NTNU – Trondheim
Norwegian University of
Science and Technology

Explicit, A Priori Constrained Model Parameterization for Inverse Problems, Applied on Geophysical CSEM Data

Ole-Johan Skrede

Master of Science in Physics and Mathematics
Submission date: August 2014
Supervisor: Elena Celledoni, MATH
Co-supervisor: Jan Petter Morten, EMGS

Norwegian University of Science and Technology
Department of Mathematical Sciences

ABSTRACT

This thesis introduce a new parameterization of the model space in global inversion problems. The parameterization provides an explicit representation of the model space with a basis constrained on a priori information about the problem at hand. It is able to represent complex model structures with few parameters, and thereby enhancing the speed of the inversion, as the number of iterations needed to converge is heavily scaled with the number of parameters in stochastic, global inversion methods. A standard Simulated Annealing optimization routine is implemented, and further extended to be able to optimize for a dynamically varying number of variables. The method is applied on inversion of marine CSEM data, and inverts both synthetic and real data sets and is able to recover resistivity profiles that demonstrate good resemblance with provided well bore log data. The trans-dimensional, self-parameterizing Simulated Annealing algorithm which is introduced in this thesis proves to be superior to the regular algorithm with fixed parameter dimensions.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

INTRODUCTION

In this thesis we introduce a new parameterization of the model parameter space in an inversion problem. It is applied on inversion of geophysical CSEM (controlled source electromagnetic) data, but the general idea also applies to other problem instances. Generally speaking, a set d^{obs} of data is observed, the task is then to find a model m that through the relation $d = \mathcal{G}(m)$ reproduce this observed data, where $\mathcal{G}$ is an operator which describe the explicit relationship between the model and the resulting data. The best model is found by an optimization routine which for some norm $||\cdot||$, minimizes the distance $||d - d^{obs}||$.

The model m is in this case the resistivity at a given set of locations below the seabed. The governing equations, represented by the operator $\mathcal{G}$, that relates the model to the data is *Maxwell's equations*. The observable physical phenomenon is the resulting electromagnetic field from which the data d is obtained.

Assume that the model consists of n locations at which the resistivity is to be decided, assume further that each location is treated independently such that in order to find the best model, one need to optimize n parameters. Now, assume that the stochastic optimization routine has a complexity such that it requires $\mathcal{O}(n \log n)$ function evaluations before termination. As a modest example, consider a 1D model which consist of $n = 100$ locations where the resistivity is to be decided. A 2D square model with similar grid spacing would then consist of $n = 10^4$ unknowns, whereas a cubic 3D model would consist of 10^6 unknowns to be decided. If one assume that the function evaluation is of $\mathcal{O}(1)$ (extremely modest) and that the elapsed time in the 1D inversion is 1000 s, then the elapsed time of the full 3D inversion would be about 347 days.

Now, imagine that you could represent the model in terms of $\tilde{n} << n$ parameters, say $\tilde{n} = 10$ in the 1D case, $\tilde{n} = 100$ in the 2D case and $\tilde{n} = 1000$ parameters in the 3D case. Then the full 3D inversion would be finished in about 4 hours, or 0.05 % of the original time. Although just a superficial thought experiment, it alludes that a reduction in the parameter dimensionality has a great potential in terms of reducing computational costs. In addition, function evaluations is much more costly in 3D inversion than in 1D inversion. For a global inversion in three dimensions to be computationally feasible, the number of function evaluations need to be reduced. It is worth mentioning that the analysis above applies to stochastic or global inversion schemes as the number of iterations needed scales heavily with the number of parameters. Local inversion schemes with gradient based opti-

mization routines, which is normally used in 3D CSEM inverion, does not scale that heavily with the number of parameters.

This dramatic reduction of parameters is the main motivation behind this parameterization, and what makes it different from plane layered model representations used in e.g. (Roth and Zach, 2007) and (Ray and Key, 2012). Models which in a plain layered model is represented with about 30 parameters, could be represented with less than 10 parameters with this new parameterization. Also the geologically restricted basis makes this representation intuitive and physically meaningful, which is a desired property.

This thesis propose a parameterization of the model space which in an intuitive and explicit way, utilize *a priori* information of the original model to create a restricted basis on which the parameterization is built. With this, the objective is to significantly reduce the number of parameters and still maintain a good representation of the model.

This parameterization is applied on inversion of CSEM data. Marine CSEM is a method which maps the resistivity profile beneath the seabed and will be further explained in Chapter 2. A 1D model of the sub seabed is assumed, and the actual parameterization of this one dimensional sub seabed model is presented in Chapter 3. Chapter 4 discusses the underlying physics given this model assumption, and derives the governing equations which relates the resistivity model to the resulting electromagnetic field. In order to find the models which minimizes the misfit, a *Simulated Annealing* optimization routine is chosen, the discussion of which is addressed to Chapter 5. This concludes the first part, which aims to establish a sufficient theoretical background to cover the topics presented in this thesis.

While part one lay the theoretical foundation of what is discussed in this thesis, part two is devoted to how the task is carried out. Practical issues concerning implementation and other solution procedure issues is discussed in Chapter 6, while results and accompanying discussion is presented in Chapter 7.

The appendix will provide necessary information in terms of different algorithmic parameter values which goes in the attached program code, in order to be able to completely reproduce some of the results presented in this thesis.

Part I

THEORETICAL BACKGROUND

2

THE CONTROLLED SOURCE ELECTROMAGNETIC TECHNIQUE

The marine *controlled source electromagnetic* (CSEM) method is a geophysical technique with many applications. It was first used to measure basalt properties in the Pacific ocean. In the petroleum industry, its main purpose is to expose accumulations of hydrocarbon by mapping the resistivity below the seabed. The resistivity in the various sedimentary layers below the seabed is obtained through inversion of CSEM data. Supplemented with seismic techniques, it serves as a great tool for subsurface hydrocarbon exploration. Commercially, in the petroleum industry, it is a relative new field, starting around year 2000 (Eidsmo et al., 2002). Since then, a great number of surveys have been successfully conducted (Constable, 2010). For a historical overview and development of the method, the reader is referred to (Constable, 2010) and (Constable and Srnka, 2007).

A detailed description of the sub seabed resistivity profile is interesting because of the difference in resistivity in the different sub seabed formation layers. It is particularly interesting because of the large resistivity contrast between hydrocarbon saturated reservoirs and the surrounding sedimentary layers saturated with aqueous saline fluids. The marine CSEM method utilizes this fact and make it a valuable tool in the petroleum industry.

Sea water is quite conductive with a conductivity of about 4 S/m, whereas the formation beneath the seabed has variable resistivity values of order 1 Ωm to 10 Ωm (remember the reciprocal relationship between the resistivity ρ, and the conductivity σ, $\rho = 1/\sigma$). Hydrocarbon saturated reservoirs will have resistivity of order 10^2 Ωm, significantly more resistive than the surrounding sediments. Air is also is very resistive with a resistivity of order 10^{16} Ωm, and is therefore a major contributor to the field response detected by the receivers.

However, the sub seabed resistivity profile is not easy to obtain, but can be mapped indirectly by inversion of CSEM data. In summary, the electromagnetic field responses are measured at the desired locations. Then, a guess is made on how the resistivity profile is, and calculates the field response at the same locations due to this guessed model. With both the observed field data and the synthetic field data, a comparison is made of those two and difference is calculated using some misfit function. An optimization routine is used to minimize this misfit function, and this optimization routine will guide the search of new models to a final model which produce an electromagnetic field response with minimum misfit. This minimum misfit

model is then accepted as the best approximation to the true model based on the observed field.

2.1 CSEM DATA SAMPLING

Before any data analysis and interpretation can be done, one need to gather data. Two different electromagnetic surveying methods are the magnetotelluric (MT) method and the controlled-source electromagnetic (CSEM) method. In the MT method, the naturally occurring electric and magnetic field is measured. This low-frequency electromagnetic field is generated by the solar winds as it interacts with the earth's magnetic field. It has low resolution, but penetrates greater depths than the CSEM method (Constable and Weiss, 2006). Contrary to the naturally occurring electromagnetic field used in the MT method, the CSEM method uses an electromagnetic field generated from a controlled source.

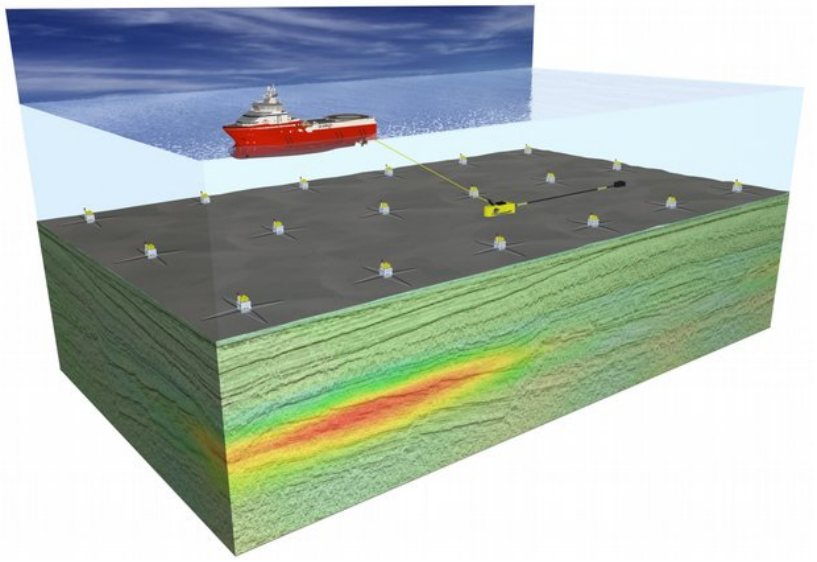


Figure 2.1: Illustration of CSEM-sampling.

Figure 2.1 illustrates the basic concept. A vessel is towing an electromagnetic dipole antenna (source) over a grid of detectors (receivers) stationed at the seabed. The dipole source emits low frequency electromagnetic waves into the surroundings, both up towards the sea surface and down into the seabed. The the electromagnetic waves penetrate and interact with the matter it propagates through, and the resulting field is detected at the seafloor by receiving dipoles (Eidsmo et al., 2002).

The source is a dipole antenna which is towed after the vessel about 30 m above the grid of receivers. This source emit a continuous signal, and the source positions referred to in this paper corresponds to the mid of the time interval in the fast Fourier transform.

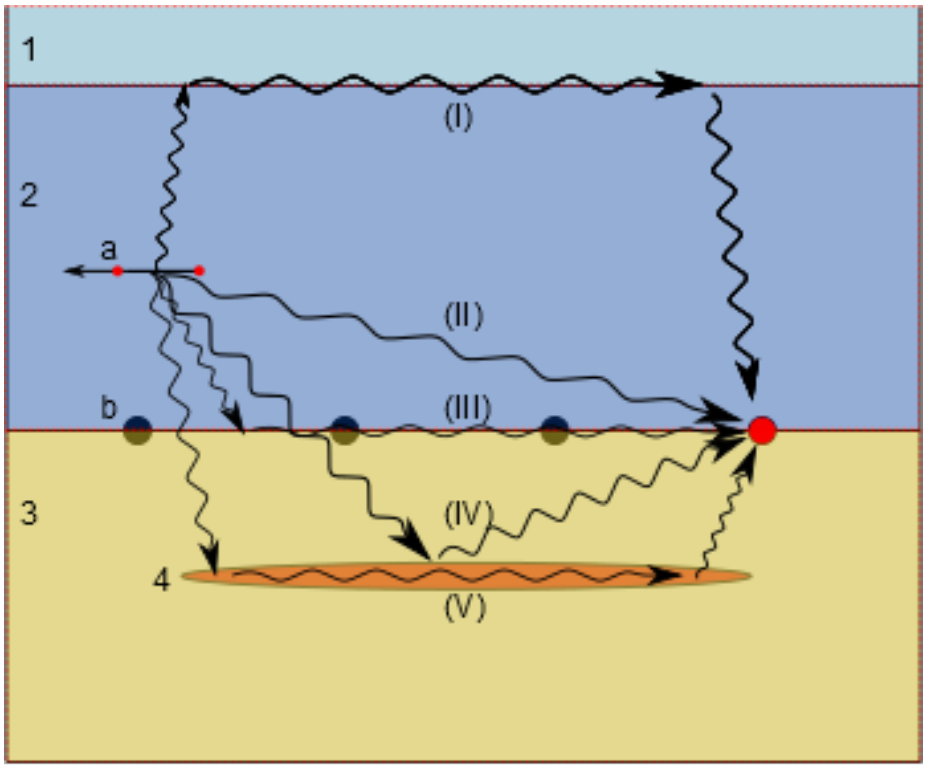


Figure 2.2: A 2D Illustration of signal propagation from dipole source (a) to receiver (b, red). A dipole antenna (a) is towed through the sea water, emitting electromagnetic radiation which propagates through and interact with different media before the response is observed by the receiver (b, red)). This figure illustrate the main contributions to the total response; sea-surface propagation (I), direct field (II), seabed (III), target reflection (IV), and target guiding (V). The different media present is air (1), sea water (2), sub sea formation (3) and the resistor (4).

Different signal propagation paths from the source to one of the receivers are illustrated in Figure 2.2. With reference to this figure, the different contributions are: the sea surface response (I), the direct field (II), the lateral wave along the seabed (III), the ray reflected from the thin layer (IV), and the guided wave in the resistive reservoir (V). These are the main contributors, but their contribution varies in different situations. At short offsets (< 2 km) between the source and receiver, the direct field (III) dominates the received signal, whereas at longer offsets the dominating contributions are those from the sea surface (I) and the resistor (V) (Løseth, 2007). If the reservoir has a low resistivity, the subsurface waves tend to pass through the reservoir instead of being refracted back up to the receivers. If the resistor (often termed the target) is not present, there will be no guiding effect of the signal. A model without a resistive reservoir is often termed the background model.

The ability to detect a hydrocarbon saturated reservoir is highly dependent on the thickness and resistivity of the reservoir. Sometimes it is diffi-

cult to differ between the two contributions, this is in situations when the detection ability is very sensitive to the product of the two parameters. This product is sometimes termed *transverse resistance* or simply resistivity-thickness-product (Ansari et al., 2012).

2.2 INVERSION

The overall objective is to map the resistivity profile below the seabed, something that is not directly obtained by measuring the field. One have to find a resistivity model that sufficiently reproduce the observed data, and the assumption is that this model is a good candidate for the true model. Finding this model is done by inversion, and this problem can be stated as: given the observed data, find a model that best explain the observed data. In general, there is several solutions to this problem, i.e. the solution is not unique and the problem is therefore ill-posed. The next paragraph will try to give a simple and superficial description of the process.

One begin with an initial model, that is, a guess of the resistivity profile below the seabed. The number of locations where one wants to calculate the resistivity is thus the number of unknowns in this problem, and throughout this project, this number is denoted N_l with the subscript l because of the plane layered model assumption used in this thesis (see Chapter 4). This model is then used as input in the so called forward processing engine which solves a system of differential equations in order to determine the field response given this model. This field is often labeled F^{Synth}, as opposed to the observed field which is labeled F^{Obs}, here F is chosen as a generic representation of the different field components. The exact nature of the field F, is explained below, but this general field is now used in the derivation for simplicity. In the synthetic model used in the simulation, the set up, i.e. source and receiver positions, transmitter frequency etc. is the same as when the observed data was acquired. The created synthetic data is then compared with the observed data, and depending on the fit, the model is accepted or rejected. The fit, or equivalently, the misfit is measured with a misfit function, and an optimization routine is used to minimize this misfit function. If rejected, the optimization routine determines a new resistivity model which in turn is given as input in the forward engine. This loop continues until a satisfactory model is found.

In the following chapters, details concerning the model parameterization, the forward solver and the optimization will be further explained.

3

MODEL PARAMETERIZATION

This chapter explains how the sub seabed resistivity profile is parameterized. That is, how to represent the unknown model which is to be determined. This discussion will be about a 1D model, but the arguments is easily extended to 2D or 3D. The goal is to determine the resistivity at a given number of locations below the seabed. One approach would be to determine the resistivity at a evenly spaced set of locations. Say, from $z = 500$ m to $z = 2500$ m with a spacing $h = 10$ m would result in $N = 201$ parameters to be determined. In two dimensions, a 2000 m $\times$ 2000 m square with a 10 m squared grid spacing would result in 40401 free parameters, and for a similar cube, 8120601 free parameters. As is evident, this is a vast number of unknowns to be calculated at each iteration of the optimization, so a smart parameterization of the model space could be highly beneficial with respect to lowering the computation cost.

The idea presented in this thesis is to represent the model space as a function $\rho(z)$, where the parameters needed to determine this function are the free parameters to be optimized. The total number of free parameters used in this parameterization is denoted N. The form of the function is based on prior knowledge about the general geology beneath the seabed, with the purpose of keeping the number of parameters low whilst maintaining a good representation of the resistivity profile.

The potential of a significant reduction of forward calls in the inversion is the main motivation behind this idea, a reduction which would certainly increase with the model dimensionality. Also, since the basis is restricted by a priori information about the sub seabed formation, it gives a intuitive and physically meaningful representation of the geology.

3.1 RESISTIVITY PROFILE FUNCTION

This parameterization of the sub seabed consists of three types of terms. A polynomial term, a set of *Heaviside step functions* and a set of *Gauss functions*

$$\rho(z;\mathbf{x}) = \rho_p(z;\mathbf{x}_p) + \rho_r(z;\mathbf{x}_r) + \rho_g(z;\mathbf{x}_g). \tag{3.1}$$

3.1.1 *Polynomial term*

The polynomial will have the form

$$\rho_p(z;\mathbf{x}_p) = \sum_{n_p=0}^{N_p} c_{n_p}(z - z_{wd})^{n_p}, \quad N_p \geq 0, z > z_{wd}, \tag{3.2}$$

where the fixed constant z_{wd} is the water depth and $\mathbf{x}_p = (c_0, \ldots, c_{N_p})^\intercal$. The coefficients c_{n_p} are thus the free parameters, resulting in $N_p + 1$ parameters to be determined. This polynomial is included mainly because of the contribution of the two first terms in the sum. The constant term to shift the resistivity in the whole model, and the 1st order term to represent a possible *compression trend* which is that due to gravity, the resistivity is increasing with increasing depth. This compression effect is due to the fact that the increase of pressure will generally lead to an increase in resistivity (Zhang, 2011). Higher order terms is also possible to include, and together, these polynomial terms will represent the trending background model. For practical purposes $N_p \geq 0$ which imply that the constant term is always included. This constant c_0, is bounded below by some reasonable positive value, e.g. the water resistivity. If included, the higher order terms are allowed to attain negative values. Although not necessarily geologically meaningful, the first order term is also allowed to attain negative values for the sake of generality.

3.1.2 *Resistor term*

$$\rho_r(z;\mathbf{x}_r) = \sum_{n_r=1}^{N_r} \rho^r_{n_r}[H(z - z^r_{n_r}) - H(z - z^r_{n_r} - t)], \quad N_r \geq 0, z > z_{wd} \tag{3.3}$$

where $\rho^r_{n_r}$ is the resistivity, $z^r_{n_r}$ is the upper boundary of resistor number n_r and $\mathbf{x}_r = (\rho^r_1, \ldots, \rho^r_{N_r}, z^r_1, \ldots, z^r_{N_r})^\intercal$. H is the *Heaviside step function* and t is the thickness of the layers, which is fixed. One could of course also optimize for the thickness of the resistor layer, however, because of the transverse resistivity relation discussed earlier, this is deprecated since it would just complicate the problem without adding particularly useful information. This yields $2N_r$ free parameters to be determined. These terms are included to model possible thin, hydrocarbon saturated layers in the sub seabed formation. By thin it is meant that we make the fixed parameter t small comparable to typical hydrocarbon target thickness. The Heaviside function is appropriate since resistivity can differ with several orders of magnitude within just a few centimetres if one encounters a resistive layer. The number of resistor terms to include is N_r and could be zero, in which case this term would completely disappear. The resistivity parameters is constrained as $\rho^r_{n_r} \in [\rho^{\min}_r, \rho^{\max}_r]$, where $\rho^{\min}_r$ is some small, positive value, and $\rho^{\max}_r$ is

some reasonable upper bound. The depth parameters is constrained by the model size such that it can only attain values that is physically meaningful within the framework of the physical 1D model space.

3.1.3 *Gauss function term*

$$\rho_g(z;\mathbf{x}_g) = \sum_{n_g=1}^{N_g} \beta_{n_g} \frac{1}{\sqrt{2\pi w_{n_g}^2}} \exp\left\{ -\frac{1}{2w_{n_g}^2}(z - z_{n_g}^g)^2 \right\}, \quad N_g \geq 0, z > z_{wd} \tag{3.4}$$

where β_{n_g} is an amplitude amplifying factor, w_{n_g} is a measure of the width of the bell curve, and $z_{n_g}^g$ is the depth at which the bell curve is centered, $\mathbf{x}_g = (\beta_1, \ldots, \beta_{N_g}, z_1^g, \ldots, z_{N_g}^g, w_1, \ldots, w_{N_g})$, that is, $3N_g$ free parameters. The Gauss terms is included for smoothing purposes, but also to model layers of the geology which resistivity differs from the background. As for the resistor terms, one could choose not to include any Gaussian terms. The amplifying factors β_{n_g} are unbounded, but for practical reasons, some lower (possibly negative) and upper bounds are set. The width is naturally bounded below by some positive value, ideally greater than t, but this was not implemented. For practical reasons, it is also bounded above by some large value. The center of the Gauss function is bounded by some reasonable bounds, but contrary to the resistor depth, these parameters can attain values outside the boundaries of the physical 1D model, the reason for this is simply to be able to include small parts of one of the tails of the bell curve in the resistivity profile.

3.1.4 *Final representation*

The resulting resistivity profile is then governed by

$$\begin{aligned} \rho(z;\mathbf{x}) &= \rho_p(z;\mathbf{x}_p) + \rho_r(z;\mathbf{x}_r) + \rho_g(z;\mathbf{x}_g) \\ &= \sum_{n_p=0}^{N_p} c_{n_p}(z - z_{wd})^{n_p} + \sum_{n_r=1}^{N_r} \rho_{n_r}^r [H(z - z_{n_r}^r) - H(z - z_{n_r}^r - t)] \\ &+ \sum_{n_g=1}^{N_g} \beta_{n_g} \frac{1}{\sqrt{2\pi w_{n_g}^2}} \exp\left\{ -\frac{1}{2w_{n_g}^2}(z - z_{n_g}^g)^2 \right\}, z > z_{wd} \end{aligned} \tag{3.5}$$

with the following constraints

$$0 \leq N_p \leq N_p^{\max} \tag{3.6a}$$
$$0 < c_0^{\min} \leq c_0 \leq c_0^{\max} \tag{3.6b}$$
$$c_p^{\min} \leq c_{n_p} \leq c_p^{\max}, \quad 1 \leq n_p \leq N_p \tag{3.6c}$$
$$0 \leq N_r \leq N_r^{\max} \tag{3.6d}$$
$$0 < \rho_r^{\min} \leq \rho_{n_r}^r \leq \rho_r^{\max}, \quad 1 \leq n_r \leq N_r \tag{3.6e}$$
$$z_1 \leq z_r^{\min} \leq z_{n_r}^r \leq z_r^{\max} < z_{N_l}, \quad 1 \leq n_r \leq N_r \tag{3.6f}$$
$$0 \leq N_g \leq N_g^{\max} \tag{3.6g}$$
$$\beta_g^{\min} \leq \beta_{n_g} \leq \beta_g^{\max}, \quad 1 \leq n_g \leq N_g \tag{3.6h}$$
$$z_g^{\min} \leq z_{n_g}^g \leq z_g^{\max}, \quad 1 \leq n_g \leq N_g \tag{3.6i}$$
$$0 < w_g^{\min} \leq w_{n_g} \leq w_g^{\max}, \quad 1 \leq n_g \leq N_g \tag{3.6j}$$
$$0 < \rho^{\min} \leq \rho(z;\mathbf{x}). \tag{3.6k}$$

The last restriction follows from the physical restriction of positive resistivity.

In total, this parameterization results in $N = 1 + N_p + 2N_r + 3N_g$ free parameters which values is to be determined by the optimization routine.

3.1.5 *Transdimensional parameterization*

In the above presentation, the parameters N_p, N_r and N_g are assumed fixed. It simplifies the implementation as most optimization routines, originally are designed for a fixed number of dimensions in the parameter space.

If initial information about the sub seabed topology is available, one can adjust the number of parameters accordingly. Say, one expect to see one hydrocarbon saturated layer and one layer with rocks more resistive than the background, in addition one expects there to be a compression trend. Then it would be natural to include one resistor term and one Gaussian term in addition to the first two polynomial terms. One would then need to determine the best value for 7 parameters.

However, in many situations one cannot make such assumptions. A more general approach is therefore to also optimize for the parameters N_p, N_r and N_g. With this, one could hope that the algorithm itself would adjust to the best suited number of parameters. In this thesis, both a fixed dimensional and trans-dimensional approach is explored.

4

FORWARD SOLVER

This chapter aims to derive an analytical expression of the electromagnetic field components from a horizontal electric dipole source in a transverse isotropic layered model, illustrated in Figure 4.1. Because of this simplified model and additional simplifying assumptions, it is possible to attain a fast quasi-analytic evaluation of the resulting field, which is implemented in the forward solver used to calculate the resulting field in this thesis. Note that, contrary to everything else treated in this thesis, this modeling tool which calculates these field expressions derived in this chapter is not implemented from scratch by the author. An already existing software developed at EMGS (*Electromagnetic Geoservices ASA*) is used, but the theory behind is included here for completeness.

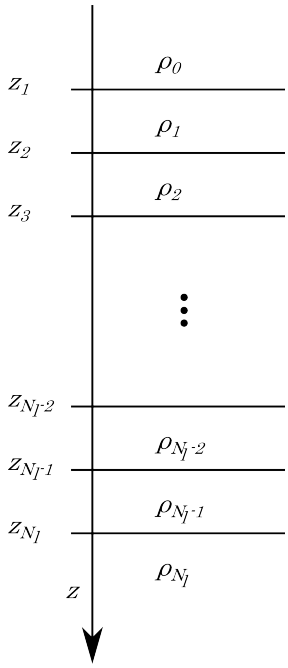


Figure 4.1: Simple illustration of one dimensional plane layer model.

This derivation is an excerpt of a more rigorous treatment of this field which can be found in (Løseth and Ursin, 2007). In addition to (Løseth and Ursin, 2007), multiple references is made to (Ursin, 1983) and (Stovas and Ursin, 2003), for more information about this topic, the reader are referred to these papers and the references therein.

4.1 MAXWELL'S EQUATIONS

Maxwell's equations in the time domain (Griffiths, 1999) are

$$\nabla \cdot \mathbf{D}(\mathbf{x},t) = \gamma \tag{4.1a}$$

$$\nabla \cdot \mathbf{B}(\mathbf{x},t) = 0 \tag{4.1b}$$

$$\nabla \times \mathbf{E}(\mathbf{x},t) + \frac{\partial}{\partial t}\mathbf{B}(\mathbf{x},t) = 0 \tag{4.1c}$$

$$\nabla \times \mathbf{H}(\mathbf{x},t) - \frac{\partial}{\partial t}\mathbf{D}(\mathbf{x},t) = \mathbf{J}(\mathbf{x},t). \tag{4.1d}$$

Here, $\mathbf{E}$ is the complex electric field vector and $\mathbf{H}$ is the complex auxiliary magnetic field vector. The density of free charges is given by γ (commonly symbolized as ρ, but in this thesis, ρ is reserved to resistivity), and

$\omega = 2\pi f$ is the angular frequency. $\mathbf{D}$ is the electric displacement vector, $\mathbf{B}$ the magnetic field vector and $\mathbf{J}$ the current density vector (Løseth and Ursin, 2007). The del operator is in the time domain the vector differential operator $\nabla = \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{pmatrix}^\top$, and when applied on a vector field $\mathbf{F}$, $\nabla \cdot \mathbf{F}$ is the *divergence* and $\nabla \times \mathbf{F}$ the *curl* of the field. In this chapter, $\mathbf{x}$ is used to denote the physical location such that $\mathbf{E}(\mathbf{x})$ is the electric field at $\mathbf{x}$. Anywhere else in this thesis, $\mathbf{x}$ is used to denote the vector of free parameters that is to be optimized. It should be clear out of context which is which, and should not cause any confusion.

By introducing the *Fourier transform* pair

$$\hat{\mathbf{F}}(\mathbf{x},\omega) = \mathscr{F}(\mathbf{F}(\mathbf{x},t))(\omega) = \int_{-\infty}^{\infty} \mathbf{F}(\mathbf{x},t) e^{-\imath\omega t}\, dt, \tag{4.2a}$$

$$\mathbf{F}(\mathbf{x},t) = \mathscr{F}^{-1}(\hat{\mathbf{F}}(\mathbf{x},\omega))(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\mathbf{F}}(\mathbf{x},\omega) e^{\imath\omega t}\, d\omega. \tag{4.2b}$$

for $\mathbf{F} \in \{\mathbf{E}, \mathbf{H}\}$ one can transform Maxwell's equations from the time domain to the frequency domain. Applying the Fourier transform on Equation 4.1 yields Maxwell's equations in the frequency domain

$$\nabla \cdot \mathbf{D}(\mathbf{x},\omega) = \gamma \tag{4.3a}$$

$$\nabla \cdot \mathbf{B}(\mathbf{x},\omega) = 0 \tag{4.3b}$$

$$\nabla \times \mathbf{E}(\mathbf{x},\omega) + \imath\omega \mathbf{B}(\mathbf{x},\omega) = 0 \tag{4.3c}$$

$$\nabla \times \mathbf{H}(\mathbf{x},\omega) - \imath\omega \mathbf{D}(\mathbf{x},\omega) = (\mathbf{J}_0 + \mathbf{J}_c)(\mathbf{x},\omega). \tag{4.3d}$$

This is so since the Fourier transform is linear, and

$$\begin{aligned} \frac{\partial}{\partial t}\mathbf{F}(\mathbf{x},t) &= \frac{\partial}{\partial t}\left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\mathbf{F}}(\mathbf{x},\omega) e^{\imath\omega t}\, d\omega \right) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \imath\omega \hat{\mathbf{F}}(\mathbf{x},\omega) e^{\imath\omega t}\, d\omega \\ &= \mathscr{F}^{-1}(\imath\omega \hat{\mathbf{F}}(\mathbf{x},\omega)) \end{aligned}$$

such that

$$\mathscr{F}\left(\frac{\partial}{\partial t}\mathbf{F}(\mathbf{x},t) \right) = \mathscr{F}\left(\mathscr{F}^{-1}(\imath\omega \hat{\mathbf{F}}(\mathbf{x},\omega)) \right) = \imath\omega \hat{\mathbf{F}}(\mathbf{x},\omega).$$

Note that, although the expressions in Equation 4.3 are in the frequency domain, the hat notation $\hat{\mathbf{F}}$ from the Fourier transform is omitted for convenience as it should not cause any misunderstanding. In Equation 4.3d, the current density $\mathbf{J}$ is split into an external source term $\mathbf{J}_s$ and a conduction current density term $\mathbf{J}_c$ (Griffiths, 1999).

In macroscopic media, the constitutive relations between the field quantities can be very complicated. The following simplifying assumptions is therefore applied. Nonlinear effects found in ferro-electric and ferromagnetic materials are not taken into account. Only (piecewise) homogeneous regions are considered, and possible non-local effects in space of the material parameters are ignored (Løseth and Ursin, 2007). The constitutive relations may then be written as

$$\mathbf{D}(\mathbf{x},\omega) = \boldsymbol{\varepsilon}(\omega)\mathbf{E}(\mathbf{x},\omega)$$
$$\mathbf{B}(\mathbf{x},\omega) = \boldsymbol{\mu}(\omega)\mathbf{H}(\mathbf{x},\omega)$$

and in the conductive media the relation between the conduction current density $\mathbf{J}_c$ and the electric field is given by *Ohm's law*

$$\mathbf{J}_c(\mathbf{x},\omega) = \boldsymbol{\sigma}(\omega)\mathbf{E}(\mathbf{x},\omega), \tag{4.5}$$

where $\boldsymbol{\varepsilon}$ is the electric permittivity and $\boldsymbol{\mu}$ is the magnetic permeability of the medium, $\boldsymbol{\sigma}$ is the electric conductivity. The parameters $\boldsymbol{\varepsilon}$, $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are tensors of order two and rank one.

As a source for the electromagnetic field, an electric dipole antenna is used. An infinitesimal electric dipole with general orientation can be represented by a periodic line current of length $\mathbf{l} = (l_x \quad l_y \quad l_z)^\intercal$ with current amplitude $I(\omega)$, yielding

$$\mathbf{J}_s(\mathbf{x},\omega) = I(\omega)(l_x \quad l_y \quad l_z)^\intercal \delta(\mathbf{r})$$

where $\mathbf{r} = \mathbf{x} - \mathbf{x}_s$ is the distance from the source, and δ is the *Dirac delta*.

In planar stratified media the electromagnetic properties vary in one direction only. Consider the model assumption illustrated in Figure 4.1 and let the medium properties vary in the z-direction. A transformation between the the time domain and frequency wave number domain is described by the Fourier transform pair

$$\hat{\mathbf{F}}(k_x,k_y,z,\omega) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{F}(\mathbf{x},t)e^{-\imath(k_x x + k_y y - \omega t)}\,dx\,dy\,dt, \tag{4.6a}$$

$$\mathbf{F}(\mathbf{x},t) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\mathbf{F}}(k_x,k_y,z,\omega)e^{\imath(k_x x + k_y y - \omega t)}\,dk_x\,dk_y\,d\omega. \tag{4.6b}$$

Analogous to the transformation to the frequency domain, $\frac{\partial}{\partial x} \rightarrow \imath k_x$, $\frac{\partial}{\partial y} \rightarrow \imath k_y$ and $\frac{\partial}{\partial t} \rightarrow -\imath\omega$ which yield Maxwell's equations in the frequency wave number domain

$$\nabla \cdot \mathbf{D}(k_x,k_y,z,\omega) = \gamma \tag{4.7a}$$
$$\nabla \cdot \mathbf{B}(k_x,k_y,z,\omega) = 0 \tag{4.7b}$$
$$\nabla \times \mathbf{E}(k_x,k_y,z,\omega) - \imath\omega\mathbf{B}(k_x,k_y,z,\omega) = 0 \tag{4.7c}$$
$$\nabla \times \mathbf{H}(k_x,k_y,z,\omega) + \imath\omega\mathbf{D}(k_x,k_y,z,\omega) = (\mathbf{J}_0 + \mathbf{J}_c)(k_x,k_y,z,\omega). \tag{4.7d}$$

Here, in the wave number domain, the del operator is $\nabla = \begin{pmatrix} \imath k_x & \imath k_y & \frac{\partial}{\partial z} \end{pmatrix}$, analogous to the time domain differential operator. Note that also here the hat notation is omitted, also the explicit notation showing dependency (k_x, k_y, z, ω) will be left out in the rest of this chapter. The quantities will be in the frequency wave number domain unless otherwise stated.

From Equation 4.7c with the magnetic source term ($\nabla \times \mathbf{E} - \imath\omega\mathbf{B} = -\mathbf{J}_M$), and Equation 4.7d one obtain the following set of equations

$$\imath k_y E_z - \frac{\partial E_y}{\partial z} - \imath\omega B_x = -J_x^M \tag{4.8a}$$

$$\frac{\partial E_x}{\partial z} - \imath k_x E_z - \imath\omega B_y = -J_y^M \tag{4.8b}$$

$$\imath k_x E_y - \imath k_y E_x - \imath\omega B_z = -J_z^M \tag{4.8c}$$

$$\imath k_y H_z - \frac{\partial H_y}{\partial z} + \imath\omega D_x - J_x^c = J_x^s \tag{4.8d}$$

$$\frac{\partial H_x}{\partial z} - \imath k_x H_z + \imath\omega D_y - J_y^c = J_y^s \tag{4.8e}$$

$$\imath k_x H_y - \imath k_y H_x + \imath\omega D_z - J_z^c = J_z^s. \tag{4.8f}$$

4.1.1 *Transverse isotropy in the vertical direction*

Transverse isotropy in the vertical direction (TIV) means that the electromagnetic properties of a medium are rotationally symmetric about the vertical z-axis. In layers where this is the case, the material constants simplify to

$$\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_h & 0 & 0 \\ 0 & \varepsilon_h & 0 \\ 0 & 0 & \varepsilon_v \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_h & 0 & 0 \\ 0 & \mu_h & 0 \\ 0 & 0 & \mu_v \end{pmatrix}, \quad \boldsymbol{\sigma} = \begin{pmatrix} \sigma_h & 0 & 0 \\ 0 & \sigma_h & 0 \\ 0 & 0 & \sigma_v \end{pmatrix}$$

where the subscript h and v stands for horizontal and vertical directions respectively. The conduction current density is given by Ohm's law (Equation 4.5) and, considering only a horizontal electric dipole oriented along the x-axis, the y and z components of the source term is zero while the x component is given as

$$J_x^s = I(\omega) l_x \delta(z - z_s) \tag{4.9}$$

and the source is located in $\mathbf{x}_s = (0 \quad 0 \quad z_s)^\intercal$. This infinitesimal dipole is a good approximation for a physical dipole with finite length when the length is much smaller than the wavelength ($l << \lambda$) and when one is far away from the dipole ($|\mathbf{r} >> \lambda$), where λ is used to denote the emitted wavelength. Introduce the complex valued constant $\tilde{\boldsymbol{\varepsilon}} = \boldsymbol{\varepsilon} + \imath\boldsymbol{\sigma}/\omega$ and let $p_x = k_x/\omega$ and

$p_y = k_y/\omega$. Then, by algebraically eliminating E_z and H_z, the above system in Equation 4.8 reduces to the coupled system of ordinary differential equations

$$\frac{\partial E_x}{\partial z} + \imath\omega\left(-\mu_h H_y + \frac{p_x^2}{\tilde{\varepsilon}_v} H_y - \frac{p_x p_y}{\tilde{\varepsilon}_v} H_x\right) = 0 \tag{4.10a}$$

$$\frac{\partial E_y}{\partial z} + \imath\omega\left(\frac{p_x p_y}{\tilde{\varepsilon}_v} H_y + \mu_h H_x - \frac{p_y^2}{\tilde{\varepsilon}_v} H_x\right) = 0 \tag{4.10b}$$

$$-\frac{\partial H_y}{\partial z} + \imath\omega\left(\tilde{\varepsilon}_h E_x - \frac{p_y^2}{\mu_v} E_x + \frac{p_y p_x}{\mu_v} E_y\right) = J_x \tag{4.10c}$$

$$\frac{\partial H_x}{\partial z} + \imath\omega\left(\frac{p_x p_y}{\mu_v} E_x + \tilde{\varepsilon}_h E_y - \frac{p_x^2}{\mu_v} E_y\right) = 0 \tag{4.10d}$$

or

$$\left(\frac{d}{dz} I + \imath\omega A\right)\mathbf{b} = \mathbf{s} \tag{4.11}$$

where I is the 4×4 identity matrix, and the system matrix is given as

$$A = \begin{pmatrix} 0 & A_1 \\ A_2 & 0 \end{pmatrix}$$

with

$$A_1 = \frac{1}{\tilde{\varepsilon}_v}\begin{pmatrix} \mu_h\tilde{\varepsilon}_v - p_x^2 & -p_x p_y \\ -p_x p_y & \mu_h\tilde{\varepsilon}_v - p_y^2 \end{pmatrix}$$

and

$$A_2 = \frac{1}{\mu_v}\begin{pmatrix} \mu_v\tilde{\varepsilon}_h - p_y^2 & p_x p_y \\ p_x p_y & \mu_v\tilde{\varepsilon}_h - p_x^2 \end{pmatrix}.$$

The field components is gathered in $\mathbf{b} = (E_x \quad E_y \quad -H_y \quad H_x)^\intercal$ and the source vector is given as $\mathbf{s} = (0 \quad 0 \quad J_x^0 \quad 0)^\intercal$. When assuming that the source and receiver layers enjoy this TIV property, the expressions above can be used to calculate the resulting horizontal field at the receiver. In addition, the vertical components follows directly from Equation 4.10, and are given as

$$E_z = \frac{1}{\tilde{\varepsilon}_v}\left(-p_x H_y + p_y H_x\right) \tag{4.12a}$$

$$H_z = \frac{1}{\mu_v}\left(p_x E_y - p_y E_x\right). \tag{4.12b}$$

4.2 DECOMPOSITION INTO UP-GOING AND DOWN-GOING FIELDS

In order to solve Equation 4.10 it is convenient to decompose the wave vector **b** into up- and down-going waves via a propagation vector **w**. In order to achieve this, just a simple transformation of Equation 4.10 is needed (Ursin, 1983).

The system matrix A is diagonalizable if it has four linearly independent eigenvectors. In this case, the spectral decomposition of A is

$$AQ = Q\Lambda,$$

with Λ being a diagonal matrix containing the eigenvalues of A along its diagonal, and Q is a matrix which columns is the corresponding eigenvectors. Introduce the mode field vector

$$\mathbf{w} = Q^{-1}\mathbf{b} = \begin{pmatrix} \mathbf{u} \\ \mathbf{d} \end{pmatrix} \tag{4.13}$$

which can be expressed as a vector consisting of vectors describing up-going and down-going waves (Ursin, 1983). Each of these vectors consists of a *s-polarized* and *p-polarized* component. The polarization of the wave is commonly referred to as the direction at which the electric field oscillates. When the wave is polarized such that the electric component is perpendicular to the *plane of incidence* it is called s-polarized. When the electric component is parallel with the plane of incidence, it is referred to as p-polarization. The plane of incidence is the plane spanned by the wave propagation vector and a vector normal to the boundary surface. The electromagnetic field can be decomposed into a *TE-mode* (transverse electric mode) and *TM-mode* (transverse magnetic mode) and this is what is done in Equation 4.13, hence the name mode field vector for **w**. The electromagnetic field component where the electric field is normal to the plane of incidence is termed as the TE-mode, and likewise the electromagnetic field component where the magnetic field is normal to the plane of incidence is termed the TM-mode. This is similar to s- and p-polarization, and a wave in TE-mode is the same as s-polarization while TM-mode is the same as p-polarization. Modes are often used to describe wave propagation in waveguides where electric and magnetic fields in the direction of propagation can occur. TE-mode is thus the mode where there are no *electric field* in the direction of propagation, and the TM-mode is when there are no *magnetic field* in the direction of propagation. Different disciplines will attach different meanings to these terms, and in this thesis the explanation above is used. Also, both the polarization terminology and mode terminology is used throughout the rest of this report.

With this, the differential equation in Equation 4.11 transforms to the system

$$\frac{d\mathbf{w}}{dz} = -\left(\iota\omega\Lambda + Q^{-1}\frac{dQ}{dz}\right)\mathbf{w} + Q^{-1}\mathbf{s}. \tag{4.14}$$

The term with the derivative of the eigenvector matrix describes possible medium variations. If one consider a homogeneous medium with no vertical change in the medium properties this term vanish. One can also discretize the medium variations into piecewise homogeneous regions which are connected by applying appropriate boundary conditions. In this case, the term also vanish, which yield the decoupled system

$$\frac{d\mathbf{w}}{dz} = -\imath\omega\Lambda\mathbf{w} + Q^{-1}\mathbf{s} \tag{4.15}$$

with solution of in a source free region

$$w_n(z) = w_n(z_0)\exp\{-\imath\omega\lambda_n(z-z_0)\}, \quad n = 1,2,3,4, \tag{4.16}$$

since the eigenvalue matrix is diagonal

$$\Lambda = \begin{pmatrix} \hat{\lambda} & 0 \\ 0 & -\check{\lambda} \end{pmatrix}, \quad \text{with } \hat{\lambda} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{and } \check{\lambda} = \begin{pmatrix} -\lambda_3 & 0 \\ 0 & -\lambda_4 \end{pmatrix}. \tag{4.17}$$

The eigenvalues can be organized such that $\lambda_1 = -\lambda_3$ and $\lambda_2 = -\lambda_4$, this reveals that the mode-field vector $\mathbf{w}$ consists of up-going ($\mathbf{u}$) and down-going ($\mathbf{d}$) field constituents. This propagation vector can be used to describe propagation of electromagnetic fields in homogeneous regions, across interfaces, and in a system of layers (Amundsen et al., 2006), (Løseth and Ursin, 2007). Represent propagation of $\mathbf{w}$ downwards in the positive z-direction as $\check{P}$ and the propagation upwards as $\hat{P}$ (Løseth and Ursin, 2007). In this way, for $z_b > z_a$

$$\mathbf{w}(z_b) = \check{P}(z_b, z_a)\mathbf{w}(z_a), \tag{4.18a}$$

$$\mathbf{w}(z_a) = \hat{P}(z_a, z_b)\mathbf{w}(z_b), \tag{4.18b}$$

which imply that

$$\hat{P}(z_a, z_b) = \check{P}^{-1}(z_b, z_a). \tag{4.19}$$

That is, the propagation matrix $\check{P}(z_b, z_a)$ relates the up-going and down-going field components propagating downwards at z_b to the up-going and down-going field components propagating downwards at z_a. Similarly, the propagation matrix $\hat{P}(z_a, z_b)$ relates the up-going and down-going field components propagating upwards at z_a to the up-going and down-going field components propagating upwards at z_b (Orfanidis, 2008). In the homogeneous source-free region which solution is given in Equation 4.16, the downward propagator matrix is thus

$$\check{P}(z_b, z_a) = \begin{pmatrix} e^{-\imath\omega\hat{\lambda}(z_b - z_a)} & 0 \\ 0 & e^{\imath\omega\check{\lambda}(z_b - z_a)} \end{pmatrix}. \tag{4.20}$$

For the upward propagator matrix, this is obtained using Equation 4.19.

When propagating across a source, $\mathbf{w}$ is discontinuous. Denote the transformed source in Equation 4.14 as

$$Q^{-1}\mathbf{s} = \Sigma = \begin{pmatrix} \hat{\Sigma} \\ \check{\Sigma} \end{pmatrix} \tag{4.21}$$

where $\hat{\Sigma}$ denotes the upward radiation and $\check{\Sigma}$ the downward radiation. Let z_s be the source position, and let z_s^+ and z_s^- be the position just below and above the source, respectively. Then, propagation across a source can be described as

$$\mathbf{w}(z_s^+) = \Sigma(z_s) + \mathbf{w}(z_s^-) \tag{4.22}$$

(Løseth and Ursin, 2007).

4.2.1 *Reflection and transmission response*

In order to derive expressions for the propagator matrices, properties of the electromagnetic field when propagating through boundaries will be utilized. When an incident electromagnetic wave (say $\mathbf{E}_I$ and $\mathbf{H}_I$) propagates towards and interacts with the boundary between two layers with their respective material constants, it gives rise to a reflected wave ($\mathbf{E}_R$ and $\mathbf{H}_R$) and a transmitted wave ($\mathbf{E}_T$ and $\mathbf{H}_T$). By applying appropriate boundary conditions, one can relate the reflected and transmitted field amplitudes to the amplitude of the incident field. These boundary conditions can be deduced from the integral representation of Maxwell's equations (Equation 4.1), and are given as

$$D_1^\perp - D_2^\perp = \gamma^f \tag{4.23a}$$
$$B_1^\perp - B_2^\perp = 0 \tag{4.23b}$$
$$\mathbf{E}_1^\parallel - \mathbf{E}_2^\parallel = 0 \tag{4.23c}$$
$$\mathbf{H}_1^\parallel - \mathbf{H}_2^\parallel = \mathbf{J}_c^f \times \mathbf{n}. \tag{4.23d}$$

These are the boundary conditions at a boundary between two different media, medium 1 and medium 2. The unit normal vector $\mathbf{n}$ is perpendicular to the surface between the two, and is pointing in the direction from medium 2 to medium 1. The surface carries a surface density γ^f of free charges and surface density $\mathbf{J}_c$ of free currents. The superscripts $()^\perp$ and $()^\parallel$ denotes the field components perpendicular to and parallel with the surface respectively (Griffiths, 1999). The reflection and transmission relations between the incident, reflected and transmitted field is denoted R and T respectively. When considering propagation through several layers with a more general anisotropy than TIV (remember that up until now, only the

source and receiver layers is assumed TIV), R and T will in this derivation be 2×2 matrices

$$R = \begin{pmatrix} R_{ss} & R_{sp} \\ R_{ps} & R_{pp} \end{pmatrix}, \quad T = \begin{pmatrix} T_{ss} & T_{sp} \\ T_{ps} & T_{pp} \end{pmatrix} \tag{4.24}$$

which consists of reflection and transmission coefficients relating the reflection and transmission amplitudes to the amplitude of a s-polarized and p-polarized incident wave, hence the subscripts. Capitalized R and T will be used to denote reflection and transmission responses from propagation through several layers whilst r and t will be used considering reflection and transmission through a single interface. At first, R and T will be used to derive expressions of the propagation matrices P. How to calculate R, T, r and t will be discussed in *Section* 4.2.7 and *Section* 4.2.8.

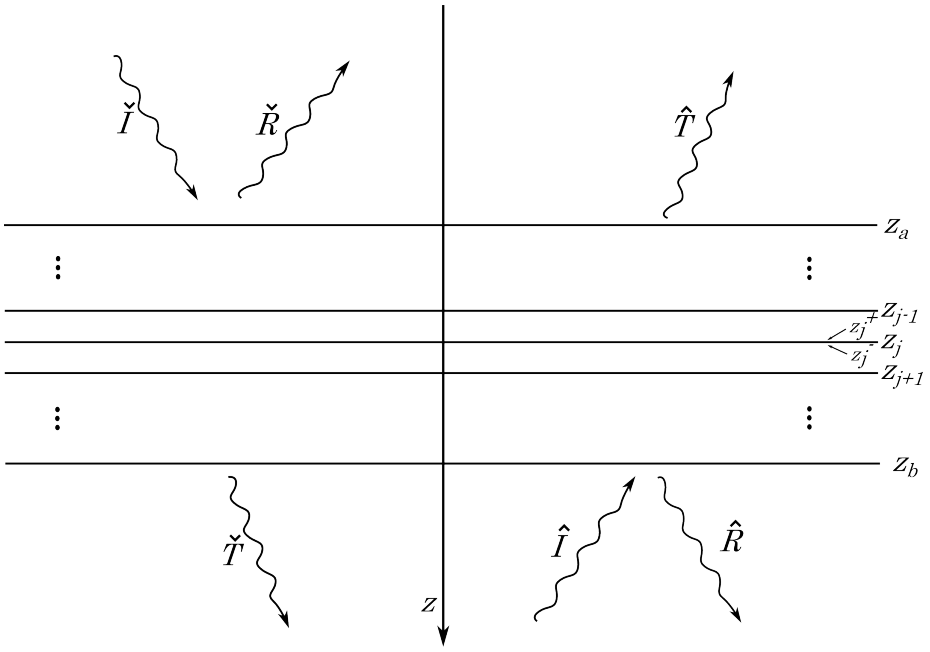


Figure 4.2: Reflection and transmission from an unit incident field.

Consider a unit incident field propagating downwards (and upwards) through an isotropic upper (and lower) half-space as illustrated in the left (and right) of Figure 4.2, respectively. Figure 4.2 then describes the reflection and transmission fields due to this incident field in the region between the interfaces z_a and z_b, $z_b > z_a$. This region between z_a and z_b may contain several layers or just one single interface, but no source (Ursin, 1983).

4.2.2 *Downwards propagation*

At first, downward propagation from z_1 to z_{N_l} is considered. The mode field at z_{N_l} relates to the mode field at z_1 via the propagator matrix as

$$\mathbf{w}(z_b) = \check{P}(z_b, z_a)\mathbf{w}(z_a). \tag{4.25}$$

Referring to the left of Figure 4.2, the layer below the interface at z_b there is only down-going field components, so that in terms of the incident field

$$\mathbf{w}(z_b) = \begin{pmatrix} \mathbf{0} \\ \mathbf{d}(z_b) \end{pmatrix} = \begin{pmatrix} 0 \\ \check{T} \end{pmatrix} \begin{pmatrix} \check{I}^s \\ \check{I}^p \end{pmatrix} \tag{4.26}$$

and at the layer above z_a

$$\mathbf{w}(z_a) = \begin{pmatrix} \mathbf{u}(z_a) \\ \mathbf{d}(z_a) \end{pmatrix} = \begin{pmatrix} \check{R} \\ I \end{pmatrix} \begin{pmatrix} \check{I}^s \\ \check{I}^p \end{pmatrix}. \tag{4.27}$$

where $\check{I}^s$ and $\check{I}^p$ are s- and p-polarized unit incident wave amplitudes corresponding to the up/down decomposition. Now, inserting Equation 4.26 and Equation 4.27 into Equation 4.25 yields

$$\begin{pmatrix} 0 \\ \check{T} \end{pmatrix} = \begin{pmatrix} \check{P}_{11} & \check{P}_{12} \\ \check{P}_{21} & \check{P}_{22} \end{pmatrix} \begin{pmatrix} \check{R} \\ I \end{pmatrix} \tag{4.28}$$

where $\check{P}_{11}$ and $\check{P}_{22}$ are the upper left and lower right 2×2 sub-matrices in $\check{P}$ respectively. $\check{P}_{12}$ and $\check{P}_{21}$ are accordingly the upper right and lower left 2×2 sub-matrices.

With the same reasoning in the right case of Figure 4.2

$$\mathbf{w}(z_b) = \begin{pmatrix} \mathbf{u}(z_b) \\ \mathbf{d}(z_b) \end{pmatrix} = \begin{pmatrix} I \\ \hat{R} \end{pmatrix} \begin{pmatrix} \hat{I}^s \\ \hat{I}^p \end{pmatrix}, \tag{4.29}$$

and

$$\mathbf{w}(z_a) = \begin{pmatrix} \mathbf{u}(z_a) \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \hat{T} \\ 0 \end{pmatrix} \begin{pmatrix} \hat{I}^s \\ \hat{I}^t \end{pmatrix}. \tag{4.30}$$

Inserting Equation 4.29 and Equation 4.30 into Equation 4.25 yields

$$\begin{pmatrix} I \\ \hat{R} \end{pmatrix} = \begin{pmatrix} \check{P}_{11} & \check{P}_{12} \\ \check{P}_{21} & \check{P}_{22} \end{pmatrix} \begin{pmatrix} \hat{T} \\ 0 \end{pmatrix}. \tag{4.31}$$

Solving Equation 4.28 and Equation 4.31 for the reflection and transmission matrices yields

$$\hat{T} = \check{P}_{11}^{-1} \tag{4.32a}$$

$$\hat{R} = \check{P}_{21}\check{P}_{11}^{-1} \tag{4.32b}$$

$$\check{R} = -\check{P}_{11}^{-1}\check{P}_{12} \tag{4.32c}$$

$$\check{T} = \check{P}_{22} - \check{P}_{21}\check{P}_{11}^{-1}\check{P}_{12}. \tag{4.32d}$$

4.2.3 *Upwards propagation*

In a similar fashion, expressions for the reflection and transmission matrices can be found when considering propagation upwards in the negative z-direction. The mode field above z_1 relates to the mode field below z_{N_l} via the propagator matrix as

$$\mathbf{w}(z_a) = \hat{P}(z_a, z_b)\mathbf{w}(z_a). \tag{4.33}$$

The relations between the incident, reflected and transmitted field is the same as in the upward propagation case, so referring to the left of Figure 4.2, inserting Equation 4.26 and Equation 4.27 into Equation 4.33, yields

$$\begin{pmatrix} \check{R} \\ I \end{pmatrix} = \begin{pmatrix} \hat{P}_{11} & \hat{P}_{12} \\ \hat{P}_{21} & \hat{P}_{22} \end{pmatrix} \begin{pmatrix} 0 \\ \check{T} \end{pmatrix} \tag{4.34}$$

and in the right case, inserting Equation 4.29 and Equation 4.30 into Equation 4.33

$$\begin{pmatrix} \hat{T} \\ 0 \end{pmatrix} = \begin{pmatrix} \hat{P}_{11} & \hat{P}_{12} \\ \hat{P}_{21} & \hat{P}_{22} \end{pmatrix} \begin{pmatrix} I \\ \hat{R} \end{pmatrix}. \tag{4.35}$$

Combining Equation 4.34 and Equation 4.35 yields the following expressions

$$\check{T} = \hat{P}_{22}^{-1} \tag{4.36a}$$
$$\check{R} = \hat{P}_{12}\hat{P}_{22}^{-1} \tag{4.36b}$$
$$\hat{R} = -\hat{P}_{22}^{-1}\hat{P}_{21} \tag{4.36c}$$
$$\hat{T} = \hat{P}_{11} - \hat{P}_{12}\hat{P}_{22}^{-1}\hat{P}_{21}. \tag{4.36d}$$

4.2.4 *Propagation matrices expressions*

From Equation 4.32 and Equation 4.36 the following expressions for the downward and upward propagator matrices follows

$$\check{P}(z_b, z_a) = \begin{pmatrix} \hat{T}^{-1} & -\hat{T}^{-1}\check{R} \\ \hat{R}\hat{T}^{-1} & \check{T} - \hat{R}\hat{T}^{-1}\check{R} \end{pmatrix}, \tag{4.37a}$$

$$\hat{P}(z_a, z_b) = \begin{pmatrix} \hat{T} - \check{R}\check{T}^{-1}\hat{R} & \check{R}\check{T}^{-1} \\ -\check{T}^{-1}\hat{R} & \check{T}^{-1} \end{pmatrix}. \tag{4.37b}$$

or, in terms of the mode field vector

$$\mathbf{w}(z_b) = \begin{pmatrix} \hat{T}^{-1} & -\hat{T}^{-1}\check{R} \\ \hat{R}\hat{T}^{-1} & \check{T} - \hat{R}\hat{T}^{-1}\check{R} \end{pmatrix} \mathbf{w}(z_a), \tag{4.38a}$$

$$\mathbf{w}(z_a) = \begin{pmatrix} \hat{T} - \check{R}\check{T}^{-1}\hat{R} & \check{R}\check{T}^{-1} \\ -\check{T}^{-1}\hat{R} & \check{T}^{-1} \end{pmatrix} \mathbf{w}(z_b). \tag{4.38b}$$

4.2.5 *System of layers containing a source*

In this section we will derive expressions for the mode-field vector **w** in the layers above and underneath the source. Use Figure 4.3 and the subsections above for reference on the different expressions.

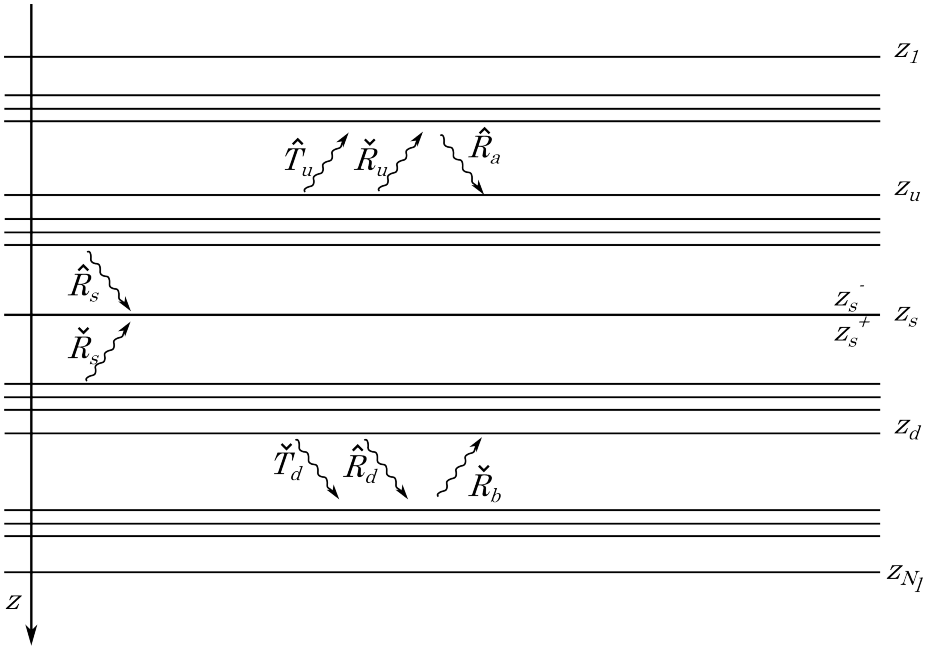


Figure 4.3: Reflection and transmission in a multilayered system. $\hat{R}_s$ (and $\check{R}_s$) is the reflection in the source layer from all interfaces above (and below) the source due to incident up-going (and down-going) radiation from the source. $\hat{T}_u$ (and $\check{T}_d$) is the transmission at layer u (and d) from all interfaces between layer u (and d) and the source due to incident up-going (and down-going) radiation from the source. $hatR_a$ (and $\check{R}_b$) is the reflection from all interfaces above (and below) layer u (and d) due to up-going (and down-going) radiation from the source. $\check{R}_u$ (and $\hat{R}_d$) is the reflection from all the interfaces below (and above) layer u (and d) due down-going (and up-going) radiation at the source.

First, the mode field above the source $\mathbf{w}(z), z < z_s$ will be derived. Consider downward propagation from z anywhere above the source down to the source at z_s^-

$$\mathbf{w}(z_s^-) = \check{P}(z_s^-, z)\mathbf{w}(z). \tag{4.39}$$

By comparing Figure 4.3 with Figure 4.2 and inserting the corresponding expressions into Equation 4.37a one get the following expression for the up-going fields just above the source

$$\mathbf{u}(z_s^-) = \hat{T}_u^{-1}\mathbf{u}(z) - \hat{T}_u^{-1}\check{R}_u\mathbf{d}(z). \tag{4.40}$$

Now, since there is only one source, and this is below z, the down-going fields at z will only be the reflected up-going fields at z, so $\mathbf{d}(z) = \hat{R}_a\mathbf{u}(z)$, with this

$$\mathbf{u}(z_s^-) = \hat{T}_u^{-1}(I - \check{R}_u\hat{R}_a)\mathbf{u}(z). \tag{4.41}$$

Inverting this yields the following expression for the mode-field

$$\mathbf{w}(z) = \begin{pmatrix} I \\ \hat{R}_a \end{pmatrix}\mathbf{u}(z) = \begin{pmatrix} I \\ \hat{R}_a \end{pmatrix}(I - \check{R}_u\hat{R}_a)^{-1}\hat{T}_u\mathbf{u}(z_s^-). \tag{4.42}$$

In order to derive an expression for $\mathbf{u}(z_s^-)$, consider downward propagation from the source to the bottom most layer

$$\mathbf{w}(z_{N_l}) = \check{P}(z_{N_l}, z_s^+)\mathbf{w}(z_s^+) = \check{P}(z_{N_l}, z_s^+)[\mathbf{w}(z_s^-) + \Sigma(z_s)] \tag{4.43}$$

where the last relation follows from Equation 4.22.

In the lower half space, the region below z_{N_l}, there are only down-going fields. This is so since it is assumed to be only one source, which lay above z_{N_l}, and since there are no interfaces below z_{N_l} to reflect on. Furthermore, the down-going field just above the source is just the reflected up-going field $\mathbf{d}(z_s^-) = \hat{R}_s\mathbf{u}(z_s^-)$, again because there is only one source layer, all down-going fields above the source must be reflected up-going fields. With this, Equation 4.43 can be written as

$$\begin{pmatrix} \mathbf{0} \\ \mathbf{d}(z_{N_l}) \end{pmatrix} = \check{P}_{11}\begin{pmatrix} I & -\check{R}_s \\ \check{P}_{11}^{-1}\check{P}_{21} & \check{P}_{11}^{-1}\check{P}_{22} \end{pmatrix}\left[\begin{pmatrix} \mathbf{u}(z_s^-) \\ \hat{R}_s\mathbf{u}(z_s^-) \end{pmatrix} + \begin{pmatrix} \hat{\Sigma}(z_s) \\ \check{\Sigma}(z_s) \end{pmatrix}\right], \tag{4.44}$$

where the relation $\check{R}_s = -\check{P}_{11}^{-1}\check{P}_{12}$ from Equation 4.32c is used. From this,

$$\mathbf{u}(z_s^-) = (I - \check{R}_s\hat{R}_s)^{-1}(-\hat{\Sigma} + \check{R}_s\check{\Sigma}),$$

and inserting this into Equation 4.42, yields the expression of the total mode-field above the source

$$\mathbf{w}(z) = \begin{pmatrix} I \\ \hat{R}_a \end{pmatrix}(I - \check{R}_u\hat{R}_a)^{-1}\hat{T}_u(I - \check{R}_s\hat{R}_s)^{-1}(-\hat{\Sigma} + \check{R}_s\check{\Sigma}), \quad z < z_s. \tag{4.45}$$

Now, the mode-field below the source $\mathbf{w}(z), z > z_s$ will be derived with the same procedure as with the mode-field above the source. Consider upward propagation from z anywhere below the source up to the source at z_s^+

$$\mathbf{w}(z_s^+) = \hat{P}(z_s^+, z)\mathbf{w}(z). \tag{4.46}$$

Again, by comparing Figure 4.3 with Figure 4.2 and inserting the corresponding expressions into Equation 4.37b yields the following expression for the down-going fields just above the source

$$\mathbf{d}(z_s^+) = -\check{T}_d^{-1}\hat{R}_d\mathbf{u}(z) - \check{T}_d^{-1}\mathbf{d}(z). \tag{4.47}$$

Since there is only one source, and this is above z, the up-going fields at z will only be the reflected down-going fields at z, so $\mathbf{u}(z) = \check{R}_b\mathbf{d}(z)$, with this

$$\mathbf{d}(z_s^+) = \check{T}_d^{-1}(I - \hat{R}_d\check{R}_b)\mathbf{d}(z). \tag{4.48}$$

Inverting this yields the following expression for the mode-field

$$\mathbf{w}(z) = \begin{pmatrix} \check{R}_b \\ I \end{pmatrix}\mathbf{d}(z) = \begin{pmatrix} \check{R}_b \\ I \end{pmatrix}(I - \hat{R}_d\check{R}_b)^{-1}\check{T}_d\mathbf{d}(z_s^+). \tag{4.49}$$

In order to find an expression for $\mathbf{d}(z_s^+)$, consider the upwards propagation from the source to the topmost interface z_1

$$\mathbf{w}(z_1) = \hat{P}(z_1, z_s^-)\mathbf{w}(z_s^-) = \hat{P}(z_1, z_s^-)[\mathbf{w}(z_s^+) - \Sigma(z_s)].$$

In the upper half-space, the region above z_1, there are only up-going fields, and the up-going field at the source equals the reflected response from the down-going field $\mathbf{u}(z_s^+) = \check{R}_s\mathbf{d}(z_s^+)$. Corresponding to Equation 4.44, this leads to

$$\begin{pmatrix} \mathbf{u}(z_1) \\ \mathbf{0} \end{pmatrix} = \hat{P}_{22}\begin{pmatrix} \hat{P}_{22}^{-1}\hat{P}_{11} & \hat{P}_{22}^{-1}\hat{P}_{12} \\ -\hat{R}_s & I \end{pmatrix}\left[\begin{pmatrix} \check{R}_s\mathbf{d}(z_s^+) \\ \mathbf{d}(z_s^+) \end{pmatrix} - \begin{pmatrix} \hat{\Sigma}(z_s) \\ \check{\Sigma}(z_s) \end{pmatrix}\right] \tag{4.50}$$

where $\hat{R}_s = -\hat{P}_{22}^{-1}\hat{P}_{21}$ is from Equation 4.36c. The down-going field vector just below the source now becomes

$$\mathbf{d}(z_s^+) = (I - \hat{R}_s\check{R}_s)^{-1}(\check{\Sigma} + \hat{R}_s\hat{\Sigma}),$$

and the mode-field vector anywhere below the source $z > z_s$, is found by inserting this into Equation 4.49

$$\mathbf{w}(z) = \begin{pmatrix} \check{R}_b \\ I \end{pmatrix}(I - \hat{R}_d\check{R}_b)^{-1}\check{T}_d(I - \hat{R}_s\check{R}_s)^{-1}(\check{\Sigma} - \hat{R}_s\hat{\Sigma}), \quad z > z_s. \tag{4.51}$$

4.2.6 *Eigenvalues and eigenvectors*

Now, consider again the system in Equation 4.11, restated here for convenience

$$\left(\frac{d}{dz}I + \imath\omega A\right)\mathbf{b} = \mathbf{s}$$

which for TIV media has the system matrix

$$A = \begin{pmatrix} 0 & A_1 \\ A_2 & 0 \end{pmatrix}$$

with

$$A_1 = \frac{1}{\tilde{\varepsilon}_v}\begin{pmatrix} \mu_h\tilde{\varepsilon}_v - p_x^2 & -p_x p_y \\ -p_x p_y & \mu_h\tilde{\varepsilon}_v - p_y^2 \end{pmatrix}$$

and

$$A_2 = \frac{1}{\mu_v}\begin{pmatrix} \mu_v\tilde{\varepsilon}_h - p_y^2 & p_x p_y \\ p_x p_y & \mu_v\tilde{\varepsilon}_h - p_x^2. \end{pmatrix}$$

The field components is given in $\mathbf{b} = (E_x \quad E_y \quad -H_y \quad H_x)^\intercal$ and the source vector for a horizontal electric dipole is given as $\mathbf{s} = (0 \quad 0 \quad J_x^0 \quad 0)^\intercal$. The eigenvalue matrix Λ and the eigenvector matrix Q in the spectral decomposition of A

$$AQ = Q\Lambda$$

is stated. $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ where

$$\lambda_1 = -\lambda_3 = \sqrt{\mu_h\tilde{\varepsilon}_h - \frac{\mu_h}{\mu_v}(p_x^2 + p_y^2)}, \tag{4.52a}$$

$$\lambda_2 = -\lambda_4 = \sqrt{\mu_h\tilde{\varepsilon}_h - \frac{\tilde{\varepsilon}_h}{\tilde{\varepsilon}_v}(p_x^2 + p_y^2)}. \tag{4.52b}$$

Because of symmetry in the up-going and down-going field constituents

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} Q_E & Q_E \\ Q_H & -Q_H \end{pmatrix} \tag{4.53}$$

and

$$Q^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} Q_H^\intercal & Q_E^\intercal \\ Q_H^\intercal & -Q_E^\intercal. \end{pmatrix}. \tag{4.54}$$

If the medium enjoy the TIV property, the eigenvector sub-matrices are given by

$$Q_E = \frac{1}{p_r}\begin{pmatrix} p_y\sqrt{\frac{\mu_h}{\lambda_1}} & p_x\sqrt{\frac{\lambda_2}{\tilde{\varepsilon}_h}} \\ -p_x\sqrt{\frac{\mu_h}{\lambda_1}} & p_y\sqrt{\frac{\lambda_2}{\tilde{\varepsilon}_h}} \end{pmatrix}, \quad \text{and} \quad Q_H = \frac{1}{p_r}\begin{pmatrix} p_y\sqrt{\frac{\lambda_1}{\mu_h}} & p_x\sqrt{\frac{\tilde{\varepsilon}_h}{\lambda_2}} \\ -p_x\sqrt{\frac{\lambda_1}{\mu_h}} & p_y\sqrt{\frac{\tilde{\varepsilon}_h}{\lambda_2}} \end{pmatrix}. \tag{4.55}$$

where $p_r = \sqrt{p_x^2 + p_y^2}$.

4.2.7 *Reflection and transmission at single interfaces*

Before calculating the expressions of R and T, expressions of r and t will be derived. In the same way as R and T represents reflection and transmission through multiple layers, r and t represents reflection and transmission through one single interface. From the boundary conditions of Maxwell's equations (Equation 4.23) one can see from Equation 4.23c and Equation 4.23 that the horizontal components given by $\mathbf{b} = Q\mathbf{w}$ are continuous across an interface. When there are no free currents at the surface, this means that $Q(z_j^-)\mathbf{w}(z_j^-) = Q(z_j^+)\mathbf{w}(z_j^+)$, where z_j^- and z_j^+ are the z-coordinates on each side of the interface at z_j. Thus the propagator matrices across the interface becomes

$$\check{P}(z_j^+, z_j^-) = Q^{-1}(z_j^+)Q(z_j^-), \text{ and } \hat{P}(z_j^-, z_j^+) = Q^{-1}(z_j^-)Q(z_j^+). \tag{4.56}$$

The expressions in for the reflection and transmission matrices is given in Equation 4.32 and Equation 4.36. Together with the expressions of the propagator matrices in Equation 4.56 and the expressions of the eigenvalue matrices given in Equation 4.53 and Equation 4.54, the expressions of reflection and transmission over one interface follows

$$\hat{t} = \check{P}_{11}^{-1} = 2\left[(\hat{Q}_E^+)^\intercal \hat{Q}_H^- + (\hat{Q}_H^+)^\intercal \hat{Q}_E^-\right]^{-1} \tag{4.57a}$$

$$\hat{r} = -\hat{P}_{22}^{-1}\hat{P}_{21} = -\left[(\check{Q}_H^-)^\intercal \check{Q}_E^+ + (\check{Q}_E^-)^\intercal \check{Q}_H^+\right]^{-1}\left[(\check{Q}_H^-)^\intercal \hat{Q}_E^+ - (\check{Q}_E^-)^\intercal \hat{Q}_H^+\right] \tag{4.57b}$$

$$\check{r} = -\check{P}_{11}^{-1}\check{P}_{12} = -\left[(\hat{Q}_E^+)^\intercal \hat{Q}_H^- + (\hat{Q}_H^+)^\intercal \hat{Q}_E^-\right]^{-1}\left[(\hat{Q}_E^+)^\intercal \check{Q}_E^- - (\hat{Q}_H^+)^\intercal \check{Q}_E^-\right] \tag{4.57c}$$

$$\check{t} = \hat{P}_{22}^{-1} = 2\left[(\check{Q}_H^-)^\intercal \check{Q}_E^+ + (\check{Q}_E^-)^\intercal \check{Q}_H^+\right]^{-1}. \tag{4.57d}$$

Here, the superscripts $()^-$ and $()^+$ refers to the eigenvector matrices in the upper and lower homogeneous layers, respectively, the other notations follow the same nomenclature defined earlier.

In the case of TIV, one can get the following explicit representations from Equation 4.55

$$\check{t} = \begin{pmatrix} 2\frac{\sqrt{\mu_h^-\mu_h^+\lambda_1^-\lambda_1^+}}{\mu_h^+\lambda_1^- + \mu_h^-\lambda_1^+} & 0 \\ 0 & 2\frac{\sqrt{\tilde{\varepsilon}_h^-\tilde{\varepsilon}_h^+\lambda_2^-\lambda_2^+}}{\tilde{\varepsilon}_h^-\lambda_2^+ + \tilde{\varepsilon}_h^+\lambda_1^-} \end{pmatrix}, \quad \check{r} = \begin{pmatrix} \frac{\mu_h^+\lambda_1^- - \mu_h^-\lambda_1^+}{\mu_h^+\lambda_1^- + \mu_h^-\lambda_1^+} & 0 \\ 0 & \frac{\tilde{\varepsilon}_h^-\lambda_2^+ - \tilde{\varepsilon}_h^+\lambda_2^-}{\tilde{\varepsilon}_h^-\lambda_2^+ + \tilde{\varepsilon}_h^+\lambda_1^-} \end{pmatrix}, \tag{4.58}$$

$\hat{t} = \check{t}$ and $\hat{r} = -\check{r}$.

4.2.8 *Recursive calculation of R and T*

The reflection and transmission response from a stack of layers can be calculated by multiplying together the propagator matrices of the respective

layers. In simulations, this procedure may be numerically unstable, so a recursive calculation is proposed. This procedure can be used to calculate reflection and transmission responses in a stack of piece-wise homogeneous layers between the interfaces z_a and z_b ($z_a < z_b$), which do not contain any sources, see Figure 4.2. Appropriate initial values will be provided at the respective boundary interfaces z_a and z_b.

First, the downward reflection and transmission matrices $\check{R}$ and $\check{T}$ will be derived. Consider propagation from z_b upwards through a stack of layers to z_j^+, from there, through the interface at z_j up through one layer to z_{j-1}^+. In terms of the propagator matrix, this becomes

$$\hat{P}(z_{j-1}^+, z_b) = \hat{P}(z_{j-1}^+, z_j^-)\hat{P}(z_j^-, z_j^+)\hat{P}(z_j^+, z_b) \tag{4.59}$$

For $\hat{P}(z_j^+, z_b)$, the reflection and transmission matrices in the right column of Equation 4.37b are denoted $\check{R}_{j+1}$ and $\check{T}_{j+1}$, respectively. Correspondingly, the reflection and transmission matrices in $\hat{P}(z_{j-1}^+, z_b)$ is denoted $\check{R}_j$ and $\check{T}_j$, respectively. The propagator across the single interface at z_j is also given by Equation 4.37b, but here the reflection and transmission matrices will be denoted r_j and t_j since it is only through one interface. The propagation through the homogeneous layer is given by the inverse of the propagator matrix in Equation 4.20. After some calculations, the following recursion formulas are obtained

$$\check{R}_j = e^{\imath\omega\hat{\lambda}h_j}\left[\hat{t}_j\check{R}_{j+1}\left(I - \hat{r}_j\check{R}_{j+1}\right)^{-1}\right)\check{t}_j + \check{r}_j\Big]e^{\imath\omega\check{\lambda}h_j} \tag{4.60a}$$

$$\check{T}_j = \check{T}_{j+1}\left(I - \hat{r}_j\check{R}_{j+1}\right)^{-1}\check{t}_j e^{\imath\omega\check{\lambda}h_j}, \tag{4.60b}$$

where $h_j = z_j - z_{j-1}$, $j \in \{a, a+1, \ldots, b\}$, and $\hat{\lambda}$ and $\check{\lambda}$ are the eigenvalue sub-matrices in layer j. The initial values are $\check{R}_b = 0$ and $\check{T}_b = e^{\imath\omega\check{\lambda}h_b}$.

In a similar manner, down-going propagators can be used to derive recursive formulas for $\hat{R}_j$ and $\hat{T}_j$,

$$\hat{R}_j = e^{\imath\omega\check{\lambda}h_{j+1}}\left[\hat{t}_j\hat{R}_{j-1}\left(I - \check{r}_j\hat{R}_{j-1}\right)^{-1}\right)\hat{t}_j + \hat{r}_j\Big]e^{\imath\omega\hat{\lambda}h_{j+1}} \tag{4.61a}$$

$$\hat{T}_j = \hat{T}_{j-1}\left(I - \check{r}_j\hat{R}_{j-1}\right)^{-1}\hat{t}_j e^{\imath\omega\hat{\lambda}h_{j+1}}, \tag{4.61b}$$

for $j \in \{a, a+1, \ldots, b\}$. Here, $\hat{R}_{j-1}$ and $\hat{T}_{j-1}$ describe the reflection and transmission at z_j^- whilst $\hat{R}_j$ and $\hat{T}_j$ describe the reflection and transmission at z_{j+1}^-. The initial values are given as $\hat{R}_a = 0$ and $\hat{T}_a = e^{\imath\omega\hat{\lambda}h_a}$ and the eigenvalue sub-matrices are those in layer $j+1$.

As is further shown in (Løseth and Ursin, 2007), these recursive expressions simplify in TIV media, $\check{R} = \hat{R}^\intercal$ and $\check{T} = \hat{T}^\intercal$. Furthermore, the off-diagonal elements, referring to coupling of the TE- and TM-modes, are zero, which imply that $\check{R} = \hat{R}$ and $\check{T} = \hat{T}$. The following recursive relation is valid

for both the TE-mode and the TM-mode and will therefore only be denoted by the scalar R and T,

$$R_j = \frac{r_j + R_{j+1}}{1 + r_j R_{j+1}} e^{2\imath\omega\lambda h_j} \tag{4.62a}$$

$$T_j = \frac{t_j T_{j+1}}{1 + r_j R_{j+1}} e^{\imath\omega\lambda h_j}. \tag{4.62b}$$

The initial conditions is $R_b = 0$ and $T_b = e^{\imath\omega\lambda h_b}$, and the validity is as in the general case. The expressions are equal for the two polarizations with exception of the eigenvalues, which for the TE-mode is given by the scalar $\lambda = \lambda_1$ and for the TM-mode by the scalar $\lambda = \lambda_2$.

4.3 FINAL FIELD COMPONENT EXPRESSIONS

We are now in the position to derive the final expressions for the field components from a horizontal electrical dipole in TIV medium. For a horizontal electric dipole

$$\check{\Sigma} = -\frac{1}{\sqrt{2}}\hat{\Sigma} = -\frac{1}{\sqrt{2}} Q_E^{\intercal} \mathbf{s}_E^h$$

where the source vector is split as $\mathbf{s} = \mathbf{s}_E + \mathbf{s}_M$ and $\mathbf{s}_E = \begin{pmatrix} \mathbf{s}_E^v \\ \mathbf{s}_E^h \end{pmatrix}$. This follows from Equation 4.21 and that for a horizontal electric dipole, $\mathbf{s} = (0 \quad 0 \quad J_x \quad 0)^{\intercal}$, that is $\mathbf{s}_M = \mathbf{0}$ and $\mathbf{s}_E^v = \mathbf{0}$. Furthermore J_x is given in Equation 4.9 which yield

$$\check{\Sigma} = -\hat{\Sigma} = -\frac{I l_x}{\sqrt{2} p_r} \begin{pmatrix} p_y \mathscr{E}_s \\ p_x \mathscr{M}_s^{-1} \end{pmatrix},$$

$$\mathscr{E} = \sqrt{\frac{\mu_h}{\lambda_1}}, \quad \text{and} \quad \mathscr{M} = \sqrt{\frac{\tilde{\varepsilon}_h}{\lambda_2}}$$

and the subscript s means that the value of the material constants are those of the source layer. Using Equation 4.45 we get for $z < z_s$

$$\mathbf{b} = \begin{pmatrix} \mathbf{b}_E \\ \mathbf{b}_H \end{pmatrix} = Q\mathbf{w}(z < z_s) \tag{4.63a}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} Q_E (I + \hat{R}_a)(I - \check{R}_u \hat{R}_a)^{-1} \hat{T}_u (I - \check{R}_s \hat{R}_s)^{-1} (I + \check{R}_s)\check{\Sigma} \\ Q_H (I - \hat{R}_a)(I - \check{R}_u \hat{R}_a)^{-1} \hat{T}_u (I - \check{R}_s \hat{R}_s)^{-1} (I + \check{R}_s)\check{\Sigma} \end{pmatrix} \tag{4.63b}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} Q_E \hat{\mathscr{R}}^A \\ Q_H \hat{\mathscr{R}}^D \end{pmatrix} \check{\Sigma} \tag{4.63c}$$

and for $z > z_s$, using Equation 4.51

$$\mathbf{b} = \begin{pmatrix} \mathbf{b}_E \\ \mathbf{b}_H \end{pmatrix} = Q\mathbf{w}(z > z_s) \tag{4.64a}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} Q_E(\check{R}_b + I)(I - \hat{R}_d\check{R}_b)^{-1}\check{T}_d(I - \hat{R}_s\check{R}_s)^{-1}(I + \hat{R}_s)\check{\Sigma} \\ Q_H(\check{R}_b - I)(I - \hat{R}_d\check{R}_b)^{-1}\check{T}_d(I - \hat{R}_s\check{R}_s)^{-1}(I + \hat{R}_s)\check{\Sigma} \end{pmatrix} \tag{4.64b}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} Q_E\check{\mathscr{R}}^A \\ Q_H\check{\mathscr{R}}^D \end{pmatrix} \check{\Sigma}. \tag{4.64c}$$

Written explicitly, the electric and magnetic field components in the frequency wave number domain, when the source and receiver layers are transverse vertical isotropic and the source is a horizontal electric dipole, becomes

$$E_x = -\frac{Il_x}{2p_r^2}\left[\mathscr{E}\mathscr{E}_s p_y^2 \mathscr{R}^A_{11} + \frac{\mathscr{E}}{\mathscr{M}_s} p_y p_x \mathscr{R}^A_{12} + \frac{\mathscr{E}_s}{\mathscr{M}} p_x p_y \mathscr{R}^A_{21} + \frac{1}{\mathscr{M}\mathscr{M}_s} p_x^2 \mathscr{R}^A_{22}\right] \tag{4.65a}$$

$$E_y = -\frac{Il_x}{2p_r^2}\left[-\mathscr{E}\mathscr{E}_s p_x p_y \mathscr{R}^A_{11} - \frac{\mathscr{E}}{\mathscr{M}_s} p_x^2 \mathscr{R}^A_{12} + \frac{\mathscr{E}_s}{\mathscr{M}} p_y^2 \mathscr{R}^A_{21} + \frac{1}{\mathscr{M}\mathscr{M}_s} p_y p_x \mathscr{R}^A_{22}\right] \tag{4.65b}$$

$$H_x = +\frac{Il_x}{2p_r^2}\left[\frac{\mathscr{E}_s}{\mathscr{E}} p_y^2 \mathscr{R}^D_{11} + \frac{1}{\mathscr{E}\mathscr{M}_s} p_y p_x \mathscr{R}^D_{12} + \mathscr{M}\mathscr{E}_s p_x p_y \mathscr{R}^D_{21} + \frac{\mathscr{M}}{\mathscr{M}_s} p_x^2 \mathscr{R}^D_{22}\right] \tag{4.65c}$$

$$H_y = -\frac{Il_x}{2p_r^2}\left[-\frac{\mathscr{E}_s}{\mathscr{E}} p_x p_y \mathscr{R}^D_{11} - \frac{1}{\mathscr{E}\mathscr{M}_s} p_x^2 \mathscr{R}^D_{12} + \mathscr{M}\mathscr{E}_s p_y^2 \mathscr{R}^D_{21} + \frac{\mathscr{M}}{\mathscr{M}_s} p_y p_x \mathscr{R}^D_{22}\right] \tag{4.65d}$$

and the horizontal components follows from Equation 4.12b

$$E_z = -\frac{Il_x}{2\tilde{\varepsilon}_v}\left[\mathscr{M}\mathscr{E}_s p_y \mathscr{R}^D_{21} + \frac{\mathscr{M}}{\mathscr{M}_s} p_x \mathscr{R}^D_{22}\right] \tag{4.66a}$$

$$H_z = \frac{Il_x}{2\mu_v}\left[\mathscr{E}\mathscr{E}_s p_y \mathscr{R}^A_{11} + \frac{\mathscr{E}}{\mathscr{M}_s} p_x \mathscr{R}^A_{12}\right]. \tag{4.66b}$$

Here, $\mathscr{R} \to \check{\mathscr{R}}$ for $z < z_s$ and $\mathscr{R} \to \hat{\mathscr{R}}$ for $z > z_s$.

4.3.1 *TIV or isotropy in all layers*

In the case where all layers are TIV or isotropic, the reflection and transmission response simplifies. Consider the matrix expressions in Equation 4.63 and Equation 4.64

$$\mathscr{R} = \begin{pmatrix} \mathscr{R}_{11} & \mathscr{R}_{12} \\ \mathscr{R}_{21} & \mathscr{R}_{22} \end{pmatrix}, \tag{4.67}$$

where $\mathscr{R}$ refers to all $\mathscr{R}$ matrices in Equation 4.63 and Equation 4.64. Let Equation 4.67 represent the reflection and transmission response in a TIV or isotropic layer from an anisotropic stack of layers due to a source in a TIV or isotropic medium. Then the subscript 1 is related to the TE-mode and the subscript 2 is related to the TM-mode such that $\mathscr{R}_{11}$ represent the emitted TE-polarization and resulting TE-mode, $\mathscr{R}_{12}$ represent emitted TE-mode and resulting TM-mode. Likewise $\mathscr{R}_{21}$ represents emitted TM-mode and resulting TE-mode while $\mathscr{R}_{22}$ represents both emitted and resulting field is in TM-mode.

As mentioned above, at interfaces with more complicated anisotropy than TIV, the off-diagonal in the reflection and transmission matrices, R and T respectively, are nonzero. If, however the off-diagonal elements are zero, $\mathscr{R}$ is diagonal. Moreover, if all layers in the model are TIV or isotropic, the responses for the TE-mode and TM-mode are given by the same scalar expression. Written out this means that $\mathscr{R}_{12} = \mathscr{R}_{21} = 0$, $\mathscr{R}^A_{11} = \mathscr{R}^A_{22} = \mathscr{R}_A$ and $\mathscr{R}^D_{11} = \mathscr{R}^D_{22} = \mathscr{R}_D$ in our particular case (Løseth and Ursin, 2007). Although given by the same expression, let $\mathscr{R}_{11}$ be denoted by $\mathscr{R}^E$ and let $\mathscr{R}_{22}$ be denoted by $\mathscr{R}^M$.

For $z < z_s$ this is given explicitly as

$$\hat{\mathscr{R}}_A = \hat{T}_u \frac{1+\hat{R}_a}{1-\check{R}_u\hat{R}_a}\frac{1+\check{R}_s}{1-\check{R}_s\hat{R}_s}$$

$$\hat{\mathscr{R}}_D = \hat{T}_u \frac{1-\hat{R}_a}{1-\check{R}_u\hat{R}_a}\frac{1+\check{R}_s}{1-\check{R}_s\hat{R}_s}$$

and for $z > z_s$

$$\check{\mathscr{R}}_A = \check{T}_d \frac{\check{R}_b+1}{1-\hat{R}_d\check{R}_b}\frac{1+\hat{R}_s}{1-\hat{R}_s\check{R}_s}$$

$$\check{\mathscr{R}}_D = \check{T}_d \frac{\check{R}_b-1}{1-\hat{R}_d\check{R}_b}\frac{1+\hat{R}_s}{1-\hat{R}_s\check{R}_s}$$

This reduces the expression of the horizontal field components to

$$E_x = -\frac{Il_x}{2p_r^2}\left[\mathscr{E}\mathscr{E}_s p_y^2\mathscr{R}_A + \frac{1}{\mathscr{M}\mathscr{M}_s}p_x^2\mathscr{R}_A\right] \tag{4.70a}$$

$$E_y = -\frac{Il_x}{2p_r^2}\left[-\mathscr{E}\mathscr{E}_s p_x p_y\mathscr{R}_A + \frac{1}{\mathscr{M}\mathscr{M}_s}p_y p_x\mathscr{R}_A\right] \tag{4.70b}$$

$$H_x = +\frac{Il_x}{2p_r^2}\left[\frac{\mathscr{E}_s}{\mathscr{E}}p_y^2\mathscr{R}_D + \frac{\mathscr{M}}{\mathscr{M}_s}p_x^2\mathscr{R}_D\right] \tag{4.70c}$$

$$H_y = -\frac{Il_x}{2p_r^2}\left[-\frac{\mathscr{E}_s}{\mathscr{E}}p_x p_y\mathscr{R}_D + \frac{\mathscr{M}}{\mathscr{M}_s}p_y p_x\mathscr{R}_D\right] \tag{4.70d}$$

and for the vertical components

$$E_z = -\frac{Il_x}{2\tilde{\varepsilon}_v}\frac{\mathscr{M}}{\mathscr{M}_s}p_x\mathscr{R}_D \tag{4.71a}$$

$$H_z = \frac{Il_x}{2\mu_v}\mathscr{E}\mathscr{E}_s p_y\mathscr{R}_A. \tag{4.71b}$$

In order to get the quantities the in Equation 4.70 back to the spatial domain, the inverse Fourier transform given in Equation 4.6b is applied. In case of TIV or isotropy in all layers, the double Fourier integral over the horizontal wave numbers k_x and k_y can be simplified to a single integral in terms of *Bessel* functions and a polar horizontal wave number k_r. Introduce polar coordinates as

$$\begin{aligned} r &= \sqrt{x^2+y^2}, & x &= r\cos\beta, & y &= r\sin\beta, \\ k_r &= \sqrt{k_x^2+k_y^2}, & k_x &= k_r\cos\alpha, & k_y &= k_r\sin\alpha. \end{aligned}$$

where α is the polar angle in the wave number domain and β is the angle in the spatial domain and r is the polar radius. With this, the double integral from Equation 4.6b without the time and frequency part becomes

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathbf{F}(k_x,k_y)e^{\imath(k_x x+k_y y)}dk_x dk_y = \int_0^{\infty}\int_0^{2\pi}k_r\mathbf{F}(k_r,\phi)e^{\imath(k_r r\sin\phi)}d\phi dk_r,$$

where $\phi=\alpha-\beta+\pi/2$. And thus the Fourier transform from the wave number domain to the spatial domain is given as

$$F(r,\beta)=\frac{1}{(2\pi)^2}\int_0^{\infty}\int_0^{\infty}k_r\hat{F}(k_r,\phi)e^{\imath k_r r\sin\phi}d\phi dk_r. \tag{4.73}$$

Note that no we are still in the frequency domain, no transform to time domain is made. In what follows, the expression of the final polar spatial field components from a horizontal electric dipole in TIV or isotropic media $E_r(r,\beta)$, is derived. The derivation of the other field components $E_\beta(r,\beta)$, $E_z(r,\beta)$, $H_r(r,\beta)$, $H_\beta(r,\beta)$, $H_z(r,\beta)$ is similar.

Let

$$\begin{aligned} g_A^E &= \mathscr{E}\mathscr{E}_s\mathscr{R}_A^E, & g_A^M &= \frac{1}{\mathscr{M}}\frac{1}{\mathscr{M}_s}\mathscr{R}_A^M \\ g_D^E &= -\frac{\mathscr{E}_s}{\mathscr{E}}\mathscr{R}_D^E, & g_D^E &= -\frac{\mathscr{M}}{\mathscr{M}_s}\mathscr{R}_D^E \end{aligned}$$

such that the field components from Equation 4.70 becomes

$$\begin{aligned} E_x(k_r,\alpha) &= -\frac{Il_x}{2p_r^2}[g_A^E p_y^2+g_A^M p_x^2] \\ &= -\frac{Il_x}{2}[g_A^E\sin^2\alpha+g_A^M\cos^2\alpha] \\ E_y(k_r,\alpha) &= -\frac{Il_x}{2p_r^2}[-g_A^E p_x p_y+g_A^M p_x p_y] \\ &= -\frac{Il_x}{2}[-g_A^E\sin\alpha\cos\alpha+g_A^M\sin\alpha\cos\alpha] \end{aligned}$$

since $p_x = k_x/\omega$, $p_y = k_y/\omega$ and $p_r = k_r/\omega$. Furthermore

$$\begin{aligned}
\sin^2\alpha &= \sin^2(\phi+\beta-\pi/2) \\
&= \frac{1}{2}[\cos 2\phi\cos 2\beta - \sin 2\phi\sin 2\beta + 1] \\
\cos^2\alpha &= \cos^2(\phi+\beta-\pi/2) \\
&= \frac{1}{2}[-\cos 2\phi\cos 2\beta + \sin 2\phi\sin 2\beta + 1] \\
\sin\alpha\cos\alpha &= \sin(\phi+\beta-\pi/2)\cos(\phi+\beta-\pi/2) \\
&= -\frac{1}{2}[\cos 2\phi\sin 2\beta + \sin 2\phi\cos 2\beta]
\end{aligned}$$

Now, introduce the *Bessel functions*, with the integral representation

$$J_n(z) = \frac{(-1)^n}{2\pi}\int_0^{2\pi} e^{\imath n\phi}e^{\imath z\sin\phi}\,d\phi. \tag{4.77}$$

Since

$$\begin{aligned}
\int_0^{2\pi}\cos\phi e^{\imath z\sin\phi}\,d\phi = 0 \quad \forall z, \\
\int_0^{2\pi}\sin 2\phi e^{\imath z\sin\phi}\,d\phi = 0 \quad \forall z,
\end{aligned}$$

the first three functions can be described as

$$\begin{aligned}
J_0(z) &= \frac{1}{2\pi}\int_0^{2\pi} e^{\imath z\sin\phi}\,d\phi \\
J_1(z) &= \frac{-1}{2\pi}\int_0^{2\pi}[\cos\phi + \imath\sin\phi]e^{\imath z\sin\phi}\,d\phi \\
&= \frac{1}{2\pi\imath}\int_0^{2\pi}\sin\phi e^{\imath z\sin\phi}\,d\phi \\
J_2(z) &= \frac{-1}{2\pi}\int_0^{2\pi}[\cos 2\phi + \imath\sin 2\phi]e^{\imath z\sin\phi}\,d\phi \\
&= \frac{1}{2\pi}\int_0^{2\pi}\cos 2\phi e^{\imath z\sin\phi}\,d\phi.
\end{aligned}$$

Since the terms $\cos\phi$ and $\sin 2\phi$ do not contribute to the Bessel expansion, this means that in practice

$$\begin{aligned}
\sin^2\alpha &\to \frac{1}{2}[\cos 2\phi\cos 2\beta + 1] \\
\cos^2\alpha &\to \frac{1}{2}[-\cos 2\phi\cos 2\beta + 1] \\
\cos\alpha\sin\alpha &\to -\frac{1}{2}\cos 2\phi\sin 2\beta
\end{aligned}$$

and

$$E_x(k_r,\alpha) = -\frac{Il_x}{4}[g_A^E \cos 2\phi \cos 2\beta + g_A^E - g_A^M \cos 2\phi \cos 2\beta + g_A^M]$$
$$E_y(k_r,\alpha) = -\frac{Il_x}{4}[g_A^E \cos 2\phi \sin 2\beta - g_A^E \cos 2\phi \sin 2\beta].$$

Transforming these using Equation 4.73 gives the following representation in the spatial domain

$$E_x(r,\beta) = -\frac{Il_x}{8\pi}\int_0^\infty k_r[g_A^E J_2(rk_r)\cos 2\beta + g_A^E J_0(rk_r) - g_A^M J_2(rk_r)\cos 2\beta + g_A^M J_0(rk_r)]\,dk_r$$
$$E_y(r,\beta) = -\frac{Il_x}{8\pi}\int_0^\infty k_r[g_A^E J_2(rk_r)\sin 2\beta - g_A^M J_2(rk_r)\sin 2\beta]\,dk_r$$

The cylindrical components are obtained through the rotation

$$\begin{pmatrix} F_r \\ F_\beta \\ F_z \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta & 0 \\ -\sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix}.$$

This, together with the property of the Bessel functions $J_0(z) + J_2(z) = \frac{2}{z}J_1(z)$, yields the final expression for the field components

$$E_r = -\frac{Il_x}{4\pi}\cos\beta\left[\mathcal{I}_{A0}^M + \frac{1}{r}(\mathcal{I}_{A1}^E - \mathcal{I}_{A1}^M)\right] \tag{4.83a}$$
$$E_\beta = -\frac{Il_x}{4\pi}\sin\beta\left[-\mathcal{I}_{A0}^E + \frac{1}{r}(\mathcal{I}_{A1}^E - \mathcal{I}_{A1}^M)\right] \tag{4.83b}$$
$$H_r = \frac{Il_x}{4\pi}\sin\beta\left[-\mathcal{I}_{D0}^E + \frac{1}{r}(\mathcal{I}_{D1}^E - \mathcal{I}_{D1}^M)\right] \tag{4.83c}$$
$$H_\beta = -\frac{Il_x}{4\pi}\cos\beta\left[\mathcal{I}_{D0}^M + \frac{1}{r}(\mathcal{I}_{D1}^E - \mathcal{I}_{D1}^M)\right] \tag{4.83d}$$
$$E_z = \frac{Il_x}{4\pi}\frac{\imath\cos\beta}{\omega\tilde{\varepsilon}_v}\int_0^\infty k_r^2 g_D^M J_1(rk_r)\,dk_r \tag{4.83e}$$
$$H_z = \frac{Il_x}{4\pi}\frac{\imath\sin\beta}{\omega\mu_v}\int_0^\infty k_r^2 g_D^E J_1(rk_r)\,dk_r \tag{4.83f}$$

where the *Hankel transforms* are given as

$$\mathcal{I}_{A0}^E = \int_0^\infty k_r g_A^E J_0(rk_r)dk_r, \quad \mathcal{I}_{A1}^E = \int_0^\infty g_A^E J_1(rk_r)\,dk_r \tag{4.84a}$$
$$\mathcal{I}_{A0}^M = \int_0^\infty k_r g_A^M J_0(rk_r)dk_r, \quad \mathcal{I}_{A1}^M = \int_0^\infty g_A^M J_1(rk_r)\,dk_r \tag{4.84b}$$
$$\mathcal{I}_{D0}^E = \int_0^\infty k_r g_D^E J_0(rk_r)dk_r, \quad \mathcal{I}_{D1}^E = \int_0^\infty g_D^E J_1(rk_r)\,dk_r \tag{4.84c}$$
$$\mathcal{I}_{D0}^M = \int_0^\infty k_r g_D^M J_0(rk_r)dk_r, \quad \mathcal{I}_{D1}^M = \int_0^\infty g_D^M J_1(rk_r)\,dk_r. \tag{4.84d}$$

can now be solved with a numerical integration method (Løseth and Ursin, 2007).

4.4 FAST HANKEL TRANSFORM

In order to solve Equation 4.83, the integrals in Equation 4.84 needs to be evaluated numerically. This section describes a method called the *Fast Hankel Transform* used to evaluate Hankel transforms, which are integrals of the type

$$F(r) = \int_0^\infty f(\lambda) J_\nu(\lambda r)\, d\lambda \tag{4.85}$$

where J_ν is an νth order Bessel function of the first kind (Johansen and Sørensen, 1979). Due to the oscillatory behaviour of J_ν (and possibly f), standard quadrature methods applied to these integrals tends to have converge, or they may not converge at all if the integral is divergent. The Fast Hankel Transform implemented with digital filters which was proposed in (Ghosh, 1971) tends to handle some of the above issues in a favourable way, making it a frequently used tool in EM geophysics discipline. In problems concerned in this thesis, where the model is quite small, a numerical evaluation of Equation 4.83 is performed once for each forward call in the inversion, and one inversion typically consist of 10000 to 50000 forward calls. Other, larger modeling applications may require to evaluate integrals of the type in Equation 4.85 millions of times(Key, 2012). Evidently, this sets high requirements on the precision and speed of the evaluation method.

Now, let

$$\lambda = e^{-u} \text{ and } r = e^{v}, \quad u, v \in (-\infty, \infty)$$

and define the new functions K and G on $(-\infty, \infty)$ as

$$G(v) := e^{v} F(e^{v}) \text{ and } K(u) := f(e^{-u}).$$

By this substitution, Equation 4.85 can be reformulated as

$$G(v) = \int_{-\infty}^{\infty} K(u) e^{v-u} J_\nu(e^{v-u})\, du.$$

This is recognised as the convolution

$$G(v) = \int_{-\infty}^{\infty} K(u) H_\nu(v-u)\, du = \int_{-\infty}^{\infty} K(v-u) H_\nu(u)\, du = K * H$$

where

$$H_\nu(v) := e^{v} J_\nu e^{v})$$

is the *kernel* of the transformation. The idea is now to approximate G numerically with some function $\widetilde{G}$ that is computable such that $|G(v) - \widetilde{G}(v)| < \epsilon, \quad \forall v \in S \subset \mathbb{R}$, for some small, positive ϵ larger than the computational numerical precision. Before $\widetilde{G}$ is defined, discretize u into equidistant points $n\Delta$ for integer values n and some spacing δ, and define the interpolation

$$\widetilde{K}(u) = \sum_{n=-\infty}^{\infty} K(n\delta) \psi\left(\frac{u}{\delta} - n\right)$$

where ψ is some interpolating function (e.g. the normalized sinc function) such that $\psi(0) = 1$ and $\psi(n) = 0$ for all integer values $n \neq 0$. Then introduce $\widetilde{G}$ as

$$\widetilde{G}(v) := \int_{-\infty}^{\infty} \widetilde{K}(u) H_v(v-u)\,du \tag{4.86a}$$

$$= \sum_{n=-\infty}^{\infty} K(n\Delta)\widetilde{H}_v(v-n\Delta) \tag{4.86b}$$

where

$$\widetilde{H}_v(v) := \int_{-\infty}^{\infty} \psi\left(\frac{u}{\Delta}\right) H_v(v-u)\,du \tag{4.87}$$

is the ψ-response of the H_v-transform. Now, $\widetilde{H}_v$ is still difficult to evaluate, and in order to evaluate $\widetilde{G}(v)$ for arbitrary values of v would require that $\widetilde{H}(v)$ is known at the corresponding set of values (Johansen and Sørensen, 1979). Instead, discretize v into a set of equidistant points $m\Delta$ for integer m, such that, from Equation 4.86, $\widetilde{G}(v)$ can be expressed in these points as

$$\widetilde{G}(m\Delta) = \sum_{-\infty}^{\infty} K(n\Delta)\widetilde{H}[(m-n)\Delta]. \tag{4.88}$$

Now, one could interpolate between them once more with an approximation $\widetilde{\widetilde{G}}$ of $\widetilde{G}$ as

$$\widetilde{\widetilde{G}}(v) = \sum_{m=-\infty}^{\infty} \widetilde{G}(m\Delta)\phi\left(\frac{v}{\Delta} - m\right),$$

where ϕ may or may not be the same interpolation function as ψ. The problem of computing values of $\widetilde{H}$ remains, and is in general a difficult task (Johansen and Sørensen, 1979).

It is recognised in (Ghosh, 1971) that $\widetilde{H}$ evaluated at a certain distinct points is essentially a vector of linear filter coefficients that could be predetermined and applied together with arbitrary functions K (or f). Optimal filter coefficients for finite length filters can be found by solving Equation 4.88 for a given length filter using known integral transform pairs for G and f where the known f share similar characteristics as the functions the resulting method is applied on. With this, the resulting T-point digital filter approximation is

$$G(v) \approx \sum_{k=-l}^{l} K(v-k\Delta)\widetilde{H}_k$$

or

$$rF(r) \approx \sum_{k=-l}^{l} f(b_i/r)\widetilde{H}_k$$

with the logarithmic spaced filter abscissas $b_k = \lambda_k r = e^{k\Delta}$ for $k = -l, -l + 1, \ldots, l$ and $T = 2l + 1$ (Key, 2012). When applied in modeling tools, the different values of H_k is usually already calculated, so only f needs to be evaluated, which still can be costly. But since it only needs to be evaluated at T points, and this filter length usually lies in a range of tens or hundreds, the digital filter method is usually much faster than simple, direct quadrature methods (Key, 2012).

5

OPTIMIZATION OF THE MISFIT FUNCTION

At each iteration of the inversion, a new candidate model is proposed by changing the parameter values in the parameterization discussed in Chapter 3. The problem of finding the best configuration of parameter values is stated as an optimization problem

$$\rho(x,y,z;\mathbf{x}) = \arg\min_{\mathbf{x}\in D} \varepsilon(\rho(\mathbf{x}))$$

where ρ represents the model, e.i. the resistivity at locations (x,y,z) in the physical space. $\mathbf{x}$ represents the parameterization of the model space, and is a collection of variables which is to be optimized, these are contained in some space $D \subset \mathbb{R}^N$.

In this project the stochastic optimization routine named *simulated annealing* is chosen. The reason for this choice is mainly the easy implementation and robust convergence properties, that is, the ability to locate global minima with minimal knowledge of the solution space topology, in addition to the possibility of escaping local minima. Also, stochastic inversion methods are the main application of the parameterization introduced in this thesis, since the computational cost explodes in higher dimensions with many parameters.

It is assumed that the number of parameters N, is fixed. The trans-dimensional case requires some extra attention and is discussed at the end of this chapter.

5.1 MISFIT FUNCTIONAL

The misfit functional ε, which is to be minimized is the complex squared, weighted L^2-norm of the difference between the the observed and synthetic (modeled) electromagnetic data

$$\varepsilon(\rho) = \sum_{s,r,f,F} W^2(\mathbf{x}_s,\mathbf{x}_r,f)\Delta F^*(\mathbf{x}_s,\mathbf{x}_r,f|\rho)\Delta F(\mathbf{x}_s,\mathbf{x}_r,f|\rho) \tag{5.1}$$

where $(\cdot)^*$ denotes the complex conjugate and the difference field is

$$\Delta F(\mathbf{x}_s,\mathbf{x}_r,f|\rho) = F^{Obs}(\mathbf{x}_s,\mathbf{x}_r,f) - F^{Synth}(\mathbf{x}_s,\mathbf{x}_r,f|\rho). \tag{5.2}$$

This is a sum over every source position $\mathbf{x}_s$, receiver position $\mathbf{x}_r$, transmitter frequency f, and field component F, which, in general can be the electric and magnetic field in both x and y directions, relative to a coordinate system where the z is the vertical axis. As is common, these field components

are denoted by E_x,E_y,H_x and H_y, respectively. These are thus the data sample variables, and they represent the amount of information about the problem. The total number of data parameters is denoted M. The notation is chosen to emphasize that the synthetic data is dependent on the current proposed resistivity model ρ, whilst the observed (and therefore, the weights, see Equation 5.3) are not.

As weights, the following is used

$$W(x_s, x_r, f) = \frac{1}{\delta E^{Obs}(x_s, x_r, f)}, \tag{5.3}$$

in order to include measurement inaccuracy. This expression follows from the assumption of Gaussian noise, that we are minimizing the L^2 norm, and that we seek the *maximum likelihood estimator*. The predicted data uncertainty can be estimated using error propagation analysis (Morten et al., 2009), and is given by

$$(\delta E^{Obs}(\mathbf{x}_s, \mathbf{x}_r, f))^2 = \sum_k |\frac{\partial E^{Obs}(\mathbf{x}_s, \mathbf{x}_r, f)}{\partial p_k} \delta p_k|^2. \tag{5.4a}$$

$$\approx \alpha^2 |E_x^{Obs}(\mathbf{x}_s, \mathbf{x}_r, f)|^2 + \eta^2, \tag{5.4b}$$

under the assumption of independent and random contributions. p_k denotes an experimental parameter with the associated measurement uncertainty δp_k (Mittet and Morten, 2013). The approximation in Equation 5.4b is analysed and justified in (Morten and Mittet, 2012), and is the expression used in this project. α is a constant, relative error in the field amplitude measurement, mainly caused by instrument calibration inaccuracy, whereas η represent the background noise (Morten et al., 2009). α is of order around 10^{-2} and η around $1 \times 10^{-16}\,\mathrm{V\,m^{-1}}$. These weights reflects that terms with field observations dominated by noise should contribute less to the misfit sum.

5.2 SIMULATED ANNEALING

The simulated annealing method is a stochastic optimization method proposed by *Kirkpatric*, *Gelett* and *Vecci* (Kirkpatrick et al., 1983). It is inspired by a physical annealing process where a solid is heated above its critical temperature, before it is slowly cooled, with the intention that the solid will end up in a lower energy state.

Although, first applied to discrete combinatorial problems (see e.g. (Kirkpatrick et al., 1983), (Hajec, 1988)) it is also applied to optimization on continuous variable spaces (see e.g. (Vanderbilt and Louiem, 1984), see also (Locatelli, 2000) and the references therein for a great summary on the field). There are several different approaches, but the main idea of the method can be summarized as

1. $\varepsilon(\mathbf{x})$ is the objective function that is to be minimized.
2. Start with a initial point $\mathbf{x}_0$ in the feasible set D.
3. Sample a new point $\mathbf{y}_{k+1}$ from some proposal distribution.
4. Set the new point $\mathbf{x}_{k+1} = \mathbf{y}_{k+1}$ with probability $\alpha(\varepsilon(\mathbf{x}_k), \varepsilon(\mathbf{y}_{k+1}), T_k)$, otherwise, set $\mathbf{x}_{k+1} = \mathbf{x}_k$.
5. Update the *temperature* T_{k+1} according to some cooling schedule.
6. Check termination criteria, and continue from step 3 if it is not met.

The rest of this section will focus on simulated annealing on a constrained continuous global optimization problem. Let the parameter space be denoted D and assume for generality that it is compact in $\mathbb{R}^N$. Consider a probability space $(\Omega, \mathscr{F}, P)$ where $\omega \in \Omega$ is elements in the sample space, $\mathscr{F}$ is a σ-algebra containing subsets of Ω and P is some probability measure on $\mathscr{F}$. Let g be a strictly positive continuous probability density on D and let π be the absolutely continuous probability measure called the *target distribution* with density g

$$\pi(A) = \int_A g(\mathbf{x})\, d\mathbf{x}, \quad \forall A \in \mathscr{B} \tag{5.5}$$

where $\mathscr{B}$ is the *Borel σ-algebra* on D (Romeijn and Smith, 1994).

Define the function $\alpha : D \times D \to (\epsilon, 1]$ for some $\epsilon > 0$ to be jointly measurable on $D \times D$ and that it satisfies $\alpha(\mathbf{x}, \mathbf{y}) g(\mathbf{x}) = \alpha(\mathbf{y}, \mathbf{x}) g(\mathbf{y})$, for all $\mathbf{x}, \mathbf{y} \in D$, call this function the *acceptance probability function*.

The idea now is to sample the proposed next solution from some proposal distribution. This can be sampled from the whole of D independent of the current state $\mathbf{x}_k$ or, maybe more common, sampled from some neighbourhood of $\mathbf{x}_k$.

Continuing this iterative process generates a sequence of random variables $\{\mathbf{X}_k\}_{k \geq 0}$ which constitutes a Markov chain on D with stationary transition probabilities. Moreover, it can be shown that the target distribution π is a stationary distribution for this Markov chain and that for every starting point $\mathbf{x}_0 \in D$, the Markov chain converges in distribution to π (Romeijn and Smith, 1994).

The standard density used in the literature is

$$g_T(\mathbf{x}) = \frac{1}{Z_T} \exp\left\{-\frac{\varepsilon(\mathbf{x})}{T}\right\}, \tag{5.6}$$

where $Z_T = \int_D \exp\left\{-\frac{\varepsilon(\mathbf{z})}{T}\right\} d\mathbf{z}$ is a normalizing constant. The acceptance probability is set to the Metropolis criterion

$$\alpha_T(\mathbf{x}, \mathbf{y}) = \min\left\{1, \frac{g_T(\mathbf{y})}{g_T(\mathbf{x})}\right\} = \min\left\{1, \exp\left\{-\frac{\varepsilon(\mathbf{y}) - \varepsilon(\mathbf{x})}{T}\right\}\right\}.$$

A common procedure and a key point in simulated annealing is to introduce a non-increasing sequence of temperatures $\{T_0, T_1, \ldots\}$, in order to concentrate the target distribution around small function values, as can be seen from Equation 5.6. More formally, the temperatures are determined through a sequence of functions $\{t_0, t_1, \ldots\}$ called *cooling schedule* where $T_k = t_k(\cdot)$. There are several different schedules in the literature (see e.g. (Kirkpatrick et al., 1983), (Vanderbilt and Louiem, 1984), (Aarts and van Laarhoven, 1985), (Lundy and Mees, 1986), (Hajec, 1988), (Cohn and Fielding, 1999)). A common class of schedules are the ones that are deterministic, that is, they are fully determined before the algorithm starts. A second class are the ones of *adaptive* schedules, which are dependent on the search trajectory of the specific run of the algorithm. Formally one say that a process $h(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}^N$ is $\mathcal{N}_t$ adapted if, for an increasing family of σ-algebras $\{\mathcal{N}_t\}_{t\geq 0}$, and for each $t \geq 0$, the mapping $\omega \mapsto h(t,\omega)$ is $\mathcal{N}_t$-measurable (Øksendal, 2010). In this case the process is the cooling schedule and the σ-algebras are the one generated by the sequence $\{\mathbf{X}_k\}_{k\geq 0}$, meaning that at iteration k, the schedule is dependent on the history of the iterates up to k, or $T_k = t_k(\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_k)$.

5.2.1 *Convergence*

We will now show that for a decreasing temperature, the simulated annealing will eventually converge to the minimum function value ε^*.

Define the δ-level set

$$D_\delta = \{x \in D; \varepsilon(\mathbf{x}) \leq \varepsilon^* + \delta\}$$

where ε^* denotes the optimal function value. Then, for all $\delta > 0$

$$\lim_{T\to 0} \pi_T(D_\delta) = 1. \tag{5.7}$$

To show why this is so, fix $\delta > 0$, then

$$\begin{aligned}
\pi_T(D_\delta) &= 1 - \pi_T(D \setminus D_\delta) \\
&= 1 - \frac{\int_{D\setminus D_\delta} e^{-\varepsilon(\mathbf{x})/T}\, d\mathbf{x}}{\int_D e^{-\varepsilon(\mathbf{z})/T}\, d\mathbf{z}} \\
&\geq 1 - \frac{\int_{D\setminus D_\delta} e^{-\varepsilon(\mathbf{x})/T}\, d\mathbf{x}}{\int_{D_{\delta/2}} e^{-\varepsilon(\mathbf{z})/T}\, d\mathbf{z}} \\
&\geq 1 - \frac{e^{-(\varepsilon^*+\delta)/T}\varphi(D \setminus D_\delta)}{e^{-(\varepsilon^*+\delta/2)/T}\varphi(D_{\delta/2})} \\
&= 1 - e^{-\frac{\delta/2}{T}} \frac{\varphi(D \setminus D_\delta)}{\varphi(D_{\delta/2})},
\end{aligned}$$

where φ denotes the *Lebesgue-measure* on $\mathbb{R}^N$. Note that the assumptions made on D and ε guarantees $\varphi(D_{\delta/2}) > 0$. The first inequality follows from

the fact that $D_{\delta/2} \subset D$ and the integration is over a non-negative function. The second inequality is valid since $\varepsilon^* + \delta \leq \varepsilon(\mathbf{x})$ on $D \setminus D_\delta$ and $\varepsilon^* + \delta/2 \geq \varepsilon(\mathbf{x})$ on $D_{\delta/2}$. Thus it follows that $\lim_{T \to 0} \pi_T(D_\delta) = 1$.

It can be shown (Romeijn and Smith, 1994), that for every starting point $\mathbf{x}_0$ in D, the Markov chain $\{\mathbf{X}_k; k \geq 0\}$ generated by algorithm 1 converges in total variation to the target distribution π, i.e.

$$\lim_{k \to \infty} P(\mathbf{X}_k \in B | \mathbf{X}_0 = \mathbf{x}_0) = \pi(B), \quad \forall \mathbf{x}_0 \in D, \forall B \in \mathscr{B}. \tag{5.9}$$

Now, fix $T > 0$ and hence the distribution π_T, by Equation 5.9 one can now by algorithm 1 generate a sequence $\{\mathbf{X}_k; k \geq 0\}$ with the property

$$\lim_{k \to \infty} P(\mathbf{X}_k(T) \in D_\delta | \mathbf{X}_0 = \mathbf{x}_0) = \pi_T(D_\delta)$$

which together with the result from Equation 5.7 yields

$$\lim_{T \to 0} \lim_{k \to \infty} P(\mathbf{X}_k(T) \in D_\delta | \mathbf{X}_0 = \mathbf{x}_0) = 1, \quad \forall \delta > 0.$$

This result is promising, but as several authors mentions, it is not always guaranteed that a convergent algorithm outperforms a non-convergent one, (Cohn and Fielding, 1999) even show that an increasing temperature schedule (so-called *boiling*) is to be preferred in some cases. Often is a faster convergence to a near optimum state preferred.

5.3 IMPLEMENTATION

Although simple in theory, many different choices can be made within the simulated annealing framework which will affect the performance of the algorithm. The aspects discussed are aimed for the case where the number of parameters is fixed, but many of the principles also apply for the trans-dimensional case.

The algorithm can be stated as follows (minimization problem)

5.3.1 *Sampling of a new state*

The proposed new solution $\mathbf{y}_{k+1}$ is sampled from some proposal distribution $p(\mathbf{x}|\mathbf{x}_k)$. There are several choices available in the literature (see e.g. Romeijn and Smith (1994)), (Vanderbilt and Louiem, 1984)), and which to choose is problem dependent. In addition to choosing *how* to sample a new solution, there is also the question of *which* parameters to update. The different options on which parameters to update at each iteration is presented below.

(a) *Full sampling.* Every parameter $x_i, i = 1, \ldots, N$.

(b) *Block sampling.* Some parameters $x_i, i \in \mathscr{I} \subset \{1, \ldots, N\}$.

Algorithm 1 Simulated annealing

Initialize $\mathbf{x}_0 \in D$ and temperature T_0.
while termination criteria not met **do**
    Set $\mathbf{y}_{k+1} \leftarrow \text{perturb}(\mathbf{x}_k)$.
    Evaluate $\varepsilon(\mathbf{y}_{k+1})$.
    if $\varepsilon(\mathbf{y}_{k+1}) \leq \varepsilon(\mathbf{x}_k)$ **then**
        $\mathbf{x}_{k+1} \leftarrow \mathbf{y}_{k+1}$.
    else
        Sample $u \sim \mathcal{U}([0,1])$.
        $\alpha \leftarrow \exp\left\{-\frac{\varepsilon(\mathbf{y}_{k+1})-\varepsilon(\mathbf{x}_k)}{T_k}\right\}$.
        if $u \leq \alpha$ **then**
            $\mathbf{x}_{k+1} \leftarrow \mathbf{y}_{k+1}$.
        else
            $\mathbf{x}_{k+1} \leftarrow \mathbf{x}_k$.
        end if
    end if
    Check termination criteria.
    Update T.
    Increment k.
end while

(c) *Single sampling.* Just a single parameter x_i.

Suppose that the feasible set D is bounded by linear constraints, $D = \{\mathbf{x} \in \mathbb{R}^N; \quad l_i \leq x_i \leq u_i, \quad i = 1,\ldots,N\}$. This is the case in this project as can be seen from the parameter constraints in Equation 3.6. The most naive way to sample a new solution is simply to sample from a uniform distribution on the feasible set $y_i \sim \mathcal{U}(l_i, u_i), \quad i = 1,\ldots,N$ (where the obvious iteration index $k+1$ is omitted for brevity). A more adaptive approach is to sample a direction θ_i and a step length λ_i and set $y_i = x_i + \lambda_i\theta_i$ where x_i is the current state of parameter i, $i = 1,\ldots,N$. In addition to the lower and upper bounds on each parameter, the immediate exploration of the solution space is controlled by introducing a minimum threshold $\lambda_i^{\min}$, and a maximum threshold $\lambda_i^{\max}$, on the step lengths for each parameter. With this, a simple approach is to sample a positive or negative unity direction and then sample the step length uniformly on the interval restricted by the feasible boundaries or the step length thresholds. Sampling of the step length is easy in this particular project because of the simple topography of D. For a more exotic topology, one could try to create a simpler space $\tilde{D} \supset D$ and sample from D with a rejection method on $\tilde{D}$.

Algorithm 2 Sample new proposed solution

Given current state $\mathbf{X}_k = \mathbf{x}_k$.
for parameter $i \in \mathscr{I} \subseteq \{1,\dots,N\}$ **do**
  Sample direction $\theta_i \sim \mathscr{U}(\{-1,1\})$.
  if $\theta_i == -1$ and $x_i - l_i < \lambda_i^{\min}$ **then**
    $\theta_i = 1$.
  else if $\theta_i == 1$ and $u_i - x_i < \lambda_i^{\min}$ **then**
    $\theta_i = -1$.
  end if
  if $\theta_i == -1$ **then**
    Sample $\lambda_i \sim \mathscr{U}([\lambda_i^{\min}, \max\{x_i - l_i, \lambda_i^{\max}\}])$.
  else
    Sample $\lambda_i \sim \mathscr{U}([\lambda_i^{\min}, \max\{u_i - x_i, \lambda_i^{\max}\}])$.
  end if
end for

5.3.2 *Cumulative step probabilities*

Consider an arbitrary, feasible parameter value x_i, obtained after the k'th iteration. Since it is feasible, $l_i \le x_i \le u_i$ for some lower bound l_i and upper bound u_i. For some value y, the probability $Pr(x_i + \lambda_i\theta_i \le y)$ can be calculated in the following way.

Using total probability

$$\begin{aligned}\Pr(x_i + \lambda\theta_i \le y) = \Pr(x_i - \lambda_i \le y | \theta_i = -1)\Pr(\theta_i = -1) \\ + \Pr(x_i + \lambda_i \le y | \theta_i = 1)\Pr(\theta_i = 1).\end{aligned}$$

Since θ_i is uniformly either 1 or -1, $\Pr(\theta_i = -1) = \Pr(\theta_i = 1) = 1/2$. Now, let $m_i^l = \max\{x_i - l_i, \lambda_i^{\max}\}$ and $m_i^u = \max\{u_i - x_i, \lambda_i^{\max}\}$ such that $\lambda_i \sim \mathscr{U}([\lambda_i^{\min}, m_i^l\}])$ if $\theta_i = -1$ and $\lambda_i \sim \mathscr{U}([\lambda_i^{\min}, m_i^u\}])$ if $\theta_i = 1$. Then

$$\begin{aligned}\Pr(x_i + \lambda_i\theta_i \le y | \theta_i = -1) &= \Pr(x_i - \lambda \le y) \\ &= 1 - \Pr(\lambda_i \le x_i - y) \\ &= 1 - \begin{cases} 0 & , x_i - y \le \lambda_i^{\min} \\ \frac{x_i^k - y - \lambda_i^{\min}}{m_i^l - \lambda_i^{\min}} & , \lambda_i^{\min} \le x_i - y \le m_i^l \\ 1 & , x_i - y \ge m_i^l \end{cases} \\ &= \begin{cases} 1 & , x_i - y \le \lambda_i^{\min} \\ \frac{m_i^l - (x_i^k - y)}{m_i^l - \lambda_i^{\min}} & , \lambda_i^{\min} \le x_i - y \le m_i^l \\ 0 & , x_i - y \ge m_i^l \end{cases}\end{aligned}$$

Likewise

$$\begin{aligned}\Pr(x_i+\lambda_i\theta_i \le y|\theta_i=1) &= \Pr(x_i+\lambda \le y)\\ &= \Pr(\lambda_i \le y-x_i)\\ &= \begin{cases} 0 & ,y-x_i \le \lambda_i^{\min}\\ \frac{y-x_i-\lambda_i^{\min}}{m_i^u-\lambda_i^{\min}} & ,\lambda_i^{\min} \le y-x_i \le m_i^u\\ 1 & ,y-x_i \ge m_i^u \end{cases}\end{aligned}$$

This yields the total step selection cumulative probability

$$Pr(x_i^k+\lambda_i\theta_i \le y) = \begin{cases} 0 & ,y \le x_i - m_i^l\\ \frac{1}{2}\frac{m_i^l-(x_i-y)}{m_i^l-\lambda_i^{\min}} & ,x_i-m_i^l \le y \le x_i-\lambda_i^{\min}\\ \frac{1}{2} & ,x_i-\lambda_i^{\min} \le y \le x_i+\lambda_i^{\min}\\ \frac{1}{2}\frac{m_i^u+y-x_i-2\lambda_i^{\min}}{m_i^u-\lambda_i^{\min}} & ,x_i+\lambda_i^{\min} \le y \le x_i+m_i^u\\ 1 & ,y \ge x_i+m_i^u. \end{cases} \quad (5.12)$$

5.3.3 *Temperature schedule*

The pace at which the temperature T is decreasing is determined by the temperature schedule. One of the most common is the geometric schedule

$$T_{k+1} = \nu_g T_k = \nu_g^{k+1} T_0 \quad (5.13)$$

for some $\nu_g \in (0,1)$, (Kirkpatrick et al., 1983). Another schedule that decay more rapidly in the beginning is

$$T_{k+1} = \frac{T_0}{1+k\nu_l T_0}, \quad (5.14)$$

as proposed in (Lundy and Mees, 1986). Here $\nu_l > 0$ and in both cases, k is the iteration number when the temperature is updated and T_0 is the initial temperature. These are both deterministic schedules, and both the initial temperature and the tuning constant need to be set before the algorithm is run. This constants is problem dependent, and often some fine tuning and experience is needed in order to set them appropriate. Another way to calculate them is presented in the end of this subsection.

An adaptive schedule is also implemented, this is developed in (Romeijn and Smith, 1994) with the intention to inherit some favourable properties of a so-called *Pure Adaptive Search* introduced in (Patel et al., 1988).

$$t_k(\mathbf{x}_0,\ldots,\mathbf{x}_k) = \tau(\min_{0\le j\le k} \varepsilon(\mathbf{x}_j)), \tag{5.15a}$$

$$\tau(g) = 2\frac{g-\varepsilon^*}{\chi^2_{1-p}(N)}. \tag{5.15b}$$

Here, χ^2_{1-p} is the $100(1-p)$ percentile point of the chi-squared distribution with N degrees of freedom, (Romeijn and Smith, 1994) suggests $p = 0.01$. ε^* is the function value at the global minimum, which is assumed known. If this is not the case, it is suggested to replace this value with some estimate $\tilde{\varepsilon}(\mathbf{x}_0,\ldots,\mathbf{x}_k)$, of the minimum value. With this schedule, the temperature is scaled with the function values obtained throughout the run.

Another important aspect is *when* to update temperature. For every fixed temperature T, the chain of states generated by the algorithm converges to the limiting distribution π_T. In order not to violate the sufficient conditions of a (weak) ergodic search, the temperature need to decrease slow enough (Ingber, 1995). Every time the temperature is updated, the equilibrium is interrupted, and therefore it is advised to wait until this equilibrium is reached (or at least almost reached). However, in order to increase efficiency, (Romeijn and Smith, 1994) suggests to update the temperature each time a new state is accepted. (Singh et al., 2008) introduce an epoch constant E and updates the temperature every $E \cdot N$ iterations. It is proposed that $E = 15$, something that is also suggested in (Vanderbilt and Louiem, 1984).

A final aspect to consider is how to decide the initial temperature. From the adapted schedule in Equation 5.15 the initial temperature is obtained as

$$T_0 = 2\frac{\varepsilon(\mathbf{x}_0)-\varepsilon^*}{\chi^2_{1-p}(N)}$$

and this temperature can of course be used in the other schedules as well. Another approach is to select the initial temperature based on the initial function value and acceptance probability. Define $\Delta\varepsilon$ as a positive jump, and let p_0 be the probability of accepting this jump. With this

$$T_0 = -\frac{\Delta\varepsilon_0}{\log p_0} \tag{5.16}$$

and is in that way adjusted to the probability of accepting a jump with increasing function value $\Delta\varepsilon_0$. Typically this probability is set quite high in order to be able to explore greater parts of the solution space at initial iterations, but this is of course dependent on the initial jump. In (Singh

et al., 2008) it is proposed that $\Delta\varepsilon = \max\{\tilde{\varepsilon}_1,\ldots,\tilde{\varepsilon}_d\} - \min\{\tilde{\varepsilon}_1,\ldots,\tilde{\varepsilon}_d\}$ where $\{\tilde{\varepsilon}_1,\ldots,\tilde{\varepsilon}_d\}$ is a set of d function values from randomly sampled solutions.

Using the same principle as presented above, it is possible to calculate the tuning constants in the deterministic schedules. Let $\Delta\varepsilon_K$ be a jump with increasing function value at iteration K, which can be the terminal iteration. Let p_K be the probability of accepting this step at iteration K, then

$$T_K = -\frac{\Delta\varepsilon_K}{\log p_K}. \tag{5.17}$$

Contrary to a large value for p_0, p_K is set quite low and $\Delta\varepsilon$ quite small to avoid acception of steps with increasing function value at the end of the algorithm. With both T_0 and T_K set, one can calculate the constants in the deterministic schedules as

$$\nu_g = \left(\frac{T_K}{T_0}\right)^{\frac{1}{K}}, \quad \nu_l = \frac{1}{KT_0}\left(\frac{T_0}{T_K} - 1\right). \tag{5.18}$$

5.3.4 *Termination*

When to terminate the algorithm is also an important subject that need to be covered. In this project there is implemented three different ones.

Terminating the search after a maximum number of iterations K is reached is perhaps the most obvious criterion, and this should be implemented in every algorithm of this kind. Also by imposing a maximum number of iterations, one can fit the cooling schedules and calculate the constants based on starting temperature and end temperature as was discussed in the previous subsection.

Another obvious criterion is to terminate the search when the global minimum is reached, or at least when the current solution is within some percentage of the alleged global minimum value. This percentage depends on the problem and how close to the optimal solution one need to get. A tendency in solving this problem with simulated annealing is that the decrease in the accepted function values is rapid in the beginning of the search and significantly slower at the end, as can be seen in Figure 5.1. Often one is satisfied with a solution that is close enough to the optimal, and therefore one need to consider if it is necessary to keep the search alive.

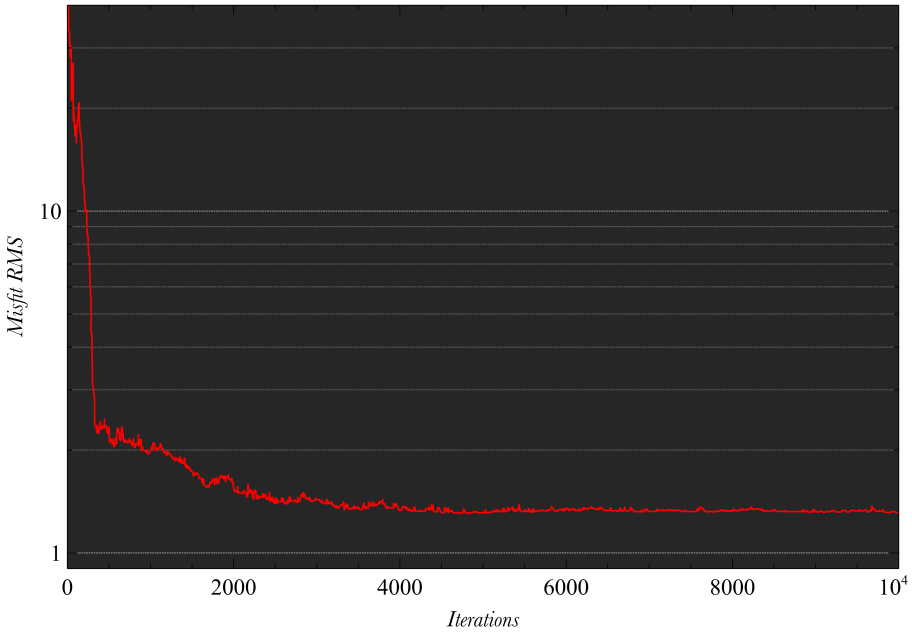


Figure 5.1: Misfit RMS trace plot example.

Often there is no information about the global minimum value, and different methods needs to be used. One could of course estimate the minimum value based on experience, but one could also exploit other qualities about the global minimum. The idea is that close to the global minima, the relative difference in accepted solutions will differ very little, as can be seen from Figure 5.1. It is suggested in (Vanderbilt and Louiem, 1984) to store every accepted function value in a segment of $E \cdot N$ iterations, where E is the same epoch constant introduced above. At the end of each segment, the mean value $\bar{\varepsilon}$ and minimum value ε_{min} of the function values obtained in that segment is computed, and the algorithm is terminated if

$$\frac{\bar{\varepsilon} - \varepsilon_{min}}{\bar{\varepsilon}} < \xi. \tag{5.19}$$

It is further suggested that $\xi = 10^{-3}$, but this is highly problem dependent. Also, because of the ill-posed nature of the problem, there is no unique solution to the problem. The solution space topography close to the minimum values is very flat because of this, so the termination criteria posed in Equation 5.19 hinder the algorithm from searching further when a satisfactory solution is found.

For this particular project, it was discovered that using the stopping criterion in Equation 5.19 resulted in a too early termination. This has to do with the topology of the solution space and that few solutions is accepted close to the global minimum. This tendency is utilized by concentrating the search every time the stopping criterion in Equation 5.19 is true. That is, every time the stopping criterion is imposed, the maximum allowed steplength is decreased with some factor, say 1/2. This limits the exploration

of the solution space, but close to the optimal solution this is a desired property as it reduces the number of deprecated proposed solutions. The algorithm is terminated when the above criterion is imposed a given number of times.

5.4 TRANS-DIMENSIONAL SIMULATED ANNEALING

As discussed in Section 3.1.5, the parameterization of the sub seabed resistivity

$$\begin{aligned}\rho(z;\mathbf{x}) &= \rho_p(z;\mathbf{x}_p)+\rho_r(z;\mathbf{x}_r)+\rho_g(z;\mathbf{x}_g)\\ &= \sum_{n_p=0}^{N_p} c_{n_p}(z-z_{wd})^{n_p} + \sum_{n_r=1}^{N_r} \rho^r_{n_r}[H(z^r_{n_r})-H(z^r_{n_r}+t)]\\ &+ \sum_{n_g=1}^{N_g} \beta_{n_g}\frac{1}{\sqrt{2\pi w^2_{n_g}}}\exp\left\{-\frac{1}{2w_{n_g}}(z-z^g_{n_g})^2\right\}.\end{aligned}$$

is implemented with the option of dynamic dimensions N_p, N_r and N_g. This with the intention that the algorithm would adjust itself to the most appropriate number of parameters.

Define a set of models $\mu = \{\mu_0,\ldots,\mu_{N_m-1}\}$ where a model refers to a configuration of polynomial, resistor and Gaussian terms respectively, (N_p,N_r,N_g). In this project the model index mapping $(N_p,N_r,N_g)\rightarrow j$ is as follows. μ_0 corresponds to a model with just the constant, zeroth order polynomial term, that is $(0,0,0)$. μ_1 corresponds to a model with one additional Gaussian term $(0,0,1)$. This continues, and the mapping from the triplet (N_p,N_r,N_g) to the model index is $j = N_p(N_r^{\max}+1)(N_g^{\max}+1)+N_r(N_g^{\max}+1)+N_g$. Here $0\le N_p\le N_p^{\max}$,$0\le N_r\le N_r^{\max}$ and $0\le N_g\le N_g^{\max}$ so that the number of different models is $N_m = (N_p^{\max}+1)(N_r^{\max}+1)(N_g^{\max}+1)$. Now, let D_j be the solution space of model μ_j, then the objective is to find the model μ^* and corresponding variable values $\mathbf{x}\in D^*$ that minimizes the misfit function $\varepsilon(\mathbf{x})$.

Two different approaches to solving this trans-dimensional problem is implemented; one simple and naive approach, and one approach based on a method suggested in (Singh et al., 2008). For practical reasons, the constant, zeroth order polynomial term cannot be removed.

5.4.1 *Equal treatment*

In this approach the idea is to treat the parameter dimensions almost as ordinary parameters meaning that in each iteration, there is a probability that the dimensions is altered, that is, to change model. More specifically, before the next solution is proposed, there is a certain probability of increasing one of the parameter dimensions N_p, N_r or N_g and a certain probability

of decreasing one of them. There is also a certain probability to keep the current dimensionality, that is, not change any of them. If change is imposed, the number of terms can only be increased or decreased by one. When reducing the number of dimensions, two different methods is implemented. Either remove the last term (the newest added term) or remove a random term. The first one makes the different solution vectors into so-called last-in-first-out (lifo) lists and has the advantage that the latest included terms have greater probability of being removed than terms added earlier. The consequence is that earlier terms have more time to be fitted properly. A downside is that if a less favourable term has been added before a desired one, the latest added term will not be removed due to its good fit, and hence will not the less favourable term be removed. The last issue is resolved by removing every term with equal probability.

Algorithm 3 trans-dimensional simulated annealing with equal treatment parameter dimension and parameter values.

Initialize $\mathbf{x}_0 \in D \subseteq \mathbb{R}^N$ and temperature T_0
Determine $p_0, \ldots, p_6$ such that $\sum_{j=0}^{6} = 1$
while termination criteria not met **do**
    Do *one* of the following marked with $\star$.
    $\star$ Do not change the dimension with probability p_0.
    $\star$ $N_p^{k+1} = N_p^k + 1$ with probability p_1.
    $\star$ $N_p^{k+1} = N_p^k - 1$ with probability p_2.
    $\star$ $N_r^{k+1} = N_r^k + 1$ with probability p_3.
    $\star$ $N_r^{k+1} = N_r^k - 1$ with probability p_4.
    $\star$ $N_g^{k+1} = N_g^k + 1$ with probability p_5.
    $\star$ $N_g^{k+1} = N_g^k - 1$ with probability p_6.
    Sample $\mathbf{y}_{k+1}$ from some proposal distribution.
    Evaluate $\varepsilon(\mathbf{y}_{k+1})$.
    if $\varepsilon(\mathbf{y}_{k+1}) \leq \varepsilon(\mathbf{x}_k)$ **then**
        $\mathbf{x}_{k+1} \leftarrow \mathbf{y}_{k+1}$.
    else
        Sample $u \sim \mathcal{U}([0,1])$.
        $\alpha \leftarrow \exp\left\{-\frac{\varepsilon(\mathbf{y}_{k+1}) - \varepsilon(\mathbf{x}_k)}{T_k}\right\}$.
        if $u \leq \alpha$ **then**
            $\mathbf{x}_{k+1} \leftarrow \mathbf{y}_{k+1}$.
        else
            $\mathbf{x}_{k+1} \leftarrow \mathbf{x}_k$.
        end if
    end if
    Check termination criteria.
    Update T.
    Increment k
end while

As can be seen from algorithm 3, the method is as in algorithm 1 except from the part where the dimensions is proposed. The dimensions of the proposed next state $\mathbf{y}_{k+1}$ changes accordingly.

If no change is imposed, the next state is chosen as discussed in Section 5.3.1. If, however the dimensionality is increased or decreased will be explained in terms of an example.

Suppose that one additional Gaussian term is added, that is $N_g^{k+1} = N_g^k + 1$. Then 3 new parameters are created: $\beta_{N_g^k+1}$, $z_{N_g^k+1}$ and $w_{N_g^k+1}$. Every one of them need to be assigned a value from some proposal distribution and inserted into the proposed next state $\mathbf{y}_{k+1}$ which now has dimension $dim(\mathbf{y}_{k+1}) = N_k + 3$. The respective parameters can be sampled from different distributions, e.g. uniformly on the feasible domain of the particular parameter or normally around a predetermined mean with a predetermined standard deviation. The remaining parameter values is inherited from the previous accepted solution. In that way only the added term contribute to the alteration of the proposed solution. This new term could of course be added into the parameterization where the remaining parameters was perturbed as discussed in Section 5.3.1.

Now, suppose that a Gaussian term is removed, that is $N_g^{k+1} = N_g^k - 1$. Then 3 parameters from the previous accepted solution is removed: β_j, z_j and w_j. The two choices implemented is to remove elements indexed with $j = N_g^k$ or at random $j \in \{1, \ldots N_g^k\}$. As in the addition case, the remaining $N-3$ parameters could either be perturbed as discussed in Section 5.3.1, or remain unchanged in the new proposed solution $\mathbf{y}_{k+1}$.

This method is presented with no references to external literature. This is because it was created specifically for the purpose of dealing with dynamic parameter dimensionality in this particular application. As it seems like an intuitive extension to the ordinary simulated annealing with fixed parameter values, I am sure that similar approaches has been used on similar problems before, but I did not succeed in finding any literature on the subject using this approach. Therefore, another method using a different approach to solving trans-dimensional optimization problems with simulated annealing is implemented and presented in the next subsection.

5.4.2 *Fitness based model exploration*

The method presented here is based on a method proposed in (Singh et al., 2008) for solving single objective, trans-dimensional optimization problems. It is included here for reference purposes as it is a well documented technique applied on similar problems, whilst the approach discussed in Section 5.4.1 is not, as far as the author of this thesis knows. In this method, all models μ are considered and explored in a more systematic way than in the previous method.

Let, as before, N_m be the number of models and define an *archive* $A = \{A_0, \ldots, A_{N_m-1}\}$ which is a collection of lists A_j containing all accepted function values from the corresponding model, μ_j. Then, let $\tilde{F} = \{\tilde{F}_0, \ldots, \tilde{F}_{N_m-1}\}$ where $\tilde{F}_j$ is the average of the smallest 20% function values contained in A_j (this is the percentage recommended in (Singh et al., 2008)). Define the *fitness* of a model μ_j as

$$F_j = \frac{\max\{\tilde{F}\} - \tilde{F}_j}{\max\{\tilde{F}\} - \min\{\tilde{F}\}}$$

and collect them in $F = \{F_0, \ldots, F_{N_m-1}\}$. With this, the models with the lowest 20% function values are given the highest fitness value. Initially, a function value from a random solution from each of the models are stored in the archive A.

Now, choose one of the models μ_j. The model μ_j is chosen using one of the two following options.

1) With probability p_f, choose a model from a *fitness proportionate selection* based on the fitness values F such that the probability that model μ_j is chosen is

$$p_j = \frac{F_j}{\sum_{i=0}^{N_m-1} F_i}, \quad j = 0, \ldots, N_m - 1.$$

2) Completely at random with probability $(1-p_f)$, such that the probability that model μ_j is chosen is

$$p_j = \frac{1}{N_m - 1}, \quad j = 0, \ldots, N_m - 1.$$

In total, the probability of selecting model μ_j is

$$p_j = \frac{F_j}{\sum_{i=0}^{N_m-1} F_i} p_f + \frac{1}{N_m - 1}(1 - p_f)$$

for each model $\mu_j, j = 0, \ldots, N_m - 1$. It is recommended that p_f is quite high, so that models with greater fitness is explored more frequently (Singh et al., 2008).

Then, when a model μ_j is chosen, run an ordinary simulated annealing routine for $E \cdot N_j$ iterations, where E is the epoch constant defined earlier and N_j is the number of parameters in model μ_j. Store every accepted function value in its respective bin A_j in the archive and at the end of each epoch, update the fitness F, the temperature T, and check the termination criteria. The method is summarized in algorithm 4.

As mentioned above, this method is included for comparison with the method proposed in Section 5.4.1, but in all fairness, it is not very well suited for this particular problem. In (Singh et al., 2008), it is applied

Algorithm 4 Simulated annealing

Initialize T, A and pick $\mathbf{x}_0 \in A$ randomly.
while termination criteria not met **do**
    Calculate F
    Select model based on fitness with probability p_f or at random with probability $(1 - p_f)$
    Calculate epoch length $E \cdot N$
    for $j = 1 \rightarrow E \cdot N$ **do**
        Set $\mathbf{y}_{k+1} \leftarrow \text{perturb}(\mathbf{x}_k)$
        Evaluate $\varepsilon(\mathbf{y}_{k+1})$
        if $\varepsilon(\mathbf{y}_{k+1}) \leq \varepsilon(\mathbf{x}_k)$ **then**
            $\mathbf{x}_{k+1} \leftarrow \mathbf{y}_{k+1}$
        else
            Sample $u \sim \mathcal{U}([0,1])$
            $\alpha \leftarrow \exp\left\{-\frac{\varepsilon(\mathbf{y}_{k+1}) - \varepsilon(\mathbf{x}_k)}{T_k}\right\}$
            if $u \leq \alpha$ **then**
                $\mathbf{x}_{k+1} \leftarrow \mathbf{y}_{k+1}$
            else
                $\mathbf{x}_{k+1} \leftarrow \mathbf{x}_k$
            end if
        end if
        Increment k
    end for
    Check termination criteria.
    Update T and A
end while

on combinatorial problems with 5 and 15 different models and still needs 30000 function evaluations. As the models is defined above, the number of models in this problem is $(N_p^{\max}+1)\times(N_r^{\max}+1)\times(N_r^{\max}+1)$, which quickly becomes quite large if one want some flexibility with the model dimensions. As a comparison, the algorithm presented in Section 5.4.1 is implemented with $N_p^{\max} = N_r^{\max} = N_g^{\max} = 10$, and a comparable dimensionality is therefore not practically feasible, as the whole purpose of this thesis is to reduce number of function evaluations in an inversion problem.

Part II

PROCEDURE AND RESULTS

6

METHODOLOGY

This chapter aims to elucidate details about the actual realization of the problem task, and how the theory is incorporated in order to obtain the results presented in Chapter 7.

6.1 MODELING

The parameter solution vector $\mathbf{x}$ determine the resistivity profile function $\rho(z;\mathbf{x}_k)$ completely. Details concerning the modelling of this resistivity profile function such that it is in accordance with the plane layer model discussed in Chapter 4 is presented Section 6.1.1.

The methods presented in this thesis was tested on different data sets, both synthetic and real. Elements of the models behind the resulting data is discussed in Section 6.1.2 for the synthetic data and Section 6.1.3 for the real data.

6.1.1 *Discretization of the model space*

The model space, that is, the physical 1D space below the sea surface is discretized in a set of layers separated by interfaces.

This layering of the model space is equal to the one discussed in Chapter 4. The number of layers at which the resistivity is to be determined is termed N_l, excluding the water layer where the resistivity is assumed known. The number of interfaces is thus N_l, including the interface at the sea bed. This system of layers and interfaces is illustrated in Figure 6.1. Because of the parameterization there is N number of unknowns to be determined by the optimization routine, but in practice, the resistivity is to be determined at N_l layers. The parameterization simplifies the optimization in the way that it optimizes for fewer variables, but the number of evaluation points N_l is still the same, and the computational cost at each forward call is not reduced (the parametrization reduces hopefully the number of forward calls though). In mod-

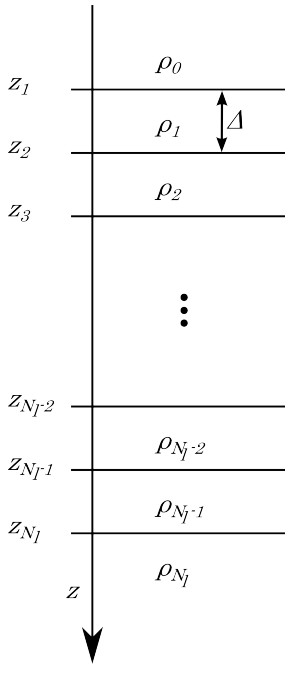


Figure 6.1: Illustration of one dimensional model plane layered model.

eling the model space, one is free to chose how many layers the model should consist of, and a denser grid of layers means a more accurate representation of the resistivity profile and is preferable. However, one want to balance this with the increasing computational cost of adding more layers. In this thesis, both the water resistivity ρ_0 and the water depth z_1 is measurable and assumed known. The other interfaces is placed down in the formation as can be seen in Figure 6.1. In this thesis, they are placed with an increasing spacing between the interfaces as

$$z_k = z_1 + (k-1)\Delta + \frac{(k-1)(k-2)}{(N_l-1)(N_l-2)}[z_{N_l} - z_1 - (N_l-1)\Delta] \tag{6.1}$$

for $1 \leq k \leq N_l$, where Δ is the spacing between the first and second interface. z_1, z_{N_l} and N_l are fixed, predetermined modeling constants. Let

$$\zeta = \frac{z_{N_l} - z_1}{\Delta(N_l - 1)}, \tag{6.2}$$

then ζ in Equation 6.2 determines if the spacing is increasing ($\zeta > 1$), decreasing ($\zeta < 1$) or equal ($\zeta = 1$). In order to decrease computational cost some increase in the spacing is used in this thesis. The initial spacing is set to $\Delta = 5$ m, the last interface is placed at a depth of $z_{N_l} = 2500$ m, and the number of interfaces is set to $N_l = 150$. The water depth vary from model to model, and will be specified when the different models are presented in **??**. The increased spacing can be justified by the fact that the resolution of the EM field is decreasing with increasing depth. The EM field's attenuation with propagation distance in conductive media can be described by the skin depth

$$\delta \approx \sqrt{\frac{\rho}{\pi \mu_0 f}}$$

which is a measure of the depth from the surface of the conductor where the EM field will be reduced by a factor e^{-1} in amplitude (Amundsen et al., 2006). Here, the magnetic permeability μ is set to that of free space $\mu = \mu_0 = 4\pi \times 10^{-7}$ H/m, and with a water resistivity of $\rho = 0.271$ Ωm and emittance frequency of $f = 0.25$ Hz, the skin depth is $\delta \approx 524$ m.

6.1.2 *Synthetic model*

With the same modeling tools used in the inversion, a synthetic model is made. That is, a sub seabed resistivity profile is created, and resulting field components is computed for a specified set of source points and emitting frequencies.

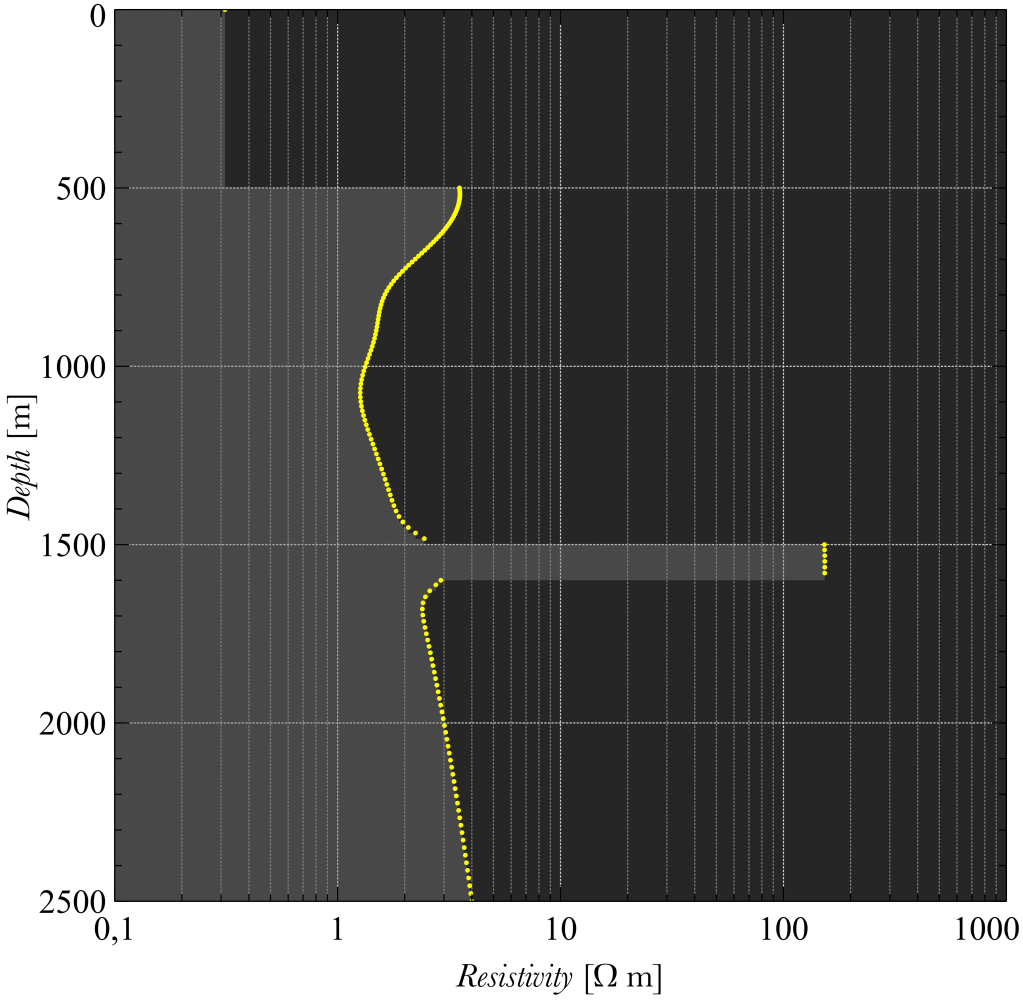


Figure 6.2: Resistivity profile of synthetic model.

The model consists of a 500 m deep water column with resistivity of 0.3125 Ωm and a resistive target resistor at 1500 m. This is summarized in Table 6.1 with the additional resistivity function parameters used to create the model.

PARAMETERS	VALUES
ρ_w	0.3125 Ω m
z_{wd}	500 m
ζ_0	0.0
ζ_1	4.0
ρ_1^r	150.0 Ω m
z_1^r	1500.0 m
β_1	1500.0 Ω m^2
β_2	200.0 Ω m^2
β_3	500.0 Ω m^2
z_1^g	500.0 m
z_2^g	900.0 m
z_3^g	1550.0 m
w_1	170.0 m
w_2	100.0 m
w_3	50.0 m

Table 6.1: Parameter values for the synthetic model.

6.1.3 *Troll West Oil Province*

The method is also tested on real data which is collected from a survey EMGS made on the *Troll West Oil Province* in 2009 (Gabrielsen et al., 2009). This field is located to the west of Bergen, Norway as can be seen in Figure 6.3.

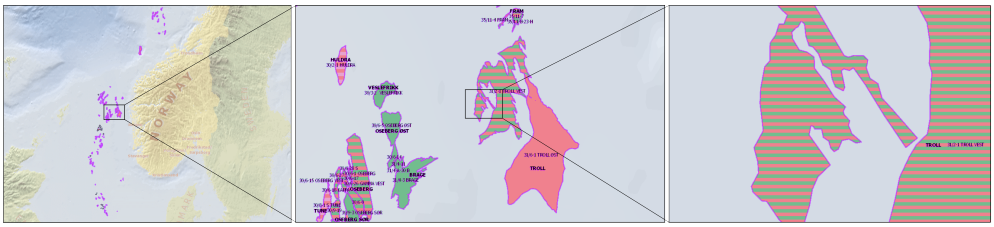


Figure 6.3: Location of the Troll West Field.

The oil province has a top reservoir at about 1500 m and consists of a thin gas column of approximately 20 m followed by a oil column of approximately 20 m. A well log showing the measured horizontal resistivity is provided in Figure 6.4.

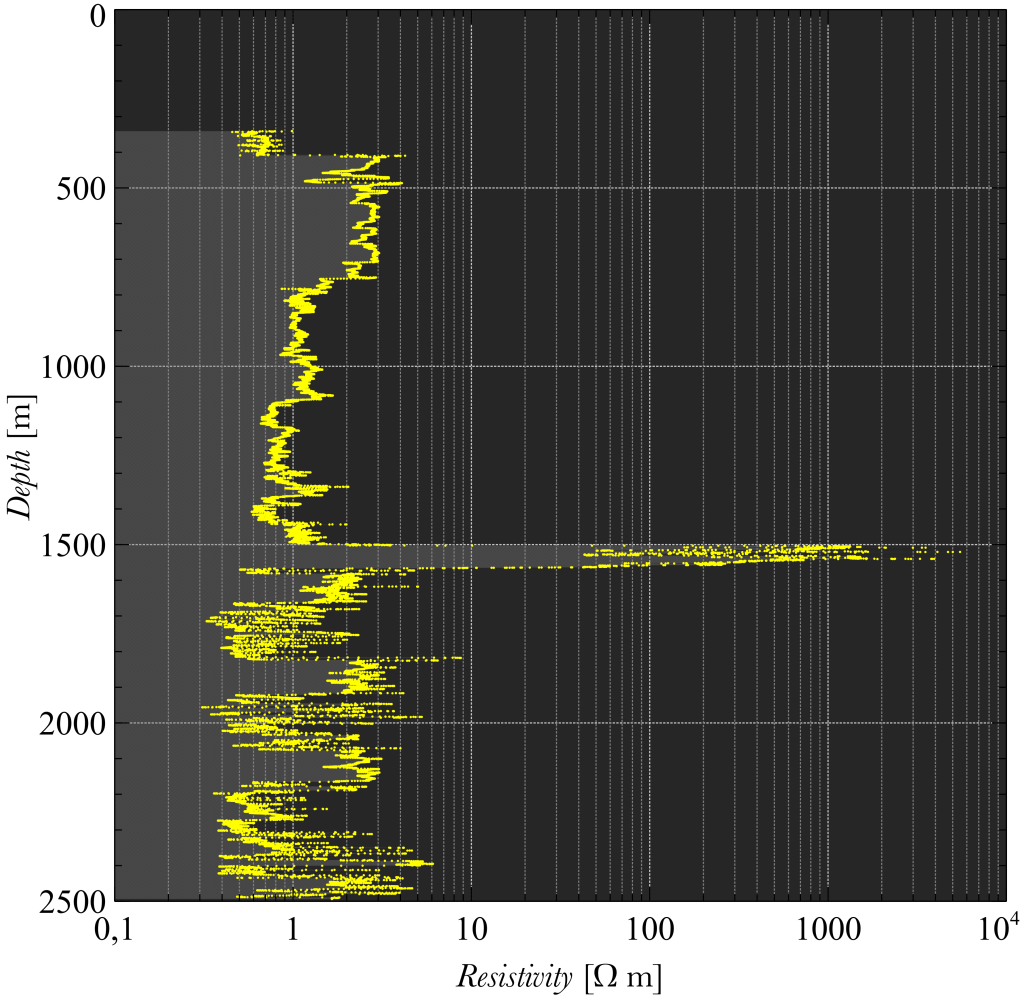


Figure 6.4: Processed resistivity data from well log from well 31/2 - 5, which location can be seen in Figure 6.6. Measurements starts at about 340 m below the sea surface and continues to about 2500 m below sea surface. The sea depth is about 407 m. In lithostratigraphic terms, the *Nordland group* starts at about 365 m down to the *Hordaland group* from 788 m down to the *Rogaland group* at 1346 m which continues down to the *Shetland* and *Viking group* at 1533 m and 1536 m respectively. The *Sognefjord formation* starts at 1536 m and this is where the hydrocarbon saturated layer, or resistor is (Norwegian Petroleum Directorate, 2014). Note that in order to use the measured data from the well, it is usually processed by experienced analysts before it appears as in this figure.

The 1D inversions made in this thesis finds the *vertical resistivity*, whilst the well log in Figure 6.4 shows measured *horizontal resistivity* on a small length scale. To get an indication of the vertical resistivity in a given depth

interval of length $|\mathcal{Z}|$, one could obtain an approximation $< \rho_v >_{\mathcal{Z}}$ of the vertical resistivity in this interval through

$$< \rho_v >_{\mathcal{Z}} = \frac{\Delta h}{|\mathcal{Z}|} \sum_{z_i \in \mathcal{Z}} \rho_h(z_i),$$

where $\mathcal{Z}$ is the collection of measured horizontal resistivity values in this interval, and Δh is a typical spacing between the layers in the plane layered model. The well bore log with this vertical resistivity approximation is showed in Figure 6.5.

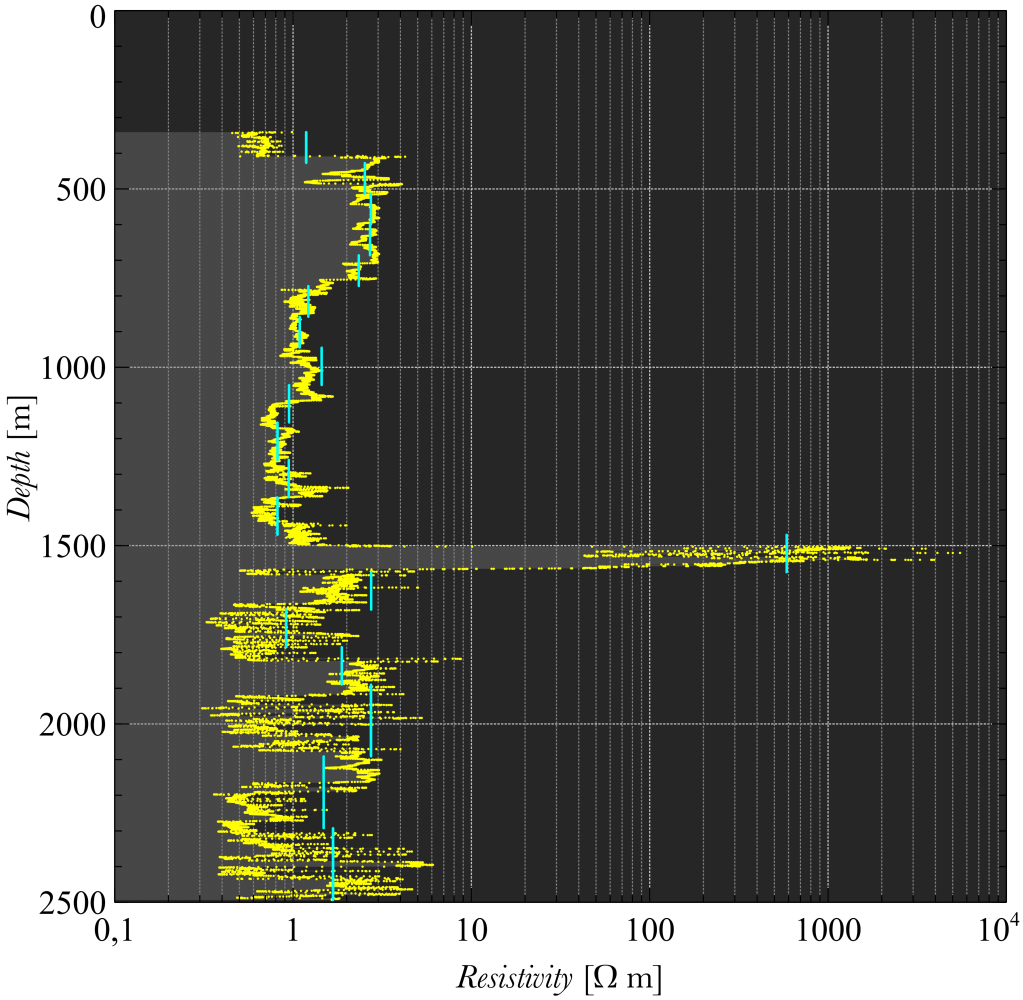


Figure 6.5: Well bore log from the Troll survey showing measured horizontal resistivity (yellow) together with approximated vertical resistivity (cyan).

The transverse resistivity in the hydrocarbon saturated reservoir can be calculated by integrating the well log over the reservoir interval, with this well log, it is then calculated to be about 35640 Ωm.

The observed real data used in this thesis is from a 3D survey on the Troll West Oil Province conducted by EMGS. The data is acquired with a

dense grid with nominal line and receiver spacing of 1.25 km and additional towlines in the east-west direction as illustrated in Figure 6.6.

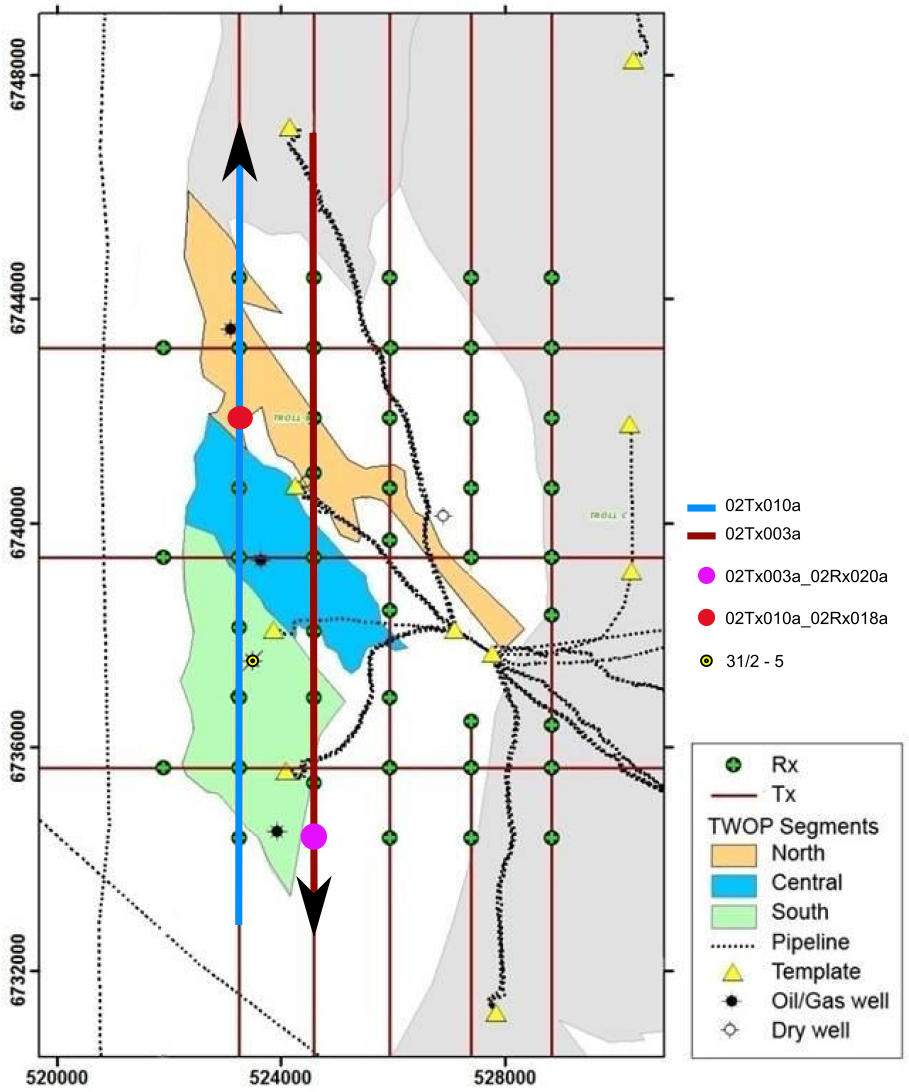


Figure 6.6: Troll West Oil Province surveys with towlines and receiver positions.

Data from two different receivers are inverted, and they are termed `02Rx020a` and `02Rx018a` in accordance with the EMGS survey nomenclature. Each have a corresponding receiver towline from which the signal is emitted, and data gathered from when the source is towed *towards the receiver* (termed *intow*) is inverted separately from data obtained from when the source is towed *away from the receiver* (termed *outtow*).

The location of receiver `02Rx020a` is pictured with the magenta colored dot and its corresponding towline directed south, is pictured as the burgundy colored line in Figure 6.6. The resulting field response is plotted in Figure 6.7. The location of receiver `02Rx018a` is pictured with the red dot and its corresponding towline directed north, is pictured as the blue line in Figure 6.6. The field response is plotted in Figure 6.8

This, and other modeling data are summarized in Table 6.2. The names attached to each model is kept throughout this thesis to refer to the different models from which the data is acquired.

	TROLL 1	TROLL 2	TROLL 3	TROLL 4
Receiver	`02Rx020a`	`02Rx020a`	`02Rx018a`	`02Rx018a`
Towline	`02Tx003a`	`02Tx003a`	`02Tx010a`	`02Tx010a`
Towing direction	South	South	North	North
Source direction	Intow	Outtow	Intow	Outtow
ρ_w	0.271 Ω m	0.271 Ω m	0.271 Ω m	0.271 Ω m
z_{wd}	325.6 m	325.6 m	333.6 m	333.6 m

Table 6.2: Parameter values Troll models.

6.2 DATA SETS

The models discussed in Section 6.1.2 and Section 6.1.3 can give rise to several different sets of data since one can choose to invert for different configurations of field components, frequencies and source points. In what follows, the configurations used in this thesis is presented and explained.

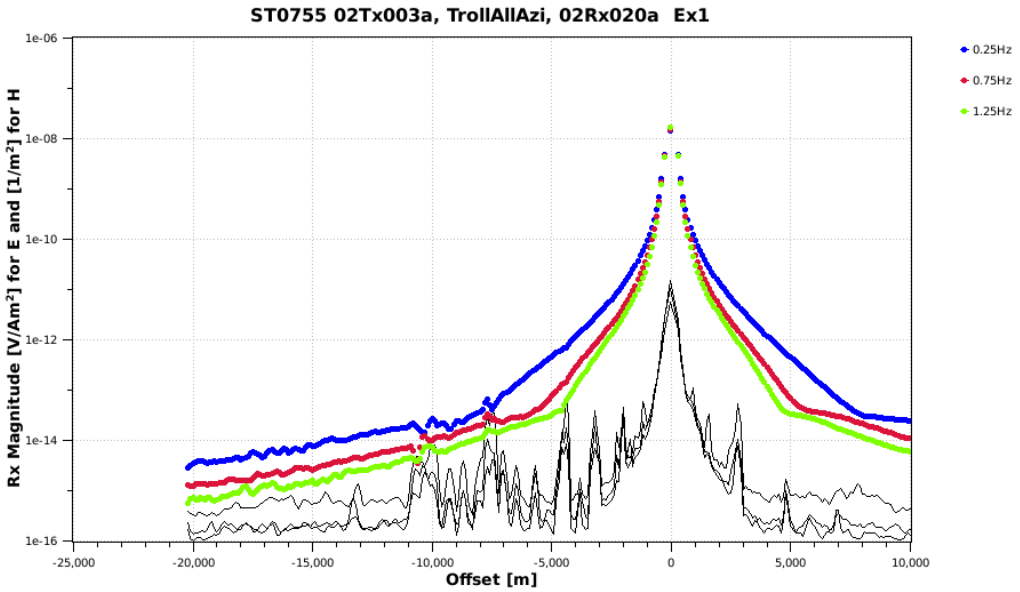
The number of gathered data M, is determined by four factors: M =(number of source points)×(number of receivers)×(number of frequencies)×(number of field components). This is the number of known data in the system, and a smaller data sample means less information about the problem and a more difficult problem to solve and a less accurate solution. This must however, be balanced with the increasing computational cost of increasing the number of sampled data. As an example, the *Jacobian matrix* in a gradient based optimization routine is of size $M \times N$.

Three emitting frequencies is used, {0.25 Hz, 1.00 Hz, 2.00 Hz} in the synthetic model, and {0.25 Hz, 0.75 Hz, 1.25 Hz} for the troll models. These are included because they are likely to pick up the target (Gabrielsen et al., 2009). The source points included lies in the range of 1000 m to 10000 m, with a shot spacing of 100 m, this is so for both the synthetic and real models. This offset interval is chosen since this is the interval in where the most desired field contributions is, as discussed in Section 2.1. Only one of the field components, namely E_x is included, since this is the component with the most significant field response in this case. This results in a total number of data $M = 91 \times 1 \times 3 \times 1 = 273$.

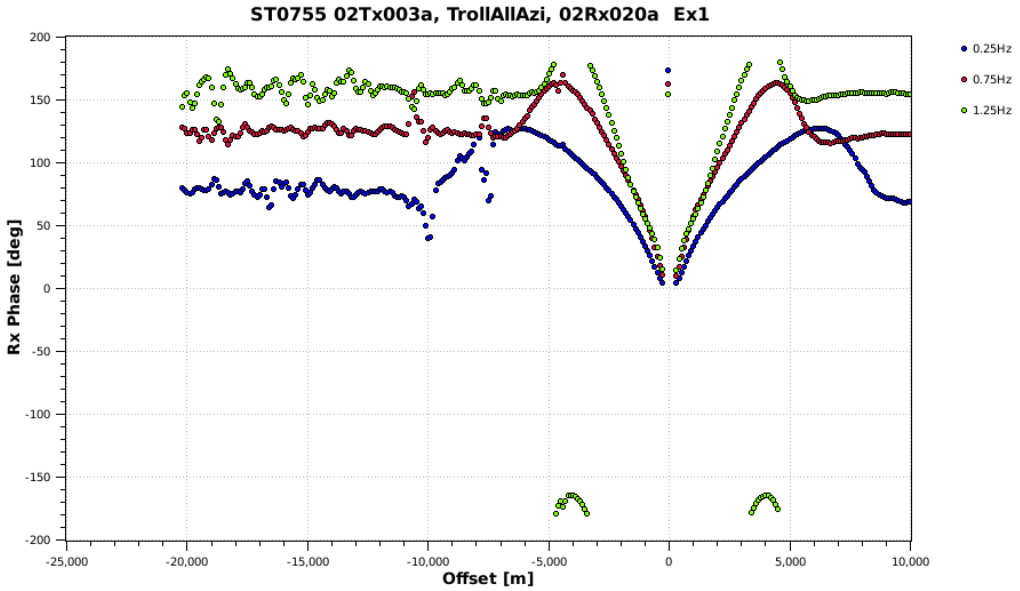
273 is the number of data included in the synthetic data case, with real data however, there is fewer. This is because one can chose to skip data where the field magnitude is less than some threshold. Also there is the possibility to skip data where the signal to noise ratio (snr) is less than

some threshold. In this thesis, thresholds is set to $1 \times 10^{-15}\,\mathrm{V\,m^{-1}}$ and 20 respectively. As can be seen from Figure 6.7(a) and Figure 6.8(a), some data is below the previously mentioned thresholds, resulting in the following number of included data points for the previous discussed data sets. Troll 1: M = 144, Troll 2: M = 190, Troll 3: M = 190, Troll 4: M = 190.

The resulting field response observed by the two receivers displayed in Figure 6.7 and Figure 6.8 respectively. The data inverted in Troll 1 is the field response with negative offset in Figure 6.7 whilst the data inverted in Troll 2 is the field response with positive offset in the same figure. In a similar manner, the data inverted in Troll 3 is the field response with negative offset in Figure 6.8 and for Troll 4 it is the data with positive offset.

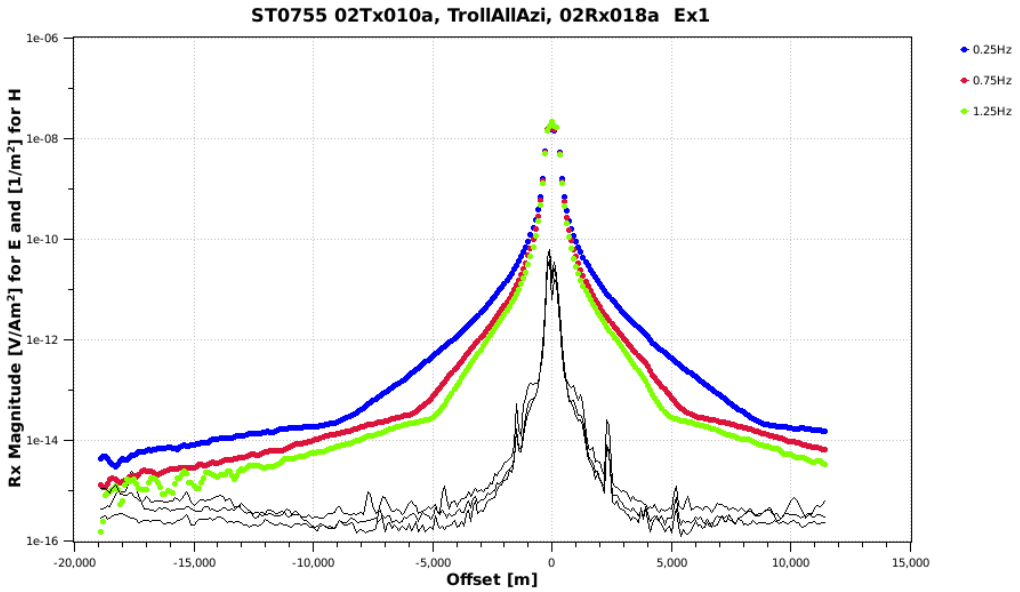


(a) Field magnitude with noise

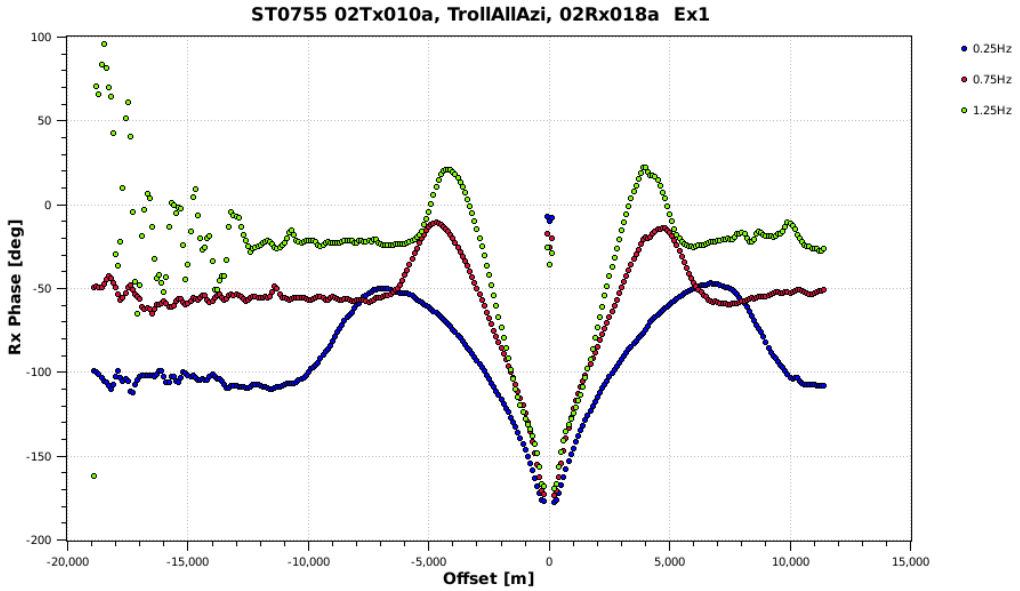


(b) Field phase

Figure 6.7: Towline: `02Tx003a`, receiver: `02Rx020a`.



(a) Field magnitude with noise



(b) Field phase

Figure 6.8: Towline: `02Tx010a`, receiver: `02Rx018a`

6.3 CONTAMINATED DATA

Assume that a noisy observed signal for one field component F, could be modeled as

$$F_j^{Obs} = F_j^{Sim}\left[1 + (w_j^{\alpha} + \imath\tilde{w}_j^{\alpha})\right] + \frac{1}{\sqrt{2}}[w_j^{\eta} + \imath w_j^{\eta}] \tag{6.3a}$$

$$:= F_j^{Sim} + \zeta_j, \tag{6.3b}$$

and the where the index j represent a configuration of the triplet $(\mathbf{x}_s, \mathbf{x}_r, f)$, and in this way $j = 1,\ldots,M$, where M is the number of observed data. The noise is standard white Gaussian noise, $w_j^{\alpha} \overset{i.i.d.}{\sim} \mathcal{N}(0,\alpha^2)$ and $w_j^{\eta} \overset{i.i.d.}{\sim} \mathcal{N}(0,\eta^2)$ for each $j = 1,\ldots,M$. Furthermore, $\tilde{w}_j^{\alpha}$ is defined such that

$$\Re\{F_j^{Sim}\}w_j^{\alpha} + \imath\Im\{F_j^{Sim}\}w_j^{\alpha} = F_j^{Sim}[w_j^{\alpha} + \imath\tilde{w}_j^{\alpha}].$$

Consider the misfit functional ε, and introduce the perturbed signal such that Equation 5.1 takes the form

$$\varepsilon = \sum_{j=1}^{M} \frac{\left(F_j^{Sim} + \zeta_j - F_j^{Sim}\right)\left(F_j^{Sim} + \zeta_j - F_j^{Sim}\right)^*}{(\delta F)^2}. \tag{6.4}$$

where $(\cdot)^*$ is the complex conjugate and δF_j is as defined in Equation 5.4b. In a simple model without the data uncertainty term, $\varepsilon \to 0$ when $F_j^{Obs} \to F_j^{Sim}$, $j = 1,\ldots,M$. However, with this term present, it can be shown that $E[\varepsilon] \to M$ when $F_j^{Obs} \to F_j^{Sim}$, $j = 1,\ldots,M$. This can be seen by first realizing that when $F_j^{Obs} = F_j^{Sim}$, $j = 1,\ldots,M$, Equation 6.4 becomes

$$\varepsilon = \sum_{j=1}^{M} \frac{\zeta_j\zeta_j^*}{(\delta F)^2}. \tag{6.5}$$

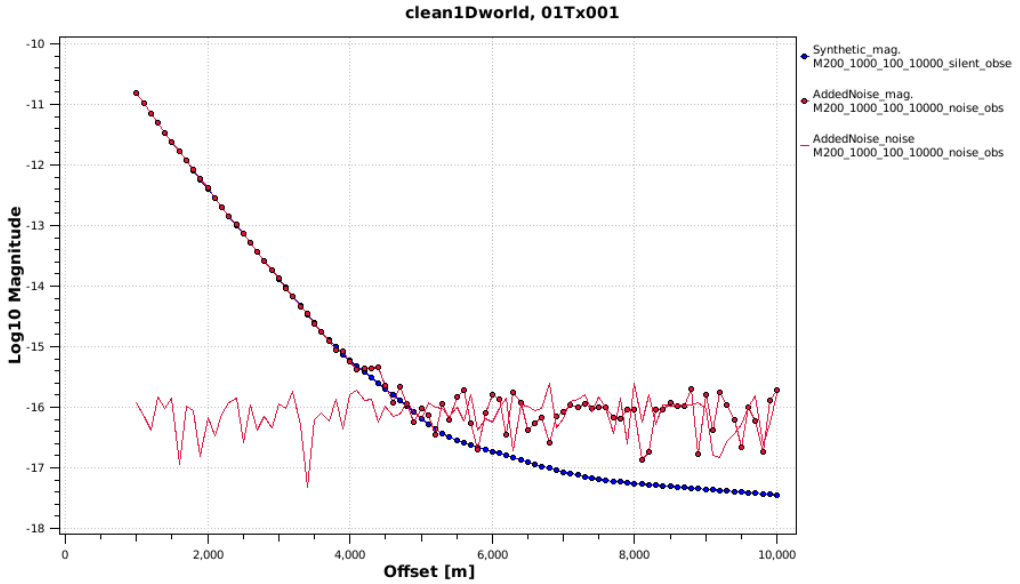
Let, for simplicity, $\psi_j := \Re\{F_j^{Sim}\}w_j^{\alpha}$ and $\phi_j := \Im\{F_j^{Sim}\}w_j^{\alpha}$ such that $Var\{\psi_j + \imath\phi_j\} = \alpha^2|F_j^{Sim}|^2$. Then, the expectation

$$\begin{aligned}
E[\varepsilon] &= \sum_{j=1}^{M} E\left[\frac{\zeta_j\zeta_j^*}{(\delta F)^2}\right] \\
&= \sum_{j=1}^{M} \frac{E\left[\psi_j^2 + \phi_j^2 + (w_j^{\eta})^2\right]}{(\delta F)^2} \\
&= \sum_{j=1}^{M} \frac{E\left[(\psi_j + \imath\phi_j)(\psi_j + \imath\phi_j)^*\right] + E\left[(w_j^{\eta})^2\right]}{(\delta F)^2} \\
&= \sum_{j=1}^{M} \frac{\alpha^2|F_j^{Sim}|^2 + \eta^2}{(\delta F)^2} \\
&= M.
\end{aligned}$$

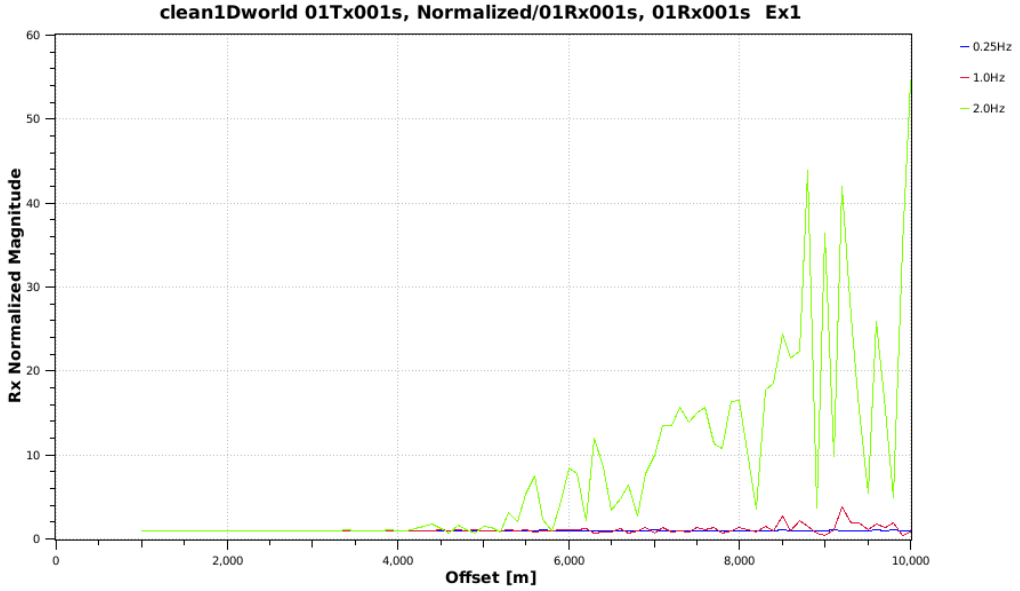
This follows since w_j^{α} and w_j^{η} is assumed independent with zero mean for every j. This means that $E[RMS(\varepsilon)] \to 1$ when $F_j^{Obs} \to F_j^{Sim}$, $j = 1,\ldots,M$. With noise present, the solution is no longer unique as there is many synthetic models which produce a $RMS(\varepsilon(\sigma))$ close to unity.

In this project, the values $\alpha = 0.02$ and $\eta = 10^{-16}$ is used, and is the same constants used when creating the weights in Equation 5.4. Figure 6.9 illustrate the effects of η noise contributor. The α contribution is easier to spot

in Figure 6.9(b), where the noisy signal is normalized on a signal without noise. Notice that higher frequencies is more effected by the contamination of the data. The reason is that the noise is more significant for weaker signals, that is, signal with a lower magnitude.



(a) Magnitude of observed field versus offset before (dotted blue) and after (dotted red) the signal has been contaminated with noise (red). Here $\alpha = 0.02$ and $\eta = 10^{-16}$, the frequency used is f =2.00 Hz



(b) Magnitude of noisy signal normalized on a signal without noise for different frequencies

Figure 6.9: Illustration of a modeled signal contaminated with noise.

6.4 IMPLEMENTATION OF THE RESISTIVITY FUNCTION

This subsection will deal with the practical implementation of the resistivity profile function, restated here for convenience

$$\begin{aligned}\rho(z) &= \rho_p(z) + \rho_r(z) + \rho_g(z) \\ &= \sum_{n_p=0}^{N_p} c_{n_p}(z - z_{wd})^{n_p} + \sum_{n_r=1}^{N_r} \rho_{n_r}[H(z - z_{n_r}) - H(z - z_{n_r} - t)] \\ &+ \sum_{n_g=1}^{N_g} \beta_{n_g} \frac{1}{\sqrt{2\pi w_{n_g}^2}} \exp\left\{ -\frac{1}{2w_{n_g}}(z - z_{n_g})^2 \right\}, z > z_{wd}\end{aligned}$$

where the parameters is restricted as

$$\begin{aligned}0 \leq N_p &\leq N_p^{\max} \\ 0 < c_0^{\min} &\leq c_0 \leq c_0^{\max} \\ c_p^{\min} &\leq c_{n_p} \leq c_p^{\max}, && 1 \leq n_p \leq N_p \\ 0 \leq N_r &\leq N_r^{\max} \\ 0 < \rho_r^{\min} &\leq \rho_{n_r} \leq \rho_r^{\max}, && 1 \leq n_r \leq N_r \\ z_1 \leq z_r^{\min} &\leq z_{n_r} \leq z_r^{\max} < z_{N_l}, && 1 \leq n_r \leq N_r \\ 0 \leq N_g &\leq N_g^{\max} \\ \beta_g^{\min} &\leq \beta_{n_g} \leq \beta_g^{\max}, && 1 \leq n_g \leq N_g \\ z_g^{\min} &\leq z_{n_g} \leq z_g^{\max}, && 1 \leq n_g \leq N_g \\ 0 < w_g^{\min} &\leq w_{n_g} \leq w_g^{\max}, && 1 \leq n_g \leq N_g \\ \rho^{\min} &\leq \rho(z;\mathbf{x}), && \forall z \in z_1, \ldots, z_{N_l}\end{aligned}$$

6.4.1 *Constraint treatment*

The linear parameter constraints are dealt with by the optimization routine and therefore discussed in Section 5.3.1, the positive resistivity function constraint is not and therefore discussed here. Since the higher order polynomial coefficients and the Gaussian amplifying coefficients can attain negative values, some extra restrictions must be made in order to keep the overall resistivity function $\rho(z;\mathbf{x})$ positive for all feasible z. This is solved by a simple rejection method by sampling a new proposed solution $\mathbf{y}_{k+1}$ from the feasible set D. Then, if this proposed function results in a negative overall resistivity at some z, then this proposed solution is rejected and a new one is sampled. This continues until a valid solution is proposed.

6.4.2 *Sensitivity analysis of the resistivity function*

This sensitivity analysis aims to cast light on the different parameters contribution to the resistivity profile function $\rho(z;\mathbf{x})$. It is included here because it reveals some issues with the parameterization which will be handled in the next section. More precisely, it reveals how much a perturbation in the different parameters $\Delta x_i, i = 1,\ldots,N$, effects the resulting resistivity function $\rho(z;\mathbf{x})$.

The perturbation in the resistivity function due to a perturbation Δx_i is labeled $\Delta_i\rho(z)$ and the following relationship is proposed

$$\Delta_i\rho(z;\mathbf{x}) = \frac{\partial\rho(z;\mathbf{x})}{\partial x_i}\Delta x_i. \tag{6.9}$$

The depth parameters z_{n_r} and z_{n_g} are not considered as they only shifts their respective terms in the z-direction which is irrelevant in this discussion.

PARTIAL DERIVATIVES
$\frac{\partial\rho}{\partial c_{n_p}} = (z - z_1)^{n_p}$
$\frac{\partial\rho}{\partial\rho_{n_r}} = H(z - z_{n_r}) - H(z - z_{n_r} + t)$
$\frac{\partial\rho}{\partial\beta_{n_g}} = \frac{1}{\sqrt{2\pi}w_{n_g}}\exp\left\{-\frac{1}{2w_{n_g}^2}(z - z_{n_g})^2\right\}$
$\frac{\partial\rho}{\partial w_{n_g}} = \frac{\beta_{n_g}}{\sqrt{2\pi}}\frac{(z-z_{n_g})^2 - w_{n_g}^2}{w_{n_g}^4}\exp\left\{-\frac{1}{2w_{n_g}^2}(z - z_{n_g})^2\right\}$

Table 6.3: Some partial derivatives for the resistivity profile function (Equation 3.5), and their respective maximum values w.r.t. z.

This illustrates the impact a small change in the polynomial coefficients will have on the overall resistivity function. One unit perturbation of c_{n_g} results in a maximum perturbation

$$\Delta_{n_p}\rho(z;\mathbf{x})^{\max} = (z_{N_l} - z_1)^{n_g}, \quad \text{at } z = z_{N_l}$$

of the overall resistivity function. For the resistor term, the resistivity value is only inflicted at the interval $z \in (z_{n_p}, z_{n_p} + t)$, and there a unit change in ρ_{n_g} results in a unit change in the resistivity function. For the Gaussian term, one unit change in β_{N_g} results in a maximum change in the resistivity function of

$$\Delta_{n_p}\rho(z;\mathbf{x})^{\max} = \frac{1}{\sqrt{2\pi}w_{n_g}}, \quad \text{at } z = z_{n_g}.$$

Likewise, an unit change in w_{n_g} results in a maximum change of the resistivity function of

$$\Delta_{n_p}\rho(z;\mathbf{x})^{\max} = \frac{2\beta_{n_g}}{\sqrt{2\pi}w_{n_g}^2}e^{-3/2}, \quad \text{at } z = z_{n_g} + \sqrt{3}w_{n_g}$$

From this, it is quite obvious that the perturbation of the polynomial coefficients affects the overall resistivity function the most. Issues concerning this will be discussed in Section 6.4.3.

6.4.3 *Polynomial term*

From the sensitivity analysis in Section 6.4.2, there is no doubt that the resistivity profile function is most sensitive to changes in the polynomial coefficients. The problem is that, because of the scaling and sampling choices, it is difficult to optimize these coefficients as they appear in Equation 3.5. Because of the scaling of the parameters and the way that both the direction and the step length is chosen, it is very difficult for the parameter to climb from a state with a low value. This will be shown with an example. Say the value of this parameter at iteration k is $c_1 = 10^{-5}$, then there is a fifty-fifty chance for the algorithm to propose a smaller value. So say it choose a positive direction. The probability that the step taken is to large is quite dominating. Say $l_i = 0$ and $u_i = 1$ (although it seems like a narrow interval, remember that $c_1 = 1$ would contribute with 1000 Ωm at a depth of 1000 m below the sea bed). From the reasons above it is quite unlikely that a value of $c_1^{k+1} \geq 0.01$ would result in an accepted step. So let us calculate the probability for this happening, given that the direction is towards a greater value. Let also the true value be 10^{-3}, which is larger than 10^{-5}. Then, from the expressions in Section 5.3.2

$$\Pr(10^{-5} + \lambda_1\theta_1 \geq 10^{-2} | \theta_i = 1) = 1 - \frac{10^{-2} - 10^{-5}}{1} \approx 0.99$$

given that $m_1^u = 1$ and $\lambda_1^{\min} = 0$. The consequence is that if the coefficient value is small, very few steps in the correct direction is accepted whilst steps in the wrong direction would likely be accepted (as a change from say, 10^{-5} to 10^{-6} would not contribute significantly to the overall resistivity function). This would then result in a slow convergence.

This issue discussed here is perhaps quite obvious and could be solved in many different ways. A logarithmic step length sampler is implemented, which means that instead of sampling λ_i uniformly from some interval, one let $\lambda_i = 10^{\xi_i}$ and then samples ξ_i from some interval. This solved the issue, but a more intuitive approach is to normalize the polynomial coefficients and then keep the original sampling method. Instead of optimizing the coefficients c_{n_p} directly, let

$$c_{n_p} = \frac{\zeta_{n_p}}{(z_{N_l} - z_1)^{n_p}}, \quad 0 \leq n_p \leq N_p$$

and optimize for the unit free ζ_{n_p}. In practice, this alters the polynomial term from Equation 3.2 to

$$\rho_p(z; \mathbf{x}_p) = \sum_{n_p=0}^{N_p} \zeta_{n_p} \frac{(z - z_{wd})^{n_p}}{(z_{N_l} - z_1)^{n_p}}, \quad N_p \geq 0, \tag{6.10}$$

but no change in the optimization routine is required.

6.5 OPTIMIZATION DECISIONS

Every aspect discussed in Chapter 5 is implemented in the programming code created to carry out the tasks in this project. There are a huge number of different possibilities, and much time has gone into finding the right values for the different hyperparameters, and to figure out which choices yields the best results. The performance of the algorithm depends on these decisions, and the values presented below are used to generate the results in Chapter 7. It is possible that there are better choices, and therefore every made decision is explained thoroughly.

6.5.1 *Simulated Annealing*

Here, parameters specific to the Simulated Annealing method is discussed, and the reader is referred to the theory presented in Section 5.3.

Every result is generated using a single parameter updating scheme when a new state is proposed. When compared to inversions which used a full parameter sampling approach, the latter tended to converge slower. This is because fever proposed solutions was accepted than when just one parameter was updated at a time. Since step lengths are treated independently for each parameter, a single parameter updating approach also provides an explicit connection with the overall search direction and step length. It could be argued that this approach would require more iterations than a full sampling approach. As it would require at least N iterations to make the same step as is done in one iteration by the full sampling method, this intuitively makes sense. However, as both approaches was extensively tested, the single sample approach tended to propose new states which was more likely accepted than the proposed new full sample states. This would guide the search slowly but surely nearer the optimal values.

As with the other choices explained in this section, the reason for this is explained by the solution space topology in this optimization problem. The problem is ill-posed, and the minimal misfit values are non-unique. Using physical space analogous terms, the topography near these minima is vast and flat surrounded by large walls with high misfit values. The optimization search tends to quickly settle in this flat valley, and therefore searches that explore this bottom of the valley most efficiently is preferable.

When it comes to the temperature schedule, there are a huge freedom of choice. The initial and terminal temperature, which schedule to choose, and which values should be set for the different tuning parameters are all factors that will affect the performance of the algorithm. The temperature schedule in Equation 5.14 is chosen, and this, together with the geometric temperature schedule in Equation 5.13 is shown in Figure 6.10. The

initial temperature is set to $T_0 = 10000$ and this is chosen such that an initial misfit RMS increase of 1 is accepted with a probability of about 0.1. Based on experience, the initial misfit RMS is about 40 which then from Equation 5.16 yields an initial temperature of about 10000. The terminal temperature is chosen from Equation 5.17 such that the probability of selecting a state which result in an increase in the RMS value of 0.01 is about 0.005 after 20000 iterations. From Equation 5.18, $\nu_l = 5 \cdot 10^{-5}$, and this is the temperature schedule shown in Figure 6.10 together with the geometric temperature schedule which tuning constant is calculated accordingly. The epoch constant introduced Section 5.3.3 is set to 15 which means that the temperature is updated for every $15N$th iteration.

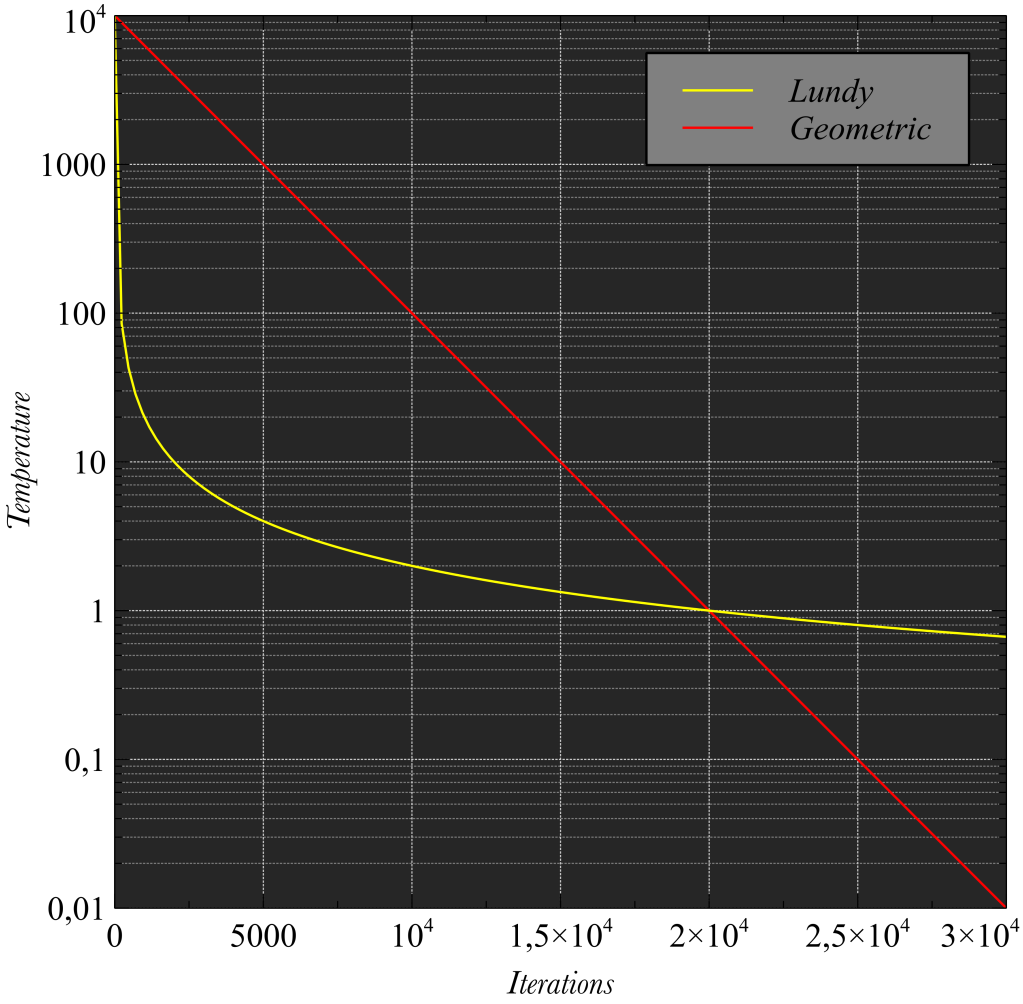


Figure 6.10: The different temperature schedules from Equation 5.14 (lundy) and Equation 5.13 (geometric) with initial temperature $T_0 = 10000$ and tuning constants $\nu_l = e \cdot 10^{-5}$ and $\nu_g = 0.9995$ respectively.

Other schedules and constant values has been tried out in an extensive degree before settling this choice. It seems like the rapid initial decrease in temperature is beneficial in this problem as it quickly guides the search

to the flat bottom of the solution space. As can be seen in Figure 5.1 and in the other trace plots presented in Chapter 7, the search quickly reaches an RMS value of about 5, and then, further descent of the misfit is slow, but steady. The slow decrease in temperature, combined with its low value at this stage is the reason for this slow, steady decrease in the misfit function value. As can be seen from Figure 6.10, the geometric temperature has much higher temperature values than the Lundy schedule in a large iteration interval. This makes the algorithm accept greater leaps with increasing misfit value than the Lundy function. Also, after the temperatures intersect each other, the temperature is decreasing much faster to zero than the Lundy schedule. The consequence of this is that after said number of iterations, only a few more temperature updates is needed before the acception of even a very small increase in the misfit function is practically impossible. In effect, it is not long before the search stops completely. The Lundy schedule however, will accept small jumps with increasing value for a great number of iterations, such that even if the step length is small, a greater number of steps is accepted which keeps the search alive. The great number of iterations at these small, but almost constant temperature values, makes the decrease slow, but steady towards a minimum. This discussion has to be seen in light of the solution space topology in this problem. What makes the Lundy temperature favourable is that it guides the search quickly down in this flat valley from the analogy above, and allows the search to traverse this valley with small steps. With this, the flat valley is explored extensively while the rest of the solution space consisting of uninteresting solutions is poorly explored, but in this problem this is a favourable. Other problems may require a greater exploration of the whole solution space, making a schedule with a less rapid initial decrease more attractive.

Maximal and minimal step length reduction is imposed such that the step length is reduced bey a factor of 1/2 every time the inequality in Equation 5.19 is true for a value of $\xi = 10^{-3}$. This inequality is checked every $15N$th iterations, where the value 15 stems from the epoch constant.

When the step length is reduced more than 5 times, the search terminates. This combination of hyperparameter values creates a good balance of whether to terminate or not, and fulfills its purpose adequately. It terminates searches which are trapped in local minima quite early, and also terminates searches which has found sufficient minimas. Searches that continues to sample descending misfit function values are allowed to continue.

A maximal number of iterations is set to 20000 iterations as this seems to be enough iterations to find satisfactory solutions. Also, searches that need more iterations is not that interesting as it defeats the purpose of this thesis which is to maintain a low number of forward calls. That being said, some runs have been allowed to continue for 50000 iterations, and some of these are also presented in Chapter 7.

The search is terminated when a misfit RMS value below 1 found. This is chosen in light of what minimum misfit values one could expect to find. This discussion is only relevant for the data generated from the real models discussed in defined in Table 6.2. Although the real data obtained from the Troll survey is 3D-data, it is suited for 1D inversion. It is not ideal of course, but the lithostratigraphy is quite similar in the field area. That is, the 1D assumption of transverse isotropy in the vertical direction (see Section 4.1.1) is not that erroneous as one could expect, which makes the Troll data sufficiently suited for 1D inversion. However, one could expect to see some 3D-effects that 1D inversion cannot describe in a satisfactory way. This will in general add some to the expected minimum RMS value, how much added is not known and will probably vary between each of the different data sets. Also, one could suspect that the when 3D data is inverted in a 1D model, the algorithm will try to fit data which explains 3D phenomena which does not exist in a 1D model, making it futile. Also, since the data is noisy, it is explained in Section 6.3 that one could expect a misfit RMS value of 1 when the observed field matches the simulated field.

For the algorithm with dynamic parameter dimensions, the probability of keeping the dimensionality in an iteration is set to $p_0 = 0.7$, while $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = 0.05$ where notation follows that of algorithm 3. Furthermore, when adding terms, the new parameter values is sampled uniformly from their respective feasible intervals.

6.6 INITIAL PARAMETER VALUES AND PARAMETER CONSTRAINTS

It is important that the parameterization and the method is general. By general it is meant that it could be applied on a wide range or problems, and that it would generate useful solutions with little or no prior information about the problem. Therefore, unless otherwise stated, every result presented in Chapter 7 is obtained from a flat resistivity profile with a constant resistivity (e.g. the the measured sea-water resistivity if this is available). With this, no simplifying assumptions is made, and although the well bore log give an indication of favourable initial solution values, this is not exploited for the sake generality. In order to obtain these initial conditions, $\zeta_0 = \rho_w$, $\zeta_{n_p} = 0.0, n_p \in \{1,\dots,N_p\}$, $\rho^r_{n_r} = 0.0, n_r \in \{1,\dots,N_r\}$ and $\beta_{n_g} = 0.0, n_g \in \{1,\dots,N_g\}$. With this set up, the values of z_{n_r}, z_{n_g} and w_{n_g} does not contribute to the resistivity profile and is yet to be determined. Even though their initial values could not obstruct the flat initial resistivity profile, some initial values could still be more beneficial than others.

Therefore, to maintain generality, their initial values is uniformly sampled from their respective feasible domains. This is summarized as

$$\zeta_0 = \rho_w \tag{6.11a}$$
$$\zeta_{n_p} = 0, \qquad n_p = 1,\dots,N_p \tag{6.11b}$$
$$\rho^r_{n_r} = 0, \qquad n_r = 1,\dots,N_r \tag{6.11c}$$
$$z^r_{n_r} \sim \mathcal{U}(z^{min}_r, z^{max}_r) \qquad n_r = 1,\dots,N_r \tag{6.11d}$$
$$\beta_{n_g} = 0, \qquad n_g = 1,\dots,N_g \tag{6.11e}$$
$$z^g_{n_g} \sim \mathcal{U}(z^{min}_g, z^{max}_g) \qquad n_g = 1,\dots,N_g \tag{6.11f}$$
$$w_{n_g} \sim \mathcal{U}(w^{min}_g, w^{max}_g) \qquad n_g = 1,\dots,N_g. \tag{6.11g}$$

For the solutions with a fixed number of

The set of constraints defined in Equation 3.6 are set as shown in Equation 6.12. Constraints that are freely chosen, are chosen so large or small that they are rarely touched by the search trajectory. With this, the only real influence these constraints have on the algorithm is that they defines the intervals at which new parameter values are sampled from. The constraint intervals should therefore not be to narrow, as this could guide the search, which is deprecated. But not unrealistically wide either, as this would increase the probability of sampling parameter values which are not physically consequential.

$$N^{\min}_p = 0, \qquad N^{\max}_p = 10 \tag{6.12a}$$
$$\zeta^{\min}_0 = 0, \qquad \zeta^{\max}_0 = 20 \tag{6.12b}$$
$$\zeta^{\min}_p = -20, \qquad \zeta^{\max}_p = 20 \tag{6.12c}$$
$$N^{\min}_r = 0, \qquad \zeta^{\max}_p = 20 \tag{6.12d}$$
$$\rho^{\min}_r = 0, \qquad N^{\max}_r = 10 \tag{6.12e}$$
$$z^{\min}_r = z_{wd}, \qquad \rho^{\max}_r = 500\,\Omega\mathrm{m} \tag{6.12f}$$
$$N^{\min}_g = 0, \qquad z^{\max}_r = z_{N_l} - z_{wd} \tag{6.12g}$$
$$\beta^{\min}_g = -10000\,\Omega\,\mathrm{m}^2, \qquad \beta^{\max}_g = 10000\,\Omega\,\mathrm{m}^2 \tag{6.12h}$$
$$z^{\min}_g = z_{wd}, \qquad z^{\max}_g = z_{N_l} - z_{wd} \tag{6.12i}$$
$$w^{\min}_g = 1\ \mathrm{m}, \qquad w^{\max}_g = 1000\ \mathrm{m} \tag{6.12j}$$
$$\rho^{\min} = \rho_w, \tag{6.12k}$$

The sampling of step lengths which is shown in algorithm 2 are dependent on minimal and maximal step length values which are given in Equation 6.13. These are the values which are reduced if the step length reduction criteria given in Equation 5.19 is imposed. The maximal step length boundary imposes naturally a restriction on the immediate exploration of

the solution space. The values in Equation 6.13 is set such that the search adequately explore the solution space in the beginning of the search. As the initial parameter values are randomly chosen, it is important that these maximal step length boundaries are large enough as the temperature schedule is quickly decreasing in the beginning of the search. When the search has approached this flat valley which is mentioned above, one would like to concentrate the search. These constraint values are a bit too large for this purpose and cause slower convergence since too large steps are proposed just to be rejected. It is exactly for this purpose the step length reduction is implemented.

$$\lambda(\zeta_0)^{\min} = 10^{-3}, \qquad \lambda(\zeta_0)^{\max} = 10, \tag{6.13a}$$

$$\lambda(\zeta_{n_p})^{\min} = 10^{-3}, \qquad \lambda(\zeta_{n_p})^{\max} = 10, \qquad n_p = 1,\ldots,N_p \tag{6.13b}$$

$$\lambda(\rho^r_{n_r})^{\min} = 10^{-2}\,\Omega\mathrm{m}, \qquad \lambda(\rho^r_{n_r})^{\max} = 100\,\Omega\mathrm{m}, \qquad n_r = 1,\ldots,N_r \tag{6.13c}$$

$$\lambda(z^r_{n_r})^{\min} = 10^{-2}\,\mathrm{m}, \qquad \lambda(z^r_{n_r})^{\max} = 100\,\mathrm{m}, \qquad n_r = 1,\ldots,N_r \tag{6.13d}$$

$$\lambda(\beta_{n_g})^{\min} = 10^{-2}\Omega\,\mathrm{m}^2, \qquad \lambda(\beta_{n_g})^{\max} = 1000\Omega\,\mathrm{m}^2, \qquad n_g = 1,\ldots,N_g \tag{6.13e}$$

$$\lambda(z^g_{n_g})^{\min} = 10^{-2}\,\mathrm{m}, \qquad \lambda(z^g_{n_g})^{\max} = 100\,\mathrm{m}, \qquad n_g = 1,\ldots,N_g \tag{6.13f}$$

$$\lambda(w_{n_g})^{\min} = 10^{-2}\,\mathrm{m}, \qquad \lambda(w_{n_g})^{\max} = 100\,\mathrm{m}, \qquad n_g = 1,\ldots,N_g. \tag{6.13g}$$

The computing code is written in `C++`, and executed on *Linux* with a 2.90 GHz *Intel® Xeon® E5-2690* CPU.

7

RESULTS

Each of the different optimization routines was tested thoroughly on the different data sets presented in Section 6.2. In order to get an impression of the new parameterization presented in this thesis, several evaluation criteria was imposed.

Since Simulated Annealing is a stochastic optimization routine, multiple inversions should be performed with the same initial values and different seeds for the pseudo-random number generator implemented in the algorithms. With this, eventual bias towards different seeds will be reduced. Also, there should be performed multiple inversions with different, random initial values, in order to reduce the bias towards certain initial values.

The above discussions is as much about optimization as it is about the particular parameterization presented in this thesis. Since this is a new parameterization, the results should also exhibit how well the parameterization can represent the model responsible for the observed data. The parameterization should be independent of the optimization routine, and it is not impossible to imagine that a better optimization routine could be chosen. Because of this, also the best inversions, that is, the inversions with the lowest misfit values, are also presented. These are often obtained with more ideal initial parameter values, and is not that interesting as a measure of the performance of the method per se. But as a measure of the parameterization, this is interesting, and it could also be used as a reference for the performance in the general case. In the synthetic case, this is not interesting, as the synthetic model is simulated with this parameterization, but with the real data, the location and the minimum misfit value is not known.

By ideal initial parameter values, it is meant that they are somewhat close to the alleged optimal solution. This optimal solution is not known, but from the well bore log, one could expect one resistor at about 1500 m depth at the reservoir. One could also expect a certain compression effect, and some more resistive layers at shallow deep. From these expectations, one could adjust the parameters so that they resembles this, and run the inversion with e.g. smaller step lengths to concentrate the search around this alleged global minimum. Note that, even with these assumptions, attempts to rig the parameters to ideal values has not been able to produced lower misfit values than the inversion provides.

For each data set, multiple inversions was carried out and the results of these inversions is presented in this chapter. The methods concerned is the regular Simulated Annealing algorithm with a fixed number of parameters

$N_p = 1$, $N_r = 1$ and $N_g = 2$, and the trans-dimensional Simulated Annealing algoritm presented in Section 5.4.1. Results from inversions done by the other trans-dimensional algorithm, presented in Section 5.4.2 is only included for the Troll 1 model. The reason for this is that it is only a reference method in this thesis, and that it is not ideal for solving this kind of problem. Therefore, more emphasis is made on the two other methods.

The set up in every inversion is as explained in Chapter 6 unless otherwise stated. Statistics for inversions with a terminal misfit RMS less than 5.0 is collected in tables for the respective data sets. The ratio of inversions that managed this, and thus is used to create the statistics is showed in the header of the respective tables. In addition, excerpts of the inversions are illustrated explicitly with a traceplot of the misfit RMS versus iterations, and the terminal resistivity profile which is plotted against well bore log data (Figure 6.5) as a reference, where the indicated vertical resistivity is included. This, together with the transverse resistance, which was calculated to be 35640 Ωm, will be used as quality indications for the results.

7.1 INVERTED CSEM DATA FROM A SYNTHETIC MODEL

This is the results of the inversion of the CSEM data gathered from the synthetic model presented in Section 6.1.2. As the data is not contaminated with noise, one unique global minima with misfit 0 exists. The different methods did not manage to attain perfect fit, and also the inversions was terminated when reaching a misfit RMS of less than 1.0, a termination criterion that should have been disabled in this case.

As can be seen in Table 7.1, the inversions with a fixed number of parameters perform better than the ones from the dynamic dimension algorithm. As the fixed parameters is tailored to be able to fit such a resistivity profile, this is perhaps not that surprising.

	FIXED DIMENSIONS		DYNAMIC DIMENSIONS	
Accepted	9/10		9/10	
	Mean	*Std. dev.*	*Mean*	*Std. dev.*
Iterations	16866.89	3547.09	18870.66	1756.60
Elapsed time[s]	101414.57	14576.86	131657.78	10335.56
ε	1162.56	1680.33	1092.54	301.02
$RMS(\varepsilon)$	1.78	1.10	1.98	0.28
Accept ratio	0.25	0.09	0.25	0.06
N_p	1.00	0.00	5.22	3.15
N_r	1.00	0.00	1.22	0.44
N_g	2.00	0.00	6.00	2.50

Table 7.1: Statistics from inversions of CSEM-data from a synthetic model.

7.1.1 *Fixed number of parameters*

As can be seen from Figure 7.1 and Figure 7.1, there algorithm manage to recover the modeled resistivity to a great extent, even though they where terminated when reaching a misfit RMS of 1.0.

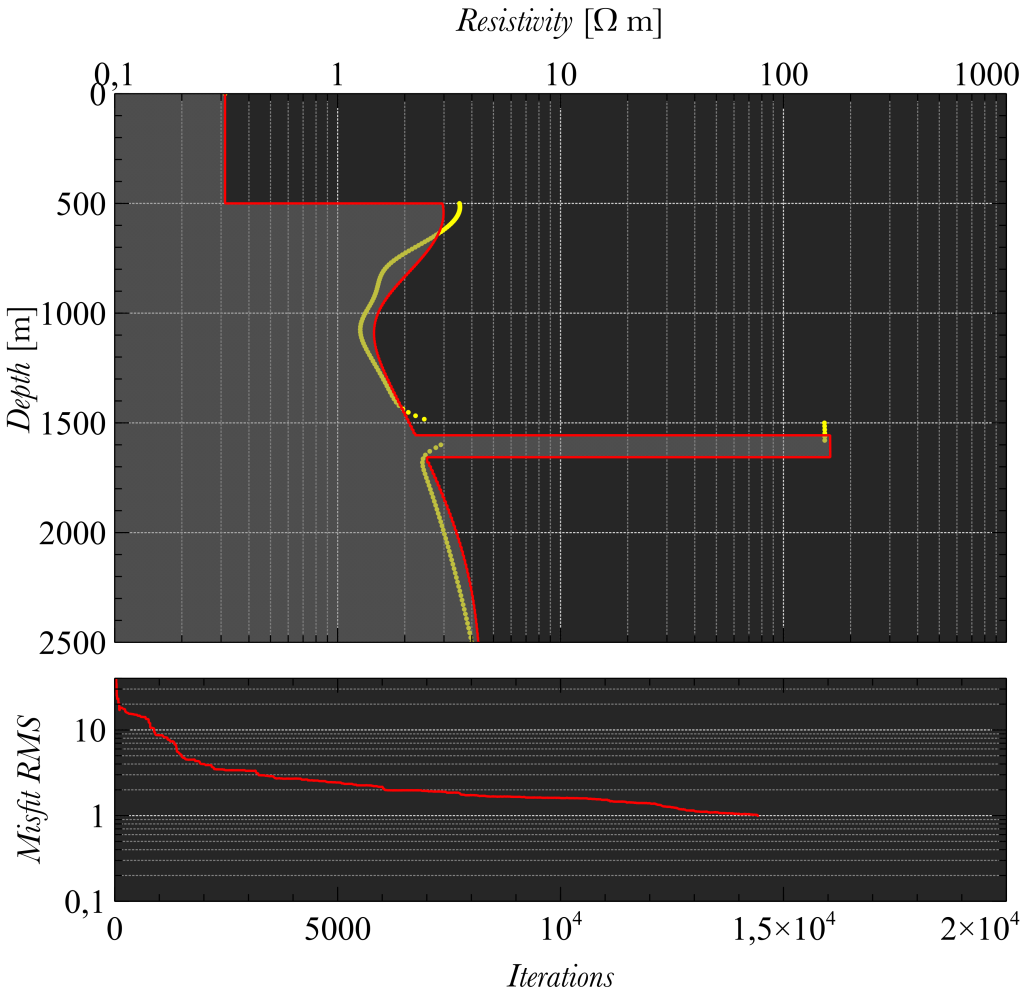


Figure 7.1: Inversion of synthetic data with fixed number of parameters. Final misfit RMS = 1.00.

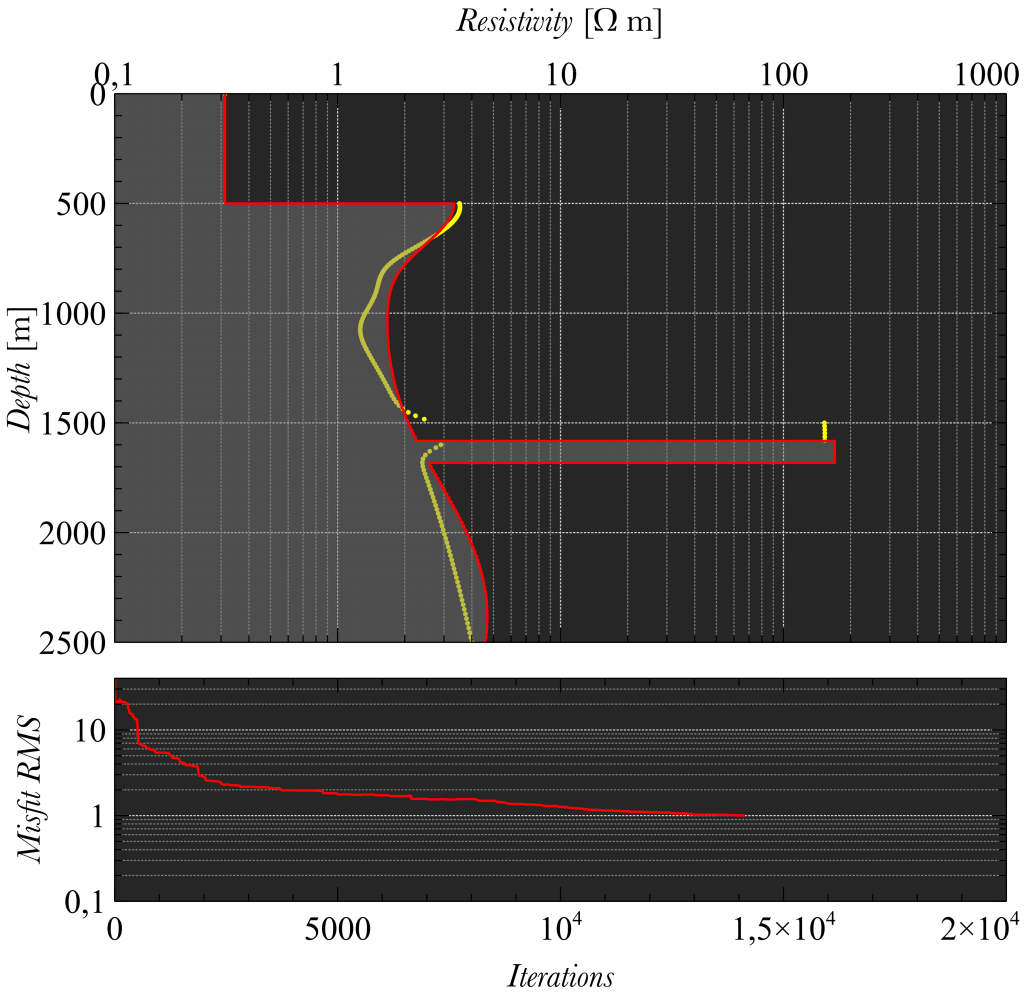


Figure 7.2: Inversion of synthetic data with fixed number of parameters. Final misfit RMS = 1.00.

The inversion which produced the profile in Figure 7.3 was obtained with the misfit termination criterion turned off, and the maximum iteration termination criterion was risen to a maximum of 50000 iterations.

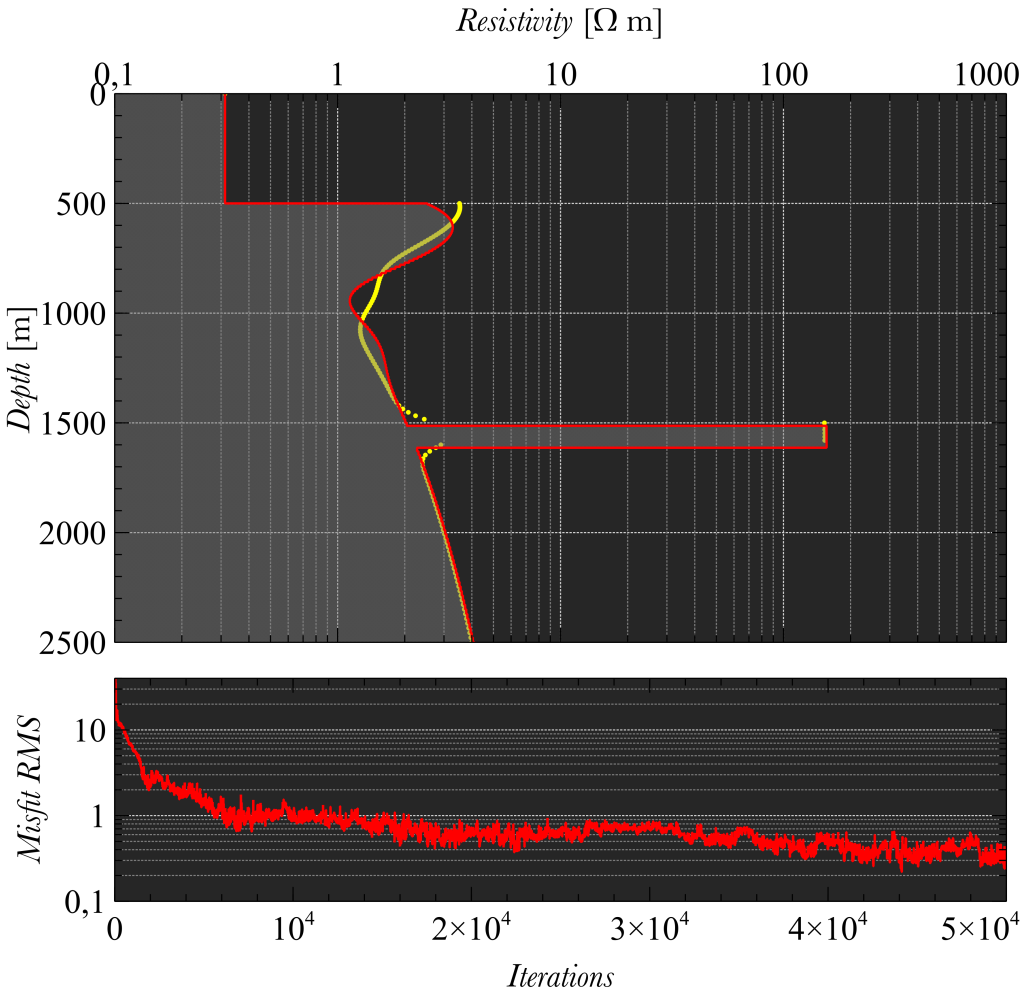


Figure 7.3: Inversion of synthetic data with fixed number of parameters. Final misfit RMS = 0.302.

7.1.2 *Dynamic number of parameters*

The inversion which result is pictured in Figure 7.4 did manage to recover the location and resistivity of the hydrocarbon resistor to a great extent. The background is not perfectly recovered, but it seems to capture the essentials, also note the logarithmic axis which exaggerate the difference at small resistivity scales compared to the larger scale.

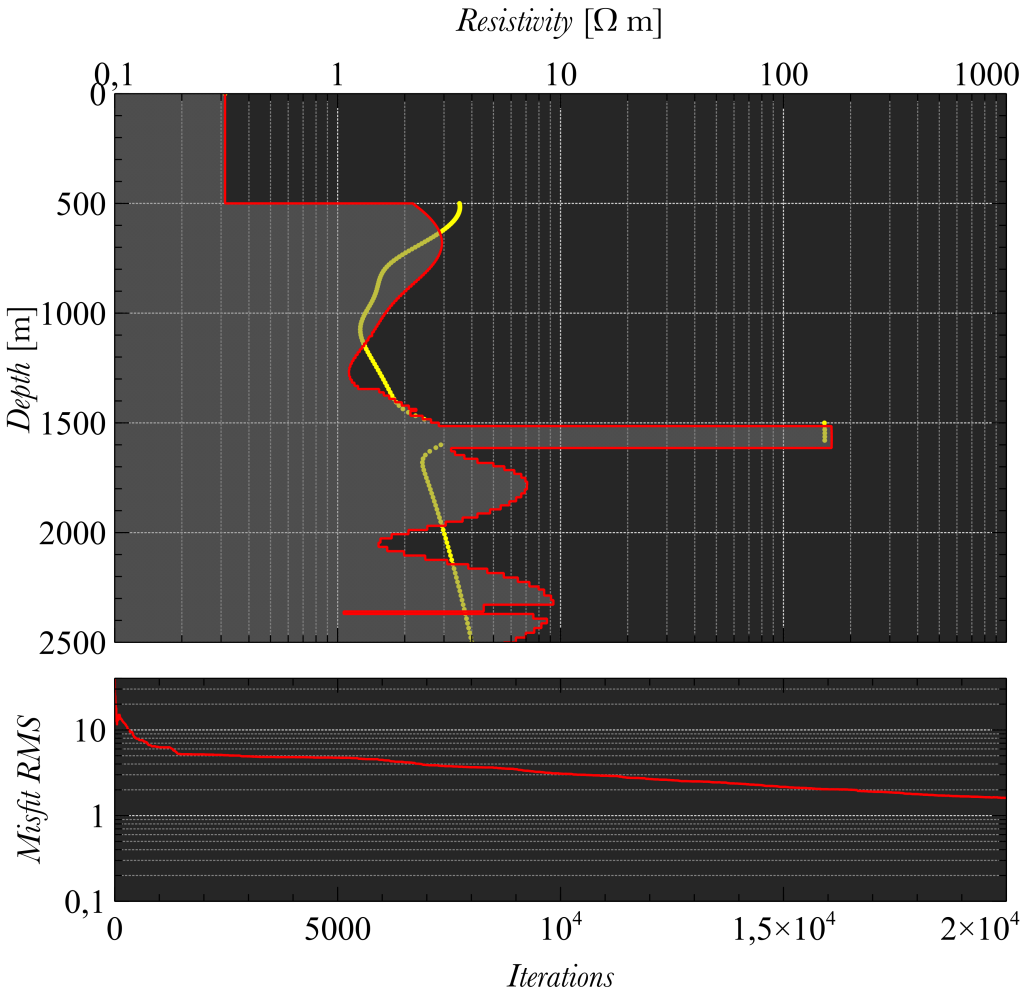


Figure 7.4: Misfit RMS traceplot after inversion of synthetic data with fixed number of parameters. Final misfit RMS = 1.610.

7.2 INVERTED CSEM DATA FROM THE TROLL 1 MODEL

This section presents inversions from the Troll 1 data set defined in Table 6.2 and Section 6.2. Inversions done by the reference trans-dimensional Simulated Annealing method from (Singh et al., 2008), presented in Section 5.4.2 is also included, in addition to inversions from the fixed parameter dimension algorithm and the dynamic parameter dimension algorithm. Statistics from the inversions by this method is gathered in the columns under the heading with marked with an asterisks at the right in Table 7.2. As is explained in Section 5.4.2, this method is not ideal for this problem, so in order to obtain the presentable results, the maximum parameter dimensions was limited compared to the trans-dimensional algorithm presented in Section 5.4.1. These are set to $N_p^{\max} = 2$, $N_r^{\max} = 3$ and $N_g^{\max} = 3$. As a comparison, the maximum parameter dimensions used in the method from Section 5.4.1 is $N_p^{\max} = 10$, $N_r^{\max} = 10$ and $N_g^{\max} = 10$.

The maximum number of iteration is increased to 50000 iterations for 2 of the inversions with fixed parameter dimensions, and for 25 of the inversions carried out by the trans-dimensional algorithm from (Singh et al., 2008). The rest is limited to 20000 iterations.

As expected, the trans-dimensional algorithm performs better on average than both the one with fixed dimensions and the one from (Singh et al., 2008).

	FIXED		DYNAMIC		DYNAMIC*	
Accepted	22/22		20/20		33/39	
	Mean	*Std. dev.*	*Mean*	*Std. dev.*	*Mean*	*Std. dev.*
Iterations	21206.91	9753.00	19254.85	1951.70	26529.17	14195.88
El. time [s]	99285.76	41739.87	91741.44	8774.91	110041.21	40507.49
ε	1423.64	547.01	863.21	447.00	1911.98	704.30
$RMS(\varepsilon)$	3.09	0.58	2.39	0.52	3.59	0.66
Accept ratio	0.17	0.11	0.31	0.13	0.19	0.16
N_p	1.00	0.00	5.80	3.32	1.03	0.77
N_r	1.00	0.00	2.45	1.67	2.24	1.17
N_g	2.00	0.00	5.53	2.29	2.32	1.46

Table 7.2: Statistics from inversions of CSEM-data from the Troll 1 model.

7.2.1 *Fixed number of parameters*

The results in Figure 7.5 and Figure 7.6 show that there seems to be an agreement about the recovered profile. A resistive formation that is a couple of hundred meters wide with resistivity of about $5\,\Omega\,\mathrm{m}$ centered around a depth of about 600 m. This is followed by a decrease in resistivity to about $0.5\,\Omega\,\mathrm{m}$ at a depth of about 800 m. The resistor is found at about 1500 m depth with a transverse resistance of about $40000\,\Omega\,\mathrm{m}^2$. This is in agreement with the well log and the measured transverse resistance of the resistor. Also, the inversion resulting in Figure 7.7, which was carried out with rigged initial values shows agreement.

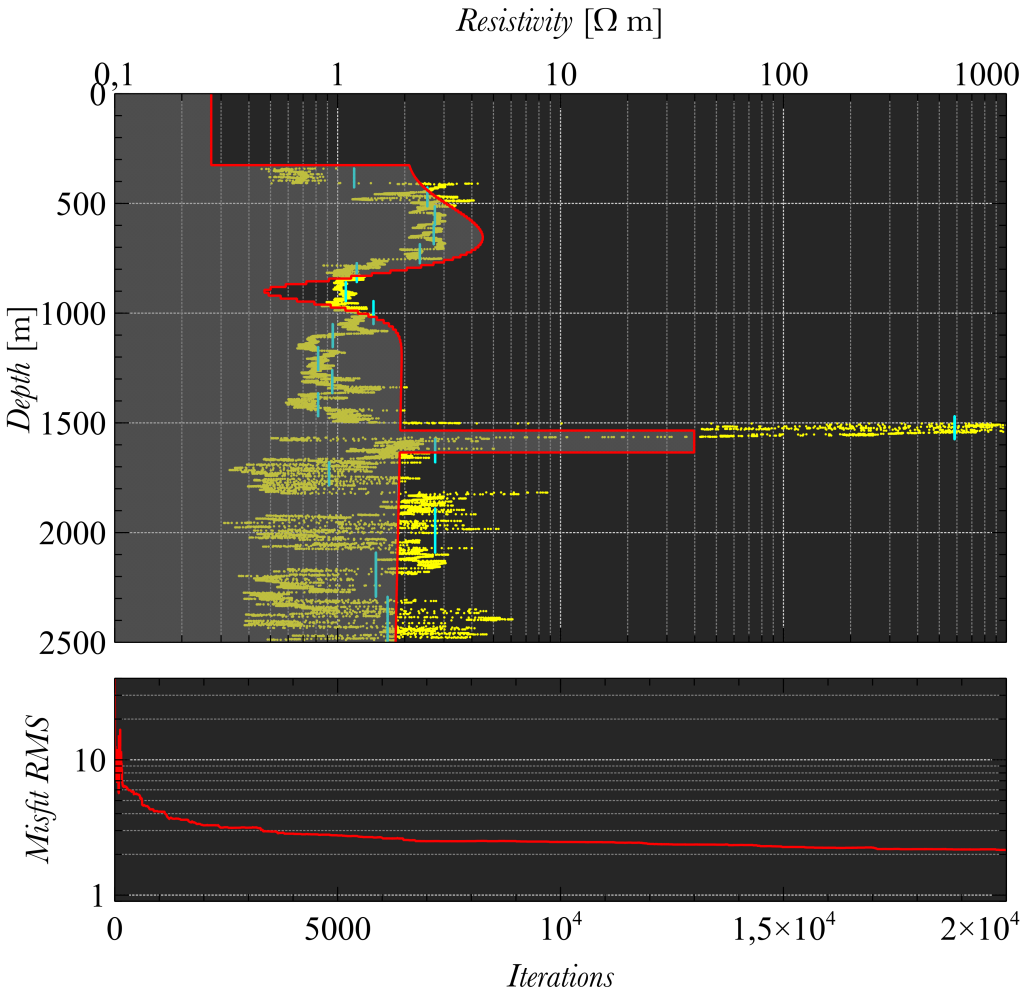


Figure 7.5: Resistivity profile after inversion of troll 1 data with fixed number of parameters. Final misfit RMS = 2.157.

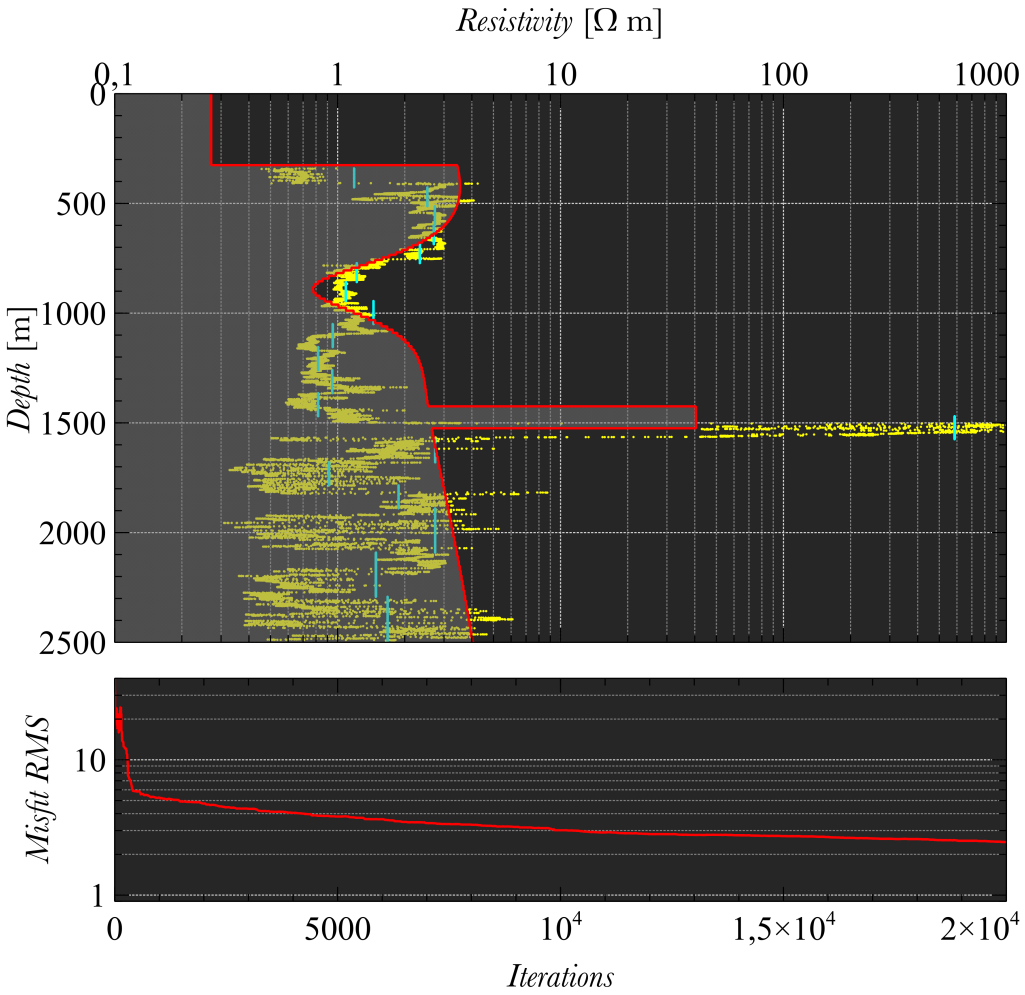


Figure 7.6: Resistivity profile after inversion of troll 1 data with fixed number of parameters. Final misfit RMS = 2.467.

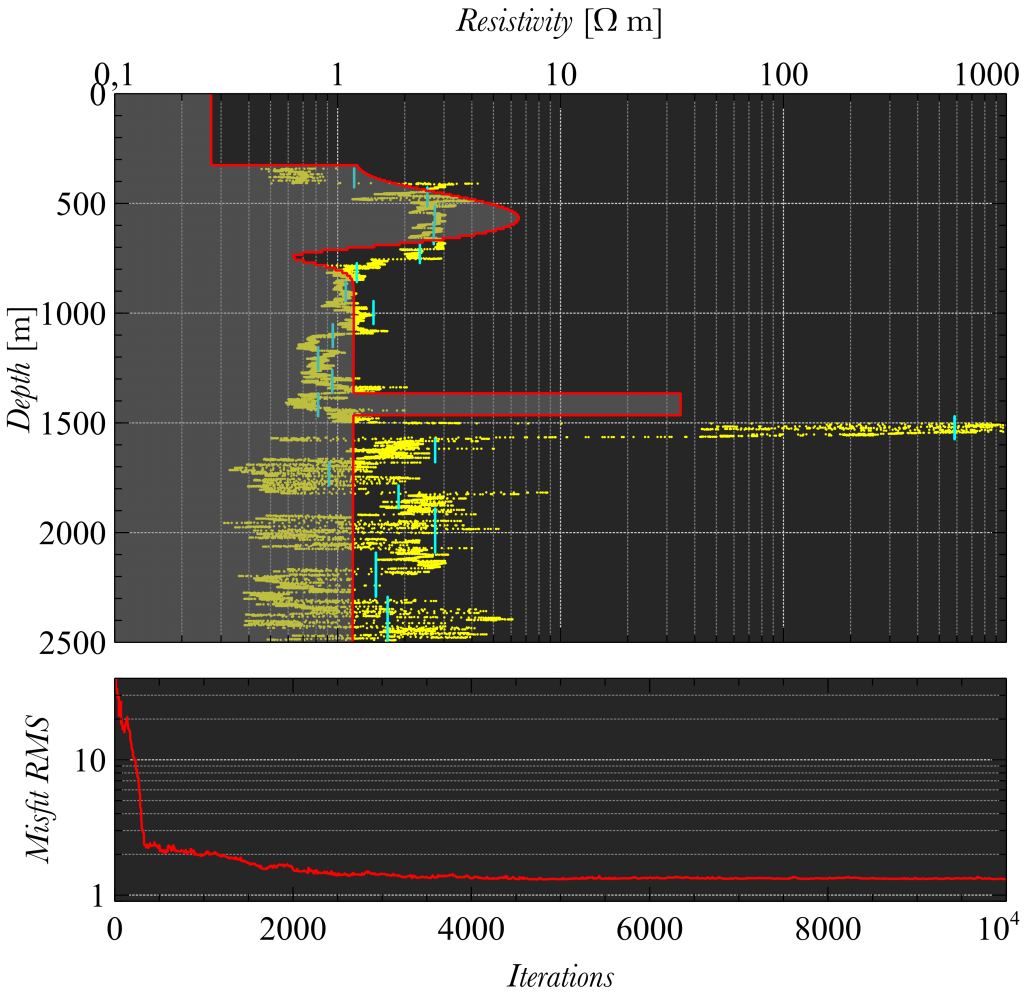


Figure 7.7: Resistivity profile after inversion of troll 1 data with fixed number of parameters. Final misfit RMS = 1.313.

7.2.2 *Dynamic dimensions I*

The results shown in Figure 7.8 and Figure 7.9 is in agreement with the results in Section 7.2.1 from inversion with a fixed number of parameters. They place the resistor a little deeper, and also manages to recover a conductive layer at about 2200 m which is in agreement with the well log data. This was not captured with the fixed dimension method.

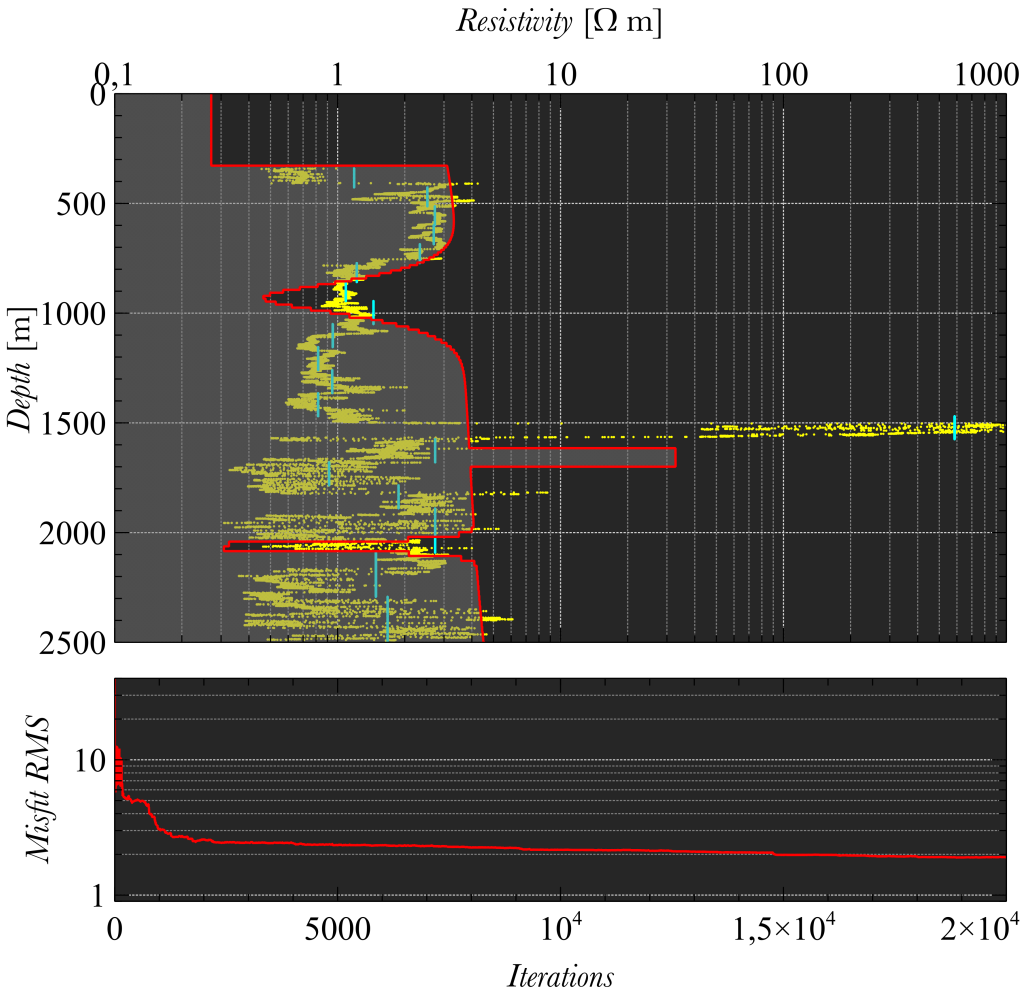


Figure 7.8: Resistivity profile after inversion of troll 1 data with dynamic number of parameters. Final misfit RMS = 1.908.

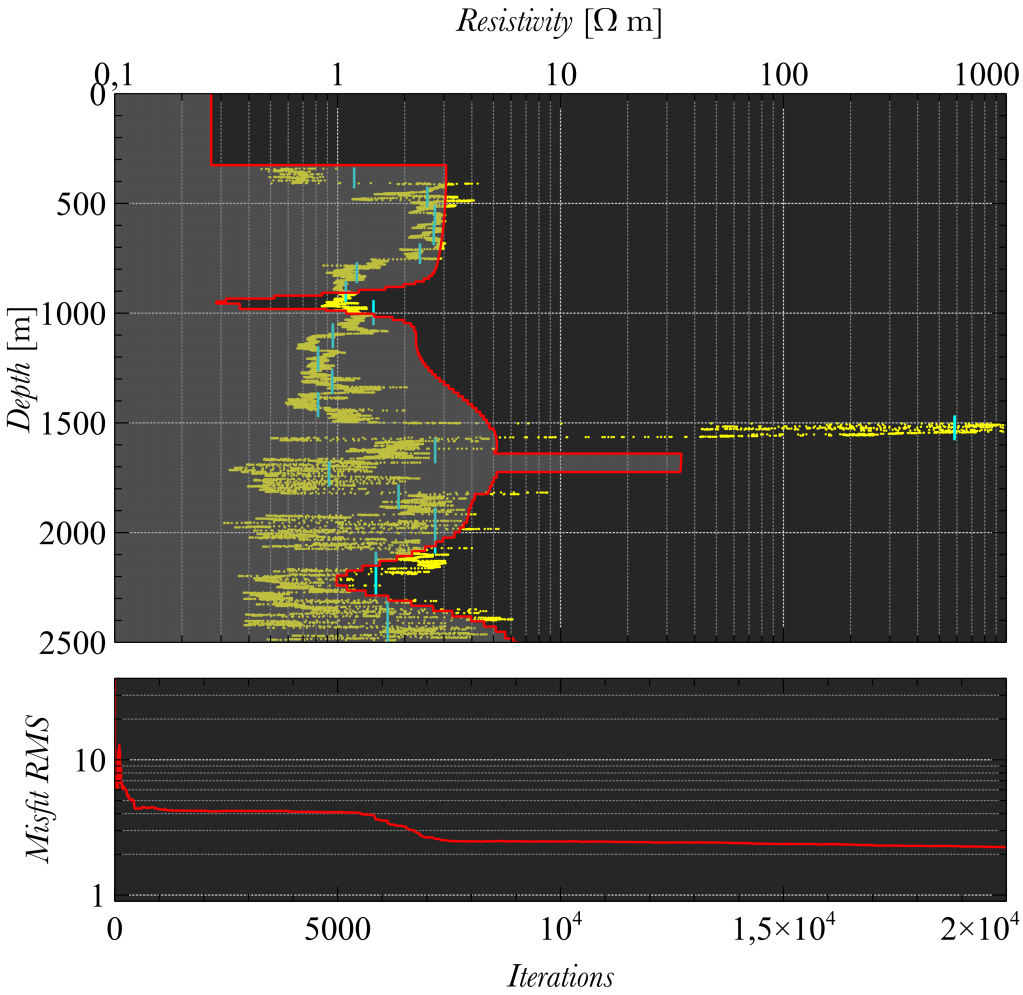


Figure 7.9: Resistivity profile after inversion of troll 1 data with dynamic number of parameters. Final misfit RMS = 2.258.

7.2.3 *Dynamic dimensions II*

These results are the ones with the smallest misfit value from the respective inversions. There also seems to be general agreement with the discussion in the previous subsections. As can be seen in Figure 7.10, this has tried to fit the reservoir with a Gaussian function in stead of a Heaviside step function.

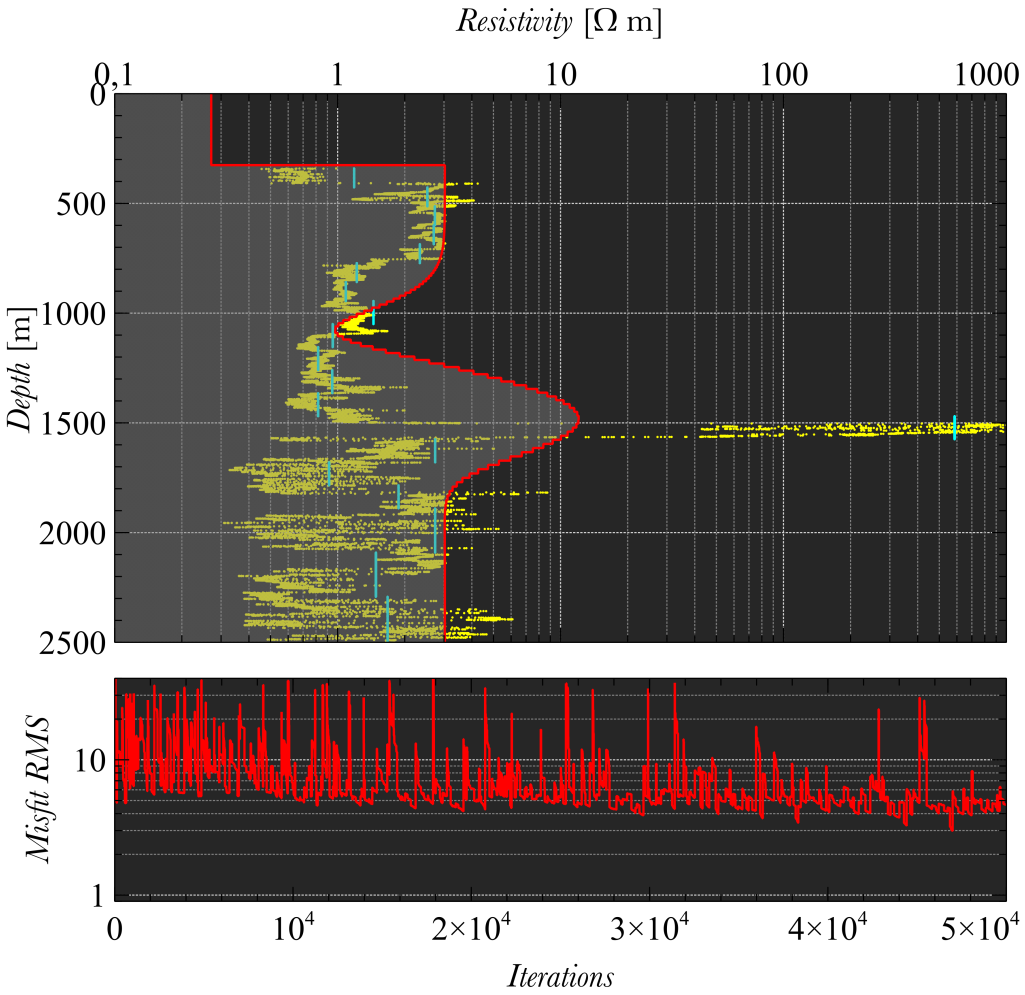


Figure 7.10: Resistivity profile after inversion of troll 1 data with trans-dimensional Simulated Annealing from (Singh et al., 2008). Best misfit RMS = 2.981.

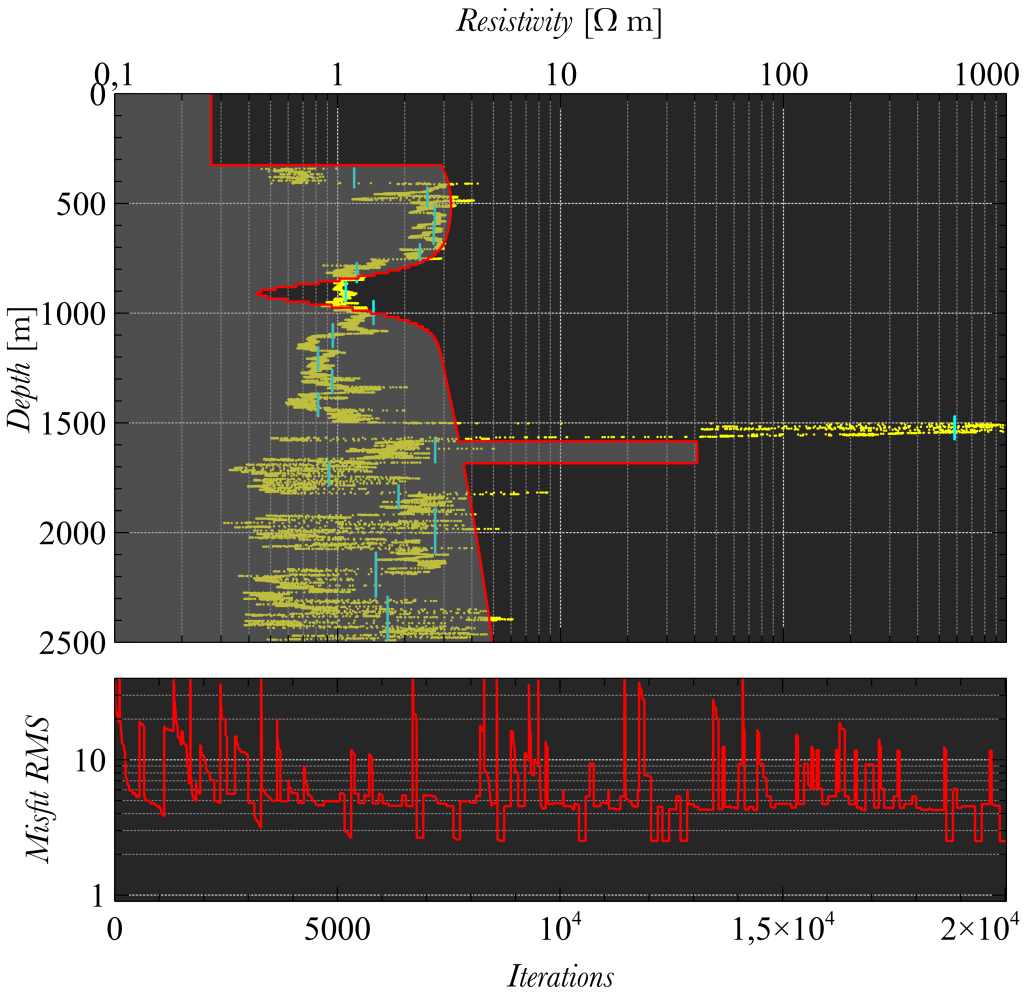


Figure 7.11: Resistivity profile after inversion of Troll 1 data with trans-dimensional Simulated Annealing from (Singh et al., 2008). Best misfit RMS = 2.503.

7.3 INVERTED CSEM DATA FROM THE TROLL 2 MODEL

Again, the self-parameterizing algorithm outperforms the one with fixed parameter dimensions. From Figure 6.6 it it can be seen that the data from this model is stems from sources which is not above the reservoir. It was therefore not certain what to expect in terms of if the methods would recover the target reservoir or not. The fact that the towline is partially not covering the target could be a reason why the different recovered profiles seems to differ more from each other than was the case for the towing towards the receiver which is the case in Section 7.2.

Accepted	FIXED DIMENSIONS 12/20		DYNAMIC DIMENSIONS 15/20	
	Mean	*Std. dev.*	*Mean*	*Std. dev.*
Iterations	10298.92	3679.12	17179.47	4466.24
Elapsed time [s]	97403.16	14731.89	148405.28	29432.60
ε	3356.51	1082.68	1586.24	1225.03
$RMS(\varepsilon)$	4.12	0.85	2.71	1.05
Accept ratio	0.33	0.07	0.31	0.11
N_p	1.00	0.00	3.80	3.32
N_r	1.00	0.00	1.87	1.46
N_g	2.00	0.00	5.53	2.29

Table 7.3: Statistics for inversion of troll 2 data.

7.3.1 *Fixed number of parameters*

The resistive top layer is recovered in both Figure 7.12 and Figure 7.13, and there is also agreement of the placement and resistivity of the resistor. The recovered resistivity is about half of what is found in Section 7.2, which seems appropriate in light of the placement of the towline. The inversion with rigged initial values presented in Figure 7.14 also seems to agree in terms of the resistive top layer and the target. The formation beneath the target differs quite much between the two models, and the one in Figure 7.13 seems to be more in agreement with the well log and Figure 7.14 although it is obvious from the terminal misfit and number of iterations, that it has settled in a less fortunate local minimum.

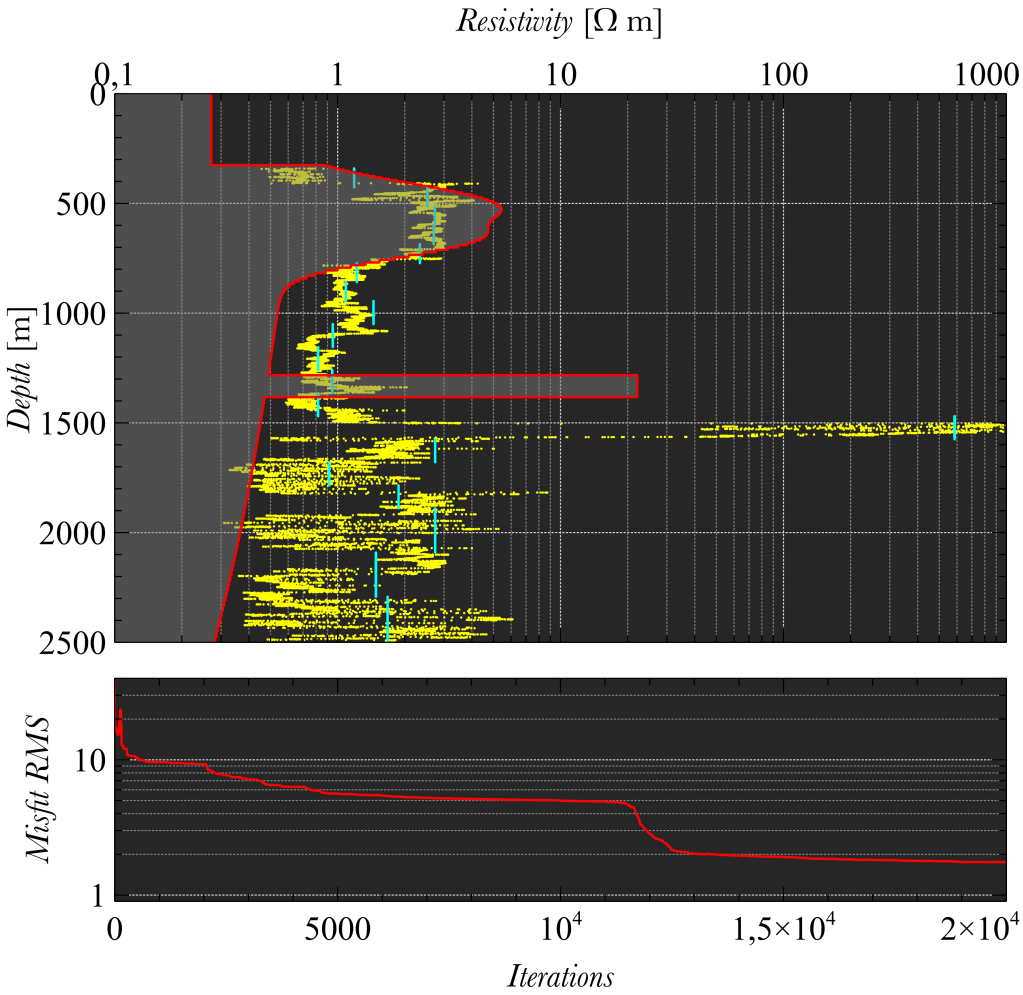


Figure 7.12: Inversion of troll 2 data with fixed number of parameters. Final misfit RMS = 1.751.

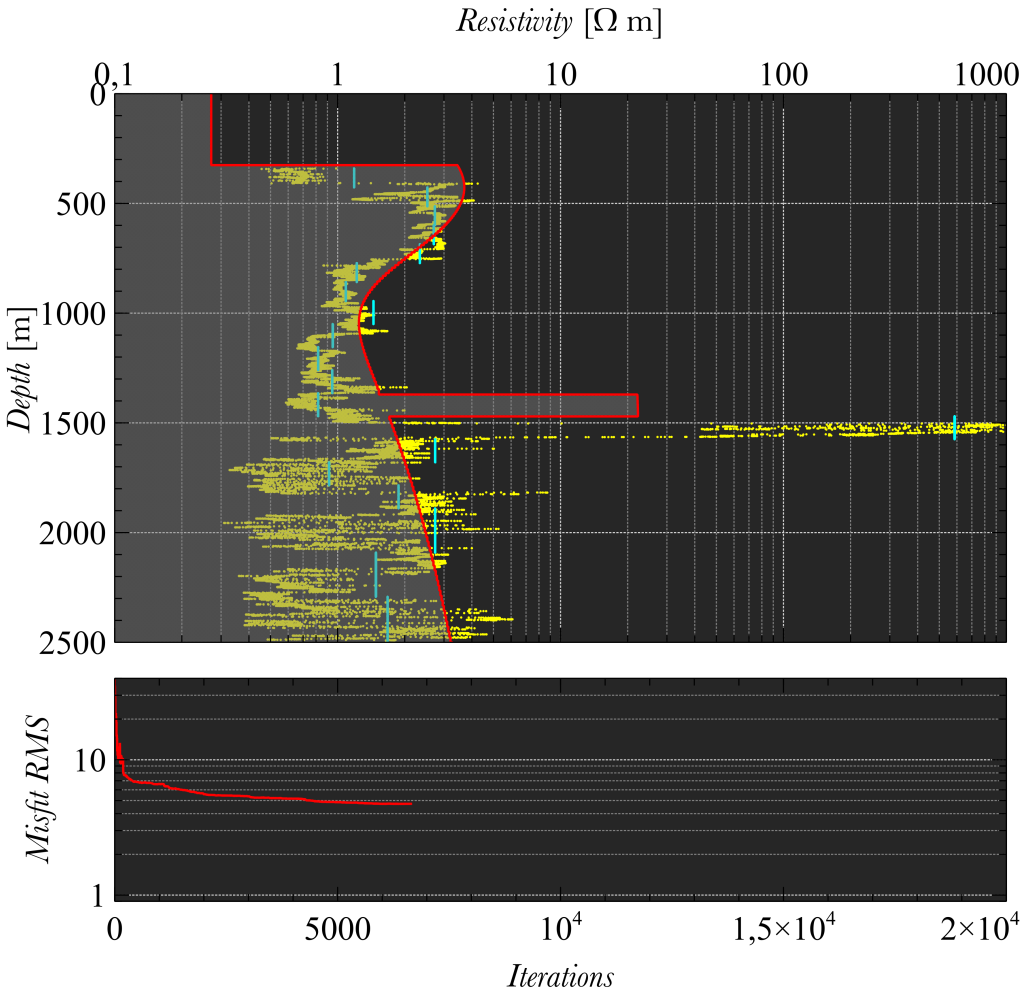


Figure 7.13: Inversion of troll 2 data with fixed number of parameters. Final misfit RMS = 4.730.

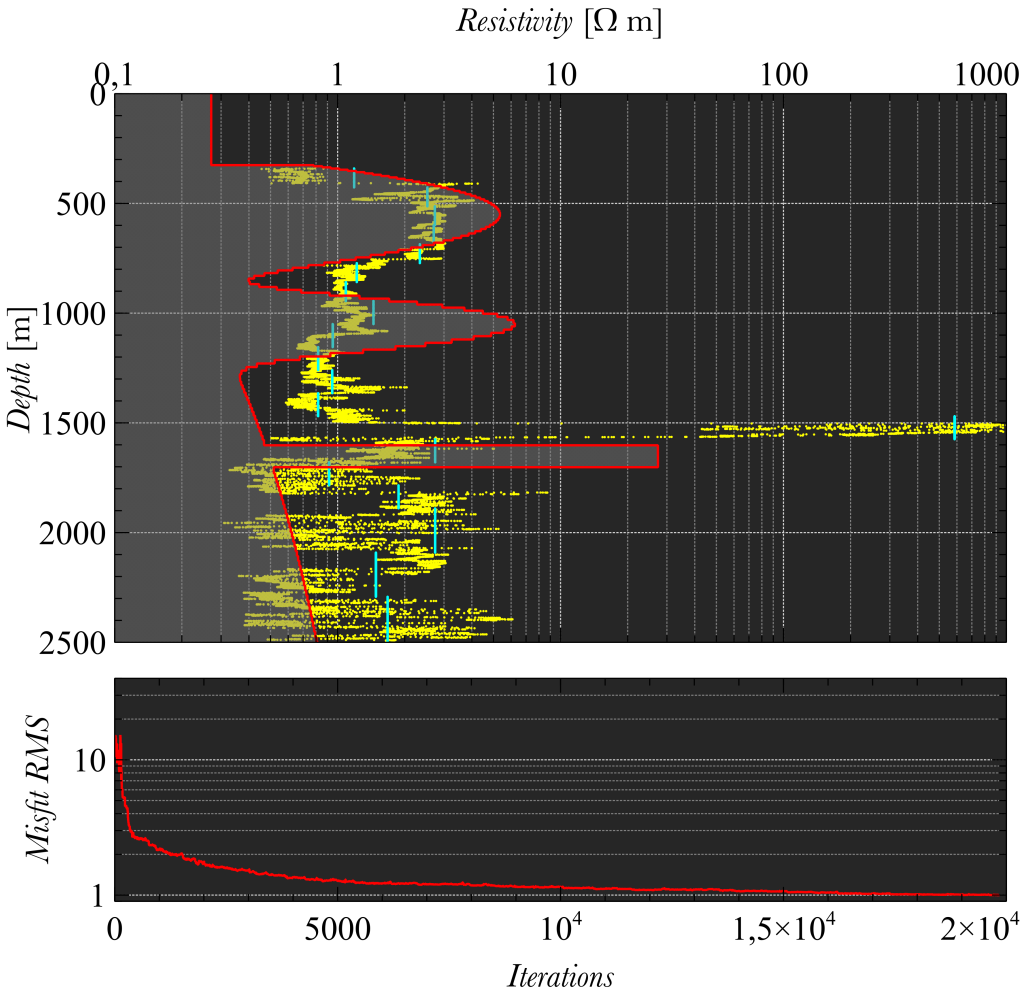


Figure 7.14: Misfit RMS traceplot after inversion of troll 2 data with fixed number of parameters. Final misfit RMS = 1.000.

7.3.2 *Dynamic dimensions*

The resistive top layer and the resistor is in accordance with what was found in Section 7.2.1. Also an increasing resistivity in the region below the resistor can be spotted in both Figure 7.15 and Figure 7.16, although recovered with different functions.

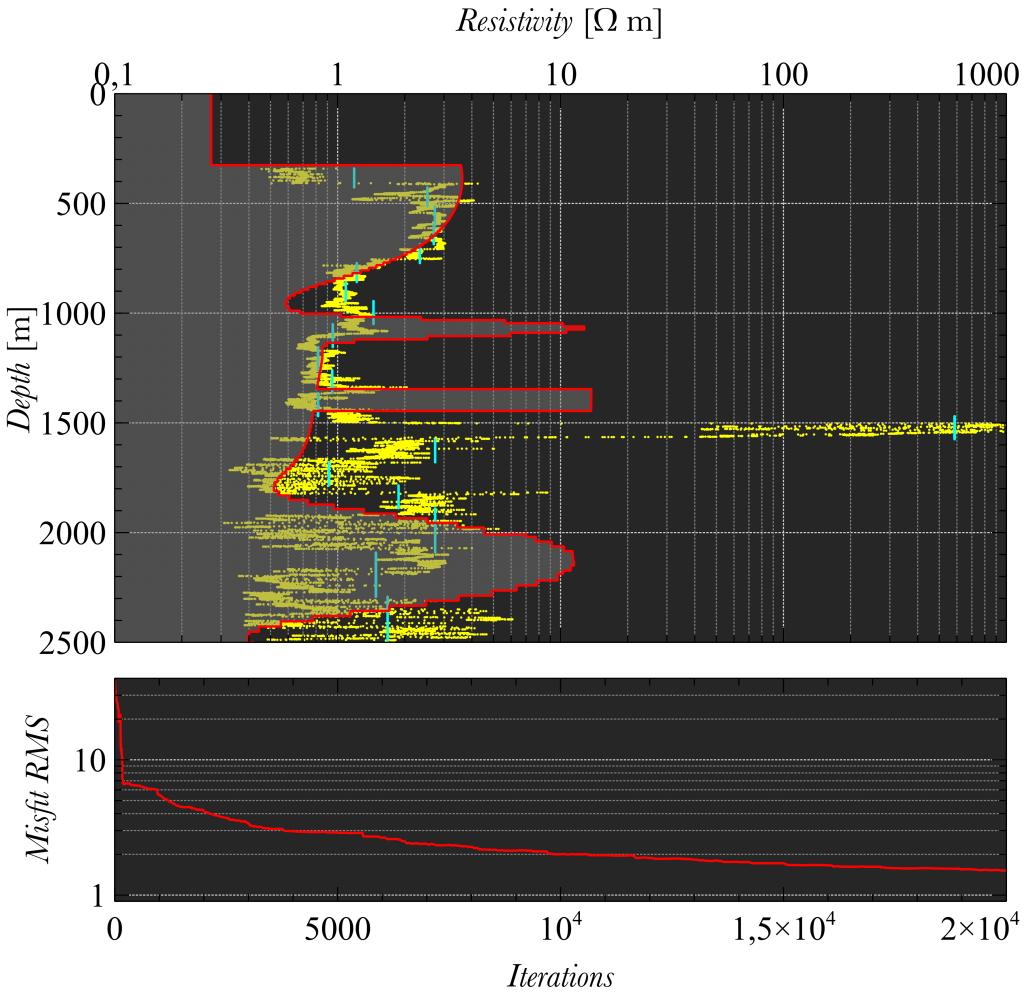


Figure 7.15: Misfit RMS traceplot after inversion of troll 2 data with fixed number of parameters. Final misfit RMS = 1.499.

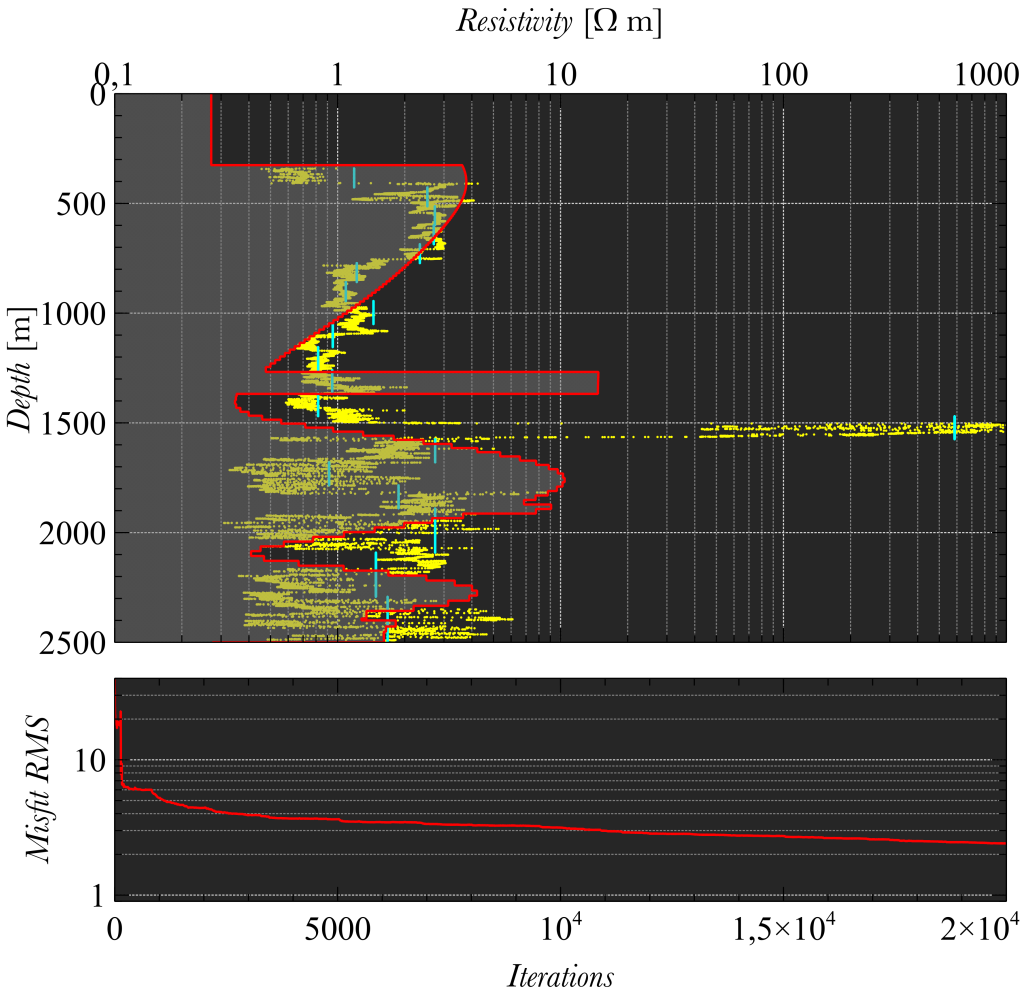


Figure 7.16: Misfit RMS traceplot after inversion of troll 2 data with fixed number of parameters. Final misfit RMS = 2.410.

7.4 INVERTED CSEM DATA FROM THE TROLL 3 MODEL

As can be in Table 7.4, fewer inversions was carried out in this data set than in the previous, but there should still be enough information to analyse. The trans-dimensional method deliver better results on average than the method with fixed number of parameters in terms of misfit value. The CSEM data from this model is sampled with the source towline on top of the reservoir. The effects of this can be seen from a greater consensus between the recovered profiles than when the source is towed outside the field. Also, the transverse resistance is about doubled, and close to the one measured from the well bore log.

	FIXED DIMENSIONS		DYNAMIC DIMENSIONS	
Accepted	5/6		6/6	
	Mean	*Std. dev.*	*Mean*	*Std. dev.*
Iterations	14962.8	5340.97	16509.33	2690.46
Elapsed time [s]	94612.14	31046.57	109663.65	17991.63
ε	2414.51	1531.50	1391.45	309.61
$RMS(\varepsilon)$	3.42	1.14	2.69	0.31
Accept ratio	0.36	0.10	0.29	0.05
N_p	1.00	0.00	3.50	2.51
N_r	1.00	0.00	1.67	0.82
N_g	2.00	0.00	7.00	3.22

Table 7.4: Statistics for inversion of troll 3 data.

7.4.1 *Fixed number of parameters*

The recovered profiles in Figure 7.17 and Figure 7.18 is similar in terms of the resistive top layer, the following thin conductive layer and the location and shape of the target. The reference which was inverted from rigged initial values places the reservoir about 250 m higher and has a more modest thin conductive layer in addition to a more conductive background in general.

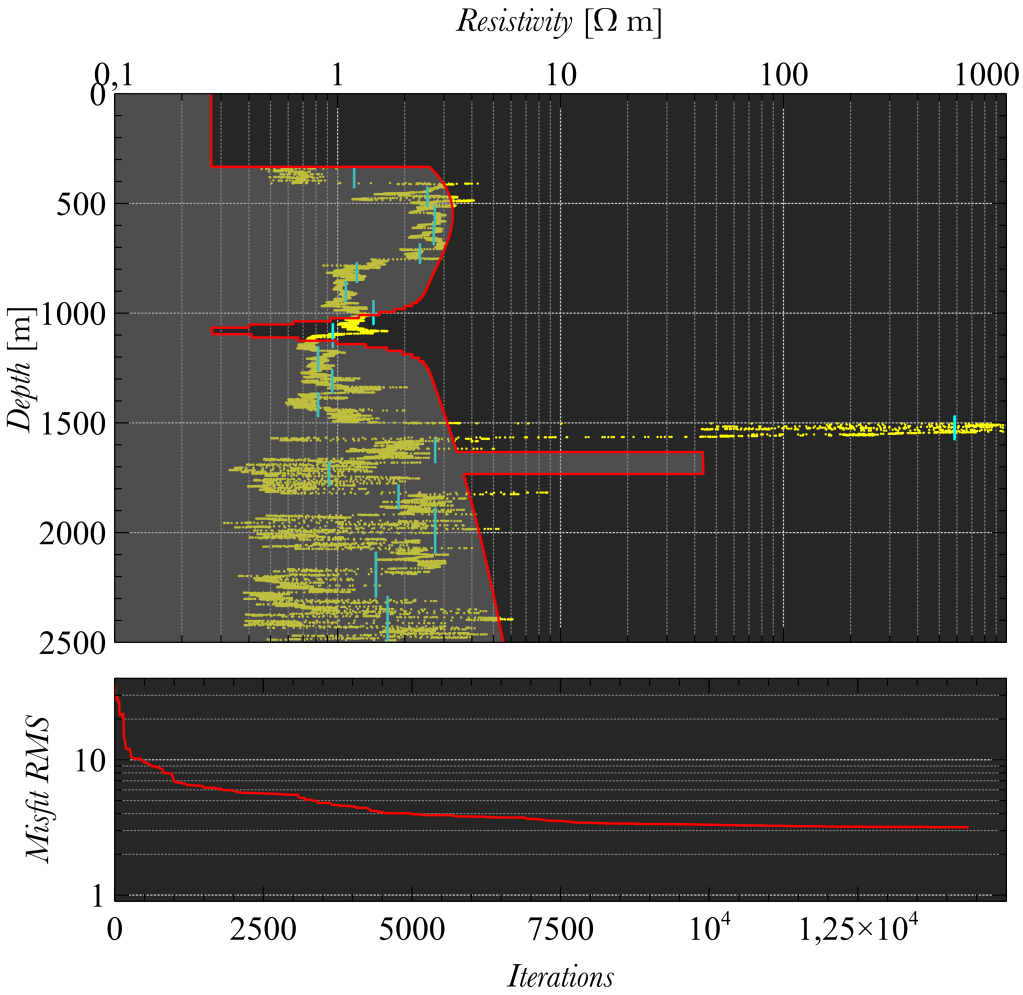


Figure 7.17: Inversion of troll 3 data with fixed number of parameters. Final misfit RMS = 3.176.

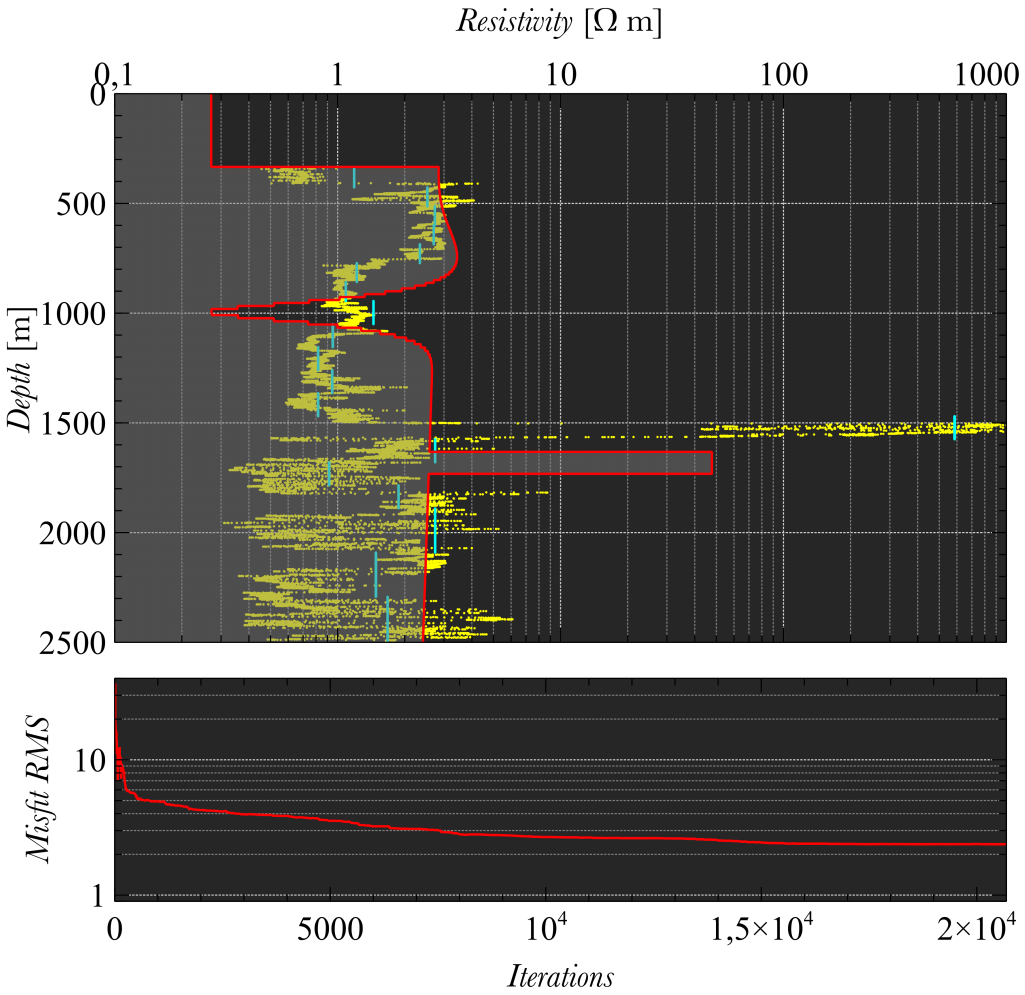


Figure 7.18: Inversion of troll 3 data with fixed number of parameters. Final misfit RMS = 2.376.

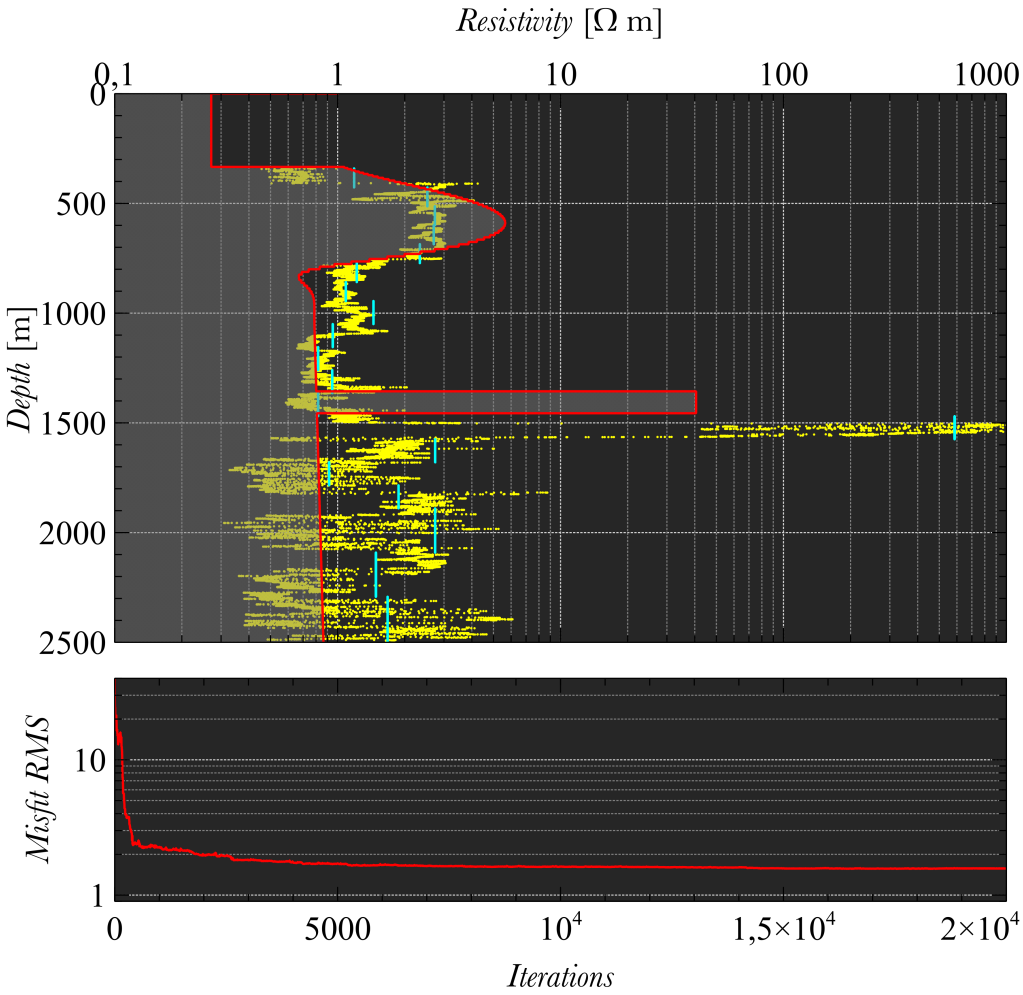


Figure 7.19: Inversion of troll 3 data with fixed number of parameters. Final misfit RMS = 1.572.

7.4.2 *Dynamic dimensions*

Aside from the top resistive layer, the two recovered profiles illustrated in Figure 7.20 and Figure 7.21 differ quite a lot relative to the other inversions discussed above.

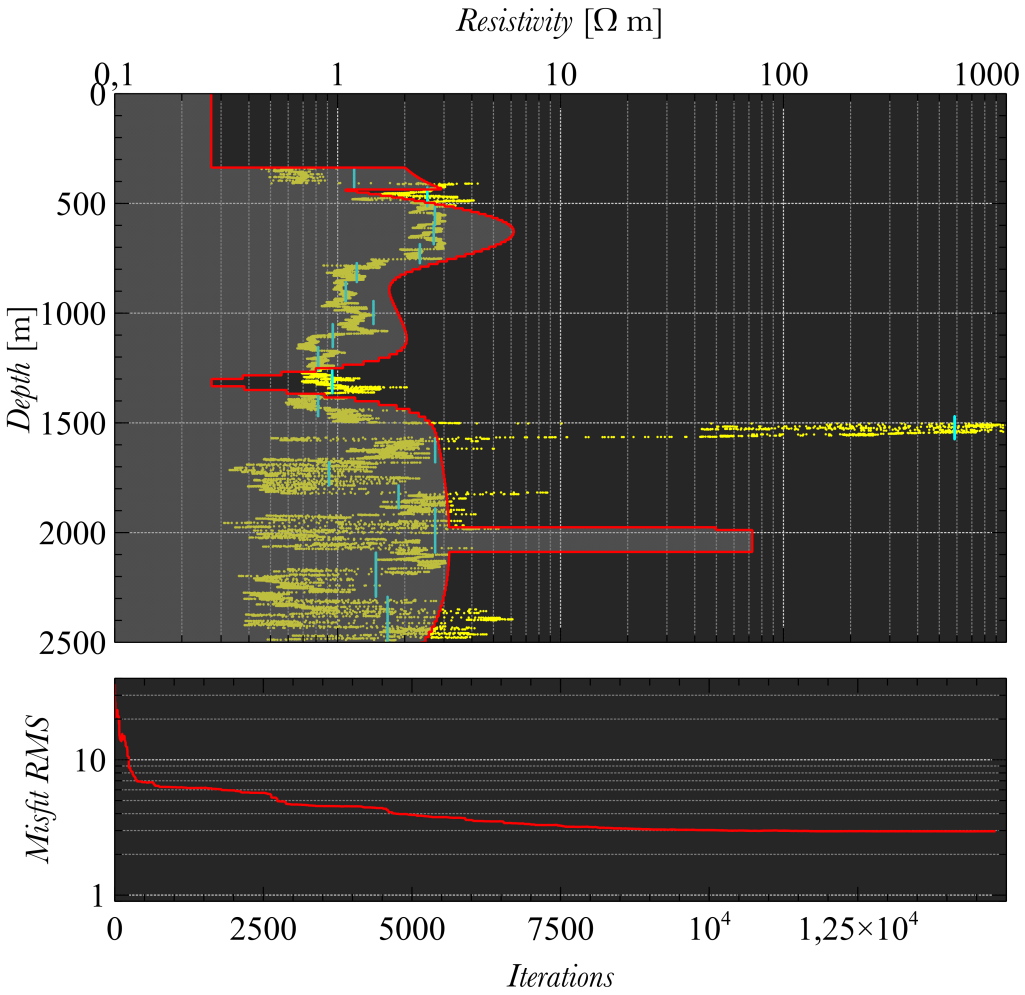


Figure 7.20: Inversion of troll 3 data with dynamic number of parameters. Final misfit RMS = 2.963.

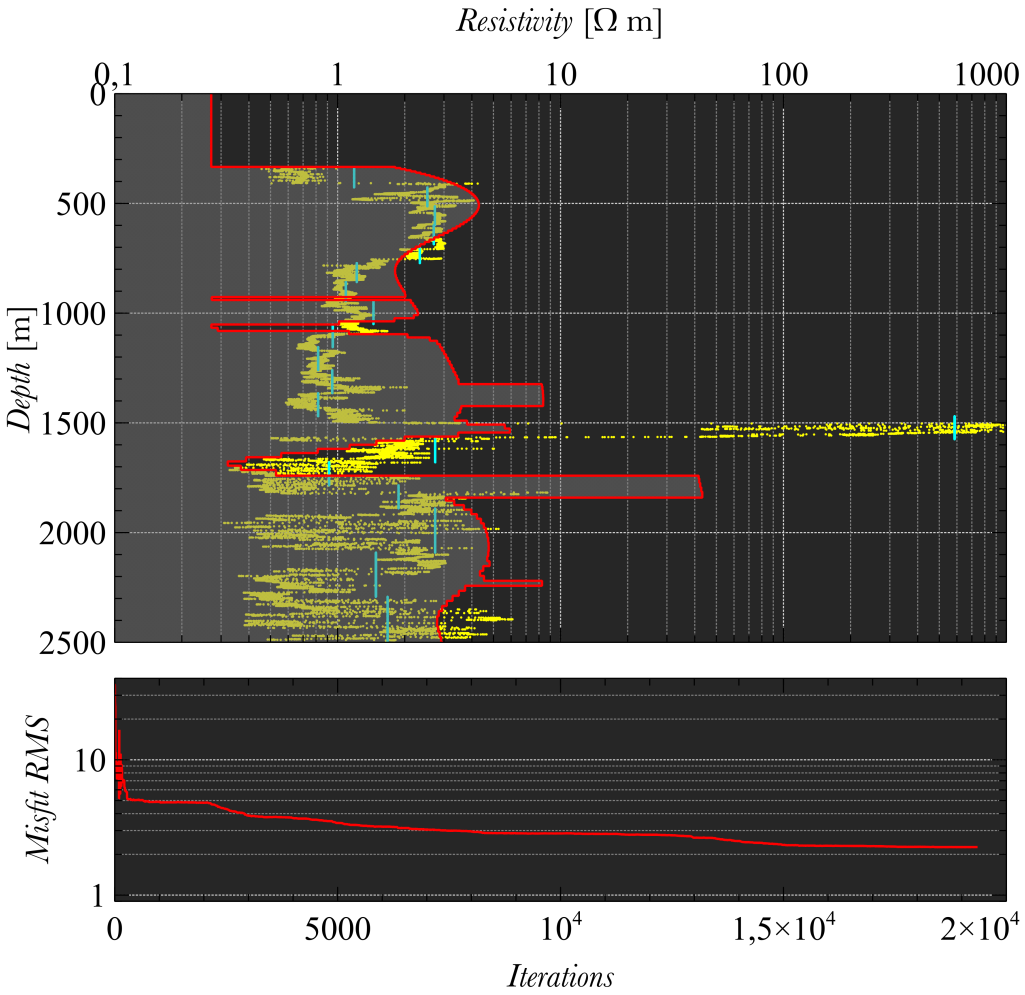


Figure 7.21: Inversion of troll 3 data with dynamic number of parameters. Final misfit RMS = 2.264.

7.5 OBSERVED DATA FROM TROLL 4

The source in this model is also towed partially outside the reservoir, which can be seen in Figure 6.6. It could therefore be expected to see some of the same tendencies as was discussed in Section 7.3. As was the case in the other models, the trans-dimensional self-parameterizing Simulated Annealing algorithm performs better than the regular Simulated Annealing algorithm with fixed dimensions.

	FIXED DIMENSIONS		DYNAMIC DIMENSIONS	
Accepted	5/6		6/6	
	Mean	*Std. dev.*	*Mean*	*Std. dev.*
Iterations	11023.00	5092.75	17250.67	5613.99
Elapsed time [s]	82179.7	35084.50	113011.03	30223.60
ε	2931.08	1119.06	1714.16	799.09
$RMS(\varepsilon)$	3.87	0.74	2.95	0.63
Accept ratio	0.29	0.05	0.36	0.07
N_p	1.00	0.00	2.83	1.94
N_r	1.00	0.00	2.33	1.37
N_g	2.00	0.00	6.67	2.16

Table 7.5: Statistics for inversion of troll 4 data.

7.5.1 *Fixed number of parameters*

Both Figure 7.22 and Figure 7.23 recover the top resistive layer, and seems to agree on an increasing resistivity in the lower layers. They also seems to recover a resistive layer at a depth of about 1000 m, and this seems to be in accordance with the results in Figure 7.24. However, as can be seen in Figure 7.24, the reference inversion places a very resistive layer at about 1700 m below the seabed, and obtains an overall better misfit.

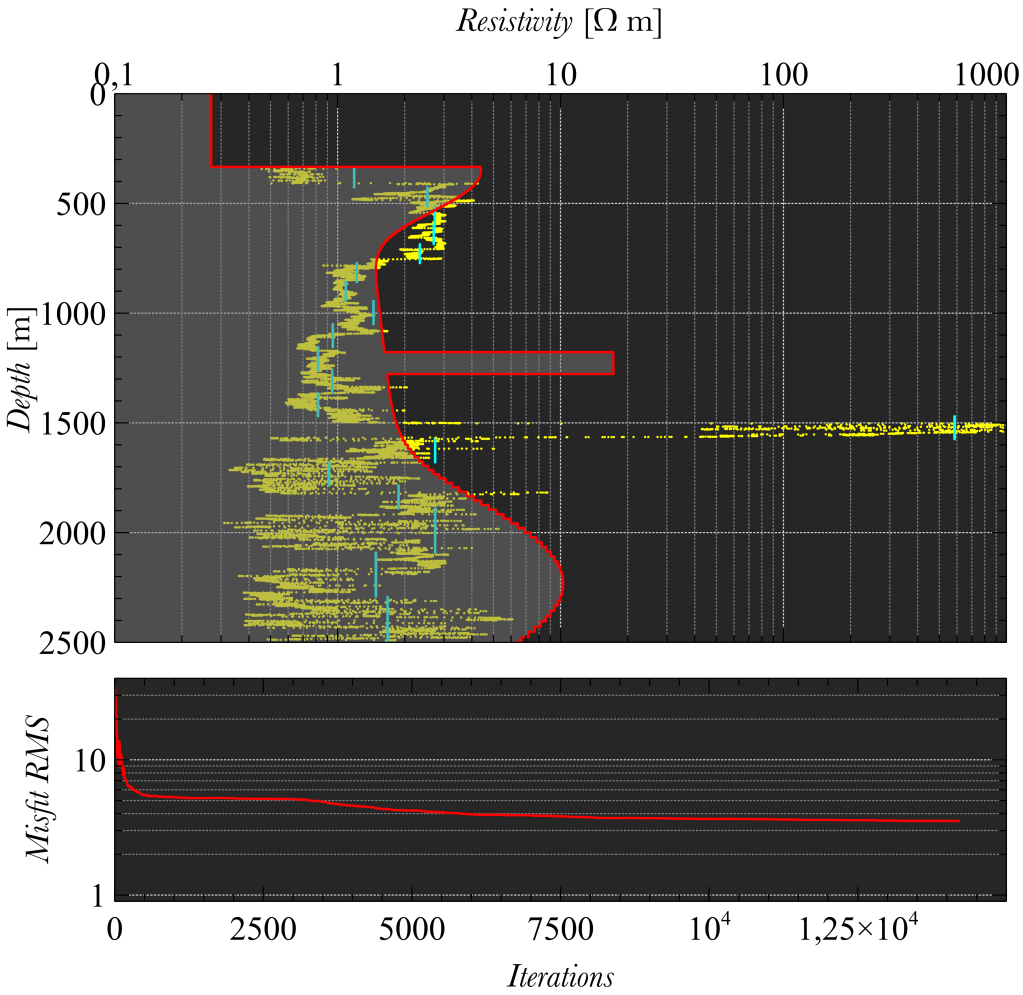


Figure 7.22: Inversion of troll 4 data with fixed number of parameters. Final misfit RMS = 3.533.

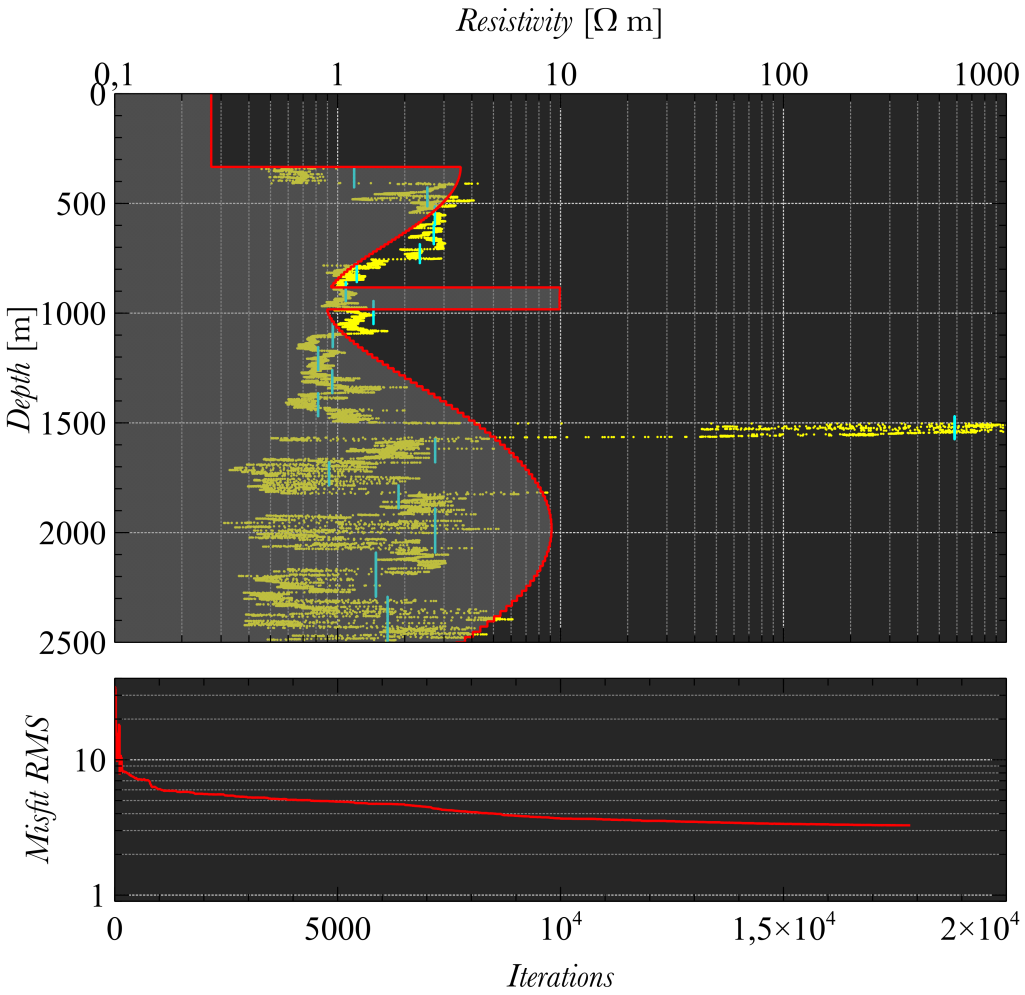


Figure 7.23: Inversion of troll 4 data with fixed number of parameters. Final misfit RMS = 3.279.

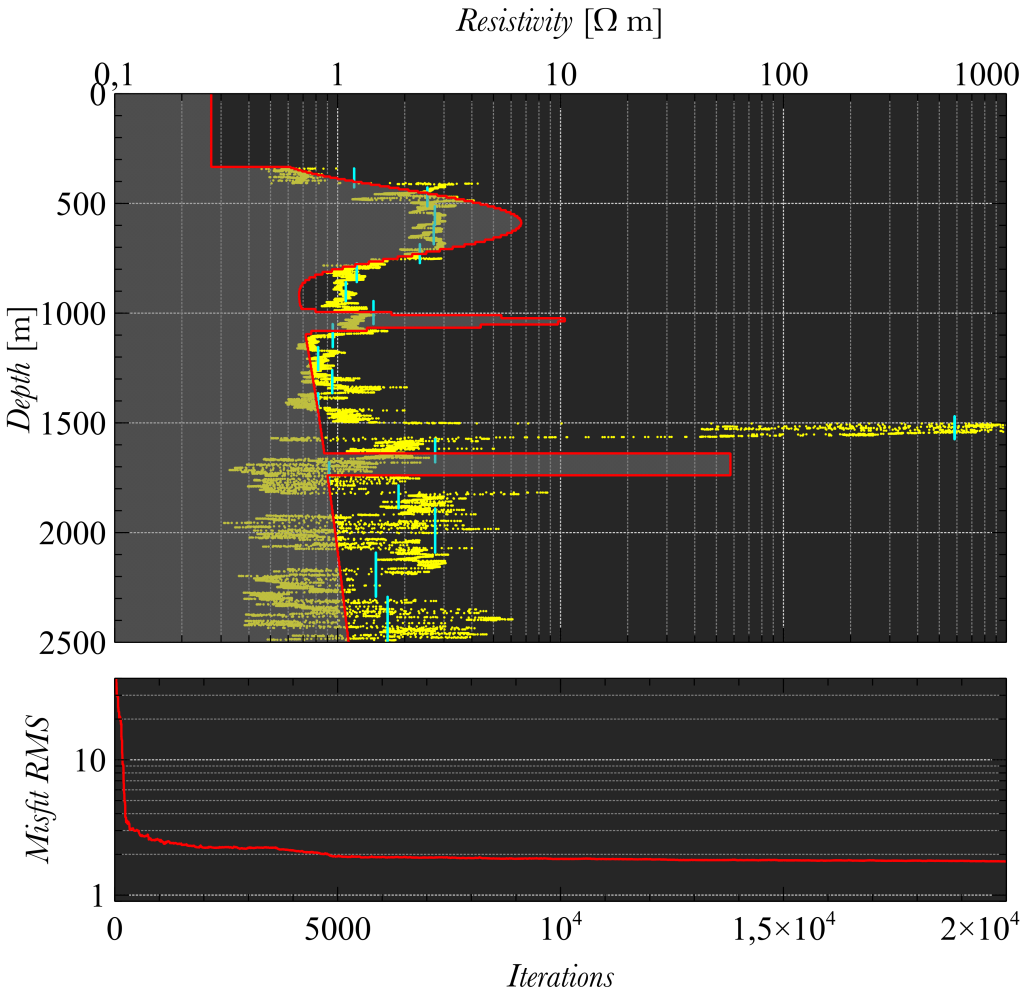


Figure 7.24: Inversion of troll 4 data with fixed number of parameters. Final misfit RMS = 1.774.

7.5.2 *Dynamic dimensions*

As can be seen in Figure 7.25 and Figure 7.26, there is an overall consensus between the two resulting resistivity profiles considered. With an exception of the top resistive layer, the resulting resistivity profile is quite different from the ones recovered with a fixed number of parameters in Section 7.5.1.

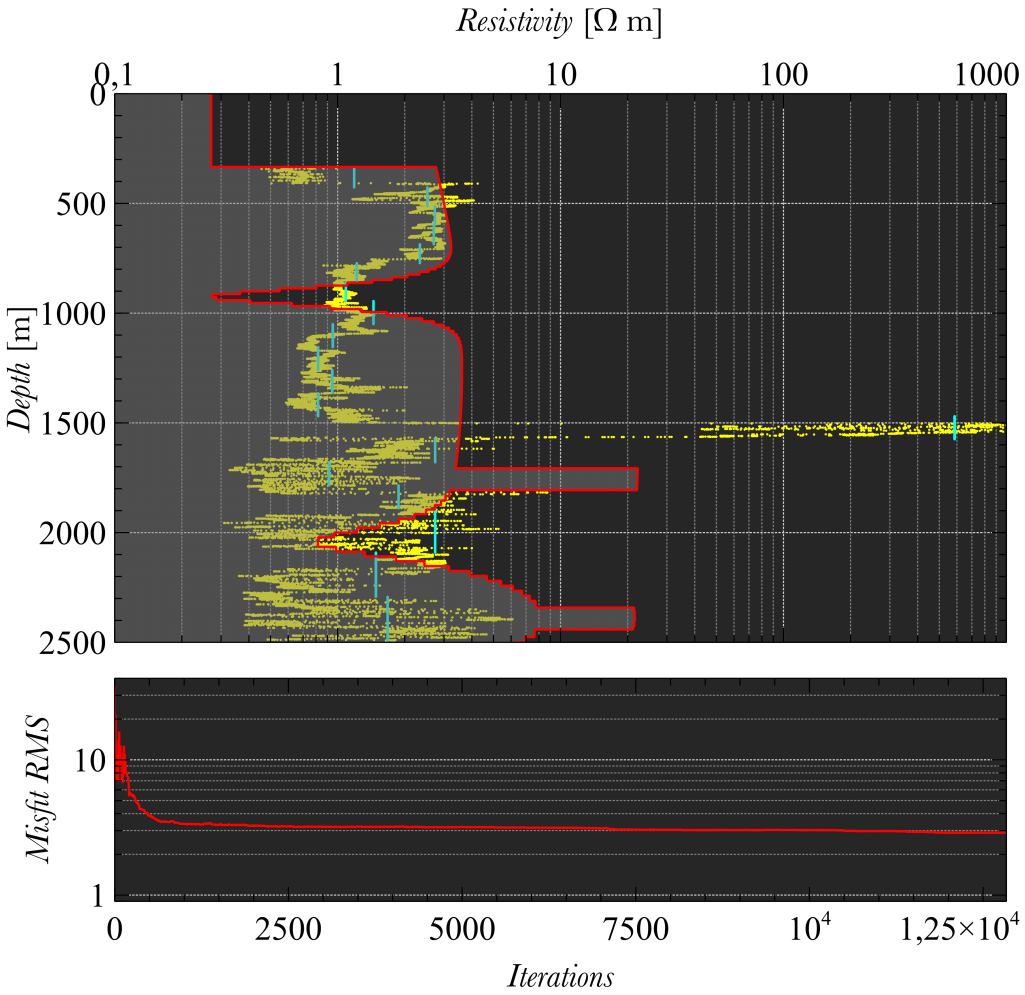


Figure 7.25: Inversion of troll 4 data with dynamic number of parameters. Final misfit RMS = 2.892.

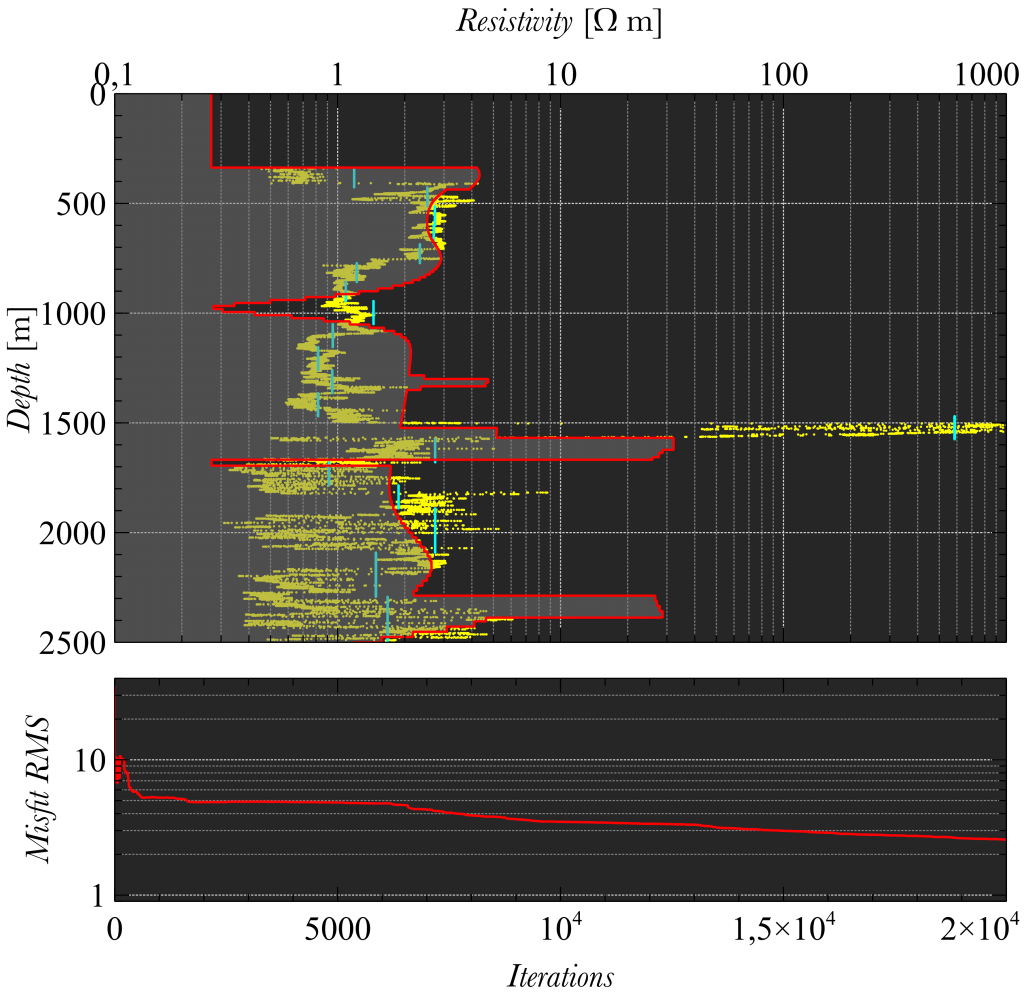


Figure 7.26: Inversion of troll 4 data with dynamic number of parameters. Final misfit RMS = 2.569.

7.6 ADDITIONAL REMARKS

The different data sets results in different resistivity profiles, and the location of the source towline seems to be of importance. When the source is towed directly above the reservoir (models Troll 1 and Troll 3), there is more consensus between the different inversions. The inversions also recovers profiles that is more in accordance with the measurements from the well bore log. This is not surprising, as it can be seen from Figure 6.6 that the well is located closer to these towlines than the towlines in the models Troll 2 and Troll 4. In these models, the source is towed away from the reservoir, and this results in a less resistive recovered target in addition to more dissension between the different inversions. The inverted data from the Troll 1 model presented in Section 7.2 gives the best results in terms of average misfit. There is also a greater consensus about the resistivity profile between the different inversions in this model than in the other models.

Also, the inversions done with the trans-dimensional algorithm gives in general a lower misfit than the one where the parameter dimensions are fixed. This is so, even though the terms included in the fixed dimensional case was included after the resistivity profile was known and thus adapted to be able to fit this profile. The trans-dimensional method is less dependent on such model assumptions, hence the advantage of the trans-dimensional algorithm is expected to increase when less or no information about the sub-seabed resistivity is known. This is as expected as it can include more terms in the resistivity profile function. The results reveals that the trans-dimensional algorithm includes more terms on average than what is set in the fixed dimensions case, but settles with less parameters than the maximum allowed.

In the different models, the number of terms used varies much, perhaps with the exception of N_r, which is apparent from the standard deviations provided. However, there seems to be consensus between the different models in terms of both mean values and their standard deviations. This is illustrated in Table 7.6, which gives mean and standard deviation values for the mean and standard deviation values of the different Troll models.

	MEAN VALUES		STANDARD DEVIATIONS	
	Mean	*Std. dev.*	*Mean*	*Std. dev.*
N_p	3.983	1.278	2.773	0.674
N_r	2.080	0.370	1.330	0.362
N_g	6.183	0.765	2.49	0.491

Table 7.6: Statistics for average parameter dimension values and standard deviations from the different Troll models.

It is interesting that the averaged parameter dimensions differs so little from the different models. The parameter dimension statistics from the synthetic model also seems to agree with the mean values in Table 7.6, which can be explained by that the synthetic model and the Troll models are quite similar. This means that the trans-dimensional simulated annealing, on average, settles on the same number of terms in the resistivity profile function. This further indicates that the algorithm is able to extract the required number of parameters needed to fit the observed data, which was a desired property when designing this method.

When compared to other methods applied on this type of problems (see e.g. (Ray and Key, 2012), (Key, 2009) (Hermanrud, 2009) and (Roth and Zach, 2007)), the parameterization introduced in this thesis is able to reproduce complex resistivity profiles with significantly less parameters than the comparable methods would have been able to do.

8 CONCLUSION

This thesis introduces a new parameterization of the model space in a one dimensional global inversion scheme, with the intention of reducing the number of parameters, and by that, reducing the number of iterations needed in stochastic, global inversion methods. The parameterization provides an explicit representation of the model space with a basis constrained on the problem at hand.

The parameterization is applied on inversion of marine CSEM data and is stated in Equation 3.5. The necessary theory to completely explain the whole inversion process is derived in Chapter 2, Chapter 3 and Chapter 4.

The standard stochastic optimization routine Simulated Annealing is used in the inversion. In order to handle parameter sets with a varying number of parameters, a new self-parameterizing trans-dimensional version of the Simulated Annealing is introduced in Section 5.4.1. The whole inversion process is implemented from scratch in C++, with the exception of the routines handling the forward modeling based on the theory presented in Chapter 4. The whole program consist of several thousand lines of code, and therefore it is not appended on paper. The source code files is therefore instead attached to this thesis in a separate container. As a considerable amount of time was spent on writing this code and getting it to work appropriately, it is the most important part of this thesis. Files containing the complete set of input variable values used to create the results displayed in Chapter 7 is also attached. With this, every result presented is can be reproduced exactly.

The results from inversion of five different data sets using methods with both fixed and dynamic number of parameters is presented in Chapter 7 and concludes this master thesis.

8.1 CONCLUSION

The parameterization introduced in this thesis is able to represent complex resistivity profiles with significantly less parameters than comparable methods. When implemented into a global inversion routine and tested on both synthetic and real data sets, it is able to recover resistivity profiles that demonstrate good resemblance with provided well bore log data of the real data sets.

There is a general tendency that the trans-dimensional, self-parameterizing Simulated Annealing algorithm introduced in Section 5.4.1 performs better than the regular Simulated Annealing algorithm with fixed parameter di-

mensions. It achieves in general lower absolute misfit both on average and in single inversions.

8.2 OUTLOOK

Since the method parameterize a one dimensional resistivity profile, it is of mostly academic interest. The essence is however possible to extend to higher dimensions, and a future goal would be to present an analogue parameterization of a 2D or 3D model which can be implemented into stochastic inversion schemes, and hopefully make these computationally feasible.

BIBLIOGRAPHY

E.H.L. Aarts and P.J.M. van Laarhoven. Statistical cooling: a general approach to combinatorial optimization problems. *Philips Journal of Research*, 40(4):193–226, 1985.

Lasse Amundsen, Lars Løseth, Rune Mittet, Svein Ellingsrud, and Bjørn Ursin. Decomposition of electromagnetic fields into upgoing and downgoing components. *Geophysics*, 71(5):G211–G223, 2006.

Adeel Ansari, Afza Bt. Shafie, and Abas B. Md. Said. Relationship of resistivity contrast and thickness depth of hydrocarbon for seabed logging application. *IJCSI International Journal of Computer Science Issues*, 9 (3):214–221, 2012.

Harry Cohn and Mark Fielding. Simulated annealing: Searching for an optimal temperature schedule. *SIAM Journal on Optimization*, 9(3):779–802, 1999.

Steven Constable. Ten years of marine csem for hydrocarbon exploration. *Geophysics*, 75(5):75A67–75A81, 2010.

Steven Constable and Leonard J. Srnka. An introduction to marine controlled-source electromagnetic methods for hydrocarbon exploration. *Geophysics*, 72(2):WA3–WA12, 2007.

Steven Constable and Chester J. Weiss. Mapping thin resistors and hydrocarbons with marine csem methods: Insights from 1d modeling. *Geophysics*, 71(2):G43–G51, 2006.

T. Eidsmo, S. Ellingsrud, L.M. MacGregor, S. Constable, M.C. Sinha, S. Johansen, F.N. Kong, and H. Westerdahl. Sea bed logging (sbl), a new method for remote and direct identification of hydrocarbon filled layers in deepwater areas. *First Break*, 20:144–152, 2002.

Pål T. Gabrielsen, Ivar Brevik, Rune Mittet, and Lars O. Løseth. Investing the exploration potential for 3d csem using a calibration survey over the troll field. *First Break*, 27:67–75, 2009.

D. P. Ghosh. The application of linear filter theory to the direct interpretation of geoelectrical resistivity sounding measurements. *Geophysical Prospectiong*, 19:192–217, 1971.

David J. Griffiths. *Introduction to Electrodynamics*. Prentice Hall, Reed College, 3rd edition, 1999.

Bruce Hajec. Cooling schedules for optimal annealing. *Mathematics of Operations Research*, 13(2):311–329, 1988.

Kristine Hermanrud. Monte carlo methods for assessment of the probability distributions of sub sea resistivity models. Technical report, Norwegian Institute of Science and Technology, 2009.

Lester Ingber. Adaptive simulated annealing (asa): Lessons learned. *Control and Cybernetics*, 1995.

H. K. Johansen and K. Sørensen. Fast hankel transforms. *Geophysical Prospecting*, 27:876–901, 1979.

Kerry Key. 1d inversion of multicomponent, multifrequency marine csem data: Methodology and synthetic studies for resolving thin resistive layers. *Geophysics*, 74(2):F9–F20, 2009.

Kerry Key. Is the fast hankel transform faster than quadrature? *Geophysics*, 77(3):F21–F30, 2012.

S. Kirkpatrick, C.D. Gelatt, and M.P. Vecci. Optimization by simulated annealing. *Science, New Series*, 220(4590):671–680, 1983.

Bernt Øksendal. *Stochastic Differential Equations*. Springer, University of Oslo, P.O. Box 1053 Blindern, 0316 Oslo, 5th edition, 2010.

M. Locatelli. Simulated annealing algorithms for comtinuous global optimization, 2000.

Lars Ole Løseth. *Modelling of Controlled Source Electromagnetic Data*. PhD thesis, Norwegian University of Science and Technology, Department of Physics, 2007.

L.O. Løseth and B. Ursin. Electromagnetic fields in planarly layered anisotropic media. *Geophysical Journal International*, 170:44–80, 2007.

M. Lundy and A. Mees. Convergence of an annealing algorithm. *Mathematical programming*, 34:111–124, 1986.

Rune Mittet and Jan Petter Morten. The marine controlled-source electromagnetic method in shallow water. *Geophysics*, 78(2):E67–E77, 2013.

Jan Petter Morten and Rune Mittet. Detecting and imaging sensitivity of the marine csem method. *Geophysics*, 77(6):E411–E425, 2012.

Jan Petter Morten, Astrid K. Bjørke, and Trude Støren. Csem data uncertainty analysis for 3d inversion. *SEG Houston 2009 International exposition and annual meeting*, pages 724–728, 2009.

Norwegian Petroleum Directorate. 31/2-5 wellbore, 2014. URL http://factpages.npd.no/factpages/default.aspx?culture=en.

Sophocles J. Orfanidis. *Electromagnetic Waves and Antennas*. Rutgers University, Rutgers University, 1st edition, 2008.

N. R. Patel, R. L. Smith, and Z. B. Zabinsky. Pure adaptive search in monte carlo optimization. *Mathematical programming*, 43:317–328, 1988.

Anandaroop Ray and Kerry Key. Bayesian inversion of marine csem data with a trans-dimensional self parameterizing algorithm. *Geophysical Journal International*, 191:1135–1151, 2012.

H. Edwin Romeijn and Robert L. Smith. Simulated annealing for constrained global optimization. *Journal of Global Optimization*, 5:101–126, 1994.

Friedrich Roth and Jürgen J. Zach. Inversion of marine csem data using up-down wavefield separation and simulated annealing, 2007. URL www.emgs.con/technical_papers.

Hemant Singh, Amitay Isaacs, Tapabrata Ray, and Warren Smith. A simulated annealing algorithm for single objective trans-dimensional optimization problems. *Eighth International Conference on Hybrid Intelligent Systems*, pages 19–24, 2008.

Alexey Stovas and Bjørn Ursin. Reflection and transmission responses of layered transversely isotropic viscoelastic media. *Geophysical Prospecting*, 51:447–477, 2003.

Bjørn Ursin. Review of elastic and electromagnetic wave propagation in horizontally layered media. *Geophysics*, 48(8):1063–1081, 1983.

David Vanderbilt and Stephen G. Louiem. A monte carlo simulated annealing approach to optimization over continuous variables. *Journal of Computational Physics*, 56:259–271, 1984.

Jincai Zhang. Pore pressure prediction from well logs: Methods, modifications, and new approaches. *Earth-Science Reviews*, 108:50–63, 2011.