered here the values $f_1 = 1.241$, $f_2 = 0.987$ and $f_3 = 0.795$ are obtained. The initial concentration of the parent, which we consider to be incorporated within the first $5cm$ of the soil, is computed from the actual amount of substance applied $m_p = 2.45 \cdot 10^{-3}\ [g\, m^{-3}]$ according to

$$c_0 = \frac{m_p}{V_w R_P} = 0.579 \quad [g\, m^{-3}]. \tag{5.46}$$

Here, $V_w = \theta V_s$ where $V_s = 8.25 \cdot 10^{-4}\ [m^3]$ is the volume of the soil cylinder considered and $\theta = 0.35\ [1]$ is the water content at time $t = 0$. Due to equilibrium sorption only a part of the substance applied is available in the dissolved phase. In order to obtain the concentration in the dissolved phase, the quantity m_p/V_w has to be divided by the retardation factor $R = 1 + \frac{\rho}{\theta} K_{d,P}$ with $\rho = 1.62 \cdot 10^6\ [g\, m^{-3}]$.

Unknown parameters are the linear sorption coefficients for the metabolites, K_{d,M_1} and K_{d,M_2}, the degradation rates of the parent to the metabolites M_1 and M_2 and to other metabolites and CO_2, k_1, k_2, k_3, the degradation rates of the metabolites M_1 and M_2 to other metabolites and CO_2, k_{el1} and k_{el2}, the dispersion length λ and the parameter b in the apparent dispersion coefficient D_s. In summary, 9 parameters have to be estimated on the basis of leachate data for the parent and the two metabolites M_1 and M_2.

5.5.3 Results

In a first step, we investigate the performance of the initial guesses for the unknown parameters as given in Table 5.19, which are collected from literature or are obtained from prior (laboratory) studies (Horn, private communication). In Figure 5.24 the simulation results for the parent and the two main metabolites using the initial guesses are compared with the measured leachate concentrations. While the simulated results for the parent are in accordance with the measurements, the ones for the metabolites differ by a least one order of magnitude from the measured leachate concentrations.

In particular, note that at the time point $t = 84\ [d]$ no measurement data is available. Due to a drought of two weeks and a missing irrigation (see Figure 5.23), not enough outflow is sampled to enable a reliable analysis. In order to guarantee measurement data for the last sampling point $t = 98\ [d]$ (23 August 1999), the mini-lysimeters are irrigated during the last week with up to $10\, mm$ water per day.

The parameter estimation problem for all 9 parameters, however, is singular, which is not surprising considering the measurement data available. In order to identify the parameters, the optimization problem is regularized by estimating only some of the parameters simultaneously while keeping the other parameters fixed. The final results are presented in Table 5.19. The

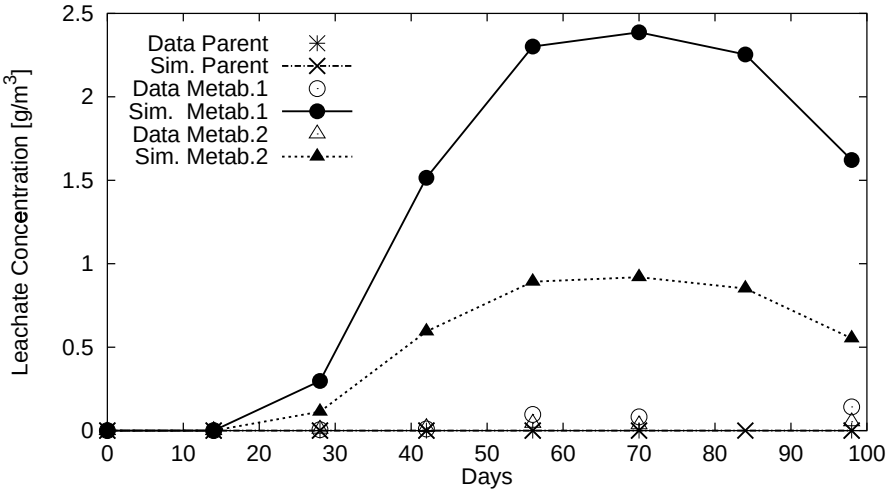

Figure 5.24: Simulation with initial parameter guesses.

	Initial Guesses	Estimates	Unit
K_{d,M_1}	$0.12 \cdot 10^{-6}$	$1.18 \cdot 10^{-6} \pm 0.176 \cdot 10^{-6}$	$[m^3 g^{-1}]$
K_{d,M_2}	$0.14 \cdot 10^{-6}$	$0.96 \cdot 10^{-6} \pm 0.168 \cdot 10^{-6}$	$[m^3 g^{-1}]$
k_r	0.0335	0.23 ± 0.06	$[d^{-1}]$
k_1	$1.1 \cdot 10^{-2}$	$0.89 \cdot 10^{-2} \pm 0.16 \cdot 10^{-2}$	$[d^{-1}]$
k_2	$7.85 \cdot 10^{-3}$	$6.05 \cdot 10^{-3} \pm 0.12 \cdot 10^{-2}$	$[d^{-1}]$
λ	0.054	0.144 ± 0.02	$[m]$
k_{el1}	$5.31 \cdot 10^{-3}$	$5.37 \cdot 10^{-3} \pm 0.64 \cdot 10^{-3}$	$[d^{-1}]$
k_{el2}	$1.32 \cdot 10^{-2}$	$1.34 \cdot 10^{-2} \pm 0.15 \cdot 10^{-2}$	$[d^{-1}]$
b	10.0	4.83 ± 4.89	$[1]$

Table 5.19: Initial Guesses, Estimates and 95 % Confidence Intervals

groups of parameters estimated together are separated by horizontal lines. Accordingly, the 95% confidence intervals only refer to the corresponding groups. Table 5.19 shows that the initial guesses for the degradation rates k_1, k_2, k_{el1} and k_{el2} are essentially confirmed by the estimates except for k_r which is increased by a factor of 7. The initial guesses for the linear sorption coefficients K_{d,M_1} and K_{d,M_2} are significantly reduced by optimization, while the estimate for the dispersion length λ is approximately three times the value of the initial guess. The value for the parameter b in the apparent dispersion coefficient, however, is ill-determined. Even though the estimation results partially differ considerably from the initial guesses, which are obtained from literature or laboratory experiments, all estimates lie within reasonable ranges (see Horn, 1999 [62]).

As Figure 5.25 demontrates, a satisfying match of simulated and measured leachate concentrations is obtained, bearing in mind the approximations made for the water contents and the

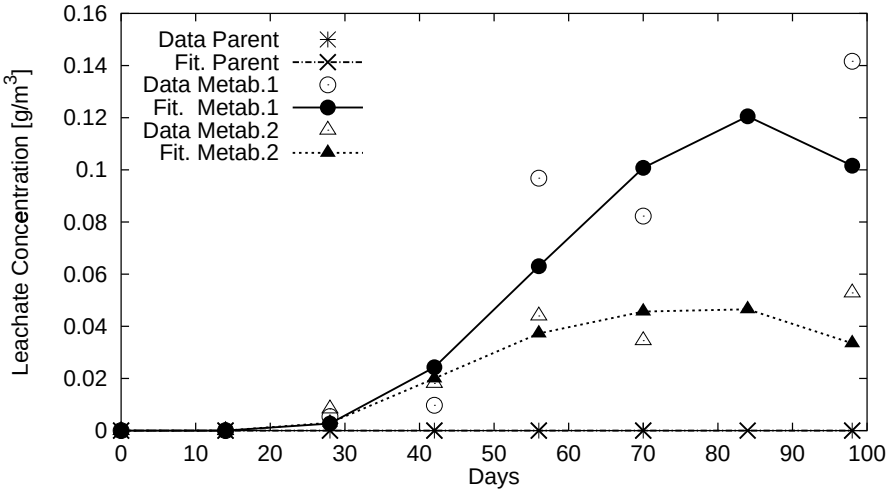

Figure 5.25: Simulation with estimated parameters.

water flux densities. The solution surfaces illustrating the transport and degradation behavior of the parent compound and the two metabolites are shown in the Figures 5.26, 5.27 and 5.28, respectively.

This example, however, demonstrates that the measured data is by far not optimal for parameter estimation - a case that often occurs in practice. Up till now, experiments have generally not been designed with the objective of parameter estimation. Commonly, the main focus of such studies is on investigating the transport and degradation behavior of substances under normal climatic conditions. The data obtained from these experiments is then re-used for parameter estimation leading frequently to ill-posed or singular problems.

The results discussed above reveal that parameter estimation requires designs that guarantee data which contains enough information in order to reliably determine the unknown parameters. In the next section we will outline how optimal experimental designs for this type of outflow experiments can be obtained and how they improve the quality of the estimation results.

5.6 Optimization of Experimental Conditions in Column Outflow Experiments

In this section an optimal experimental design problem for a hypothetical column outflow experiment is outlined to demonstrate the potential of ECOPLAN. The computed optimal experimental design shows that it is possible to reliably estimate both soil hydraulic and environmental parameters in one experiment solely on the basis of averaged leachate concentrations (Dieses et al., 2000 [42]).

In the following several scenarios are discussed. In the first scenario, the sampling scheme

[1]Different scale than for the metabolites M_1 and M_2

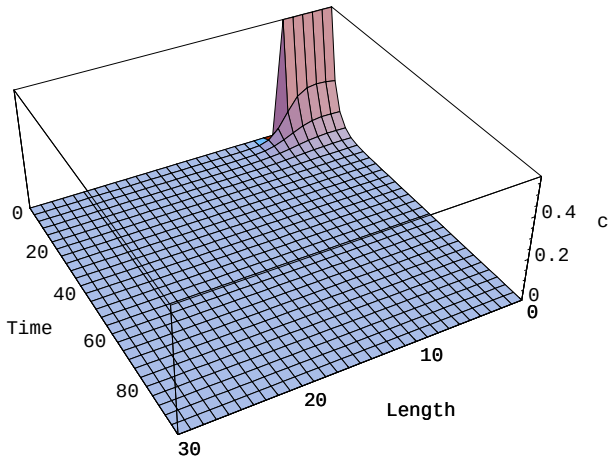

Figure 5.26: Resident solute concentration of the parent $\mathrm{compound}^1$.

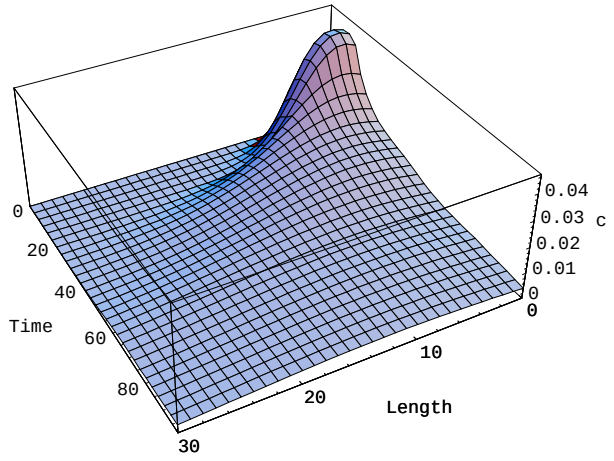

Figure 5.27: Resident solute concentration of metabolite 1.

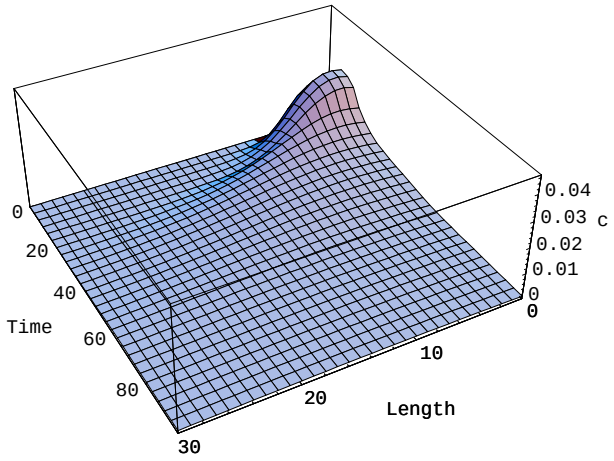

Figure 5.28: Resident solute concentration of metabolite 2.

is given and kept fixed. Here, the aim is to optimize the experimental (boundary) conditions, in particular the irrigation scheme and the substance concentration in the irrigation water. Moreover, the influence of different types of measurement errors with different standard deviations is investigated. In the second scenario, both the sampling scheme and the experimental conditions are optimized simultaneously. The last section is devoted to parameter sensitivities of the resulting optimal designs.

5.6.1 Model Equations

For both scenarios the system is described by the following coupled PDEs and initial and boundary conditions:

Water transport:

$$C(\psi_m)\frac{\partial \psi_m}{\partial t} = \frac{\partial}{\partial z}\left[K(\psi_m)\frac{\partial}{\partial z}(\psi_m - \bar{w}z)\right] \tag{5.47}$$

$$\begin{aligned} \psi_m(0,z) &= 670, \qquad z \geq 0 && (5.48)\\ -K(\psi_m(t,0))\frac{\partial}{\partial z}\big(\psi_m(t,0) - \bar{w}z\big) &= q(t,0) && (5.49)\\ \frac{\partial \psi_m(t,\infty)}{\partial z} &= 0, \qquad t \geq 0. && (5.50) \end{aligned}$$

Solute transport:

$$\frac{\partial(\theta c)}{\partial t} = \frac{\partial}{\partial z}\big(\theta D_h(\theta)\frac{\partial c}{\partial z}\big) - \frac{\partial(qc)}{\partial z} - \theta k c \tag{5.51}$$

$$\begin{aligned} c(0,z) &= 0, \qquad z \geq 0 && (5.52)\\ -D_h(\theta(t,0))\frac{\partial c(t,0)}{\partial z} + \frac{q(t,0)}{\theta(t,0)}c(t,0) &= \frac{q(t,0)}{\theta(t,0)}c_0(t,0) && (5.53)\\ \frac{\partial c(t,\infty)}{\partial z} &= 0, \qquad t \geq 0 && (5.54) \end{aligned}$$

Altogether 6 parameters are unknown: the van Genuchten parameters n, α and K_s, the degradation rate k, and the parameters b and D_m in the hydrodynamic dispersion term of the form:

$$D_h(\theta) = \frac{a e^{b\theta}}{\theta} + D_m. \tag{5.55}$$

The length of the soil column (L_e) is $31.5\,cm$. To approximate the infinite lower boundary condition a hypothetical column of $90\,cm$ length is computed.

	Original	Scaled	Design A	Design B
A-Criterion			3.88	0.75
n	1.2	1.0	$\pm$ 0.1604	$\pm$ 0.0800
α	0.0102	1.0	$\pm$ 1.5968	$\pm$ 0.8600
K_s	10.0	1.0	$\pm$ 3.3751	$\pm$ 1.6735
k	0.05	1.0	$\pm$ 0.0124	$\pm$ 0.0083
b	10.0	1.0	$\pm$ 3.0460	$\pm$ 0.9712
D_m	20.0	1.0	$\pm$ 0.2644	$\pm$ 0.1032

Table 5.20: Start Designs A and B: A-Criterion Values and Standard Deviations

5.6.2 Scenario 1: Optimization of Experimental Conditions

According to common measurement practice for outflow experiments, we set up the following scenario: Leachate is sampled and analyzed on a prescribed equidistant time grid. Two types of measurements are carried out. During the experimental period of 12 days leachate is sampled twice a day and substance concentrations in the leachates are analyzed. At the end of the experiment the column is sliced and resident concentrations in 4.5, 9.0, 13.5, 18.0, 22.5 and $27\,cm$ depth are determined. The standard deviation of the leachate concentration measurements is assumed to be $\sigma = 0.01$. For the profile resident concentration measurements a standard deviation of $\sigma = 0.1$ is used.

The experimental conditions in this scenario are optimized while keeping the sampling design fixed. Here, the two control functions are the water flux density $q(t, 0)$ (5.49) and the added substance concentration $c_0(t, 0)$ (5.53) at the upper boundary $z = 0$. Both control functions are parameterized by piecewise constant functions with values in the intervals [0, 0.6] and [50.0, 200.0], respectively. It is assumed that the controls may change once a day within the prescribed ranges.

Start designs

As the optimization problem considered is nonlinear, an initial guess has to be provided. In the context of our experimental design problem this implies that a start design for both parameterized control functions is needed.

Start designs are given in Figure 5.29 and 5.30. For example in start design A, the input water flux $q(t, 0)$ is reduced step by step while keeping the substance input concentration $c_0(t, 0)$ constant at a medium level. In start design B, we combine the maximum feasible water flux during the complete experimental period with a constant medium substance input concentration. Within the given constraints for the control functions the start designs A and B seem to be good choices.

Before presenting the optimal experimental design results for this problem, we discuss the quality of the parameter estimates one would obtain carrying out start design A or B. For the

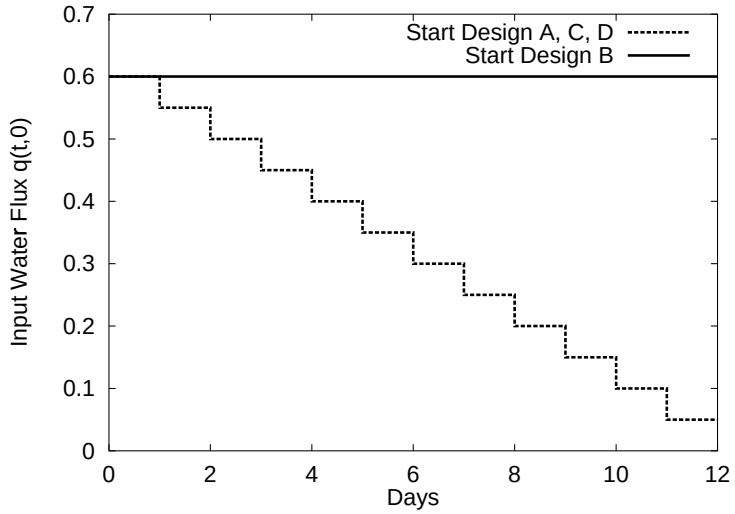

Figure 5.29: Piecewise constantly parameterized start designs A-D for input water flux $q(t, 0)$.

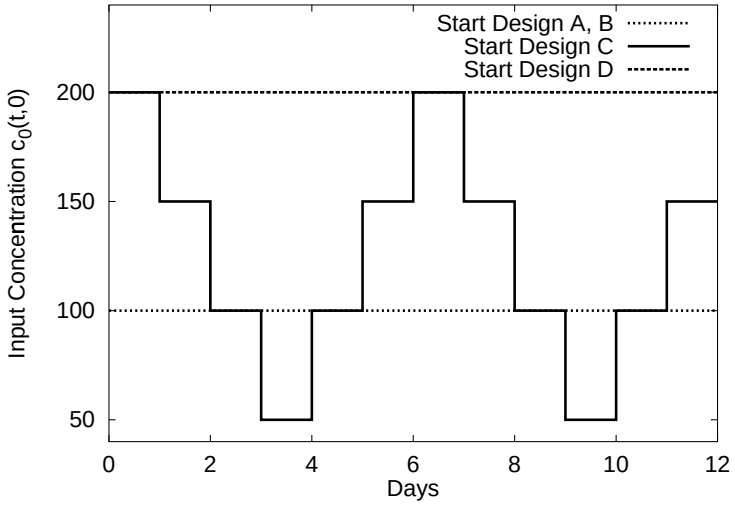

Figure 5.30: Piecewise constantly parameterized start designs A-D for input concentration $c_0(t, 0)$.

	α	K_s	k	b	D_m
n	-0.79	-0.99	0.80	-0.41	0.43
α	-	0.85	-0.29	-0.23	0.21
K_s	-	-	-0.74	0.32	-0.35
k	-	-	-	-0.84	0.87
b	-	-	-	-	-0.99

Table 5.21: Start Design A: Correlation Matrix

	α	K_s	k	b	D_m
n	-0.91	-0.99	-0.12	0.03	0.0 6
α	-	0.87	-0.27	-0.43	0.36
K_s	-	-	0.21	0.06	-0.15
k	-	-	-	0.97	-0.98
b	-	-	-	-	-0.99

Table 5.22: Start Design B: Correlation Matrix

calculations the original parameter values are scaled to 1.0 (Table 5.20). Even though both designs enable parameter estimates, some of the standard deviations are not acceptable.

For start design A with an A-criterion value of 3.88 the standard deviations for α, K_s and b are 1.5 respectively 3.0 times the parameter values (Table 5.20). Estimates of this accuracy are unsatisfactory for any practical purposes. The correlation matrix, Table 5.21, reveals that the estimation problem is nearly singular with two entries in the correlation matrix of modulus 0.99.

Start design B seems to be more promising with an A-criterion value of 0.75 being 5 times smaller than for design A. Nevertheless, the standard deviations for α, K_s and b are still too large (Table 5.20). Additionally, the correlation matrix, Table 5.22, contains five entries with modulus greater than 0.90 and two of them being 0.99.

Optimal experimental design

Starting the optimization from start design A, B, or one of the other start designs given in Figure 5.29 and 5.30, ECOPLAN delivers in all cases the same optimal experimental design with an A-criterion value of 0.047 which is a factor of 19 better than start design B. This optimal design delivers sufficiently small standard deviations for all parameter estimates (Table 5.23) with the same experimental effort.

The optimal designs for the input water flux $q(t, 0)$ and the substance concentration $c_0(t, 0)$ are shown in Figure 5.31 and 5.32. The water flux (Figure 5.31) starts with the maximum feasible rate and then the infiltration is stopped for 3 days. During the following 3 days a medium rate is chosen which is followed again by 4 days of maximum input. For the substance input concentration (Figure 5.32) it is found that a change between maximum, minimum and again maximum rates is optimal. It should be noted that - as a side effect - the correlations are reduced such that there occur no entries of modulus 0.99 any more (Table 5.24).

To summarize we can state that by mathematical optimization we obtain experimental conditions that allow for parameter estimates of significantly higher accuracy compared to straightforward designs.

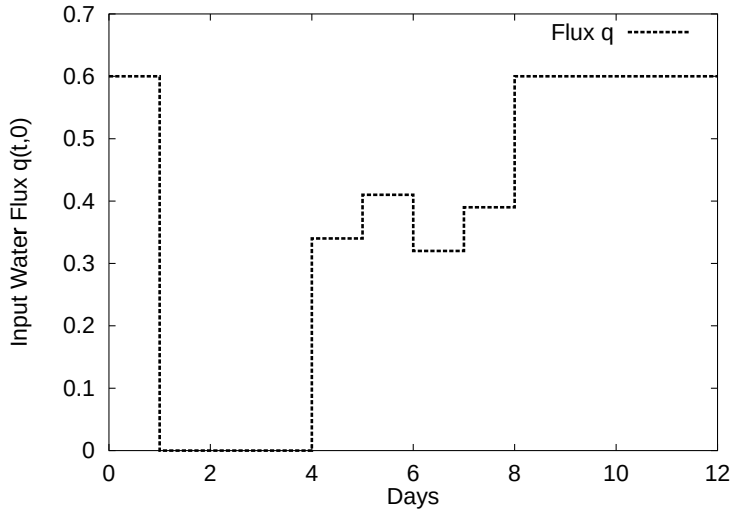

Figure 5.31: Scenario 1: Optimal design for input water flux $q(t, 0)$.

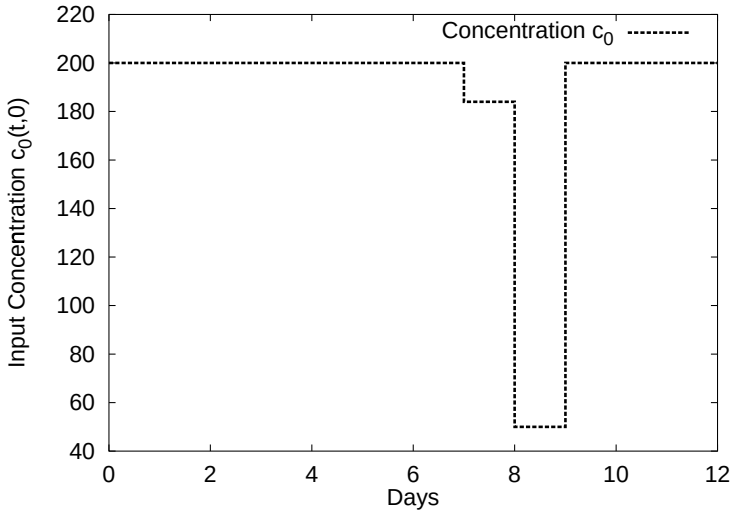

Figure 5.32: Scenario 1: Optimal design for input concentration $c_0(t, 0)$.

A-Criterion	0.047		
n	1.0	$\pm$	0.0191
α	1.0	$\pm$	0.2438
K_s	1.0	$\pm$	0.4348
k	1.0	$\pm$	0.0086
b	1.0	$\pm$	0.1852
D_m	1.0	$\pm$	0.0108

Table 5.23: Scenario 1 (Example 1): Optimal Design: A-Criterion Value and Standard Deviations

	α	K_s	k	b	D_m
n	-0.98	-0.93	0.19	0.77	-0 .87
α	-	0.98	-0.04	-0.63	0.76
K_s	-	-	-0.17	-0.49	0.64
k	-	-	-	0.71	-0.60
b	-	-	-	-	-0.98

Table 5.24: Scenario 1 (Example 1): Optimal Design: Correlation Matrix

	Original	Scaled	Exp. 1	Exp. 2	Exp. 3
A-Criterion			0.047	0.075	0.0038
n	1.2	1.0	$\pm$ 0.0191	$\pm$ 0.0283	$\pm$ 0.0047
α	0.0102	1.0	$\pm$ 0.2438	$\pm$ 0.2850	$\pm$ 0.0747
K_s	10.0	1.0	$\pm$ 0.4348	$\pm$ 0.4735	$\pm$ 0.1126
k	0.05	1.0	$\pm$ 0.00857	$\pm$ 0.0121	$\pm$ 0.00007
b	10.0	1.0	$\pm$ 0.1852	$\pm$ 0.3764	$\pm$ 0.0691
D_m	20.0	1.0	$\pm$ 0.0108	$\pm$ 0.0351	$\pm$ 0.0064

Table 5.25: Scenario 1 (Experiments 1-3): A-Criterion Values and Standard Deviations

Modified experiments

As we have seen it is possible to find a design by optimization that enables reliable parameter estimates on the basis of 24 leachate measurements and 6 profile measurements. In the following this set-up is referred to as Experiment 1.

However, the question may arise, whether the 6 profile observations obtained by slicing the column are essential or whether they contain in principle redundant information. How 'good' would an optimal design be without these profile data compared to the results of the previous section?

To answer these questions we compute the optimal experimental design for a second experiment, Experiment 2, that only takes the leachate measurements into account. As in Experiment 1, we assume a standard deviation of $\sigma = 0.01$ for these measurements. Using the same start designs for the control functions as before, ECOPLAN converges again in all cases to the same minimum with an A-criterion value of 0.075. But the A-criterion value has nearly doubled. Table 5.25 shows that important information, in particular for n, b and D_m, is lost when neglecting the profile data.

In addition, these results confirm well known difficulties with a precise estimation of K_s. Experiment 1 and 2 demonstrate that the profile data contains paramount information even though they are a factor of 10 less accurate. Suppose we could measure the profile concentrations with the same accuracy as the leachate concentrations ($\sigma = 0.01$), how would this affect the estimation results? The computation of the optimal experimental design for this set-up,

	Scaled	Scen. 1 (Exp. 1)	Scen. 2
A-Criterion		0.047	0.026
n	1.0	$\pm$ 0.0191	$\pm$ 0.0215
α	1.0	$\pm$ 0.2438	$\pm$ 0.203
K_s	1.0	$\pm$ 0.4348	$\pm$ 0.2872
k	1.0	$\pm$ 0.00857	$\pm$ 0.0050
b	1.0	$\pm$ 0.1852	$\pm$ 0.1876
D_m	1.0	$\pm$ 0.0108	$\pm$ 0.0217

Table 5.26: Scenario 1 and 2 (Example 1): A-Criterion Values and Standard Deviations

denoted by Experiment 3, reveals the significant impact of good profile measurements on the reliability of the parameter estimates (Table 5.25).

5.6.3 Scenario 2: Simultaneous Optimization of Sampling Scheme and Experimental Conditions

In accordance with common measurement practice leachates are sampled in Scenario 1 on a fixed equidistant time grid consisting of 2 measurements per day over a period of 12 days (altogether 24 measurements). But it seems most likely that a non-equidistant sampling grid optimized together with the experimental conditions could further improve the reliability of estimates.

Assuming that leachate sampling is generally possible every 6 hours, a feasible set of 48 measurement points Ξ is defined. Within this set 24 measurements, the same number as in Scenario 1, have to be chosen. For the two control functions $q(t,0)$ and $c_0(t,0)$ the same start designs as in Scenario 1, Figure 5.29 and 5.30, are used.

As expected, by the optimization of both the sampling scheme and the experimental conditions, again a considerable reduction of the A-criterion value from 0.047 in Scenario 1 (Example 1) to 0.026 can be achieved (Table 5.26). It should be noted that this improvement is obtained without increasing the number of measurements. The optimal designs for the control functions $q(t,0)$ and $c_0(t,0)$ are given in Figure 5.33 and 5.34.

Figure 5.35 shows the computed cumulated leachate concentrations and the optimized distribution of measurement points, i.e. the weights w, which are marked by vertical bars. Note, that the integrality constraints for the weights are satisfied. So no rounding strategy was necessary.

5.6.4 Parameter Sensitivity of Optimal Designs

At the first glance one drawback of designs optimized to the classical experimental design criteria (A-, D-, E-optimality) for nonlinear problems could be their dependence on the (unknown)

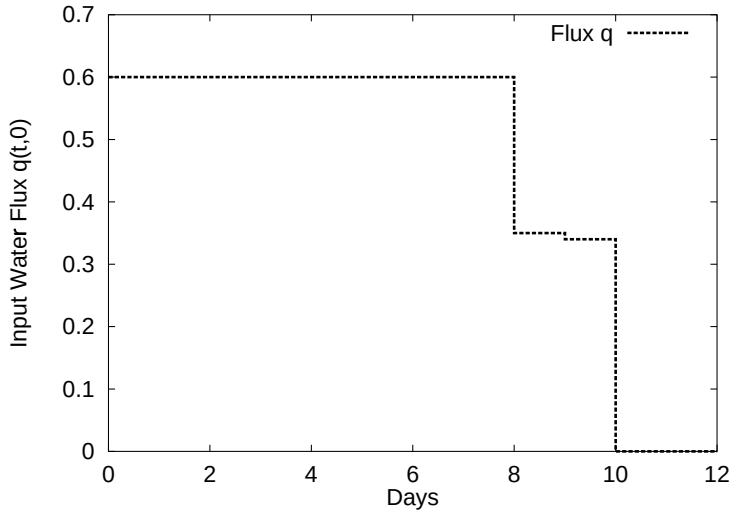

Figure 5.33: Scenario 2: Optimal design for input water flux $q(t, 0)$.

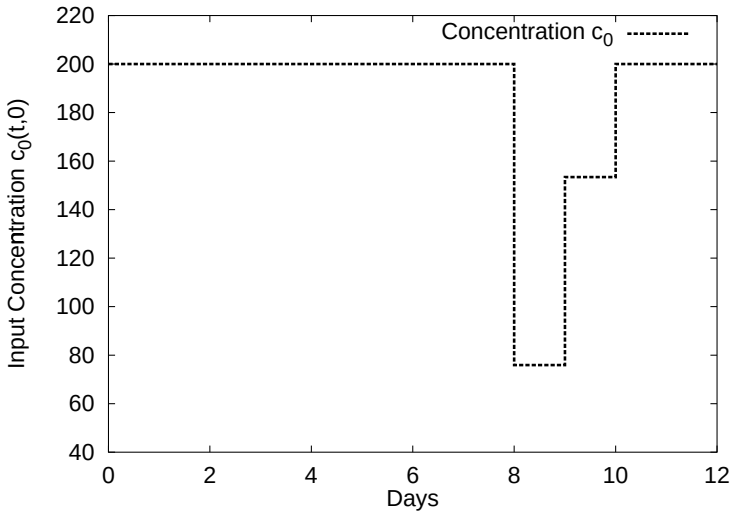

Figure 5.34: Scenario 2: Optimal design for input concentration $c_0(t, 0)$.

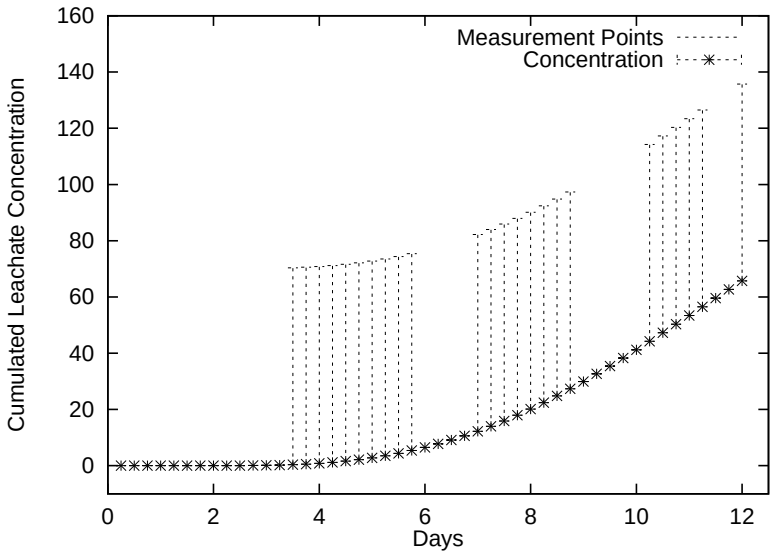

Figure 5.35: Scenario 2: Optimal sampling scheme and cumulated leachate concentrations. The 24 chosen measurement points out of the 48 feasible ones are marked by vertical bars.

parameters. As stated before, different assumptions are made for parameter estimation and optimal experimental design. While for parameter estimation controls and weights are fixed and parameters are optimized, this is vice versa for optimal experimental design where controls and weights are free variables with parameters being fixed.

However, the true values of these parameters are, of course, generally unknown. Thus, we are facing the paradoxical situation that the optimal design relies on the guesses of those parameters which in fact the experiment is being designed to identify. As these guesses might considerably differ from the true values, the optimized design should also perform well over a reasonable range of parameter values.

To study the effect how well suited a design that was optimal for a specific parameter set is for other values of parameters, A-criterion values for Example 1 (Scenario 1) are evaluated by successively shifting one parameter within a plausible range while keeping the other parameters fixed. For perturbations up to 100 $\%$ in both directions the A-criterion values are plotted in Figure 5.36. Here, the (scaled) parameter values 1.0 correspond to the values for which the design was optimized.

For some shifts the A-criterion value is reduced, but for others, for example when doubling the value of K_s, it is increased up to a factor of 5. At first, this does not seem very promising.

But comparing these results with the A-criterion values obtained using the start designs as given in Figure 5.29 and 5.30, the tremendous potential of optimal experimental design becomes obvious. Figure 5.37 reveals that even though the A-criterion values might be a factor of 5 or 10 worse for the disturbed parameters, they are drastically better than for any of the straightforward start designs. This indicates the robustness of the optimized experimental design against poor parameter guesses.

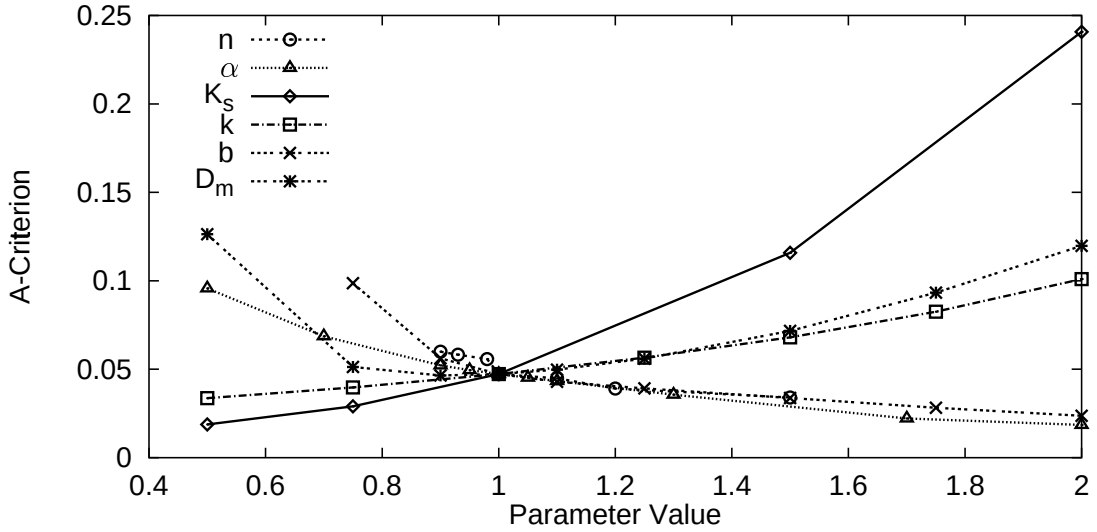

Figure 5.36: Sensitivity of A-criterion values for shifting one parameter while keeping the other parameters fixed. Parameters of value 1.0 correspond to the optimal design (Scenario 1, Example 1).

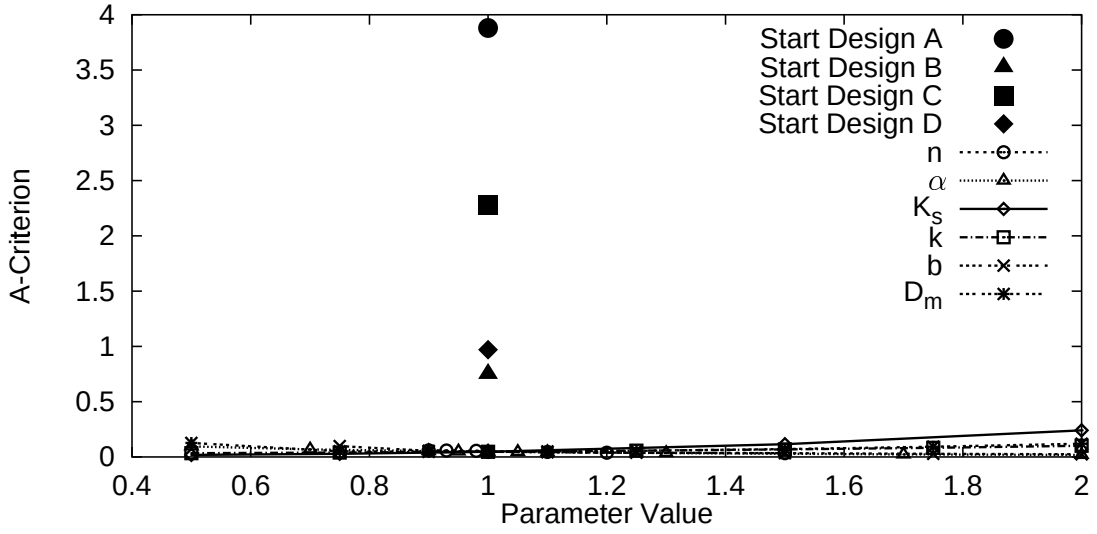

Figure 5.37: A-criterion values for start designs, optimal design and disturbed optimal designs (Scenario 1, Example 1). In contrast to Figure 5.36 a different scaling is used in order to include also the A-criterion values for the start designs. This demonstrates the potential of optimal experimental design and its robustness against poor initial parameter guesses.

Conclusions and Outlook

In this last chapter, we a give a brief summary of the developed methods for parameter estimation and optimal experimental design with regard to their performance in application problems. Finally, we discuss possible extensions and developments of our methods as well as further areas of application.

Summary and Conclusions

In this work we have presented efficient and reliable methods for both parameter estimation and optimal experimental design in water flow and reactive solute transport processes in soils described by instationary partial differential equations coupled with ordinary differential equations. In particular, the tools ECOFIT and ECOPLAN have been developed supporting the inverse modeling approach for the purpose of registration. The regulatory framework of European authorities has been taken into account as well.

ECOFIT has so far been the only tool which provides, even for highly resolved spatial grids, a reliable solution for inverse modeling in water flow and reactive solute transport processes within a range of 1-2 hours CPU time. This reduction of the computational effort for the problem class considered was essentially achieved by means of the following strategies:

- **A reduced Generalized Gauss-Newton method (Chapter 2)**
  We exploit the fact that the systems under consideration have only few degrees of freedom, namely as many as the model has unknown parameters. The explicit computation and storage of the Jacobian of the nonlinear constrained least-squares problem is avoided. Instead, the linear systems are simultaneously evaluated and decomposed using directional derivatives. This reduces the computational effort for the generation of derivatives essentially to the one required by a single shooting method while maintaining the advantages of the multiple shooting method.

- **A modified Newton method for the computation of directional derivatives (Chapter 3)**
  As a result of the reduced Generalized Gauss-Newton method only few directional derivatives have to be computed. Thus, instead of solving the linear systems arising from the computation of the corresponding variational differential equation directly, a modified

Newton method is used to avoid the expensive computation of the right hand side f with respect to the states y (f_y) in every BDF-step of the integrator.

- **A specially tailored strategy for the computation of f_y (Chapter 3)**
  We exploit the fact that as a result of the use of fixed spatial grids, which arise from the discretization of the PDEs in space, the sparsity pattern of f_y remains unchanged in the course of the reduced Generalized Gauss-Newton method. Applying matrix compression techniques in the framework of automatic differentiation, the effort to compute f_y is reduced to the same complexity order as the evaluation of the right hand side f.

The robustness and reliability of ECOFIT has been tested and demonstrated by means of hypothetical column experiments using a controlled scenario (Section 5.1 and 5.2). Its use for estimating hydraulic and environmental fate parameters from field (Section 5.3) and mini-lysimeter data (Section 5.4 and 5.5) has proven ECOFIT to be a powerful tool for inverse modeling.

Secondly, we have developed ECOPLAN (Chapter 4), which is the first tool for optimal experimental design of column and lysimeter experiments that allows to simultaneously optimize both

- the experimental conditions, such as boundary conditions describing irrigation and application schemes, and
- the sampling schemes, such as the allocation of measurement points in time and space.

In this work, ECOPLAN has been used to optimize the experimental design of a typical column outflow experiment (Section 5.6). The different features have been demonstrated. Recently, we have determined an optimal experimental design for soil column experiments by ECOPLAN which will be carried out by the group “Terrestrial Systems” of Kurt Roth at the Institute of Environmental Physics, University of Heidelberg.

Several institutes as well regulatory agencies have expressed their interest in ECOFIT and ECOPLAN. ECOFIT has already been used for inverse modeling studies by BASF Agricultural Center Limburgerhof .

Outlook and future work

In the following, we outline further promising directions of research for parameter estimation and optimal experimental design building on ECOFIT and ECOPLAN:

- Incorporation of adaptive spatial discretization methods that are suitable for parameter estimation in the context of the reduced Generalized Gauss-Newton method.
- Extension of the developed methods for parameter estimation and optimal experimental design to two and three dimensions, e.g. in order to study heterogeneous soils.

- Development of methods for model discrimination in transport processes described by instationary partial differential equations.
- Online optimization of experimental designs.
- Extension of the methods for the treatment of parameter estimation and optimal experimental design problems over random parameter fields.

In this work the primary focus has been upon water flow and reactive solute transport processes in soils considering in particular column, mini-lysimeter and field experiments. In general, the developed methods can be applied to a by far broader class of transport processes in porous media, as they arise, for example, in chemical engineering.

Appendix A

Spatial Discretization Routines

A.1 Weighting Coefficients for DSS004 and DSS020

The routine DSS004 (Schiesser, 1991 [107]), based on an equidistant centered difference scheme of order four, is appropriate for parabolic PDEs. The respective weighting coefficients can be derived by linear combinations of Taylor series expansions (Fornberg, 1988 [51]):

$$
\begin{array}{rl}
du_1/dz & =1/(24\Delta z)(-50\ u_1\ +96\ u_2\ -72\ u_3\ +32\ u_4\ -6\ u_5\ )+\mathcal{O}(\Delta z^4)\\
du_2/dz & =1/(24\Delta z)(-\ 6\ u_1\ -20\ u_2\ +36\ u_3\ -12\ u_4\ +2\ u_5\ )+\mathcal{O}(\Delta z^4)\\
du_i/dz & =1/(24\Delta z)(\ \ 2\ u_{i-2}-16\,u_{i-1}+0\ \ u_i\ +16\,u_{i+1}-2\,u_{i+2})+\mathcal{O}(\Delta z^4)\\
du_{N-1}/dz & =1/(24\Delta z)(-\ 2\,u_{N-4}+12u_{N-3}-36u_{N-2}+20u_{N-1}+\ 6\ u_N\ )+\mathcal{O}(\Delta z^4)\\
du_N/dz & =1/(24\Delta z)(\ \ 6\,u_{N-4}-32u_{N-3}+72u_{N-2}-96u_{N-1}+50\,u_N\ )+\mathcal{O}(\Delta z^4)
\end{array}
$$

Table A.1: DSS04: Weighting Coefficients for u_1, u_2, u_i $(i = 3, \ldots, N-2)$, u_{N-1} and u_N.

Note, that in order to avoid fictitious points for the grid points $i = 1, 2, N-1, N$ non-centered approximations are used. The derivation of weighting coefficients for unequally spaced grids is usually based on Lagrange interpolation polynomials.

The routine DSS020 (Schiesser, 1991 [107]) which takes into account the direction of flow is suitable for convective systems modeled by first-order hyperbolic PDEs. The *five point biased upwind* finite difference scheme of order four derived by Carver and Hinds (1978 [30]) combines both centered and biased upwind approximations to reduce numerical oscillation and numerical diffusion. The corresponding weighting coefficients are given in Table A.2.

A.2 Comparison of Numerical and Analytical Solutions

In order to check the accuracy achieved by using the spatial discretization routines DSS004 and DSS020 numerical simulation results are compared with analytical solutions evaluated by Mathematica [140]. Similar to Stock (1995 [122]), the following convection-dispersion

$$
\begin{array}{l}
du_1/dz = 1/(12\Delta z)(-25\ u_1 + 48\ u_2 - 36\ u_3 + 16\ u_4 - 3\ u_5\ ) + \mathcal{O}(\Delta z^4) \\
du_2/dz = 1/(12\Delta z)(-3\ u_1 - 10\ u_2 + 18\ u_3 - 6\ u_4 + 1\ u_5\ ) + \mathcal{O}(\Delta z^4) \\
du_3/dz = 1/(12\Delta z)(\ \ 1\ u_1 - 8\ u_2 + 0\ u_3 + 8\ u_4 - 1\ u_5\ ) + \mathcal{O}(\Delta z^4) \\
du_i/dz = 1/(12\Delta z)(-1\ u_{i-3} + 6\ u_{i-2} - 18\, u_{i-1} + 10\ u_i + 3\, u_{i+1}) + \mathcal{O}(\Delta z^4) \\
du_N/dz = 1/(12\Delta z)(\ \ 3\, u_{N-4} - 16 u_{N-3} + 36 u_{N-2} - 48 u_{N-1} + 25\, u_N\ ) + \mathcal{O}(\Delta z^4)
\end{array}
$$

Table A.2: DSS020: Weighting Coefficients for u_1, u_2, u_3, u_i $(i = 4, \ldots, N-1)$ and u_N.

equation for which an analytical solution (van Genuchten and Alves, 1982 [130]) is available is used as a reference problem:

$$R\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial z^2} - v\frac{\partial c}{\partial z} \tag{A.1}$$

$$
\begin{aligned}
c(z,0) &= 0 \\
-D\frac{\partial c(0,t)}{\partial z} + vc(0,t) &= \begin{cases} vc_0 & 0 < t \le t_0 \\ 0 & t > t_0 \end{cases} \\
\frac{\partial c(\infty,t)}{\partial z} &= 0.
\end{aligned} \tag{A.2}
$$

The associated analytical solution is of the form:

$$c(z,t) = \begin{cases} c_i + (c_0 - c_i)A(z,t) & 0 < t \le t_0 \\ c_i + (c_0 - c_i)A(z,t) - c_0 A(z, t - t_0) & t > t_0, \end{cases} \tag{A.3}$$

where

$$
\begin{aligned}
A(z,t) = {} & \frac{1}{2}\mathrm{erfc}\Big[\frac{Rz - vt}{2\sqrt{DRt}}\Big] + \sqrt{\frac{v^2 t}{\pi DR}} \cdot \exp\Big[-\frac{(Rz - vt)^2}{4DRt}\Big] \\
& - \frac{1}{2}\Big(1 + \frac{vz}{D} + \frac{v^2 t}{DR}\Big)\exp\Big(\frac{vz}{D}\Big)\mathrm{erfc}\Big[\frac{Rz + vt}{2\sqrt{DRt}}\Big].
\end{aligned} \tag{A.4}
$$

The simulation is carried out for the values $R = 1.0$, $v = 1.0$, $D = 1.0$, $t_0 = 0.5$, $c_0 = 1$ and $c_i = 0$. In Figure A.1 concentration profiles in depth are given for the time points $t = 2.5$, 5, 10, 15 and 20. Comparing the numerical and the analytical solutions, Figure A.2, reveals a good agreement for the use of the routines DSS004 and DSS020 based on 101 spatial nodes. Plotted is the difference between the analytical and the numerical solution with respect to depth z for different points in time. Note, that different scalings for the concentrations c is used.

In accordance to the theory, by refining the grid, i.e. by increasing the number of spatial nodes, the spatial discretization error is reduced. For example, using a grid with 401 nodes, the maximal absolute error becomes less than 6.0×10^{-7}.

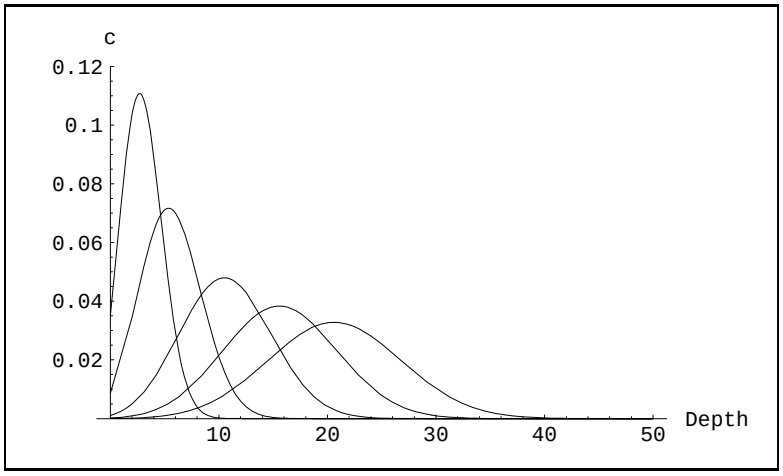

Figure A.1: Concentration profiles in depth at $t = 2.5$, 5, 10, 15 and 20.

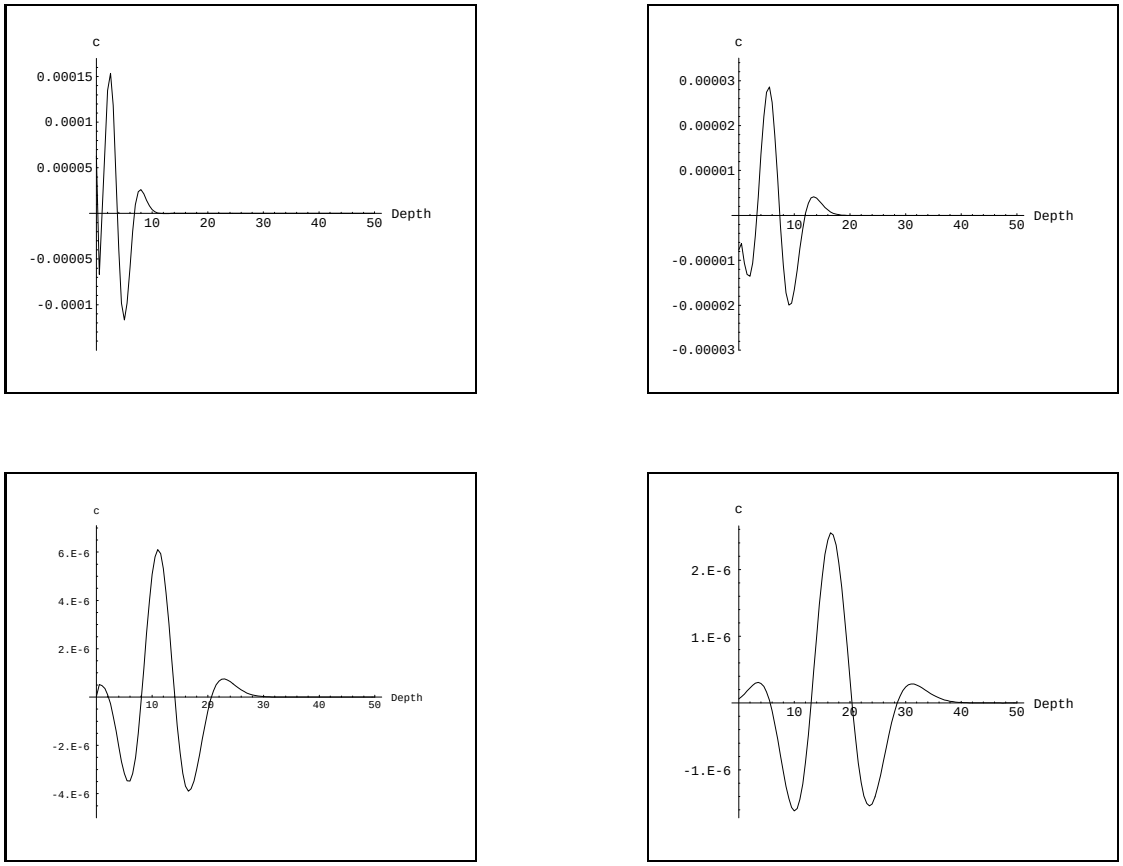

Figure A.2: Difference function between the analytical and the numerical solution for 101 nodes for $t = 2.5,\ 5,\ 10,\ 15$. Different scalings for c are used.

Bibliography

[1] K. C. Abbaspour, M. Th. van Genuchten, R. Schulin, and E. Schläppi. A sequential uncertainty domain inverse procedure for estimating subsurface flow and transport parameters. *Water Resour. Res.*, 33(8):1879–1892, 1997.

[2] K. Aden. *In-Situ Prozeß- und Parameteridentifikation für den Abbau und Transport von Metazachlor im Boden*. PhD thesis, TU Braunschweig, in preparation.

[3] K. Aden, O. Richter, and B. Gottesbüren. Parameter estimation using pesticide degradation experiments under varying soil temperatures and soil water contents. In A. A. M. Del Re, C. Brown, E. Capri, G. Errera, S. P. Evans, and M. Trevisan, editors, *Human and Environmental Exposure to Xenobiotics*, Proceedings of the XI Symposium Pesticide Chemistry, Sept 11-15 1999, Cremona, Italy, pages 89–99. La Goliardica Pavese, 1999.

[4] M. Arning. *Lösung des Inversproblems von partiellen Differentialgleichungen beim Wassertransport im Boden*. PhD thesis, TU Braunschweig, 1994.

[5] A. Atkinson and A. Donev. *Optimum Experimental Designs*. Oxford Statistical Science Series Number 8. Oxford University Press, Oxford, 1992.

[6] B. M. Averick, J. J. Moré, C. H. Bischof, A. Carlé, and A. Griewank. Computing large sparse Jacobian matrices using automatic differentiation. *SIAM J. Sci. Comput.*, 15(2):285–294, 1994.

[7] Y. Bard. *Nonlinear Parameter Estimation*. Academic Press, Inc., San Diego, 1974.

[8] I. Bauer. *Numerische Verfahren zur Lösung von Anfangswertaufgaben und zur Generierung von ersten und zweiten Ableitungen mit Anwendungen bei Optimierungsaufgaben in Chemie und Verfahrenstechnik*. PhD thesis, University of Heidelberg, 2000.

[9] I. Bauer, H. G. Bock, S. Körkel, and J. P. Schlöder. Numerical methods for initial value problems and derivative generation for DAE models with application to optimum experimental design of chemical processes. In F. Keil, W. Mackens, H. Voss, and J. Werther, editors, *Scientific Computing in Chemical Engineering II*, volume 2: Simulation, Image Processing, Optimization, and Control, pages 282–289. Springer-Verlag, Berlin, Heidelberg, 1999.

[10] I. Bauer, H. G. Bock, S. Körkel, and J. P. Schlöder. Numerical methods for optimum experimental design in DAE systems. *J. Comput. Appl. Math.*, 120(1-2):1–15, 2000.

[11] I. Bauer, H. G. Bock, and J. P. Schlöder. DAESOL – a BDF-code for the numerical solution of differential algebraic equations. Internal report, IWR, SFB 359, University of Heidelberg, 1999.

[12] M. Berzins, P. J. Capon, and P. K. Jimack. On spatial adaptivity and interpolation when using the method of lines. *Appl. Numer. Math.*, 26:117–133, 1998.

[13] M. Berzins and J. M. Ware. Solving convection and convection-reaction problems using the method of lines. *Appl. Numer. Math.*, 20:83–99, 1996.

[14] S. Beulke. *Untersuchung und mathematische Beschreibung des Abbaus von Herbiziden im Boden in Abhängigkeit von Wirkstoffverfügbarkeit, mikrobieller Biomasse und Aktivität*. Agrarwissenschaft. Shaker Verlag, 1998.

[15] C. H. Bischof, A. Carle, G. Corliss, A. Griewank, and P. Hovland. ADIFOR Generating derivative codes from Fortran programs. *Scientific Programming*, 1:11–29, 1992.

[16] C. H. Bischof, A. Carle, P. M. Khademi, and A. Mauer. The ADIFOR 2.0 System for the Automatic Differentiation of Fortran 77 Programs. CRPC-TR94491, Center for Research on Parallel Computation, Rice University, 1994.

[17] C. H. Bischof, A. Carle, P. M. Khademi, A. Mauer, and P. Hovland. ADIFOR 2.0 User's Guide. Technical Memorandum No. 192, Mathematics and Computer Science Division, 1998.

[18] C. H. Bischof, A. Carle, P. M. Khademi, and A. Maurer. ADIFOR 2.0: Automatic differentiation of Fortran 77 programs. *IEEE Computational Science & Engineering*, 3(3):18–32, 1996.

[19] H. G. Bock. Numerical treatment of inverse problems in chemical reaction kinetics. In K. H. Ebert, P. Deuflhard, and W. Jäger, editors, *Modelling of Chemical Reaction Systems*, volume 18 of *Springer Series in Chemical Physics 18*, pages 102–125, Heidelberg, 1981.

[20] H. G. Bock. Recent advances in parameter identification techniques for ODE. In P. Deuflhard and E. Hairer, editors, *Numerical Treatment of Inverse Problems in Differential and Integral Equations*, volume 2 of *Progress in Scientific Computing*, pages 95–121, Birkhäuser, Boston, 1983.

[21] H. G. Bock. Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen. *Bonner Mathematische Schriften 183*, 1987.

[22] H. G. Bock, E. Kostina, and J. P. Schlöder. On the role of natural level functions to achieve global convergence for damped Newton methods. In M. J. D. Powell, editor, *Proc. 19 IFIP TC 7 Conf. on System Modelling and Optimization*. Kluwer Academic Publishers, in press (2000).

[23] H. G. Bock, J. P. Schlöder, and V. H. Schulz. Numerik großer Differentiell-Algebraischer Gleichungen. Simulation und Optimierung. In H. Schuler, editor, *Prozeßsimulation*, pages 35–80. Verlag Chemie, Weinheim, 1995.

[24] J. J. T. I. Boesten, R. L. Jones, M. Businelli, A.-B. Delmas, B. Gottesbüren, K. Hanze, T. Jarvis, M. Klein, A. M. A. van der Linden, W.-M. Maier, S. Rekolainen, H. Resseler, M. Styczen, K. Travis, and M. Vanclooster. The development of FOCUS scenarios for assessing pesticide leaching to groundwater in EU registration. In A. A. M. Del Re, C. Brown, E. Capri, G. Errera, S. P. Evans, and M. Trevisan, editors, *Human and Environmental Exposure to Xenobiotics*, Proceedings of the XI Symposium Pesticide Chemistry, Sept 11-15 1999, Cremona, Italy, pages 795–808. La Goliardica Pavese, 1999.

[25] G. E. P. Box and H. L. Lucas. Design of experiments in non-linear situations. *Biometrika*, 46:77–90, 1959.

[26] H. Braden. The model AMBETI - a detailed description of a soil-plant-atmosphere model. Berichte des Deutschen Wetterdienstes No. 195, Deutscher Wetterdienst, Offenbach, Germany, 1995.

[27] R. H. Brooks and A. T. Corey. Hydraulic properties of porous media. *Hydrology Paper*, 3:22–27, 1964.

[28] N. T. Burdine. Relative permeability calculation from size distribution data. *Trans. AIME*, 198:71–78, 1953.

[29] R. F. Carsel, J. C. Imhoff, P. R. Hummel, J. M. Cheplick, and A. S. Donigian. PRZM-3, A model for predicting pesticide and nitrogen fate in the crop root and unsaturated soil zones. User Manual for Release 3.0, GA 30605-2720, National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Athens, 1998.

[30] M. B. Carver and H. W. Hinds. The method of lines and the advective equation. *Simulation*, pages 59–69, 1978.

[31] T. F. Coleman, B. S. Garbow, and J. J. Moré. Fortran subroutines for estimating sparse Jacobian matrices. *ACM Trans. Math. Software*, 10:346–347, 1984.

[32] T. F. Coleman, B. S. Garbow, and J. J. Moré. Software for estimating sparse Jacobian matrices. *ACM Trans. Math. Software*, 10:329–345, 1984.

[33] T. F. Coleman and J. J. Moré. Estimation of sparse Jacobian matrices and graph coloring problems. *SIAM J. Numer. Anal.*, 20:187–209, 1983.

[34] A. R. Curtis, M. J. D. Powell, and J. K. Reid. On the estimation of sparse Jacobian matrices. *J. Inst. Math. Appl.*, 13:117–119, 1974.

[35] J. H. Dane and S. Hruska. In-situ determination of soil hydraulic properties during drainage. *Soil Sci. Soc. Am. J.*, 47:619–624, 1983.

[36] B. Diekkrüger. *Standort- und Gebietsmodelle zur Simulation der Wasserbewegung in Agrarökosystemen*. PhD thesis, TU Braunschweig, 1992.

[37] B. Diekkrüger, P. Nörtersheuser, and O. Richter. Modeling pesticide dynamics of a loam site catchment using HERBSIM and SIMULAT. *Ecological Modelling*, 81:111–119, 1995.

[38] J. Diels. *A validation procedure accounting for model input uncertainty: methodology and application to the SWATRER model*. PhD thesis, Catholic University of Leuven, 1994.

[39] J. Dierckx, C. Belmans, and P. Pauwels. SWATRER. A computer package for modelling the field water balance. Reference Manual, Soil and Water Engng. Lab, Catholic University of Leuven, Leuven, Belgium, 1986.

[40] A. E. Dieses. Numerische Verfahren zur Diskriminierung nichtlinearer Modelle für dynamische chemische Prozesse. Master's thesis, IWR, University of Heidelberg, 1997.

[41] A. E. Dieses, J. P. Schlöder, H. G. Bock, and O. Richter. Parameter estimation for nonlinear transport and degradation processes of xenobiotica in soil. In F. Keil, W. Mackens, H. Voss, and J. Werther, editors, *Scientific Computing in Chemical Engineering II*, volume 2: Simulation, Image Processing, Optimization, and Control, pages 290–297. Springer-Verlag, Berlin, Heidelberg, 1999.

[42] A. E. Dieses, J. P. Schlöder, H. G. Bock, and O. Richter. Optimal experimental design for parameter estimation in column outflow experiments. Preprint 17, IWR/SFB 359, University of Heidelberg, 2000.

[43] A. E. Dieses, J. P. Schlöder, H. G. Bock, O. Richter, K. Aden, and B. Gottesbüren. A parameter estimation tool for nonlinear transport and degradation processes of xenobiotics in soil. In A. A. M. Del Re, C. Brown, E. Capri, G. Errera, S. P. Evans, and M. Trevisan, editors, *Human and Environmental Exposure to Xenobiotics*, Proceedings of the XI Symposium Pesticide Chemistry, Sept 11-1 5 1999, Cremona, Italy, pages 171–180. La Goliardica Pavese, 1999.

[44] J. Doherty. *PEST*. Watermark Computing, Corinda, Australia, 1994.

[45] W. Durner, B. Schultze, and T. Zurmühl. Transient flow experiments for the determination of soil hydraulic properties - An evaluation of different experimental boundary conditions. In W. Durner, J. Halbertsma, and M. Cislerova, editors, *European Workshop on Advanced Methods to Determine Hydraulic Properties of Soils*, pages 85–88, Department of Hydrology, University of Bayreuth, 1996.

[46] W. Durner, B. Schultze, and T. Zurmühl. State-of-the-art in inverse modeling of inflow/outflow experiments. In M. Th. van Genuchten, F. J. Leij, and L. Wu, editors, *Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media*, Proc. Int. Workshop, pages 661–681, University of California, Riverside, CA, 1999.

[47] V. V. Fedorov. *Theory of Optimal Experiments*. Academic Press, New York, 1972.

[48] V. V. Fedorov. Optimal experimental designs for discriminating two rival regression models. In J. N. Srivastava, editor, *A Survey of Statistical Design and Linear Models*. North-Holland Publishing Co., Amsterdam, 1975.

[49] V. V. Fedorov and M. B. Malyutov. Optimal designs in regression problems. *Math. Operationsforschung and Statistik*, 3:281–308, 1972.

[50] G. Fent. Determination of effective K_d-values of a substance X on undisturbed soil columns from three different sites. Confidential report, Staatliche Lehr- und Forschungsanstalt für Landwirtschaft, Weinbau und Gartenbau (SLFA), Department of Ecology, Neustadt/Weinstr., Germany, 1999.

[51] B. Fornberg. Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of Computation*, 51(184):699–706, 1988.

[52] R. M. Furzeland, J. G. Verwer, and P. A. Zegeling. A numerical study of three moving-grid methods for one-dimensional partial differential equations which are based on the method of lines. *J. Comput. Phys.*, 89:349–388, 1990.

[53] J. V. Gallitzendörfer. *Parallele Algorithmen für Optimierungsrandwertprobleme*. Fortschritt-Berichte, Reihe 10, Nr. 514. VDI-Verlag, Düsseldorf, 1997.

[54] W. Gander and J. Hrebicek. *Solving problems in scientific computing using MAPLE and MATLAB*. Springer, Berlin, 1993.

[55] Ph. E. Gill, W. Murray, and M. A. Saunders. SNOPT: an SQP algorithm for large-scale constrained optimization. Technical report, NA 97-2, Department of Mathematics, University of California, San Diego, and Report SOL 97-3, Dept. of EESOR, Stanford University, 1997.

[56] A. Grace. Optimization Toolbox User's Guide. Technical report, The MathWorks, Inc., Natick, MA, 1992.

[57] M. M. Gribb. Parameter estimation for determining hydraulic properties of a fine sand from transient flow measurements. *Water Resour. Res.*, 32(7):1965–1974, 1996.

[58] C. Großmann and H.-G. Roos. *Numerik partieller Differentialgleichungen*. Teubner, Stuttgart, 1992.

[59] K.-D. Hilf. *Optimale Versuchsplanung zur dynamischen Roboterkalibrierung*. Fortschritt-Berichte, Reihe 8, Nr. 590. VDI-Verlag, Düsseldorf, 1996.

[60] M. C. Hill. Methods and guidelines for effective model calibration. Water Resources Investigations Report 98-4005, U.S. Geological Survey, Denver, CO, 1998.

[61] J. W. Hopmans and J. Šimůnek. Review of inverse estimation of soil hydraulic properties. In M. Th. van Genuchten, F. J. Leij, and L. Wu, editors, *Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media*, Proc. Int. Workshop, pages 643–659, University of California, Riverside, CA, 1999.

[62] A. Horn. Leaching behaviour of S-Metolachlor and its main soil metabolites in different minilysimeters filled with various soils. Master's thesis, Technical University of Braunschweig, 1999.

[63] U. Hornung. Identification of nonlinear soil physical parameters from an input-output experiments. In P. Deuflhard and E. Hairer, editors, *Numerical Treatment of Inverse Problems in Differential and Integral Equations*, volume 2 of *Progress in Scientific Computing*, pages 227–237, Birkhäuser, Boston, 1983.

[64] N.-S. Hsu and W. W.-G. Yeh. Optimum experimental design for parameter identification in groundwater hydrology. *Water Resour. Res.*, 25(5):1025–1040, 1989.

[65] W. Hunter and A. Reiner. Designs for discriminating between two rival models. *Technometrics*, 7:307–323, 1965.

[66] M. Inoue, J. Šimůnek, J. W. Hopmans, and V. Clausnitzer. In situ estimation of soil hydraulic functions using a multistep soil-water extraction technique. *Water Resour. Res.*, 34(5):1035–1050, 1998.

[67] N. Jarvis. MACRO - A model of water movement and solute transport in macroporous soils. Reports and Dissertations 9, Department of Soil Science, Swedish University of Agriculture Sciences, Uppsala, Sweden, 1991.

[68] N. Jarvis and M. Larsson. The Macro Model (Version 4.1). Technical Description, Department of Soil Science, Swedish University of Agriculture Sciences, Uppsala, Sweden, http://www.mv.slu.se/macro/doc, 1998.

[69] B. Jene. PELMO 3.00. Manual Extension, Staatliche Lehr- und Forschungsanstalt für Landwirtschaft, Weinbau und Gartenbau (SLFA), Department of Ecology, Neustadt/Weinstr., Germany, 1998.

[70] Y. Jin and W. A. Jury. Characterizing the dependence of gas diffusion coefficient on soil properties. *Soil Sci. Soc. Am. J.*, 60:66–71, 1996.

[71] W. D. Kemper and J. D. van Schaik. *Soil Sci.*, 104:314–322, 1966.

[72] M. Klein. PELMO Pesticide Leaching Model, Version 2.01. User Manual, Frauenhofer Institut für Umweltchemie und Ökotoxikologie, Schmallenberg, Germany, 1995.

[73] R. Kloskowski, R. Fischer, R. Binner, and R. Winkler. Draft guidance on the calculation of predicted environmental concentration values (PEC) of plant protection products for soil, ground water, surface water and sediment. In A. A. M. Del Re, C. Brown, E. Capri, G. Errera, S. P. Evans, and M. Trevisan, editors, *Human and Environmental Exposure to Xenobiotics*, Proceedings of the XI Symposium Pesticide Chemistry, Sept 11-15 1999, Cremona, Italy, pages 835–850. La Goliardica Pavese, 1999.

[74] D. S. Knopman and C. I. Voss. Behavior of sensitivities in the one-dimensional advection-dispersion equation: Implications for parameter estimation and sampling design. *Water Resour. Res.*, 23(2):253–272, 1987.

[75] D. S. Knopman and C. I. Voss. Discrimination among one-dimensional models of solute transport in porous media: Implications for sampling design. *Water Resour. Res.*, 24(11):1859–1876, 1988.

[76] D. S. Knopman and C. I. Voss. Multiobjective sampling design for parameter estimation and model discrimination in groundwater solute transport. *Water Resour. Res.*, 25(10):2245–2258, 1989.

[77] D. S. Knopman, C. I. Voss, and S. P. Garabedian. Sampling design for groundwater solute transport: Tests of methods and analysis of Cape Cod tracer test data. *Water Resour. Res.*, 27(5):925–949, 1991.

[78] J. B. Kool and J. C. Parker. Development and evaluation of closed-form expressions for hysteretic soil hydraulic properties. *Water Resour. Res.*, 23:105–114, 1987.

[79] J. B. Kool and J. C. Parker. Analysis of the inverse problem for transient unsaturated flow. *Water Resour. Res.*, 24(6):817–830, 1988.

[80] J. B. Kool, J. C. Parker, and M. Th. van Genuchten. Determining soil hydraulic properties from one-step outflow experiments by parameter estimation: I. Theory and numerical studies. *Soil Sci. Soc. Am. J.*, 49:1348–1354, 1985.

[81] J. B. Kool, J. C. Parker, and M. Th. van Genuchten. Parameter estimation for unsaturated flow and transport models - A review. *J. Hydrol.*, 91:255–293, 1987.

[82] S. Körkel. *Numerische Methoden für Optimale Versuchsplanungsprobleme bei chemischen Reaktionssystemen*. PhD thesis, University of Heidelberg, in preparation.

[83] S. Körkel, I. Bauer, H. G. Bock, and J. P. Schlöder. A sequential approach for nonlinear optimum experimental design in DAE systems. In F. Keil, W. Mackens, H. Voss, and J. Werther, editors, *Scientific Computing in Chemical Engineering II*, volume 2: Simulation, Image Processing, Optimization, and Control, pages 338–345. Springer-Verlag, Berlin, Heidelberg, 1999.

[84] D. B. Leineweber. *Efficient reduced SQP methods for the optimization of chemical processes described by large sparse DAE models*. Fortschritt-Berichte, Reihe 3, Nr. 613. VDI-Verlag, Düsseldorf, 1999.

[85] S. Li, L. Petzold, and Y. Ren. Stability of moving mesh systems of partial differential equations. *SIAM J. Sci. Comput.*, 20(2):719–738, 1998.

[86] Th. W. Lohmann. *Ein numerisches Verfahren zur Berechnung optimaler Versuchspläne für beschränkte Parameteridentifizierungsprobleme*. Reihe Informatik. Verlag Shaker, Aachen, 1993.

[87] Th. W. Lohmann, H. G. Bock, and J. P. Schlöder. Numerical methods for parameter estimation and optimal experiment design in chemical reaction systems. *Ind. Eng. Chem. Res.*, 31:54–57, 1992.

[88] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.*, 11:431 – 441, 1963.

[89] J. T. McCord. Application of second-type boundaries in unsaturated flow modeling. *Water Resour. Res.*, 27(12):3257–3260, 1991.

[90] A. Medina and J. Carrera. Coupled estimation of flow and solute transport parameters. *Water Resour. Res.*, 32(10):3063–3076, 1996.

[91] R. J. Millington and J. P. Quirk. Permeability of porous solids. *Trans. Faraday Soc.*, 57:1200–1207, 1961.

[92] S. Mishra and J. C. Parker. Parameter estimation for coupled unsaturated flow and transport. *Water Resour. Res.*, 25(3):385–396, 1989.

[93] J. L. Monteith. Evaporation and environment. *Symp. Soc. Exp. Biol.*, 19:205–224, 1965.

[94] Y. Mualem. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.*, 12:513–522, 1976.

[95] T. Nishikawa and W. W.-G. Yeh. Optimal pumping test design for the parameter identification of groundwater systems. *Water Resour. Res.*, 25(7):1737–1747, 1989.

[96] P. Nörtersheuser. *Aufbau von Modellen zur Beschreibung des Verhaltens von Pflanzenschutzmitteln im Boden und Anwendung am Beispiel des Herbizids Quinmerac*. PhD thesis, TU Braunschweig, 1993.

[97] R. V. O'Neill, R. A. Oldstein, H. H. Shugart, and J. B. Makin. Terrestrial ecosystem energy model. Technical report, Eastern deciduous forest biome memo report No. 72-19, 1972.

[98] J. C. Parker, J. B. Kool, and M. Th. van Genuchten. Determining soil hydraulic properties from one-step outflow experiments by parameter estimation: II. Experimental studies. *Soil Sci. Soc. Am. J.*, 49:1354–1359, 1985.

[99] J. C. Parker and M. Th. van Genuchten. Flux-averaged and volume-averaged concentrations in continuum approaches to solute transport. *Water Resour. Res.*, 20(7):866–872, 1984.

[100] H. L. Penman. Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. London*, 193:120–145, 1948.

[101] J. R. Plimmer. Past, present and future of pesticides. In A. A. M. Del Re, C. Brown, E. Capri, G. Errera, S. P. Evans, and M. Trevisan, editors, *Human and Environmental Exposure to Xenobiotics*, Proceedings of the XI Symposium Pesticide Chemistry, Sept 11-15 1999, Cremona, Italy, pages 1–10. La Goliardica Pavese, 1999.

[102] E. P. Poeter and M. C. Hill. Documentation of UCODE, A Computer Code for Universal Inverse Modeling. Water Resources Investigations Report 98-4080, U.S. Geological Survey, Denver, CO, 1998.

[103] F. Pukelsheim. *Optimal Designs of Experiments*. Wiley series in probability and mathematical statistics. John Wiley and Sons, New York, 1993.

[104] O. Richter, B. Diekkrüger, and P. Nörtersheuser. *Environmental Fate Modelling of Pesticides*. VCH, Weinheim, 1996.

[105] O. Richter, P. Nörtersheuser, and B. Diekkrüger. Modeling reactions and movement of organic chemicals in soils by coupling of biological and physical processes. *Modeling Geo-Bioshere Processes*, 1:95–114, 1992.

[106] K. Roth. *Lecture Notes in Soil Physics*. University of Hohenheim, Institute of Soil Science, 1996.

[107] W. Schiesser. *The Numerical Method of Lines – Integration of Partial Differential Equations*. Academic Press, San Diego, New York, Boston, 1991.

[108] J. P. Schlöder. Numerische Methoden zur Behandlung hochdimensionaler Aufgaben der Parameteridentifizierung. *Bonner Mathematische Schriften 187*, 1988.

[109] J. P. Schlöder and H. G. Bock. Identification of rate constants in bistable chemical reactions. In P. Deuflhard and E. Hairer, editors, *Numerical Treatment of Inverse Problems in Differential and Integral Equations, Progress in Scientific Computing*, pages 27–47. Birkhäuser, Boston, 1983.

[110] H. R. Schwarz. *Numerische Mathematik*. Teubner, Stuttgart, 1997.

[111] S. Silvey. *Optimal Design*. Chapman and Hall, London, 1980.

[112] J. Šimůnek, R. Angulo-Jaramillo, M. G. Schaap, J.-P. Vandervaere, and M. Th. van Genuchten. Using an inverse method to estimate the hydraulic properties of crusted soils from tension-disc infiltrometer data. *Geoderma*, 86:61–81, 1998.

[113] J. Šimůnek, R. Kodešová, M. M Gribb, and M. Th. van Genuchten. Estimating hysteresis in the soil water retention function from cone permeameter experiments. *Water Resour. Res.*, 35(5):1329–1345, 1999.

[114] J. Šimůnek and M. Th. van Genuchten. Estimating unsaturated soil hydraulic properties from tension disc infiltrometer data by numerical inversion. *Water Resour. Res.*, 32(9):2683–2696, 1996.

[115] J. Šimůnek and M. Th. van Genuchten. Estimating unsaturated soil hydraulic properties from multiple tension disc infiltrometer data. *Soil Science*, 162(6):383–398, 1997.

[116] J. Šimůnek and M. Th. van Genuchten. Using the HYDRUS-1D and HYDRUS-2D codes for estimating unsaturated soil hydraulic and solute transport parameters. In M. Th. van Genuchten, F. J. Leij, and L. Wu, editors, *Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media*, Proc. Int. Workshop, pages 1523–1536, University of California, Riverside, CA, 1999.

[117] J. Šimůnek, T. Vogel, and M. Th. van Genuchten. The SWMS_2D code for simulating water flow and solute transport in two-dimensional variably saturated media, version 1.2. Res. Rep. No 132, U.S. Salinity Lab., U.S. Dep. of Agric., Riverside, CA, 1994.

[118] J. Šimůnek, M. Šejna, and M. Th. van Genuchten. The HYDRUS-2D software package for simulating water flow and solute transport in two-dimensional variably saturated media, version 1.0. IGWMC-TPS-53, Int. Ground Water Modeling Cent., Colo. School of Mines, Golden, CO, 1996.

[119] J. Šimůnek, M. Šejna, and M. Th. van Genuchten. The HYDRUS-1D software package for simulating water flow and solute transport in two-dimensional variably saturated media, version 2.0. IGWMC-TPS-70, Int. Ground Water Modeling Cent., Colo. School of Mines, Golden, CO, 1998.

[120] J. Šimůnek, O. Wendroth, and M. Th. van Genuchten. Parameter estimation analysis of the evaporation method for determining soil hydraulic properties. *Soil Sci. Soc. Am. J.*, 62:894–905, 1998.

[121] J. Šimůnek, O. Wendroth, and M. Th. van Genuchten. Estimating unsaturated soil hydraulic properties from laboratory tension disc infiltrometer experiments. *Water Resour. Res.*, 35(10):2965–2979, 1999.

[122] A. Stock. *Untersuchung der räumlichen Variabilität des Stofftransportes in der wasserungesättigten Bodenzone*. PhD thesis, TU Braunschweig, 1995.

[123] N.-Z. Sun. *Inverse Problems in Groundwater Modeling*, volume 6 of *Theory and Applications of Transport in Porous Media*. Kluwer Academic Publishers, 1994.

[124] N.-Z. Sun and W. W.-G. Yeh. Coupled inverse problems in groundwater modeling, 2. Identifiability and experimental design. *Water Resour. Res.*, 26(10):2527–2540, 1990.

[125] A. Tiktak, F. van den Berg, J. J. T. I. Boesten, M. Leistra, A. M. A. van der Linden, D. van Kraalingen, and J. te Roller. Pesticide emission assessment at regional and local scales: User manual of PEARL, version 1.1. RIVM report 711401 008, RIVM, Bilthoven, The Netherlands, 1999.

[126] A. F. Toorman, P. J. Wierenga, and R. G. Hills. Parameter estimation of hydraulic properties from one-step outflow data. *Water Resour. Res.*, 28(11):3021–3028, 1992.

[127] J. C. van Dam, J. N. M. Stricker, and P. Droogers. Inverse method for determining soil hydraulic functions from one-step outflow experiments. *Soil Sci. Soc. Am. J.*, 56:1042–1050, 1992.

[128] J. C. van Dam, J. N. M. Stricker, and P. Droogers. Inverse method to determine soil hydraulic functions from multistep outflow experiments. *Soil Sci. Soc. Am. J.*, 58:647–652, 1994.

[129] M. Th. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.*, 44:892–898, 1980.

[130] M. Th. van Genuchten and W. J. Alves. Analytical solutions of the one-dimensional convective-dispersive solute transport equation. Technical Bulletin Nr. 1661, U.S. Department of Agriculture, Agricultural Research Service, Washington, D.C., 1982.

[131] M. Th. van Genuchten and J. C. Parker. Boundary conditions for displacement experiments through short laboratory soil columns. *Soil Sci. Soc. Am. J.*, 48:703–708, 1984.

[132] J. G. Verwer, J. G. Blom, and J. M. Sanz-Serna. An adaptive moving grid method for one-dimensional systems of partial differential equations. *J. Comput. Phys.*, 82:454–486, 1989.

[133] J. P. M. Vink, P. Nörtersheuser, O. Richter, B. Diekkrüger, and K. P. Groen. Modeling the microbial breakdown of pesticides in soil using a parameter estimation technique. *Pesticide Science*, 40:285–292, 1994.

[134] N. von Götz. *Regionalisierung des Herbizidabbaus im Boden unter besonderer Berücksichtigung von Bentazon*. PhD thesis, TU Braunschweig, 1997.

[135] M. von Schwerin. *Numerische Methoden zur Schätzung von Reaktionsgeschwindigkeiten bei der katalytischen Methankonversion und Optimierung von Essigsäure– und Methanprozessen*. PhD thesis, University of Heidelberg, 1998.

[136] B. J. Wagner. Sampling design methods for groundwater modeling under uncertainty. *Water Resour. Res.*, 31(10):2581–2591, 1995.

[137] A. Walker and R. Allen. Influence of soil and environmental factors on pesticide persistence. *Symp. Soils Crop Protect. Chem.*, 27:89–100, 1984.

[138] E. Walter and L. Prozato. Robust experiment design: Between qualitative and quantitative identifiabilities. In E. Walter, editor, *Identifiability of Parametric Models*. Pergamon, New York, 1987.

[139] R. Weiss and L. Smith. Parameter space methods in joint parameter estimation for groundwater flow models. *Water Resour. Res.*, 34(4):647–661, 1998.

[140] S. Wolfram. *MATHEMATICA*. Cambridge University Press, 1997.

[141] J. H. M. Wösten, G. H. Veerman, and J. Stolte. Water retention and hydraulic conductivity functions of top- and subsoils in the Netherlands. The Staring Series. SC-DLO Technical Document No. 18, RIVM, Wageningen, The Netherlands, 1994.

[142] W. W.-G. Yeh. Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resour. Res.*, 22(2):95–108, 1986.

[143] D. W. Zachmann, P. C. DuChateau, and A. Klute. The calibration of the Richards flow equation for a draining column by parameter identification. *Soil Sci. Soc. Am. J.*, 45:1012–1015, 1981.

[144] E. Zauderer. *Partial Differential Equations of Applied Mathematics*. Wiley, New York, 1989.

[145] M. Zieße, H. G. Bock, J. Gallitzendörfer, and J. Schlöder. Parameter estimation in multispecies transport reaction systems using parallel algorithms. *Parameter Identification and Inverse Problems in Hydrology, Geology and Ecology*, 7:273–282, 1996.

[146] T. Zurmühl and W. Durner. Determination of parameters for bimodal hydraulic functions by inverse modeling. *Soil Sci. Soc. Am. J.*, 62:874–880, 1998.