ERFASSUNG UND ANALYSE VON TRANSPORTPROZESSEN IN BÖDEN MIT ERT

DIPLOMARBEIT

vorgelegt von

Tobias Pfaff

Heidelberg,	den 14.10.2007
1. Gutachter:	Prof. Dr. Peter Nielaba
2. Gutachter:	Prof. Dr. Wolfgang Dieterich

UNIVERSITÄT KONSTANZ

Contents

Zusammenfassung

Hydrologische Fragestellungen erhalten in Zeiten steigender Wasserknappheit und Verschmutzung von Böden und Grundwasser eine zunehmende Bedeutung. Methoden zur großräumigen Bestimmung des Wassergehalts von Böden sowie zur Untersuchung der Ausbreitung von Schadstoffen und Düngemitteln sind längst nicht mehr nur für Anwendungen im Agrarbereich von Interesse, sondern besitzen zunehmend Relevanz für die Sicherstellung der Trinkwasserversorgung. Auch für die Ankopplung von Klimamodellen ist ein Verständnis von Wassertransport in Böden von großem Interesse.

Die genannten Anwendungen setzten die Untersuchung hydrologischer Prozesse auf Feldskala, d.h. von mehreren Metern bis vielen Kilometern, voraus. Die klassischen Messmethoden der Hydrologie, die auf Probennahme basieren, erreichen allerdings nur eine sehr begrenzte räumliche und zeitliche Auflösung. Außerdem kann durch die Probennahme das natürliche Flußsystem gestört werden. Geophysikalische Messungen sind hingegen größtenteils nicht invasiv und zudem in der Lage, räumlich aufgelöste Information über den Boden zu liefern. Der Einsatz geophysikalischer Methoden zur Bearbeitung hydrologischer Fragestellungen ist daher vielversprechend. Aus einem Zusammenspiel von Geophysik, Hydrologie und Bodenphysik ist daher das noch junge Feld der Hydrogeophysik entstanden, in dem auch diese Arbeit angesiedelt ist.

Bei der Behandlung größerer hydrologischer Probleme ist in der Regel eine Kombination aus Modellierung und Messung nötig. Für die in dieser Arbeit betrachteten Stofftransportprozesse wird beispielsweise ein Modell aus der Bodenphysik für den lokalen Wasser- sowie den Stofftransport angewandt. Die Herausforderung liegt hierbei neben der Auswahl geeigneter Modelle in deren Parametrisierung. Letztere soll durch eine möglichst geringe Anzahl geophysikalischer Messungen realisierbar sein. Da sich viele Modellparameter nicht direkt messen lassen, erfordert die Formulierung der Schnittstelle zwischen den Messdaten einerseits und den Modellparametern andererseits besondere Aufmerksamkeit. Diese Arbeit beschäftigt sich unter anderem mit diesen Schnittstellen und zielt auf eine bessere Ankopplung diverser geophysikalischer Methoden untereinander sowie auch an hydrologische Modelle. Dabei wird insbesondere die geophysikalische Methode der Elektrischen Widerstandstomographie (ERT, aus dem Englischen „Electrical Resistivity Tomography") betrachtet.

ERT ist eine Methode zur Bestimmung der Verteilung des elektrischen Widerstands im Boden durch Injektion von Gleichströmen. Aus der Widerstandsverteilung können Information

über geologische Schichtungen, gelöste Stoffe im Boden und den Wassergehalt abgeleitet werden. Schon im Jahre 1920 wurde diese Methode von Schlumberger für eindimensionale geologische Modelle eingesetzt. Heute findet ERT in sehr unterschiedlichen Bereichen Anwendung: Vom Aufspüren vergrabener Bauwerke in der Archäologie (*Noel and Walker* [1991]) über die Untersuchung von Mülldeponien (*Chambers et al.* [2006]) bis hin zu dem für diese Arbeit relevanten Fall der Beobachtung von Transportprozessen in der Hydrologie und Bodenphysik (*Benson et al.* [1997], *Michot et al.* [2003]).

ERT-Systeme bestehen aus einer Anordnung von Elektroden, über die Strom in den Boden gespeist und das entstehende Potentialfeld gemessen werden kann. Üblicherweise wird für jeden Datenpunkt ein Strom zwischen zwei Elektroden angelegt, und die Potentialdifferenz zwischen zwei anderen Elektroden gemessen. Der so entstehende Datensatz kann allerdings nicht direkt in ein räumliches Widerstandsmodell transformiert werden. Das inverse Problem, das durch die Simulation von Messdaten unter Vorgabe eines Widerstandmodells gegeben ist, ist hingegen durch Lösung der DC-elektrischen Differentialgleichung behandelbar. Daher kann durch eine iterative numerische Inversion auch das ursprüngliche Problem gelöst werden. Allerdings sind ERT-Probleme in der Regel zumindest teilweise unterbestimmt, so dass Uneindeutigkeiten entstehen. Diese müssen durch die Definition von Nebenbedingungen ausgeräumt werden.

Im Rahmen dieser Arbeit wurde zunächst ein Differentialgleichungslöser für die DC-elektrische Gleichung implementiert. Die Modellierung arbeitet auf Basis von Finiten Differenzen. Die Korrektheit und Qualität der berechneten Potentialfelder wurde durch Vergleiche mit analytischen Lösungen getestet. Auch eine Funktion zur numerischen Berechnung von Sensitivitäten wurde implementiert.
Für die Inversion wurde das Softwarepaket „Dc2dInvRes" von T. Günther verwendet, dessen Programmtext im Rahmen einer Kooperation eingesehen werden konnte. Dadurch war die Einbindung in ein Framework zum Datenaustausch mit dem Differentialgleichungslöser und den anderen in dieser Arbeit entwickelten Programmen möglich.

Eine Methode zum Stabilisieren der ERT-Inversion besteht darin, Messdaten anderer geophysikalischen Verfahren über die Nebenbedingungen in die Inversion einfließen zu lassen. Diese Methode wurde von *Günther et al.* [2006] für die Einbindung seismischer Daten angewandt. In dieser Arbeit wurde sie für Ground Penetrating Radar (GPR) und Time Domain Reflectometry (TDR) adaptiert und implementiert.

Desweiteren wurde mit einem einfachen Modell ein Indikator zur Bewertung von numerischen Fehlern durch die finite Elektrodeneindringtiefe entwickelt.

Ein weiterer Teil dieser Arbeit bestand in der Kombination von ERT mit der Modellierung von Transportprozessen. Zu diesem Zweck wurde eine Schnittstelle der bestehenden Programme zu der freien Simulations-Software SWMS-2D von Simunek erstellt. SWMS-2D enthält einen Richards-Löser zur Simulation des Wassertransports, und einen CDE-basierten

Löser für Stofftransport. Ein neues Konzept zur inversen Bestimmung von Transportparametern mithilfe mehrerer geophysikalischer Verfahren wurde entwickelt. Dabei wurde die Inversion von Transportprozessen mit der ERT-Inversion zu einer integrierten Inversion gekoppelt. Für den Inversionsschritt wurde die Methode der konjugierten Gradienten verwendet. Synthetische Tests zur Parameterschätzung konnten erfolgreich durchgeführt werden.

Um die entwickelten Methoden anhand von Messdaten testen und anwenden zu können, wurde ein Tracer-Feldexperiment durchgeführt. Als Testfeld wurde eine Grasfläche in landwirtschaftlich genutzter Umgebung ausgewählt. Meteorologische Daten und Temperaturprofile konnten von einer ebenfalls auf dem Testfeld vorhandenen Wetterstation gewonnen werden. Auf diesem Testfeld wurden ein Profil mit TDR-Sonden zur Wassergehalts- und Leitfähigkeitsmessung sowie Elektroden für 2D-Oberflächen-ERT-Messungen installiert. Die ERT-Messkonfiguration wurde so optimiert, dass im Tracerbereich eine möglichst hohe Auflösung erzielt werden konnte. In diesem Kontext wurde ein einfacher Algorithmus für die Optimierung der Messelektrodenkonfiguration entwickelt. Für die untersuchten Fälle reichen die Ergebnisse bei signifikanter Rechenzeitverkürzung an den Algorithmus von *Stummer et al.* [2004] heran, so dass die Anwendung auch für unser größere Elektrodensysteme, wie das im Experiment zum Einsatz gekommene 96-Elektrodensystem möglich wurde. Desweiteren wurde die Multikanalausnutzung des ERT-Messinstrument SyscalPro von IRIS Instruments, das für die Feldmessungen verwendet wurde, optimiert. Das Gerät ermöglichst zwar Mehrkanalmessungen, diese werden von der proprietären Software aber nur mangelhaft unterstützt. Daher wurde auf Basis der Graphentheorie ein Algorithmus entworfen, der die Messkonfigurationen derart umordnet, dass eine optimierte Multikanalausnutzung erreicht werden kann. Zum Datentransfer mit dem Gerät wurde das Übertragungsprotokoll entschlüsselt und ein neues Übertragungsprogramm verfasst.

Nach Aufbringung des Tracers ($CaCl_2$) im Februar 2007 wurde dessen natürliche Ausbreitung mittels Aufnahme von ERT-, GPR- und TDR-Zeitreihen beobachtet. Kombinierte Messdaten liegen vom Zeitpunkt der Aufbringung bis Anfang Juni 2007 vor.
In der Einzelauswertung der Daten ist mit Ausnahme eines vermutlichen ERT-Inversionsartefakts eine befriedigende Übereinstimmung von ERT und TDR zu beobachten,. Die Stabilisierung von ERT-Daten durch TDR-Daten war in der ersten Zeithälfte des Versuchs erfolgreich, in der zweiten Zeithälfte konnte auf diese Weise jedoch keine Stabilisierung erreicht werden.
Eine Auswertung der GPR-Daten war wegen des lehmhaltigen Bodens auf dem Testfeld nur in bestimmten Bereichen möglich. In diesen Bereichen zeigte sich eine sehr gute Übereinstimmung der Einzelauswertung mit den ERT-Daten. Auch die Stabilisierung der ERT-Daten durch GPR-Daten zeigte in diesen Bereichen sehr gute Ergebnisse.

Schließlich wurden mittels der entwickelten integrierten Inversion eine Bestimmung der Transportparameter vorgenommen. Die auftretenden Abweichungen lassen sich weitgehend auf fehlende Simulation von Evaporation und Transpiration durch Pflanzen zurückführen. Eine Erweiterung des Transportmodells um diese Prozesse sollte zu einer weiteren Verbesserung der Ergebnisse beitragen.

Zusammenfassend wurden in dieser Arbeit diverse aufeinander abgestimmte Algorithmen und Dienstprogrammen entwickelt, um ERT-Daten als einzelne Methode und in Kombination mit anderen geophysikalischen Methoden und hydrologischen Modellen auszuwerten. Die entwickelten Methoden wurden anhand eines Tracer-Feldexperiments angewandt und getestet.

Chapter 1

Introduction

The topic of this thesis is situated in the domain of hydrogeophysics. Hydrogeophysics is still a very young field, and emerged from the fields soil physics, hydrology and geophysics. It addresses questions in hydrology, especially water resources, with the methods of soil physics and geophysics. Today, these questions are becoming more and more important, as pollution and fresh water shortage increasingly affect our society – therefore today, research in this areas is considered as vital.

The questions posed are therefore very concrete ones, for example :

- *How can we reliably assess the water distribution in an area, and predict its future development ?*
- *How fast and far do harmful substances spread, when leaking into the ground ?*
- *How can the intrusion of salt- or brackish water be monitored ?*

To answer these questions, two topics have to be addressed. Firstly, measurements and monitoring of water and solute content are necessary. While investigation methods on a small scale or in the lab are well developed, one often has to deal with huge areas in order to approach the problems. Measurements have to be done on a large scale, and preferably non-invasive and cheap. This is not possible using traditional techniques.

Secondly, in order to predict future trends, simulations are necessary. For the problems addressed above, these simulations have to describe transport of water and solutes. The question is therefore, which models to use, how to parametrize them and especially, how the model parameters can be determined.

Actually, both topics are interconnected. To derive model parameters, a set of specialized measurements is necessary, and since it is often not possible to directly measure the complete area of interest, one has to resort to modeling.

While hydrology provides the broad setting and soil physics describes the dynamics, geophysics comes into play when measurements are concerned. The traditional method of hydrological investigations is based on sampling, which cannot provide the spatial and temporal resolution desired. In addition, the soil system is disturbed by sampling. Geophysical methods on the other hand are mostly non-invasive and can map larger areas. Therefore, its application in hydrology stands to reason.

In my thesis, I will study the geophysical method of Electrical Resistivity Tomography (ERT), and its application to hydrogeophysics. In particular, the interface between ERT, other geophysical methods and hydrology will be examined.

Introduction to ERT DC Geoelectric measurements are a non-invasive method for determining the subsurface electrical resistivity distribution. The functional principle is generally as follows : A DC current is injected through current electrodes. In the emergent potential field in the ground, voltages are measured by reception electrodes. This is repeated for different electrode arrangements, each providing a data point. Determining a resistivity model from this data set is an inverse problem, with the DC-electric equation as a forward operator. The first geoelectrical measurements were conducted by Schlumberger in 1920. He used an 1D-layered resistivy model (resistivity sounding), to which the DC equation could be solved analytically. The measured data was then compared to a band of analytic curves to obtain model parameters for each layer. Modern geoelectrical surveys use high resolution 2D or 3D resistivity models, in combination with large, automatically acquired data sets. Here, numerical forward solvers are used, and the inverse problem is also solved automatically. This methods are called Electrical Resistivity Tomography (ERT). ERT is a frequently used method in geophysics, and can be used for mapping soil and rock layers, inspecting landfill sites (*Chambers et al.* [2006]), locating contaminated ground (*Vogelsang* [1997]) or detecting buried structures in archeology (*Noel and Walker* [1991]). Recently, it is also becoming a popular method in soil science and hydrology, where it can be used for transient monitoring of soil water content(*Zhou et al.* [2001], *Michot et al.* [2003]) or contaminant and tracer propagation (*Benson et al.* [1997]).

ERT and hydrogeophysics In hydrogeophysics, one of the most interesting questions is how to automate and integrate geophysical data processing with water and solute transport models. In recent years, significant progress has been made in relating ERT measurements of solute tracers with transport simulations. *Singha and Gorelick* [2006] monitored tracer propagation between wells with a controlled water flux to test applicability of simple petrophysical models. A. Kemna conducted a large-scale ERT tracer-study with borehole ERT (*Kemna et al.* [2002]). Parameters for a hydraulic model could be derived from breakthrough curves. Further research in the same setting was performed by (*Vanderorght et al.* [2005]), who presented a method to derive transport parameters by comparing ERT inversion results with hydraulic models.

On the other hand, joint evaluation of ERT and other geophysical methods as Ground Penetrating Radar (GPR) and Time Domain Reflectometry (TDR) are increasingly studied. *Günther et al.* [2006] presented a joint inversion of ERT and seismic data, which stabilizes the ERT inversion and produces consistent subsurface structure information. *Linde et al.* [2006] was able to combine cross-borehole ERT and GPR into a joint inversion.

This thesis aims at the direct integration of ERT with water and solute transport. Optimally, such a model would not include only ERT, but also other geophysical data – in a first step, it is aimed at TDR data.

Structure In the first chapter, the theory of ERT data evaluation is discussed. Also, a short summary of the other methods used as well as an introduction to soil physics is given.
In chapter 2 the implementation of the ERT algorithms developed for this thesis are explained. Also, concepts on integrating geophysical and hydrological data processing are transferred to ERT, and a new concept is presented.
Finally, chapter 3 presents the tracer experiment conducted during this thesis. The methods presented before are applied and the results discussed.

Chapter 2

Geoelectric theory

In this section, the theory of Electrical Resistivity Tomography and other theory which is used in this thesis is introduced and briefly explained.

2.1 DC Geoelectrics

Measuring data Today, ERT measurements are usually performed using a fixed, large set of electrodes (50 – 200). In this thesis, the focus lies on surface ERT, which means these electrodes are inserted in the surface, forming a line for 2D measurements, and a grid for 3D measurements[1].

Mostly four-electrode measurements are performed. Current is injected via a pair of current electrode, denoted by A and B, and the emergent potential field is measured differentially at two potential electrodes M and N. This method has the advantage that contact resistance of the electrodes to the ground are less relevant. For theoretical considerations,

---

[1] The other popular type, borehole ERT, uses electrodes buried at several depths in boreholes. The mathematical and numerical treatment is almost identical, with the addition of a source term mirrored at the surface

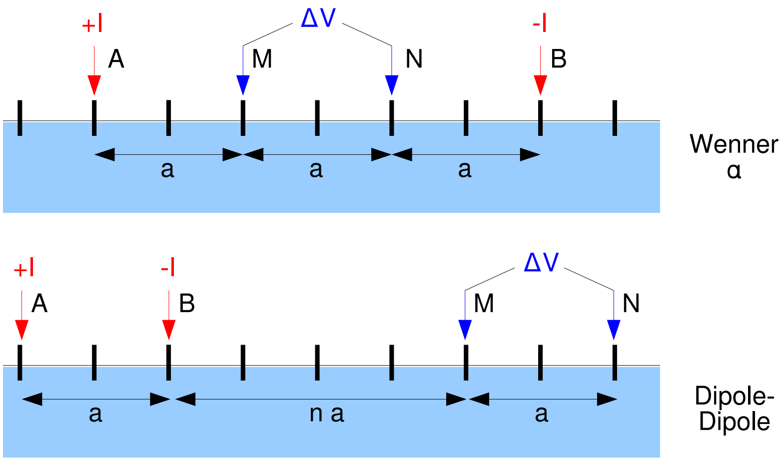


Figure 2.1: Wenner-α array (a), and Dipole-Dipole array (b) schematics

the simpler so-called Pole-Pole measurements are usually considered, with one current and one potential electrode, referencing to the zero ground.[2]. Due to linearity, results for Pole-Pole arrays can be easily transformed to four-electrode measurements by superpositioning. Therefore in this chapter, mainly pole-pole configuration is assumed.

While traditionally, two electrodes for current injection and two electrodes for voltage measurement were placed and connected manually for each data point, on modern systems all electrodes are connected to a switching unit, which automatically assign current and potential electrodes each step. With a large number of electrodes, not all combinations of current and potential electrodes are measured, but a subset called "measurement array". Typical arrays for 2D line measurements include Wenner-α and Dipole-Dipole, which are depicted in fig. 2.1. An array is divided into data levels, which share a common set of parameters (a,n), but feature shifted starting points.

Apparent resistivity Often, the measured data is given not as potential difference, but as apparent resistivity

$$\rho_a = k\frac{\Delta V}{I} \qquad .$$

The configuration factor k is defined so that $\rho = \rho_a$ for a homogeneous halfspace resistivity model (2.7), so that

$$k = \frac{2\pi}{\frac{1}{|\mathbf{r_A}-\mathbf{r_M}|} - \frac{1}{|\mathbf{r_A}-\mathbf{r_N}|} - \frac{1}{|\mathbf{r_B}-\mathbf{r_M}|} + \frac{1}{|\mathbf{r_B}-\mathbf{r_N}|}}$$

for a surface electrodes setting without surface topology. Apparent resistivity can be used to get a first impression of the resistivity distribution to be expected, but is no suitable base for any qualitative analysis.

Data evaluation The transformation from the measured data set to the real subsurface resistivity model is the problematic step in ERT measurements. While the reverse direction can be described easily by the DC differential equation ("forward operator"), an easy mapping from data to model is not possible. Adding to the problem, this problem is partly under- and partly overdetermined, in an irregular fashion determined by the chosen measurement array and resistivity model discretization. Solutions therefore use the inverse approach, in which one or multiple resistivity models are estimated, the forward operator is applied, and the generated data ("forward response") is compared to the measured data set. Then, the models are updated and the procedure is repeated, until a satisfying model is found.

In the following chapters, these steps are described in detail.

---

[2] While Pole-Pole measurements can be experimentally approximated by fixing one current and potential electrode at the far limit, this method tends to provide low quality data.

2.2 Underlying model

For a current density $\boldsymbol{j}$,charge density ρ and the E/M-fields (E, D, B, H), Maxwell's equations state

$$\begin{aligned}\boldsymbol{\nabla}\mathbf{D} &= \rho\\ \boldsymbol{\nabla}\mathbf{B} &= 0\\ \boldsymbol{\nabla}\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial t}\\ \boldsymbol{\nabla}\times\mathbf{H} &= \mathbf{j}+\frac{\partial\mathbf{D}}{\partial t}\end{aligned}$$

We can then deduce

$$0 = \boldsymbol{\nabla}(\boldsymbol{\nabla}\times\mathbf{H}) = \boldsymbol{\nabla}\mathbf{j}+\frac{\partial\rho}{\partial t} \tag{2.1}$$

In geoelectrical measurements, a DC current is injected into the ground via electrodes. Since electrode size is small compared to the investigated ground array, the electrodes are usually assumed to be a point current source. With an electrode located at $\boldsymbol{r}_{\text{inj}}$ injecting a current $I = const$, (2.1) is

$$\boldsymbol{\nabla}\mathbf{j} = -I\delta(\mathbf{r}-\mathbf{r}_{\text{inj}}) \tag{2.2}$$

Assuming a scalar potential field Φ that vanishes for infinite distances, we can state

$$\mathbf{E} = -\boldsymbol{\nabla}\Phi \tag{2.3}$$

Usually, an isotropic resistivity model following Ohm's law is assumed for the ground, giving

$$\mathbf{j} = \sigma\mathbf{E} \tag{2.4}$$

where σ is the electrical conductivity. Equation (2.2) - (2.4) can be combined to form the governing geoelectrical equation

$$\boldsymbol{\nabla}(\sigma\boldsymbol{\nabla}\Phi) = -I\delta(\mathbf{r}-\mathbf{r}_{\text{inj}}) \tag{2.5}$$

(2.5) is an elliptic partial differential equation (PDE). It is defined for $\Phi \in C^2$, i.e. twice differentiable, and $\sigma \in C^1$ within the domain Ω. Unique solutions require the specification of boundary conditions at the domain boundary. Analytic solutions for (2.5) exist for a number of conductivity models (see chapter 2.3), while for generic conductivity distributions numerical calculations have to be performed. Numerical solutions to elliptic PDE have been widely studied, and the equation is considered to be rather easy to treat numerically. The singular source term on the right hand poses the main problem for numerical solving.

2.3 Analytic solutions

An analytic or semi-analytic solution is known for a number of conductivity distributions. Though analytic solutions are not used any more for data interpretation, they can be used as

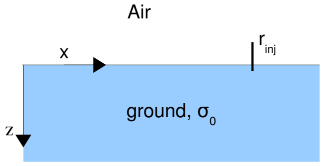


Figure 2.2: Homogeneous half-space with injecting electrode.

starting models or primary fields for numerical calculations, as well as for accuracy testing of the discrete solvers. The two most relevant problems are the homogeneous half-space, and the horizontal two-layered models. The analytic treatment to these and other problems is discussed in details in *Wait* [1982]. Here, the results for the potential distribution are presented for these cases in a brief manner.

2.3.1 Real space

Homogeneous half-space For a homogeneous half-space model with the conductivity σ_0 of the ground ($z > 0$) and a surface current source at the origin, (fig. 2.2) (2.5) simplifies to the regular possion equation

$$\sigma_0 \triangle \Phi = -I \nabla^2(\mathbf{r}) \tag{2.6}$$

The solution for the lower half-space is known to be

$$\Phi_0(\mathbf{r}) = \frac{I}{2\pi\sigma_0|\mathbf{r}|} \tag{2.7}$$

For a subsurface source at $\mathbf{r}_{\text{inj}} = (\mathbf{x_s}, \mathbf{y_s}, \mathbf{z_s})$ the effect of the earth's surface has to be considered. The solution is superposed with the solution of the source mirrored at the surface at $\mathbf{r}_{\text{mr}} = (\mathbf{x_s}, \mathbf{y_s}, -\mathbf{z_s})$.

$$\Phi_0(\mathbf{r}) = \frac{I}{4\pi\sigma_0}\left(\frac{1}{|\mathbf{r}-\mathbf{r}_{\text{inj}}|} + \frac{1}{|\mathbf{r}-\mathbf{r}_{\text{mr}}|}\right) \tag{2.8}$$

In the following, the current source is always assumed to be at the origin for brevity. For the case of a surface current source at $\mathbf{r}_{\text{inj}} = (\mathbf{x_s}, \mathbf{y_s}, \mathbf{0})$ the vector $\mathbf{r}$ obviously just has to be shifted. When dealing with subsurface current electrodes, an additional mirror term as in (2.8) has to be introduced.

Horizontal two-layered model For a two-layered media

$$\sigma(z) = \begin{cases} 0 & , z < 0 \\ \sigma_1 & , 0 < z < h \\ \sigma_2 & , z \geq h \end{cases}$$

it can be shown (*Wait* [1982]) that the potential field is defined by

$$\Phi(x,y,z) = \Phi_{\sigma_1}(x,y,z) + \frac{I}{2\pi\sigma_1}\sum_{n=0}^{\infty} \kappa^{n+1} T_n(\sqrt{x^2+y^2}, z) \tag{2.9}$$

where Φ_{σ_1} is the homogeneous halfspace solution with $\sigma = \sigma_1$, and the factor $\kappa = \frac{\sigma_1+\sigma_2}{\sigma_2-\sigma_1}$. The function T_n is defined as

$$T_n(r,z) = \left(r^2 + (2h(n+1)+z)^2\right)^{-\frac{1}{2}} + \begin{cases} \left(r^2 + (2h(n+1)-z)^2\right)^{-\frac{1}{2}} & , 0 < z < h \\ \left(r^2 + (2h(n+1)+z)^2\right)^{-\frac{1}{2}} & , z > h \end{cases}$$

The summation in (2.9) can be cut off at a value n_{thres} for an approximation to the potential.

2.3.2 Fourier space

As will be described in chapter 2.4.2, numerical finite difference calculations in 2D are performed in a cosine-Fourier transformed space. In order to test accuracy of the inner solver, or pose primary fields, semi-analytic expressions in Fourier space are beneficial. Therefore, as a fetch-ahead, the potentials for the two basic cases given above are also derived in cosine-Fourier space. To do so, the expressions derived in the previous chapter are transformed by

$$\tilde{\Phi}(x,\lambda,z) = \int_0^\infty \Phi(x,y,z)\cos(\lambda y)\,\mathrm{d}y \qquad .$$

Using the relation

$$\int_0^\infty \frac{\cos(x)}{\sqrt{x^2+a^2}}\mathrm{d}x = \mathrm{K}_0(\sqrt{a^2+x^2})$$

the transformed equations can be easily expressed by Bessel functions K_0.

Homogeneous half-space (2.7) translates to

$$\tilde{\Phi}_0(x,\lambda,z) = \frac{I}{2\pi\sigma_0}\mathrm{K}_0(\lambda|x^2+z^2|) \qquad . \tag{2.10}$$

Horizontal two-layered model (2.9) translates to

$$\tilde{\Phi}(x,\lambda,z) = \tilde{\Phi}_{\sigma_1}(x,\lambda,z) + \frac{I}{2\pi\sigma_1}\sum_{n=0}^{\infty}\kappa^{n+1}\tilde{T}_n(x,\lambda,z) \tag{2.11}$$

where $\tilde{\Phi}_{\sigma_1}$ is the transformed homogeneous halfspace solution. The function $\tilde{T}_n$ is defined as

$$\tilde{T}_n(x,\lambda,z) = \mathrm{K}_0(\lambda\sqrt{x^2+(2h(n+1)+z)^2} + \begin{cases} \mathrm{K}_0(\lambda\sqrt{x^2+(2h(n+1)-z)^2} & , 0 < z < h \\ \mathrm{K}_0(\lambda\sqrt{x^2+(2hn+z)^2} & , z > h \end{cases}$$

2.4 Numerical forward solving

The forward operator of (2.5) is defined as the transformation from model space $\{\mathbf{m} = \sigma(\mathbf{r})\}$[3] into data space $\{\mathbf{d} = \Phi(\mathbf{r})\}$. For bodies of closed geometry, boundary integral

---

[3] Strictly speaking, the current source location $\mathbf{r}_{\text{inj}}$ and strength I also belong into model space. This is omitted in the following for brevity.

methods *Xu et al.* [1988] can be used to describe the forward operator. For some special cases, e.g. layered structures of constant conductivity, a semi-analytic solution exists. For a general conductivity distribution however, a numerical calculation on a discretized parameter space is usually required. The discrete sampling of the continuous potential distribution is examined best in Fourier space

$$\tilde{V}(\lambda) = \frac{1}{(2\pi)^{3/2}} \int_{\mathbf{R}^3} V(\mathbf{r}) \mathrm{e}^{-i\lambda\mathbf{r}} \mathrm{d}\mathbf{r} \qquad \Longleftrightarrow \qquad V(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int_{\mathbf{R}^3} \tilde{V}(\lambda) \mathrm{e}^{-i\lambda\mathbf{r}} \mathrm{d}\lambda$$

where λ is the Fourier variable vector. Discretization on an equidistant grid with lattice constant h means the integrals are turned into finite summation. Only components that can be described by the transformed distribution $\tilde{V}(\lambda) < \frac{1}{2h}$ are represented correctly as the remaining terms vanish.

The methods most commonly used to approximate the forward operator are finite differences and finite elements.

- Finite differences are computationally fast and straightforward to implement. However, this method requires a regular grid in e.g. cylindrical or Cartesian coordinate system. It is used when surface elevation can be neglected, and regular spacing is adequate.
- Finite elements are more expensive, but also more versatile. They can be used for explicit modelling of surface topography, or for a more versatile discretization focusing/boundary matching in areas of high contrast. The generation of a problem-adapted finite elements grid is however non-trivial. See *Sasaki* [1994] for the application of FE to Geoelectrics.

In *Li and Spitzer* [2002], a comparison of these methods can be found. There are also other approaches on the problem, e.g. using charge densities *Boulanger and Chouteau* [2005]. For the type of experiments conducted, finite difference are adequate, so I will concentrate on FD methods here.

2.4.1 3D FD operator

Forward operator For 3D finite difference modelling, the domain is discretized by a grid with nodes x_i $(i \in 1 \ldots n)$, y_j $(j \in 1 \ldots m)$, z_j $(j \in 1 \ldots l)$. In the following, the discretization scheme of *Dey and Morrison* [1979b] is used. Fig. 2.3 shows a subsection of the grid around node (i,j,k). The potential Φ_{ijk} is defined at the grid node (i,j,k), while the conductivity σ_{ijk} is defined in the area bounded by the nodes (i, i+1, j, j+1, k, k+1). By integrating (2.5) over the domain Ω_{ijk} around node (i,j,k), we obtain

$$\iiint_{\Omega_{ijk}} \boldsymbol{\nabla}(\sigma \boldsymbol{\nabla} \Phi)\, \mathrm{d}^3\mathbf{r} = -\iiint_{\Omega_{ijk}} I\delta(\mathbf{r} - \mathbf{r}_{\mathrm{inj}})\, \mathrm{d}^3\mathbf{r} = -I_{ijk} \tag{2.12}$$

With Γ_{ijk} denoting the surface boundary of Ω_{ijk} and η the normal vector, application of Gauss' theorem yields

$$\iint_{\Gamma_{ijk}} \sigma \frac{\partial \Phi}{\partial \eta} \mathrm{d}\Gamma = -I_{ijk} \tag{2.13}$$

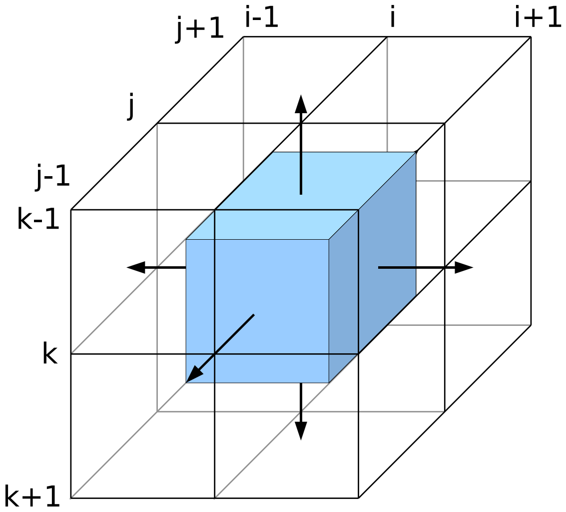


Figure 2.3: 3D-FD grid around node (i,j,k). The domain Ω_{ijk} is drawn in blue.

We now apply finite differences to the equation. For each of the six surface boundaries, $\frac{\partial\Phi}{\partial\eta}$ is approximated by a central difference, e.g. for the upper surface $(\frac{\partial\Phi}{\partial\eta})_{\text{up}} \approx \frac{\Phi_{i,j,k-1}-\Phi_{i,j,k}}{z_k-z_{k-1}}$. A weighted mean is used for the integration of σ – in the case of the upper surface, this means

$$\begin{aligned}\overline{\sigma}^{\text{up}}_{i,j,k} = \frac{1}{4}\big[&(x_i - x_{i-1})(y_j - y_{j-1})\sigma_{i-1,j-1,k-1} + (x_{i+1} - x_i)(y_j - y_{j-1})\sigma_{i,j-1,k-1} + \\ &(x_i - x_{i-1})(y_{j+1} - y_j)\sigma_{i-1,j,k-1} + (x_{i+1} - x_i)(y_{j+1} - y_j)\sigma_{i,j,k-1}\big] \qquad (2.14)\end{aligned}$$

The approximated integral (2.13) can therefore be factorized as

$$\begin{aligned}-I_{i,j,k} =& C_{\text{up}}\,\Phi_{i,j,k-1} + C_{\text{down}}\,\Phi_{i,j,k+1} + C_{\text{left}}\,\Phi_{i-1,j,k} + C_{\text{right}}\,\Phi_{i+1,j,k} + \\ & C_{\text{front}}\,\Phi_{i,j-1,k} + C_{\text{back}}\,\Phi_{i,j+1,k} + C_{\text{self}}\,\Phi_{i,j,k} \qquad (2.15)\end{aligned}$$

where the upper example coupling coefficient C_{up} is

$$C_{\text{up}} = \overline{\sigma}^{\text{up}}_{i,j,k}\,\frac{\Phi_{i,j,k-1} - \Phi_{i,j,k}}{z_k - z_{k-1}} \qquad .$$

The self coupling coefficient C_{self} is the negative sum of the 6 other coefficients

$$C_{\text{self}} = -(C_{\text{up}} + C_{\text{down}} + C_{\text{left}} + C_{\text{right}} + C_{\text{front}} + C_{\text{back}})$$

Forming (2.15) for every inner node lead to a system of linear equations in Φ

$$\mathbf{L}_\sigma \boldsymbol{\Phi} = \mathbf{b} \qquad (2.16)$$

where $\mathbf{L}_\sigma$ is the finite difference forward operator, and $\mathbf{b}$ contains the current source terms.

Boundary conditions Since finite differences can only handle finite domains, boundary conditions have to be introduced at the domain boundaries. There are generally 3 types of conditions

- **Dirichlet boundary conditions** specify fixed potential values at the boundaries. The exact values are not known for non-trivial conductivity models. However, it can be made use of the fact that potential vanishes for infinite distances. A common technique is to add some prolonged grid cells around the modeling domain, and then specifying the potential of these new "far" boundaries as zero. Implementation is as easy as replacing (2.15) by $\Phi_{i,j,k} = c$ for boundary cells.
- **Neumann boundary conditions** fix the potentials derivative at the boundaries. As the derivative essentially specifies the flux, Neumann conditions can be used to prevent current flow through the ground surface, by enforcing a derivative of zero. They are therefore called "no-flow boundary conditions". It can implemented by omitting the coefficient pointing outwards at the boundary layer (i.e. $C_{\text{up}} = 0$ for ground surface).
- **Mixed boundary conditions** relate the potential and its derivative, enforcing a potential characteristic at the boundary, instead of an absolute value. This is a better approximation than Dirichlet or Neumann conditions for most problems. Mixed boundary conditions have been proposed first by *Dey and Morrison* [1979b]. The derivative of the homogeneous halfspace (2.7) by the boundary's outward normal η can be written as
  $$\frac{\partial \Phi_{\text{homo}}}{\partial \eta} = -\frac{\cos(\theta)}{r}\Phi_{\text{homo}}$$
  where θ is the angle between the source vector and the outward normal vector[4]. At the vertical boundary for instance, $\cos(\theta) = \frac{x}{r}$. Alternatively, other characteristics based on e.g. the analytic solution for a layered conductivity model could also be applied. Mixed boundary conditions have to be calculated in advance for each source electrode configuration, but usually the source is approximated to be in the middle for all configurations, which is a good approximation for long distances. Implementation is done as for Dirichlet conditions.

Singularity removal As stated before, the singularity at the current source constitutes the main challenge in forward solving (2.5). In order the minimize the error through finite difference approximation, a very high discretization around the injection point would be required, as this area is obviously the main error source. This is not really viable however, since current is injected at some 50-200 different points in a typical measurement. *Lowry et al.* [1989] was first to apply secondary field techniques to Geoelectrics. The approach is to split the potential $\Phi = \Phi_p + \Phi_s$ into a primary (normal) potential Φ_p and a secondary (abnormal) potential Φ_s, so that the secondary potential remains free from singularities. The primary potential is chosen to be an easy-to-calculate potential, which satisfies

$$\boldsymbol{\nabla}(\sigma_p \boldsymbol{\nabla}\Phi_p) = -I\delta(\mathbf{r} - \mathbf{r}_{\text{inj}}) \tag{2.17}$$

---

[4] The injection point is assumed to be at the origin here.

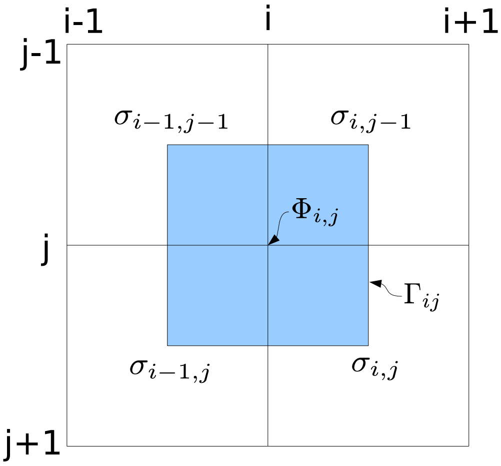


Figure 2.4: 2D-FD grid around node (i,j). The domain Ω_{ij} is drawn in blue.

Usually, the homogeneous half space model with $\sigma_p = \mathrm{const}$ is used. In some cases, layered primary models can produce better result. We concentrate on the first case here. The total potential Φ has to satisfy (2.5) :

$$\boldsymbol{\nabla}(\sigma\boldsymbol{\nabla}(\Phi_p + \Phi_s)) = \boldsymbol{\nabla}(\sigma\boldsymbol{\nabla}\Phi_p) + \boldsymbol{\nabla}(\sigma\Phi_s) = -I\delta(\mathbf{r} - \mathbf{r}_{\mathrm{inj}}) \tag{2.18}$$

Substituting (2.17) in (2.18) yields

$$\boldsymbol{\nabla}(\sigma\boldsymbol{\nabla}(\Phi_s)) = \boldsymbol{\nabla}((\sigma_p - \sigma)\boldsymbol{\nabla}\Phi_p) \qquad , \tag{2.19}$$

therefore extending (2.16) to

$$\mathbf{L}_\sigma\mathbf{\Phi_s} = \mathbf{L}_{(\sigma_\mathbf{p}-\sigma)}\mathbf{\Phi_p} \qquad . \tag{2.20}$$

As stated by *Zhao and Yedlin* [1996], obviously the singularity is removed only if $\sigma_p = \sigma$ at the injection point. Thus, for each current electrode position, a different σ_p has to be used. To prevent the recalculation of the L_{σ_p}'s, (2.20) can be written as

$$\mathbf{L}_\sigma\mathbf{\Phi_s} = (\sigma_p\mathbf{L_1} - \mathbf{L}_\sigma)\mathbf{\Phi_p} = \sigma_p\mathbf{L_1}\mathbf{\Phi_p} - \mathbf{L}_\sigma\mathbf{\Phi_p} \tag{2.21}$$

due to linearity.

When using singularity removal, it is often a good approximation to set the secondary potential to zero at the boundaries, while the enforcing of a potential characteristic makes no longer sense. Therefore, Dirichlet boundary conditions are preferred over mixed-type conditions in this case.

2.4.2 2D FD operator

2.5D space The underlying principles of 3D numerical forward modeling (2.4.1) hold for 2D modeling as well. Here, the conductivities $\sigma(x, y, z) = \sigma(x, z)$ are assumed to be constant in y-direction, the potential distribution however has to be still calculated in three dimensions. To prevent having to solve the full 3D problem, so-called 2.5D-solving can be applied. Essentially, this means working with a cosine-Fourier transformed potential *Dey and Morrison* [1979a].

$$\tilde{\Phi}(x,\lambda,z) = \int_0^\infty \Phi(x,y,z)\cos(\lambda y)\,\mathrm{d}y \tag{2.22}$$

Cosine-Fourier transformation of (2.5) yields

$$\boldsymbol{\nabla}(\sigma\boldsymbol{\nabla}\tilde{\Phi}) - \lambda^2\sigma\tilde{\Phi}(x,\lambda,z) = \frac{I}{2}\delta(x-x_{\mathrm{inj}})\delta(z-z_{\mathrm{inj}}) \qquad . \tag{2.23}$$

$\boldsymbol{\nabla}$ denotes the two-dimensional divergence/gradient here. The 2D problem (2.23) has to be solved for $\lambda = 0\ldots\infty$ and $\tilde{\Phi}$ has to be Fourier-backtransformed. Since only the potential at $y=0$ is needed, backtransformation reduces to an integration

$$\Phi(x,y=0,z) = \frac{2}{\pi}\int_0^\infty \tilde{\Phi}(x,\lambda,z)\,\mathrm{d}\lambda \tag{2.24}$$

Forward operator Again, we discretize our domain by a grid with nodes x_i ($i \in 1\ldots n$), z_j ($j \in 1\ldots l$). Fig. 2.4 shows a subsection of the grid around node (i,j). The potential Φ_{ij} is defined at the grid node (i,j), while the conductivity σ_{ij} is defined in the area bounded by the nodes (i, i+1, j, j+1). To solve (2.23), finite difference approximation is performed in the same way as in the 3D-case (chapter 2.4.1). (2.23) is integrated over the domain Ω_{ij} Gauss' theorem is applied to obtain

$$\oint_{\Gamma_{ij}} \sigma\frac{\partial\tilde{\Phi}}{\partial\eta}\mathrm{d}\Gamma - \lambda^2\iint_{\Omega_{ij}}\sigma\,\tilde{\Phi}\,\mathrm{d}\mathbf{r} = \frac{1}{2}I_{ij} \qquad . \tag{2.25}$$

Here, η is the boundary's outwards normal. For each of the four boundary lines in Γ_{ij}, $\frac{\partial\tilde{\Phi}}{\partial\eta}$ is approximated by a central difference, e.g. for the upper line $(\frac{\partial\tilde{\Phi}}{\partial\eta})_{\mathrm{up}} \approx \frac{\tilde{\Phi}_{i,j-1}-\Phi_{i,j}}{z_j-z_{j-1}}$. A weighted mean along the integration line is to approximate σ. For the upper line, this means

$$\overline{\sigma}^{\mathrm{up}}_{i,j} = \frac{1}{2}\big[(x_i-x_{i-1})\sigma_{i-1,j-1} + (x_{i+1}-x_i)\sigma_{i,j-1}\big] \tag{2.26}$$

The area integral is approximated by using an weighted area mean for σ, and assuming $\tilde{\Phi}$ to be constant over the integration area.

$$\begin{aligned}\overline{\sigma}_{\mathrm{int}} = \iint_{\Omega_{ij}}\sigma\tilde{\Phi}\mathrm{d}\mathbf{r} \approx & \frac{1}{4}\big[(z_{j+1}-z_j)(x_{i+1}-x_i)\sigma_{i,j} + (z_{j+1}-z_j)(x_i-x_{i-1})\sigma_{i-1,j} + \\ & (z_j-z_{j-1})(x_{i+1}-x_i)\sigma_{i,j-1} + (z_j-z_{j-1})(x_i-x_{i-1})\sigma_{i-1,j-1}\big]\end{aligned} \tag{2.27}$$

The approximated equation (2.25) can be factorized as

$$-I_{i,j} = C_{\mathrm{up}}\,\tilde{\Phi}_{i,j-1} + C_{\mathrm{down}}\,\tilde{\Phi}_{i,j+1} + C_{\mathrm{left}}\,\tilde{\Phi}_{i-1,j} + C_{\mathrm{right}}\,\tilde{\Phi}_{i+1,j} + C_{\mathrm{self}}\,\tilde{\Phi}_{i,j} \tag{2.28}$$

where the example upper coupling coefficient C_{up} is

$$C_{\text{up}} = \overline{\sigma}^{\text{up}}_{i,j} \frac{\tilde{\Phi}_{i,j-1} - \tilde{\Phi}_{i,j}}{z_j - z_{j-1}} \qquad .$$

The self coupling coefficient C_{self} is the negative sum of the 4 other coefficients and the area integral

$$C_{\text{self}} = -(C_{\text{up}} + C_{\text{down}} + C_{\text{left}} + C_{\text{right}}) + \lambda^2 \overline{\sigma}_{\text{int}}$$

Forming (2.28) for every node lead to a system of linear equations in $\tilde{\Phi}$

$$\mathbf{L}_\sigma \tilde{\mathbf{\Phi}} = \mathbf{b} \tag{2.29}$$

where $\mathbf{L}_\sigma$ is the finite difference forward operator.

Boundary conditions The types of boundary conditions as well as their advantages and disadvantages are identical to the 3D case (eq. 2.4.1). Implementation of mixed boundary conditions however differs, as a consequence of the Fourier space. Here, usually (2.10) is used for a characteristic. This yields

$$\frac{\partial \tilde{\Phi}_{\text{homo}}}{\partial \eta} = -\lambda \cos(\theta) \frac{\mathrm{K}_1(\lambda |\mathbf{r}|)}{\mathrm{K}_0(\lambda |\mathbf{r}|)}$$

where θ is the angle between the source vector and the outward normal vector η[5]. Again, due to the function's fast decay, $\mathbf{r}$ can be approximated as the distance to the electrode array middle instead of the real injection point. The boundary condition can then be precalculated and used for every configuration.

Singularity removal Again, the results of the 3D singularity removal can be re-used for 2.5D by Fourier-transforming (2.19).

$$\mathbf{\nabla}(\sigma \mathbf{\nabla} \tilde{\Phi}_s) - \lambda^2 \sigma \tilde{\Phi}_s = -\mathbf{\nabla}\left((\sigma - \sigma_p)\, \mathbf{\nabla} \tilde{\Phi}_p \right) + \lambda^2 (\sigma - \sigma_p) \tilde{\Phi}_p \tag{2.30}$$

The primary potential is thus again the homogeneous halfspace, however in Fourier space (2.10). With the 2D forward operator $\mathbf{L}$, this can be written as

$$\mathbf{L}_\sigma \tilde{\mathbf{\Phi}}_{\mathbf{s}} = \sigma_p \mathbf{L}_{\mathbf{1}} \tilde{\mathbf{\Phi}}_{\mathbf{p}} - \mathbf{L}_\sigma \tilde{\mathbf{\Phi}}_{\mathbf{p}} \qquad .$$

2.4.3 Equation system solvers

(2.29) and (2.16) are systems of linear equations. which have at most one solution. The matrix $\mathbf{L}_\sigma$ is positive, semi-definite and regular (except for the case of all-no-flux boundary conditions and nonzero current injection, where $\mathbf{L}_\sigma$ is singular). Due to the FD structure, $\mathbf{L}_\sigma$ is a sparse, banded matrix, with a bandwidth of $2n + 1$ for 2D, and $2mn + 1$ for the 3D

---

[5] The injection point is assumed to be at the origin here.

case. The system can be solved by methods like Gauss-elimination, QR-decomposition or Cholesky decomposition together with the matching solver, which are in the complexity class $\mathcal{O}(N^3)$ (*Locher* [1997]). Assuming d steps for λ, this means two-dimensional solving can be done in $\mathcal{O}(d\,(n\,l)^3)$, compared to $\mathcal{O}((n\,m\,l)^3)$ for full three-dimensional solving. Of the exact solvers, Cholesky decomposition $\mathbf{L}_\sigma = \mathbf{M}\,\mathbf{M}^T$ makes use best of the matrix' structure. The incomplete decomposition for Cholesky or LU converge much faster, and can also be applied.

A preconditioner can be used to reduce the condition number, which can be rather high especially when irregular grid spacing is used. Instead of the original system,

$$\mathbf{P}^{-1}\,\mathbf{L}_\sigma\,\mathbf{\Phi} = \mathbf{P}^{-1}\,\mathbf{b}$$

is solved. For $\mathbf{P}$, an incomplete chalky preconditioner ($\mathbf{L}_\sigma \approx \mathbf{P}^T\,\mathbf{P}$) shows the best results. A comparison of different solvers can be found in *Spitzer and Wurmstich* [1995].

2.5 Sensitivity

The sensitivity S_V is defined as the change of the model response $f(m(\mathbf{r}))$ under the change of the model m in the volume V.

$$S_V = \iiint_V \frac{\partial f(m(\mathbf{r}))}{\partial m}\,\mathrm{d}^3\mathbf{r}$$

Sensitivities are needed for gradient-based inversion schemes like Gauss-Newton, and also useful for evaluation of resolution or array optimization. In the case of DC Geoelectrics, $m(\mathbf{r}) = \rho(\mathbf{r})$ denotes the model resistivities and f is the forward operator. The model response here is defined as the potential Φ at the electrodes for a pole-pole measurement.

2.5.1 Homogeneous halfspace sensitivity

3D sensitivity The sensitivity for a homogeneous halfspace can be calculated semi-analytically. It is widely used for homogeneous inversion starting models and as approximations to the real sensitivities of more complicated resistivity distributions. Without loss of generality, we assume the current electrode at $\mathbf{r_a} = (0,0,0)$ and the potential electrode at $\mathbf{r_m} = (d,0,0)$. *Park and Van* [1991] showed that the sensitivity in the volume V can be calculated by the reciprocity principle

$$S_V(0,d) = \iiint_V \frac{\partial \Phi_a(\mathbf{r_m})}{\partial \rho}\mathrm{d}^3 r = \iiint_V \frac{1}{\rho^2}\frac{\nabla\Phi_a(\mathbf{r})}{I_a}\,\frac{\nabla\Phi_m(\mathbf{r})}{I_m}\mathrm{d}^3\mathbf{r} \tag{2.31}$$

where Φ_a is the potential caused by injection of a current I_a at $\mathbf{r_a}$, while Φ_m is the potential that would be caused by injection of a current I_m at $\mathbf{r_m}$. Inserting the derivative of (2.7) results in

$$S_V^h(0,d) = \frac{I_0}{4\pi^2}\iiint_V \frac{x(x-a)+y^2+z^2}{(x^2+y^2+z^2)^{\frac{3}{2}}((x-a)^2+y^2+z^2)^{\frac{3}{2}}}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \qquad . \tag{2.32}$$

This equation can be integrated numerically by Gauss-Legendre integration or adaptive integration methods.

2D sensitivity For 2D sensitivity, the y-integration in (2.32) is replaced by an infinite integration $y = -\infty \ldots + \infty$. This integration step can evaluated analytically, as shown by *Loke and Barker* [1995][6].

$$S_V^{2d}(0, a) = \iint_V \frac{I_0}{4\pi^2} G(x, z, a) \mathrm{d}x \, \mathrm{d}z \tag{2.33}$$

for $x < \frac{a}{2}$

$$G(x, z, y) = \frac{2}{\beta} \left[\frac{\mathrm{K}(q_1) - \mathrm{E}(q_1)}{\beta^2 - \alpha^2} - a\,x\, \frac{(\beta^2 + \alpha^2)\mathrm{E}(q_1) - 2\alpha^2 \mathrm{K}(q_1)}{\alpha^2(\beta^2 - \alpha^2)^2} \right]$$

for $x = \frac{a}{2}$

$$G(x, z, y) = \frac{\pi}{2} \left[\frac{1}{\alpha^3} - \frac{3x^2}{2\alpha^5} \right]$$

for $x > \frac{a}{2}$

$$G(x, z, y) = \frac{2}{\alpha\beta^2} \left[\frac{\alpha^2 \mathrm{E}(q_2) - \beta^2 \mathrm{K}(q_2)}{\alpha^2 - \beta^2} - a\,x\, \frac{(\alpha^2 + \beta^2)\mathrm{E}(q_2) - 2\beta^2 \mathrm{K}(q_2)}{(\alpha^2 - \beta^2)^2} \right]$$

where $\alpha^2 = x^2 + z^2$, $\beta^2 = (x - a)^2 + z^2$, $q_1 = \frac{1}{\beta}\sqrt{\beta^2 - \alpha^2}$, $q_2 = \frac{1}{\alpha}\sqrt{\alpha^2 - \beta^2}$, and K and E are the complete elliptical integrals of first and second kind. Again, (2.33) can be integrated numerically by Gauss-Legendre integration or adaptive integration methods.

2.5.2 4-electrode sensitivities

Since the potential of a four-electrode configuration (x_A, x_B, x_M, x_N) is a superposition of pole-pole potentials, 4-electrode sensitivities can be calculated as

$$S_V^4(x_A, x_B, x_M, x_N) = S_{V_A}(0, x_M) - S_{V_A}(0, x_N) - S_{V_B}(0, x_M) + S_{V_B}(0, x_N)$$

where V_A is V shifted by $-x_A$, V_B is V shifted by $-x_B$.

2.5.3 Numerical calculation for inhomogeneous models

The sensitivities for inhomogeneous models are calculated on a discretization grid. In the discrete case, sensitivity is a matrix $\mathbf{S_{ij}}$, where $i = 1 \ldots N$ denotes the measurement, and $j = 1 \ldots M$ denotes the changed model cell

$$\mathbf{S_{ij}}(\rho) = \frac{\partial \Phi_{[i]}}{\partial \rho_j} \quad .$$

There are basically three ways of calculating sensitivities numerically

---

[6] The elliptic integrals' arguments given there have misprints, though.

- **Direct forward calculation**
  Consider the system of equations obtained by finite differences forward calculation, e.g. (2.16). Following *Spitzer* [1998], this equation can be differentiated by the resistivity cells ρ_i
  $$\mathbf{L}_\sigma \frac{\partial \mathbf{\Phi}}{\partial \rho_i} + \frac{\partial \mathbf{L}_\sigma}{\partial \rho_i} \mathbf{\Phi} = 0 \qquad .$$
  Therefore, by solving the equation system for the vector $\frac{\partial \mathbf{\Phi}}{\partial \rho_i}$, the sensitivity can be obtained. However, since each model cell and each measurement have an differing source term, $N\,M$ equations have to be solved. Although the numerical cost can be reduced by preconditioning $\mathbf{L}_\sigma$, this method is seldom used, as model cell counts are high for most problems.

- **Perturbation**
  A direct approximation of the sensitivity can be obtained by finite differences
  $$\mathbf{S_{ij}} \approx \frac{f(\rho + \Delta\rho\, \mathbf{e}_j)_i - f(\rho)_i}{\Delta\rho}$$
  with $\Delta\rho$ being a small perturbation and $\mathrm{e_j}$ defined as $(e_j)_k = \delta_{j,k}$. This means, $(n\,m\,l)$ additional forward calculations with all measurements have to be calculated. For small numbers of model parameters, this method is often used due to simplicity, for full ERT sensitivity calculations it is however not useful due to its cost.

- **Approximation**
  Equation (2.31) can also be evaluated numerically. The partial derivatives are again approximated by finite differences e.g.
  $$\begin{aligned}\frac{\partial \Phi_{i+0.5,j,k}}{\partial x} \approx \frac{1}{4\Delta x} \big[& \Phi_{i+1,j,k} + \Phi_{i+1,j+1,k} + \Phi_{i+1,j,k+1} + \Phi_{i+1,j+1,k+1} - \\ & \Phi_{i,j,k} - \Phi_{i,j+1,k} - \Phi_{i,j,k+1} - \Phi_{i,j+1,k+1} \big]\end{aligned}$$
  This method has the advantage that only one forward calculation is needed. It however suffers from the inaccuracies of the finite difference approximation especially near the current injection point.

When using Quasi-Gauss-Newton inversion scheme, sensitivity is not recalculated explicitly each step. Instead, homogeneous sensitivities are used for the first iteration, and Broyden updates (see 2.41) are used as a means to approximate the next steps.

In inventing a model we may assume what we wish, but should avoid impossibilities. -Aristotle

2.6 Inversion

In Geoelectrics, a model $\mathbf{m}$ of the subsurface has to be determined on the base of a limited number of measurements $\mathbf{d}$, which are in addition contaminated with noise. In most cases, it is not possible to obtain all model parameters by a direct evaluation of the measurement data. Therefore, a model whose forward response $\mathbf{f}(\mathbf{m})$ is near the measured data vector is sought. This is called an inverse problem. Linear inverse theory, i.e. problems with a linear forward operator $\mathbf{f}$, is a well-investigated topic. A good source here is e.g. *Menke* [1989]. However, with the notable exception of gravimetry and vertical seismic profiling, most geophysical problems, including Geoelectrics, are non-linear. Non-linear theory is not nearly as well covered – A common approach for these non-linear problems is linearization. This means iteratively recalculating the gradient and solving the linear inverse problem until a threshold is met. (Quasi-) Newton type techniques *Loke and Barker* [1996] are used most commonly and successfully for most problems, and will be described in the following section. A comprehensive description of Newton-type and different other methods like non-linear conjugate gradients or SIRT can be found in *Günther* [2004]. Instead of iterative inversion, statistical approaches *Yeh et al.* [2002] or global optimization algorithms like genetic algorithms *Schwarzbach* [2003] and simulated annealing could be used. They have the advantage that an indicator for the solution goodness and uniqueness may be derived. The computational cost is however still too high for application to realistic problems.

2.6.1 Gauss-Newton method

Inversion is essentially a minimization problem. For a data vector $\mathbf{d}$, a model parameter vector m and a forward operator $\mathbf{f}$ the objective function to minimize is chosen as

$$\Psi(\mathbf{m}) = ||\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}))||^2 \qquad . \tag{2.34}$$

The matrix $\mathbf{D} = \mathrm{diag}(\epsilon^{-\mathbf{1}})$ with data error vector ϵ is used here to weight the data misfit. The error values can be determined from the measured data (standard deviation via stacking, or reciprocity error) or be derived from an error propagation model. The L_2-norm is chosen here as the noise is assumed to be Gaussian. Sometimes L_1 inversion is also desirable, as it is less sensitive to outliers. While a direct L_1 inversion is not stable for ERT, quasi L_1 iterative data reweighting is commonly used (robust inversion, see chapter 2.6.3). As stated before, Newton-type methods are iterative. Starting from a initial model $\mathbf{m}^0$, the model $\mathbf{m}^k$ is updated each iteration step k by applying a model update difference, i.e.

$$\mathbf{m}^{k+1} = \mathbf{m}^k + \Delta\mathbf{m}^k$$

In order to minimize (2.34) for the updated model m^{k+1}, we first consider the 2^{nd}-order Taylor series

$$\Psi(\mathbf{m}^k + \Delta\mathbf{m}^k) \approx \Psi(\mathbf{m}^k) + (\boldsymbol{\nabla}\Psi(\mathbf{m}^k))^T\Delta\mathbf{m}^k + \frac{1}{2}\Delta(\mathbf{m}^k)^T(\boldsymbol{\nabla}^2\Psi(\mathbf{m}^k))^T\Delta\mathbf{m}^k \tag{2.35}$$

In finding minima, forcing the first derivative in $\Delta\mathbf{m}^k$ to zero is sufficient, as the positivity of the Hessian can be shown. This results in the linear equation system

$$(\boldsymbol{\nabla}^2\Psi(\mathbf{m}^k))\Delta\mathbf{m}^k = -\boldsymbol{\nabla}\Psi(\mathbf{m}^k) \tag{2.36}$$

For the Newton objective function (2.34), this equation reads

$$\underbrace{\left[(\mathbf{DS})^T(\mathbf{DS}) + (\boldsymbol{\nabla}^T\mathbf{S}^T)\mathbf{D}^T\mathbf{D}(\mathbf{f}(\mathbf{m}^k) - \mathbf{d})\right]}_{\mathbf{H}^k} \Delta\mathbf{m}^k = (\mathbf{DS})^T\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}^k)) \qquad (2.37)$$

with the sensitivity matrix $\mathbf{S}_{ij} = \frac{\partial \mathbf{f}(\mathbf{m}^k)_i}{\partial m_j}$ (see 2.5). As the second term of the Hessian $\mathbf{H}^k$ is computationally complex to calculate, it is neglected in the Gauss-Newton method. (2.37) reduces to

$$\underbrace{(\mathbf{DS})^T(\mathbf{DS})}_{\mathbf{H}^k} \Delta\mathbf{m}^k = (\mathbf{DS})^T\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}^k)) \qquad . \qquad (2.38)$$

This lowers the convergence property, especially for strong non-linear problems, but yields a better overall execution time of the algorithm.

During each iteration, the right-hand side gradient in (2.36) as well as the $\mathbf{H}^k$ have to be calculated, and the linear problem (2.36) solved for the update vector $\Delta\mathbf{m}^k$. The inversion is stopped if the root mean square (RMS) of the data misfit, or the RMS decrease rate lies under a threshold value. Note that since the minima search is only local, the inversion might get stuck in a local minimum that is far from the global minimum. As the problem is however strong multidimensional and ambiguous, local minima will be largely near to the global optimum.

2.6.2 Regularization

ERT problems are known to be severely ill-posed, due to the large number of model parameters compared to independent data points. This non-uniqueness is an inherent problem of the ERT method. Minimizing the data misfit alone by an inversion method results in a model which might or might not be representative of the true model. Often the application of e.g. pure Gauss-Newton yields a highly oscillating resulting model. Therefore, additional constraints have to be provided to the inversion, favoring e.g. feasibility, or after Occam, “simplicity” of a model. This concept mainly bases on the work *Constable et al.* [1987]. On the mathematical side, this process is a regularization of the matrix $\mathbf{H}^k$ in (2.38). This can be implemented by adding an additional penalty function to the objective function (2.34)

$$\Psi(\mathbf{m}) = ||\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}))||^2 + \gamma||\mathbf{C}(\mathbf{m} - \mathbf{m}_0)||^2 \qquad . \qquad (2.39)$$

$\mathbf{m}_0$ is here a reference model, which can a homogeneous model, a model known from previous investigations or the previous iterations’ model, depending on the desired regularization scheme. If the reference model from the previous iteration is used, the regularization scheme is called local regularization, else global regularization.

The regularization parameter γ controls the regularization strength, and is normally hand-tuned for best results. (There are also approaches for automatically obtaining an ideal γ, as the L-curve algorithm *Hansen and O’Leary* [1993]). Finally, the $\mathbf{C}$ matrix controls the type of regularization

- **Minimum length regularization** minimizes the difference between model and reference model. This is useful for example in time series, for enforcing small changes between the measurements. Another case is ridge regression (Levenberg-Marquardt), which limits the model update changes to prevent oscillating behavior. For Minimum length regularization, the C-Matrix is an identity, or weighted diagonal matrix $\mathbf{C} = \mathrm{diag}(\mathbf{w_C})$.

- **Smoothness constraining** is the most commonly used constraint for general problems. It bases on the assumption that spatial variations in the resistivity model are slow, thus enforcing a smooth model. This is done by using a first-order differential operator for $\mathbf{C}$. Using finite differences, $\mathbf{C}$ has the form

  $$\mathbf{C} = \begin{pmatrix} -1 & 1 & 0 & \dots \\ 0 & -1 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

  for an equally spaced one-dimensional model space. For 3D, the total operator can be composed from the respective x,y,z-Operators as

  $$\mathbf{C} = \begin{pmatrix} \alpha_x \mathbf{C_x} \\ \alpha_y \mathbf{C_y} \\ \alpha_z \mathbf{C_z} \end{pmatrix}$$

  Tuning of the α's allows for supporting lateral or vertical structure, which can be useful if e.g. layering is expected. For boundaries with high resistivity contrasts, robust constraining (i.e. using the L_1 norm) can be useful. L_1 allows for sharper boundaries than L_2 smoothness constraints, therefore the typical large gradient around the boundaries can be avoided.

- **Higher-order contraining**, especially second-order differential operators are also sometimes used in the same way as first-order smoothness constraints. They are useful when dealing with abnormalities with a high resistivity contrast.

Applying the Gauss-Newton procedure in chapter 2.6.1 to the new objective function (2.39) yields the final equation system

$$\underbrace{\left((\mathbf{DS})^T(\mathbf{DS}) + \gamma \mathbf{C}^T\mathbf{C}\right)}_{\mathbf{H}_c^k} \Delta \mathbf{m}^k = \mathbf{S}^T\mathbf{D}^T\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}^k)) - \gamma \mathbf{C}^T\mathbf{C}(\mathbf{m}^k - \mathbf{m}_0) \qquad (2.40)$$

It should be noted that most the problems inherent to ERT inversion can be attributed to its ill-posedness and the therefore necessary strong regularization. Therefore, the choice of regularization and starting model has a great impact on the inversion result. Techniques that stabilize the inversion with additional hard data (see 3.2.3), allow for weaker regularization and can produce more trustworthy results.

2.6.3 Additional techniques

Here, some commonly used techniques to improve the inversion quality or speed will be introduced shortly.

- **Quasi-Gauss-Newton inversion**
  As $\mathbf{H}^k$ contains the sensitivity matrix which needs to be recalculated as well, often $\mathbf{H}^k$ is not evaluated each step, but approximately updated. The procedure is called Broyden update and was adapted to Geoelectrics by *Loke and Barker* [1996]. The Broyden equation is

$$\mathbf{H}^{k+1} \approx \mathbf{H}^k + \frac{(\mathbf{f}(\mathbf{m}^{k+1}) - \mathbf{f}(\mathbf{m}^k) - \mathbf{H}^k \Delta\mathbf{m}^k)\Delta(\mathbf{m}^k)^T}{(\Delta\mathbf{m}^k)^T \Delta\mathbf{m}^k} \quad . \tag{2.41}$$

- **Line search**
  To improve convergence for strong non-linear problems, a line search procedure may be implemented. Here, a parameter τ is introduces, and the optimality of a scaled model update $\mathbf{m}^k + \tau\Delta\mathbf{m}^k$ is tested. This can be done by testing/interpolating $\Psi(\mathbf{m}^k + \tau\Delta\mathbf{m}^k)$ for different τ. τ lies generally between $0 \ldots 1$, with 1 being the linear case (*Press et al.* [2002]).

- **Robust inversion**
  As stated above, sometimes a L_1 inversion is desired for its error characteristic. As it is however hard to cope with numerically, a quasi-L_1 inversion scheme called “Robust inversion” can be applied. It works by reweighting the data values, so that a L_1 norm is simulated. At each inversion step, a weight

$$\frac{\frac{|\Delta d_i|}{\|\Delta\mathbf{d}\|^2}}{\frac{|\Delta d_i|^2}{\|\Delta\mathbf{d}\|^2}}$$

  is calculated for each data value d_i. $\Delta d_i = \frac{1}{\epsilon_i}(f(\mathbf{m})_i - d_i)$ is the data misfit vector of the current iteration. Together with the error weighting, the data weighting matrix $\mathbf{D^k}$ at iteration k can be calculated as

$$\begin{aligned}
w_i^0 &= \frac{1}{\epsilon_i} \\
w_i^k &= w_i^{k-1} \frac{1}{|\Delta d_i|} \frac{\sum_n \Delta d_n^2}{\sum_n |\Delta d_n|^2} \\
\mathbf{D^k} &= \mathrm{diag}(\mathbf{w}^k)
\end{aligned}$$

2.6.4 Equation system solvers

Each iteration step, the equation system (2.40) has to be solved. As the sensitivity matrix S is not sparse and mostly rather big ($\mathrm{S} \in \mathbb{R}^{M \times N}$ with the number of data N and the number of model parameters M), care has to be taken when solving, especially products like $\mathrm{S}^T\mathrm{S}$ should not be formed explicitly. Determining the exact solution is often too costly, and approximate Conjugate-Gradient-based algorithms (*Hestenes and Stiefel* [1952]) show the best result. An adapted conjugate gradient method (weighted least square conjugate gradients, CGLSCD) capable of solving equation of the type (2.40) with non-sparse matrices $\mathbf{S}$ is described in *Günther* [2004].

2.6.5 Resolution

The quality and resolution of an ERT survey depends on a number of factors – the used measurement array geometry, the model discretization and not at least also the inversion and regularization parameters. To quantitatively evaluate the quality of different configurations, resolution analysis can be applied. Linear resolution theory is widely discussed, e.g. *Menke* [1989], *Meju* [1994]. Non-linear resolution analysis for ERT has recently been investigated by e.g. *Friedel* [2003]. Again, a comprehensive overview can be found in *Günther* [2004].

Consider the data vector as a result of the true model's forward response and a noise vector n

$$\mathbf{d} = \mathbf{f}(\mathbf{m}^{\text{true}}) + \mathbf{n} \tag{2.42}$$

This equation is Taylor-linearized at a model iteration $\mathbf{m}^k$ close to the true model

$$\mathbf{d} \approx \mathbf{f}(\mathbf{m}^k) + \mathbf{S}(\mathbf{m}^{\text{true}} - \mathbf{m}^k) + \mathbf{m} \tag{2.43}$$

Next, generalized inverse are specified for the matrices $(\mathbf{DS})$ and $\mathbf{C}$,

$$(\mathbf{DS})^G = \left[(\mathbf{DS})^T(\mathbf{DS}) + \gamma\mathbf{C}^T\mathbf{C}\right]^{-1} (\mathbf{DS})^T \tag{2.44}$$

$$\mathbf{C}^G = \gamma \left[(\mathbf{DS})^T(\mathbf{DS}) + \gamma\mathbf{C}^T\mathbf{C}\right]^{-1} \mathbf{C}^T \tag{2.45}$$

so the inversion scheme (2.40) for a global regularization can be expressed as

$$\mathbf{m}^{k+1} = \mathbf{m}^k + (\mathbf{DS})^G\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}^k)) - \mathbf{C}^G\mathbf{C}(\mathbf{m}^k - \mathbf{m}_0) \tag{2.46}$$

Assuming $k+1$ to be the final step, and inserting (2.43) yields

$$\mathbf{m}^{\text{est}} = \mathbf{m}^k + (\mathbf{DS})^G\mathbf{SD}(\mathbf{m}^{\text{true}} - \mathbf{m}^k) - \mathbf{C}^G\mathbf{C}(\mathbf{m}^k - \mathbf{m}_0) + (\mathbf{DS})^G\mathbf{Dn} \tag{2.47}$$

$$= \mathbf{Rm}^{\text{true}} + (\mathbf{I} - \mathbf{R})\mathbf{m}_0 + (\mathbf{DS})^G\mathbf{Dn} \qquad . \tag{2.48}$$

Here, $\mathbf{R} = (\mathbf{DS})^G(\mathbf{DS})$ is the resolution matrix. For a perfect resolution, $\mathbf{R}$ would be the identity matrix. The more $\mathbf{R}$ is localized in the diagonals, the better is the correlation between $\mathbf{m}^{\text{true}}$ and $\mathbf{m}^{\text{est}}$. Therefore, the definition of an information content

$$IC = \sum_{i=1}^{M} R_{ii} \tag{2.49}$$

makes sense. Divided by the number of measurements N an information efficiency can be introduced

$$IE = \frac{IC}{N} \tag{2.50}$$

Similarly, *Friedel* [2003] motivated the concept of the model cell resolution radii, which quantifies the resolution radius of a cell i of volume ΔV_i as

$$res_i = \sqrt[3]{\frac{3\Delta V_i}{4\pi R_{ii}}} \tag{2.51}$$

The other information to obtain from (2.47) is that in areas of low resolution $(\mathbf{I} - \mathbf{R})$, the model is strongly determined by the starting model. This explains the strong influence of the starting model in the lower model section. Numerically, information content can be calculated more efficiently by the means of Singular value decomposition (SVD). This approach is described in *Günther* [2004].

2.7 Related theory

For my experiments in the context of hydrogeophysics, not only Geoelectrics, but also other geophysical methods are used. Also, an understanding of fluid dynamics and solute transport in soil systems is needed. I will therefore present a basic introduction to some methods I used, as well as provide the central equations. However, since the focus of my work lies on Geoelectrics, a detailed description would be beyond the scope of this thesis.
At the end of each section I will give references to more detailed summaries to these topics.

2.7.1 Fluid dynamics in porous media

Soil is essentially a porous media, with its fluid dynamics governed by the pore space geometry. However, since this geometry is structured at different scales from micrometer to centimeter scale for soils, a direction representation of the geometry is nearly impossible. Therefore, a simple macroscopic description is sought. This is achieved by averaging state parameters over a volume ΔV. For example, the macroscopic water content is defined as

$$\theta_w := \frac{\Delta V_w}{\Delta V} \quad ,$$

where ΔV_w is volume of the water phase inside ΔV. ΔV has to be chosen large enough to include all scales, so that the averaged parameters are stable under enlargement of the volume. Such a volume is a called a Representative Elementary Volume (REV), and its existence is vital for a macroscopic description of the system. The potential energy to move a fluid (here water is assumed) from reference position z_0 with pressure p_0 to the state at position r is (for the isothermal case)

$$\psi_w(\mathbf{r}) = \underbrace{p_w(\mathbf{r}) - p_0}_{\psi_p} - \underbrace{\int_{z_0}^{z} \rho_w(\xi) g \,\mathrm{d}\xi}_{\psi_m} \tag{2.52}$$

with the density ρ_w and gravitational constant g. ψ_p, which is the potential induced by pressure difference, is linear for incompressible materials : $\psi_p(\mathbf{r}) = (z - z_0)\rho g$. The gravitational part ψ_m is called matric potential, and dependant on the soil matrix structure.

Darcy's law Generally, fluid dynamics can be described by the Navier-Stokes equation. Under the prerequisite of slow external forcing (stationary flow), small flux (neglectable inertia), small dimensions of the system (gravity neglectable to viscosity) and incompressible Newtonian fluids, Darcy's law

$$\mathbf{j} = -\mathbf{K}\boldsymbol{\nabla} p$$

can be deduced from the Navier-Stokes equation. $\mathbf{j}$ is the flux, and $\mathbf{K}$ the hydraulic conductivity tensor. $\mathbf{K}$ reduces to a scalar for homogeneous media. When considering more than one fluid phase, Buckingham's empirical extension to Darcy's law can be applied. For a fluid i it states

$$\mathbf{j_i} = -\mathbf{K_i}(\theta_i)\boldsymbol{\nabla}\psi_i \tag{2.53}$$

which is equivalent to Darcy's law with an dependency of $\mathbf{K}$ of the fluid content θ_i. As the conductivity $\mathbf{K}$ depends on the flow channel radius ($j \propto r^2$), it rapidly decreases with decreasing fluid content.

Richards equation and parametrization Combining (2.53), (2.52) with mass balance law $\frac{\partial}{\partial t}\theta_w + \boldsymbol{\nabla}\mathbf{j_w} = 0$ yields Richards equation

$$\frac{\partial\theta_w}{\partial\psi_m}\frac{\partial\psi_m}{\partial t} - \boldsymbol{\nabla}\left(\mathbf{K}(\psi_m)(\boldsymbol{\nabla}\psi_m - \rho_w\mathbf{g})\right) = 0 \tag{2.54}$$

Most of the hydraulic simulation code for water movement in soils today bases on Richards equation. As the functional relations $\theta_w(\psi_m)$ and $\mathbf{K}(\psi_m)$ can not be deduced by first principles, a parametrization is sought, which can be fit to experimental data. Most commonly, the w3van Genuchten parametrization is used *van Genuchten* [1980]

$$\Theta(\psi_m) = \left(1 + (\alpha\frac{\psi_m}{\rho_m g})^n\right)^{-m} \tag{2.55}$$

with the parameters $\alpha > 0$, $n > 1$, $m > 0$. $\Theta = \frac{\theta_w - \theta_R}{\theta_S - \theta_R}$ is the water saturation, with the saturated water content θ_S (which is equivalent to porosity) and the residual water content θ_R. *Mualem* [1976] uses connected stacks of capillary bundles as a soil model for description of the conductivity. Together with (2.55), it forms the Mualem-van-Genuchten equation

$$K(\Theta) = K_0\Theta^\alpha\left[1 - \left(1 - \Theta^{\frac{n}{n-1}}\right)^{1-\frac{1}{n}}\right]^2$$

A discussion in detail is given in *Roth* [2006].

2.7.2 Solute transport

We are now looking at transport of solutes in a porous media. The mechanics of convection and dispersion for solute transport can be derived at pore-scale, and then again be averaged over an REV to obtain a macroscopic continuum description. The parameter of interest here is the averaged concentration C_w of a given solute in water

$$C_w = \frac{\Delta M}{\Delta V_w}$$

with the solute mass ΔM and the REV water volume ΔV_w. In the following, a steady flow field in an uniform porous media is assumed. Generally, the solute transport in porous media at pore scale level is determined by convection, molecular dispersion, Taylor-Aris dispersion and hydrodynamic dispersion. At macroscopic level, the flux law for the solute flux $\mathbf{j_s}$ is of the form

$$\mathbf{j_s} = \theta \mathcal{V} C_w \qquad ,$$

where $\mathcal{V}$ is the generalized velocity operator. Basically two transport models can be identified:

- **The Stochastic Convection (SC) model** assumes that the macroscopic dispersion is only due to the flow field variability, hydrodynamic dispersion is neglected. A solute particle entering the flow field (e.g. represented by stream tubes) keeps a constant velocity. The flux law in SC reads

  $$\mathbf{j_s} = \theta \mathbf{v} C_w \tag{2.56}$$

  where $\mathbf{v}$ is the velocity vector. In this approximation, the variance of transport distance behaves as $\mathbf{var}(z) \propto t^2$. SC is a good approximation only for the near field, i.e. short transport times.

- **The convection dispersion (CD) model** is represented similarly, with the difference that the solute particle may change channels in the velocity field. The particles are assumed to "see" the complete flow field, therefore CD is only reasonable in the far field, i.e. long transport times. The flux law consists of a convective and dispersive part,

  $$\mathbf{j_s} = \mathbf{j_w} C_w - \theta \mathbf{D}_{\text{eff}} \boldsymbol{\nabla} C_w \qquad . \tag{2.57}$$

  The effective dispersion coefficient $\mathbf{D}_{\text{eff}}$ includes molecular and hydrodynamic dispersion. (2.57) can be combined with mass conservation law to yield the Convection Dispersion Equation (CDE)

  $$\frac{\partial}{\partial t}(\theta C_w) + \nabla(\theta \mathbf{v} C_w) - \nabla(\theta \mathbf{D}_{\text{eff}} \boldsymbol{\nabla} C_w) \tag{2.58}$$

  In the CD regime, $\mathbf{var}(z) \propto t$ holds for the variance of the transport distance. An extension to the CD model is the Mobile-Immobile-Model (MIM), which takes into account that velocities in big inter-aggregate pores are considerably faster than the inner-aggregate velocities. The fluid phase is therefore divided into a mobile and an immobile phase.

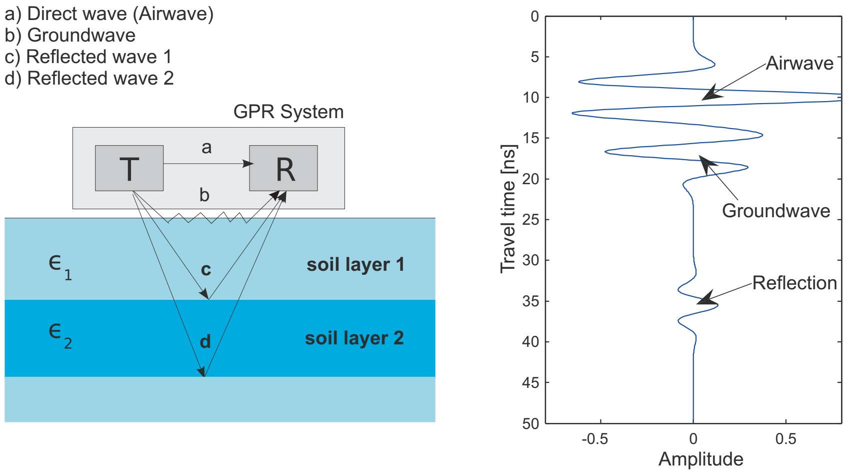


Figure 2.5: (left)GPR system with different categories of reflections, (right) example synthetic GPR trace

When considering real soil systems, the steady flow condition is almost never met. In these cases, the fluid dynamics are calculated by Richards equation (2.54). In the regime of long distances, the CDE can be applied as well (with the local flow field) to calculate the solute dynamics. However, $\mathbf{D}_{\text{eff}}$ is generally a function of water flux and water content here. It is usually written as $\mathbf{D}_{\text{eff}} = \theta\mathbf{D_0}(\theta) + \lambda(\theta)\mathbf{j_w}$. $\mathbf{D_0}(\theta)$ is the molecur dispersion coefficient, and $\lambda(\theta)$ the materials' dispersivity.

Again, *Roth* [2006] and *Vogel* [2004] may be referred to for a more complete treatment of the topic.

2.7.3 Electromagnetic techniques

In my experiments, ERT data was compared with data from Ground Penetrating Radar (GPR) and Time Domain Reflectometry (TDR) probe measurements. The parameters of interest were the volumetric water content θ, the electrical conductivity σ and structural information on the soil layers. Both GPR and TDR are electromagnetic techniques, and allow estimation of the bulk dielectric permittivity ϵ_B. This parameter is of interest, as it is strongly influenced by the soil water content.

Dielectric permittivity The bulk permittivity of soil as a 3-phase system can expressed by the CRIM equation *Roth et al.*

$$\epsilon_B = [\theta\epsilon_w^\alpha + (1 - \theta_S)\epsilon_s^\alpha + (\theta_S - \theta)\epsilon_a^\alpha]^{\frac{1}{\alpha}} \tag{2.59}$$

where ϵ_a,ϵ_w,ϵ_s denote the permittivity of the air phase, liquid phase and soil matrix respectively. α is a factor on the orientation of the electric field ($\alpha = 0.5$ for an isotropic medium) while θ_S is the soil porosity. Since permittivity of air ($\epsilon_a \approx 1$) and rocks ($\epsilon_s \approx 3 \dots 10$) are rather constant and small compared to the permittivity of water ($\epsilon_w \approx 80$), (2.59) can be used to determine the water content, if θ_S and α are known.

Ground-penetrating radar Ground-penetrating radar (GPR) is a geophysical method that can be used to determine subsurface structure as well as the dielectric permittivity ϵ. A GPR system consists of at least one transmitter and one receiver antenna, which is moved on the ground (see fig. 2.5). The transmitter circuit produces a short high voltage pulse which is radiated into the ground via the transmitter antenna. The reflected signal is then received by the receiver antenna, amplified and passed to the control unit.

For GPR, usually frequencies in the range of 10–1000 MHz are used, depending on the application. The choice of center frequency and band width influences the spatial resolution and penetration depth. Generally, higher frequencies allow for a higher spatial resolution while offering smaller penetration depth. The ground material is also an important factor determining penetration depth, as the radiation is for example absorbed by conducting material (salty soil etc.).

The measured signal is voltage as a function of the signal travel time (see fig 2.5). Since there are many possible paths for waves from transmitter to the receiver, the observed voltage peaks have to be associated with the origin of the wave. There is usually a direct wave, arising from the straight connection between transmitter and receiver, and a ground wave, coupled with the ground surface (see fig 2.5). There are also reflected waves at different soil layers, as well as multiple reflected waves (also called ringing, not shown in the figure).
The reflected waves contain information on reflector depth as well as the dielectric permittivity ϵ of the traversed layers, as for the velocity v applies

$$v = \frac{c_0}{\sqrt{\epsilon}} \tag{2.60}$$

(at least for low-loss geological materials). The ground wave holds information about ϵ at the ground/air interface.

We have to note that the depth of objects and layers can only be obtained at interfaces with a large enough dielectric permittivity contrast. For a normal incident signal at a layer boundary, with permittivities ϵ_1 and ϵ_2 respectively, the reflection coefficient is defined by Fresnel's law as

$$R = \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2} \quad .$$

Using (2.60), a relation of reflector depth and permittivity is obtained, so assumptions about one of these variables have to be made in order to obtain the other. This limitation can be avoided by CMP or multichannel measurements.
The idea with both methods is that by measuring at least two linearly independent dataset per position, the equation can be solved for both unknowns. Linearly independent data is obtained by different reflection angles, i.e. different antenna separations. In CMP, the point to measure, the midpoint between the sender and receiver, remains constants while both antennas are moved in opposite directions. While this provides a continuous range of reflection angles, CMP is obviously very laborious to conduct for more than a few data points.
Multichannel measurements on the other hand make use of multiple senders and/or receivers[7]. The distances between the different sender/receiver signal paths are inherently

---

[7] Usually, one sender and receiver antenna are fixed in an antenna system, so at least two of them would have to be used.

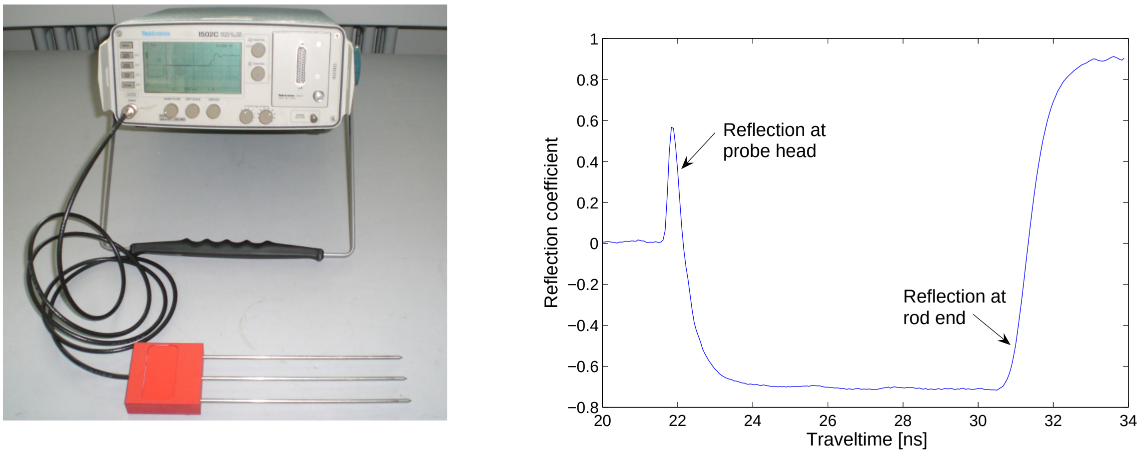


Figure 2.6: 3-rod TDR probe and Tektronics cable tester (left) and time domain signal for a probe in water (right)

different, so with each data sample taken, multiple angles are covered. The complete system of antennas, with fixed separations, can be pulled over the ground recording a continuous profile, as in the case of a single channel measurement. A picture of our multichannel GPR setup is shown in the appendix (fig. 2).

By this ray-based approach, for each layer only a mean permittivity can be obtained in z-direction. To obtain the spatial distribution of ϵ, forward modeling together with inversion techniques might be used. This approach has however been only successfully applied so far to straight-line radar measurements (A fullwave inversion of borehole-borehole radar was achieved by *Ernst et al.* [2007]).

A more detailled introduction to Georadar in soil physics applications can be found in *Huisman et al.* [2003],*Davis and Annan* [1998].

Time Domain Reflectometry Time Domain Reflectometry (TDR) is another way of determining the bulk dielectric permittivity and also the electric conductivity of soils. It again uses the fact that the propagation speed v of electromagnetic waves depends on the mediums permittivity (2.60). A TDR probe consists of two or three parallel rods of length l, which form a waveguide (fig. 2.6). The TDR measurement instrument, for example a cable tester, generates a pulse of voltage U_0, which is transmitted through the probe. The end of the rods poses an impedance jump, which causes a reflection. Reflections also occur at the interface between the coaxial cable and the probe rods. The signal travel time on the rods is therefore

$$t = \frac{2l\sqrt{\epsilon}}{c} \qquad ,$$

so measuring the travel time between head reflection and rod end reflection allows the determination of the permittivity surrounding the probe. A measured time domain reflection signal $\zeta(t) = \frac{U(t)}{U_0}$ (TDR trace) is shown in fig. 2.6.

On the other hand, is has been shown that electrical conductivity σ can be obtained by

$$\sigma = \frac{K_p}{Z_c}\frac{1-\zeta(\infty)}{1+\zeta(\infty)}$$

where $\zeta(\infty)$ is the reflection coefficient in the limit of long travel times, Z_c is characteristic impedance of the measurement system and K_p is the probes' cell constant. K_p depends on the probe geometry and can be calculated, but is best obtained by calibration.

The quality of these measurement values depends on the quality of the TDR trace. For example, since a high electrical conductivity in the soil lowers the reflection coefficient, determination of the reflection point at the end of the rods gets harder. Therefore, permittivity can not be determined accurately for high conductivities.

You might refer to *Robinson et al.* [2003] and *Heimovaara et al.* [1995] for a detailed discussion on TDR.

2.8 Hydrogeophysics

Introduction In hydrology and soil physics, an understanding of water and solute dynamics in porous media is sought. Hence, monitoring the spatial and temporal variation of state variables like temperature, water content θ or solute concentration c is necessary. In the field, this is achieved traditionally by either applying tracers and probing at different locations and times, or by instrumenting a profile with sensors. However, these methods only provide point information in time and space, and are invasive, so the system might get disturbed. On the other hand, geophysical methods are mostly non-invasive, and can provide spatially resolved data. The application of geophysical methods to the monitoring of transport processes in soils therefore stands to reason. This emergent field is called hydrogeophysics. Here, I will concentrate on a description of water and solute transport dynamics at the field scale.

Methods The geophysical and classical monitoring methods can be classified into three categories

- **Point measurements :** Instrumented boreholes or profiles constitute this category. The sensors will usually be temperate sensors, TDR probes for water and solute content or tensiometers for pressure head measurement. Once installed, they can provide with hard data in high temporal resolution, but they represent only one point in space, and care has to be taken to limit disturbances of the natural flow regime during installation.

- **Large-scale measurements :** Geophysical methods like Geoelectrics, Seismics and Georadar fall in this category. With the exception of borehole techniques they are non-invasive. While these methods can provide a spatially resolved image of the subsurface, state variables are determined indirectly (“soft” data). For example, Georadar data can be related to water content, but only the mean water content between reflectors can be obtained directly (see chapter 2.7.3). ERT on the other hand is limited by the uncertainties introduced in the inversion process. The temporal resolution is limited by the measurements duration (especially with ERT) and the manual effort to conduct the measurements (especially with GPR).

- **Sampling** The classical method to determine soil characteristics is sampling. These soil samples from excavations can be examined in lab experiments to obtain soil characteristics. Often, an undisturbed samples have to be taken, which is a time-consuming task. Also, extensive excavation will distort the soil system.

The challenge in hydrogeophysics is therefore to take advantage of the strengths of the different methods, and obtain a consistent data set of optimal resolution.

Modeling Even with geophysical methods, a direct observation of transport processes is often not possible, as the processes of interest are often on a very small scale or happen too fast for the monitoring. Also, not only the dynamics of observable state variables θ, C_w for a specific setting are of interest, but also characteristic values for the soil system are sought, which would allow to transfer results to other settings. Here, numerical modeling can be employed. The structure of the soil system is approximated by a model, and the transport described by a set of dynamic equations. Instead of measuring the state variables directly, parameters to the dynamic equations are determined. The state variables are then predicted by numerical forward solving in high resolution. The quality of this prediction obviously depends on the applicability of the dynamics model, as well as the accuracy of the measured parameters.

A numerical transport model consists of several parts

- **Governing equations** An exact description of the transport dynamics can be arbitrarily complex. Therefore, a simplified model is used, which consists of a set of partial differential equations. For water transport at the field scale, most commonly Richards equation (see chapter 2.7.1) is employed, while solute transport is described usually by the CD model (see chapter 2.7.2). The limitations of these equations (e.g. short transport distances for CDE) have to be considered.

- **Structural model** The heterogeneity of the soil system is approximated by dividing the model into several sections. In each section, the material is assumed to be homogeneous, so the differential equations are defined with an individual set of parameters in each section. These sections can be e.g. different soil and geologic layers.

- **Parameters** For a Richards model the parameters consist of the Mualem-van-Genuchten parameters. When coupled with CDE, the dispersivity λ is required as an additional parameter. Some of these parameters are rather easy to determine, e.g. the soil porosity can be derived easily by weighing of a dried soil sample. Others, as the hydraulic conductivity, are hard to derive directly, at least for the vadose zone. Theoretically all parameters can be obtained by Multistep-outflow (MSO) experiments (*Schneider* [2005]) on an undisturbed REV soil sample. However, the extraction of an undisturbed REV sample might not be possible in all cases, boundary effects are neglected, and the extraction of undisturbed samples is very time-consuming.

- **Boundary conditions** As boundary conditions for a Richards/CD model, an initial state for the water content and solute concentration as well as the precipitation over time are needed.

Given the structural description of the field site (obtained by geophysical investigations or probing), the material parameters (from soil samples, etc.) and boundary conditions (e.g. geophysics/TDR for initial condition, rain gauge), simulations can be run. The results can be validated by point measurements or geophysical methods.

Inverse modeling While modeling can produce high-resolution visualization of transport processes, its quality depends on the accuracy of the material model. As pointed out, some of these parameters are hard to obtain by direct measurements. However, the state variables themselves are accessible through various measurements, and can be used to verify simulation results. This process can be automatized by an inversion. Here, starting parameters are obtained as above. Then a modeling step (forward calculation) is conducted, providing a spatial and temporal map of $\theta, C_w, \ldots$. The resulting data is checked against measurement data. This can be done efficiently for hard data, e.g. point TDR data, as shown by *Ritter et al.* [2003]. For soft data as ERT, more advanced methods have to be sought, see chapter 3.2. As with the ERT inversion, an update of the material parameters is calculated based on the difference gradient, and iterated until a threshold is met. In contrast to ERT, the number of parameters to determine is much lower, as only material parameters are inverted, not a whole spatial model. Therefore, a simple inversion without constraints is sufficient to the problem.

The process is explained in more detail in chapter 3.2.

Chapter 3

Concepts and Implementation

In this section, new concepts developed for this thesis are presented, and the software implementation of various algorithms is described.

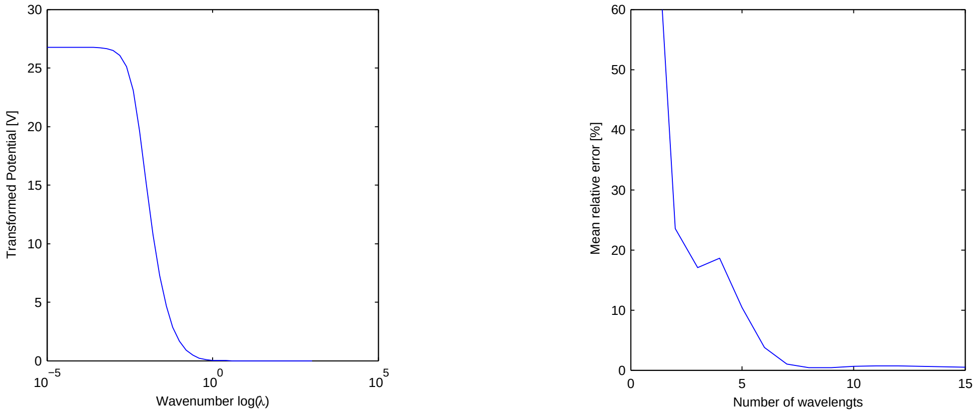


Figure 3.1: Left: Transferred potential in logarithmic λ-space (example). Right: the relative potential error as a function of the wavenumbers used in the backtransformation

3.1 ERT evaluation software

3.1.1 Introduction

The are a number of commercially available software packages for ERT, the most well-known being the Res2dInv program by M.Loke. For routine processing of data sets, usage of one of these packages might be the easiest solution. In this work, however, a direct access to the solver was required. On the one hand, it was necessary to batch process test series, implement additional features and provide direct data exchange with other tools. On the other hand, ERT results strongly depend on the inversion properties, due to the ambiguity of the inverse process. Therefore, an exact understanding what steps the inversion and forward solver are performing was needed, which is difficult without source code access.

For this purpose, an ERT evaluation toolkit was developed in the course of this thesis. A forward solver and several tools and routines were implemented – In this chapter, some results will be briefly presented. At the time of developing the inversion code, access to a sophisticated inversion package and its source code was obtained by an agreement of cooperation (see 3.1.4).

3.1.2 Forward solver

Implementation A forward solver (see chapter 2.4 for the scheme) for 3D and 2.5D was implemented in MATLAB. It supports singularity removal and the boundary conditions mentioned in 2.4. As shown in the next paragraph, singularity removal together with Dirichlet boundary conditions and grid prolonging shows the best results and will be used as a standard. I will focus on the 2.5D solver here, as the 3D case is an easier subproblem. Using the scheme (2.21), the forward operator matrix $\mathbf{L}$ has to be build twice per wavenumber. To build $\mathbf{L}$, the C-coefficients for all cells are calculated vectorized in order to avoid for-loops, which perform bad in interpreted languages like MATLAB. First the inner body of the model space is constructed this way, then boundaries are applied. The sparse forward matrix is then constructed efficiently from the coefficient vector using the MATLAB command `spdiags`. The most time-consuming step is the calculation of the Bessel functions for the primary potential for each wavenumber. However, as it also depends on the model grid and the wavenumber λ, it can be stored and reused when recalculating, e.g. each step of an inversion. The same holds for $\mathbf{L_1}$, so that only one forward matrix has to be created after the first calculation. Decomposition of this matrix is also limited to the first run. The linear equation system is solved by decomposition using MATLABs Cholesky routine and backsubstitution. Finally, backtransformation (2.24) is performed using Gauss-quadrature. As can be seen in fig. 3.1a, the potential is hard to numerically integrate, as it very steep for small wavenumbers. For the steep part, Gauss-Legendre integration is used, while the tail is integrated by Gauss-Laguerre. The transition point between these two methods is approximated as

$$t = \frac{1}{2d} \qquad ,$$

where d is the smallest spacing in the model grid. For both integration schemes a number of wavelengths n_λ is used. Fig. 3.1b shows the backintegration quality for different numbers n_λ. Here, the mean relative surface potential error of a vertical two-layered model is given[1]. $n_\lambda = 8$ is chosen as standard parameter, as for higher numbers no significant improvement is reached.

Performance The quality of the solver for different parameters was tested. Here, a vertical two-layered model with a resistivity contrast of 1:10 and a boundary at 1m was used for testing. A current of 1A is injected at x=11m. A lattice constant of 10cm was used with

---

[1] The layers have a resistivity contrast of 1:100, the boundary is located at 1m. (2.11) was integrated, while (2.9) was taken as real space reference

a (25m × 3m) core model, with 5 logarithmically prolonged cells added at the boundaries. The potential field with and without singularity removal is shown in fig. 3.2.

The semi-analytical potential (2.9) is calculated and compared with forward calculations as above. The relative potential error at the surface is shown in fig. 3.3. It can be seen that the calculation is very sensitive to these parameters. Dirichlet boundary conditions together with singularity removal on a prolonged grid pose the best option.

On a 1GHz machine, calculation for a model of 400 × 100 cells takes 1s per wavenumber without, and 6s/3s for the first/subsequent runs with singularity removal.

3.1.3 Sensitivities

Routines to calculate 2D/3D sensitivities have also been implemented in MATLAB. Homogeneous half-space sensitivities are only dependant on the model geometry and can thus be precalculated. Therefore, accuracy, not speed is desired here. The analytic expressions (2.33) or (2.32) are numerically integrated over the single cells of the model grid by MATLABs adaptive quadrature routine to an accuracy of 10^{-6}. A pole-pole sensitivity diagram calculated this way is shown in fig. 3.4(a).

For calculating sensitivities for a given resistivity model, a routine for numerical approximation based on forward potentials (see 2.5.3) is implemented. A comparison of the analytic sensitivity to the numerically calculated homogeneous sensitivity at the surface is shown in fig. 3.4b. A grid with lattice constant of 2.5cm is used for the calculation, the mean relative error lies at around 5%.

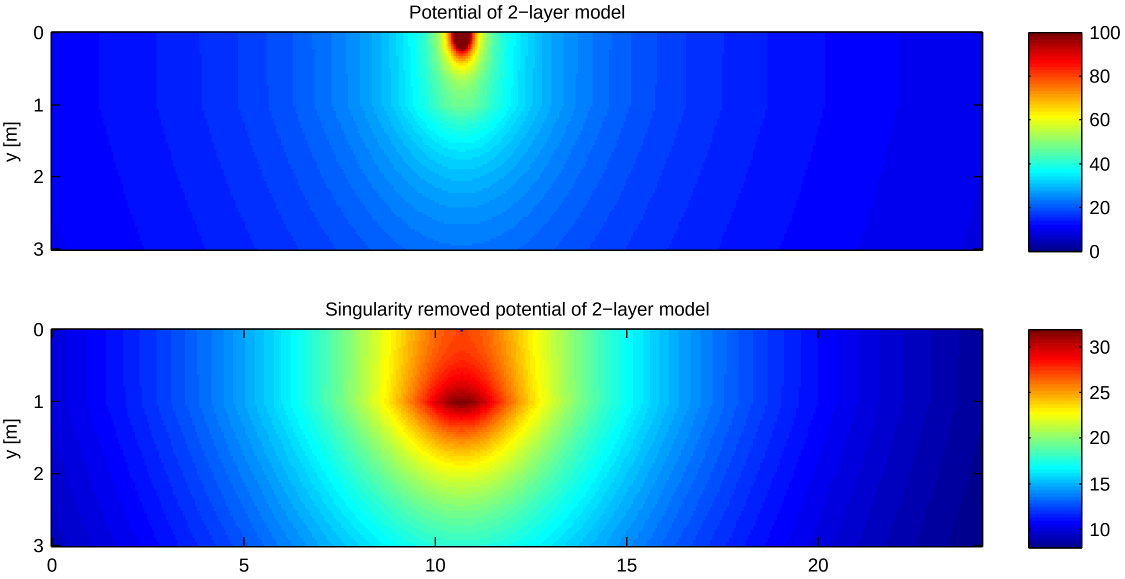


Figure 3.2: The potential field of a vertical two-layered model, with and without singularity removal (in Volts)

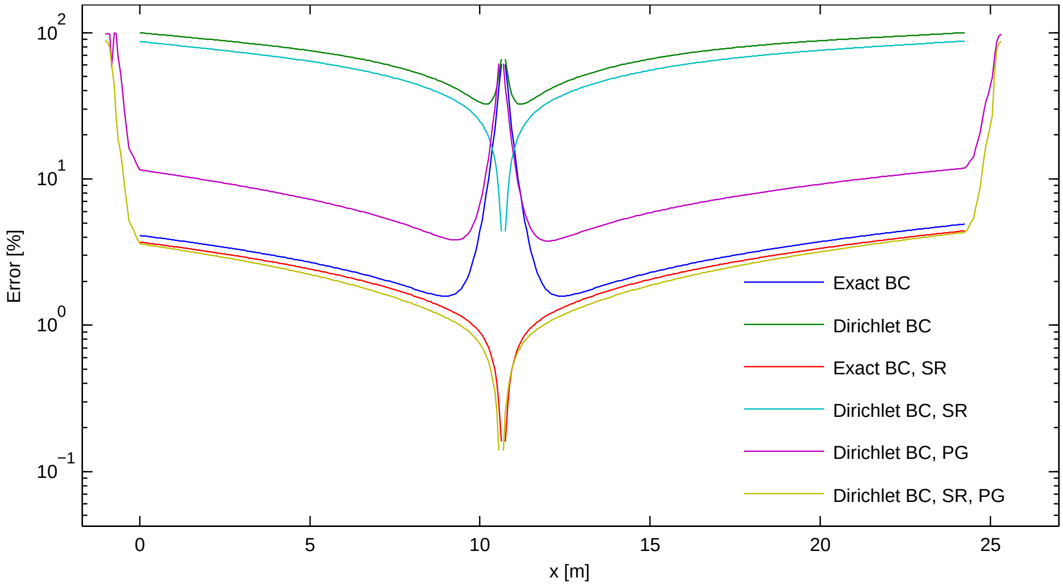


Figure 3.3: Surface potential deviation for different boundary conditions (BC), with/without prolonged grid (PG), singularity removal (SR)

3.1.4 Inversion

While at first the development of our own inversion software was planned, it seemed more efficient at some point to join another code basis. While the implementation of a pure Gauss-Newton inversion routine (see chapter 2.6.1) is rather straightforward, the amount of time needed for testing and tweaking to get it running as stable and fast as established software seemed unreasonable. Therefore, a cooperation with Thomas Günther, the author of the sophisticated inversion package Dc2dInvRes, was initiated, in which T. Günter kindly granted full source code access. Dc2dInvRes is written mainly in MATLAB[2], so the connection with the tools and routines developed in this thesis were straightforward. Apart from Gauss-Newton inversion, it also supports SIRT, truncated SVD and conjugate-gradient based inversion as well as various regularization schemes.

3.1.5 Optimal acquisition

Introduction

In order to maximize the information about the subsurface gained in an ERT survey, both data processing and data acquisition are subject to optimization. While a wide range of publications studied properties of data processing, i.e. the forward solver and inversion scheme, only recently the optimization of data acquisition has gained some attention. *Friedel* [2003], showed the application of nonlinear resolution theory to ERT data, and *Stummer et al.* [2004] provides an algorithm for obtaining arrays with optimal resolution properties. As these publications, I will consider optimal acquisition in terms of determining measurement arrays

---

[2] Some tasks are performed by external C and Fortran libraries for better performance

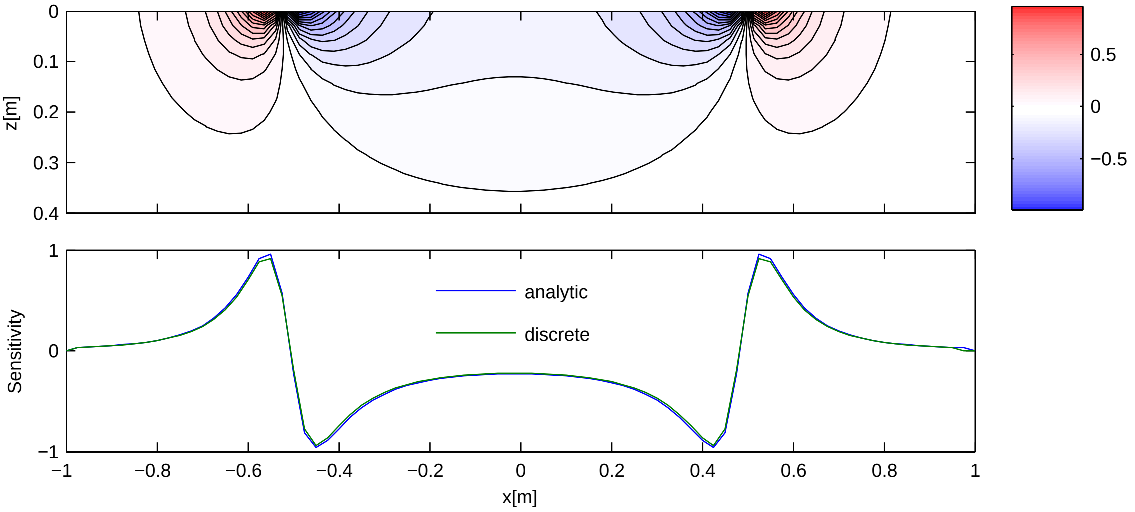


Figure 3.4: Top: Analytical pole-pole sensitivity. Bottom: Comparison of analytical and numerical sensitivity

with optimal resolution for a given set of electrodes. The application will be a equidistantly spaced 2D surface profile, but the theory is not limited to this case.

The aim was to find an optimal acquisition scheme for the Grenzhof tracer experiment (see ch. 4). Here, not a global optimal resolution, but a maximum spatial resolution in the area of tracer application was required, which covers 2 m of the 24 m profile. Also, in contrast to many other applications of ERT, a shallow investigation depth is sufficient, as the tracer propagates into the ground slowly. Thus, I will present a simple approach to focus resolution in an area of interest.

Techniques for resolution focusing

The model resolution, i.e. the elements R_{ii} of the resolution matrix (see chapter 2.6.5) is the best indicator for spatial information content of a given configuration. The optimal approach therefore would obviously be to evaluate the resolution matrix for all possible combinations of four-electrode measurements, and chose the one with

$$IC_f = \sum_{i=1}^{M} A_i R_{ii} = max$$

where A_i is the area of interest, which specifies a weight $0 \ldots 1$ for each model cell. In the case discussed here, setting a weight of 1 in the tracer area and 0 elsewhere is sufficient. As the calculation of R is very time consuming, it is impossible to evaluate for all combinations of four-electrode measurements. This motivated *Stummer et al.* [2004] to develop a heuristic for the approximate determination of the optimum, without a direct calculation of the resolution matrix for every combination. Starting from a dipole-dipole base set, they evaluate all possible measurements i to add to the set by the goodness function

$$GF(i) = \sum_{j=1}^{M} \frac{|S_{ij}^{\mathrm{add}}|}{S_j^{\mathrm{sum}}} \left(1 - \frac{R_{jj}^{\mathrm{base}}}{R_{jj}^{\mathrm{com}}}\right) \tag{3.1}$$

with the sensitivity normalization term

$$S_j^{\mathrm{sum}} = \frac{1}{N} \sum_{i=1}^{N} |S_{ij}^{\mathrm{com}}| \quad .$$

Here, S_{ij}^{com} refers to the complete set of measurements, R_{jj}^{base} to the base and previously chosen measurements and S_{ij}^{add} to the part of the complete set not yet chosen. The idea is to add measurements in each step which focus their sensitivity in areas where the resolution is still low. This process is iterated until the desired number of measurements is reached. The method could easily be adapted to focusing by introducting a vector A_i as above in the goodness function. However, while this approach works fine for 30 electrodes as demonstrated by the article, it is still too computationally expensive for the 96 electrode system used in the experiment[3]. Therefore, other indicators have to be found.

For a single measurement, the model cell resolution is basically governed by the absolute sensitivity of that measurement. Therefore, taking the sensitivity coverage (sum of the normalized sensitivities) in the region of interest as an indicator for the resolution suggests itself.

$$\tau = \sum_{i=1}^{M} \sum_{j=1}^{N} A_i \frac{|S_{ij}|}{S_j^{\mathrm{sum}}} \quad . \tag{3.2}$$

An optimization can be done like above by iteratively adding measurements to a base pool, maximizing τ. However, this tends to produce an unwanted inhomogeneous sensitivity coverage. A way to circumvent this problem is to replace the linear summation by a function which favors the first increases, so that coverage is raised simultaneously over all cells. This can be e.g. the logarithm function,

$$\tau^m = \sum_{i=1}^{M} \log \left(\sum_{j=1}^{N} A_i \frac{|S_{ij}|}{S_j^{\mathrm{sum}}} \right) \quad . \tag{3.3}$$

Validation

In order to show that τ as defined above is a good indicator for the resolution, a simulation has been run. To be able to compare the results with *Stummer et al.* [2004], a setting of 30 electrodes with 5 m spacing have been used. The discretization grid is also chosen similar, with 30x21 cells up to a depth of 30 m of homogeneous material. For different arrays, the coverage and the resolution matrix have been calculated. For the comprehensive array, the

---

[3] 96 electrodes ⇒ 79727040 possible measurements, whereas 30 electrodes ⇒ 657720 possible measurements. The evaluation of $>$ 1000000 measurements as well as the calculation of the complete resolution matrix is not feasible for memory and CPU time complexity.

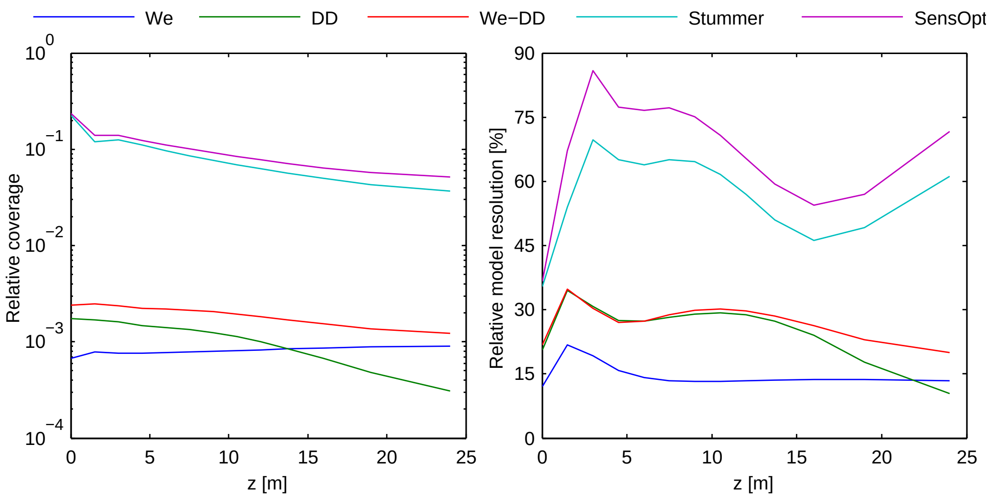


Figure 3.5: The logarithmic relative coverage (left), and relative model resolution (right), horizontal mean, for Wenner, standard Dipole, combined Wenner-Dipole and two 670-point optimized arrays. 30 electrodes, 1m spacing.

same restrictions have been imposed. Fig. 3.5 shows the horizontal coverage and resolution mean for different arrays. It can be seen that the arrays with higher model resolution also provide higher sensitivity coverage, so coverage as an indicator is able to distinguish arrays with good from arrays with bad resolution in this case.

The figure also shows the performance of the optimization algorithm. As in *Stummer et al.* [2004], the standard dipole-dipole set has been taken as a base set. Then, iteratively measurements are added to the set until 670 data points are reached. It can be seen that in this configuration, the array even has slightly higher model resolution than the array given in *Stummer et al.* [2004]. In another configuration tested, the performance was slightly below the Stummer array. The same result was obtained for 20 and 40 electrodes.
In short, it can be found that the coverage-based optimization is comparable to the Stummer method, at least for small numbers of electrodes where a direct comparison is possible.

Implementation for a large set

In this chapter, the implementation of the optimization algorithm is presented, and special strategies to deal with the large number of possible measurements, induced by the large number of electrodes used, are discussed.

Reducing the comprehensive set Even for the computationally simple approach used here, the high number of possible measurements poses a problem. The comprehensive set can be reduced significantly without significant loss of resolution by the application of the following concepts :

- First of all, measurements with the two current electrodes switched carry the same information; the same is true for the reception electrodes. Additionally, the result is

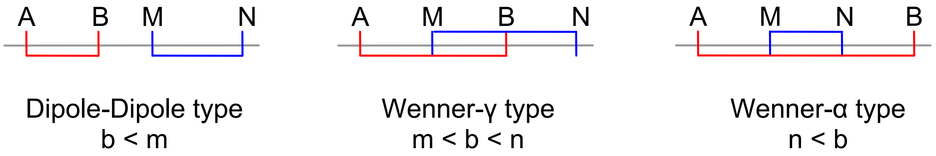


Figure 3.6: Three categories of measurements. Only the permutation a¡b, m¡n, b¡m is considered here.

invariant under exchange of the sender and receiver dipole, as stated by the reciprocity principle *Park and Van* [1991]. Thus only one electrode permutation (here: a¡b, m¡n, b¡m) has to be included in the comprehensive set.

- Measurements can be characterized into generalized Dipole-Dipole, Wenner-α and Wenner-γ type groups, which as illustrated in fig. 3.6. The overlapping Wenner-γ configurations are generally unwanted, as they may destabilize the inversion. These measurements can be excluded as well.
- Measurements with a large configuration factor are also not endorsed, as these measurement show small voltage differences and are therefore more susceptible to noise[4]. Here, all measurements with an geometric factor $k > 1000\,a$ are excluded, which is equivalent to $n > 6$ for Dipole-Dipole arrays.

By performing these reductions, the comprehensive set can be reduced from 79727040 to 5488721 measurements (factor 14).

Implementation The optimization algorithm is an iterative (greedy) optimizer, basing on equation 3.3 as rating function. The following steps are executed :

(i) $\Gamma := \Gamma_0$

(ii) $\Lambda := \Lambda_0 \setminus \Gamma_0$

(iii) for every $\lambda \in \Lambda$

  (a) $S_\Gamma := S(\Gamma \cup \lambda)$

  (b) $sc(\lambda) := \tau^m(S_\Gamma)$

(iv) $\Gamma := \Gamma \cup \{\lambda \in \Lambda \| sc(\lambda) = max\}$

(v) if num˙el(Γ) $< n_f$, go to 2.

Λ_0 is the reduced comprehensive set obtained as in the previous paragraph, while Γ_0 is the base set. The algorithm calculates τ^m for every possible addition to the base set, and adds the element with the best score to the base set, until a number of n_f is met. The problematic step here is the sensitivity calculation in step 4, which is very time consuming. On the

---

[4] This is the reason for the common advice that values of $n > 6$ should not be used for standard Dipole-Dipole arrays.

other hand, a precalculation for the complete comprehensive set is not possible for memory reasons. However, if homogeneous sensitivities are used, a precalculation of the pole-pole sensitivities for all possible electrode distances on a prolonged grid is sufficient. Only one half of the result has to be saved because pole-pole sensitivities are symmetrical. In step 4, the sensitivities can be easily constructed by horizontally shifting and superpositioning of the precalculated sensitivity images..

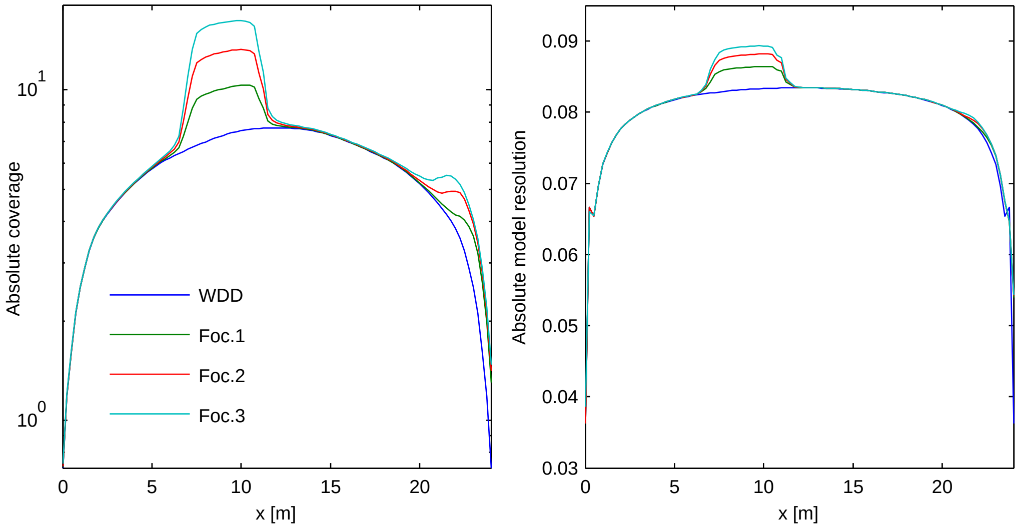


Figure 3.7: The logarithmic absolute coverage (left), and absolute model resolution (right), vertical mean, for Wenner-Dipole and two focused arrays. 96 electrodes, 25cm spacing.

Results For the ERT measurements in the tracer study, it is desirable to obtain maximum resolution on a 3m-wide vertical band, in which the tracer will propagate downwards (see chapter 4.2). In this area (around 7% of the total model cells), focusing is applied, i.e. $A_i = 1$ for model cells inside the band, 0 for cells outside. The optimization is carried out as above in three stages of expansion, with 891[5], 2×892 and 3×892 additional measurements, hereafter called Foc.1, Foc.2 and Foc.3. As a base set, a modified Wenner-Dipole-Dipole set with 3947 points is taken.

Fig. 3.7 shows the coverage and resolution for the base set and Foc.1-3 in a vertical mean. Absolute values are given, because the resolution matrix of the comprehensive set cannot be calculated. In comparison to fig. 3.5, the resolution enhancement achieved by optimization is less pronounced, as the base set already has a very high information content.

Outlook

Optimization of data acquisition is currently only discussed in the context of an optimal array configuration for a given electrode line. An optimization which includes the electrode positions could produce better results. This is especially true for measurements that include

---

[5] As the tracer area constist of 891 model cells in refined forward mode

3D and subsurface electrodes, as in contrast to the 2D-line, no thouroughly tested classical arrays exist for these. A sensitivity-based approach could be applied to this problem, too. However, a direct evaluation is nearly impossible with variable electrode positions, so optimizers like genetic algorithms would have to be used.

Generally, in the 3D case optimization will have a larger impact as in 2D, as information efficiency is a more important issue.

Other interesting topics for further research could include a systematic testing for 3D effects, as the 2D hypothesis is widely used in ERT surveys and seldom verified. For a 2D line, optimization could provide a minimal set of additional out-of-line measurements. By the use of such a set, the existence of 3D artifacts could be reduced.

3.2 Joint approaches for hydrogeophysics

3.2.1 Introduction

As described in chapter 2.8, one of the main challenges in hydrogeophysics is the determination of transport parameters for modeling water and solute transport. As pointed out, a direct parameter estimation is often not possible, at least not at field scale. Therefore, inverse methods have to be used. Inverse modeling has been done for single data sources. In this chapter, existing methods are summarized and a new integrated approach for ERT and TDR is developed. A focus is given to multimethods, i.e. combined information from multiple geophysical methods.

We will look at the problem class of monitoring solute transport in the vadose zone. Therefore, the variables to monitor are water content θ and solute concentration C_W. Water and solute transport are modeled using Richards equation and CDE, respectively. The material parameters are the Mualem-van-Genuchten parameters and the solute dispersivity λ. A detailed description of the target study can be found in 4. The methods described here are not limited to this special case, though. Similarly, while the methods available at the Grenzhof field site, namely TDR, GPR and ERT, are considered here, results can easily be transferred to other geophysical methods.

3.2.2 Conventional method

Traditionally, hydrogeophysical studies are mostly limited to one geophysical method, which is validated by ground truth data, i.e. hard information obtained by excavation or drilling. If multiple methods are available, overlapping data can be used for validation, and results combined taking into account the different sensitivities and reliabilities of the respective method. The state variables θ and C_W are obtained independently by TDR, GPR and ERT. The information obtained by the different methods has the following characteristics.

- *TDR* provides hard point information on the dielectric permittivity ϵ as well the electrical conductivity σ. Using the CRIM equation (2.59), θ can be obtained from ϵ.

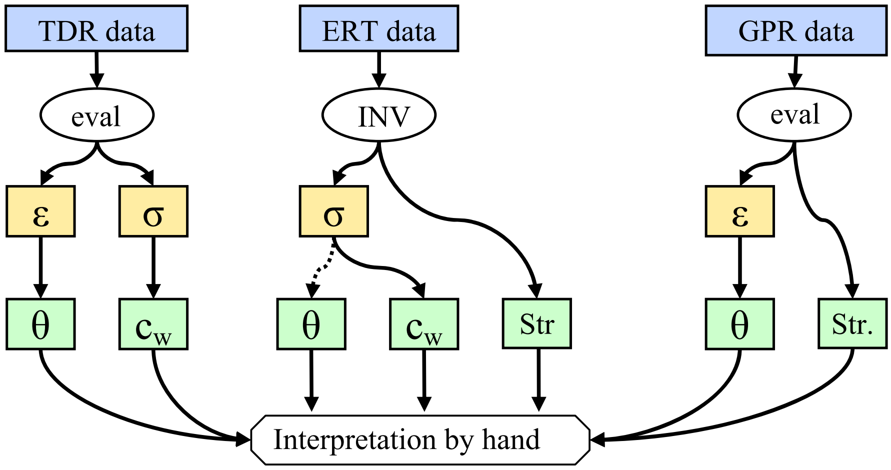


Figure 3.8: Conventional data processing with manual interpretation. Str. denotes structure.

The relation between σ and C_w is rather complex (refer to ch. 4.3.2), and has to be calibrated. TDR data can be measured with high temporal resolution.

- *ERT* produces a spatially resolved image of the electrical conductivity. Again, this can be related to the solute concentration C_w by calibration. There are approaches to relate σ to θ in areas with a constant ion concentration, but they can only be applied only in special cases with careful calibration. Apart from the direct spatial distribution, independent structure information on layer boundaries etc. can be extracted.

- *GPR* can provide spatially resolved information about ϵ, so that water content can be derived. While the resolution along the surface direction can be very high, resolution and penetration depth into the ground vary depending on the subsurface material. Using the commonly used ray-based evaluation, only the position of reflectors at layer boundaries are obtained, and ϵ can be determined as a mean value over the respective layers only.

Fig. 3.8 shows this conventional way of data interpretation. The obvious disadvantage of this method is that this interpretation has to be done manually, which requires experience with this kind of data, and it is time consuming. Also spatial resolution is limited, as it depends on such factors as the inversion process for ERT and the presence of reflectors for GPR.

3.2.3 Inverse method

As pointed out in 2.8, inverse modeling may be used to obtain a consistent, high-resolved image of the state variables. Here, the inverse process consists of both a Richards and CDE solver, but will be referred to as "transport inversion" thereafter for simplicity. The model parameter vector is formed by a number of layers, each characterized by the Mualem-van-Genuchten-parameters and λ. A full starting condition, as well as the atmospheric boundary

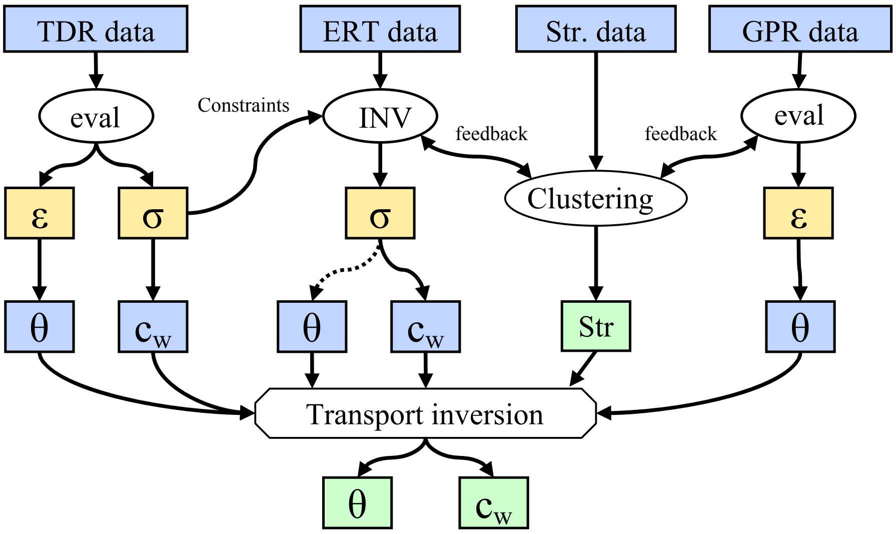


Figure 3.9: Data processing using inverse modeling

condition (precipitation) is estimated from the available data. In the forward step, a data vector is generated to match the existing data. That is, for every point in time and space where water content/solute concentration measurement data exists, a forward response is generated. This forward response can be directly compared to the measurement data, and fed back in the inversion. Of course, this procedure has to deal with the problems associated with all inversion methods, namely non-uniqueness and stability. The data vector has to be significant enough in order to enable stable results. Also, it only works if the subsurface structure is simple enough (i.e. sharp boundaries, etc.) and can be determined with the available methods. The process is shown in figure 3.9. While to my knowledge, no such attempt is documented for ERT, it has been successfully used with other single data sources. For TDR, *Ritter et al.* [2003] presented an onedimensional inverse approach on water content, yielding transport parameters. *Kowalsky et al.* [2005] performed a joint inversion of the twodimensional Richards equation with a GPR forward solver for borehole-borehole radar data.

However, this procedure is especially advantageous when multiple geophysical methods are used, as the manual interpretation step can be avoided by the automated integration of the data in the inversion.

Overlapping data that does not agree, especially if produced by "soft" methods like ERT, can pose a problem though. While a consistent image can usually be derived by hand-tuning, considering the different reliability of data, this is hard to automatize. Therefore the question arises on how to use overlapping data to produce a stable combined image.

One approach is to use hard data to constrain soft data, which means posing further constraints to the ERT inversion. This is explained in the next paragraph.

Constraining There are several possibilities to introduce additional information as constraints:

- **Hard constraining :** If hard resistivity data is present at defined areas or points, these values may be fixed in the inversion. This is achieved by excluding the affected model cell from the inversion process, and using a predefined value for them. This procedure is especially useful for incorporating TDR data. Ideally, the discretization grid should be matched with the constrained areas.

- **Soft constraining :** Another option is to use regularization (see chapter 2.6.2) for constraining. If a minimum length regularization scheme is used, then data can simply be set in the reference model m_0. Then, a weight can be given to the constrained area by setting the respective elements of w_c. If a smoothness constraint is applied, an additional regularization term has to be used[6].

  $$\Psi(\mathbf{m}) = \|\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}))\|^2 + \gamma_1 \|\mathbf{C_s m}\|^2 \qquad + \gamma_2 . \|\mathbf{C_f}(\mathbf{m} - \mathbf{m_f})\|^2 \tag{3.4}$$

  Here, $\mathbf{C_s}$ is the smoothness matrix, while $\mathbf{C_f} = \mathrm{diag}(\mathbf{w_f})$ is the new constraint matrix. $\mathbf{w_f}$ contains the weights, $\mathbf{m_f}$ the constraining resistivity values for the model cells to constrain. They are 0 for all other cells.

- **Constraining structure :** Different geophysical methods measure different properties, but often see the same structure. This is especially true for ERT and GPR, as a jump in water content between different layers also results in a jump in electrical conductivity. Therefore, GPR reflections from layers can often be identified with layers seen in ERT. As ray-based multichannel GPR evaluation can provide relatively exact depth information of the layer boundaries, this information could be used to stabilize ERT layer boundaries. This is especially useful with layers with an high resistivity contrast, as ERT smoothness regularization introduces a wide gradient around the boundary. As shown by *Günther et al.* [2006] for ERT and Seismics, this can be achieved by weighing the smoothness constraint matrix C (see chapter 2.6.2). For a mainly horizontal layering for example, rows of the derivative matrix in z-direction C_z belonging to model cells on the layer boundary are multiplied by a value < 1, or set to zero entirely. This enables the ERT inversion to draw a sharp boundary along the line. It is worth noting that the transition is not enforced, which means if ERT data does not allow for a clear transition, none is taken. Therefore the method is robust against reflectors seen by GPR which do not have a counterpart in the conductivity, e.g. sand layering.

  If two methods with inner inversion (e.g. ERT and full-wave GPR evaluation) are to be coupled, the structure information can be extracted in each inversion step and fed back into the other method as a constraint, thus iteratively establishing a common structure.

In fig. 3.9 the inverse method including constraining is presented. TDR data is used to stabilize the ERT inversion using hard constraining and structure data from GPR and excavations/drilling is used to provide sharp layer boundaries in the ERT image. The combined

---

[6] While it is possible to combine different constraints into the same constraint matrix C, this does not work here, as the sharp boundaries introducted into the reference model would produce unwanted terms in the differential smoothness matrix.

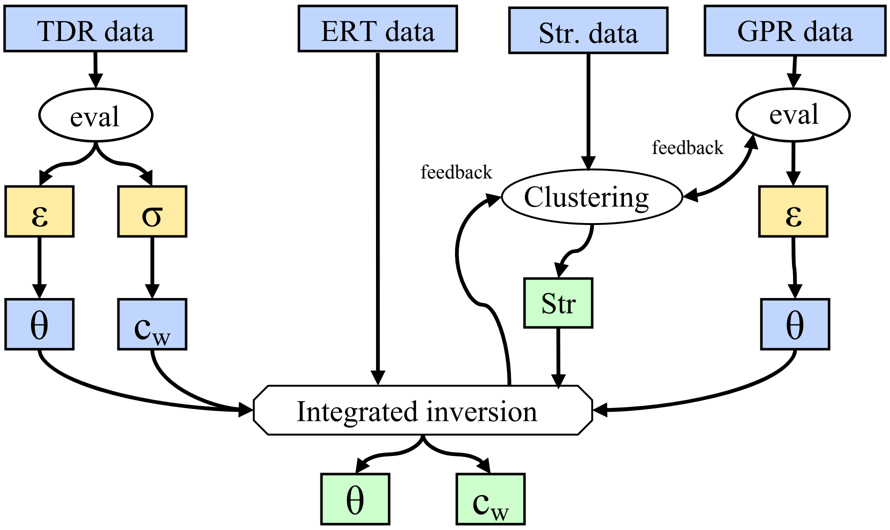


Figure 3.10: Data processing using an integrated inversion

structure information of these 3 sources can then be used as an input to the transport inversion.

3.2.4 Integrated inversion

As pointed out before, one main problem in combining data from different sources is the difficulty on how to handle soft data, even if the ERT inversion can be stabilized by constraining as described above. An approach to eliminate this problem is presented here. The basic idea is to integrate the ERT inversion with the transport inversion. The ERT forward step operates on a resistivity model obtained by Richards (2.54) and CD equation (2.58). Therefore, the inversion does not operate on a complete spatially resolved resistivity model, but only on the set of material parameters. This greatly simplifies the inverse problem, and reduces the need for regularization with the associated problems. The high-level schematics are shown in fig. 3.10, the integrated inversion procedure is shown in detail in fig. 3.11. $m_{\mathrm{Hyd.par}}$ denote the material parameter set for Richards and CD equation. The combined forward operator of these two equation is determined by f_{Hyd}. The output data vector $\mathbf{d}_{\mathrm{Hyd}}$ consists of a spatial and temporal resolved image of water content and solute concentration. The solute concentration is transformed into a conductivity in the next step, and a background conductivity is added. This model is the base model for the ERT forward operator f_{ERT}, which is run for all points of time in which an ERT measurement was performed. The resulting data vectors $\mathbf{d}_{\mathrm{ERT}}$ can be compared to the measured data, and a model update step can be performed. Similarly, water content and electrical conductivity information is extracted for all points in time and space in which TDR measurement data is present. This is again the base for a model update step.

In the algorithm, there are not two but one combined model update step. The different contributions from TDR and ERT data are weighted according to reliability.

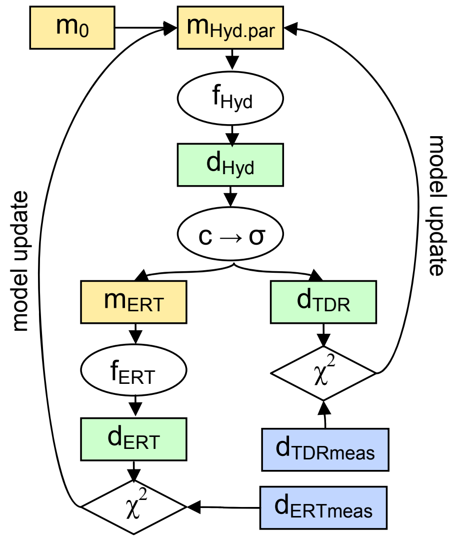


Figure 3.11: Schematics of the integrated inversion. For clarity, only TDR and ERT paths are shown, GPR is treated analogical.

3.2.5 Implementation

The techniques described above have been applied to the Grenzhof tracer study. The tools necessary for separate inversion and evaluation have already been described in previous chapters, while constraining can be added without much effort the ERT inversion code. Therefore, these methods are applied directly to experimental data. The results can be found in chapter 4.3.2. Integrated inversion on the other hand is a new concept with a nontrivial implementation, therefore implementation is described here in more detail. As shown in fig. 3.11, the algorithm focuses on TDR and ERT data.

Data The data vector is formed from three data sources; ERT apparent resistivities ρ_{app} at a number of times $t_{E_1} \dots t_{E_l}$ and TDR electrical conductivity σ_{TDR} and water content θ_{TDR} at a number of times $t_{T_1} \dots t_{T_k}$:

$$\mathbf{d} = \begin{pmatrix} \frac{\beta_1}{N}\, \rho_{\text{app}}(t_{E_1} \dots t_{E_l}) \\ \frac{\beta_2}{m}\, \sigma_{\text{TDR}}(t_{T_1} \dots t_{T_k}) \\ \frac{\beta_3}{m}\, \theta_{\text{TDR}}(t_{T_1} \dots t_{T_k}) \end{pmatrix}$$

λ_i β is a weighting factor to balance the influence of the three methods. Given l ERT measurements with N data points and k TDR measurements with m probes, the data vector has a total of $l\,N + 2\,k\,m$ data points.

Algorithm The algorithm is shown in fig. 3.11. The main process was implemented in MATLAB, creating input files for all other required subprocesses, reading their output and performing the inversion. For the transport forward operator (Richards/CDE), the open SWMS-2D package (*Simunek and Vogel* [1994]) was used. It requires a discretized grid, a

division into sections with model parameters for each section, and the atmospheric boundary conditions as an input. The grid can be synthesized from layer boundary information, which may be fixed or part of model parameter vector to invert. The model parameters include residual water content θ_R, porosity θ_S, van-Genuchten parameters α, n, the hydraulic conductivity K and dispersivity λ. Again, part of these parameters are known and therefore fixed, while others are subject to inversion.
For solute transport, boundary conditions have to be provided. To simulate a tracer transport study, an initial model with concentrated solutes at the surface layer at $t = 0$ is provided.

SWMS outputs spatial water content and solute concentration distributions at the required times $t_E \cup t_T$. From this output, TDR data and the ERT model is generated. The relation $\sigma \leftrightarrow c$ is assumed to be linear here for simplicity. The ERT model rescales the grid and adds a background conductivity model, which is obtained by a real reference measurement. The ERT forward step is the same as in the single ERT inversion.

For least square optimization, the MATLAB implementation of preconditioned conjugate gradients was chosen. A Levenberg-Marquardt type optimization is also feasible, but the large-scale optimizer was chosen for faster convergence. The Jacobian needed is calculated by finite differences

$$J_{ij} = \frac{\partial f(\mathbf{m})_i}{\partial m_j} \approx \frac{f(\mathbf{m} + \mathbf{e_i}\Delta m_i) - f()}{\Delta m_i}$$

which causes $p+1$ transport forward steps per inversion step, where p is the number of model parameters to invert.

Distributed computing One problem arising is the high numerical effort involved in this process. Each iteration, $p+1$ transport forward steps and $l\,(p+1)$ ERT forward steps have to be calculated. With a mean running time of several minutes for ERT and up to one hour for the hydraulic forward operator, computation on a single processor is not feasible. Therefore, the procedure was partly parallelized to run on a Linux computing cluster.

SWMS itself does not support parallel computing, so one SWMS forward step constitutes the bottleneck operation. Still, a parallelization factor of $p + 1$ can be achieved for the SWMS step, and a much higher one for the ERT step. SWMS is available as FORTRAN code, which was recompiled on the cluster. The input files are generated by the main MATLAB process residing on a control computer, and transferred between these computers by a scheduler. The ERT forward operator was ported and compiled on the cluster, and a file interface was implemented for data exchange.

To efficiently distribute SWMS/ERT tasks on the cluster nodes without concurrency problems, a task scheduler was implemented (C++). It consists of a server program residing on the cluster control node, and a command tool residing available to every computer involved. Using the command tool, the user can add or remove nodes from the processing pool at any time by command line, the main MATLAB program can add tasks to a queue, and the subprocesses can signalize their termination to the server. The tasks in queue are scheduled to the next free node in the pool, and appropriate file structures are created on the respective cluster node.

Thus, in each inversion step, the main process feeds $p + 1$ SWMS processes to the scheduler, and switch to waiting mode. As soon as a process is finished, the main process gets notified, reads the results and posts the appropriate ERT tasks to the scheduler. When all ERT tasks are finished, data is read out and the forward data vector can be constructed.

3.2.6 Case study

Before applying the algorithm on real data, a case study with synthetic data has been conducted. As the target application the the Grenzhof tracer study (chapter 4), parameters have been chosen as expected for this experiment.

Setting The model grid for transport forward modeling is a 6 m long and 2 m deep profile, with a discretization of 5 cm × 2 cm. The salt concentration is initially concentrated in the uppermost layer between 2 m and 4 m. The layering is assumed to be horizontal[7]. 4 Layers analog to fig. 4.1 are expected. Soil porosity etc. for these layers are fixed to the values determine by soil sampling, while α, n, K and λ are unknowns. Starting values for these parameters are taken from an hydraulic inverse model by Ute Wollschläger. It was generated from TDR water contents using the commercial HYDRUS package. For λ approximate values are estimated. The ERT forward operator uses a 25cm spaced 96-electrode Wenner-Dipole configuration as described in ch. 4.1. The background conductivity model is the same presented there.
All parameters are listed in table 4.1.

Response surfaces Before running the inversion, response surfaces, i.e. images of the objective function in parameter space, are calculated for the model as defined above. Due to the high time complexity of each forward step, response surfaces can be rasterized only as a function of 2 or at maximum 3 parameters. A selection of 2D slices for different parameter combinations is shown in fig. 3.12, separate for each of the three methods. It can be seen that on a large scale, only one minimum exists in a 2D transect and the response surface structure is rather simple. On a smaller scale however, there is a tendency to local minima and complex surface geometry. This can be seen for example in the K_1/α_1 section in fig. 3.12. This observation holds true for all examined 2D and 3D sections. The full multidimensional function is expected to have a more complex structure.

For ERT data, the different gradients tend to be of the same magnitude, while for TDR methods the gradients are further apart. For ERT data, this leads to faster converging and lowers susceptibility to noise, but increases the chance of getting stuck in a local minimum. By using a weighted sum of the three methods as objective function, these effects can be finetuned by specifying the β values.

In conclusion, for the scanned area, the response surfaces results look good enough that optimization might work, while problems with small-scale local minima might arise

---

[7] Which is a good approximation for this patch, as shown by GPR and excavation

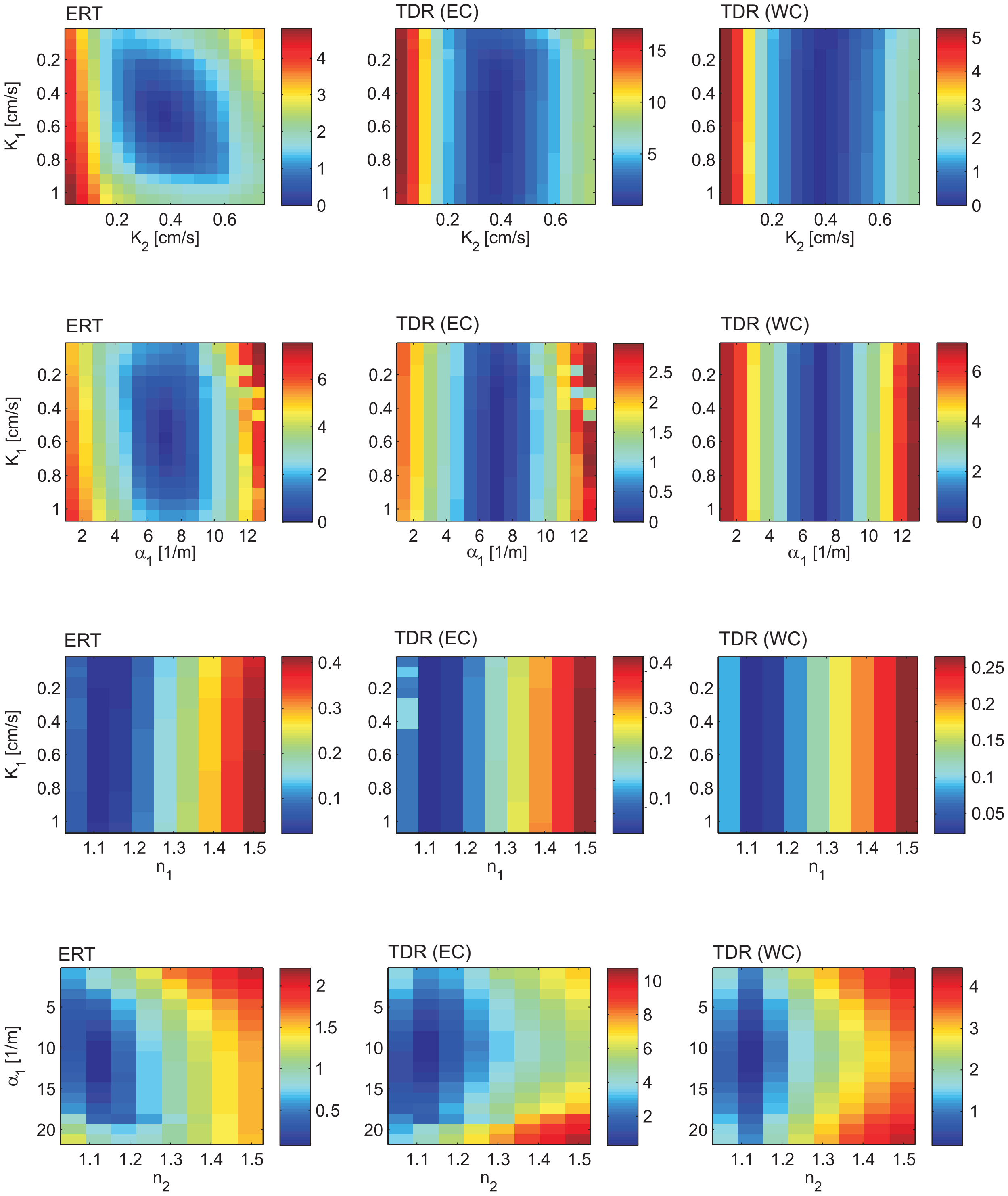


Figure 3.12: Response surface $|\chi|$ for different parameters (Arbitrary units).

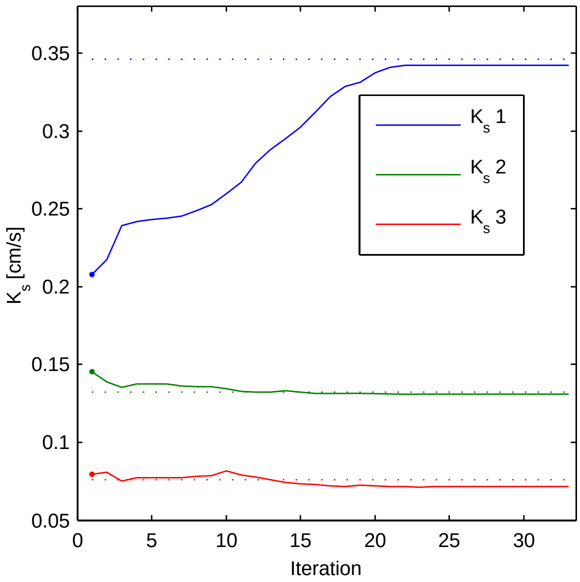


Figure 3.13: Integrated inversion for K_1,K_2 and K_3. The target parameters are marked by dotted lines.

depending on parameter space position. This analysis can however not rule out a potentially highly complex surface when looking at more dimensions.

Inversion results Next, the inversion is tested for two different parameter sets. First, only the hydraulic conductivities K_1, K_2 and K_3 are inverted. A synthetic data set is created with changed K_i's. K_1 is varied strongly, K_2 and K_3 slightly from the reference model. The other parameters are set as in tab. 4.1. The results are shown in fig. 3.13. After 20 iterations, the inverted parameters stabilize near the target parameters.

This test is repeated with a larger set of inverted parameters, namely n and α for all four layers. Fig. 3.14 shows the development of the parameter values. Again, they stabilize near the target parameters, however a larger number of iterations is required. This is most likely due to the differing sensitivity of the objective function regarding n and α.
This especially leads to very slow convergence when K, n and α are inverted together, as the the sensitivity range is even larger there. The different sensitivities can be also observed in the response surfaces.

3.3 Other implementations

3.3.1 Electrode spacing

Introduction Usually, ERT is used with large electrode spacings ($\geq$ 1m). In my experiment I wanted to resolve a rather shallow area ($\leq$ 2m) with a maximum spatial resolution, so a small spacing is desired. On the other hand, the measurement becomes more susceptible to geometric errors with decreasing electrode spacings. A well-known effect is the suscepti-

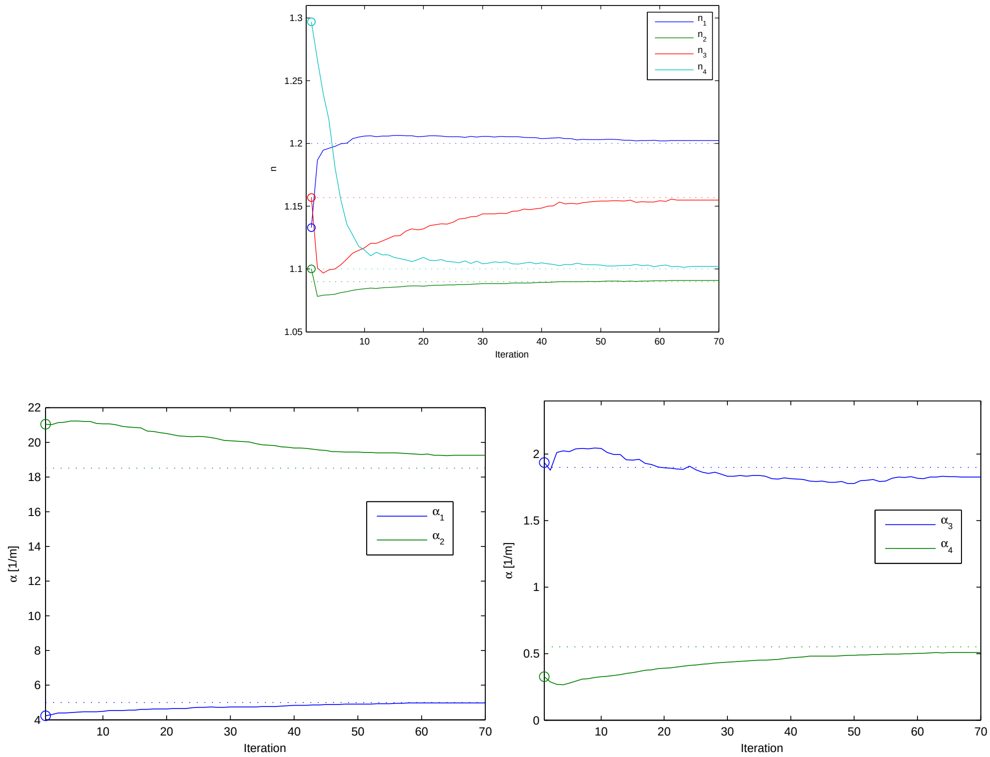


Figure 3.14: Integrated inversion for α and n of all layers. The target parameters are marked by dotted lines.

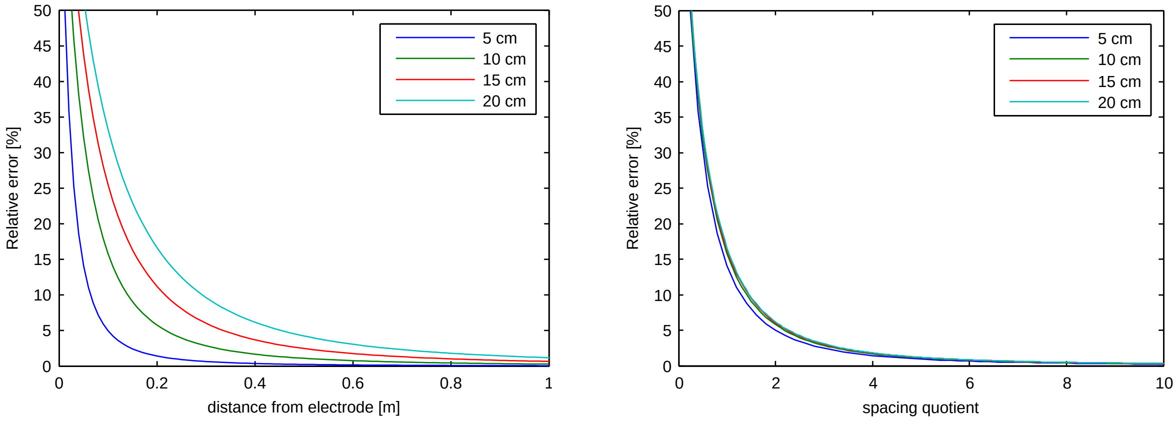


Figure 3.15: Relative error of the point-source potential to the finite-intrusion length potential for different electrode lengths

bility to electrode positioning error, so care has to be taken during electrode set up.
A neglected aspect here is the electrode intrusion depth in the ground. To provide good ground contact, the electrode usually is inserted rather deep in the ground. The forward solver however assumes the current source to be a point source, an assumption that may no longer be good for small electrode distances. The effect of this is discussed in this chapter.

Error estimation To quantify the errors induced by this effect, a simple simulation has been conducted. The potential of an electrode with intrusion depth l is discretely approximated by a number of point sources with $z \in [0, l]$, i.e.

$$\Phi^f(x,z) \approx \frac{p}{l} \sum_{i=0}^{l/p} \Phi_0(x, z + p\,i) \qquad .$$

p is the discretization lattice constant. Note that the expression for subsurface sources (2.8) has to be used for Φ_0. In the same way, the potential seen by a receiver electrode of finite size is assumed to be the mean of potential along its intrusion depth. A receiver electrode in the distance d therefore sees the potential

$$\Phi^r_d \approx \frac{p^2}{l^2} \sum_{i=0,j=0}^{l/p} \Phi_0(d, p\,(i+j)) \qquad .$$

For $p \rightarrow 0$ Φ^r converges against the continuous solution for a line electrode of finite size.

The "real" potential Φ^r is compared against the idealized point-source potential in fig. 3.15 for different electrode lengths. Here, p = 1cm is used. Naturally, the error diminishes for large distances between current and receiver electrode. The right diagram shows that a spacing quotient

$$q = \frac{d}{l}$$

is a good parameter here. It can be concluded that $q > 2$ should always be ensured for the described experiments. For the desired spacing of 0.25 m, the intrusion depth should therfore be limited to less than 12.5 cm, which is achievable.

Error avoidance For cases in which a big q is not achievable, the forward solver can be adapted to take the electrode length into account. This can be done by identifying p with the FD grid lattice constant, and proceeding as above. In the FD code, this procedure uses multiple source terms, and an additional averaging step after the calculation. When using singularity removal, a semi-analytic expression for the line electrode could be used as a primary field to achieve better results.

3.3.2 Multichannel optimization

Introduction

Our measurement instrument IRIS Syscal-Pro supports 10-channel in-line measurements. For an quadrupole array, this means for every current injection dipole (a, b), up to 11 voltages

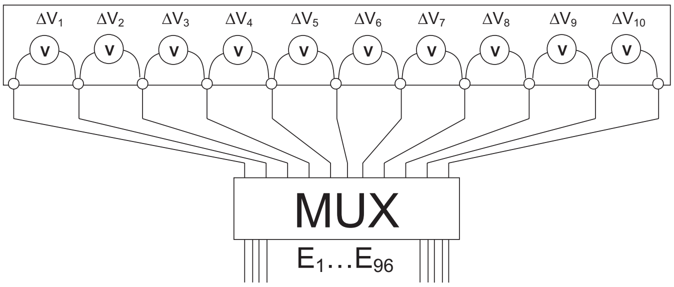


Figure 3.16: 10-channel in-line measurement schematics

and therefore 10 potential differences (m_i, n_i) can be measured simultaneously (3.16) in one multichannel measurement (MCM). In the best case, this can reduce measurement time up to a factor of 10. However, for a given measurement array, usually only a much smaller factor can be achieved. This is due to

- often, fewer than 10 receiver dipoles are to be measured per injection dipole, e.g. only one in the extreme case of a Wenner array
- these receiver dipoles are generally not connected, e.g. the dipoles $(1-2), (3-4)$ use up 4 of 11 voltage channels, the difference $(2-3)$ remains unused
- each electrode can be used only twice, e.g. the dipoles $(1-2), (1-3), (1-4)$ have to be measured in 2 sets: $(2-1)(1-3)$ and $(1-4)$

The IRIS toolset features only a basic support for multichannel measurements. Therefore, I devised an own multichannel optimization algorithm. In order to maximize the multichannel usage, there are basically two strategies:

(i) *Multichannel set generation*
When generating a custom array, i.e. using a focusing or array optimization algorithm, a full 10-channel set is directly created. This can be achieved by not choosing one best-fit quadrupole, but a best-fit set of one injection dipole and 10 connected receiver dipoles each optimization step. This technique is described in detail in (3.1.5).

(ii) *Multichannel set optimization*
Set optimization tries to maximize the efficiency of a given measurement array by optimal allocation of the 11 voltage channels for each injection dipole set. In contrast to multichannel set generation, this is a non-trivial problem. The approach is described in detail in the next chapter.

Set optimization

For a given measurement array, optimization tries to reduce the number of measurements for each injection dipole by efficiently allocating the receiver dipoles to the multiple channels. Since Syscal-Pro expects multichannel measurements to be written sequentially in line,

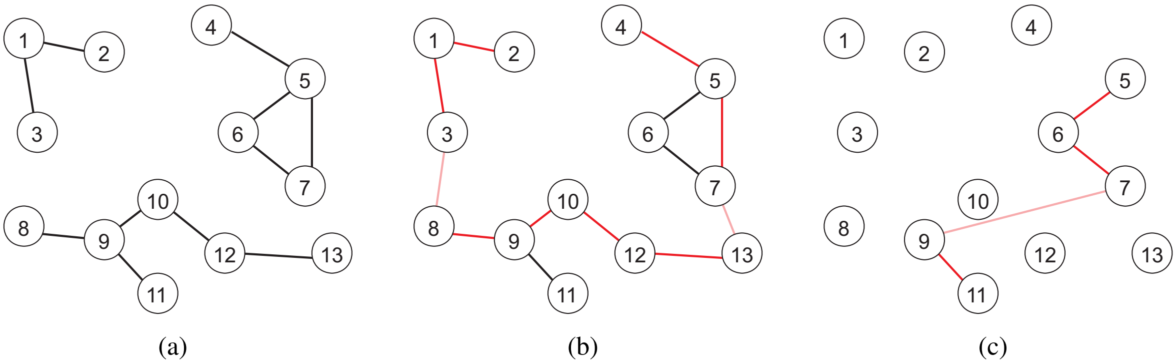


Figure 3.17: An example problem is shown in (a). (b) shows the first multichannel set chosen by the algorithm in red. Added connection dipoles are shown in pink. (c) shows the next algorithm step, the first set has been removed and the next one identified.

dummy dipoles have to be inserted between non-connected dipoles, e.g. $(1-2),(1-3),(4-5) \Rightarrow (2-1)(1-3)(\mathbf{3}-\mathbf{4})(4-5)$. In the following, the optimization for the measurements belonging to one given current dipole is discussed. The process has to be repeated for every current dipole.

This problem is essentially a graph problem. In the graph $G = (E, N)$, the nodes N represent electrodes and edges E represent measurement dipoles. Fig. 3.17(a) shows an example receiver dipole graph for a given current dipole. A multichannel measurement is represented by a set of up to 10 edges which may or may not be in E.

$$\Pi = \{(\epsilon_1, \dots, \epsilon_m) \in E' | m \leq 10\}$$

where E' refers to the edges of the fully connected graph $G' = (E', N)$. Edges $\in E'$, but $\notin E$ are the dummy dipole mentioned above. The optimization procedure is therefore to find a minimum set of multichannel measurements which cover all measurements, i.e. to find a set

$$\Sigma = \{(\Pi_1, \dots, \Pi_n) | E \subseteq \bigcup_{i=1}^{n} \Pi_i\}$$

with minimal n. For small numbers of edges, this problem can be solved by brute force (testing all possible permutations for optimality). However, as this method scales exponentially, more advanced algorithms have to be used for larger numbers of edges. The developed algorithm contains an heuristic to find a near-optimal solution, which will in many cases be identical to the optimal solution.

Algorithm

(i) Build the graph G from input

(ii) G is partitioned in its connected components $C_1 \dots C_n$.

(iii) Initialize a set of paths $S = \{\}$

(iv) For every component C_i

- (a) Find all nodes N_O in the component with a odd number of connections
- (b) Add all unique paths between the nodes in N_O to S (this can be implemented using a simple depth search)

(v) Initialize a high score $v_h = 0$

(vi) Backtrack recursively through the power set of S. At each step

- For the current branch $s_1 \ldots s_n \in S$ calculate

$$\begin{aligned} l_n &= (n-1) + \sum_{i=1}^{n} \operatorname{len}(s_i) \\ v &= \begin{cases} \frac{l_n}{\lceil l_n/10 \rceil} & \text{if } l_n > 10 \\ l_n - 1.5n & \text{else} \end{cases} \end{aligned} \tag{3.5}$$

- Stop the branch if $l_n > 10$.
- If v v_h, store new highscore.

(vii) Write the branch with the highscore as multichannel measurement to output, and remove the edges from G

(viii) Continue with (2) until the graph is empty

The algorithm essentially performs a greedy optimization, with the optimization function (3.5). It is defined so that the score rises with the number of edges $\in E$ per multichannel operation, and has a penalty for unnecessary connection edges $\notin E$. So two paths of length 4, with one connection edge between the paths would receive a higher score that one path of length 7, but a lower score than one path of length 8. Paths longer than 10 have to be splitted, and receive an analogous partial score. The elementary operation unit of this algorithm is not a single node, but all paths between odd-connected nodes. This takes advantage of the fact the most optimizable arrays (i.e. Dipole-Dipole, Gradient, ...) have a structure of long connected lines. Even-connected nodes with a connection number > 2 are interpreted as crossings of paths, not path junctions. It turns out that this assumption performs good for the investigated arrays. Between these paths, a brute-force evaluation for the highest score per path is conducted. This performed fast enough for all arrays investigated (some seconds for 3-digit measurement points). In the case of an array with largely non-connected edges, the inner step of the algorithm would converge to the speed of a brute-force per edge algorithm. For dealing with such problems, more advanced evaluation algorithms would have to be sought.

3.17 shows an example problem that is solved by the algorithm into 2 multichannel sets.

The algorithm was implemented in C++. The graph partitioning in step 2 could have been implemented by calculation of the convex hull of the graph, but here the open Boost

Array	Measurements	Num MCMs	scaling factor
Dipole-Dipole (a=1)	4371	480	9.1
ext. Dipole-Dipole (s=4)	958	245	3.91
Wenner	1488	1488	1
Gradient (s=15,t=1)	1185	158	7.5
Wdd+Opt3	6620	2973	2.23

Table 3.1: Performance of the algorithm for some array types.

library was used for speed and simplicity and efficiency. The performance of the algorithm for a number of chosen arrays is shown in table 3.1 (It can easily be shown to be optimal for Dipole-Dipole, Wenner and Gradient arrays). The execution time for these arrays ranked from 1s to 3s on a 1GHz machine.

Also, a tool was developed to directly transfer the optimized arrays to the Syscal instrument by serial port[8]. For this purpose, the serial communication to the device was eavesdropped by a serial monitor, and the binary communication protocol decoded. It is implemented in both C++ and Matlab for connectivity to the other tools.

---

[8] This eliminates the need to transform the array files into a IRIS-compatible format, and uploading it using the proprietary IRIS toolset.

Chapter 4

Experiment

To be able to test and apply the methods developed in this thesis, a field experiment was carried out. As we were especially interested in the capability of ERT to monitor solute transport, a tracer study has been conducted. The setup and results are presented in this chapter.

4.1 Experiment description

Instruments The following instruments were used for the experiment :

- The ERT measurement instrument used was a Syscal Pro 96 by IRIS Instruments (fig. 1). The instrument features :
  - A voltage transformer for current injection with an external source voltage 12V and adaptive output voltage up to 1000V.
  - A 12-bit AC/DC converter for measuring potential differences, with an adaptive voltage range. 10 channels can be measured simultaneously.
  - Automatic injection ranging, i.e. the injected current can be adaptively increased if measured voltage differences are below a threshold, to improve measurement accuracy.
  - Full switching support for 96 electrodes on 11 reception channels + 2 current channels. The electrodes are connected to the device by two 48-core cables.
- GPR measurements were performed using two 250-MHz RAMAC/GPR antennas from Malå Geoscience, with a multichannel control unit. Both antennas are connected with a rope or a wooden frame to keep a fixed antenna separation (fig. 2).
- For TDR measurements, custom-built 3-rod TDR probes are used (fig. 2.6). Manual testing is done using a Tektronics 1502C cable tester, while for monitoring a Campbell Scientific TDR-100 device is used.

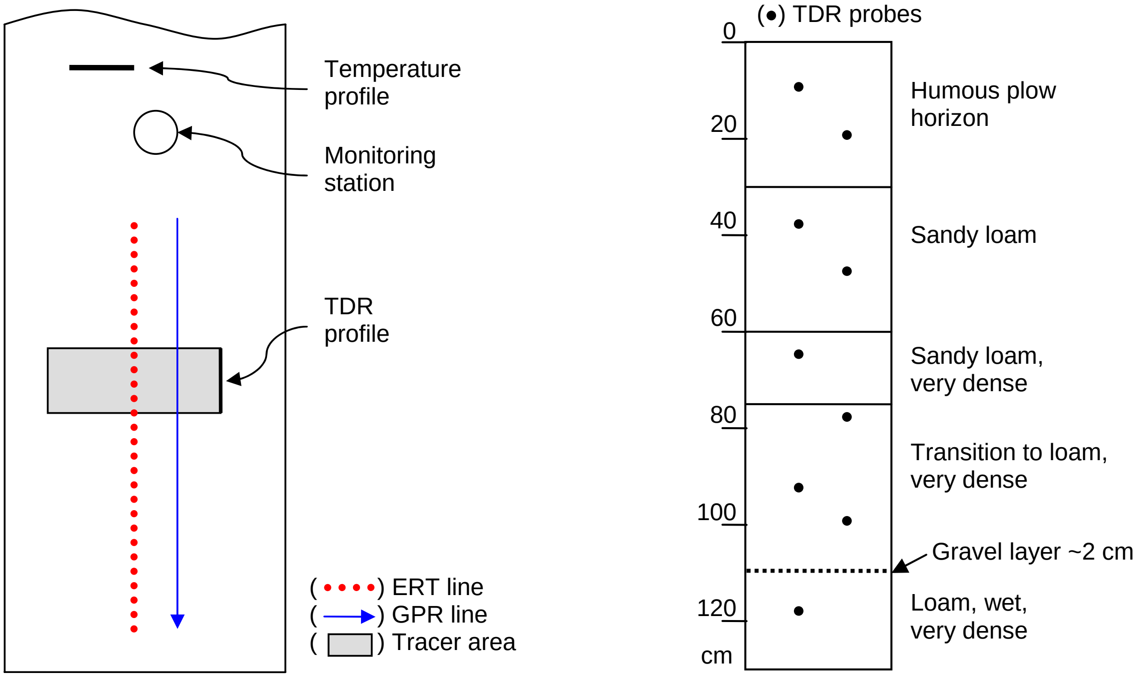


Figure 4.1: Left: Sketch of the Grenzhof test site (from above). Right: TDR profile (front)

- 107 Thermistor temperature probes by Campbell are installed in a soil profile, and monitored by a Campbell CR10X datalogger.

Setting and Installation The experiment took place at the Grenzhof test site, which is a grassy patch of land (9x200 m) located in the farmslands west of Heidelberg-Eppelheim (for a photograph see appendix 3). A monitoring station has been set up in February 2004, which provides meteorological data (precipitation, wind speed and direction, radiation, air temperature and pressure) as well as soil temperature and water content in a profile next to the station. Details can be found in *Bretthauer* [2004]. Details on the ground structure of the site can be found in *Wollschläger and Roth* [2005].

A 2D surface ERT line with 96 electrodes and 25 cm spacing has been permanently installed (see sketch fig. 4.1). ERT measurements and multichannel GPR measurements along the ERT line were conducted on a weekly basis. At one border of the area of tracer application (see next paragraph), a profile has been excavated before the start of the experiment for soil sampling and installation of TDR probes. The probes measure water content and electrical conductivity, which is useful for comparison with ERT data. The profile has a distance of 2 m to the ERT line to minimize the impact of the probes and disturbed ground on the ERT measurements, but still obtain similar tracer concentrations as ERT.

Fig. 4.1 shows a schematic drawing of the profile, including instrumentation. Five different layers can be identified :

(i) Humous plow horizon, roots (0–29 cm)

(ii) Sandy loam, homogeneous (29–60 cm)

(iii) Sandy loam, dense (60–75 cm)

(iv) Transition to loam, very dense and hard (75–108 cm)

(v) Thin gravel layer (108–110 cm)

(vi) Loam, very dense, wet (110–125 cm)

(vii) Gravels (125 cm –)

In total, 10 TDR probes (30 cm rod length) have been installed in the profile at the depths of 9.5 cm, 19.5 cm, 37.5 cm, 48 cm, 64.5 cm, 79 cm, 92 cm, 99 cm and 119 cm. At 9.5 cm depth, an additional probe is installed in a side wall of the profile not facing the tracer area. This enables water content estimation even when the high salt tracer concentration is too high to reliably estimate water content from the top-most probe in the tracer area.

The TDR probes are controlled by a Campbell TDR100 device present in the monitoring station. An additional 8-channel analog multiplexer was installed, and the stations data logger reprogrammed to record TDR traces for water content and electrical conductivity estimation in a three-hourly interval. The probes were calibrated in air, distilled water and water with different salt concentrations.

Soil samples At the excavated profile, soil samples have been taken every 10 cm to determine material properties, in particular soil porosity. Also, four large undisturbed soil samples have been taken. Using evaporation-based MSO, these samples can be examined to try to determine hydraulic conductivity and van-Genuchten-parameters. The examination is still underway at the time of writing of this thesis, *Appel* [2007] will describes this analysis in detail.

Tracer application Before tracer application, several ERT, TDR and GPR measurements were taken to provide reference information. The tracer was applied on a rectangular area perpendicular to the ERT line. The extent was 2 m in profile direction, and 4 m in the normal direction to provide an pseudo-twodimensional setting for the ERT inversion and minimize 3D artifacts.

As a tracer, $CaCl_2$ salt was used. It was applied as a powder with a Cl^--ion concentration of 500 g/m^2 onto the the rectangular tracer area. A previous study (*Ulbricht* [2005]) showed that this concentration is a good compromise between good visibility for ERT due to high conductivity and side effects originating from high conductivity contrasts. No artificial irrigation was provided, so that the tracer is propagated only by natural precipitation. By applying the tracer as a powder, a homogeneous distribution could be achieved easier than with the commonly used fluid tracers. Also, by avoiding additional water influx at the tracer area, the boundary conditions for the hydraulic simulation are simplified.

Layer	θ_R	θ_S	$\alpha\,[1/\mathrm{m}]$	n	$K\,[10^{-3}\mathrm{cm/s}]$	a	$\lambda\,[10^{-4}1/\mathrm{cm}]$
0 - 29 cm	0.045	0.37	4.22	1.13	207.6	0.5	3.0
29 - 60 cm	0.045	0.35	21.05	1.10	145.4	0.5	1.0
60 - 75 cm	0.045	0.36	1.94	1.16	79.3	0.5	1.0
75 - 200 cm	0.045	0.38	0.33	1.30	0.09	0.5	1.5

Table 4.1: Approximated material parameters. K, α, n are obtained by TDR inverse modeling, θ_S is determined from soil samples while a, λ and θ_R are estimates.

4.2 ERT configuration

Measuring configuration The ERT survey was chosen as a 2D surface line profile. While borehole ERT may provide a better vertical resolution, we were interested in the performace of non-invasive ERT, to be used alone or in conjunction with an existing TDR profile. The subsurface structure in the surrounding area is estimated by GPR measurements to be rather uniform, which means 3D artifacts will be limited. A as high as possible spatial resolution is desired, so an electrode spacing of 25 cm is chosen. Smaller spacings would introduce bigger errors as accurate electrode positioning below the cm-scale is unrealistic. To maintain a q value of less than 2 (chapter 3.3.1), electrode length in ground should not exceed 12.5 cm, which can be achieved while still guaranteeing a good ground contact.

A modified Wenner-Dipole-Dipole array is used because of its high information content. As the tracer propagation is very slow, only the depth range up to 1.5 m is of interest. Therefore, the pseudodepth levels of the dipole area are concentrated in this region. The array has a total of 3947 measurements. Additionally, a focusing is applied to the region around the tracer application. It is presented in detail in chapter 3.1.5. Together, a total of 4839, 5731 and 6623 data points result for focusing stages one to three. With multichannel optimization (chapter 3.3.2), this figure is reduced to 1641, 2095, 2517, 2973 multichannel measurements for the base set and the focused arrays respectively.

Automatic injection ranging was used with a trigger value of 200 mV[1], and 6 stack levels[2] were chosen.

Data quality Due to good ground contact, the small spacing and the use of injection ranging a very high signal to noise ratio can be achieved. Between two electrodes, line resistances of 1-4 kΩ are measured. The standard deviation of the stacking was around 0.2%, while a maximum deviation of 2% was observed for typical measurements in this setting.
Additionally, a reciprocal test, i.e. a measurement with current/reception dipoles exchanged, was conducted. It showed a mean reciprocal error of 0.9%, and a maximum error of 4.5%. Systematic errors due to electrode mispositioning will be present. As the electrode positions were metered by measuring tape, a positioning error of $\approx$ 2% (array length 24.5 m $\pm$ 5 cm)

---

[1] Injection current is increased automatically until a voltage difference of at least 200 mV is measured at the reception dipole, or the maximum injection voltage is reached.

[2] Half of these step levels are measured with positive, the other half with negative polarity in order to minimize polarization effects.

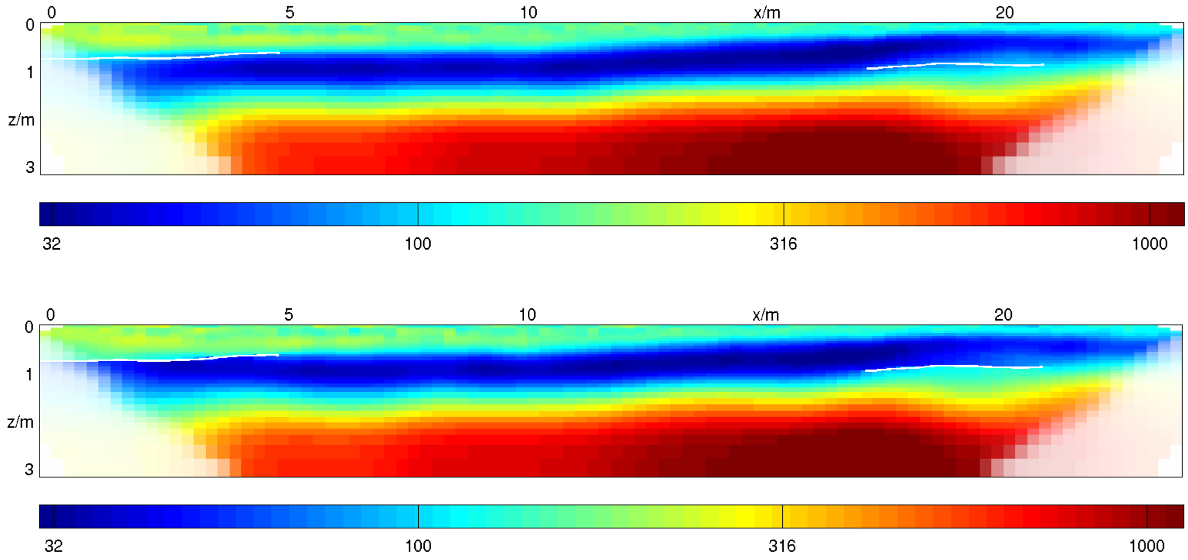


Figure 4.2: Reference ERT resistivity image [Ω] from standard inversion (top), and structure-constrained inversion (bottom). GPR reflectors are drawn as white lines

seems realistic. This error propagates linearly, introducing a 2% uncertainty in the depth evaluation.

It can be concluded that data quality very good for an ERT experiment — One needs to keep in mind though that the main errors in ERT images originate from the inversion process, discretization and assumptions as 2D flatness, not from poor quality data.

4.3 Results

4.3.1 Reference image

As a reference measurement, the last measurement taken before the excavation was chosen.

It was processed by a standard Gauss-Newton inversion with smoothness constraint. A γ value of 55 showed the best results (in terms of smallest RMS error and no visible artifacts), and was used for all inversions. The inversion discretization grid was linearly set to one model cell per electrode (25 cm) in x direction. In z direction, 32 logarithmically spaced division up to 3.2 m are used, which are modified to match the TDR locations. For forward calculations, these cells are refined by factor 4, and 5 large prolonging divisions are added at each boundary.
The inversion result can be seen in fig. 4.2 (top). A χ^2 value of 1.5, and an RMS of 1.4% were reached.

At the same time, a GPR multichannel measurement was carried out along the ERT line. However, the Grenzhof site is only partly suitable for GPR measurement, as the high loam content in the upper layers leads to strong signal attenuation. Thus, a clear GPR signal for multichannel evaluation was only present at some windows. Two reflectors visible with GPR are drawn as white lines in fig. 4.2. It can be observed that the ERT layer boundaries

and the GPR reflectors match rather well. However, due to the strong smoothing constraint, especially at the right reflector, ERT draws a very smooth gradient. It is normally hard to tell from ERT images alone if such a gradient is really present in the ground, or just an artifact of the smoothness constraining. In the second case, the position of the real boundary transition within the gradient remains unclear[3]. Assuming that GPR reflectors also denote resistivity changes, the upper transition to the gradient can now be specified by structure constraining(ch. 3.2.3). The result is shown in the bottom image of fig. 4.2. The left transition is slightly adjusted to match the GPR reflector, while at the right the upper part of the gradient is replaced by a transition at the GPR reflector level.
In this case, the changes introduced by constraining with GPR are rather small due to limited GPR data, but it shows the potential of the method.

4.3.2 Time series

Starting with the tracer application on February 21st, 2007, regular time series of ERT and TDR have been measured. The ERT series was interrupted from Mid June to early July because of a hardware failure and ended in the end of July. From mid-May, probes 2 and 3 cannot provide reliable water content due to high electrical conductivity

Atmospheric conditions Soon after the tracer application, moderate precipitation is observed. In the end of March, a heavy rain event occurred, after which a period with virtually no rainfall until mid-May can be registered. During this dry period, air and soil temperatures increase strongly. From June on, steady precipitation with high temperatures are observed.

TDR series For the duration of the experiment, TDR data is recorded every three hours. For the 10 installed TDR probes, only 8 multiplexer ports were available. So the two probes at 92 cm and 119 cm are measured manually once a week. These two probes were chosen as the first tends to provide unreliable data, and the second is the deepest probe, were dynamics are very low. Fig. 4.3 shows the precipitation, TDR and temperature data starting from the probe installation on February 16th. Electrical conductivity has been corrected to the reference temperature of 25 °C by

$$\sigma^{25} = \sigma(T) + \alpha_T \sigma(T)(T - 25\,\mathrm{K})$$

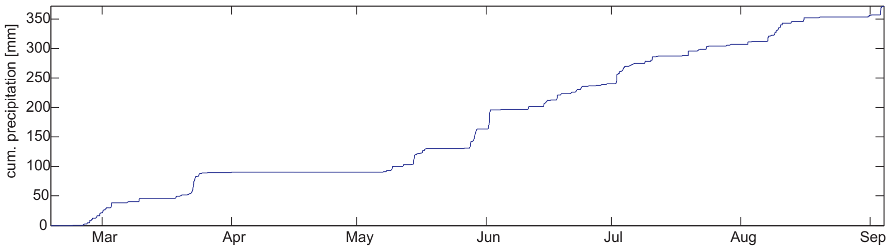
with the constant $\alpha_T = 0.0191$ (refer to *Heimovaara et al.* [1995]). The tracer is initially not visible, and appears in the conductivity plot after the first heavy rainfall around March 1st. From April to mid May, temperatures are rising with virtually no rainfall, causing the upper soil layers to dry out. The tracer seems to disappear from the conductivity plot in this time. This can be explained by

(i) the measured conductivity depends at least linearly from the water content, and

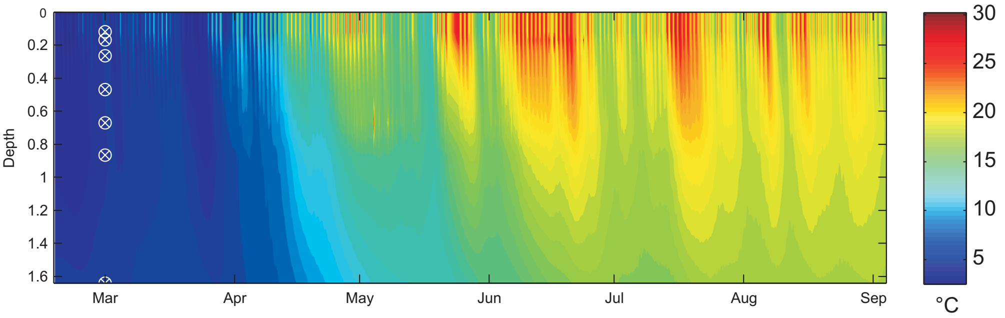
---

[3] Here, it is known from the excavation that actually both is true : While several layer transitions can be identified, there is also gradual change within the lower layers.

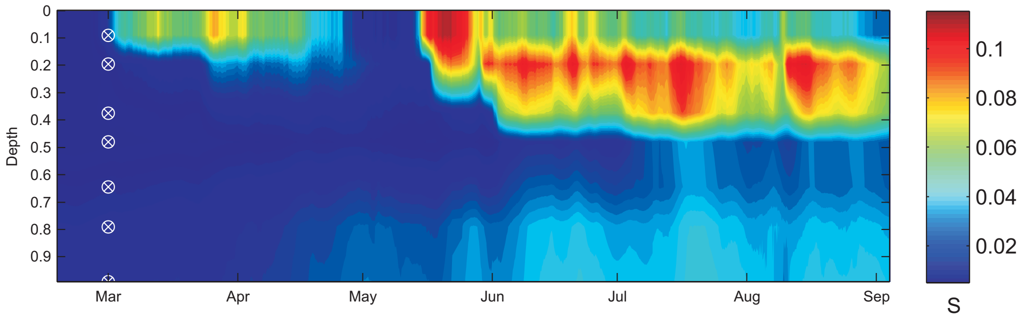
a) Cumulative precipitation



b) Soil temperature



c) Electrical conductivity



d) Water content

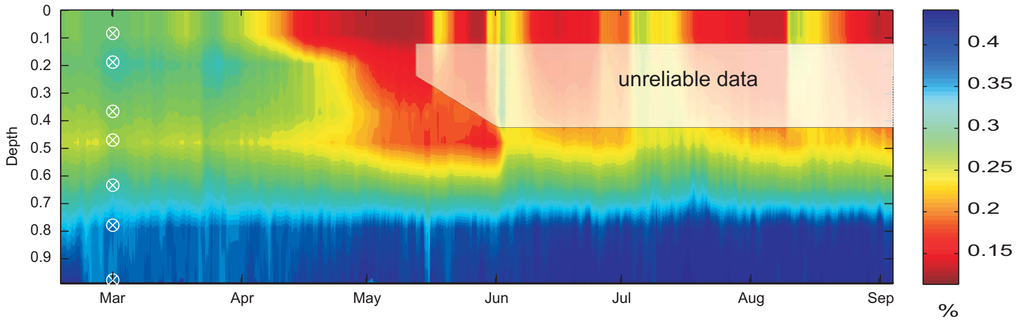


Figure 4.3: Cumulative precipitation, TDR water content, TDR conductivity [corrected to 25° C] and soil temperature profile data as time series. The depths of the respective probes are denoted with a circle. Water content data of probes 2 and 3 is unreliable from mid-May on due to high conductivity.

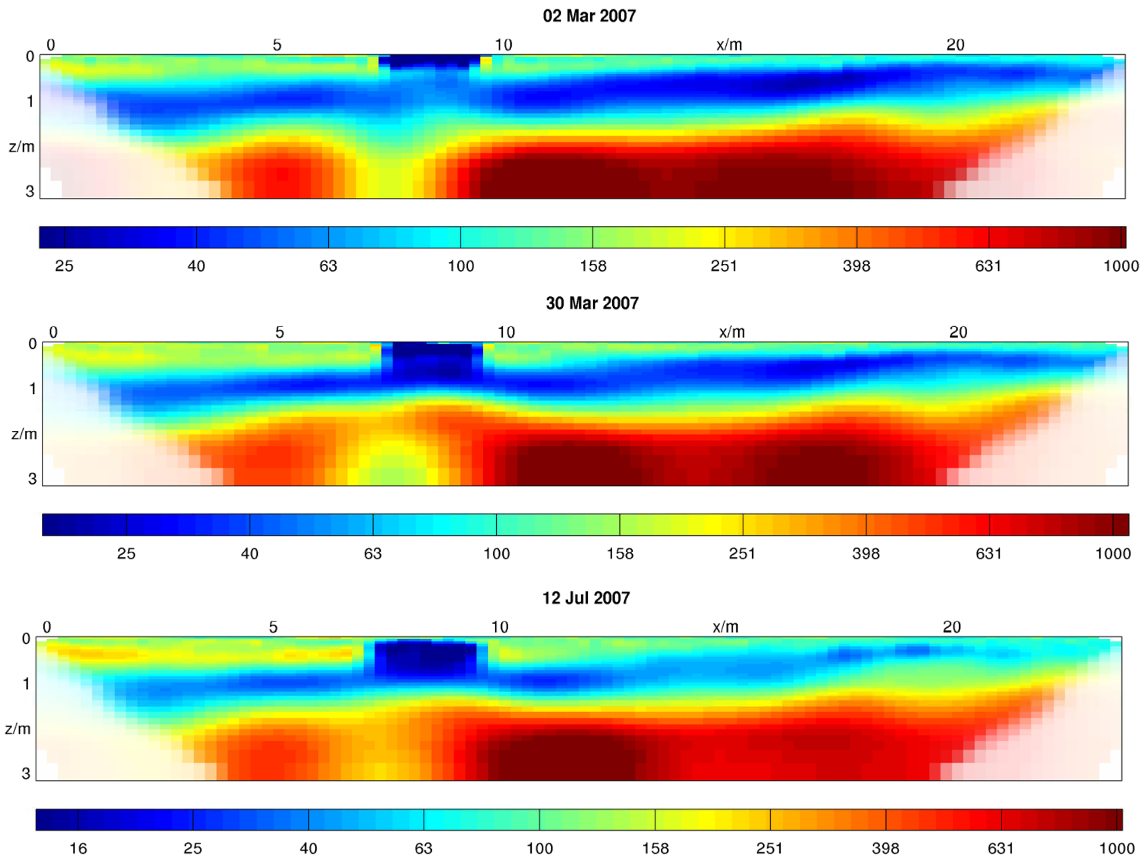


Figure 4.4: Three inverted ERT resistivity images [Ωm^{-1}] at different times.

(ii) the evaporating/transpirated water may actually cause a net flux upwards, taking tracer salt back into areas where it is not detected by the probes.

At the end of May to mid June, heavy rainfall occurred propagating the tracer further. From end May on, the second and third TDR probe cannot provide reliable water content information, as the tracer concentration is too high[4].

After that, tracer propagation slows down, most likely due to different soil material. Also, for the high summer temperatures, a partial water flux upwards due to evaporation and plant transpiration has to be taken into account.

ERT series ERT time series data was recorded twice a week initially, and weekly after. The data was evaluated by reference inversion. The image from fig. 4.2 is taken as a reference to a smoothness-constrained Gauss-Newton inversion. For the first time steps, a minimum length constraint actually showed better results, as the difference to the reference image is small. For the later images results were very unstable, so smoothness constraining was used for the complete series for consistency. The same is true for previous-step reference

---

[4] This problem does not affect the topmost probe, as it is doubly installed, with one probe residing inside and one probe outside the tracer area

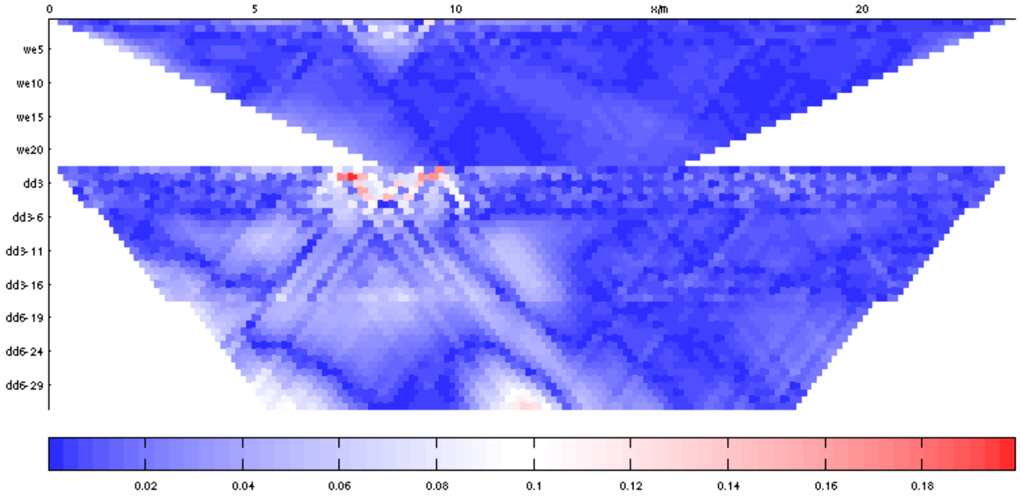


Figure 4.5: Misfit of forward response and measured data for a WDD data set. The upper part shows the Wenner pseudosection, the lower part the Dipole-Dipole pseudosection.

inversion, i.e. taking the result of the previous time step as a reference image. As the drastic changes happen very fast, caused e.g. by heavy rainfalls, this method was shown to pose no advantage over a fixed reference image, and only introduced additional instability.

A selection of ERT resistivity images is depicted in fig. 4.4. It can be seen that while the tracer is clearly visible under the background of the upper layers, the transition to the loamy layer in the middle tends to merge with the tracer, which is unrealistic when compared with TDR data. Also, when displaying the misfit vector as a pseudosection, it can be seen that most of the misfit originates from the area beneath the tracer (fig 4.5).

While this feature might be real, e.g. introduced by a preferential flow channel which is not present in the TDR area, this is rather unlikely, as through the ERT 3D-integration the channel would have to be large to be noticed. Therefore, it is assumed to be an inversion instability. Two effects play an important role here :

First, the smoothness constraint discourages the formation of a small band of high resistivity between the tracer and the clay layer with high conductivity. This effect is hard to deter; it has been tried to lower the constraint weight in the tracer area, but which leads to infeasible oscillations. Robust constraining helps a little, but also enforces rectangular shaping which leads to other artifacts.[5].

Second, the information density in the affected area is very low, as the current flow concentrates in the tracer and the middle layer. Here, a borehole-borehole configuration could provide a better data situation.

---

[5] It should be noted it is possible to get the expected results by some engineering, as seen in some ERT literature : Applying an engineered constraint weight and/or selective robust constraining in the problematic area does produce results as expected. However, such constraining approaches which are not based on hard data are discouraged The method of ERT constraining derives its credibility from the “simplicity” assumption, which is infringed here, leading to a situation were almost every expected result can be engineered.

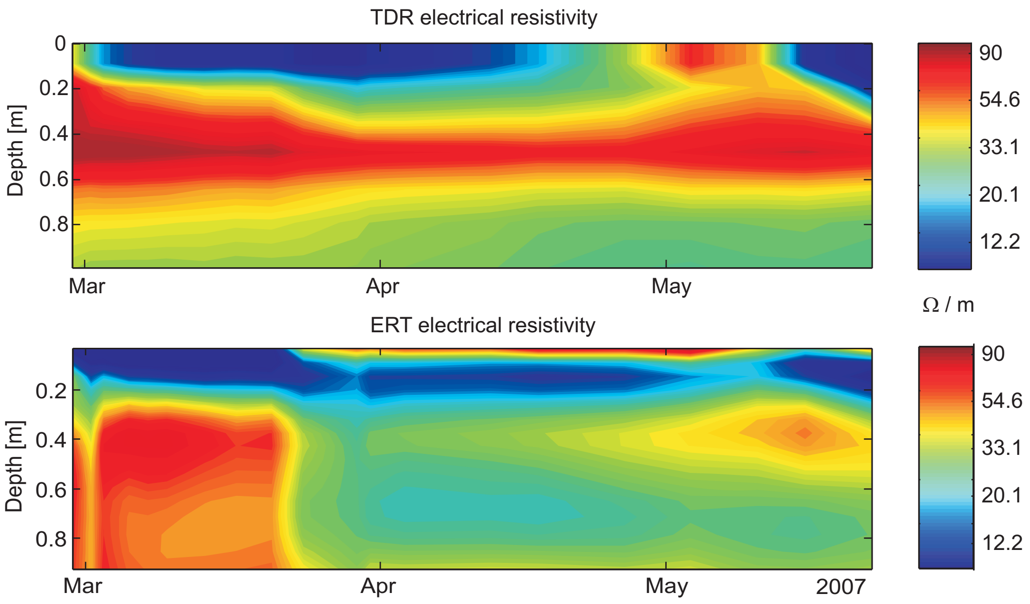


Figure 4.6: TDR and ERT resistivity [Ωm^{-1}], averaged over the tracer region, in the time window of continuous ERT measurements.

Another effect that can be observed in the ERT images is the “hole” in the bottom layer beneath the tracer area, originating from the TDR profile 2 m away. While the TDR probes themselves are unlikely to produce any measureable influence on the ERT data, their connection to a common grounding effectively induces a shortcircuit in the ERT potential field. As the profile is in 2 m distance, mainly the measurements with a large separation are affected, which are also provide the information for the lower layers. The upper layers remain undisturbed, so the tracer monitoring is not affected.

TDR / ERT comparison In the next step, consistency of ERT and TDR is tested. As the ERT grid is matched to the TDR probe locations, the measured TDR resistivities and the respective inverted ERT resistivities can be compared directly. For 7 TDR probes and the dates of ERT measurement, these values are crossplotted in fig. 4.7. It can be seen that while most probes fit acceptably with the ERT data, especially the probe at 37.5 cm shows a strong deviation. This matched the observation that the ERT inversion has problems forming the area below the tracer, as described in the previous paragraph. With a change in the ERT inversion parameters, this crossplot changes dramatically, indicating the instability of the ERT inversion in the tracer area.

To examine the full temporal process, ERT and TDR resistivity sections have been plotted in the time windows of regular ERT measurements in fig. 4.6. The image is focused on the upper 1 m. The ERT data is averaged in x-direction over the tracer area. Three major discrepancies can be found between the ERT and the TDR data :

(i) The shallow high-resistivy layer in the ERT data starting in late March cannot be observed by TDR. This is simply because the first TDR probe sits in 10 cm depth, so TDR cannot see it.

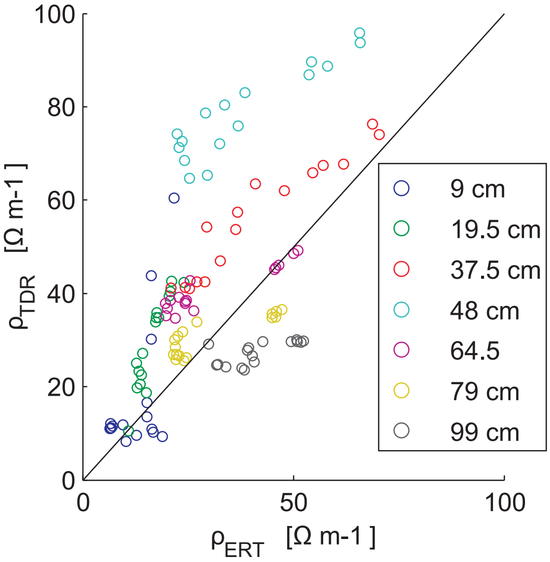


Figure 4.7: ERT and TDR resistivity crossplot.

(ii) The second feature is the high-resistivity layer between 40 cm and 60 cm, which is not present in ERT data starting from late March. This instability has been discussed in the previous paragraph.

(iii) The high resistivity break-in in early May is less pronounced in ERT as it is in TDR. One explanation could be that water content in the upper layers is smaller at the TDR probe location as the ERT area, raising the effective resistivity for TDR. This is plausible, as most of the vegetation in the tracer area died due to high salt concentration. The TDR probes are located on the border of the tracer area, and are partly affected by the transpiration of the surrounding grass. The uppermost TDR probe for water content is complete outside the tracer area, facing significantly more transpiration.

TDR Constraining TDR data can be used to constrain the ERT inversion further, by fixing TDR resistivity values at the respective cells (see paragraph 3.2.3). Two example pictures can be seen in fig. 4.8. These images match the fist two images in fig. 4.4. It can be noted that in the upper image an improvement in stability is achieved by the constraining. In the lower image, constraining in the high-resistivity band is not sufficient though, which can be seen in that the constrained cells (turquoise strip) don't merge with the area left and right. To quantify the quality improvement by constraining, $\chi 2$ is plotted for the TDR-constrained and unconstrained case over the entire time series (fig 4.9). Here, also the effect of Focusing is tested: In the curves denoted with WDD, the data has been stripped to the base Wenner-Dipole-Dipole set before the inversion. It can be seen that the focused set with TDR constraining shows the smallest misfit, as expected.

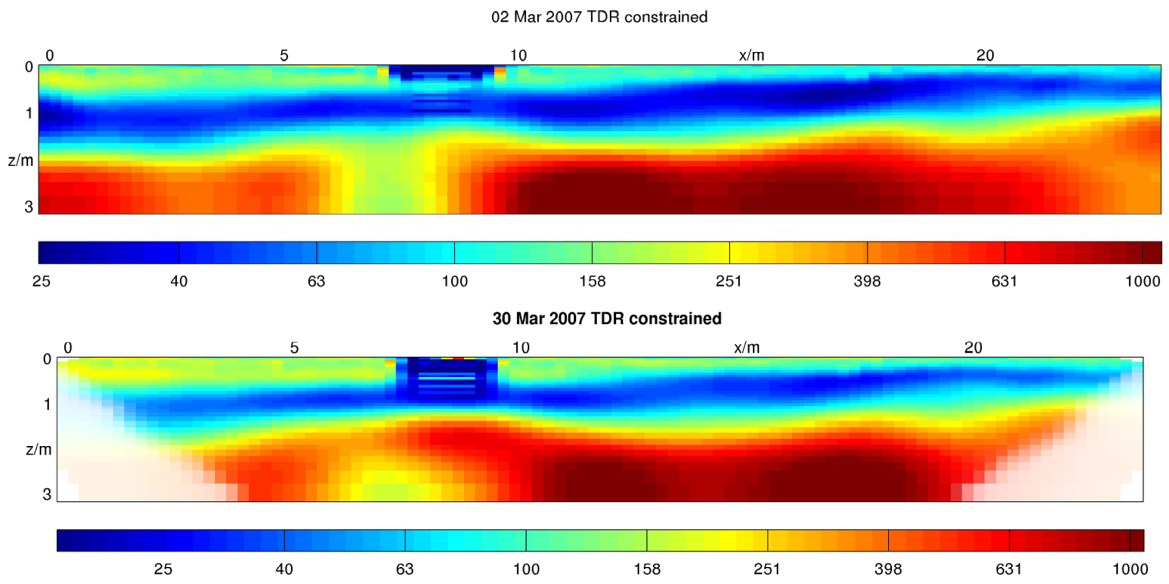


Figure 4.8: Two example inverted ERT images [Ωm^{-1}], using hard TDR constraining.

Modeling

As the ERT inversion leads to several problems, integrated inversion (see ch. 3.2.4) was applied, and shown to succeed as a proof of principle. It was calibrated using a simple transfer function, further improvement will be possible by thourough testing and calibration.

A simple transfer function The coupling between SWMS output and ERT, TDR input, i.e. the transfer function from calculated solute concentration to measured electrical conductivity, is problematic in pratice. While there are empirical and model-derived descriptions for this function, they could only be validated with good results in lab experiments. In the field, the complex dynamics of temperature, water content and the different electrical transport processes often lead to different descriptions, depending on the site *Boike and Roth*. As pointed out *Boike and Roth*, a calibrated linear transfer model might actually provide better results than the multivariable equations by van Loon or Mualem-Friedmann (*Mualem and Friedmann* [1991]) in the field. Therefore,

$$\sigma_T = A\theta\sigma_w = B\theta c_w$$

is used to describe the tracer. Together with the background conductivity σ_B not related to the tracer, the function reads

$$\sigma = \sigma_B + A\theta\sigma_w = \sigma_B + B\theta c_w \qquad .$$

Calibration will be done for the reference temperature of 25°, so θ_B and σ will have to be temperature corrected by

$$\sigma^{25} = \sigma(T) + \alpha_T\sigma(T)(T - 25\,\mathrm{K}) \qquad .$$

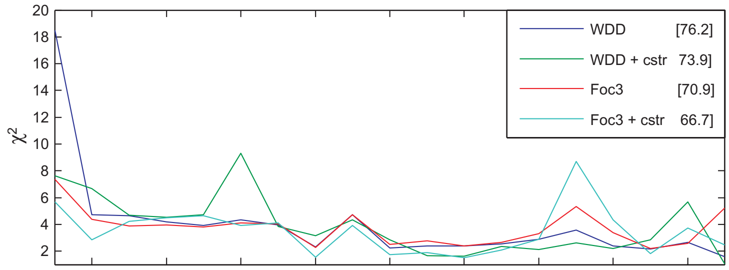


Figure 4.9: χ^2 of different configurations for the complete ERT time series.

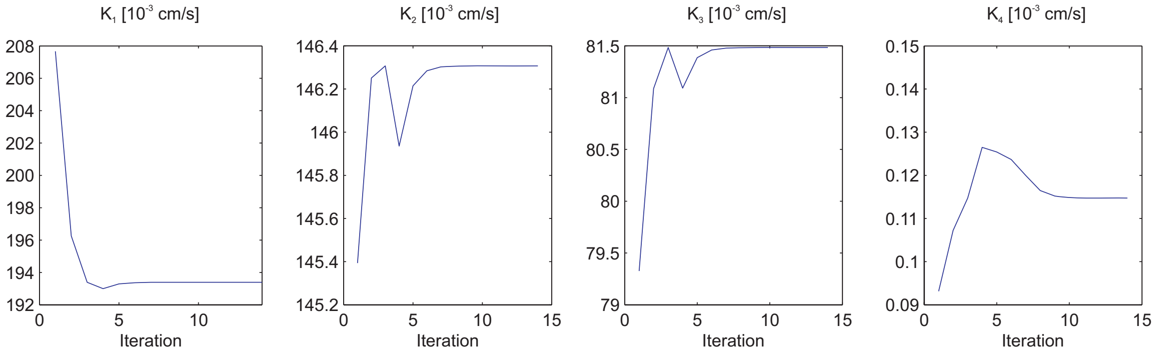


Figure 4.10: Development of the hydraulic conductivities during the inversion process.

For simplicity, the constant reference model is used as a background conductivity model. The error introduced by this assumption is constant for all configurations, so this should not affect the inversion. The calibration constant B is estimated using the ERT set where the tracer peaks at 20 cm, and a matching SWMS forward calculation. This way, B is estimated as

$$B \approx 4.65\,10^{-4}\frac{Sm^4}{g} \qquad .$$

Inversion of hydraulic conductivity For testing, only the hydraulic conductivities were inverted, as they have a significant influence on the dynamics, while being some of the hardest parameters to determine. Here, the other parameters are assumed to be correct, or at least close enough to the real parameters. With the addition of the calibrated transfer function, the same setup and parameters as in ch. 3.2.5 are used.

The result can be seen in fig. 4.10. After a few iterations, the hydraulic conductivities stabilize. The evaluate the quality of the result, the calculated configuration was compared with the measured data for TDR water content and electrical conductivity in fig. 4.11. It can be observed that the water content is represented quite well except that the decrease in the last few days is not observed in the simulated model. Likewise, the tracer propagation looks similar but happens faster in the simulation results. One possible explanation for this behavior is that the water content decrease is mainly caused by evapo-transpiration which

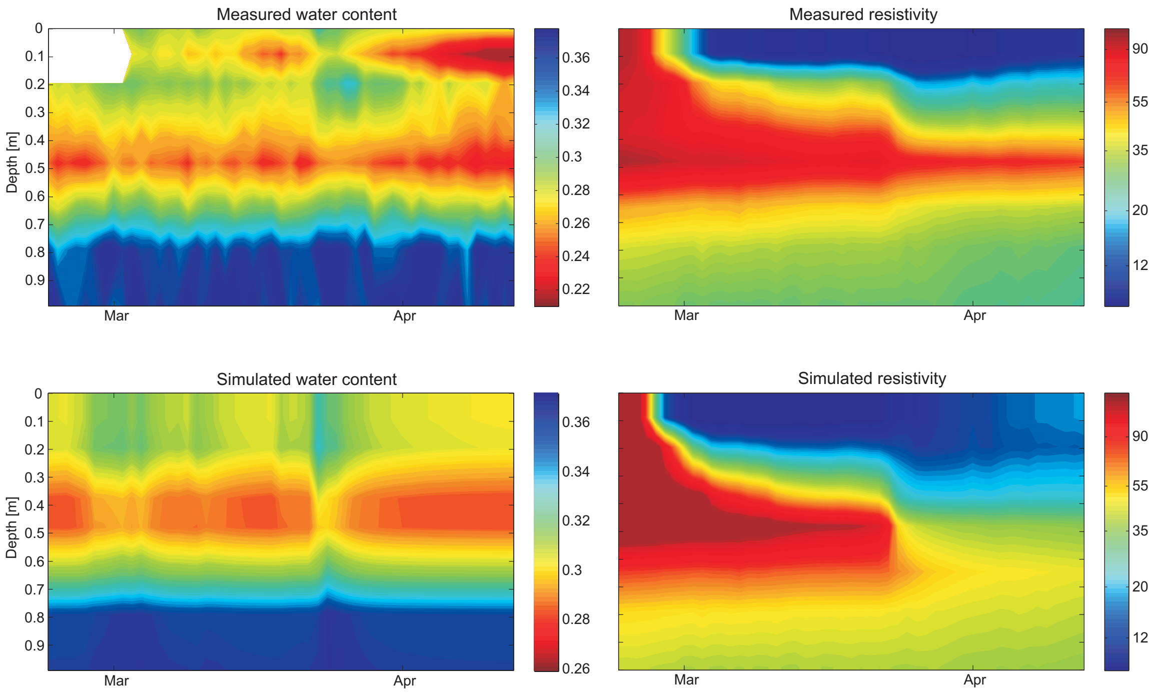


Figure 4.11: Comparison of measured and inverted results for TDR water content and el. conductivity

is not yet accounted for in the numerical model[6]. The missing upwards flux component causes a faster water flow, and therefore faster tracer propagation in the simulation. Another point is that the used time range is limited. If a longer time window, or more precisely, a larger amount of total precipitation is are used, a better data significance can be reached. In this case however, extending the time range to April/May would add only little precipitation while intensifying the evapo-transpiration-induced problems. In this sense, the chosen time window is a compromise.

To conclude, the first results look very promising. It has been shown that the method can be applied to real field data with stable results. In contrast to the TDR water content only-inversion, the tracer itself is monitored, and therefore the propagation can be controlled more reliably[7]. Adding support for evapo-transpiration and a more advanced transfer function is most likely key to further improving the quality.

---

[6] Evapo-Transpiration is the only effect that can create a flux upwards in the model, therefore a better match is not possible with the given set.

[7] In fact, the TDR only inversion showed an even faster tracer propagation for this data set

Chapter 5

Conclusion and Summary

In this thesis, a toolset of methods for the use of ERT in the context of hydrogeophysics was developed, including new algorithms for ERT optimization and integration of multimethods data into an hydraulic inversion. The toolset has been tested with synthetic and experimental data.

ERT data processing As a core element for ERT-based evaluations, a numerical forward solver was implemented. It solves the DC electric PDE using finite differences, and was tested against analytical solutions. Around the open inversion package Res2DInv, a toolset was build, to allow the automation of ERT data processing and data exchange with other applications.
To identify datasets with optimal resolution, measurement array optimization was studied, and an own lightweight algorithm implemented. The performance of this algorithm was tested against the algorithm of Stummer, with comparable results for the tested problems at significantly less computational complexity. This allows us to apply optimization to a large electrode set of 96 electrodes.

To be able to transfer optimized arrays to our measurement instrument, a tool was created. It also features an advanced self-devised algorithm to rearrage arrays for optimal usage of the multichannel capability.
Also, studies on the hitherto neglected problem of electrode spacing was conducted. The q-value as an indicator for the expected error level is introduced.

Hydrogeophysics Multimethods, i.e. the fusion of data from various geophysical methods, together with an hydraulic model, is identified as a key technique in hydrogeophysics. Various existing techniques have been categorized and transferred to a multimethods setting. Especially the question on how to constrain soft data with hard data has been addressed. Then, a completely new approach for an integrated hydraulic inversion is presented. It has the advantage that the ERT inversion can be eliminated altogether, therefore excluding one major source of instability. While it does hold the inherent problems of all inverse methods, it possesses maximal flexibility. Any weighted combination of data from different instruments on different time windows can be specified as input.

In a synthetic study, the new approach looks promising, and can resolve the hydraulic conductivities of an example model.

All methods mentioned in the chapter are implemented and integrated into the ERT toolset.

Experiment The methods developed are applied to a field experiment. On the Grenzhof test site, a tracer study was conducted. An optimized ERT array was set up, an ERT, TDR and GPR time series were recorded. An accuracy gain by ERT focussing could be observed, but was limit due to the high information content of the base array.

The layer depths obtained by multichannel GPR evaluation showed to be, where present, consistant with ERT. The method for ERT/GPR structure constraining was successfully tested on this data set.

The comparison of ERT and TDR showed partial agreement, the exception most likely being an ERT inversion artifact. The artifact could be mostly eliminated by hard constraining with TDR data for the first half of the time window. For the later half, the constraining was not sufficient.

Finally integrated hydraulic inversion was applied to the field data, with promising results. However, tracer propagation speed was slightly overrated, which can be linked to lacking evapo-transpiration support.

Outlook The topics covered in this thesis offer various possibilities for further research and refinement. The most promising extensions will be specified here.

Refinement of the integrated inversion The calibration conducted in this thesis could be refined, and support for evapo-transpiration could be added. It is likely that a significant improvement could be reached in this way. The addition of other data sources also seems interesting, especially the addition of a fullwave GPR forward solver.

3D ERT The current studies focused in 2D ERT measurements. While treatment of 3D ERT and hydraulic simulation is analogical to the 2D case in theory, in practice new aspects arise. For example, with growing measurement numbers, optimization will have a much greater impact as in 2D. Also, due to increasing measurement duration, the different temperatures on the single data points, affecting the conductivity, have to be taken into account. In the integrated inversion, this could be treated by timestamping.

Time-lapse inversion Another possibility to stabilize ERT inversion is time constraining. Here, the differences in time are minimized. To do this, the complete series has to be inverted in parallel.

appendix

Figure 1: IRIS Syscal Pro 96 (Picture from www.irisinstruments.com)

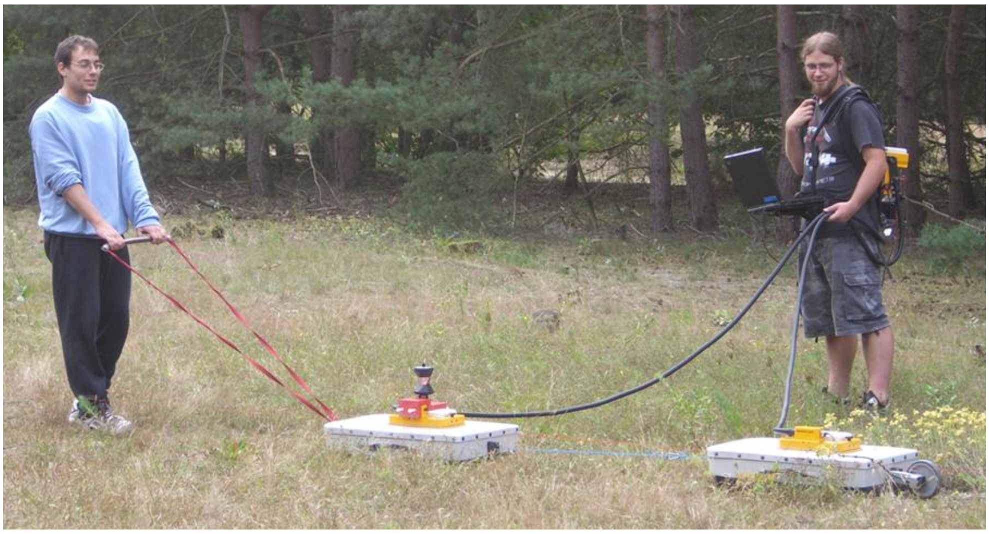


Figure 2: Multichannel GPR measurement

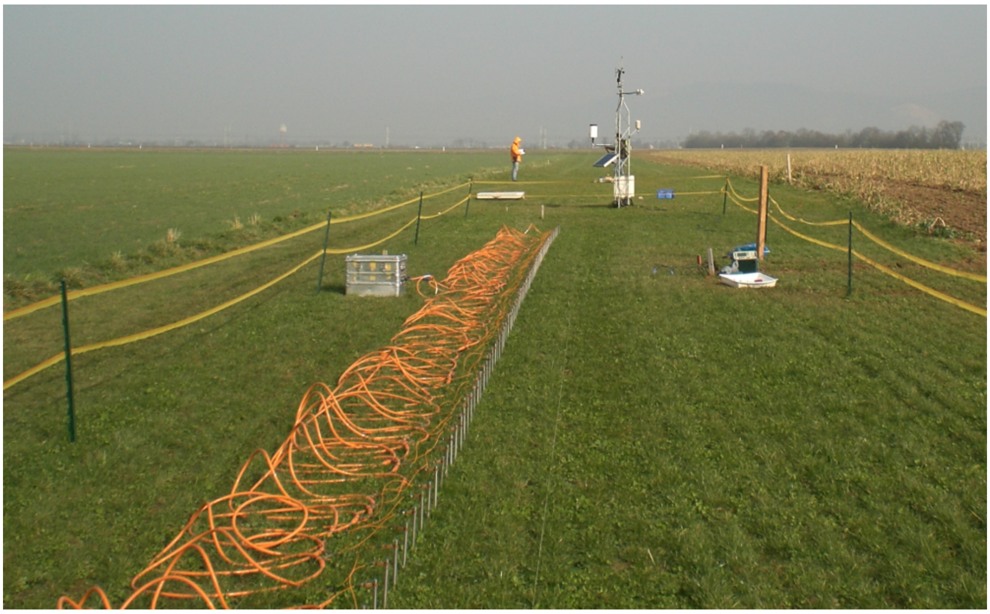


Figure 3: Grenzhof test site

Figure 4: Salt tracer application

Bibliography

Appel, D.: Transport Phenomena in Soils, Master's thesis, Institute for Environmental Physics, University of Heidelberg, 2007.

Benson, A. K., L. K. Payne, and A. M. Stubben: Mapping groundwater contamination using dc resistivity and VLF geophysical methods : A case study, *Geophysics* **62**, 80–86, 1997.

Boike, J., and K. Roth: Time Domain Reflectometry as a Field Method for Measuring Water Content and Soil Water Electrical Conductivity at a Continous Permafrost Site, *Permafrost and Periglacial Processes* **8**, 359–370.

Boulanger, O., and M. Chouteau: 3D modelling and sensitivity in DC resistivity using charge density, *Geophysical prospecting* **53**, 579–617, 2005.

Bretthauer, K.: Aufbau und Kalibrierung einer automatischen Boden-Wetterstation, Master's thesis, Institute for Environmental Physics, University of Heidelberg, 2004.

Chambers, J., O. Kuras, P. Meldrum, and R. Ogilvy: Electrical resistivity tomography applied to geologic, hydrogeologic, and engineering investigations at a former waste-disposal site, *Geophysics* **71**, 321–327, 2006.

Constable, S. C., R. L. Parker, and C. Constable: Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics* **52**, 289–300, 1987.

Davis, J. L., and A. P. Annan: Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy, *Geophysical prospecting* **37**, 531–551, 1998.

Dey, A., and H. F. Morrison: Resistivity modeling for arbitrarily shaped two-dimensional structures, *Geophysical Prospecting* **27**, 106–136, 1979a.

Dey, A., and H. F. Morrison: Resistivity modeling for arbitrarily shaped three-dimensional structures, *Geophysics* **44**, 753–780, 1979b.

Ernst, J. R., A. Green, H. Maurer, and K. Hollinger: Application of a new 2D time-domain full-waveform inversion scheme to crosshole radar data, *Geophysics* **72**, 153–164, 2007.

Friedel, S.: Resolution, stability and efficiency of resistivity tomography estimated from a generalized inverse approach, *Geophysics Journal Int.* **153**, 305–316, 2003.

Günther, T.: Inversion methods and resolution analysis for the 2D/3D reconstruction of resistivity structures from DC measurements, Ph.D. thesis, TU Freiberg, 2004.

Günther, T., L. Bentley, and M. Hirsch: A new joint inversion algorithm applied to the interpretation of dc resistivity and refraction data, in *Ext. Abstract, Workshop for Computational Methods in Water Resources, 18.-22.6.06 Kopenhagen(Denmark).*, 2006.

Hansen, P. C., and D. P. O'Leary: The use of the L-curve in the regularization of discrete ill-posed problems, *SIAM Journal of Scientific Computing* **14**, 1487–1503, 1993.

Heimovaara, T., A. G. Focke, W. Bouten, and J. M. Verstraten: Assessing temporal variations in soil water composition with time domain reflectometry, *Soil Science Society of America Journal* **59**, 689–698, 1995.

Hestenes, C. M. R., and E. Stiefel: Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur Stand* **49**, 409–436, 1952.

Huisman, J. A., S. S. Hubbard, J. D. Redman, and A. P. Annan: Measuring soil water content with ground penetrating radar : A review, *Vadose Zone Journal* **2**, 476–491, 2003.

Kemna, A., J. Vanderborght, B. Kulessa, and H. Vereecken: Imaging and characterization of subsurface solute transport using electrical tomography (ERT) and equivalent transport models, *Journal of Hydrology* **267**, 125–146, 2002.

Kowalsky, M. B., S. Finsterle, J. Peterson, S. Hubbard, Y. Rubin, E. Majer, A. Ward, and G. Gee: Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data, *Water resources research* **41**, 11,425, 2005.

Li, Y., and K. Spitzer: Three-dimensional DC resistivity forward modelling using finite elements in comparison with finite-difference solutions, *Geophysical Journal Int.* **151**, 924–934, 2002.

Linde, N., A. Binley, A. Tryggvason, L. B. Pedersen, and A. Revil: Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltime data, *Water Resources Research* **42**, 12,404, 2006.

Locher, F.: Vorlesung Mathematik für Informatiker, lecture notes, 1997.

Loke, M. H., and R. Barker: Least-square deconvolution of apparent resistivity pseudosections, *Geophysics* **60**, 1682–1690, 1995.

Loke, M. H., and R. Barker: Rapid least-square inversion of apparent resistivity pseudosections by a quasi-Newton method, *Geophysical Prospecting* **44**, 131–152, 1996.

Lowry, T., M. B. Allen, and P. N. Shive: Singularity removal: A refinement of resistivity modeling techniques, *Geophysics* **54**, 766–774, 1989.

Meju, M. A.: *Geophysical data analysis: Understanding inverse problem theory and practice*, Society of Exploration Geophysicists, 1994.

Menke, W.: *Geophysical data analysis: Discrete inverse theory*, Academic Press, New York, 1989.

Michot, D., Y. Benderitter, A. Dorigny, and A. Tabbagh: Spatial and temporal monitoring of soil water content with an irrigated corn crop cover using surface electrical resistivity tomography, *Water Resources Research* **39**, 1138–1154, 2003.

Mualem, Y.: A new model for predicting the hydraulic conductivity of unsaturated porous media, *Water Resources Research* **12**, 513–522, 1976.

Mualem, Y., and S. P. Friedmann: Theoretical prediction of electrical conductivity in saturated and unsaturated soil, *Water Resources Research* **27**, 2771–2777, 1991.

Noel, M., and R. Walker: Imaging archaeology by electrical resistivity tomography: a preliminary study, *Archaeological Sciences* 1991.

Park, S. K., and G. P. Van: Inversion of pole-pole data for 3D resistivity structure beneath arrays of electrodes, *Geophysics* **56**, 951–960, 1991.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery: *Numerical Recipes in C++: The Art of Scientific Computing*, Cambridge University Press, 2002.

Ritter, A., F. Hupet, R. M. noz Carpena, S. Lambot, and M. Vanclooster: Using inverse methods for estimating soil hydraulic properties from field data as an alternative to direct methods, *Agricultural Water Management* **59**, 77–96, 2003.

Robinson, D. A., S. B. Jones, J. M. Wraith, D. Or, and S. P. Friedman: A review of advances in dielectric and electrical conductivity measurement in soils using TDR, *Vadose Zone Journal* **2**, 444–475, 2003.

Roth, K.: *Soil physics. Lecture notes*, 2006.

Roth, K., R. Schulin, H. Flühler, and W. Attinger: Calibration of time domain reflectometry for water content measurements using a composite dielectric approach, *Water resources research* **26**, 2267–2273.

Sasaki, Y.: 3-D resistivity inversion using the finite-element method, *Geophysics* **11**, 1839–1848, 1994.

Schneider, K.: Novel evaporation experiment to measure soil hydraulic properites, Master's thesis, Institute for Environmental Physics, University of Heidelberg, 2005.

Schwarzbach, C.: Entwicklung eines genetischen Inversionsalgorithmus für die Geolektrik, Master's thesis, TU Bergakademie Freiberg, 2003.

Simunek, J., and T. Vogel: *The SWMS-2D code for simulating water and solute transport in two-dimensional variably saturated media*, 1994.

Singha, K., and S. M. Gorelick: Effects of spatially variable resolution on field-scale estimates of tracer concentration from electrical inversions using Archie's law, *Geophysics* **71**, 83–91, 2006.

Spitzer, K.: The three-dimensional DC sensitivity for surface and subsurface sources, *Geophysical Journal International* **134**, 736–746, 1998.

Spitzer, K., and B. Wurmstich: Speed and accuracy in 3d resistivity modeling, in *3D EM Symposium*, Ridgefield, USA, 1995.

Stummer, P., H. Maurer, and A. Green: Experimental design: Electrical resistivity data sets that provide optimum subsurface information, *Geophysics* **69**, 2004.

Ulbricht, C.: Monitoring Field Tracer Experiment with Ground Penetrating Radar and Time Domain Reflectometry, Master's thesis, Institute for Environmental Physics, University of Heidelberg, 2005.

van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Science Society of America Journal* **44**, 647–652, 1980.

Vanderorght, J., A. Kemna, H. Hardelauf, and H. Vereecken: Potential of electrical resistivity tomography to infer aquifer transport characteristics from tracer studies, *Water Resources Research* **41**, 6013–6025, 2005.

Vogel, H.-J.: *Transport gelöster Stoffe im Boden*, 2004.

Vogelsang, D.: *Geophysik an Altlasten*, Springer, Berlin, 1997.

Wait, J. R.: *Geo-electromagnetism*, Academic Press, New York, 1982.

Wollschläger, U., and K. Roth: Estimation of Temporal Changes of Volumetric Soil Water Content from Ground-Penetrating Radar Reflections, *Subsurface Sensing Technologies and Applications* **6**, 207–218, 2005.

Xu, S. Z., Z. C. Gao, and S. Zao: An integral formulation for three-dimensional terrain modeling for resistivity surveys, *Geophysics* **53**, 546–552, 1988.

Yeh, T.-C. J., S. Liu, R. J. Glass, K. Baker, J. Brainard, and D. Alumbaugh: A geostatistically based inverse model for electrical resistivity surveys and its application to vadose zone hydrology, *Water resources research* **38**, 1278–1290, 2002.

Zhao, S., and M. J. Yedlin: Some refinements on the finite-difference method for 3D dc resistivity modeling, *Geophysics* **61**, 1301–1307, 1996.

Zhou, Q. Y., J. Shimada, and A. Sato: Three-dimensional spatial and temporal monitoring of soil water content using electrical resistivity tomography, *Water Resources Research* **37**, 273–286, 2001.

Acknowledgements

First of all I want to thank Prof. Kurt Roth for the supervision of my project, many inspiring insights and though-provoking discussions in physics and beyond. He had faith in me and my strange ideas when I almost lost it. I also want to thank him for giving me the unique chance to join in the Xinjiang field campaign. I wish you'll be able to realize many of your ideas in the new job ! My gratitude also goes to Prof. P. Nielaba and Prof. W. Dieterich, who kindly agreed to review my thesis, even if the topic is not their field of study.

Thomas Günther owes my thanks for granting me access to his inversion code, which formed the base for all my calculations, and his brilliant dissertation which soon became my ERT bible. Also, it was nice to meet you in Heidelberg and Aachen !
Thanks also to Johannes Köstel, without whom we'd still be scratching our heads over the ERT/TDR artifacts.

I am grateful for the support I received from Ute Wollschläger, who helped me with the experimental stuff and soil-physics, that tended to drive me nuts.
Further thanks go out to the Linux-grandmaster Klaus for complicated solutions to complicated problems, the always helpful king of Taekwondo ‚Holger[1], Ms. Levin for the kind provision of her saucers :), the soil physics coffee connection and the Grenzhof troop – especially Angelika, who was always open to help, even with the boring stuff !

I also want to thank Julia, Julia & Hannes for keeping the rats out of my room when I was pushing night shifts again, and for enduring my experimental cooking.
Another factor that helped keep me alive was my theater group, for the direly needed distraction on Monday evening. Bühnenrausch, anyone ?

Finally, thank you Isa for being what you are (butsu!).

---

[1] whom I still can't visualize hitting people