Advances in Magnetotelluric Modelling: Time-Lapse Inversion, Bayesian Inversion and Machine Learning

A dissertation presented by

Dennis Conway

In fulfilment of the requirements
for the degree of
Doctor of Philosophy

THE UNIVERSITY
of ADELAIDE

Submitted to the
Department of Earth Sciences,
School of Physical Sciences, Faculty of Sciences

Adelaide, September 2018

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This thesis presents advancements to the area of magnetotelluric (MT) modelling. There are three main aims to this work. The first aim is to implement an inversion to model time-lapse MT data in a temporal dimension. The algorithm considers the entire dataset at once, with penalisations for model roughness in both the spatial and temporal dimensions. The inversion is tested on synthetic data, as well as a case-study from a coal-seam gas dewatering survey. Second is to explore the problem of non-uniqueness in MT data inversion by implementing a 1D Bayesian inversion using an efficient sampler. The implemented model includes a novel way of regularising MT inversion by allowing the strength of smoothing to vary between different models. The Bayesian inversion is tested on synthetic and case-study datasets with results matching known data. The third aim is to implement a proxy function for the 3D MT forward function based on artificial neural networks. This allows for rapid evaluation of the forward function and the use of evolutionary algorithms to invert for resistivity structures. The evolutionary search algorithm is tested on synthetic data sets and a case-study data set from the Curnamona Province, South Australia. Together, these three novel algorithms and software implementations represent a contribution to the toolkit of MT modelling.

Thesis Supervisors: Graham Heinson, Stephan Thiel, Derrick Hasterock

STATEMENT OF ORIGINALITY

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed

Date

ACKNOWLEDGEMENTS

I would like to thank my supervisor Graham Heinson for his continual guidance throughout my PhD and the support, encouragement and vision he has provided. I would also like to thank Brad Alexander from the School of Computer Science for his guidance in machine learning. A special thanks also to Lars Krieger and Yohannes Didana for their friendship and readiness to help with any question.

I would also like to thank my fellow geophysics post-graduate students, both past and present, for providing a wonderful atmosphere and friendly work environment. I would particularly like to thank my office-mates Nigel Rees and Joe Rugari for continual motivation, discussions and games of mini-pool which made coming to the office a pleasure. Thank you too to Goran Boren for organising field-trips and tending to the MT instruments.

Finally, I would like to thank my friends and family. Thank you to my parents and sister for their support, and to Myall in the School of Biological Sciences for the many, many coffee breaks. Lastly, I would like to thank Sonia for her constant encouragement, support and confidence.

CHAPTER

ONE

INTRODUCTION

1.1 Introduction

The magnetotelluric method (MT) is a geophysical technique where passively sourced, time-varying magnetic field and induced electrical field data are collected at the Earth's surface. These data are sensitive to subsurface conductivity. The ability to accurately produce images of the Earth's subsurface based on these data, a process known as inversion, is an essential step in the interpretation of MT data.

The interpretation of MT data is important for a number of applications, and often the inversion processes will be tailored to the particular application. For example, MT is used in hydrocarbon exploration, often to complement seismic data (Constable, 1998; Mitsuhata et al., 1999; Zhang and Yan, 2015). As such, extensive research has been conducted into joint inversions of seismic and magnetotelluric data (Yang et al., 2002; Jegen et al., 2009; Ogaya et al., 2016; Lan et al., 2018). This has the ability to combine the layer boundary sensitivity of a seismic technique with the resistivity sensitivity of MT.

Other applications include exploration in the minerals industry, where MT can be used to image ore-bodies with resistivity contrasts. Compared to other methods, MT has the ability to image much deeper areas of the

crust, which may also be of interest in mineral exploration (Bedrosian and Feucht, 2014). For example, Heinson et al. (2006) have shown that the Olympic Dam deposit in South Australia has large areas of conductivity underneath the ore-body, which are interpreted as fluid pathways from the mantle. The identification of similar structures elsewhere may lead to the identification of new areas of prospectivity (Robertson et al., 2016). Inversions of these deep and complex structures, however, require 3D MT inversion codes, which can be slow to run (Newman and Alumbaugh, 2000).

A relatively new application of MT data has been to monitor the temporal development of dynamic systems. After the injection of 3.1 ML of saline fluids into an enhanced geothermal system (EGS) reservoir in Paralana, South Australia, Peacock et al. (2012) were able to record changes in the magnetotelluric data which indicated a drop in the resistivity of the target area. A similar experiment was undertaken at the Haberno EGS reservoir, South Australia, where Didana et al. (2017) detected a slight change in resistivity after the injection of 36.5 ML of fluid. Similar experiments have successfully imaged resistivity changes during coal-seam gas dewatering (Rees et al., 2016b), shale-gas hydraulic fracturing (Rees et al., 2016a), and carbon sequestration (Ogaya et al., 2013; Vilamajó et al., 2013).

With resistivity monitoring proving a new application for MT, there is the need for new inversion algorithms that can produce dynamic Earth models. Recent advancements include a probabilistic three-dimensional (3D) inversion (Rosas-Carbajal et al., 2015) and a 1D layer-stripping approach which has been applied to carbon sequestration (Ogaya et al., 2016). Standard MT inversions can also be adapted, for example Rees et al. (2016b) use cascading two-dimensional (2D) inversions to model a coal-seam gas depressurisation, where the results from the pre-injection inversion are used as a constraint for the post-injection inversion. Further research can draw inspiration from other geophysical techniques, such as electrical resistivity tomography (ERT), where time-lapse inversion has been extensively researched due to the common environmental application of the technique (Hayley et al., 2011).

Parallel to the development of inversion algorithms for particular applications, an increase in the computational power available has led to an increasing interest in algorithms for Bayesian inversion. A typical inversion

will only resolve a single resistivity model, however a Bayesian inversion will consider all models which are supported by the data, and rate how well they fit the data.

Bayesian inversion of 1D MT data has been successfully applied using MCMC algorithms, where the calculation of the forward algorithm is simple (e.g. Grandis et al., 1999; Guo et al., 2011; Mandolesi et al., 2018). Several challenges exist which make 2D and 3D inversions computationally difficult, however. One compromise is to reduce the number of parameters in the inversion using a reparameterisations (Chen et al., 2012; Rosas-Carbajal et al., 2015). The first full pixel-based inversion has been implemented for 2D radio magnetotelluric data by Rosas-Carbajal et al. (2013). The authors present a successful inversion of radio-magnetotelluric data, with best results when combined with electrical resistivity tomography in a joint inversion. For a pixel-based 3D MT inversion the process becomes more difficult, with Grandis (2006) concluding from preliminary studies that a very coarse mesh would be required to keep computation times feasible.

The process of MT inversion may also benefit from machine learning. Recent advancements in deep-learning algorithms (e.g. Krizhevsky et al., 2012; Dahl et al., 2013; Srivastava et al., 2014; Silver et al., 2016) and the use of highly efficient graphical processing units (GPUs) (see Oh and Jung, 2004) for model training has led to a wider application of machine learning in many fields. Once trained, machine learning models have the benefit of rapid evaluation.

For this reason, artificial neural networks (ANNs) has been suggested to circumvent the computational burden of MT inversion in high dimensions. Inversions of MT data using ANNs have been developed in 1D and 2D in limited model ranges (Zhang and Paulson, 1997); for simplified 3D models (Spichak and Popova, 2000); for a broader range of model classes in 2D (Shimelevich et al., 2007); for azimuthally anisotropic resistivities (Montahaei and Oskooi, 2014); and for a limited class of 3D models representing Kimberlite pipes (Shimelevich et al., 2017).

1.2 Aims and Objectives

The aim of this dissertation is to investigate new techniques which may be applied to MT modelling and inversion. There are three main contributions to this field which are presented. Firstly, in the developing field of MT monitoring, there is a need for new algorithms to leverage the new paradigm of inverting for a time-changing resistivity model. To this end, an algorithm is presented which simultaneously inverts time-lapse MT data to produce a model which is smooth both spatially and temporally. The efficacy of the algorithm is tested on both synthetic and case-study datasets.

Secondly, in the growing field of probabilistic modelling of MT data, a new approach to sampling from the probability distribution of likely resistivity models is presented. Probabilistic modelling of MT data is a computationally difficult problem for 2D and 3D model spaces, made difficult by the high covariance between model parameters and long evaluation times of the MT forward function. To alleviate these problems a sampler is implemented which utilises the gradient of the probability space to select samples with maximum efficiency, thereby reducing computational cost.

Finally, the possibility of using machine learning techniques to bypass the costly MT forward function altogether is investigated. A 3D resistivity model is considered and an artificial neural network is trained to learn the relationship between the model and the forward response. The trained proxy function is simple to evaluate and would result in a drastic speed-up in computation time.

1.3 Thesis outline

The thesis is presented in five chapters excluding this introductory chapter. In Chapter 2, the theory of the MT method is introduced, as well as geophysical inversion. A literature review of MT inversion is presented, for both the deterministic and Bayesian cases. The chapter also includes an introduction to machine learning and a literature review of its applications in the geosciences. Chapter 3 describes a deterministic inversion

of time-lapse data using simultaneous inversion. It is has been published in the journal *Earth, Planets and Space*. In Chapter 4 a new Bayesian inversion is presented for the probabilistic modelling of 1D MT data. It has been accepted for publication in the journal *Pure and Applied Geophysics*. Chapter 5 describes the use of ANNs to approximate the 3D MT forward function. It has been submitted to *Computers & Geosciences*. Finally, Chapter 6 provides an overview of the main contributions of the dissertation, and discusses future avenues for research in MT inversion.

CHAPTER

TWO

BACKGROUND

2.1 Magnetotelluric theory

The magnetotelluric method (MT) is a plane-wave electromagnetic (EM) technique used to investigate conductivity of the Earth (Tikhonov, 1986; Cagniard, 1953). As with all EM methods, the physics governing the MT method is described by Maxwell's Equations:

$$\begin{aligned} \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J} &= \nabla \times \boldsymbol{H} && \text{Faraday's Law,} \\ \frac{\partial \boldsymbol{B}}{\partial t} &= -\nabla \times \boldsymbol{E} && \text{Ampère's Law,} \\ \nabla \cdot \boldsymbol{B} &= 0 && \text{Gauss' Law,} \\ \nabla \cdot \boldsymbol{D} &= q_e && \text{Coulomb's Law,} \end{aligned} \tag{2.1}$$

where $\boldsymbol{D}$ is dielectric displacement, $\boldsymbol{J}$ is the electric current density due to free charges, $\boldsymbol{H}$ is the magnetic field intensity, $\boldsymbol{B}$ is the magnetic induction, $\boldsymbol{E}$ is electric field intensity and q_e is the electric charge density due to free charges.

In addition, for isotropic, linear media, the following relationships are true:

$$\boldsymbol{B} = \mu \boldsymbol{H}, \tag{2.2}$$

$$\boldsymbol{D} = \epsilon \boldsymbol{E}, \quad \text{and} \tag{2.3}$$

$$\boldsymbol{J} = \sigma \boldsymbol{E}. \tag{2.4}$$

For Earth resistivity studies, μ_0 and ϵ_0 are used for μ and ϵ, as the variations in electrical permittivites and magnetic permeabilities are negligible compared to those of bulk rock conductivities.

MT is a plane-wave EM method, which means that the source fields are sufficiently distant that they can be approximated as uniform, plane-polarised waves with vertical incidence to the Earth's surface, and that displacement currents are negligible.

These assumptions allow for the formation of the following relationship between the horizontal electric ($\boldsymbol{E}$) and horizontal source magnetic fields ($\boldsymbol{H}$) for a given angular frequency ω at the Earth's surface:

$$\boldsymbol{E}(\omega) = \boldsymbol{Z}(\omega)\boldsymbol{H}(\omega), \tag{2.5}$$

where $\boldsymbol{Z}$ is a 2×2 complex matrix known as the impedance tensor. The impedance tensor is calculated from MT field data, and contains information about the resistive and conductive parts of the Earth's structure. The complex components can be physically interpreted by their phase angle ϕ and scaled amplitude, commonly known as apparent resistivity, ρ_a. These are calculated as

$$\rho_{a,ij}(\omega) = \frac{1}{\mu_0 \omega} \| Z_{ij}(\omega) \|^2, \quad \text{and} \tag{2.6}$$

$$\phi_{ij} = \arctan \frac{\mathrm{Im}(Z_{ij})}{\mathrm{Re}(Z_{ij})}, \tag{2.7}$$

for each ij-th component of $\boldsymbol{Z}$.

Typically an MT survey is conducted over range of frequencies, as different depths of the Earth structure can be investigated by different frequencies. The depth at which the source signal has decayed to 1/e of its original strength, known as the skin-depth δ, can be used as a guide for the penetration-depth of a survey. It is approximated at a given frequency f by

$$\delta = 500\sqrt{\rho T}, \tag{2.8}$$

where ρ is the volumetric apparent resistivity of the Earth through which an EM wave is propogating and T is the period of the signal (inverse frequency).

The magnetic source fields for MT are created from an array of natural phenomena. At frequencies above 1 Hz, the dominant source is lightning activity. Lightning strikes produce electromagnetic pulses known as sferics, which propagate within the waveguide formed between the ground and ionosphere (Nagano et al., 2007). These pulses dominate the frequency band from 3 Hz to 30 kHz. Source fields at lower frequencies are predominantly formed by the interaction of solar winds on the Earth's magnetosphere. This is the source for frequencies 10^{-1} Hz to 10^{-4} Hz. The gap between these two sources, in the range of 0.1 Hz to 5 Hz, is known as the "dead-band" due to the lack of a natural source field. Magnetotelluric data from this band is typically of a lower quality. Data quality can be improved by using an active grounded dipole source in a technique known as the controlled source audio magnetotelluric method (Goldstein and Strangway, 1975).

2.2 Geophysical inversion

Geophysical inversion is the process of modelling the Earth's subsurface using data typically collected at the Earth's surface. The inversion is a problem with two halves: finding an earth model which fits the data, and finding all the models which fit the data (Backus, 1988). These are known as the existence problem and the uniqueness problem, respectively.

Integral to both of these problems is the use of a forward function, F, which produces a synthetic response for a given earth model $\boldsymbol{m}$. For observed geophysical data $\boldsymbol{d_{obs}}$ with corresponding errors $\boldsymbol{\epsilon}$, the inverse problem is defined by solving for $\boldsymbol{m}$ such that

$$\boldsymbol{d_{obs}} = F(\boldsymbol{m}) + \boldsymbol{\epsilon}. \tag{2.9}$$

Usually the inverse of F does not exist or is discontinuous, so some form of regularisation is required to find $\boldsymbol{m}$. The regularisation will depend on the type of geophysical data and the goal of the inversion. This review predominantly focuses on magnetotelluric inversion, a non-linear problem which has been regularly used as an example for advances in geophysical theory (Constable et al., 1987; Tarits et al., 1994; Grandis et al., 1999; Guo et al., 2011), however the general methods are applicable to other geophysical inverse problems as well.

2.2.1 Existence problem

The inversion of MT data to a single resistivity model is an under-determined problem and requires some form of regularisation. This means that there will be many models which can provide a equivalent misfit χ^2 between the model response $F(\boldsymbol{m}) = \boldsymbol{z}$ and the observed data $\boldsymbol{z}_{obs}$ given data errors $\boldsymbol{\sigma}$. The data misfit over n frequencies is given by

$$\chi^2 = \sum^{n} \frac{\left\| \mathbf{z}_n - \mathbf{z}_{obs_n} \right\|^2}{\sigma_n^2}. \tag{2.10}$$

The root-mean-square error, or RMS, is commonly used instead of χ^2 to describe the fit of models as it is not dependent on the size of the data. It is given by

$$RMS_e = \sqrt{\chi^2/n}. \tag{2.11}$$

In a 1-D MT inversion, one approach to dealing with the under-determined problem is to invert with the class of models known as simple layered models. In a simple layered model with a fixed number of layers the algorithm

will find the single layered model with the lowest χ^2. The concept has also been extended to 2-D MT inversion by deGroot Hedlin and Constable (2004), who created an inversion which inverts for a fixed number of resistivities and their boundaries.

The main alternative to layered inversion is the many-layered smooth-inversion of Constable et al. (1987), which is used in 1-D, 2-D and 3-D MT inversions. In a smooth-inversion, an objective function o is constructed by combining the misfit and a roughness term, R, weighted by a smoothness parameter λ. The resulting sum,

$$o(\boldsymbol{m}) = \chi^2(\boldsymbol{m}) + \lambda R(\boldsymbol{m}), \tag{2.12}$$

is minimised. This is known as Tikhonov regularisation (Tikhonov, 1963). The model roughness is defined as either the first or second derivative smoothness of the model parameters, i.e. for a 1-D earth with resistivity varying in the z direction

$$R_1(\boldsymbol{m}) = \int (dm/dz)^2 dz, \tag{2.13}$$

or

$$R_2(\boldsymbol{m}) = \int (d^2m/dz^2)^2 dz. \tag{2.14}$$

For a 2-D or 3-D inversion the roughness is calculated in each dimension and R is the summation of these roughnesses.

The user specifies either the smoothness parameter λ or else a target RMS is given and the optimal λ is found by the inversion. A large value of λ results in more weight given to the roughness parameter and a smoother model overall, and a small λ would include more structure as the inversion focuses on the data-misfit. Minimisation of the objective function is achieved using Newton-type iterative methods.

The smooth-inversion principle has been applied to 2-D inversion (deGroot Hedlin and Constable, 1990) and 3-D inversion (Mackie et al., 2001). Inversion of MT data in 3-D becomes very computationally intensive due to the large number of model parameters. Various advancements have been made to the inversions to improve efficiency, including an algorithm by

Siripunvaraporn et al. (2005) with that operates in the data-space. This makes the inversion time largely dependent on the number of data points instead of model parameters. More efficient algorithms can also decrease inversion time, for example Avdeev and Avdeeva (2006) use a Limited Memory Quasi-Newton technique which avoids costly calculations of second derivatives.

2.2.2 Uniqueness problem

There are several methods which are used to explore parameter space and solve the uniqueness problem of geophysical inversion. One example is analysing the model covariance matrix (MCM), which allows for the direct empirical analysis of spatial variation from a linear inversion (Alumbaugh and Newman, 2000). Care must be applied however in non-linear problems such as MT inversion.

An alternative to MCM analysis is most-squares inversion (e.g. Meju and Hutton, 1992; Jackson, 1976). This inversion solves for the extremal cases of models with low and high model resistivity which fit the observed data, and these act as bounding models for the inversion. Another option is to use bootstrap resampling of the data $\boldsymbol{d}$, where data are inverted using a set of data resamplings (e.g. Schnaidt and Heinson, 2015). In both cases a qualitative uncertainty in model parameters may be made by anaylsing the range of inversions.

In order to quantify our understanding of parameter uncertainty, Bayesian methods are required. These methods stem from Bayes' theorem

$$P(\mathbf{m}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{m})P(\mathbf{m})}{P(\mathbf{d})}. \tag{2.15}$$

In this equation, $P(\mathbf{m}|\mathbf{d})$ is the posterior probability distribution function (pdf) of the model conditional on the data, $P(\mathbf{d}|\mathbf{m})$ is the pdf of the data conditional on the model (it is also called the likelihood function, $L(\mathbf{m}|\mathbf{d})$), and $P(\mathbf{m})$ and $P(\mathbf{d})$ are the probabilities of observing $\mathbf{m}$ and $\mathbf{d}$ without taking each-other into account. $P(\mathbf{m})$ is known as the prior.

In the case of MT inversion, the likelihood function is given by how well the responses from the proposed model fit the data compared with the

measured data and the Gaussian distribution of data errors over a range of n frequencies:

$$L(\mathbf{m}|\mathbf{d}) = \prod_{i=1}^{n} \frac{1}{2\pi\sigma_n} \exp\left(- \frac{\|\mathbf{Z}_n - \mathbf{Z}_{obs_n}\|^2}{2\sigma_n^2} \right), \tag{2.16}$$

where the data $\boldsymbol{d}$ are in the form of MT impedances $\boldsymbol{Z}_{obs}$ and associated errors $\boldsymbol{\sigma}$. The quantity $\boldsymbol{Z}$ is obtained from a model $\boldsymbol{m}$ through the use of the forward operator F.

A prior distribution must be given on the inversion parameters, which acts as regularisation for the inversion. Similar to deterministic inversion, the regularisation may take different forms. Common techniques include restricting the number of layers (e.g. Guo et al., 2011); calculating the roughness and penalising rougher models with lower probabilities (e.g. Grandis et al., 1999; Rosas-Carbajal et al., 2013); or having variable number of parameters and a prior which penalises more complex models (e.g. Mandolesi et al., 2018). In some probabilistic inversions, data uncertainty can also be treated as probabilistic parameter (e.g. Guo et al., 2011; Rosas-Carbajal et al., 2013).

Finding a direct solution to 2.15 is usually not possible. Instead the Markov Chain Monte Carlo (MCMC) method is commonly used to find samples from the posterior distribution. In MCMC, a Markov chain is constructed which has its equilibrium distribution as the posterior distribution. This is achieved in different ways by different MCMC samplers, the most common of which is the Metropolis-Hastings (MH) algorithm.

In the MH algorithm, chains are arbitrarily initialised at x_0, with a new sample at each step t drawn from a jumping distribution $g(x_{t*}|x_t)$. The jumping distribution is commonly a Gaussian distribution centred at x_t, but may be any symmetric probability density (that is any distribution g where $g(x_{t^*}|x_t) = g(x_t|x_{t^*})$ is satisfied). An acceptance ratio α is calculated for the new sample,

$$\alpha = \frac{f(x_{t^*})}{f(x_t)}, \tag{2.17}$$

where f is a function proportional to the posterior distribution, in this case $f = bayes$. If $\alpha > 1$ then the new sample is automatically accepted, otherwise it is accepted with probability α. If the sample is accepted then

the chain moves to $x_{t+1} = x_{t^*}$, if it is rejected then $x_{t+1} = x_t$. This eventually leads to a Markov chain converged with a equilibrium distribution which matches 2.15.

The MH algorithm is commonly applied to Bayesian MT inversion (e.g. Guo et al., 2011; Buland and Kolbjørnsen, 2012; Liu et al., 2011). Other samplers include Gibbs sampling (e.g. Grandis et al., 1999; Cerv et al., 2007), DREAM (e.g. Rosas-Carbajal et al., 2015) and NUTS (e.g. Conway et al., 2018). It is also possible to sample using neighbourhood sampling, however although efficient, this method leads to biased results (Guo et al., 2011).

Markov Chain Monte Carlo is traditionally applied to 1D-MT due to the computational difficulties of 2D and 3D problems (Grandis, 2006). Recently however a full pixel 2D-MT MCMC inversion has been implemented by Rosas-Carbajal et al. (2013), and the technique has also been used by parameterised inversions, e.g. inverting for a resistivity change parameterised as a plume body (Rosas-Carbajal et al., 2015).

2.3 Machine learning

Machine learning is a field of computer science centred on algorithms which are able to learn from experience. A program is said to be learning if it improves its ability to perform a task with increasing exposure to new data. Machine learning algorithms may be broadly grouped into three categories: supervised learning, unsupervised learning and reinforcement learning, however the latter is not prevalent in the geosciences.

In supervised learning an algorithm is tasked with finding a function $f : X \mapsto Y$ given a labelled set of data inputs x_i and outputs $y_i = f(x_i)$. If Y may take only a finite number of values then the problem is called a classification problem, otherwise it is called regression. The algorithm will have some measure of fitness, J, known as a loss-function or cost-function, which is calculated on the misfit of $f(x_i)$ against y_i. The algorithm will then seek to minimise J. The representation of f will vary between machine learning techniques. Some of the most commonly algorithms are support vector

machines (SVMs), artificial neural networks (ANNs), decision trees, and ensemble learning.

In contrast, an unsupervised learning algorithm is given only an input dataset x_i. There are various unsupervised problems, however the most common in geosciences is clustering. A clustering algorithm seeks to find a set of k clusters such that for each x_i either belongs to a single cluster or belongs to each cluster with some probability. These are called hard clustering and soft clustering, respectively. These clusters can then be analysed, or used to extract features from the data. Common clustering algorithms include k-means clustering and hierarchical clustering.

2.3.1 Machine learning in geology and geophysics

Regression algorithms have been extensively used in bore-hole geophysics, where general algorithms can replace simplified empirical models. Artificial neural networks are commonly employed to functions mapping bore-hole data to various parameters, including thermal conductivity (e.g. Goutorbe et al., 2006; Gąsior and Przelaskowska, 2014), porosity in crystalline rock (e.g. Konaté et al., 2015b), static formation temperatures (e.g. Bassam et al., 2010), basalt characterisation (e.g. Asfahani and Ghani, 2012), and petroleum reservoir characterisation (e.g. Mohaghegh et al., 1996; Saemi et al., 2007; Karimpouli et al., 2010; Aminian and Ameri, 2005; Helle et al., 2001; Tatar et al., 2015; Gholami et al., 2012).

Classification algorithms are widely used in geoscience to determine lithology. It is common practice to use SVMs or ANNs to classify bore-hole lithology based on down-hole parameters (e.g. Benaouda et al., 1999; Briqueu et al., 2002; Maiti et al., 2007; Wang and Zhang, 2008; Schmitt et al., 2012; Moazzeni and Haffar, 2015; Konaté et al., 2015a; Mahmoodi and Smith, 2015; Karmakar et al., 2017; Deng et al., 2017). The same methods have been used to create geological maps of lithologies based on a combination of airborne geophysical data and satellite data (e.g. Oommen et al., 2008; Leverington, 2010; Yu et al., 2012). Other methods such as random forests, naive Bayes, and k-Nearest neighbors have also been tested by Cracknell and Reading (2013), who identified random forests as the best tool for lithological mapping.

Unsupervised learning algorithms are an ideal tool for analysing geophysical data where there is no prior knowledge, and hence no labelled data to train on. Various clustering algorithms are used in geophysics. One of the most popular, self-organising maps (SOMs), allows for the intuitive interpretation of high-dimensional data. The technique has been applied to seismic facies analysis (de Matos et al., 2006), interpreting multidimensional geophysical data (Klose, 2006), analysing seismic datasets with multiple reflections (Essenreiter et al., 2001), event detection in continuous seismic records (Köhler et al., 2010) and semi-automatic geological mapping from airborne geophysics (de Carvalho Carneiro et al., 2012). Other clustering techniques have been variously used in fault zone imaging (Di Giuseppe et al., 2014), joint inversion (e.g. Paasche and Tronicke, 2007; Sun and Li, 2012; Carter-McAuslan et al., 2014) and geological mapping from gravity, magnetic, and radiometric data (Paasche and Eberle, 2009).

CHAPTER

THREE

TIME-LAPSE INVERSION OF ONE-DIMENSIONAL MAGNETOTELLURIC DATA

Conway D.[1], Heinson G.[1], Rees N.[1], and Rugari J.[1].

[1]School of Physical Sciences,
University of Adelaide, Adelaide SA 5005, Australia

Published in *Earth, Planets and Space*
https://doi.org/10.1186/s40623-018-0791-y

Statement of Authorship

Title of Paper	Time-lapse inversion of one-dimensional magnetotelluric data
Publication Status	Published
Publication Details	Conway, D., Heinson, G., Rees, N., Rugari, J. (2018), Time-lapse inversion of one-dimensional magnetotelluric data, Earth, Planets and Space, 70(1), 27

Principal Author

Name	Dennis Conway
Contribution	Conception of ideas, programming, modelling, visualisation, manuscript writing and academic correspondence
Overall percentage	85%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.

Signature _____ Date 10/5/18

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

(i) the candidate's stated contribution to the publication is accurate (as detailed above);

(ii) permission is granted for the candidate in include the publication in the thesis; and

(iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name	Graham Heinson
Contribution	Assistance with construction of test models, feedback on manuscript and edits

Signature _____ Date 10/5/18

Name	Nigel Rees
Contribution	Provision of data, feedback on case-study models

Signature _____ Date 16/5/18

Name	Joseph Rugari
Contribution	Feedback on manuscript and edits, assistance with conception

Signature _____ Date 15/05/18

Abstract

We present a new tool for modelling time-lapse magnetotelluric (MT) data, an emerging technique for monitoring changes in subsurface electrical resistivity. Time-lapse MT data have been acquired in various settings, including sites of hydraulic fracturing, dewatering and sequestration. It has been shown in other geophysical techniques that the most effective way to model time-lapse data is with simultaneous inversion, which uses information from all time-steps to produce models with higher accuracy and fewer artefacts. We introduce this method to model time-lapse 1D MT data. As with a standard MT inversion our routine penalises spatial roughness at each time-step, however we also introduce temporal regularisation. The inversion is simple to apply, requiring only the ratio between regularisation parameters and the desired level of misfit from the user. The algorithm is tested on both synthetic data, and a case study. We find that in the synthetic example our inversion successfully retrieves the main characteristics of the test model and introduces minimal artefacts, even in the presence of significant noise. We also test the effect of changing the ratio of regularisation parameters. In the case-study we produce an easily interpretable model that compares favourably with previous inversions of the synthetic data. We conclude that time-lapse modelling of 1D MT data can be a valuable tool for imaging subsurface change.

3.1 Introduction

In recent years the magnetotelluric method (MT) has been increasingly used as a cost-effective technique for subsurface resistivity monitoring (see Rees et al., 2016b, for a general introduction). As this is a relatively new application for MT, there are few modelling tools available. Peacock et al. (2012) have used differences in MT phase tensors to qualitatively interpret resistivity changes due to fluid injection in an enhanced geothermal system (EGS) at Paralana, South Australia. This technique has also be used by Didana et al. (2017) interpreting change at the Habernero EGS. Recent work includes parameterising resistivity changes as a three-dimensional (3-D) plume structure and inverting using Markov Chain Monte Carlo

(Rosas-Carbajal et al., 2015). Another approach has been to use 1D layer-stripping, where the effect of overlying structures is removed to model the time-varying magnetotelluric responses at depth (Ogaya et al., 2016). Standard MT modelling tools such as inversion can also be adapted, for example Rees et al. (2016b) use cascading two-dimensional (2D) inversions to model a coal-seam gas depressurisation, where the results from the pre-injection inversion are used as a prior-model for the post-injection. An improvement on cascaded inversion has been implemented by Rosas-Carbajal et al. (2012), where differenced 2D MT inversions have achieved high accuracy by subtracting the prior-model response from all data and reducing the data error. High accuracy simultaneous time-lapse 1D MT inversions have also been approximated using 2D MT codes (e.g. Rees et al., 2016b; Didana et al., 2017), however this approximation leads to small errors in the calculation of the forward-model and hence inversion result.

The problem of inverting time-lapse data is fairly well researched in other geophysical methods. Hayley et al. (2011) has compared the various techniques used for inverting electrical resistivity tomography (ERT) time-lapse data. They showed that a simultaneous inversion of time-lapse ERT data gives a superior result compared to independent inversions, cascading inversions or differenced inversions. Their preferred technique was a special case of the 4D algorithm of Kim et al. (2009), which has also successfully been applied to gravity (Karaoulis et al., 2013a), induced polarisation (Karaoulis et al., 2013b), and seismic tomography (Karaoulis et al., 2015). This technique has not yet been presented for magnetotelluric inversion.

We present an implementation of a simple time-lapse algorithm to simultaneously invert 1D MT monitoring data. Firstly, we present the results from synthetic inversions of 1D magnetotelluric data to test the validity of the algorithm. We then present inversions from a case study of data from a coal-seam gas production survey (Rees et al., 2016b).

3.2 Inversion Method

The algorithm is an Occam style inversion (Constable et al., 1987), seeking to find the model which fits the data to a desired level while introducing

the least amount of structure - in this case in both temporal and spatial domains. We seek to find the model ($\mathbf{m}$) which minimises the model roughness ($R(\mathbf{m})$) subject to the constraint that the model misfit ($\chi^2(\mathbf{m})$) is less than a desired target misfit (χ^2_{target}). That is,

$$\underset{\mathbf{m}}{\text{minimize}} \quad R(\mathbf{m})$$

$$\text{subject to} \quad \chi^2(\mathbf{m}) \leq \chi^2_{target}.$$

χ^2 misfit is given by the weighted differences of n 1D MT forward responses ($F(\mathbf{m})$) with the data (d) and data errors (σ):

$$\chi^2 = \sum_{i=1}^{n} \frac{(F(\mathbf{m}) - d)^2}{\sigma^2}. \tag{3.1}$$

It is more convenient to express the misfit in terms of root-mean-square error (RMS), given by

$$RMS = \sqrt{\frac{\chi^2}{n}}. \tag{3.2}$$

The model roughness is the sum of regularisation in both spatial (S) and temporal (T) directions, weighted by a factor β:

$$R(\mathbf{m}) = S(\mathbf{m}) + \beta T(\mathbf{m}). \tag{3.3}$$

The temporal regularisation is given by the squared sum of the model changes in time:

$$T(\mathbf{m}) = \sum_{time=i} (\mathbf{m}_i - \mathbf{m}_{i-1})^2, \tag{3.4}$$

and the spatial regularisation by the sum of second derivatives at each time slice:

$$S(\mathbf{m}) = \sum_{layer=i} (\mathbf{m}_{i-1} - 2\mathbf{m}_i + \mathbf{m}_{i+1})^2. \tag{3.5}$$

The optimisation is performed using the Sequential Least SQuares Programming (SLSQP) algorithm available in `scipy` (Jones et al., 2001), which is based on the algorithm by Kraft and Schnepper (1989). The algorithm does not require any tradeoff parameter between data misfit and smoothing, as this is encapsulated in the target RMS. Data are read into the program using the `mtpy` module (Krieger and Peacock, 2014).

Due to the efficiency of the 1D MT algorithm, which is calculated analytically, the model readily reaches an acceptable misfit within a few iterations. Subsequent iterations reduce the roughness of the model.

3.3 Synthetic study

Numerical data are first used to study the results of the time-lapse MT algorithm. The data are produced from a model of an idealised resistivity change at a single site for a coal-seam gas pump into a shallow aquifer. The initial resistivity is a 10 Ωm halfspace. At survey day 4 a sharp change occurs at a depth of 665 m - 958 m, with the resistivity dropping to 2 Ωm. The input model for the synthetic study is shown in Figure 1.

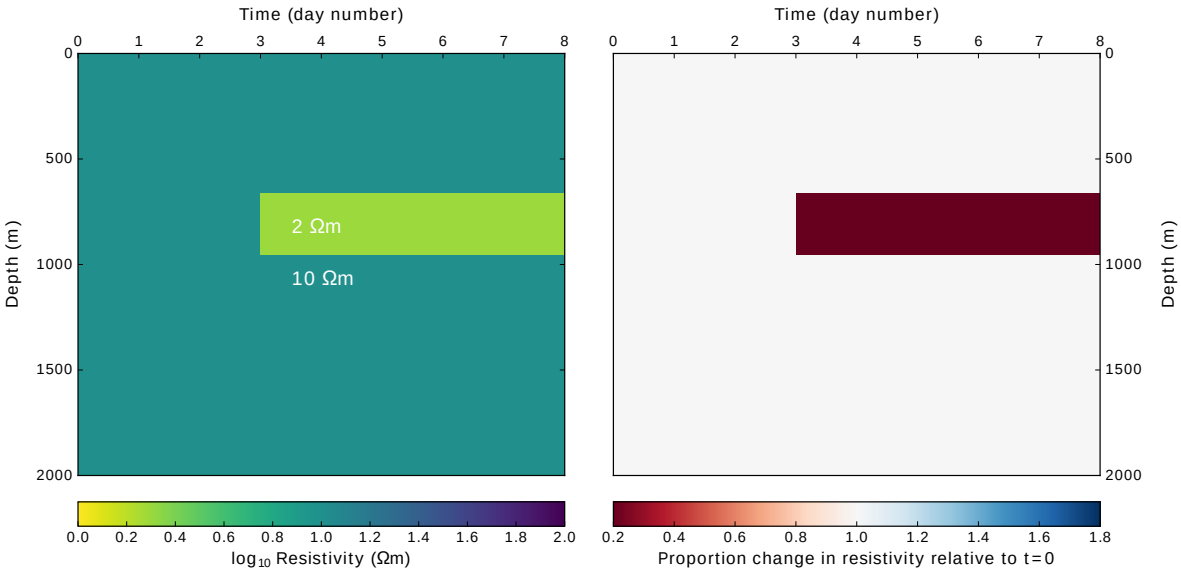

Figure 3.1: The model used to generate input data for the synthetic test study. The first layer is at 30 m depth and all subsequent layers are at a factor of 1.2 greater depths. The model is broken into 10 equal time periods.

Magnetotelluric impedance data are produced at 17 frequencies ranging from 0.14 Hz - 6.26 Hz. The data were computed using the standard analytic 1D forward algorithm. The 1D assumption is valid if the resistivity changes are sufficiently laterally continuous, which we would expect for changes in a lateral coal-bed. The assumption is not perfect however, as 3D effects would be present during the development of the resistivity

front. The data were then contaminated with 5% Gaussian random noise on the impedance tensor, which is a reasonably high estimate of the errors expected in field operations. The data errors were fixed at 5%.

The inversion was conducted with a RMS_{target} of 1.0. Smaller target RMS values result in overfitted models which mapped noise in the data as temporal change, whilst larger target RMS values underfit the change in the data. The β value, which weights the relative importance of temporal and spatial change, was varied between in three separate inversions with values of β = 100, β = 1000 and β = 10000 in order to show the effect of this parameter.

For each inversion the target RMS was achieved within 5 iterations, however additional smoothing iterations were run, with a total of 100 iterations for each inversion.

3.3.1 Results and Discussion

Figure 2 shows the resistivity for three synthetic inversions with varying ratios of temporal smoothing to spatial smoothing. On the left panels the absolute resistivity is shown, and on the right panels the change in resistivity relative to the initial resistivity, is displayed.

The main feature in common to all three figures is the strong decrease in conductivity from day 2-3 onwards. In each inversion the change is diffusively spread out across multiple layers, with the amount of spread dependent on the value for β. The diffusive spreading is most evident with the smallest β value in Figure 2a, and the largest value of β in Figure 2c has the changes confined to almost a single layer.

Compared to the input model, each inversion model underestimates the change in the data. The relative change in the input model is an 80% drop over two layers, whereas the three inversion models have drops in the order of 50% - 60%. The extent of the conductive change is overestimated by the β = 100 model and underestimated by the β = 10000 model, with the middle β = 1000 model providing a fairly accurate image. With increasing β the models also become more rough spatially. In the β = 10000 rough, for example, the model changes from 1 Ωm to 100 Ωm between three layers.

Figure 3.2: Inversion results for synthetic inversions using the parameters a) β = 100, b) β = 1000 and c) β = 10000. The left panels show the final resistivity models for the inversions at each day. The right panels show the proportional difference between the linear resistivities of the models at t_i and t_0. Dotted lines show the extent of the resistivity changes in Figure 1.

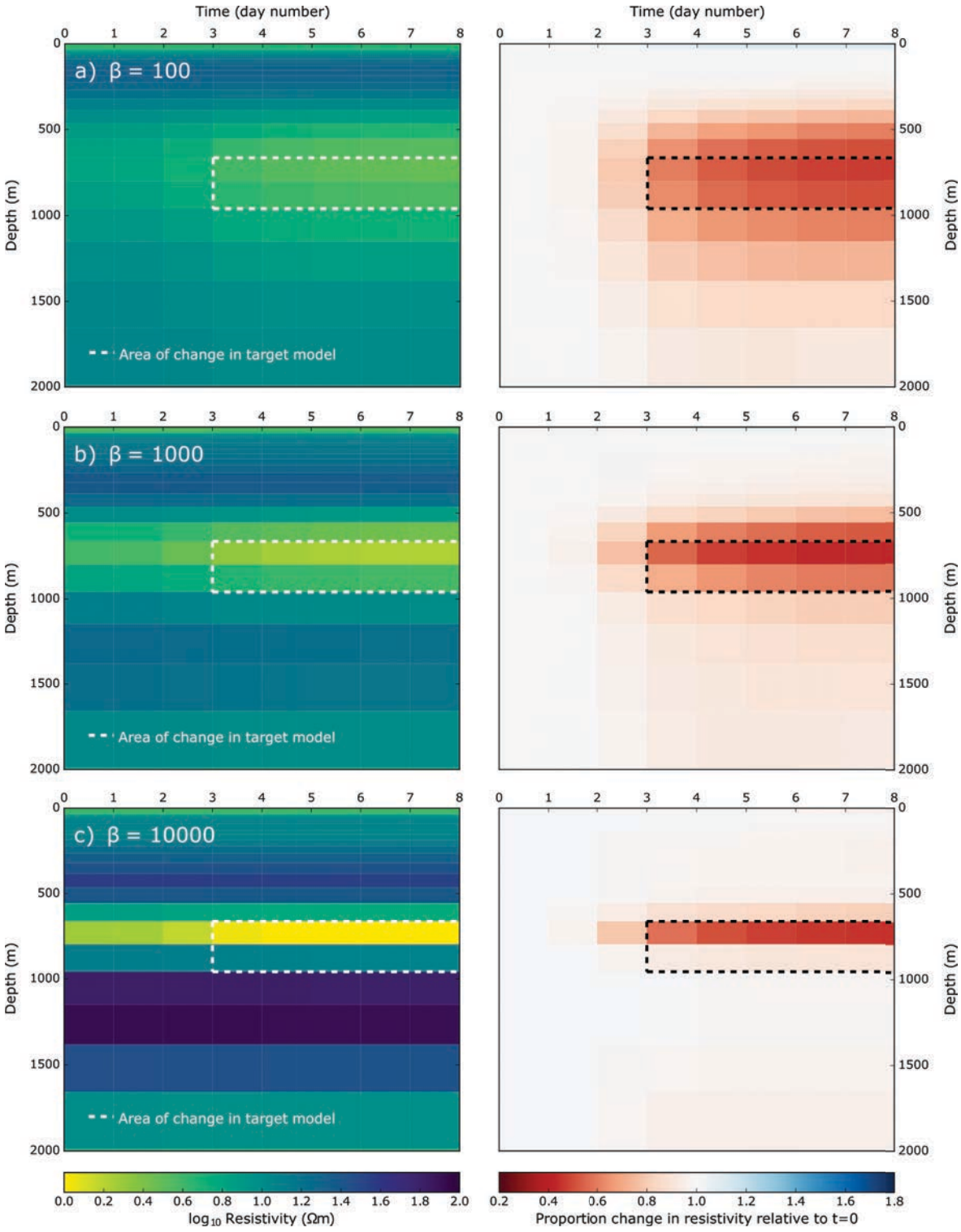

Figure 3.3: Response curves showing the difference between the data fits at t_0 (blue) and the final time slice of synthetic data, t_{10} (green). The data used for inversion are shown as points with error bars attached, and then the smooth curves show the model responses obtained from the inversion for $\beta = 1000$.

Example model fit curves from the first and final days are overlain in Figure 3 for the $\beta = 1000$ inversion, plotted alongside the synthetic data used in the tests. The curves show that a good fit to the data is achieved at convergence. They also highlight the substantial change in the synthetic data due to the resistivity changes.

The inversions were reasonably accurate in terms of the absolute resistivities and the relative changes in resistivity. The smaller β value, however, spread the changes out over a large spatial area. The algorithm will preference changes over as many layers as possible for two reasons. Firstly, the temporal roughness is calculated on the square of the change in any one layer, which means that it is preferable to spread changes over multiple layers if possible. Secondly, and more importantly, the entire model is smoothed in the spatial dimension. The lowest β parameter resulted in a less concentrated area of change, with smaller changes spread over a greater area. The highest β weighting gave a stronger concentration of change and a smaller overall change in resistivity. The model, however, became rougher in the spatial dimension leading to improbable absolute resistivities.

It is worth noting that the inverted models accurately pick the time when the resistivity change occurs. There is a slight change in resistivity before the inversion, however this is minimal and these changes unavoidable as the small data misfit which it introduces is offset by the improved temporal smoothing. If *a priori* information exists about the time t_c when a resistivity change occurs it can be included by restricting all temporal change for $t < t_c$.

3.4 Case study

An example from a coal-seam gas field is used to further examine the algorithm. A coal-seam gas monitoring experiment is ideal to test the 1D time-lapse MT algorithm, as changes occur at shallow depths where the MT response is better approximated by a 1D algorithm, and we expect any changes to be laterally continuous through the coal seam.

We use time-lapse MT data from a 2013/2014 survey of a coal-seam gas field in the Surat Basin, Queensland, Australia (Rees et al., 2016b). Data for this survey were collected along two lines near several production wells. The wells extracted both water and gas from depths of 400 m - 700 m, and a resulting resistivity change was observed by Rees et al. (2016b) in differenced MT inversions.

A schematic of the survey design is shown in Figure 4, which also features production well locations. Site 105 was used for our inversion, which is the same site used for the 1D inversions in Rees et al. (2016b). Data were used from 10 days between January 23, 2014 and February 19, 2014. The site is the near to the most active production well and had reasonable data quality during these days. MT data are taken from the YX mode as these are of higher quality than the XY mode data.

Identical to the synthetic study, the data have 17 frequencies ranging from 0.14 Hz - 6.26 Hz. We invert these data using the presented time-lapse algorithm. We use parameters $RMS_{target} = 1.0$, $\beta = 1000$ and an error floor of 2%. The inversion is run for 100 iterations and completed within 20 minutes on a 4 core machine.

3.4.1 Results and Discussion

The inversion resistivity model is shown in Figure 5, with the resistivity differences shown in the same figure. A slight reduction in resistivity between 500 m and 700 m begins on day 9 and slowly increase with time, culminating with a drop of roughly 15 % in resistivity by the final day. There is also a slight ($\approx$ 5%) increase in resistivity in the area between 100 m and 500 m. Finally, there is also a reduction in resistivity at depth,

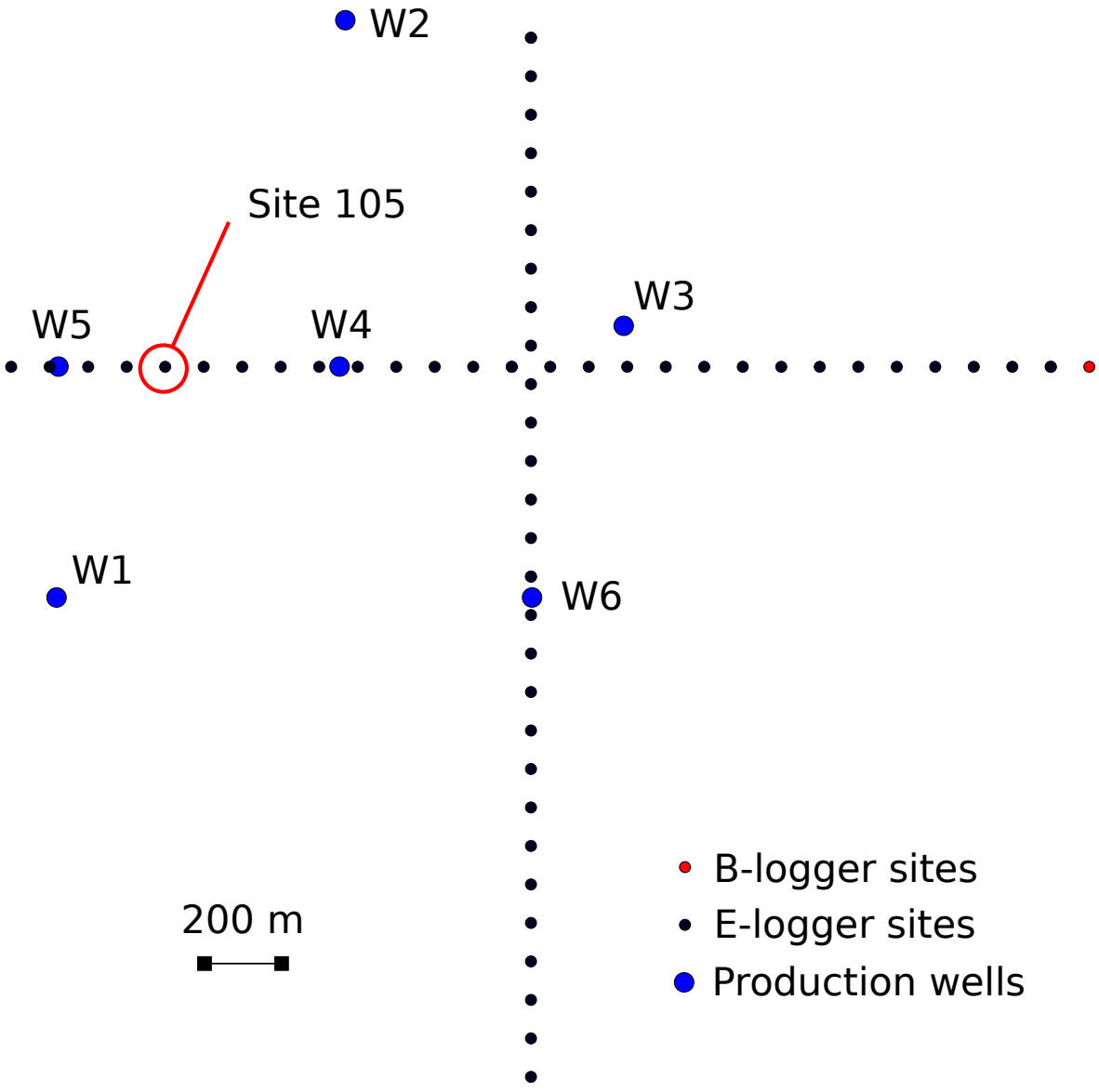

Figure 3.4: A schematic of the site used for inversion in the case-study. Site 105 (highlighted) was one of the closest sites to the active pump well W4, and the most likely candidate for resistivity change. Electric field data were collected at each e-logger site, and MT responses calculated using the magnetic field data at the single b-logger site.

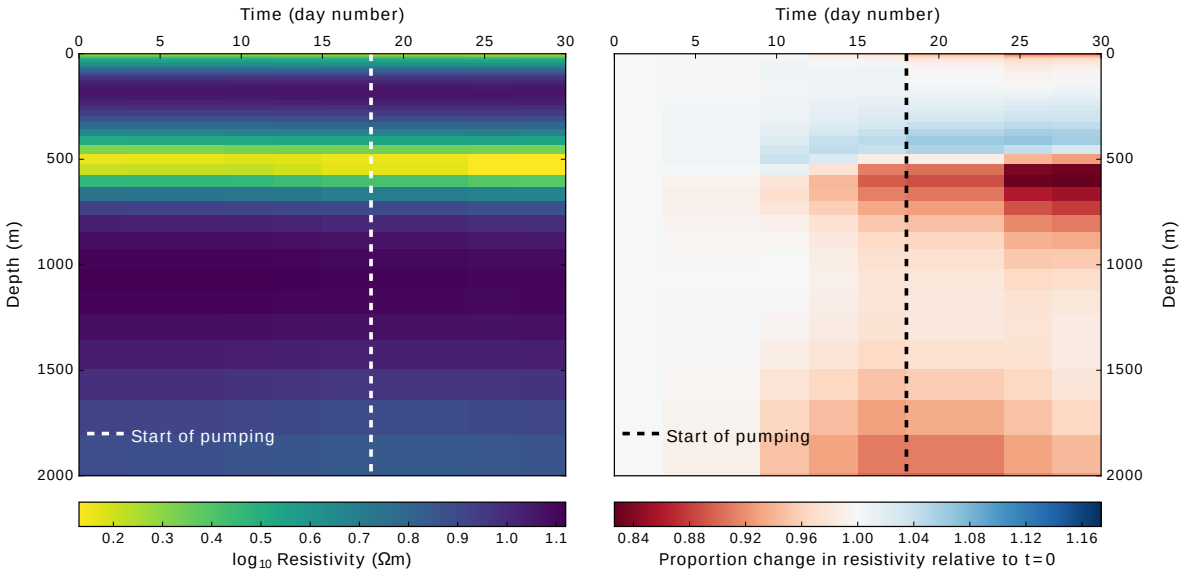

Figure 3.5: a) The final resistivity model for the case-study inversion at each day. b) The proportional difference between the model at t_i and t_0.

however this is smaller than the resistivity drop in the higher zone. Notably this would be near the penetration depth of the data and would be less constrained than the other areas of the model. The entire model resistivity is between 1.25 Ω.m and 16 Ω.m. When considering the commencement date of pumping, shown as a dashed line in Figure 5, we see that the resistivity changes occur slightly before the commencement of dewatering. In the synthetic model, temporal changes were smoothed to dates before the change in the target model, and we would suggest that the same phenomenon is occurring here.

Looking at the fit of the response curves in Figure 6, we see that the model slightly underfits some of the changes in the MT data. This is expected as the algorithm introduces the minimum change required to fit the data. This would place the estimations of 15% change as conservative, and shows the importance of obtaining high quality data, as tighter error bars would lead to stronger changes in resistivity. Notably there is a significant dip in the model fit curves between the first (in blue) and final (in green) time slices.

The most important area to consider for interpretation is the area of the resistivity decrease between 500 m and 700 m, as the other resistivity changes in the model are much smaller and should not be overinterpreted.

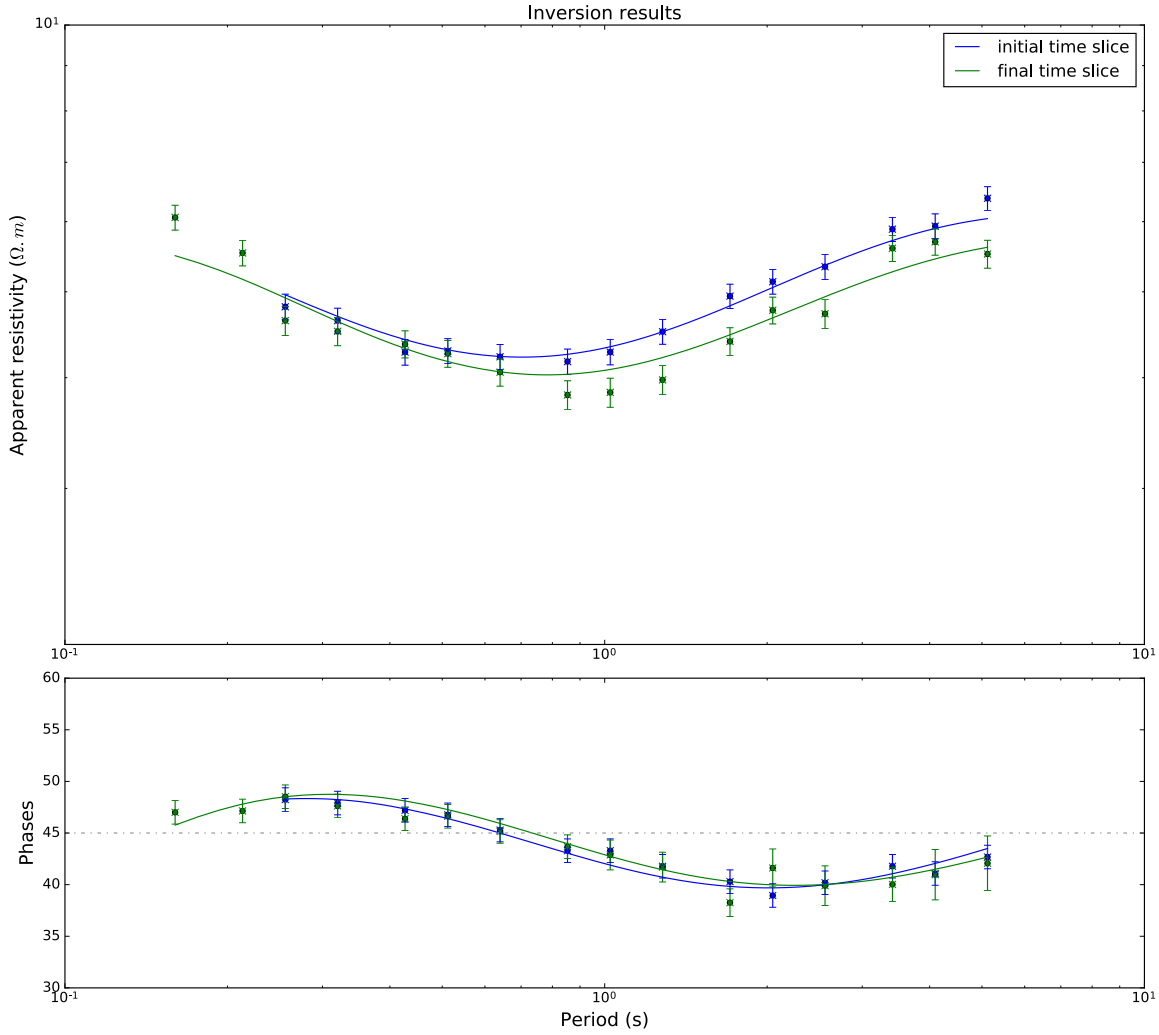

Figure 3.6: Response curves showing the difference between the data fits on January 23, 2014 (blue) and the final day of inversion, February 19, 2014 (green). The data used for inversion are shown as points with error bars attached, and then the smooth curves show the model responses obtained from inversion.

There seems to be a strong link between the start of the resistivity change and the dewatering event at W4. There is no gas extraction during this time. Hence, our new model would agree with the interpretation in Rees et al. (2016b) that the changes are due to the increased permeability of dewatered coal-seams, potentially due to the reduction of the coal matrix due to the release of gas is released from the matrix.

Unlike the synthetic example, the changes in the model were slow, with a build up to the maximum change in the final day included in the inversion. This would lead us to believe that resistivity changes resulted from coal-seam gas pumping are related not just to the rate of pumping, but also the amount of material that has been removed from the pores.

3.5 Conclusion

We presented a simple time-lapse inversion of 1D MT data, which we tested on synthetic data and case-study data. Both inversions resulted in defined areas of resistivity change with few artefacts in the models. The synthetic inversion obtained resistivity changes which slightly underestimated the changes in the input model, however it accurately retrieved spatial and temporal locations, as well as absolute resistivities. We investigated the effect of the weighting parameter β between spatial and temporal smoothing, and showed that there is a tradeoff between models which are overly sharp spatially and models which overestimate the extent of the resistivity change. We also showed results from applying the algorithm to a case-study with coal-seam gas MT data. The inversion resulted in small changes in the area of the coal-seams.

Compared to other modelling techniques our algorithm has several advantages. Full 3D modelling is highly computationally intensive. A 1D approximation is simple to compute, which leads to rapidly converging model. This allows the modeller to trial several parameters and select the best model. It also allows a large amount of data to be inverted at once - it is feasible to invert data from each day for several months, which would result in an extremely large model space in 3D. Compared to independent inversions, or time-lapse inversions using 2D MT codes as approximations,

it has previously been shown that simultaneous inversion results in a model with fewer artefacts and stronger constraints on areas of change.

One of the disadvantages of the technique is that it is unable to fit any 2D or 3D aspects of the data. Further work could expand the technique into higher spatial dimensions, or to deal with anisotropy. It would be useful to extend the inversion into 3D space and incorporate all data into the inversion, however the long processing times of higher-dimensional MT forward codes would make the resulting code extremely computationally-intensive.

Acknowledgements

We would like to thank QGC Pty Limited for their funding of the data collection, as well The Australian Geophysics Observing System and AuScope for providing the instrumentation. We thank S. Rana, G. Hicks, P. McKelvey and A. Smart for their help with field logistics, and S. Schnaidt, G. Boren, A. Tullis, B. Warren, S. MacMillan and C. Matthews for their assistance in field work. Finally, S. Carter and O. Putland and thanked for their help with data processing.

CHAPTER

FOUR

PROBABILISTIC MAGNETOTELLURIC INVERSION WITH ADAPTIVE REGULARISATION USING THE NO-U-TURNS SAMPLER

Conway D.[1], Simpson J.[2], Didana Y.[1], Rugari J.[1]., and Heinson G.[1]

[1]School of Physical Sciences,
University of Adelaide, Adelaide, Australia
[2]Geological Survey of Queensland,
Department of Natural Resources and Mines, Brisbane, Australia

Published in *Pure and Applied Geophysics*
https://doi.org/10.1007/s00024-018-1870-5

Statement of Authorship

Title of Paper	Probabilistic magnetotelluric inversion with adaptive regularisation using the No-U-Turns sampler
Publication Status	Published
Publication Details	Conway, D., Simpson, J., Didana, Y., Rugari, J., Heinson, G. (2018), Probabilistic magnetotelluric inversion with adaptive regularisation using the No-U-Turns sampler, Pure and Applied Geophysics, 1-17

Principal Author

Name	Dennis Conway
Contribution	Conception of ideas, programming, modelling, visualisation, manuscript writing and academic correspondence
Overall percentage	80%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.

Signature _____ Date 10/5/18

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

(i) the candidate's stated contribution to the publication is accurate (as detailed above);

(ii) permission is granted for the candidate in include the publication in the thesis; and

(iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name	Janelle Simpson
Contribution	Provision of test data and feedback on modelling

Signature _____ Date 21/5/18

Name	Yohannes Didana
Contribution	Assistance with manuscript writing and edits

Signature _____ Date 10/5/18

Name	Joseph Rugari
Contribution	Assistance with conception of ideas

Signature _____ Date 15/05/18

Name	Graham Heinson
Contribution	Feedback on manuscript and edits, assistance with conception

Signature _____ Date 10/5/18

Abstract

We present the first inversion of magnetotelluric (MT) data using a Hamiltonian Monte Carlo algorithm. The inversion of MT data is an underdetermined problem which leads to an ensemble of feasible models for a given dataset. A standard approach in MT inversion is to perform a deterministic search for the single solution which is maximally smooth for a given data-fit threshold. An alternative approach is to use Markov Chain Monte Carlo (MCMC) methods, which have been used in MT inversion to explore the entire solution space and produce a suite of likely models. This approach has the advantage of assigning confidence to resistivity models, leading to better geological interpretations. Recent advances in MCMC techniques include the No-U-Turns Sampler (NUTS), an efficient and rapidly converging method which is based on Hamiltonian Monte Carlo. We have implemented a 1D MT inversion which uses the NUTS algorithm. Our model includes a fixed number of layers of variable thickness and resistivity, as well as probabilistic smoothing constraints which allow sharp and smooth transitions. We present the results of a synthetic study and show the accuracy of the technique, as well as the fast convergence, independence of starting models, and sampling efficiency. Finally we test our technique on MT data collected from a site in Boulia, Queensland, Australia to show its utility in geological interpretation and ability to provide probabilistic estimates of features such as depth to basement.

4.1 Introduction

The aim of magnetotelluric (MT) inversion is to model the subsurface physical properties of the earth based on observations taken at the surface. Geophysical inversion in general is a problem with two halves: an existence problem, where the data are fit by a single earth model which best fits the data, and a uniqueness problem, where all models which fit the data are sought (Backus, 1988). In a Bayesian inversion, the second side of this problem is explored with a model which treats each parameter in the inversion probabilistically, and seeks to find their probability den-

sity functions (PDFs) (Tarantola, 2005; Tarits et al., 1994). This task is typically achieved with Markov Chain Monte Carlo (MCMC) methods.

The inversion of MT data is an under-determined problem and requires some form of regularisation (Grandis et al., 1999; Schott et al., 1999). In Bayesian inversion a prior distribution is given for the inversion parameters, which acts to regularise the inversion. Similar to deterministic inversion, the regularisation may take different forms. Common techniques include restricting the number of layers (Guo et al., 2011); calculating the roughness and penalising rougher models with lower probabilities (Grandis et al., 1999; Rosas-Carbajal et al., 2013); or having a variable number of parameters and a penalisation for more complex models (Minsley, 2011; Mandolesi et al., 2018).

Bayesian inversion of 1D MT data has been successfully applied using MCMC algorithms, however several challenges exist which make 2D and 3D inversions computationally difficult. These include the reduced speed of the forward algorithms, the slower convergence of the MCMC algorithm with increased number of parameters, and the high covariance of parameters. One solution is to reduce the number of parameters in inversion, e.g. Chen et al. (2012) who used sharp inversions in 2D with a fixed number of nodes. A sharp inversion has the added advantage of avoiding model regularisation through smoothness constraints, however lacks the expressivity of a pixel-based inversion. The first full pixel-based inversion has been implemented for 2D radio magnetotelluric data by Rosas-Carbajal et al. (2013) using the MT-DREAM(ZS) algorithm (Laloy and Vrugt, 2012). The authors show that the algorithm can successfully invert the RMT data, with particularly good results when combined with electrical resistivity tomography in a joint inversion. The inversion is computationally expensive however, requiring up to 130000 iterations to reach convergence for 228 model cells, even with prior constraints on resistivity ranges and model roughness. Bayesian inversion has also been used in reduced parameter 3D inversions by Rosas-Carbajal et al. (2015) who used MCMC to invert 7 parameters describing time-lapse changes in resistivity. For a pixel-based 3D MT inversion the process becomes more difficult, as Grandis (2006) concludes from preliminary studies that a very coarse model mesh would be required to keep computation times reasonable.

One solution to these issues is the use of gradient-based MCMC algorithms based on Hamiltonian mechanics, known as Hamiltonian Monte Carlo (HMC) or Hydrid Monte Carlo. These methods operate by taking a series of steps informed by first-order gradient information (Hoffman and Gelman, 2014). This makes them much more efficient than standard MCMC sampling methods such as random-walk Metropolis (Metropolis et al., 1953) and Gibbs sampling (Geman and Geman, 1984), with the cost per independent sample from a target distribution of dimensionality D reduced from $\mathcal{O}(D^2)$ to $\mathcal{O}(D^{5/4})$ (Creutz, 1988). This is clearly advantageous in the case of multi-dimensional MT MCMC inversion, where the number of parameters can run into the thousands.

One of the main obstacles of the HMC algorithms has been the requirement to specify at least two additional parameters to the algorithm, which can have serious effects on sampling efficiency if poorly chosen, and require additional runs to tune correctly. This problem has been solved by the No-U-Turns algorithm (Hoffman and Gelman, 2014), an algorithm based on HMC which eliminates the need for specifying tuning parameters. The NUTS algorithm is under active development and has been successfully used in a number of studies in various fields. Some examples include a study of energy consumption from buildings (Chong et al., 2017), ecological studies investigating populations of various plants and animals (Monnahan et al., 2017) and a supernova study which modelled 9176 parameters (Sanders et al., 2015).

We present the first implementation of an MT inversion using the NUTS algorithm. The inversion uses a novel adaptive regularisation scheme which preferences models with smoothly changing resistivities while allowing sharp changes. Firstly, we show the results of a 1D MCMC inversion using NUTS on synthetic data, and highlight the sampling efficiency and model accuracy. Secondly, we apply the same inversion to MT data obtained from Boulia, Queensland, Australia and show the practical use of the algorithm in geological interpretation. The results show a promising Bayesian inversion which can be upscaled to 2D and 3D MT problems.

4.2 Bayesian Model

4.2.1 Bayesian inversion and the NUTS sampler

The goal of a Bayesian geophysical inversion is to find the probability distribution of model parameters ($\mathbf{m}$) given fixed geophysical data ($\mathbf{d}$), known as the posterior probability distribution. Bayes' theorem states that this distribution, $P(\mathbf{m}|\mathbf{d})$, is proportional to the likelihood of $\mathbf{m}$ given $\mathbf{d}$, $L(\mathbf{m}|\mathbf{d})$, multiplied by the prior probability of $\mathbf{m}$, $P(\mathbf{m})$. That is,

$$P(\mathbf{m}|\mathbf{d}) \propto L(\mathbf{m}|\mathbf{d})P(\mathbf{m}). \tag{4.1}$$

Although we cannot tractably solve for the posterior probability distribution directly, the MCMC method allows us to sample $P(\mathbf{m}|\mathbf{d})$. From these samples we may build a probability distribution.

In MCMC, a Markov chain is constructed that has its equilibrium distribution as the posterior distribution. The way in which this is achieved varies between sampling techniques. In Metropolis-Hastings sampling, a random walk proposal generates new samples from a Gaussian distribution centred on the current sample (Chib and Greenberg, 1995). The new sample is accepted or rejected with probability proportional to the ratio of the new probability density to the previous probability density. This method can be inefficient for high dimensional spaces, where the number of directions to explore increases exponentially, and many may yield low likelihoods.

An alternative approach is the Hamiltonian Monte Carlo method (HMC). In HMC the state of the MCMC sampler is treated as a particle in a Hamiltonian system, with an energy calculated from the potential energy (calculated from the probability density of its position) and a randomised momentum, or kinetic energy (Betancourt, 2017). The system then evolves according to Hamiltonian dynamics, leading to a preference to explore new positions with lower potential energy and hence higher probability density. This results in better convergence and more efficient exploration of the posterior probability distribution.

To simulate the Hamiltonian dynamics in an HMC iteration, the Störmer-Verlet, or "leapfrog", integrator is used. Time in the system is discretely

incremented by a step-size of ϵ for a series of L steps, and the position and momentum of the particle is updated at each time. The final position and momentum of the particle can then be accepted or rejected as with the Metropolis-Hasting algorithm. An HMC proposal, however, has the dual advantages of typically higher acceptance ratios, and therefore more efficient sampling, and also a greater distance from the previous location, leading to better exploration of the model space.

The choice of L and ϵ can be difficult in HMC. If ϵ is too small then the simulation will run slowly with unnecessary calculation, however if it is too large then the Hamiltonian simulation will become inaccurate. A choice of L which is too low will result in poor exploration and the sampler will exhibit random-walk behaviour, however if L is too large then the particle trajectory will begin to loop back around on itself, resulting in both unnecessary calculation and poor exploration. Fortunately the issue of tuning these parameters is solved by the No-U-Turns Sampler (NUTS) (Hoffman and Gelman, 2014). The NUTS algorithm tunes L by preventing the particle from doubling back on itself and tunes ϵ using a dual averaging schemed based on work by Nesterov (2009).

4.2.2 1D magnetotelluric model

We create a model $\boldsymbol{m}$ to probabilistically invert MT data consisting of n frequencies of complex impedance data $\mathbf{Z}_{obs_n}$ and corresponding data errors, σ_n.

The model consists of N layers of variable thicknesses $\boldsymbol{t} \in [10\ m, 1500\ m]$ and log_{10} unconstrained resistivities $\boldsymbol{\rho}$. The thicknesses have been constrained to prevent the sampler becoming stuck in regions with layer sizes tending to infinity or zero. The layered resistivity model is then transformed into n complex MT impedances $\mathbf{Z}_n$ using the Wait recursion formula (Wait, 1962).

To calculate the likelihood function of $\mathbf{m}$ given $\mathbf{d}$, we approximate the distribution of $\mathbf{Z}_{obs_n}$ as two independent normal distributions with means $\mathrm{Re}(\mathbf{Z}_n)$ and $\mathrm{Im}(\mathbf{Z}_n)$ and a shared standard deviation σ_n. That is,

$$L(\mathbf{m}|\mathbf{d}) = \prod_{i=1}^{n} \frac{1}{2\pi\sigma_n} \exp\left(- \frac{\|\mathbf{Z}_n - \mathbf{Z}_{obs_n}\|^2}{2\sigma_n^2} \right). \tag{4.2}$$

We also establish a constraint on the roughness of the model. For each layer i above the basement, we construct a prior on the resistivity ρ_i such that it is normally distributed with expected value ρ_{i-1}. The standard deviation is given by β_i, which is a parameter of the prior distribution, known as a hyper-parameter. It is useful to parameterise the model this way so that the posterior distribution is not overly dependent on the choice of prior. To summarise,

$$P(\rho_i) = \frac{1}{\sqrt{2\pi\beta_i}} \exp\left(-\frac{(\rho_i - \rho_{i-1})^2}{\beta_i^2} \right). \tag{4.3}$$

Notably this differs from a typical smooth inversion in that the degree of roughness is adaptive for each layer. A sharp change between two layers is allowed if supported by the data, otherwise consecutive layers must be similar in resistivity. An exponential prior is given to β_i in the form of

$$P(\beta_i) = \lambda \exp(-\lambda\beta_i), \tag{4.4}$$

with λ set by the user; we have found a value of 0.5 suitable for our models.

The final model is not heavily dependent on λ and there is flexibility in its choice. A good smoothness constraint will allow the model to accommodate sharp changes if they greatly improve the model likelihood, while penalising complex models which do not improve the model likelihood. Similarly the number of layers is not vital, as long as enough layers are chosen to sufficiently express the model complexity. The resistivity of additional layers will converge to the resistivity of the previous layer if they do improve the model fit due to the smoothness constraint.

4.2.3 Sampling

We use `Stan`, a probabilistic programming language used for creating statistical models to implement our model (Carpenter et al., 2016). The model

$\boldsymbol{m} = (\boldsymbol{t}, \boldsymbol{\rho}, \boldsymbol{\beta})$, including the 1D MT forward function, is coded into a Stan file, which is compiled and interfaced by the python module `PyStan`. Magnetotelluric data are read into the model using the `mtpy` python module (Krieger and Peacock, 2014). The user supplies the number of layers N and smoothness parameter λ. The model is compiled by Stan and run with multiple chains using the NUTS algorithm. An initial warm-up phase is run for tuning the NUTS parameters ϵ and L and calculating the mass matrix M, which is analogous to the covariance matrix in Metropolis-Hastings samplers. After warm-up, samples are drawn from the posterior probability distribution. The samples for each parameter are then analysed for their convergence using the Gelman-Rubin $\hat{R}$ statistic, also known as the potential scale reduction statistic (Gelman and Rubin, 1992). If the value of $\hat{R}$ is greater than 1.1 then the chains may not have adequately converged (Kruschke, 2014). In this case it would be recommended to recommence sampling with a longer warm-up phase.

4.3 Synthetic study

We conduct a synthetic study to test the effectiveness of our 1D MT inversion model with adaptive regularisation. We produce synthetic data from a 1D vertical resistivity profile shown in Figure 1. The profile is composed of a 300 m top-layer of 10^2 Ωm, a second 100 m layer of 10^0 Ωm and then a linear transition down to 10^4 Ωm at 2000 m. Magnetotelluric impedance data are generated at 32 logarithmically spaced frequencies from 0.01 Hz to 100 Hz. Gaussian distributed noise is added to the complex-valued data at each frequency to simulate errors, with a standard deviation of 5% of the absolute value of the data. The data-errors σ are also set at 5% of the absolute value.

Four different inversion models are run. The first three investigate the effect of varying the number of model layers. These use the NUTS sampler with N set to 3, 4, 5 for models A, B, and C, respectively. The final model, D, is to test the efficiency of the NUTS algorithm. It has N set to 4 but uses the MT-DREAM(ZS) sampler. In each model λ is held constant at 0.5. Each of the three NUTS models are run in three independent chains with random initialisations. The 4 layer model was the easiest to

sample from, and satisfactorily converged after 500 warm-up samples. The 3 and 5 layer models did not converge after 500 samples so were run for 15000 warm-up samples each to ensure convergence. After warm-up, 500 samples were drawn from each model. The total running time for the 4 layer model was 9 minutes running on 3 cores. Model D was sampled using the PyDREAM implementation of MT-DREAM(ZS) with 10 chains run in parallel (Shockley et al., 2017). At each interval of 5000 samples the second half of the samples were analysed using the Gelman-Rubin statistic and convergence was declared once this measure was as good as, or better, than achieved with the NUTS algorithm with the same number of layers. This occurred after 50000 samples. Total run-time was 1 hour 57 minutes running on 10 cores. A summary comparison between the Model B and Model D inversions are shown in Table 1. Traces from the three chains in Model B are shown in Figure 2 for the parameters ρ_1 and ρ_4

The samples from each inversion are then transformed from layer thickness and resistivity into samples of resistivity as a function of depth (Z), $\rho(Z)$, by calculating the resistivity at one-metre intervals for each sample. A kernel-density estimator is then used at each depth interval to find the resistivity probability density (Scott, 1979). This is shown in Figure 1 alongside the original resistivity profile which we are inverting for.

A similar process is used to analyse the distribution of the MT data. Impedance data $\boldsymbol{Z}$ from each sample are taken and transformed into apparent resistivity (ρ_a) and phase (ϕ) at each frequency f using the formulae

$$\rho_a = \frac{\|\boldsymbol{Z}\|^2}{\mu\omega}, \quad \text{and} \quad \phi = \arctan\frac{\mathrm{Im}(\boldsymbol{Z})}{\mathrm{Re}(\boldsymbol{Z})}, \quad (\omega = 2\pi f), \tag{4.5}$$

where the magnetic permeability, μ, is approximated by the permeability of free space μ_0. These samples are used to generate probability distributions using a kernel-density estimator. The results are shown in Figure 3 for Model B.

4.3.1 Results and discussion

We will focus our analysis on Model B, which appears to provide the best fit to the data. Model A and Model C will be used to investigate the

	mean	mean_{se}	sd	2.5%	50%	97.5%	n_{eff}	$\hat{R}$
NUTS sampler - Model B								
ρ_1	2.00	3.4e-3	0.07	1.89	1.99	2.16	421	1.00
ρ_2	-0.11	0.02	0.27	-0.73	-0.06	0.29	255	1.00
ρ_3	3.57	0.11	1.40	0.58	3.71	6.54	170	1.02
ρ_4	3.96	3.2e-3	0.09	3.80	3.96	4.15	793	1.00
t_1	306	0.88	15.4	274	307	333	302	1.00
t_2	90.1	2.69	47.1	17.7	85.9	195	306	1.01
t_3	682	17.6	448	34.0	635	1451	650	1.01
β_1	2.65	0.05	1.55	0.91	2.27	7.003	1002	1.00
β_2	2.81	0.09	2.05	0.90	3.19	8.48	493	1.01
β_3	1.69	0.08	1.67	0.06	1.19	6.32	403	1.01
DREAM sampler - Model D								
ρ_1	1.98	2.1e-3	0.09	1.96	1.98	2.00	2074	1.00
ρ_2	-0.51	7.9e-3	0.22	-0.88	-0.50	-0.10	736	1.00
ρ_3	3.07	0.03	1.20	1.50	3.04	5.87	1940	1.00
ρ_4	3.68	5.9e-3	0.35	3.34	3.66	4.10	3554	1.00
t_1	355	0.23	6.02	344	356	362	731	1.00
t_2	29.5	0.87	22.1	10.9	25.6	64.7	649	1.00
t_3	822	7.60	442	52.4	866	1470	3201	1.00
β_1	2.98	3.6e-2	1.85	1.03	2.49	7.77	2602	1.00
β_2	3.56	3.8e-2	1.90	1.16	3.12	8.53	2549	1.00
β_3	1.75	3.1e-2	1.51	0.06	1.41	5.74	2311	1.00

Table 4.1: A comparison summary of the sampled parameters from the synthetic probabilistic inversions of the 4 layer model with NUTS sampler (Model B) and 4 layer model with MT-DREAM(ZS) sampler (Model D). Subscripts indicate increasing layer number. Summaries included are the mean of the distribution; standard error of the mean; standard deviation; percentiles at 2.5%, 50%, 97.5% (the median and the bounds for the 95% confidence interval); effective sample size estimate; $\hat{R}$, the potential scale reduction statistic.

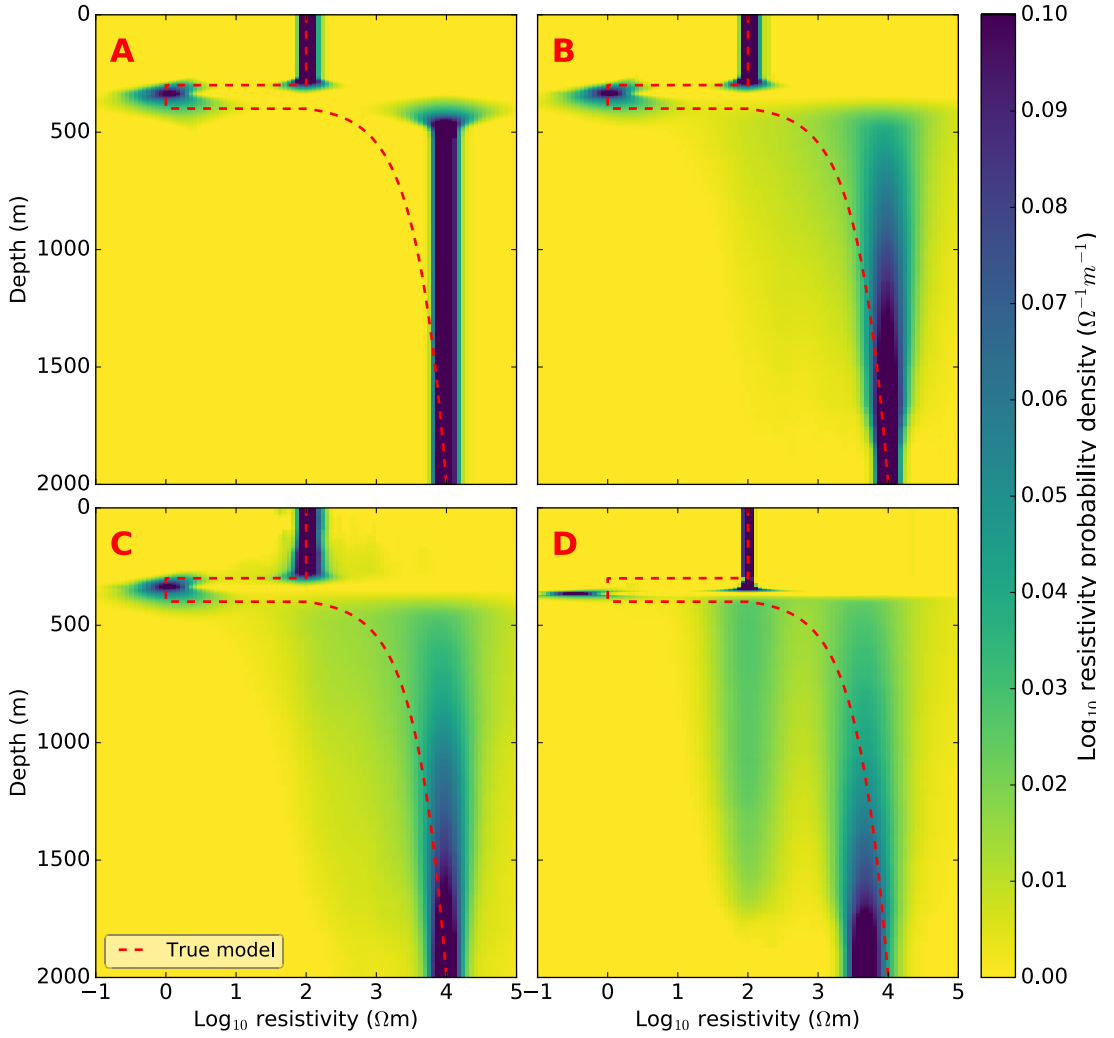

Figure 4.1: Probability densities generated from the four synthetic probabilistic inversions, each generated using 1500 samples. Shown are four different models; A. NUTS sampler, 3 layers B. NUTS sampler, 4 layers, C. NUTS sampler, 5 layers and D. MT-DREAM(ZS) sampler, 4 layers. The dotted red line shows the resistivity profile used to generate the synthetic data. Probability densities are with respect to log_{10} resistivity. Densities have been clipped at 0.1 $\Omega^{-1}m^{-1}$ to allow better visualisation of the lower density areas.

effect of changing the number of model layers, and Model D will provide a reference for the efficiency of the NUTS algorithm.

The traces of MCMC samples in Figure 2 for the variables ρ_1 and ρ_4 show that the parameters have converged to similar distributions across the three chains in Model B. If the chains were noticeably different it would indicate poor convergence. This visual analysis is supported by the $\hat{R}$ values in Table 1 which are very close to 1.00. The $\hat{R}$ value is a measure of

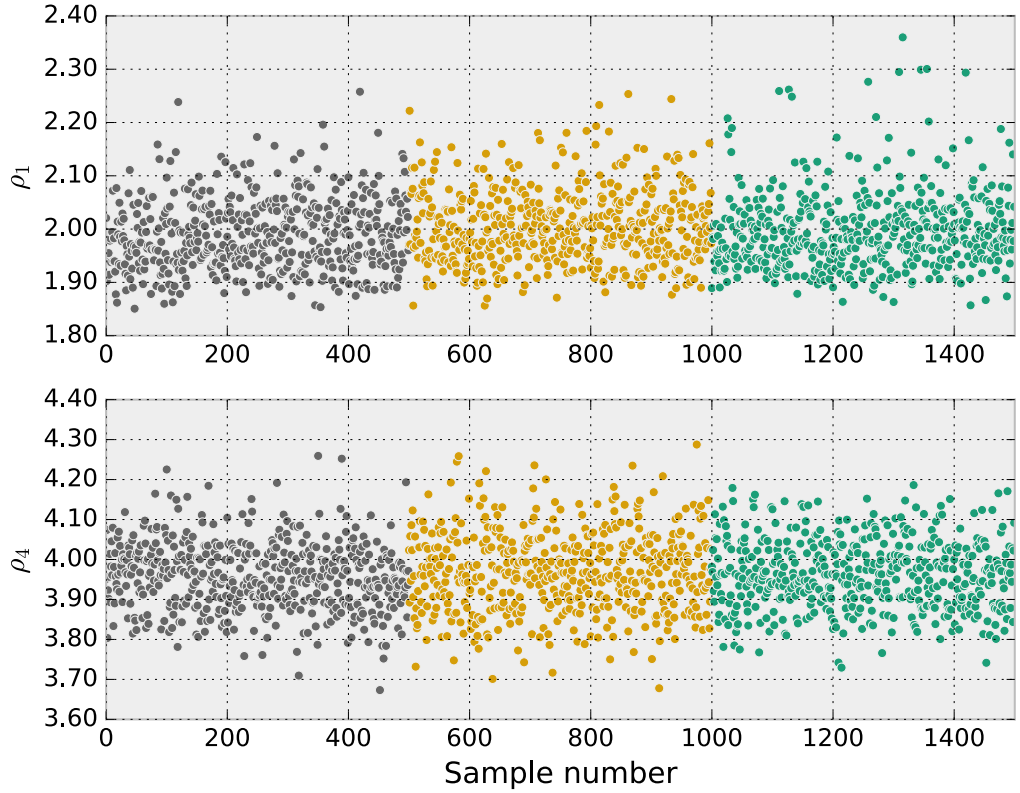

Figure 4.2: Parameter traces from the synthetic inversion for Model B, showing the sampled log_{10} resistivities of the top layer, ρ_1, and bottom layer ρ_4. The traces are split into their three chains by colour. Each chain is independent and is given a different random initialisation. Similar distributions between chains is a sign of convergence.

how similar the separate chains are, with an ideal value of 1.00 for chains which have converged to the same distribution. The fact that several of the variables have $\hat{R}$ values > 1.00 shows how difficult it is to sample a non-linear MT inversion model, however the maximum value of $\hat{R}$ = 1.02 is small enough that we can still be confident of convergence.

The $\boldsymbol{\rho}$ values for Model B have converged to distributions with means very close to the actual resistivity profile. The resistivity values for ρ_1, ρ_2 and ρ_4 have means and standard deviations of $10^{2.00\pm0.07}$ $\Omega.m$, $10^{-0.11\pm0.27}$ $\Omega.m$ and $10^{3.96\pm0.09}$ $\Omega.m$ with target values of $10^{2.0}$ $\Omega.m$, 10^{0} $\Omega.m$ and 10^{4} $\Omega.m$ respectively. The layer ρ_3 is a smooth transition between ρ_2 and ρ_4 in the target profile, and as such has no fixed target value.

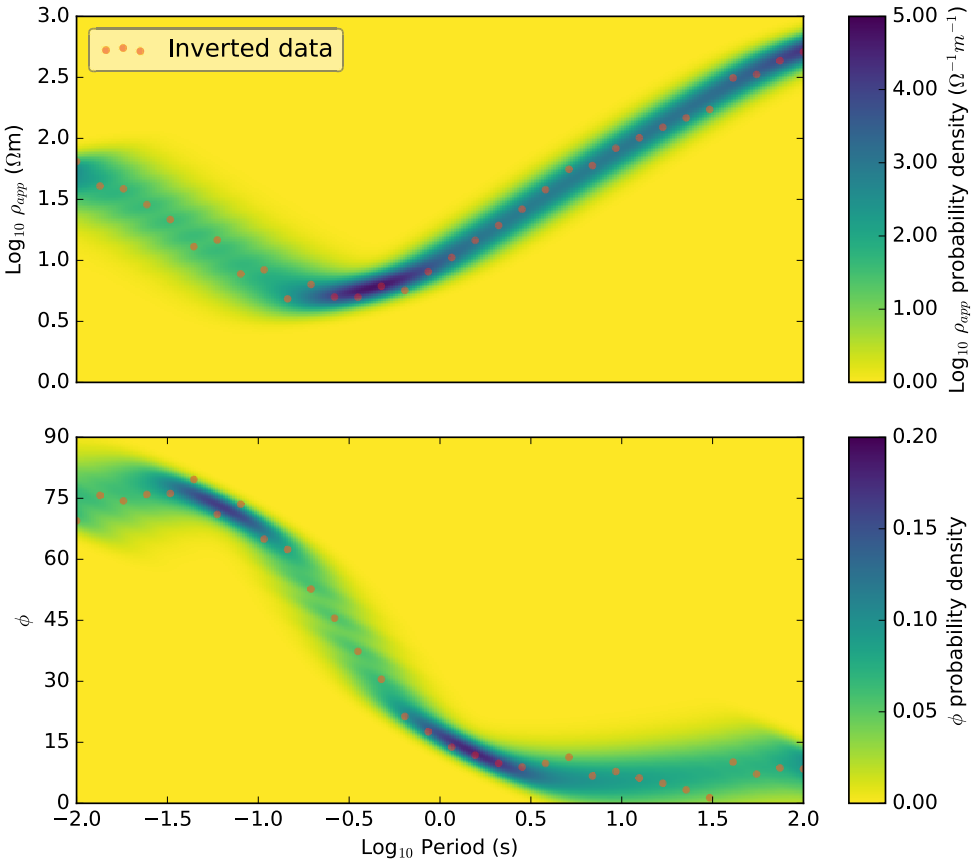

Figure 4.3: Probability distributions for response curves from the synthetic probabilistic inversion of Model B. Distributions are period independent; the top plot shows probability per $log_{10}(\rho_{app})$, the bottom plot probability per phase angle (unit-less). Overlain in red are the data points used for the inversion.

The $\boldsymbol{t}$ values given by Model B are also within the ranges of the target model. t_1 and t_2 have distributions of 307 ± 15 m and 88 ± 45 m, with target values of 300 m and 100 m respectively.

The $\boldsymbol{\beta}$ values in the model can serve as a proxy for the probability of a sharp transition in resistivity. The values of β_1 and β_2 are large enough that they do not introduce much smoothing. This allows for sharp changes between the layers ρ_1 and ρ_2, as well as between ρ_2 and ρ_3. Notably the 2.5 percentile values for β_1 and β_2 are 0.91 $\Omega.m$ and 0.90 $\Omega.m$, which indicates that there is a low probability that the layers ρ_1 and ρ_2 have the same resistivities. In contrast, the β_3 2.5 percentile is 0.08 $\Omega.m$, which

means that there exist models where ρ_3 and ρ_4 are very close, i.e. a three-layer model. The reasonably high mean of 1.73 $\Omega.m$ indicate however that there are 4-layer models which can better fit the data. The end distribution is a continuous mixture of the two cases.

These results are shown graphically in Figure 1. The inversion accurately maps the top two layers in the profile, with a strong degree of confidence in the boundary location. The second layer is also mapped correctly in thickness and resistivity. The inversion has a large degree of uncertainty underneath the second layer. Here the model can either fit with the basement resistivity immediately, or with a better fitting but more complex model which smoothly changes. Once again, both possibilities have a non-negligible probability, however the three-layer model has the higher probability.

Figure 3 shows the distribution of apparent resistivity and phase for the 1500 samples. Although each data point is contaminated with an equal magnitude of noise, the model clearly has more confidence in the apparent resistivity and phase at certain periods. The apparent resistivity is most tightly constrained near its minima between 0.1 s and 1 s, as well as at 100 s, and the phase has probability density peaks between 0.03 s and 0.1 s as well as between 1 s and 3 s. These distributions could be used to export new data values and error values as a form of 1D data smoothing.

Comparing Model B with Model A in Figure 1, we can see that the 3-layer Model A does not have enough flexibility to model the smooth change between the second layer and the basement. It still manages to provide accurate resistivites for the top two layers and basement, however. It is not surprising that a 3-layer model can provide a good fit to the model, as the maximum a posteriori model from the 4-layered Model B is a 3-layered model. The probability densities in Figure 1 from Model A however have a negligible probability in the area where there is a smooth change from the second layer to basement, so it is not as accurate as Model B for modelling this dataset.

The 5-layer inversion in Model C however provides a resistivity probability density in Figure 1 very similar to Model B. The top two layers are identical between the two models, however there is a slight difference in the smooth

transition area, with Model C favouring slightly more structure. Due to the fact that the regularisation is applied on a layer by layer case rather than looking at the overall structure, increasing the number of layers would allow for slightly rougher models. This could be mitigated by increasing the λ parameter for models with more layers. The results from Model A and Model D show that the final model is largely independent of the number of layers in the model, as long as there are enough layers to introduce sufficient structure into the model. If in doubt, it would be reasonable to err on the side of too many layers, as doing so appears to have few deleterious effects apart from a reduction in sampling efficiency. In 2D and 3D smooth inversions the layer geometries are fixed throughout the inversion, so choosing the number of layers is not an issue.

The results from Model D have a much greater contrast with those of Model B. Despite the excellent $\hat{R}$ convergence statistics for Model D in Table 1, we can see that the distributions are not as accurate as those in Model B. The upper limit of the 95% confidence interval for ρ_1 is exactly the true value of $10^{2.0}$ $\Omega.m$, however the true values of ρ_2, t_1 and t_2 all fall outside of their 95% confidence intervals. The basement resistivity ρ_4 is more accurate, however. In the resistivity density plot for Model D shown in Figure 1, we see probability densities in small areas for the top two layers. This is characteristic of a poor exploration of the model a posteriori, which is common for models with highly correlated parameters. We also see that in the smooth transition zone there appears to be a bimodal distribution between a 3-layer model and a 4-layer model. This characteristic is also present in Model B, however in there is a smooth reduction is probability with increasing model complexity rather than two strong modes. Coupled with the faster compute time of Model B, we would conclude from these results that NUTS is an excellent choice to sample MT resistivity structures in our 1D model. We would further expect the NUTS algorithm to perform well in 2D and 3D spaces, as the sampling efficiency scales well with increased dimensionality. However, further investigation is required to compare it with the MT-DREAM(ZS) algorithm in 2D and 3D, as the MT-DREAM(ZS) algorithm is designed particularly for higher dimensional problems and is not expected to perform optimally when inverting fewer parameters. This may cause the relative efficiencies to change with increased parameters.

4.4 Case Study

To further test the 1D MT inversion model, we also apply it to a two MT case-study datasets from near Boulia, Queensland, Australia. The first site, IEA03242, has a known geology from nearby drilling: 0 - 14 m calcrete; 14 - 72 m weathered Eromanga Basin; 72 - 560 m banded limestone; 560 - 609 m interbedded limestone and sandstone; 609 - m basement. The basement is a meta-dolerite with a tested resistivity of $> 10^4$ $\Omega.m$. The second site, IEB1006A, also has an adjacent drill-hole with shows the following geology: 0 -16 m weathered material; 16 - 1115 m limestone of various composition; 1115 - m basement.

For the inversion of data from IEA03242, a total of 41 frequencies are used from 1 Hz to 10 kHz. After some trial and error the number of layers is fixed at five, the parameter λ is set to 0.5 and error floors are set at 5%. The model is run for 10000 iterations in warm-up and then 5000 sampling iterations for each three chains.

The data from IEB1006A contain 20 frequencies from 10 Hz to 320 Hz. Similarly, these are inverted with the number of layers fixed at five, the parameter λ is set to 0.5 and error floors are set at 5%. Once again, the model is run for 10000 iterations in warm-up and then 5000 sampling iterations for each three chains.

An overview of the results from the two sites tabulated in Table 2. As with the synthetic example, the results are converted into resistivity-depth probability distribution maps (Figure 4 and Figure 5).

Of particular interest in this dataset is locating the depth to basement. We calculate this as the shallowest depth at which there is no greater depth with a resistivity lower than a threshold ρ_{base}. For site IEA03242 we use 10^3 $\Omega.m$ for ρ_{base} and for site IEB1006A we use $10^{2.6}$ $\Omega.m$. The choice of these resistivities will depend on the resistivity of the basement rock in the area of investigation and the range of basement resistivities encountered by the MCMC sampling. The depth to basement is calculated for both sites for each sample from the posterior distribution and a plot of the resulting probability distribution and sample frequency is given in Figure 6 for site IEA03242 and Figure 7 for site IEB1006A.

	mean	mean_{se}	sd	2.5%	50%	97.5%	n_{eff}	$\hat{R}$
Site IEA03242								
ρ_1	1.56	1.6e-3	0.09	1.41	1.56	1.62	3593	1.00
ρ_2	0.74	3.5e-4	0.02	0.69	0.74	0.78	3830	1.00
ρ_3	1.95	5.6e-3	0.26	1.48	1.94	2.40	2182	1.00
ρ_4	3.87	0.04	1.31	1.75	3.97	6.65	879	1.00
ρ_5	4.70	8.1e-3	0.44	4.13	4.62	5.80	2974	1.00
t_1	11.7	0.02	1.01	10.1	11.6	13.9	3267	1.00
t_2	72.1	0.18	8.18	55.7	72.5	86.3	2171	1.00
t_3	509	5.90	223	60.5	497	1010	1430	1.00
t_4	664	9.92	423	38.2	612	1450	1820	1.00
β_1	1.73	0.04	1.38	0.40	1.32	5.50	1239	1.00
β_2	1.98	0.03	1.42	0.52	1.56	5.88	3140	1.00
β_3	2.46	0.04	1.82	0.18	2.05	7.24	2340	1.00
β_4	1.85	0.04	1.66	0.07	1.45	6.18	1475	1.00
Site IEB1006A								
ρ_1	1.50	3.8e-3	0.16	1.11	1.53	1.72	1812	1.00
ρ_2	2.23	0.01	0.40	2.04	2.11	3.41	1065	1.00
ρ_3	3.49	0.06	1.94	1.48	3.04	8.59	938	1.00
ρ_4	4.72	0.08	2.21	1.85	4.22	10.95	852	1.00
ρ_5	5.18	0.09	2.51	3.37	4.34	12.32	781	1.00
t_1	59.2	1.08	37.9	13.0	49.5	156.5	1224	1.00
t_2	840	15.8	398	66.4	1050	1270	634	1.00
t_3	602	7.57	398	38.7	540	1420	2755	1.00
t_4	707	5.31	434	37.3	681	1459	6684	1.00
β_1	1.57	0.02	1.32	0.31	1.15	5.13	6269	1.00
β_2	2.15	0.04	1.89	0.07	1.66	7.12	2296	1.00
β_3	2.06	0.04	1.81	0.08	1.58	6.80	2467	1.00
β_4	1.70	0.03	1.67	0.05	1.18	6.14	3649	1.00

Table 4.2: A summary of the sampled parameters from the case-study probabilistic inversions for sites IEA03242 and IEB1006A. Subscripts indicate increasing layer number. Summaries included are the mean of the distribution; standard error of the mean; standard deviation; percentiles at 2.5%, 50%, 97.5% (the median, and bounds for 95% confidence interval); number of effective samples from `stan`'s effective sample size estimator; $\hat{R}$, the potential scale reduction statistic (Gelman and Rubin, 1992).

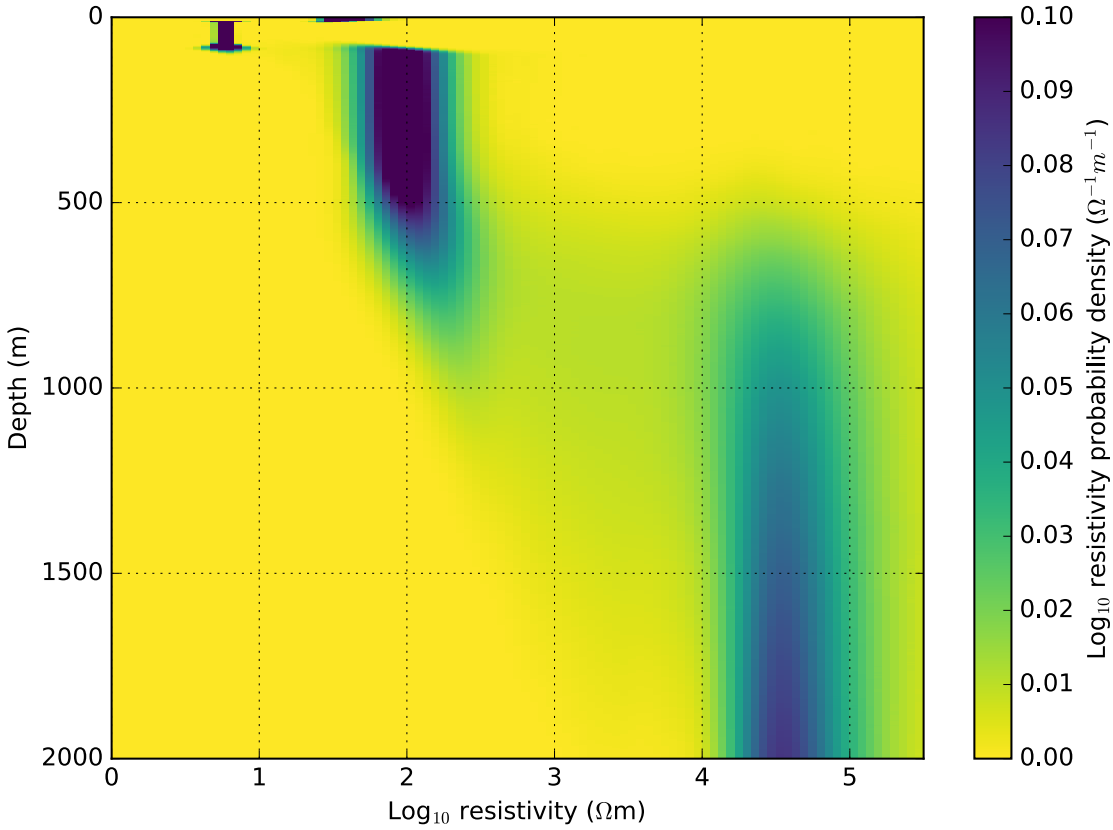

Figure 4.4: Probability densities generated from the site IEA03242 case-study probabilistic inversion, using 15000 samples. Densities have been clipped at 0.1 $\Omega^{-1}m^{-1}$ to allow better visualisation of the lower density areas.

4.4.1 Results and discussion

The $\hat{R}$ values in Table 2 show excellent convergence for both of the inverted sites. An $\hat{R}$ value of 1.00 is achieved for each parameter in each sites. This indicates that the distributions in the three independent chains are indistinguishable and have converged to the same stationary distribution. It is possible that it was easier to achieve convergence in the case-study data compared to the synthetic study as error floors were enforced in the case-study, which makes it easier to fit the data.

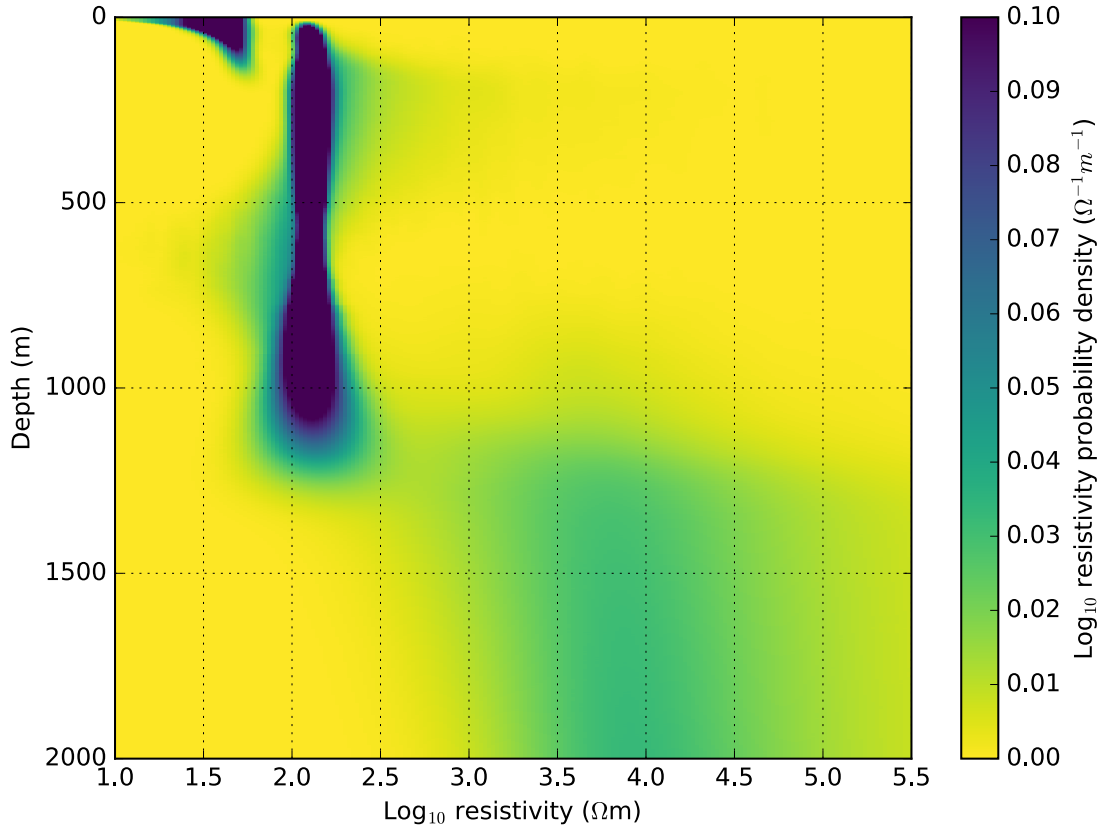

Figure 4.5: Probability densities generated from the site IEB1006A case-study probabilistic inversion, using 15000 samples. Densities have been clipped at 0.1 $\Omega^{-1}m^{-1}$ to allow better visualisation of the lower density areas.

4.4.1.1 Case study one: site IEA03242

From the resistivity-depth profile in Figure 5, a geological interpretation can be made. There is a thin 11.7 ± 1 m resistive layer at the surface, which could represent a calcrete cap. Underneath is a conductive layer of thickness 72 ± 8 m, which corresponds well with the weathered rock layer found in the drill-hole. Beneath this is a layer of resistivity $10^{1.46\pm0.27}$ $\Omega.m$, however a relatively unknown thickness. The layer extends until at least 450 m and after this the structure becomes ambiguous. There most likely model is a straight jump to basement with a four layer model, however there may also be a smooth change down to basement. The basement has a resistivity of $10^{4.4\pm0.4}$ $\Omega.m$, which fits with the laboratory measured resistivity of $10^{4.7}$ $\Omega.m$ for this formation.

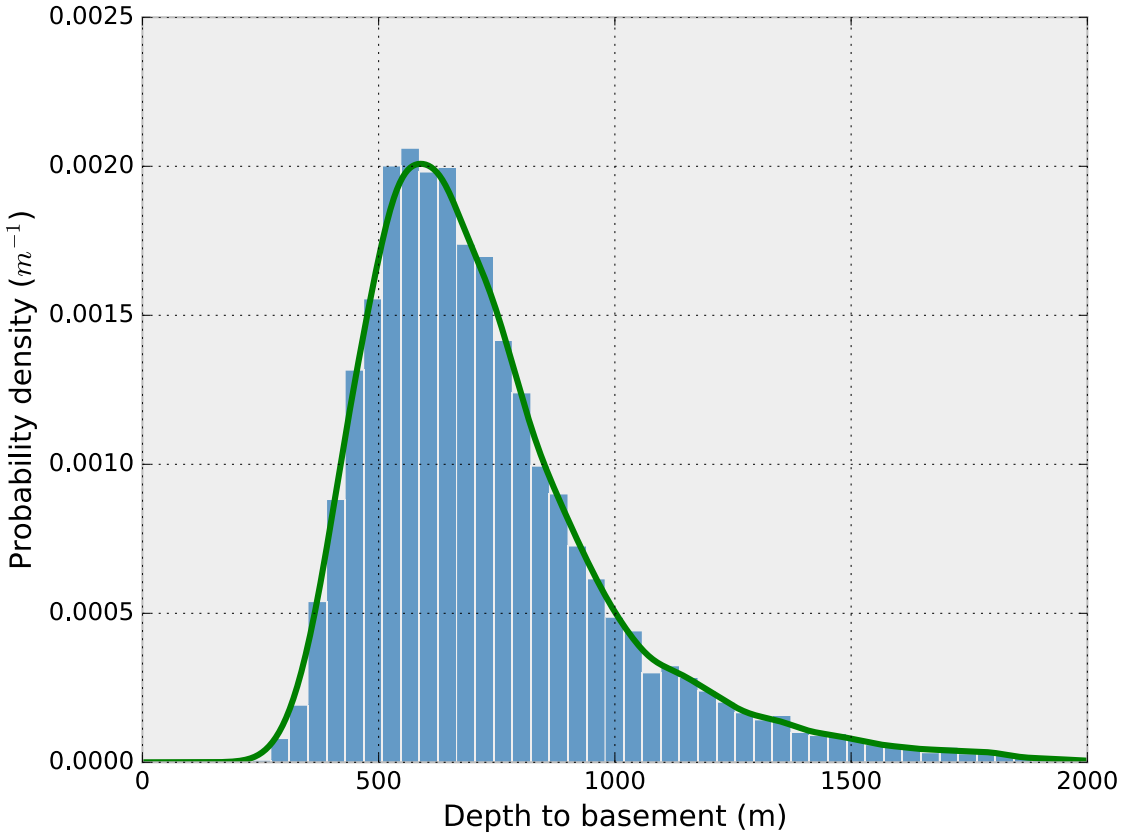

Figure 4.6: Probability distribution with overlain histogram for the estimated depth to basement for site IEA03242. In this case basement is defined as the depth such that all greater depths are more resistive than 10^3 $\Omega.m$. The site has a known depth to basement of 609 m. The 15000 samples have a mean of 723 m, median 667 m, and the probability distribution has a maximum probability at 590 m.

Due to the uncertain nature of the model at depth, it is useful to refer to the depth to basement probability density graph in Figure 7. The probability distribution peaks at 590 m, however the distribution is positively skewed with a mean depth of 723 m. These figures agree well with the known depth to basement of 609 m.

4.4.1.2 Case study two: site IEB1006A

Similarly, Figure 6 allows us to make geological interpretations of the survey area near site IEB1006A. Due to the narrower bandwidth of the data from this site compared to IEA03242, the top segment is relatively unconstrained. The 95% confidence interval for the thickness of the top

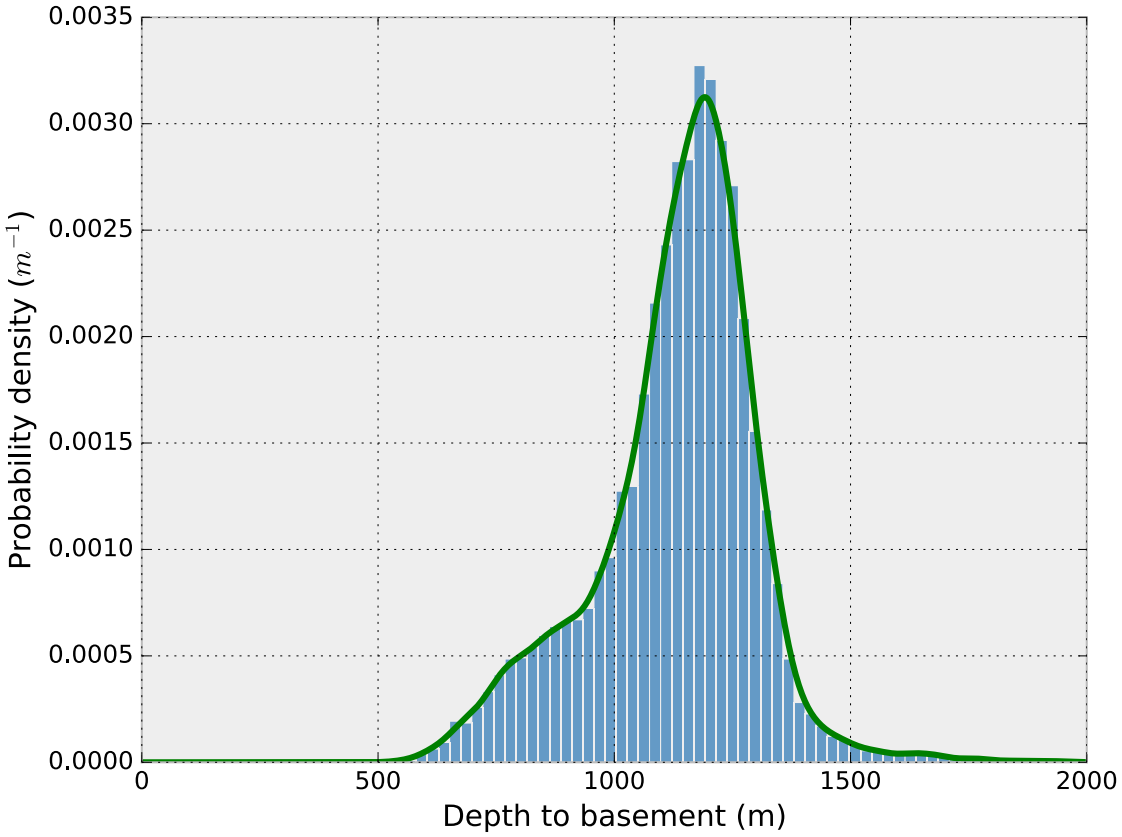

Figure 4.7: Probability distribution with overlain histogram for the estimated depth to basement for site IEB1006A. In this case basement is defined as the depth such that all greater depths are more resistive than $10^{2.7}$ $\Omega.m$. The site has a known depth to basement of 1150 m. The 15000 samples have a mean of 1131 m, median 1156 m, and the probability distribution has a maximum probability at 1195 m.

conductive layer is 13 - 157 m, and we expect a conductive layer of 16 m thickness corresponding to the weathered sediment in the drill-hole. Measurements of this precision are also hampered by the dipole length in the original survey.

Below this weathered layer is a limestone layer which continues down until basement. The inversion model shows this as most likely a single layer with resistivity $\approx 10^{2.05}$ $\Omega.m$, however there is also some probability that there are two layers in here, with a more resistive layer up until 400 m followed by a conductive layer.

Below this section is the basement, which has a relatively unconstrained resistivity of $10^{5.18\pm2.51}$ $\Omega.m$. The high uncertainty of this measurement is

likely due to a lack of penetration depth from the minimum frequency of 10 Hz.

The probability distribution for the depth to basement for site IEB1006A is shown in Figure 8. Compared to the known drill-hole depth to basement of 1150 m, the probabilistic model performs well, with a mean of 1131 m and peak probability at 1195 m.

4.5 Conclusion

We have presented the first geophysical inversion using the NUTS algorithm, implementing an MT inversion with adaptive regularisation. The results of the synthetic probabilistic inversion show that the model can accurately map abrupt jumps in resistivity as well as smooth changes. We have shown that the model is robust to different numbers of layers, provided that enough layers are present to fit the data. Further, we have provided a comparison with the MT-DREAM(ZS) algorithm and shown the NUTS algorithm to be efficient both in speed and ability to explore the parameter space.

The utility of the model has been shown in two case-study sites. For the first site, the model was able to map three layers in the model with strong confidence, and provided an accurate measure of the basement resistivity. The model also gave a probability distribution for the depth to basement which was compatible with the measured depth-to-basement. Similarly for the second site the model was able to map the subsurface resistivity and label the degree of confidence, as well as provide accurate intervals for the depth to basement.

This algorithm is promising for future work in MT probabilistic inversion, where the rapid convergence and efficient exploration will be well suited to dealing with the high number of dimensions involved in a 2D or 3D MT inversion.

Acknowledgements

We would like to thank The Geological Survey of Queensland for supplying the data used in the case study inversions. We would also like to thank the users of the discourse.mc-stan.org community for their guidance in implementing the inversion model. Finally, we would like to thank the two anonymous reviewers who provided feedback to improve this manuscript.

CHAPTER

FIVE

INVERTING MAGNETOTELLURIC RESPONSES IN A THREE-DIMENSIONAL EARTH USING FAST APPROXIMATIONS BASED ON ARTIFICIAL NEURAL NETWORKS

Conway D.[1], Alexander B.[2], King M.[2], Heinson G.[1]., and Kee Y.[2]

[1]School of Physical Sciences,
University of Adelaide, Adelaide, Australia
[2]School of Computer Science,
University of Adelaide, Adelaide, Australia

Submitted to *Computers & Geosciences*

Statement of Authorship

Title of Paper	Inverting magnetotelluric responses in a three-dimensional earth using fast approximations based on artificial neural networks
Publication Status	Submitted for publication
Publication Details	Conway, D., Alexander, B., King, M., Heinson, G., Kee, Y. (2018). Inverting magnetotelluric responses in a three-dimensional earth using fast approximations based on artificial neural networks, submitted to Computers & Geosciences

Principal Author

Name	Dennis Conway
Contribution	Contribution to ideas, programming, modelling, visualisation, manuscript writing and academic correspondence
Overall percentage	75%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.

Signature _____ Date 8/5/18

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

(i) the candidate's stated contribution to the publication is accurate (as detailed above);

(ii) permission is granted for the candidate in include the publication in the thesis; and

(iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name	Brad Alexander
Contribution	Conception of ideas, assistance with machine learning, manuscript supervision

Signature _____ Date 8/05/18

Name	Michael King
Contribution	Generation of training data, assistance with figures

Signature _____ Date 11/5/18

Name	Graham Heinson
Contribution	Assistance with test models and interpretation

Signature _____ Date 10/5/18

Name	Yang Kee
Contribution	Conception of parameterisation technique

Signature _____ Date 11 May 2018

Abstract

The most computationally intensive step in 3D magnetotelluric (MT) inversion is the calculation of the forward response. This fact makes any modelling which requires many function evaluations, including genetic algorithms and Markov Chain Monte Carlo inversion, extremely time consuming. Using Artificial Neural Networks (ANNs) it is possible to approximate these expensive forward functions with rapidly evaluated alternatives. We apply this technique to approximate the 3D MT forward function for a limited subset of resistivity models created in a simple parameterisation. The trained ANN is able to reproduce forward responses with accuracy similar to the level of typical data errors. To evaluate the accuracy of the models, we show that these forward responses may be used to successfully invert MT data in an evolutionary framework. Examples are shown in both synthetic and real-world scenarios, and results are compared with those from traditional inversion algorithms. We conclude that the trained ANN inversion has a fraction of the run-time of a traditional inversion, and is successful at modelling the space of its limited parameterisation.

5.1 Introduction

Magnetotelluric inversion is the process of delineating subsurface resistivities based on surface measurements of passively sourced electric and magnetic fields (Chave and Jones, 2012). This is an under-determined problem, meaning that there are an arbitrary number of resistivity structures which can fit the MT data to equivalent level (Tikhonov, 1963). The most common solution to limit the model space is to regularise the problem using a smoothness constraint (Constable et al., 1987). This approach is efficient, as Fréchet derivatives can be used to continually refine the model. Other methods, such as evolutionary or genetic algorithms, have the advantage of exploring non-differentiable model spaces (Gallagher et al., 1991; Sambridge and Mosegaard, 2002). This approach has been shown to be useful in, for example, Pareto-optimal joint inversions (Schnaidt et al., 2018). Other methods seek to not just find a single model, but to explore the space of viable models using Markov Chain Monte Carlo meth-

ods (Grandis et al., 1999; Guo et al., 2011; Rosas-Carbajal et al., 2013). The majority of these methods, however, have to date been applied only to the 1D model space, where the MT forward algorithm is an analytic calculation, and therefore the efficiency of gradient search methods is not required. Only recently have these methods extended into 2D model space (Rosas-Carbajal et al., 2013), and 3D inversions are realistically impractical without a gradient-search algorithm (Grandis, 2006) or a simplified parameterisation (Rosas-Carbajal et al., 2015).

The use of artificial neural networks (ANNs) has been suggested to alleviate the computational burden of MT inversion (Raiche, 1991). ANNs have the ability to learn complex functions and represent them in an approximate and compact form, which can be rapidly evaluated. Complete MT inversions using ANNs have been developed variously by Zhang and Paulson (1997) in 1D and 2D in limited model ranges; Spichak and Popova (2000) for simplified 3D models; Shimelevich et al. (2007) for a broader range of model classes in 2D; (Montahaei and Oskooi, 2014) for azimuthally anisotropic resistivities; and Shimelevich et al. (2017) for a limited class of 3D models representing Kimberlite pipes.

An alternative to modelling the entire inversion process with an ANN is to model the computationally expensive forward function used by modelling algorithms. This approach is more flexible than a direct inversion, as it allows the user to easily vary the survey design and trial different inversion schemes and regularisations. Learning forward models using ANNs have been employed in various fields. Kello and Plaut (2004) used ANNs to learn the forward function for the relationship between the shape of the vocal tract and the acoustic energy that it emits; Wang et al. (1997) applied the technique for eddy current forward modelling in finding the size and shape of flaws in metal such as heat exchanger tubes in nuclear power plants; and Campisi (2015) has approximated the analytical solution of the diffusion equation using neural networks. There have also been applications within the geosciences, with Krasnopolsky and Schiller (2003) using ANNs to approximate forward operators used in remote sensing. More recently, Hansen and Cordua (2017) have used an ANN to approximate the forward function for crosshole ground penetrating radar (GPR) travel time data. This allowed them to probabilistically invert GPR data using Monte Carlo

simulations, including for errors introduced by the use of an approximate forward function.

This paper will demonstrate the use of an ANN to approximate the MT response from a 3D Earth. We utilise a compact parameterisation described by Alexander and Lee (2018) to generate resistivity structures based on nine parameters: 3D position in space (three parameters); 3D extent in space (three parameters); background resistivity; anomaly resistivity; and attenuation. The synthetic forward responses are then generated with existing forward codes. A regression ANN is then trained to map these parameters to their expected forward responses for the off-diagonal elements of the MT tensor, at an arbitrary point on.

We show that the network is able to achieve a high level of accuracy in its forward response. We additionally show that the ANNs forward response can be used in an evolutionary search inversion with rapid convergence. The inversion is tested on 3D MT synthetic data and on a case-study from the Curnamona Province, South Australia. Both synthetic and real MT data inversions are compared to gradient-based inversions, and are found to have much faster convergence at the expense of model detail.

5.2 Background

In geophysical modelling, the forward function g, relates a given Earth model $\mathbf{m}$ to geophysical data $\mathbf{d}$ via the equation

$$\mathbf{d} = \mathbf{g}(\mathbf{m}). \tag{5.1}$$

The process of inversion is then to find the model $\mathbf{m}$ which best describes a set of observed data $\mathbf{d_{obs}}$ with corresponding error $\sigma_{\mathbf{obs}}$.

For the 3D MT case, we seek to model the function g using an ANN. An ANN is composed of a network of elements known as neurons, which are arranged into layers. This network can be thought of as a generalised function f which maps a set of inputs X to a set of outputs Y: $f : X \rightarrow Y$. Given at least three layers and sufficient neurons and training data, a neural network may be trained to approximate any given function. Artificial

neural networks are trained by inputting corresponding sets of inputs $\mathbf{x}$, and outputs $\mathbf{y}$ originating from the function that is being modelled. During training, the inputs $\mathbf{x}$ are continually evaluated by the ANN and network weights are adjusted to reduce the misfit between the output of the ANN and the target output $\mathbf{y}$. When learning a geophysical forward model, training data are generated by forward modelling a proposed model using existing codes to find the forward response.

To constrain the training data to simpler 3D structures, we used the 3D variant of the reduced-parameter representation described in Alexander and Lee (2018). This paper introduced the parameterisation of a 3D images space as a background value, b, combined with a set of overlapping diffuse ellipse functions f_i. The ellipse functions, referred to as "blobs", allow smooth and simple overlapping large-scale structures to be approximated, whilst having the flexibility to compactly express arbitrary shapes. Each blob is described by a set of parameters: δ, the local resistivity; s, strength (local dominance); α, attenuation; x_{pos}, y_{pos} and z_{pos}, the position in 3D space; x_{rad}, y_{rad} and z_{rad}, the radial size in each dimension; and optionally rotational parameters x_r, y_r and z_r.

For the training set we choose a simplified single ellipse model which reduces the complexity of the model space. The strength property, which relates to the dominance of ellipses over each-other, can be ignored in this case. For the single blob model, we also set rotational parameters to zero.

Each blob parameter is in the range between $[0.0, 1.0]$, and the interpretation of this value depends on the parameter being represented. For parameters representing resistivity: b and δ, the interval $[0.0, 1.0]$ maps on a log-scale to the range: $\rho \in [10^{-2}\ \Omega m, 10^{5}\ \Omega m]$. For spatial position of blobs the range $[0.0, 1.0]$ spans the entirety of the given model space in all three dimensions. For blob size a the parameter represents the radius of the blob as a proportion of the model size in that dimension.

To illustrate the type of models that can be expressed under this scheme, Figure 5.1 shows examples of three different models with the parameters given in Table 5.1.

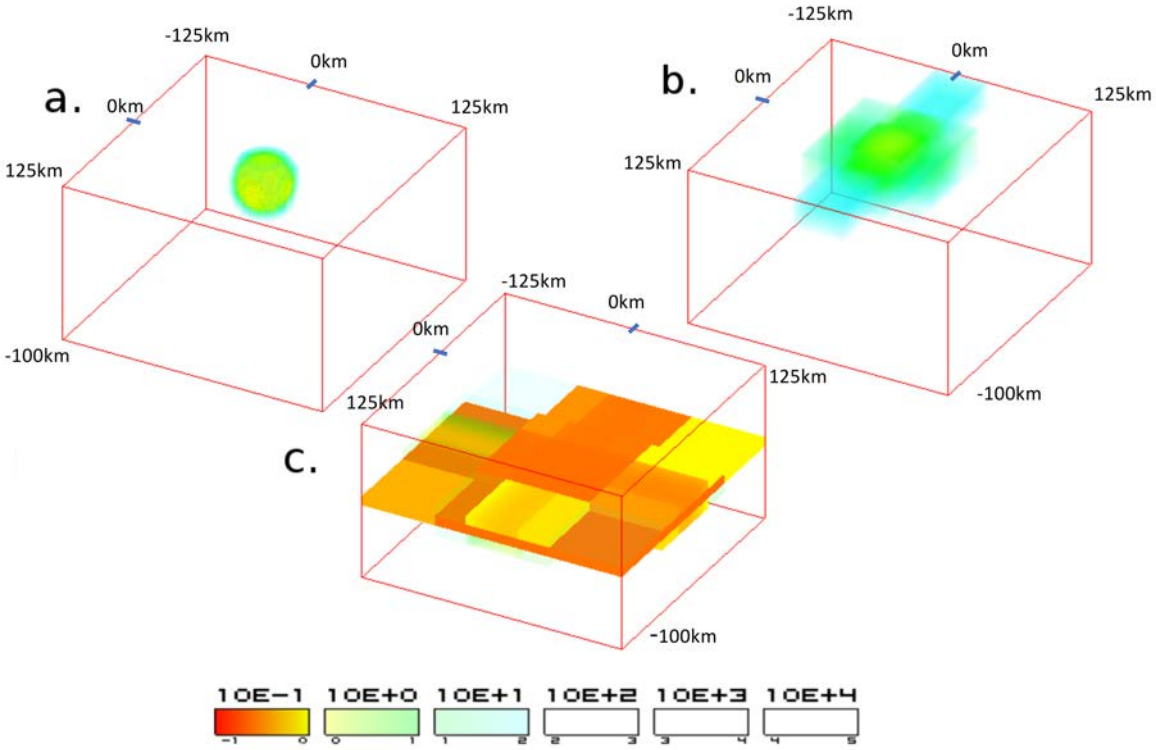

Figure 5.1: Three different realisations of a resistivity structure using a single blob parameterisation and no rotation. The parameters used in the generation of each image are given in Table 5.1

	a.	b.	c.
b	0.7	0.7	0.7
δ	0.1	0.3	0.0
α	0.5	0.3	0.9
x_{pos}	0.5	0.5	0.45
y_{pos}	0.5	0.5	0.55
x_{rad}	0.1	0.3	0.5
y_{rad}	0.1	0.2	0.6
z_{pos}	0.1	0.05	0.3
z_{rad}	0.1	0.1	0.05

Table 5.1: Parameters used to generate images in Figure 5.1.

b	X~U(0.1, 0.9)
δ	X~U(0, 1)
α	X~U(0.02, 0.99)
x_{pos}	X~U(0.4, 0.6)
y_{pos}	X~U(0.4, 0.6)
x_{rad}	X~U(0.05, 0.2)
y_{rad}	X~U(0.05, 0.2)
z_{pos}	X~U(0, 1), $0.4X^{1.2}$
z_{rad}	X~U(0.05, 0.2)

Table 5.2: Distributions used to generate training samples for the ANN for each variable X. $U(a_1, a_2)$ refers to a uniform distribution bounded between a_1 and a_2. Note that the final distribution for z_{pos} has been transformed and is bounded between 0 and 0.4. Distributions have been chosen to focus resistivity structure into areas of interest given the mesh structure.

5.3 Method

To generate a set of MT responses for ANN training, blob parameters were randomly generated by sampling the distributions shown in Table 5.2. The blob resistivity model was discretised onto a hexahedral mesh compatible with the WSINV3D MT forward algorithm (Siripunvaraporn et al., 2005). Mesh geometry was kept relatively simple in order to minimise the time taken for forward evaluations. The mesh, shown in Figure 5.2, contains a dense area of 50 km by 50 km by 100 km in the x, y and z directions, with added padding of 100 km in the x and y directions, and 40 km in the z direction. The dense grid is discretised into 16 by 16 blocks in the x and y directions, giving a total of 256 synthetic site locations at the surface. In the z direction, 20 blocks are used with a initial thickness of 420 m and thicknesses increasing at a factor of ≈ 1.3 per layer.

A total of 50000 sets of blob parameters were generated, converted into a hexahedral mesh, and their corresponding MT responses evaluated using the WSINV3D forward algorithm with an error tolerance of 10^{-7}. MT responses come in the form of the complex valued 2×2 MT tensor, $\mathbf{Z}$, which describe the relationship between the horizontal components of the measured electric and magnetic fields at different periods. We calculated the MT response at five periods: 0.5 s, 5 s, 10 s, 100 s and 1000 s. These frequencies were chosen to give a wide range of the MT response, with a particular focus on the deep MT response which is more likely to be 3D

in nature (Chave and Jones, 2012). The $\boldsymbol{Z}$ outputs were converted into apparent resistivity ρ_a and phase ϕ. Responses were only included for the off-diagonal elements, which are the simplest to model and the most important during inversion. This gives a total of 20 responses for each site location. The total time taken to generate the dataset was roughly 2-weeks using a 16 core machine.

The base-10 logarithm of the apparent resistivities were taken and their values normalised to values bounded by $[0, 1]$. Similarly, the phases were scaled to a $[0, 1]$ range. For each set of blob parameters, the MT tensor responses were calculated for 256 site locations spaces evenly on the dense 50 km by 50 km grid at the centre of the mesh in Figure 2. This gives 256 training samples for each blob configuration, with the x and y grid locations added to the 9 blob parameters to give a total of 11 input parameters. Output parameters are the off-diagonal apparent resistivities and phases, ρ_{xy}, ρ_{yx}, ϕ_{xy}, ϕ_{yx} for each of the periods calculated, which totals 20 outputs. The size of the sample set is 1.28 million.

An ANN is trained using this set of blob parameters and MT responses, using Keras (Chollet et al., 2015), a python wrapper, with the Tensorflow back-end (Abadi et al., 2016). The ANN has the following topology: an input shape of 11 parameters; 7 sequential layers of 100 neurons with rectified linear activation units; and an output layer of 20 neurons with sigmoid activations. This gives a total number of trainable parameters as 63920. The network topology is summarised in Table 5.3. The ANN was trained with a mean average error loss function, and optimised using the Nadam gradient descent optimisation algorithm (Dozat, 2016). A split of 20% of the dataset was saved to use as validation, and the remaining 80% was used during training. The ANN had fully converged after 100 epochs (forward and backward passes of all training samples), with a training time of $\approx$ 5 minutes on a NVIDIA Quadro K5000 GPU. The final training loss was 0.0274 and the validation loss was 0.0277.

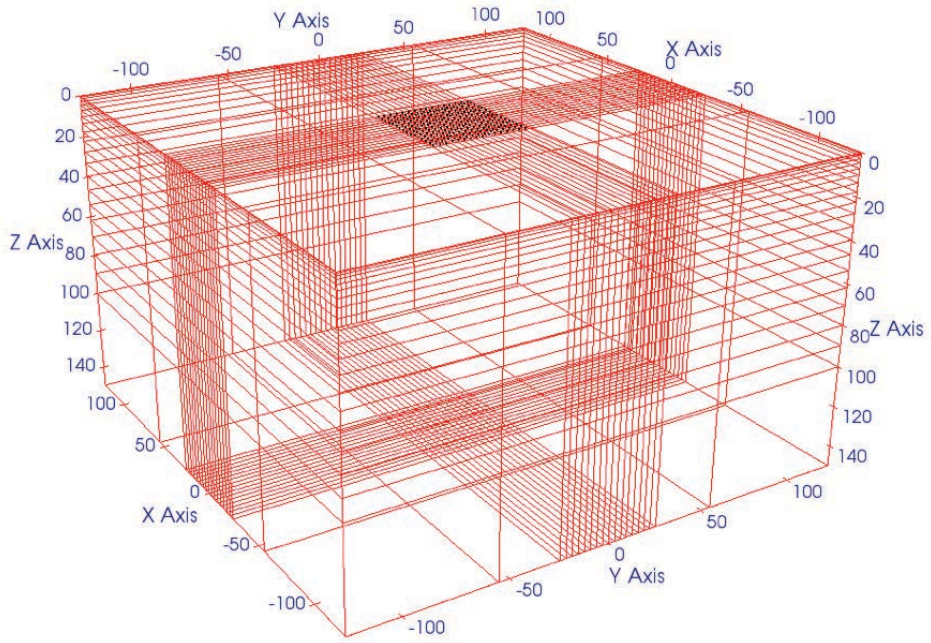

Figure 5.2: Figure showing geometry of the mesh used to generate the forward functions. Axes labels are shown in kilometres.

Layer	Inputs	Outputs	Activation function
1	11	100	rectified linear unit
2	100	100	rectified linear unit
3	100	100	rectified linear unit
4	100	100	rectified linear unit
5	100	100	rectified linear unit
6	100	100	rectified linear unit
7	100	20	sigmoid

Table 5.3: Network structure for the trained artificial neural network. All layers are fully connected to the previous layer.

b	δ	α	x_{loc}	y_{loc}	x_{rad}	y_{rad}	z_{loc}	z_{rad}
0.82	0.206	0.510	0.5	0.5	0.065	0.069	0.09	0.069

Table 5.4: Blob parameters used to create forward responses

5.4 Results

5.4.1 Forward model

To illustrate the differences between the WSINV3D forward model and the ANN proxy algorithm, a test example is shown. A set of blob parameters shown in Table 5.4 were selected to showcase a synthetic example with a strong conductivity contrast.

For the true response, the blob parameters were converted into a hexahedral mesh. The resistivity structure was evaluated using the WSINV3D forward function to give MT impedance responses, which were converted to apparent resistivities and phases. The ANN proxy model was then used to evaluate the same blob parameters, directly outputting the apparent resistivities and phases for the off-diagonal elements. Results are plotted in Figure 5.3, showing the xy responses for apparent resistivity and phase at periods 0.5 s, 10 s and 1000 s. The differenced responses are shown in Figure 5.4. In this relatively extreme example - featuring close to the maximum conductivity contrast supported by the parameterisation - we see that the error introduced by using the ANN is at most 5 degrees in phase and roughly 30 % in apparent resistivity. These are considerable errors, particularly in the apparent resistivity. There is a strong spatial correlation to the errors in the ANN proxy function, for example in the 0.5 s plot in apparent resistivity we see that the ANN overestimates the spatial extent of the anomaly, resulting in a halo of reduced apparent resistivity around the centre.

Importantly, the evaluation of the ANN forward algorithm, once trained, is much faster than the WSINV3D forward function. On an NVIDIA Quadro K5000 GPU and using a batch size of 2048 and random blob parameters, the average time per function call was 6.84×10^{-5} s. Comparatively, averaged over 1000 runs, the time per run for the WSINV3D forward model

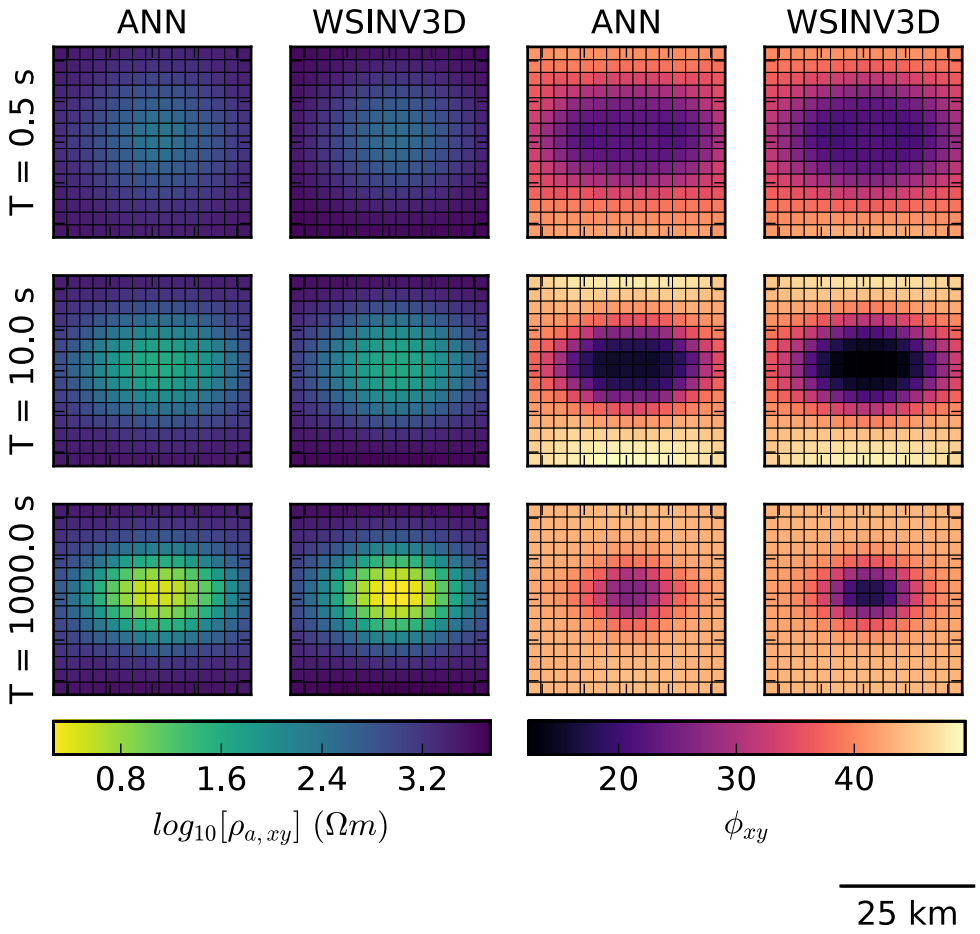

Figure 5.3: Cross sections of surface measured apparent resistivity and phase responses in the xy-orientation at three different periods for each blob, with comparison to the true response from WSINV3D.

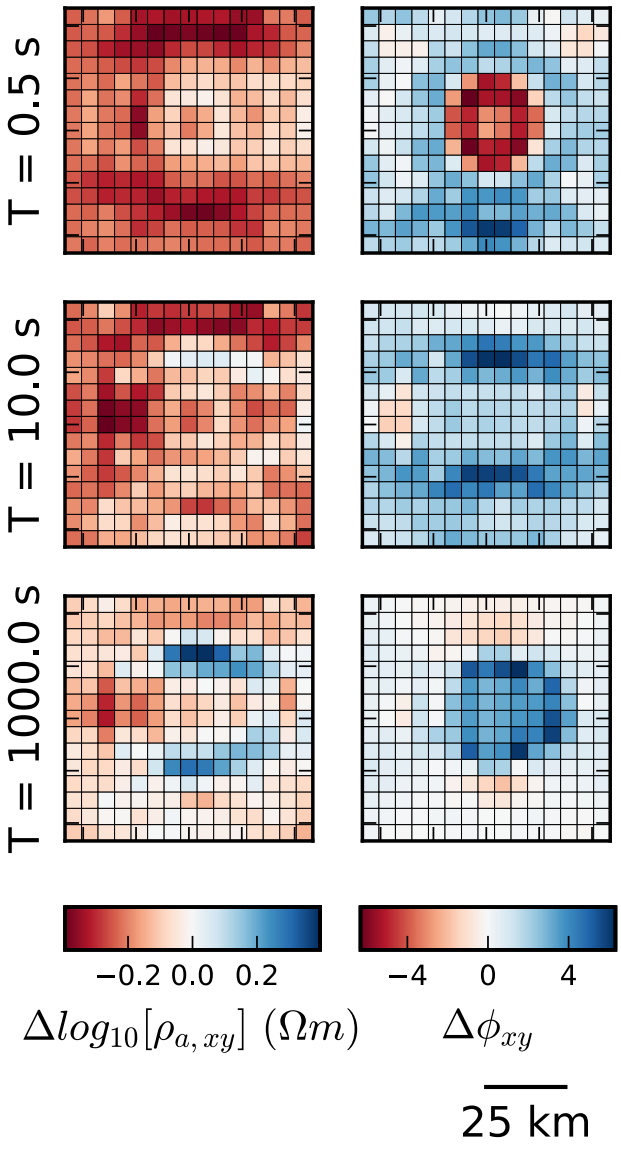

Figure 5.4: Cross sections of surface measured apparent resistivity and phase responses from the ANN proxy minus the response from WSINV3D.

with 6 frequencies was 8.17 s running on a single core of an Intel i7-2600 CPU. This is a speed-up of 5 orders of magnitude. Although it is not as accurate as the WSINV3D forward function, the ANN model allows us to use a wider range of inversion strategies due to its fast evaluation.

5.4.2 Synthetic inversion

5.4.2.1 Method

Using the same blob model in Section 5.4.1, we take the true response from the WSINV3D forward function and apply 5 % Gaussian distributed noise to the impedance tensor data to create a synthetic data-set for inversion. Two inversions are undertaken: a gradient-based, regularised inversion, WSINV3D (Siripunvaraporn et al., 2005), and an optimisation using the Covariance Matrix Adaption Evolution Strategy (CMA-ES) algorithm (Hansen et al., 2003). Both inversions use a chi-squared misfit to measure the goodness of data-fit, which is normalised by number of data-points and expressed as an RMS. An RMS of 1 means that the model is fitting the data to the level of the noise.

The WSINV3D inversion was run with a target RMS of 1, and an initial model of a 1000 $\Omega.m$ half-space. The mesh geometry was the same as used for training in Figure 5.2. The inversion converged after 6 iterations, taking 8 hours to complete running on a single core, with a final RMS of 0.95. The inverted model is shown in Figure 5.5 alongside the true model and ANN inversion.

The ANN inversion was run using the CMA-ES algorithm to find the set of blob parameters which best fit the synthetic data. The CMA-ES algorithm is a stochastic, derivative-free optimiser for continuous variables which is well suited to non-linear or non-convex optimisation problems. At each iteration of the CMA-ES algorithm, a population set of n_b candidate blob parameter solutions are evaluated using the ANN proxy forward function.

For this inversion, convergence was achieved after 400 iterations with a population size of 50, and total running time was less than 10 minutes. The final blob parameters from the best fitting candidate solution were

converted to a hexahedral mesh, and plotted alongside the true model and WSINV3D inversion in Figure 5.5.

In order to ascertain the accuracy of the ANN inversion, the best fitting solution from each iteration was converted into a hexahedral mesh and evaluated using WSINV3D to determine true RMS. A plot of the RMS from the ANN inversion at each iteration is plotted alongside the true RMS in Figure 5.6.

Finally, in order to investigate the generality of this approach, we repeated the ANN inversion using 100 different sets of randomly generated blob parameters. After convergence, the best fitting solution from each of these inversions was then evaluated with the WSINV3D forward algorithm to determine the true RMS. The true RMS was then plotted against the ANN RMS in Figure 5.7.

5.4.2.2 Discussion

Figure 5.5 shows that the ANN inversion provides an excellent fit to the synthetic data. The ANN inversion provides a model which is structurally very similar to the true model. The background resistivity, blob resistivity, blob location and blob x and y radii are all close to the true model. The inversion does overestimate the vertical extent of the anomaly, however, as a diffusive technique, MT in general is not so sensitive to structures beneath conductors. The WSINV3D inversion does better at constraining the vertical extent of the anomaly, but it underestimates both spatial extent and the conductivity. This is expected, as the WSINV3D is a minimum structure inversion. In comparison, the ANN inversion includes much more structure than the WSINV3D model, better recovering the resistivity and extent of the anomaly. It should be noted that we would expect the ANN model to perform well in these tests, as the inverted space for the ANN is the same as the space used to produce the synthetic models.

In Figure 5.6, we test the legitimacy of using the evolutionary search process with our ANN forward function. Plotted in the blue is the RMS given by the ANN forward function on the current best model at each iteration, which is then verified by checking the true RMS with WSINV3D, shown

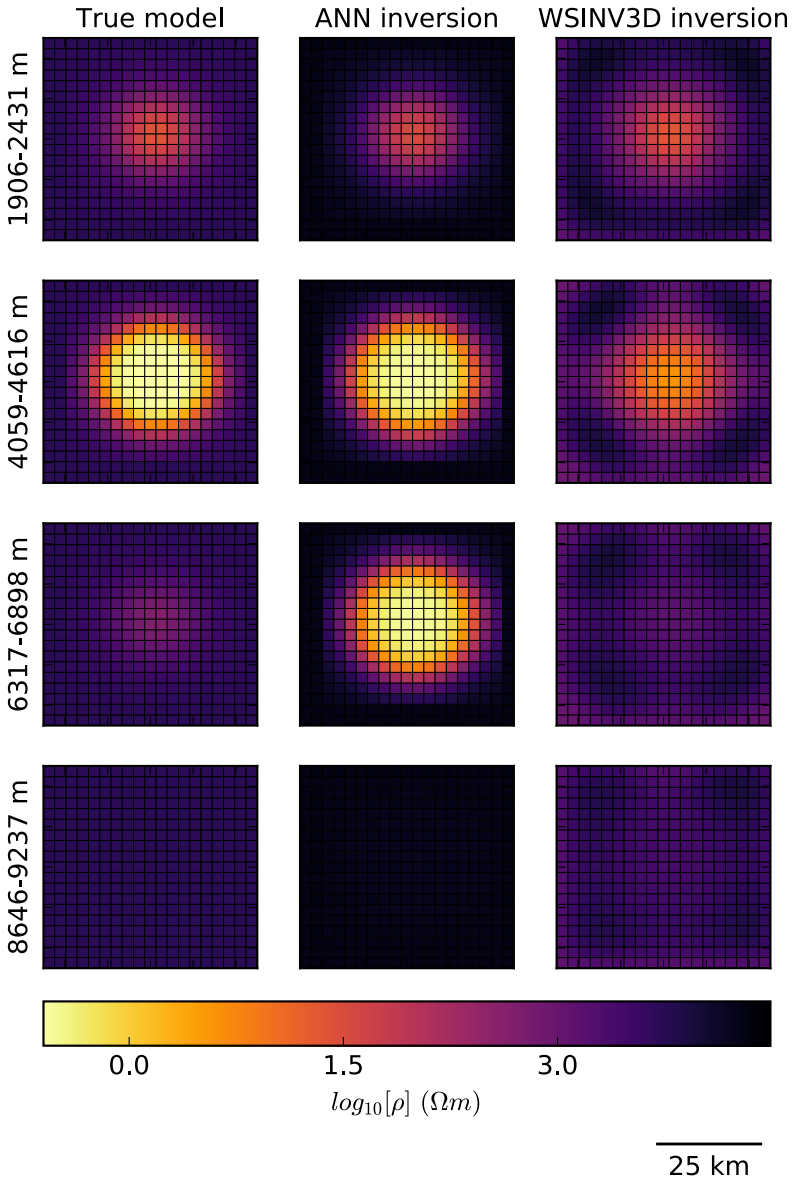

Figure 5.5: Depth slices of inversion results targetting a synthetic model (left) using the ANN proxy function (middle) and the WSINV3D inversion package (right). Each cell represents a station location used in the inversion.

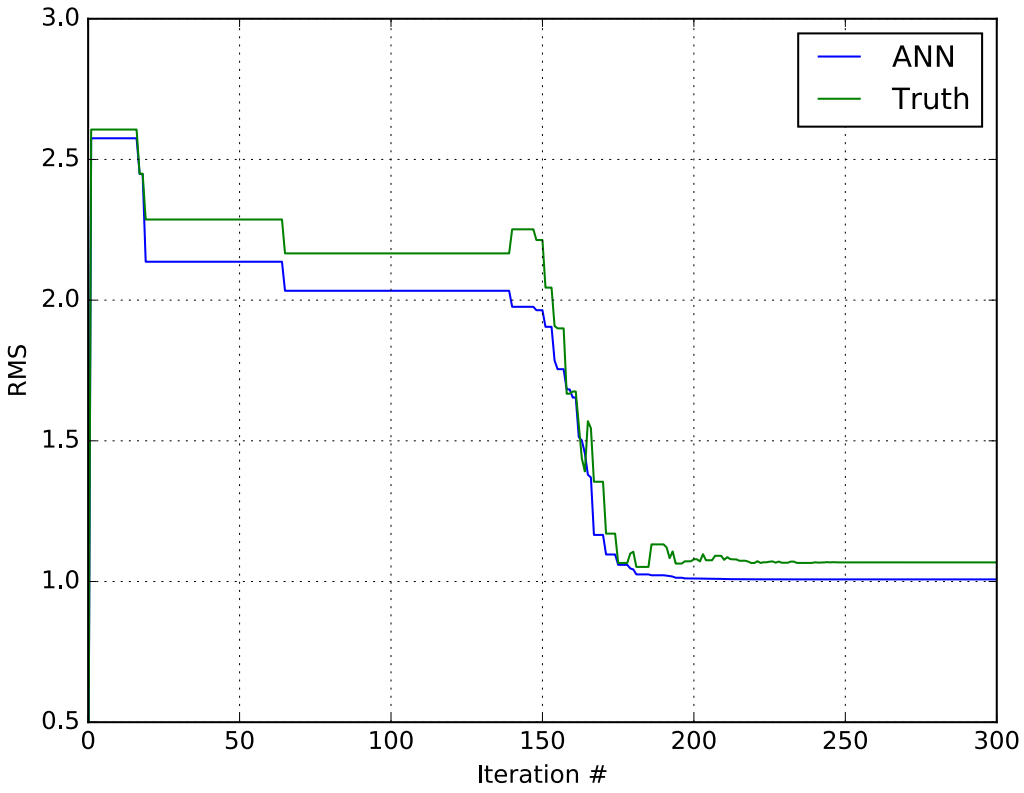

Figure 5.6: A plot of the RMS from the best solution in the population of the evolutionary search at each step of the inversion process, as calculated by the ANN (blue) and the WS-INV3D forward model (green). An RMS of 1 indicates that the model is fitting the data to the noise.

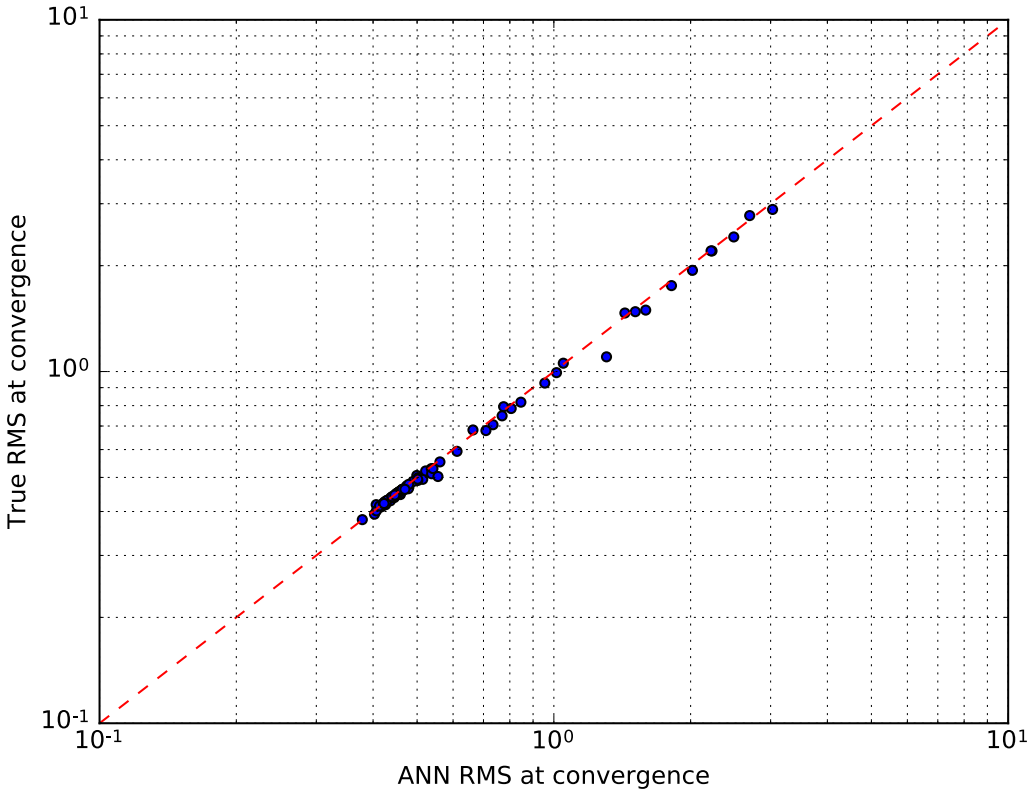

Figure 5.7: The results of 100 inversions of randomly generated blobs. The inversions are undertaken using the ANN evolutionary search method, and then the RMS calculated with both WSINV3D and the ANN.

in green. From the graph, it is clear that there is a strong relationship between the RMS calculated from the proxy function and true RMS, with both calculations in step up to convergence after 225 iterations. The graph also shows that the ranking of models given by the ANN is very similar to the true ranking. This means that as the CMA-ES inversion progresses, the algorithm is progressively finding models which fit the data better.

The plot in Figure 5.7 allows us to further evaluate the effectiveness of the ANN inversion using CMA-ES. It is clear that there is a strong linear relationship between the RMS given by the ANN inversion and the RMS given by WSINV3D, with very little deviation from the red line which shows equivalence of WSINV3D and the proxy function. This shows that the inversion using the ANN proxy function is providing trustworthy fits to the data, which is important for rank based inversion techniques and probabilistic inversions. Of the 100 models, only 11 have RMS values greater than 1. This means that the inverted blob parameters were unable to model the synthetic data according to the noise. A more expressive model would be able to model these data. The majority of the data fits, however, have RMS values less than 1. This means that the inversion is fitting the data better than the error bars. As the blob parameters were randomly generated, many of the models will have low resistivity contrasts. These models will be easier to overfit, as only the background parameter is required to fit the data.

5.5 Case-study inversion

5.5.1 Method

A case-study inversion using MT data collected from the Curnamona Province, South Australia, was undertaken to showcase the utility of the new ANN forward algorithm and CMA-ES inversion. The survey data was originally collected as part of the AusLAMP program, with modelling and interpretation made by Robertson et al. (2017) using the ModEM inversion code (Kelbert et al., 2014). The survey included 74 sites with a period range of 2 s to 17000 s. The sites are spaced roughly within 50 km of each-other.

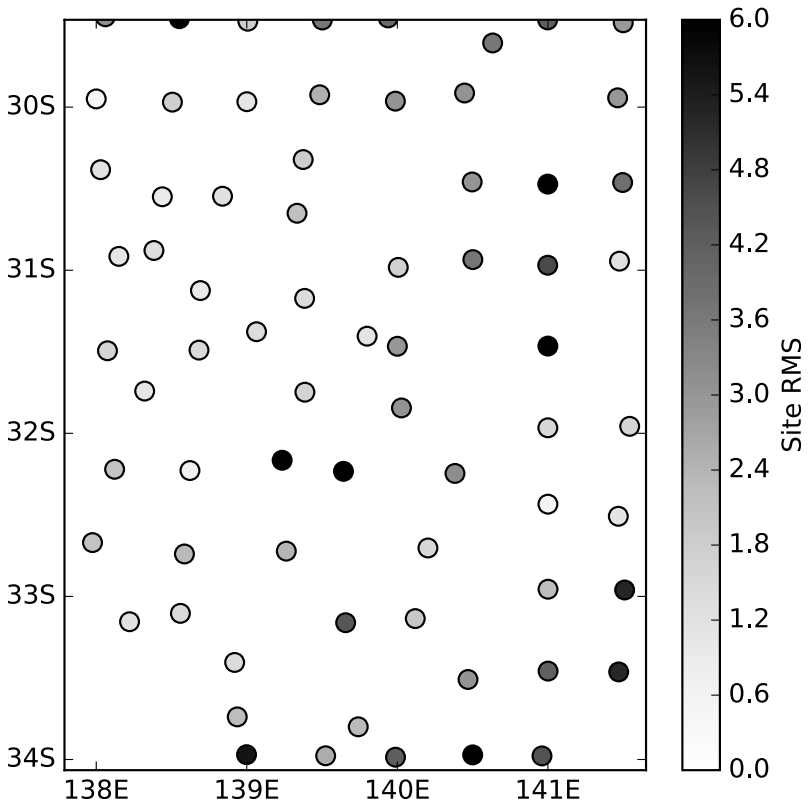

Figure 5.8: Site locations are shown as circles with colouring proportional to the RMS attained by the independent ANN inversion. Locations are expressed as longitudes and latitudes.

Due to the large distance between the sites, we are unable to calculate the forward model of multiple sites using the same blob parameters from our ANN proxy function. Instead, we undertook 74 separate independent inversions at each site location using the same technique as established in Section 5.4.2. For this inversion, we only consider the periods of 10 s, 100 s and 1000 s, as these are the only points for which we have data at. We enforce an error floor of 5% on the data errors. The 74 independent inversions finished within 10 minutes, and the overall RMS for the inversion was 3.76. The RMS values for each inversion is plotted in Figure 5.8, showing the spatial distribution of the model fits.

The resultant inverted blob parameters are then converted to resistivities on a hexahedral mesh. These resistivities are then interpolated onto the

same mesh geometry used for the comparison inversion of Robertson et al. (2017) by setting each point at $\boldsymbol{r} = (x, y, z)$ to an average of the resistivity given by the closest point from each site i, located at $\boldsymbol{x_i} = (x_i, y_i, 0)$, weighted by cube of the distance between the points and the and the quality of fit attained by the inversion as follows:

$$w_{\rho(\boldsymbol{r})} = \frac{\|\boldsymbol{r} - \boldsymbol{x_i}\|^3}{RMS_i}, \tag{5.2}$$

where RMS_i is the RMS error of the inversion at site i. This allows the site nearest to the interpolated cell to dominate the resistivity weights, while giving a preference to the better fitting sites when the cell has multiple adjacent sites. Depth slices from this inversion are plotted in Figure 5.9.

For comparison, the inverted result from Robertson et al. (2017) is also plotted alongside the ANN inversion in Figure 5.9. This inversion was run using the ModEM MT inversion software. It included full impedance tensor data from 23 periods ranging from 2 s to 17000 s, and tipper data from 21 periods from 1 s to 8000 s. The modellers used error floors of 5% for the impedance tensor and 3% for the tipper error. The final RMS for the inversion was 1.33.

5.5.2 Discussion

In Robertson et al. (2017) there are three main structural regions which are imaged in the slice shown in Figure 5.9. The shallowest, the Curnamona Conductor (CC) is present is most prevalent as the strong conductor in the north-east section of the 7 km slice of the ModEM inversion in Figure 5.9. This structure is first seen in the model at 5 km and continues through the 30 km slice to a final depth of roughly 40 km. The inversion also images two other strong conductors, called the Western Nackara Arc Conductor (WNAC) and the Eastern Nackara Arc Conductor (ENAC). These two conductors are roughly parallel to each-other and follow the curve of the Nackara Arc of the Ilkara-Flinders Ranges. The WNAC is present in the model from 20 to 80 km and the ENAC from 20 to 60-70 km.

Robertson et al. (2017) use these resistivity models to create interpretations of the major conductive structures by considering their structures

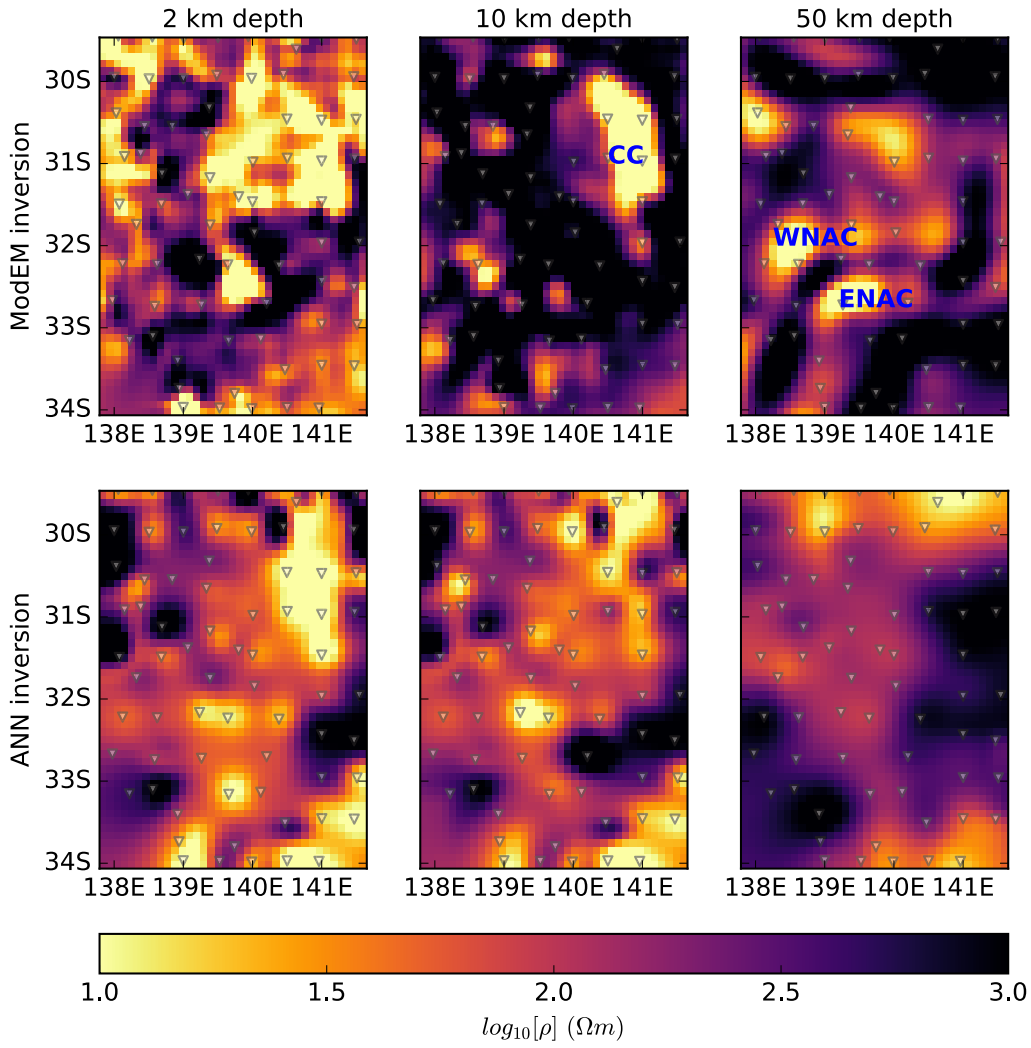

Figure 5.9: Depth slices of the inversion from Robertson et al. (2017) using ModEM (top) compared with the ANN evolutionary search (bottom). Site locations are shown as triangles. Locations are expressed as longitudes and latitudes. The anomalies Western Nackara Arc Conductor (WNAC), Eastern Nackara Arc Conductor (ENAC) and Curnamona Conductor (CC) have been labelled in the figure.

in conjunction with isotope analyses, seismicity studies, elemental compositions and previous geological studies. The Curnamona Conductor is interpreted to support conclusions of the presence of crustal conduits for fluids in the Curnamona Province. Similarly, the Eastern Nackara Arc Conductor is interpreted as the result of fluids or partial-melts sourced from the mantle forming alterations and depositing conductive phases. Additionally, the authors argue that the Western Nackara Arc Conductor could represent the preservation of an active rift conductor.

The results from the ANN inversion have some similarities with the ModEM inversion of Robertson et al. (2017). In the 3 km and 7 km slices there is clear evidence of the Curnamona Conductor in the north-east sector. This is shallower than in the ModEM inversion, which does not have clear evidence of the Curnamona Conductor in the 3 km slice. Structurally, the 3 km and 7 km slices in the ANN inversion are very similar, however in the ModEM inversion there are more regions of high resistivity in the 7 km model. In the 7 km slice of the ANN inversion the majority of the inversion area is roughly 1-10 Ωm, whereas the ModEM inversion features many regions of greater than 1000 Ωm. In the 30 km comparison, the ANN inversion has very similar average resistivities to the ModEM inversion, however the ANN inversion fails to delineate the same structures present in the ModEM inversion. In particular, the Western Nackara Arc Conductor is less prevalent in the ANN inversion, which features a much broader area of conductivity. Similarly, the Eastern Nackara Arc Conductor is not a strong feature in the ANN inversion.

When considering the RMS errors shown in Figure 5.8, we can see that there a relationship between how well the ANN inversion was able to fit the MT data, and the agreement with the ModEM inversion. Particularly, the north-west corner of the inversion features the highest density of well-fitting sites, and these have a strong agreement with the ModEM inversion. One of the most difficult locations for the ANN inversion was the north-east corner, where the majority of data is unable to be fitted using a simple one-blob model. Similarly, the south-east corner contains many sites with a poor model fit. This area was modelled as a region of high conductivity by the ANN inversion, which does not agree with the model provided by ModEM.

Although the ANN inversion is unable to provide the same level of regional detail as the ModEM inversion, it is clear that there is still a strong similarity between the two models. The ANN is able to resolve the key conductive feature in the area, the Curnamona Conductor.

Further work is required to improve the ANN proxy function to a similar level to standard 3D MT forward function codes. Multiple blobs can be used in the same model to produce a wider range of resistivity structures than just a single blob. Training could use the values of the hexahedral mesh directly, allowing multiple models to easily be stitched together. This would also give the forward function a more direct access to the spatial layout of the resistivity structure below. Instead of responses with fixed resistivities, randomly generated frequency responses could be evaluated and treated as function inputs, allowing the user flexibility in frequency evaluations. Additionally, it would be desirable to output the full MT tensor, and potentially tipper responses, rather than just the off-diagonal elements. This added complexity will however require a much larger training set compared to the simple case presented here, adding considerably to the overall computation time to generate the forward responses used for training. Some of the extra training complexity could be alleviated with the use of 3D convolutional neural network layers to discern which parts of a hexahedral mesh input are most important for forward response inference. We would suggest however that the current ANN inversion could be used for the rapid inversion of simple 3D structures, as a more informative starting model for a ModEM inversion, or for the efficient exploration of model space using Monte Carlo sampling techniques.

5.6 Conclusion

In this study, we have demonstrated the use of an artificial neural network (ANN) as proxy function for the MT forward response to facilitate rapid inversion. We have successfully trained an ANN to take a compact parameterisation of the Earth's resistivity and return the MT forward response for the off-diagonal elements of the impedance tensor at five periods. Our results show a good level of accuracy in forward response applied to simple geologies within the compact parameterisation, and importantly a speedup

of five orders of magnitude compared to the WSINV3D forward code. Applied to inversion using the CMA-ES optimisation algorithm, we were able to utilise the ANN forward algorithm to achieve comparable inversion results to WSINV3D, with convergence achieved in 1/50th of the time. By verifying a set of 100 inversions with the WSINV3D forward function, we were also able to confirm that the ANN inversion consistently converges to well-fitting models. The ANN proxy function was further tested by inverting a case-study dataset from the Curnamona Province, Australia. The ANN inversion resolved rapidly, converging in under 10 minutes, however it lacked the same level of detail as the comparison ModEM inversion. The ANN inversion did however delineate some of the key geological structures present in the ModEM inversion.

5.7 Acknowledgements

Thank you to Weerachai Siripunvaraporn for the provisioning of the WSINV3D codes used in the forward modelling and inversion. We would like to thank Kate Robertson for providing us with access to the ModEM output files from the Curnamona inversion, as well as the people and agencies involved with the collection of this data, including Goran Boren, AuScope, and the Geological Survey of South Australia. We would also like to thank Daniel Frasier for creating the visualisation software used for the 3D ellipse plots.

CHAPTER

SIX

DISCUSSION

6.1 Summary

This thesis presents new tools for the inversion of MT data to address areas where gaps in the modelling toolkit exist. There are three broad goals to this work. Firstly is to implement an inversion to model time-lapse MT data in a temporal dimension. Secondly is to explore the problem of non-uniqueness in MT data inversion by implementing a 1D Bayesian inversion using an efficient sampler which is readily extendable to 2D and 3D Bayesian inversion. The third aim is to explore the possibility of approximating the 3D MT forward function with a proxy function based on artificial neural networks.

A time-lapse inversion of 1-D MT data is implemented in Chapter 3. The algorithm considers the entire dataset at once, with penalisations for model roughness in both the spatial and temporal dimensions. The inversion is simple to apply, requiring only a desired level of misfit and the ratio between the temporal and spatial regularisations. The model is tested on synthetic MT data and it successfully retrieves the main characteristics of the test model while introducing minimal artefacts, even in the presence of significant noise in the data. A time-lapse inversion of data from an active coal-seam gas field in the Surat Basin, Queensland, Australia is used as

a case-study. The inversion images small changes in resistivity consistent with the location of layers targetted by the dewatering. The results also agree with previous inversions of the data.

In Chapter 4, a Bayesian inversion of MT data is presented using the NUTS MCMC sampler. The implemented model includes a novel way of regularising MT inversion by allowing the strength of smoothing to vary between different models. This allows for a model which can image sharp boundaries as well as smooth changes. The MCMC sampling shows extremely good convergence and an independence of the final model on the initial model. When tested on noisy synthetic data, the algorithm converges to a model to a maximum *a posteriori* likelihood which is extremely close to the target resistivity curve, and a reasonable approximation of the uncertainty of model parameters. Compared to other sampling techniques, the NUTS sampler provides a more accurate *a posteriori* probability distribution with a considerably shorter run-time. The effect of increasing and decreasing the number of layer numbers is also investigated. It is found that a less-than-sufficient number of model parameters leads to a model which is unable to resolve detail in the model, however the effect of too many layers is negligible. The inversion is then tested against two datasets from Boulia, Queensland, where the inversion result is used to infer the lithology of sedimentary units as well as the depth to basement. The final results are highly consistent with drill hole data in the area, and shows that the inversion is a viable tool for the probabilistic modelling of depth-to-basement.

Finally, in Chapter 5, a proxy function is presented which approximates the MT 3D forward function using ANNs. The function is trained on a parameterised model space and the corresponding MT response from the WSINV3D forward function. The network is able to achieve a reasonable degree of accuracy, with results being within 5% of the correct response. Due to the highly parallelisable nature of ANNs, the forward function is able to evaluate 5×10^5 models per second, which is roughly six orders of magnitude faster than the corresponding analytical model. This allows us to use robust search algorithms which contain large numbers of function calls, such as the evolutionary algorithm CMA-ES, to efficiently model 3D MT datasets. Results from a synthetic inversion show comparable results

between the ANN inversion and standard inversion codes for a simplified 3D model. Importantly, the RMS provided by the ANN proxy function throughout the inversion closely mirrors the true RMS as calculated with the WSINV3D forward function. This result is replicated with another set of 100 inversions. Additionally, these inversions show that the ANN consistently over-estimates the RMS of the model, which is preferable to generating false positives. A case-study inversion is also presented from the Curnamona Province, South Australia. These results show many of the same broad features present in a published ModEM inversion from the same dataset. The ANN inversion, however, is unable to resolve the same level of detail as the ModEM inversion, and misses some of the conductive features present in the ModEM inversion which are essential for intepretation.

The advancements made in this thesis belong in a wider context of MT modelling research. Here a brief overview is given for further development in the work presented in this thesis.

In Chapter 3 an algorithm was presented to invert time-lapse magnetotelluric data in 1D space, motivated by the work of Kim et al. (2009) in time-lapse DC resistivity modelling. The obvious extension of this work would be to extend this work to 2D and 3D model space. Concurrent research has already provided this inversion in 3D space, with Nam et al. (2017) applying the same time-smoothing regularisation as Kim et al. (2009) to successfully invert for a dynamic geothermal reservoir. Other interesting avenues include continuing the work of Rosas-Carbajal et al. (2015), who used a parameterisation of a fluid injection utilising plume physics applied to an injection in a geothermal reservoir. Other parameterisations could include running joint magnetotelluric and hydrological models, for example utilising the relationship between fracture permeability and electrical properties explored by (Kirkby and Heinson, 2017).

Similarly, the 1D MCMC sampler in Chapter 4 can also be extended to 2D and 3D. One of the main advantages of using the NUTS sampler is that it is extremely efficient at sampling in high-dimensional spaces. The sampler is also well suited to correlated parameters. Unfortunately, the `stan` programming language that the algorithm in Chapter 4 is coded in does not support external functions, which would be required to pass the

forward response from 2D and 3D forward models. One alternative is to use the `pymc3` package, which allows for "black-box" functions to be passed along with their gradient (Salvatier et al., 2016).

Finally, the approach in Chapter 5 provides a prelimenary study for much future work using machine learning in geophysical inversion. As with many machine learning problems, the accuracy of the ANN presented would improve with a larger training set. A larger training set would enable features such as variable frequency and responses from the diagonal elements of the MT tensor to be learned. It would also give the possibility of running all resistivity values of the hexahedral mesh as inputs to the ANN, which would allow much more expressive models. This could be achieved by generating random resistivity meshes using multiple randomly generated blobs, and using 3D convolutional neural networks (CNNs) to learn the resistivity features which lead to certain MT responses. The use of 3D CNNs is common in medical imaging for volumetric medical imaging (Kamnitsas et al., 2015; Kleesiek et al., 2016; Milletari et al., 2016) and generally for video imaging (Ji et al., 2013; Molchanov et al., 2016). In a seperate development, the CMA-ES inversion presented in Chapter 5 can be supplemented by a MCMC inversion. As applied for (Hansen and Cordua, 2017) for GPR data, the combination of an ANN proxy function and MCMC allows the modeller to make fast MCMC sampling and also directly account for the error introduced by the proxy function by the introduction of an additional error term into the likelihood function.

To conclude, this thesis presents three advancements in the field of MT modelling. The three achievements are to implement conventional MT inversion to the unconvential setting of time-lapse MT; to further advance Bayesian inversion of MT data using new algorithms and a new model conception; and finally, to create a fast proxy function for the 3D MT forward function to facilitate rapid 3D inversion using evolutionary algorithms. Each of these three areas represent the creation of a new tool for the MT community to use in modelling, extending the quality of models available in the interpretation of MT data. This, in turn, can help geoscientists make better decisions in the many applications of MT.

APPENDIX

A

Schnaidt, S., Conway, D., Krieger, L., & Heinson, G. (2018).
Pareto-Optimal Multi-objective Inversion of Geophysical Data.
Pure and Applied Geophysics, 1-16.

In this paper, further development is made to uncertainty analysis in geophysical inversion. An algorithm is showcased which simultaneously inverts multiple geophysical datasets and provides information regarding the compatibility of the datasets and the probability distribution of the parameters. Personal involvement in the paper included the synthesis of previous work into a single manuscript, editing of figures and tables, and acting as corresponding author for the submission of the manuscript.

Pure Appl. Geophys.

https://doi.org/10.1007/s00024-018-1784-2

Pure and Applied Geophysics

Pareto-Optimal Multi-objective Inversion of Geophysical Data

Sebastian Schnaidt,[1] Dennis Conway,[2] Lars Krieger,[2,3] and Graham Heinson[2]

Abstract—In the process of modelling geophysical properties, jointly inverting different data sets can greatly improve model results, provided that the data sets are compatible, i.e., sensitive to similar features. Such a joint inversion requires a relationship between the different data sets, which can either be analytic or structural. Classically, the joint problem is expressed as a scalar objective function that combines the misfit functions of multiple data sets and a joint term which accounts for the assumed connection between the data sets. This approach suffers from two major disadvantages: first, it can be difficult to assess the compatibility of the data sets and second, the aggregation of misfit terms introduces a weighting of the data sets. We present a pareto-optimal multi-objective joint inversion approach based on an existing genetic algorithm. The algorithm treats each data set as a separate objective, avoiding forced weighting and generating curves of the trade-off between the different objectives. These curves are analysed by their shape and evolution to evaluate data set compatibility. Furthermore, the statistical analysis of the generated solution population provides valuable estimates of model uncertainty.

Key words: Multi-objective optimisation, joint-inversion, data set compatibility, model uncertainty, magnetotellurics.

1. Introduction

Geophysical models can benefit greatly from the combined inversion of multiple data sets. Different methods are sensitive to different petrophysical parameters and different parts of the subsurface, and they usually have uncorrelated noise components. Even the use of multiple data sets from the same method can be beneficial, as the noise components of data sets collected at different times are also likely to be uncorrelated. Thus, additional information available for inversion will improve the quality of the resulting model by reducing solution non-uniqueness (Muñoz and Rath 2006). Standard joint inversion approaches are generally used for data that are sensitive to the same petrophysical parameter, such as electrical and electromagnetic resistivity (Yang and Tong 1988; Abubakar et al. 2011) and seismic velocities (Julià et al. 2000), or methods that are sensitive to different physical parameters, but have a structural connection (Gallardo and Meju 2003, 2007; Commer and Newman 2009; Jegen et al. 2009; Moorkamp et al. 2011).

The classical approach to the joint inversion problem is based on a scalar objective function that combines misfit measures for all data sets and also includes a joint term that connects the different data sets (Haber and Oldenburg 1997; De Stefano et al. 2011). Weighting has to be employed to aggregate all misfits into one objective function. Data sets may be weighted equally (Dobróka et al. 1991; de Nardis et al. 2005), have individual weightings (Julià et al. 2000; Mota and Santos 2006), or use sophisticated techniques such as fuzzy c-means coupling for the joint inversion (Carter-McAuslan et al. 2014). The choice of weights can vary between problems (Treitel and Lines 1999), and the choice of inappropriate weights can lead to bias in the results (De Stefano et al. 2011). A set of guidelines for setting weights is given by Marler and Arora (2010).

The use of a combined objective function also makes it difficult to judge the compatibility of data sets: it is important to determine whether data sets are sensitive to similar features and if the assumed relationship between the data sets is valid. Forcing

[1] Deep Exploration Technologies Cooperative Research Centre, Adelaide, Australia. E-mail: sebastian.schnaidt@adelaide.edu.au

[2] Electrical Earth Imaging Group, Department of Earth Sciences, School of Physical Sciences, University of Adelaide, Adelaide, SA 5005, Australia. E-mail: dennis.conway@adelaide.edu.au; krieger@igem-energie.de; graham.heinson@adelaide.edu.au

[3] *Present Address*: Institute for Geothermal Resource Management, Berlinstrasse 107a, 55411 Bingen, Germany.

Published online: 29 January 2018

incompatible data sets into a joint model may yield a model that is worse than the corresponding single data set models, because an inversion algorithm will produce unnecessary artefacts trying to compensate for an underlying incompatibility.

One alternative to the conventional approaches is the group of multi-objective evolutionary algorithms, which mimic natural evolution processes (Holland 1975). Such algorithms treat each data set as a separate objective rather than aggregating them into a single objective function, which circumvents forced weighting. Calculating individual objective values allows for detailed statistical analysis. For example, it leads to the creation of trade-off surfaces, which allow inference of data set compatibility. These methods are direct search methods (Lewis et al. 2000), which do not require linearisation approximations or any gradient information. They create an ensemble of solutions rather than a single best fit result, which has the added advantage that the solution ensemble can be evaluated to infer qualitative estimates of model uncertainty.

Multi-objective evolutionary algorithms have demonstrated potential to solve problems in engineering, computer sciences, and finance (Coello et al. 2007; Zhou et al. 2011), but they have been sparsely used in the geophysics community. Kozlovskaya et al. (2007) compared conventional and multi-objective methods for seismic anisotropy investigations, but used a neighbourhood algorithm (Sambridge 1999a, b) instead of an evolutionary algorithm. The earliest applications of multi-objective evolutionary algorithms in geophysics included (Moorkamp et al. 2007, 2010), to jointly invert teleseismic receiver functions and magnetotelluric data, as well as receiver functions, surface wave dispersion curves, and magnetotelluric data. Other work has been done on seismic data (Giancarlo 2010), magnetic resonance and vertical electric soundings (Akca et al. 2014), cross-borehole tomography (Paasche and Tronicke 2014), and reservoir modelling (Emami Niri and Lumley 2015).

We present here a multi-objective joint optimisation algorithm, which is based on the Borg multi-objective evolutionary algorithm by Hadka and Reed (2013). In this work, we focus on the application of the algorithm to quantify data set compatibility and also produce a solution ensemble. We will first explain the algorithm in detail and show how the solution ensemble can be used to generate reliable models. We will then demonstrate the functionality of our data set compatibility measure in synthetic model tests and evaluate influences of noise and data error estimates. In our study, we focus on two sets of magnetotelluric data; however, the concept may be extended to any pair of geophysical data.

2. Theory

2.1. Definition of Multi-dimensional Pareto-Optimality

When dealing with multiple conflicting objectives, it is impossible to define a single best solution without introducing weighting of the objectives. In combination with solution non-uniqueness, this is the reason that conventional approaches, which search for a single best fit solution to a joint-inversion problem, produce biased results.

To mitigate this problem, an alternative way to define optimality has to be employed. In the field of multi-objective optimisation, the most widely used concept to rate solution quality is that of *pareto-optimality*, which was first introduced by Edgeworth (1881) and Pareto (1896). A solution is considered pareto-optimal if there is no other feasible solution that can improve an objective without deteriorating any other objective, and the entirety of solutions fulfilling this criterion is called the *pareto-optimal set*. When the pareto-optimal set is projected onto a surface, it is referred to as the *pareto-front*, which comprises a trade-off surface between the different objectives.

The objective value vectors of the pareto-optimal solutions are *pareto-non-dominated*. For a minimisation problem with N objectives, the objective vector $\mathbf{x}^* = (\mathbf{x_1^*}, \mathbf{x_2^*}, \ldots, \mathbf{x_N^*})$, containing the N objective function values for a given solution, is defined to pareto-dominate another vector $\mathbf{x} = (\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N})$ if and only if:

$$x_i^* \leq x_i \quad \forall i \in \{1, 2, \ldots, N\} \wedge \exists j \in \{1, 2, \ldots, N\} : x_j^* < x_j, \tag{1}$$

which is denoted by $\mathbf{x}^* \prec_{\mathbf{p}} \mathbf{x}$ (see, e.g., Coello et al. 2007, p. 10–11).

In a pareto sense, all non-dominated solutions are rated as optimal and no non-dominated solution is considered better than any of the others. In our case, pareto-optimality is a minimal optimality condition that will not always produce physically meaningful results, but rating of the solutions using pareto-efficiency allows for solving the optimation free of weighting biases.

2.2. *Multi-objective Evolutionary Algorithm (MOEA)*

The multi-objective joint optimisation algorithm is a stochastic approach to yield an ensemble of model solutions to an inversion problem. It is based on the auto-adaptive Borg Multiobjective Evolutionary Algorithm (Hadka and Reed 2013).

The Borg algorithm was chosen as it is a state-of-the-art multi-objective evolutionary algorithm capable of adapting to various problems. Multi-objective evolutionary algorithms generally deteriorate in performance for more than three objectives (Ishibuchi et al. 2008; Zhou et al. 2011); however, the Borg algorithm performs well on problems with many objectives (Hadka and Reed 2013). Other advantages of the algorithm include good convergence and high solution diversity of the solution ensemble, which is necessary to infer model ranges and generate reliable information on the compatibility of different objectives.

Evolutionary algorithms are direct search methods that do not require computation of Frechet derivatives. Such methods require significantly more function evaluations than conventional inversion algorithms, but parallelisation of codes is often possible and enhanced computing power is readily available. The stochastic component inherent in evolutionary algorithms makes them very robust against local minima.

The workflow is illustrated in Fig. 1. A starting population is initiated with random parameters inside predetermined parameter thresholds. All member solutions of the population are then evaluated against the measured data sets and objective values calculated for every objective. This is followed by an evaluation of the domination status of each solution. The objective values are usually expressed as root mean square (RMS) deviations δ, the misfit of the forward calculated response of a set of model parameters $\mathbf{m}$ to a set of n observed data points $\mathbf{d}$, normalised by the errors of the observed data points σ_d:

$$\delta = \sqrt{\frac{1}{n}\sum_{i=0}^{n}\left(\frac{d_{\mathrm{i}} - F(\mathbf{m})_i}{\sigma_{d_i}}\right)^2}. \qquad (2)$$

The algorithm also allows the user to set misfit constraints, which effectively limits the feasible region of objective space. Solutions outside the feasible region are treated as invalid.

In addition to the misfit functions, a regularisation measure has to be defined to stabilise the inversion. This measure is treated as separate objective, resulting in pareto-fronts between the model misfits and model complexity. This provides stability by making solutions with lower model complexity outrank solutions with higher complexity for an equal model misfit. The calculation of the regularisation measure is customisable and depends on the model parameters and geometries. In a conventional inversion scheme, the regularisation functional is part of the objective function and its influence in comparison with the misfit measure(s) is determined by a weighting factor, which has to be determined appropriately. Treating the regularisation functional separately from the objective-functions eliminates the need to find this weight factor.

New population members are created via recombination operators after the solutions are evaluated and their domination status is determined. The solutions to be used for recombination are chosen via tournament selection (Miller and Goldberg 1995). There are a variety of different recombination operators available, but usually, only one is implemented in a given algorithm. Different kinds of operators have different degrees of effectiveness, depending on the type and nature of each individual search problem. This led to the proposal of adaptive operators (Vrugt and Robinson 2007; Vrugt et al. 2009). Hadka and Reed (2013) implemented the Borg algorithm with the capability to auto-adaptively

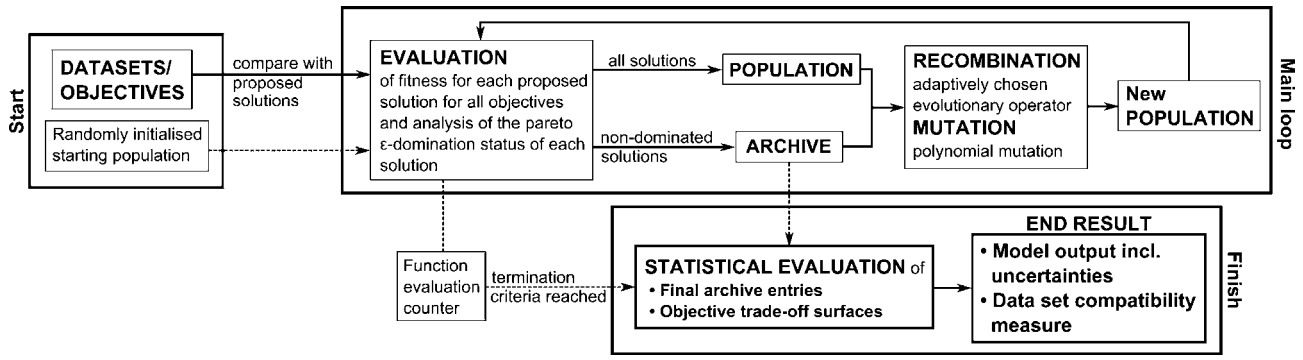

Figure 1
Flowchart of the algorithm's functionality. A starting population is initiated with random parameters and objective values are calculated. After the domination status for each solution is determined, new population members are created via recombination based on the current population. The new population is then evaluated and the loop is repeated until a termination criterion is reached. After search termination, the results are analysed statistically

select from six different recombination operators: simulated binary crossover (Deb and Agarwal 1994), differential evolution (Storn and Price 1997), parent-centric recombination (Deb et al. 2002), unimodal normal distribution crossover (Kita et al. 1999; Deb et al. 2002), simplex crossover (Tsutsui et al. 1999; Higuchi et al. 2000), and uniform mutation (Syswerda 1989). The algorithm adapts the probability of a given operator to be used according to its success rate in producing solutions in non-dominated solutions. For a given problem, generally, one of the operators will be dominant (Hadka and Reed 2013). New solutions produced by all recombination operators, except for the uniform mutation operator, are subjected to polynomial mutation (Deb and Goyal 1996). Mutation operators randomly mutate a given parameter of a solution and add a stochastic component to the search, ensuring better search space exploration and robustness of the search against local minima.

The new population produced by the recombination and mutation process is then evaluated and the loop is repeated until a termination criterion—usually a maximum number of solution evaluations—is reached.

It is important to retain optimal solutions during the search to ensure optimisation success and convergence of the search (Zitzler 1999; Zitzler et al. 2000). Borg exercises this so-called *elitism* by keeping an archive of the non-dominated solutions. When using pareto-efficiency as the optimality criterion for a multi-objective optimisation approach, one has to ensure that the calculated pareto-front is as complete and as close to the real pareto-front as possible. As population and archive cannot be of infinite size, a multi-objective evolutionary algorithm will eventually eliminate solutions, even though they might be non-dominated, known as deterioration of the pareto-front (Hanne 1999). Preventing the pareto-front from deteriorating requires active diversity management (Purshouse and Fleming 2007). Borg employs a modified version of ε-dominance (Hanne 1999; Laumanns et al. 2002) to ensure solution diversity.

The N-dimensional objective space is discretised by dividing it into hyper-rectangles (Coxeter 1973) with side lengths $\varepsilon > 0$ (Fig. 2). Using the notation $\left\lfloor \frac{\mathbf{x}}{\varepsilon} \right\rfloor = \left(\left\lfloor \frac{x_1}{\varepsilon} \right\rfloor, \left\lfloor \frac{x_2}{\varepsilon} \right\rfloor, \ldots, \left\lfloor \frac{x_N}{\varepsilon} \right\rfloor \right)$ ($\lfloor \cdot \rfloor$ denotes the floor function) for a ε-box index vector for an N-objective problem, dominance [Eq. (1)] is redefined as discrete ε-box dominance. An objective vector $\mathbf{x}^* = (\mathbf{x_1^*}, \mathbf{x_2^*}, \ldots, \mathbf{x_N^*})$ is defined to ε-box dominate a vector $\mathbf{x} = (\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N})$ if and only if one of the following equivalent conditions holds:

$$\left\lfloor \frac{\mathbf{x}^*}{\varepsilon} \right\rfloor \prec_p \left\lfloor \frac{\mathbf{x}}{\varepsilon} \right\rfloor, \quad \text{or} \tag{3a}$$

$$\left\lfloor \frac{\mathbf{x}^*}{\varepsilon} \right\rfloor = \left\lfloor \frac{\mathbf{x}}{\varepsilon} \right\rfloor \wedge \left|\left| \mathbf{x}^* - \varepsilon \left\lfloor \frac{\mathbf{x}^*}{\varepsilon} \right\rfloor \right|\right| < \left|\left| \mathbf{x} - \varepsilon \left\lfloor \frac{\mathbf{x}}{\varepsilon} \right\rfloor \right|\right|, \tag{3b}$$

which is denoted by $\mathbf{x}^* \prec_\varepsilon \mathbf{x}$ (after Hadka and Reed 2013). The algorithm also allows for individual

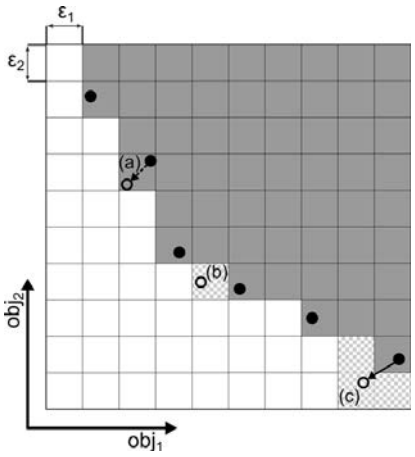

Figure 2
Illustration of ε dominance and ε progress for a hypothetical two objective case. Filled circles mark existing archive members, open circles mark solutions that are newly added to the archive, and grey ε boxes mark the area dominated by the existing archive members. Solutions (**a**) and (**c**) will replace existing archive members, solutions (**b**) and (**c**) also satisfy the conditions for ε progress, and the ε boxes marked with a chequerboard pattern are newly dominated. Modified from Hadka and Reed (2013)

$\varepsilon_i > 0 \quad \forall i, i = \{1, \ldots, N\}$ to be assigned for each objective.

Only one solution per ε box is added to the archive. If a new solution is found that ε box dominates another solution in the same ε box, the former solution will be replaced with the new one.

The ε-box criterion is also used to monitor search progress. The so-called ε progress is achieved if a new-found solution not only ε dominates at least on existing archive entry, but is also located in a previously unoccupied ε box. ε progress is checked sporadically and search restarts will be triggered if search stagnation is detected. If a restart is triggered, the size of the main population is adjusted in relation with the current archive size, according to a predetermined population-to-archive ratio and the population is purged and refilled with new solutions. These new solutions are generally made up of (mutated) archive entries, or new randomly initialised solutions. Maintaining a constant population-to-archive ratio can assist in the avoidance of local minima (Tang et al. 2006). This constant ratio also means that the ε values limit the archive and population sizes and the ε values can be chosen to control these.

We have adapted the Borg algorithm to jointly invert multiple geophysical data sets, such as electromagnetic resistivity well-logs, and seismic. Each data set is treated as a separate objective represented by its own objective function (see Eq. 2). We have added modules for the statistical evaluation of the resulting solution ensembles of the final archive and intermediate archives, to calculate model statistics and uncertainties, and to determine data set compatibilities.

2.3. *Solution Ensemble Appraisal*

The $n_{\text{arch.}}$ solutions contained in the final archive represent the full range of pareto-optimal solutions found by the algorithm before the termination criterion was reached. A pareto-set exists whether or not the data are compatible, but the shape of the distribution of pareto-set members in conjunction with the evolution of this distribution during the optimisation process is dependent on the degree of compatibility. This final solution ensemble can be used to analyse the variability of the model parameters across all solutions to estimate parameter uncertainties. An ideal point in objective space is determined and the solutions close to the ideal point are evaluated to determine the variability of these solutions in parameter space, which indicates parameter uncertainties (Kozlovskaya et al. 2007). The solution with the smallest Euclidean distance to the ideal point is taken as the optimal solution found by the algorithm. This point is chosen as the ideal point under the assumption that with correctly estimated data errors, the normalised misfit will reach a value of $\delta_i^j = 1$ for the optimal solution.

In our tests, we will consider the hypothetical solution with a misfit of $\boldsymbol{\delta} = \mathbf{1}$ in all objectives as the ideal solution or ideal point for our tests, with

$$\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_N)^{\mathrm{T}}$$
$$\mathbf{1} = (1, 1, \ldots, 1)^{\mathrm{T}}.$$

Achieving a misfit of unity is reliant on correct error estimation, and the ideal point will need to be changed if there is reason to believe that error estimates are systematically higher or lower than the given

values. Individual misfits are normalised relative to their ideal point, such that

$$\delta^j_{i;\text{norm.}} = \frac{\delta^j_i}{\left(\boldsymbol{\delta}^j\right)}. \tag{4}$$

Weighted means $\bar{x}$ and the corresponding variances σ_x^2 are calculated for all parameters $\{x_k\}_{k=1\ldots n_{\text{arch.}}}$:

$$\bar{x} = \frac{\sum_{k=1}^{n_{\text{arch.}}} w_k \cdot x_k}{\sum_{k=1}^{n_{\text{arch.}}} w_k}, \tag{5a}$$

$$\sigma_x^2 = \frac{\left(\sum_{k=1}^{n_{\text{arch.}}} w_k \cdot (x_k - \bar{x})^2\right) \cdot \sum_{k=1}^{n_{\text{arch.}}} w_k}{\left(\sum_{k=1}^{n_{\text{arch.}}} w_k\right)^2 - \sum_{k=1}^{n_{\text{arch.}}} w_k^2}. \tag{5b}$$

The weights $\{w_k\}$ are chosen as the distance of a given solution k to the ideal solution in objective space:

$$w_k = ||(\boldsymbol{\delta}_k - \mathbf{1})||_2, \tag{6}$$

to ensure that solutions closest to the ideal point have the largest influence on the result. The regularisation objective is not included in the computation of the weights, as it is not calculated as a misfit-function. The solution's distance from the ideal point is also used to assess the convergence of the population during an inversion by calculating the median of the distances of all analysed solutions.

2.4. Data Set Compatibility

The concept of data-set compatibility is closely related to the concept of conflicting objectives and tries to quantify the degree of conflict. Pareto-front objective trade-off surfaces can be used to analyse compatibility of the different conflicting objectives.

Identical data sets are considered maximally compatible. Hence, for any solution, the misfits $\{\boldsymbol{\delta}_k\}_{k=1\ldots n_{\text{arch.}}}$ for perfectly compatible data sets would be identical across all N objectives and would be distributed in objective misfit space along $\delta_{k,1} = \delta_{k,2} = \cdots = \delta_{k,N} \forall k$. Therefore, in two-objective misfit space, the ideal fit is equivalent to a line with slope $m_{\text{ideal}} = 1$.

To assess the pairwise compatibility of any two objectives, we calculate a linear fit for the solutions in the 2-D plane of objective misfit space of the objectives in question. The deviation of this fit from the ideal line with slope 1 gives information about the degree of compatibility between the two data sets. This scheme is illustrated in Fig. 3.

The standard linear least squares regression (Lawson and Hanson 1974) is a non-robust measure (McKean 2004). We choose the robust Theil–Sen estimator (Theil 1950; Sen 1968) as a regression method to avoid bias from outliers without needing to analyse the data set for outliers and remove them. This estimator for a set of Q 2-D points $\{(x_i, y_i)|i = 1\ldots Q\}$ is calculated as the median $\tilde{m}$ of the slopes $\{m_{i,j}|i,j = 1\ldots Q\}$ calculated between every possible two point combination:

$$m_{i,j} = \frac{y_j - y_i}{x_j - x_i} \quad \forall \quad i \neq j, \quad i > j; \quad \tilde{m} = \text{median}\{m_{i,j}\}. \tag{7}$$

The opening angle γ between the ideal line and the fitted line is assessed to make the analysis independent of objective misfit scale choice, and we assess

$$\tan\gamma = \left|\frac{\tilde{m} - m_{\text{ideal}}}{1 + \tilde{m} \cdot m_{\text{ideal}}}\right| = \left|\frac{\tilde{m} - 1}{1 + \tilde{m}}\right|. \tag{8}$$

Representing the ideal line and fitted line graphically, and using identically scaled axes, perfect compatibility results in a deviation angle from the ideal line of $\gamma = 0°$, and maximum incompatibility results in a deviation angle of $\gamma = 90°$. Deviation angles of $\gamma < 45°$ indicate data compatibility, whereas deviation

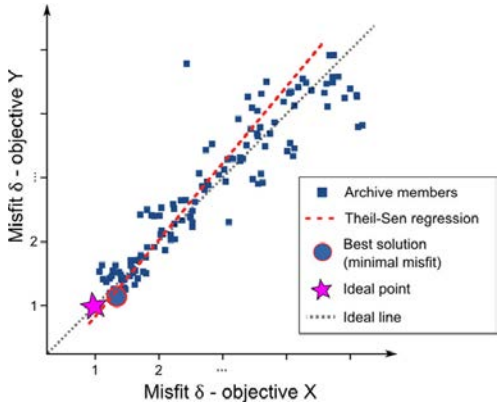

Figure 3
Conceptual misfit visualisation of two objectives for a hypothetical archive of two compatible data sets. The archive members of the pareto-optimal set are scattered around the ideal line with slope 1. The optimal solution is defined as the archive member with the smallest norm deviation from the point **1** in the space of normalised misfits

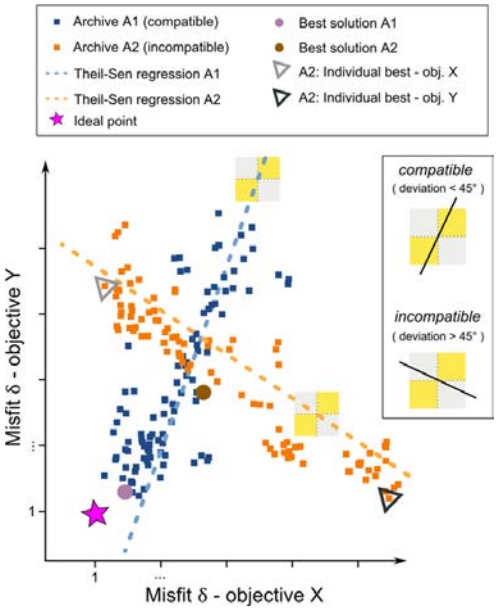

Figure 4
Conceptual misfit visualisations for two hypothetical pairs of data sets: one pair of compatible data sets (blue) and one pair of incompatible data sets (orange). The slopes of the Theil–Sen regressions through both archives are indicated by the labelled 'compatible' and 'incompatible' regions

angles of $\gamma > 45°$ indicate incompatibility. Figure 4 demonstrates the conceptual differences between the misfits of solutions for compatible and incompatible data sets, respectively.

For real-world data sets, perfect compatibility can never be achieved due to a variety of reasons, which will have different manifestations in the way the pareto-fronts deviate from the ideal line: different methods can have different sensitivities and resolution, different depth of investigation, or data sets might have different levels of data error. Different sensitivities or different depth of investigation can cause data sets to neither be fully compatible nor incompatible, but rather partially compatible or disconnected. The pareto-front surfaces for disconnected or partially compatible data sets will have different characteristics than fronts of truly incompatible data sets.

3. Synthetic Tests

We demonstrate the functionality of our approach using sets of synthetic data. We use simulated 1-D magnetotellurics (MT) data sets and resistivity well-logs, which will be inverted for isotropic resistivity and layer thickness.

Using 1-D MT data, we ensure complete controllability of the compatibility of the data sets, while still being able to simulate a variety of different compatibility situations, such as partially compatible data sets with different depths of sensitivity (penetration depth is proportional to the root of signal period). The choice of 1-D data sets also enables easy implementation and greatly reduces the runtime of the algorithm, allowing for intensive testing.

The misfit for the vth frequency is calculated as

$$\delta F_v = \sqrt{\left(\frac{d_v - F(\mathbf{m})_v}{\sigma_{d_v}}\right)^2}. \quad (9)$$

To assess partial compatibility, we analyse the misfits for each individual recording frequency, in addition to the standard misfits, calculated from the sum of all individual misfits.

There are a variety of different regularisation functionals with different characteristics (Pek and Santos 2006, p. 144) of which we use the discretised version (discretisation h) of the total variation functional (Rudin et al. 1992)

$$\sum_{i=1}^{n_{\text{layers}}} \sqrt{(\mathbf{m}_i - \mathbf{m}_{i-1})^2 + \beta^2} \overset{h\to 0}{\longrightarrow} \int_0^\infty |\nabla \mathbf{m}(z)|\,\mathrm{d}z, \quad (10)$$

with a small regularisation constant $\beta > 0$ for numerical stabilisation. We chose the total variation as it can conserve sharp contrast in the model. This is advantageous, as sharp contrasts are often required in layered models.

We created two different synthetic resistivity models (Fig. 5). Model I is the reference model with a low resistivity anomaly between 500 and 600 m and Model II has been designed to generate data incompatible to the first set. Model II has higher resisitivities than Model I in the top 1290 m of the model and lower resisitivities below that depth.

3.1. Data Set Properties

For each of the models, two MT data sets with different frequency ranges are created using Wait's

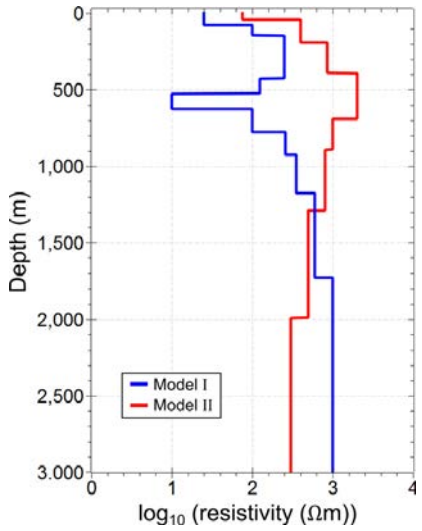

Figure 5
Synthetic 9-layer model Model I (blue) and the synthetic 7-layer model Model II (red)

recursion formula (Wait 1954). This allows us to assess how the compatibility measures behave for data with different depths of sensitivity. The MT data sets have 17 frequencies each, with a frequency range of 6–1448 Hz [broadband (BB) data set], and 128–32768 Hz, respectively [audio-magnetotelluric (AMT) data set]. Eight data points of each of the two different types of MT data lie within the overlapping frequency range of 128–1448 Hz. Everything deeper than the penetration depth corresponding to a 128 Hz signal can, therefore, only be detected by the simulated BB MT measurements. In addition, a resistivity well-log was built for each model, ranging from a depth of 150–1000 m, with a 0.25 m sample interval.

Gaussian noise with a standard deviation equivalent to 3% of the impedance tensor amplitude is added to both the MT data types. Accordingly, error estimates equal to 3% of the impedance tensor amplitudes are assigned. Gaussian noise of 5% is added to the well-log data and error estimates equal to 5% of the parameter values are assigned.

All tests were run for 250,000 solution evaluations each. For the data set compatibility analysis, intermediate solution archives are extracted after 1000, 10,000, and 100,000 solution evaluations in addition to the final archive. As all non-dominated solutions are retained during the processing, the same solutions can be contained in multiple iterations of the archive. Hence, whenever solutions from multiple archives are analysed together, only unique solutions are considered to avoid skewing the statistical analysis.

3.2. Resulting Model Distribution

During the model building, the layer thicknesses are variable, but the number of layers n_{layers} is held constant across all models. The resulting ensemble of models is evaluated with regard to the geometry and the resistivity of the layers. To account for the inherently different data sensitivities and resolutions of different geophysical data, as well as to increase comparability between the different archive solutions, the depth interval between the surface and the deepest overall estimate for the bottom of the last layer is evenly divided into small discrete model segments of constant thickness. A layer-interface can occur at the top of each individual segment. For each solution, the parameter values at a certain depth are mapped to the corresponding segments for each solution, transferring all solutions into a unified segment space. For example, for MT data, the sensitivity decreases with depth dependent on the frequency range. By keeping the segment thickness constant, it is guaranteed that no information is lost when jointly working with data sets from different methods, which have varying sensitivities and resolution.

The segment resistivities are analysed by calculating weighted averages across all extracted solutions.

The layer geometry is evaluated by computing the probability for an interface to be located in a specific segment. This is calculated by using the number of archive solutions that have an interface in a given segment c_i and the total number of final archive solutions $n_{\text{arch.}}$:

$$p_i^{\text{interface}} = \frac{c_i}{n_{\text{arch.}}}, i \in \{1, 2, \ldots, n_{\text{seg.}}\} \tag{11a}$$

$$\Rightarrow \sum_i^{n_{\text{seg.}}} p_i^{\text{interface}} = n_{\text{layers}}. \tag{11b}$$

If all solutions have an interface in the same segment, the interface probability at that segment will be 1. Including the top interface of the first layer, which

is assigned a probability of 1, the sum of all interface probabilities over all segments equals the number of model layers.

In addition to the standard misfit for all data points as defined in Eq. 2, for the MT data sets, we calculate the cumulative misfit over the eight overlapping frequencies 128–1448 Hz (Eq. 9), to allow for a detailed comparison of the regions of equal sensitivity for the different MT data types:

$$\delta^{\mathrm{CMOF}} = \sqrt{\frac{1}{8}\sum_{v=0}^{8} \delta F_v{}^2}. \qquad (12)$$

4. Modelling

To demonstrate the feasibility of the approach, we analyse the example data sets successively and in detail to illustrate the influences of the various parameters. First, we demonstrate the overall functionality and present the algorithm's outputs using compatible data sets. Then, we characterise incompatible data sets, and extend the concepts from two to multiple objectives. Finally, we discuss the negative influence of ill posed problems and the lack of adequate regularisation.

4.1. Two Objectives—Compatible Data Sets

We will first evaluate a simple case with two compatible objectives to introduce the concepts of the method. The objectives are built from AMT and BB MT data sets, combined with regularisation. This compatible data example uses the MT data sets that both have been calculated from Model I.

Figure 6 shows the best solution and the average result for the compatible case calculated from the final solution archive. The optimal solution achieved misfits of $\delta_{\mathrm{AMTI}} = 1.3$ and $\delta_{\mathrm{BBI}} = 1.5$ and was at a distance of 0.5 from the ideal point. The average model exhibits an average standard deviation of 21% relative to the segment values. There is an overestimation of the resistivity in the low resistivity zone, which reaches values of 30Ωm for the best solution and 43Ωm for the average model, as opposed to the 10Ωm of the true model. There is also an

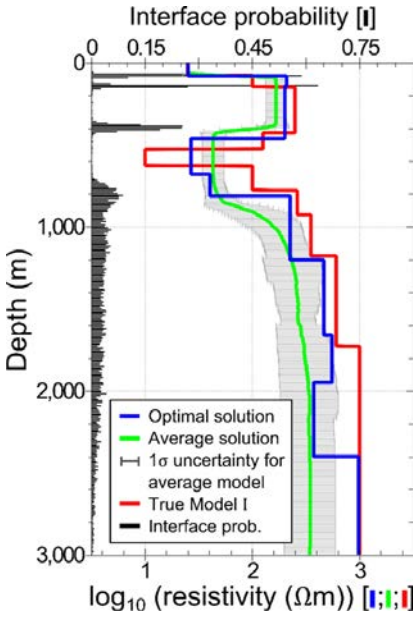

Figure 6
Model results from the two-objective case with 3% noise. The red line shows the true synthetic model and the green line represents the weighted average model based on the 1142 solutions of the final archive, including the model uncertainties in grey. This are calculated using Eqs. (5a) and (5b). The optimal solution is presented in blue. It achieved misfits of $\delta_{\mathrm{AMTI}} = 1.3$ and $\delta_{\mathrm{BBI}} = 1.5$ and was at a distance of 0.5 from the ideal point. The average model exhibits an average standard deviation of 21% relative to the segment values. The interface probabilities are presented in black calculated from the final archive solutions

underestimation of the resistivities at greater depths, with the best solution showing a closer fit than the average solution. The locations of layer interfaces are well determined at low depths, but are subject to higher uncertainty at larger depths. The depth interval of 700–800 m is jointly constrained by the two data sets, resulting in well constrained layer boundaries. Below this depth, the model is only constrained layer boundaries. Below this depth, the model is only sensitive to the broadband data, which relies on lower frequencies and, therefore, has a lower resolution, making it incapable of determining well constrained interfaces.

The CMOF are shown in Fig. 7a. The solutions are distributed along the ideal line. The linear fit deviates from the ideal line by 3°.

Figure 7b displays the locations in objective space of all the solutions extracted from the archives. The different depths of investigation of the two data set cause the solutions to be distributed in a cone shape,

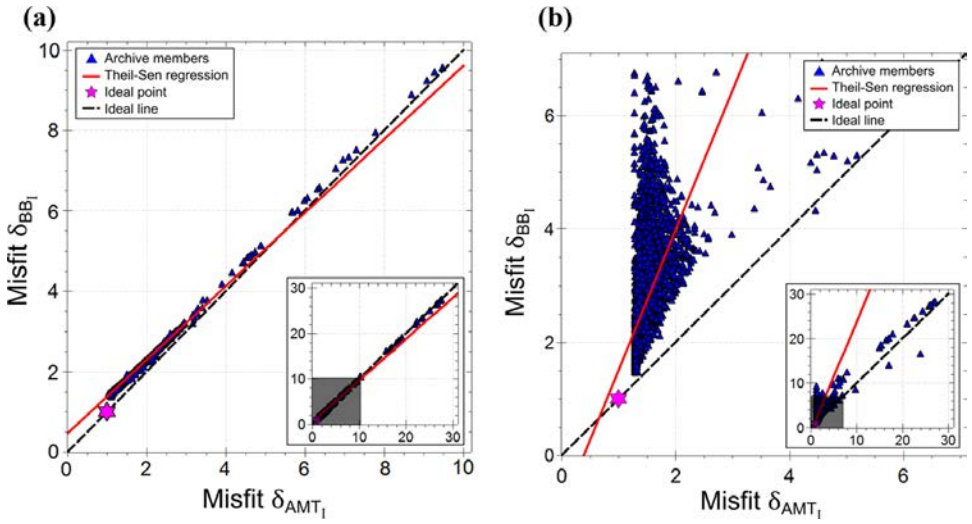

Figure 7
Archive solutions for AMT objective and the BB-objective in the compatible two-objective case with 3% noise. The 2440 unique solutions combined from the archives after 1000, 10,000, 100,000, and 250,000 solutions evaluations are displayed, as well as the corresponding Theil–Sen regression, the ideal line, and the ideal point. **a** Cumulative objective mists over the overlapping frequencies. **b** Objective mists over all frequencies

shifted towards higher BB misfits. The main cluster extends over a misfit of about 1.3–2.4 in the direction of the AMT misfit and from about 1.5–6.8 in direction of the BB misfit, with the optimal solution found with misfits of $\delta_{\mathrm{AMTI}} = 1.3$ and $\delta_{\mathrm{BBI}} = 1.5$, at a distance of 1.9 from the ideal point. The deviation from the ideal line of the linear fit is 23° (Table 1).

It should be noted that in Fig. 7b, a combination of archive members after 1000, 10,000, 100,000, and 250,000 evaluations is plotted. Although each individual archive does not contain dominated solutions, earlier archive solutions are likely dominated by members of later archives. The dominated solutions of earlier archives are included to capture the evolution of the solution distribution, which is a major indicator of the objective compatibility. Therefore, all results are included during the compatibility analysis; however, in the final consideration of a representative model, dominated solutions should be discarded.

Table 1

Analysis of the deviation from the ideal line and median distance from the ideal point (1,1) for the compatible two-objective case (with 3% noise). The analysis is performed for archives at different stages of the inversion run, as well as for all extracted archive members combined and the CMOF of the combined archive members

Solution evaluations	Deviation from ideal line	Median distance from ideal point
1000	27°	5.6
10,000	9°	2.3
100,000	26°	1.9
250,000	23°	1.9
Combined	23°	–
CMOF	3°	0.8

4.2. *Two Objectives—Incompatible Data Sets*

We have established how resulting model distributions behave for compatible data. Now, we explore the results of the algorithm for incompatible data. The AMT data set is built from Model I and the BB data set is calculated using Model II to simulate data incompatibility.

The resulting pareto-fronts are shown in Fig. 8a. The CMOF are distributed along a line with a deviation of 65° from the ideal line and a median distance from the ideal point of 20.64 (Table 2), which contrasts the analysis of compatible data. The main cluster of solutions covers AMT misfits of 4–350 and BB misfits of 5–170. These differences in misfit ranges are caused by the fact that the models for Model I and Model II exhibit greater similarity at depth than close to the surface. Hence, the misfits of the lower frequency BB data set are smaller.

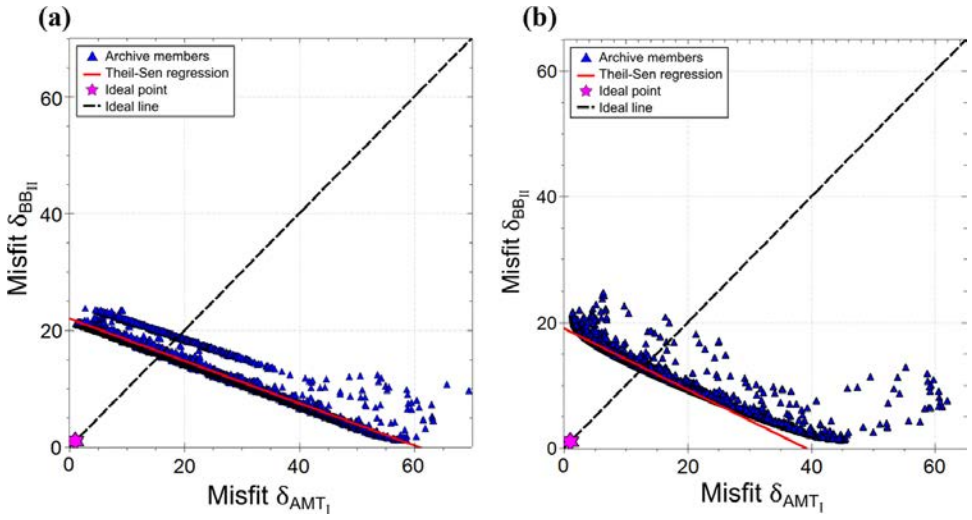

Figure 8
Archive members for AMT objective and the BB objective in the incompatible two-objective case with 3% noise. As there is no compatibility between the objectives, there is no pareto front in this case. The 22,148 unique solutions combined from the archives after 1000, 10,000, 100,000, and 250,000 solution evaluations are displayed, as well as the corresponding Theil–Sen regression, the ideal line, and the ideal point. **a** Cumulative objective misfits over the over lapping frequencies. **b** Objective mists over all frequencies

Table 2

Analysis of the deviation from the ideal line and median distance from the ideal point (1,1) for the incompatible data two-objective case with 3% Gaussian noise on the data. The analysis is performed for archives at different stages of the inversion run, as well as for all extracted archive members combined and the CMOF of the combined archive members

Solution evaluations	Deviation from ideal line	Median distance from ideal point
1000	69°	20.1
10,000	70°	17.8
100,000	71°	16.7
250,000	71°	16.4
Combined	71°	–
CMOF	65°	20.64

The same pattern can be observed for the full frequency range misfits (Fig. 8b). Compared to the CMOF the line shows a higher degree of scatter, a slight curvature, and exhibits a deviation from the ideal line of 71°. This curvature is caused by the different frequency ranges of the two data sets.

The distances from the ideal solution and the large deviation from the ideal line illustrate that the algorithm is able to find solutions with low misfits for each of the objectives individually, but it is impossible to find a solution that reaches acceptable misfits for both objectives at the same time.

4.3. Multiple Objectives

We perform two test runs with three objectives to investigate the behaviour of the compatibility measures for compatible and incompatible cases with more objectives. Both tests use the AMT and the BB data set based on Model I. The test simulating compatible data sets uses the synthetic resistivity well-log based on Model I and the test for incompatible data uses the Model II resistivity well-log. Both well-logs cover depths of 150–1000 m.

Figure 9 shows the best and average results for the compatible three-objective case. The added information from the well-log helps to better define the position and resistivity of the low resistivity anomaly compared to the two-objective case (Fig. 6). The anomaly is identified at the true location and has a resistivity of 11.8 Ωm for the best found solution and 24.5 Ωm for the average solution. The benefit of the constraints added by the well-log is also reflected in the smaller error bars of the average solution, as compared to the two-objective case, with the average model exhibiting an average standard deviation of 18% relative to the segment values.

As there are three objectives competing in this test, the compatibility analysis is performed pairwise for each of the three possible two-objective combinations. In the case of compatible data sets, the linear

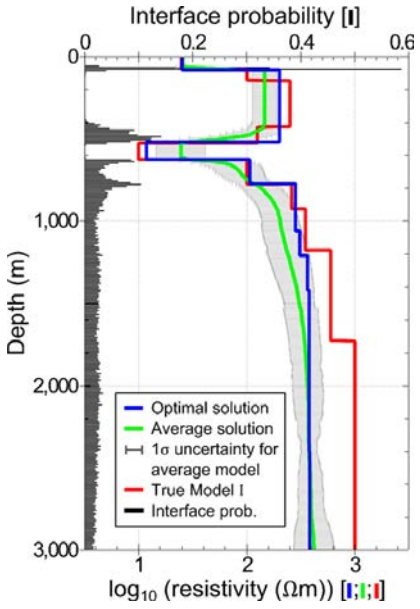

Figure 9
Model results from the multi-objective case. The red line shows the true synthetic model and the green line represents the weighted average model based on the 6771 solutions of the final archive, including the model uncertainties in grey. The optimal solution is presented in blue. It achieved misfits of $\delta_{\mathrm{AMTI}} = 1.3$, $\delta_{\mathrm{BBI}} = 1.7$, and $\delta_{\mathrm{WELLI}} = 1.4$ and was at a distance of 0.8 from the ideal point. The average model exhibits an average standard deviation of 18% relative to the segment values. The interface probabilities are presented in black calculated from the final archive solutions

fits for extracted solutions exhibit deviations from the ideal line of $\gamma_{\mathrm{comp;AMT-BB}} = 5°$, $\gamma_{\mathrm{comp;AMT-WELL}} = 6°$ and $\gamma_{\mathrm{comp;BB-WELL}} = 15°$, indicating good compatibility between all objectives (Table 3). The optimal found solution has objective values of $\delta_{\mathrm{AMTI}} = 1.3$, $\delta_{\mathrm{BBI}} = 1.7$ and $\delta_{\mathrm{WELLI}} = 1.4$, and is at a distance of 0.8 from the ideal point. The median distance from the ideal point achieved by the solutions from the final archive is 4.85.

Good objective compatibility is also indicated for the two MT objectives in the case of incompatible data, with the linear fit for all solutions deviating by $\gamma_{\mathrm{incomp;AMT-BB}} = 3°$. The two objective combinations featuring the well-log data on the other hand show clear signs of incompatibility. The linear fit of the solutions projected onto the objective space plane of the AMT misfit and the well-log misfit exhibits a deviation of $\gamma_{\mathrm{incomp;AMT-WELL}} = 83°$ from the ideal line, and for the combination of BB MT data set and well-log the deviation is $\gamma_{\mathrm{incomp;BB-WELL}} = 57°$ (Table 3). This smaller deviation for the BB-WELL projection compared to the AMT-WELL combination is caused by the

Table 3
Analysis of the deviation from the ideal line and median distance from the ideal point (1,1,1) for the multi-objective case inversion runs. The analysis is performed for archives at different stages of the inversion run, as well as for all extracted archive members combined and the CMOF of the combined archive members. The three datasets are audiomagnetotelluric (A), broadband magnetotelluric (BB), and well-log (W)

Solution evals.	Deviation from ideal line						Median distance from ideal point	
	Compatible data			Incompatible data			Comp.	Incomp.
	A-BB	A-W	BB-W	A-BB	A-W	BB-W		
1000	28°	58°	64°	4°	49°	50°	11.6	11.1
10,000	1°	32°	23°	20°	56°	53°	5.3	12.6
100,00	2°	1°	15°	2°	86°	56°	5.0	9.8
250,000	9°	13°	13°	6°	80°	59°	4.9	9.1
Combined	5°	6°	15°	3°	83°	57°	–	–

larger penetration depth of the BB data that exceeds the depth range constrained by the well-log, whereas most of the depth range that the AMT data are sensitive to is constrained by the well-log. The median distance from the ideal point is 9.1, and as such significantly larger than for the compatible data.

The clear separation into compatible and incompatible data apparent from the analysis of the deviations of the linear fits from the ideal line is less obvious from a visual inspection of the solution distributions (Fig. 10). The objective combinations including the well-log show similar distributions for the compatible and the incompatible case. In each case, the main solution clusters have a width of about 10 in direction of the well-log objectives and a width of 40–90 in direction of the MT objectives. This asymmetry is caused by the fact that the well-log only constrains part of the model, so that models fitting the well-log can still vary significantly in the misfit of the MT data sets.

5. Discussion

The evaluation of jointly inverted or jointly interpreted geophysical data is complicated, and it is vital to assess if information from different data sets can be jointly analysed in the first place. We have demonstrated that the output of the algorithm can be interpreted as a measure for the mutual compatibility of multiple data sets.

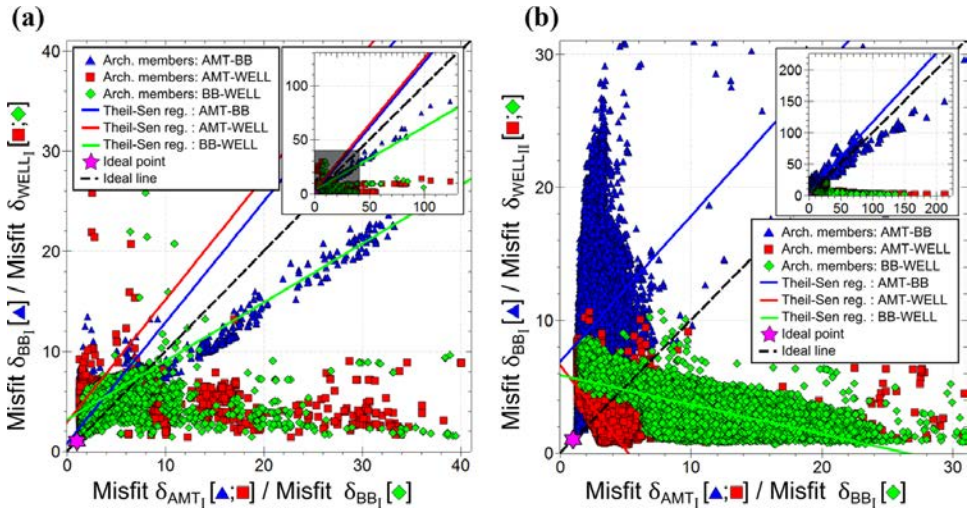

Figure 10
Pareto-fronts for pairwise AMT-BB objectives (blue triangles), AMT-WELL objectives (red squares), and BB-WELL objectives (green diamonds) with the corresponding Theil–Sen regressions in the same colour for the multi-objective case for both the **a** compatible and **b** incompatible cases. The ideal point is represented by a magenta star and the broken black line represents the ideal line. For both cases, the objective misfits are shown over all frequencies. The calculated pareto-front is actually 3D, but here, we visualise the 2D components. Inset in both cases are zoomed out versions of the graph showing the overall structure of the regressions

Using a linear regression allows us to make direct meaningful analysis of the geometry of the solution space. The chosen tool, Theil–Sen regression, is also very robust with respect to outliers. The slope of the Theil–Sen regression to the projection of the solution distribution into 2-D objective space is a good indicator for objective compatibility. Incompatible objectives generally show deviations of $\gamma > 45°$ and compatible objectives exhibit deviations of $\gamma < 45°$.

The results for the deviation angles are consistent across individual archives, but cases can occur were the deviation angle results based on different intermediate archives vary significantly. Analysing only individual archives could, therefore, lead to false conclusions about the level of objective compatibility. Archives from the early stage of an inversion in particular often contain only a small number of solutions, yielding misleading results. Hence, a maximal number of solutions should be extracted during inversion runs to be analysed together.

The necessity for a statistical analysis of the solution distributions is illustrated by the multi-objective tests. This case demonstrates that visual inspection can be deceiving and Theil–Sen analysis is required. Inspecting the solution distributions of the MT–well-log projections, the distributions look very similar for the compatible and the incompatible case, but the Theil–Sen analysis detects major differences in the distributions and correctly indicates the compatibility in both cases. The close clustering of a large number of solutions can especially lead to false interpretations, as distribution patterns may be obscured.

The deviation angles show values just above the compatibility threshold, whereas the distances from the ideal point are very low and indicate that acceptable misfits are reached for both objectives. Only the analysis of the misfits for the coinciding frequencies yields a deviation angle below the threshold, with $\gamma = 25°$. These are the kind of solution distributions that also have been found to represent compatible objectives by Moorkamp et al. (2007). In cases like this a thorough visual and numerical analysis of the solution distributions has to be performed. This has to be done carefully, as close clustering of many solutions can give false impressions. In the test situation, the linear fit is dominated by solutions that extend along the δ_{BBI}-direction. This is caused by the BB MT data sets fully constraining the AMT data set, as the BB data have a higher penetration depth, which is expressed in the solution distribution being shifted towards higher BB data misfits. These types of shifted distribution can also be observed in Moorkamp et al. (2010), especially for the combination of Rayleigh wave dispersion data and MT data.

For the assessment of the balance of mutual data constraint, it is important that the misfits of the objectives are of comparable magnitude. The normalisation of the data misfits by the data error (Eq. (2)) ensures that the misfits become comparable to some extend and also guarantees that fits below the error level are expressed in misfits of $\delta < 1$ regardless of the type of data. Nevertheless, different data types and/or varying error levels can influence the deviation measure and may bias the assessment of the mutual constraint balance.

In addition to assessing the compatibility of inverted data sets, the generated solution ensembles can be used to estimate average models and model errors. However, in our case, these are of qualitative rather than objective nature, as it is not statistically possible to extract robust estimates of model covariance from a single solution ensemble generated by a genetic algorithm. The trade-off for fast convergence of genetic algorithms compared to Monte Carlo methods is that the final ensemble is not generated completely independently, but often depends on good models from the early iterations. Robust statistical model averages and model errors can be determined by performing several inversion runs (Stoffa and Sen 1991) or resampling the final solution ensemble (Sambridge 1999b).

The inversion runs for this study had run times of 1–20 min for 100,000–250,000 solution evaluations. General run-time analysis for multi-objective evolutionary algorithms has been performed (Laumanns et al. 2004), but precise run-time predictions are difficult as they are highly problem dependent. The run times are dependent on the number of objectives and the degree of compatibility between the objectives, as compatible objectives make it easier to find well fitting solutions and, therefore, show accelerated convergence. The adaptive nature of the Borg algorithm makes run-time predictions especially difficult, as the variable population and archive sizes and the search restarts performed to mitigate search stagnation can not be projected.

6. Conclusions

Multiple approaches to joint inversion modelling of geophysical data exist, but the application of evolutionary algorithms is not common in this field. The ability to jointly invert a number of data sets without the need for data weighting, while providing model uncertainty and data set compatibility information makes multi-objective approaches advantageous over conventional linearised schemes.

We have developed and implemented a pareto-optimal multi-objective inversion algorithm for the analysis of geophysical data, the advantages of which are as follows. The use of an evolutionary algorithm allows the evaluation of a distribution of solution models. This distribution can be analysed with regard to the physical implications of the model parameters and with respect to the quality of the data. Potential contained ambiguities and resolution restrictions of the data can be expressed in terms of data set compatibility. We have presented a scheme to effectively assess this compatibility. This analysis can be applied independent of the actual modelling part, and it can be combined with other (multi-objective) inversion and modelling software to independently assess data set quality. This can improve the overall data and model interpretation, and it, therefore, is a valuable addition to the general toolbox for geophysical data inversion modelling.

We have demonstrated the capabilities of this algorithm by applying it to synthetic data. By defining different objective functions, the application of the algorithm to other data sets, both synthetic and real, is a straight forward process and does not require major alterations of the code.

Acknowledgements

We thank the developers of the Borg Algorithm, David Hadka and Patrick Reed from the Pennsylvania State University, for making their algorithm available to us (http://borgmoea.org). The work has been supported by the Deep Exploration Technologies Cooperative Research Centre whose activities are funded by the Australian Government's Cooperative Research Centre Programme. We would especially like to thank Rodrigo Bijani for their extensive help refining this work through the review process. This is DET CRC Document (2016/857).

References

Abubakar, A., Li, M., Pan, G., Liu, J., & Habashy, T. M. (2011). Joint MT and CSEM data inversion using a multiplicative cost function approach. *Geophysics*, *76*(3), F203–F214. https://doi.org/10.1190/1.3560898.

Akca, I., Günther, T., Müller-Petke, M., Başokur, A. T., & Yaramanci, U. (2014). Joint parameter estimation from magnetic resonance and vertical electric soundings using a multi-objective genetic algorithm. *Geophysical Prospecting*, *62*(2), 364–376.

Carter-McAuslan, A., Lelièvre, P. G., & Farquharson, C. G. (2014). A study of fuzzy c-means coupling for joint inversion, using seismic tomography and gravity data test scenarios. *Geophysics*, *80*(1), W1–W15.

Coello, C. A. C., Lamont, G. B., & Veldhuizen, D. A. V. (2007). *Evolutionary algorithms for solving multi-objective problems* (2nd ed.). Berlin: Springer.

Commer, M., & Newman, G. A. (2009). Three-dimensional controlled-source electromagnetic and magnetotelluric joint inversion. *Geophysical Journal International*, *178*, 1305–1316. https://doi.org/10.1111/j.1365-246X.2009.04216.x.

Coxeter, H. (1973). *Regular polytopes. Dover books on mathematics series*. Dover Publications, New York. http://books.google.com.au/books?id=iWvXsVInpgMC.

De Stefano, M., Andreasi, F. G., Re, S., Virgilio, M., & Snyder, F. F. (2011). Multiple-domain, simultaneous joint inversion of geophysical data with application to subsalt imaging. *Geophysics*, *76*(3), R69–R80. https://doi.org/10.1190/1.3554652.

Deb, K., Agarwal, R. B. (1994). Simulated binary crossover for continuous search space. Tech. Rep. ITK/ME/SMD-94027, Indian Institute of Technology, Kanpur, UP, India.

Deb, K., & Goyal, M. (1996). A combined genetic adaptive search (GeneAS) for engineering design. *Computer Science and Informatics*, *26*(4), 30–45.

Deb, K., Joshi, D., Anand, A. (2002). Real-coded evolutionary algorithms with parent-centric recombination. In: *Proceedings of the world congress on computational intelligence*, pp. 61–66.

de Nardis, R., Cardarelli, E., & Dobróka, M. (2005). Quasi-2D hybrid joint inversion of seismic and geoelectric data. *Geophysical Prospecting*, *53*, 705–716.

Dobróka, M., Gyulai, Á., Ormos, T., Csókás, J., & Dresen, L. (1991). 2D sections of porosity and water saturation percent from combined resistivity and seismic surveys for hydrogeologic studies. *Geophysical Prospecting*, *39*, 643–665.

Edgeworth, F. Y. (1881). *Mathematical physics: An essay on the application of mathematics to the moral sciences*. Reprints of economic classics, Paul, C. K. . http://books.google.com.au/books?id=s7cJAAAAIAAJ.

Emami Niri, M., & Lumley, D. E. (2015). Simultaneous optimization of multiple objective functions for reservoir modeling. *Geophysics*, *80*(5), M53–M67.

Gallardo, L. A., & Meju, M. A. (2003). Characterization of heterogeneous near-surface materials by joint 2D inversion of dc resistivity and seismic data. *Geophys Res Lett*, *30*(13), 1658. https://doi.org/10.1029/2003GL017370.

Gallardo, L. A., & Meju, M. A. (2007). Joint two-dimensional cross-gradient imaging of magnetotelluric and seismic traveltime data for structural and lithological classification. *Geophys J Int*, *169*, 1261–1272. https://doi.org/10.1111/j.1365-246X.2007.03366.x.

Giancarlo, D. M. (2010). Insights on surface wave dispersion and HVSR: Joint analysis via pareto optimality. *Journal of Applied Geophysics*, *72*(2), 129–140.

Haber, E., & Oldenburg, D. (1997). Joint inversion: A structural approach. *Inverse Problems*, *13*(1), 63.

Hadka, D., & Reed, P. (2013). Borg: An auto-adaptive many-objective evolutionary computing framework. *Evolutionary Computation*, *21*(2), 231–259.

Hanne, T. (1999). On the convergence of multiobjective evolutionary algorithms. *European Journal of Operational Research*, *117*(3), 553–564. https://doi.org/10.1016/S0377-2217(98)00262-8.

Higuchi, T., Tsutsui, S., & Yamamura, M. (2000). Theoretical analysis of simplex crossover for real-coded genetic algorithms. In: *Parallel problem solving from nature PPSN VI, lecture notes in computer science*, vol. 1917. Berlin: Springer, pp. 365 – 374. https://doi.org/10.1007/3-540-45356-3_36.

Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. University of Michigan Press. http://books.google.com.au/books?id=JE5RAAAAMAAJ.

Ishibuchi, H., Tsukamoto, N., Hitotsuyanagi, Y., & Nojima, Y. (2008). Effectiveness of scalability improvement attempts on the performance of NSGA-II for many-objective problems. In: Keijzer, M., Antoniol, G., Congdon, C. B., Deb, K., Doerr, B., Hansen, N., Holmes, J. H., Hornby, G. S., Howard, D., Kennedy, J., Kumar, S., Lobo, F. G., Miller, J. F., Moore, J., Neumann, F., Pelikan, M., Pollack, J., Sastry, K., Stanley, K., Stoica, A., Talbi, E. G., & Wegener, I. (eds) *GECCO '08: Proceedings of the 10th annual conference on Genetic and evolutionary computation, ACM, Atlanta, GA, USA*, pp. 649–656. http://www.cs.bham.ac.uk/∼wbl/biblio/gecco2008/docs/p649.pdf.

Jegen, M. D., Hobbs, R. W., Tarits, P., & Chave, A. (2009). Joint inversion of marine magnetotelluric and gravity data incorporating seismic constraints—preliminary results of sub-basalt imaging off the Faroe Shelf. *Earth and Planetary Science Letters*, *282*, 47–55. https://doi.org/10.1016/j.epsl.2009.02.018.

Julià, J., Ammon, C. J., Herrmann, R. B., & Correig, A. M. (2000). Joint inversion of receiver function and surface wave dispersion observations. *Geophysical Journal International*, *143*, 99–112.

Kita, H., Ono, I., & Kobayashi, S. (1999). Multi-parental extension of the unimodal normal distribution crossover for real-coded genetic algorithms. In: *Proceedings of the 1999 congress on evolutionary computation*, pp. 1581–1588.

Kozlovskaya, E., Vecsey, L., Plomerová, J., & Raita, T. (2007). Joint inversion of multiple data types with the use of multiobjective optimization: Problem formulation and application to the seismic anisotropy investigations. *Geophysical Journal International*, *171*(2), 761–779. https://doi.org/10.1111/j.1365-246X.2007.03540.x.

Laumanns, M., Thiele, L., Deb, K., & Zitzler, E. (2002). Combining convergence and diversity in evolutionary multi-objective optimization. *Evolutionary Computation*, *10*(3), 263–282.

Laumanns, M., Thiele, L., & Zitzler, E. (2004). Running time analysis of multiobjective evolutionary algorithms on pseudo-boolean functions. *Evolutionary Computation IEEE Transactions on*, *8*(2), 170–182. https://doi.org/10.1109/TEVC.2004.823470.

Lawson, C. L., & Hanson, R. J. (1974). Solving least squares problems. Classics in Applied Mathematics, Society for Industrial and Applied Mathematics. http://books.google.com.au/books?id=ROw4hU85nz8C.

Lewis, R. M., Torczon, V., & Trosset, M. W. (2000). Direct search methods: Then and now. *Journal of Computational and Applied Mathematics*, *124*(1–2), 191–207. https://doi.org/10.1016/S0377-0427(00)00423-4.

Marler, R. T., & Arora, J. S. (2010). The weighted sum method for multi-objective optimization: New insights. *Structural and Multidisciplinary Optimization*, *41*(6), 853–862.

McKean, J. W. (2004). Robust analysis of linear models. *Statistical Science*, *19*(4), 562–570. https://doi.org/10.1214/088342304000000549.

Miller, B. L., & Goldberg, D. E. (1995). Genetic algorithms, tournament selection, and the effects of noise. *Complex Systems*, *9*, 193–212.

Moorkamp, M., Jones, A. G., & Eaton, D. W. (2007). Joint inversion of teleseismic receiver functions and magnetotelluric data using a genetic algorithm: Are seismic velocities and electrical conductivities compatible? *Geophysical Research Letters*, *34*(L16311), https://doi.org/10.1029/2007GL030519.

Moorkamp, M., Jones, A. G., & Fishwick, S. (2010). Joint inversion of receiver functions, surface wave dispersion, and magnetotelluric data. Journal of Geophysical Research. https://doi.org/10.1029/2009JB006369.

Moorkamp, M., Heincke, B., Jegen, M., Roberts, A. W., & Hobbs, R. W. (2011). A framework for 3-D joint inversion of MT, gravity and seismic refraction data. *Geophysical Journal International*, *184*, 477–493. https://doi.org/10.1111/j.1365-246X.2010.04856.x.

Mota, R., & Santos, F. M. D. (2006). 2D sections of porosity and water saturation percent from combined resistivity and seismic surveys for hydrogeologic studies. *The Leading Edge*, *25*(6), 735–737.

Muñoz, G., & Rath, V. (2006). Beyond smooth inversion: The use of nullspace projection for the exploration of non-uniqueness in MT. *Geophysical Journal International*, *164*, 301–311. https://doi.org/10.1111/j.1365-246X.2005.02825.x.

Paasche, H., & Tronicke, J. (2014). Nonlinear joint inversion of tomographic data using swarm intelligence. *Geophysics*, *79*(4), R133–R149.

Pareto, V. (1896). Cours d'Économie Politique, vol I and II. F. Rouge, Lausanne.

Pek, J., & Santos, F. A. (2006). Magnetotelluric inversion for anisotropic conductivities in layered media. *Physics of the Earth and Planetary Interiors*, *158*(2–4), 139–158. https://doi.org/10.1016/j.pepi.2006.03.023.

Purshouse, R. C., & Fleming, P. J. (2007). On the evolutionary optimization of many conflicting objectives. *IEEE Transactions on Evolutionary Computation*, *11*(6), 239–245.

Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D Nonlinear Phenomena*, *60*(1–4), 259–268. https://doi.org/10.1016/0167-2789(92)90242-F.

Sambridge, M. (1999a). Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space. *Geophysical Journal International*, *138*(2), 479–494. https://doi.org/10.1046/j.1365-246X.1999.00876.x.

Sambridge, M. (1999b). Geophysical inversion with a neighbourhood algorithm—II. Appraising the ensemble. *Geophysical Journal International*, *138*(3), 727–746. https://doi.org/10.1046/j.1365-246x.1999.00900.x.

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, *63*(324), 1379–1389.

Stoffa, P. L., & Sen, M. K. (1991). Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms. *Geophysics*, *56*(11), 1794–1810.

Storn, R., & Price, K. (1997). Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Geophysical Journal International*, *11*(4), 341–359.

Syswerda, G. (1989). Uniform crossover in genetic algorithms. In: Schaffer, D. J. (ed) *Proceedings of the third international conference on genetic algorithms*, pp. 2–9.

Tang, Y., Reed, P., & Wagener, T. (2006). How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? *Hydrology and Earth System Sciences*, *10*(2), 289–307. https://doi.org/10.5194/hess-10-289-2006.

Theil H (1950) A rank-invariant method of linear and polynomial regression analysis. I. Proceedings of Koninalijke Nederlandse Akademie 53:386–392. Indagationes Math. 12, 85–91.

Treitel, S., & Lines, L. R. (1999). *Past, present and future of geophysical inversion—a y2k analysis*. CREWES Research Report Volume 11, Indian Institute of Technology.

Tsutsui, S., Yamamura, M., & Higuchi, T. (1999). Multi-parent recombination with simplex crossover in real coded genetic algorithms. In: *Proceedings of the genetic and evolutionary computation conference (GECCO 1999)*, pp. 657–664.

Vrugt, J. A., & Robinson, B. A. (2007). Improved evolutionary optimization from genetically adaptive multimethod search. *Proceedings of the National Academy of Sciences*, *104*(3), 708–711.

Vrugt, J. A., Robinson, B. A., & Hyman, J. M. (2009). Self-adaptive multimethod search for global optimization in real-parameter spaces. *IEEE Transactions on Evolutionary Computation*, *13*(2), 243–259.

Wait, J. R. (1954). On the relation between telluric currents and the earth's magnetic field. *Geophysics*, *19*(2), 281–289. https://doi.org/10.1190/1.1437994.

Yang, C. H., & Tong, L. T. (1988). Joint inversion of DC, TEM, and MT data. In: *58th annual international meeting*, Taiwan, Society of Exploration Geophysicists: National Central University, pp. 408–410.

Zhou, A., Qu, B. Y., Li, H., Zhao, S. Z., Suganthan, P. N., & Zhang, Q. (2011). Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, *1*(1), 32–49. https://doi.org/10.1016/j.swevo.2011.03.001.

Zitzler, E. (1999). Evolutionary algorithms for multiobjective optimization: Methods and applications. PhD Thesis, ETH Zurich, Switzerland.

Zitzler, E., Deb, K., & Thiele, L. (2000). Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary Computation*, *8*(2), 173–195.

(Received May 31, 2017, revised January 10, 2018, accepted January 22, 2018)

APPENDIX
B

Rees, N., Heinson, G., & Conway, D. (2018).
Monitoring hydraulic stimulation using telluric sounding.
Earth, Planets and Space, 70(1), 7.

This paper details a new technique for analysing monitoring data from the telluric sounding method using singular value decomposition and eigenvalue analysis. Personal contributions include assistance with modelling and manuscript editing.

Rees *et al. Earth, Planets and Space* (2018) 70:7
https://doi.org/10.1186/s40623-017-0767-3

Earth, Planets and Space

FULL PAPER Open Access

Monitoring hydraulic stimulation using telluric sounding

Nigel Rees[1*], Graham Heinson[2] and Dennis Conway[2]

Abstract

The telluric sounding (TS) method is introduced as a potential tool for monitoring hydraulic fracturing at depth. The advantage of this technique is that it requires only the measurement of electric fields, which are cheap and easy when compared with magnetotelluric measurements. Additionally, the transfer function between electric fields from two locations is essentially the identity matrix for a 1D Earth no matter what the vertical structure. Therefore, changes in the earth resulting from the introduction of conductive bodies underneath one of these sites can be associated with deviations away from the identity matrix, with static shift appearing as a galvanic multiplier at all periods. Singular value decomposition and eigenvalue analysis can reduce the complexity of the resulting telluric distortion matrix to simpler parameters that can be visualised in the form of Mohr circles. This technique would be useful in constraining the lateral extent of resistivity changes. We test the viability of utilising the TS method for monitoring on both a synthetic dataset and for a hydraulic stimulation of an enhanced geothermal system case study conducted in Paralana, South Australia. The synthetic data example shows small but consistent changes in the transfer functions associated with hydraulic stimulation, with grids of Mohr circles introduced as a useful diagnostic tool for visualising the extent of fluid movement. The Paralana electric field data were relatively noisy and affected by the dead band making the analysis of transfer functions difficult. However, changes in the order of 5% were observed from 5 s to longer periods. We conclude that deep monitoring using the TS method is marginal at depths in the order of 4 km and that in order to have meaningful interpretations, electric field data need to be of a high quality with low levels of site noise.

Keywords: Telluric sounding, Hydraulic stimulation, Monitoring, Transfer functions

Introduction

The telluric sounding (TS) method was introduced in the 1960s and involves simultaneously recording the horizontal components of electric fields (**E**) at different sites (Berdičevskij and Keller 1965; Yungul 1966). The measured **E** can be affected by galvanic distortion, which is caused by gradients in electrical conductivity associated with near-surface heterogeneities (Chave and Smith 1994). The accumulation of charge at conductivity boundaries strongly alters **E** (Groom and Bailey 1989). This local distortion of **E** can be described by a real-valued second-rank tensor **D** that relates the electric field measured at a local ($\mathbf{E}_M$) and regional ($\mathbf{E}_B$) site according to (Chave and Jones 2012)

$$\mathbf{E}_M = \mathbf{D}\mathbf{E}_B \tag{1}$$

where

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix}. \tag{2}$$

In this case, **D** is relative to axes X and Y, typically north and east. Introducing a rotation matrix $\mathbf{R}(\theta)$,

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \tag{3}$$

and the transpose $\mathbf{R}^{\mathbf{T}}(\theta)$, the value of **D** relative to axes X' and Y' (which are rotated θ' clockwise from north and east) is given by the matrix $\mathbf{D}'$ (Lilley 2015)

$$\begin{bmatrix} D'_{xx} & D'_{xy} \\ D'_{yx} & D'_{yy} \end{bmatrix} = \mathbf{R}(-\theta') \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix} \mathbf{R}(\theta'). \tag{4}$$

*Correspondence: nigel.rees@anu.edu.au
[1] The Australian National University, Canberra, ACT 0200, Australia
Full list of author information is available at the end of the article

For a 1D Earth, the distortion matrix will be the identity matrix. The presence of galvanic distortion will be manifest as an amplitude shift, and a twist and shear operation (Lilley 2015). For 2D structures, the distortion matrix can be rotated to reflect the changes in **E** along and across strike.

Targets for hydraulic stimulation are generally in sedimentary basins that are laterally extensive, and for most of the bandwidth the responses are 1D. The injection of conductive fluids into the subsurface will alter the telluric distortion matrix which can be mapped to show the lateral constraints of fluid migration. Lilley (2015) proposed the application of both eigenvalue analysis (EA) and singular value decomposition (SVD) on the telluric distortion matrix, with Mohr diagrams introduced as a versatile way of visualising properties of the matrix. These diagrams can be used to determine the extent to which the matrix is diagnostic of 1D, 2D or 3D geological structure as well as determining a strike direction (with 90° ambiguity) and a relative amplitude change. Grids of Mohr circles can show where the greater amplitude changes occur at depth and may be used to determine the extent to which fluid has migrated from the injection point.

Following from Lilley (2015), EA of the telluric distortion matrix involves finding a direction of $\mathbf{E}_M$ for which the change in $\mathbf{E}_B$ is in the same direction, with the eigenvalue of the direction giving the gain of the process. When real eigenvectors exist, the characteristic equation for **D**,

$$\zeta^2 - (D_{xx} + D_{yy})\zeta + D_{xx}D_{yy} - D_{xy}D_{yx} = 0, \tag{5}$$

can be solved for the two eigenvalues ζ_1 and ζ_2 with solutions

$$\zeta_1, \zeta_2 = \frac{1}{2}(D_{xx} + D_{yy}) \pm \frac{1}{2}\sqrt{(D_{xx} + D_{yy})^2 + 4(D_{xy}D_{yx} - D_{xx}D_{yy})}. \tag{6}$$

Eigenvectors can then be found corresponding to these eigenvalues. The solutions to EA can be plotted on a Mohr circle with centre $[(D_{xx} + D_{yy})/2, (D_{xy} - D_{yx})/2]$ and radius $r = \sqrt{(D_{xy} + D_{yx})^2 + (D_{xx} - D_{yy})^2}$. Axes for D'_{yx} and D'_{yy} can also be plotted to display the variation with axis rotation of all components of $\mathbf{D}'$. Figure 1a shows an example of such a Mohr diagram, where P represents the observed point, H and J mark the EA positions and C represents the centre of the circle. The eigenvalues are the D'_{xx} axis values of H and J.

The SVD of **D** produces a rotation for the axes at the regional site B and a different rotation for the axes at local site M. Therefore, a change in **E** at rotated site B produces a change in **E** along the corresponding but differently rotated axis at site M, with the change generally amplified or attenuated (Lilley 2015). This rotation of the local and regional axes reduces the telluric distortion tensor to an ideal 2D form. The results of SVD can also be displayed on a Mohr diagram as shown in Fig. 1b. Here, OG and OF represent the greater and lesser singular values, P is the observed point and C is the centre. Grids of SVD Mohr circles can show the magnitudes and directions of the injected near surface anomaly. For a detailed explanation on the theory of EA, SVD and Mohr diagrams related to the telluric distortion matrix, the reader is referred to Lilley (1993, 2012, 2015).

The goal of this study is to test the viability of utilising the TS method for monitoring hydraulic stimulation at depth. Ideally we would like to be able to constrain the spatial and temporal dimensions of resistivity changes. Spatial changes can be constrained laterally from using multiple sites, and depths can be estimated given resistivity data to map periods to depths. One advantage of the TS method is that it is relatively easy to measure **E** with many dipoles and multi-channel systems and therefore **E** arrays could be deployed for continuous monitoring. Additionally, hydraulic stimulation targets are generally laterally extensive sedimentary basins where **E** transfer functions are essentially the identity matrix. Therefore, monitoring would involve plotting deviations relative to the identity matrix, with static shift appearing as a galvanic multiplier at all periods. The impedance on the other hand has real and imaginary components that vary with frequency. Finally, the TS method is relatively low cost when compared with traditional magnetotelluric measurements and may prove a simple and favourable method for monitoring fluid movement.

3D feasibility study

A 3D feasibility study was conducted to test the viability of the TS method for monitoring conductive change at depth. The 3D forward modelling code of Mackie et al. (1993) was used to create a baseline resistivity structure as shown in Fig. 2. The baseline resistivity model consisted of a 10 Ωm conductive layer down to 0.8 km, a 50 Ω m layer from 0.8 to 2 km, a 100 Ωm layer from 2 to 4.5 km and finally a 1000 Ωm layer from 4.5 to 1200 km. The stimulation model introduces a 1 Ωm conductive block at 3.6 km depth, with volumetric dimensions of $3 \times 1 \times 0.4$ km.

3D forward modelling for both base and stimulation resistivity structures produced classical MT responses (see Additional file 1). From the forward modelling code and the following equation, we calculate **D**

$$\mathbf{Z}_M = \mathbf{D}\mathbf{Z}_B \tag{7}$$

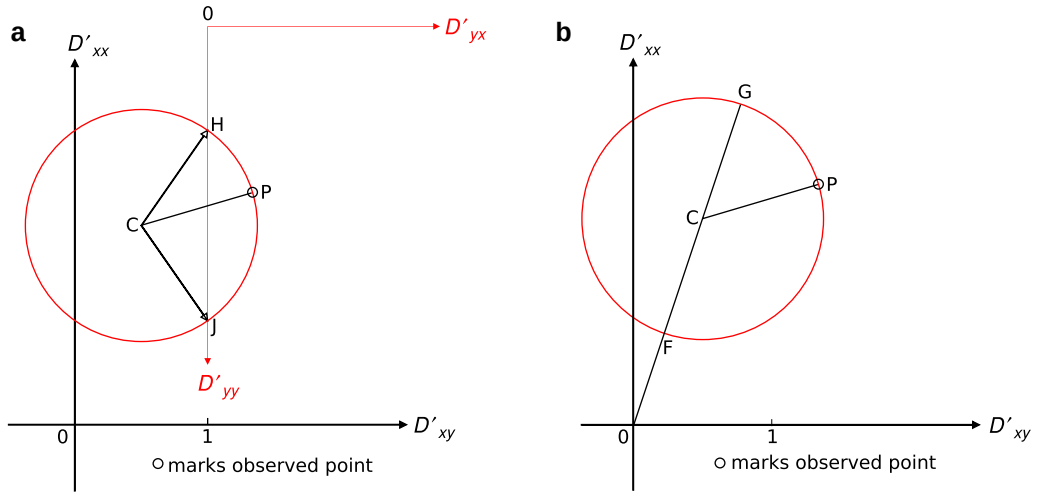

Fig. 1 **a** Example eigananalysis Mohr diagram adapted from Lilley (2015). Points H and J mark the eigenanalysis positions, and the eigenvalues are the D'_{xx} values of points H and J. **b** Example Mohr diagram for singular value decomposition adapted from Lilley (2015). OG and OF represent the greater and lesser singular values

where $\mathbf{Z}_M$ and $\mathbf{Z}_B$ are the local and regional site impedance tensors, respectively, and $\mathbf{D}$ is a 2×2 matrix. If we multiply both sides by a regional and uniform $\mathbf{B}$ field (i.e. $\mathbf{B}_B = \mathbf{B}_M$), then

$$\mathbf{Z}_M\mathbf{B} = \mathbf{D}\mathbf{Z}_B\mathbf{B}. \tag{8}$$

As $\mathbf{E} = \mathbf{Z}\mathbf{B}$, Eq. 8 can be written as

$$\mathbf{E}_M = \mathbf{D}\mathbf{E}_B. \tag{9}$$

Therefore, the distortion matrix calculated from the impedance tensors is the same as the impedance matrix calculated from the electric fields.

Figures 3, 4, 5 and 6 show each component of the resultant transfer functions (D_{xx}, D_{xy}, D_{yx}, D_{yy}; here the first subscript represents the local site and the second the reference site) between the $\mathbf{E}_B$ and $\mathbf{E}_M$ sites for the six lines in the synthetic grid.

The baseline transfer functions represent the identity matrix, where D_{xx} and D_{yy} are 1 and D_{xy} and D_{yx} are 0 for all periods. The stimulation D_{xx} and D_{yy} show changes of between 1 and 2%, with these changes occurring from about 5 s and continuing to longer periods. The changes in D_{yy} are slightly larger when compared with D_{xx}, and the stations furthest away from the injection (e.g. station 59) show negligible changes. The changes in D_{xy} and D_{yx} are generally less than 1%.

Changes in transfer functions (and hence telluric distortion matrices) can be analysed using SVD and EA and plotted onto Mohr circles grids [see Lilley (2015)]. Such a grid is shown in Fig. 7, where SVD analysis was performed on telluric distortion matrices comparing $\mathbf{E}_B$ and $\mathbf{E}_M$ along Line 3 during stimulation. The grey vertical and horizontal lines represent the D'_{xx} and D'_{xy} axes, respectively. The black radius interacts with the red circle at the observed point. The Mohr circles increase in diameter for stations surrounding the 1 Ω m conductive body (stations 32–34) at periods of 3 s and greater. The size of the Mohr circles progressively become smaller and approach the identity matrix moving further east from station 34. Similar grids can be plotted using EA as shown in Fig. 8. Here the dashed red lines are the D'_{yy} axis and the two black radii represent the eigenvalues. Notice that for the sites surrounding the conductive zone, the eigenvalues are real and different and the eigenvectors are orthogonal which is a representation of the 2D case [see Fig. 4 from Lilley (2015) for a detailed explanation on Mohr diagram eigenanalysis using different matrices as examples].

Another interesting way of viewing telluric distortion matrix changes is by mapping Mohr circles at specific periods for each site in the array. Such a representation is shown in Fig. 9, where SVD Mohr circles are drawn for every site at a period of 5 s. The 1 Ωm conductive body lies between Lines 3 and 4, and the Mohr circles surrounding the body increase in size and amplitude. Line 6 is the furthest away from the conductive body, and all Mohr circles on this line are the identity matrix. The edge of the conductive body is horizontally well marked along lines 2 and 3. However, the boundary is not as well marked when you look vertically in Fig. 9, where the effects of the conductive block are also seen on Lines 1 and 4. This is also noticeable in Fig. 6 where the transfer functions show poorer cross-line constraints on the lateral extent. Figure 3 shows slightly better cross-line

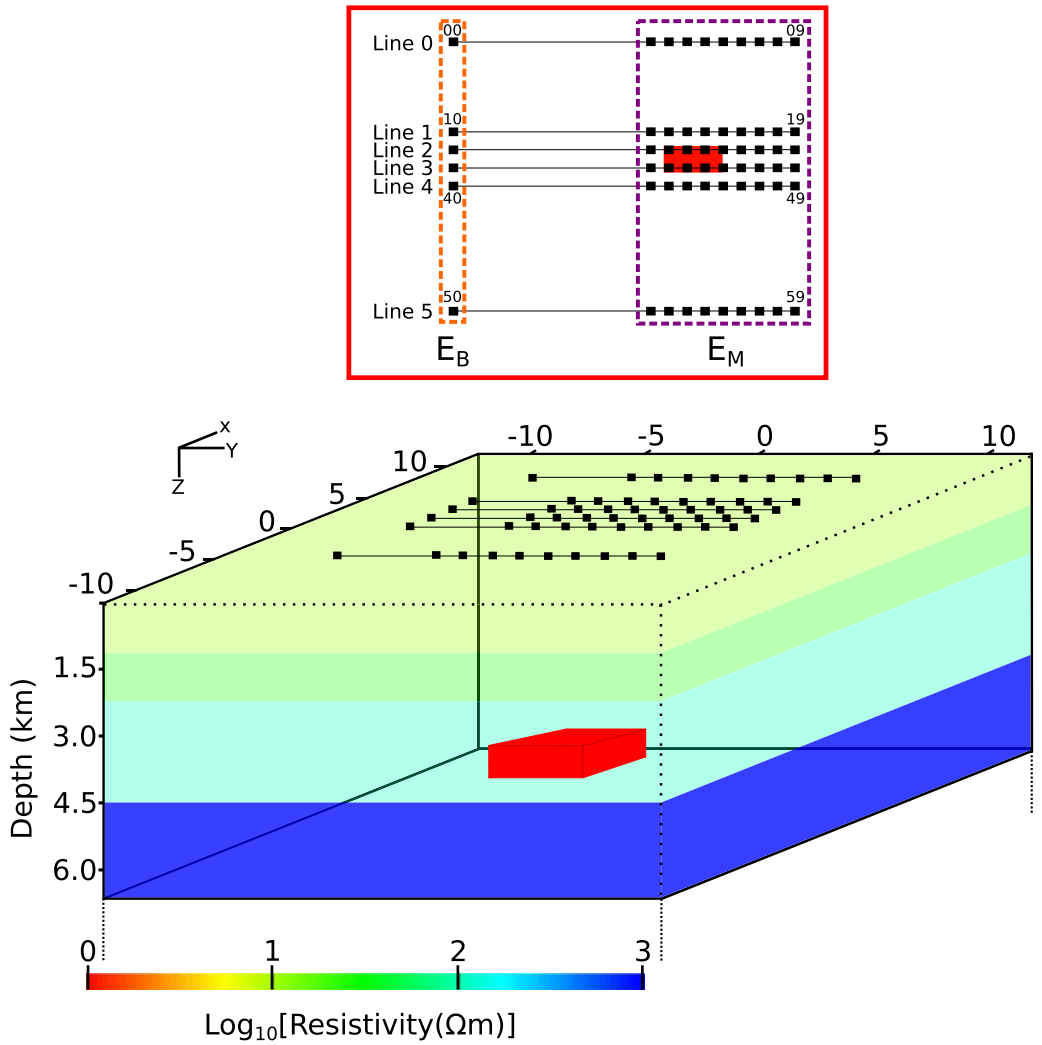

Fig. 2 Resistivity model used for the 3D synthetic forward modelling study. The red box shows a birds eye view of the six lines, with each black square representing one site. The orange dashed box shows the respective base sites ($\mathbf{E}_B$) used for each line, and the dashed purple box shows the measured sites ($\mathbf{E}_M$). The location of the red conductive block at 3.6 km depth is shown

constraints, where Lines 1 and 4 show reduced maximums in transfer functions when compared with Lines 2 and 3.

This synthetic study demonstrates different techniques for mapping changes in transfer functions resulting from changes in electric fields caused by hydraulic stimulation. The changes in transfer functions were observed to be in the order of 1–2%, and occurred in D_{xx} and D_{yy}. The changes in D_{xy} and D_{yx} were generally less than 1%. The resultant telluric distortion matrix changes can be analysed using SVD and EA and mapped onto Mohr circle grids. The Mohr circles were found to increase in diameter at periods and areas associated with the conductive anomaly. In order to sufficiently monitor hydraulic stimulation, surveys should be designed such that lines of instruments extend sufficiently beyond the stimulation zone in both horizontal and (especially) vertical directions. Interpretations regarding the extent of the stimulation should be conservative and take into account the lack of cross-line constraint. In general, changes in telluric distortion resulting from hydraulic fracturing are small and for a real survey it may be difficult to see such changes due to noise in $\mathbf{E}_B$ and $\mathbf{E}_M$.

Example from Paralana, South Australia

In July 2011, Peacock et al. (2013) conducted a magnetotelluric (MT) survey in Paralana, South Australia, with the aim of continuously monitoring changes in MT responses associated with the introduction of saline hydraulic fracturing fluids at depth. The test site was a geothermal reservoir containing hot granites, with the heat source coming from radiogenic elements within the

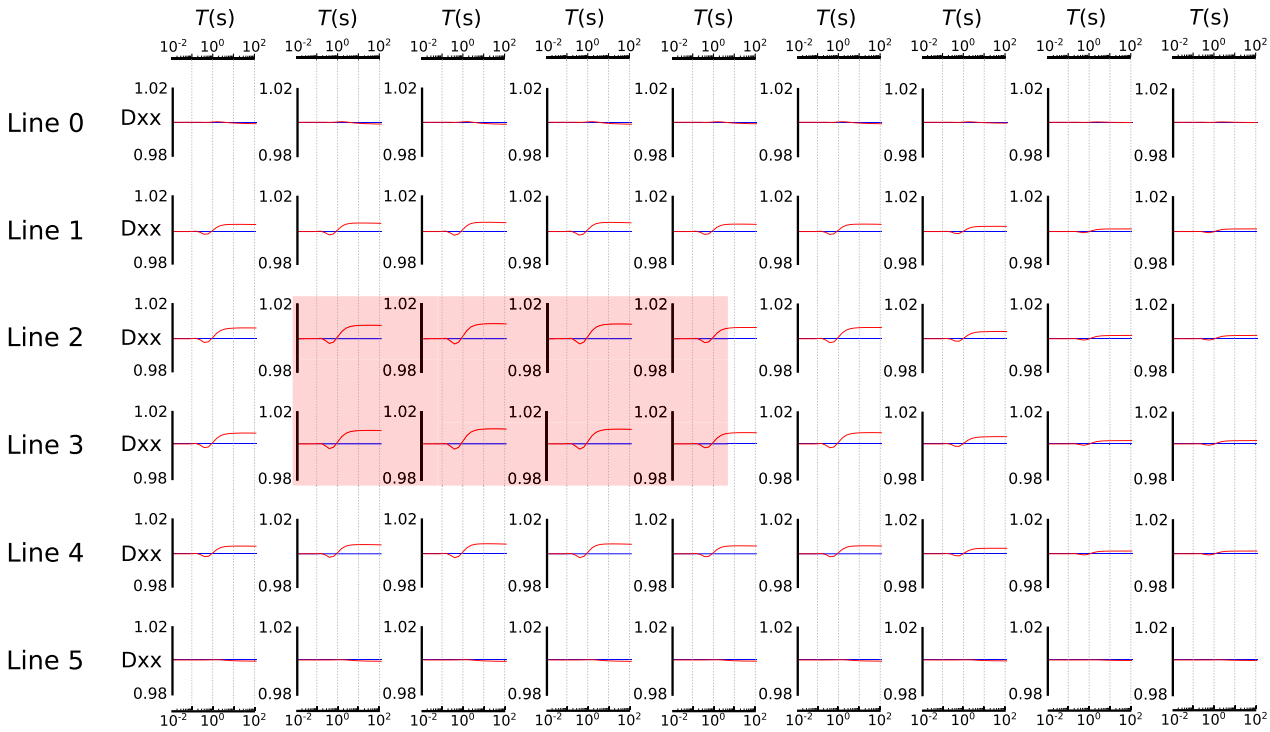

Fig. 3 Synthetic transfer functions ($\mathbf{D}_{xx}$) comparing base (orange dashed box in Fig. 2) and measured (purple dashed box in Fig. 2) sites for Lines 1–6. The *x*-axis is the period in seconds. The blue transfer functions represent the baseline model, and red transfer functions result from the introduction of the 1 Ωm conductive body at 3.6 km depth (represented by the red shaded rectangle). Note that only the real components of the transfer functions are plotted

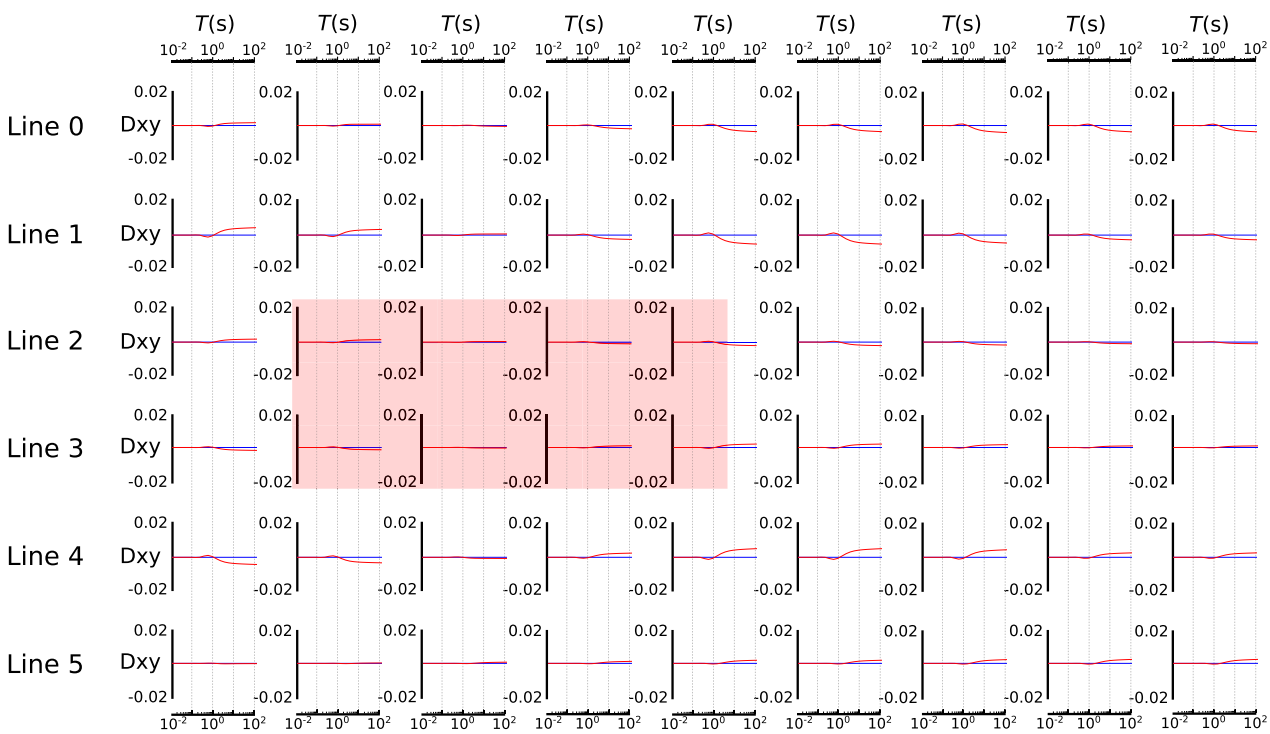

Fig. 4 Synthetic transfer functions ($\mathbf{D}_{xy}$) for the six lines

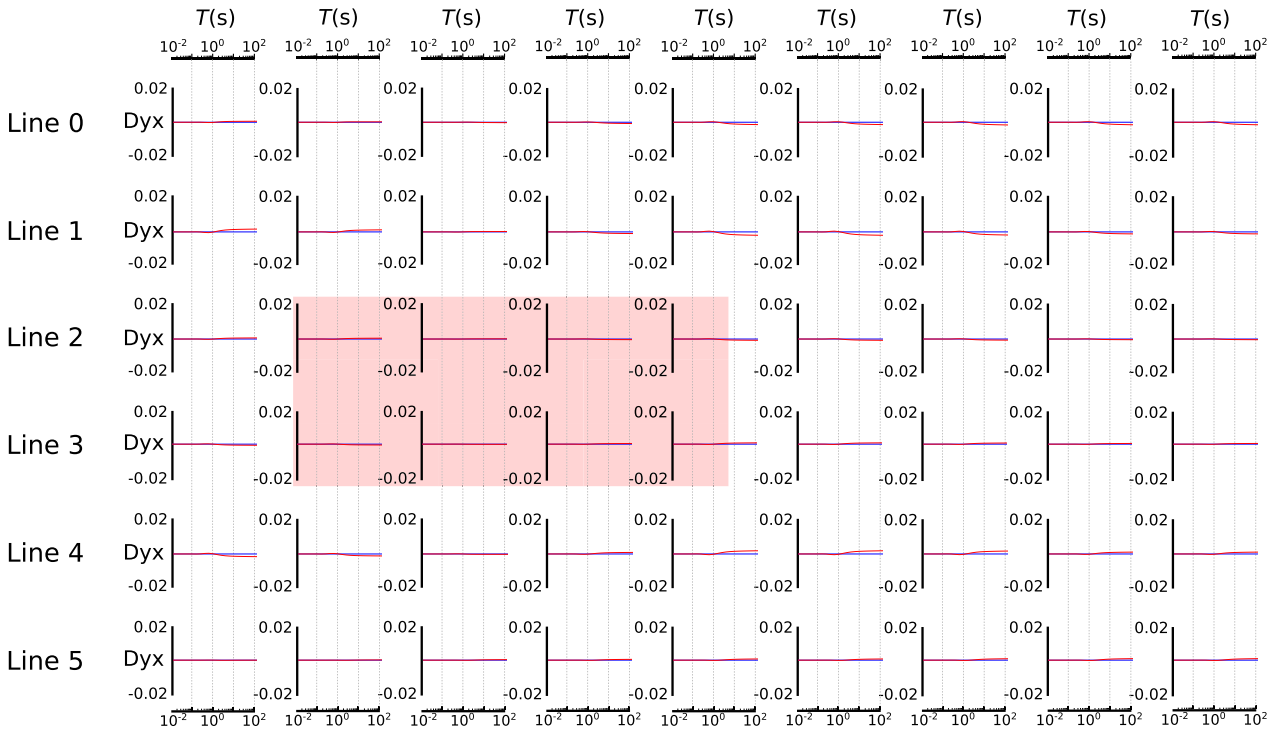

Fig. 5 Synthetic transfer functions ($\mathbf{D}_{yx}$) for the six lines

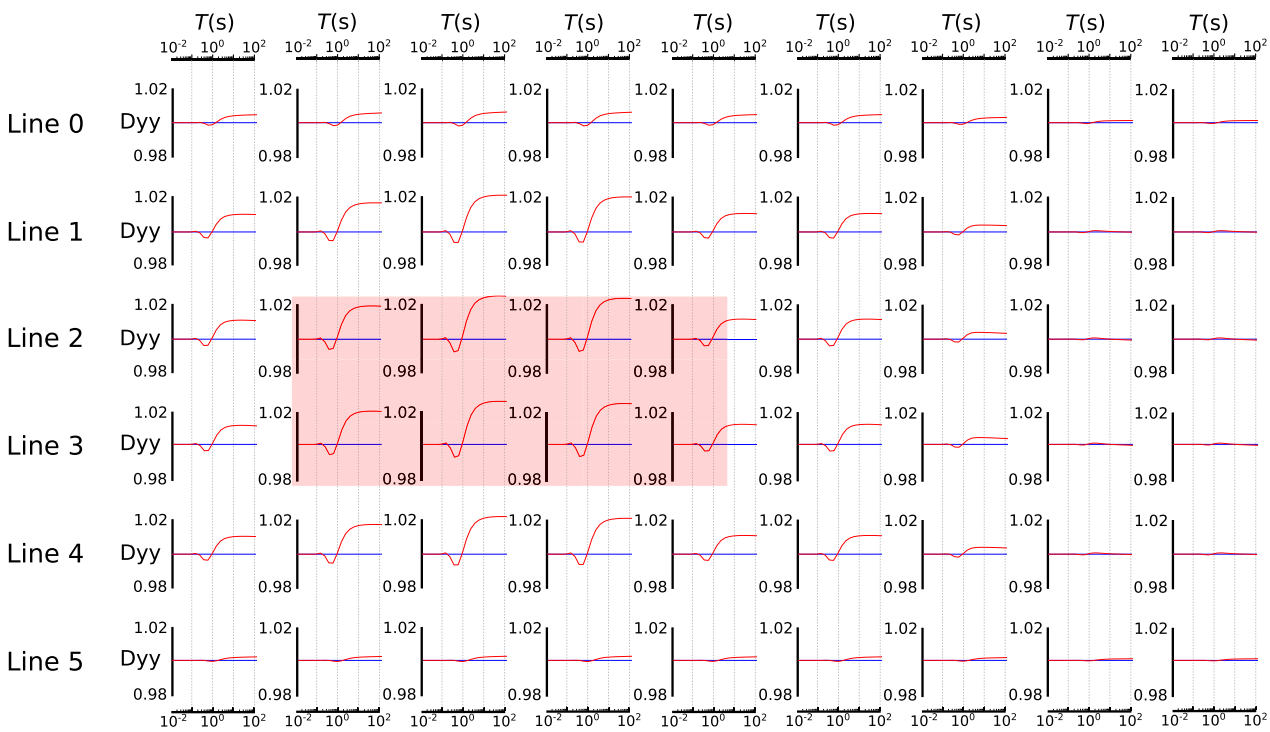

Fig. 6 Synthetic transfer functions ($\mathbf{D}_{yy}$) for the six lines

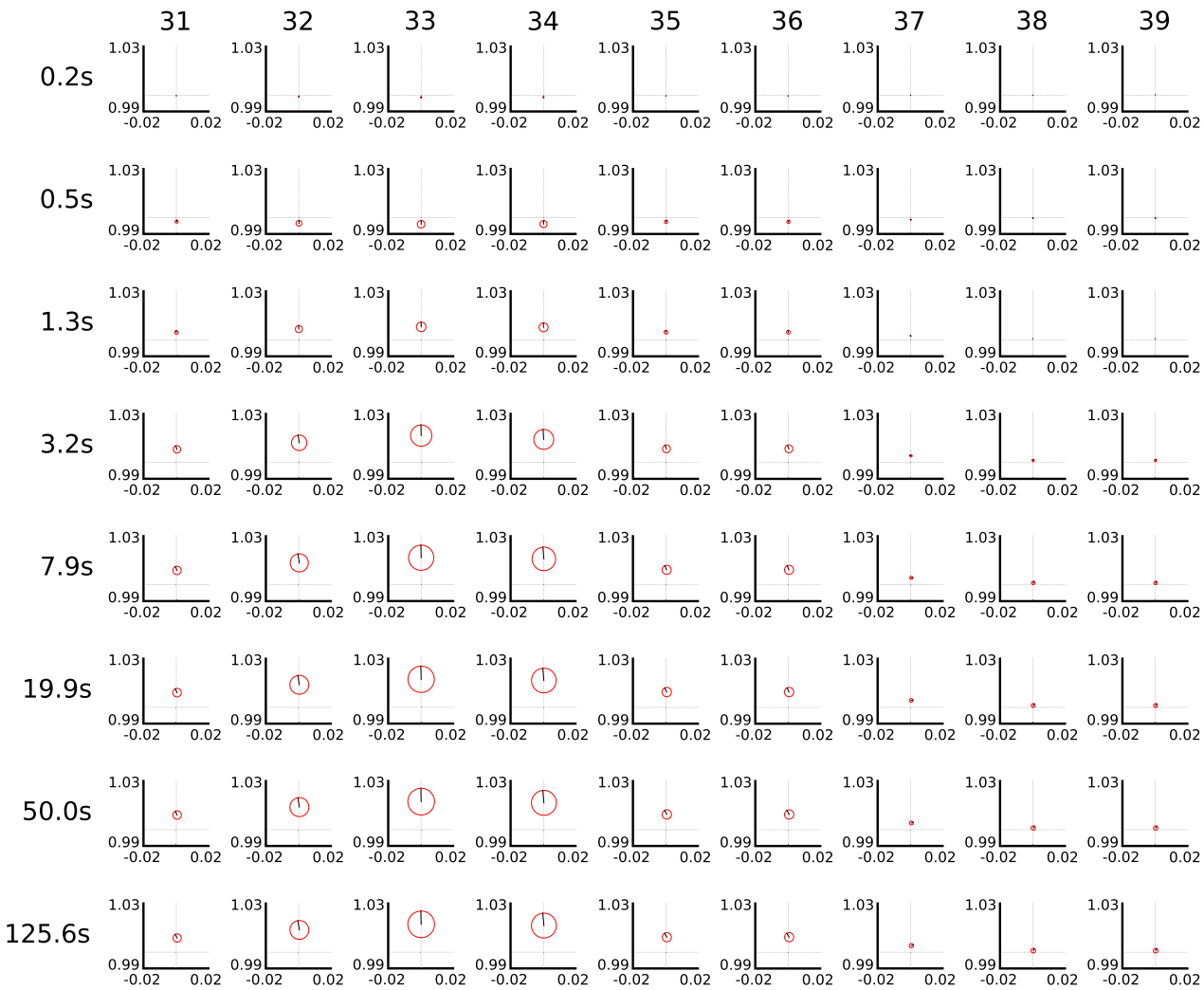

Fig. 7 Synthetic SVD Mohr circles [see Lilley (2015)] for the Line 3 telluric distortion matrices. Grey vertical lines represent the D'_{xx} axes, and grey horizontal dashed lines represent D'_{xy} axes. The interaction of the black radius with the red circle marks the observed point

Paleoproterozoic to Mesoproterozoic gneiss, granites and metasediments of the Mount Painter Domain that underlay the Flinders Rangers (McLaren et al. 2003; Brugger et al. 2005). 3.1 million litres of saline water (resistivity of 0.3 Ωm) was injected at a depth of 3680 m, with the injection beginning on day 193 at 0400 universal time (UT) and taking 4 days to complete (Peacock et al. 2012). The survey layout is shown in Fig. 10, with the microseismic cloud visible in the bottom right corner.

Phase tensor (PT) and resistivity tensor (RT) residuals were used as a diagnostic tool for determining directional fluid migration. The changes observed in residual PT and RT were interpreted as fracturing fluids migrating towards the northeast of the injection well along an existing fault system trending north-northeast. This section attempts to analyse the impedance matrix **D** (calculated from the electric fields) from the Paralana experiment to determine whether electric fields alone can be used to monitor hydraulic stimulation at depth.

For this study, station 04 is used as the measured electric field ($\mathbf{E}_M$) and stations 13 and 27 as the base electric fields ($\mathbf{E}_B$). Each station recorded responses during the hydraulic stimulation from days 193–197 (see Additional File 1. Note that day 195 was excluded as the signal strength was low and hence the responses are dominated by noise). Transfer functions comparing stations 13 ($\mathbf{E}_B$) and 04 ($\mathbf{E}_M$) over the stimulation time interval are shown in Fig. 11. The grey shaded area represents the MT dead band, where signal power is naturally low, especially from 2009 to 2011 (Peacock et al. 2013). The blue horizontal lines represent the theoretical baseline transfer functions, with the pink shaded area representing a 1% noise uncertainty.

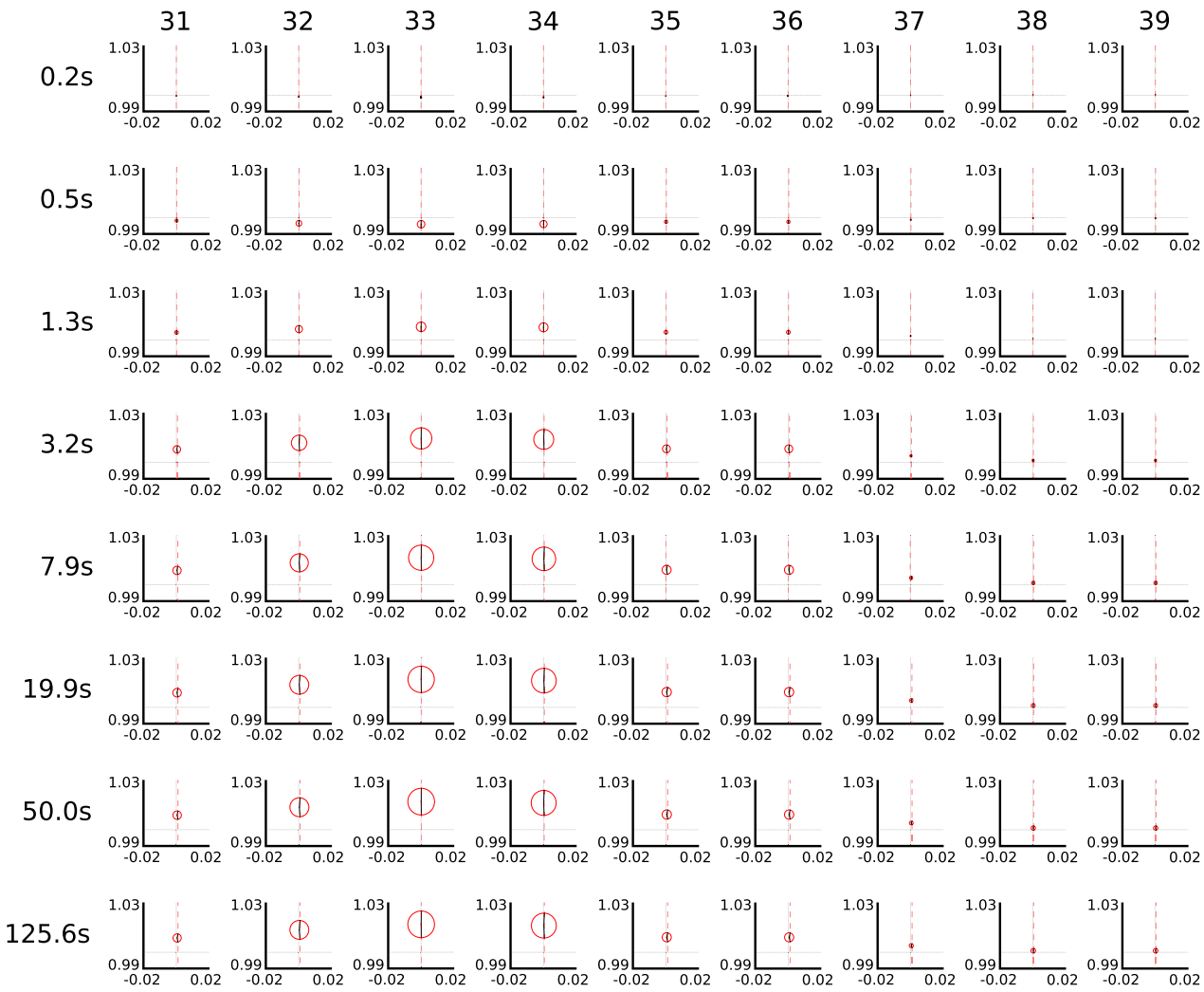

Fig. 8 Synthetic EA Mohr circles [see Lilley (2015)] for the Line 3 telluric distortion matrices

Note that only the real components of the transfer functions are plotted. D_{yy} shows similar behaviour to the feasibility study, where increases of approximately 2% above the noise window occur from 5 s and continue to longer periods for all days. D_{yx} is mostly within the $\pm 1\%$ noise window for all days. D_{xy} shows an almost 5% increase on day 194, with the other days mostly within the noise window. Finally, D_{xx} decreases by $< 1\%$ on days 193 and 196.

Transfer functions comparing stations 27 ($\mathbf{E}_B$) and 04 ($\mathbf{E}_M$) over the same time interval are shown in Fig. 12. Similar trends to Fig. 11 are observed, where an increase in D_{yy} exists in the order of 5% from 5 s to longer periods for all days. D_{xy} shows a 2% increase over a similar period range to D_{yy}. Both D_{xx} and D_{xy} are within the noise windows for most days.

SVD and EA is performed on the telluric distortion matrices for the four days, with the results plotted onto Mohr circle grids as shown in Figs. 13 and 14, respectively. Although the site noise and dead band add complexities to the analysis, the general trend shows the Mohr circles increase in diameter for periods greater than 5 s. When comparing stations 04 and 27, an interesting trend exists for periods between 2 and 14 s, where the Mohr circles progressively increase in diameter over the length of the pumping interval. EA shows eigenvalues that are real and different, with eigenvectors trending orthogonally indicating the telluric distortion matrix is close to 2D. This differs when comparing stations 04 and 13 where the eigenvectors are real but not orthogonal, which may be indicative of a general 3D structure.

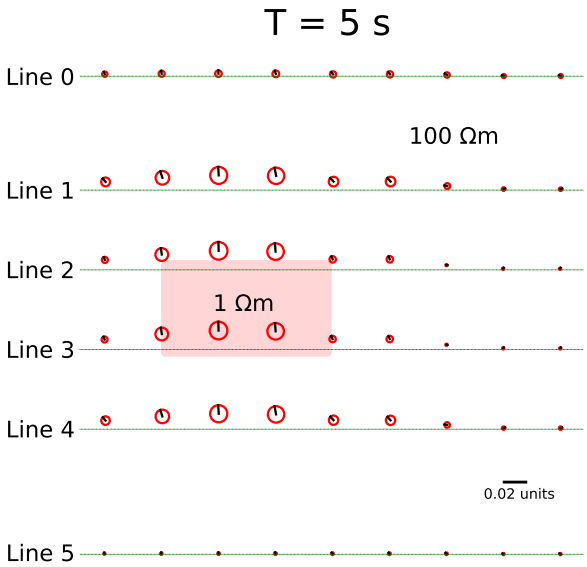

Fig. 9 Example of SVD Mohr circles for all sites at a period of 5 s. Red rectangle represents the 1 Ωm conductive block, and the background resistivity is 100 Ωm. Green horizontal lines represent the D'_{xy} axes

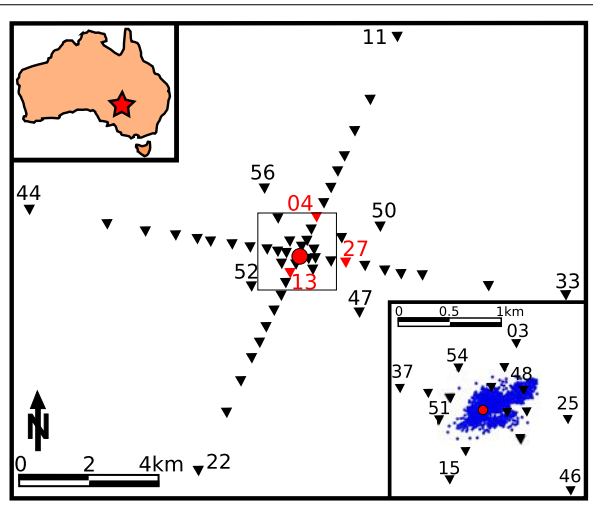

Fig. 10 Survey layout for the Paralana MT monitoring experiment [adapted from Peacock et al. (2013)]. The black triangles are the MT stations and the red triangles represent the stations used in our study. The red circle represents the Paralana 2 wellbore. The rectangle surrounding Paralana 2 is inflated in the bottom right, with the micro-seismic events shown as small blue circles

Discussion and conclusion

We have designed and conducted a feasibility study to determine the potential for using the TS method to monitor changes in resistivity structure of the Earth resulting from hydraulic stimulation. Three-dimensional forward modelling of our base and stimulation resistivity structures showed changes in **D** of between 1 and 2% as a result of introducing a $3 \times 1 \times 0.4$ km conductive body at 3.6 km depth. The transfer functions were mapped onto a grid by comparing each measured station to a reference (or base) station. Eigenanalysis and singular value decomposition were performed on the telluric distortion matrices from each station at various periods, and plotted onto Mohr circle grids. The resultant grids showed circles increasing in diameter at sites and periods associated with the conductive body. Moving away from the conductive body, the circles progressively approach the identity matrix indicating no telluric distortion had occurred.

The Paralana electric field data were relatively noisy and affected by the dead band making interpretations difficult. D_{xx} and D_{yx} components showed little change from the theoretical transfer functions. D_{xy} and D_{yy} were increasing by 2–5% for periods greater than 5 s. Changes in the

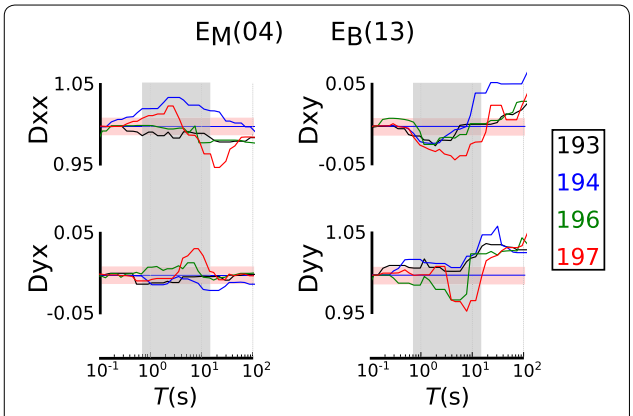

Fig. 11 Transfer functions comparing stations 13 ($\mathbf{E}_B$) and 04 ($\mathbf{E}_M$) between days 193 and 197. The *x*-axis shows the period in seconds, and the grey shaded area represents the approximate dead band. The blue horizontal lines represent the identity matrix transfer functions. The pink shaded region represents a noise window of ± 1%

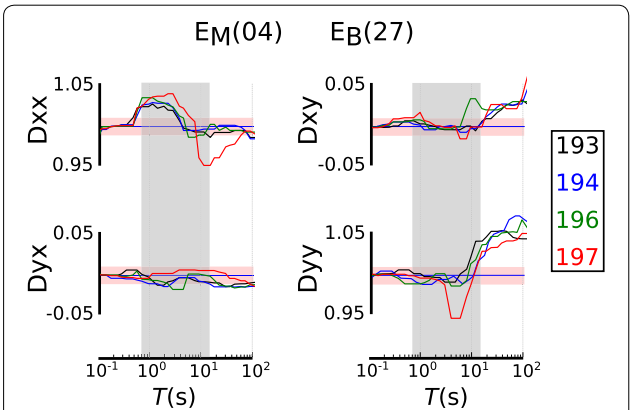

Fig. 12 Transfer functions comparing stations 27 ($\mathbf{E}_B$) and 04 ($\mathbf{E}_M$)

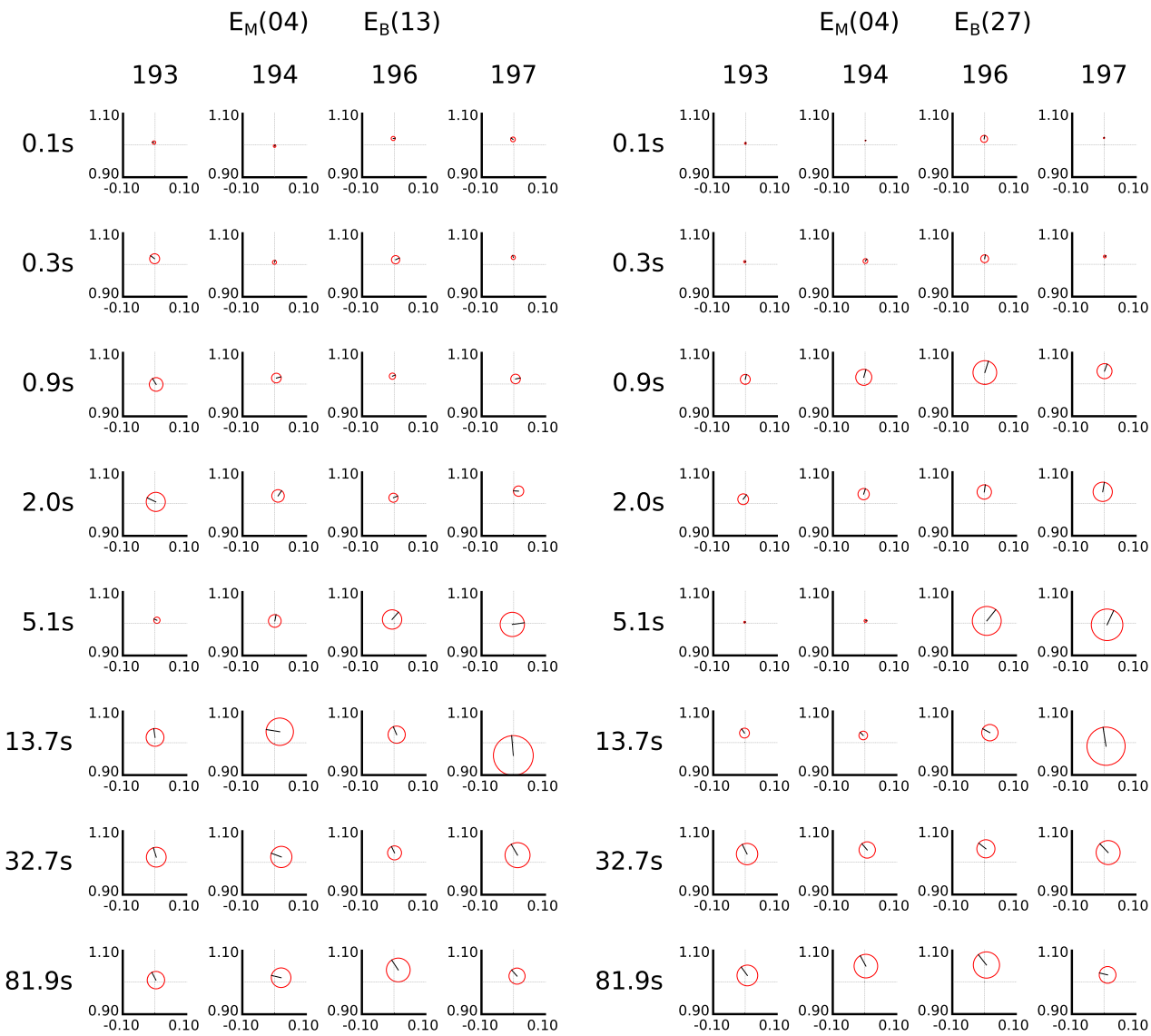

Fig. 13 SVD Mohr circles [see Lilley (2015)] for the telluric distortion matrices presented in Figs. 11 (left) and 12 (right)

y-direction are significant but in the *x*-direction are not, suggesting the response is more complex and possibly anisotropic (MacFarlane et al. 2014). Mohr circle grids visualised these changes in transfer function, with the circle diameters increasing at periods greater than 5 s.

This study has shown some potential for utilising the TS method for monitoring subsurface fluid movement. The advantage of this technique is that electric fields are relatively cheap and easy to measure when compared with traditional magnetotelluric surveys. Additionally, this technique can be used for near real-time monitoring from a remote location. However, electric fields are more complex and noisier than magnetics making interpretation more challenging. The impedance matrix transfer functions are largely real rather than being complex numbers and, in principal for a 1D Earth, are the identity matrix at all periods ($\pm$ static shift). Therefore changes in the earth resulting from the introduction of conductive bodies can be associated with deviations away from the identity matrix. SVD and EA can reduce the complexity of the telluric distortion matrix to simpler parameters that can be visualised in the form of Mohr circles. Grids of Mohr circles may be a useful diagnostic tool for understanding the extent of fluid movement resulting from hydraulic stimulation at depth. However, deep monitoring is always going to be marginal and for depths in the order of 4 km, changes in transfer functions are

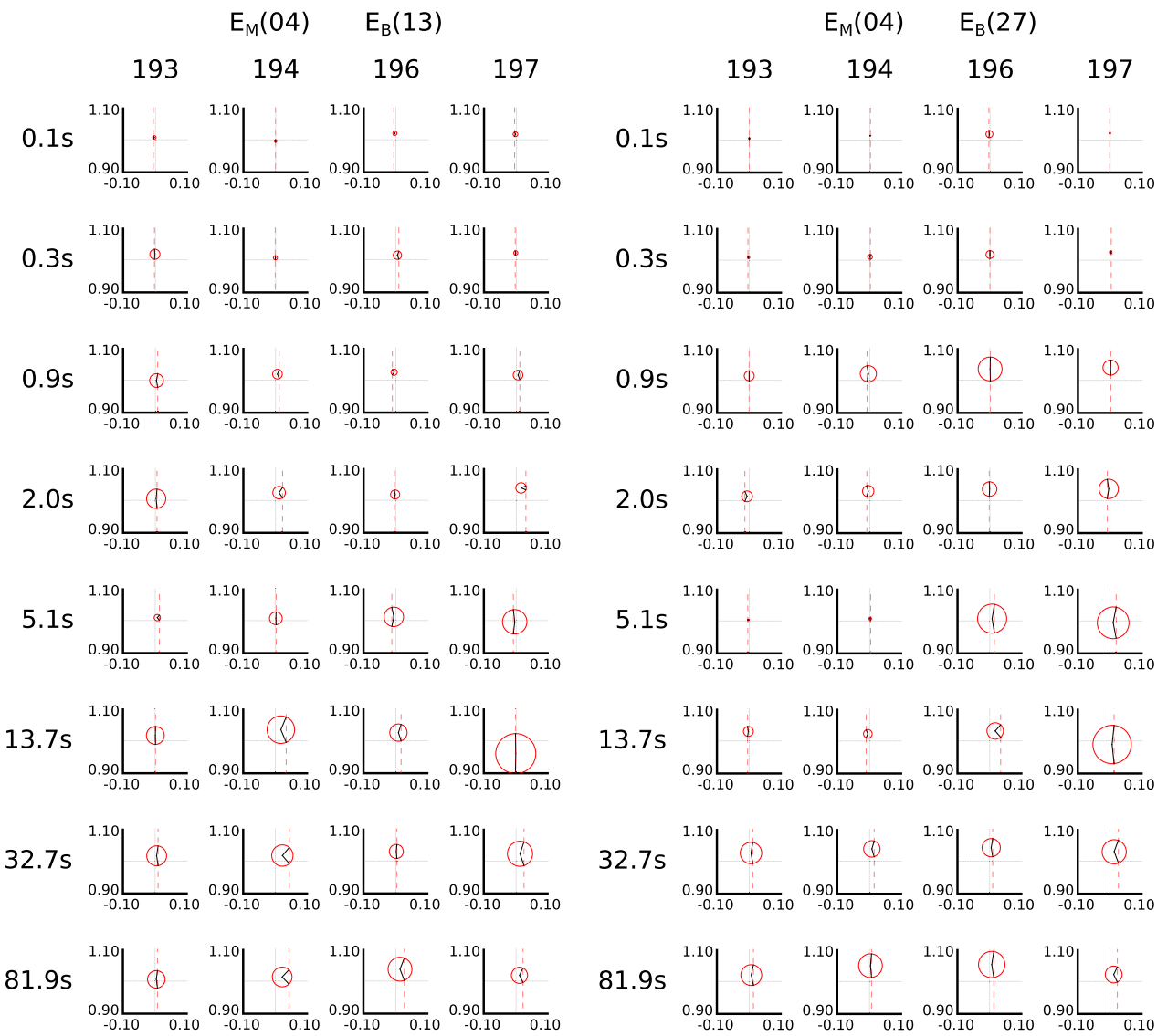

Fig. 14 EA Mohr circles [see Lilley (2015)] for the telluric distortion matrices presented in Figs. 11 (left) and 12 (right)

small. Therefore, in order to have meaningful interpretations, electric field data need to be of a high quality with low levels of site noise. Additionally, the dead band affects the transfer functions significantly so a controlled source may be useful for monitoring at depths incorporating dead band periods. Real data examples from depths shallower than the dead band may prove more suitable for determining distinguishable changes in the telluric distortion matrix resulting from the introduction of conductive fluids.

Additional files

Additional file 1. EDI files for both the 3D feasibility study and Paralana example.

Authors' contributions
All authors were involved in the design of the study and development of the methodology. NR wrote the manuscript. All authors read and approved the final manuscript.

Author details
[1] The Australian National University, Canberra, ACT 0200, Australia. [2] The University of Adelaide, Adelaide, SA 5005, Australia.

Acknowledgements
We would like to acknowledge Jared Peacock for his assistance working with the Paralana dataset.

Competing interests
The authors declare that they have no competing interests.

Availability of data and materials
Included in Additional file.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Publisher's Note

Received: 27 April 2017 Accepted: 19 December 2017
Published online: 11 January 2018

References

Berdičevskij MN, Keller GV (1965) Electrical prospecting with the telluric current method. Colorado School of Mines, Colorado
Brugger J, Long N, McPhail D, Plimer I (2005) An active amagmatic hydrothermal system: the Paralana hot springs, northern flinders ranges, South Australia. Chem Geol 222:35–64
Chave AD, Jones AG (2012) The magnetotelluric method: theory and practice. Cambridge University Press, Cambridge
Chave AD, Smith JT (1994) On electric and magnetic galvanic distortion tensor decompositions. J Geophys Res Solid Earth 99:4669–4682
Groom RW, Bailey RC (1989) Decomposition of magnetotelluric impedance tensors in the presence of local three-dimensional galvanic distortion. J Geophys Res Solid Earth 94:1913–1925
Lilley F (1993) Magnetotelluric analysis using Mohr circles. Geophysics 58:1498–1506
Lilley F (2012) Magnetotelluric tensor decomposition: insights from linear algebra and Mohr diagrams. In: Lim HS (ed) New achievements in geoscience. InTech Open Science, Rijeka, pp 81–106
Lilley F (2015) The distortion tensor of magnetotellurics: a tutorial on some properties. Explor Geophys 47:85–99
MacFarlane J, Thiel S, Pek J, Peacock J, Heinson G (2014) Characterisation of induced fracture networks within an enhanced geothermal system using anisotropic electromagnetic modelling. J Volcanol Geotherm Res 288:1–7
Mackie RL, Madden TR, Wannamaker PE (1993) Three-dimensional magnetotelluric modeling using difference equations-theory and comparisons to integral equation solutions. Geophysics 58:215–226
McLaren S, Sandiford M, Hand M, Neumann N, Wyborn L, Bastrakova I (2003) The hot southern continent: heat flow and heat production in Australian Proterozoic terranes. Geol Soc Am Spec Papers 372:157–167
Peacock J, Thiel S, Reid P, Heinson G (2012) Magnetotelluric monitoring of a fluid injection: example from an enhanced geothermal system. Geophys Res Lett 39:L18403. https://doi.org/10.1029/2012GL053080
Peacock JR, Thiel S, Heinson GS, Reid P (2013) Time-lapse magnetotelluric monitoring of an enhanced geothermal system. Geophysics 78:B121–B130
Yungul S (1966) Telluric sounding—a magnetotelluric method without magnetic measurements. Geophysics 31:185–191

APPENDIX C

Rees, N., Carter, S., Heinson, G., Krieger, L., Conway, D., Boren, G., & Matthews, C. (2016).
Magnetotelluric monitoring of coal-seam gas and shale-gas resource development in Australia.
The Leading Edge, 35(1), 64-70.

This paper demonstrates the use of MT as a geophysical monitoring tool. The method is applied to case-study datasets investigating changes in resistivity due to coal-seam gas dewatering and shale-gas hydraulic fracturing. Personal involvement in the research includes the collection and processing of MT data, assistance with modelling, manuscript feedback and development of software and hardware solutions for data processing.

Magnetotelluric monitoring of coal-seam gas and shale-gas resource development in Australia

Nigel Rees[1], Simon Carter[1], Graham Heinson[1], Lars Krieger[1], Dennis Conway[1], Goran Boren[1], and Chris Matthews[1]

Abstract

Extraction of unconventional energy has become a major global industry in the last decade and is driven by changes in technology and increasing demand. One of the key factors for the success of gas extraction is establishing sufficient permeability in otherwise low-porosity and low-permeability formations. Permeability can be established through hydraulic stimulation of deep formations, either through existing fracture networks or by creating new pathways for fluids to flow, and through depressurization of coalbeds by extracting existing subsurface fluids. Geophysical monitoring of hydraulic stimulation and depressurization can be used to determine lateral and vertical constraints on fluid movements in the target lithologies. Such constraints help to optimize production and well placement. In addition, independent verification is critical for social and environmental regulation, to ensure that hydraulic stimulations and depressurization do not interact with overlying aquifers. To date, the primary and most successful geophysical technique has been microseismic, which measures small seismic events associated with rock fractures from arrays of surface and downhole geophones. The microseismic approach has been used widely for many types of unconventional energy-resource development. The magnetotelluric (MT) method is an alternative approach to monitoring hydraulic stimulations and depressurization. In contrast to microseismic, which delineates the locations of rock fractures, MT is sensitive directly to the presence of fluid as measured by the earth's bulk electrical resistivity, which is dependent on permeability. MT is sensitive to the direction of fluid connection, so it might yield important information on how fluids migrate with time. Because subsurface fluids conduct electrical current dependent on the porosity, connectivity, and ionic saturation of the fluid, it follows that the introduction or removal of fluids will change the electrical resistivity of the formation. The physics of the approach is outlined, and the feasibility of the MT method for monitoring unconventional energy-resource development is demonstrated. Two case studies are conducted, one for a shallow (CSG) depressurization and the second for a deep hydraulic stimulation of a shale-gas reservoir.

Introduction — Magnetotellurics

The magnetotelluric (MT) method is a well-established technique that can determine information about earth's electrical resistivity in three dimensions (Chave and Jones, 2012). The approach uses natural sources of external magnetic-field variations that induce electrical currents in the subsurface. Sources of such magnetic-field variations are frequency dependent. At frequencies greater than 1 Hz, the sources are high-amplitude lightning strikes around the globe that occur mainly at equatorial latitudes.

Variations in the magnetic field propagate around the globe approximately at the speed of light. Close to the lightning strike, the fields have curvature, but at greater distances (typically 1000 km or more), the fields can be approximated as a plane-wave source. At lower frequencies of 10^{-1} to 10^{-4} Hz, the source of magnetic-field variations is the interaction of the solar wind with earth's magnetosphere, generating large flows in electrical current in the conducting ionosphere approximately 100 km above ground. Because the interaction of the solar wind is on a global scale, again, the signals can be approximated as plane waves. In the bandwidths of 1 to 10 Hz and 800 to 2000 Hz, there is a so-called dead band at which there is little naturally occurring signal.

Denoting the magnetic-induction horizontal components as B_x and B_y (with x being in a geographic north direction and y being the orthogonal east direction) and the resulting electric fields being E_x and E_y, Figure 1 shows a small window of time-series data sampled every 2 ms. Fields are shown with arbitrary scaled units to illustrate the inductive coupling between orthogonal components. In Figure 1a, which shows three minutes of data, most of the abrupt changes are caused by distant lightning strikes around the globe, known as sferics. It is evident that these events have different amplitude and phase in the two orthogonal magnetic inductions, depending on the source location and path of the wavefront. The inset (Figure 1b) shows a 1 s window, highlighting one particular event, with strong correlations between the orthogonal components of electric and magnetic fields.

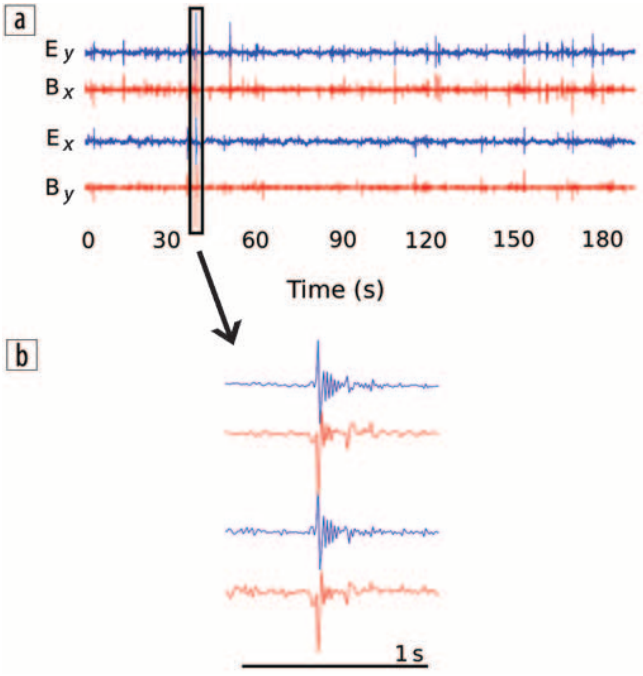

Figure 1. Example of a time series for orthogonal components of magnetic and electric fields. The fields are scaled arbitrarily in amplitude to illustrate the causal link between orthogonal components.

[1]University of Adelaide.

http://dx.doi.org/10.1190/tle35010064.1.

Typically, time series of the orthogonal electric and magnetic fields are recorded over intervals of hours to several days. Windows of time-series data then are Fourier-transformed (Chave and Thomson, 2004) to generate power spectral estimates over a range of frequencies.

In the frequency domain, we can relate the magnetic fields and electric fields through an electrical impedance tensor **Z**:

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{bmatrix} \begin{bmatrix} B_x \\ B_y \end{bmatrix}. \quad (1)$$

These impedances have both amplitude and phase and thus are represented by a complex number. From Maxwell's equations, the impedance tensor **Z** usually is dominated by its off-diagonal elements Z_{xy} and Z_{yx} that relate orthogonal components of electric and magnetic fields which are highly correlated, as shown in Figure 1. This is particularly true for sedimentary geologic formations in which the geology is mostly horizontally continuous. The impedance elements can be related to electrical resistivity through the relationship

$$\rho = \frac{1}{\mu_0 \omega} |Z_{ij}|^2. \quad (2)$$

Similarly, the phase of the complex impedance tensor is given by

$$\phi = \tan^{-1}\left(\frac{\Im(Z)}{\Re(Z)}\right). \quad (3)$$

Equation 2 can be interpreted as a measure of the depth-weighted electrical resistivity ρ, which is dependent on the depth of propagation of the time-varying magnetic field, and equation 3 is the phase difference between the electric and magnetic fields. In these equations, μ_0 is the magnetic permeability of free space, which has an SI value of $4\pi \times 10^{-7}$ Hm^{-1}, and ω is the angular frequency ($rad \cdot s^{-1}$).

A measure of the depth of investigation, an electromagnetic skin depth δ, can be defined as

$$\delta(T) = \sqrt{\frac{\rho}{\pi \mu f}} \approx 0.5 \sqrt{\frac{\rho}{f}} \text{ km}. \quad (4)$$

Skin depth is defined mathematically as the depth at which the inducing magnetic field reduces to a factor of e^{-1} (~ 37%) of its surface amplitude in an earth of resistivity ρ for frequency *f*. From equation 4, it is evident that skin depth increases in more electrically resistive materials and at lower frequencies. It is important to note, however, that unlike in seismic-reflection methods, skin depth does not denote the depth at which the MT method is sensitive but rather the approximate depth above which ~ 63% of the inductive effects occur.

The two off-diagonal components provide a measure of how the earth's impedance varies laterally. The impedance tensor in equation 1 can be rotated mathematically to any arbitrary coordinate systems that can maximize the difference. If the substructure is relatively laterally homogenous and varies only with depth, then the two off-diagonal elements are approximately equal in magnitude and phase.

On the other hand, a difference in the magnitudes and phases of the impedance elements indicates that the subsurface varies laterally. This might be on a macroscale (for example, the margins of a sedimentary basin) or on a lithologic scale, which is evident as an anisotropic property. Such anisotropy, for example, might be caused by pervasive fracture networks with a dominant strike from the regional stress field (Kirkby et al., 2015).

Electrical resistivity of unconventional-energy resources

The propagation of electromagnetic fields through the earth is governed by diffusion, meaning that high resolution of particular layers is intrinsically not possible. Most modeling approaches involve generating smooth resistivity responses that reflect the resolution of the MT method. However, the advantage of monitoring is that the zone of altered resistivity is well constrained, given that we know the depth, volume, and timing of injected fluid in an otherwise unchanging earth.

The electrical resistivity of subsurface lithologies is dependent on the resistivity of the lithologic matrix and the resistivity of pore fluids. For the matrix materials, silica-dominated clastic sediments and basement formations are typically resistive (> 1000 Ωm) and thus contribute little to the overall conduction. However, clay-rich lithology can have significantly lower resistivity (1 to 10 Ωm), with conduction dominated by surface mobility of ions.

Interstitial fluids generally have a much more significant role. Ionic conduction through dissolved salts can have low resistivity (0.1 to 10 Ωm), depending on ionic concentration. For deep formations at depths of 3 km or more, high temperatures of > 100°C will further reduce the fluid resistivity by about an order of magnitude (Nesbitt, 1993).

The final factor in determining bulk resistivity is in how fluids are connected. Such connectivity depends on porosity and the way pores are interconnected, which is a function of permeability. For primary porosity in clastic sediments, porosity and permeability are relatively isotropic, meaning that the properties do not vary in any specific orientation. However, secondary porosity caused by fractures, faults, bedding planes, and solution channels in carbonates might result in highly anisotropic permeability and hence anisotropic resistivity properties (MacFarlane et al., 2014; Kirkby et al., 2015).

The effect of hydraulic stimulation on bulk resistivity at depth depends on several factors. These include the resistivity of the fluids used for the stimulation; total volume of fluid and rate of pumping; temperature of the formation; the regional stress field and presence of existing fracture networks; and the presence of overpressured interstitial fluids. Peacock et al. (2012) and Peacock et al. (2013) show a significant change of more than 5% in the MT responses for a stimulation of a geothermal target with 3.1 million liters of fluid of resistivity of 0.3 Ωm at 20°C injected at 3.8 km depth.

On the other hand, Y. L. Didana (personal communication, 2015) reports a much smaller MT response change for 36.5

million liters of fluid of resistivity of 13 Ωm at 20°C injected at 4 km depth.

Although the total volumes of stimulation fluids are relatively small, opening existing and new fracture networks might connect significantly larger volumes of preexisting fluids. Didana (personal communication, 2015) notes that naturally occurring overpressured fluids in the granites had resistivities of < 0.3 Ωm. Microseismic events spanned a much larger volume (approximately 1.25 km^3) than might be expected for 36.5 million liters of injected fluid, suggesting that bulk resistivity might be affected more by connecting existing fluids than by the stimulation fluid alone.

To demonstrate the feasibility of MT monitoring of a hydraulic stimulation of a shale formation at a depth of 2900 m, simple 1D forward-modeling studies were generated. Although these are unrealistic in that only a finite volume is stimulated rather than a layer of infinite extent, the models indicate the magnitude and bandwidth of the MT response change. Didana (personal communication, 2015) shows that a 3D volumetric model of stimulation also can be detected, and moreover, the MT responses at the surface yield information on the extent and orientation of the stimulated region.

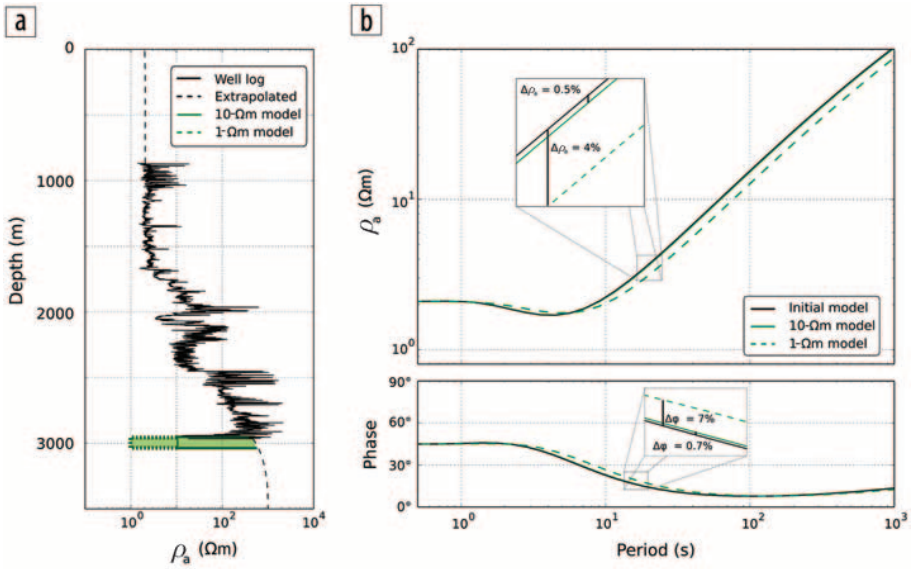

Figure 2. (a) Wireline-log resistivity plot that has been extrapolated to the surface and to greater depth (dashed line). The green regions indicate two scenarios of injected fluid that reduces the resistivity to 1 Ωm (dashed green) and 10 Ωm (solid green). (b) Forward-modeled apparent resistivity and phase for a typical bandwidth of period 1 to 1000 s (1 to 10^{-3} Hz). The black curve shows the response prior to stimulation, and the dashed green and solid green curves indicate the change in response to the two scenarios.

Wireline logs from a nearby well were used to generate a background resistivity model over the length of the logs (869 to 2964 m). Resistivity was extrapolated for depths above and below the well log, based on simple trends. An analytic 1D forward model was used to generate MT responses at 50 frequencies over a range of 10^{-3} to 10^2 Hz. Two additional models then were constructed to simulate the resistivity structure after hydraulic stimulation. These models included a 100 m thick layer of 1 Ωm or 10 Ωm below 2900 m. Figure 2 shows the model and the MT responses for these three scenarios.

For the 10 Ωm layer, the MT responses change by less than 1%, but the 1 Ωm layer yields an MT response change of as much as 7%, which are detectable (Peacock et al., 2013). The apparent-resistivity data show the most difference at frequencies lower than 0.1 Hz, whereas the phase has maximum difference between 1 and 0.1 Hz. We note that this is also the dead band of low signal, which introduces some noise issues. However, such problems generally can be reduced by using sufficiently long time windows (two to three days in many cases) to improve the signal-to-noise response.

MT monitoring survey layout

Because most MT monitoring programs are over a relatively small spatial extent (typically less than 5 km in any lateral dimension), the inducing magnetic fields can be considered to be spatially uniform over the extent of the array. Any change in subsurface resistivity is largely evident in a change in the induced electric fields. Thus, our MT monitoring arrays consist of a large number

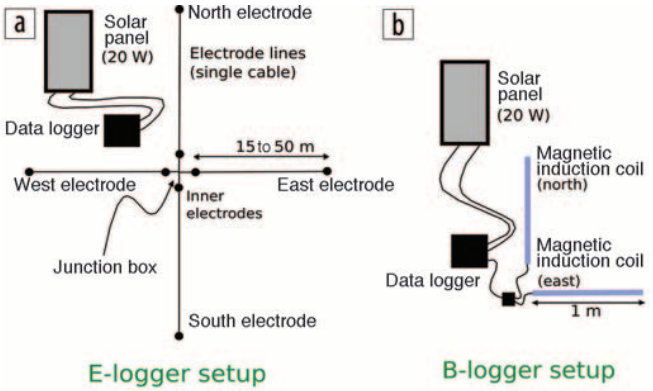

Figure 3. Schematic of the layout of (a) electric-field and (b) magnetic-field loggers. For the electric-field loggers, two sets of dipoles are deployed in a cross-formation with cable lengths of 15 to 50 m for each dipole. The redundancy of dipole in each orientation is important to mitigate against problems of animal interferences with cables and changing characteristics of the electrodes. In practice, for deployments over several months, the incidents of damage are high. Twenty-watt solar panels were used to trickle-charge batteries during daylight hours.

of electric-field loggers (typically 50 or more), but with only three to five sets of magnetic-field loggers (as shown in Figure 3).

The advantage of deploying only electric-field loggers is one of time efficiency (a crew of two people can deploy 10 instruments per day) and the spatial logistics of equipment in the field. With three to five sets of magnetic-field loggers, the magnetic-field time series can be checked for consistency across the array and can be used for various combinations of local and remote reference observations.

We record data at 651 Hz (sampling interval of 1536 μs) with GPS time synchronization at 1 s intervals to correct for clock

drift. Electrodes used are typically lead-lead chloride, buried 30 cm below the surface, with wires generally left on the surface. The magnetic-induction coils used are broadband LEMI120 from the Lviv Centre of the Institute for Space Research, Ukraine. Such broadband sensors have an approximately flat response of 1 mV/nT in the bandwidth of 1000 to 1 Hz, with a 10 dB/decade of frequency roll-off at lower frequencies. Data are recorded on 32-GB USB cards with storage of as long as four months. In practice, instruments were checked every two months for repairs, maintenance, and data transfer. The instrument pool was provided by the AuScope National MT Facility, funded by the Australian Government.

A major logistical benefit of natural-source MT is that data loggers are passive receivers, and thus no active transmitting source of electromagnetic (EM) signal is required. This allows for long deployments with low maintenance requirements and few safety issues other than manual handling. An additional advantage of long deployments is that the instrument setup logistics are the same for the whole experiment. Thus, continuous deployment ensures that an estimate of time-changing MT responses can be made under the same conditions.

MT monitoring of coal-seam gas depressurization

Coal-seam gas (CSG), also known as coalbed methane (CBM), is a form of natural gas with a predominantly methane component. CSG exists in underground coal seams, where it is adsorbed into the coal matrix and held in place by in situ fluid pressure. To extract adsorbed gas, depressurization is undertaken. A well is drilled into the coal-measure formations, and groundwater is extracted to reduce pore pressure within the coal measures, allowing the trapped gas to desorb (Garthwaite et al., 2015).

One of the key questions surrounding the CSG industry is whether one can effectively monitor the movement of fluids and changes in the subsurface resulting from depressurization. CSG monitoring bores are an expensive and inefficient means of monitoring fluctuations in groundwater level and quality. Monitoring bores provide information only at a single location and are not capable of measuring hydraulic connectivity of subsurface fluids. In this study, we demonstrate the utility of MT as a new approach to provide a cost-effective alternative for monitoring CSG depressurization.

From December 2013 to April 2014, an MT survey was conducted at a GSG production testing site in the Surat Basin, Queensland, Australia. Fifty-two electric-field loggers and three magnetic-field loggers were deployed to produce MT responses along two survey lines, each 2.5 km in extent and crossing several producing wells (Figure 4). Depressurization was undertaken at several wells at different times over the depth interval of 400 to 700 m.

Magnetotelluric response functions were generated from the time-series data as a function of both time during the survey and location in the array. For each site, a window of 72 hours of downsampled 100-Hz data was extracted centered on each day of the survey, and site-specific electric and magnetic fields were extracted to obtain the impedance-tensor and error estimates in the bandwidth of 10 to 0.01 Hz.

Line 1 of the test site crosses three production wells (Wells 3, 4, and 5). Unconstrained 2D inversions (Rodi and Mackie, 2001) of line 1 were conducted using the MT modeling software Winglink. MT responses from 8 December 2013 and 21 March 2014 were used for the pre- and postproduction models, respectively.

Figures 5a and 5b show results of pre- and postproduction inversions, with a color scale ranging from 2 to 5 Ωm. Approximate depths of the dominant stratigraphic units are shown, taken from Well 4 wireline logs. The location of Well 4 also is shown. That well produced by far the most water and gas during this time interval. Figure 5c shows the difference in resistivity between pre- and postproduction inversions, represented as a percentage change in resistivity.

Results from the 2D inversions suggest that the change in resistivity is confined to the upper and lower coal measures, with the larger effect west of Well 4. The change in resistivity to the west of Well 4 is on the order of 8% to 10%, whereas to the east of Well 4, it is 4% to 6%. This suggests greater permeability to the west of Well 4 as the coal measures are depressurized and the coal matrix shrinks (e.g., Palmer and Mansoori, 1996).

As an alternative to 2D time-lapse inversions that show spatial variations in resistivity associated with production, it is possible to undertake 1D inversions of data from a single site to show temporal variations in subsurface resistivity. We present one such example site from line 1 (site 1, 600 m to the west of Well 4). Water production from Well 4 is an order of magnitude higher than at other producing wells, and for the time interval shown in Figure 6, it was the only well producing water. Results from the 2D inversions suggest that sites to the west of this well should show the largest change in resistivity.

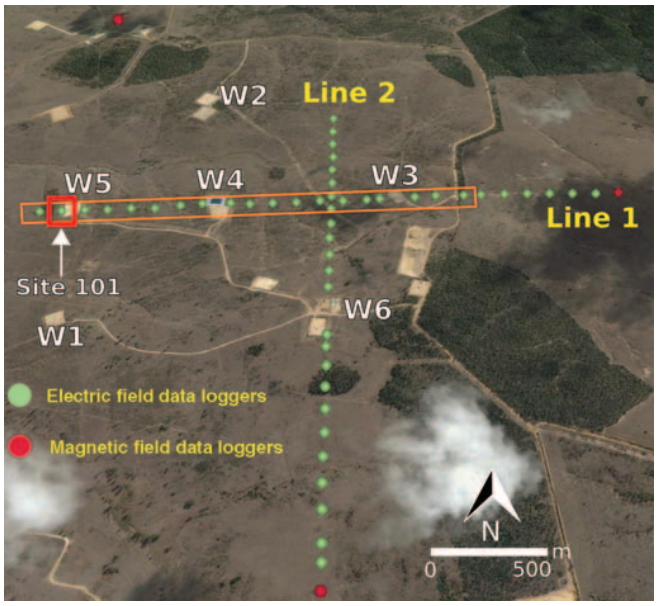

Figure 4. Field setup of 52 electric-field loggers (green) and three magnetic-field loggers (red). Each survey line was 2.5 km long, and the remote-reference station was in the northwest corner. W1 through W6 represent locations of operational wells. Depressurization was undertaken at different wells over the depth interval of 400 to 700 m. Instruments were left in the field for four months of continuous monitoring. The red box represents the site used for temporal 1D inversions. The orange box represents the sites used for time-lapse 2D modeling.

Figure 6 shows the time-lapse inversion from site 1 between 23 January 2014 and 3 March 2014 over the coal-seam interval. Water production from depressurization began on 11 February 2014, and gas production began on 20 February 2014. There is a decrease in resistivity in the coal measures (~ 350 to 650 m) approximately three days after water production began. The resistivity is approximately 3 Ωm from 23 January 2014 to 14 February 2014, after which the coal measures become more conductive, dropping to < 2.5 Ωm. This drop in resistivity is on the order of 10% to 15% and might be caused by enhanced permeability as the gas is released from the coal formation and the coal matrix slightly reduces in volume (Harpalani and Schraufnagel, 1990; Mirta et al., 2012).

MT monitoring of shale-gas hydraulic stimulation

Onshore shale-gas extraction has experienced a rapid increase in the past decade, given newly developed techniques coupled with the existence of large global reserves. One of the challenges of shale-gas extraction is the typically low permeability of shale. Hydraulic stimulation involves inducing geomechanical failure to stimulate economic hydraulic-conductivity rates (Cipolla et al., 2010). Although several related methods, loosely termed *hydraulic fracturing* ("fracking") are now commonplace, there are several associated environmental concerns such as gas leakage and loss of injection fluids into aquifer systems (Bierman et al., 2011).

A magnetotelluric survey was undertaken in the Cooper Basin, South Australia, in conjunction with a hydraulic stimulation conducted by Santos. A well was drilled vertically to a depth of 2900 m at the center of the profile (Figure 7) and then was deviated to horizontal, extending 1000 m south. Fluid injections began on 29 May 2014, with spatial and temporal sectioning of fluid pumping lasting 12 days. Pumping was for 10 frack stages beginning at the southernmost location and proceeding north. Pumping ceased 12 days later, marking the beginning of the fluid-flowback phase.

Hydraulic fracturing was monitored using 40 electric-field loggers and three magnetic-field loggers to record MT responses over a 65-day window — from day 120 (20 May 2014) until day 185 (22 June 2014). Electric-field loggers were deployed in a cross-array, with 200 m site separation, as shown in Figure 7.

A key attribute of MT monitoring is to identify how the full-impedance tensor changes with time. Peacock et al. (2012), Peacock et al. (2013), and Didana (personal communication,

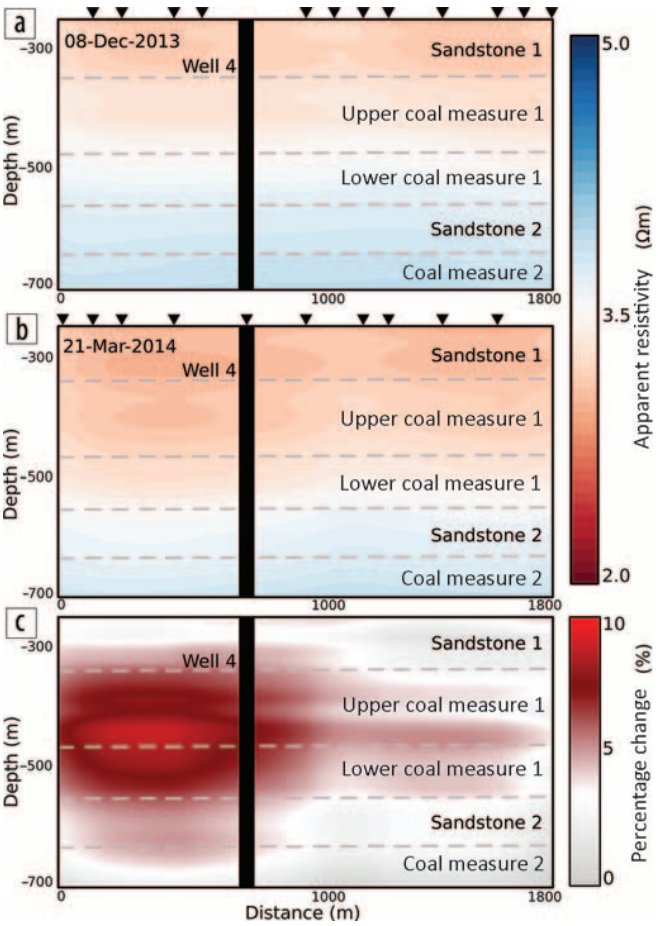

Figure 5. (a) Inversion of preproduction along line 1 on 8 December 2013. (b) Inversion of postproduction along line 1 on 21 March 2014. (c) Percentage change of March profile relative to December profile. The major change is in the upper and lower coal measures to the west of Well 4 and is on the order of 8% to 10% difference.

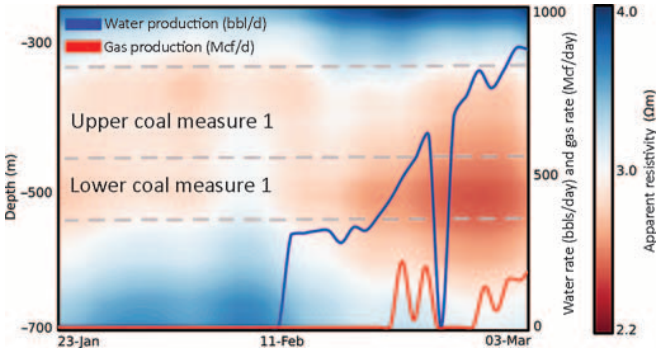

Figure 6. Time-lapse inversion showing temporal variation in the target coal-seam resistivity beneath site 1 on line 1 from 23 January 2014 to 3 March 2014. The blue line represents the water-production rate in barrels per day (bbl/d), and the red line represents the gas production rate in million cubic feet per day (Mcf/d) from Well 4.

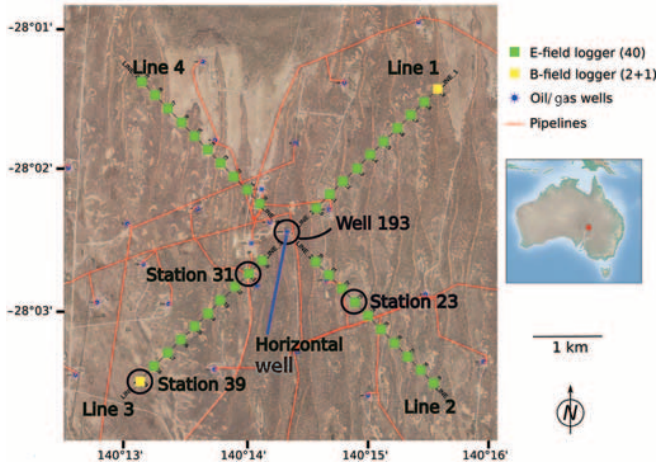

Figure 7. Field setup of 40 electric-field loggers (green) and three magnetic-field loggers (yellow). Each survey line was 4.4 km long, and the remote reference station was at the end of two lines, with an additional remote 10 km away. Examples of data from sites 23, 31, and 39 are shown in this article.

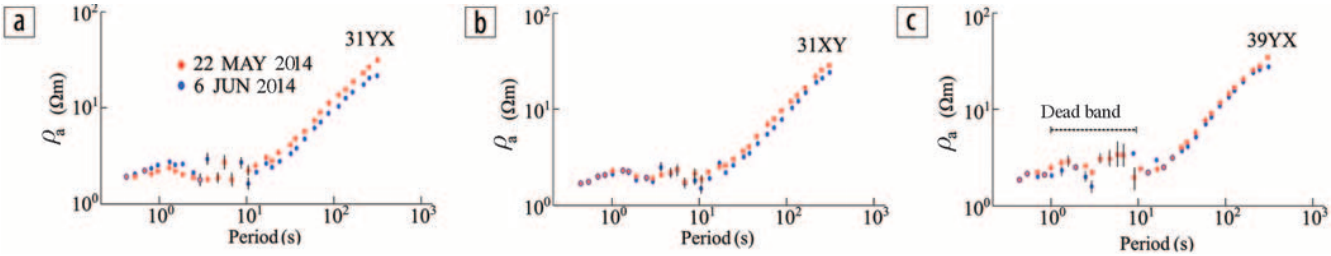

Figure 8. Differences in apparent-resistivity responses for a preproduction day (22 May 2014) and a postproduction day (6 June 2014). For site 31, two modes are shown, YX with the electric field perpendicular to line 3 and XY with the electric field parallel to line 3. For site 39, neither mode exhibited a major difference in response. Only the YX component is shown; it has a slightly larger change than XY.

2015) demonstrate this with some success using difference in phase tensors (Caldwell et al., 2004) across a large array. One advantage of this approach is that the observed data can be interpreted directly without the problems of nonunique modeling.

Figure 8 shows the apparent-resistivity response curves for site 31 (close to injection) and site 39 (at the end of the line and distant from the injection) for a preinjection day (22 May 2014) and postinjection day (6 June 2014). The impedance tensors were rotated clockwise 45° so that the XY component had the electric field parallel to line 3 and YX with the electric field perpendicular to line 3. The phases are not shown because the bandwidth at which a difference might be expected lies within the dead band.

Site 31 indicates that there is a clear difference for each of the modes, slightly larger in effect for the YX component; this is consistent with the feasibility study (Figure 2). Vector information from the impedance tensor can be used to help determine the areal extent of the injected fluids and/or preferential flow paths. On the other hand, site 39 (Figure 8c) exhibits a much a smaller change in both modes (only the YX is shown; the XY has a similarly small difference). Thus, we demonstrate a quantitative difference between proximal and distal sites that can be used to infer some characteristics of the injection without having to model. We note that similar response differences were observed along line 2, indicating that the stimulation was relatively symmetrical around the horizontal well.

In Figure 9, we show a 1D inversion from site 23 of the MT response rotated 45° from geographic coordinates with the electric field orientated perpendicular to line 2. These 1D inversions have a time-smoothing constraint to reduce modeling variations caused by noise. In Figure 9, the models are normalized to the prestimulation resistivity profile to show the magnitude of the modeled resistivity change as a function of depth and time.

Without any temporal or spatial constraints, the resistivity change caused by stimulation fluids injected at 2900 m is smoothed over a much larger depth interval of 2500 to > 5000 m in the plot. However, Figure 9 does show how resistivity changes over the entire time interval, from initial stimulation to flowback.

Figure 10 shows the total electrical conductance (defined as the sum of the thickness of each layer divided by its resistivity) between 2500 and 5000 m as a function of time. Electrical conductance is resolved better than electrical resistivity for 1D models (Parker, 1980). The increase in conductance is small (on the order of 1 S) and reflects the introduction of the stimulation fluid. Of note is that the conductance reverts to the preinjection value after

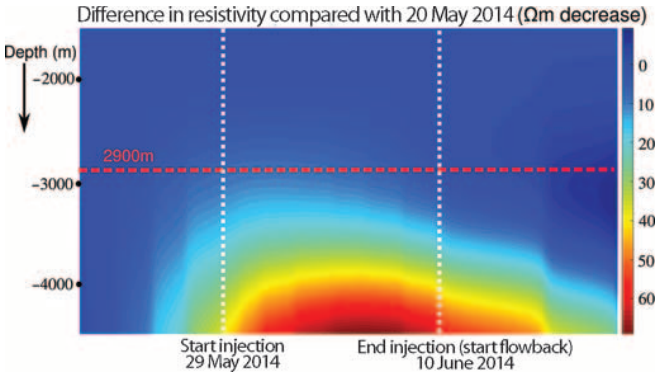

Figure 9. Temporal changes in 1D resistivity beneath site 23 for a time window that spans the stimulation and flowback phases. Such 1D inversions with depth are smoothed with time to dampen temporal changes in noise day by day. The resulting section is normalized with respect to the preinjection profile so that the plot shows resistivity differences.

flowback. This does not necessarily mean all the stimulation fluid has been removed, but enhancements in fluid-filled permeability are not maintained, and compressive stress reduces permeability as soon as stimulation has been concluded.

Conclusion

The utility of magnetotelluric monitoring of unconventional-resource development is at an early stage, with relatively few projects to date. Results from surveys in this article demonstrate that the MT method has great promise, but it is a new technology that will require some time to improve outcomes. There is scope for great innovation in the design of field surveys, potentially integrating the MT and microseismic methods. In particular, it is important to incorporate all known constraints in space (such as resistivity logs as a priori background and depth of fluid stimulation), time (commencement and rate of fluid injection), and fluid characteristics (fluid resistivity, formation temperature, and fluid volume). A significantly constrained model will be much more successful in defining regions of fluid injection or depressurization than a smooth unconstrained model. TLE

Acknowledgments

We thank many people at QGC and Santos for making these projects possible and for their enthusiasm for trying out a new approach. Santos and QGC staff members facilitated

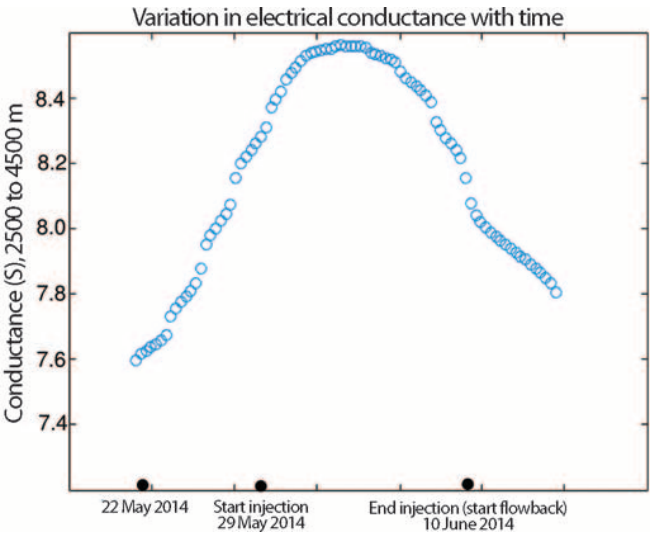

Figure 10. Electrical conductance of the interval 2500 to 5000 m as a function of time, derived from the resistivity section in Figure 9, indicating the movement of stimulating fluids into and out of the formation.

all aspects of access, logistics, and background data, without which the project would not have occurred. Magnetotelluric equipment was made available through AuScope and the Australian Geophysical Observing System (AGOS) through the Australian Department of Education and Training, and we are grateful for access to a world-class facility. AGOS is thanked for funding the shale-gas MT monitoring project as part of the subsurface observatory program. Many data-processing and modeling codes used in this project have been kindly provided through the academic community and are referenced within. The University of Adelaide is thanked for supporting many students and staff members on these projects, both financially and logistically.

Corresponding author: nigel.rees@adelaide.edu.au

References

Bierman, J., C. Kulp, and J. B. Foote, 2011, Reviews of science for science librarians: Hydraulic fracturing: Geological, engineering and environmental literature: Science & Technology Libraries, **30**, no. 4, 326–342, http://dx.doi.org/10.1080/0194262X.2011.626336.

Caldwell, T. G., H. M. Bibby, and C. Brown, 2004, The magnetotelluric phase tensor: Geophysical Journal International, **158**, no. 2, 457–469, http://dx.doi.org/10.1111/j.1365-246X.2004.02281.x.

Chave, A. D., and A. G. Jones, eds., 2012, The magnetotelluric method: Theory and practice: Cambridge University Press, http://dx.doi.org/10.1017/CBO9781139020138.

Chave, A. D., and D. J. Thomson, 2004, Bounded influence magnetotelluric response function estimation: Geophysical Journal International, **157**, no. 3, 988–1006, http://dx.doi.org/10.1111/j.1365-246X.2004.02203.x.

Cipolla, C. L., E. P. Lolen, J. C. Erdle, and B. Rubin, 2010, Reservoir modeling in shale-gas reservoirs: SPE Reservoir Evaluation & Engineering, **13**, no. 4, 638–653, http://dx.doi.org/10.2118/125530-PA.

Garthwaite, M., M. Hazelwood, S. Nancarrow, A. Hislop, and J. Dawson, 2015, A regional geodetic network to monitor ground surface response to resource extraction in the northern Surat Basin, Queensland: Australian Journal of Earth Sciences, **62**, no. 4, 469–477, http://dx.doi.org/10.1080/08120099.2015.1040073.

Harpalani, S., and R. R. Schraufnagel, 1990, Shrinkage of coal matrix with release of gas and its impact on permeability of coal: Fuel, **69**, no. 5, 551–556, http://dx.doi.org/10.1016/0016-2361(90)90137-F.

Kirkby, A., G. Heinson, S. Holford, and S. Thiel, 2015, Mapping fractures using 1D anisotropic modelling of magnetotelluric data: A case study from the Otway Basin, Victoria, Australia: Geophysical Journal International, **201**, no. 3, 1961–1976, http://dx.doi.org/10.1093/gji/ggv116.

MacFarlane, J., S. Thiel, J. Pek, J. Peacock, and G. Heinson, 2014, Characterisation of induced fracture networks within an enhanced geothermal system using anisotropic electromagnetic modelling: Journal of Volcanology and Geothermal Research, **288**, 1–7, http://dx.doi.org/10.1016/j.jvolgeores.2014.10.002.

Mirta, A., S. Harpalani, and S. Liu, 2012, Laboratory measurement and modeling of coal permeability with continued methane production: Part 1 — Laboratory results: Fuel, **94**, 110–116, http://dx.doi.org/10.1016/j.fuel.2011.10.052.

Nesbitt, B. E., 1993, Electrical resistivities of crustal fluids: Journal of Geophysical Research: Solid Earth, **98**, no. B3, 4301–4310, http://dx.doi.org/10.1029/92jb02576.

Palmer, I., and J. Mansoori, 1996, How permeability depends on stress and pore pressure in coalbeds: A new model: Annual Technical Conference and Exhibition, SPE conference paper 36737-MS, http://dx.doi.org/10.2523/36737-MS.

Parker, R. L., 1980, The inverse problem of electromagnetic induction: Existence and construction of solutions based on incomplete data: Journal of Geophysical Research: Solid Earth, **85**, no. B8, 4421–4428, http://dx.doi.org/10.1029/JB085iB08p04421.

Peacock, J. R., S. Thiel, G. S. Heinson, and P. Reid, 2013, Time-lapse magnetotelluric monitoring of an enhanced geothermal system: Geophysics, **78**, no. 3, B121–B130, http://dx.doi.org/10.1190/geo2012-0275.1.

Peacock, J. R., S. Thiel, P. Reid, and G. Heinson, 2012, Magnetotelluric monitoring of a fluid injection: Example from an enhanced geothermal system: Geophysical Research Letters, **39**, no. 18, L18403, http://dx.doi.org/10.1029/2012GL053080.

Rodi, W., and R. L. Mackie, 2001, Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion: Geophysics, **66**, no. 1, 174–187, http://dx.doi.org/10.1190/1.1444893.

BIBLIOGRAPHY

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). Tensorflow: A system for large-scale machine learning. In *OSDI*, volume 16, pages 265–283.

Alexander, B. and Lee, Y. H. (2018). A cell-division search technique for inversion with application to picture-discovery and magnetotellurics. *arXiv:1804.07887*.

Alumbaugh, D. L. and Newman, G. A. (2000). Image appraisal for 2-D and 3-D electromagnetic inversion. *Geophysics*, 65(5):1455–1467.

Aminian, K. and Ameri, S. (2005). Application of artificial neural networks for reservoir characterization with limited data. *Journal of Petroleum Science and Engineering*, 49(3):212–222.

Asfahani, J. and Ghani, B. A. (2012). Automated interpretation of nuclear and electrical well loggings for basalt characterization (case study from Southern Syria). *Applied Radiation and Isotopes*, 70(10):2500–2506.

Avdeev, D. B. and Avdeeva, A. D. (2006). A rigorous three-dimensional magnetotelluric inversion. *Progress In Electromagnetics Research*, 62:41–48.

Backus, G. E. (1988). Bayesian inference in geomagnetism. *Geophysical Journal International*, 92(1):125–142.

Bassam, A., Santoyo, E., Andaverde, J., Hernández, J., and Espinoza-Ojeda, O. (2010). Estimation of static formation temperatures in geothermal wells by using an artificial neural network approach. *Computers & Geosciences*, 36(9):1191–1199.

Bedrosian, P. A. and Feucht, D. W. (2014). Structure and tectonics of the northwestern United States from Earthscope USArray magnetotelluric data. *Earth and Planetary Science Letters*, 402:275–289.

Benaouda, D., Wadge, G., Whitmarsh, R., Rothwell, R., and MacLeod, C. (1999). Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs: an example from the ocean drilling program. *Geophysical Journal International*, 136(2):477–491.

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. *arXiv:1701.02434*.

Briqueu, L., Gottlib-Zeh, S., Ramadan, M., and Brulhet, J. (2002). Traitement des diagraphies à l'aide d'un réseau de neurones du type «carte auto-organisatrice»: application à l'étude lithologique de la couche silteuse de marcoule (Gard, France). *Comptes Rendus Geoscience*, 334(5):331–337.

Buland, A. and Kolbjørnsen, O. (2012). Bayesian inversion of CSEM and magnetotelluric data. *Geophysics*, 77(1):E33–E42.

Cagniard, L. (1953). Basic theory of the magneto-telluric method of geophysical prospecting. *Geophysics*, 18(3):605–635.

Campisi, L. D. (2015). Multilayer perceptrons as function approximators for analytical solutions of the diffusion equation. *Computational Geosciences*, 19(4):769–780.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., and Riddell, A. (2016). Stan: A probabilistic programming language. *Journal of Statistical Software*, 20:1–37.

Carter-McAuslan, A., Lelièvre, P. G., and Farquharson, C. G. (2014). A study of fuzzy c-means coupling for joint inversion, using seismic tomography and gravity data test scenarios. *Geophysics*, 80(1):W1–W15.

Cerv, V., Menvielle, M., and Pek, J. (2007). Stochastic interpretation of magnetotelluric data, comparison of methods. *Annals of Geophysics*, 50(1).

Chave, A. D. and Jones, A. G. (2012). *The Magnetotelluric Method: Theory and Practice*. Cambridge University Press.

Chen, J., Hoversten, G. M., Key, K., Nordquist, G., and Cumming, W. (2012). Stochastic inversion of magnetotelluric data using a sharp boundary parameterization and application to a geothermal site. *Geophysics*, 77(4):E265–E279.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4):327–335.

Chollet, F. et al. (2015). Keras.

Chong, A., Lam, K. P., Pozzi, M., and Yang, J. (2017). Bayesian calibration of building energy models with large datasets. *Energy and Buildings*.

Constable, S. C. (1998). Seafloor magnetotelluric system and method for oil exploration. US Patent 5,770,945.

Constable, S. C., Parker, R. L., and Constable, C. G. (1987). Occams inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, 52(3):289–300.

Conway, D., Simpson, J., Didana, Y., Rugari, J., and Heinson, G. (2018). Probabilistic magnetotelluric inversion with adaptive regularisation using the no-u-turns sampler. *Pure and Applied Geophysics*, pages 1–14.

Cracknell, M. J. and Reading, A. M. (2013). The upside of uncertainty: Identification of lithology contact zones from airborne geophysics and satellite data using random forests and support vector machines. *Geophysics*, 78(3):WB113–WB126.

Creutz, M. (1988). Global Monte Carlo algorithms for many-fermion systems. *Physical Review D*, 38(4):1228.

Dahl, G. E., Sainath, T. N., and Hinton, G. E. (2013). Improving deep neural networks for LVCSR using rectified linear units and dropout. In *Acoustics, Speech and Signal Processing (ICASSP)*, pages 8609–8613. IEEE.

de Carvalho Carneiro, C., Fraser, S. J., Crósta, A. P., Silva, A. M., and de Mesquita Barros, C. E. (2012). Semiautomated geologic mapping using self-organizing maps and airborne geophysics in the Brazilian Amazon. *Geophysics*.

de Matos, M. C., Osorio, P. L., and Johann, P. R. (2006). Unsupervised seismic facies analysis using wavelet transform and self-organizing maps. *Geophysics*.

deGroot Hedlin, C. and Constable, S. (1990). Occams inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics*, 55(12):1613–1624.

deGroot Hedlin, C. and Constable, S. (2004). Inversion of magnetotelluric data for 2D structure with sharp resistivity contrasts. *Geophysics*, 69(1):78–86.

Deng, C., Pan, H., Fang, S., Konaté, A. A., and Qin, R. (2017). Support vector machine as an alternative method for lithology classification of crystalline rocks. *Journal of Geophysics and Engineering*, 14(2):341.

Di Giuseppe, M. G., Troiano, A., Troise, C., and De Natale, G. (2014). k-Means clustering as tool for multivariate geophysical data analysis. an application to shallow fault zone imaging. *Journal of Applied Geophysics*, 101:108–115.

Didana, Y. L., Heinson, G., Thiel, S., and Krieger, L. (2017). Magnetotelluric monitoring of permeability enhancement at enhanced geothermal system project. *Geothermics*, 66:23–38.

Dozat, T. (2016). Incorporating Nesterov momentum into Adam.

Essenreiter, R., Karrenbach, M., and Treitel, S. (2001). Identification and classification of multiple reflections with self-organizing maps. *Geophysical Prospecting*, 49(3):341–352.

Gallagher, K., Sambridge, M., and Drijkoningen, G. (1991). Genetic algorithms: An evolution from Monte Carlo methods for strongly nonlinear geophysical optimization problems. *Geophysical Research Letters*, 18(12):2177–2180.

Gąsior, I. and Przelaskowska, A. (2014). Estimating thermal conductivity from core and well log data. *Acta Geophysica*, 62(4):785–801.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, pages 457–472.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):721–741.

Gholami, R., Shahraki, A., and Jamali Paghaleh, M. (2012). Prediction of hydrocarbon reservoirs permeability using support vector machine. *Mathematical Problems in Engineering*, 2012.

Goldstein, M. and Strangway, D. (1975). Audio-frequency magnetotellurics with a grounded electric dipole source. *Geophysics*, 40(4):669–683.

Goutorbe, B., Lucazeau, F., and Bonneville, A. (2006). Using neural networks to predict thermal conductivity from geophysical well logs. *Geophysical Journal International*, 166(1):115–125.

Grandis, H. (2006). Magnetotelluric (MT) inversion for 3-D conductivity model resolution using Markov Chain Monte Carlo (MCMC) algorithm: Preliminary results. *Jurnal Geofisika*, 1.

Grandis, H., Menvielle, M., and Roussignol, M. (1999). Bayesian inversion with markov chainsi. the magnetotelluric one-dimensional case. *Geophysical Journal International*, 138(3):757–768.

Guo, R., Dosso, S. E., Liu, J., Dettmer, J., and Tong, X. (2011). Nonlinearity in Bayesian 1-D magnetotelluric inversion. *Geophysical Journal International*, 185(2):663–675.

Hansen, N., Müller, S. D., and Koumoutsakos, P. (2003). Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1–18.

Hansen, T. M. and Cordua, K. S. (2017). Efficient Monte Carlo sampling of inverse problems using a neural network-based forward-applied to GPR crosshole traveltime inversion. *Geophysical Journal International*, 211(3):1524–1533.

Hayley, K., Pidlisecky, A., and Bentley, L. (2011). Simultaneous time-lapse electrical resistivity inversion. *Journal of Applied Geophysics*, 75(2):401–411.

Heinson, G. S., Direen, N. G., and Gill, R. M. (2006). Magnetotelluric evidence for a deep-crustal mineralizing system beneath the Olympic Dam iron oxide copper-gold deposit, southern Australia. *Geology*, 34(7):573–576.

Helle, H. B., Bhatt, A., and Ursin, B. (2001). Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study. *Geophysical Prospecting*, 49(4):431–444.

Hoffman, M. D. and Gelman, A. (2014). The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(1):1593–1623.

Jackson, D. D. (1976). Most squares inversion. *Journal of Geophysical Research*, 81(5):1027–1030.

Jegen, M. D., Hobbs, R. W., Tarits, P., and Chave, A. (2009). Joint inversion of marine magnetotelluric and gravity data incorporating seismic constraints: Preliminary results of sub-basalt imaging off the Faroe Shelf. *Earth and Planetary Science Letters*, 282(1-4):47–55.

Ji, S., Xu, W., Yang, M., and Yu, K. (2013). 3D convolutional neural networks for human action recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):221–231.

Jones, E., Oliphant, T., Peterson, P., and Others (2001). {SciPy}: Open source scientific tools for {Python}.

Kamnitsas, K., Chen, L., Ledig, C., Rueckert, D., and Glocker, B. (2015). Multi-scale 3D convolutional neural networks for lesion segmentation in brain MRI. *Ischemic Stroke Lesion Segmentation*, 13:46.

Karaoulis, M., Revil, A., Minsley, B., Todesco, M., Zhang, J., and Werkema, D. D. (2013a). Time-lapse gravity inversion with an active time constraint. *Geophysical Journal International*, 196(2):748–759.

Karaoulis, M., Revil, A., Tsourlos, P., Werkema, D. D., and Minsley, B. J. (2013b). IP4DI: A software for time-lapse 2D/3D DC-resistivity and induced polarization tomography. *Computers and Geosciences*, 54:164–170.

Karaoulis, M., Werkema, D. D., and Revil, A. (2015). 2D time-lapse seismic tomography using an active time-constraint (ATC) approach. *The Leading Edge*, 34(2):206–212.

Karimpouli, S., Fathianpour, N., and Roohi, J. (2010). A new approach to improve neural networks' algorithm in permeability prediction of petroleum reservoirs using supervised committee machine neural network (SCMNN). *Journal of Petroleum Science and Engineering*, 73(3):227–232.

Karmakar, M., Maiti, S., Singh, A., Ojha, M., and Maity, B. S. (2017). Mapping of rock types using a joint approach by combining the multivariate statistics, self-organizing map and Bayesian neural networks: an example from IODP 323 site. *Marine Geophysical Research*, pages 1–13.

Kelbert, A., Meqbel, N., Egbert, G. D., and Tandon, K. (2014). ModEM: A modular system for inversion of electromagnetic geophysical data. *Computers & Geosciences*, 66:40–53.

Kello, C. T. and Plaut, D. C. (2004). A neural network model of the articulatory-acoustic forward mapping trained on recordings of articulatory parameters. *The Journal of the Acoustical Society of America*, 116(4).

Kim, J.-H., Yi, M.-J., Park, S.-G., and Kim, J. G. (2009). 4-D inversion of DC resistivity monitoring data acquired over a dynamically changing earth model. *Journal of Applied Geophysics*, 68(4):522 – 532.

Kirkby, A. and Heinson, G. (2017). Three-dimensional resistor network modeling of the resistivity and permeability of fractured rocks. *Journal of Geophysical Research: Solid Earth*, 122(4):2653–2669.

Kleesiek, J., Urban, G., Hubert, A., Schwarz, D., Maier-Hein, K., Bendszus, M., and Biller, A. (2016). Deep MRI brain extraction: a 3D convolutional neural network for skull stripping. *NeuroImage*, 129:460–469.

Klose, C. D. (2006). Self-organizing maps for geoscientific data analysis: geological interpretation of multidimensional geophysical data. *Computational Geosciences*, 10(3):265–277.

Köhler, A., Ohrnberger, M., and Scherbaum, F. (2010). Unsupervised pattern recognition in continuous seismic wavefield records using self-organizing maps. *Geophysical Journal International*, 182(3):1619–1630.

Konaté, A. A., Pan, H., Fang, S., Asim, S., Ziggah, Y. Y., Deng, C., and Khan, N. (2015a). Capability of self-organizing map neural network in geophysical log data classification: case study from the CCSD-MH. *Journal of Applied Geophysics*, 118:37–46.

Konaté, A. A., Pan, H., Khan, N., and Ziggah, Y. Y. (2015b). Prediction of porosity in crystalline rocks using artificial neural networks: an example from the Chinese continental scientific drilling main hole. *Studia Geophysica et Geodaetica*, 59(1):113–136.

Kraft, D. and Schnepper, K. (1989). SLSQPA Nonlinear Programming Method with Quadratic Programming Subproblems. *DLR, Oberpfaffenhofen*.

Krasnopolsky, V. M. and Schiller, H. (2003). Some neural network applications in environmental sciences. Part I: forward and inverse problems in geophysical remote measurements. *Neural Networks*, 16(3-4):321–334.

Krieger, L. and Peacock, J. R. (2014). MTpy: A Python toolbox for magnetotellurics. *Computers & Geosciences*, 72:167–175.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 1097–1105.

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Laloy, E. and Vrugt, J. A. (2012). High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing. *Water Resources Research*, 48(1).

Lan, T., Liu, H., Liu, N., Li, J., Han, F., and Liu, Q. H. (2018). Joint inversion of electromagnetic and seismic data based on structural constraints using variational born iteration method. *IEEE Transactions on Geoscience and Remote Sensing*.

Leverington, D. W. (2010). Discrimination of sedimentary lithologies using Hyperion and Landsat Thematic Mapper data: a case study at Melville Island, Canadian High Arctic. *International Journal of Remote Sensing*, 31(1):233–260.

Liu, J., Guo, R., Tong, X., Liu, C., and Liu, Y. (2011). A decorrelation-based hybrid global search algorithm for inversion of 1D magnetotelluric data. *Journal of Geophysics and Engineering*, 8(2):225.

Mackie, R. L., Rodi, W., and Watts, M. D. (2001). 3-D magnetotelluric inversion for resource exploration. In *SEG Technical Program Expanded Abstracts 2001*, pages 1501–1504. Society of Exploration Geophysicists.

Mahmoodi, O. and Smith, R. (2015). Clustering of downhole physical property measurements at the Victoria property, Sudbury for the purpose of extracting lithological information. *Journal of Applied Geophysics*, 118:145–154.

Maiti, S., Krishna Tiwari, R., and Kümpel, H.-J. (2007). Neural network modelling and classification of lithofacies using well log data: A case study from KTB borehole site. *Geophysical Journal International*, 169(2):733–746.

Mandolesi, E., Ogaya, X., Campanyà, J., and Agostinetti, N. P. (2018). A reversible-jump Markov chain Monte Carlo algorithm for 1D inversion of magnetotelluric data. *Computers & Geosciences*, 113:94–105.

Meju, M. and Hutton, V. (1992). Iterative most-squares inversion: application to magnetotelluric data. *Geophysical Journal International*, 108(3):758–766.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *3D Vision (3DV), 2016 Fourth International Conference on*, pages 565–571. IEEE.

Minsley, B. J. (2011). A trans-dimensional Bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data. *Geophysical Journal International*, 187(1):252–272.

Mitsuhata, Y., Matsuo, K., and Minegishi, M. (1999). Magnetotelluric survey for exploration of a volcanic-rock reservoir in the Yurihara oil and gas field, Japan. *Geophysical Prospecting*, 47(2):195–218.

Moazzeni, A. and Haffar, M. A. (2015). Artificial intelligence for lithology identification through real-time drilling data. *Journal of Earth Science & Climatic Change*, 6(3):1.

Mohaghegh, S., Arefi, R., Ameri, S., Aminiand, K., and Nutter, R. (1996). Petroleum reservoir characterization with the aid of artificial neural networks. *Journal of Petroleum Science and Engineering*, 16(4):263–274.

Molchanov, P., Yang, X., Gupta, S., Kim, K., Tyree, S., and Kautz, J. (2016). Online detection and classification of dynamic hand gestures with recurrent 3D convolutional neural network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4207–4215.

Monnahan, C. C., Thorson, J. T., and Branch, T. A. (2017). Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. *Methods in Ecology and Evolution*, 8(3):339–348.

Montahaei, M. and Oskooi, B. (2014). Magnetotelluric inversion for azimuthally anisotropic resistivities employing artificial neural networks. *Acta Geophysica*, 62(1):12–43.

Nagano, I., Yagitani, S., Ozaki, M., Nakamura, Y., and Miyamura, K. (2007). Estimation of lightning location from single station observations of sferics. *Electronics and Communications in Japan (Part I: Communications)*, 90(1):25–34.

Nam, M. J., Song, Y., Jang, H., and Kim, B. (2017). 4D inversion of time-lapse magnetotelluric data sets for monitoring geothermal reservoir. *Journal of Applied Geophysics*, 141:88–97.

Nesterov, Y. (2009). Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1):221–259.

Newman, G. A. and Alumbaugh, D. L. (2000). Three-dimensional magnetotelluric inversion using non-linear conjugate gradients. *Geophysical Journal International*, 140(2):410–424.

Ogaya, X., Alcalde, J., Marzán, I., Ledo, J., Queralt, P., Marcuello, A., Martí, D., Saura, E., Carbonell, R., and Benjumea, B. (2016). Joint interpretation of magnetotelluric, seismic, and well-log data in Hontomín (Spain). *Solid Earth*.

Ogaya, X., Ledo, J., Queralt, P., Marcuello, Á., and Quintà, A. (2013). First geoelectrical image of the subsurface of the Hontomín site (Spain) for CO2 geological storage: A magnetotelluric 2D characterization. *International Journal of Greenhouse Gas Control*, 13:168–179.

Oh, K.-S. and Jung, K. (2004). GPU implementation of neural networks. *Pattern Recognition*, 37(6):1311–1314.

Oommen, T., Misra, D., Twarakavi, N. K., Prakash, A., Sahoo, B., and Bandopadhyay, S. (2008). An objective analysis of support vector machine based classification for remote sensing. *Mathematical Geosciences*, 40(4):409–424.

Paasche, H. and Eberle, D. (2009). Rapid integration of large airborne geophysical data suites using a fuzzy partitioning cluster algorithm: A tool for geological mapping and mineral exploration targeting. *Exploration Geophysics*, 40(3):277–287.

Paasche, H. and Tronicke, J. (2007). Cooperative inversion of 2D geophysical data sets: A zonal approach based on fuzzy c-means cluster analysis. *Geophysics*, 72(3):A35–A39.

Peacock, J. R., Thiel, S., Reid, P., and Heinson, G. (2012). Magnetotelluric monitoring of a fluid injection: Example from an enhanced geothermal system. *Geophysical Research Letters*, 39(18).

Raiche, A. (1991). A pattern recognition approach to geophysical inversion using neural nets. *Geophysical Journal International*, 105(3):629–648.

Rees, N., Carter, S., Heinson, G., and Krieger, L. (2016a). Monitoring shale gas resources in the cooper basin using magnetotellurics. *Geophysics*, 81(6):A13–A16.

Rees, N., Heinson, G., and Krieger, L. (2016b). Magnetotelluric monitoring of coal seam gas depressurization. *Geophysics*, 81(6):E423–E432.

Robertson, K., Heinson, G., and Thiel, S. (2016). Lithospheric reworking at the Proterozoic–Phanerozoic transition of Australia imaged using AusLAMP magnetotelluric data. *Earth and Planetary Science Letters*, 452:27–35.

Robertson, K., Heinson, G., and Thiel, S. (2017). Mapping lithospheric alteration using AusLAMP MT data in the Ikara-Flinders Ranges and Curnamona Province. *MESA Journal*, 83:4–7.

Rosas-Carbajal, M., Linde, N., and Kalscheuer, T. (2012). Focused time-lapse inversion of radio and audio magnetotelluric data. *Journal of Applied Geophysics*, 84:29–38.

Rosas-Carbajal, M., Linde, N., Kalscheuer, T., and Vrugt, J. A. (2013). Two-dimensional probabilistic inversion of plane-wave electromagnetic data: methodology, model constraints and joint inversion with electrical resistivity data. *Geophysical Journal International*, 196(3):1508–1524.

Rosas-Carbajal, M., Linde, N., Peacock, J., Zyserman, F., Kalscheuer, T., and Thiel, S. (2015). Probabilistic 3-D time-lapse inversion of magnetotelluric data: application to an enhanced geothermal system. *Geophysical Journal International*, 203(3):1946–1960.

Saemi, M., Ahmadi, M., and Varjani, A. Y. (2007). Design of neural networks using genetic algorithm for the permeability estimation of the reservoir. *Journal of Petroleum Science and Engineering*, 59(1):97–105.

Salvatier, J., Wiecki, T. V., and Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 2:e55.

Sambridge, M. and Mosegaard, K. (2002). Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40(3).

Sanders, N. E., Soderberg, A. M., Gezari, S., Betancourt, M., Chornock, R., Berger, E., Foley, R., Challis, P., Drout, M., Kirshner, R., et al. (2015). Toward characterization of the type IIP supernova progenitor population: A statistical sample of light curves from Pan-STARRS1. *The Astrophysical Journal*, 799(2):208.

Schmitt, P., Veronez, M. R., Tognoli, F. M., Todt, V., Lopes, R. C., and Silva, C. A. (2012). Electrofacies modelling and lithological classification of coals and mud-bearing fine-grained siliciclastic rocks based on neural networks. *Earth Science Research*, 2(1):193.

Schnaidt, S., Conway, D., Krieger, L., and Heinson, G. (2018). Pareto-optimal multi-objective inversion of geophysical data. *Pure and Applied Geophysics*, pages 1–16.

Schnaidt, S. and Heinson, G. (2015). Bootstrap resampling as a tool for uncertainty analysis in 2-D magnetotelluric inversion modelling. *Geophysical Journal International*, 203(1):92–106.

Schott, J.-J., Roussignol, M., Menvielle, M., and Nomenjahanary, F. R. (1999). Bayesian inversion with Markov chains-II. the one-dimensional DC multilayer case. *Geophysical Journal International*, 138(3):769–783.

Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, 66(3):605–610.

Shimelevich, M., Obornev, E., Obornev, I., and Rodionov, E. (2017). The neural network approximation method for solving multidimensional nonlinear inverse problems of geophysics. *Izvestiya, Physics of the Solid Earth*, 53(4):588–597.

Shimelevich, M. I., Obornev, M., and Gavryushov, S. (2007). Rapid neuronet inversion of 2D magnetotelluric data for monitoring of geoelectrical section parameters. *Annals of Geophysics*, 50(1).

Shockley, E. M., Vrugt, J. A., and Lopez, C. F. (2017). PyDREAM: high-dimensional parameter inference for biological models in python. *Bioinformatics*.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489.

Siripunvaraporn, W., Egbert, G., Lenbury, Y., and Uyeshima, M. (2005). Three-dimensional magnetotelluric inversion: data-space method. *Physics of the Earth and Planetary Interiors*, 150(1):3–14.

Spichak, V. and Popova, I. (2000). Artificial neural network inversion of magnetotelluric data in terms of three-dimensional earth macroparameters. *Geophysical Journal International*, 142(1):15–26.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958.

Sun, J. and Li, Y. (2012). Joint inversion of multiple geophysical data: A petrophysical approach using guided fuzzy c-means clustering. In *SEG Technical Program Expanded Abstracts 2012*, pages 1–5. Society of Exploration Geophysicists.

Tarantola, A. (2005). *Inverse Problem Theory and Methods For Model Parameter Estimation*. SIAM.

Tarits, P., Jouanne, V., Menvielle, M., and Roussignol, M. (1994). Bayesian statistics of non-linear inverse problems: example of the magnetotelluric 1-D inverse problem. *Geophysical Journal International*, 119(2):353–368.

Tatar, A., Naseri, S., Sirach, N., Lee, M., and Bahadori, A. (2015). Prediction of reservoir brine properties using radial basis function (RBF) neural network. *Petroleum*, 1(4):349–357.

Tikhonov, A. (1963). On the solution of incorrectly put problems and the regularisation method. *Outlines Joint Sympos. Partial Differential Equations (Novosibirsk, 1963)*, pages 261–265.

Tikhonov, A. (1986). On determining electrical characteristics of the deep layers of the earth's crust. *Magnetotelluric Methods*, pages 2–3.

Vilamajó, E., Queralt, P., Ledo, J., and Marcuello, A. (2013). Feasibility of monitoring the Hontomín (Burgos, Spain) CO2 storage site using a deep EM source. *Surveys in Geophysics*, 34(4):441–461.

Wait, J. R. (1962). Theory of magnetotelluric fields. *J. Res. NBS*, 66(5):509–541.

Wang, B., Basart, J. P., and Moulder, J. C. (1997). *Fast Eddy Current Forward Models Using Artificial Neural Networks*, pages 217–224. Springer US, Boston, MA.

Wang, K. and Zhang, L. (2008). Predicting formation lithology from log data by using a neural network. *Petroleum Science*, 5(3):242–246.

Yang, H., Wang, J.-L., Wu, J.-S., Yu, P., and Wang, X.-M. (2002). Constrained joint inversion of magneto-telluric and seismic data using simulated annealing algorithm. *Chinese Journal of Geophysics*, 45(5):764–776.

Yu, L., Porwal, A., Holden, E.-J., and Dentith, M. C. (2012). Towards automatic lithological classification from remote sensing data using support vector machines. *Computers & Geosciences*, 45:229–239.

Zhang, K. and Yan, J. (2015). A 3D magnetotelluric inversion system with a theoretical assessment in oil and gas exploration. *Acta Geologica Sinica (English Edition)*, 89(s1):434–436.

Zhang, Y. and Paulson, K. (1997). Magnetotelluric inversion using regularized Hopfield neural networks. *Geophysical prospecting*, 45(5):725–743.