DISSERTATION

submitted
to the
Combined Faculties for the Natural Sciences and for Mathematics
of the
Ruperto-Carola University Heidelberg, Germany

for the degree of
Doctor of Natural Sciences

Put forward by

Dipl.-Math. Ulrike Angela Niederle

born in Hadamar, Germany

Oral examination: _____

Time-Series Projection of Surface-Measured Quantities into Depth

Advisor: Prof. Dr. Peter Bastian

Abstract

This thesis deals with projection methods of soil temperature and soil moisture into depth by a forward model based on near-surface time-series. In addition, thermal as well as hydraulic soil parameters are estimated by using the Levenberg-Marquardt algorithm.
For soil temperature, two analytical projection methods are compared which use the transfer function and the Fourier transform approach, respectively. In each case, additional mathematical strategies are required to improve the projection results, e. g. by adding an integral over the initial profile or applying the Tukey window on the time-series. The resulting projected temperature matches field measurements very well, especially for the transfer function method a residuum down to $\pm 0.05\,^{\circ}\mathrm{C}$ is achieved. Further, the uniqueness of the parameter space is evaluated and the temporal evolution of the thermal diffusivity is estimated through both projection methods.
The projection of near-surface soil moisture is realized numerically by a finite volume scheme due to the strong non-linearity of the system. On the basis of synthetic data the conditions are explored under which accurate estimations of the hydraulic parameters are feasible. One of these conditions is found to be that the water content range should be larger than 0.5 times of the porosity. The corresponding relative parameter error is found to be some 10^{-5}. Furthermore, the study shows that an accurate estimation is more feasible for soils with a lower saturated conductivity, and steep functions of the hydraulic properties. This is typically the case for soils with a higher sand content. Data from a field site is then used to verify the findings of the synthetic study and to discuss limitations, e. g. ponding water at a soil layer interface.

Zusammenfassung

Diese Arbeit befasst sich mit Projektionsmethoden von Bodentemperaturen und Bodenfeuchten in die Tiefe durch ein Vorwärtsmodell auf der Basis von Zeitreihen nahe der Bodenoberfläche. Darauf aufbauend werden thermische und hydraulische Bodenparameter mittels des Levenberg-Marquardt Algorithmus durch einen Vergleich der Projektionsergebnisse geschätzt.
Für die Bodentemperaturen werden zwei analytische Projektionsmethoden verglichen, basierend jeweilig auf den Ansätzen der Transferfunktion und der Fourier Transformation. Durch mathematischen Strategien werden die Projektionsergebnisse weiter verbessert, z.B. durch ein zusätzliches Integral über die Tiefe oder durch Anwenden des Tukey Fensters auf die Zeitreihe. Die resultierenden Projektionsergebnisse stimmen sehr gut mit Feldmessungen überein, speziell die Transferfunktionsmethode erzielt ein Residuum von bis zu $\pm 0.05\,^{\circ}\mathrm{C}$. Schließlich werden für beide Methoden die Ergebnisse der Parameterschätzung bezüglich der Eindeutigkeit des Parameterraums und der zeitlichen Entwicklung der thermischen Diffusivität gegenübergestellt.
Die Projektion von oberflächennaher Bodenfeuchte wird wegen der starken Nichtlinearität des Systems numerisch durch eine Finite-Volumen-Methode beschrieben. In einer synthetischen Studie werden die Bedingungen untersucht, unter denen eine genaue Schätzbarkeit der Parameter möglich ist. Als eine dieser Bedingungen stellt sich heraus, dass die Wassergehaltsspanne mehr als die Hälfte der Porosität betragen sollte, um eine genaue Parameterschätzung zu ermöglichen. Der folgende relative Parameterfehler liegt in der Größenordnung von 10^{-5}. Weiterhin liefert die Studie, dass eine genaue Schätzung der Parameter für Böden mit einer niedrigen gesättigten Leitfähigkeit und steilen Funktionen der hydraulischen Eigenschaften eher möglich ist. Das ist typischerweise der Fall für Böden mit einem höheren Sandgehalt. Die daraus folgenden Beobachtungen werden schließlich an Felddaten verifiziert und die Einschränkungen u.a. durch Aufstauen von Wasser an einer Grenzschicht im Boden diskutiert.

Symbols

a	parameter of the Tukey window [−]
α	scaling parameter [m^{-1}]
χ^2	least squares value
χ^2_ν	reduced least squares value
C_h	thermal capacity [$\mathrm{J\,m^{-3}\,K^{-1}}$]
D_h	thermal diffusivity [$\mathrm{m^2\,s^{-1}}$]
f	frequency [s^{-1}]
f_c	cutoff frequency [s^{-1}]
$\mathbf{g}$	acceleration of gravity [$\mathrm{m\,s^{-2}}$]
h_m	matrix head [m]
i_{t0}	index after spin-up [−]
I_E	integral of spin-up estimation [−]
$\mathbf{j}_w$	volume flux of water [$\mathrm{m\,s^{-1}}$]
j_max	maximum of the absolute value of flux j_w [$\mathrm{m\,s^{-1}}$]
K_h	thermal conductivity [$\mathrm{W\,m^{-1}\,K^{-1}}$]
K_w	hydraulic conductivity; [$\mathrm{J\,m^5\,s^{-1}}$] in potential form or [$\mathrm{cm\,h^{-1}}$] in head form
K_0	saturated hydraulic conductivity; [$\mathrm{J\,m^5\,s^{-1}}$] in potential form or [$\mathrm{cm\,h^{-1}}$] in head form
l	parameter of the van Genuchten parameterization; [−]
n	van Genuchten parameter [−]
N_t	length of time-series [−]
Ω_t	time domain [s]
$\Omega_\mathbf{x}$	spatial domain in three-dimension [m]
Ω_z	spatial domain in vertical direction [m]
ω	angular frequency [s^{-1}]
p	estimation parameter set [−]
PL	percentage limit [−]
p_t	travel time probability density function [s^{-1}]
p_δ^IC	travel distance probability density function [m^{-1}]
$\Delta p_i'$	relative error of parameter p_i [−]
ϕ	porosity [−]
ψ_m	matrix potential [$\mathrm{J\,m^{-3}}$]
ψ_w	water potential [$\mathrm{J\,m^{-3}}$]
ρ_w	mass density of water [$\mathrm{kg\,m^{-3}}$]
θ	volumetric soil water content [−]
θ_r	residual water content [−]
θ_s	saturated water content [−]
θ_meas	measured water content [−]
θ_sim	simulated water content [−]
$\Delta\tilde{\theta}$	water content range [−]
Θ	saturation [−]

T	temperature [°C]
T_0	temperature at the surface [°C]
T_{meas}	measured soil temperature [°C]
T_{proj}	projected soil temperature at depth z_p [°C]
T_{ref}	reference temperature [°C]
ΔT	temperature difference [°C]
t	time [s]
t_0	initial time [s]
t_s	time of spin-up [s]
t_f	final time [s]
$t_{\max}$	maximal time [s]
$t_{\min}$	minimal time [s]
Δt	time difference [s]
$\mathbf{x}$	position in three-dimension [m]
z	depth [cm]
z_0	depth of the top most sensor [cm]
z_i	depth of the mirrored pluses [cm]
z_p	projection depth [cm]

Mathematical Notation

cos	cosine function
erf	Gaussian error function
exp	exponential function
sin	sine function
δ	delta distribution
Δ	Laplace operator
∇	Nabla operator
∂_t	partial derivative with respect to time
∂_z	partial derivative with respect to vertical coordinate
$\|\cdot\|_2$	L_2-norm
$\mathcal{F}$	Fourier transform
χ^2	sum of squares of the residuum
χ^2_ν	χ^2 divided by the difference of number of elements and estimation parameters
σ_i	standard deviation

Contents

1. Introduction

— It [the book of nature] is written in the language of mathematics, ...
Galilei ("II Saggiatore (The Assayer)")

For hundreds of years, humanity desired to describe their natural surroundings. They observed the nature and tried to find explanations for the observations. Galileo Galilei had an important role by applying a sparking idea (*Hund (1972)*). He linked his own observations with mathematical descriptions. His quotation above emphasizes how close he saw the connection between observation and mathematics. The main feat was to realize that mathematical symbols exist which can be mapped isomorphically to natural events.
He also developed measurement instruments to make observations, e. g. a predecessor of modern temperature sensors. Such tools in combination with his measurements were the basis for his formulation of principles. Mathematical symbols were his way to write down these principles. This kind of description gave the language of mathematics a higher significance in the context of scientific use.

The triumph of this strategy of the close connection between observation and mathematics was great. Years later lots of scientists followed this strategy and revealed further mathematical proportions in natural phenomena. This bunch of single principles was formulated in many different physical disciplines and established linear as well as more complicated connections.
By the end of the 19. century, the known natural phenomena have been translated into mathematical symbols. *Hund (1972)* reported that during this time, known descriptions were applied to many fields. Since there remained some unsolved problems in the observations that could not be described by the current models at this time, the models had to be adapted and new concepts had to be developed.
In the last century, this led to a mathematical description with much more details. The separate developed principles were combined to more complicated models to describe nature more accurately. By now, the technical tools have been improved to the point that even complicated processes can be derived within a reasonable time. The technical feasibility promote very detailed descriptions and evaluations which requires lots of computational effort.

In the context of solving complicated systems, two main ways exist to end up with a solution. As explained, the complicated system can be derived in higher resolution in order to end up with a more accurate result.
On the other hand, the complicated system can be approximated by a more simple model. Such a simplification can be used to model only the main driving processes of the complicated system. However, this approach may not describe the nature as accurately. Nevertheless, the approximation can be accurate enough in the context of the application issue.
Several questions are associated with the strategy of simplification. How much effort is required to generate acceptable results? How much effort has to be spent for a given accuracy? How efficient are simple models for the description of the complicated nature?

1.1. Projection and Parameter Estimation in Soil Physics

Natural processes have to be seen from a larger field of view to describe the human environment. The description of soil is only one aspect of the whole model issue to make predictions of huge areas. Detailed descriptions with interactions can become too elaborate to model a combination of soil, air and ocean. More simple models reduce the computational costs and can also provide the necessary accuracy. The balance of temperature and water content is one of the most important aspect within the soil and requires a deeper consideration.
Each model depend on input data which is usually provided by measurements. Based on the modeling scale, such data sets of soil temperature and water content can be measured by satellites, airplanes or local stations. But they can detect only the uppermost centimeters of the soil or the natural soil matrix is destroyed during the installation. Not only the top part of the soil gives feedback to the whole system but also deeper soil depth, e. g. in the context of water recharge.
Projection methods are required to transfer quantities from the surface into depth to get an impression of physical quantities at greater depth without any direct measurements. In this context, the question arises whether a projection into depth by a simple model with a practicable accuracy is feasible. Moreover, the soil properties are essential for a reliable projection. Hence, the parameter estimation represents a key factor. Therefore, one question to analyze is how dominant are additional processes besides the taken simple model.

In this work, projection methods will be presented for physical quantities into depth and using given time-series close to the surface. This projection of near-surface quantities will be demonstrated on two physical quantities, the soil temperature and the soil moisture. The estimation of the thermal and hydraulic soil properties will be analyzed with the focus on the feasibility, respectively. Moreover, the thermal property will be determined in the temporal evolution whereas the conditions are analyzed under which an accurate estimation is feasible for the hydraulic properties.

In both applications, the projection is based on a more simple forward model in order to describe the complicated nature. In the context of soil temperature, the heat conduction is the main driving process and yield the mathematical description. Furthermore, the soil moisture is modeled by the Richards equation together with a parameterization in porous mediums. Respectively, the heat conduction and Richards equation are simple models due to the simplified conditions. In both cases, the coupling between temperature and water content will be neglected. Furthermore, the porous medium is assumed to be homogeneous, i. e. no layers as well as other heterogeneities. Additional processes like freezing and thawing of water only have an effect on the measurements bot not on the simulation. Similarly, the complexity of the model will be reduced by neglecting plants in the model.

1.2. State of the Art

Lots of people did research in the context of modeling thermal and hydraulic processes in the soil. In the following, some of their main approaches will be presented. Since the temperature description is based on a linear system, there are more opportunities to formulate analytical solutions besides the numerical approach whereas the hydraulic in porous media follows a nonlinear system. This makes the analytical way difficult. Thus, the presentation of the main

approaches is divided into two topics, the soil temperature and the soil water content. Coupled approaches between them will be summarized in both sections.

1.2.1. Soil Temperature

In the past, there were several approaches to solve the heat conduction equation and to projected the soil temperature with this. Based on a given upper soil temperature, the deeper temperature was derived by analytical as well as numerical methods.
The major applications with a numerical approach used the finite difference or finite element method. Especially for models of permafrost soils, the numerical way was chosen to determine soil temperature like *Marchenko et al. (2008)* demonstrated at data from Alaska. The model has to account the freezing and thawing for permafrost. Thus, a coupling of water content and temperature is necessary in this case. This is also shown by *Hansson et al. (2004)* in their numerical simulations of water flow and heat transport combined with a laboratory column freezing experiment. Furthermore, *Javierre et al. (2006)* compared different numerical solution models for one-dimensional Stefan problems and concluded suitable results for the moving grid, the level set and the phase field methods for most of the tested examples. Moreover, *Riseborough et al. (2008)* summarized modelling approaches of annual soil temperature in permafrost following numerical and analytical ways. In the context of phase change, *Kane et al. (2001)* and *Van der Velde et al. (2010)* also focused on non-conductive heat transfer in soils.
There exist several approaches for the analytical way of projection , but the surface temperature will always be approximated by a sum of functions. Usually, such methods assume a homogeneous soil in a semi-infinite domain. The harmonic method is one of the most used methods in this context. There, the surface temperature is described by a Fourier's series. Its sum can be truncated after one or two summands for simplification and using only one significant frequency of the surface time-series as shown by *Horton et al. (1983)*.
de Silans et al. (1996) compared such a Fourier based method with the transfer function method. These transfer functions were derived by *Carslaw et al. (1959)* for the decomposition of the surface temperature as well as the initial profile. The applications of *de Silans et al. (1996)* generally showed similar results, but the transfer function approach required more computational time and yield better results in case of abrupt climate modifications. Moreover, the transfer function method was applied on permafrost data of Svalbard by *Roth et al. (2001)*. They showed that the projected soil temperature fitted well the measured one within frozen soils and under the same conditions, the followed winter temperature were also approximated quite accurately. As *Roth et al. (2001)* expected, their pure heat conduction model worked in pure thawed or pure frozen soils. This transfer function method and the Fourier based method is applied and compared in this work to projected soil temperature into depth.

An application of these projection methods are given in the estimation of the thermal properties of the soil. The estimation approaches can be based on a numerical inversion or on an analytical expression.
Nicolsky et al. (2009) estimated the thermal conductivity of permafrost soils by minimization the temperature difference. They summarized that their approach required certain values of heat capacity to find an unique value of the conductivity.
Most applications focused on the thermal diffusivity since it combines the thermal conductivity with the capacity. *Horton et al. (1983)* summarized six methods to estimate the thermal diffusivity. The amplitude and phase methods considered the surface temperature only as a single sine function whereas the arctangent and logarithmic methods included two sine and cosine

functions. All these methods are quite rough for the estimation of the thermal diffusivity since the surface temperature is approximated only by a few trigonometrical functions. However, the harmonic method is more accurate due to the decomposition as a Fourier series. As mentioned above, this method projects the temperature into depth and, by a minimization approach, the thermal diffusivity is estimated. Finally, *Horton et al. (1983)* described the numerical method by using finite differences. There, the diffusivity D_h was extracted by minimizing the temperature difference. Further, they mentioned that the stability of the numerical solution is ensured under the condition $D_h \Delta t/(\Delta z)^2 < 0.5$. *Horton et al. (1983)* summarized that the harmonic method is the most preferential one of the six presented methods.
Further comparisons between five of these methods and a thermal conduction-convection approach were carried out by *Gao et al. (2009)* for a data set measured in China. Their conduction-convection approach show the second best estimation results after the harmonic method. The applications from *Anandakumar et al. (2001)* also showed a reliable estimation by using the harmonic method and compared the results with the thermal conductivity and capacity estimation of *Massman (1992)*.
In different contexts. the amplitude method was applied e. g. in thawed or frozen soil or the thermal properties were tested as a proxy for soil moisture. *Hinkel (1997)* and also *Krzeminska et al. (2011)* compared the results received by the amplitude method with the finite difference approach.
In this work. the methods to project the soil temperature and to estimate the thermal diffusivity are the transfer function approach and the Fourier based method (harmonic method). As discussed above, both methods were applied to estimate the thermal diffusivity by *de Silans et al. (1996)*.

1.2.2. Soil Water Content

Several approaches were investigated in the past to calculate water content within the soil. Before the computational tools were well established, one research direction searched for an analytical solution which described the water content flow. *Knight et al. (1974)* presented one of the first approaches. They solved a non-linear diffusion equation of the water content by applying some transformations in order to transfer it into a linear equation. During the next years, *Broadbridge et al. (1988)* and *Warrick et al. (1991)* refined the analytical method for applications with constant rate infiltration and later with stepwise varying infiltration. Further development to model analytically more flexible scenarios were presented, e. g. *Pullan (1992)* for a linearized time-dependent infiltration or *Prevedello et al. (2008)* using a Boltzmann transform.

However, in the last decades the numerical approach to model soil water content became common. In general, all models deal with a water content dependent hydraulic conductivity and pressure head. The most common parameterizations of these quantities were presented by *Mualem (1976)*, *van Genuchten (1980)* and *Brooks et al. (1966)*.

Feddes et al. (1988) summarized the modeling of soil water dynamics in the unsaturated zone and went into details about thermally induced soil water flow, frozen soils and water extraction by roots. The numerical solving of these equations were discussed by introducing into finite difference and finite element methods. Further, they summarized the different types of boundary conditions and the effect of hysteresis. One of their four applications were the modeling of field water use. But other modeling scenarios of water content were applied e. g. for vertical soil profiles and a steep hillslope by *Kim et al. (2011)*, for a coupling of water, vapor and heat

transport in the vadose zone by *Saito et al. (2006)* and for non-equilibrium in soil hydraulic modeling by *Vogel et al. (2010)*. *Seneviratne et al. (2010)* provided an introduction to get an impression about the interaction between soil moisture and climate.

The data assimilation is another further huge research field of in the context of modeling soil water content. As *Reichle (2008)* defined, data assimilation aims to balance information from measurements and models in the Earth sciences into an optimal estimate of the interested field. His review gave a brief introduction to data assimilation strategies like variational techniques and Kalman filter approaches whereas *Vereecken et al. (2008)* additionally described direct insertion, statistical correction assimilation, Newtonian nudging and inverse modeling in the local and global sense.
Now, the concept of Kalman filters is widely-used and had its origin in the work of *Kalman et al. (1960)*. There he combined measurements with a linear model to predict the next state of the modeled variable. *Welch et al. (1995)* introduced the extended Kalman filter besides the normal one. The extended one improved the applicability of the Kalman filter also for non-linear description by an additional linearization. Further efforts of the development of the extended Kalman filter were made by *Walker (2006)* including constraints of the system. Exemplary, *Sabater et al. (2007)* applied the extended but also the ensemble Kalman filter in the context of root zone soil moisture. They indicated the ensemble Kalman filter as a promising technique to deal with high nonlinear systems based on near-surface SMOS data.
Evensen (1994) made the Kalman filter more efficient using a Monte Carlo method. His ensemble Kalman filter was more flexible in applications. Several people made use of the ensemble Kalman filter in the assessment of the root zone soil moisture like *Dunne et al. (2006)* and *Pan et al. (2006)*. Furthermore, *Das et al. (2006)* demonstrated the reasonable capability to generate distributions of soil moisture at the desired resolution in the context of remote sensing.
Moreover, inverse modeling is one of the strongest research fields. Related to soil science, *Vrugt et al. (2008)* summed up inversion algorithms with the focus on global ones instead of local ones. They presented one of the most specialized global ensemble optimizer AMALGAM. Based on Monte Carlo strategies, this algorithm combines coupled multiple different search strategies. In this context, a couple of studies were made about parameter estimation, e. g. *Huisman et al. (2010)* and *Steenpass et al. (2010)*, and parameter uncertainty, e. g. *Vrugt et al. (2002)* and *Loosvelt et al. (2011)*. *Scharnagl et al. (2011)* summarized their uncertainty results that the prior information about the hydraulic parameters significantly improved the parameter identifiability in Bayesian inverse modeling.
Further, genetic algorithm was applied to estimate soil moisture properties. *Ines et al. (2008)* compared the near-surface soil moisture and found out that the soil properties could be identified better than for a medium-textured soil for the coarser and finer soil examples. In their layered scenarios, the parameters of the bottom layer were not identifiable. *Ines et al. (2009)* applied their method using airborne remote sensing.
Steenpass et al. (2010) simulated a coupled thermal and hydraulic model by using near-surface data. They compared a single- and multi-objective estimation approach for the hydraulic parameters using a global optimizer. The multi-objective approach had more data information and fitted the given temperature and water content slightly better.
Durner et al. (2008) analyzed the influence of heterogeneities onto effective hydraulic parameters by using a global inversion strategy. They found out that the accurate estimation of hydraulic parameters in two layers decreases by increasing the heterogeneity of the soil.

Traditionally, a lot of local optimization strategies of non-linear problems exist. One of the

most common one is the Levenberg-Marquardt algorithm which is applied in this work. The basis of this algorithm was invented by *Levenberg (1944)* and *Marquardt (1963)*. Later, *More (1978)* described their idea by weighing between the Gauss-Newton and steepest descent strategy.
Wollschläger et al. (2009) estimated the hydraulic properties in four soil layers on the field scale . Their inversion by the Levenberg-Marquardt algorithm showed useful results but they concluded that a better representation of the upper boundary condition of the model would be necessary.
A lot of further applications with the Levenberg-Marquardt algorithm were done. To name just a few of them: *Simunek et al. (1998)* estimated hydraulic properties by three field methods in unsaturated soils from transient flow processes whereas *Schwen et al. (2011)* focused more on the time-variable parameters for different tillage methods.
A further study area is the class of multistep outflow experiments. Lots of their optimization work was done by the Levenberg-Marquardt algorithm. *Ippisch et al. (2006)* showed that under saturated conditions the classical Mualem-van Genuchten model produced unrealistic conductivity curves when the van Genuchten parameter n was smaller 2 comparing such experiments with simulations including either van Genuchten or Brooks-Corey parameterization.

1.3. Outline

In this work different methods are presented to project soil temperature and soil moisture into depth within homogeneous porous medium. These methods are used to estimate thermal and hydraulic properties, respectively. The coupling between temperature and soil moisture and other processes like freezing are neglected. Thus, the projection and estimation is analyzed separately. In any case, the projection idea is to calculate the physical quantity at depth from a given surface time-series.
First, the theory of the projection and estimation methods is described for soil temperature and soil moisture. Afterwards, the applications of a real measured data set will be discussed for both cases, the linear system of the soil temperature and the non-linear system for the soil moisture. In both cases, a forward model will be presented for the projection and an inversion problem will be solved by the Levenberg-Marquardt algorithm to estimate the soil properties.

Chapter 2: Projection Methods of Soil Temperature.

The projection of soil temperature is realized by two different approaches, the transfer function and the Fourier based method. Since the simplified heat equation is linear, this advantage is taken to determine an analytical solution of the presented heat model.
In the transfer function method, the surface temperature is decomposed into a superposition of δ-functions under a constant initial condition. A variable initial profile is taken into account by adding a spin-up phase and a further integral to improve the projection results.
In the Fourier based approach, the decomposition of the surface time-series into complex exponential functions results that the projection is derived in the frequency space. The strategy is adapted to avoid some error source by applying an additional window.
The thermal diffusivity is estimated by comparing the projected and measured temperature for both methods.

Chapter 3: Projection Method of Soil Moisture.
A numerical problem is used to solve the Richards equation together with the van Genuchten parameterization by a finite volume method. When the soil parameters are known, the projected water content is produced a numerical forward run. A main focus is on the parameter estimation based only on the near-surface information which is demonstrated at a first synthetic example.

Chapter 4: Applications of Estimation and Projection to Measured Soil Temperature.
The projection and estimation is presented for measured temperatures for both methods. Step-by-step, the additional tools are included to demonstrate the improvement. In each case, the parameter space is scanned to analyze the minimization problem. The temporal evolution of the thermal diffusivity is determined by a moving time interval and discussed. Finally, both methods are compared to study the pros and cons of the approaches.

Chapter 5: Estimation of Hydraulic Properties - Synthetic Study.
A pure synthetic study is presented for feasibility to estimate the hydraulic parameters. A synthetic as well as a measured upper boundary condition is analyzed for the three instrument descriptions - the point sensor, TDR and GPR. In all cases, the infiltration flux is scaled. That changes the estimation condition so that for high fluxes, the estimation error for the parameters is very accurate. A transition from feasible to non-feasible accurate estimations can be observed scanning over the given soil types.

Chapter 6: Applications of Estimation and Projection to Measured TDR-Data.
The found conditions for a feasible accurate estimation are tested at a measured data set of soil moisture. The estimation shows a sensitive behavior due to measurement errors as well as model simplifications. The water content is projected into depth using the finial estimated parameters. The non-linearity of the model prevents a projection into depth around layer interfaces due to ponding water in the measurements.

2. Projection Methods of Soil Temperature

The aim of projection methods in soil is generally to use time-series at the surface or close to it and transport these into depth. From this we can get information in the soil without taking direct measurements in greater depth, e. g. digging holes. On the other hand the soil material properties can be deduced by parameter estimation: If we have available additional time-series in the depth we can estimate an effective parameter of the soil.
In a first step, we focus in energy transport in the soil and build a mathematical model of the temperature. We take a closer look with that at two projection approaches using the transfer function and the Fourier transform. They are based on two extremes of orthonormal systems in the time domain: On one hand, the family of δ-functions $\{\delta(\cdot - t_k)\}_k$ and on the other hand the family of complex exponential functions $\{\exp(\mathrm{i}\omega_k \cdot)\}_k$. Localization and global expansion in time characterize these two approaches respectively.
In the following the heat equation is introduced together with initial and boundary conditions under certain assumptions. Then the projected time-series is given by the analytical solution of this mathematical model. This solution is described by transfer function and in the other approach by periodic functions using Fourier transform. Mathematical tools are applied to get rid of a more realistic description of the initial conditions for each approach.

2.1. General Model

The first step to determine a projection relation for soil temperature is to formulate a mathematical model which describes the dynamics of heat transport in a soil profile. The basic description of the thermal dynamics forms the general heat conduction equation

$$\partial_t \left[C_h \left(\theta \right) \, T \right] - \nabla \cdot \left[K_h \left(\theta \right) \, \nabla T \right] = 0 \qquad \text{in} \quad \Omega_{\mathbf{x}} \times \Omega_t \tag{2.1}$$

with soil temperature T. The source term is neglected. This linear differential equation is based on two main principles, the conservation of thermal energy and an empirical flux law, also known as Fourier's law. This general form in the space $\Omega_{\mathbf{x}} \times \Omega_t$ is characterized by a time derivative and spatial derivatives with respect to the three spatial directions. Furthermore, it is important to take into consideration that the soil properties, the thermal capacity C_h and the thermal conductivity K_h, in general depend explicitly not only on time t and space $\mathbf{x}$, but also vary with the soil water content θ.
Due to the dependence of the soil properties on the soil water content, equation (2.1) is generally not analytically solvable. Thus we consider special cases. One important assumption is that the soil should be homogeneous and therefore, the soil properties C_h and K_h are no longer dependent on space $\mathbf{x}$. We further assume that the soil properties do not change within the time domain. Additionally, the soil water content θ is considered to be constant during projection in time and space. Finally we end up with the special heat equation, a parabolic partial differential equation,

$$\partial_t T - D_h \Delta T = 0 \tag{2.2}$$

and define the effective thermal diffusivity

$$D_h = \frac{K_h}{C_h}\,. \tag{2.3}$$

The final simplification yields an one-dimensional, differential equation in depth z. This implies that the main axis of temperature propagation is the z-axis. We can neglect the other two directions, since the main issue in our case is the depth. It is assumed isotropic in horizontal direction. Hence, we finally receive

$$\partial_t T - D_h \partial_z^2 T = 0 \qquad \text{in} \quad \Omega_z \times \Omega_t\,. \tag{2.4}$$

Usually, we count the diffusivity D_h in m^2/s, the temperature T in $°C$, the depth z in m and the time in s if the unit is not specified otherwise. In this context we assume to have no phase transition due to freezing and thawing.
Now we have to complete our model, which consists of the special heat equation and the following conditions. The upper boundary condition is constituted by the soil temperature T_0 of the surface. At the lower boundary the temperature vanishes in increasing depth. Therefore, the spatial domain is the lower half-space $[0, \infty)$. It is specified by the positive values of z. Later, we will extend the following model with an initial condition

$$\begin{aligned} \partial_t T - D_h \partial_z^2 T &= 0\,, \\ T(0,t) &= T_0(t) \qquad \forall t \in \Omega_t\,, \\ \lim_{z\to\infty} T(z,t) &= 0 \qquad \forall t \in \Omega_t \end{aligned} \tag{2.5}$$

with a spatial and a temporal domain $\Omega_z := [0, \infty)$ and $\Omega_t := [-\infty, \infty]$.
We solve this model to get an analytical expression of the projected temperature in the soil. This step will be the focus of the next section, using the *transfer function*, also known as *Green's function*, before the Fourier approach will be discussed afterwards.

2.2. Projection by Transfer Function

Starting from the mathematical and physical model of thermal energy transport in soils we take a closer look at the theory of soil temperature projection and how we can combine it with the transfer function approach. The general concept is also shown in *Roth et al. (2001)*.

2.2.1. General Concept

For the transfer function concept, we completed the model (2.5) with a constant initial condition and set this to zero, i. e.

$$T(z,0) = 0 \qquad \forall z \in \Omega_z\,. \tag{2.6}$$

The main idea for solving model (2.5) is to decompose the surface temperature $T_0(t)$ into a basis of δ-functions in time

$$T(0,t) = \int_{-\infty}^{t} T_0(\tau)\,\delta(t-\tau)\,\mathrm{d}\tau\,. \tag{2.7}$$

At each point in time the temperature is represented as a convolution of an amplitude function, the surface temperature $T_0(\cdot)$, and a normalized basis function, the δ-function $\{\delta(\cdot - \tau)\}_\tau$.

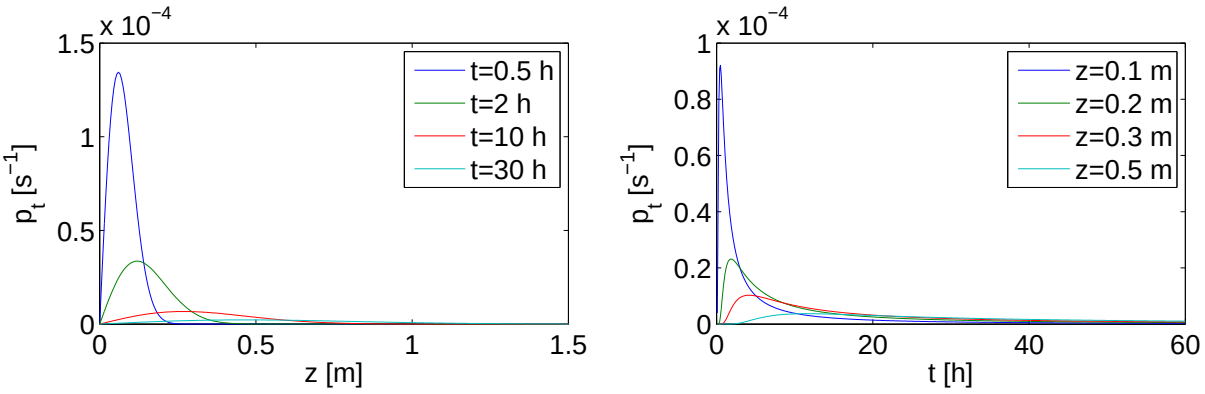


Figure 2.1.: Temporal evolution of the travel time pdf p_t (2.10) with unit $[p_t] = \mathrm{s}^{-1}$ for diffusivity $D_h = 10^{-6}\,\mathrm{m}^2/\mathrm{s}$ at time $t = \{0.5, 2, 10, 30\}\,\mathrm{h}$ in increasing time order: blue, red, yellow, green

Approach (2.7) shows that we can solve a special form of our model analytically by combining equation (2.7) with model (2.5). The transfer function p_t solves the resulting mathematical model

$$\begin{aligned} \partial_t p_t - D_h \partial_z^2 p_t &= 0\,, & & \\ p_t(0,t) &= \delta(t) & \forall t \in \Omega_t\,, \\ p_t(z,0) &= 0 & \forall z \in \Omega_z\,, \\ \lim_{z\to\infty} p_t(z,t) &= 0 & \forall t \in \Omega_t \end{aligned} \tag{2.8}$$

and is called *travel time pdf* (probability density function). This pdf describes the thermal dynamic of a single δ-pulse propagating through the soil. Due to linearity of the model (2.8) every solution of (2.5) can be expressed by this kernel p_t. Therefore, we can solve representatively model (2.8) for all other upper Dirichlet boundary functions $T_0(\cdot)$.
Hence, a Laplace and Fourier transform in space and time yields the solution

$$p_t(k,\omega) = \frac{1}{\mathrm{i}\,\omega + D_h k^2} \tag{2.9}$$

0 in the corresponding k- and ω- domains with complex unit i. When we transform solution (2.9) back to normal space and time and calculate the travel time pdf (Figure 2.1), it gets the form

$$p_t(z,t) = \frac{z}{2\sqrt{\pi D_h t^3}} \exp\left(-\frac{z^2}{4D_h t}\right). \tag{2.10}$$

The last step to determine the projected soil temperature is to combine surface temperature T_0 and travel time pdf p_t. This can be done by weighting each travel time pdf with the corresponding surface temperature and integrating over the whole past due to causality, i. e. no temperature signal of the future can influence the one of the past. In mathematical terms this yields a convolution of the function T_0 and p_t

$$(T_0 * p_t)(z,t) = \int_{-\infty}^{t} T_0(\tau)\, p_t(z, t-\tau)\,\mathrm{d}\tau \tag{2.11}$$

as the solution of the model (2.5).
Moreover, for applications of the projection on measurement data we have to consider that the

measured time-series is finite. Therefore, we restrict the convolution (2.11) to the finite time interval $[0, t_f]$. Additionally, we have to deal with an initial condition which is not equal zero. Due to the linearity of the system, we can extent the expression (2.11) to a solution of the mathematical problem with a non-zero initial condition by an additional term $T_{\text{init}}(z,t)$. In a first approximation, we choose the initial measured temperature $T(z,0)$ in the depth. This yields

$$T_{\text{proj}}(z,t) = T(z,0) + \int_0^t T_0(\tau)\, p_t(z, t-\tau)\, \mathrm{d}\tau \tag{2.12}$$

for a given homogeneous, time and space independent medium. Later, we discuss in more detail how we can further determine initial soil temperature $T_{\text{init}}(z,t)$. This will become important in the context of estimating soil properties.

2.2.2. Estimation of Effective Soil Thermal Diffusivity

In section 2.2.1 we derived the soil temperature projection for a given soil with effective thermal diffusivity D_h. In the following, we consider the inverse case where the soil properties are unknown and the effective diffusivity D_h has to be estimated. In this case we need a time-series of soil temperatures at projection depth z_p as additional information. For that purpose, we define the residuum $\triangle T$

$$\triangle T(t) := \left(T_{\text{proj}}(t;z) - T_{\text{meas}}(t;z)\right)\big|_{z=z_p} \tag{2.13}$$

as the difference of projected and measured temperature T_{proj} and T_{meas} in projection depth z_p. This objective function is a measure for the approximation error of the model to real data. If this error is minimal, the assumed soil property is a good representation of the real soil.
A minimization algorithm adapts the thermal diffusivity D_h, but additionally, the initial part $T_{\text{init}}(z,t)$ of the projected temperature, approximated by a constant in time, the reference temperature T_{ref°. If only one temperature sensor in depth is assumed to be given, the measured initial temperature $T(z_p,0)$ at projection depth is rough but practicable first estimation of $T_{\text{init}}(z,t)$. Hence, we assume that the reference temperature is close to $T(z_p,0)$. Both the effective diffusivity D_h and the reference temperature T_{ref} will be estimate in the following minimization problem

$$\begin{aligned} &\min_{D_h, T_{\text{ref}}} \parallel \triangle T \parallel_2^2, \\ &\triangle T(t) = T_{\text{proj}}(t; z_p, D_h, T_{\text{ref}^\circ}) - T_{\text{meas}}(t; z_p) \qquad \forall t \in \Omega_t \end{aligned} \tag{2.14}$$

with the L_2-norm $\parallel \cdot \parallel_2$. This approach of least-squares over the approximation error is founded on the concept of a maximum likelihood estimator, i. e. the search of a most likely parameter set which approximates the given data set in L_2-sense.
The projected temperature is given by equation (2.12), and the measured temperature in depth z_p has to be an input quantity. We can use this optimization problem to find the value for the effective diffusivity D_h that fits best the measured temperatures and therefore the given soil. For solving this optimization problem, we applied the Levenberg-Marquardt algorithm implemented in *MATLAB (2008)*.

2.2.3. Initial Condition

In the following, we focus on the initial condition of the model in comparison to measurements. The current model (2.5) implies a zero initial condition, i.e. at initial time $t_0 = 0$ the soil

Figure 2.2.: Spin-up: the two main parts of the time interval, spin-up phase and minimization time interval.

temperature is assumed to be zero in each depth. If we deal with real soils, i.e. measurement data, this condition is very unrealistic. Real soils always contain a certain amount of thermal energy. Due to this fact, the soil has a certain temperature distribution over depth z which is not equal to zero.
Thus, in a more general case, the initial soil temperature distribution has to be adapted so that the distribution can change by varying the input values of the projection function. This variation can occur in the values of position and length of the time interval, of different projection depths or of different soils. A spin-up phase, an additional z-integral or a combination of both realize this adaption.

Spin-up

First, we consider the approach of a spin-up phase. Therefore, we divide the time interval $[0, t_f]$ into a spin-up phase $[0, t_s]$ and an estimation interval $[t_s, t_f]$ (Figure 2.2). Then we project during both intervals $[0, t_s]$ and $[t_s, t_f]$. However, the minimization uses only the second interval $[t_s, t_f]$ to adapt the effective diffusivity D_h and reference temperature $T_{\mathrm{ref}\circ}$. With such a spin-up phase we get an initial calculated temperature profile at intermediate time t_s which is different from zero and depends on spin-up time t_s, projection depth z_p and the thermal diffusivity D_h.
One important parameter for the accurate representation of a real initial condition is the spin-up time t_s. Here, we can estimate a limit for its values. We take a deeper look on a single δ-pulse and its transfer function in order to explore a suitable spin-up time for this projection method. During propagation into depth, this single δ-pulse is smeared and damped as given from our diffusion model (Figure 2.1). Therefore, the thermal energy started on top reaches a certain depth with different intensities over time and describe this thermal energy intensity caused by a δ-pulse in space and time as the integral of the corresponding travel time pdf over time:

$$I_{\mathrm{E}}(z,t) = \int_0^t p_t(z,\tau)\,\mathrm{d}\tau\,. \tag{2.15}$$

Here, one should note that the limit of this integral for infinite time is normalized in each depth, mathematically due to the role of p_t as an orthonormal system, and physically due to conservation of energy and pure diffusion model. Therefore, the percentage of arrived thermal energy in a certain depth is represented by this integral I_{E} and yields information about reliability of the projection results at the end of the spin-up phase, shown in Figure 2.3.
If we count the minimal time, when $I_{\mathrm{E}}(z_p, t) \geq PL$ for a given projection depth z_p, percentage limit PL and soil, this $t_{\min}$ is our measure for the quality of the chosen spin-up phase. This minimal time is drawn in Figure 2.3 against depth z. For instance, if we project into 1 m depth for a homogeneous soil with $D_h = 10^{-6}\,\mathrm{m^2/s}$, then 50 % of the initial surface pulse has reached 1 m after $t_{\min} = 12\,\mathrm{d}$ (red cross), 60 % after 21 d (blue cross) and 70 % after 39 d (blue cross). Observation shows that a limit lower than $PL = 50\,\%$ is not practicable. Points of time, which include more 70 % energy, are too many. Then the calculation time of the spin-up phase is

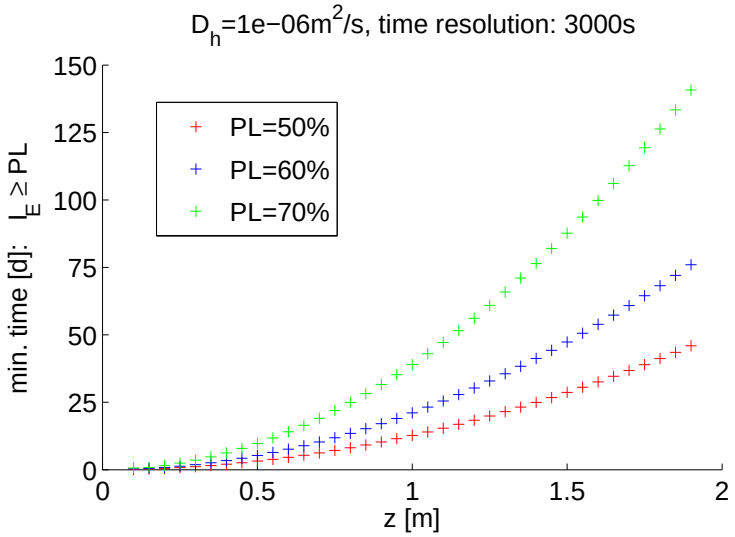


Figure 2.3.: Estimation of the spin-up duration: To depth z the minimal time is counted when the integrated travel time pdf reaches a given percentage limit PL of the initial total energy. Here, the thermal diffusivity is $10^{-6}\,\mathrm{m}^2/\mathrm{s}$ and the integral is numerically evaluated each 3000 s.

numerically too expensive due to the evaluation of the convolution integral in (2.12).
An alternative for a very rough estimation for the lower limit of t_s is based on the evolution of the pdf maximum. It reaches depth z_p at time

$$t_{\max} = \frac{z_p^2}{6D_h}\,, \tag{2.16}$$

but with at most 50% of the thermal energy of a δ-pulse. Therefore $t_{\max}$ is too rough as lower bound.

z-Integral

A second approach to include a more realistic initial condition is based on an additional integral. The idea is to determine the thermal energy in the initial profile (gray hatched area in Figure 2.4 left) and its temporal evolution.
To create a mathematical model for this case, the same approach is chosen as before. Thus we decompose the temperature profile at initial time $t_0 = 0$ into δ-functions in z and realize the model on the basis of the method of mirror charges (Figure 2.4 right) in the following way

$$\begin{aligned} \partial_t p - D_h \partial_z^2 p &= 0\,, & & \\ p(0,t) &= 0 & &\forall t \in \Omega_t\,, \\ p(z,0) &= \delta(z - z_i) - \delta(z + z_i) & &\forall z \in \Omega_z\,, \\ \lim_{z\to\infty} p(z,t) &= 0 & &\forall t \in \Omega_t\,. \end{aligned} \tag{2.17}$$

At the surface, the temperature is equal to zero. At initial time one δ-pulse is located in depth z_i and one in $-z_i$ with an opposite sign to provide the surface condition. Model (2.17) can be solved independently from model (2.8), since the heat conduction is a linear process and can be added later due to the linearity of our model description.

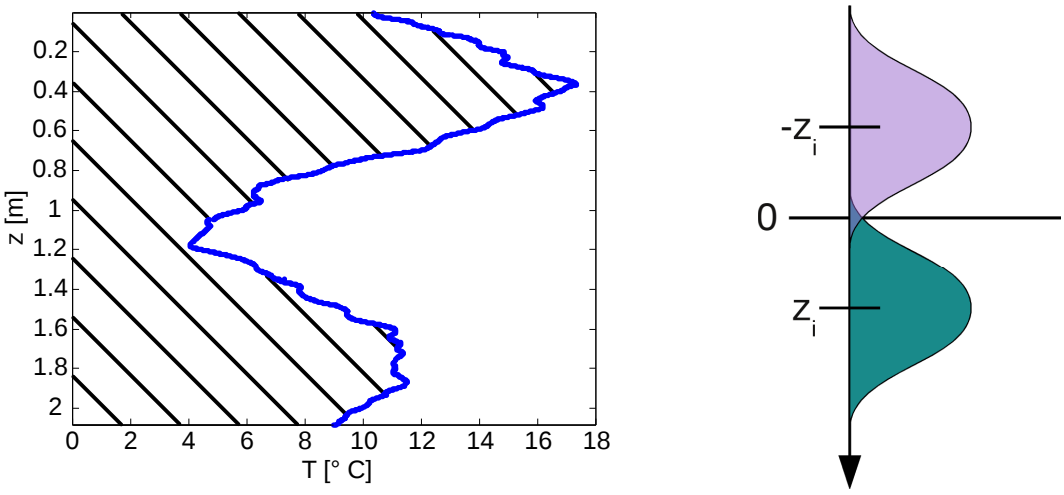


Figure 2.4.: Left: Sketch of a possible initial soil temperature $T_{\text{init}}(z;0)$. Blue: Initial soil temperature. Grey hatched area: Amount of thermal energy which is within the soil at initial state. Right: Idea to solve the initial value problem (2.17) using the concept of mirror charges: A δ-pulse is placed at depth $z = z_i$ and its corresponding mirrored pulse at $z = -z_i$ with an opposite sign (different colors), i. e. the two pulses vanish at the surface $z = 0$ to zero. Image by *Ludin (2010)*.

Again, by applying the Fourier transform in space and time we end up with the solution in k- and ω-domain

$$p_\delta^{\text{IC}}(k,\omega,z_i) = \frac{\exp(-kz_i) - \exp(kz_i)}{\mathrm{i}\omega + D_h k^2} \tag{2.18}$$

and in z- and t-domain

$$p_\delta^{\text{IC}}(z,t,z_i) = \frac{1}{2\sqrt{\pi D_h t}} \left[\exp\left(-\frac{(z-z_i)^2}{4D_h t} \right) - \exp\left(-\frac{(z+z_i)^2}{4D_h t} \right) \right] . \tag{2.19}$$

This pdf yields the temporal and spatial expansion of a δ-pulse in a soil which starts at depth z_0 and interacts with the surface $z = 0$. The temporal propagation of this pdf and the dependency with respect to z_0 is shown in Figure 2.5. The pdf (2.19) fulfills the node boundary condition at the surface in (2.17). Physically, this means that thermal energy runs through the surface and is extracted from the soil.

Now we weigh the pdf with the initial soil temperature and integrate over depth in the lower half-space. Then this integral gives the amount of thermal energy which is already in the soil at initial time t_0 and propagates through it. This results the final expression for the total projected temperature expression

$$T_{\text{proj}}(z,t) = T_{\text{ref}} + \int_0^t T_0(\tau)\, p_t(z, t-\tau)\, \mathrm{d}\tau + \int_0^\infty T(\xi,0)\, p_\delta^{\text{IC}}(z,t,\xi)\, \mathrm{d}\xi \tag{2.20}$$

with a reference temperature T_{ref}. The temperature $T_{\text{init}}(z,t)$ is replaced here by the new integral and a constant term T_{ref}. This term T_{ref} is usually close to zero, but for the numerical evaluation of the projected temperature this term estimates an offset coming from the approximation error.

The spin-up phase and the z-integral realize that an initial condition, which is different from zero and more realistic, extend the model.

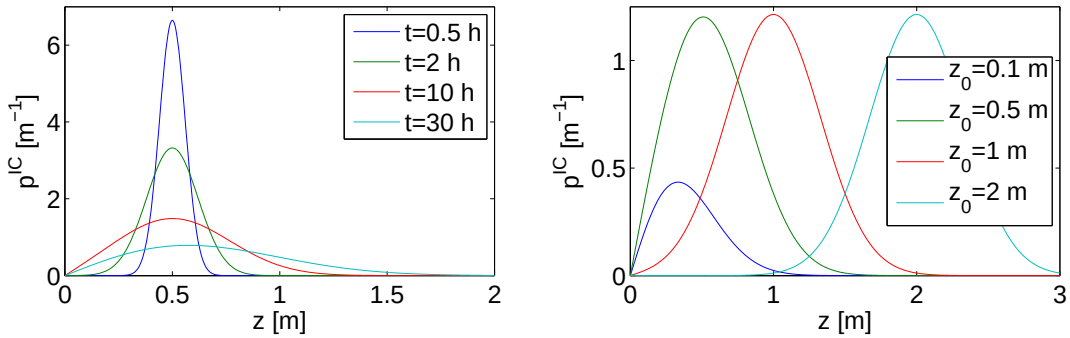


Figure 2.5.: Travel distance pdf p_δ^{IC} (2.19) with unit m^{-1} for diffusivity $D_h = 10^{-6}\,\mathrm{m}^2/\mathrm{s}$. Left: Temporal evolution with $z_0 = 0.5\,\mathrm{m}$ and time $t = \{0.5, 2, 10, 30\}\,\mathrm{h}$. Right: Different position of δ-pulse $z_0 = \{0.1, 0.5, 1, 2\}\mathrm{m}$ at time steps $t = 15\,\mathrm{h}$. Both in increasing order: blue, red, yellow, green.

In the following, we focus in more detail on how the pdf behaves when evaluated at a temporal grid.

Temporal Resolution in Small Depth

In practice, the time-series is measured at a temporal grid of some ten minutes or hour intervals. Therefore, the travel time pdf p_t in (2.10) is evaluated at the same temporal resolution as the convolution of surface temperature and travel time pdf

$$(T_0 * p_t)(z,t) = \int_{t_0}^{t} T_0(\tau)\, p_t(z, t-\tau)\,\mathrm{d}\tau\,. \tag{2.21}$$

Numerically, we resolve the integral by the trapezoidal rule (Figure 2.6) at a time grid $t_0, \ldots, t_i, \ldots, t_{N_t}$ and therefore it follows approximately for the convolution (2.21) at one temporal grid point t_M

$$(T_0 * p_t)(z,t_M) \approx \sum_{i=0}^{M-1} \frac{h_i}{2}\left(T_0(t_i)\, p_t(z, t_M - t_i) + T_0(t_{i+1})\, p_t(z, t_M - t_{i+1})\right) \tag{2.22}$$

with step size $h_i := t_{i+1} - t_i$. Figure 2.6 shows the travel time pdf in time with a temporal resolution of one hour. Usually, this resolution is fine enough to represent the pdf in an adequate way. However, in case of small depth, this resolution is too coarse to describe the peak of p_t (Figure 2.6 right). This can influence the correct weighting of p_t with the surface temperature, since the numerical evaluation of the integral (Figure 2.6 left) cannot represent the integral with a coarse resolution that is sufficiently accurate.
Now, the idea is that instead of evaluating the function

$$f(\tau) := T_0(\tau)\, p_t(z, t_M - \tau) \tag{2.23}$$

at the grid points t_i for the trapezoidal rule, we first interpolate the time-series of temperatures linearly at each neighboring interval $[t_i, t_{i+1}]$ and then solve the convolution integral analytically at each of these intervals. This means that we first take one interval of two neighboring time points t_i and t_{i+1} and interpolate a straight line through these two temperatures. It results

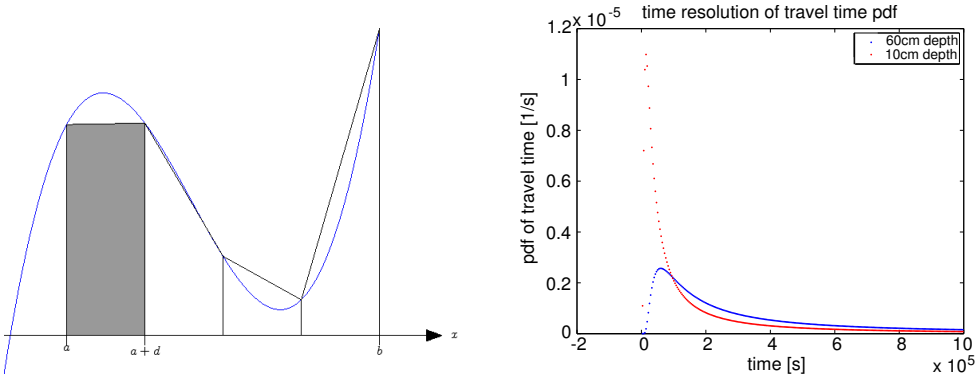


Figure 2.6.: Left: Sketch of integration of a function using trapezoidal rule. Image source: http://mo.mathematik.uni-stuttgart.de/inhalt/aussage/aussage4/ Right: Travel time pdf p_t as function of time for projection depth 0.6 m and 0.1 m and nearly 300 time points with a step size of one hour. For smaller projection depth the resolution of the peak of p_t decreases. Image adapted from *Ludin (2010)*.

that the gap between these two measurement points is filled up. In this particular interval $[t_i, t_{i+1}]$ the surface temperature T_0 has now the form

$$T_0(t) = m_i(t - t_i) + T_i \tag{2.24}$$

for all $t \in [t_i, t_{i+1}]$ with slope $m_i := (T_{i+1} - T_i)/(t_{i+1} - t_i)$ and temperature $T_i := T_0(t_i)$. Furthermore, we insert expression (2.24) in the convolution integral (2.21) and split up the integral from t_0 to t_M to a sum of integrals in the following way

$$(T_0 * p_t)(z, t_M) = \sum_{i=0}^{M-1} \int_{t_i}^{t_{i+1}} [m_i(\tau - t_i) + T_i]\, p_t(z, t_M - \tau)\, \mathrm{d}\tau\,. \tag{2.25}$$

The integral of the expression (2.25) can be solved analytically together with expression (2.10), and we end up with

$$\begin{aligned}(T_0 * p_t)(z, t_M) = \sum_{i=0}^{M-1} \frac{z}{\sqrt{\pi D_h}} m_i \left[\sqrt{\kappa_{M,i+1}} \exp\left(-\frac{z^2}{4D_h\kappa_{M,i+1}}\right) - \sqrt{\kappa_{M,i}} \exp\left(-\frac{z^2}{4D_h\kappa_{M,i}}\right)\right] \\ + \left[z^2(T_i + T_{i+1}) + 2D_h(t_M(T_i - T_{i+1}) + t_iT_{i+1} - t_{i+1}T_i)\right] \frac{g(t_{i+1}) - g(t_i)}{2D_h(t_i - t_{i+1})}\,,\end{aligned} \tag{2.26}$$

where $\kappa_{M,i} = t_M - t_i$ denotes the time difference.

$$g(\tau) := \operatorname{erf}\left(\frac{z}{2\sqrt{D_h(t_M - \tau)}}\right) \tag{2.27}$$

is the Gaussian error function. Equation (2.26) represents the analytical solution of the convolution integral by using piecewise interpolation between each measurement point (t_i, T_i). This solution should be use at smaller depths, where the resolution of the peak of travel time pdf p_t is very coarse.

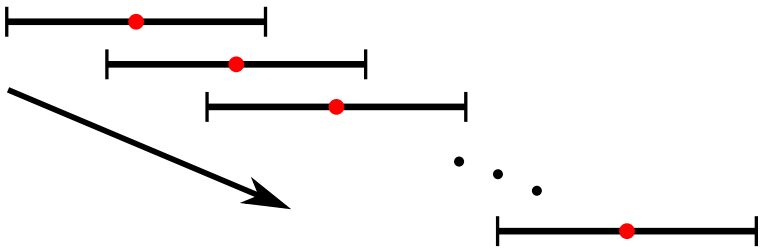


Figure 2.7.: Duplication and shifting of time intervals $I_1, \ldots, I_{N_I}$: All intervals I_i have the same number of points and different starting values. The arrow denotes the direction of the temporal shift and the red dot the mean value of each interval.

Since the evaluation of the error function is very expensive in computational time, we approximate it by a rational and exponential expression (*Abramowitz (1965)*)

$$\operatorname{erf}(x) \approx 1 - \left(a_1 y + a_2 y^2 + a_3 y^3\right) e^{-x^2} \tag{2.28}$$

with $y = (1 + px)^{-1}$, $p = 0.47047$, $a_1 = 0.3480242$, $a_2 = -0.0958798$, $a_3 = 0.7478556$ and an approximation error of the error function less than 2.5×10^{-5}.
Summarizing, we could find an analytical approximation of the transfer integral. Due to the Gaussian error function, this approximation needs a lot of computational time, but it has the advantage that the accuracy of the projection is improved.

2.2.4. Temporal Evolution of Effective Thermal Diffusivity

The real effective thermal diffusivity D_h, can vary in time due to changes of soil moisture, e. g. precipitation, evaporation or freezing. However, we assume that there is no temporal variation in the model of the soil properties within the time interval of interest. Therefore, we concentrate on shorter time-series and hence reduce the error due to this assumption. Thus, we follow the idea of constant soil properties in shorter intervals, but they may vary between different intervals.
This way, we study variations in the effective diffusivity D_h with respect to time. Afterwards, we take a closer look at the accuracy of the estimations and projections.

At first, we need a small time interval at the beginning of a larger time-series. This first time interval I_1 should have at least a length of three days without a spin-up phase, i.e. large enough that the signal has some structure in it like a daily cycle. The minimal interval length is an empirical value and depends strongly on how dense the measured values are and on how significant the structure varies.
When the vector length of spin-up and estimation time interval I_1 is chosen for fixed depth z_p, I_1 will be duplicated N_I times and produces time intervals $I_1, \ldots, I_{N_I}$. Each of these intervals I_i is shifted in time, i.e. they have all the same vector length, but different starting values (Figure 2.7). The shift is measured in terms of the number of points in the time-series which refers to a certain duration. If the total time intervals overlap, the temporal resolution will be more dense. This overlap can occur within the estimation as well as within the spin-up interval. Due to the overlap, local measurement artifact like data gaps or outliers within measurement data influence more than one interval and can lead to larger variations of the estimated parameter D_h.
Furthermore, we represent the total time interval in the plot of temporal evaluation as the

mean value of all time values with the interval I_i. This corresponds to the red points in Figure 2.7.

Accuracy of Estimation

A next issue is to have a measure for the quality of a projection, i. e. which strategies are available to compute how close the projection is to the measured data.
One approach needs very little computational time since the value χ^2 is used and calculated anyway. The usual χ^2-value

$$\chi^2 = \sum_{i=1}^{N_t} \left(\frac{T_{\text{proj}}(z_p, t_i) - T_{\text{meas}}(z_p, t_i)}{\sigma_i} \right)^2 \tag{2.29}$$

with constant variance $\sigma_i^2 = 1$ is the objective function value using the Levenberg-Marquardt algorithm implemented in *MATLAB (2008)*. Further, for each optimized parameter we calculate the reduced χ^2_ν-value

$$\chi^2_\nu = \frac{\chi^2}{\nu}\,, \tag{2.30}$$

which is also based on the method of least squares. This scaled value is comparable for all time-series since it is independent on their length just by dividing by the factor ν. The factor $\nu := N_t - m$ describes the number of degrees of freedom and consists of the number of data points N_t and parameters m. Since we estimate the effective diffusivity D_h and the reference temperature, the number of parameters m is equal two.
This approach gives an overview over the whole residuum by averaging. Local, larger peaks of the residuum influence only in weighted way the χ^2_ν-value. The quality of the projection can be measured with χ^2_ν globally, but it does not resolve locally the quality. In general, it is sufficient enough for applications.

Summarizing this section, we showed how we can use the transfer function to project temperature into depth. Moreover, we discussed two approaches to improve the initial distribution of soil temperature, the spin-up phase and the z-integral. For smaller depths, we introduced one possibility to increase the temporal resolution. This was necessary as the pdf was evaluated at a coarser grid than its maximum needs to still be resolvable. The last issue was how we measure the quality of the projection result, and we defined the scale χ^2-value.

2.3. Projection by Fourier Transform

In the last section, we discussed the transfer function method for solving the heat equation. In the following, we are going to take a closer look at another method to project soil temperature into depth. This Fourier transform approach has advantages especially in case of large time-series where periodicity in time can dominate.

2.3.1. General Concept

The basis of the projection is again the model (2.5). In section 2.1 we deduced this model with its conditions and discussed the underlying assumptions.
In contrast to the transfer function approach with the δ-functions in time, we decompose

the surface temperature T_0 into periodic functions, i.e. complex exponential functions, and correspond them to the Fourier transform

$$T(0,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} T_0(\omega)\exp(\mathrm{i}\omega t)\,\mathrm{d}\omega \tag{2.31}$$

with frequency $\omega = 2\pi f$. This representation shows a multiplication of amplitude of the surface temperature and a normalized basis function for the whole frequency domain. Here, the basis is chosen as the set of complex exponential functions.
First, we solve the model (2.5) for one basis function as forcing at the upper boundary due to linearity of the system. We obtain the model

$$\begin{aligned} \partial_t q - D_h\partial_z^2 q &= 0\,, \\ q(0;t) &= \exp(\mathrm{i}\omega t) \qquad \forall t \in \Omega_t \end{aligned} \tag{2.32}$$

with an arbitrary but fixed frequency ω. Separation of variables yields a general solution of the heat equation (2.2)

$$q(z,t) = A\exp(\gamma t)\left[B_1\exp\left(-\sqrt{\gamma/D_h}z\right) + B_2\exp\left(\sqrt{\gamma/D_h}z\right)\right] \tag{2.33}$$

with the amplitudes A, B_1 and B_2 and a factor γ. If we make use of the boundary conditions, then B_2 will vanish due to the lower boundary condition

$$\lim_{z\to\infty} q(z,t) = 0\,. \tag{2.34}$$

In other words the part of soil temperature due to the periodic forcing of equation (2.32) has to vanish if we are deep enough. We can see that the space-dependent factor of the solution has its maximum at the surface $z = 0$ and decreases continuously with depth z to zero.
Moreover, if we consider periodicity in time with frequency ω, the constant γ will become $\mathrm{i}\omega$. Together with expression

$$2\mathrm{i} = (1+\mathrm{i})^2 \tag{2.35}$$

the solution of model (2.32) with its conditions takes the form

$$q(z,t) = \alpha\exp\left(-\sqrt{\omega/2D_h}z\right)\exp\left(\mathrm{i}\left(-\sqrt{\omega/2D_h}z + \omega t\right)\right) \tag{2.36}$$

with an amplitude factor α. This factor also depends on the frequency ω and is proportional to the surface temperature $\tilde{T}_0(\omega)$. Furthermore, we note, that at initial time $t_0 = 0$ the temperature profile is a damping exponential function, in general unequal to zero.
These last discussions were made only for one arbitrary, but fixed frequency ω. Therefore, in the following we consider how the solution changes in case of many different frequencies. Due to the linearity of the heat equation (2.2) we can linearly superpose the solution (2.36) of one periodic function with the ones of other frequencies. In cases of time-series, the indefinite integral in (2.37) is defined and exists. Therefore, we can make use of the linearity of the integral and end up with the solution of the thermal model (2.5) for frequencies $\omega \in \mathbb{R}$ and in projection depth z

$$T(z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{T}(0,\omega)\exp\left(-\sqrt{\omega/2D_h}z\right)\exp\left(\mathrm{i}\left(-\sqrt{\omega/2D_h}z + \omega t\right)\right)\,\mathrm{d}\omega\,. \tag{2.37}$$

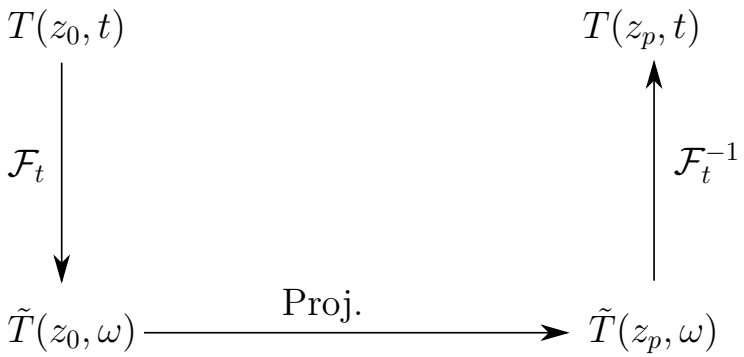


Figure 2.8.: Sketch of the sequence of projection and Fourier transform $\mathcal{F}$ with surface depth z_0 and projection depth $z = z_p$.

Furthermore, the expression (2.37) can also be written in terms of the inverse Fourier transform, denoted by $\mathcal{F}^{-1}$ (Def. C.2),

$$T(z,t) = \mathcal{F}^{-1}\left[\tilde{T}(0,\omega)\exp\left(-\sqrt{\omega/2D_h}z(1+\mathrm{i})\right)\right] \tag{2.38}$$

with the surface temperature also given by the Fourier transform (Def. C.1)

$$\tilde{T}(0,\omega) = \mathcal{F}\left[T(0,t)\right] . \tag{2.39}$$

Expression 2.38 solves the thermal model (2.5). But the initial temperature of the profile need not coincide with the real measured soil temperature. Thus, we introduce a correction function $T_\mathrm{r}(z,t)$ which describes the difference between the real initial temperature profile and the modeled one (by evaluating (2.37) at $t = 0$). Then, we approximate this function by its initial value and use the formulation of the projected temperature

$$T_\mathrm{proj}(z,t) = T_\mathrm{r}(z,t_0) + \mathcal{F}^{-1}\left[\tilde{T}(0,\omega)\exp\left(-\sqrt{\omega/2D_h}z(1+\mathrm{i})\right)\right] . \tag{2.40}$$

This expression describes the soil temperature at depth z in the course of time by decomposition of the surface time-series into periodic functions for a homogeneous soil.

Moreover, the projected temperature in time can be interpreted as a projection in frequency domain as shown in Figure 2.8. Instead of direct projection from top to depth z in time domain, we transform the time-series of surface temperature into the frequency domain . Then the temperature $\tilde{T}(0,\omega)$ is projected into depth z by multiplying the temperature with the complex function as in equation (2.38)

$$\tilde{T}(z,\omega) = \tilde{T}(0,\omega)\exp\left(-\sqrt{\omega/2D_h}z(1+\mathrm{i})\right) . \tag{2.41}$$

Finally, the projected soil temperature is transfered back to the time domain by the inverse Fourier transform.

With this solution, we can now create an optimization problem, which will be discussed in the following section.

2.3.2. Estimation of Effective Soil Thermal Diffusivity

In the last section, we found an expression for the solution of the heat equation with constraints by using the Fourier transform and we interpret this as a projection in frequency domain. Now

we continue with the estimation of the soil thermal property D_h and the reference temperature in a way analogous to section 2.2.2. However, this time we make use of the projection solution (2.38) instead of the transfer integral as before.
The projected temperature should reproduce the measurements in the same depth. Therefore, we search for an effective thermal diffusivity D_h for the heat equation (2.2) and a reference temperature T_{ref} in

$$T_{\mathrm{proj}}(z,t;D_h,T_{\mathrm{ref}}) = T_{\mathrm{ref}} + \mathcal{F}^{-1}\left[\tilde{T}(0,\omega)\exp\left(-\sqrt{\omega/2D_h}z(1+\mathrm{i})\right)\right], \tag{2.42}$$

which fit the model best. We obtain it again by considering the difference between projected and measured temperature. In the ideal case these two time-series are identical and the norm of the difference is zero. Hence, if this norm gets a small value, the projection is said to reproduce the measurement data in depth z_p. Thus, in the same way as in section 2.2.2 the minimization problem is given in the form

$$\begin{aligned} &\min_{D_h,T_{\mathrm{ref}}} \parallel \triangle T \parallel_2^2, \\ &\triangle T(t) = T_{\mathrm{proj}}(t;z_p,D_h,T_{\mathrm{ref}}) - T_{\mathrm{meas}}(t;z_p) \qquad \forall t \in \Omega_t \end{aligned} \tag{2.43}$$

with the L_2-norm $\parallel \cdot \parallel_2$ and the fixed projection depth z_p. Nevertheless, T_{proj} is the projected temperature here determined by the Fourier transform as in expression (2.38). As before, this optimization problem is solved by the Levenberg-Marquardt algorithm implemented in *MATLAB (2008)*.
We end up with an approximated effective value of the thermal diffusivity and a measure for the quality of the approximation. The least squares of the residuum produces the weighted χ^2-quantity by weighting by the number of temporal grid points and estimation parameters. This is helpful for checking the improvements of the projection independent of the time-series length.

2.3.3. Initial Condition

A next step is to take a look at the thermal profile at initial time and examine how it can influence the total approximation error in (2.43).
If we evaluate the projected temperature (2.37) at initial time, we can see that the initial temperature profile is distributed like a superposition of frequency-dependent surface temperatures which are damped and phase-shifted. But a real data set will behave different to it, e. g. not necessary monotonically decreasing with depth. This initial thermal energy will also propagate through the soil profile like in Figure 2.9 by an example of a single initial δ-peak. If such a short pulse starts in a certain depth, the surrounding area will be heated up. In Figure 2.9 a surface sensor can detect it after a certain time, but then the reflection is included into the projection procedure since it contains in the surface time-series.
But what happens with the thermal energy traveling downwards? Again, after 20 min a small amount of energy has reached 20 cm depth. Nevertheless, the reflected part does not yet influence a sensor in 20 cm. Therefore, differences exist between measured and projected time-series due to a realistic initial condition. After a warm-up phase, this error vanishes. Such a local error has a global impact on the time-series and can expressed by in the estimated value of D_h, since the Fourier approach has no chronological order like the transfer approach. By adaption of D_h and T_{ref}, this error is arranged over the whole spectrum. Consequently, an initial interval should be excluded for the minimization problem.

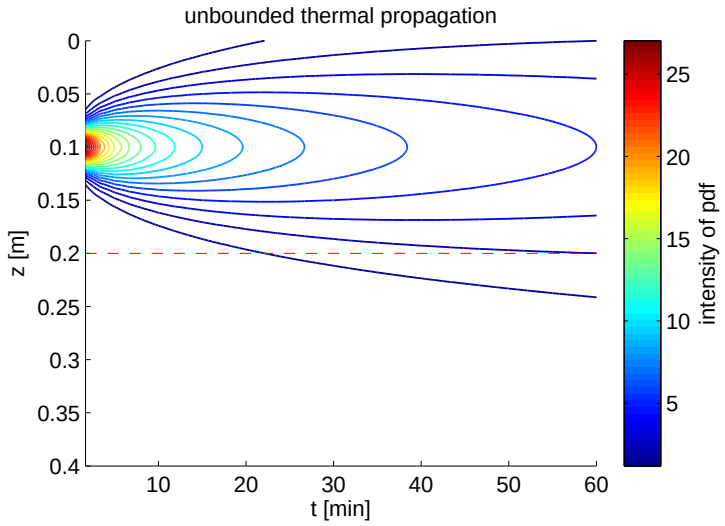


Figure 2.9.: Probability density function of a δ-pulse in 0.1 m depth for an unbounded thermal problem. Equipotential lines are shown within a 0.4 m soil profile with $D_h = 10^{-6}\,\mathrm{m}^2/\mathrm{s}$ during the first 1 h. One equipotential line reaches a depth of 0.2 m (red dashed line) after 20 min and a later one around 1 h.

Alternatively, we could derive the energy input due to the difference between real and modeled initial profile by means of the z-integral. This needs additional computational time and the present initial distribution has to be considered. In a later section, we find another reason why it is useful to deduct at least some starting points. We introduce the discrete Fourier transform in the following to prepare for this aspect.

2.3.4. Discrete Fourier Projection

The continuous description is not useful anymore in real application. Either a measurement unit detects observations in regular and time limited ways or an evaluation of the Fourier transform with a computer is limited to a finite number of data points. Then the infinite integrals of the Fourier transform can be only approximated numerically, except for analytical results of the Fourier transform.
Therefore, a discrete version of the Fourier theory was developed corresponding to the continuous one und is applied in many numerical software tools. Following *Strampp et al. (2003)*, the transition from continuous to discrete case starts with an infinite time domain $\mathbb{R}$, which is then sampled to a finite series t_n. We can build a distribution $\sum_n f(t_n)\delta(t-t_n)$ defined on the whole time domain evaluating a given f at these finite values. Finally, the application of a continuous Fourier transform in t results

$$\mathcal{F}\left(\sum_{n=0}^{N_t-1} f(t_n)\delta(t-t_n)\right)(\omega) = \sum_{n=0}^{N_t-1} f(t_n)\exp(-\mathrm{i}t_n\omega)\,. \tag{2.44}$$

The corresponding sampling in frequency domain with ω_k results in a discrete Fourier transform $\mathcal{DF}$:

$$\mathcal{F}\left(\sum_{n=0}^{N_t-1} f(t_n)\delta(t-t_n)\right)(\omega_k) = \mathcal{DF}(f(t_n))(k)\,. \tag{2.45}$$

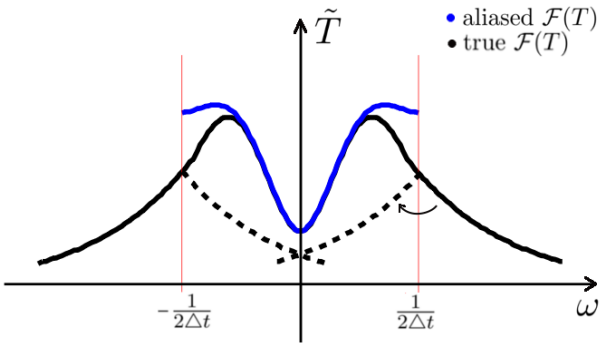


Figure 2.10.: In frequency domain, a true Fourier transformed function is shown in black. Its aliased function is blue in the band-limited domain $[-\frac{1}{2\triangle t}, \frac{1}{2\triangle t}]$. In dashed line, the part of the sample which is falsely transferred into the finite domain.

The time-series is considered as infinitely periodic extended finite series in this discrete description with sample index k. Thus, some additional aspects for the Fourier transform arise in general, but also in the context of our projection.

Sampling

In case of a finite time-series, a finite upper boundary for the frequency exists. This critical frequency is defined as one half of the inverse of the constant temporal grid size (*Press et al. (1988)*). Therefore, the frequency domain is reduced to $f \in [-f_c, f_c]$ with the so called *cutoff frequency*

$$f_c = \frac{1}{2 \triangle t} = (1\,\mathrm{h})^{-1} \ , \tag{2.46}$$

since soil temperature of our data set is measured every 30 min. We can only reproduce signals with a frequency lower than the cutoff frequency. This means that if the underlying continuous temperature function is band-limited to frequencies $|f| < f_c$, this function can be determined completely by the amplitudes in frequency domain due to the *sampling theorem* (*Press et al. (1988)*). All other signals larger than this frequency f_c, which could belong to real processes, oscillate within two consecutive points in time and cannot be resolved anymore with the sampling frequency $\triangle t^{-1}$. In this case, the so-called *aliasing* (Figure 2.10) becomes important, i. e. a signal with frequency $|f| > f_c$ will be translated in a wrong way into the bounded range. Therefore, it is preferable to look for a given sample whether it is more likely band-limited and in which range. In the application of this work we see in a later chapter that the soil temperature time-series can be assumed to be band-limited with cutoff frequency $f_c = (1\,\mathrm{h})^{-1}$, since the amplitude spectrum is very close to zero for many frequencies in the neighborhood of f_c.
After checking if we mismatch or lose information of the temperature time-series, we focus on the finiteness of samples in theory and praxis. This leads to a correction during the projection procedure.

Finiteness

We were just presented effects which can occur in dependency of the grid size. Now, we take a deeper look at the finiteness of the time-series.

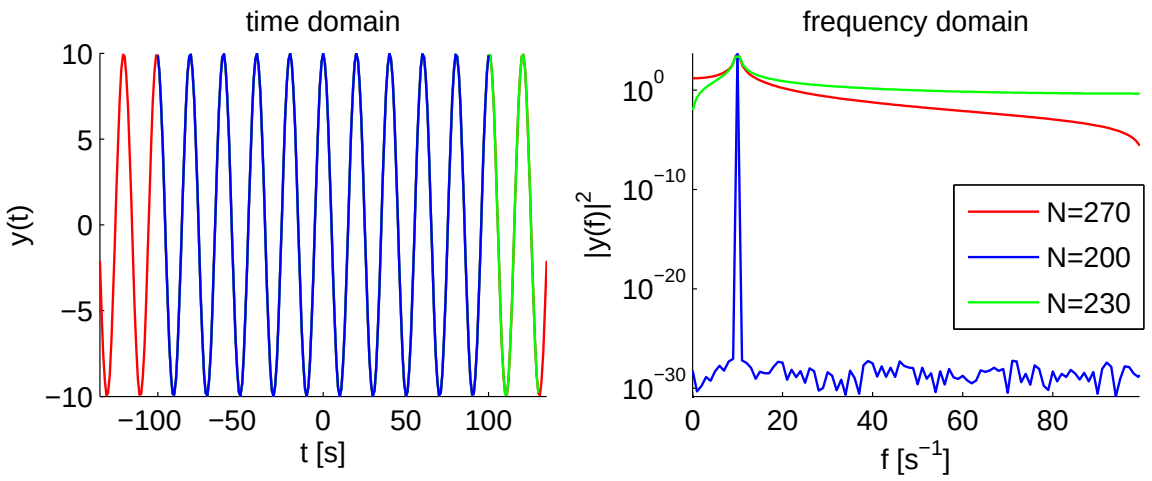


Figure 2.11.: Left: cosine functions with frequency $f = 10\,\mathrm{s}^{-1}$ and different length N. The curves in blue and green are starting at the same time. Right: The corresponding periodograms of the left function in logarithmic scale (see color code).

Each time-series that we receive from measurements as well as synthetic produced data sets has a finite length. The discrete Fourier analysis assumes periodic functions in the finite domain and extended infinitely to the whole domain $\mathbb{R}$. This will produce a discontinuity at the extension points, if these end-points of the finite time-series do not end at the same height. It occurs in case of mismatching signal periods with the time-series duration and it leads to the so-called *leakage*. This means that besides the true frequency entries, additional frequencies exist for reproducing this jump. Leakage does not occur for series, whose length is an integral multiple of their wave length for all frequency entries.
We consider a cosine of three different lengths in Figure 2.11 to illustrate such discontinuity effects using FFT. As expected, the periodogram (right) shows a peak in the logarithmic scale at $f = 10\,\mathrm{s}^{-1}$. The shortest cosine (blue) is periodic in its domain. With the implicit periodic extension of the discrete Fourier transform, an ideal signal to identify true components of frequency signals. Whereas the red cosine ends with same start and end points. This yields much higher entries for the whole frequency domain and again the peak at cosine frequency. But here, the extension produces a kink at the transition point. Higher frequencies are needed to model such a kink with periodic functions. Therefore, their amplitudes increase. On the other hand, the lowest entry for f_0, which corresponds to the mean value of the time-series, becomes higher, since the mean value is not equal or close to zero anymore. Similar to this, the green cosine produces a f_0-entry higher than the blue one, but lower than the green one. This is reasonable, since the green end point does not reach the last minimum. The total mean value gets an entry different to zero, but less than in the red case. However, the green cosine ends with a different value as it starts and causes in FFT a discontinuity which lets the amplitudes of higher frequencies raise much more than for only a kink.

During application of projection, we deal with the Fourier transform and its inverse to temperature time-series. For a clear separation of real frequencies in the time-series, it is important to have a periodic, continuous and differentiable signal, as we learned from the cosine example (Figure 2.11). It is not practicable to select members of a sub-series which fulfill these requirements due to superposition of functions with some frequencies. Therefore, we have to ensure that each true frequency peak will overtop the frequency level of such extension effects at the

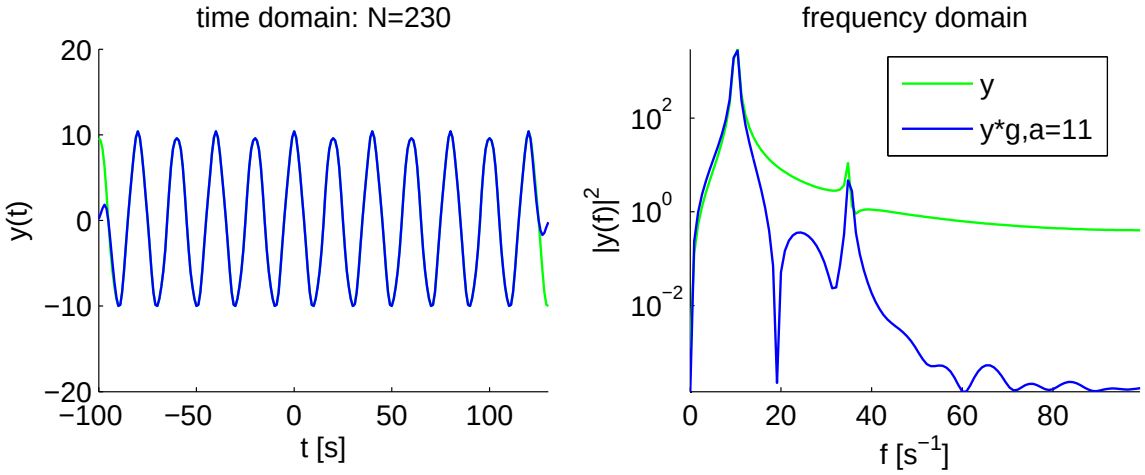


Figure 2.12.: Cosine function y (equation (2.48), green) and with the multiplication of the Tukey window (blue): Left: time domain. Right: frequency domain with logarithmic scale.

bounds and assume the time-series as centered around zero to prevent too high f_0-entries. Consequently, we avoid such jumps and kinks by multiplying the time-series with a window function. A continuous zero-decrease to the sides is required for the window. We assume to has an advantage with a constant part in between. The *cosine-tapered window* fulfills these properties and is known as *Tukey window*:

$$g(t_j) = \begin{cases} \frac{1}{2}\left(1 - \cos\left(\frac{j}{a}\pi\right)\right), & j = 1, \ldots, a, \\ 1, & \text{else}, \\ \frac{1}{2}\left(1 - \cos\left(\frac{N_t+1-j}{a}\pi\right)\right), & j = N_t - a, \ldots, N_t \end{cases} \tag{2.47}$$

with an arbitrary parameter $a \in [1, N_t/2)$. *Schlittgen et al. (1999)* approximates a rectangle window by this cosine-tapered one and compares their properties. In addition to it, *Harris (1978)* compiles some more common windows and summarizes its characteristics.
The influences of the cosine-tapered window on an example time-series

$$y(t) = 10\cos\left(2\pi\, 10\,\mathrm{s}^{-1} \cdot t\right) + 0.4\cos\left(2\pi\, 35\,\mathrm{s}^{-1} \cdot t\right) \tag{2.48}$$

are shown in Figure 2.12. Besides the frequency of $10\,\mathrm{s}^{-1}$, there is a much lower one at $35\,\mathrm{s}^{-1}$. The second frequency is not strongly pronounced after the pure transformation as in the windowed case. The window has the advantage that the neighboring frequencies are damped by some orders whereas the true peak stays closely the same. Thus, using windows the true peaks can be identified more significantly as such, and are not falsely interpreted as an error or noise. In Figure 2.12 the blue curve features the sidelobe structure, typically for the leakage effect. Observations show, the larger the parameter a, i. e. the non-constant area, the more narrow are the sidelobes of the tapered spectrum. The region of the jump adaptation is broadened by the window and increasing of a.
In a next step, we make use of the advantage of the cosine-tapered window for our projection schema und prevent possible errors in the projected soil temperature caused by the finiteness of the series. Before our known projection procedure starts, the surface soil temperature T will be multiplied pointwise by our window g, as pictured in Figure 2.13. We apply the Fourier

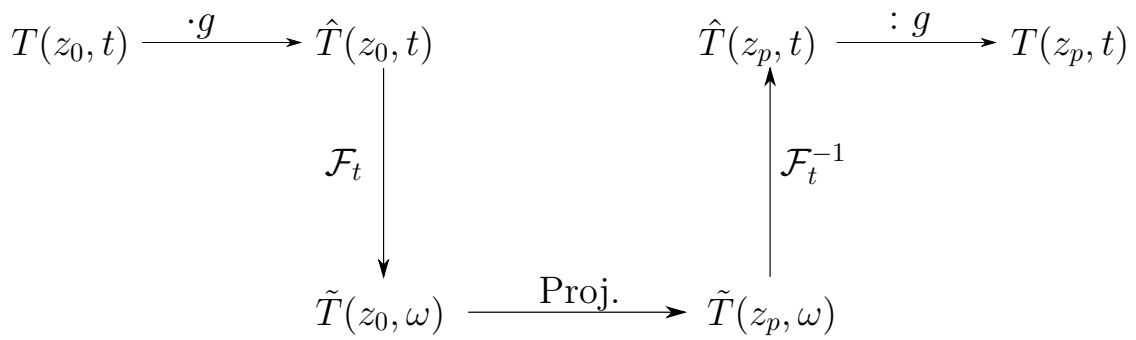


Figure 2.13.: Sketch 2.8 is extended by first applying the window and finally inverting the correction.

transform on this corrected soil temperature $\hat{T}$, the projection in the frequency domain and its inverse Fourier transform. After the inverse Fourier transform, we divide it by the window function and get the projected soil temperature for the signal corresponding to the different starting and ending values.

Under the assumption of $g \neq 0$, the inverting of the window can be done by dividing it by g. In general, the time interval $[t_0 + \triangle t_a, t_f - \triangle t_a]$, where g is constant one, fulfills this condition and we chose this smaller interval for our comparison between projected and measured soil temperature in depth. Then, the factor a gets an additional interpretation as a measure for how many data points are skipped due to the projection procedure.

In summary, we explained the general idea of the projection method by using Fourier transform to project soil temperature in depth. Furthermore, we formulated a minimization problem for estimating mainly an effective value for the soil thermal diffusivity. After discussing the initial condition, we transferred this method to the discrete case and considered same aspects which are important in the context of time-series. The last and for the quality of the projection important step was the correction by the cosine-tapered window.

3. Projection Method of Soil Moisture

Comparable to chapter 2, the goal, in the case the soil moisture, is the projection of a time-series into depth. This means we are searching for a mathematical method to predict the water content deeper in the soil for a given time-series of water content at the surface or close to it. However, in the case of water flow in porous media a major challenge is that the partial differential equation (PDE) used for the description is strongly non-linear due to the coupling of the hydraulic properties with the water content itself. In contrast to the linear system in chapter 2, we were able to simplify the PDE there and to derive an analytical solution. For soil moisture, the non-linear terms are too strong to neglect them. It prevents an analytical description which is flexible to variations of the boundary conditions. Thus, a numerical approach is chosen to simulate to time-series in the depth.
Further, an optimization problem is solved to estimate the hydraulic parameters based on the numerical solution of the PDE. The numerical forward simulation and the parameter estimation are the basic tools to analyze the conditions when it is feasible to estimate the hydraulic parameters accurately. A further question to address is which additional information are essential for the mathematical model to project reliably water content into depth.

3.1. General Model

The dynamic of fluid water in porous media is described by the so-called *Richards equation*

$$\partial_t \theta - \nabla \cdot [K_w(\theta) [\nabla \psi_m - \rho_w \mathbf{g}]] = 0 \tag{3.1}$$

with water content θ, hydraulic conductivity K_w and matrix potential ψ_m. This equation is based on the main principle, *conservation of water volume*, and the empirical law, *Buckingham-Darcy law*

$$\mathbf{j}_w = -K_w(\theta) \nabla \psi_w = -K_w(\theta) [\nabla \psi_m - \rho_w \mathbf{g}] \ . \tag{3.2}$$

In comparison to other flux laws, one should notice that the hydraulic conductivity K_w is a function of water content θ. Referring to this relation and the one of $\theta(\psi_m)$ the strong non-linearity of the Richards equation is represented. More details are explained in *Roth (2007)*.
Further, the Richards equation is a degenerated parabolic PDE and can get effectively hyperbolic by dependency on the material properties, *Ippisch (2009)*.

Parameterizations

Due to the induced non-linearity of (3.1) we have to study the behavior of the soil water characteristic function $\theta(\psi_m)$ and the hydraulic conductivity function $K_w(\theta)$. There are some empirically found parameterizations. *Van Genuchten* and *Brooks-Corey* parameterizations are the most common ones for the soil water characteristic $\theta(\psi_m)$ and *Mualem* for the hydraulic conductivity function which can be combined with the first two parameterizations.
So *van Genuchten (1980)* describes the simplified relation of saturation $\Theta := (\theta - \theta_r)/(\theta_s - \theta_r)$

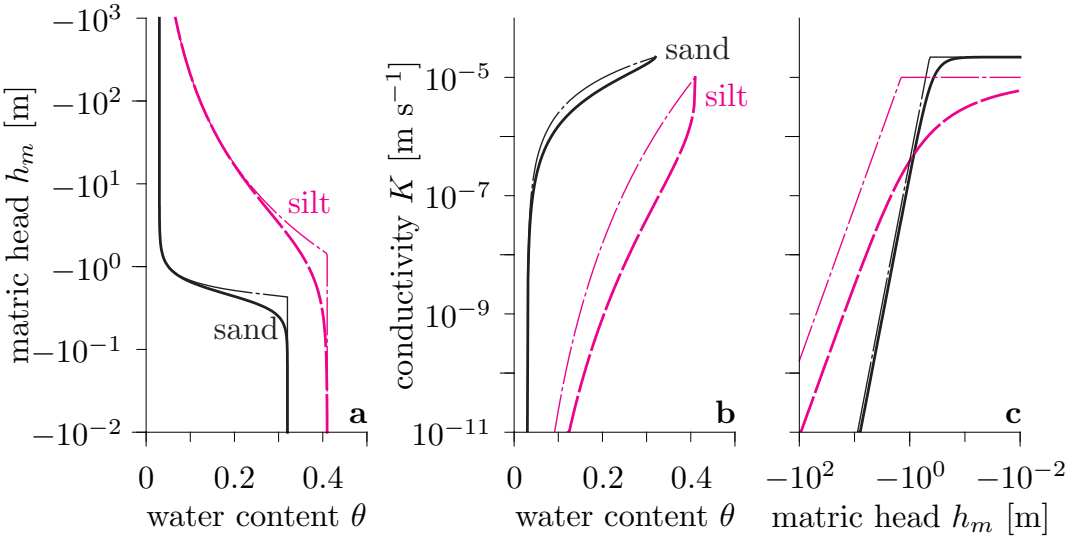


Figure 3.1.: Mualem-van Genuchten (MvG; thick lines) and Brooks-Corey (thin lines) parameterization for two soil types, sand (black) and silt (magenta). a) Soil water characteristic function $\theta(h_m)$. b) Hydraulic conductivity function $K_w(\theta)$ and c) $K_w(h_m)$. The chosen MvG parameters are for sand $(n, \alpha, K_0, \theta_s, \theta_r, a) = (4.17, -2.3\,\mathrm{m}^{-1}, 2.2 \times 10^{-5}\,\mathrm{m/s}, 0.32, 0.03, -1.1)$ and for silt $(n, \alpha, K_0, \theta_s, \theta_r, a) = (1.3, -0.7\,\mathrm{m}^{-1}, 10^{-5}\,\mathrm{m/s}, 0.41, 0.01, 0)$. The figure is taken from *Roth (2007)*.

and matrix head $h_m := \psi_m/(\rho_w g)$ as

$$\Theta(h_m) = [1 + [\alpha h_m]^n]^{-1+1/n} \tag{3.3}$$

with scaling parameter $\alpha < 0$ [m^{-1}], form parameter $n > 1$, residual θ_r and saturated water content θ_s, *Roth (2007)*.

In Figure 3.1 a), the soil water characteristic for two types of soils sand and silt is shown. Both curves vary over some order of magnitudes in matrix head h_m. The solid lines illustrate the continuous van Genuchten curve. The dashed line shows the Brooks-Corey parameterization. In Figure 3.1 a), the Brooks-Corey relation is based on a power-law and has a discontinuity at air entry value $h_m = h_0$ (*Brooks et al. (1966)*). In the following, we chose the parameterization of *van Genuchten (1980)* and note that for lower water contents, both parameterizations are more similar.

Mualem's description of the conductivity together with van Genuchten's description in (3.3) yields the so-called *Mualem-van Genuchten* parametrization

$$K_w(\Theta) = K_0 \Theta^l \left[1 - \left[1 - \Theta^{n/(n-1)}\right]^{1-1/n}\right]^2 \tag{3.4}$$

with saturated conductivity $K_0 > 0$ [cm/h] and parameter l. In Figure 3.1 b), this conductivity function is shown with logarithmic scale. This wide range with its steep slopes together with the ones of $\theta(\psi_m)$ introduces the non-linearity of the Richards equation and depends on the hydraulic soil parameters n, α, K_0, θ_s, θ_r and l.

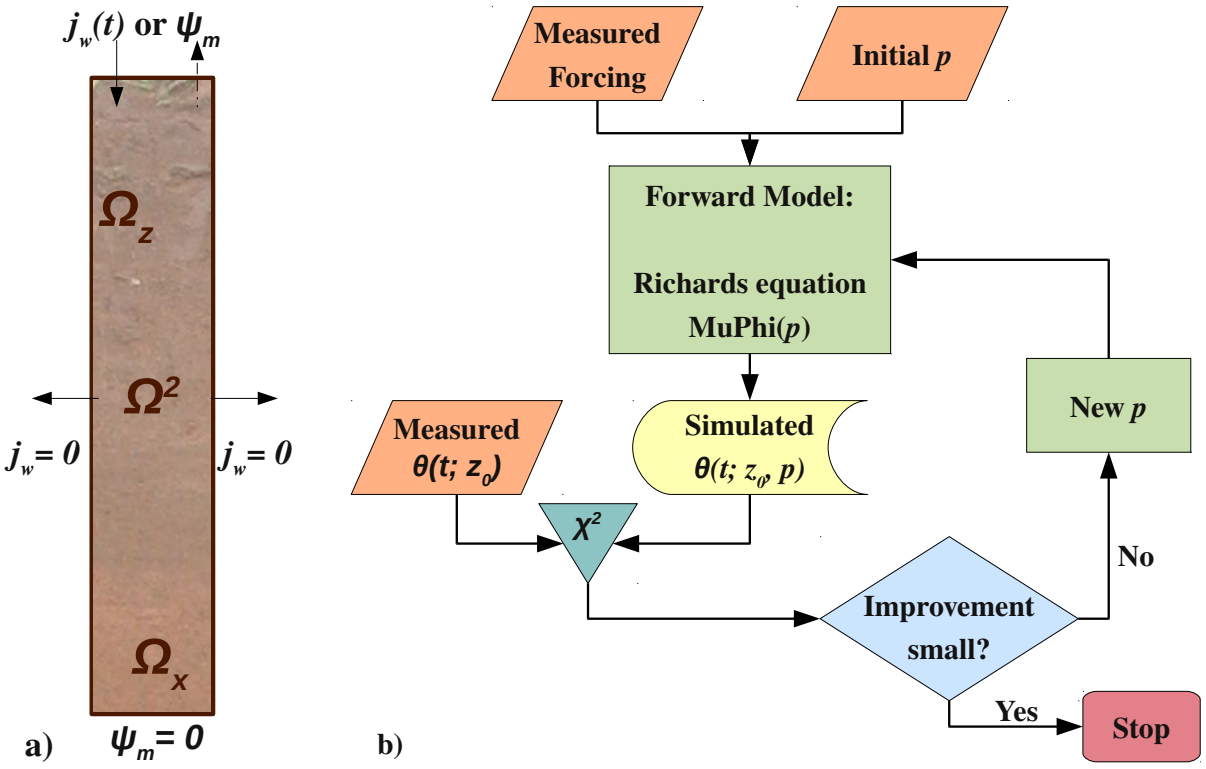


Figure 3.2.: a) Sketch of the spatial domain $\Omega^2 = \Omega_x \times \Omega_z$ and the boundary conditions in model (3.5) b) Flow chart of our input data and inversion procedure for estimating the hydraulic soil parameters p based on measured and simulated soil moisture $\theta(t; z_0)$ close to the surface.

Boundary Conditions

To solve the Richards equation (3.1) numerically, we complete our model with some boundary conditions. In the following, we consider a two-dimensional vertical profile $\Omega^2 := \Omega_x \times \Omega_z$ in the soil (Figure 3.2 a)).
We assume a zero-flux condition (Neumann condition) at the left and right vertical bounds since the soil is assumed to be horizontal isotropic. Hence, the infiltrating water flows through the vertical profile. We lose no water on the left and right.
Moreover, we set the matrix potential (Dirichlet condition) to zero at the lower boundary. This means that the groundwater table is fixed at the lower boundary. Then drainage water flows through the lower bound.
The soil surface is at the top boundary and can have both Dirichlet and Neumann conditions. This depends on the simulation scenario. Infiltration of water is implemented as a Neumann flux, whereas evaporation is realized by a Dirichlet condition with a given matrix potential.

Initial Condition

At initial time, we give the matrix potential ψ_m and hold it horizontally constant. But we assume a linear distribution in vertical direction, i. e. ψ_m is zero at the bottom for the groundwater table. The physical system starts from hydraulic equilibrium.

Model

Summarizing all these conditions, we can create the model which we intend to solve afterwards. The two-dimensional profile $\Omega^2 := \Omega_x \times \Omega_z := [x_l, x_r] \times [0, z_{\text{low}}]$ is counted vertically in x, horizontally in z increasing with depth z and constant in the time domain $\Omega_t := [t_0, t_f]$

$$\begin{aligned}
\partial_t \theta - \nabla \cdot [K_w(\theta)\,[\nabla \psi_m - \rho_w \mathbf{g}]] = 0 \quad & \text{in} \quad \Omega^2 \times \Omega_t\,, \\
\left.\begin{aligned} j_w(x,0,t) = j_{\text{in}}(t)\,, \quad & \text{if infiltration} \\ \psi_m(x,0,t) = \psi_m(t)\,, \quad & \text{if evaporation} \end{aligned}\right\} \quad & \forall t \in \Omega_t\ \forall x \in \Omega_x\,, \\
\psi_m(x, z_{\text{low}}, t) = 0 \quad & \forall t \in \Omega_t\ \forall x \in \Omega_x\,, \\
j_w(x_l, z, t) = 0 \quad & \forall t \in \Omega_t\ \forall z \in \Omega_z\,, \\
j_w(x_r, z, t) = 0 \quad & \forall t \in \Omega_t\ \forall z \in \Omega_z\,, \\
\psi_m(x, z, t_0) = \rho_w g (z - z_{\text{low}}) \quad & \forall (x,z) \in \Omega^2
\end{aligned} \tag{3.5}$$

with the Mualem-van Genuchten parameterizations (3.3) and (3.4).

3.2. Projection of Soil Moisture

Our aim in this section is to find an accurate description of soil moisture in depth by using a time-series of soil moisture $\theta_0(t)$ close to the surface. In our domain of horizontal isotropic soil Ω this time-series is represented as an one point evaluation in time. Later, we will discuss further examples where the time-series is averaged vertically over an area. It has the advantage that measurements coming from TDR or GPR can be represented in the model in a better way. In contrast to the soil temperature case, model (3.5) is not analytically solvable. The non-linearity in the Richards equation does not permit an analytical solution in its general form. Under some assumptions and for special scenarios, analytical solutions have been found, e. g. *Knight et al. (1974)*, *Broadbridge et al. (1988)* and *Warrick et al. (1991)*. But these solutions are too complicated and specialized. Hence, for this work it is more useful to derive a numerical solution which needs less computation time and is more flexible in comparison to the analytical one.
We can directly determine the numerical time-series of soil moisture in a certain depth within Ω with our model (3.5) and knowledge about the soil. Referring to the upper boundary condition in (3.5), we chose the flux of precipitation and the potential of evaporation, respectively. These quantities have to be given together with the near-surface water content from measurement data. The hydraulic soil parameters are estimated from this water content since in general the soil properties are not known.
In summary, this means for the projection procedure of soil water content that if we estimate the hydraulic soil parameters $p := (n, \alpha, K_0, \theta_s, \theta_r, a)$ accurately enough, we will get the projected time-series with a single forward simulation.

3.3. Estimation of Hydraulic Parameters

The basis of the estimation of the hydraulic parameters is the temporal observations at the surface. For a given upper boundary condition, i. e. forcing at the surface, the hydraulic parameters are adapted so that the simulated water content coincides with the observed one. In this context, the questions on which we want to focus are under which conditions an accurate

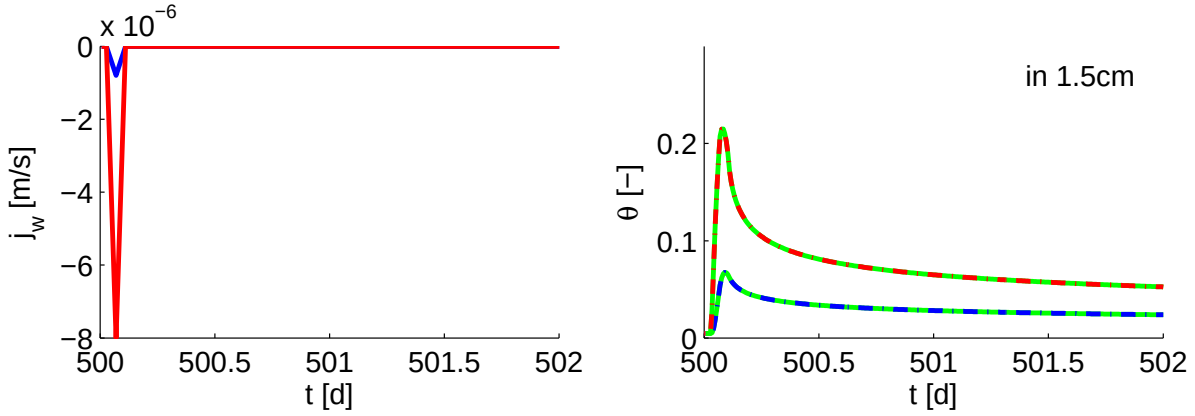


Figure 3.3.: Simulation of two infiltration scenarios for a soil with parameters $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. Left: The upper boundary condition as a Neumann condition is counted negative for incoming fluxes into the soil. Here the rain flux starts after 30 min, ends after further 2 h and linear increase and decrease till its maximum. Right: Modeled near-surface water content θ at 1.5 cm depth corresponding to the two UBC's. The green dashed lines are the estimated θ by inversion, respectively.

estimation of p feasible and under which conditions the input information is too low for estimations. We will investigate it on the basis of simulated time-series with our forward model (3.5). A further aspect is studied in terms of measured field data. Here, we want to address the prospect of the applicability to heterogeneous soils and also dependent on depth.

The hydraulic soil parameters p describe the dynamic behavior. Hence, their correct representation is substantial for the later projection. In Figure 3.2 b), a flow chart is shown to illustrate the inversion steps. The measured forcing (upper BC), the initial parameter set p_{start} and the measured water content $\theta_{\mathrm{meas}}(t; z_0)$ with a small z_0 are required as input data. The first ones flow into the forward model (3.5) and the last one into the objective function for the inversion. The numerical forward simulation of our Richards model (3.5) is calculated by MuPhi (*Ippisch et al. (2006)*). Then the produced output, the simulated water content $\theta_{\mathrm{sim}}(t; z_0, p)$ which belongs to the current parameters, is compared to $\theta_{\mathrm{meas}}(t; z_0)$ in the objective function

$$\chi^2(p) = \sum_i \left(\frac{\theta_{\mathrm{meas}}(t_i; z_0) - \theta_{\mathrm{sim}}(t_i; z_0, p)}{\sigma_i} \right)^2 \tag{3.6}$$

in a least-squares approach with standard deviation σ_i. Here, we assume time independent σ_i and further we choose all σ_i equal one.

Afterwards, the improvement of the current to the previous χ^2-value determines whether a new update of the current parameter set is calculated or the parameter step size is changed. Here, this inversion strategy is realized by the Levenberg-Marquardt algorithm (Software: FitPhi by *Ippisch et al. (2006)*). The inversion stops due to two conditions. Either the initial χ^2-value is reduced by a relative limit or the χ^2-improvement due to the current parameter is below an absolute limit. The relative and absolute limit in the stop criterion are set both to 10^{-8}.

We take a look at the six parameters n, α, K_0, θ_s, θ_r and l before applying this to a first scenario. The van Genuchten parameter n strongly influences the steepness of the van Genuchten parameterization, whereas α scales the curve and has therefore an effect on the air entry value. The saturated hydraulic conductivity K_0 specifies the maximum value in the MvG parametrization.

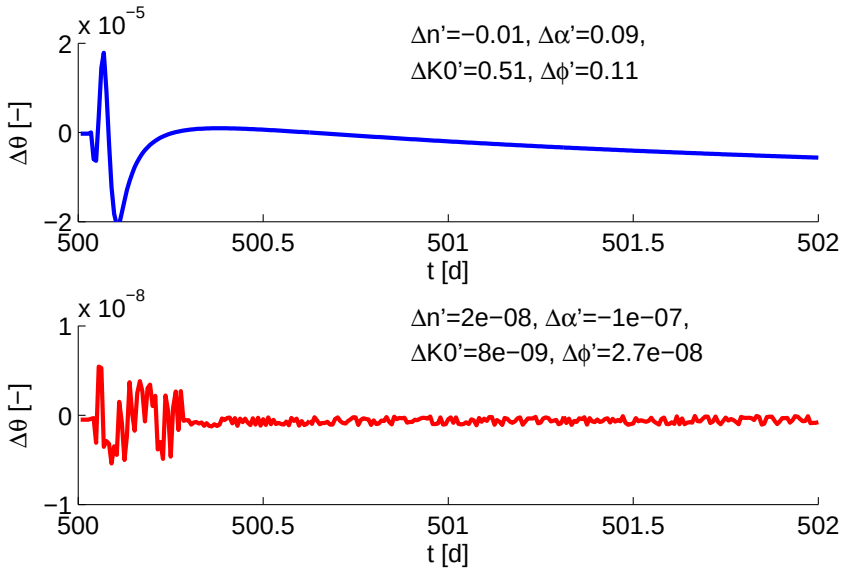
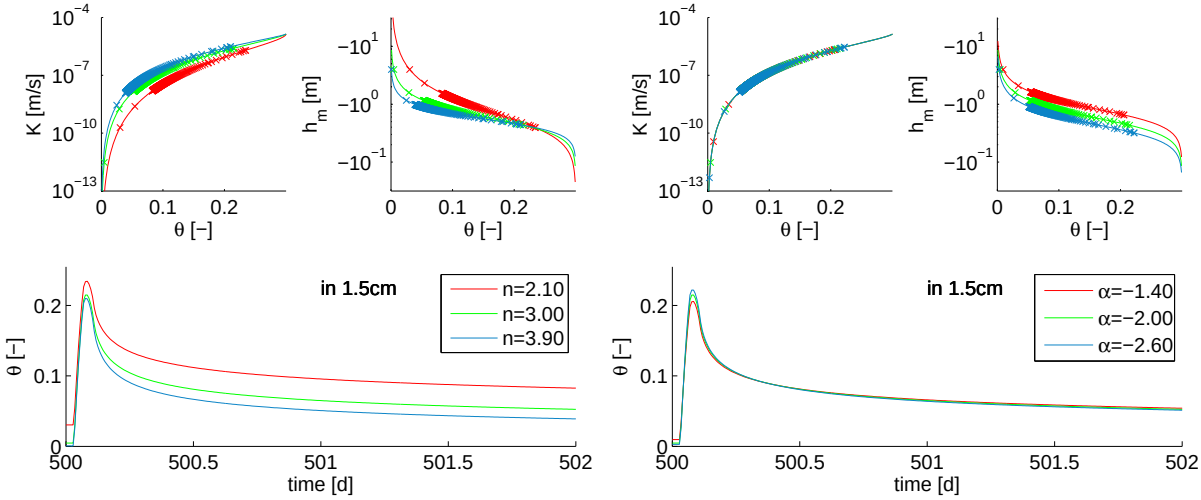


Figure 3.4.: Difference between water content produced by forward model and by inversion. The colors belong to the scenarios of Figure 3.3 with maximal infiltration rate 8×10^{-7} m/s (blue) and 8×10^{-6} m/s (red). The relative errors in each parameter are given after the inversion.

Moreover, θ_s is estimated as the porosity ϕ and is estimated. Since the parameterizations are given by the saturation Θ, the width of the water content domain $\theta_s - \theta_r$ is more important than θ_s and θ_r, respectively. So we restrict the estimation to porosity ϕ and hold $\theta_r = 0$ fixed. Further, *Mualem (1976)* recommends 0.5 as an useful value for the parameter l which is a measure for the pore connectivity.
In conclusion, we restrict the parameter space to four dimensions and estimate the parameter set $p = (n, \alpha, K_0, \phi)$ with the fixed quantities $\theta_r = 0$ and $l = 0.5$.

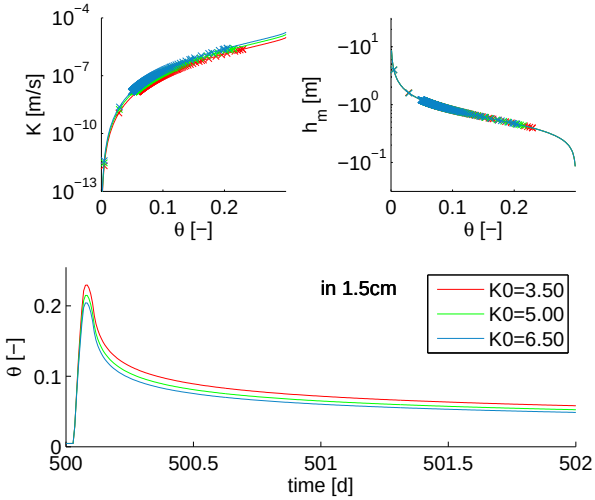
3.4. Infiltrating Hat Function

Our next step is the application of the estimation procedure onto a first example. For that purpose, we chose a hat function as an infiltration flux at the upper boundary. After 30 min, the incoming flux increases linearly over 1 h, reaches its maximum value $\max |j_w(t)|$ and decreases again over 1 h (Figure 3.3 left). After one forward simulation, the resulting time-series of water content at 1.5 cm depth is shown in the right Figure 3.3 for a given soil with the set $p_{\text{true}} = (3, -2\,\text{m}^{-1}, 5\,\text{cm/h}, 0.3)$. The blue and red coloring belongs to different amplitude values of j_w.
The flux j_w and the water content $\theta(t; 1.5\,\text{cm})$ are assumed to be given as synthetically produced measurement data for this synthetic case. The parameter set should be estimated from this information as described in section 3.3. The reconstructed $\theta_{\text{model}}(t; 1.5\,\text{cm})$ is drawn as green dashed lines in Figure 3.3. On a first look, they seem to be similar to its measured water content. However, in Figure 3.4, the residual plot $\Delta\theta(t)$ illustrate some differences. The first thing to note is that the maximum residual amplitudes vary over some magnitudes from 2×10^{-5} till less than 10^{-8} for the infiltration extrema -8×10^{-7} m/s (blue) and -8×10^{-6} m/s (red), respectively.

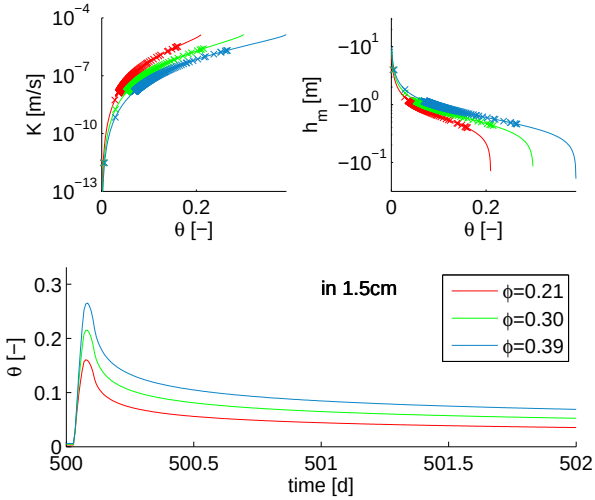


(a) Variation in n

(b) Variation in α



(c) Variation in K_0



(d) Variation in ϕ

Figure 3.5.: Variation in all four parameters (a)-(d) each with factors 0.7, 1 and 1.3. Base of the soil parameters is $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ with the same scenario as in Figure 3.3 (red curves). The crosses are the measurement points of time-series (bottom). In each block, the line colors correspond to the legend of time-series (bottom): Top Left: MvG parameterization. Top Right: vG parameterization.

In case of higher infiltration rates (red curve), the residual is close to the range of computational accuracy. Here, the small structural error during the infiltration peak is comparable to the computational noise level in the whole time range. However, for small flux rates (blue), the error during the peak is so dominating that we cannot notice the remaining noise level.
The same fact can be seen in the corresponding errors of the parameters p_i

$$\Delta p_i' := \frac{p_{i,\mathrm{est}} - p_{i,\mathrm{true}}}{p_{i,\mathrm{true}}} \qquad \forall p_i \in \{n, \alpha, K_0, \phi\}\,. \tag{3.7}$$

Whereas the error $\Delta p_i'$ for a larger amplitude $|j_w|$ is very low in orders of 10^{-8} (Figure 3.4). For smaller fluxes, $\Delta p_i'$ increases much stronger than the residual and ends with error values

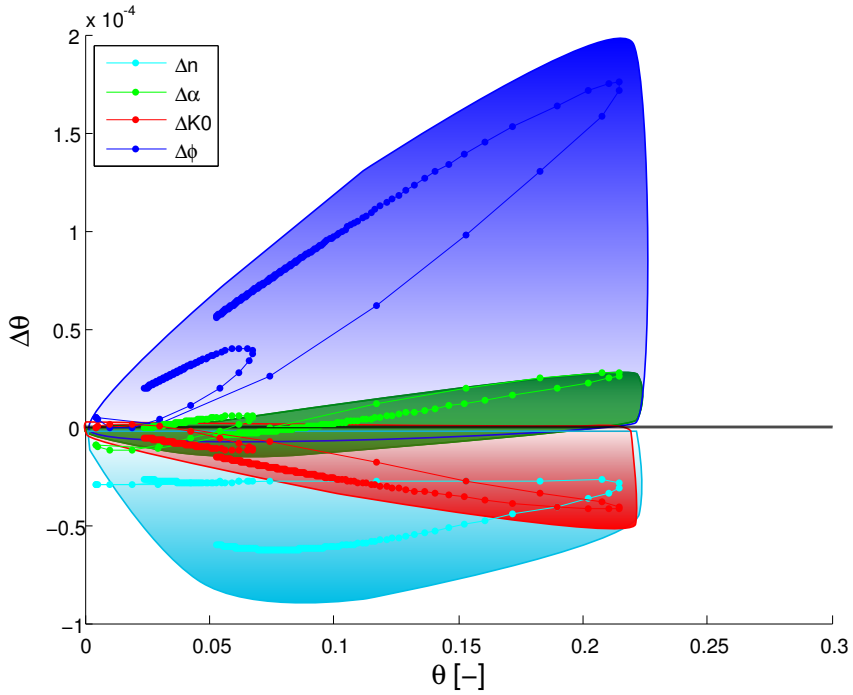


Figure 3.6.: Sketch of the sensitivity due to variations in the Mualem-van Genuchten parameters (colored areas). The points are calculated from a variation of each parameter by 0.1 % for the red (smaller θ-range) and blue (larger θ-range) scenario in Figure 3.3.

around $1 - 50\,\%$.

Sensitivity of Hydraulic Parameters

We have a discrepancy of around one order of magnitude between each error in the four parameters from a more general perspective. This wider range in parameter error $\Delta p_i'$ can be understand more easily if we look at the response of $\theta(t)$ due to variations in one of the four parameters in more detail. Together with the hydraulic conductivity and soil water characteristic, these results are drawn in Figure 3.5 (a)-(d).
The plots for n and ϕ, i. e. (a) and (d), show the largest variations in water content due to $\pm 30\,\%$ deviation in the parameters, respectively. Nevertheless, for different porosities ϕ the change does only apply with increasing water contents, whereas for n the whole θ-range is affected. Hence, the smallest estimation errors in p were expected for the van Genuchten parameter n since large sensitivities make it easier to determine accurate parameters.
Moreover, the saturated hydraulic conductivity K_0 is the least sensitive parameter of the four since the variations in θ are weaker, especially for low θ, and an error in ϕ can be transferred into one in K_0. A comparison of the conductivity functions in Figure 3.5 (c) and (d) shows that a displacement in $K(\theta)$ due to ϕ can be compensated by a much larger deviation in K_0. Also the saturated conductivity K_0 usually varies over some orders of magnitudes for different soil types. Hence, it is not astonishing that K_0 leads to higher errors as in our example in Figure 3.4. Furthermore, the parameter α is expected to be as error-prone as the porosity ϕ. Since variations in α influence the total θ-range with small changes but for ϕ, θ varies stronger with less influence in case of small θ-values.
Further test cases in chapter 5 also demonstrate that the order of the estimation error in the

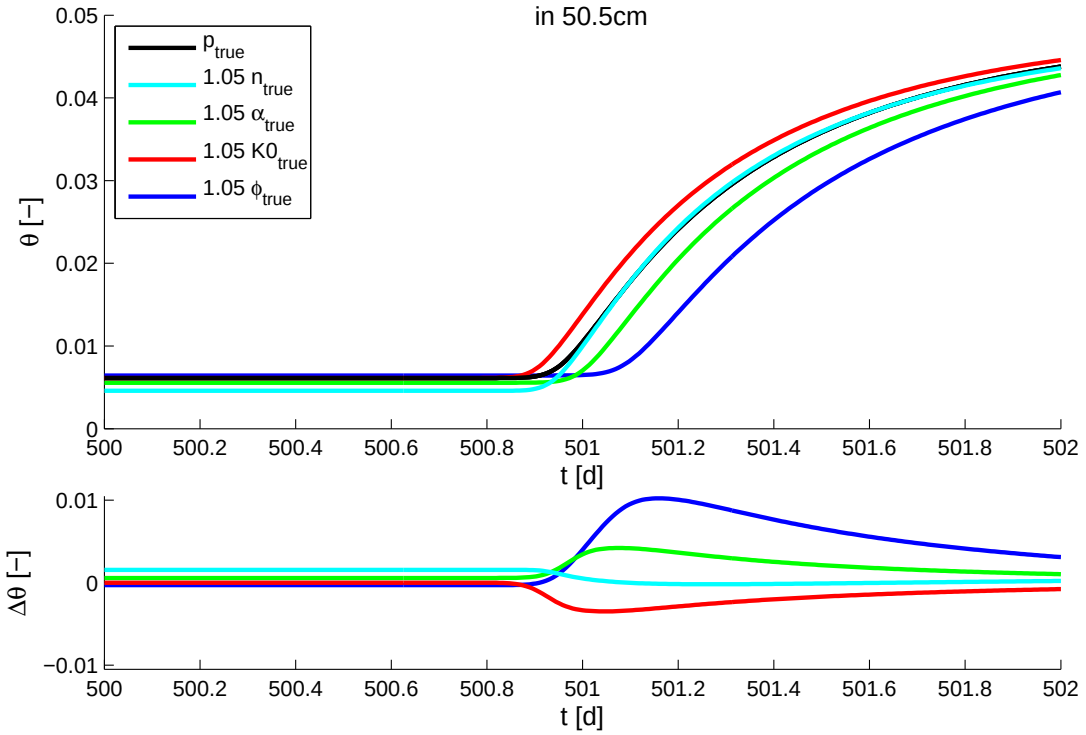


Figure 3.7.: Projected water content at 50.5 cm depth is shown with the true parameter set (green) $p = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ and disturbed ones with $\pm 5\,\% p_{\mathrm{true}}$ (colored curves). Bottom: Residual of the disturbed water content (colored curves) to the true one (black).

four parameter is as follows:

$$\Delta n' \leq \Delta\alpha' \approx \Delta\phi' \leq \Delta K_0' \,. \tag{3.8}$$

Figure 3.6 illustrates more explicitly for which water contents θ each parameter is most sensitive. Here, the colored points denote the difference in water content $\Delta\theta = \theta(p + \Delta p_i' \hat{e}_i) - \theta(p)$ for a change in the parameter by $\Delta p_i' = 0.1\,\% \, p_i$ and $\hat{e}_i$ is the i-th unit vector in $\mathbb{R}^4$. Further, this sensitivity $\Delta\theta$ forced by $\Delta p_i'$ is given for our scenario in Figure 3.3 for both infiltration amplitudes. The colored areas to the parameters sketch the sensitivity areas of p. This makes it clearer that n and α are sensitive in the whole θ-range, whereas K_0 and ϕ have nearly a linear dependency on the maximum infiltration rate of $|j_w|$.

Projection to Depth

As a next step, we clarify the question how the effects are on the projected time-series in depth if we assume a parameter error of 5 % in each parameter. For that purpose, the time-series is projected into 50.5 cm depth with maximal incoming flux $8 \times 10^{-6}\,\mathrm{m/s}$ in Figure 3.3.
In Figure 3.7, the black curve shows the real projected $\theta(t)$ with the true parameters $p = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. The real curve and the four disturbed time-series change the water content after nearly one days. The first two hours are needed for infiltrating the water, then it drains through the soil. After $0.8\,\mathrm{d} \approx 20\,\mathrm{h}$ the water content rises in 50.5 cm. This duration is reasonable for an amount of infiltrating water of $8 \times 10^{-6}\,\mathrm{m/s} \cdot 1\,\mathrm{h}$.

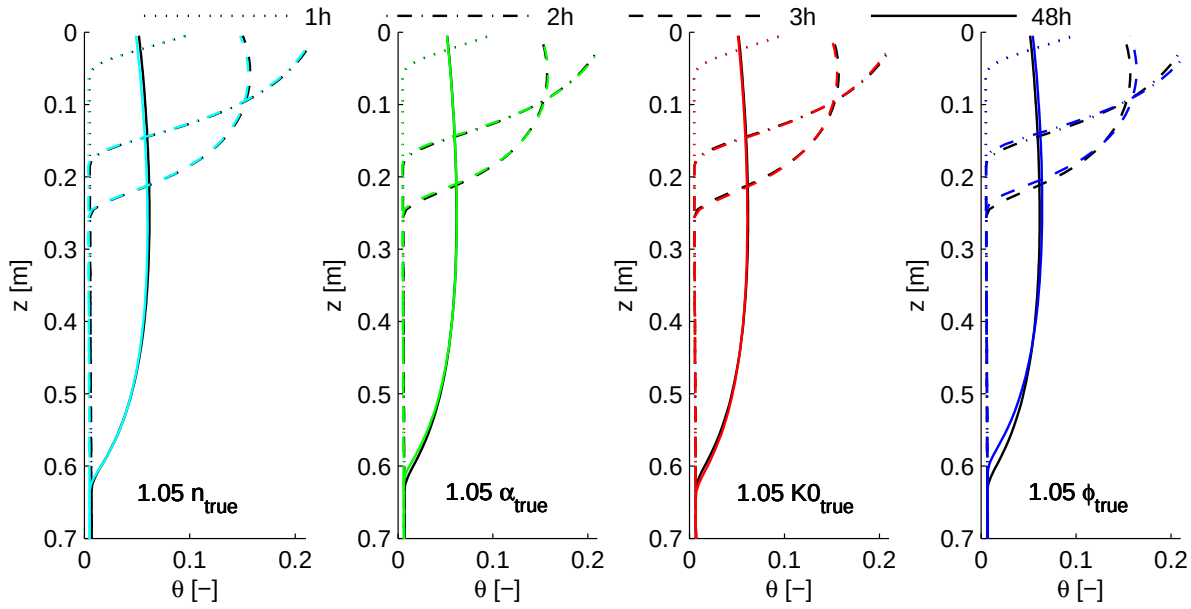


Figure 3.8.: The infiltrating hat function j_w (Figure 3.3 red) propagates through homogeneous and initially dry soil with $p = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. The black lines denote different states in time. The variations of the function $\theta(z;t)$ are shown due to 5 % variation in each parameter. The colors correspond to each parameter of n, α, K_0 and ϕ like in Figure 3.7.

The disturbance of n by $\pm 5\,\%$ yield the light blue curve. Especially in the residual plot below, we observe a difference between the true and the light blue projected curve. This difference is based on the initial state, which is hydrostatic equilibrium in this simulation scenario. It means that in the soil water characteristic in Figure 3.5 we start the time-series at the certain height h_m belonging to the groundwater level at 4 m depth and the projection depth 50.5 cm. Then the corresponding water content is for the larger $n = 1.05\,n_{\mathrm{true}}$ lower than for the smaller $n = n_{\mathrm{true}}$. This deviation is shown before the arrival of the infiltration peak in the projection depth. After the arrival the slope of the water content curve will change if we compare the black with the light blue curve. By changing the slope of the parameterization curve by n, also the slope of the infiltration front is influenced. This can be seen in Figure 3.8. The first 70 cm of the soil profile is shown at different time steps with the same given soil and the same infiltration. The black lines always belong to the true parameter set. Each colored line represents the parameters with the same color as in Figure 3.5. The curvature of these profile curves are also slightly different due to the slope change by n.
If we consider now the green curves of θ and $\Delta\theta$ in Figure 3.7, we will see again the deviation of water content at initial time and later due to a $\pm 5\,\%$-change in α. Analogous to the discussion for n, the reason can be seen in Figure 3.5 b) for the different values due to $\Delta\alpha'$ at initial time. If α becomes more negative, the initial water content decreases according to the initial matrix head. Furthermore, the simulation with $\alpha = 1.05\,\alpha_{\mathrm{true}}$ starts with a lower hydraulic conductivity than the one with α_{true} with the arrival of the infiltration peak, since the water content is lower as shown in function $K(\theta)$ of Figure 3.5 b). This leads to a slower increase of the water content for $1.05\,\alpha_{\mathrm{true}}$ than for α_{true}.
Moreover, the variation of K_0 by 5 % does not lead to a change of the initial water content (red curve) since K_0 has no influence on the soil water characteristic $h_m(\theta)$. Hence, the true and

the disturbed time-series start with the same water content but with different conductivities. Figure 3.5 c) illustrates it and therefore, the infiltration front with the higher K_0 reaches the projection depth quickly. The earlier arrival of the front with $1.05\,K_{0\text{true}}$ can also be observed in Figure 3.8 with the red curves.
Finally, the variation of the parameter ϕ (dark blue curves) leads to an increase of water content, also at initial time as shown in Figure 3.5 d). The water content is larger for a higher porosity there for the same initial matrix head. At initial time, the conductivities in both cases are similar as shown in Figure 3.5 d), but the slope of the function $K(\theta)$ for the lower porosity is stronger than the one for $1.05\,\phi_{\text{true}}$. Hence, the system with the lower ϕ reaches higher conductivities earlier and the infiltration front travels faster through the soil for the true parameter set compared to the parameter set with $1.05\,\phi_{\text{true}}$.
All four disturbed curves show a decrease in the difference to the true curve. The water content is still high at the end of the simulation interval. We go back to Figure 3.8 to understand this. In all four plots, the effect is similar but in last one for the porosity the effect is most obvious. There, we can observe that the small deformation of the parameterization curves due to the parameter disturbance also leads to a deformation of the infiltration fronts. Thus, the profile shows after 48 h that the water content in greater depth is higher for the true parameters, whereas θ for lower depth is smaller for p_{true} than for the ϕ-disturbed parameter set. These two lines intersect each other in a certain depth. This happens a little bit above the projection depth of 50.5 cm at time $t = 48$ h. Hence, the difference of water content due to the parameter disturbance decreases around this time.

Summary

In this chapter, we constructed our model to describe an infiltration or evaporation event onto porous media. In this context, we introduced the non-linear PDE the so-called Richards equation together with the Mualem-van Genuchten parameterization to describe the dynamical water flow in the porous medium.
A next important step was to define our projection strategy via a numerical forward solution. This required the knowledge of the hydraulic soil parameters. These could be estimated based on near-surface time-series of soil moisture using the Levenberg-Marquardt algorithm combined with the forward solver.
The parameter estimation was applied in a first example with two different intensities of infiltration. The relative errors of the parameters varied by some order of magnitudes depending on the infiltration flux rate. Together with a sensitivity analysis we studied the influences of the parameter variations to the projected water content. We observed that a variation of 5 % in one parameter can produce a projection error of water content up to one volume percent at 50.5 cm depth.

4. Application of Estimation and Projection to Measured Soil Temperature

In chapter 2, we introduced two analytical methods to project soil temperature from the surface into depth. In this chapter, both projection methods are applied to a real data set to analyze their applicability. Each method improves the description of the initial condition by mathematical tools. In the case of the transfer function, an addition integral counts the initial thermal energy using the transfer function approach within the profile. In the Fourier based method, the initial condition is improved by a spin-up phase which are compatible with the applied Tukey window to decrease leakage effects. These mathematical tools are introduced stepwise to demonstrate their importance.

Both projection methods were deduced from the heat equation together with some assumptions. In this context, the main points were the horizontal isotropy in x- and y-directions, the spatial and temporal independence of the soil properties and a constant water content in the whole domain $\Omega_z \times \Omega_t$.

It is important to distinguish between the depth z and projection depth z_p. In chapter 2, the depth z is defined as $z = z_p + z_0$, i. e. the projection depth describes a relative relation between the top and the lower sensor. The depth z_0 represents the depth of the highest sensor and thus, the displacement of the modeled top sensor from the surface. Further, an important point is that the model and the measurement are different in the position of the highest sensor. The model assumes the sensor at the surface and measurement z_0 below. But the projection depth z_p is the same for the model and the measurements.

4.1. Transfer Function

In the following, we apply in detail the concept of the transfer function to a real data set. Soil temperatures in different depth are used for the projection. They were measured at the Grenzhof test site close to Heidelberg. Detailed information about this former agriculture field and its profile structure can be found in appendix A.

The projection theory was verified by taking a synthetic data set using transfer function from chapter 2. *Ludin (2010)* studied the dependency of the effective thermal diffusivity D_h on the projection depth within a homogeneous and also a layered medium. The variations of D_h became smaller further away from boundaries, e. g. surface or layer interfaces. Further, she found out that the variations of D_h increased if the top temperature sensor was placed some centimeters below the soil surface. Moreover, in the example at a two-layered medium, *Ludin (2010)* explained that the ratio of amplitude and phase shift changed for different soils for the transformation into depth.
The application of the transfer function method was done in a permafrost region by *Ludin (2010)*. The soil temperature was projected in frozen soil for time-series with several months

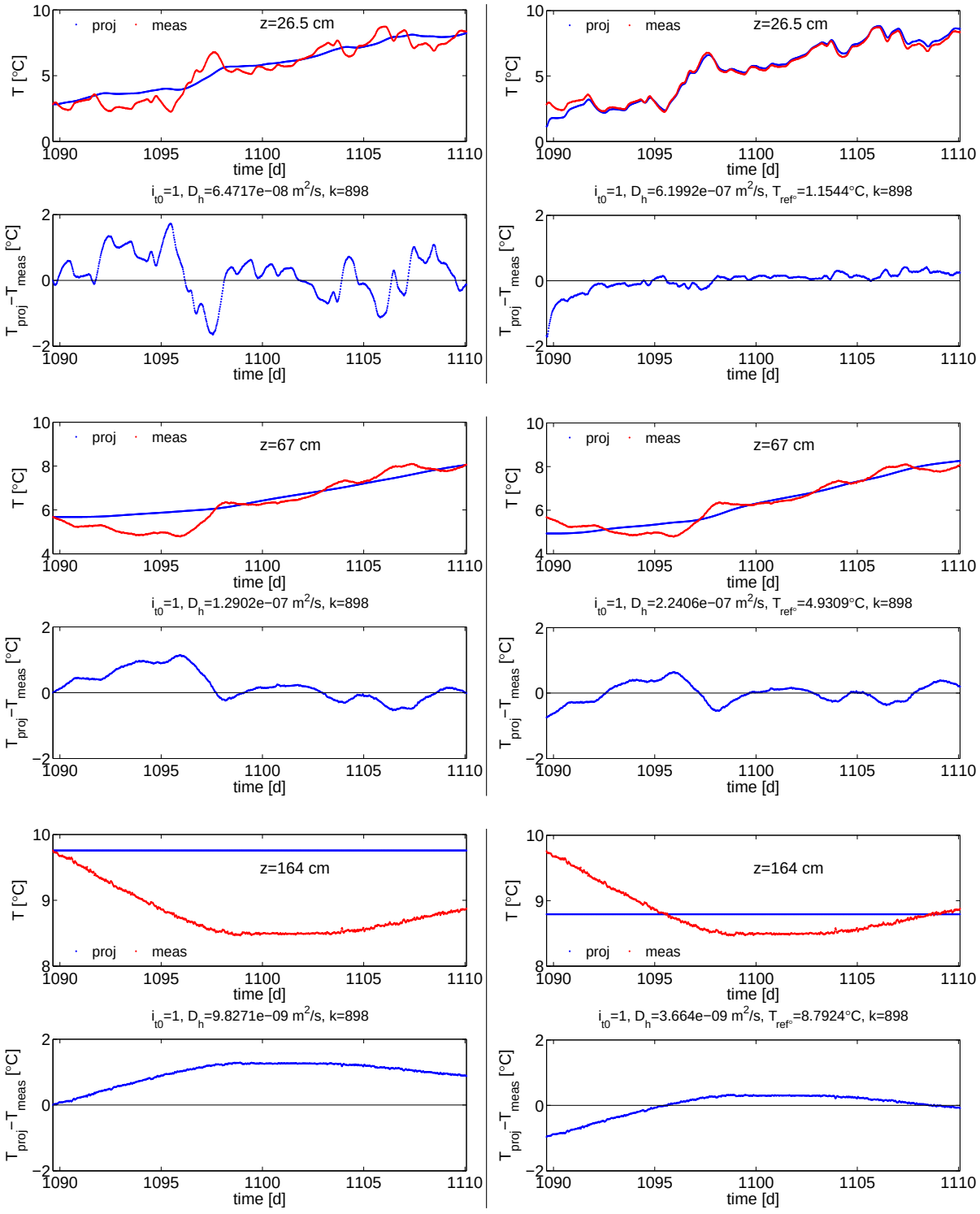


Figure 4.1.: Projection (blue dots) of soil temperature from $z_0 = 12\,\text{cm}$ to three different depths $\{26.5\,\text{cm}, 67\,\text{cm}, 164\,\text{cm}\}$ in each line. The red points denote the measured time-series in the corresponding projection depth. Below, the difference between the projected and the measured soil temperature is drawn in blue. Left: Initial temperature in projection depth is chosen by the measurement at this time and depth. Right: Initial temperature in z_p is the estimated T_{ref°.

length. For such long time-series the projection worked with quite low residuals of $\pm 0.25\,^{\circ}\mathrm{C}$.

4.1.1. Additional Tools for the Transfer Function Approach

In section 2.2, we formulated the model (2.5) in order to calculate the thermal distribution within the soil. First, we take a look at the pure transfer integral expression for the projected temperature

$$T_{\mathrm{proj}}(t, z_p; D_h) = T(z_p, t_0) + \int_{t_0}^{t} T_0(\tau)\, p_t(t-\tau, z_p; D_h)\, \mathrm{d}\tau \tag{4.1}$$

with initial soil temperature $T(z_p, t_0)$ in projection depth z_p. Here, the convolution of the surface temperature T_0 and the travel time pdf p_t (2.10) depends strongly on the effective thermal diffusivity D_h. We can directly determine the projected soil temperature at depth applying (4.1) if parameter D_h is known, i. e. the soil is given.
Now, we follow the inverse problem (2.14) to extract the effective thermal diffusivity D_h. Since at Grenzhof the soil temperature is measured at certain depths (Figure A.2), the left hand side of equation (4.1) is given approximatively. As mentioned in section 2.2.2, the remaining free parameter D_h can be estimated by comparison of projected and measured temperature in the same depth. The optimization problem (2.14) is solved with the Levenberg-Marquardt algorithm implemented in *MATLAB (2008)*.

Two-dimensional Parameter Space

The result of this optimization process for one estimation parameter is shown in the first column of Figure 4.1. Each row represents one depth, i. e. we project from the top sensor $z_0 = 12\,\mathrm{cm}$ into depths $z \in \{26.5\,\mathrm{cm}, 67\,\mathrm{cm}, 164\,\mathrm{cm}\}$ for a time interval of about 20 days. The red points denote the measured time-series $T_{\mathrm{meas},t}$ in the projection depth $z_p = z - z_0 \in \{14.5\,\mathrm{cm}, 55\,\mathrm{cm}, 152\,\mathrm{cm}\}$, respectively. If we compare the measured time-series with the projected ones for the upper two depths, the projection seems to smooth the measurement strongly. Especially for the upper sensor with $z = 26.5\,\mathrm{cm}$, we observe that the main trend of the measurement is reproduced but the amplitudes are damped and thus the time-series $T_{\mathrm{proj},t}$ does not include too much structure of the original time-series $T_{0,t}$. In all these three examples, we observe the same effect. The effective thermal diffusivity is underestimated and therefore leads to a too large damping factor in our travel time pdf.
Furthermore, the time-series shows no detailed structure anymore except the measurement uncertainty in the projection into greatest depth $z = 164\,\mathrm{cm}$, so the parameter D_h cannot be estimated correctly enough to get at least the global trend. The estimated diffusivity in this case is, with a value of $9.8 \times 10^{-9}\,\mathrm{m}^2/\mathrm{s}$, one to two orders of magnitude lower than for the other two projection depths. Here, the damping is much higher due to two reasons: On the one hand, the estimated D_h is much smaller and on the other hand the projection depth $z_p = 152\,\mathrm{cm}$ is over one meter larger. The second reason is a natural one, i. e. it always holds for the modeled and the measured time-series. But the first one can be improved.
In Figure 4.1, the figures of the temperature residues indicate high errors up to $\pm 2\,^{\circ}\mathrm{C}$. This is not an effect of a bad determination of the Levenberg-Marquardt algorithm, or rather the influence of a discrepancy of initial condition in the model and the measurement. In the first step, we chose the initial temperature $T(z_p, t_0)$ in our model (2.5) as the measured temperature at this time and depth. Due to the measurement uncertainty, this choice of the initial temperature is not practicable. Hence, we adapted some improvements in section 2.2 on which we focus now, e. g. an estimation of a second parameter, the reference temperature T_{ref°.

The projected temperature becomes then

$$T_{\text{proj}}(t, z_p; D_h, T_{\text{ref}\circ}) = T_{\text{ref}\circ} + \int_{t_0}^{t} T_0(\tau)\, p_t(t-\tau, z_p; D_h)\, \mathrm{d}\tau \tag{4.2}$$

with the two estimation parameters D_h and $T_{\text{ref}\circ}$. In contrast to the case with one parameter, we do not fix the soil temperature $T(z_p, t_0)$ at the measurement at this point and time. The convolution part of the projection (4.2) is a solution of the heat equation (2.2) and generates all solutions up to a constant term. This term is theoretically given by the initial condition, but in practice it is advantageous to estimate this temperature. The pure convolution integral as the projected temperature does not take into account that the system may be in a dynamic state. Such a difference between measurement and model in the initial state can be estimated in the first order by the estimation of a constant term $T_{\text{ref}\circ}$. Furthermore, the additional degree of freedom is motivated by the uncertainty of the measurements. Thus, we allow the starting value to change and bring more flexibility into our system.
The impact of this variation can be seen in the right column of Figure 4.1. Here, the estimation of the two parameters yields the system one more degree of freedom so that the estimation of D_h can be more accurate. Therefore, we can observe that the residuum decreases for all three depths and its average is around zero. In an obvious case, we can see this fact for depth $z = 26.5\,\text{cm}$. Here, the effective diffusivity grows by a factor of 10 to $6 \times 10^{-7}\,\text{m}^2/\text{s}$. This yields less damping of the time-series and a better structural adaption of the measurement. Further, we can observe that the structural variation is just around $\pm 0.3\,^\circ\text{C}$, but the global trend looks like an exponential progression with a large error at the beginning of the time-series. This error is mainly produced by the difference of the given initial conditions.
The diffusivity D_h changes only by a factor of 2 and 3 due to the additional estimation parameter in the other two depths, respectively. The projection into 67 cm is slightly improved, whereas the one into 164 cm stays worse and D_h decreases. In such a range of the diffusivity with $10^{-9}\,\text{m}^2/\text{s}$ it is not very important if D_h varies a little since the damping is very strong anyway. The largest approximation error in this case is also based on the implementation of the initial condition as we see in the following step.

Initial Condition

A next step to improve the projection result was realized in section 2.2.3 using a spin-up phase. Here, the main idea was to equilibrate the thermal system during a first time interval and after inserting the forcing at the soil surface. In the first column of Figure 4.2, the projection is shown with the additional spin-up phase and is again estimating the two parameters. As before, the results for the same three depths $z \in \{26.5\,\text{cm}, 67\,\text{cm}, 164\,\text{cm}\}$ are drawn by the lines. The spin-up time is chosen around 10 days, i. e. 500 points in time and closely half of the time-series in the examples. This is quite reasonable since our spin-up approximation in section 2.2.3 in Figure 2.3 showed for our example soil with $D_h = 10^{-6}\,\text{m}^2/\text{s}$ that approximatively 50 % of the incoming energy reached a depth of 90 cm after around 10 days spin-up. 60 % reached a depth of 70 cm and 70 % a depth of 45 cm. This example calculation is reliable as spin-up approximation since the value of D_h has the same order than the estimated D_h in the right column of Figure 4.1 for the two higher depths. Applying this to our projection with spin-up in Figure 4.2 (left column), we know that more than 70 % of the entering thermal energy has reached the depth of 26.5 cm after the 10 days spin-up. This should be enough to reduce the

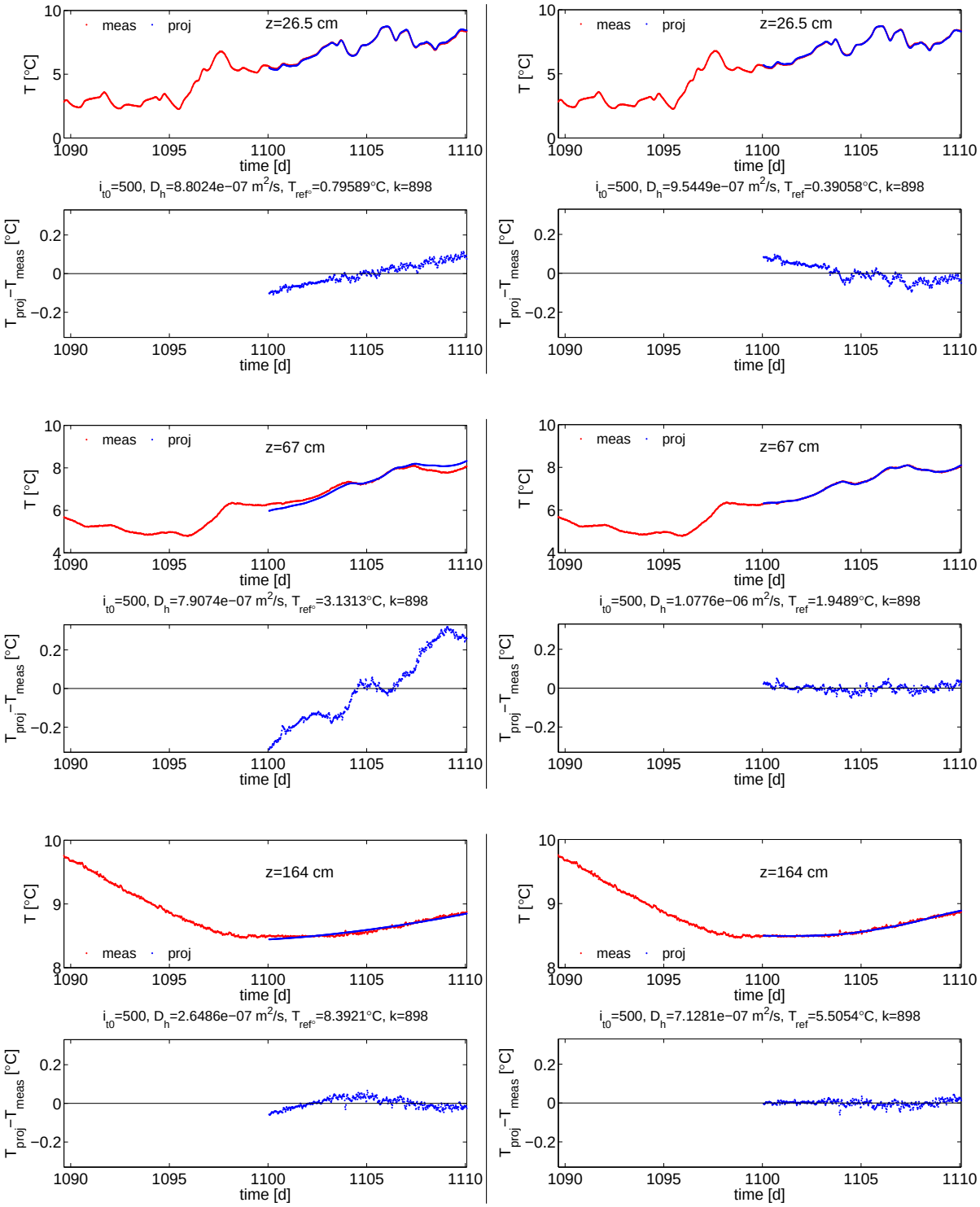


Figure 4.2.: Improved initial condition: Projection (blue dots) of soil temperature from $z_0 = 12\,\text{cm}$ to three different depths $\{26.5\,\text{cm}, 67\,\text{cm}, 164\,\text{cm}\}$ in each line. The red points denote the measured time-series in the corresponding projection depth. Below, the difference between the projected and the measured soil temperature is drawn in blue and is discussed in the text. Left: Projection including spin-up of 500 points, i. e. 10.4 d. Right: Projection including spin-up and z-integral.

estimation error, which can be observed in Figure 4.2. There is still a trend, but now within the bounds of ±0.1 °C.
Including the spin-up phase improves the projection into 67 cm depth, but not as strong as for the lower depth. Nevertheless, after the spin-up, approximately 60 % of the thermal energy comes to this projection depth. This is useful to work with, if a residuum of ±0.4 °C is accurate enough. But it can be improved if we spend more computation time by using a longer spin-up or including the z-integral, which is discussed later.
The greatest projection depth in Figure 4.2 (left column) fits the global trend quite well with this. The estimated thermal diffusivity is increased by two orders of magnitude due to the spin-up phase. Therefore, the damping is not as strong anymore, a straight is fitted through the measurements as in the case without the spin-up. Much less than 50 % of the thermal energy comes into the depth of 164 cm after the spin-up. However, the residuum is below ±0.1 °C. The reason for this is based on the less temporal variations of the measurements in such deep depth. The time-series does not contain much information about the surface signal anymore, e. g. the daily signal is mainly damped out in such depths. Beside a look at the temperature data at 164 cm depth, this can be read in the frequency analysis of measured soil temperature from *Schenk (2011)*. Therefore, an approximation of the right order of the D_h is enough to get the correct trend. This can be further improved if the thermal energy in this depth is better represented.
Till now, we mainly consider the thermal forcing from the top of the soil in the model. It changes the temperature within the soil and warms it up over time to the real thermal distribution in the profile. On the one hand, we can extend the spin-up phase to improve the projection further which costs more computational time due to the hierarchical structure of the transfer integral. On the other hand, the z-integral which we introduced in section 2.2.3 improves the representation of the profile temperature especially around the initial time. The z-integral describes the propagation of the initial temperature profile over time. Further, it is damped out over time due to the diffusion process. This shortens the necessary spin-up time for an accurate estimation.
Thus, the projected soil temperature is given by the expression (4.2) including the z-integral

$$T_{\mathrm{proj}}(t, z_p; D_h, T_{\mathrm{ref}}) = T_{\mathrm{ref}} + \int_{t_0}^{t} T_0(\tau)\, p_t(t-\tau, z_p; D_h)\, \mathrm{d}\tau + \int_0^{z_{\max}} T(\xi; t_0)\, p_\delta^{\mathrm{IC}}(z; t; \xi)\, \mathrm{d}\xi \quad (4.3)$$

with maximal depth $z_{\max}$. We implemented the basic idea of mirror charges to solve the mathematical model in the half space Ω_z. This concept is now realized in Figure 4.2 at the right column. We can observe a reduction of the errors to almost ±0.1 °C for all three depths. The approximation error is mainly dominated by the measurement uncertainty for depth 67 cm and 164 cm and yield very accurate projection results. In both cases, the trend is reduced and almost vanished. Therefore, the estimated thermal diffusivities are very reliable. The D_h-values of $1.1 \times 10^{-6}\,\mathrm{m^2/s}$ and $7.1 \times 10^{-7}\,\mathrm{m^2/s}$ for the two depths 67 cm and 164 cm are different by less than a factor of 2. This variation is based on the layering since the estimated D_h describes the soil between the top most and the projection sensor. If there are more layers within the projection catchment for greater depth, the influence of the layering onto the estimated D_h will be increased since the model assumes only an uniform medium. Details to the influence of layers to the D_h-estimation is analyzed by *Ludin (2010)*.
The projection into depth 26.5 cm needs a deeper look on the assumption. Although in this shallow depth the influences of the spin-up and the z-integral onto the projection results should be very good due to the small projection depth, the projected temperature has still a trend within the small bounds of ±0.1 °C. This behavior can occur due to the position of the top

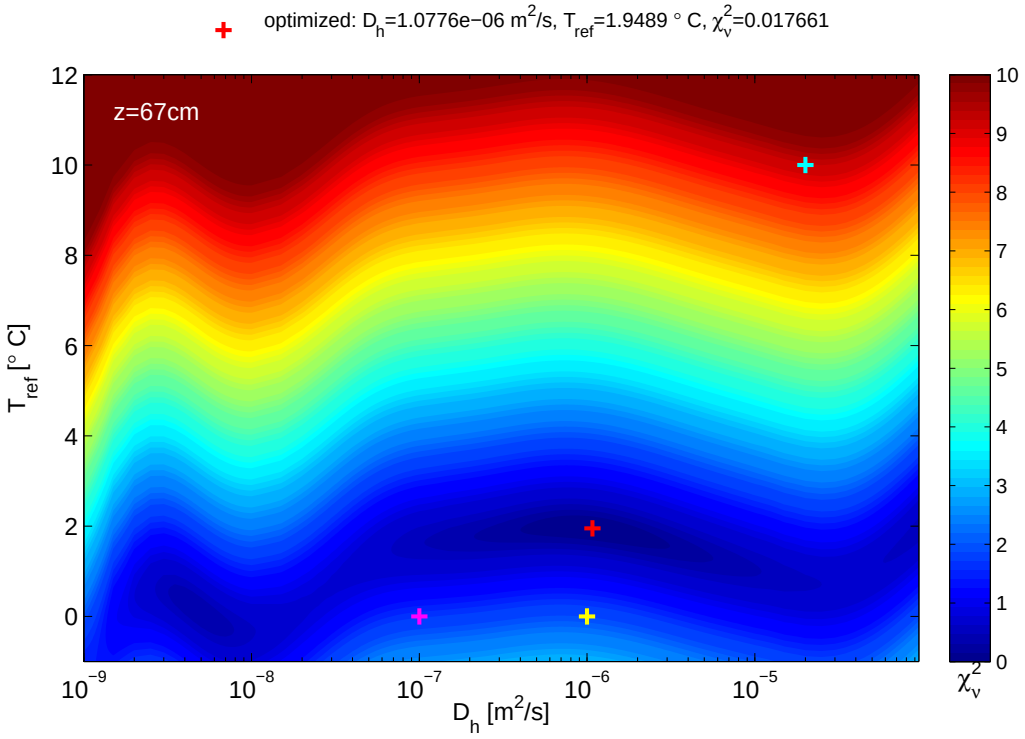


Figure 4.3.: Parameter space of the minimization problem (2.14) for depth $z = 67\,\text{cm}$, 500 spin-up points i_{t0} and with z-integral as in Figure 4.2 right column. The optimized parameter set is marked with a red cross and the corresponding values are given above. The crosses in yellow, cyan and magenta mark different starting points of the parameters. The colorbar denotes the χ^2_ν value described in section 2.2.4. In general, the structured grid is chosen equidistant in $\log_{10} D_h$ and T_{ref}, but partly with a finer grid around the optimal value.

most sensor in the real soil and in the model. The sensor is located at 12 cm depth which is a similar depth as the projection depth $(26.5 - 12)\,\text{cm}$ itself. We know that we produce an error if the real sensor is not located directly at the surface, but we assume that this error can be neglected more and more with increasing projection depth. This error can be observed within the small error bounds of $\pm 0.1\,°\text{C}$ but is acceptably small.
Altogether we can estimate an effective thermal diffusivity with a very accurate residuum with the transfer function method and the additional improvements. Here, we used a combination of spin-up phase and z-integral to realize it. Further, we observe that the z-integral can reduce the trend of the temperature residuum remarkably and leads to a more accurate estimation.

4.1.2. Scanning of the Parameter Space

After the analysis how the additional tools to the transfer function method improved the projection result, we take a look at the parameter space of our optimization problem (2.14). The shape of the parameter space is very important for considering the convergence of the minimization algorithm. Essential questions are about the shape of the space in this context, if there is an unique minimum or many local ones, and further, if there are some local minimum pots or very narrow valleys. Some of these cases can lead to numerical problems during the

optimization.

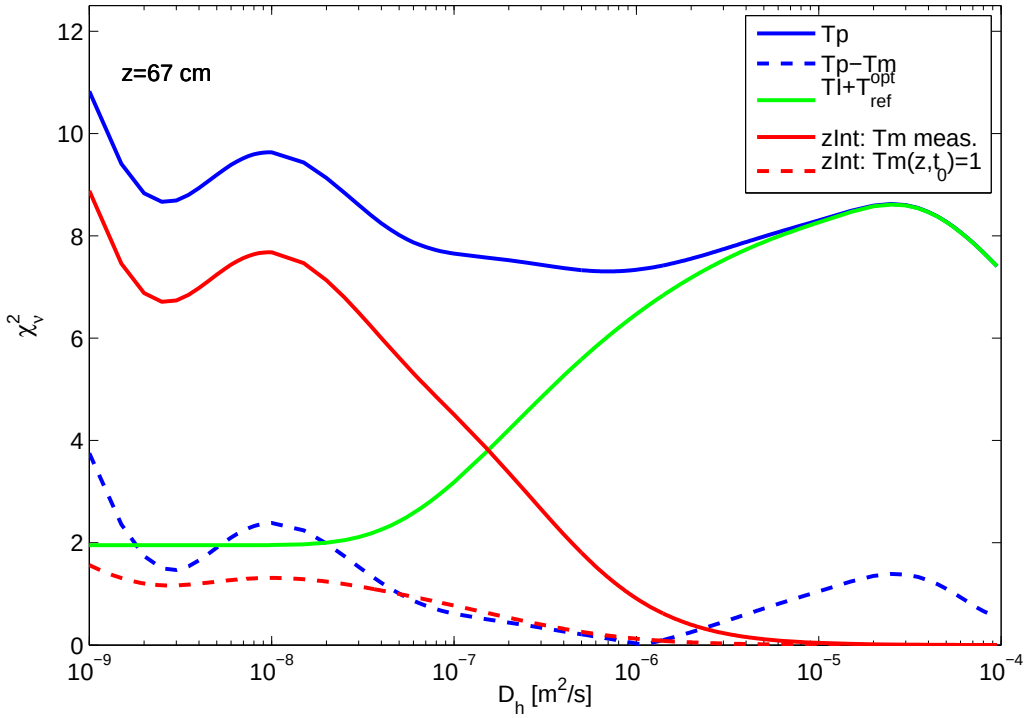


Figure 4.4.: $\chi^2_\nu \sim \|T_\mathrm{p} - T_\mathrm{m}\|^2$ over time-series in dependency of the effective thermal diffusivity D_h. For depth $z = 67\,\mathrm{cm}$ and fixed reference temperature $T^\mathrm{opt}_\mathrm{ref}$, the time-series are as solid blue line: Projected temperature $T_\mathrm{p} = \mathrm{TI} + T^\mathrm{opt}_\mathrm{ref} + \mathrm{zInt}(T_\mathrm{m})$ (4.3); dashed blue line: approximation error belongs to the difference between projected and measured temperature T_m; green: transfer integral (TI) as the convolution (4.2); solid red: z-Integral with measured initial temperature; dashed red: z-Integral with initial temperature identical to 1.

First, we study the parameter space of the projection example in Figure 4.2 at 67 cm depth and with the best improved initial condition. There, the approximation error was in the range of the measurement uncertainty and thus, our model can accurately describe the temperature distribution. In the following, this projection is our basis for the determination of the parameter space in Figure 4.3. We should notice in this figure that the diffusivity is drawn in logarithmic scale.

The scaled $\chi^2_\nu(\triangle T)$ of the residuum $\triangle T$ (2.13) which is connected to the objective function in the Levenberg-Marquardt algorithm, shows more than one minimum in Figure 4.3. At the red cross, we find the deeper minimum compared to the one around $D_h \sim 10^{-8}\,\mathrm{m}^2/\mathrm{s}$. Further, we would expect an additional local minimum some orders of magnitude higher around $10^{-3}\,\mathrm{m}^2/\mathrm{s}$ since χ^2_ν decreases for D_h larger than $3 \times 10^{-5}\,\mathrm{m}^2/\mathrm{s}$.

It is obvious for a fixed diffusivity D_h and variations in T_ref that there is an unique minimum since the reference temperature $T_\mathrm{ref} = \tilde{T}$ is an additional constant to the projection integrals in (2.20). Therefore, it yields an additive term to the objective function (2.13) whereas for a given D_h, the integrals and the measurement data do not vary with T_ref. If we go from very large values for T_ref to smaller once, e. g. in Figure 4.2, the projected temperature becomes deeper vertically to the measurement data and the structure of the projected temperature stays the same. Due to the additive term, the norm of the objective function is smaller. When

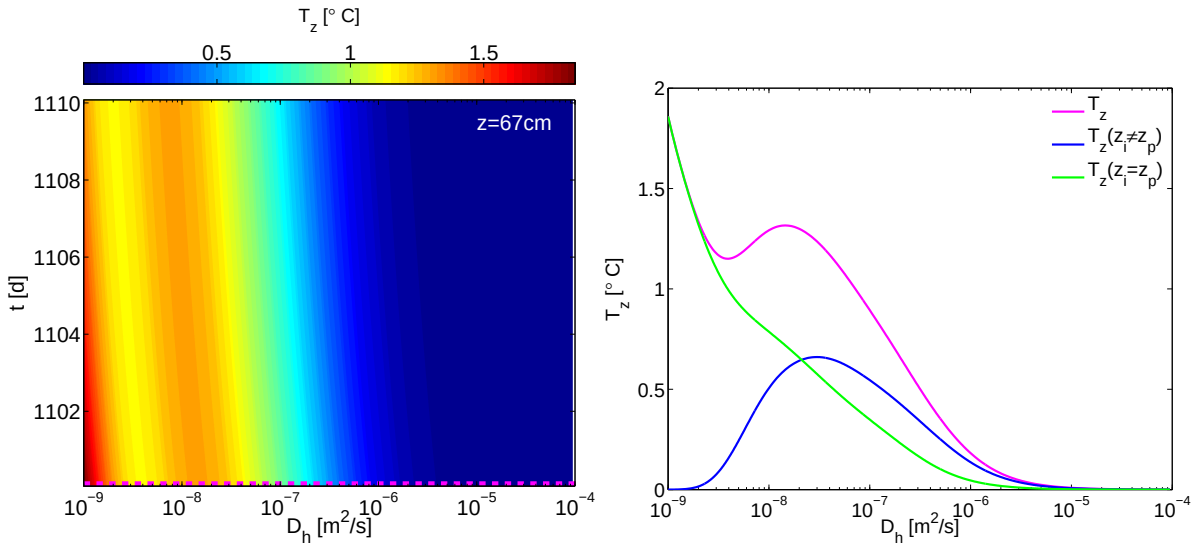


Figure 4.5.: Left: Integral over the travel distance pdf (4.4) as a function of time and diffusivity observed in the depth $z = 67\,\mathrm{cm}$. The magenta dashed line denotes the location of the cross-section for the right figure. Right: Horizontal cross section (magenta) through the left figure at time $t = 1100.1\,\mathrm{d}$. In magenta, this integral T_z is split into spatial integrals $T_z(z_i \neq z_p)$ (blue) for depths z_i strictly greater and smaller than $z = 67\,\mathrm{cm}$ and $T_z(z_i = z_p)$ (green) for depth equal z_p as defined in equation (4.5).

we decrease the reference temperature further so that the projected temperature crosses the measured ones, we find a point with a minimal residuum. The residuum increases for still smaller values of T_{ref} since the projected and the measured temperatures move towards each other.

However, we observe some local minima in direction of D_h. The structure for fixed reference temperatures is based on the projected temperature in dependency of the diffusivity as in Figure 4.4. The dashed blue line of Figure 4.4 refers to a horizontal cross section through the optimal red point in Figure 4.3 with the constant reference temperature $T_{\mathrm{ref}}^{\mathrm{opt}} = 1.95\,^{\circ}\mathrm{C}$. Instead of the residuum $\triangle T$, we look at the projected temperature T_{p} itself, or more precise $\chi^2_\nu(T_{\mathrm{p}})$, which is approximatively vertically shifted from $T_{\mathrm{p}} - T_{\mathrm{m}}$ in Figure 4.4. As we know from expression (4.3), the projected temperature consists of the transfer integral (green slid line) together with the constant T_{ref} and the z-integral (red solid line). Therefore, the superposition of the green and red solid lines produce the blue projection line. The z-integral is very small for high values of D_h and the transfer integral dominates, whereas the transfer integral clearly decreases for small diffusivities. The structure of the z-integral gives the projected temperature curve its characteristic form. Our minimum which fits best the measured data is in the transition zone between these dominating parts of the transfer integral and the z-integral.

Moreover, it is important to notice that the structure of the z-integral curve does not come from the measurement data, which are included in the determination of the z-integral. Therefore, we calculated the z-integral with a constant initial temperature profile $T(z; t_0)$ identical to 1:

$$T_z(t, z; D_h) = \int_0^{z_{\max}} p_\delta^{\mathrm{IC}}(z; t; \xi)\,\mathrm{d}\xi\,. \tag{4.4}$$

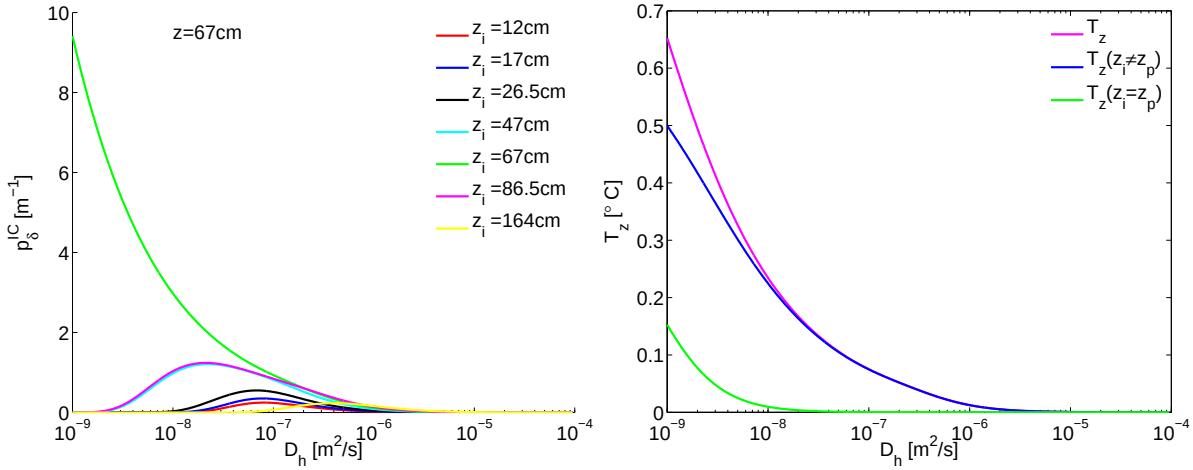


Figure 4.6.: Left: Travel distance pdf p_δ^{IC} (2.19) plotted as a function of thermal diffusivity observed in the depth $z = 67\,\mathrm{cm}$ at time $t = 1100.1\,\mathrm{d}$. The colored lines are different starting points of an initial δ-pulses. Right: Analogue to Figure 4.5 right, but here with a finer grid: the grid size is 1 cm instead of the sensor distribution in the measurement hole.

Thus, the dashed red line is without reference to any measurement data and is purely determined by numerical integration of the travel distance pdf over depth. In comparison with the solid red line, the main structure stays, but is shrunk. The reason for this is that the weights of the travel distance pdf's are uniformly distributed for the dashed case whereas in the solid case, the distribution of the weights follows the measured temperature profile.
Starting from the sum of squared time-series of the z-integral over time (dashed red line in Figure 4.4) with initial temperature identical to 1, we consider the pure z-integral as a time-series in the dependency of thermal diffusivity D_h (Figure 4.5 left) in the following. This means that we observe the propagation of an initial temperature amount through further time (vertical axis). It can be compared for many values of D_h (horizontal axis). Moreover, to get a better overview, the summation over time for a fixed D_h yields again the dashed red curve in Figure 4.4. But we want to focus on a horizontal cross-section (magenta dashed line) through the T_z-plot at time 1100.1 d in the left Figure 4.5. This cross-section is drawn in magenta at the right hand side.
Further, we divide this integral T_z into smaller ones

$$T_z(t,z;D_h) = \int_0^{z_{p-1}} p_\delta^{\mathrm{IC}}(z;t;\xi)\,\mathrm{d}\xi + \int_{z_{p+1}}^{z_{\max}} p_\delta^{\mathrm{IC}}(z;t;\xi)\,\mathrm{d}\xi + \int_{z_{p-1}}^{z_{p+1}} p_\delta^{\mathrm{IC}}(z;t;\xi)\,\mathrm{d}\xi \tag{4.5}$$

$$=: T_z(z_i \neq z_p) + T_z(z_i = z_p) \tag{4.6}$$

to look closer into the structure of the integral (4.4) and the magenta curve in the right Figure 4.5 We define the first two integrals as the integral $T_z(z_i \neq z_p)$ over all travel distance pdf's starting strictly higher or deeper than the one in the projection depth z_p. The third integral of equation 4.5 summarizes the volume directly around z_p. These two new defined integrals are illustrated in Figure 4.5 on the right in blue and green evaluated by the trapezoidal rule. If we compare the magenta and the blue curve, we find a maximum close to each other in each case whereas the integral $T_z(z_i = z_p)$ (green) decreases strictly monotonically. Further, we want to analyze the source of the maxima in the following Figure 4.6 (left). There, the same behavior

can be observed where only each travel distance pdf is drawn instead of the integral for all measurement depths z_i looking from the projection depth z_p. These are the integrands of the integrals, i. e. the integration of all colored curves leads to the integral $T_z(z_i \neq z_p)$ without the green one (blue curve in the right Figure 4.5).
This left Figure 4.6 shows clearly if we observe the thermal system from the depth $z = 67\,\text{cm}$ and if we have one δ-pulse in each measurement depth z_i. These pulses will propagate through the whole soil profile. After a certain time, these pulses reach the observation depth $z = 67\,\text{cm}$ and yield their rate to the z-integral as in equation (4.4). In this context, we can associate changes of the diffusivity as ones of time since the explicit expression (2.19) of the travel distance pdf only depends on the product $D_h t$. Therefore, if we fixed t and vary D_h by a factor, it will influence the pdf in the same way as we fixed D_h and vary the time t.
Thus, we can finally reduce the source of the local minimum around $D_h \sim 6 \times 10^{-9}\,\text{m}^2/\text{s}$ in the parameter space (Figure 4.3) to the weighted superposition of the travel distance pdf (Figure 4.6 left). We know that each of them describes a propagation of thermal energy starting at z_i, but from the view of the projection depth z_p. The superposition of them leads to the local structure of extrema in the parameter space.
However, what would happen if we had more temperature sensors within the soil and therefore a finer grid of sensors? In this case, the number of grid points for the integral $T_z(z_i \neq z_p)$ would increase whereas the number of grid points for the integral $T_z(z_i = z_p)$ would stay the same. Hence, the difference between the two integrals of lines (magenta and blue) in the right Figure 4.5 decreases. We can observe it in the right Figure 4.6 where a much finer sensor distribution with 1 cm is assumed compared to the real measurement profile with approximately 20 cm sensor distance. The integral $T_z(z_i = z_p)$ (green) has still its monotonic form and the other two integrals also get strictly monotonic decreasing with larger D_h. This means that a finer sensor distribution, especially around the projection depth, would vanish the local structure of extrema of T_z since $T_z(z_i \neq z_p)$ would include more travel distance pdf's with its maxima smaller than $10^{-8}\,\text{m}^2/\text{s}$. A grid refinement of the depth induces a need of measurement values at more depths. An alternative is the interpolation of the initial soil temperature between the existing depths and to refine the grid onto the interpolated basis. This interpolation is only an approximation of the temperature of the profile. Its quality depends on the temperature variation within the profile and the number of available sensors.
Summarizing, we saw that the ambiguity of the minimum in the direction of small values of the thermal diffusivity is caused by the z-integral and especially there by the coarse superposition of the travel distance pdf's. Further calculations show that a much finer spatial grid will remove these artefacts for small diffusivities. However, such small grid sizes are inefficient for measurements. Therefore, an interpolation of the measured initial soil temperatures in space is one possibility to approximate the z-integral with a finer grid. In total, the grid refinement would lead to a more smoother catchment of the parameter space around the optimal point. In our example, there is no convergence problem for physically realistic D_h-starting points approximatively from $2 \times 10^{-8}\,\text{m}^2/\text{s}$ to $2 \times 10^{-5}\,\text{m}^2/\text{s}$. But this widening of the well-posed catchment can be more robust to the choice of a D_h-starting value for other applications, but it costs more computational effort.

Response to Variations of Diffusivity

Further, we want to answer the question how large the residuum of the projection will be by varying the estimation parameter D_h around its optimal set. Therefore, we look at the horizontal cross-section through the optimal parameter point in D_h-direction in Figure 4.4.

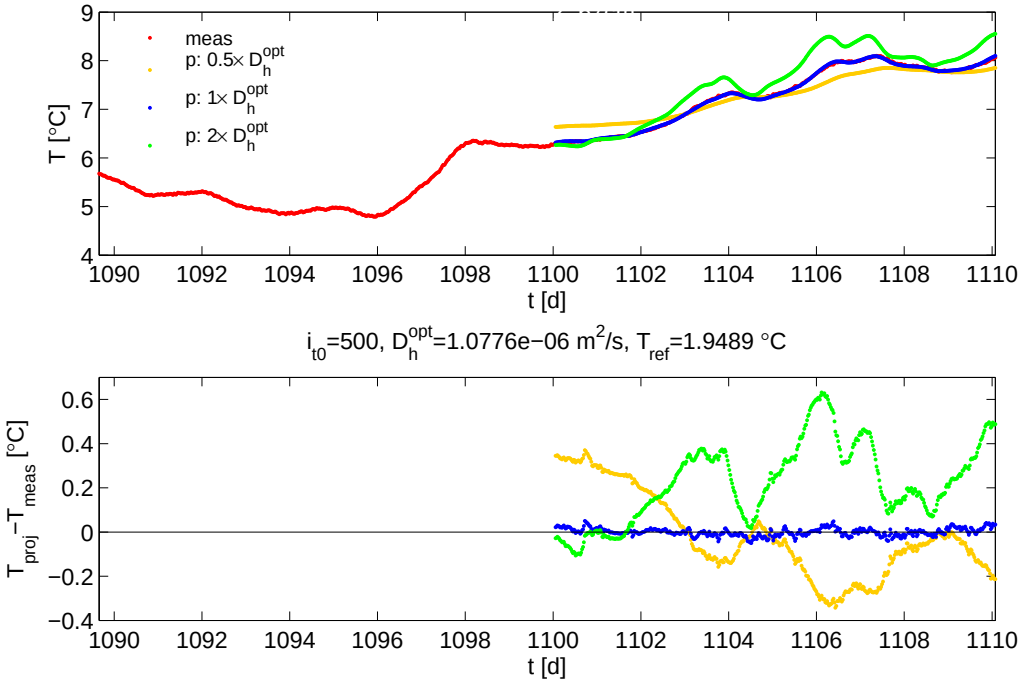


Figure 4.7.: Projection of soil temperature into depth $z = 67\,\mathrm{cm}$ with improved initial condition, the same example as in Figure 4.2 right. Blue denotes the optimized D_h^{opt}, orange and green one half and two times of D_h^{opt}, respectively. In each case, the optimized reference temperature is used.

Since the direction of the reference temperature influences only the offset of the projected temperature, we keep T_{ref} constant at its optimized value. Thus, we vary our parameter D_h only on the dashed blue curve in Figure 4.4.

In the following, we study two cases of variation in which the effective thermal diffusivity is 2 or $\frac{1}{2}$ times the optimized value $D_h^{\mathrm{opt}} = 1.0776 \times 10^{-6}\,\mathrm{m}^2/\mathrm{s}$. The response of these variations of D_h to the projected temperature is shown in green and yellow in Figure 4.7. As in previous plots the projection is colored in blue with a good accuracy and measurement in red. If we increase the effective diffusivity by a factor 2, the damping factor decreases proportional to $\frac{1}{D_h}$. Thus, the calculated projected temperature (green) still includes too much of the surface temperature structure. Beside the damping, the phase shift is influenced further by the diffusivity. Since the phase shift changes also with $\frac{1}{D_h}$, it is too little to produce the extrema at the same time as the measurement time-series.

If we now consider the other possibility and decrease D_h by a factor of $\frac{1}{2}$, the damping and the phase shift increases. The structure of the projected time-series in yellow is too little to reproduce the measurement correctly enough. In this case, the result looks more like a moving average.

The approximation error grows as shown in both cases in the bottom Figure 4.7. The changes of the residuum in both D_h-directions with the factors 2 and $\frac{1}{2}$ seem to be globally symmetric since in the first case, the time-series structure is overestimated and underestimated in the second case. But the error amplitudes is different despite the symmetric D_h-variation. This non-symmetry of the residual is caused by the nonlinear dependency of the travel time pdf from the D_h. Hence, this is also transferred to a non symmetrical, but smooth parameter space

(4.3) in D_h-direction as well as in its cross-section in Figure 4.4.
We find a smooth surface in the parameter space representatively for all parameter spaces of the minimization problem (2.14). This surface is formed by integrals of the weighted pdf's as in equation (4.3) and therefore, at least one time differentiable.

Convergence of Levenberg-Marquardt Algorithm

In the context of optimization problems, it is important to choose suitable starting values for the optimization parameters. In practice, we start with an effective thermal diffusivity $D_h = 10^{-6}\,\mathrm{m}^2/\mathrm{s}$ which is within the range of a realistic value. The initial reference temperature is chosen to zero in the projection depth since we include the z-integral to the projection for a realistic initial profile. This initial parameter set is drawn in yellow in Figure 4.3 with two further initial parameter sets. These are summarized together with their estimation results in Table 4.1.
Since we observed more than one minimum in direction of D_h, a good choice of D_h is important in general. The Levenberg-Marquardt (LM) algorithm is an efficient solver for this optimization problem within the range of $10^{-8}\,\mathrm{m}^2/\mathrm{s}$ till $10^{-5}\,\mathrm{m}^2/\mathrm{s}$ for D_h. As we saw previously, the range can be extended by refinements of the initial profile grid for the z-integral. The smooth shape of response surface in Figure 4.3 and the construction of the LM algorithm, as a combination of Gauss-Newton (GN) and steepest descent method, ensure the convergence to a local minima. Further, we detect similar behaviors with the testing of the algorithm by starting from different parameter sets. In the example (Figure 4.3), the LM algorithm always reaches the same minimum after less suitable iteration steps, often in the range of up to 20 steps. If we start at one point of the yellow, cyan or magenta crosses, the optimal point (red cross) will be reached very closely. We end up with values around $1.0776 \times 10^{-6}\,\mathrm{m}^2/\mathrm{s}$ and $1.9488\,^\circ\mathrm{C}$ for the different initial parameters D_h^{init} and $T_{\mathrm{ref}}^{\mathrm{init}}$. The optimized points vary in a very low range under 0.013%. The estimation of the reference temperature is one order of magnitude more accurate than the one of the effective diffusivity. These variations are a lot below the accuracy under consideration of the measurement errors within the projection determination.
Furthermore, if we take a look at the residuum or rather the χ^2_ν-value, we analogously observe very small changes for this quantity around 1.77×10^{-2}. This leads to clear changes under 10^{-6} which corresponds to the implemented tolerance value for the objective function. It confirms

Table 4.1.: Initial parameters for the Levenberg-Marquardt algorithm together with their corresponding optimized parameters D_h^{opt} and $T_{\mathrm{ref}}^{\mathrm{opt}}$. χ^2_ν is a measure for the objective function. The cross colors refer to Figure 4.3 and sign different initial points in the parameter space. The percentage changes for the certain values are compared to the yellow marker, i. e. $D_h^{\mathrm{init}} = 10^{-7}\,\mathrm{m}^2/\mathrm{s}$ and $T_{\mathrm{ref}}^{\mathrm{init}} = 0\,^\circ\mathrm{C}$. Each of these optimization runs stops due to the change of the objective function which is less than the tolerance 10^{-6}.

cross color	D_h^{init} $[m^2/s]$	$T_{\mathrm{ref}}^{\mathrm{init}}$ $[^\circ C]$	D_h^{opt} $[m^2/s]$	$T_{\mathrm{ref}}^{\mathrm{opt}}$ $[^\circ C]$	χ^2_ν
yellow	10^{-7}	0	1.07762×10^{-6}	1.9488700	1.7660784×10^{-2}
cyan	10^{-6}	0	1.07759×10^{-6}	1.9488781	1.7660793×10^{-2}
magenta	2.5×10^{-5}	10	1.07776×10^{-6}	1.9488431	1.7660786×10^{-2}
yellow change			0%	0%	0%
cyan change			0.003%	0.0004%	5×10^{-5}%
magenta change			0.013%	0.0014%	1×10^{-5}%

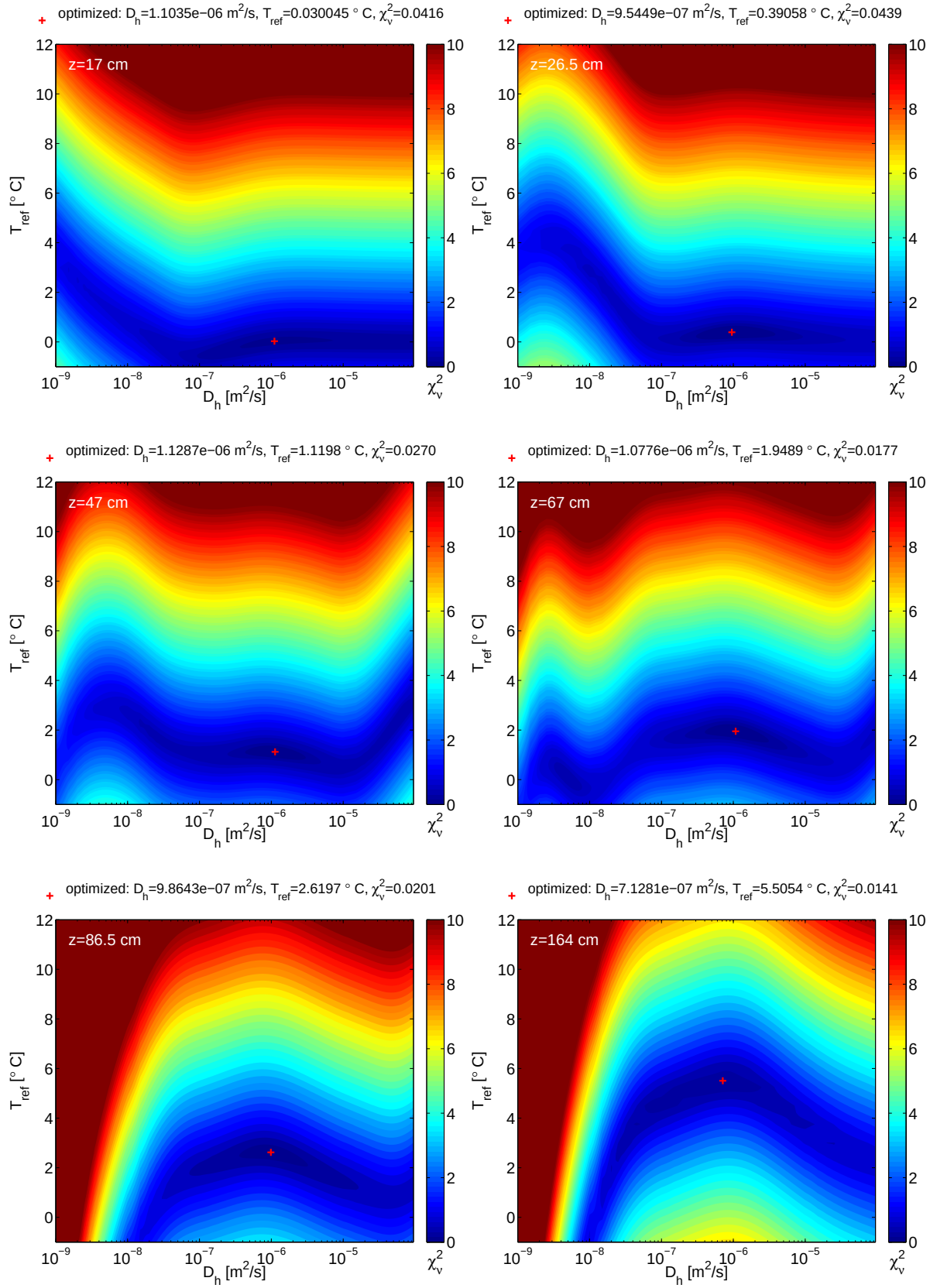


Figure 4.8.: Parameter space for depths $z = 17\,\text{cm}, 26.5\,\text{cm}, 47\,\text{cm}, 67\,\text{cm}, 86.5\,\text{cm}, 164\,\text{cm}$ (line-by-line) calculated with projected temperature (4.3), ~ 10 days spin-up phase, 10 days projection interval, compare Figure 4.2.

that the variations are all within the assumed tolerance range in the region with dominant transfer integral content.
In conclusion, these results do not show problems of convergence within the range of physically realistic values due to the choice of the starting parameters. A grid refinement of the profile can increase this suitable range. Moreover, the parameter space and the estimation information, e. g. the number of inversion steps, do not yield features of a difficult shape around the minimum, for instance a very narrow valley close to the optimal point.

Spatial Variations

Previously, we studied the structure of the two-dimensional parameter space at the example projection depth 67 cm and the 10 d-time interval. In the following, we consider the spatial variation of the parameter space, i. e. for different projection depths.
In Figure 4.8, the projection of the soil temperature is shown from depth 12 cm into the greater depths $z \in \{17\,\text{cm}, 26.5\,\text{cm}, 47\,\text{cm}, 67\,\text{cm}, 86.5\,\text{cm}, 164\,\text{cm}\}$. The parameter space into 67 cm we discussed in detail. For the other five subfigures, all conditions stay the same, except the projection depths.
In general, we observe a similar structure of the parameter space for all depths compared to the one at 67 cm, i. e. the parameter space is continuous and shows multiple minima. Further, the figure evinces also the uniqueness of the minimum and its monotonicity beside the minimum in T_{ref}-direction.
Furthermore, we observe an increase of the χ^2-surface around the minimum with an interference of the extrema structure for each depth in D_h-direction due to the coarser grid resolution for the z-integral. The χ^2-surface becomes steeper for smaller diffusivities With greater depth. There, the uneven shape of the parameter space is reduced with deeper projection depth since the influence of the z-integral is weaker. The z-integral includes only the measurement depth, i. e. to 164 cm depth. If we look back to Figure 4.6 with the travel distance pdf's against D_h, only the pdf with $z_i = z_p$ has a larger value for a small D_h. The reason for this is the longer travel time of thermal energy which starts beside the projection depth towards z_p. Thus, the travel distance pdf's with $z_i \neq z_p$ are decreased a lot with lowering the projection depth. The only dominant entry to the z-integral is the monotonic decreasing entry with $z_i = z_p$. Therefore, this surface of the parameter space increases strongly for smaller thermal diffusivities.
The optimal points, marked by red crosses, vary only slightly with depth. These variations are mainly caused by the layered profile structure and so by the different soil properties. Further, the soil is characterized in the appendix A.1 to get an impression over the soil structure to these data.
Moreover, the optimal reference temperature rises with each depth. We should remember the meaning of T_{ref} to understand the reason for this increase. The reference temperature originally stands for the measured initial temperature value at the projection depth, but reduced by the z-integral. Thus, the reduction by the z-integral is smaller with lowering the influences of the z-integral with depth. All together it increases the optimal reference temperature.

4.1.3. Temporal Evolution of Effective Thermal Diffusivity

After analyzing how accurate the transfer function method can project soil temperature into depth and how the parameter space of our optimization problem (2.14) is shaped due to the different projection terms in (4.3), we focus on the temporal component. In a next step, we use this projection method with all of our improvements to estimated the effective thermal

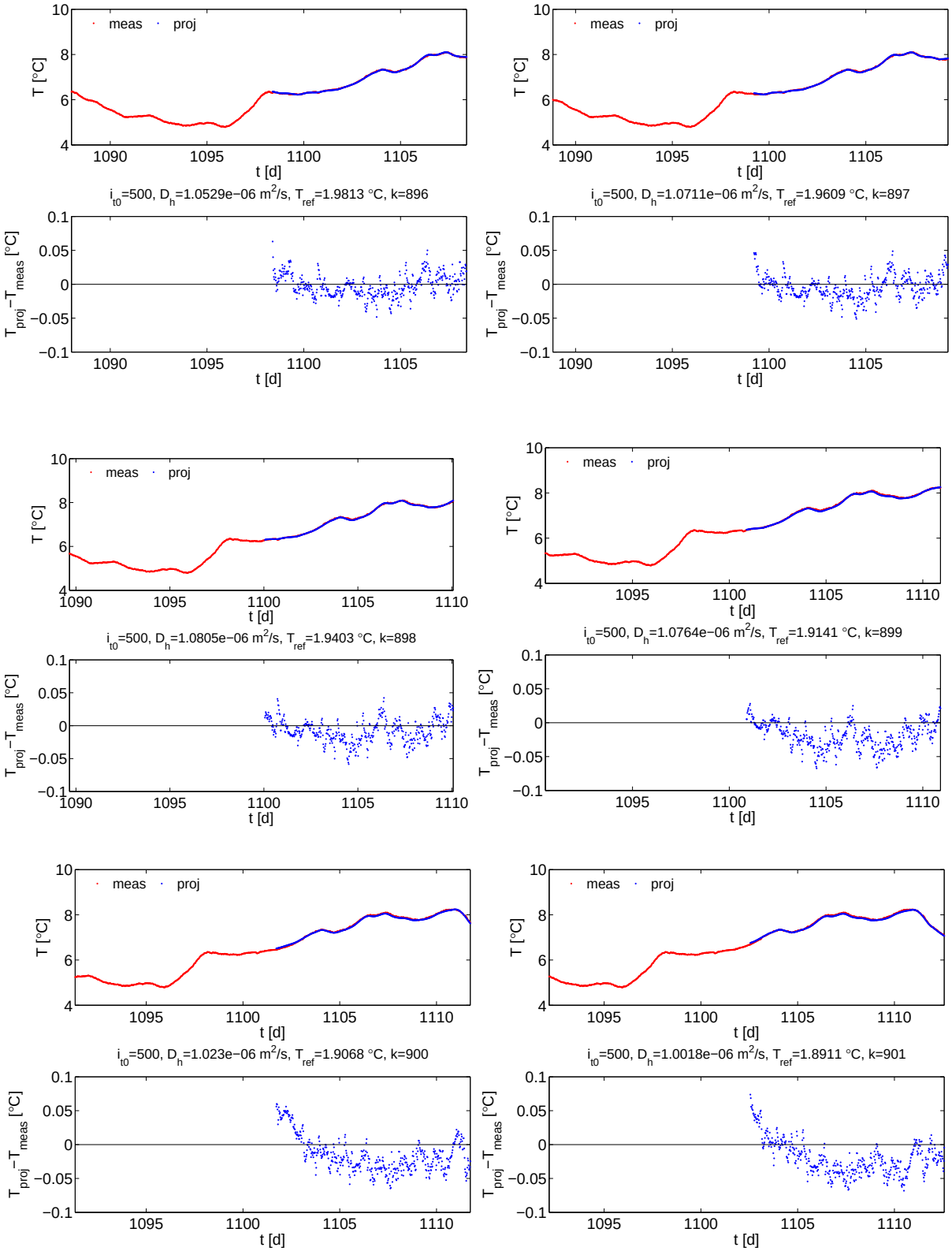
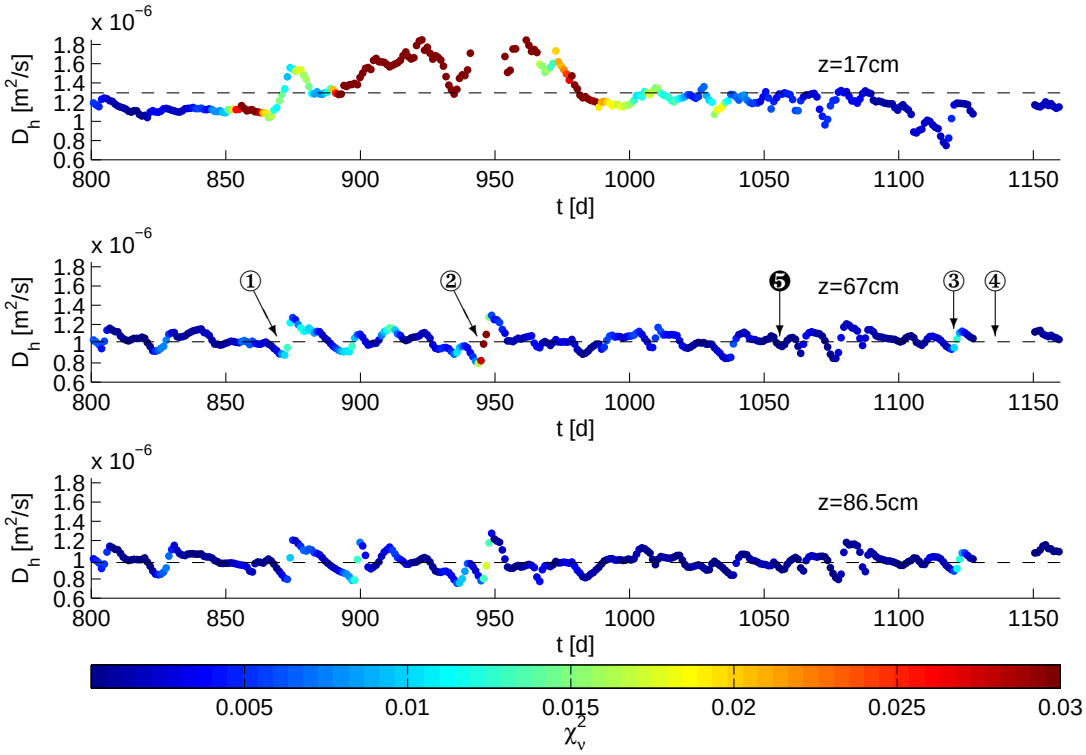


Figure 4.9.: Projection of soil temperature (blue) with spin-up phase (10.4 days) and z-integral from $z_0 = 12\,\mathrm{cm}$ to depth $z = 67\,\mathrm{cm}$. Line-by-line, the time interval is shifted by one day. The estimation interval has a length of 10 days.

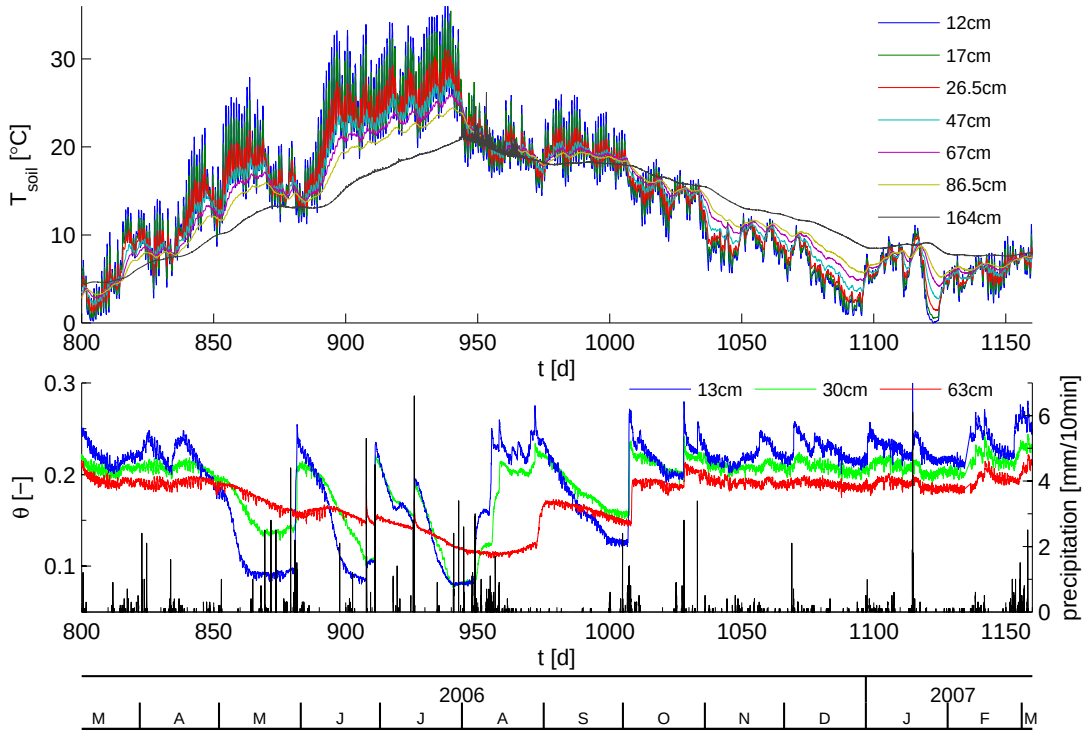
diffusivity D_h in its temporal evolution. As theoretically discussed in section 2.2.4, we take moving time intervals to realize this temporal evolution of the effective thermal diffusivity $D_h(t)$. The time interval is divided into two subintervals, the spin-up phase with 10.4 days, i. e. 500 data points, and the estimation interval 10 days, i. e. 480 data points in Figure 4.9. One should remember that the estimation of D_h and T_{ref} only belongs to the second interval, the so-called estimation interval, whereas the calculation of the projection runs over the whole time interval. In Figure 4.9, an estimation sequence is shown with these moving time intervals for the projection into 67 cm depth. Line-by-line, the time interval is shifted further by 1 day. The projected temperature reproduces the measurements very accurately for all these six estimations. The last two time intervals show a small trend despite their low residual. Thus, this is also given in the scaled χ^2_ν-value of the approximation error. They increase continuously with each time interval from $\chi^2_\nu = 2.6 \times 10^{-4}$ to 6.3×10^{-4}. It corresponds to an averaged approximation error smaller than $\pm 0.05\,^\circ\text{C}$ which is quite accurate and in the range of the measurement uncertainty.
In the last two estimation intervals, a stronger decrease in soil temperature is visible after day 1110. During this time, no rain is falling, but the surface temperature is dropping down to 1 °C and forces the soil temperature to follow. This stronger curvature gives the time-series more characteristic structure sensitive to the D_h-values. During the inversion, parts with stronger curvature are the dominant entries of the objective function due to the sum of squares of the residuum. Thus, it is adjusted in a stronger way. Since the structure after 1105 days is represented quite well, the slightly lower estimated diffusivity can also be realistic here. The small trend of 0.1 °C could be caused by the z-integral since lowering of D_h slowed down the transport of thermal energy. In such sensitive cases, the refinement of the sensor grid or the extension of the spin-up interval can be useful.
We observe a variation of the effective thermal diffusivity D_h from these six time intervals. First, D_h increases and then decreases within a low range from $D_h = 1.0018 \times 10^{-6}\,\text{m}^2/\text{s}$ to $1.0805 \times 10^{-6}\,\text{m}^2/\text{s}$. In the following, we extent these calculations of Figure 4.9 to a greater time range and we analyze how the estimated diffusivity evolves over the year.

Longterm Evaluation of D_h

In our model, we assume time-independent soil properties and constant water content within the simulation interval, i. e. during spin-up and estimation. Further, the effective thermal diffusivity D_h is only determined in the estimation interval and is restricted to the simulation interval. Thus, we can generate a temporal evolution of D_h over longer periods, e. g. one year, if we choose short estimation intervals and shift these forward in time. In this context, key questions can be how sensitive the estimated D_h on the conditions of the seasonal process is and are the variations of D_h stronger than the uncertainty due to the estimation.
The mid point of each estimation interval is plotted as a colored dot to display the estimation results of D_h. Further, the χ^2_ν-value defines the color of the dot to indicate the quality of the estimation. Three projection depth are chosen at 17 cm, 67 cm and 86.5 cm depth to illustrate the temporal evolution of D_h. Figure 4.10 shows one year from March 2006 till March 2007, a) the estimated thermal diffusivity over time and b) the measured soil temperatures, soil water content and precipitation. These quantities demonstrate the profile conditions over time. The same model is used for each point for the D_h-time-series. The only change is the shifted time window of the projection. These quantities in b) contribute to understand the differences of the diffusivity $D_h(t)$.
In Figure 4.10 a), the first observation is that the majority of the estimations have a low



(a) Time-series of the effective thermal diffusivity D_h estimated in moving time intervals: The colored dots represent the midpoint of the estimation interval and are colorized by the χ^2_ν-value. The mean of $D_h(t)$ is given by the black dashed line.



(b) Measurement time-series: The soil temperature is using for the inversion. Below the soil water content and the precipitation are measured corresponding to the temperature profile.

Figure 4.10.: Temporal evolution of D_h in projection depth $z \in \{17\,\text{cm}, 67\,\text{cm}, 86.5\,\text{cm}\}$. Below measurements of soil temperature, water content and precipitation.

χ^2_ν-value denoted by the blue color of the dots. All estimations with χ^2_ν lower than 0.01 are very accurate, i. e. the averaged residuum is lower than 0.1 °C. Such estimation results are of the order of measurement uncertainty.
The time-series $D_h(t)$ at depth 67 cm and 86.5 cm behave similar over time, whereas $D_h(t)$ at depth 17 cm varies much more in D_h and in χ^2_ν. The variations at depth 17 cm are based on measurement noise. Some measurement values are at depth 17 cm between days 889 and 966 which differ $1-3$ °C from the expected curve. During these warm months July and August, plenty of measurements points are large outliers for the temperature sensor at depth 17 cm due to temporal disturbance of this sensor. They are excluded by a 3σ-filtering and replaced by a linear interpolated value of their neighbors. But the uncertainty during this time is around 2 °C for the remaining points. Around the days 860 and 980, the same problem occurs but with a lower uncertainty of around 1 °C. This affects not only the residuum which increases, but also the estimated value of the diffusivity. In the case with lower measurement error, the influence on the estimated D_h is less than for the higher measurement noise since the number of noisy data points is much smaller. Therefore, we observe in Figure 4.10 a) a large increase in D_h. There, the noise is too strong to be sensitive to the few reliable measurement point in the objective function.
Representative for the two greater depths, we analyize the 67 cm depth. During most of time the χ^2_ν-value is quite low with an averaged residuum below 0.1 °C, colorized by blue to dark blue. In the following, we focus more on the times with higher temperature residuals which are marked by the numbers ①-④. A projection example to each of these four time zones is given in the appendix B.1. Usually, the residuum in these zones is quite high around ±0.4 °C and shows a clear trend during the 10 days of estimation time.
The first higher error in D_h, marked by ①, is produced by overestimating the trend in the projection (Figure B.1). The reason for this trend is based on the different thermal and hydraulic states at initial condition compared to the estimation interval. We know from the observations in Figure b) that at initial time, the soil is more wet with $\theta = 0.15$, the temperature profile is warmer with $15-18$ °C and no precipitation occurs. On the other hand, the soil is dryer with a constant $\theta = 0.10$ during the estimation and the temperature profile gets colder due to some rain events. But the water content does not rise since the roots of the plants take up this additional water or it is evaporated. These different conditions effect the thermal diffusivity since D_h is lower in dry soils than in less dry to more wet ones. During the estimation interval, the water content is constant in each depth, but θ varies with greater depth from 0.1 to 0.2. Consequently, the thermal energy for the measurement data travels through these different water content regions and thus with different diffusivities. But, the model assumes homogeneity of the water content distribution and fits one effective D_h to all of these heterogeneous water content regions. Further, the conditions for the z-integral changed since the whole profile is more wet at initial time. Due to the roots, the water content decreases and changes the D_h over time till the estimation interval. For both times, the same D_h is estimated and cannot fit both conditions well enough to reduce the residuum lower than ±0.2 °C.
For the next projection interval, marked by ②, the residuum is the highest one with ±0.4 °C in the D_h-time-series (Figure B.2). Similar to the previous example, the trend of the residuum is caused by the different condition at initial time compared to the estimation interval. Again, the initial condition is much more wet than later, when a part of the water evaporates and the rest is taken up by the plants from the first layer. But in this case, the water content in the estimation interval is not constant. After some rain events, not all water is used by the plants and the water content rises by 0.1 and 0.05 at depths 13 cm and 30 cm. This changes the diffusivity in the real profile and hence, also the estimated effective D_h. These changing conditions

in the estimation interval intensify the trend of the residuum. Therefore, the residuum of ② is in the bounds of $\pm 0.4\,^{\circ}\mathrm{C}$ instead of $\pm 0.2\,^{\circ}\mathrm{C}$ like in ①.
For the example ③, the condition is a lot different since the estimation interval is during winter instead of summer. During this time, the water content in all depths is high and constant around 0.22. The initial condition is comparable with the one during estimation time. But the soil temperature is low and becomes very close to zero in the uppermost sensor depth. In fact, the cooling affects also the deeper layer, but the thermal properties are changed when freezing starts. During these two days of the estimation interval, the additional physical processes cannot be described with the diffusion model. The effective diffusivity of all ten days averages over all states of the soil and produce a residuum of $\pm 0.2\,^{\circ}\mathrm{C}$.
Moreover, the data gap of $D_h(t)$ is marked by ④. Here, a data gap of two days exists and is linearly interpolated not only at 67 cm depth but also for the highest sensor at 12 cm depth. There, the model cannot match the measurement during this interpolated part. This does not yield reliable estimations of the D_h and is excluded.
However, the higher χ^2_ν-value can be observed more often during spring and summer when the water content changes clearly due to water uptake of plants, evaporation or strong rain events. Therefore, we focus on autumn and winter where the water content is quite constant. Over several months, the χ^2_ν-value is low and indicates accurate estimations, especially around the marker ❺ in Figure 4.10 a). There, a smaller dip of $D_h(t)$ is located in colored dark blue around the mean value of $D_h(t)$. At the time of this drop, the water content in the upper layer rises by 0.02 due to some rain. This variation is lower than the ones during summer. Here, the water flows into greater depths during December and is not taken out of the system due to plants. Therefore, a lower disturbance of the thermal and hydraulic conditions occurs due to the wettening. Further, the thermal conductivity K_h is less sensitive for wet soil than for dryer one. Since K_h rises slower than the thermal capacity C_h increases, the diffusivity increases for dry soils and decreases for wet soils (*Roth (2007)*). The time window with the increased water content is really within the estimation interval at the dip with marker ❺. In previous and later times, the time window with increased water content is only partly represented in the estimation interval.
This description of the D_h behavior indicates that the variation in $D_h(t)$ is based on the surrounding condition within the soil. However, the interpretation of $D_h(t)$ will become more difficult if we choose a time sequence with results closer to χ^2_ν-values of 0.005. it is really important for a feasible interpretation of $D_h(t)$ to have very low and comparable χ^2_ν-values. Otherwise, the very small variations in D_h get lost due to fluctuations of the estimation accuracy.

4.1.4. Summary

In this section, we applied the projection method using the transfer functions to measured soil temperature data from Grenzhof test site. Step-by-step, the mathematical tools were added to improve the initial description and thus the projection results. All these tools together reduced the error of the estimation result to the order of the measurement uncertainty.
Further, we analyzed the parameter space at one example. In direction of the estimated reference temperature, the minimum was unique. But in direction of the thermal diffusivity, we observed several minima. The analysis for small diffusivities showed that the local minimum can be avoided with a finer spatial resolution of the z-integral at depth. The z-integral changed the slopes of the parameter space for different projection depths. But the main structure stayed similar to other depths.

Moreover, the sensitivity of the projection result were studied referring to variations in the diffusivity. Doubling of the diffusivities scaled the residuum much more than a factor of two.
In order to test the convergence of the inversion, we started with different initial parameter sets. All these were chosen within a physically realistic range of the parameter values. The different runs came very close to the minimum within this range.
Moreover, we produced a time-series of estimated thermal diffusivities by shifting of the time interval. The time-series of the effective thermal diffusivity has a very low variation despite of quite strong changes of the water content. If the conditions within the soil varies too much, especially for involved additional processes like freezing, evaporation or water uptake by plants, the estimation results become more error-prone. Also noisy measurement data or interpolated data gaps are weak points for an accurate projection result. But there are variations of the diffusivity which have a very low estimation error and can be related to changes of hydraulic conditions, especially during autumn and winter. Moreover, the conditions within the soil should be similar to be able to interpret the variations of D_h, i. e. the water content can change, but only slightly. Furthermore, the number of active processes should be minimal for the interpretation of ΔD_h to relate the variation of D_h to changes of e. g. water content due to precipitation. Otherwise, the correlation is not significant enough or the estimation is too bad due to dominant non-conductive processes.

4.2. Fourier Transform

In chapter 2, we introduced the projection method based on Fourier transform besides the transfer function approach. In the following, the Fourier approach is considered to project soil temperature into depth. There, the basic idea is to decompose the surface time-series into a basis of periodic functions and project functions within the frequency domain into depth. The basis is chosen to be a set of complex exponential functions. This is a different choice of the basis compared to the transfer function approach with its δ-functions. Further, we know from the theory section 2.3 that the amplitude is damped exponentially and the phase is shifted linearly with larger depth. Finally, the projected temperature will be transferred into time domain (Figure 2.8). The estimation of the thermal diffusivity is realized by comparing the projected and the measured temperature at projection depth in the objective function. This is described by the minimization problem (2.43).
In the following, the Tukey window is included stepwise to demonstrate the improvement of the projection due to the application of the window.

4.2.1. Additional Tool for the Fourier Approach

In the following, we apply the project method by Fourier transform on a time-series of measured soil temperature. The data is measured at the Grenzhof test site in several depth (appendix A). The highest temperature sensor is located at 12 cm depth and represents the starting point of the projection as the surface-nearest time-series. At this point, we should mention that it introduces an approximation since the input time-series is assumed at the surface in the theory. This is analogue to our discussion for the transfer function where we used the same approximation. The time-series of soil temperature started in August 2004 for our measurement example and ended in January 2012. These nearly eight years corresponds to 2^{17} data points which are sampled each half hour.
In a first step, we apply only the basic procedure of the projection onto the near-surface data set (Sketch 2.8). Later, the influences of the window become obvious. In this context, we also

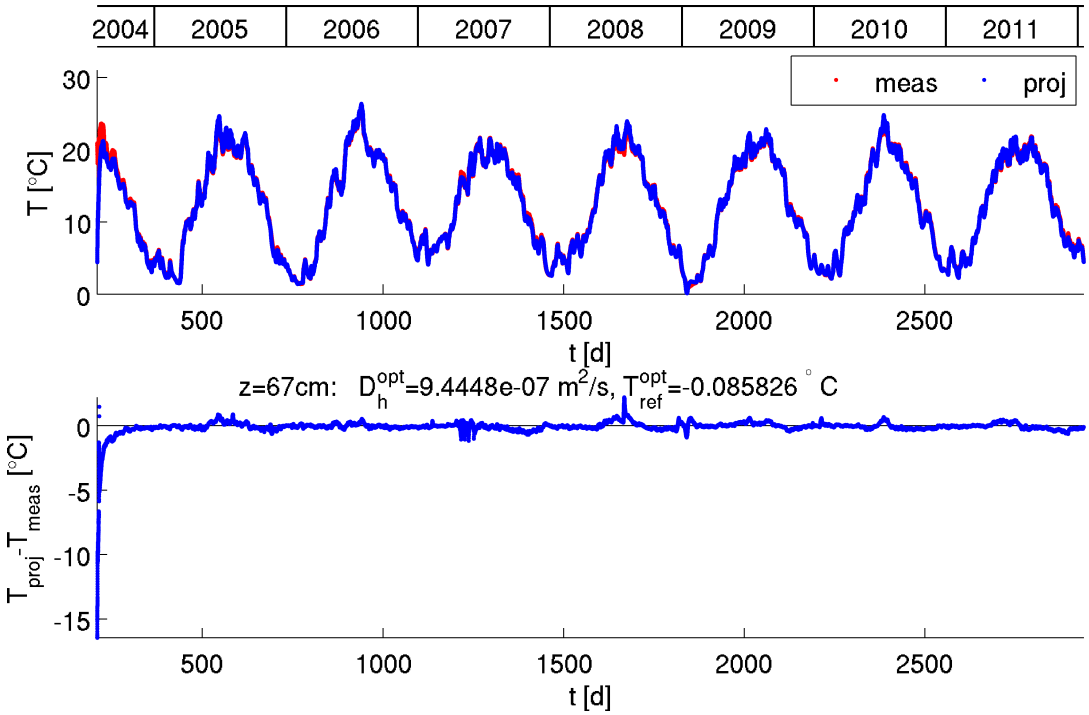


Figure 4.11.: Projection results: Top: The simulated soil temperature (blue) is projected from 12 cm into 67 cm depth using Fourier transform without any additional window. As reference for the estimation the measured soil temperature (red) is used. Bottom: The residuum between projected and measured temperature is shown together with the optimized parameters D_h^{opt} and $T_{\text{ref}}^{\text{opt}}$.

analyze the spectrum of the time-series. But for now, the focus is only on the projection result at 67 cm depth in Figure 4.11.

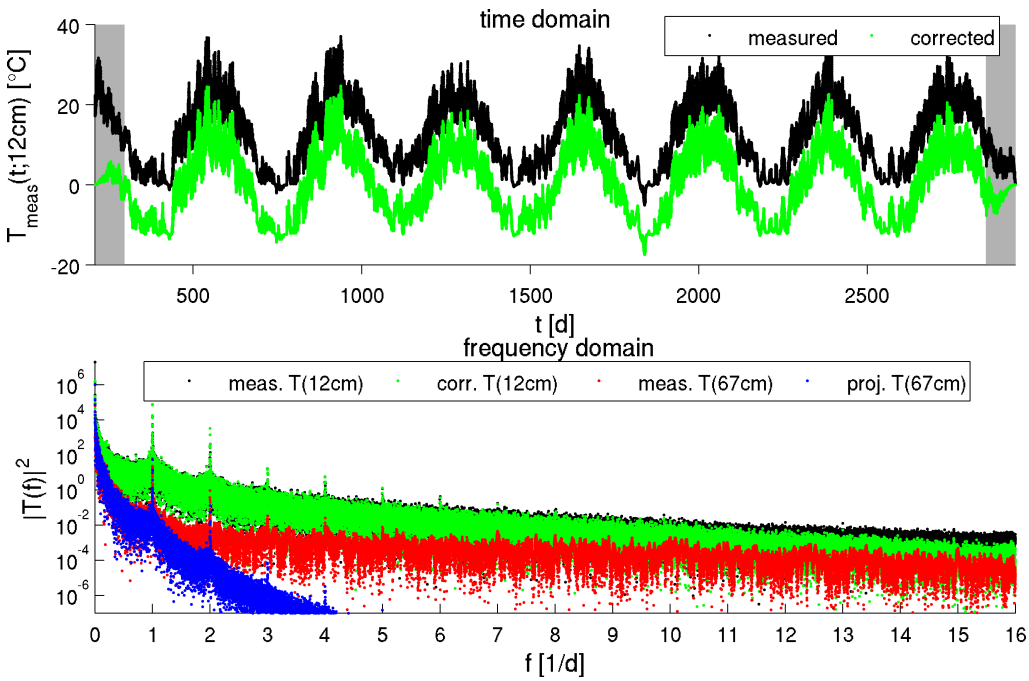
In the top Figure 4.11, the measured time-series at 67 cm depth is shown in red over almost eight years. The optimized projected temperature in the same depth overlaps with the measurement. The main characteristics of both time-series are equal like the cycles of the years and the days. These years show a higher error at a few times as given below in the residuum plot. The most dominant error is obviously during the first months in 2004. There, the difference between projection and measurement is over 15 °C. This error occurs due to the different heights of temperatures between the start and end time. Simulated time-series is calculated by discrete Fourier transform. This means that a finite periodic function is assumed with an infinitely extension in time. As explained in section 2.3, this produces a discontinuity at the bounds of the finite time-series and also has to be reproduced by periodic functions with primarily higher frequencies. However, these higher frequencies are damped more by the projection since the damping factor $\exp\left(-\sqrt{\omega_k/(2D_h)}z_p\right)$ reduces the amplitudes more for higher frequencies $\omega_k = 2\pi f_k$ than for low one. The consequence of this is that the jump at the boundary is not represented anymore after the retransform into the time domain. The start and end points of the time-series is then at a similar height.

This kind of error is based on the calculation and not on the measurement itself. Therefore, the estimated parameters are not reliable and they are adapted so that the whole residuum is minimized. This means that a local error in time produces an error in D_h and with that, a global mismatching of the whole time series. Finally, it results in a global trend in the residuum

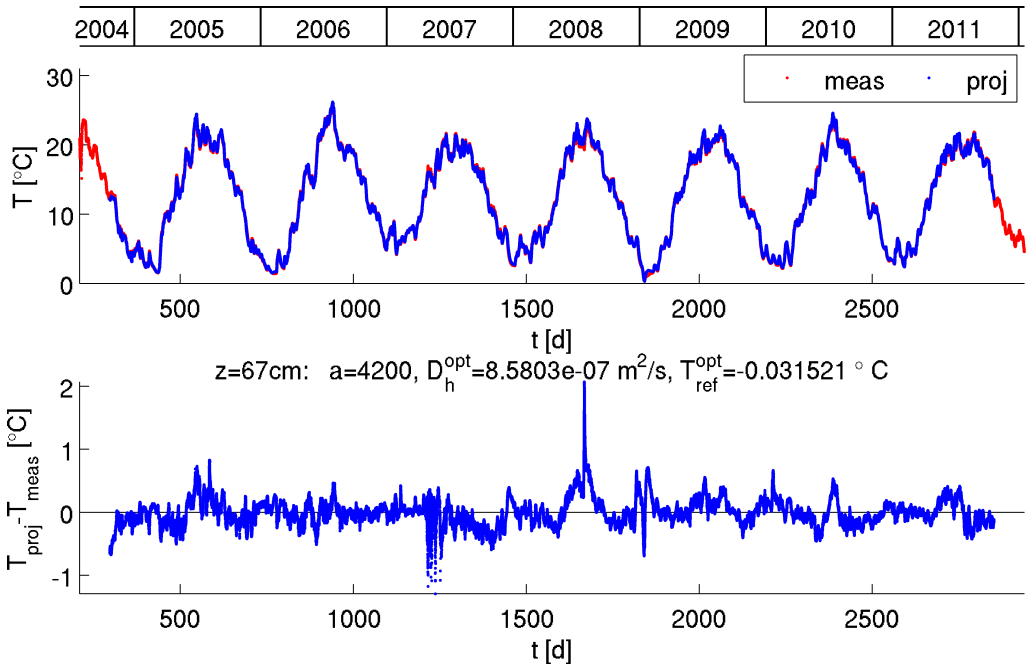
plot (Figure 4.11 bottom).
Therefore, the length should be chosen so that no jump or kink occurs at the time boundaries or we apply a window on the time-series before the transform. It allows the time-series to decay to zero at the sides.

Tukey Window

Above, we showed that without any further improvements of the projection strategy, the estimation error is too large and thus, the estimated D_h is not reliable enough. We will apply the Tukey window on the time-series before the Fourier transform to be flexible on the chosen length of the time-series as described in section 2.3.4.
Figure 4.12 a) displays the near-surface time-series on top. The black curve is the same which we used above for the projection. However, the green curve is corrected from the measured black time-series. This means that the mean value is subtracted to avoid a too high temperature entry at the smallest frequency later in the frequency domain. Finally, the Tukey window is multiplied. From section 2.3.4, we know that this window is constant to 1 except at the sides. There, it decays to zero like a cosine function. This decay interval is given by the parameter a which counts the number of points at one side within this decay interval. In this example, we choose a equal to 4200, i. e. 87.5 days. Later, we discuss the choice of the parameter a.
Moreover, we recognize the decay to zero towards the time bounds at the corrected time-series (green). The influenced interval is marked by the gray areas. Obviously, the subtraction of the mean value is given by the shift towards zero.
The Fourier transform brings the near-surface time-series to the frequency domain which is drawn with a logarithmic scale in Figure 4.12 a) below. The coloring of the different amplitude spectra corresponds to colors of the time-series. Generally, for all spectra we detect significant entries at frequency $1\,\mathrm{d}^{-1}$ and its decaying overtones. The amplitude becomes larger towards zero. Close to zero, we can find the highest entry at frequency $2.7 \times 10^{-3}\,\mathrm{d}^{-1}$ which belongs to the annual cycle. Since the spectrum decreases with larger frequencies, the frequency domain is zoomed in to the domain $[0, \frac{2}{3} f_c]$ with cutoff frequency $f_c = 24\,\mathrm{d}^{-1}$.
Moreover, the green spectrum overlaps with the black one for lower frequencies, i. e. the measured and the corrected spectrum for 12 cm depth is similar to each other in the logarithmic scale. But we also observe that the black spectrum has higher entries than the green corrected one, especially if we are looking from the linear perspective. The entries of the large peaks are further dominant but to the sides of the peaks, the spectrum decreases. The noise level is lowered with higher frequencies. The largest decrease occurs for the lowest frequencies where the amplitudes become lower by two orders of magnitude. Due to the application of the window, there is no discontinuity in the time domain anymore. There, the frequency components which represent this jump, have no contributions to the spectrum.
The other two spectra in blue and red belong to the temperature at projection depth. The red one denotes the measured temperature in the frequency domain at depth 67 cm. The first observation is that the amplitudes are much lower than for the near-surface quantities, especially the noise level which is spread over the whole frequency domain. Moreover, the increase of amplitudes towards zero occurs for much lower frequencies than in the black and green spectra. This indicates that the amplitudes are more damped for deeper time-series. Assuming white noise for the measurements, the noise is uniformly distributed (*Schlittgen et al. (1999)*). Larger amplitudes than this level correspond to physical information of the time-series. If we compare



(a) Top: Time domain: The measured near-surface temperature T_{meas} at 12 cm (black) is corrected to $T_{\text{meas}} \cdot g$ (green) by subtracting the mean value and applying the Tukey window, given by equation (2.47). The gray areas denote the non-constant parts of the Tukey window counted by $a = 4200$. Bottom: Frequency domain: The spectrum of the original measured (black) and corrected (green) time-series for 12 cm depth are drawn in logarithmic scale. Additionally, the spectrum of the measured (red) and projected time-series are shown corresponding to b) at 67 cm depth. The colors are chosen equally to the time-series in a) and b)



(b) Top: The simulated soil temperature (blue) is projected from 12 cm into 67 cm depth using Fourier transform and applying the Tukey window with parameter a, equation (2.47). As reference for the estimation, the measured soil temperature (red) is used. Bottom: The residuum between projected and measured temperature is shown together with D_h^{opt} and $T_{\text{ref}}^{\text{opt}}$.

Figure 4.12.: Projection results from depth 12 cm into 67 cm with optimized parameters using Fourier transform with Tukey window.

the measured time-series at 12 cm (black) and 67 cm (red) depth in Figure 4.12 a) and b), the daily fluctuations of the temperature and also the annual change are smaller for the two different depth. This is the damping which we observe in the spectra, e. g. the lower daily fluctuation at frequency $1\,\mathrm{d}^{-1}$.
Now, the corrected spectrum (green) is multiplied by the complex exponential function into depth (section 2.3.1) applying the projection in the frequency domain. The projection result in the frequency space is given in blue in Figure 4.12 a). The ratio of this projection (blue) and the corrected measurement (green) is the exponential function by construction. Here, the ratio is linearly dependent on the parameter

$$\tilde{\gamma} = \sqrt{\frac{\omega}{2D_h}} \tag{4.7}$$

in this logarithmic amplitude spectrum. $\tilde{\gamma}$ measures the strength of damping. Thus, the variation of the two spectra changes with the square root of the frequency in this logarithmic scale. The damping becomes stronger with higher frequencies which we can also observe in Figure 4.12 a).
Afterwards, the projected time-series will be fitted to the measurement time-series by optimizing the thermal diffusivity and the reference temperature, Figure 4.12 b). All in all, the estimation error is strongly reduced compared to the previous estimation without the window. Except one smaller interval, the residuum is within $\pm 1\,^{\circ}\mathrm{C}$.
However, what happens in the frequency space with the two compared quantities? Back in Figure 4.12 a), the blue and the red spectra are similar for low frequencies and differ a lot with increasing frequencies. In a first step, let us consider only frequencies below $1\,\mathrm{d}^{-1}$. There, the overlapping is quite good. Both spectra decay with a similar slope where the most important frequency entries are located, i. e. the most dominant ones like the cycles from years to days. They define the rough structure of the time-series and make sure that the two time-series fit quite well with a residuum which is more than 10 times lower than the original time-series.
In the second part of the frequency space, the two spectra differ a lot from each other. As mentioned above, the spectrum of the measurement is a superposition of the real signal with a measurement noise. Assuming this noise is white noise, its spectrum is constant. This is the difference to the spectrum of the projection. There, the signal is damped dependent on the frequency and it also holds for the noise of the near-surface signal. This means, the noise is damped due to the projection whereas the noise stays for the measurement signal at 67 cm depth. Therefore, variations between the blue and red spectra occur. Furthermore, the high frequency contributions later affect only the fine structure of the time-series.
Considering the residuum of the projection in Figure 4.12 b), a remaining error structure is still there around $\pm 1\,^{\circ}\mathrm{C}$ and with one larger error peak. This structure is based on several error sources which also affects the parameter estimation. Therefore, we focus on the error sources in more detail.

Error Sources

The total error of our projection is based on several kind of error sources, e. g. coming from the model description, the evaluation of each calculation step and the measurement itself. All these influence the projection result and thus, also the parameter estimation of the effective thermal diffusivity. Hence, we look in more detail at these kind of error types.
One error we have discussed already is the modeling of the initial condition. In our model, we have surface soil temperature and the temperature decreases with depth. Generally, this

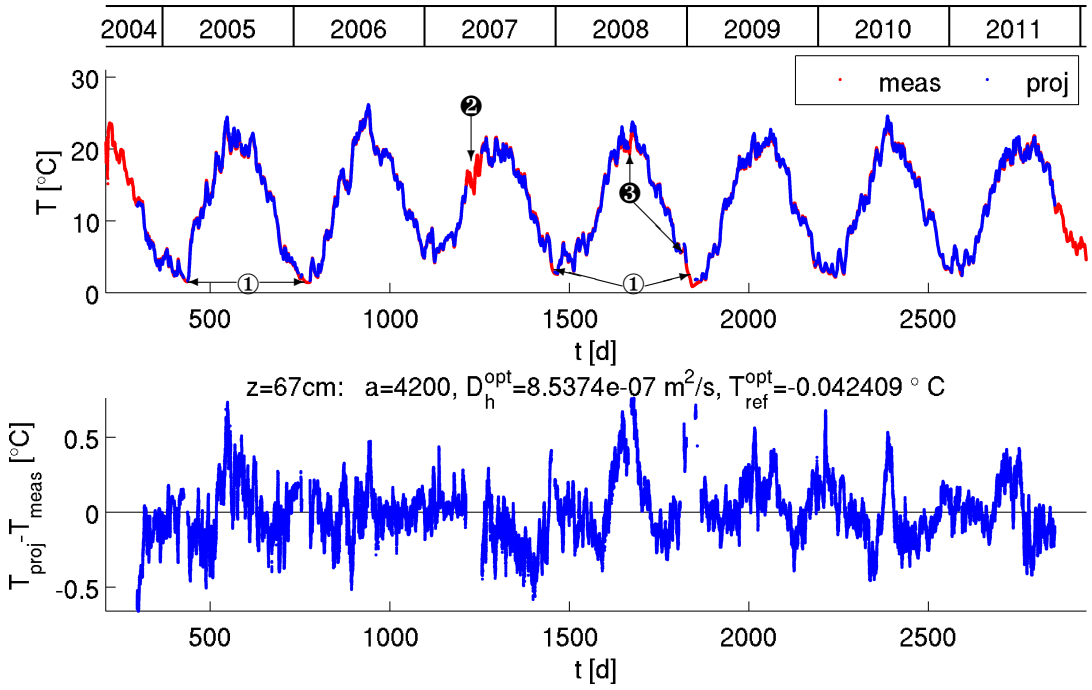


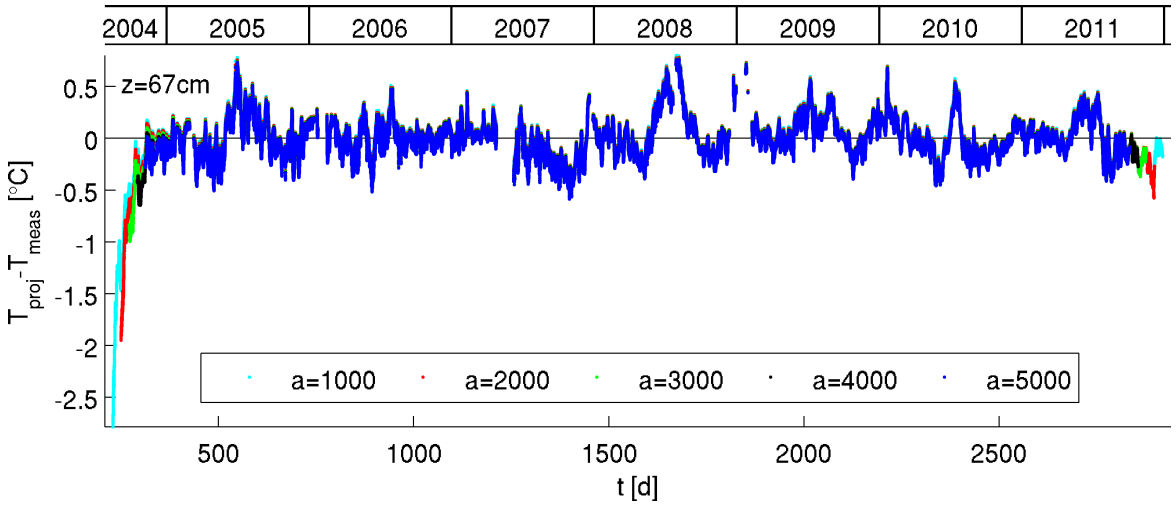
Figure 4.13.: Projection results: Top: The simulated soil temperature (blue) is projected from 12 cm into 67 cm depth using Fourier transform. The Tukey window is applied. In the objective function, time intervals are excluded which contain frost ① at 12 cm depth, high measurement uncertainty ❷ or linear interpolated measurement gaps ❸. As reference for the estimation, the measured soil temperature (red) is used. Bottom: The residuum between projected and measured temperature is shown together with the optimized parameters D_h^{opt} and $T_{\text{ref}}^{\text{opt}}$.

is not always the case in real soils. Even more, the initial temperature profile is not exactly. Therefore, the error is the largest at initial time combined with error due to the finiteness of the time-series in the context of Fourier transform.
Further, the previously introduced window has two advantages. On the one hand, the cutting of points at the temporal boundaries excludes the points with more errors from the initial condition. And on the other hand, noise is reduced especially for the higher frequencies and with that, real signals become more dominant. But the window decreases to zero with the decay towards the sides, also the important frequency entries. However, the noise reduction outweighs the decrease of important signals. All in all, the real information of the time-series is emphasized with using the window.
As we saw from the initial condition, there are further aspects which indicate differences between the model and the measurements. We assume only heat conduction in our model for the projection, but in real soils, there are also non-conductive processes. The freezing is one and there, thermal energy is used to change the aggregate state of the soil moisture. Therefore, only a part of the incoming heat will be transported conductively. Thus, if the soil is frozen near the surface but not in the projection depth, the soil properties will change during this time and the projection receives a larger error. This error is again tried to be minimized in the estimation and changes D_h so that the diffusivities are not correct in all parts of the time-series.
Therefore, we exclude all data points for which the soil temperature is zero or negative at 12 cm

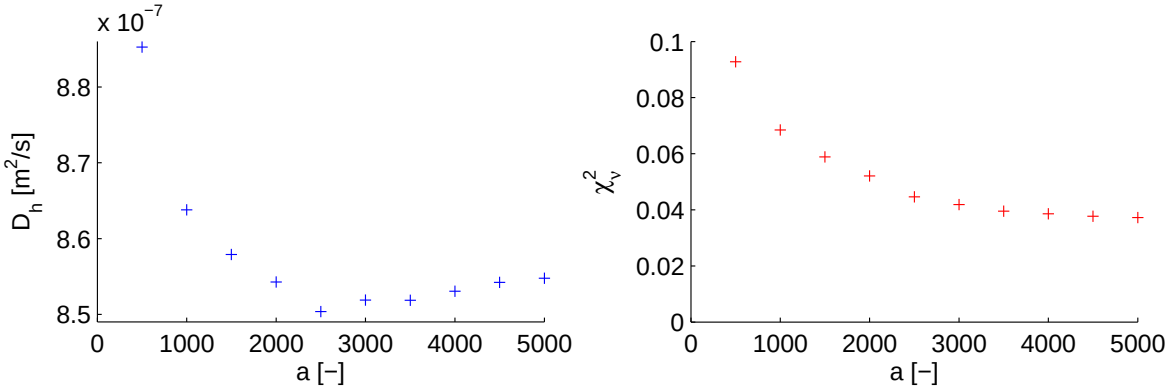
depth. In Figure 4.13, the projection result is shown excluding times with error sources. The marker ① displays the freezing times in several winters. Especially during the winter 2008/2009, the frost stayed several days at 12 cm and reached greater depth but not up to the projection depth. However, during this time the thermal transport did behave not only conductively. A relevant contribution to the total residuum was effected.
Similar to the winter, the residuum of the summer is more increased than for spring or autumn. Between the seasons, the conditions within the soil change. Whereas the layer distribution at Grenzhof stays the same, the layers in different depth contains more or less water over the seasons. Especially, the top most sensor at 12 cm depth measured the soil temperature in different hydraulic states but our model assumes constant soil properties in space and time. This discrepancy also contributes to the residuum in Figure 4.13. In contrast to projection using the transfer function approach, the assumed constant model conditions have to be valid over the whole 8 years instead of some days.
A further error source is that uncertainty is entered with measurements. At the example in Figure 4.12, we analyzed the noise level of the measured time-series in the frequency space. Generally, the measurement noise have a certain level over the whole measurement time. But additional measurement uncertainty can occur in time, e. g. with the excluded interval of over 40 days marked by ❷. There, a lot of outliers occurred which varied by over 20 °C only at 67 cm depth. A filtering of the time-series by excepting measurement with the 3σ-interval excluded a lot of points. But the remaining variation during these 40 days was still around $1-2\,°C$ which is visible larger than the general measurement noise of around 0.1 °C.
And finally, data gaps of the time-series can be a further error source. It can effect the spectrum of the time-series to work with non-equidistant distributed time-series. Several persons like *Jones (1971)* or *Parzen (1963)* studied time-series especially with missing data. However, we used linearly interpolated data for our projection to prevent such errors. The linear interpolation is a reliable interpolation for small gaps of some hours. On the other hand, the filled gaps produce remarkable errors for larger gaps with several days or weeks. An example for this is the large excluded error peak during summer 2008 marked by ❸ in Figure 4.13. Due to the non-linear damping of the amplitudes in the frequency space, the behavior of the projected time-series becomes non-linear within the previous linearly interpolated time-series whereas the measured and interpolated time-series at projection depth is linear by construction in the filled gap. As we observe at ❸, such an error referring to the interpolation can be around 1 °C. Therefore, we exclude this interpolated interval only for the estimation.
All together, these error sources influence the projection result and thus, the parameter estimation. Some of these errors or their influence can be reduced. Therefore, some time intervals are excluded from the objective function but the forward model itself uses the whole data set. This data exclusion changes the estimated diffusivity by 0.5 %. The reduced χ^2_ν-value is decreased from 0.046 to 0.038.

Impact of parameter a

After excluding the some error sources from the time-series in the objective function, we focus on the parameter a of the Tukey window in equation (2.47). This parameter counts the number of points which are used to decrease continuously from 1 to 0, respectively for each side of the total time interval. Further, we choose this number of points to exclude them from the objection function.
Several estimations are taken varying the value of a and the residuum of five of these estimations are drawn in Figure 4.14 a). Here, we observe a clear reduction of the error with higher values



(a) Residuum between projected and measured temperature at 67 cm depth: Parameter a is varied from 1000 to 5000. For each choice of a, the parameters are estimated.



(b) The estimated effective thermal diffusivity and the reduced χ^2_ν-value are drawn for different values of a.

Figure 4.14.: Variation of parameter a: This parameter a denotes the number of time points where the Tukey window is not constant one, respectively at each boundary.

of a during the year 2004. This is caused by the cutting of more data points from the estimation interval.

However, the slope of the residuum changes with variations of a. This change illustrates different diffusivities which are estimated under different conditions of a. As discussed previously, at the beginning of the time-series the projection error with any window is strongest since a discontinuity occurs at the bounds for the discrete Fourier transform as well as a mismatching of the initial condition between model and measurement. Such a large residuum at the beginning has to be reduced by estimating D_h if a is chosen too small. Figure 4.14 b) presents the estimated thermal diffusivities for some more values of a as well as the corresponding χ^2_ν-values. The changing number of points in the objective function is considered to make the χ^2_ν-values comparable for them. The error for early times becomes less important for the estimation with increasing of a since less data points with higher error entries are summed up in the objective function. Then, this can also be observed for the χ^2-value which decays in a similar curvature like the residuum at initial time.

Moreover, the thermal diffusivity decreases for a from 500 to 2500. Then, the damping in the

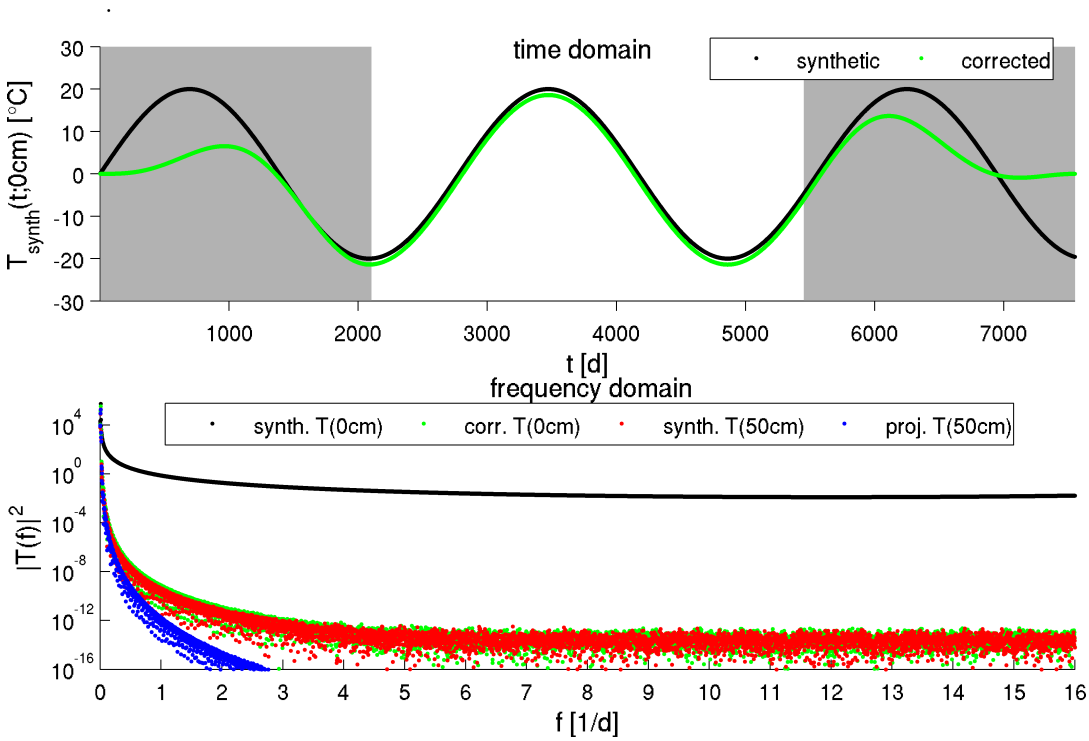
projection becomes larger due to equation (4.7) and decreases the amplitude of the projection. In total, the residuum $T_{\text{proj}} - T_{\text{meas}}$ is more negative. But why is the estimated D_h largest for the smallest a? We have to focus on the days of 2004 to analyze this relation. These days are most sensitive on the choice of D_h since the annual cycle is most disturbed there. If the D_h for $a = 500$ were smaller, then the enclosed area (Figure 4.14 a)) between the residuum and the zero-line starting at initial time till end of 2004 would give a large contribution to the objective function. D_h is increased to minimize it since the damping of the signals decreases.
The residuum at the beginning becomes a comparable magnitude for values of a larger than 2500. With that the choice of a is less important. But first, the sensitivity of the residuum referring to varying a is less recognizable for over 5000 data points. Our choice with a equal 4200 in our previous examples is chosen so that the initial error is excluded. Many data points are excluded and the χ^2_ν is varying not as strong any more with the choice of a. The parameter a can be chosen with a smaller value for other time-series, for which difference of height at the boundaries is not such large. But a should be at least as large that an error in the initial profile is vanished. Due to the consistence of concept with transfer functions and with Fourier transform we can use the spin-up estimation as an approximation for the time till the initial error vanishes from section 2.2.3. From Figure 2.3, we know for our projection depth that 70 % of the thermal energy reaches the depth after at least 10 days. We here use much more with 87.5 d due to the error of the discontinuity.

Synthetic Data Set

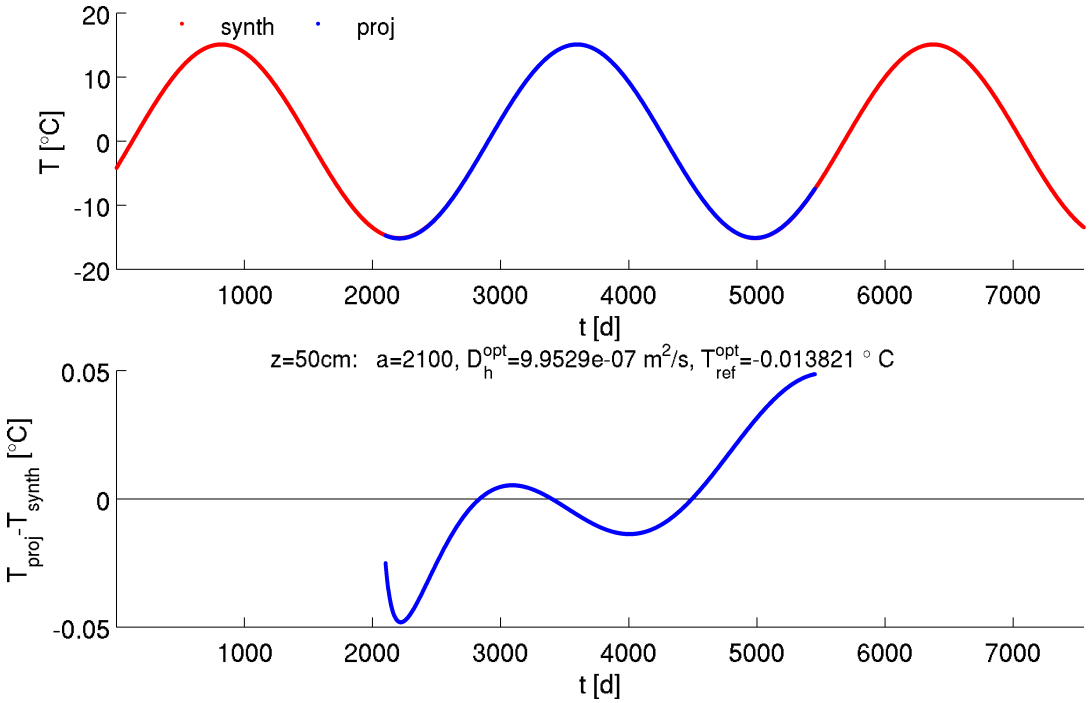
In order to confirm a general accurate estimation by applying the Fourier approach, we analyze a synthetic data set in the following. The synthetic data is the same data as *Ludin (2010)* used for the synthetic analysis with the transfer function for a homogeneous medium. The surface time-series is given by

$$T_{\text{synth}}(0, t) = A \sin(2\pi f t) \tag{4.8}$$

with an amplitude $A = 20\,^\circ\text{C}$ and frequency $f = 10^{-7}\,\text{Hz} \approx 8.64 \times 10^{-3}\,\text{d}^{-1}$. This sine function is drawn in black in Figure 4.15 a). Since the resolution of the synthetic data is only one hour, we reduce the Tukey parameter a to 2100. After subtracting the averaged temperature and applying of the window, it results the corrected time-series in green. This curve is shifted slightly due to the subtraction. With the gray area defined by a, the time-series is damped to zero towards the sides by the window cosine.
The transform of both time-series into the frequency space yields the black and green spectra in Figure 4.15 a). Here, we observe that the black spectrum of the synthetic time-series is much higher than all other spectra. The reason for this is based on the discontinuity of the finite time-series since this time-series is used as infinitely periodic extended for the discrete Fourier transform.
However, all spectra have a significant peak for the given frequency in equation (4.8) which is located beside the zero-frequency in Figure 4.15 a). Other values of frequency are not dominant, especially in the spectrum of the corrected temperature. This is slightly higher than the red spectrum of the corrected synthetic time-series at 50 cm depth. This small difference is the damping of the signal with depth for an assumed soil with $D_h = 10^{-6}\,\text{m}^2/\text{s}$. It becomes more dominant with smaller frequencies. For larger frequencies, both reach a similar level in the amplitude spectrum of around 10^{-13} to 10^{-14}. Since we calculate with temperatures values which are given up to the seventh decimal place, the low level of the spectra refers to the computational accuracy. This would be sufficient for real data since the measurement noise is



(a) Synthetic surface soil temperature (black) at 0 cm depth and the time-series after subtracting the mean value of T_{synth} and the application of the Tukey window with $a = 2100$. Below, the amplitude spectra of all time-series in a) and b) is summarized.



(b) Projected and synthetic soil temperature at 50 cm depth and its residuum below.

Figure 4.15.: Projection of synthetic temperature with a temporal resolution of one hour. For the production of the synthetic data, a soil is assumed with $D_h = 10^{-6}\,\text{m}^2/\text{s}$ and the surface temperature is given by the sine function of equation 4.8.

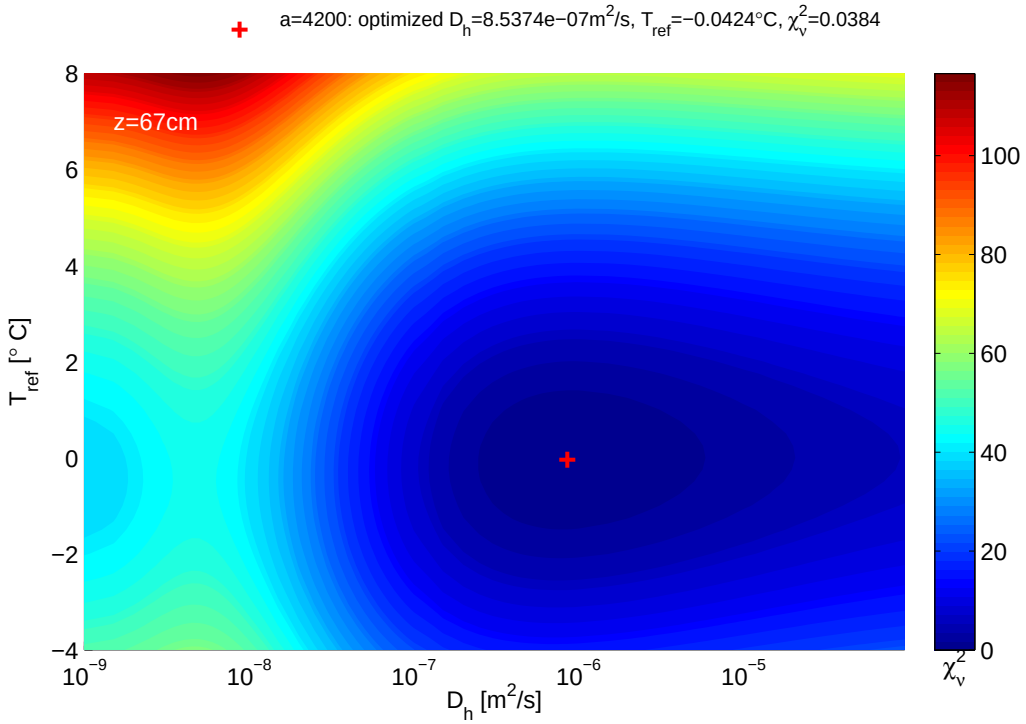


Figure 4.16.: Parameter space of the minimization problem (2.43) using the χ^2_ν-value based on the Fourier approach. The reduced time-series with almost eight years is used like in Figure 4.13 for a projection into $z = 67\,\text{cm}$ depth and a parameter $a = 4200$. The optimized parameter set is marked with a red cross and the corresponding optimized parameters are given together with its χ^2_ν-value above.

some orders higher of magnitude.
The corrected surface temperature (green) results by multiplication of the complex exponential function to blue spectrum in the frequency space. The damping is quite small for very low frequencies but it fits the red spectrum very well. However, both spectra spread for increasing frequencies. Here, the projected spectrum is damped below 10^{-14} since the influence of the exponential function scales down the blue spectrum. This is done with a higher computational accuracy.
The resulting projection itself is transferred back into the time domain in Figure 4.15 b). This projected temperature matches very well the synthetic data at 50 cm depth. Their residuum is around $\pm 0.05\,°\text{C}$ which is ten times smaller than in the case of real measured data. The small trend comes in the residuum due to the estimation error of the thermal diffusivity. This small underestimation of D_h can be based on the use of the window. Each application of a window influences the resulting spectrum. Usually, they are used to strongly reduce the noise level but this can also slightly change the real signal. All in all, the error of the projected temperature and of the parameter is quite low.

4.2.2. Scanning of Parameter Space

In the following, we look in more detail onto the shape of our parameter space to get an impression of the possible starting parameters. The question of uniqueness of the estimated minimum is connected with this analysis. For that reason, the analyzed quantities are the objective func-

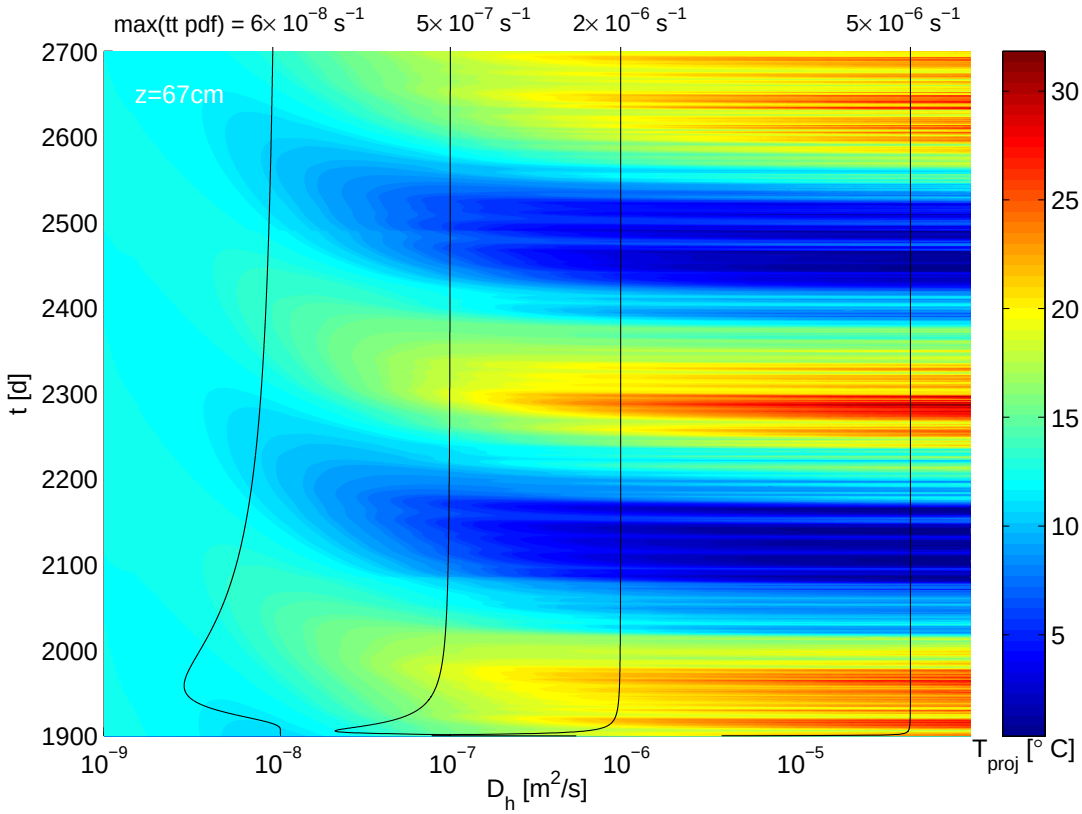


Figure 4.17.: Projected temperature T_{proj} (colorbar) is shown as a function in D_h and time with the fixed $T_{\text{ref}}^{\text{opt}} = -0.0424\,^{\circ}\text{C}$. The diffusivity is counted logarithmically and the time interval is zoomed in. A vertical cross section through the two-dimensional plot yields the projection in Figure 4.13. The 4 black curves denote the travel time pdf's for the thermal diffusivities $D_h \in \{10^{-8}\,\text{m}^2/\text{s}, 10^{-7}\,\text{m}^2/\text{s}, 10^{-6}\,\text{m}^2/\text{s}, 5\times 10^{-5}\,\text{m}^2/\text{s}\}$. These transfer functions are drawn at different scale with the maximal amplitudes $6\times 10^{-8}\,\text{s}^{-1}, 5\times 10^{-7}\,\text{s}^{-1}, 2\times 10^{-6}\,\text{s}^{-1}$ and $5\times 10^{-6}\,\text{s}^{-1}$.

tion for the minimization problem (2.43) scaled by the difference of the length of the time-series and the number of parameters. In this case, the scale is equal since the example in Figure 4.13 is taken with a fixed number of exclusions.

The determined surface of the parameter space is given for projection depth 67 cm in Figure 4.16 with the diffusivity in a logarithmic scale. The optimized parameter set with $D_h^{\text{opt}} = 8.5374\times 10^{-7}\,\text{m}^2/\text{s}$ and $T_{\text{ref}}^{\text{opt}} = -0.0424\,^{\circ}\text{C}$ is located within the lowest observed region of the parameter space. Thus, the optimization algorithm reached at least a local minimum.

In the direction of the reference temperature, the behavior of the χ^2_ν-surface is almost symmetric and monotonically decreasing towards the optimized T_{ref}. In T_{ref}-direction, the minimum seems to be unique. This can be confirmed if the linear dependency of the reference temperature and the projected temperature is considered. The T_{ref} is an additive term and a change of it causes a quadratic variation of the objective function due to the least squares approach. We expect the uniqueness of the minimum in direction of T_{ref} with these considerations.

Furthermore, the uniqueness and the monotonicity is not clear in the direction of the thermal diffusivity D_h. Even in a low neighborhood of the optimized D_h between $10^{-7}\,\text{m}^2/\text{s}$ and $10^{-5}\,\text{m}^2/\text{s}$, the behavior is not symmetric anymore since D_h depends non-linear on the pro-

jecting exponential function. Therefore, the uniqueness of the minimum is not given automatically. We observe it for low diffusivities. A saddle point in the χ^2_ν-surface can be detect around $10^{-8}\,\mathrm{m}^2/\mathrm{s}$ and for still smaller D_h the χ^2_ν-values again decreases. We have to analyze further the projected temperature referring to variations of D_h to understand this behavior.
Therefore in the following, we study only the projected temperature T_{proj} and the measurement data is exclude from the analysis. The time-series of the projected temperature is vertically plotted in Figure 4.17 for the $T^{\mathrm{opt}}_{\mathrm{ref}}$. Into 800 days of almost eight years of the time-series, it is zoomed in in order to observe more details. In horizontal direction, we vary D_h on a logarithmic scale. We detect the annual cycle in more than two runs with the coloring of the projected temperature. The more detailed vertical strips within the more blue or red regions are points in time and the change between them represents the daily fluctuations.
Further, our optimized projected temperature is located slightly below $10^{-6}\,\mathrm{m}^2/\mathrm{s}$. In this region, the observed strips start to decrease visibly. The damping of the temperature amplitudes becomes stronger with lowering the diffusivity like the damping factor $\tilde{\gamma}$ (4.7) of the exponential expression shows. This damping of the amplitudes is expressed in Figure 4.17 by a decrease of the colors vertically.
Besides this damping, a phase shift of the decomposed complex functions occurs as given in equation (2.37) by $-\tilde{\gamma}z$. Since the damping factor $\tilde{\gamma}$ increases with lower diffusivities, the phase shift becomes more dominant with lower D_h. This phase shift is responsible for the colored stripes curves for low diffusivities towards higher times. Thereby, the shape of the curvature of the phase shifted stripes follows the behavior of $\tilde{\gamma}$ with D_h.
Another perspective realizes the transfer functions or the travel time pdf's which are drawn for 4 different values of D_h as the additional black curves in Figure 4.17. The theory of the pure transfer integral and the Fourier approach are based on the same mathematical model and assumption to the soil. Therefore, both approaches evince the same structure in the resulting projection.
The transfer functions with $D_h = 10^{-8}\,\mathrm{m}^2/\mathrm{s}$, $10^{-7}\,\mathrm{m}^2/\mathrm{s}$, $10^{-6}\,\mathrm{m}^2/\mathrm{s}$ and $5\times10^{-5}\,\mathrm{m}^2/\mathrm{s}$ show similarities to the previous analysis. The maximum of these functions runs clearly further in time when D_h strongly decreases. We can understand this by remembering the background of the travel time pdf in chapter 2. This function describes the distribution over time of a δ-pulse which is initiated at the surface and travels through the soil. Now, we observe different soils in one fixed depth with different diffusivities.
Analogue, we can interpret these four functions in Figure 4.17 in another way. Since the travel time pdf in (2.10) depends on the ratio z^2/D_h, the four functions can be related to one fitted soil in four different depths, i. e. small values of D_h can be interpreted as high depth and vise versa. Hence, the amount of thermal energy needs more time to travel into greater depth and therefore, the maximum of p_t occurs at a later time which we can see at the curve located at $D_h = 10^{-8}\,\mathrm{m}^2/\mathrm{s}$. If we compare these four pdf's with the colored plot, the same structure can be observe, i. e. the same curvature in the maximum curves. A change of the behavior occurs at the same values of D_h around $10^{-7}\,\mathrm{m}^2/\mathrm{s}$ for the colored plot as well as the pdf's. There, the almost constant behavior for large D_h values varies to one changing in time for low D_h. The phase shift becomes more and more important for values of D_h lower this boundary. The pdf's also show the damping effect by their maximal amplitudes which are printed at the top of Figure 4.17.

Summary of the Parameter Space

Our analysis of the parameter space showed that the parameter space is well-formed within the physically relevant interval of the thermal diffusivity so that a convergence to the one local minimum is given for the starting parameters with this domain. Further, we found an unique minimum in direction of the reference temperature. Only in the direction of D_h, the objective function did not only increase monotonically away from a local minimum. A further analysis of the behavior of D_h showed that additional local decrease of χ^2_ν for low diffusivities is caused by the dominant effect of the phase shift of the projected periodic functions.

4.2.3. Temporal Evolution of Effective Thermal Diffusivity

After the analysis of the projection for time-series over some years, we apply the projection method by using Fourier transform onto an estimation interval of six months. As previously, we focus on how the estimated thermal diffusivity behaves in the temporal evolution with a moving interval like in section 4.1.3.
Further, we assume that the diffusivity is temporally and spatially constant for the whole estimation interval. This is a strong assumption for the time interval of eight years. If we look back to Figure 4.13, we observe that especially in winter and summer the residuum gets more errors in a range of $\pm 0.6\,^\circ\mathrm{C}$. Thus, separate estimations within one year can be useful in order to analyze the temporal variations of D_h.
The results of the six month moving interval are given in the first row of Figure 4.18. At the mid points of the intervals, the estimate D_h is drawn colored by its χ^2_ν-value. Although χ^2_ν is normalized with respect to the length of the time-series, here the χ^2_ν-values are three times larger than for the transfer function method. The reason for this is that the much longer time intervals cannot fulfill the assumption of constant properties in the soil in so an accurate way. Thus, the residuum for these six months intervals can also be around three times larger than for the ten days interval with the transfer function.
The estimated diffusivity for the interval of eight years is drawn as a blue dashed line in Figure 4.18 for comparison. This value is similar to the mean of all single estimations. Considering all years, a periodicity of this time-series of estimated D_h is observable. The vertical lines mark the transition between years and highlight this periodicity of half a year. We zoom into year 2006 and 2007 to analyze it in more detail which is drawn in the second row of Figure 4.18.
The points with the black crosses ①-③ are exemplary for all other estimated points of D_h. Each of them stands for an interval of six months (e. g. black dashed lines) and the $a = 4200$ points to each side due to the Tukey window. All these three examples have a similar order of the error with $\pm 0.4\,^\circ\mathrm{C}$ as observed in the appendix B.2.
A first interpretation of the temporal evolution of D_h shows a seasonal behavior of the D_h with time. This behavior is strong for the used time-series length since the frequency $0.5\,\mathrm{a}^{-1}$ has a high entry in the spectrum and frequency $1\,\mathrm{a}^{-1}$ is not resolved within the spectrum. During winter and during late summer, D_h is lower whereas in spring and autumn D_h is larger. If we look at the estimation interval ②, the main trend of the measured temperature is monotonically decreasing as shown in the third row of Figure 4.18. Following in the same time interval one row below, we observe the measured water content at three depths and the measured precipitation. There, we recognize that the water content is wet approximatively around 0.2 volume fraction or more after rain events within the interval ②. The difference of θ between the measurement depth is only 0.05. This indicates a higher and similar thermal conductivity in all three

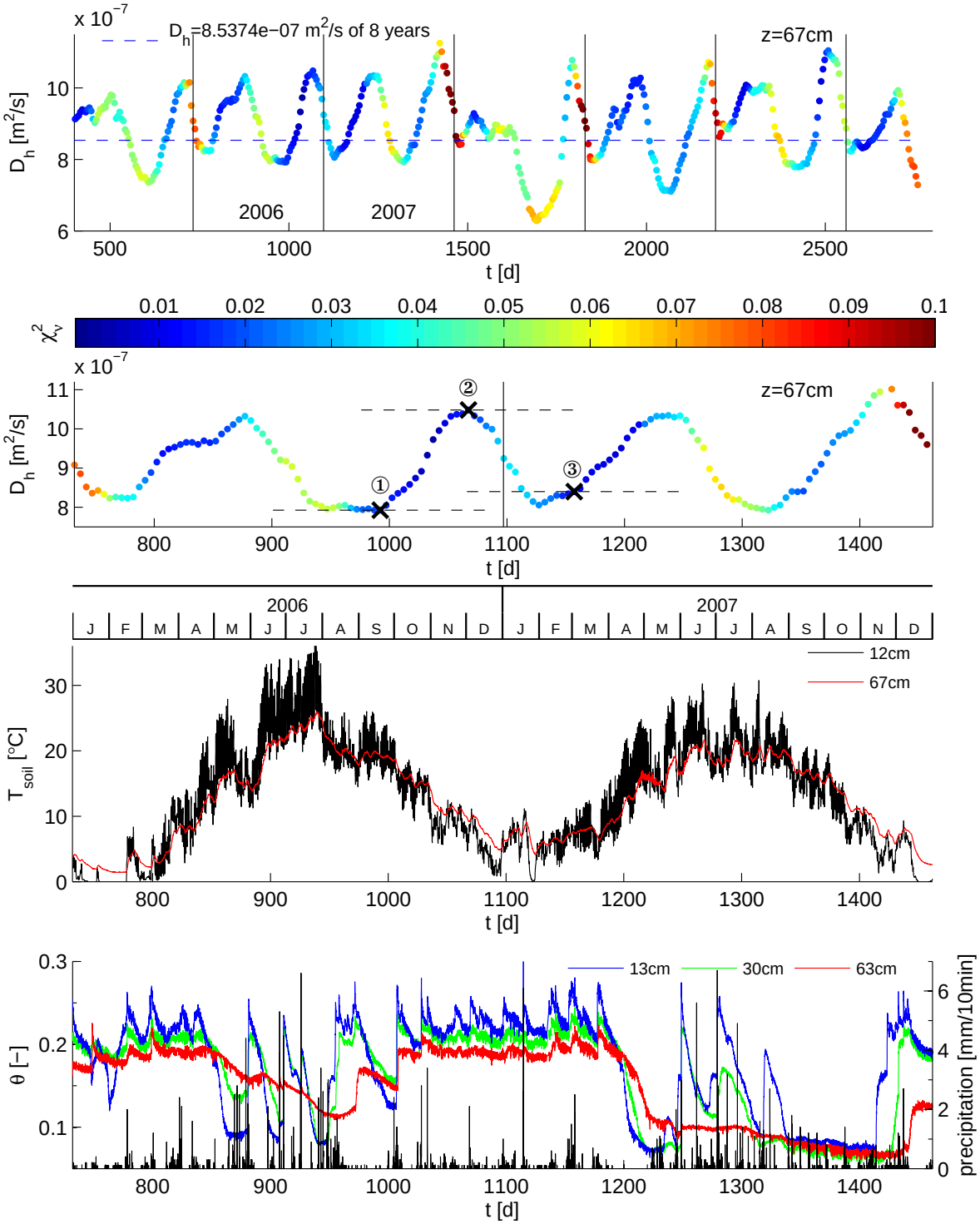


Figure 4.18.: 1. Row: Temporal evolution of the estimated D_h: Each colored dot denotes the mid point of an interval of six months and the color is given by the χ^2_ν-value. The optimized D_h for 8 years from Figure 4.13 represents the dashed blue line. 2. Row: Zoom of the first row into years 2006 and 2007. The six month interval is shown as a dashed line for the points with black crosses. 3. Row: Measured soil temperature at 12 cm and 67 cm depth for the same two years. 4. Row: In colored curves the measured water content θ is drawn for depths 13 cm, 30 cm and 63 cm. The measured precipitation (black) is counted at the right axis.

measurement depths which yield a higher diffusivity.
Going to previous times at marker ①, only the second part of the interval contains the higher water contents. During the first part from July till September 2006, the water content varies a lot not only in time but also with depth. The fraction of times with lower water content is risen from interval ② to ①. With that, the effective diffusivity falls.
Analogue, the interval with marker ③ also contains a larger variation of the water content since during April 2007, the soil dried out and after May, the plants prevented the water recharge of deeper layers. This lower and spatially varying water content caused the effective thermal diffusivity to decrease.
The seasonal behavior of D_h vanished by shortening the simulation intervals. Further estimations showed that the seasonal structure still existed for 3 months but for 1.5 months, variations were observed but on much smaller time scales. Due to the vanishing to shorter times, it cannot be observed in our transfer function example of 10 days.
Summarizing, we observe a variation of the effective thermal diffusivity with a periodicity of half a year. However, the error of the estimation depends on the temporal conditions e. g. of the soil and the measurement as the coloring with the χ^2_ν-values illustrates. The conditions of the soil changes within such time intervals so that residuum is slightly below the one of the eight years' projection but not around the measurement uncertainty.

4.2.4. Summary

In this section, we applied the Fourier approach onto a long real data set. The discrete Fourier transform required the use of a window in order to reduce the noise level of the spectrum. We chose the Tukey window since it scales only the side of the near-surface time-series. These influenced points in time were excluded for the optimization procedure but not for the forward model itself.
Furthermore, we discussed several sources of error like freezing in the soil, measurement uncertainty or linear interpolated data gaps. The choice of Tukey's parameter a affects the error at the beginning of the time-series for lower values of a. Thus, it then influences the estimated effective thermal diffusivity. Moreover, the analysis of a synthetic data set showed that the residuum can be more than ten times lower if measurement errors are not included.
The analyzing of the parameter space showed that a local minimum is observed in D_h-direction and within the physically relevant region from $10^{-8}\,\mathrm{m}^2/\mathrm{s}$ to $10^{-4}\,\mathrm{m}^2/\mathrm{s}$, it is unique. We discussed that the minimum is unique for the direction of the reference temperature.
Finally, the temporal evolution of D_h for a moving interval showed a periodicity of six months. We saw that this periodicity is effected by the soil moisture conditions but also by the dominant entries of the spectrum. But also the error of the estimation varies with these changing conditions.

4.3. Comparison of Methods

In the two previous sections, we applied the projection theory of the transfer function as well as the Fourier approach on measured soil temperatures from a test site close to Heidelberg. Both methods follow the approach of first solving the special heat equation analytically and then evaluating the solution numerically. However, the methods use different ideas to solve the differential equation.
For the transfer function, the surface time-series is decomposed into a basis of δ-functions. Each of these travel time pdf's starts at different points in time and describes the propagation

of the δ-function within the soil and in time. This means that the δ-function spreads and its amplitude is damped with time and depth.
By the Fourier approach, the surface time-series is represented as a basis of complex exponential functions. Each of these globally distributed functions in time produces a sharp local entry in the frequency space. The projection itself is applying in the frequency space by multiplying of an exponential function. Finally, the inverse Fourier transform yields the projected time-series.
Both methods start with a decomposition of the surface time-series. But in the case of real soils, the initial profile is usually different from the model assumption and it can vary with time. Forward in time, the contribution of the initial profile to the projected temperature decreases. Hence, approaches exist for each method to reduce the initial error. In case of the transfer function method, a combination of two approaches shows the best results, a spin-up phase and the z-integral. During spin-up, the calculation of the projection is done but it does not contribute to the estimation. This phase goes on till the initial error is not dominant anymore. Furthermore, the z-integral determines the amount of thermal energy at the initial profile. Temperature measurements are needed for it at initial time and over several depths. The spatial resolution of the measurements in the profile strongly influences the quality of the initial description. Therefore, a combination of the spin-up phase and the z-integral is very useful in order to reduce the error due to the initial condition.
On the other hand, we have to deal with an error for the Fourier approach due to the initial profile as well as due to a discontinuity of the time-series between starting and final time. For both reasons, the window is introduced and make the time-series smooth to the sides. Afterwards, the main effected sides of the time-series are ignored for the estimation. This can be interpreted as a similar approach to the spin-up of the transfer method. Including this tool improved the projection result and the parameter estimation.
All together, the Fourier method calculates the projected temperature much faster than the transfer function method, especially for longer time-series. Short time-series can be determined in a similar order of time. The reason for this is based on the fact that the projection integral of transfer function method has to be evaluated chronologically in contrast to the Fourier method. The transfer integral or the numerical approximation of the integral has to be evaluated for each time point and its numerical sum grows with each point. Therefore, the calculation of the sum becomes more expensive in computational time. However, the Fourier transform can be determined very efficiently by the fast Fourier transform. Moreover, the decomposition into exponential functions does not need chronological determination. Therefore, errors locally in time can influence the projection globally by introducing errors into the frequency space.
Otherwise, a local error of the surface time-series disturbs the projected time-series only locally. Due to the chronology, this means that the local temporal error reaches the projected time-series with a time shift and this error decays with time.
Moreover, these aspects have no significant influences onto the parameter space within the physically interesting domain. This means we observe a local minimum for both methods. The parameter space shows an unique one within the region of D_h from $10^{-8}\,\mathrm{m}^2/\mathrm{s}$ to $2\times10^{-5}\,\mathrm{m}^2/\mathrm{s}$.
Furthermore, both methods have to deal with the approximation that in reality the sensor does not measure exactly at the surface. Practically, the sensor is located a few centimeters below the surface. But after some time, thermal energy which travels towards the surface and is partly reflected, is detected later. Therefore, a spin-up phase or the window also deals with this problem, respectively for both methods.
The analysis of both methods shows that the transfer function method has more advantages for short time-series of several days. The Fourier method is more useful for longer time-series from some months till years. The main restriction of the transfer function approach is the compu-

tation time in practice. On the one hand, the evaluation of the transfer integral increases with longer time-series, on the other hand, the error increases referring to assumed constant soil properties. An used simplification is that the soil is assumed to be uniform. In reality, there are soil layers and they behave differently to incoming rain water. This water has also affects on the thermal dynamic within the soil. All processes which have influence on the soil temperature and cannot be described by conductive heat transport, show their error contribution in the residuum. Especially, the soil moisture varies for larger time interval of the simulation. The transfer function method as well as the Fourier approach have to deal with such varying conditions. If the total error increases for long time-series due to varying condition within the soil, then the Fourier approach will be of advantage.
The estimation of the interval of almost eight years by the Fourier approach runs faster than the 10-days estimation using the transfer function method. But the residuum of the 8-year estimation is significantly larger than the measurement noise with a values between $\pm 0.5\,^{\circ}\mathrm{C}$. We expect to have errors at least up to $\pm 0.5\,^{\circ}\mathrm{C}$ due to physical and methodical reasons with the Fourier method. The physical ones contain the varying condition within the soil but also local increased measurement noise in time. The filtering of outliers of such measured time-series leads to interpolation of data gaps. So methodical errors arises in both cases, interpolating to an equidistant time grid or keeping the gaps to a non-equidistant grid. Further, the window decreases the noise of the spectrum but also slightly changes the real information. All together, they influence the final error of the projected temperature.
Moreover, we have additional challenges with the Fourier approach in the context of short time-series if the most important frequencies cannot be resolved by given time-series. For instance, a time-series of a few days can show a trend belonging to an annual cycle. But the frequency entry belonging to this cycle is not contained in the spectrum.
The estimation interval should be similar to compare the estimated effective D_h from both methods. Thus, we average the $D_h(t)$ of the transfer function method over the same six months as in the estimation with the Fourier approach. The results are given in Table 4.2. The two markers ① and ② of Figure 4.18 illustrate the extrema of the $D_h(t)$-values and the means of their D_h is representative for the averaged D_h in the whole year. The variation of the estimated D_h with the Fourier approach are stronger than the small changes using the transfer function method. These stronger variations are based on the cycle of the six months which we discussed above. By averaging these two times, the means of D_h become similar but not

Table 4.2.: Comparison of the estimated D_h for both methods, transfer function and Fourier transform approach: The markers ① and ② belong to Figure 4.18 and represents typical D_h-estimations with maximum and minimum.

method	transfer function	Fourier transform
$\langle D_h \rangle_t$ ① $[10^{-6}\,\mathrm{m}^2/\mathrm{s}]$	1.029	0.793
length of average interval [d]	181.7	182.5
number of D_h-points	183	1
index of the intervals	$394 - 576$	123
$\langle D_h \rangle_t$ ② $[10^{-6}\,\mathrm{m}^2/\mathrm{s}]$	1.026	1.048
length of average interval [d]	181.7	182.5
number of D_h-points	183	1
induces of the intervals	$473 - 651$	138
mean of D_h between ① and ② $[10^{-6}\,\mathrm{m}^2/\mathrm{s}]$	1.028	0.921

equal. The offset between the mean values of D_h between both methods is still $10^{-7}\,\mathrm{m}^2/\mathrm{s}$. At this point, the residuals of both methods have to be compared. The residuum from the Fourier approach is at least $\pm 0.5\,^{\circ}\mathrm{C}$ whereas the residuum from the transfer function method is down to $\pm 0.05\,^{\circ}\mathrm{C}$ in the shorter time intervals. Thus, if the residuum is larger during a longer time interval then the estimated D_h includes more varying states and compensates it by a changed value during the estimation. Hence, we would expect to have similar but not equal estimated values of the effective thermal diffusivity.

Summarizing, both methods can be used to project soil temperature into depth. The final error depends on how accurate the assumptions of the soil properties are fulfilled and how accurate the given data set is. A final error was reached of the order of the measurement noise for the transfer function method. These low errors could only be ensured by including the additional mathematical tools. Further, the projection results were produced also by projecting over layer boundaries. This shows that the simplification of an uniform medium can be a suitable approximation. Even the variation of soil moisture yields residual of order $\pm 0.5\,^{\circ}\mathrm{C}$ for both methods.
Further, the Fourier approach is more powerful for long time-series whereas the transfer function method has more advantages for short time-series. Both methods showed advantages. The choice of the method for a given application depends on the purpose of use. The Fourier approach is more useful with long time-series for global modeling where the soil description is only one part of the whole due to the fast computation and the practicable residuum at depth. This approximation error of the soil temperature went down to $\pm 0.5\,^{\circ}\mathrm{C}$ and can be quite practicable in some application issues. The transfer function approach can be very accurate at shorter time intervals. This has its advantage for studying the thermal processes within the soil. It is more important for such applications to be very accurate and to detect other processes in the evolution of time.

5. Estimation of Hydraulic Properties - Synthetic Study

As discussed in chapter 3, one of the most required quantities are the hydraulic soil parameters $p = (n, \alpha, K_0, \phi)$ for description of the hydraulic soil dynamic. A correct representation of p will ensure an accurate simulation of the soil water flow through the soil into depth and thus a reliable projected soil moisture result. Thus, the parameter estimation plays an important role in this chapter.
In the following sections, forward simulations and parameter estimation based on a least squares approach are applied to some synthetic scenarios. We will analyze with these techniques the question under which conditions hydraulic parameters are feasible to be estimate accurately and in which situation it is not the case. Moreover, the estimation results are analyzed with respect to the influences of the measurement instrument onto the feasibility of an accurate estimation.
When using pure synthetic view, we investigate the question which external forcing and which soil types are more helpful for the estimation of the hydraulic soil parameters $p = (n, \alpha, K_0, \phi)$. Further, we distinguish between measurement instruments, i. e. point sensor, TDR and GPR. The last two instruments are based on averaging over some area in depth. *Robinson et al. (2003)* and *Huisman et al. (2003)* introduce the TDR and GPR with more details, respectively, for the measurement principles.
Here, the numerical forward algorithm deals with a grid resolution of 1 cm in each direction. The spatial domain is given by a $1\,\text{cm} \times 400\,\text{cm}$ area. If nothing else is mentioned, the temporal resolution of input and output data will be 10 min.

5.1. Point Sensor

The most simple description of a measurement instrument is the point sensor and therefore, our discussion starts with it. In the following, we will describe as a point sensor a cell-center of the underlying finite volume grid element. The depth of the point sensor is chosen to be $z_0 = 1.5\,\text{cm}$ below the surface. This corresponds to the center of the second cell counted from top.
Three scenarios will be presented, one infiltrating hat function and two real measured infiltration fluxes. One will have strong and short infiltration fronts and the other lower infiltrating peaks but distributed more uniformly.

5.1.1. Infiltrating Hat Function

First, we consider the scenario with the infiltrating hat function (Figure 3.3) introduced in the theory chapter 3. There we compared estimations with two different flux amplitudes and realized that for the stronger flux the residuum and the relative error (3.7) in all parameters is very small, but for the tenth smaller flux the resulting error is quite high. We could understand it by analyzing the sensitivity of $\theta(t)$ due to variations in each parameter p_i. Further, we found

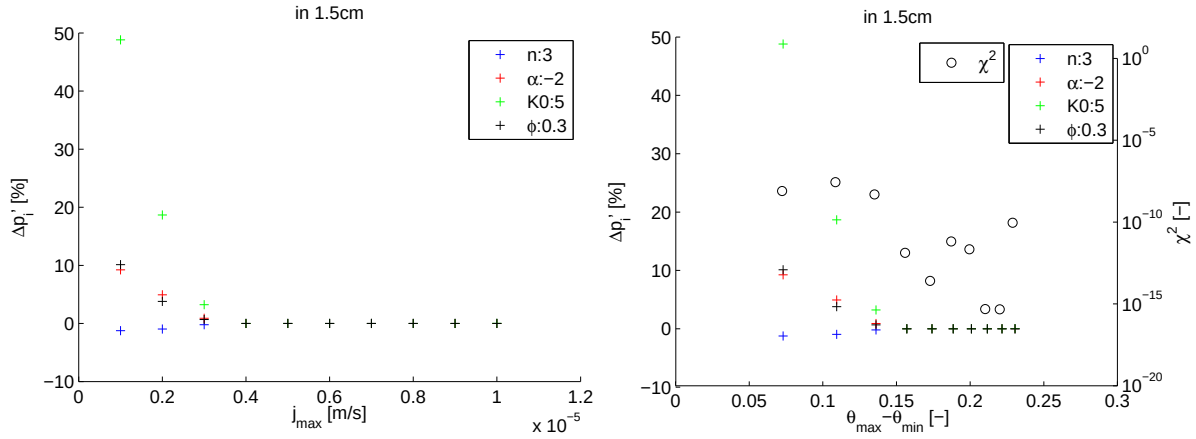


Figure 5.1.: Relative estimation error (3.7) [%] of the MvG parameters for a given soil with $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ at 1.5 cm depth. The starting parameter is $1.25 \cdot p_{\mathrm{true}}$. Left: Relative error is plotted against the maximal rain flux intensity. Right: Relative error is plotted against the water content range $\Delta\tilde{\theta} := \theta_{\max} - \theta_{\min}$. Additionally, axis the χ^2-value is shown in logarithmic scale on the right vertical.

influence regions in $[\theta_r, \phi]$ for each p_i and were able to give an expected ordering for the error (3.8) of the hydraulic parameters.
Under the assumption that a time-series $\theta(t)$ at depth z_0 changes continuously with the maximal flux $j_{\max}$ of $|j_w|$, we expected to have a transition flux $j_{\mathrm{trans}}(p)$ for our given soil. This transition flux should divide the domain of maximal rain flux $j_{\max}$ into two main areas. The estimation produces high errors for lower j_{trans}. The estimated parameters are accurate for higher j_{trans}. Therefore, we discuss the question how large such a transition flux is and how it is influenced. Ten linearly increasing and, maximal infiltration flux rates were analyzed for the same given soil as in chapter 3. The estimation results for a starting parameter set $p_{\mathrm{start}} = 1.25 \cdot p_{\mathrm{true}}$ are shown in the dependency of $j_{\max}$ in Figure 5.1 (left). Following the coloring, we observe larger estimation errors for low flux rates than for higher rates. In addition, the order of the error in each parameter is as expected. The van Genuchten parameter n fits best. The saturated hydraulic conductivity occupies the highest error. The other two parameters are lying in between with their estimation errors circa five-times smaller than the ones of K_0.
Moreover, the relative estimation error in each of the hydraulic parameter decreases to zero with increasing the infiltration flux. Thus, the transition flux is between $3 - 4 \times 10^{-6}\,\mathrm{m/s}$ since for higher flux rates, all estimated errors are much smaller than 0.1 %.
The same result can be presented from the perspective of the time-series $\theta(t)$. In the right Figure 5.1, the error of the parameters is plotted against $\Delta\tilde{\theta} = \theta_{\max} - \theta_{\min}$, the maximal change of water content in the time-series. Here, the transition to the feasibility to estimate accurate parameters is around $\Delta\tilde{\theta} = 0.15$, i. e. one half of the porosity

$$\Delta\tilde{\theta}_{\mathrm{trans}} \approx \frac{\phi}{2}\,. \tag{5.1}$$

Further, the right axis denotes the χ^2-value between the estimated and true time-series in logarithmic scale. This is as well the result of our objective function and confirms the estimation results. All estimations with χ^2 lower than 10^{-10} yield accurate results. The others show significant errors. Hence, we choose empirically the χ^2-limit for an accurate estimation at 10^{-10}. However, the accurate estimations have wide spread χ^2-values from orders 10^{-10} to

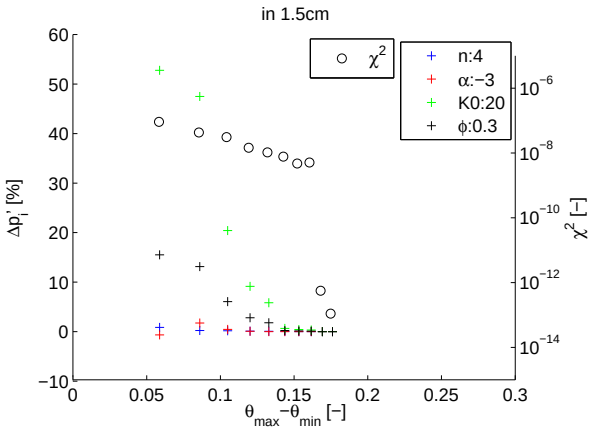


Figure 5.2.: Relative estimation error (3.7) [%] of the MvG parameter is plotted against the water content range $\theta_{\max} - \theta_{\min}$ at 1.5 cm depth. The estimations belong to infiltration fluxes of $(1 - 10) \times 10^{-6}$ m/s. The starting parameter set is $1.25 \cdot p_{\text{true}}$. Additionally, the χ^2-value in logarithmic scale is shown on the right vertical axis. The parameters of the soil are given by $(n, \alpha, K_0, \phi) = (4, -3\,\text{m}^{-1}, 20\,\text{cm/h}, 0.3)$.

10^{-16}. But these only produce estimation errors below 10^{-4} up to 10^{-8} in each estimated parameter, which is very accurate in the context of measurement applications.
We can observe similar results if we study another given soil, e. g. $p = (4, -3\,\text{m}^{-1}, 20\,\text{cm/h}, 0.3)$ with the same scenario as before. We then end up with the estimation results in Figure 5.2 against $\Delta\tilde{\theta}$.
The first point to mention is that the water content range $\Delta\tilde{\theta}$ is decreased although the porosity ϕ keeps constant. This becomes clear as soon as we consider the variation plots (Figure 3.5) in chapter 3. The increase in n and α reduces $\theta_{\min}$, but the change in K_0 influences the reduction of $\theta_{\max}$ even more. Hence, $\theta_{\max}$ decreases by some volume percent, which in total leads to a reduction in the water content range $\Delta\tilde{\theta}$ of around 5 volume percent.
Thus, the dynamical range is lower and, e. g. for a maximal rain flux of 5×10^{-6} m/s, the time-series $\theta(t)$ is not convenient enough anymore for acceptable estimations. This is also visible in the χ^2-values for this parameter set with 10^{-8} (Figure 5.2), compared to the first soil type with 10^{-14} (Figure 5.1 left). Nevertheless, the region where the estimation error decreases is around $\phi/2$. Accordingly the transition zone for the possibility of estimating the parameters is accordingly slightly higher. This region reaches only two scenarios with the highest flux.
In conclusion, we have learned that there exists at least a transition zone for the flux and the θ-range. This zone of j_{trans} depends much stronger on the given soil p_{true} than the transition zone $\Delta\tilde{\theta}_{\text{trans}}$. Its value is still in a region of $\phi/2$ with width around 0.02 volume fraction, which we notice in Figure 5.2.
In the left Figure 5.3, we summarized the estimation results of both soils coming from Figure 5.1 and 5.2. This presentation makes is easier to compare it to the right Figure 5.3, where only the porosity ϕ changed.
Especially for soil $p = (3, -2\,\text{m}^{-1}, 5\,\text{cm/h}, 0.4)$, the transition zone at $\phi/2$ is again observed whereas for the second soil type, it is a little bit higher. The main behavior, like decreasing errors with higher fluxes and the lower total range $\Delta\tilde{\theta}$ for the second soil, are analogously to the results of Figure 5.3 (left). The major difference is the stretching of the $\Delta\tilde{\theta}$-axis. In general, this stretching is given by a nonlinear description, since the maximal and minimal water content of a time series depends on the hydraulic parameters in a non-linear way. The same applies

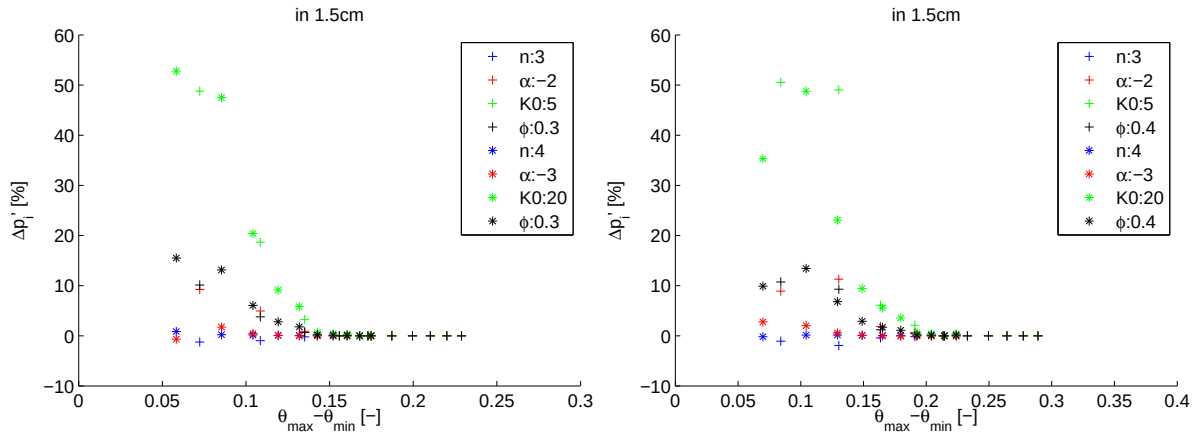


Figure 5.3.: Relative estimation error (3.7) [%] of the MvG parameter is plotted against the water content range $\theta_{\max} - \theta_{\min}$ at 1.5 cm depth. The estimations belong to infiltration fluxes of $(1-10) \times 10^{-6}$ m/s. The starting parameter is $1.25 \cdot p_{\text{true}}$. Comparison of two soil types with $(n, \alpha, K_0) = (3, -2\,\text{m}^{-1}, 5\,\text{cm/h}, \phi)$ and $(4, 3\,\text{m}^{-1}, 20\,\text{cm/h}, \phi)$. Left: $\phi = 0.3$. Right: $\phi = 0.4$.

to the porosity due to the nonlinear Richards equation. Here, we observe that the nonlinear relation is not as dominant. A linear stretching is an useful approximation.
Thus, we realize that we can scale $\theta_{\max} - \theta_{\min}$ by $1/\phi$, or more general by $1/(\phi - \theta_r)$, to a saturation quantity $\Delta\Theta$. Further, the $\Delta\Theta_{\text{trans}}$ zone where we can estimate the hydraulic parameters in an appropriate way, is slightly shifted with choosing another soil type. But it is close enough to specify a neighborhood around $\phi/2$ based on our experiences. Additionally, we noticed that not each strong rain is high enough to change the near-surface water content in such a way that trustworthy parameters p can be estimated. We observe this fact also later when we compare a lot of different soils in our estimation maps, e. g. in Figure 5.7.

Starting Parameter

In the last analysis, we always started with a starting error in each parameter of +25 %. In the following, we will look in more detail on effects due to the choice of p_{start}.
We choose a deviation of ±10 % and ±25 % in all parameters representative for other starting points. A non-symmetric disturbance for the true parameters p_{true} yields analogous results. The way of convergence from p_{start} (blue cross) to p_{true} (green cross) is shown in Figure 5.4 for the same two scenarios shown in Figure 3.3. The red and blue lines differentiate between the two infiltration amplitudes. Thus, we discover that the scenario with the larger infiltration rate converges very closely to the true parameter set from all starting points. But in case of the lower flux the last calculation steps always lie far outside of the 5 %-error bars from all starting points. Therefore, we can exclude that the good performance of convergence is strongly dependent on the choice of p_{start} for larger fluxes. Indeed the estimation error of p can vary a lot for lower flux rates, where the convergence is worse. But in these cases we cannot estimate the parameters in an appropriate way anyhow.
Another important observation in Figure 5.4 is that all paths of convergence run from the same direction towards the optimal point. In the first steps, the computed preferential direction of the Levenberg-Marquardt algorithm points more in n- and α-direction. The K_0- and ϕ-correction gets more pronounced with each further step. So it seems that there is a marked

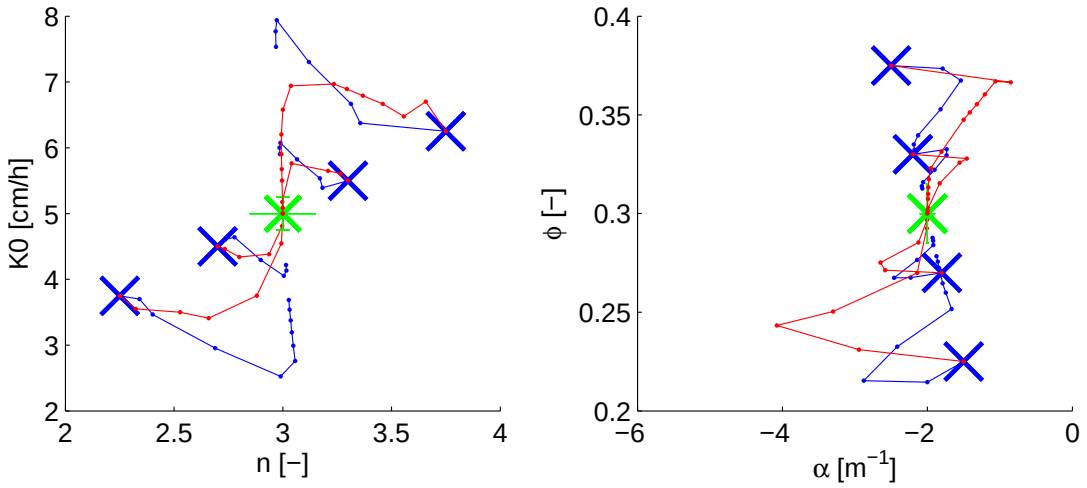


Figure 5.4.: Starting form different parameter set (blue crosses), the way of convergence is shown to the optimal parameter set $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ (green cross, Figure 3.3). The green error bars around the optimal set denote 5% error range in each direction. The blue and red curves of convergence refer to the two infiltration hat functions in Figure 3.3. Both figures together denote the four-dimensional parameter space.

valley in K_0-ϕ plane.
These observations also fit to our ordering of the estimation errors $\Delta p_i'$ in (3.8). First of all estimation parameters, the estimation of the van Genuchten parameter n converges to the region of p_{true} and is stuck in place. So if the stop criterion was fulfilled after a few steps, n would already be in a small error region whereas the saturated hydraulic conductivity K_0 could still be far off.

In summary, we saw at the example of an infiltrating hat function, that there is a transition zone for the flux and the water content, which divides the domain into region with feasible and non-feasible parameter estimations in an appropriate way. The transition flux depends on the given soil. Moreover, we identify at least a zone in $\theta_{\max} - \theta_{\min}$ where the transition zone of water content range was found in the neighborhood of $\phi/2$. Finally, the analysis of convergence showed that the choice of the starting parameter for larger flux rates has less influence on the quality of the estimation error. Further, the way to the optimal point in indicates a small valley in K_0- and ϕ-direction. This corresponds with the ordering of the estimation errors of the parameters.

5.1.2. Real Infiltration with High Peaks

In the next step, we investigate the previous results on a real measured infiltration forcing. This flux is shown in blue in Figure 5.5 and scaled by 0.5 and 0.1 in red and black. The corresponding time-series of water content $\theta(t)$ in 1.5 cm depth are shown below together with each estimated curves (green dashed lines) under the same conditions as in Figure 3.3. Dependent on the intensity of the flux, the water content rises to different heights. It is remarkable for the three θ-time-series that the fine struction of these time-series is different. The 10 %-scenario, for instance, shows a much coarser response due to the flux distribution than the 100 %-scenario.

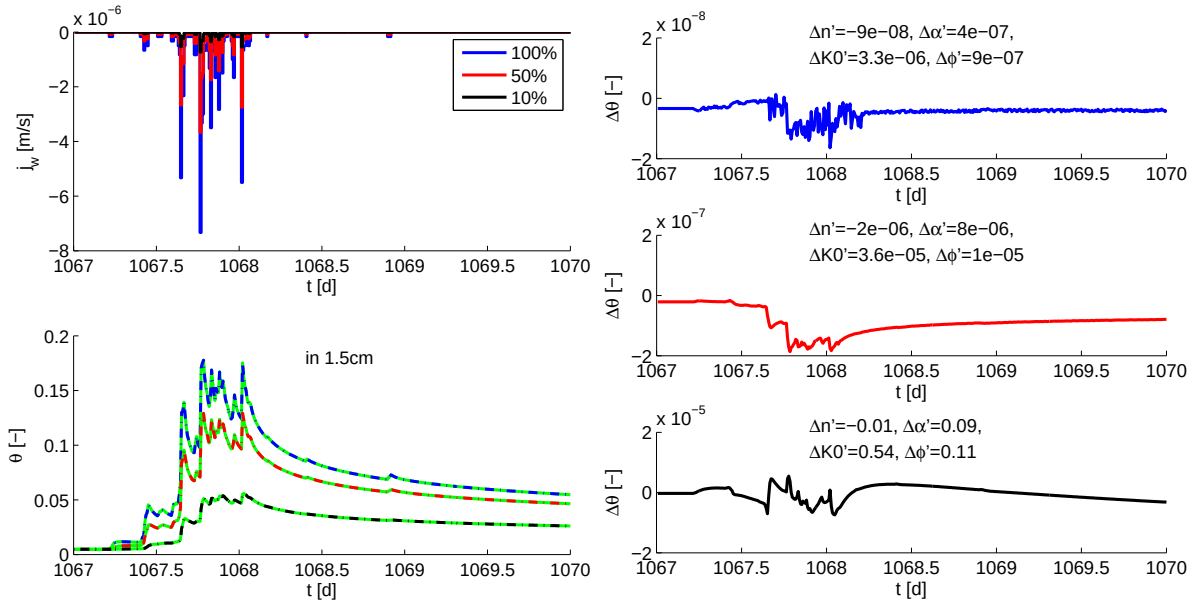


Figure 5.5.: Simulation of three infiltration scenarios for a soil with parameters $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. Left: Upper boundary condition (Neumann): The measured precipitation at Grenzhof site each 10 min results in the rain flux (blue boundary condition). The red and black colored boundary conditions are scaled by 50 % and 10 %, respectively. At the bottom, the corresponding to the boundary condition modeled θ at 1.5 cm depth. The green dashed lines are the estimated θ by inversion, respectively. Right: Difference between θ produced by forward model with p_{true} and by inversion. The colors belong to the scenarios of the top left Figure. The relative errors in each parameter are given after inversion.

The reason for this results from the dependency of the velocity of the flowing water on the current soil water content. Therefore, the reaction of the soil due to the infiltration for different flux intensities is faster in more wet soils.
Further, at the right the residuals $\Delta\theta$ between forward and inverted time-series are plotted corresponding to the previous coloring. Similar behavior as in our first scenario is found in the low error for high flux rate and the order of the parameter errors. The errors in the residuum and the parameters also increase by lowering of the flux. Due to the additional plotted 50 %-calculation, we can now observe that for a flux closer to j_{trans} the estimation is still good. However, it loses one order in the accuracy of the residuum and of the parameter error, compared to the higher flux in blue.
Further, we again find the transition for accurate and worse estimated results in Figure 5.6 if we scale the real infiltration flux down by 0.1, 0.2 up to 0.7. The two figures show the optimization results for $-25\,\%$ disturbance of the starting parameters on the left hand side and on the right hand side for $+25\,\%$ disturbance. In both cases, the χ^2-value falls below 10^{-10} for a flux scaling factor of 0.5, which is plotted red in Figure 5.5. This corresponds to a maximal infiltration flux of $3.7 \times 10^{-6}\,\mathrm{m/s}$. There we saw previously, that the accuracy is decreased compared to the first example with the infiltrating hat function.
The main difference between the plus- and minus-disturbance in p_{start} is that the estimation errors show different signs and vary slightly. But this is expected after our analysis of the starting parameters, where results coming from lower fluxes are stronger dependent on p_{start}.

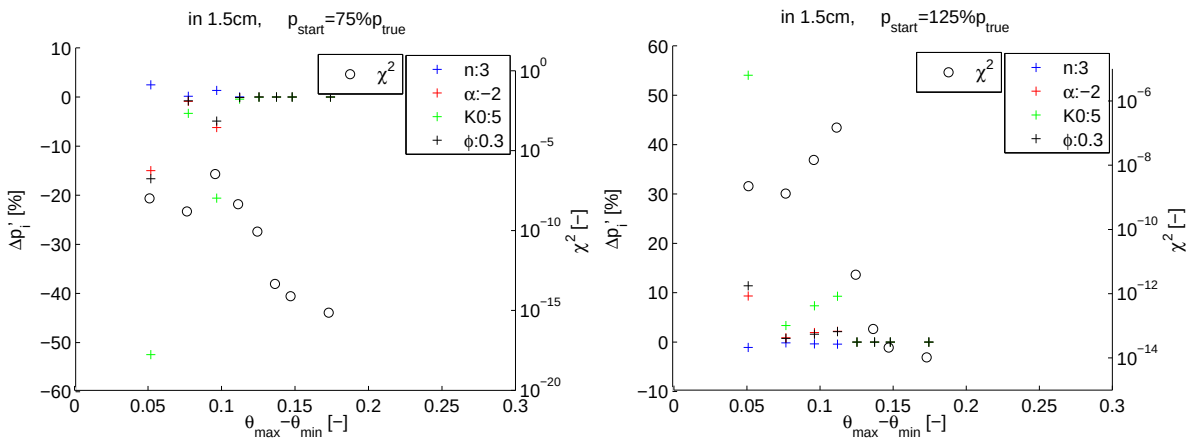


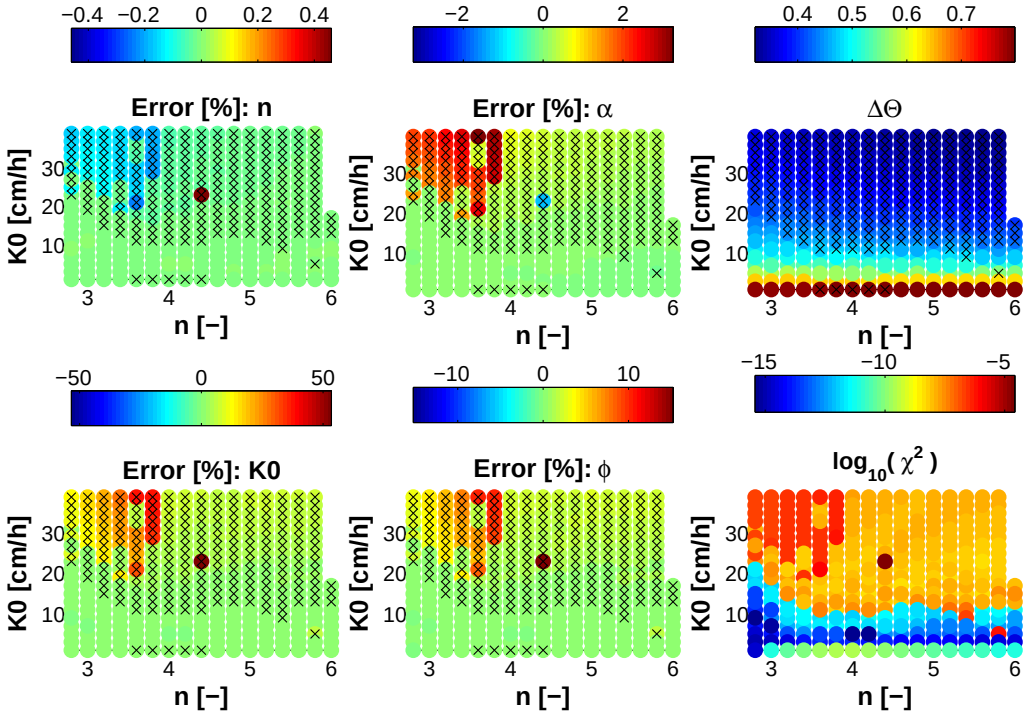
Figure 5.6.: Relative estimation error (3.7) [%] of the MvG parameters for a given soil with $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ at 1.5 cm depth. Error is plotted against $\Delta\tilde{\theta}$ for different scales of infiltration flux $0.1, 0.2, \ldots, 0.6, 0.7, 1$. Additionally, the χ^2-value is shown in logarithmic scale on the right vertical axis. Left: The starting parameter is $0.75 \cdot p_{\mathrm{true}}$. Right: The starting parameter is $1.25 \cdot p_{\mathrm{true}}$.

Also the sign of the error changes, since their way to the optimal point comes from the other side of the valley of the parameter space.
However, it is also important to notice in this case that $\Delta\tilde{\theta}_{\mathrm{trans}}$ lies below $\phi/2$. If we look at the flux and the resulting water content in more detail, we observe that the flux jumps in short time periods of 10 min very often. Thus, the reaction of the soil is too slow to be able to reach nearly the equilibrium state. This effect was not pronounced in such a strong way in our first hat function example since there the flux increases in a linear way over one hour.

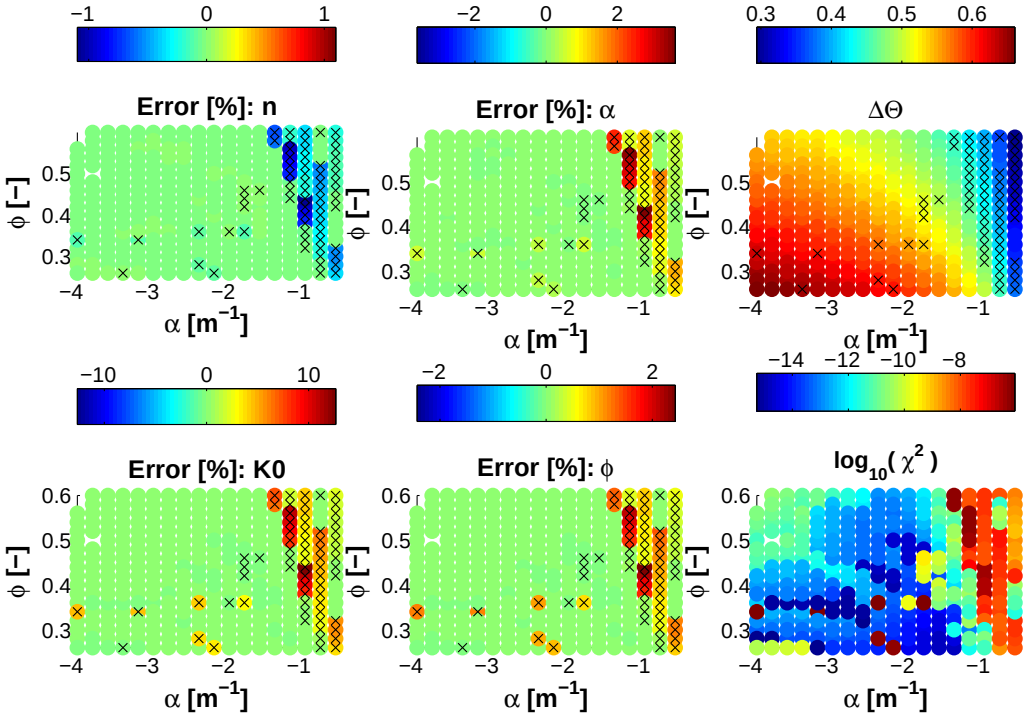
In conclusion, we identify that the transition flux j_{trans} is nearly equal for both infiltration structures for the same soil type and different upper boundary conditions. But the quantity $\Delta\tilde{\theta}_{\mathrm{trans}}$ stays nearly the same in a small region for the same upper boundary condition and different soils. Further, we noticed that the choice of the starting parameters does not effect this behavior. The different signs of the estimation errors, which are produced by different signs of the disturbance of p_{start}, we referred to the valley structure of the parameter space as shown for the point sensor in section 5.1.

Estimation Maps

Till now, we varied the flux rate for a fixed soil type. But in the following, we want to analyze the dependency of different soil onto one fixed infiltration scenario. We chose the 100 %-scenario of the realistic flux in Figure 5.5 representative for other flux distribution. For that reason, we scan over some given parameter ranges to determine some areas in the soil parameter space where the hydraulic parameter can be estimated in an accurate way.
In Figure 5.7, the results of the true parameter scan are collected. The plane of the true n and K_0 is drawn in Figure a) whereas the one of the true α and ϕ is given in b). This means in particular, that each drawn point represents one forward run to get a given synthetic time-series $\theta(t)$ in 1.5 cm and its inversion to recalculate the true parameters. The inversion results are illustrated in the six subfigures, in the left block the four relative estimation errors (relation (3.7)) in each parameter and in the right block the saturation range $\Delta\Theta$ and the logarithmic



(a) Variation of n and K_0 with constant $\alpha = -2\,\mathrm{m}^{-1}$ and $\phi = 0.3$.



(b) Variation of α and ϕ with constant $n = 3$ and $K_0 = 5\,\mathrm{cm/h}$.

Figure 5.7.: Maps varying true soil parameters p for the **100** % infiltration scenario (blue) in Figure 5.5. Each single point stands for a forward model with p_{true} and a corresponding inversion starting at $1.25\,p_{\mathrm{true}}$. The estimation error of each parameter is shown left and in the middle. At the top right $\Delta\Theta = (\theta_{\max} - \theta_{\min})/(\phi - \theta_r)$ is pictured and at the bottom right χ^2-value in logarithmic scale. The black crosses in each subfigure denote the soil with χ^2-value larger than 10^{-10}.

χ^2-value from the objective function (3.6). In all examples up to this point, the estimations, which are exact enough, have a χ^2-value below 10^{-10}. Therefore, we set a limit to that value and marked all inversions with χ^2-value larger than 10^{-10} with a black cross in Figure 5.7. This means that soils marked by such a black cross are not reliable for an accurate estimation.
For these given soils we can conclude that for this given flux, some neighboring soils can be estimated and some cannot for an appropriate precision. The estimation is accurate for given soils with small saturated hydraulic conductivity K_0 and almost all n-values. We know from the sensitivity analysis of Figure 3.5 that the water content rises by lowering the conductivity. This can also be seen in the plot for $\Delta\Theta$ whereas n shifts the water content more than shrinks it. When the water content rises for smaller K_0, the dynamical range $\Delta\tilde{\theta}$ will increase. Thus, the sensitivity of $\theta(t)$ to the parameters is stronger. The estimation can be more successful. In such a case, the given time-series $\theta(t)$ contains more characteristic information which can be fitted.
Further, for $K_0 = 1\,\mathrm{cm/h}$ a few χ^2-values are very narrow to the limit of 10^{-10}. $K_0 = 1\,\mathrm{cm/h}$ is lower than the maximum infiltration flux $j_{\max}$. Since the rain flux increases and decreases in a linear way in time, the high infiltration flux exists only for some minutes. Hence, the expected water content for a constant infiltration flux in equilibrium state is never reached. This can be noticed in the saturation range plot with the maximum $\Delta\Theta_{\max}$ of around 0.8 which is closely $\theta_{\max}$ due to the very low $\theta_{\min}$ value. Our example soil with $p = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ is located slightly higher in the maps, i. e. down left with $n = 3$ and $K_0 = 5\,\mathrm{cm/h}$.
Nevertheless, some low error points are marked with a black cross. But typically, these are soils which could reach a much higher error value if we started the inversion with another parameter set. Since they lie within the transition zone of the water content range as noticeable in the $\Delta\Theta$-figure, these points are less reliable. So it fits the fact that there is a single very bad estimation within the less reliable area. There, the water dynamic is not strong enough to ensure a similar quality of the estimation from all starting parameters.
Furthermore, we can confirm the order of estimation errors in each parameter deduced in (3.8). This means that n has the lowest error and K_0 the highest. The parameters α and ϕ are lying in between. It can also be seen in Figure 5.7 b).
In Figure 5.7 b), the true parameters α and ϕ are scanned with a fixed $n = 3$ and $K_0 = 5\,\mathrm{cm/h}$. Here the same limit $\chi^2 > 10^{-10}$ was chosen for the black crosses. In the $\Delta\Theta$-plot the χ^2-crosses match the values of $\Delta\Theta$ smaller than 0.5 for smaller negative true α-values. The four left diagrams with the errors in each estimation parameter also indicates the highest deviations in such regions.
However, some black crosses indicate bad results for higher water content ranges $\Delta\Theta$. Some of them are only marked, because their residuum is slightly too high in the order of 1.5×10^{-10} which is still acceptable. Others have a very high χ^2-value of order 10^{-7}. In such cases, the stopping criterion dominates with the absolute limit. This means that the improvement in χ^2 is faster below the given limit of 10^{-8} during the inversion step than the initial residuum is reduced by the relative limit factor 10^{-8}.
Parameter α only shows dominating problems for the estimation of higher α-values around $-1\,\mathrm{m}^{-1}$ or larger. Then the air entry value is higher and the water content to a given flux can not be as high anymore.
The porosity ϕ has the trend to lower true values that it yield a better estimation result. If we look at the $\Delta\Theta$-range in combination with Figure (3.5), then we realize that a higher porosity leads to a growth of $\theta_{\min}$ and $\theta_{\max}$. This results in an increasing range of water content $\Delta\tilde{\theta}$ by a smaller factor than for the porosity. Thus the saturation range $\Delta\Theta$ is low for higher ϕ-values.
In conclusion, we indicate regions with an accurate estimation in the error maps in Figure

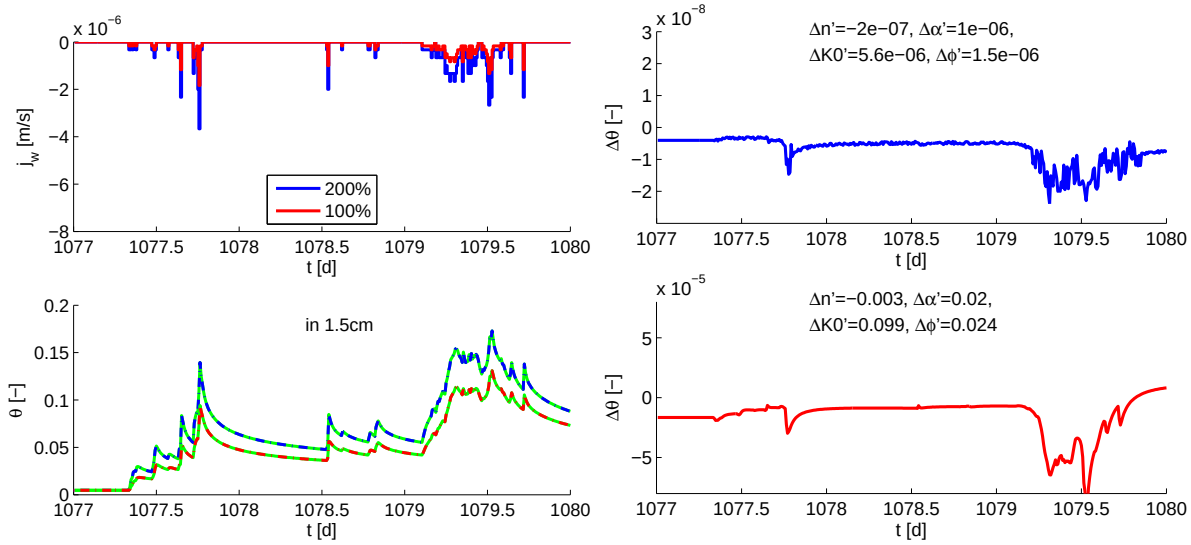


Figure 5.8.: Simulation of two infiltration scenarios for a soil with parameters $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. Left: Upper boundary condition (Neumann): The measured precipitation at Grenzhof site each 10 min results in the rain flux (red boundary condition). The blue colored boundary conditions is scaled by 200 %. At the bottom, the modeled θ is drawn at 1.5 cm depth corresponding to the boundary condition. The green dashed lines are the estimated θ by inversion starting at $1.25\,p_{\mathrm{true}}$, respectively. Right: Difference between θ produced by forward model with p_{true} and by inversion. The colors belong to the scenarios of the top left Figure. The relative errors in each parameter are given after inversion.

5.7 and regions with dynamical response that is too low for the given infiltration rate. Lower values for the saturated conductivity K_0 supports the feasibility to estimate p_{true} correctly, whereas van Genuchten parameter n shows less influence on a feasible estimation. Only for small n-values the parameter estimation is a little more successful. Further, if the value of α is negative enough, the saturation range is large enough for an accurate estimation and thus almost every value porosity ϕ of the true soil is practicable for an inversion.
Transferred to soil types it means, that a mixture of coarse and fine soil are more feasible for the estimation. Therefore, a mixture of sand and clay or loam can let the water content rise strongly enough for an inversion but also can let the water decay slow enough to a low water content.

5.1.3. Real Infiltration with more Uniform Distribution

In our last infiltration example, we again choose a real measured flux but now without a strong pronounced infiltration peak but with more stepwise increasing peaks. We compare two strengths of fluxes, 100 % and 200 % in Figure 5.8. The upper time-series is inverted with a very accurate performance whereas the lower one possesses some order of magnitude larger error in χ^2 and the parameters. Further, we can again observe that the residuals are not exactly located around zero. The reason for this is founded in the parameter errors. However, the parameter errors are small for a high infiltration, the bias error is also at a small scale. We know from Figure 3.5 that variations of the parameters n, α and ϕ can produce such a shift away from the zero-line.

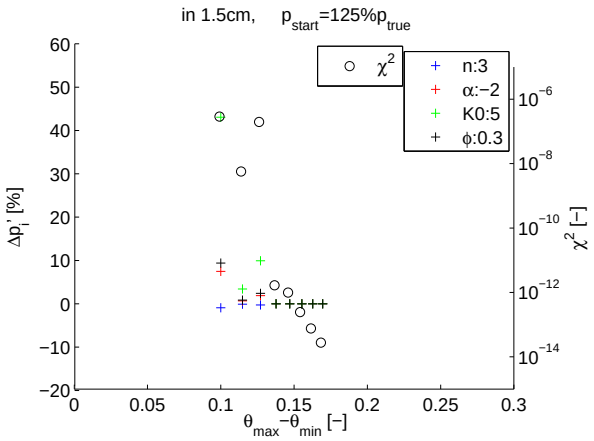
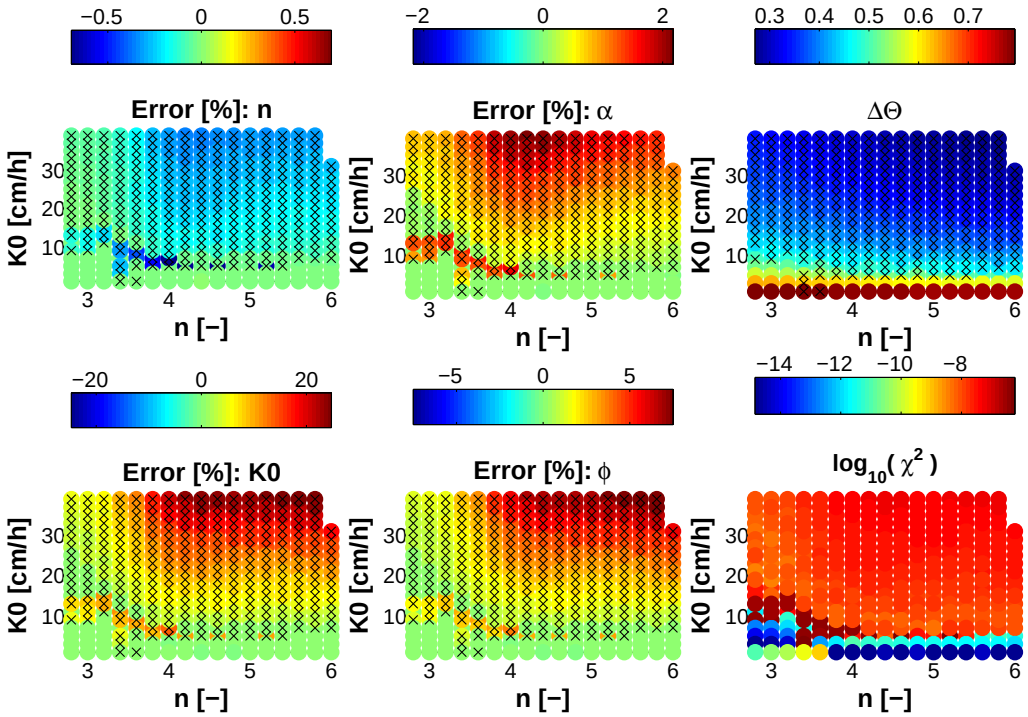


Figure 5.9.: Relative estimation error (3.7) [%] of the MvG parameters for a given soil with $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ at 1.5 cm depth. Error is plotted against $\Delta\tilde{\theta}$ for different scales of infiltration flux $0.6, 0.8, \ldots, 2$. Additionally, axis the residual in logarithmic scale is shown at the right vertical. The starting parameter is $1.25 \cdot p_{\mathrm{true}}$.

If we consider also several scaled infiltration fluxes, e. g. by $0.6, 0.8, \ldots, 2$ in Figure 5.9, we find the transition zone j_{trans} in the estimation between a scaling of 1 and 1.2 which corresponds to a maximum incoming flux peak of $3 - 4 \times 10^{-6}\,\mathrm{m/s}$. This is comparable to the previous real scenario (Figure 5.5). The χ^2-values decrease with improved parameter estimation from then on.
However, the transition in $\Delta\tilde{\theta}$ is shifted again back to larger values, not as large as in the case of the hat function but larger than in our first real infiltration scenario. A similar behavior arises for starting parameter $0.75\,p_{\mathrm{true}}$. Since the infiltration rate varies stepwise without very high jumps, the system is able to reach closer the equilibrium state that belongs to a gravity flow. Thus, the water content amplitudes become higher and let $\Delta\tilde{\theta}$ increase towards our empirical found $\phi/2$-criterion. The χ^2-values also indicate a much higher accuracy for the estimations with the given scaling factors 1.2 to 2. In this case, the least squares errors decrease with stronger rain flux.
We fix the infiltration flux scale to 200 % to take a closer look again on the estimation map (Figure 5.10). It belongs to the simulation with the largest $\Delta\tilde{\theta}$- and the lowest χ^2-value in Figure 5.9. Hence, we can be sure that there is a region with accurate estimation results.
In Figure 5.10, the estimation results are pictured in maps with varying parameters n and K_0. The maximum errors in each parameter are comparable to the other scenario of Figure 5.7. This means that the order of the error in each parameter is still the same with small $\Delta n'$ and larger $\Delta K_0'$ as mentioned in section 3.4. Further, the maximal error values (color bars) are slightly lower than in the other scenario 5.7. This fact can be caused by a distortion of the parameter space due to the different forces and therefore a varied behavior of θ.
In all these scenarios we also distinguish the two regions, where it is feasible to estimate accurate hydraulic parameters or where not. The estimation is feasible for small K_0-values and almost all n-values but the transition zone is shifted downwards compared to Figure 5.7. The reason for this is the lower infiltration flux. Here, the highest amplitude is around $4 \times 10^{-6}\,\mathrm{m/s}$, whereas the maximum incoming rate is twice as high, i. e. circa $8 \times 10^{-6}\,\mathrm{m/s}$, in the previous example.



(a) **200** % infiltration scenario (blue) in Figure 5.8.

Figure 5.10.: Maps varying true soil parameters n and K_0 with constant $\alpha = -2\,\mathrm{m}^{-1}$ and $\phi = 0.3$. Each single point stands for a forward model with p_{true} and corresponding an inversion starting at $1.25\,p_{\mathrm{true}}$. The estimation error of each parameter is shown on the left and in the middle. At the top right $\Delta\Theta = (\theta_{\max} - \theta_{\min})/(\phi - \theta_r)$ is pictured and χ^2-value at the bottom right in logarithmic scale. The black crosses in each subfigure denote the soil with χ^2-value larger than 10^{-10}.

But still our chosen soil is within the very low error region without the χ^2-crosses with $n = 3$ and $K_0 = 5\,\mathrm{cm/h}$. There, the saturation range is over 0.5 which corresponds with our $\phi/2$-criterion motivated in chapter 3.
Moreover, we notice one larger region with similar values, in the logarithmic χ^2-map, e. g. for higher n- and K_0-values. Since they are quite similar, we are able to compare the more continuous varying estimation error maps in a better way. Thus, the water content range and therefore also the saturation range is very low in the case of high n- and K_0-values. Then the parameter estimation is not as sensitive as (Figure 3.6) for soils with lower n- and K_0-values. Additionally, an error in ϕ is transferred into one in K_0 due to their strong correlation.
The error structure for the parameters n and α follows roughly the one of the other two but the location of the highest error values are shifted more to n-values of 4 to 5. The error structure is influenced by the finer χ^2-structure in this area with this fact.
Another increased error region is located in the transition zone for lower n-values in the map. Here, the least-squares value is also very high. A deeper look into the stopping criterion for such soils shows that the improvement in the last inversion step is smaller than the limit of 10^{-8}. So if we come over 0.5 in the $\Delta\Theta$-range, then the estimation is posed in such a way that the stopping criterion of the inversion dominates with the reduction factor to the initial χ^2-values. This also means that during the estimation process, the determined improvement of χ^2 is still larger than the limit and can reach closer to the minimum.

We can conclude for this last scenario that the transition zone of $\Delta\tilde{\theta}$ is between the one of the first real scenario and the hat function scenario. This fact depends on the flux distribution, i. e. small increasing steps or short intensive jumps. It also belongs to the $\phi/2$-criterion, since short jumps can not transfer the hydraulic state in at least short equilibrium states. It was clarified on the estimation maps that the transition zone of the feasible areas for estimation is shifted downwards. We explained this with the lower infiltration flux.

5.1.4. Summary to Point Sensor

We analyzed three infiltration scenarios for the point sensor description very close to the soil surface: One with a hat function structure and two with a real measured temporal time distribution. Forward simulations and the parameter estimation strategy via Levenberg-Marquardt algorithm were applied for this study. In each case, we scaled the flux and found transition zones for certain $j_{\max}$ and $\Delta\tilde{\theta}$ with a feasible accurate parameter estimation. This transition zone was located around $\phi/2$ for all three scenarios. But this varies with the flux distribution, i. e. short and intensive flux peaks showed a lower location of the transition zone whereas slowly increasing flux distributions had a higher location of $\Delta\tilde{\theta}_{\text{trans}}$.
The analysis of the starting parameters resulted in the fact that the inversion for all cases found the correct direction towards the optimum. But for lower fluxes, the optimal point was not reached. The steps of improvement were too short since the time-series $\theta(t)$ were not sensitive enough for the parameters. The optimal point was reached for large infiltration fluxes. The ways from the starting parameters to the optimum showed that there is a small valley around the optimum orientated in K_0- and ϕ-direction. This can make it harder to reach the optimum in some cases, e. g. for a low forcing.
The study of different soils in our estimation maps yield that we also found a transition zone in the given soil maps. It was dominated by the saturation range. The χ^2-marker above 10^{-10} showed similar shapes as the saturation range $\Delta\Theta$. We could illustrate the regions with feasible accurate estimations and non-feasible ones with that. Some types of soil conduced to the feasibility to estimate the hydraulic parameters, e. g. a mixture of sand with loam or clay. The properties of fast increase of water content and slow enough decay to low water content can be given in a mixture of coarse and fine soils. This behavior was similar for different flux distributions, but for the more uniform one, the transition zone in the map was shifted downwards to lower K_0 due to the lower flux amplitudes.

5.2. TDR

In the previous section we analyzed the estimation of the hydraulic parameters out of a near-surface soil water content time-series in detail. There the point sensor was represented as the evaluation of a grid cell at the cell center.
In the following, we focus on a TDR sensor to address the question under which conditions is an accurate parameter estimation feasible. Such a TDR sensor has a catchment area (*Robinson et al. (2003)*). Since the influence of the soil to the sensor decays with distance from the TDR center, we approximate the catchment area by defining a volume with $\pm 3\,\text{cm}$ around the sensor center for the following discussion. This means that we average the water content vertically over 7 grid cells with ± 3 grid cells in the arithmetic sense for our simulation domain Ω^2 with 1×400 grid cells.
In a next step, we study the first two infiltration scenarios of the previous section, but now in the context of the averaging of the water content time-series over depth.

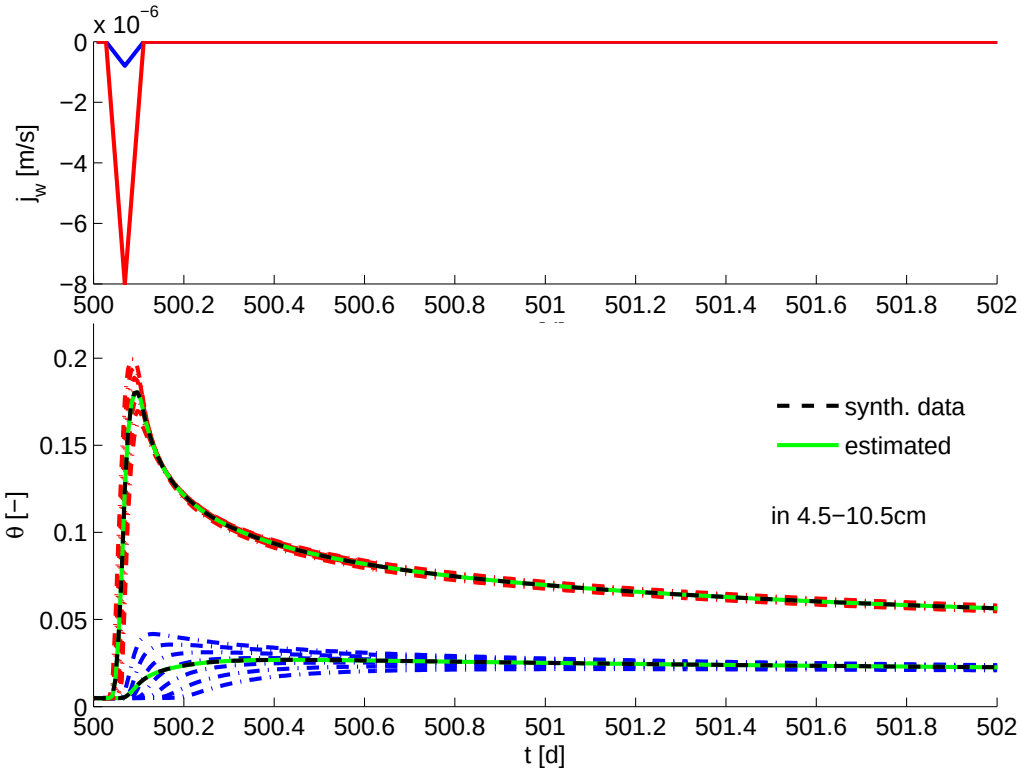


Figure 5.11.: Simulation of two infiltration scenarios for a soil with parameters $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. Top: Boundary conditions analogue to Figure 3.3: The rain flux starts after 30 min, ends after further 2 h and linear increase and decrease till zero. Bottom: Modeled near-surface water content θ in 4.5−10.5 cm depth (from top to bottom) corresponding to the two boundary conditions. The black dashed lines are the averaged θ of the θ in seven depths and overlap with the green lines which are the estimated θ from inversion.

5.2.1. Infiltrating Hat Function

We start with the infiltration flux time-series of the hat function which was described in chapter 3. There, we discussed it for a point sensor at 1.5 cm depth. Now the sensor center is shifted a little deeper to 7.5 cm. Together with the catchment area of ±3 cm, each considered time-series is drawn in Figure 5.11 with the colors corresponding to the infiltration rate. We can clearly see each time-series and their different arriving times of the infiltration front for the lower flux rate (blue). The series with the highest water content is located in the highest cell with 4.5 cm and the one with the lowest θ increases very late in 10.5 cm depth.
Since no new water infiltrates into the soil after some time, the water content decays in all depths. The velocity of the water in the soil depends on the current water content in the specific depth as given by the Mualem-van Genuchten curve (3.4) or Figure 3.5. Therefore, a region with higher water content can transport the water faster and lets the time-series coincide in the decaying part in a better way.
Additionally to the time-series in the seven depths (Figure 5.11), the averaged time-series over all these single ones is drawn in black, overlying with the green estimated one. Such an averaged curve represents the measured time-series for the following estimations. We estimate the hydraulic soil parameters from this, not from each single one in the catchment area. The

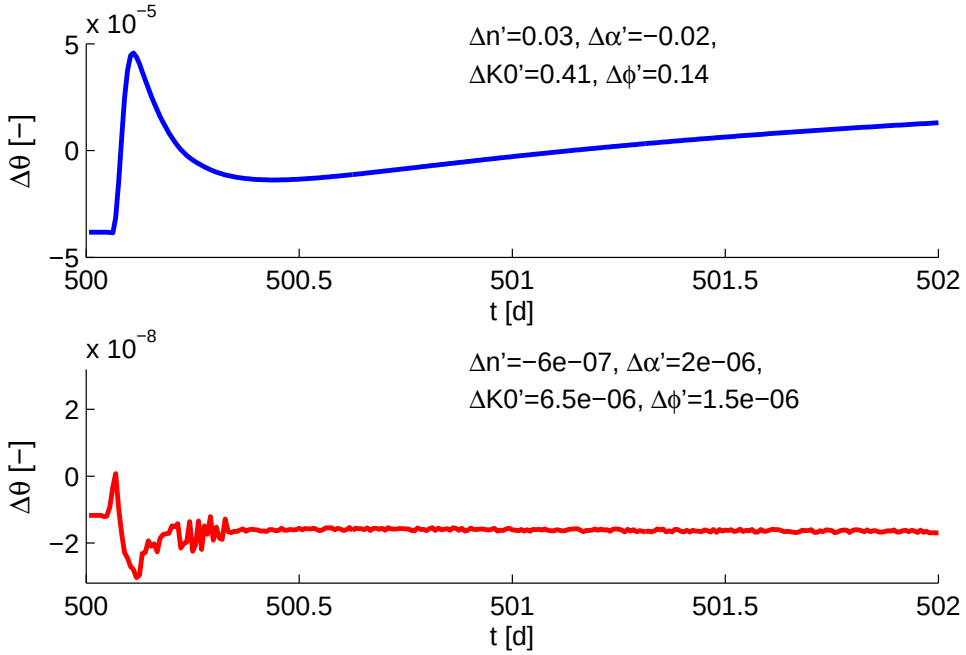


Figure 5.12.: Difference between averaged water content produced by forward model and by inversion. The colors belong to the scenarios of Figure 5.11 with maximal infiltration rate 8×10^{-7} m/s (blue) and 8×10^{-6} m/s (red). The relative error in each parameter are given after inversion.

inversion scheme is used as demonstrated in section 3.3, but with additional averaging in the forward model. Thus, we also fit an averaged version to the measured data set during the inversion.

The parameter estimation is solved for the averaged water content and both infiltration rates. The residual for both is given in Figure 5.12 with the colors and with respect to their fluxes. The results are similar to the results of the point sensor in section 5.1 for now. This means the residual error is again in the order of 10^{-5} and 10^{-8} for the maximal fluxes of $\mid j_w \mid$ 8×10^{-7} m/s and 8×10^{-6} m/s. The $\Delta\theta$-curve shows a very fine small scale structure for the higher infiltration, which indicates that we are close to the computational precision. Also the parameter error is very low with an order of 10^{-4} or lower. But for the smaller flux, the error is much higher with some orders of magnitude for $\Delta\theta$ and p. Parameter errors of 2 % to 41 % are too large to be acceptable as we saw in our projection example in Figure 3.7. We still find some characteristics in both residuals. There is a dominant trend over time for the lower infiltration flux, which is based on the larger parameter errors. If two similar time-series $\theta(t;p)$ with slightly different parameters p are compared, then it results in such a characteristic trend. Similar explanation holds for the higher flux case. But here, the effect is much lower with an order of 10^{-8} with much lower parameter errors. Thus, the effect of the trend falls close to the computational precision. We also observe the very small offset of order 1.5×10^{-8} away from the zero line. As discussed in section 5.1, this is caused by the small parameter deviations which can shift the time-series values vertically, as shown in Figure 3.5. But for the higher infiltration rate these effects are very low.

Analogue to the point sensor we assume from these two estimations that there is a transition zone from bad to accurate parameter estimation results. The feasibility to estimate is given

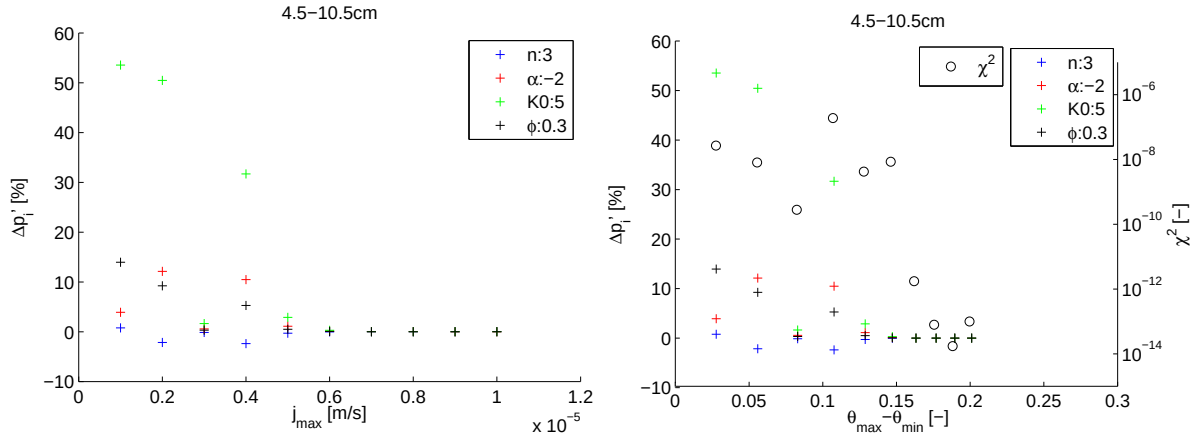


Figure 5.13.: Relative estimation error (3.7) [%] of the MvG parameter for a given soil with $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ averaged over $4.5 - 10.5\,\mathrm{cm}$ depth. The starting parameter is $1.25 \cdot p_{\mathrm{true}}$ Left: Relative error is plotted against the maximal rain flux intensity. Right: Relative error is plotted against the water content range $\Delta\tilde{\theta} := \theta_{\max} - \theta_{\min}$. Additionally, the χ^2-value is shown in logarithmic scale on the right vertical axis.

above this transition.

Therefore, we vary the maximal amplitude of the incoming flux from $10^{-6}\,\mathrm{m/s}$ to $10^{-5}\,\mathrm{m/s}$ and let one inversion run for each boundary condition. The estimation results for our test soil $p = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ are drawn in Figure 5.13 for a starting parameter of $1.25 \cdot p$. We find the transition flux around $6 \times 10^{-6}\,\mathrm{m/s}$ in the plot $j_{\max}$ against Δp considering the averaging in the forward runs. The parameter errors are higher below this and above, they are much lower. If we plot the error against the water content range $\Delta\tilde{\theta}$ in the right Figure 5.13, the estimation with this transition flux then is slightly below $\phi/2$. In this case, the χ^2-value is still high and the error is only in the order of 0.1 %. But for higher infiltration rates, the least squares values fall clearly below our chosen limit of 10^{-10} and the corresponding errors decrease by some orders.

This demonstrates that the transition water content range is still around $\phi/2$ for the TDR-averaging model, but the needed maximal flux rate is increased from $4 \times 10^{-6}\,\mathrm{m/s}$ for the point sensor to $6 - 7 \times 10^{-6}\,\mathrm{m/s}$ for the TDR. This is understandable if we look back to the averaged time-series in Figure 5.11. There, we chose seven depths with its time-series instead of a single one for the point source. And these depths are some centimeter deeper than the one in 1.5 cm. This means that the incoming water is distributed in a broader way when it reached greater depths and the water content amplitudes do not rise as high as in higher depths. Therefore, the water content in greater depth is always lower then in higher ones. Additional to this, the averaging also lets the maximal water content fall since it is averaged together with greater depths.

In total, we find lower water content values of the averaged time-series and expect to estimate accurate parameter only for higher infiltration rates as in the cases of the point sensor.

Moreover, similar behavior can be detected with other starting values for the soil parameters. So the estimated parameters are very narrow to the true ones when the flux is higher than the transition flux. Small flux rates promote a stronger dependency of the estimation error on the starting parameter, but as before all take the direction towards the optimal point. Further,

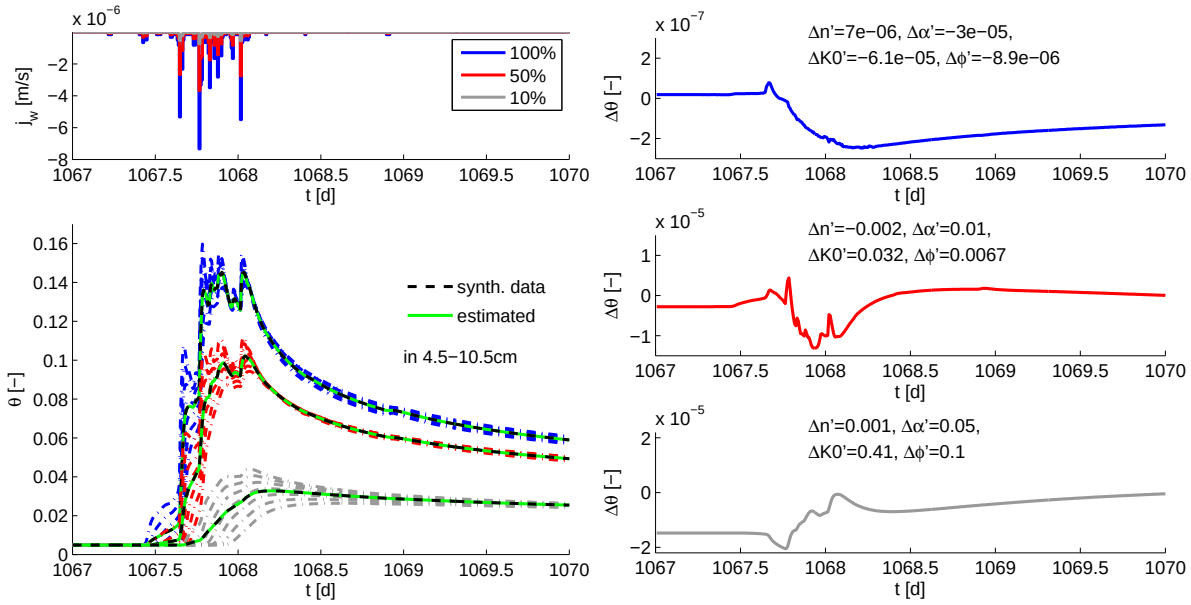


Figure 5.14.: Simulation of three infiltration scenarios for a soil with parameters $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. Left: Upper boundary condition (Neumann): The measured precipitation at Grenzhof site, each 10 min results in the rain flux (blue boundary condition). The red and black colored boundary conditions are scaled by 50 % and 10 %, respectively. At the bottom, the modeled θ in $4.5 - 10.5\,\mathrm{cm}$ depth are shown corresponding to the boundary condition. The black and green lines are the averaged and estimated θ, respectively from forward model and from inversion. Right: Difference between the averaged θ produced by forward model and by inversion. The colors belong to the scenarios of the top left Figure. The relative error in each parameter are given after the inversion.

we see also the same differences between the point sensor and the TDR with other starting parameters.

5.2.2. Real Infiltration with High Peaks

Next, we focus on a more realistic infiltration flux than the hat function. Therefore, we choose a flux time-series based on measured precipitation data from our test site at Grenzhof, the same as in section 5.1.2. In the top left Figure 5.14, this flux is drawn in blue over 3 days. Its scaled series in red and grey by a factor of 0.5 and 0.1. These are the same upper boundary case like for the point sensor in order to make these scenarios comparable.
The corresponding time-series in the depths $4.5 - 10.5\,\mathrm{cm}$ are drawn below their incoming fluxes in Figure 5.14. Again for each color, the 4.5 cm-time-series is the one with the highest water content.
One main feature is, that the blue curve has more fluctuation at a small time scale during the comparison of the blue and grey curves. This difference is based on different incoming water volumes. When the water content is low, the blue curves show a smoothed increase in water content like the grey curves. The blue time-series become steeper with more detailed fluctuations due to a new infiltration peak with a further infiltration of water and rising of

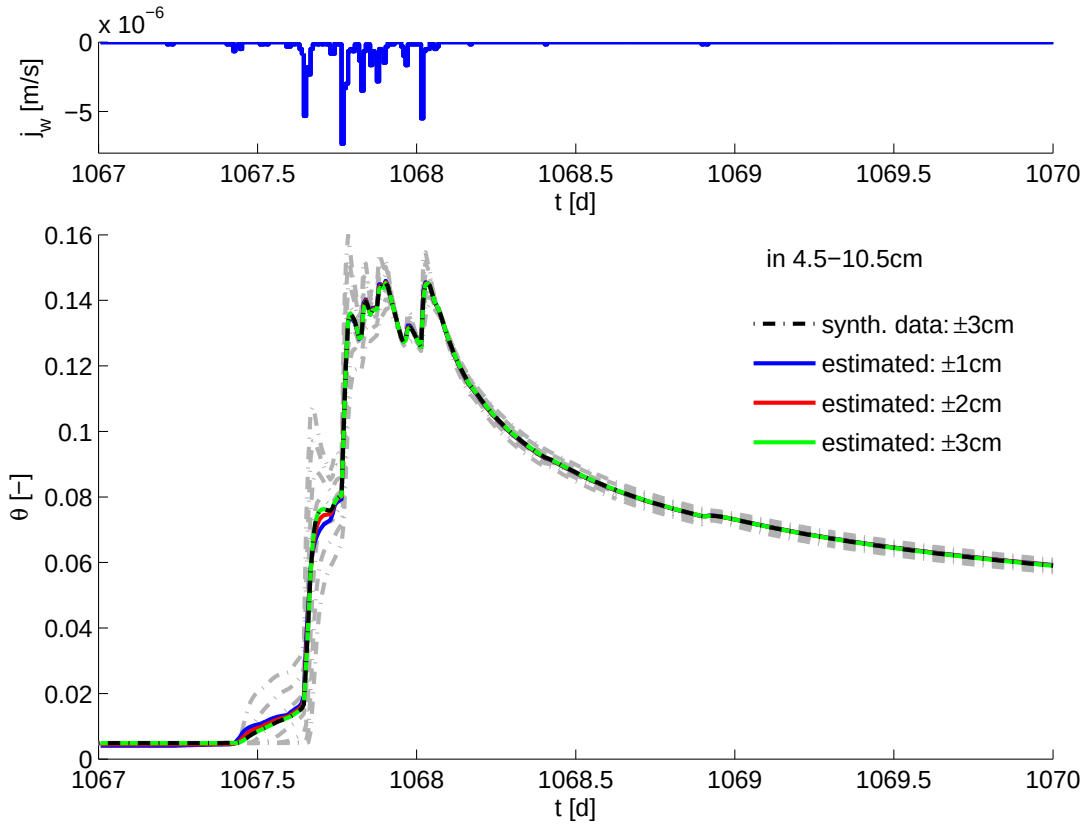


Figure 5.15.: Simulation of one infiltration scenario (top) for a soil with parameters $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. Bottom: To the boundary condition the forward modeled θ (grey) in $4.5 - 10.5\,\mathrm{cm}$ depth (top to bottom). Based on the grey lines, the black dashed line is the averaged θ and is approximated by the blue, red and green curve. An averaging in $7.5\,\mathrm{cm}$ over $\pm 1\,\mathrm{cm}$, $\pm 2\,\mathrm{cm}$ and $\pm 3\,\mathrm{cm}$ is assumed for the blue, red and green curve, respectively, during the inversion.

water content. The reasons for this is that the velocity of a water volume depends on the water content (*Roth (2007)*)

$$V_w(\theta) := \frac{\partial K_w(\theta)}{\partial \theta}\,. \tag{5.2}$$

Since it is the derivative of our hydraulic conductivity function (3.4) with respect to water content θ, we know that this velocity increases continuously with rising of θ.
The seven time-series belonging to one infiltration scenario are averaged to the black dashed curve as described in the previous example. Including this averaging also in the forward model during the inversion, we end up with the approximated time-series as the green line for each flux rate. The residuals of these inversions are pictured right in Figure 5.14. The original flux with 100 % leads to a residuum of the order 10^{-7} with an error in the parameters smaller than 10^{-5}. The other two examples are not as accurate. The parameter errors are above 0.001 to 0.67 for the scaled fluxes by 0.5 or 0.1. Here, the dynamical range in the water content domain is not strong enough to be able to estimate hydraulic parameters correctly. This means that the response of the water content due to changes in p is less sensitive.
However, if we compare scenarios with 50 %-scaled flux for a point sensor and TDR-averaged model, then the estimation accuracy differs a lot. In case of the point sensor, the parameter

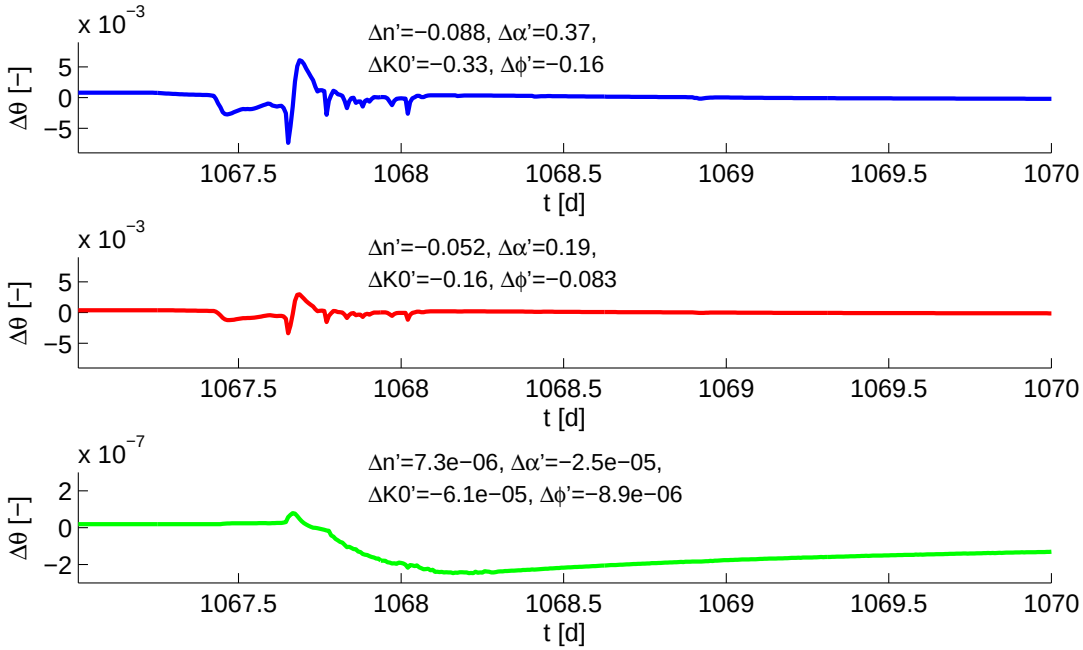


Figure 5.16.: Residual to the infiltration scenario in Figure 5.15. Differences in water content between the averaged θ over ± 3 cm and the averaged θ over ± 1 cm, ± 2 cm and ± 3 cm, respectively. The colors belong to the one in Figure 5.15. The relative error in each parameter are given after inversion.

errors are very small and also the residual is of the order of 10^{-7}. But here, the parameter error is three orders of magnitude higher and the residual around 10^{-5} for the TDR. As we learned in the first TDR scenario, the water content belonging to a fixed flux does not rise as high anymore due to the damping with depth and the averaging. Hence, the water content range shrinks and falls below the transition zone for $\theta_{\max} - \theta_{\min}$. This is the region where the time-series does not contain enough information for a reliable estimation. We also observe in the red time-series that a water content of 0.11 is only reached in maximum, whereas for the point sensor $\theta_{\max}$ is circa 0.13.

Average Volume

Till now we know the correct average volume for our TDR-sensor. But what are the effects if the average volume is not given exactly? We simulate our averaged measurement time-series (black curve) in Figure 5.15 for an averaging volume in 7.5 cm of ± 3 cm. In the next step, the forward model during inversion estimates an averaged time-series with only ± 1 cm (blue curve) or ± 2 cm (red curve). In Figure 5.15, all these curves are one above the other except the increasing part of the water content. At this stage the original time-series (grey) vary a lot in each depth which is transferred to the averaged quantities. However, the green curve fits with the given reference time-series (black) the whole time, since there, we chose the right average volume of ± 3 cm.
More details of this estimation curves are shown in Figure 5.16. The colored residuals are calculated with respect to the black reference time-series. Therefore, the blue residual with ± 1 cm averaging volume has the highest error. The corresponding error in each parameter is also the largest with $9-40\,\%$. The red residual looks like a damped blue one with a slightly more

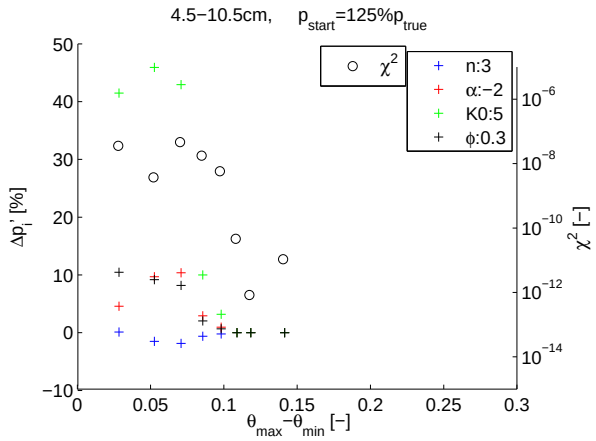


Figure 5.17.: Relative estimation error (3.7) [%] of the MvG parameter against $\Delta\tilde{\theta}$ is plotted for a given soil with $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ in $4.5 - 10.5\,\mathrm{cm}$ depth. The starting parameter is $1.25 \cdot p_{\mathrm{true}}$ and the scales of the maximal rain flux intensity varies by $0.1, 0.2, \ldots, 0.7, 1$. Additionally, the χ^2-value in logarithmic scale is shown on the right vertical axis.

correct average volume since more correct information is included for $\pm 2\,\mathrm{cm}$. The parameter errors also decrease a little to $5-20\,\%$. Only in the third case, with the average volume of $\pm 3\,\mathrm{cm}$ the estimation becomes very accurate with very low parameter errors and a residual in the order of 10^{-7}. Since the time-series in $\pm 3\,\mathrm{cm}$ depth do not grow or decrease symmetrically due to the nonlinearity of the Richards equation, we observe the differences between the averaged time-series in $\pm 3\,\mathrm{cm}$ and $\pm 2\,\mathrm{cm}$. The additional most upper and lowest time-series in $4.5\,\mathrm{cm}$ and $10.5\,\mathrm{cm}$ do not compensate each other to the averaged time-series, especially at the infiltration peaks. Due to the water content change, the nonlinearity of the dynamic is strongest at these peak and leads to larger variation in case of a change of the average volume.
In summary, the deviations of the blue and red curve in Figure 5.15 are 0.5 volume percent. This only occurs in the rising part. It would be an accurate result in case of measurements. But in the pure synthetic case, it produces quite high errors in the hydraulic parameters, as shown in Figure 5.16. Thus, a correct representation of the average volume is important for the estimation of the hydraulic parameters in case of a TDR sensor.

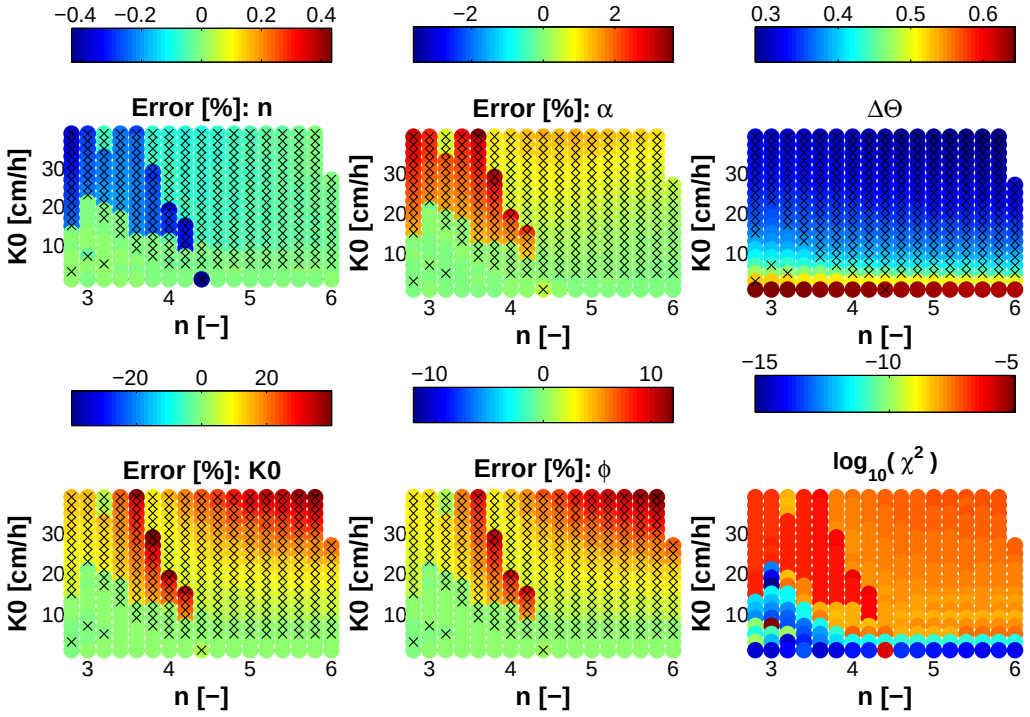
Flux Variation

For the following considerations, we assume that the average volume is known with $\pm 3\,\mathrm{cm}$ width. We can therefore concentrate on the effects of a varying infiltration. Additionally, we fix the starting parameters at $1.25 \cdot p_{\mathrm{true}}$ since no important differences to other start parameters are expected, as learned from the previous example in section 5.2.1.
In Figure 5.17, the estimation errors are illustrated against the water content range for the scenario from Figure 5.14 with a scaled infiltration flux by the factors $0.1, 0.2, \ldots, 0.7, 1$. There, the water content range $\Delta\tilde{\theta} = \theta_{\max} - \theta_{\min}$ is around $\phi/2$ for the highest flux of scale 1. A smaller scale factor of 0.6 also leads to an accurate result, as we can see concerning the low values of χ^2. As we learned from Figure 5.14, a scaling of the flux by 0.5 will be not sufficient anymore if we average over depth. This matches the observations from Figure 5.17.
But in comparison to the point source example in section 5.1.2, the water content range of each of the flux rates is decreased. In total, the increase of the transition zone cannot compensate

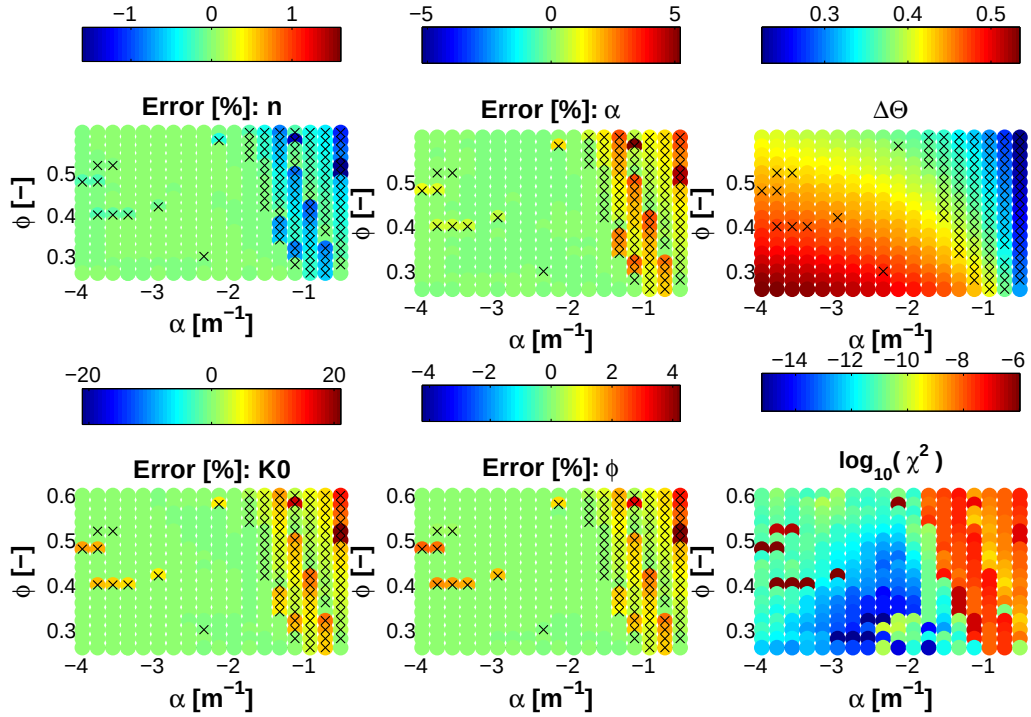
the $\Delta\tilde{\theta}$-decrease due to the averaging. The damping with depth together with the averaging can let the maximal water content fall by approximately two volume percent.

Estimation Maps

We observed for the point sensor that the transition zone for the water content range as well as the saturation range $\Delta\Theta$ is a reliable quantity for comparisons between soils. This is still the case after averaging, which can also be observed in Figure 5.18. Again, each point in these maps stands for a given soil. After a forward run to produce a synthetic data set instead of a measurement one and after the inversion, the estimation results are drawn in these maps. The maps in a) belong to a scan of true parameters n and K_0 and in b) to a scan of the true α and ϕ. The other two parameters are kept constant, respectively. In each case, the estimation error of the four parameters, the saturation range $\Delta\Theta$ and the χ^2-value are drawn in the six maps. The black crosses represent estimations with χ^2 higher than 10^{-10}, which is chosen due to our experience of the previous examples.
In the case of varying n and K_0, we observe a low least-squares value, when the saturation range is over 0.5. A few inversions also yield good estimation results for lower n and a little higher K_0 although the $\Delta\Theta$-range is nearly the same as for higher n. The reason for this is in the position where the $\Delta\Theta$-range is located at the saturation domain $[0, 1]$. We get an impression of it when going back to Figure 3.5 a). Same as here, the minimal water content $\theta_{\min}$ is given by the initial equilibrium state. If we compare $\theta_{\min}$ belonging to the initial and minimal potential $h_{\min}$, then $\theta_{\min}$ is low and keeps small for n larger than 4. However, $\theta_{\min}$ rises for smaller n since the curvature of the soil water characteristic becomes significantly smaller. The maximal water content $\theta_{\max}$ also rises and $\Delta\Theta$ compensates it. This means that a high saturation range $\Delta\Theta$ makes it more feasible to estimate accurate parameters. Covering it good enough with data points, is important at regions with a large curvature of the parameterization curves. In our example, this is always the case when the saturation range is large. We do not have to consider the cover in such cases. But the cover can have influence within the transition zone where the feasibility to estimate or not is quite closely together.
The main structure of the errors and water content ranges are quite similar to the point sensor scenario in Figure 5.7. However, there are less soils which fulfill the χ^2-criterion that χ^2 is less than 10^{-10} in Figure 5.18. The reason for that is the averaging over greater depths. The water content varies in depth for a given infiltration scenario. Further, the structure and information from top is damped in greater depths. This complicates the estimation for some soils types. If the information is damped too strong, there is nothing left to be sensitive enough for the inversion algorithm. This is the case in greater depths and by the averaging compared to the point sensor in 1.5 cm depth. This fact moves the transition zone towards lower K_0-values downwards.
Moreover, the order of the relative parameter errors is at first $\Delta n'$ and at last $\Delta K_0'$ as expected in chapter 3. It can be observed in a) and b).
Further, higher χ^2-values are calculated in regions with a low saturation range. However, the saturation range is not as wide for α lower $-2\,\mathrm{m}^{-1}$. The estimation results are accurate, however. We see a similar effect as in the maps of a). The minimal water content is very low for soils with a strong negative α. More information about the bigger curvature part of the parameterization curves is included. This makes is easier to find accurate parameters when they show a good sensitivity, e. g. n and α for low θ. The remaining structure of the $\Delta\Theta$-map for larger negative α-values, e. g. $-4\,\mathrm{m}^{-1}$, is dominated from the maximal water content of the averaged time-series.



(a) Variation of n and K_0 with constant $\alpha = -2\,\mathrm{m}^{-1}$ and $\phi = 0.3$.



(b) Variation of α and ϕ with constant $n = 3$ and $K_0 = 5\,\mathrm{cm/h}$.

Figure 5.18.: Maps varying true soil parameters p for the **100** % infiltration scenario (blue) in Figure 5.14. Each single point stands for a forward model with p_{true} and a corresponding inversion starting at $1.25\,p_{\mathrm{true}}$. The estimation error of each parameter is shown on the left and in the middle. At the top right $\Delta\Theta = (\theta_{\max} - \theta_{\min})/(\phi - \theta_r)$ is pictured and at the bottom right, χ^2-value in logarithmic scale. The black crosses in each subfigure denote the soil with χ^2-value larger than 10^{-10}.

The few crosses within the large low error region denote estimations with high χ^2-values. These inversion calculations stop due to the second stopping criterion, i. e. the improvement in the step size is lower than the set limit 10^{-8}.
Summarizing, we saw similar results for the TDR-averaged case as in the point sensor case. But due to the averaging and the additional greater depths the number of soils for which it is feasible to estimate the accurate hydraulic parameters, shrinks. There, the choice of n has less influence on the feasibility to estimate, whereas the saturated hydraulic conductivity K_0 has to be small so that an estimation is feasible. Further, more negative α-values and a lower porosity are more convenient for the estimation. As in section 5.1.2 these parameter properties belong to a mixture of coarse and fine texture soils.

5.2.3. Summary to TDR

To sum up the TDR-averaging section, the forward simulation showed by solving numerically the PDE (3.1) that with each greater depth the water content time-series becomes more damped. At the both infiltration examples, the hat function and the real flux scenario, the time-series were averaged over a certain volume and the water content range was decreased compared to our previous point sensor analysis. Therefore, the transition zone for the flux and saturation range rose which made it more difficult to estimate p.
Moreover, we discussed the influences of the average volume for the estimation. We found out that the slightly smaller average volume only yields small differences in the corresponding time-series of the order of 0.003, but these were strong enough to produce large parameter errors. Most of the increasing slopes of water content were error-prone since the time-series at different depths varied a lot. The non-linearity of the Richards equation prevents that the differences vanish by averaging.
Additionally we considered different soils in the estimation maps. There, the structure was analogous to the one of the point sensor. This means the behavior of the parameters is similar compared to the feasibility of the estimation. A mixture of coarser and finer soil supports the feasibility to estimate in terms of soil types. But the averaging leads to a decrease of feasible soils for an accurate estimation. Thus, the transition zone of the estimation maps moves downwards to low K_0 and more negative α and shrinks the number of feasible soils.

5.3. GPR

Again the question under which conditions it is feasible to estimate accurate soil parameter or not, is analyzed for the third measurement scenario. To address this question several minimization problems are solved and compared with each other.
This last measurement scenario is the ground penetrating radar (GPR). In this case electromagnetic wavelets are sent starting from the surface and its reflection at a permittivity discontinuity is detected by the receiver antenna, e. g. at an intersection between different soil layers. A main measure for the amount of water in the soil is the travel time of the wave from transmitter to the receiver. Due to the measurement principle of the GPR, the resulting time-series of water content is an averaged quantity over the volume from the surface to the reflector. Further details of the measurement principle are given by *Huisman et al. (2003)*.
Further, we assume that the mean velocity of the wave through the soil is given by the arithmetic mean of the water content over depth. Therefore, we produce our measurement time-

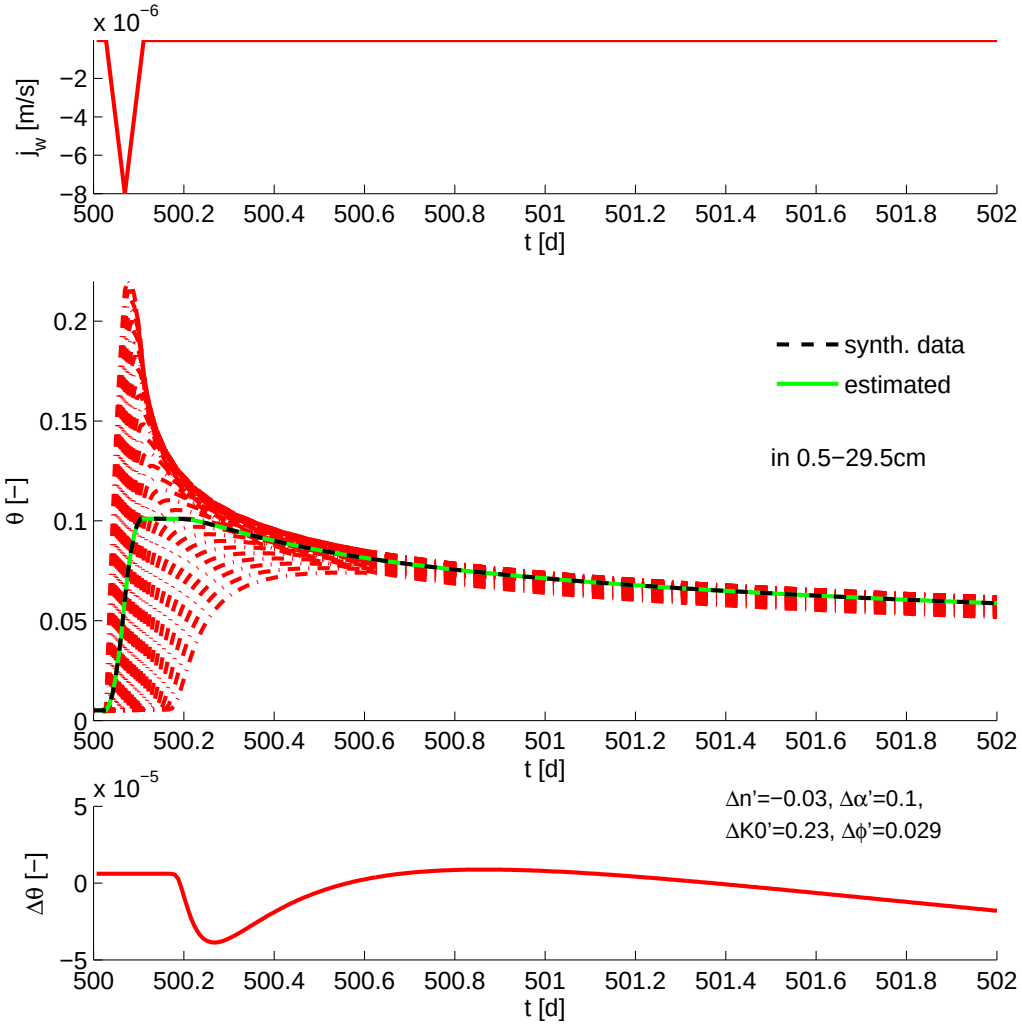


Figure 5.19.: Simulation of one infiltration scenario with $j_w = -8 \times 10^{-6}\,\mathrm{m/s}$ for a soil with parameters $p = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. Top: Boundary condition analogous to Figure 3.3: The rain flux starts after 30 min, ends after further 2 h and linear increase and decrease till zero. Middle: Modeled near-surface water content θ (red lines) in $0.5 - 29.5\,\mathrm{cm}$ depth. The black dashed and green lines are the averaged and estimated θ, respectively from forward model and from inversion. Bottom: Residual between averaged water content produced by forward model and by inversion. The relative error is given in each parameter after inversion.

series synthetically by averaging arithmetically over depth, here over the upper 30 gird cells, i. e. $0.5 - 29.5\,\mathrm{cm}$. Further, we choose our test soil with given $p_{\mathrm{true}} = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$.

5.3.1. Infiltrating Hat Function

The first scenario is the synthetic description of the infiltrating hat function from chapter 3 again. Here, we choose the infiltration example with the higher flux of $8 \times 10^{-6}\,\mathrm{m/s}$, where we found accurate estimation results in case of the point sensor.
All time-series are drawn in red within the first 30 cm in Figure 5.19. Starting with the up-

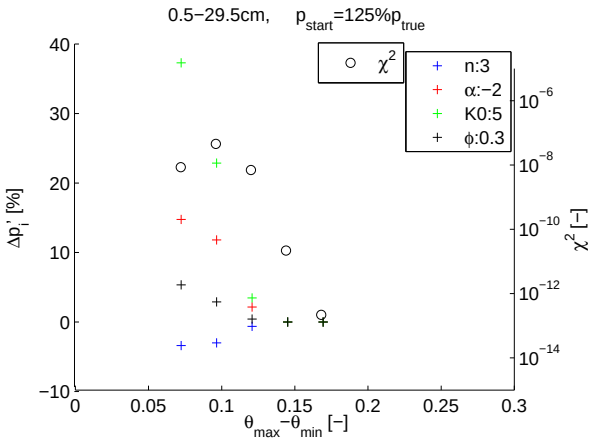


Figure 5.20.: Relative estimation error (3.7) [%] of the MvG parameter against $\Delta\tilde{\theta}$ is plotted for a given soil with $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ in $0.5-29.5\,\mathrm{cm}$ depth. The starting parameter is $1.25 \cdot p_{\mathrm{true}}$ and the infiltration flux is varied from $6 \times 10^{-6}\,\mathrm{m/s}$ to $14 \times 10^{-6}\,\mathrm{m/s}$ in steps of $2 \times 10^{-6}\,\mathrm{m/s}$. Additionally, the χ^2-value is shown in logarithmic scale on the right vertical axis.

permost depth with high water content and a pronounced peak, the water content falls to the lowest depth with a small increase and a longer decay. Further, the synthetically produced measurement time-series is shown in black and averaged over all the 30 time-series. Here, we note that the shape of the averaged time-series looks different from the single ones. This is a first indication that there is no chance to reproduce the averaged time-series with a time-series in one single depth with the true parameters. The reason for this is the nonlinear behavior of the time-series.

As we learned in section 5.2, the correct average volume is important for the quality of the parameter estimation. Therefore, we assume the correct averaging volume of $0.5 - 29.5\,\mathrm{cm}$ for the inversion calculations to focus on estimation quality without additional error sources. The resulting time-series is pictured green in Figure 5.19. It looks quite similar to the black reference time-series but the difference between both in the Figure below shows a residuum of order 10^{-5}. This is too large for an accurate estimation as the errors of each parameter points out. Thus, the relative parameter error is between $3\,\%$ and $23\,\%$ approximately. These errors are approximately four order of magnitudes larger than the errors for the same infiltration rate but a different averaging volume, e. g. for the TDR sensor. The reason for this is that we include many more greater depths as in the case of TDR. Then the time-series is a lot less sensitive, i. e. information has a much stronger weight during the averaging. Hence, the resulting averaged time-series is not sensitive enough due to variations in p. Thus, the parameter estimation gets worse. Therefore, we expect an increase of the transition zone. The maximal transition flux is above the incoming flux with $j_{\max} = 8 \times 10^{-6}\,\mathrm{m/s}$.

In the next step, we want to specify the transition zone a little more. Therefore, we vary the maximum flux from $6 \times 10^{-6}\,\mathrm{m/s}$ to $14 \times 10^{-6}\,\mathrm{m/s}$ in steps of $2 \times 10^{-6}\,\mathrm{m/s}$. The estimation errors for it are illustrated in Figure 5.20 together with their χ^2-values. Here, we again observe a decrease of the parameter errors with the increase of the infiltration flux, i. e. the water content range. We get the first reliable estimation results for a flux rate of $12 \times 10^{-6}\,\mathrm{m/s}$ with χ^2 lower than 10^{-10}. Due to the averaging, we again find an increase of the transition flux compared to the point and TDR sensor. The transition zone is around $6 - 7 \times 10^{-6}\,\mathrm{m/s}$ for TDR. This large

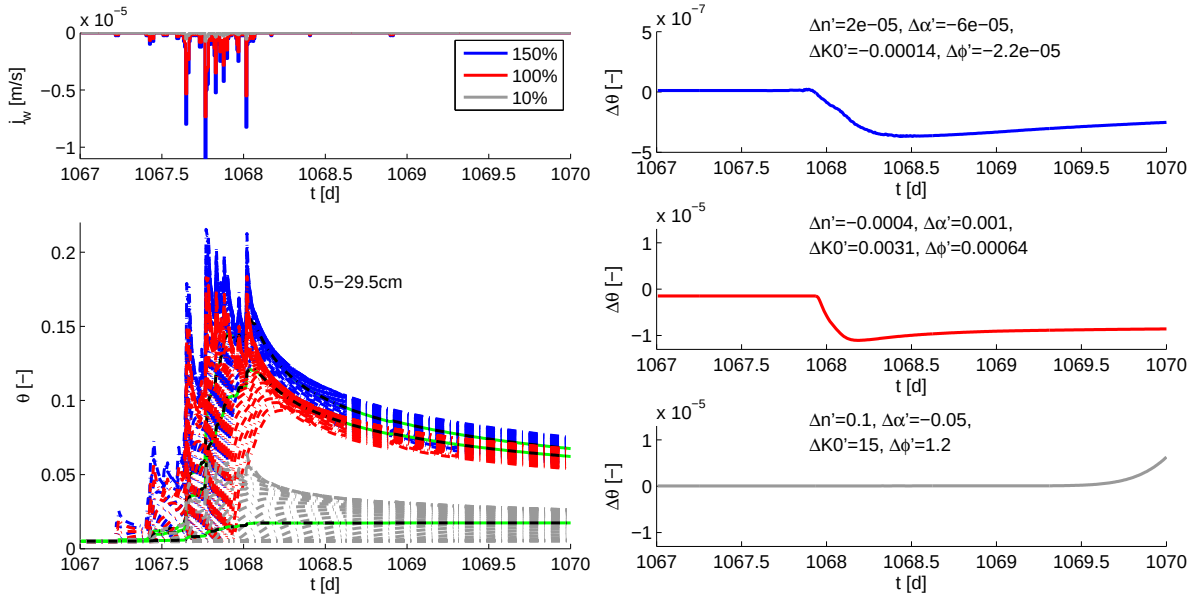


Figure 5.21.: Simulation of three infiltration scenarios for a soil with parameters $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$. Left: Upper boundary condition (Neumann): The measured precipitation at Grenzhof site each 10 min results in the rain flux (red boundary condition). The blue and grey colored boundary conditions are scaled by 150 % and 10 %, respectively. At the bottom, corresponding to the boundary condition the modeled θ in $0.5 - 29.5$ cm depth (top to bottom). The black and green dashed lines are the averaged and estimated θ, respectively from forward model and from inversion. Right: Difference between the averaged θ produced by forward model and by inversion. The colors belong to the scenarios of the top left Figure. The relative error is given in each parameter after inversion.

displacement originates in the time-series for the averaging, which is four-times more than for the TDR-averaging. More of the additional time-series are in greater depth. Therefore, the information content is more damped.
Moreover, we detect the transition zone of the water content range $\Delta\tilde{\theta}$ slightly below 0.15, i. e. $\phi/2$. This shows that $\Delta\tilde{\theta}_{\mathrm{trans}}$ is reasonably stable with respect to varying average scenarios. Similar results can be found starting parameters $0.75 \cdot p_{\mathrm{true}}$. Also further starting parameters for the scenario in Figure 5.20 show a low dependency of p_{start} on the feasibility of parameter estimation.

5.3.2. Real Infiltration with High Peaks

In the following, the flux time-series change at the upper boundary. The same realistic rain distribution is chosen instead of the synthetic hat function like in the previous sections. Similar to the TDR case, we scale the flux but now we include also a higher factor 1.5 besides the factor 0.1 into the scenario.
Respective to the colors of the fluxes, the time-series in all 30 depths are shown in the left Figure 5.21. In comparison to the TDR scenario, we can observe analogue effects like the dependence of the damping with the current water content. But one detail is more emphasized. The grey curves belong to the 0.1 scale and the highest time-series reaches an approximate

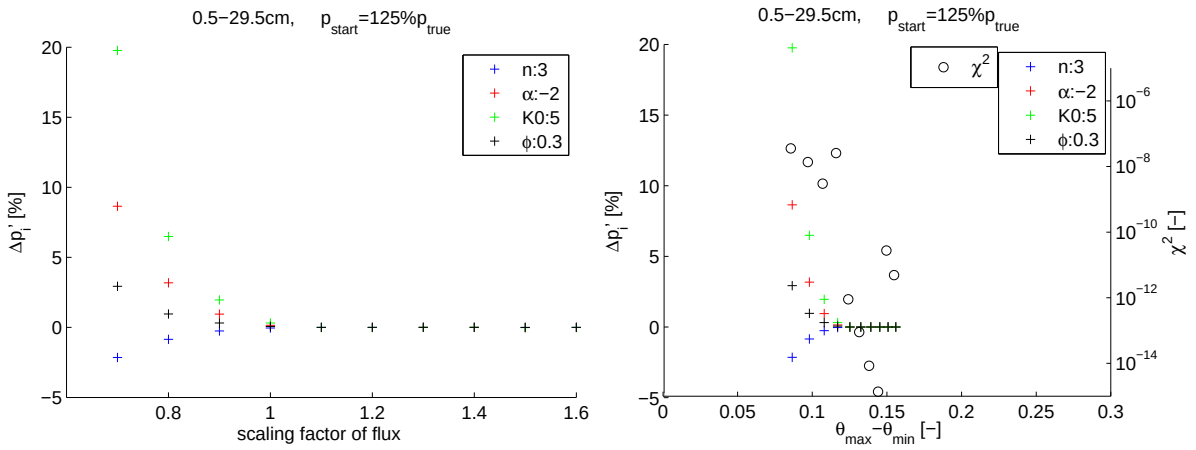


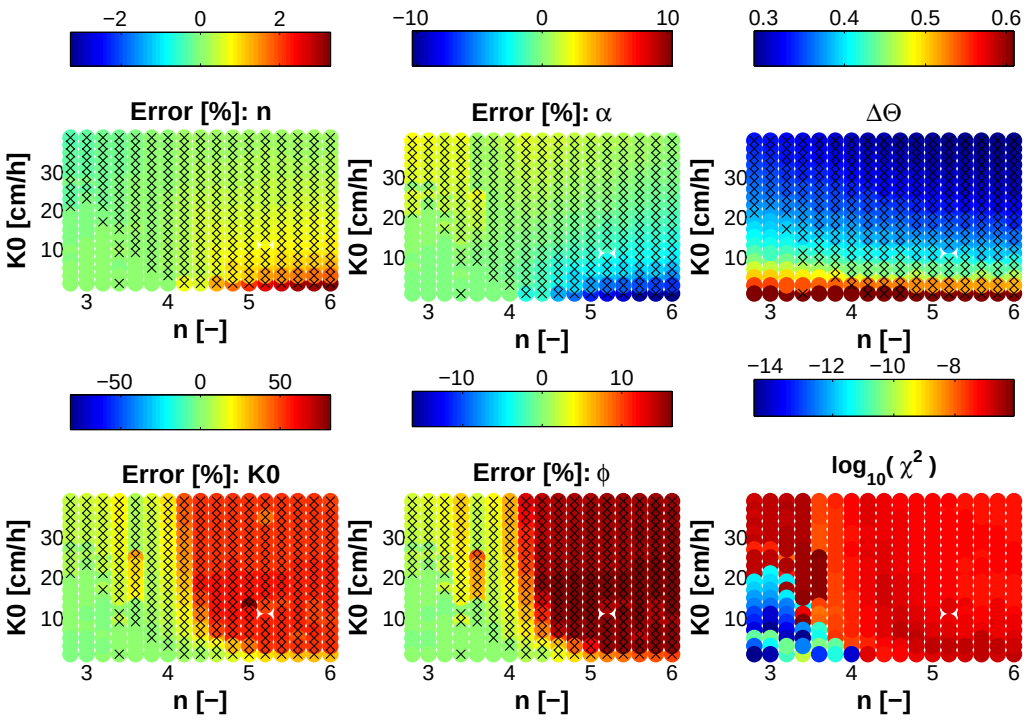
Figure 5.22.: Relative estimation error (3.7) [%] of the MvG parameter for a given soil with $(n, \alpha, K_0, \phi) = (3, -2\,\mathrm{m}^{-1}, 5\,\mathrm{cm/h}, 0.3)$ in $0.5 - 29.5\,\mathrm{cm}$ depth. The starting parameter is $1.25 \cdot p_{\mathrm{true}}$. Left: Relative error is plotted against the scales of the maximal rain flux intensity $0.7, 0.8, \ldots, 1.6$. Right: Relative error is plotted against $\Delta\tilde{\theta}$ for different scales of infiltration flux. Additionally, the χ^2-value is shown in logarithmic scale on the right vertical axis.

volume fraction of 0.07. But the corresponding averaged curve in black is 0.03 at most of the volume fraction. The reason for it is based on the many time-series which belong the greater depths. There, the water content is low and their contribution to the averaged time-series outweigh the ones of the lower depth. This will make it harder to estimate correct parameters if the weight fraction of time-series with lower information content is too high.
The fact that these estimations are more difficult is illustrated in inversion results. In Figure 5.21 left, the green curves are the estimation of the black ones. They look very similar but their differences on the right Figure 5.21 show the details. We detect the very low water content values in the grey example. It has the highest estimation errors. Especially the errors of the parameters are too high in many orders with percent values of $5 - 1500\,\%$. Although the parameters are totally off, there is still a difference in the quality between the four parameters. Since the representation of lower water content is good in this case, the parameters n and α still have a high error but one to two orders of magnitude lower than K_0 and ϕ. The parameters n and α are quite sensitive for low θ when the van Genuchten parameterization has a stronger curvature (Figure 3.5). Some data with high water content is needed to fix the parameters K_0 and ϕ since the sensitivity for them is more emphasized with higher water contents. Further, an error in one parameter can be transferred to the error of another since they are correlated with each other as explained in section 3.4.
In the red example of Figure 5.21, the residual is still of the same order as the grey example but the parameter errors are much smaller, below $0.5\,\%$. Since the deeper lying time-series reach at least maximal water content values of 0.1, the averaged time-series contains a more significant structure which has to be reproduced during the inversion. Therefore, the parameter errors decrease, but it is still not enough to estimate the parameter in a very high quality. This is managed in the third infiltration example in blue. A scaling by 1.5 yields a residuum of order 10^{-7} and parameter errors of order 10^{-5}.
From these three examples we can conclude that the transition flux is between the $100\,\%$- and $150\,\%$-scaled flux. We estimate the hydraulic parameters for some more flux scalings from 0.7

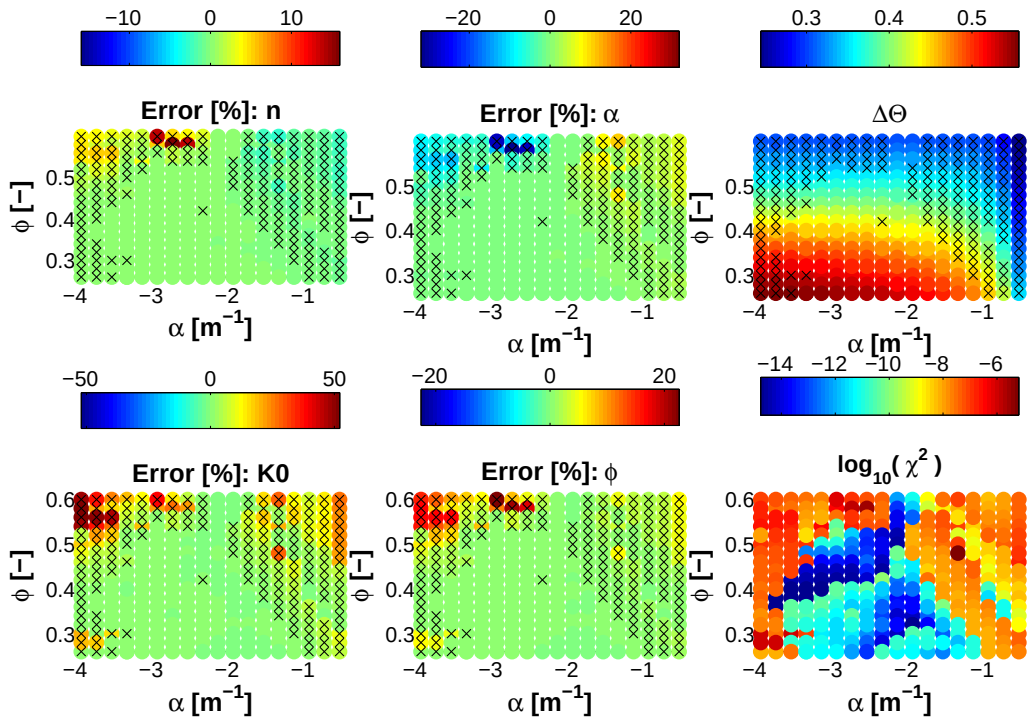
to 1.6 to specify this relation. In the left Figure 5.22, the scaling factor is plotted against the estimation error of each parameter in percent. For scale 1, the error is low, as discussed previously. The errors are still smaller for higher scale factors. The transition zone for the flux scale and also for the corresponding flux is between the scale 1 and 1.1. In Figure 5.17, a higher scaling is needed in comparison to section 5.2 with a transition scale of 0.6 due to the averaging over greater depths.
In the following, we focus on the water content range which is drawn against the parameter error on the right Figure 5.22. The transition in the water content range is around $0.12 - 0.13$. This is nearly the same in the first GPR scenario with the hat function. The averaged time-series is not as sensitive on the fluctuation of the flux distribution anymore for GPR. The time referring to the main infiltration peak within the catchment area is much shorter for the point sensor. The infiltration front reforms its specify structure in the first centimeters. Hence, more sensors and especially deeper ones can be more robust to different infiltration distributions, especially if they are averaged as in the GPR scenario.
Further, the χ^2-values also indicates very accurate estimation scenarios with values below 10^{-10}. But these only appear for water content ranges higher than 0.12 of volume fraction. All estimations below that volume fraction have a higher least-squares value and much larger errors in each parameter.

Estimation Maps

In analogous way to the previous sections, we analyze the behavior of the estimation results with respect to different types of soils. Therefore, we vary the true parameter of the soils und choose one infiltration scenario with a fixed scale. This is the case in the scenario of Figure 5.21 with the 150 %-scale (blue). Moreover, the average volume is given by the first 30 cm.
First, we start with the variation of n and K_0 in Figure 5.23 and hold α and ϕ at $-2\,\mathrm{m}^{-1}$ and 0.3 fixed, respectively. Our example soil with $n = 3$ and $K_0 = 5\,\mathrm{cm/h}$ is placed in these maps close to the lower left corner where the error in the four parameters and the χ^2-value are very low.
The region with accurate estimations is decreased compared to the previous section due to the larger average volume, although the infiltration rate is higher. This flux could not prevent this decrease, whereas the structure of the transition zone of this region stays equal. In this feasible region, the maximal water content rises as well as the minimal one due to the different curvature values n (cp. also Figure 5.18). They partly compensate each other but it slightly results an increase in the saturation range. Therefore, the water content range is shifted a little. The data distribution within the θ-domain changes. Therefore, we find similarities to the TDR-averaging case, where additional to the water content range, also the position of the range in the whole water content domain is important.
Despite the averaging, the ordering of the parameter errors is as discussed in chapter 3. Thus, the lowest error is calculated for n and the largest one for K_0 for both map types in Figure 5.23 a) and b). But the error range increases in all cases compared to the TDR-scenario due to the averaging.
In Figure 5.23 b), the α-ϕ-maps are drawn for a fixed n and K_0. As discussed in section 5.2, the transition between accurate and bad estimations is around 0.45 of the saturation range for lower negative values of α. Due to the stronger averaging, the maximal water content decreases. Therefore, the saturation range increases only slightly with increasing negative α-values. Thus, the estimation errors are larger for higher porosities since the saturation range is low.
Parameter errors of around one percent in the estimation are produced as a new feature in



(a) Variation of n and K_0 with constant $\alpha = -2\,\mathrm{m}^{-1}$ and $\phi = 0.3$.



(b) Variation of α and ϕ with constant $n = 3$ and $K_0 = 5\,\mathrm{cm/h}$.

Figure 5.23.: Maps varying true soil parameters p for the **150** % infiltration scenario (blue) in Figure 5.14. Each single point stands for a forward model with p_{true} and a corresponding inversion starting at $1.25\,p_{\mathrm{true}}$. The estimation error of each parameter is shown on the left and in the middle. At the top right $\Delta\Theta = (\theta_{\max} - \theta_{\min})/(\phi - \theta_r)$ is pictured and at the bottom right, χ^2-value in logarithmic scale. The black crosses in each subfigure denote the soil with χ^2-value larger than 10^{-10}.

these maps, although the saturation range is high, e. g. for small ϕ and more negative α. As we saw in Figure 5.4 with the different starting parameters, the way to the true parameters goes through a narrow valley. All ways enter in the same region around the optimum.
If we take a constant displacement for a fixed soil with the value 0.001 to each parameter and derive the deviation of water content due to this displacement, then we can calculate the χ^2-value for each parameter direction. The results are summarized in Table 5.1 for three different true soil parameter sets varying in α. First, we realize that the $\chi^2(\Delta p_i')$ is a large step dependent on the parameter type for this fixed displacement. The displacement of 0.001 is a very large step referring to the porosity, which is usually around $0.3 - 0.4$.In comparison the step is much smaller for K_0 since the saturated hydraulic conductivity could easily show values of $1 - 40\,\mathrm{cm/h}$.
Further, $\chi^2(\Delta n')$ is approximately 20-times larger than $\chi^2(\Delta K_0')$. This firms the observation of a narrow valley. If we compare the three soils in Table 5.1 pairwise by taking the ratio of their χ^2-values, we find that the ratio for $\Delta n'$ is higher than the one for $\Delta K_0'$. This means that the $\chi^2(\Delta n')$ increases stronger than $\chi^2(\Delta K_0')$ from the less negative α to the more negative one. Therefore, the valley will become more narrow if the true α-value is more negative. Corresponding to Figure 5.23 b), this leads to the more inaccurate estimations although the saturation range is high. If the valley is too narrow, the parameter steps will become shorter during the inversion. Thus, the optimization stops too early on the correct way to the optimum which is the case here for larger negative α. This behavior is not linear due to the underlying nonlinear dynamic, but the changes of the parameter space due to the parameters can be monotonic. The Table 5.1 gives an indication for it since the comparison of the ratios between different soils shows variations of the ratio values but no changes of the order of these ratios. The ratio of $\Delta n'$ is the highest in each case compared to the other parameters.

5.3.3. Summary to GPR

Finally, we found out that it is possible to estimate hydraulic parameter accurately in the GPR scenario. But the regions shrank due to the large average volume, where an accurate estimation is feasible. The transition zone of the flux rises to higher flux rates due to the averaging whereas the transition zone of the water content range stays similar for different infiltration scenarios. The $\Delta\tilde{\theta}_{\mathrm{trans}}$ is quite stable with respect to the averaging around $\phi/2$, but the averaging reduces the water content range for all scenarios due to damped amplitudes.
These observations can be transferred to the estimation maps. The averaging lowered the

Table 5.1.: Comparison of χ^2-values for soils with different true α_{opt} and fixed true parameters $(n, K_0, \phi) = (3, 5\,\mathrm{cm/h}, 0.3)$. The time-series in water content with the true parameter represents the measured data. The simulated θ-time-series is produced by a displacement of 0.001 to each parameter, respectively, belonging to the column of this table. $\chi^2 = \sum_i (\theta_{\mathrm{meas}}(t_i) - \theta_{\mathrm{sim}}(t_i))^2$. The last column block is the ratio between the χ^2-values of the first and second row with true α of $-4\,\mathrm{m}^{-1}$ and $-3\,\mathrm{m}^{-1}$.

α_{opt}	χ^2				$\frac{\chi^2(\alpha_{\mathrm{opt}})}{\chi^2(\alpha_{\mathrm{opt}}=-3)}$			
$[\mathrm{m}^{-1}]$	$\Delta n'$	$\Delta\alpha'$	$\Delta K_0'$	$\Delta\phi'$	$\Delta n'$	$\Delta\alpha'$	$\Delta K_0'$	$\Delta\phi'$
-4	1.770×10^{-7}	6.287×10^{-10}	9.210×10^{-9}	3.523×10^{-5}	1.086	0.546	1.060	1.068
-3.6	1.719×10^{-7}	7.683×10^{-10}	9.032×10^{-9}	3.446×10^{-5}	1.055	0.668	1.039	1.045
-3	1.630×10^{-7}	1.150×10^{-9}	8.689×10^{-9}	3.298×10^{-5}				

saturation range and the transition zone in these maps went towards more negative α-values and smaller K_0-values. The regions with non-feasible estimation increased and the region was limited for more negative α-values additionally in the GPR-case. There, it became dominant that the valley of the parameter space was more narrow in the direction of parameter n for more negative α. This restricted the feasible region more and more. But still, the most useful soils are a mixture of fine and coarse materials.

5.4. Summary

All together, forward simulations and parameter estimations were compared for several scenarios to study to feasibility of accurate parameter estimations. We analyzed three types of averaging realizations corresponding to the point sensor, TDR and GPR. For the first one, only the center of one grid cell was evaluated whereas the other two included an additional averaging over depth.
For each of these types, we found regions where is feasible to accurately estimate the hydraulic soil parameters. Such estimations could be referred to as a water content range over $\phi/2$ or at least in a zone around it, the so-called transition zone. Especially, for the point sensor we found a larger sensitivity of the transition range on the structure of the infiltration flux. This means that short and intensive rain event shifted $\Delta\tilde{\theta}_{\text{trans}}$ to higher values whereas more uniform distributed and stepwise rising fluxes decreased the transition range $\Delta\tilde{\theta}_{\text{trans}}$. This got more stable by averaging over depth like in the cases of TDR and GPR. But there, the transition zone of the water content range was more stable than the one of the maximal infiltration flux.
The dependency of the estimation result on the starting parameter were studied for all three measurement types. The parameter estimation was also stable with respect to different starting parameters when the accurate estimation of the parameters was feasible. For choices of starting parameter we always observed a convergence towards the optimal point. But for lower infiltration fluxes the optimum was not reached since the sensitivity of the water content due to parameter changes was too low and the optimization steps were too short.
Further, we observed a small valley of the parameter space which is orientated along the direction of K_0 and ϕ. Especially, for the GPR case this valley became significantly smaller such that the conditions of feasible estimation were lower for much more negative α-values. Here, a scaling of the estimation parameters could be helpful.
In general, the estimation became more difficult with increasing average volume since the dynamic variation of the time-series was damped with greater depth and by averaging. In the corresponding example of TDR, the importance of a correct averaging volume was demonstrated.
Moreover, we learned that soil with lower K_0 show a more accurate estimation in the study of different soils. But it is limited from below by the flux rate j_w. More negative values of α produced lower estimation errors whereas the value of the parameters n and ϕ show only less influence to the feasibility of estimation. Translated to soil type, this means that coarser soils with finer soil fractions supported an accurate parameter estimation to a higher extent. But the water content decayed very slow for pure fine soils like pure loam or clay so that the dynamical range of water content is very low.
Moreover, the transition zone for the soil types of the maps shrank the feasible region of good estimations by changing the three different measurement types. The feasible region was largest for the point sensor. The averaging decreased the feasible region for GPR.
Finally, we can summarize very shortly that larger and stepwise increasing infiltration flux sup-

ports an accurate parameter estimation to a greater extent. Further, coarser soil mixed with finer materials give bester hydraulic conditions to estimate accurate parameters due to their wider response of the water content.

6. Application of Estimation and Projection to Measured TDR-Data

In this chapter, the parameter estimation of the hydraulic set (n, α, K_0, ϕ) will be apply to measurement data based on the time domain reflectometry (TDR) method. A numerical forward simulation determines the modeled water content in an uniform medium. As described in section 3.3, input information is the time-series of measured water contents in a certain depth and the measured infiltration flux. Further information to the soil profile of the measurement data can be found in the appendix A.2.
Additionally, we check the developed conditions of feasible estimation (section 5.2) for this measurement set to get an impression how reliable the estimated parameter is later. Since the soil in the field is a layered medium, we comment on the question how the changes are to estimate and project for heterogeneous soils and to which depth the projection is possible. And we discuss shortly what is important for the estimation interval during other times or which role hysteresis plays in our example.

Spin-up

In all following simulations, a spin-up phase of 59 d is used to let the model run into a realistic water content distribution over depth. Since the water content depends on the hydraulic parameters during the spin-up phase, directly after the spin-up the water content profiles can vary with the current parameters during the optimization process. But after the spin-up phase the profile values should vary only little due to the initial condition. Since the error due to the initial condition depends on the hydraulic parameters and the spin-up phase one could optimize the duration of the spin-up phase by estimation. But here, we choose the spin-up duration long enough to prevent such larger effects within the estimation time interval due to the choice of the initial condition. This means that the whole interval is simulated but the spin-up phase is excluded in the objective function.
Hence, the spin-up duration is chosen to 59 d which is close to two months. In Figure A.4, we observe a longer drying out phase before day 1067 over more than 30 d. During this time, the hydraulic system can equilibrate and provide a starting condition, which excludes very large sensitivities due to the estimation parameters. But not only a drying phase is advantageous during the spin-up phase. Also same infiltration events at the beginning of the simulation interval should be included in the spin-up. These external infiltrations force the modeled water distributions to change their state into a realistic hydraulic state.

Estimation Conditions

In Figure 6.1, we choose the same estimation time interval as in the case of real infiltration for the synthetic data 5.14 and with 100 %-flux, which is shown in the top Figure 6.1. Below, the measured water content is drawn at 8 cm depth with a basic resolution of 1 h and in red like all measured quantities in this chapter. A first steep slope of the measurement time-series occurs minutes after the first higher infiltration flux of 5.3×10^{-6} m/s. After a few hours, the second

and highest infiltration peak reaches the sensor depth of 8 cm and the water content rises again with a steeper slope.
These events, corresponding to a larger deviation of the water content and a decay of θ afterwards, are the important structures for the estimation. This dynamic and the water content range of 0.112 volume fraction shows conditions that make an estimation promising. The water content range of the measurement is only 41 % of the minimal possible water content range $\phi - \theta_r$ based on the assumption with $\phi = \theta_{\max} = 0.2736$ and $\theta_r = 0$. But in the synthetic study to this infiltration scenario (Figure 5.17) the transition zone of the water content range was also lower than $\phi/2$ due to the averaging over depth within the model. Furthermore, the synthetic analysis showed in section 5.2 that this infiltration flux is strong enough for a feasible, accurate estimation for soils with lower K_0, larger negative α, nearly all n and ϕ, e. g. soils with larger sand contents. Therefore, we can conclude that the conditions given from these measurements can allow an accurate estimation, since these measurements show properties which are within the transition zone of feasible accurate estimations.

6.1. Parameter Estimation

Similar to the real infiltration in the synthetic case (Figure 5.14), we focus on the estimation of the 3 days-interval with the 59 d spin-up. The time-series of measured water content is plotted with red crosses in Figure 6.1 and is measured each hour. The optimization strategy as described in chapter 3 leads to the time series of simulated water content in Figure 6.1 denoted by the dark blue dots. Analogue to the synthetic case, we assume an averaging volume between 4.5 cm and 10.5 cm in our model since this is approximately the diameter of the TDR-sensor.
This optimized dark blue time-series shows all in all a residuum between measured and modeled θ of around ± 0.01 drawn in dark blue at the bottom of Figure 6.1. Although the water content rises over a range of 0.112 volume fraction, there is still a trend within the time-series of the residual. It can also be observed by looking at the time-series of measured and modeled water content. After the rain events the flux goes down to zero and the water content decreases in the measurement depth of 8 cm. The modeled water content shows similar behavior, but with a slightly different slope of decrease. One reason can be that the estimated parameters are not yet accurate enough since all four hydraulic parameters affect the slope of the time-series of water content. This fact can appear in a direct way like n is changing the slope or in an indirect way by the correlation of the parameters. It can be compensated up to a certain accuracy which can be quite close to the real time-series if one parameter has a larger error. In this context the principle of equifinality is well-known. *Beven et al. (2001)* has explained that there does not need to be one unique optimal parameter set. Further, he explained that several minimal points can have a similar residuum. It is comprehensible in the context of parameter estimation with real measured data sets and their measurement uncertainties.

But still we observe a decrease, then an increase and again a decrease of the measured water content during the last two days in Figure 6.1. The variation of water content is not reproduced by the modeled time-series since the model describes only the infiltration fluxes from top. However, there is no measured rain event which is strong enough to let the water content rise by 1 volume percent before the increase of water content. We can exclude a missing of precipitation measurements during this time since at another measuring weather station not far away from the Grenzhof test site has also not measured a precipitation on this day. Since the measured data is taken from the bare soil profile, roots of grass cannot effect the measured

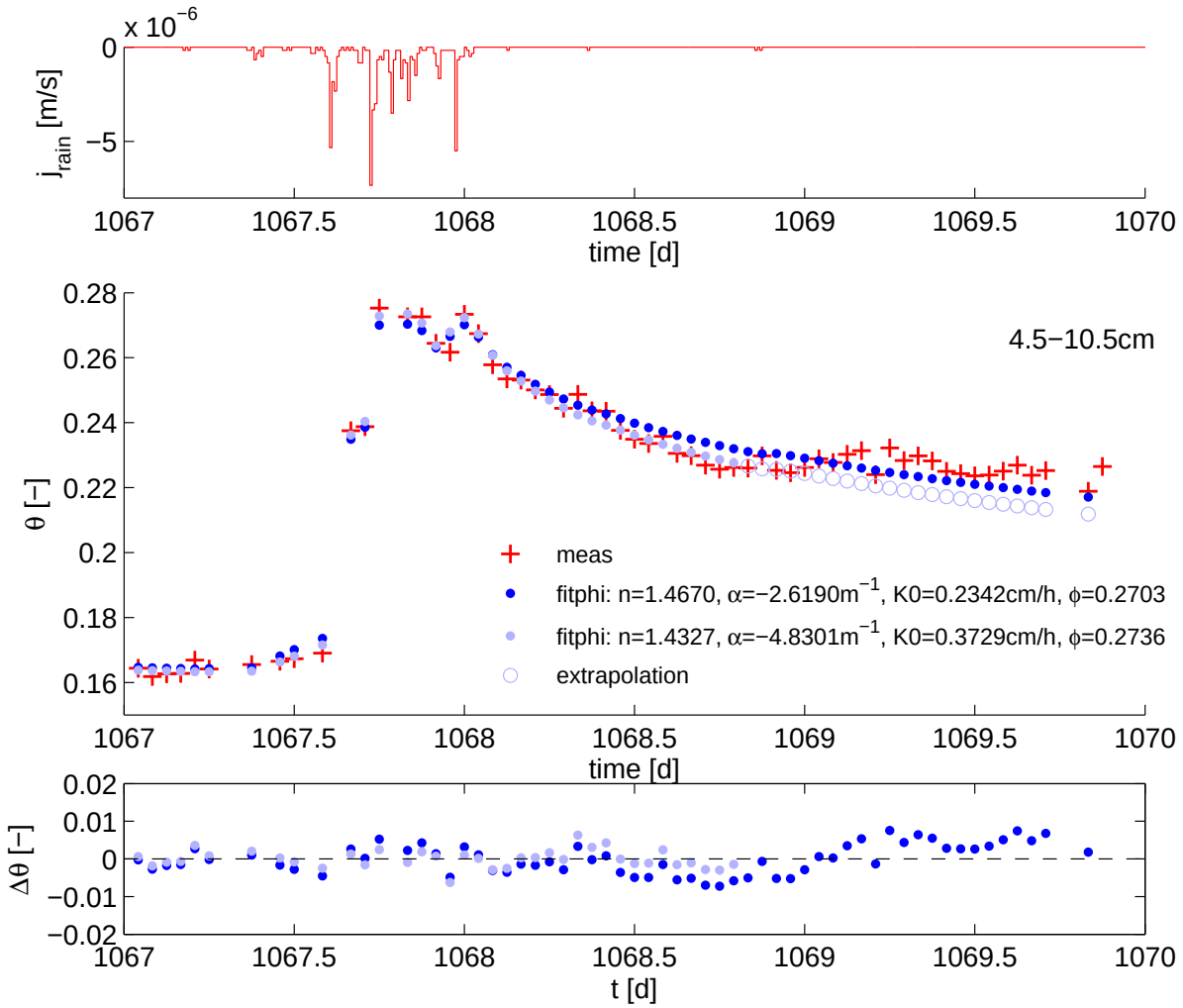


Figure 6.1.: Estimation of near-surface water content over 3 d (dark blue) and 1.84 d (light blue) with 59 d spin-up time based on Mualem-van Genuchten parameterization. Top: Infiltration flux during the estimation time interval given by precipitation measurements per 10 min. Middle: The red crosses denote the measured water content at 8 cm depth and the filled dots are the optimized water contents averaged over $4.5 - 10.5$ cm depth. θ is shown extrapolated with the optimized parameters of the 1.84 days-fit in light blue circles. Bottom: Residual between measured and simulated water content.

data. One can think about a decrease of temperature over night and with that also a change of relative humidity. This can increase the soil moisture within the top soil layer.
All in all we know that these variations of the measured water content have to be a process which is not described with our infiltration model. This means that if we include these data points, like in the dark blue time-series, the variation of the measured water content will be compensated by the estimated parameters of the modeled θ since the only quantities which can be adapted during the inversion are the four hydraulic parameters.
Therefore, we exclude these measured data point and truncate the estimation interval after 1.84 days. This leads to the light blue time-series in Figure 6.1. The time-series of water content

and the corresponding residuum match well the measurements, since the residuum varies only within ± 0.006 of the volume fraction. Also the Levenberg-Marquardt parameter λ is much lower with 10^{-7} for this 1.84 days-estimation than for the 3 days-estimation with $\lambda = 10^5$. This fact shows that the Gauss-Newton method is weighted stronger than the steepest descent for the estimation over 1.84 days whereas the final 3 days-estimation strategy with the high λ follows the steepest descent algorithm. The Gauss-Newton algorithm converges faster than the steepest descent in the neighborhood of the optimal point (*Madsen et al. (2004)*). This illustrates a better converges to the local minimum. But this method will be more robust if we start further away from the optimal point. Hence, the stopping criterion can be fulfilled, although the current parameter set is far away from the optimum. We discussed already that the 3 days-estimation includes a time period in which an additional process is active in the measurements. This is also reflected in the behavior of the optimization algorithm.
As good as the estimated time-series is for the 1.84 days, as questionable are the optimized parameter. This does not mean that the mathematical algorithm finds the wrong minimum or does not converge correctly. Optimization runs with different starting parameters and reaches all the same region in the parameter space.
In this case, the estimated parameters are not realistic physically. For example, the estimated porosity ϕ is equal to the highest measured water content within these 1.84 days. This expresses that the soil is full-saturated after reaching the second rain event which stays around day 1067.7 only for ten minutes. Afterwards, the infiltration flux and the measured water content decrease, whereas the simulated water content (light blue) stays constant till the next measurement time. Moreover, we know from the whole measured time-series over at least two years that the water content in depth can be at least $1-2$ volume percent higher and therefore also the porosity as highest possible water content.
A further questionable value is estimated for the van Genuchten parameter n. As *Ippisch et al. (2006)* showed the simplified van-Genuchten parameterization does not describe the physical behavior correctly anymore for a value of n below 2 under full-saturated conditions. *Ippisch et al. (2006)* recommends to apply a parameterization with an air-entry-value. In our case, the parameter n is 1.4 and the modeled water content reaches the value of the porosity. Under this saturation condition, the estimated value of n is critical, but in literature (*Schaap et al. (2001)* and *ROSETTA (2005)*) such values of n can occur, for instance for soils with high clay or loam content.
Further, the saturated conductivity K_0 is much lower than expected. The highest infiltration fluxes are 2.6 cm/h whereas the estimated saturated conductivity with 0.37 cm/h is smaller by almost a factor of 10. Mathematically, we apply the Buckingham-Darcy law (3.2) which relates the water flux, the hydraulic conductivity function $K_0 f(\theta)$ and the gradient of the water potential to each other

$$0 \geq j_w = -K_0 f(\theta)\, \partial_z h_w \,. \tag{6.1}$$

The function $f(\theta)$ describes the form of the parameterization curve which is given by the Mualem-van Genuchten curve (3.4) in our model. Its values are within the interval $[0, 1]$.
We consider the large negative flux j_w and a low positive saturated conductivity K_0 in this case. To hold the equation (6.1), the gradient of the water potential has to be very large by compensating the large infiltration flux. Translated to reality, it means that the whole falling rain water enters into the soil at the given time. Huge gradients of the potential are constructed. In reality, it can only occur in the case of larger infiltration events, that only part of the water infiltrates into the soil and water is ponded on top of the soil surface. In this case, the input water flux and the response time-series of water content are mismatched. Thus, we try to

compensate it by adapting the hydraulic soil parameters, which produces estimation errors of the parameters.

6.2. Estimation with Reduced Infiltration

We learned from the previous discussion that the rain flux in our application example is too strong in comparison to the response with the soil. In order to compensate this behavior we reduce the infiltration flux by 5×10^{-7} m/s, i. e. 0.3 mm with a 10 min-interval. This reduction can be interpreted as evaporated water which evaporates directly after the falling or during ponding.

6.2.1. Full Measurement Set

Moreover, we want to limit the flux reduction for very small rain events. Thus, if the originally measured infiltration flux is smaller or equal 5×10^{-7} m/s, then the infiltration flux will be set to minimal measurable infiltration flux 1.7×10^{-7} m/s, i. e. 0.1 mm with a 10 min-interval. These small infiltration events are still important for the fine-structure. It can be observed in Figure 6.1 just before the strong rising of the water content. There, the four measurement points (red crosses) around day 1067.5 increase monotonically. During this time the simulation also increases due to the very small rain events, but stronger than the measurements. For instance, the modeled water content will decay or keep constant if we neglect these small fluxes, but it will not increase like the measurements.
Furthermore, we only consider the 1.84 days-interval from above together with the 59 d spin-up for the estimation in the following. The described reduced infiltration flux is shown in Figure 6.2 in the top figure.
In the middle of Figure 6.2, the measured water content is drawn in red crosses together with the simulations in light and dark blue. First, we focus on the light blue time-series. Here, the conditions are the same as for the 1.84 days-estimation of Figure 6.1 except for the reduced rain flux.
The modeled time-series of water content still fits well with the measurements for higher water contents and the decay part during the second day with the rain reduction. The estimated hydraulic parameters are also within a realistic range. The optimized porosity ϕ is now around 2.1 volume percent higher than in the case of the pure measured flux with 29.25 volume percent. As indicated in the previous discussion about porosity, its value is much more realistic since the largest measured water content in this depth is represented now over the last two years.
Moreover, the saturated conductivity is estimated to be much higher for the reduced boundary flux. Here, the infiltration flux is smaller than the saturated conductivity which means that no huge gradients of potential are constructed anymore to fulfill the boundary condition. We expect that the gradients of potential are moderate during the rain events in the autumn- till winter-term without huge evaporation potential like in summer.
The van Genuchten parameter n is decreased and is still smaller than 2 though the estimation, but the full-saturation of the soil during the estimation time does not stay anymore. The estimated porosity is higher. The maximum of the measured water content keeps the same. Hence, the strong curvature of the Mualem-van Genuchten curve is not fully included close to saturation anymore and this relation makes the parameterization more reliable. We will discuss the shape of the estimated parameterization curves further in section 6.2.3. Furthermore, the low values of n like 1.2 are observed typically for clayey or loamy soils, see *ROSETTA (2005)*.
If we look at the residuum between measured and modeled water content, we can observe a

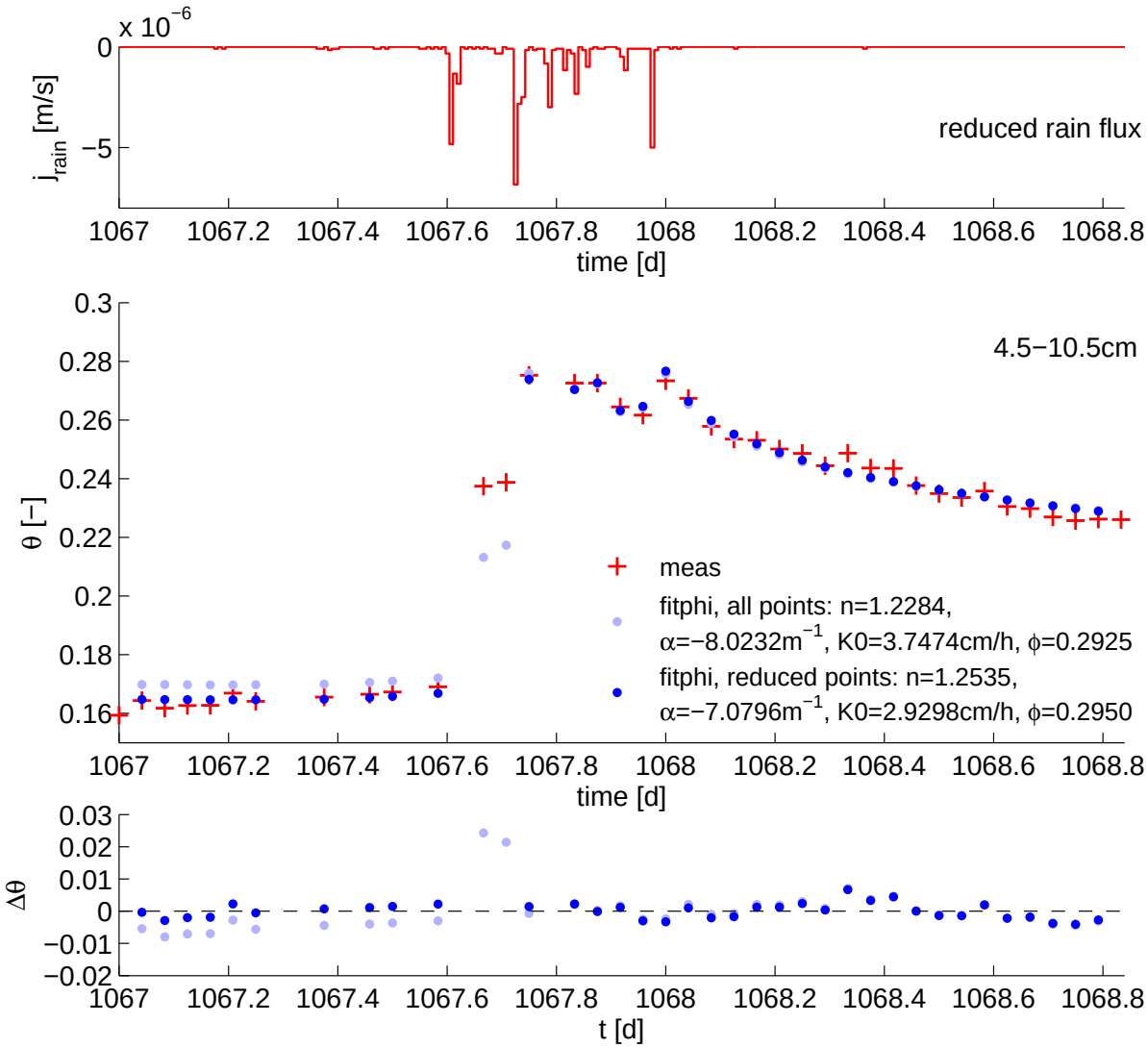


Figure 6.2.: Estimation of near-surface water content over 1.84 days with 59 d spin-up time based on Mualem-van Genuchten parameterization. Top: The reduced infiltration flux during the estimation time interval given by precipitation measurements. Middle: The red crosses denote the measured water content at 8 cm depth. The filled dots are the optimized water contents averaged over $4.5 - 10.5$ cm depth. All time points within the 1.84 days-interval are used during the optimization for the light blue time-series. The two points at the increasing slope are excluded for the dark blue time-series of water content. Bottom: Residual between measured and simulated water content.

higher difference for the increasing slope and the low water content part during the first day of the estimation interval. At the slope, the largest estimation error occurs with more than 2 volume percent. By trying to reduce this large error as a larger term within the objective function, other terms can be increased in order to reduce the total error, measured with the χ^2-value. This balance is expressed by the error during the first day of the estimation interval. Hence, we should focus on what is happening mathematically and physically during the water

content rising, when the first larger rain event penetrates through the dryer soil. Mathematically, the dynamic of the water through the unsaturated soil is described by the Richards equation together with the chosen Mualem-van Genuchten parameterization. In case of homogeneous and of horizontal layered soils, a horizontal water front is formed after the infiltration of water. This is based on the Richards model and hence on the local equilibrium assumption of all quantities for our model volume. A continuous air and water phase is required for this model, *Roth (2007)* and *Flühler et al. (2004)*.
What happens physically during an infiltration event on the other hand? There is dryer soil as a porous medium on which the rain falls. Assuming that the rain falls evenly in space, the water front is formed after a certain depth. But just in case of stronger rain events the air within the pores can be compressed if the water front penetrates faster than the physical system needs time to equilibrate. In such a case, the assumption of the Richards model of constant air pressure within the soil is not fulfilled. Some experimental work showed the so-called *fingering*, *Glass et al. (1989)* and *Rezanezhad et al. (2006)*. Moreover, a real soil is not homogeneous. Besides the heterogeneities of layers or soil lenses within a soil layer, soils can have cracks or wormholes. These represent a larger pore, through which water can infiltrate easier and faster than the global water front.
Thus, these differences of real soils to the modeled soils have a larger impact in cases of irrigation from soils that are more dry to more wet. Now, we consider our TDR-sensor at 8 cm depth with its measured time-series (red crosses) in Figure 6.2. When the infiltrated water runs through the dry soil, the catchment of the TDR-sensor measures different states of the soil moisture with time. Then the water content rises, which is reflected in the measured travel time of the electromagnetic wave within the TDR-sensor. Details are in the appendix of *Ludin (2010)*. But in the real soil, the water front is not a clear front due to the above discussed effects. Therefore, the water is not distributed uniformly over the catchment of the TDR-sensor during the irrigation process. In such a case, the non-uniform distribution of water can lead to higher uncertainties in the measured water content values during this time. The measured water content value will be stronger error-prone when the water content change is very large during short times. At time periods with small water content changes the water is closely uniformly distributed and hence, we expect to have less measurement uncertainties.
The consequence from this discussion about the mathematical and physical view during infiltration is that we rely less on the measurement data during the strong increase of water content. These two measurement points will represent a large term of the objective function if model and measurement data does not coincide. The difference between the mathematical and the physical point of view shows that a mismatch of the measured and the modeled water content can easily occur during this physical state. Consequently, we neglect these two points and exclude these in our objective function in order to provide influences to other estimation times, as in the light blue time-series during the first day (Figure 6.2).

6.2.2. Reduced Measurement Set

As discussed above, we remove the two measurement points of the increasing slope from our objective function. The optimization and the forward modeling with the reduced infiltration flux is chosen as in the previous example. The resulting optimized time-series of water content and its corresponding residuum are shown with dark blue dots in Figure 6.2. Here, we notice that the estimation regions (second day), where the previous estimation yield good results, is still accurately represented.
During the first day of the estimation interval the modeled time-series describes the measured

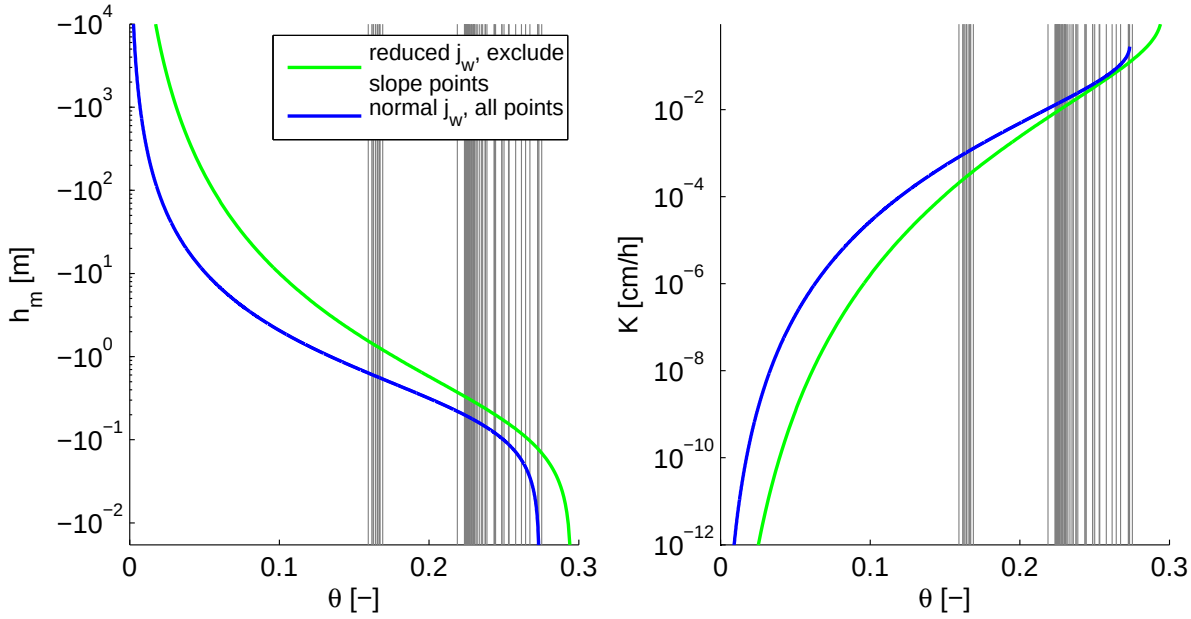


Figure 6.3.: Parameterizations for the estimated parameters in green for the reduced infiltration flux and excluding the two slope points (Figure 6.2) and, in blue, for normal infiltration flux (Figure 6.1). Both parameter sets are estimated during 1.84 days. The gray lines denote the water content at which a measurement point exists within the estimation interval. Left: Soil water characteristic $h_m(\theta)$ is based on van-Genuchten parameterization, here drawn in logarithmic scale. Right: The hydraulic conductivity function $K(\theta)$ is based on Mualem-van Genuchten parameterization, here drawn in logarithmic scale.

ones accurately and thus, the estimation error is improved. The estimation of the reliable measurements is dominant and the error-prone part of the time-series is neglected without fitting these two data points.
Further, the estimated parameters are changed but still of the same order. Furthermore, the porosity is realistic with 29.5 volume percent. The saturated conductivity is lower now, but with a value of 2.9 cm/h higher than the maximal infiltration flux of 2.6 cm/h. This is much more realistic than the estimated K_0 for the normal infiltration flux. The estimated van Genuchten parameter n is larger for excluding the slope points. The water content is reduced with increasing the value of n (Figure 3.5 a)), especially during the dryer period. The change in n has also an impact onto the other hydraulic parameters due to their correlation. Especially, the correlation between n and α are stronger pronounced and changes α.

6.2.3. Parameterization Curve

In the following, we focus on the shape of the resulting parameterization curves for the estimated parameter sets. Especially, we compare the shape for the parameter sets of two cases, both for the 1.84 days-estimation interval, the one with reduced infiltration flux and excluded slope data points (Figure 6.2 dark blue) compared to the one with the normal infiltration flux (Figure 6.1 light blue). In Figure 6.3, the soil-water-characteristic $h_m(\theta)$ (left) and the hydraulic conductivity function $K(\theta)$ (right) are plotted for these two parameter sets, in green the one for the reduced rain and in blue the one for the normal rain. The gray lines denote the

water content at which a measurement point exists within the estimation interval of 1.84 days. We first observe referring to the blue curve for the normal rain that the saturation is reached during the simulation since the blue curve ends with the highest water content close below the highest measured water content.
Further, the largest gradients of all parameterization curves are located close to the bounds of the water content domain (θ_r, ϕ) since the curves are monotonic and with an inflection point within the interior. This is valid independent of linear and logarithmic scales. These gradients get important in combination with our given measured water contents (gray lines). The region of the gray lines overlaps with the saturation region of the blue curves with the normal rain. There, the gradients of the curves are much stronger than within the interior as in the case of the reduced rain in green. As we learned from relation (6.1), a infiltration flux which is larger than the saturated conductivity produces a huge gradient of potential.
Following the parameterization curve away from the boundaries, the gradients of these curves become much smaller. This is the case for the green curve with the reduced rain within the interval of the measured water content. As discussed previously, these smaller gradients are more realistic as well as the observation that the saturation is not reached at a water content of 0.2736 volume fraction. *Ippisch et al. (2006)* explained for the Mualem-van Genuchten parameterization, that the steep form of $K(\theta)$ close to saturation is an artifact for $n < 2$. This is less important for the case with the reduced rain, since the measured data is not at saturation. The main difference between the most common parameterizations of van Genuchten and Brooks-Corey is the shape of the curve around the saturation. As mentioned in section 3.1, Brooks and Corey describe the soil water characteristic with a discontinuity around saturation whereas van Genuchten uses a continuous function. The two parameterizations are more similar for water content values further away from saturation. Due to the higher porosity of the green curve, no full saturation occurs and the different form of the parameterizations is less important for high water contents.
Moreover, the van Genuchten parameter n is estimated for the reduced rain with 1.2535 smaller than for the normal rain with $n = 1.4327$, but still both below 2. In Figure 6.3, we recognize this fact at the larger curvature of the blue curve, for which n is a little higher. In general, these values of n are realistic for clayey or loamy soils (*ROSETTA (2005)*) where also a lot of very small pore sizes exist.
The lowest measured data points with the estimation interval (gray lines) are given around 16 volume percent. In this context, it is important to remember that these measurements are taken at the end of a dry period with low infiltration fluxes below 10^{-7} m/s. Almost no water is infiltrated during 40 days and the water content depth goes down at 8 cm. Due to the stronger capillary forces in finer soils, the water content moves slowly. It can be observed in our example, where the water content only goes down to 16 volume percent and the groundwater level stays several meters under the ground.
As discussed at the beginning of of this chapter, the measurement data does not cover more than 50 % of the water content domain $[\theta_r, \phi]$. Due to our analysis in chapter 5, an estimation of this data set does not need be very accurate, but it is perfectly situated within the transition zone. A practicable estimation is possible within this zone and also with including some uncertainties due to the measurements.

6.3. Projection of Water Content

After our analysis for estimating the hydraulic parameters and its corresponding challenges, the question to answer is how good the performance of the hydraulic parameters in the context of projection into depth is. Operationally, the approach to project the water content into depth is very simple compared to the estimation procedure. We work with the forward model, the Richards solver, which we introduced in chapter 3 and used for the inversion (Figure 3.2 b). However, we assume an uniform soil whereas the measurements are taken in heterogeneous and layered soil. Hence, we have to deal with this mismatching and study what is possible for the projection despite the heterogeneity.

We should again mention at this point that our model assumes an uniform medium. The time-series of water content in Figure 6.4 are averaged over a catchment area of ± 3 cm. The middle subfigure of Figure 6.4 shows these simulated and averaged time-series of water content (gray) from depth 3.5 cm to 29.5 cm in steps of 2 cm. The top most gray time-series corresponds to the 3.5 cm. Their linear interpolation expresses a non-smooth shape whereas the interpolated curve gets smoother with greater depth. The reason for this is based on the damping of the information content with depth. This means that the diffusion part of the PDE (3.1) smears out the information of infiltration peaks with depth and time. Further, this illustrates that the 1 h-resolution is practicable for the deeper time-series, but for the higher ones the finer structures are missing. On the other hand, our interest here is only on the comparison with the measured water content (crosses), which is given at the same points in time as the simulated ones.

The blue crosses denote the measurements of the highest sensor at 8 cm depth, which we used for the estimation of the parameters. This measurement time-series corresponds to the simulation of the third highest gray curves with 7.5 cm. Except at the two slope points, which we exclude for the estimation, the measured θ in blue fits the third highest simulated θ in Figure 6.4 by construction.

Now, we compare the simulated water contents (gray) to the other measured time-series, especially the second highest time-series at 19 cm depth denoted by the red crosses. The first observation is obviously the much higher water content of the two deeper sensors compared to the measurements in blue. The difference is around 6 volume percent before the rain events and afterwards around 3 volume percent. All these three sensors are within the same soil layer (Figure A.3). Therefore, we would expect for an uniform soil far away from the capillary fringe that the water content would not increase as fast with greater depth during the wettening of the soil. Here, this is not the case, which indicates that the water cannot drain as fast as it could at the lower boundary of the top layer. The water accumulates there. In a depth of around 35 cm, the next layer interface is located. This is little below the two lower sensors. The finer soil texture of the first layer and the coarser one of the second layer can lead to an accumulation of water if the lower layer is dry and the surface flux is high (*Roth (2007)*). We know from the soil characterization in Figure A.3 that the second layer is loamy sand and the first layer is finer as sandy loam. Measurements deeper in the soil around $50 - 60$ cm also show a decrease of water content. This pushes together layer locations.

However, we can observe a difference in the shape of the measured time-series for the 8 cm and 19 cm depth in Figure 6.4. The reaching of the infiltrating water into their depths is different. Thus, the water content increases earlier for the highest sensor as expected and it reaches its maximum earlier. We cannot reproduce the offset to higher water contents with our model together with the optimized parameters from section 6.2. But we can focus on the shape of the time-series.

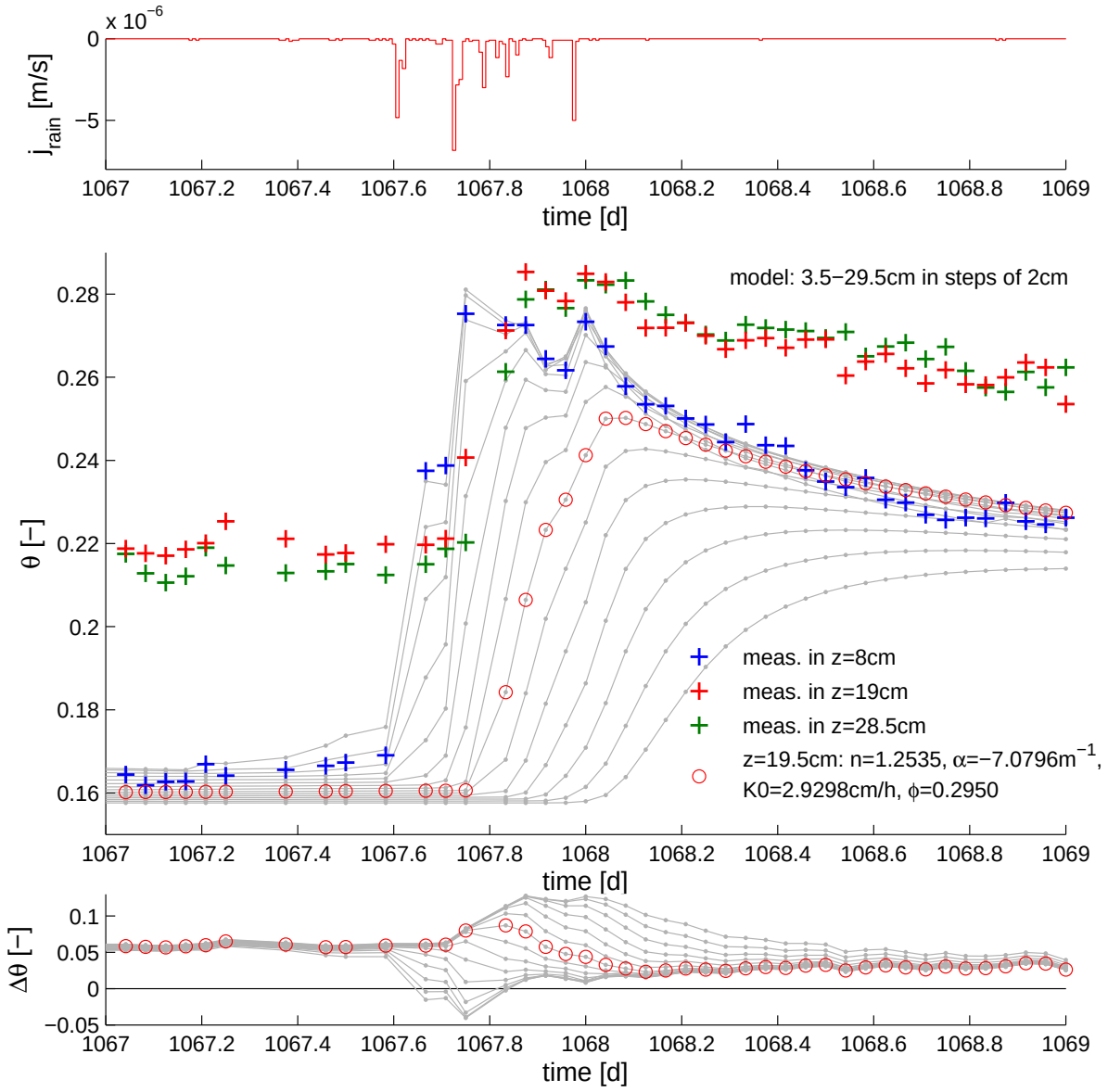


Figure 6.4.: Projection of the ±3 cm-**averaged** time-series of water content from 7.5 cm into 19.5 cm. Top: The reduced infiltration flux is given during the estimation time interval. Middle: The measured water contents at depths 8 cm, 19 cm and 28.5 cm are drawn as crosses in blue, red and green, respectively. The gray dots with linear interpolation lines show depths from 4.5 cm to 32.5 cm in steps of 2 cm for the estimated parameters from Figure 6.2. The time-series at depth 19.5 cm (red dots) minimizes the residuum at the bottom subfigure. Bottom: Residuum between the measured water content at depth 19 cm (red crosses) and the simulated θ in several depth (gray).

Hence, we derive the residuum of the measured water content at 19 cm depth (red) with respect to each of the simulated and averaged time-series (gray). At the bottom of Figure 6.4, these residues are drawn. During the dryer period, there is almost no difference between them. The longer the decay of the water content takes after the rain events, this is also the case. This will

be comprehensible if hydraulic state becomes closer to its equilibrium where the water content is almost constant and far away from the groundwater.
During the infiltration events, we observe an asymmetric deviation from a constant offset. For the shallower depth, the residuum $\Delta\theta$ becomes negative. There, the water reaches the corresponding depth earlier than the measured θ at 19 cm. Water content is larger than the measurement. The travel times of the infiltrating water increase with greater depth and become similar till they let the residuum rise. Then, the water longer penetrates into these deep depths and also the amplitudes are more damped, due to the included diffusion subprocess.
Moreover, we consider the simulated time-series at 19.5 cm depth marked with red circles in Figure 6.4. This time-series shows a larger residuum compared to other depths, but the slope during the beginning rain fall is similar. The slope of the decay also behaves in a similar way, since the offset during this time is almost constant.
Further, the simulated θ of 19.5 cm needs more time to reach its maximum since the water content level starts lower than the measurement in the similar depth. The hydraulic conductivity of the soil depends on the water content and is higher in case of higher water content. In the measurement case, the water is held back due to the layer and lets the water content increase compared to a pure uniform soil as in our simulation. Thus, if the soil is more wet, the water can flow faster due to less wettening of the soil. Therefore, the circled time-series reaches its maximum not until day 1068, whereas the measurement takes its maximum before day 1068.
Finally, the third measurement depth is plotted in green in Figure 6.4. Its water content values are quite similar over the whole time of this estimation interval. At the beginning, θ of 28.5 cm is a little lower than θ of 19 cm, but not as large as it could be for a difference at a depth of 9.5 cm. Here, we see again the influence of the layer interface. This leads to a much stronger mismatching of the simulation and measurement, each at 28.5 cm depth. The effects of the layer are not included in our model and therefore, it cannot reproduce the measurements. This effect becomes stronger as the comparison between the three depth shows.

Summary of the Projection

A projection of water content into depth is only successful if all model assumptions are fulfilled. For our example, the projection has to deal with the different assumptions of the soil structure, for the model of uniform soil and for the measurements of layered soil. Hence, there is an offset between model and measurement of order $3-6$ volume percent. The accumulation of the water in reality slightly influences the time point of the θ-increase due to the dependency of the velocity and the water content. But the slope of the measured time-series in depth can be reproduced with a time decay.
Moreover, the influence of the accumulated water increases with greater depth and avoids any type of reproduction of the correct water content within greater depths.

6.4. Further Aspects

In the following, we focus on a few aspects with which we have to deal in the context of measurement data. One approximation in our model is that soil is assumed to be homogeneous whereas real soils are heterogeneous. A further aspect is how the time interval should be chosen in order to apply the estimation on other times. Finally, we discuss effects of the hysteresis.

6.4.1. Heterogeneous Soil

Above, we applied the estimation of the hydraulic parameters and the projection of water content to a field data set at Grenzhof test site. Since there the soil has several layers, the assumption of a homogeneous medium in the model is not fulfilled.
As we discussed above, the parameter estimation itself was complicated. We verified the conditions of the model and the measurement to be able to handle the improvement and the reliability of the estimated parameters. Therefore, we excluded some points in time since the measurements are more unsure. But also the model was adapted by reducing the rain to model a more realistic scenario.
A parameter estimation is basically possible for the top centimeters also in heterogeneous medium. But in any case, the conditions of the homogeneous study for a feasible estimation (chapter 5) has to be fulfilled at least. The heterogeneity of a soil can additionally reduce the number of feasible soils and infiltration fluxes.
Durner et al. (2008) concluded from his lysimeter experiment that the hydraulic parameters can be determine uniquely for uniform soils. But they detected more unsure estimations with larger heterogeneity of the soil. This is consistent with our study in chapter 5: If the estimation conditions are fulfilled then the parameters can be estimated uniquely. And further heterogeneity of soils can restrict these conditions additionally.
But we notice that their results are based on quantities in the objective function which are deeper and in more layers up to 150 cm, whereas our estimation study handles a few top centimeters of the soil. Therefore, the influence of the heterogeneity is less on the estimation since the one time-series is only in one layer.
Nevertheless, the influences of the heterogeneity increases with depth. In our example, we could identify the equal slopes of the time-series for the model and the measurement. But ponding water prevents further reproducibility due to the intersection of the layer. Hence, a projection of water content through such layer is very hard. Then the knowledge about the important layers are required to be included into the model and to improve the whole projection result.

6.4.2. Time Interval

Exemplary, we presented the parameter estimation and the projection of water content based on measurements at one time interval. If we want to analyze further examples at different times or with different soils, we should consider a few criteria.
First, the water content range within the estimation interval should be large enough. In chapter 5, we found that an accurate estimation is feasible if the water content range is larger than approximatively one half of the porosity. The transition from non-feasible to feasible accurate estimations is a zone since the distribution of the infiltration flux influences the transition range $\Delta\tilde{\theta}_{\text{trans}}$. Thus, the water content range of the measurements should be large since the modeled data is sensitive enough also to small variations of the parameters. Then the parameters can be estimated much more correctly.
A large water content range corresponds with a larger infiltration flux. This means large precipitations should be included in the estimation interval. But here, the drawback is that not all soils can drain water as fast as a huge amount of rain can fall. Such ponding processes are represented in our model. Therefore, this process should be included in the simulation or the flux has to be reduced to a realistic value.
Moreover, the physical processes of the measurements should be describable with the mathematical model. For instance, our data set was taken from a profile with a bare surface. Thus,

the effect of plants and roots onto the time-series is negligible. If one wants to apply the estimation onto a profile with plants, the plants should be included into the model or at least the time interval should be chosen so that the plants are less active within the simulation interval. The same strategy holds for other processes like evaporation.
The spin-up interval also is important. It should be long enough to reduced the influence of the initial condition. This means, some rain events should occur to bring the modeled system into a realistic hydraulic state at the beginning of the spin-up interval. Further, a dryer interval should follow at the end of the spin-up interval. The water content is less sensitive due to variations of the hydraulic parameters in dryer soils. Then the water content is still less dependent on the total initial profile at the beginning of the estimation interval.

6.4.3. Hysteresis

In contrast to our model in real soils, hysteresis effects occur. This means in the context of hydraulic that increasing and decreasing of the water content are described by parameterization curves with different parameters. Further, this behavior can vary by every new excitation.
In our example in Figure 6.2, the measured water content follows parameterizations with different parameters during the wettening and draining times. Here, we have excluded two measurement points from the estimation. Thus, the increase of water content is not represented for the estimation and the hysteresis is less important. But still the same parameter set is used for modeling all times. Hence, the hydraulic behavior during increasing of the water content influences the simulated decreasing water content indirectly and with that, it has only a small effect on the parameter estimation.
Hysteresis occurs not only for large water content changes, but also in cases with small rise and decay of water content. For instance, around 27 volume percentage of θ, the water content varies due to change of the infiltration flux. In order to consider all these hysteresis effects, we would have to take another parameter set for the description of each increasing or decreasing part of water content.
But if the low changes of θ are around the measurement uncertainty, this hysteresis effect cannot be resolved anymore. Hence, an implementation of the hysteresis into the model would be only useful for the large variations of the water content. In our example, this would be restricted only to the increasing during the first day and the decreasing during the second day. However, the increasing part is almost excluded in Figure 6.2. Thus, the hysteresis can be neglected in this case since other aspects are more dominant like the heterogeneity of the soil.

6.5. Summary

In conclusion for the parameter estimation based on measured TDR-data, we found out that the estimated hydraulic parameters are reliable due to its values as well as due to our criterion of chapter 5. After the reduction of the rain flux, the saturated hydraulic conductivity was estimated with a larger value than the maximal infiltration flux. This made the simulation more realistic since previously, the model did not describe well the reality in the context of potential gradients and possible ponding. Further, this also effected the estimated porosity which is larger than the maximal measured water content during the year. The van Genuchten parameter n with its low value indicated that the soil is of loam or clay. Moreover, the van Genuchten parameterization was more reliable with the reduced infiltration flux since the model is not fully saturated during the estimation.
In addition, the estimation can be accurate in principle since the analysis of chapter 5 showed

that the infiltration flux - and also the reduced one - is strong enough to reproduce the parameters quite accurately. The estimated parameters are located also at the parameter maps in feasible regions for the estimation accuracy. Based on the $\phi/2$-criterion, the quality of the estimation can be plausible. It is within the transition zone.
The soil type can be loamy or clayey by value of n and by the value of the saturated conductivity with loam and sand content. Moreover, the slow decrease of water content during dry periods points more likely a finer soil. All in all, these results confirm the soil probes which showed sandy loam with a percent of humus within the upper soil layer.
Limitations of an accurate estimations will be given if the conditions of the model and the measurement do not coincide. For instance, if the amount of infiltration flux is described incorrectly, evaporating water or ponding is not taken into account. Further, the uncertainty of measurements can vary over time. The estimation model acts on the assumption that the measurement uncertainty is as low as possible due to the least-squares minimization.
The largest challenge for the projection is the heterogeneity in each layer as well as the layering. The projection of water content is strongly influenced by soil layers as we saw with the ponding water. A projection could be more successful with less heterogeneity or the knowledge of layer. Otherwise, we observed in our study that only for the lower depth, the correct slope can be reproduced. They never coincide for the greater depth measured time-series.
Further, we discussed the choice of the time interval. Here, we saw that a main criterion for choosing a suitable time interval is a large water content range. The water content range should be larger than a zone around half of the porosity. However, it should be checked if additional processes are dominant. If they are not describable with the current model, then the model should be extended or the respective data point should be excluded.
Another aspect was that hysteresis occurs in real soils in the context of the hydraulic behavior. We could neglect this effect for very small changes of water content since the measurement noise was stronger. However, the effect is stronger for the large increase and decrease of water content. But here, we excluded two measurement points during infiltration from the estimation. This reduced the hysteresis in our estimation example.

7. Summary

This work presents the projection of soil temperature and soil moisture into depth using near-surface time-series. It was distinguished between the linear process of the heat conduction and the non-linear process of the soil moisture. In each case, a coupling between temperature and water content is neglected. The thermal and the hydraulic parameters of soils are estimated by the Levenberg-Marquardt algorithm based on the projection methods as the forward model, respectively. In both studies, the conditions are simplified to study the applicability of a simple model for the projection. Thus, the soil is assumed to be uniform and the estimation parameters should not vary in time and space during one estimation time interval.

The projection of quantities plays a special role in soil physics since the projection yields information about greater depth without taking direct measurements. Many measurement techniques only observe the direct neighborhood around the measurement sensor. Others observe only the surface from the distance, like Earth observing systems. It is easier to scan larger areas with a certain resolution with these systems instead of digging a hole and installing sensors with a suitable spatial distribution.

Mathematical modeling is an efficient alternative to take direct measurements within soils and gives a guideline for the identification of the driving physical processes. The knowledge about the accuracy of the projection method is an important contribution to the modeling of soils at scales of a few centimeters till several kilometers.

Soil Temperature

Two analytical methods are applied to project soil temperature into depth. The transfer function method and the Fourier transform method are known methods which were applied by *Horton et al. (1983)*, *de Silans et al. (1996)* and *Roth et al. (2001)* in the past. Both methods are analyzed with respect to the improvements of the methods by the additional tools, the structure of the parameter space and the temporal evolution of the estimated parameter.
In both cases, the time-series of surface temperature is decomposed into a basis of functions. The transfer function approach uses δ-function whereas the Fourier transform method is based on complex exponential functions.
The chronological structure of the transfer integral makes the transfer function method less costly in computational time for shorter time-series. Moreover, in addition with the z-integral, this method incorporates a realistic initial condition and reduces the time of the warm-up of the thermal system a lot. An estimate of the spin-up phase is given and based on the transfer function itself, e. g. 50 % of the initial δ-pulse has reached 1 m depth after 12 d for an uniform soil with $D_h = 10^{-6}\,\mathrm{m}^2/\mathrm{s}$.
The Fourier based projection method is more powerful for long time-series. Due to their representation of the time-series in the frequency space, the long-term trend, e. g. annual cycle, is considered in a very efficient way. The implementation is very fast due to the use of the fast Fourier transform. Here, the initial profile of the model is determined from the surface tem-

perature. Damping of the amplitude and shifting of the phase brings the surface temperature into depth at initial time. This does not need to be a realistic description of the initial profile. Thus, a spin-up phase is used to reduce this error. Moreover, the application of the Tukey window decreases effects like leakage in the spectrum. Then, the estimation is applied on the middle part of the time-series due to the spin-up and the Tukey window.
For both methods, the projection and the estimation results are improved by the additional mathematical tools, z-integral or Tukey window, respectively. The finial observed residuum is down to 0.05 °C for the transfer function approach for an interval of 10 d and down to 0.5 °C for the Fourier method for an interval of 8 years. The main reason for the difference between the orders of the residuals is that the assumption of constant soil conditions does not hold for large time intervals. Several additional processes to the heat conduction happen during a year and their activity is not constant, e. g. change of water content, non-conductive processes due to infiltrating and evaporating water, plants activity, freezing and thawing of soil.
A scanning of the parameter space for D_h and T_{ref} roughly shows the same structure for both methods. In direction of the reference temperature, the minimum is unique since T_{ref} is a constant term of the projected temperature. In direction of D_h, the uniqueness of the minimum holds only for the restriction of the interval to a physically interesting one from $2 \times 10^{-8}\,\mathrm{m}^2/\mathrm{s}$ to $2 \times 10^{-5}\,\mathrm{m}^2/\mathrm{s}$. A further local minimum for very low diffusivities can vanish for the transfer approach by a refinement of the spatial grid in the z-integral. The slope change of the parameter space for very small diffusivities falls together with a dominant influence of the phase shift for the Fourier approach.
The thermal diffusivity is estimated over time with both methods, for the transfer method within 10 days intervals and for the Fourier approach within 6 months intervals. The estimated variations of D_h are low for the transfer function and make it hard to correlate changes of D_h with rain events since a combination of other non-conductive processes influences the quality of the estimation result.
In the context of the Fourier approach, the temporal evolution of the estimated D_h shows a cycle of 6 months with a larger variation than for the transfer approach. This cycle vanishes by decreasing of the estimation interval since the very low frequencies are not represented anymore in the spectrum. The cycle decreases since the frequency $1\,\mathrm{a}^{-1}$ is much more dominant than all other frequencies for large estimation intervals. The mean value of these estimated diffusivities is similar for both methods. But the mean value for the Fourier method is slightly lower. This can be caused by several effects, e. g. the residuals differ by a factor of 10 between both methods. This can be transferred back into the estimation.
The projection is applied across layer interfaces. This indicates a less sensitive behavior due to heterogeneities constant in time. Changes of water content cause the approximation error to increase. But the error can still be within the bounds of ±1 °C. Thus the method is applicable for coarser applications. Processes like freezing and thawing are a larger limitation of both projection methods since non-conductive processes become dominant.
The projection method should be chosen as the better one dependent on the application. The methods are different in the their computational costs. However, the more expensive transfer function method produces more accurate results. It has its advantages for the applications of the study of thermal processes at smaller scales of the soil. The Fourier approach is better for coarser applications and fast calculations, for instance, for modeling of thermal behavior within soil at large scales.

Soil Water Content

The projection of the soil water content is realized by numerically solving the Richards equation together with the Mualem-van Genuchten parameterization via finite volume methods. The main focus is on the estimation of the hydraulic parameters using only near-surface time-series. The question is addressed by comparing parameter estimation results under which conditions it is feasible to estimate accurate hydraulic parameters.
A first synthetic example is given for an infiltrating flux peak into an uniform medium. The parameter estimation results are quite different by only scaling the infiltration flux for the first to examples. A higher infiltration flux brings more water into the soil and lets the water content increase in a stronger way. The following sensitivity study shows that the parameters can be estimated more precisely for larger excitations of the water content. Further, a varying sensitivity behavior is observed for different soil types. All this influences the feasibility to accurately estimate the hydraulic parameters. Moreover, disturbing the parameter by 5 % leads to a projection error up to one volume percent of water content.
The feasibility to accurately estimate the hydraulic parameter is analyzed in detail at different distributions of the infiltration flux. Furthermore, for each of these scenarios three measurement techniques are represented in the model, i. e. the point sensor, TDR and GPR. They differ in the size of the average volume.
First, the question is synthetically analyzed how strong a maximal infiltration flux needs to be for a given soil so that the parameter estimation yields an accurate result. By scaling in several steps the infiltration flux, a transition zone for the maximal needed flux and for the water content range is determined. The transition of the water content range is observed around $\frac{\phi}{2}$ for a continuously increasing and decreasing flux scenario whereas the transition flux increases with rising of the averaged measurement volume. The transition of the water content range stays similar. Thus, the water content range can indicate whether the parameters are feasible to be estimated accurately or not.
Further studies with scaling of the flux illustrate that the dependence of the estimation result on the initial choice of the parameter becomes less by increasing the flux. This behavior is the same for all scenarios.
In a further synthetic study, the soil type is changed for given flux scenarios. A parameter scan over the true parameters obtains maps of the relative estimation error for each parameter. Based on the least-squares value, clear areas can be observed in the true parameter maps where the estimation of the parameters is not feasible. This correlates with a clear lower saturation range of 50 % of the saturated water content. Lower hydraulic conductivities supports a higher rise of the water content in a given time. An accurate estimation result is observable. The other parameters show more accurate estimation results under sandy soil conditions. The transition zone between soils which are feasible to be estimated accurately or not, varies with the intensity of the infiltration flux, as mentioned above. Further, the transition zone also changes in direction to feasible estimable soils if the average volume increases for the modeled sensor representation. There, the averaging damps the time-series of water content and make the averaged time-series less sensitive to the parameter. If the average volume is not exactly given, then the residuum has the largest error at the increasing slope of the time-series of the water content. Although the largest error of the residuum is around 0.5 volume percent, this has a large influence on the relative error of the estimated parameters around $5 - 40\,\%$.
For the averaging scenario of GPR, convergence problems occur for larger negative values of the parameter α since the valley becomes more narrow to the optimum in the parameter space. One possibility to improve this behavior would be to scale the problem in a better way since

the saturated conductivity can vary over much more orders of magnitude than the other parameters.
The application of the criterion of a feasible estimation shows for a bare soil of a test site close to Heidelberg that the characteristics of the measured data sets are typical for a soil within the transition zone. This means, an accurate parameter estimation from the given surface measurement can, but need not, be successful. The estimation results are not reliable in a first step due to the small porosity. The reason for this is based on the uncertainty between the infiltration flux and the water content. If the response of the flux does not match the measurements within the top soil, this mismatching results in an error of the estimated parameters.
The quality of the estimation is sensitive to the agreement of the model and the measurements. This also refers to the projection of the water content. The projected and the measured water content shows an offset between both time-series using the estimated parameters. Here, the simplified assumption of an uniform soil does not apply anymore since, at the layer interface, the water ponds and causes the measured water content to rise above in a catchment area around 20 cm. Thus, strong heterogeneities have to be considered in the projection model. Further, hysteresis effects also become important with repeating stronger excitations of water content changes.

All in all, the analysis of the projection methods shows that the simplified assumed models within porous medium leads to significant different applicabilities in the current state whereas the linear thermal model yields accurate projection results in the order of measurement uncertainties. The non-linear system needs further information, e. g. about heterogeneities, to project reliable water content into depth. However, the estimation of the hydraulic parameters from near-surface time-series reveals very accurate results under certain conditions. These conditions are specified for different infiltration scenarios, soil types and types of measurement instruments. This means that parameters can be estimated more accurately in regions with large rain events or regions with higher sand content combined with a finer textured soil.
Due to its linearity, the projection of soil temperature is applicable in context of small and large scale modeling. Thereby, the two thermal projection methods have a better performance dependent on their application. The non-linear hydraulic model is applicable for the estimation of effective soil parameters. Especially the assessment of the feasibility of estimation is important for a given data set and for decision whether a current data set needs to be extended for a successful estimation.

A. Grenzhof Test Site

The measurement data used in the application chapters 4 and 6 were recorded at the test site Grenzhof (49°25′ N, 8°37′ E) close to Heidelberg, Germany. This site is located in an agricultural and flat region. Since spring 2003 the surface consists of grass which is mowed in regular time steps.
Two main measurement locations on this field are logging data since 2004 and 2009, respectively. Some further information can be found in the work of *Wollschläger et al. (2009)* and *Schenk (2011)*. In this work we distinguish the two measurement locations by the type of the measurement quantity we use. In each of these profiles soil temperature and soil water content is measured, but we use the soil temperature only from location 1 and the soil water content from location 2.

A.1. Profile and Soil Temperature

The soil at the Grenzhof test site consists mainly of loam and sand. Figure A.1 shows a cross-section at the first location where the soil temperature is measured. On the left hand side a photo with the soil from the surface into around 164 cm depth and on the right hand side a structural sketch of the profile. At each side we observe five main layers with its interfaces at 28 cm, 82 cm, 144 cm and 154 cm depth.
The first layer contains humus due to plowing of the former agricultural soil. Below this there are two larger layers which both consist of sandy loam. They differ in their density, where the lower layer is more dense. Further the layers 4 and 5 contain a lot of gravel within a loamy or a sandy matrix. Within all these layers no groundwater is found. The groundwater level in this region is some meters beneath.
Moreover, in the sketch of Figure A.1 seven depths are marked. These denote the depth of the installed soil temperature sensors. Within the first layer the sensors are installed quite densely. The topmost sensor yields the time-series of measured temperature close to the surface which is our input data in chapter 4. Here one should notice that the third sensor is located close to the layer interface.
The spatial resolution of the sensors within the soil drops with depth, since in layer 2 and 3 are only two and one sensors, respectively. The fifth sensor in 67 cm depth is one of the most chosen examples in this work.
Down within the gravel our last sensor is positioned very deep in 164 cm depth. Surface signals, which reach this depth, have traveled through the heterogeneous medium.
These seven sensors measure soil temperature since spring 2004 with a temporal resolution of 10 min. The following averaging process reduces the resolution to 30 min by a running mean. The measured and averaged time-series are plotted over time and depth ($12 - 164$ cm) in Figure A.2. The horizontal lines indicate the sensor depths and between them the soil temperature is linearly interpolated. The temperature itself is color coded with the same colorbar shown below in each line. Each figure line shows two consecutive years from 2004 till 2011. The marks of each month simplify the classification of the day numbers t on the corresponding season of the year. The time is counted in days starting with the first day of year 2004 and increases

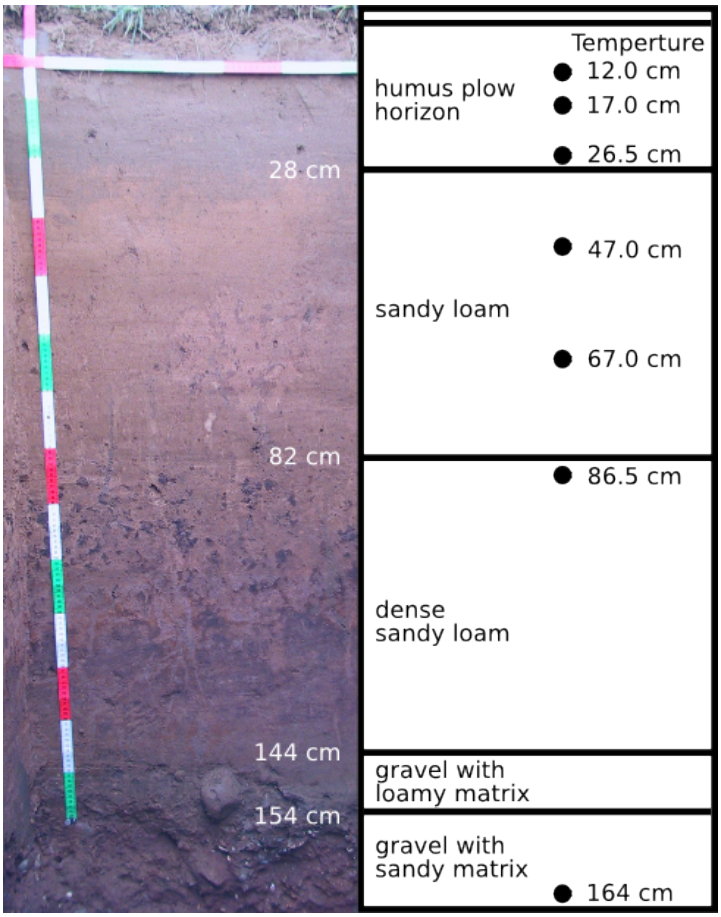


Figure A.1.: Left: Soil profile from Grenzhof test site. Right: Sketch of the soil profile with its installed temperature sensors in the certain depths. The soil type is characterized for each layer. The data is given by *Wollschläger et al. (2009)*.

steadily.
Every winter we can observe a magenta line. This line describes the isotherm of 0 °C. From this we can extract that the soil was frozen as longest and as deepest in the last eight years in year 2009. Also in year 2006 the soil was frozen over several weeks. The other isotherms are marked with black and gray lines with a resolution of 2 °C and 0.2 °C, respectively. The usual measurement uncertainty of temperature is around ±0.02 °C.
For this data set it is easy to see how soil temperature is forced by the weather over the surface. The soil temperature fluctuates over years and days. This effect is damped with deeper depth and below it is no thermal source which influences the soil temperature in this profile range.
Except the freezing time and the times with data gaps, the soil temperature data shows the behavior of a diffusion process from this global perspective. The measurement set is linearly interpolated at times with data gaps. Due to the linearity of diffusion equation the ansatz is reasonable at smaller time scales. In the summers 2004 and 2006 there are a larger number of outliers. The data set was filtered by a 3σ-standard deviation but at accumulation times of outliers there are still a few in the filtered data set. These together with a moving average lead to the observed data gaps in the year 2004.

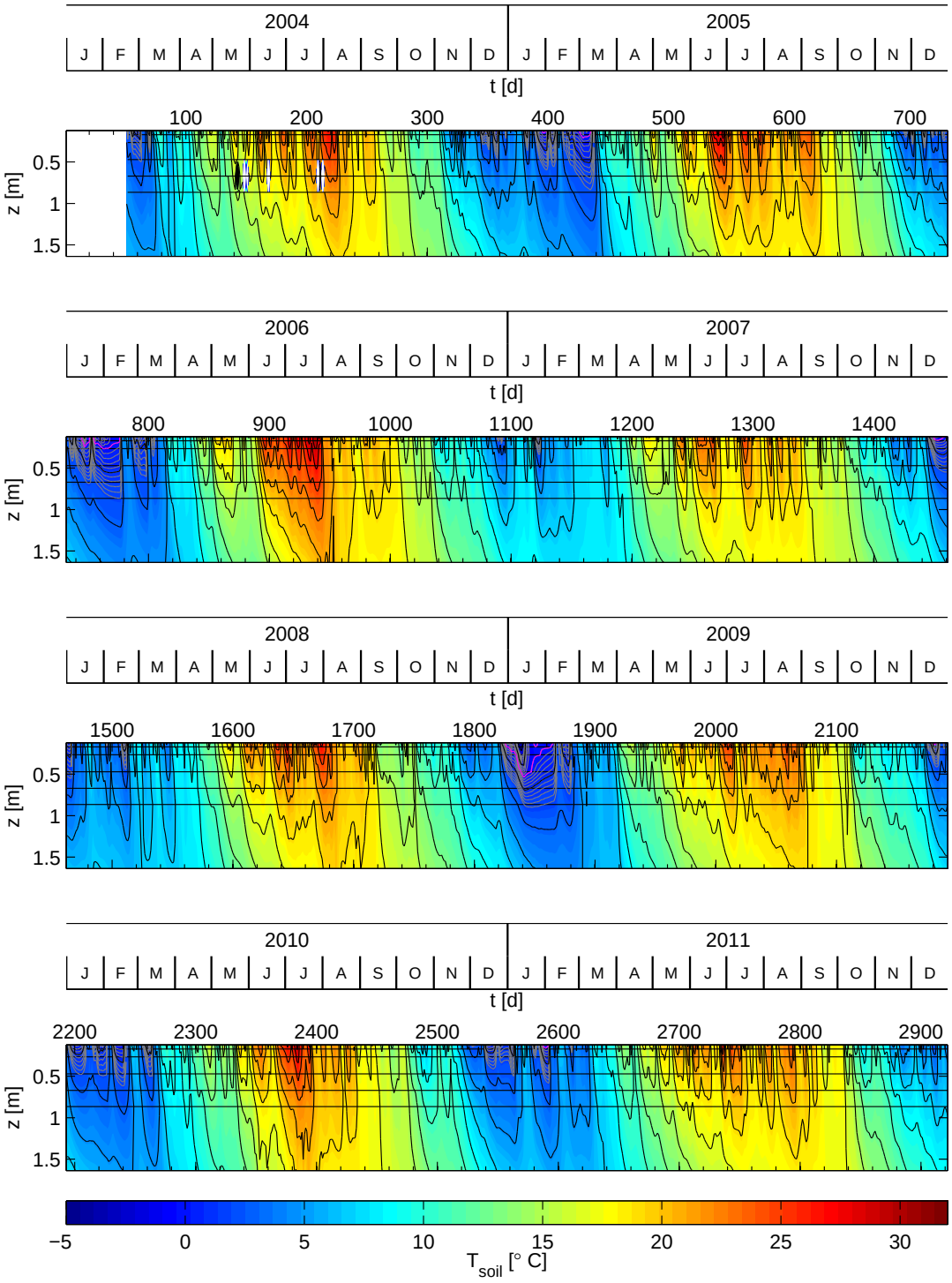


Figure A.2.: The soil temperature are measured over 8 years in several depths marked by the black horizontal lines. The isothermal lines in black and gray indicate constant temperature in steps of 2 °C and 0.5 °C, respectively, and the 0 °C-isotherm is plotted in magenta.

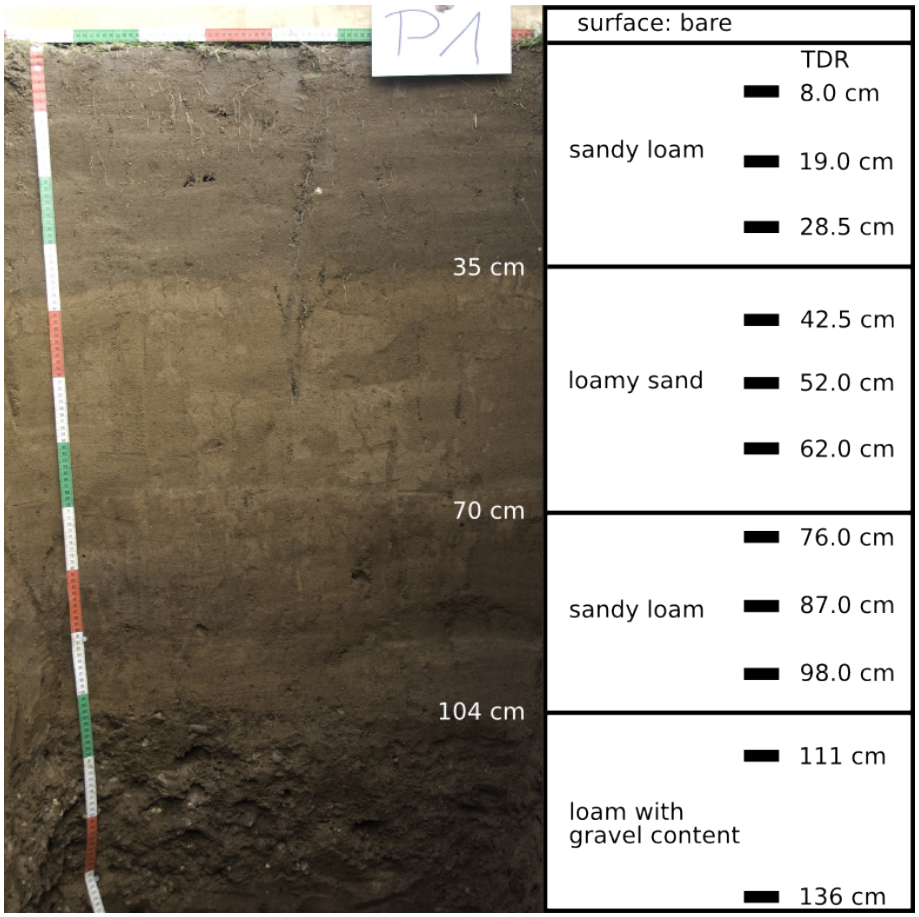


Figure A.3.: Left: Profile of the bare soil from the Grenzhof test site. Right: Sketch of the soil profile with its installed TDR-sensors in the certain depths. The soil type is characterized for each layer. Adapted from *Schenk (2011)*.

A.2. Profile and Soil Water Content

The second profile location is several meters beside the first one. From this profile we used the soil water content measurements for the application in chapter 6. In this profile the quantities are measured since spring 2009. The time-series of this profile are not as long as the ones with the temperatures. But here the applications to the soil water content do not rely strongly on very long time-series as in the temperature case with the Fourier approach. Further this profile has the advantage that the surface is bare. Since spring 2009 the grass at the surface is skimmed within an area of $5 \times 5\,\mathrm{m}^2$ to measure mainly the hydraulic behavior of the soil without the effect of plants and roots.

This profile is not as deep as the first one with a depth of around 140 cm, but the layering is similar. The first layer goes down to 35 cm at the right side of the profile, where the sensors are installed. This layer consists of loamy sand with humus due to the plowing. The second and the third layer are equally thick going to depths of 70 cm and 104 cm, respectively. There the loam and sand contents vary such that the second layer has mostly sand and the third one mostly loam. In the lowest layer the gravel content is very strong beside the loamy matrix.

The TDR-sensors are almost uniformly distributed in all layers. Each of the first three layers

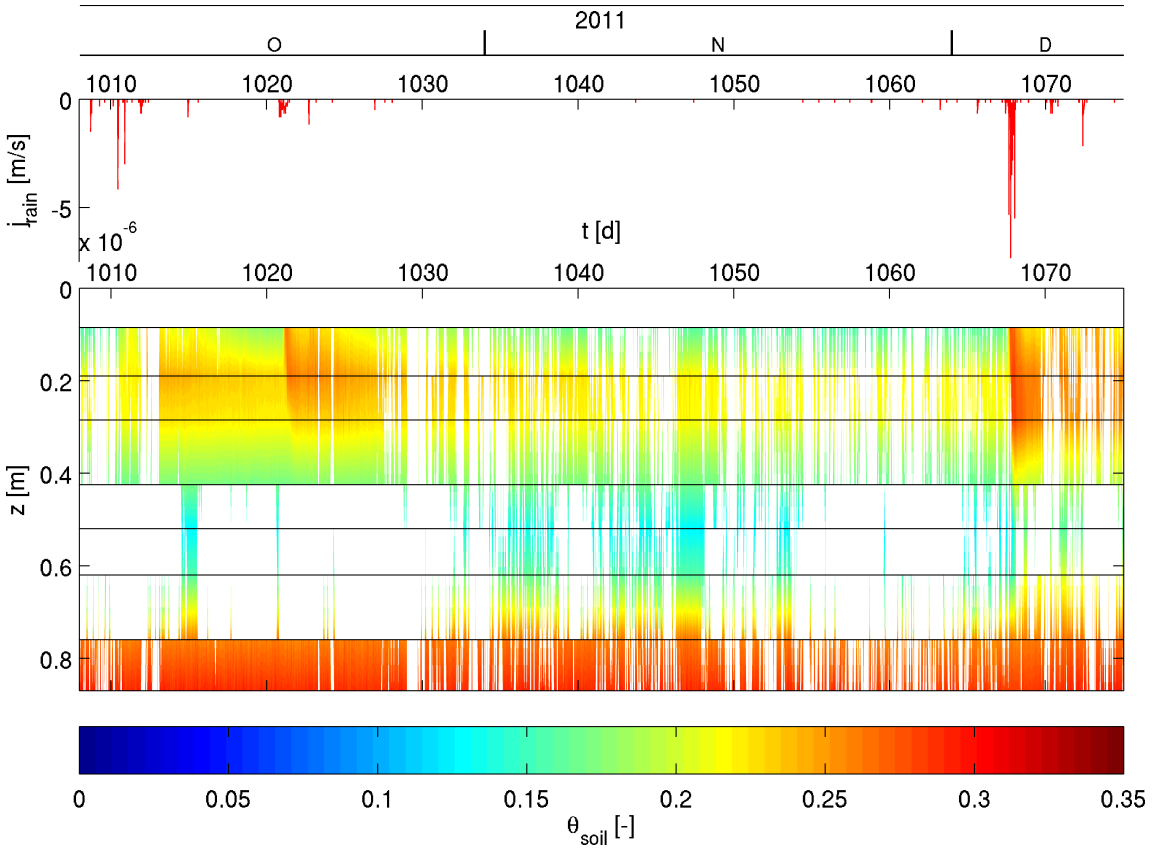


Figure A.4.: Top: Rain flux during the time interval of spin-up and estimation is used in chapter 6. The rain flux is calculated by the precipitation in millimeter per 10 min. Bottom: The soil water content is measured at the depth marked by the black lines. Between the line θ is linear interpolated in time and space. The white region denotes data gaps.

holds three TDR-sensor to measure the soil water content. The topmost sensor in 8 cm depth is taken for the parameter estimation in chapter 6. In the region of the test site the groundwater is several meters below this profile.

A cut-out of the whole time-series is drawn in Figure A.4. This is the measured soil water content profile used in chapter 6. The black lines denote the TDR-sensor depths and between them the water content is linearly interpolated. Some TDR-traces can not be well evaluated, therefore huge data gaps occur in each layer. Further information about the TDR-traces can be read in *Ludin (2010)*.

Including the spin-up phase the simulated time-series runs over 62 d. Starting at october 2011 with some infiltration events (top in red) a 40 d-drying period is taking place. At day 67 heavy rain falls occur and let the water content rise. This can be seen at the red coloring within the first layer during this time.

Around 50 cm we observe a lower soil water content. This structure is based on the layering of the soil and the different soil types comparing the higher and lower depth. The second layer consists mainly of sand which has the effect of lower water content within this layer.

B. Projection Results

B.1. Applying Transfer Function

In the following projection examples of the moving estimation interval are summarized refering to section 4.1.3 with the temporal evolution of $D_h(t)$. The projection results are shown here for some special marked time regions at 67 cm depth in Figure 4.10 a).

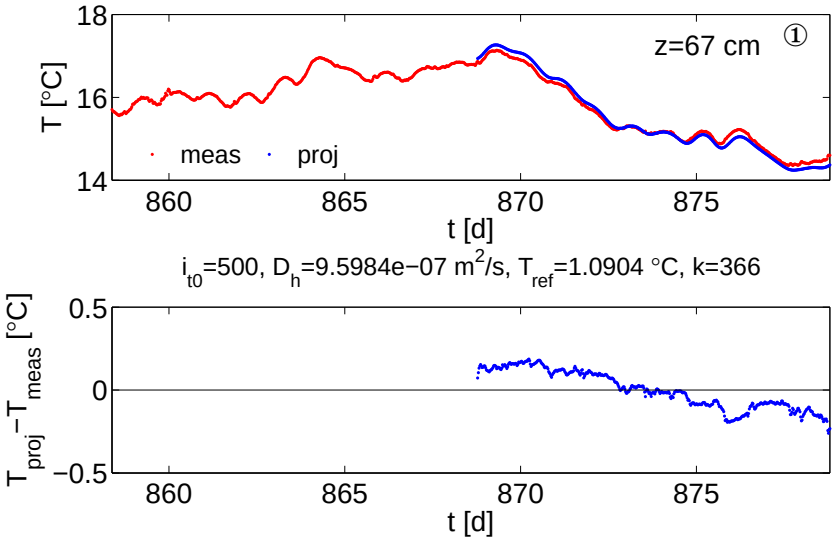


Figure B.1.: Projection of soil temperature around mean day 872.9 of the estimation interval. The time-series corresponds to time region marked by ① in Figure 4.10 a).

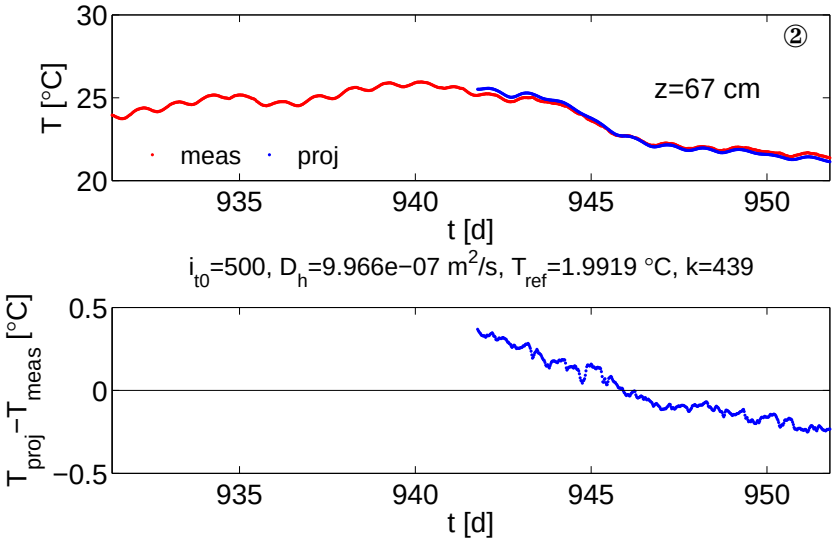


Figure B.2.: Projection of soil temperature around mean day 945.8 of the estimation interval. The time-series corresponds to time region marked by ② in Figure 4.10 a).

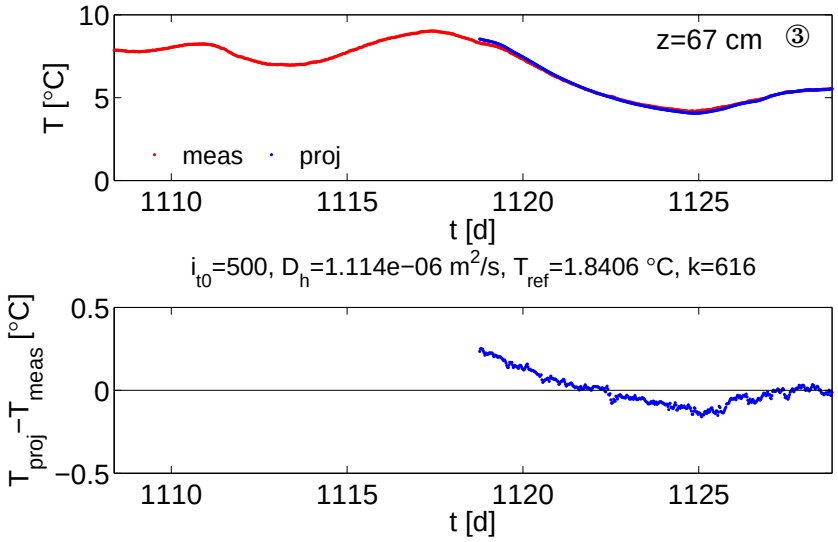


Figure B.3.: Projection of soil temperature around mean day 1122.6 of the estimation interval. The time-series corresponds to time region marked by ③ in Figure 4.10 a).

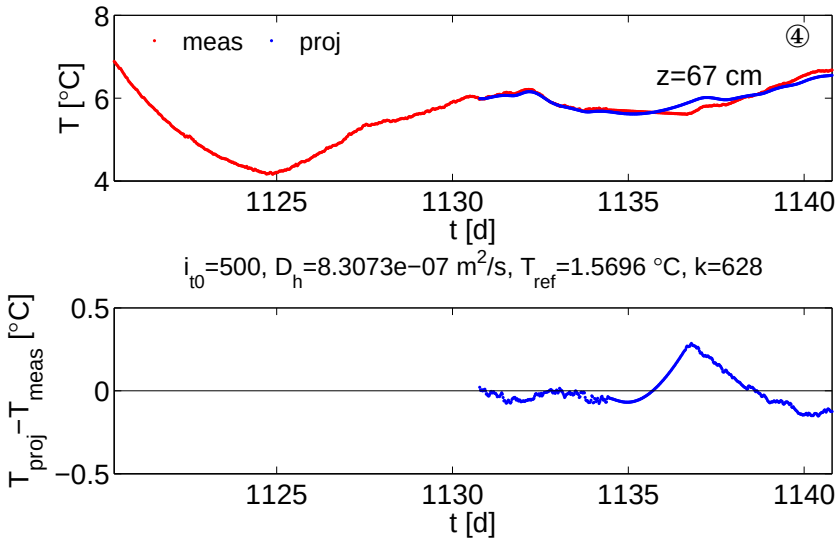
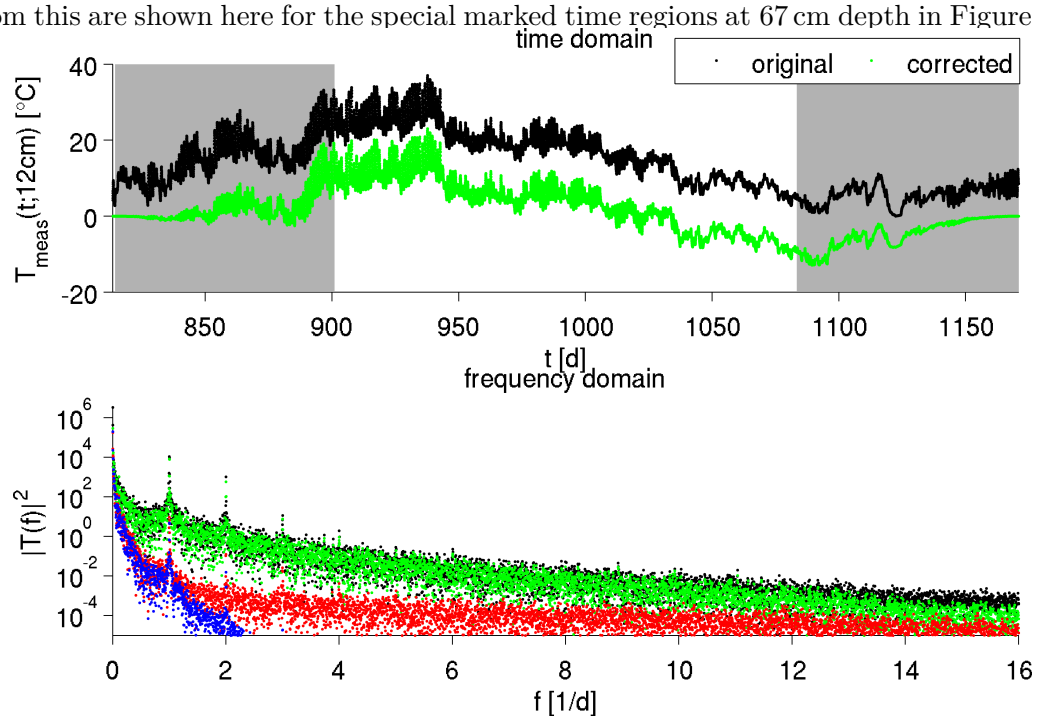


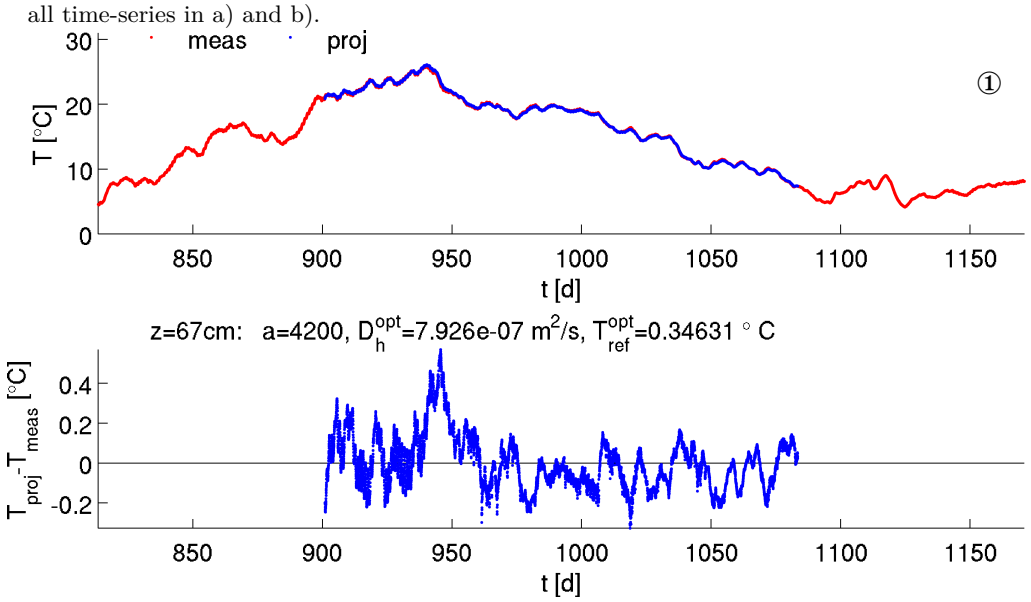
Figure B.4.: Projection of soil temperature around mean day 1134.5 of the estimation interval. The time-series corresponds to time region marked by ④ in Figure 4.10 a).

B.2. Applying Fourier Transform

The following projection examples refer to section 4.2.3. By a moving estimation interval the temporal evolution of $D_h(t)$ was realized with an estimation interval of six months. Due to the appling of the Tukey window on both sides $a = 4200$ points were added. A few projection results from this are shown here for the special marked time regions at 67 cm depth in Figure 4.18.

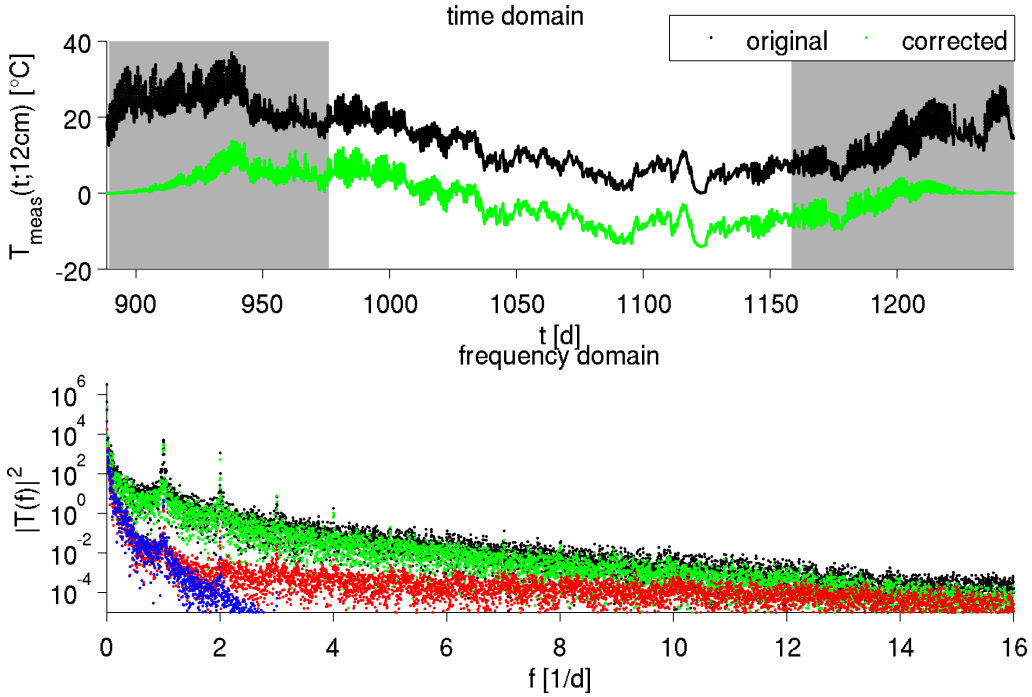


(a) Near-surface soil temperature (black) at 12 cm depth and the time-series after application of the Tukey window with $a = 4200$. Below there are summarized the amplitude spectra of all time-series in a) and b).

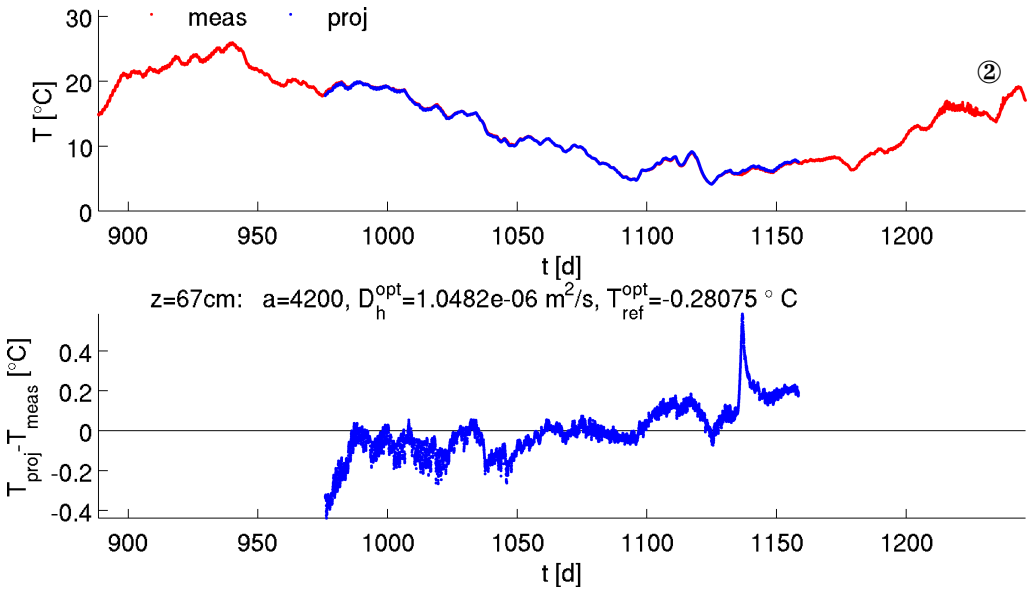


(b) Projected and measured soil temperature at 67 cm depth and its residuum below.

Figure B.5.: Projection of soil temperature around mean day 992.2 of the estimation interval. The time-series corresponds to time region marked by ① in Figure 4.18.

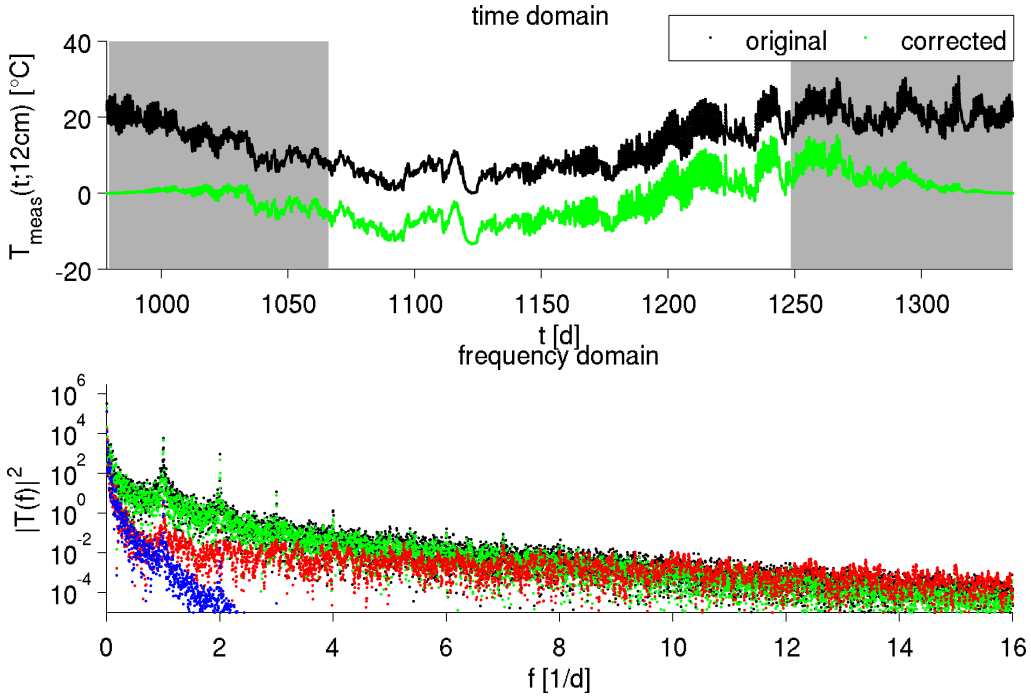


(a) Near-surface soil temperature (black) at 12 cm depth and the time-series after application of the Tukey window with $a = 4200$. Below there are summarized the amplitude spectra of all time-series in a) and b).

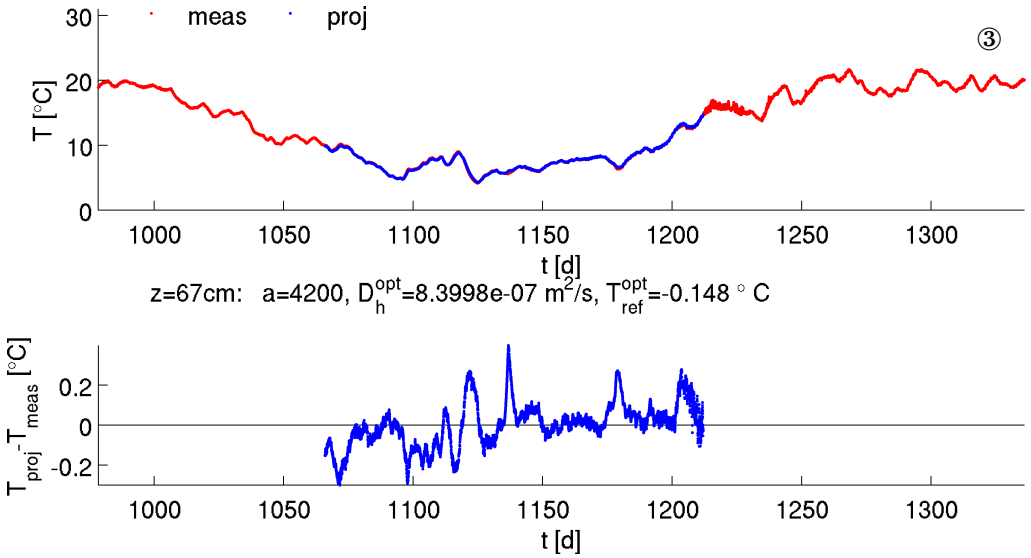


(b) Projected and measured soil temperature at 67 cm depth and its residuum below.

Figure B.6.: Projection of soil temperature around mean day 1067.2 of the estimation interval. The time-series corresponds to time region marked by ② in Figure 4.18.



(a) Near-surface soil temperature (black) at 12 cm depth and the time-series after application of the Tukey window with $a = 4200$. Below there are summarized the amplitude spectra of all time-series in a) and b).



(b) Projected and measured soil temperature at 67 cm depth and its residuum below.

Figure B.7.: Projection of soil temperature around mean day 1157.2 of the estimation interval. The time-series corresponds to time region marked by ③ in Figure 4.18.

C. Fourier Analysis

In literature the Fourier transform can be different defined by the factor before the integral and the minus sign given in forward or inverse transform. Up to this fact definitions of continuous and discrete Fourier transform can be looked up e. g. by *Press et al. (1988)* or *Strampp et al. (2003)*.

C.1. Continuous Fourier Transform

Definition *(Continuous Fourier Transform)*
Let the function f be absolutely integrable on $\mathbb{R}$. Then the mapping of $f(t)$ in $\mathbb{C}$ to $F(\omega)$ in $\mathbb{C}$ with

$$F(\omega) = \mathcal{F}(f)(\omega) = \int_{-\infty}^{\infty} f(t) \exp\left(-\mathrm{i}\omega t\right) \, \mathrm{d}t \quad \forall \omega \in \mathbb{R} \tag{C.1}$$

is called **continuous Fourier transform** of f. Its inverse is defined as

$$f(t) = \mathcal{F}^{-1}(F)(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \exp\left(+\mathrm{i}\omega t\right) \, \mathrm{d}\omega \quad \forall t \in \mathbb{R}\,. \tag{C.2}$$

C.2. Discrete Fourier Transform

Definition *(Discrete Fourier Transform)*
The mapping of a finite series of function values $\{f_1, \ldots, f_N\}$ to a finite series $\{F_1, \ldots, F_N\}$ with

$$F_k = \mathcal{DF}(f_n)(k) = \sum_{n=1}^{N} f_n \exp\left(-\mathrm{i}\frac{2\pi}{N}(k-1)(n-1)\right), \quad k = 1, \ldots, N \tag{C.3}$$

is called **discrete Fourier transform** of $\{f_n\}_n$. Its inverse is defined as

$$f_n = \mathcal{DF}^{-1}(F_k)(n) = \frac{1}{N} \sum_{k=1}^{N} F_k \exp\left(+\mathrm{i}\frac{2\pi}{N}(k-1)(n-1)\right), \quad n = 1, \ldots, N\,. \tag{C.4}$$

This definition is by analogy to MATLAB Documentation *(2008)*.

Bibliography

Abramowitz, M. (1965). *Handbook of mathematical functions*. Dover Publications, New York.

Anandakumar, K. et al. (2001). "Soil thermal properties at Kalpakkam in coastal south India". In: *Journal of Earth System Science* 110, pp. 239–245.

Beven, K. et al. (2001). "Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology". In: *Journal of Hydrology* 249.1-4, pp. 11–29.

Broadbridge, P. et al. (1988). "Constant Rate Rainfall Infilitration: A versatile Nonlinear Model 1. Analytic Solution". In: *Water Resources Research* 24, pp. 145–154.

Brooks, R. H. et al. (1966). "Properties of porous media affecting fluid flow". In: *J. Irrigation and Drainage Div., Proc. Am. Soc. Civil Eng.* 92 (IR2), pp. 61–88.

Carslaw, H. S. et al. (1959). "Conduction of heat in solids". In: *Oxford: Clarendon Press, 1959, 2nd ed.* 1.

Das, N. N. et al. (2006). "Root zone soil moisture assessment using remote sensing and vadose zone modeling". In: *Vadose Zone Journal* 5.1, p. 296.

De Silans, A. et al. (1996). "Apparent soil thermal diffusivity, a case study: HAPEX-Sahel experiment". In: *Agricultural and Forest Meteorology* 81.3-4, pp. 201–216.

Dunne, S. et al. (2006). "Land surface state and flux estimation using the ensemble Kalman smoother during the Southern Great Plains 1997 field experiment". In: *Water Resour. Res* 42.1, W01407.

Durner, W. et al. (2008). "Effective hydraulic properties of layered soils at the lysimeter scale determined by inverse modelling". In: *European Journal of Soil Science* 59.1, pp. 114–124. ISSN: 1365-2389.

Evensen, G. (1994). "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics". In: *Journal of Geophysical Research* 99, pp. 10–10.

Feddes, R. A. et al. (1988). "Modelling soil water dynamics in the unsaturated zone–State of the art". In: *Journal of Hydrology* 100.1-3, pp. 69–111.

Flühler, H. et al. (2004). *Physik der ungesättigten Zone*.

Galilei, Galileo. "II Saggiatore (The Assayer)". In: *Discoveries and Opinions of Galileo, writings translated by Stillman Drake (New York: Doubleday, 1957)*.

Gao, Z. et al. (2009). "Comparison of six algorithms to determine the soil thermal diffusivity at a site in the Loess Plateau of China". In: *HESSD* 6, pp. 2247–2274.

Glass, R. J. et al. (1989). "Mechanism for finger persistence in homogeneous, unsaturated, porous media: Theory and verification". In: *Soil Science* 148.1, p. 60.

Hansson, K. et al. (2004). "Water Flow and Heat Transport in Frozen Soil Numerical Solution and Freeze–Thaw Applications". In: *Vadose Zone Journal* 3.2, pp. 693–704.

Harris, F. J. (1978). "On the use of windows for harmonic analysis with the discrete Fourier transform". In: *Proceedings of the IEEE* 66.1, pp. 51 –83.

Hinkel, K. M. (1997). "Estimating seasonal values of thermal diffusivity in thawed and frozen soils using temperature time series". In: *Cold Regions Science and Technology* 26, pp. 1–15.

Horton, R. et al. (1983). "Evalutation of Mathods for Determining the Apparent Thermal Diffusivity of Soil Near the Surface". In: *Soil Sciene Society of America Journal* 47, pp. 25–32.

Huisman, J. A. et al. (2003). "Measuring Soil Water Content with Ground Penetrating Radar". In: *Vadose Zone J.* 2, pp. 476–491.

Huisman, J. A. et al. (2010). "Hydraulic properties of a model dike from coupled Bayesian and multi-criteria hydrogeophysical inversion". In: *Journal of Hydrology* 380.1-2, pp. 62–73.

Hund, F. (1972). *Geschichte der physikalischen Begriffe*. BI-Hochschultaschenbücher ; 543 ; BI-Hochschultaschenbücher 543. Spätere Aufl. in 2 Bd. ersch. Mannheim [u.a.]: Bibliogr. Inst., 410 S.

Ines, A. V. M. et al. (2009). "Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithms: 2. Using airborne remote sensing during SGP97 and SMEX02". In: *Water Resources Research* 45.

Ines, A.V.M. et al. (2008). "Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithm: 1. Conceptual modeling". In: *Water Resources Research* 44.6, W06422.

Ippisch, O. (2009). *Numerik von Transportprozessen in Porösen Medien*.

Ippisch, O. et al. (2006). "Validity Limits for the van Genuchten-Mualem Model and Implications for Parameter Estimation and Numerical Simulation." In: *Advances in Water Resources* 29.12, pp. 1780–1789.

Javierre, E. et al. (2006). "A comparison of numerical models for one-dimensional Stefan problems". In: *Journal of Computational and Applied Mathematics* 192, pp. 445–459.

Jones, Richard (1971). "Spectrum estimation with missing observations". In: *Annals of the Institute of Statistical Mathematics* 23 (1), pp. 387–398.

Kalman, R. E. et al. (1960). "A new approach to linear filtering and prediction problems". In: *Journal of basic Engineering* 82.1, pp. 35–45.

Kane, D. L. et al. (2001). "Non-conductive heat transfer associated with frozen soils". In: *Global and Planetary Change* 29, pp. 275–292.

Kim, S. et al. (2011). "Configuration of the relationship of soil moistures for vertical soil profiles on a steep hillslope using a vector time series model". In: *Journal of Hydrology*.

Knight, J. H. et al. (1974). "Exact Solution in nonlinear diffusion". In: *Journal of Engineering Mathematics* 8, pp. 219–227.

Krzeminska, Dominika M. et al. (2011). "High-resolution temperature observations to monitor soil thermal properties as a proxy for soil moisture condition in clay-shale landslide". In: *Hydrological Processes*. ISSN: 1099-1085.

Levenberg, K. (1944). "A method for the solution of certain problems in least squares". In: *Quarterly of applied mathematics* 2, pp. 164–168.

Loosvelt, L. et al. (2011). "Impact of soil hydraulic parameter uncertainty on soil moisture modeling". In: *Water Resources Research* 47.3, W03505.

Ludin, A. (2010). "Permafrost Temperature Analysis and Project in from Samoylov Island in the Lena River Delta, Siberia". diploma thesis. Heidelberg University.

Madsen, K. et al. (2004). *Methods for non-linear least squares problems*. 2nd ed. Technical University of Denmark.

Marchenko, S. et al. (2008). "Numerical Modeling of Spatial Permafrost Dynamics in Alaska". In: *In Proceedings of the Ninth International Conference on Permafrost* 2, pp. 1125–1130.

Marquardt, D. W. (1963). "An algorithm for least-squares estimation of nonlinear parameters". In: *Journal of the society for Industrial and Applied Mathematics* 11.2, pp. 431–441.

Massman, W. J. (1992). "Correcting errors associated with soil heat flux measurements and estimating soil thermal properties from soil temperature and heat flux plate data". In: *Agricultural and Forest Meteorology* 59, pp. 249–266.

MATLAB (2008). *version 7.7 (R2008b)*. The MathWorks Inc. URL: http://www.mathworks.com.

MATLAB Documentation (2008). The MathWorks, Inc.

More, J. (1978). "The Levenberg-Marquardt algorithm: implementation and theory". In: *Numerical analysis*, pp. 105–116.

Mualem, Y. (1976). "A New Model for Predicting the Hydraulic Conductivity of Unsaturated Porous Media". In: *Water Resources Research* 12.3, pp. 513–521.

Nicolsky, D. J. et al. (2009). "Estimation of soil thermal properties using in-situ temperature measurements in the active layer and permafrost". In: *Cold Regions Science and Technology* 55, pp. 120–129.

Pan, M. et al. (2006). "Data Assimilation for Estimating the Terrestrial Water Budget Using a Constrained Ensemble Kalman Filter". In: *Journal of Hydrometeorology* 7, 534-547.

Parzen, E. (1963). "On Spectral Analysis with Missing Observations and Amplitude Modulation". In: *Sankhya: The Indian Journal of Statistics, Series A (1961-2002)* 25.4, pp. 383–392.

Press, W.H. et al. (1988). *Numerical Recipes in C*. Cambridge University Press.

Prevedello, C. L. et al. (2008). "Analytic Solution of Boltzmann Transform for Horizontal Water Infiltration into Sand". In: *Vadose Zone Journal* 7, pp. 1170–1177.

Pullan, A. J. (1992). "Linearized Time-Dependent Infilitration From a Shallow Pond". In: *Water Resources Research* 28, pp. 1041–1046.

Reichle, R. H. (2008). "Data assimilation methods in the Earth sciences". In: *Advances in Water Resources* 31.11, pp. 1411–1418.

Rezanezhad, F. et al. (2006). "Experimental study of fingered flow through initially dry sand". In: *Hydrology and Earth System Sciences Discussions* 3.4, pp. 2595–2620.

Riseborough, D. et al. (2008). "Recent Advences in Permafrost Modelling". In: *Wiley InterScience* 19, pp. 137–156.

Robinson, D. A. et al. (2003). "A review of advances in dielectric and electrical conductivity measurement in soils using time domain reflectometry". In: *Vadose Zone Journal* 2.4, p. 444.

ROSETTA (2005). *ROSETTA Class Average Hydraulic Parameters*. Agricultural Research Service.

Roth, K. (2007). *Soil Physics Lecture Notes*.

Roth, K. et al. (2001). "Quantifying the thermal dynamics of a permafrost site near Ny-Ålesund, Svalbard". In: *Water Resources Research* 37.12, pp. 2901–2914.

Sabater, J. M. et al. (2007). "From near-surface to root-zone soil moisture using different assimilation techniques". In: *Journal of Hydrometeorology* 8.2, pp. 194–206. ISSN: 1525-7541.

Saito, H. et al. (2006). "Numerical analysis of coupled water, vapor, and heat transport in the vadose zone". In: *Vadose Zone Journal* 5.2, p. 784.

Schaap, M. G. et al. (2001). "ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions". In: *Journal of Hydrology* 251.3-4, pp. 163–176.

Scharnagl, B. et al. (2011). "Inverse modelling of in situ soil water dynamics: investigating the effect of different prior distributions of the soil hydraulic parameters". In: *Hydrology and Earth System Sciences* 15.10, pp. 3043–3059.

Schenk, G. U. (2011). "Soil-Atmosphere Coupling: Looking from within the Soil towards the Surface". PhD thesis. Ruperto-Carola University Heidelberg.

Schlittgen, R. et al. (1999). *Zeitreihenanalyse*. 8., unveränd. Aufl. München: Oldenbourg.

Schwen, A. et al. (2011). "Time-variable soil hydraulic properties in near-surface soil water simulations for different tillage methods". In: *Agricultural Water Management* 99.1, pp. 42 –50.

Seneviratne, S. I. et al. (2010). "Investigating soil moisture–climate interactions in a changing climate: A review". In: *Earth-Science Reviews* 99.3, pp. 125–161.

Simunek, J. et al. (1998). "Parameter estimation of unsaturated soil hydraulic properties from transient flow processes1". In: *Soil and Tillage Research* 47.1-2, pp. 27–36.

Steenpass, C. et al. (2010). "Estimating Soil Hydraulic Properties from Infrared Measurements of Soil Surface Temperatures and TDR Data". In: *Vadose Zone Journal* 9.4, p. 910.

Strampp, W. et al. (2003). *Mathematische Methoden der Signalverarbeitung*. Oldenbourg Wissenschaftsverlag.

Van Genuchten, M. T. (1980). "A closed-form equation for predicting the hydraulic conductivity of unsaturated soils". In: *Soil Sci. Soc. Am. J* 44.5, pp. 892–898.

Van der Velde, R. et al. (2010). "Influence of thermodynamic soil and vegetation parameterizations on the simulation of soil temperature states and surface fluxes by the Noah LSm over a Tibetan plateau site". In: *Hydrology and Earth System Sciences* 13.6, p. 759.

Vereecken, H. et al. (2008). "On the value of soil moisture measurements in vadose zone hydrology: A review". In: *Water Resources Research* 44, WW00D06.

Vogel, H. J. et al. (2010). "Non-equilibrium in soil hydraulic modelling". In: *Journal of Hydrology*.

Vrugt, J. A. et al. (2002). "Validity of first-order approximations to describe parameter uncertainty in soil hydrologic models". In: *Soil Science Society of America Journal* 66.6, pp. 1740–1751.

Vrugt, J.A. et al. (2008). "Inverse modeling of subsurface flow and transport properties: A review with new developments". In: *Vadose Zone Journal* 7.2, p. 843.

Walker, D. M. (2006). "Parameter estimation using Kalman filters with constraints". In: *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering* 16.4, p. 1067.

Warrick, A. W. et al. (1991). "An Analytical Solution to Richards' Equation for Time-Varying Infilitration". In: *Water Resources Research* 27, pp. 763–766.

Welch, G. et al. (1995). "An introduction to the Kalman filter". In: *University of North Carolina at Chapel Hill, Chapel Hill, NC* 7.1.

Wollschläger, U. et al. (2009). "Field scale effective hydraulic parameterisation obtained from TDR time series and inverse modelling". In: *Hydrology and Earth System Sciences Discussions* 6, pp. 1489–1522.

Acknowledgments

In this place, I would like to thank some special people for supporting me and making this work possible. 'Thank you!' to ...

- Prof. Peter Bastian who gave me the opportunity to do this work in mathematics. His explanatory notes motivated me for deeper reflexions.
- Prof. Kurt Roth who offered me an insight into soil physics and the application of modeling work in soil science. Thanks a lot for always being open to inspiring discussions and for the possibility to follow my own ideas.
- Olaf Ippisch who helped me through providing and supporting Muphi and Fitphi in the context of hydraulic modeling.
- Anne and Gaby for the great teamwork and for proofreading parts of this thesis. Further thanks to Andreas, Stefan, Jens W. and Christiane Haase who improved this work by valuable comments in language and thoughts.
- All office colleagues and friends – Stephan, Sabrina, Thomas, Martin, Maike and Stefan – for the friendly atmosphere, their open words and their post cards to beautify our office.
- All Bophys – Ute, Steffen, Gaby, Anne, Pan, Patrick, Andreas, Jens B., Cornelius, Stefan, Eva, Jens W., Waldemar, Felix, Klaus, Holger ... – for the discussion about science as well as for the conversations during coffee breaks, making the work at the institute more pleasant.
- My sister Astrid, her candidly words at the right time as well as Martin for their great friendship.
- My parents for their open ear which let me find my way of life and through this work independently. A further thank for encouraging words along with Sister Paulina.
- And last but not least, my boyfriend Stephan, for being at my side and supporting me on my way during the last years. Thanks for being my compass to find tracks like this ...