
Faculty of Science

Faculty Publications

A reversible-jump Markov chain Monte Carlo algorithm for 1D inversion of magnetotelluric data

Eric Mandolesi, Xenia Ogaya, Joan Campaña, Nicola Piana Agostinetti

2018

© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

This article was originally published at:

<https://doi.org/10.1016/j.cageo.2018.01.011>

Citation for this paper:

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

<https://doi.org/10.1016/j.cageo.2018.01.011>



Research paper

A reversible-jump Markov chain Monte Carlo algorithm for 1D inversion of magnetotelluric data

Eric Mandolesi^{a,b}, Xenia Ogaya^b, Joan Campanya^{b,1}, Nicola Piana Agostinetti^{b,c,*}^a University of Victoria, Victoria, BC, Canada^b Geophysics Section, Dublin Institute for Advanced Studies, DIAS, Dublin, Ireland^c Department of Sedimentology and Geodynamics, University of Vienna, Vienna, Austria

ARTICLE INFO

Keywords:
trans-d inversion
Magnetotellurics
rjMCMC
Clare basin

ABSTRACT

This paper presents a new computer code developed to solve the 1D magnetotelluric (MT) inverse problem using a Bayesian trans-dimensional Markov chain Monte Carlo algorithm. MT data are sensitive to the depth-distribution of rock electric conductivity (or its reciprocal, resistivity). The solution provided is a probability distribution - the so-called posterior probability distribution (PPD) for the conductivity at depth, together with the PPD of the interface depths. The PPD is sampled via a reversible-jump Markov Chain Monte Carlo (rjMCMC) algorithm, using a modified Metropolis-Hastings (MH) rule to accept or discard candidate models along the chains. As the optimal parameterization for the inversion process is generally unknown a trans-dimensional approach is used to allow the dataset itself to indicate the most probable number of parameters needed to sample the PPD. The algorithm is tested against two simulated datasets and a set of MT data acquired in the Clare Basin (County Clare, Ireland). For the simulated datasets the correct number of conductive layers at depth and the associated electrical conductivity values is retrieved, together with reasonable estimates of the uncertainties on the investigated parameters. Results from the inversion of field measurements are compared with results obtained using a deterministic method and with well-log data from a nearby borehole. The PPD is in good agreement with the well-log data, showing as a main structure a high conductive layer associated with the Clare Shale formation.

In this study, we demonstrate that our new code go beyond algorithms developed using a linear inversion scheme, as it can be used: (1) to by-pass the subjective choices in the 1D parameterizations, i.e. the number of horizontal layers in the 1D parameterization, and (2) to estimate realistic uncertainties on the retrieved parameters. The algorithm is implemented using a simple MPI approach, where independent chains run on isolated CPU, to take full advantage of parallel computer architectures. In case of a large number of data, a master/slave approach can be used, where the master CPU samples the parameter space and the slave CPUs compute forward solutions.

1. Introduction

Magnetotellurics is an electromagnetic (EM) passive method that infers the electrical conductivity of the subsurface structures by measuring orthogonal components of electric and magnetic fields on the Earth surface and relies on accurate data and a robust inversion algorithm. Knowledge of subsurface electrical conductivity distribution also plays a major role in exploration geophysics as well as groundwater monitoring. Most of the recent MT inversion works explore three-dimensional (3D) models, obtained by partitioning the subsurface in large regions of constant electrical conductivity and solving the inverse

problem via a linearization of the forward MT equation [e.g. Kelbert et al., 2014]. The one-dimensional (1D) MT inverse problem has been solved following a broad range of different approaches, both linearized and stochastic [e.g. Grandis et al., 1999]. However linearized approaches provides weak solutions, whilst popular stochastic approaches (e.g. Genetic Algorithms) rely on appropriate parameters (i.e. number of layers in partition, model space dimension, number of iterations) chosen ad-hoc, often not justified by the data. One way to avoid the bias in obtained models caused by a poor selection of the inversion parameters is provided by inverting the data within a Bayesian framework, so that the algorithm itself can auto-adjust the parameters based on the data. For example

* Corresponding author. Department of Geodynamics and Sedimentology, University of Vienna, Vienna, Austria.

E-mail address: piana@cp.dias.ie (N. Piana Agostinetti).

¹ Now at: School of Physics, Trinity College, Dublin, Ireland.

Malinverno (2002) use Bayesian inversion to properly select an appropriate number of layers in a direct current EM method. More recently Minsley (2011) solved a different EM inverse problem (frequency domain electromagnetic–FDEM: it makes use of a loop-loop active source.) by the adopting of a reversible jump Markov chain Monte Carlo (rjMCMC) algorithm. However, despite the description of rjMCMC algorithms in geophysics literature and the availability of several codes and libraries for its general implementation (e.g. <http://www.earth.org.au/codes/rj-MCMC/>), is not publicly available to date (to the authors knowledge) a rjMCMC implementation that solves the MT 1D inverse problem.

RjMCMC algorithm, developed at first by Green (1995), has been used to invert data from very diverse geophysics methods (see Sambridge et al. (2013) for a review). Examples exists in seismology (Bodin et al., 2012; Piana Agostinetti et al., 2015), thermochronology (Stephenson et al., 2006; Gallagher et al., 2009), stratigraphy (Charvin et al., 2009), paleoclimate records (Hopcroft et al., 2007, 2009), mixture modelling in geochronology (Jasra et al., 2006) geoacoustics (Dettmer et al., 2010, 2012; Steininger et al., 2013), marine electromagnetic inversion (Ray and Key, 2012), mantle viscosity inversion (Rudolph et al., 2015) to cite some works in tens.

The present paper develops a trans-dimensional Monte Carlo algorithm for MT 1D inversion in a Bayesian framework, as well as its computer parallel implementation. The mutual independence of the models in the Markov chains allows the parallelization of the code with minimum effort. Test are performed with either simulated data and field data from Clare Basin (Ireland).

2. Material and methods

The algorithm presented in this study is composed of two main codes: a rjMCMC sampler for extracting model according to the PPD, and a MT 1D forward solver, for computing synthetic MT responses. Here we tested our routines against two synthetic simulations, using synthetic models previously adopted in literature, and field data recorded in Western Ireland.

2.1. Simulation and field details

The models for simulations were selected from Malinverno (2002) and Cerv et al. (2007). The first synthetic model is composed of a single 1 km-thick conductive layer in a relative (one fold) resistive rock matrix. The second model, more complex, is represented by a realistic stack of conductive layers, where a thin, very conductive layer (250 m-thick with resistivity equals to $10 \Omega \cdot m^{-1}$) sits at the bottom of a stack of layer with intermediate conductivity (Cerv et al., 2007). The two models are reported in Tables 1 and 2 respectively. Field data were collected in a sedimentary basin, the Clare Basin (Co. Clare, Ireland), during summer 2014 as part of the IRECCSEM project (www.ireccsem.ie). The MT data were recorded for three days using Phoenix Geophysics audio-magnetotellurics and broadband magnetotellurics MTU_V5A systems. The processing of the MT data was carried out using Egbert (1997) and Smirnov (2003) algorithms using both the Remote Reference technique (Gamble et al., 1979a, 1979b) and the ELICIT methodology (Campanya et al., 2014). The remote reference method multiplies the equations relating the Fourier components of the electric and magnetic fields from the local site by a component of magnetic field from the remote reference site. By averaging the cross products obtain estimates of the impedance tensor that are unbiased by noise, provided there are no correlations

Table 1
Synthetic model used in the algorithm test 1. Layers are numbered top to bottom.

Layer number	Layer thickness (km)	Resistivity ($\Omega \cdot m$)
1	1.0	1000
2	1.0	100
3	∞	1000

Table 2
Synthetic model used in the algorithm test 2. Layers are numbered top to bottom.

Layer number	Layer thickness (km)	Resistivity ($\Omega \cdot m$)
1	0.60	250
2	0.40	25
3	2.00	100
4	0.25	10
5	∞	1000

between noises in the remote reference site and the local site. The remote reference method is applied to reduce the influence of correlated and uncorrelated local noise, and have the capacity to reduce bias and source effects of correlated noise. However, one of the main issues associated with this method is that the performance decrease for longer periods due to the reduced number of samples used when processing the MT data. The ELICIT method, based on the inter-station tensor relationship between electric and magnetic fields measured at different locations, combines these inter-station tensor relationships to calculate the local MT impedance tensor, which can be used to increase the number of samples when processing the MT data. The dimensionality of the geo-electrical structures was appraised evaluating the phase tensor (Caldwell et al., 2004; Rider; Farrelly et al.). The aim of this analysis was to detect and avoid the influence of more complex geoelectrical features on the MT data not possible to explain with a 1-D MT approach. Frequencies between 900 Hz and 0.43 Hz were selected for the inversion process, avoiding in this way the frequencies with bigger skew values associated with the presence of more complex geoelectrical environments, like coastal effects. The geology of the area around the MT site is characterised by five main geological formations (Farrelly et al.; Rider), from top to bottom: (1) The Central Clare Group characterised as a thin layer of a mix of sandstones, siltstone and mudstone; (2) the Gull Island Formation, which is divided into an upper section (primarily siltstones) and a lower section (siltstones interbedded with bundles of sheet sandstones); (3) the Ross Sandstone Formation, primarily composed of sandstones (around 65%) with subordinated interbedded shale and slumped horizons of mixed lithology; (4) the Clare Shale Formation, dominated by black laminated shale containing fossiliferous marine bands within it, and (5) the Visean Limestone. The depths and thicknesses of the different formations are reported by the Doonbeg borehole data, located close to the studied MT site, and are used to test the reliability and geological consistency of the obtained 1D models (Fig. 1 b,c).

2.2. A MT forward solver for a layered isotropic earth

The MT method allows the subsurface distribution of electrical conductivity (or its reciprocal, electrical resistivity) to be inferred by measuring the horizontal variations of the electromagnetic field at the Earth surface (Cagniard, 1953; Tikhonov). The electric and magnetic field variations recorded at the surface are related to each other through the MT impedance tensor Z . The impedance tensor is a second-rank tensor function of frequency that for a uniform half space can be expressed as the relation between orthogonal electric and magnetic field variations.

$$Z(\omega) = \frac{E_x(\omega)}{H_y(\omega)} = \sqrt{i\omega\mu_0\rho} \quad (1)$$

where ω is the angular frequency of the electromagnetic source, E_x is the x-component of the electric field at the Earth surface, H_y is the y-component of the magnetic field at the earth surface, i is the imaginary unit, μ_0 is the magnetic permeability in the free space and ρ is the electrical resistivity of the uniform half-space. Equivalent results would be obtained by comparing E_y and H_x .

In a layered half space, the solution of the impedance tensor can be propagated from the bottom uniform half space (bottom layer of the

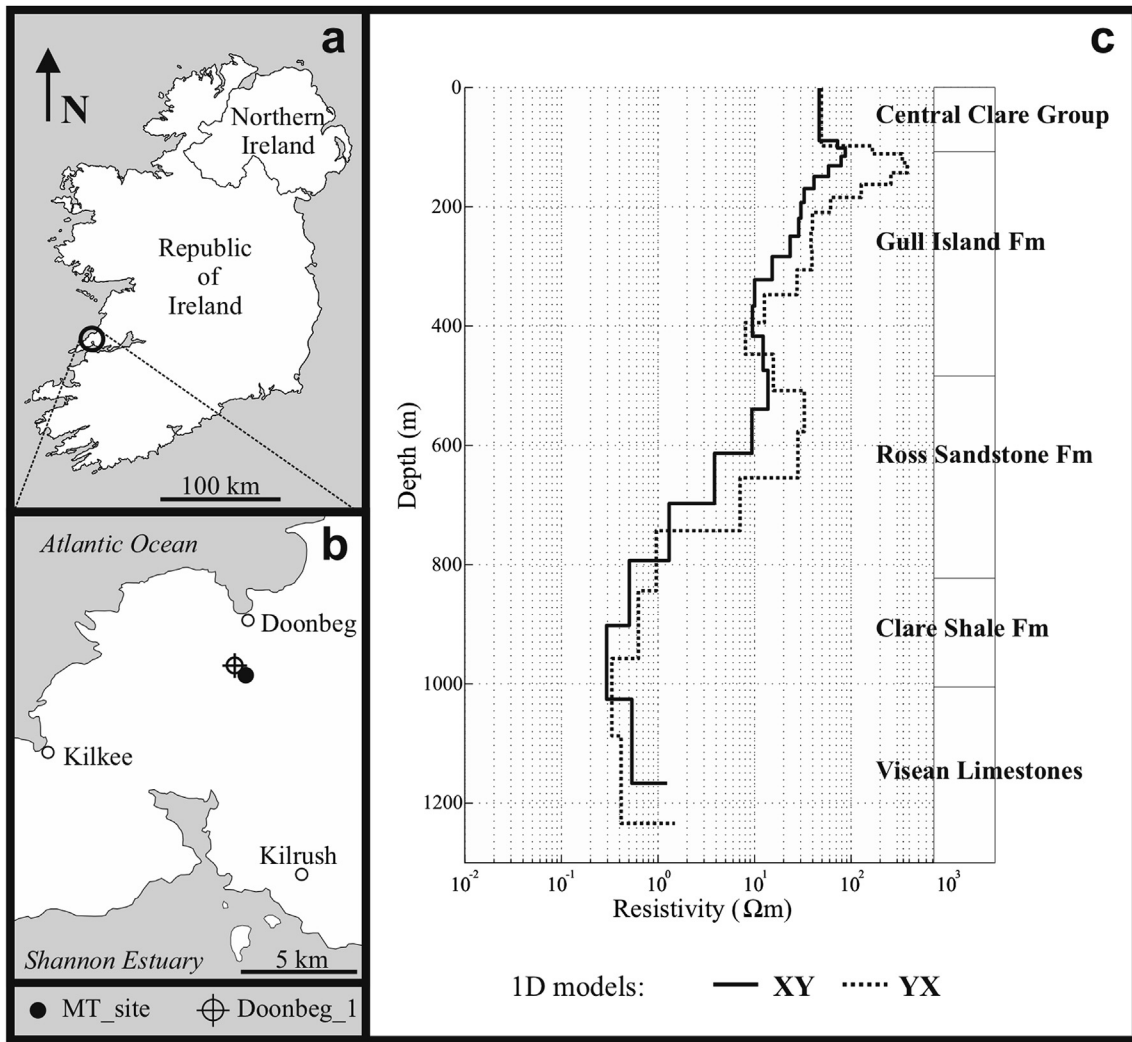


Fig. 1. a, b: position of the MT test site and the Doonbeg borehole; c: results from Occam's inversion of the MT site data compared with the geological features retrieved from the borehole. The models are reported also in Fig. 16 for comparison with posterior marginal distributions.

model) to the topmost one (Kaufman, 1981). In this way, the earth electrical impedance at the surface is effectively predicted if the thickness and electrical resistivity of each layer of the model are known. Numbering the layers in sequential order from the shallower to the deeper, the impedance of the layer j depends on the reflection coefficient R_j of the interface j

$$Z_j = \sqrt{i\mu_0\omega\rho_j} \left[1 - \frac{2R_j}{R_j + e^{2h_j\sqrt{i\mu_0\omega\rho_j}}} \right] \quad (2)$$

and R_j depends on the impedance of layer $j+1$

$$R_j = \sqrt{i\mu_0\omega\rho_j} \frac{1 - Z_{j+1}}{1 + Z_{j+1}}, \quad (3)$$

where h_j is the thickness of the layer j and $\rho_j = \sigma_j^{-1}$ its resistivity.

It is common in the MT community to display data in terms of apparent resistivity ρ_a and phase ϕ , which are related to the complex impedance Z as $\rho_a(\omega) = \|Z(\omega)\|^2 \mu_0 \omega$ and $\phi = \tan^{-1} \frac{\Im\{Z(\omega)\}}{\Re\{Z(\omega)\}}$. Nonetheless in this study the real and imaginary part of the MT impedance tensor are plotted and analyzed, as these values can be directly compared with values predicted by the MT forward solver.

2.3. *rjMCMC algorithms in the context of trans-dimensional bayesian inversion*

In the Bayesian approach, the solution of an inverse problem is a probability distribution for the model parameters given the measured data, the so-called PPD. The PPD is given by Bayes' rule as

$$p(\mathbf{m}|\mathbf{d}, \mathcal{I}) = \frac{p(\mathbf{m}|\mathcal{I})p(\mathbf{d}|\mathbf{m}, \mathcal{I})}{p(\mathbf{d}|\mathcal{I})}, \quad (4)$$

where $p(\cdot)$ is a probability distribution, $\star|\star$ denotes the conditional dependance, \mathbf{d} is the vector of measured data and \mathbf{m} is the vector that stores the model parameter values and \mathcal{I} are the non-physical information, such as the parameterization itself, also called *nuisance parameters*. With this notation, $p(\mathbf{m}|\mathcal{I})$ is the prior probability distribution, $p(\mathbf{d}|\mathbf{m}, \mathcal{I})$ is the likelihood function and $p(\mathbf{d}|\mathcal{I})$ is a normalisation factor (the so-called evidence) and can be written as

$$p(\mathbf{d}|\mathcal{I}) = \int p(\mathbf{d}|\mathbf{m}, \mathcal{I})p(\mathbf{m}|\mathcal{I})d\mathbf{m}. \quad (5)$$

As conductivity values in earth materials may on many orders of magnitude (typically between 10^{-8} and $10^2 \frac{\text{S}}{\text{m}}$ (Chave and Jones, 2012)) the conductivity distribution is sampled here in log space. The trans-dimensionality in the inversion scheme is given by the fact that the

number of parameters from which the PPD pdf is estimated is a parameter itself (i.e. the layering is included as part of the posterior to be sampled). The implicit parsimony of this approach - the algorithm preferably samples models that fit the data using parameterizations that produces the highest evidence, in relation to the data uncertainties - has been shown by other authors (Malinverno, 2002; Piana Agostinetti and Malinverno, 2010).

Despite the simple definition given in Equation (4), an analytical expression of the PPD is not generally available for non-linear problems and the PPD must be evaluated for a large number of models. A good compromise is to sample the PPD using, e.g., a Markov chain Monte Carlo algorithm. Moreover, the evidence term plays a major role in operating model selection (Sambridge et al., 2006), and its evaluation can be extremely challenging. A way to implicitly account for evidence is through the rjMCMC (Green, 1995). In geophysics, MCMC methods have been proven to be an efficient choice to sample the PPD in both classic (Mosegaard and Tarantola, 1995; Mosegaard et al., 1997) and recent (Grandis et al., 2013; Ray et al., 2014) papers.

For this study we use the popular Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) which is an MCMC implementation constituted by a two step procedure. At a first stage a candidate model \mathbf{m}' is derived by the current model \mathbf{m} of the chain. Such a candidate is chosen from a proposal distribution $q(\mathbf{m}'|\mathbf{m})$ in which the parameters of the proposed model \mathbf{m}' depend at most on the parameters of the perturbed model \mathbf{m} . At a second stage, \mathbf{m}' is accepted or rejected following the Metropolis rule, namely the probability α of acceptance of \mathbf{m}' is given by

$$\alpha = \min \left[1, \underbrace{\frac{p(\mathbf{m}'|\mathcal{I})}{p(\mathbf{m}|\mathcal{I})}}_{\text{prior ratio}} \cdot \underbrace{\frac{p(\mathbf{d}|\mathbf{m}',\mathcal{I})}{p(\mathbf{d}|\mathbf{m},\mathcal{I})}}_{\text{likelihood ratio}} \cdot \underbrace{\frac{p(\mathbf{m}|\mathbf{m}')}{p(\mathbf{m}'|\mathbf{m})}}_{\text{proposal ratio}} \right] \cdot |J| \quad (6)$$

It is important to bear in mind that the trans-dimensionality is hidden in the model arrays (i.e. \mathbf{m} and \mathbf{m}') given that the number of parameters to be estimated is also a parameter and is included in the vector \mathbf{m} as a vector element. The term $|J|$ makes reference to the Jacobian for the transformation from \mathbf{m} to \mathbf{m}' , and it is always one if \mathbf{m} and \mathbf{m}' have the same parameterization. We implement the transition from a state with n parameters to a state with n' parameters, following Piana Agostinetti and Malinverno (Piana Agostinetti and Malinverno, 2010), so that we always have $|J| = 1$ and its contribution cancels from the equation.

The details related to the implementation of the process that brings to the candidate model –the so-called “recipe” – define the algorithm flavour and its performance. While many different recipes can be configured for sampling the PPD, finding an efficient recipe requests a certain number of tentatives. In fact, if a recipe proposes a complete change in all the parameters of model, the candidate model is likely to produce a synthetic MT response totally different from the response generated by the current model, giving rise to a low probability for accepting the candidate model. In this case, the chain risks to remain stuck in the same position for a long portion of the MCMC. On the other hand, if the candidate model only displays very minor changes with respect to the current model, the acceptance probability is likely to be extremely high, but the algorithm will take a long time to reach a satisfactory exploration of the model space. Defining a reasonable acceptance ratio of candidate models along a MCMC sampling is a field of current research. Broadly, an acceptance ratio between 20% and 40% should be considered (Mosegaard, 2006). In our implementation the candidate model is obtained from the current model by the mean of perturbation of one of its parameters. A “model” is composed of a stack of k layers separated from horizontal interfaces, and it can be described as a vector with $2(k+1)$ elements, i.e. $\mathbf{m} = (k, h_1, \dots, h_k, \rho_1, \dots, \rho_{k+1})$, where k is the number of interfaces in the model. For each interface, one parameter h_i displays its current depth, while the second one ρ_i represents the

resistivity above such interface. As expected, layers are limited by two interfaces, or by one interface and the free-surface. An additional parameter ρ_{k+1} measures the resistivity of the bottom half-space.

In our “recipe”, the possible perturbations (so called *Moves*) are:

1. Perturb the value of electrical resistivity associated to an interface (i.e. within the layer above it);
2. Perturb the resistivity in the half-space;
3. Create an interface at a random position;
4. Delete an interface at random;
5. Perturb the position of one interface

The value of the resistivity in one layer is perturbed in *Move 1*. We refer to Mosegaard and Tarantola (1995) and Appendix A in Piana Agostinetti and Malinverno (Piana Agostinetti and Malinverno, 2010) for the details of the sampling procedure. Briefly, the value of resistivity related to a randomly-picked interface is perturbed according to a proposal distribution which asymptotically samples the prior probability distribution. Resistivity in the half-space is perturbed within the same scheme, in *Move 2*.

The *Move 3* and *Move 4* are referred as “birth” and “death” respectively, and are responsible for the variation of the number of dimensions. These are the so called “trans-dimensional moves”. As it is possible to step back from a model with a different number of dimensions to the previous one, this kind of implementation is also known as reversible-jump MCMC. In our case, the amplitude of the variation of the number of dimensions is one (i.e. only one interface can be added or deleted from a current model to generate the candidate model). This algorithm is also known birth-death MCMC (Geyer and Møller, 1994). During the “birth move”, a new interface is added to the model, and its depth and resistivity are sampled from the prior distributions. It is important to notice that, adopting this scheme, the proposal distribution is exactly the prior distribution (see below). Conversely, the “death move” implies the removal of a randomly selected interface.

The *Move 5* consists in a small perturbation of the depth of one interface of the model. This guarantees a small variation in the likelihood between the current and the candidate models. However, to prevent the creation of a too thin layer, if an interface is moved too close to another interface, the move is rejected. The minimum distance is below. Conversely, to enlarge the exploration of the model space, this move also includes the possibility of deleting an interface and recreating it at random in the model (i.e. operating both *Move 2.3* and *Move 3* in sequence). This second case keeps the number of parameters at the level of the current model (i.e. it is not a trans-dimensional move, see below), but allows for an enlarged exploration of the model space. We observe that the re-created interface does not keep the resistivity associated to the removed interface. Instead, both its depth and resistivity are sampled from their own prior distributions.

To speed-up the computation, we set a minimum distance between the interfaces, h_{min} , which depends on the MT frequencies considered. Adopting a minimum distance does not modify the retrieved PPD, as far as h_{min} is kept to a reasonable value (Malinverno, 2002), here set to $h_{min} = 20$ m: as rule of thumb h_{min} should be large enough to prevent extremely thin layers and small enough to be transparent to high-frequency data. The parameter h_{min} is considered during the perturbation associated to moves 3 and 5.

The inverse problem is thus reduced, now, to the estimation of the quantities needed to compute α . We opt to follow the approach of Mosegaard and Tarantola (1995), namely we build a prior sampler so that the prior and the proposal distribution are equal, and Equation (6) is simplified as

$$\alpha = \min \left[1, \frac{p(\mathbf{d}|\mathbf{m}',\mathcal{I})}{p(\mathbf{d}|\mathbf{m},\mathcal{I})} \right] \quad (7)$$

The observed data \mathbf{d}^{obs} can be written in general as

$$\mathbf{d}^{\text{obs}} = \mathbf{d}^{\text{true}} + \boldsymbol{\varepsilon} \quad (8)$$

where \mathbf{d}^{true} denotes the array of the true data and $\boldsymbol{\varepsilon}$ their relative errors. If the assumption that $\boldsymbol{\varepsilon}_i$ is distributed according to a normal multivariate distribution $\forall i$ with covariance $\overline{\mathbf{C}}_e$ and mean $E[\boldsymbol{\varepsilon}_i] = 0$ is made, the likelihood function $p(\mathbf{d}|\mathbf{m}, \mathcal{I})$ – the function that represents the goodness a specific model with a particular set of parameters fits the observed data – is computed as

$$p(\mathbf{d}|\mathbf{m}, \mathcal{I}) = \frac{e^{-\frac{1}{2}\mathbf{e}^T\overline{\mathbf{C}}_e^{-1}\mathbf{e}}}{[(2\pi)^N \det(\overline{\mathbf{C}}_e)]^{\frac{1}{2}}} \quad (9)$$

where N is the number of observed data, $\mathbf{e} = \mathbf{d}^{\text{obs}} - \mathbf{g}(\mathbf{m})$ is the difference between the N measured data \mathbf{d}^{obs} and the N data predicted by the model \mathbf{m} , $\mathbf{g}(\mathbf{m})$, $\overline{\mathbf{C}}_e$ is the covariance matrix for the data \mathbf{d}^{obs} and $\mathbf{g}(\star)$ is the forward operator that depends on the specific problem solved.

As results from Eq. (9), data covariance $\overline{\mathbf{C}}_e$ plays a key role in the likelihood computation. In MT inversions is common practice to assume that the data are uncorrelated, i.e. $\overline{\mathbf{C}}_e$ is diagonal [cf. e.g. Guo et al., 2011; Chave and Jones, 2012; Rodi and Mackie, 2001]. Nonetheless, the data frequency-correlation is a well-known effect and has been appreciated and investigated in several classic (Egbert and Booker, 1986; Chave et al., 1987) and recent (Guo et al., 2014) works. Although the effect of ignoring the data correlation in the inversion results has been investigated for some few-layer models (Guo et al., 2014), it seems to lead to underestimate the parameter uncertainties and is at the moment poorly understood. Thus, despite debated, the assumption of uncorrelated error in MT data (i.e. $\overline{\mathbf{C}}_e$ is diagonal) is common practice in MT inversions and is used in this paper. With this further assumption the argument in the exponential function of the Equation (9) is simplified in the classic L2-norm, and the maximum likelihood principle corresponds with the minimisation of the residual mean square.

It is worth to note that MCMC methods were used in other works to probabilistically adjust the error level. It is easy in theory to cast the problem using hierarchical models using an hyper-parameter as a multiplier that scale the values of $\overline{\mathbf{C}}_e$ [e.g. Bodin et al., 2012] or even sampling the likelihood in correspondence of the maximum with respect to standard deviations [e.g. Dosso and Wilmot, 2006], effectively implicit sampling data variances. These methods in principle allow for an improved estimation of the error statistics, but are often cause of the introduction of complexity. In this paper the error is considered known to avoid data over-fitting issues, these methods can be implemented as a future work.

2.4. Parallel implementation

We develop the code following a simple parallel implementation, running a number of parallel and independent rjMCMC samplings, one for each CPU available in the cluster. Due to the nature of the Monte Carlo sampling, increasing the number of chains increases the exploration of the model space and reduces the length of each single chain. However, in case of a large number of synthetic MT responses to be computed, the algorithm can be easily structured following a master/slave approach, where a master CPU defines the perturbation and sends the candidate model to the slave CPUs for computing the synthetic MT responses. In this case, for each candidate model the synthetic MT response can be computed by a number of slave CPUs, which can be a advisable feature if the algorithm should be extended to, e.g. 3D MT inversion. Due to the fact that the MT responses at different frequencies are computed separately, the CPU-time for the forward computation scale almost perfectly with the number of CPUs. At the end of the last rjMCMC, inferences about the PPD of the investigated parameters are drawn from the full ensemble of models extracted from the PPD from all the independent chains. In our 1D implementation, the CPU-time needed

for computing forward solutions at all frequency considered is limited, so that all our tests run in less than 1 h on a Linux cluster (one chain, one CPU).

Convergence of trans-D PPD may be a slow process, and the efficiency of the rjMCMC algorithm requires mixing of the Markov chains that sample the space. Parallel tempering (PT) is a method that allow some chains to collect samples from distributions "tempered" at different sampling temperature $T > 1$. That means the likelihood function is raised to the power $T > 1$, for some chains, while other chains still sample the original PPD ($T = 1$) (Geyer, 1991; Sambridge, 2014). The candidate models proposed by tempered chains ($T > 1$) are accepted with higher probability and concur to a better mixing of the chains. The main drawback of PT is that the tempered chains cannot be used efficiently to collect samples from the PPD. So, the PT algorithm increases mixing, but require a larger number of interacting chains. In our case, given the likelihood function and the recipe we have, chain mixing is achieved even without PT. PT is an advanced sampling technique and, even if not used in this work, it can improve the efficiency of trans-D algorithms in many cases, making, for example, 2D and 3D MT trans-D inversions feasible.

3. Results and discussion

In this section we present results obtained using the rjMCMC algorithm with both synthetic dataset and field data collected in Ireland. Since the sampler is implemented as a computer code, it is important to run a preliminary test to ensure the program works correctly on synthetic data with added noise before inverting field measurements.

3.1. Sampling the prior probability distributions

First of all, a test was undertaken to study how the algorithm samples the prior distribution used. This constitutes a crucial step to assure that the prior ratio and the proposal ratio simplifies in the Equation (6) and thus, that the algorithm works properly (Piana Agostinetti and Malinverno, 2010).

This test consists of sampling the prior probability distribution by imposing $p(\mathbf{d}|\mathbf{m}) = 1 \forall \mathbf{m}$ (which returns, with the described algorithm, $\alpha = 1$ for all the proposed models which are in the prior distribution). This constraint forces the algorithm to accept all the proposed models, and if the final distribution converges to the prior itself, this proves that the prior sampler is working properly.

For this test, we used uniform prior probability distributions over all the model parameters (investigated depth, number of interfaces and logarithmic value of the electrical resistivity). The investigated depth range was set between 0 and 10 km. The candidate models were allowed to have from a minimum of 1 interface to a maximum of 100 interfaces. In other words, we considered from simple half-space models to models composed of 100 layers (the air interface is always included). The log-values of the electrical resistivity were allowed to vary within the interval $[-2 \leq \log \rho \leq 5] \Omega m$.

Marginal probability distributions for interface depth and resistivity are reported in Fig. 2, the distribution of the number of interfaces is reported in Fig. 3. The prior has been sampled using 191 Markov chains with 500.000 models per chain. The distribution on the number of layer within the MCMC is almost uniform, indicating that, when data are not considered, the algorithm is not biased toward a subjective number of layers (i.e. imposed resolution). Fig. 2 display the distribution of the depth of the interfaces and the resistivity at depth. As expected, the interfaces are inserted uniformly at each depth level, and the probability distribution for the resistivity is uniform everywhere in the investigated depth range.

3.2. Test 1

A classic resistor–conductor–resistor model was considered for the

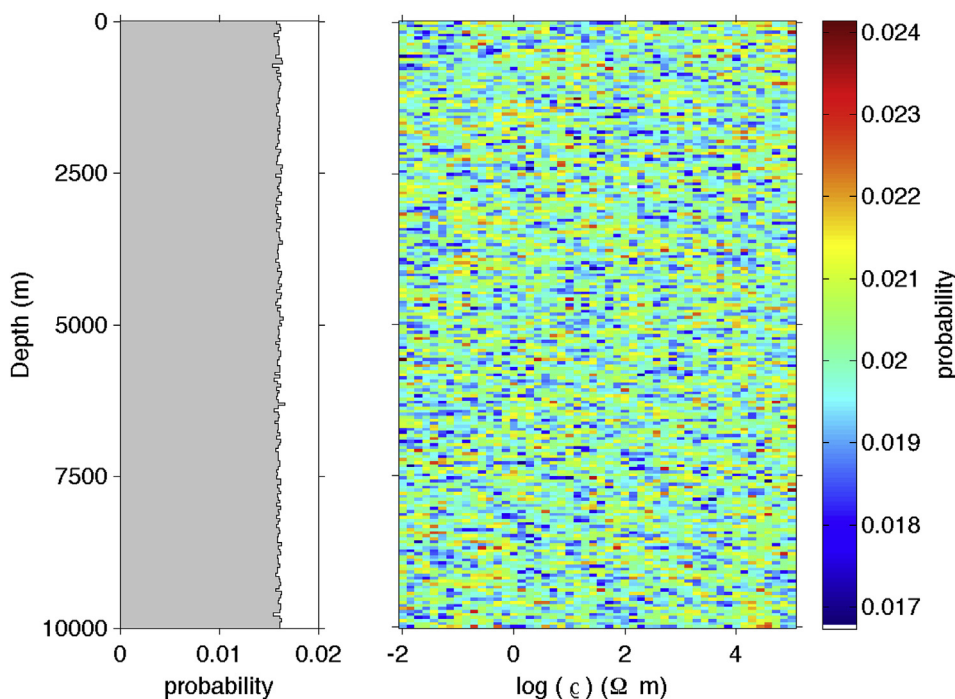


Fig. 2. Results for the prior sampling test. In this test, all candidate models are accepted, without computing the Likelihood ratio. Thus, final 1D marginal distribution of the investigated parameters is equal to the prior probability distribution. **Left panel:** prior probability distribution of the depth of layer interfaces; **right panel:** prior probability distribution for the electrical resistivity. For this test a uniform prior distribution was chosen, with resistivity values drawn in the interval $\log_{10}(\rho) \in [-2, 5]\log_{10}(\Omega\text{m})$.

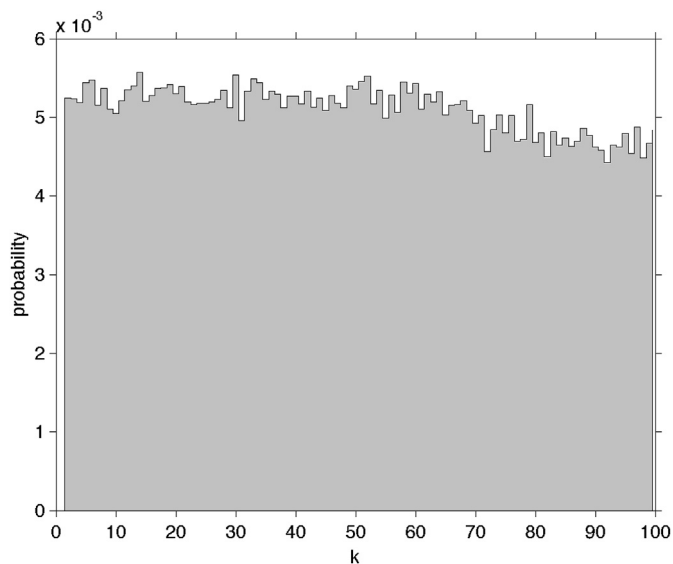


Fig. 3. Prior probability distribution of the number of layers.

first synthetic test. The model used is reported in Table 1. We chose this 3-layered Earth because the MT method can easily image the top surface of a conductive layer since it is mainly sensitive to charges and currents located on more conductive surfaces. Simulation responses were created for 41 log-uniform distributed frequencies between 10^{-1} and 10^3 Hz. Data errors were considered un-correlated (i.e. C_e is diagonal) with a 5% relative amplitude, and this source of uncertainties has been added to the synthetic responses.

The PPD inference was made with a set of 95 Markov parallel chains. These chains sampled 10^6 models each. Results from the first 10^5 models were discarded to be sure that we were sampling the PPD when the algorithm was in convergence - this phase is usually called *burn-in* (Mosegaard and Sambridge, 2002). The PPD sample was built collecting one model each 1000 accepted models along the Markov chain to

enhance the statistical independence of accepted models and reduce the sample autocorrelation within the chain.

For this first test, the algorithm easily recognises that three interfaces are needed to fit the data (Fig. 4): the air-earth interface (always present) and the two interfaces that bound the conductive layer. Their position in depth is also well defined, as it is shown in the left panel of Fig. 5. Fig. 6 shows the fit to the data. In the right panel of Fig. 5, the PPD for the $\log_{10}(\rho)$ is shown (the synthetic model is marked as dashed black line for reference). The distribution is in agreement with the test model, clearly marking that the conductive layer is most probable between 1 km and 2 km depth. The absolute resistivity below the conductor is most likely between 600 and 1500 Ωm . The position of the lower interface that bounds the conductive layer is less resolved with respect to the upper one, and translates into a broader PPD for the resistivity in the 2 – 4 km depth range. Such feature is a common characteristic of the inversion of MT data in presence of a strong conductive layer, where the top of the conductor is better constrained than the position of its bottom interface [cf. Chave and Jones, 2012].

Using the proposed approach is thus possible to successfully invert for noisy data. Moreover, the results show that it is possible to have an estimation of the credibility intervals on the retrieved parameters. Although the simple structure of the model proposed is debatable, the method is proofed to work correctly.

3.3. Test 2

The results from the previous test encouraged us to invert for a synthetic dataset generated from a more complex model. In this case we utilize a model from Cerv et al. (2007), adopting the same error statistics and the same data structure (i.e. computing the synthetic responses for the same frequencies). The true model is described in Table 2.

For this test, we run 95 parallel Markov chains, each constituted by 10^6 iterations using a uniform prior probability distribution with resistivity values drawn in the interval $\log_{10}(\rho) \in [-0.5, 3.5]\log_{10}(\Omega\text{m})$. As in the previous test, we accepted in the sampled models set one model in a set of 1000 models sampled by the rjMCMC discarding the first 10^5 sampled models to properly burn-in.

The sampled quantities are reported in Figs. 7 and 8. In Fig. 7 is shown

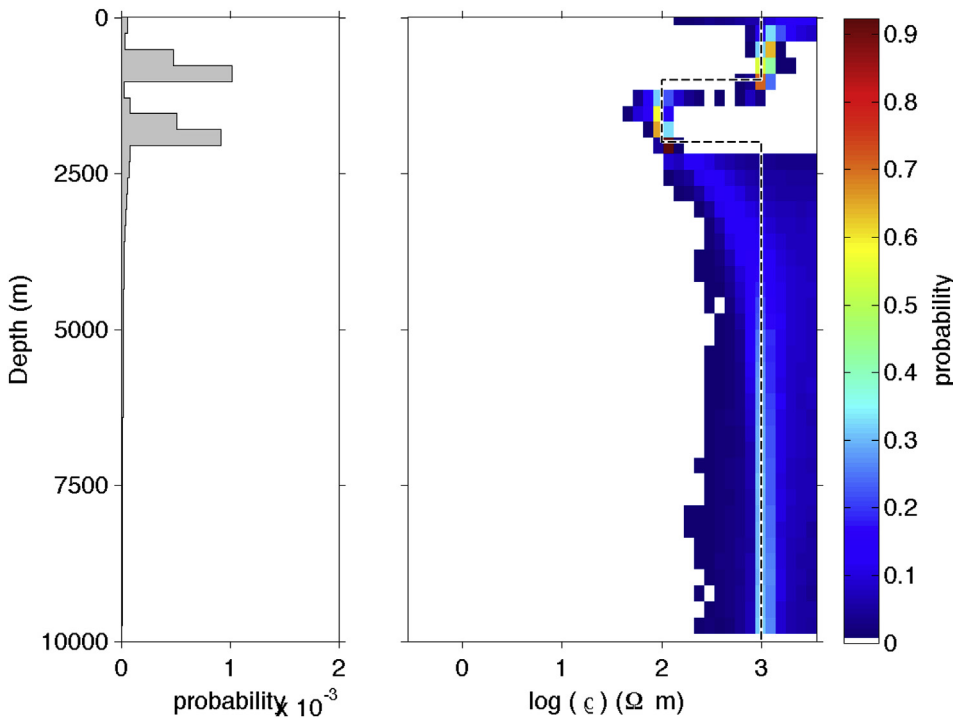


Fig. 4. Test 1. Posterior probability distribution of the number of interfaces. The dashed line marks the true model. Due to the parsimony of the algorithm, models with a limited number of interfaces are sampled more frequently, even if model with a large number of interfaces are sampled as well.

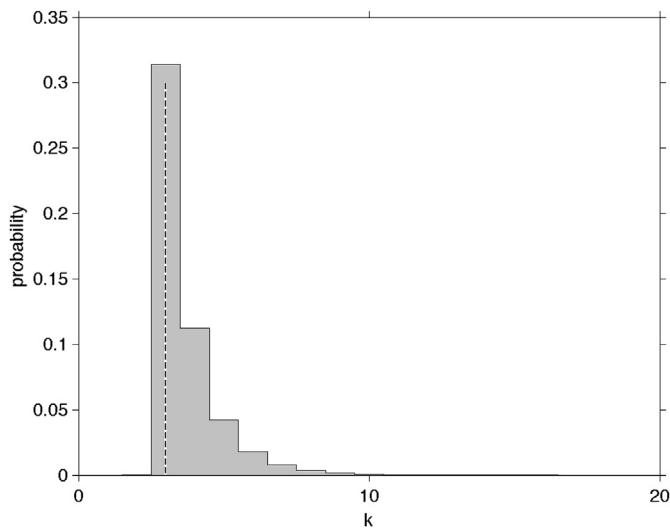


Fig. 5. Test 1. Results for the inversion of synthetic data. **Left panel:** posterior probability distribution for the depth of layer interfaces; **right panel:** posterior probability distribution for the electrical resistivity. For this test a uniform prior distribution was chosen, with resistivity values drawn in the interval $\log_{10}(\rho) \in [-0.5, 3.5]\log_{10}(\Omega\text{m})$. The dashed line outlines the “true” model used to produce the synthetic data used as input to the inversion.

the distribution of the number of layer interfaces. The most probable number of interfaces resulted to be 6 (i.e. 5 layers are needed to fit the dataset). This number is consistent with the synthetic model.

In the left panel of Fig. 8, the distribution of layer interfaces with depth is shown. The boundaries of the top-most conducting layer are resolved neatly, with the deeper interfaces being depicted correctly in the depth range of the synthetic model by the conductivity marginal. Even this test-case, the deeper interface are not characterised by probability peaks as neat as the near-surface ones.

The Fig. 8 presents the comparison of the retrieved resistivity

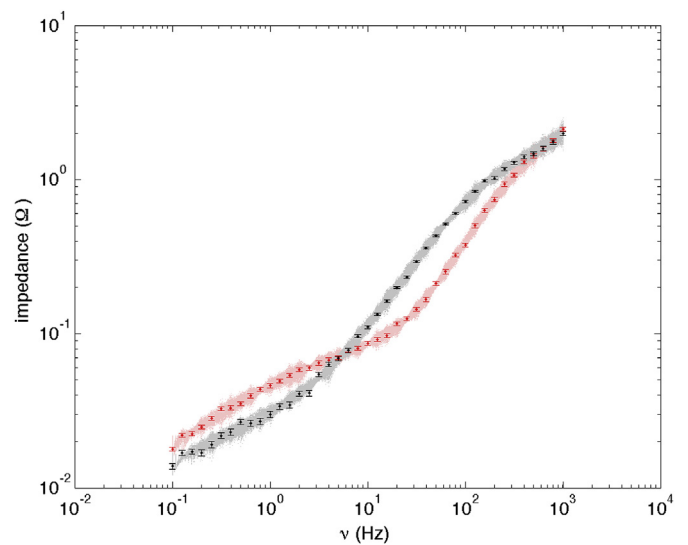


Fig. 6. Test 1. Data fit: the red and black dots marks the real and imaginary parts of the simulated impedances. The red and grey stripes marks the predicted impedances from the PPD sample. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

distribution with the true model. The synthetic model is marked as a dashed line. Overall the true model lies close to the region of high probability, even if, in the areas below the conductors, such interval is moderately broad. The lack of data sensitivity below a conductor is a well-known feature of the MT method (Chave and Jones, 2012), nonetheless the precise estimation of confidence intervals allows a better interpretation of the physical meaning of the PPD. Moreover the PPD shows that the resistivity is really well constrained even in the high-resistivity values in the topmost ≈ 1000 m of the model.

Note that Cerv et al. (2007), using deterministic algorithms, were not able to resolve the thin conductive layer below the depth of 2000 m (see

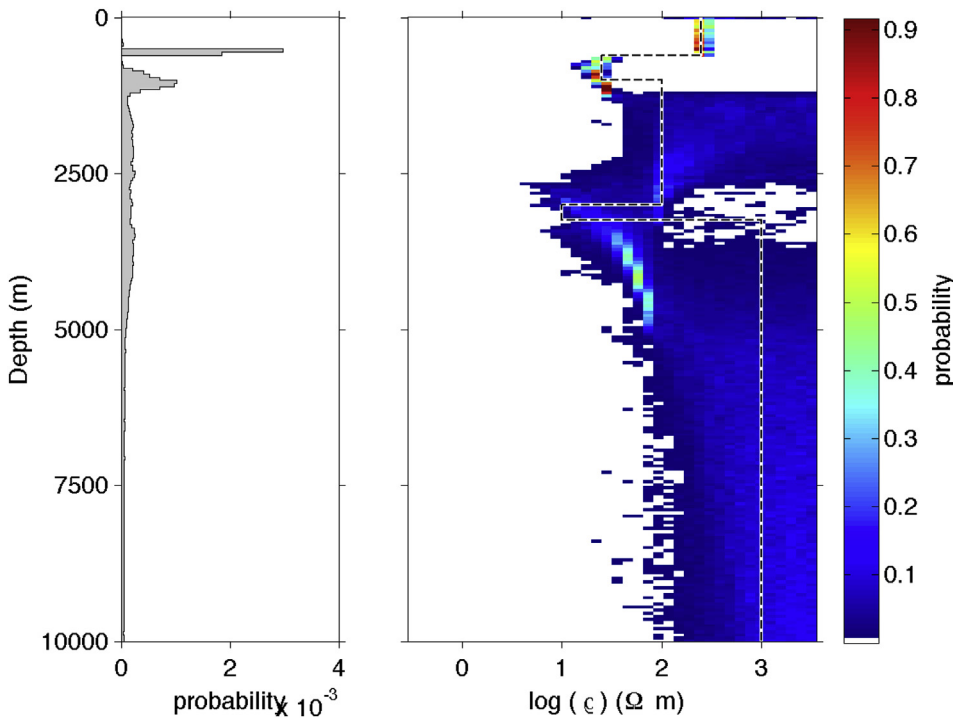


Fig. 7. Test 2. Posterior probability distribution of the number of interfaces in the sample. The dashed line marks the true one. Even in this case, where data support a more complex model with respect to Test 1, the algorithm parsimoniously sampled more frequently models with a limited number of interfaces.

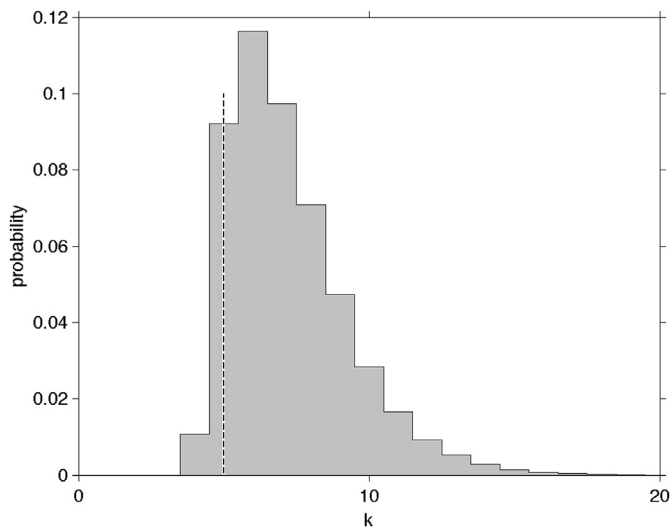


Fig. 8. Test 2. Results for the inversion of synthetic data. **Left panel:** posterior probability distribution for the depth of layer interfaces; **right panel:** posterior probability distribution for the electrical resistivity. For this test a uniform prior distribution was chosen, with resistivity values drawn in the interval $\log_{10}(\rho) \in [-0.5, 3.5]\log_{10}(\Omega\text{m})$. The dashed line outlines the “true” model used to produce the synthetic data used as input to the inversion.

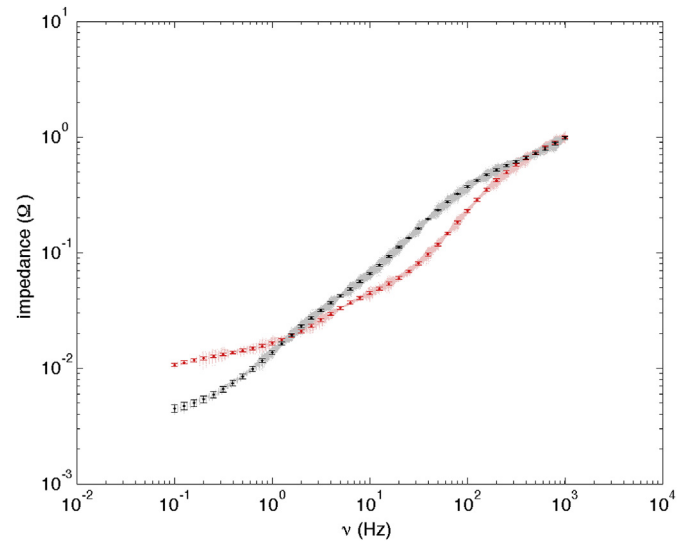


Fig. 9. Test 2. Data fit: the red and black dots marks the real and imaginary parts of the simulated impedances. The red and grey stripes marks the predicted impedances from the PPD sample. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Figs. 1 and 3 in Cerv et al. (2007)), where our algorithm suggests a discontinuity in resistivity that is distributed about the actual resistivity values of the synthetic model.

For this test the data fit is correct for both the real and the imaginary parts of the impedance tensor, as shown in Fig. 9.

3.4. Field measurements from Clare basin

We inverted the Clare Basin dataset for both polarisation using 47 Markov chains constituted of 8×10^5 iterations, discarding the first

2×10^5 models in the burn-in process. Due to the fact that, in this case, the algorithm was dealing with real data, we decided to use a less informative prior probability distribution for the resistivity, using as a prior a continuous uniform distribution that draws resistivity values in the interval $\log_{10}(\rho) \in [-2, 5]\log_{10}(\Omega\text{m})$.

In order to compare our results (Figs. 10 and 13) with the ones provided by a standard deterministic method, we inverted the Clare Basin data using a well tested Occam_1D algorithm (Constable et al., 1987). A 5% error floor of the MT impedance tensor was assumed during the inversion process, in both XY and YX polarisations.

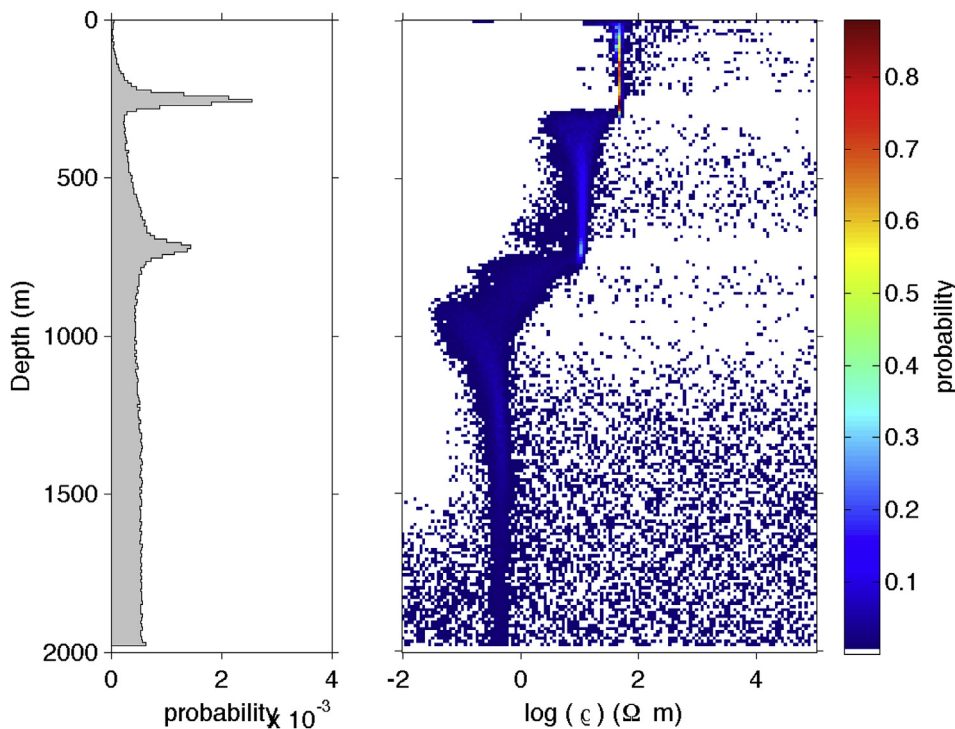


Fig. 10. Results for the inversion of field measurements, XY polarisation. **Left panel:** 1D marginal PPD the depth of layer interfaces; **right panel:** PPD for the electrical resistivity. For this test a uniform prior distribution was chosen, with resistivity values drawn in the interval $\log_{10}(\rho) \in [-2, 5]\log_{10}(\Omega m)$.

Results from the 1-D inversion show geoelectrical structures with electrical resistivity values between $0.3 \Omega m$ and $500 \Omega m$. Comparing the models obtained for XY and YX polarisation, both show similar geoelectrical patterns although some slight differences are observed at the contact between the Central Clare Group and the Gull Island Formation, and when characterising the electrical resistivity values of the Ross Sandstone Formation. The main geoelectrical features of the 1D model is the very low electrical resistivity values associated with the Shale Formation.

Results obtained using the algorithm developed in this paper are presented in Figs. 10–12 for XY polarisation and in Figs. 13–15 for YX polarisation. The most probable number of layers is in the interval [5; 7], for both polarisation, consistent with the number of layers observed in the log data (Fig. 1). The PPD for the interface depth (Figs. 10 and 13) presents sharp maxima at about 200 and 700 m, on both polarisations, with a clear maximum at about 150 m, on YX polarisation only. The position of the maxima for the interface depth does not match completely the one measured in the well logs, nonetheless both the XY and the YX modes clearly recognise the jump in resistivity associated with the Clare Shale formation.

Even more interesting is the comparison of the retrieved resistivity distributions with the models obtained using Occam's inversion. These results are shown in Fig. 16.

It is straightforward to point that the two resistivity models inferred by the Occam's inversion lie within the resistivity marginal distributions sampled by our algorithm. The Bayesian inversion produces nonetheless a further set of useful information about the model parametrisation. The marginal PPD for resistivity itself appears in both cases smoother and more regular than the correspondent Occam model. Moreover, the Occam model requires different anomalies in resistivity for the two polarisation to justify the data. In detail, the resistivity anomaly present in both polarisations at 100 m depth is not required and may be just a regularisation artefact due to the smoothness imposed to the model. In the region that connects the resistivity anomalies the Occam models enforce a smooth transition that lay in regions of the PPD that characterised by low-probability. Again, a Bayesian approach allows both discontinuities and smooth transitions in the probability, and this feature

allows to represent a broader sample of models.

Finally, we highlight two more advantages related to the use of our algorithm. First, we stress the further information returned by Bayesian inference in the form of precise estimations of credibility intervals (68% credibility intervals were estimated here, reported as solid black line in Fig. 16.). This is the main advantage of the suggested algorithm providing a crucial information to better understand how well the studied environment is seen by the geophysics methods. Second, we note that we can estimate resistivity-thickness trade-offs, not available using deterministic methods. Due to the possibility of collecting models from the PPD, it can be straightforward to compute realistic estimates of correlation between investigated parameters to point out the (likely) strong trade-off between resistivity and layer thickness. We do not explore parameters correlation as this is beyond the scope of this study. We refer to

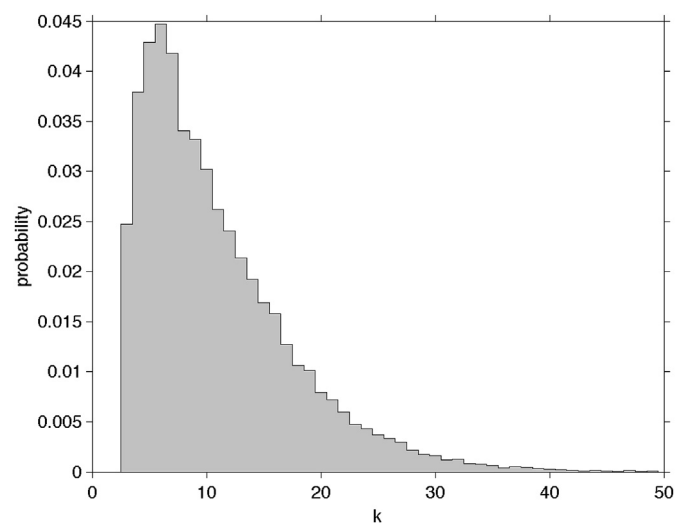


Fig. 11. Results for the inversion of field measurements, XY polarisation. Posterior probability distribution of the number of interfaces.

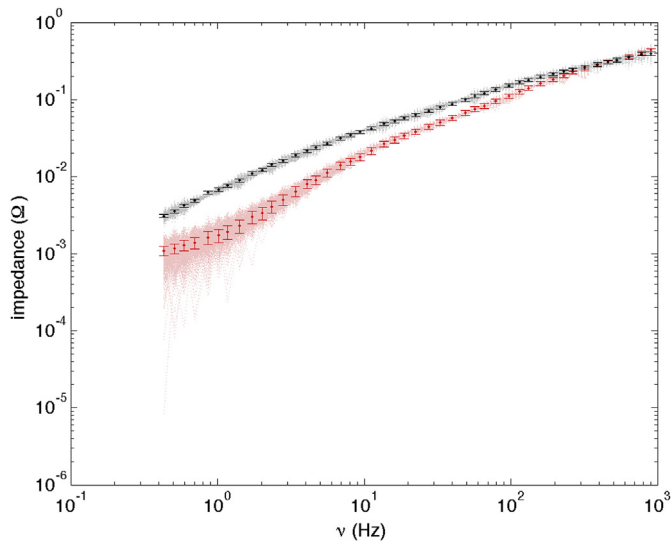


Fig. 12. Results for the inversion of field measurements, XY polarisation. Data fit: the red and black dots marks the real and imaginary parts of the simulated impedances. The red and grey stripes marks the predicted impedances from the PPD sample. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Piana Agostinetti et al. (Piana Agostinetti et al., 2015) for a discussion and example of how to investigate correlation using a RjMcMC algorithm.

Despite the named advantages, the method we propose here has the main drawback in the computational cost. In all the cases in which the forward algorithm can be solved efficiently the computation cost may be neglected, in all other cases a computational strategy must be found to work around the prohibitive computing time a single Markov chain would use to converge. From this perspective the MT method presents a great test case, as the 1D forward model has an analytic solution that can

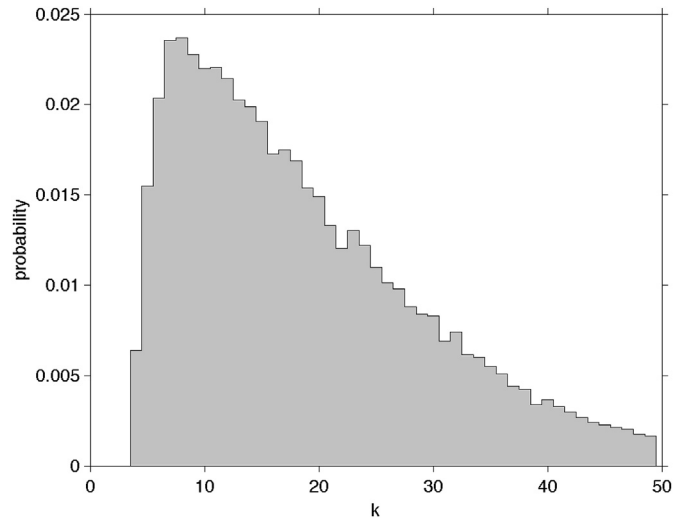


Fig. 14. Results for the inversion of field measurements, YX polarisation. Posterior probability distribution of the number of interfaces.

be computed extremely efficiently. Nonetheless we implemented a parallel version of our method that allow the joint exploration of different regions of the likelihood hyper-surface, potentially exploring a broader region of the model space.

A potential pitfall in our algorithm comes from the choice of sampling according to the priors, following the approach described in Mosegaard and Tarantola (1995). In this case, as detailed above, Eq. (6) reduces to Eq. (7). However, as we can see from Eq. (6), acceptance probability is higher when the proposal is similar to the target distribution (in our case, the PPD, the product between prior and likelihood). Hypotetically, acceptance probability for a candidate is equal to 1 if the proposal is equal to the PPD (but this means that we know the PPD, which is our target function and so it should be unknown). Clearly, the more the

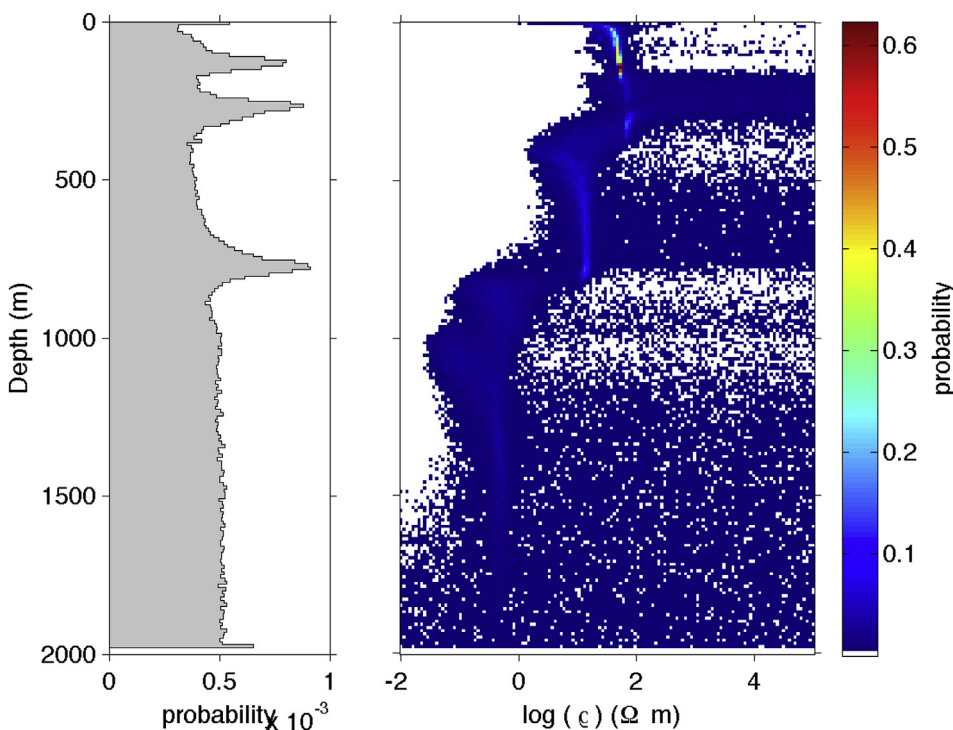


Fig. 13. Results for the inversion of field measurements, YX polarisation. **Left panel:** 1D marginal PPD for the depth of layer interfaces; **right panel:** PPD for the electrical resistivity distribution. For this test a uniform prior distribution was chosen, with resistivity values drawn in the interval $\log_{10}(\rho) \in [-2, 5] \log_{10}(\Omega m)$.

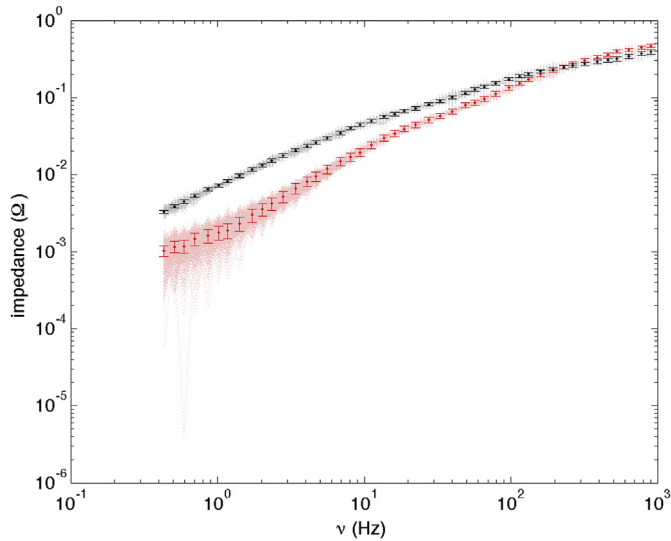


Fig. 15. Results for the inversion of field measurements, YX polarisation. Data fit: the red and black dots marks the real and imaginary parts of the simulated impedances. The red and grey stripes marks the predicated impedances from the PPD sample. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

proposal differs from the PPD, the less likely the candidate would be accepted. Going back to Mosegaard and Tarantola (1995)’s approach, where the proposal is the prior, this observation means that the more the PPD differs from the prior the less efficient would be the sampling. In other words, the more information about the investigated parameters are present in the measured data, the less efficient would be the Mosegaard and Tarantola (1995)’s approach in reconstructing the PPD. In our case,

we note that MT data are not really informative when challenged by low-conductivity values, so there are strong indication that sampling according to the priors should not strongly reduce the efficiency. Here, we have to clarify that, following the approach depicted by Mosegaard and Tarantola (1995), the value of resistivity related to a randomly-picked interface is perturbed according to a proposal distribution which asymptotically samples the prior probability distribution. This is different from proposing a candidate model replacing the value of one parameter and picking it randomly from its prior (which is much more similar to a simple Monte Carlo sampling where each model is un-correlated to the others and extracted randomly from the priors). Our algorithm asymptotically samples the prior, but itself in one iteration it perturbs the resistivity about the local value. However, from the birth/death move perspective, it has been shown by Dosso et al. (2014) that proposing parameters from the prior can be very efficient whereas the prior wight is narrow.

4. Conclusions

Recently trans-dimensional inversion has become a more popular tool to invert geophysical data-set. We developed a method to invert MT data in a trans-dimensional Bayesian framework. Previous studies shown how the Bayesian approach helps the interpretation, returning as result a model distribution rather than a single model. Moreover, the error of the model parameters are estimated directly by the posterior probability distribution, and are not estimated via the linearization of the model covariance matrix. In addition to this feature, the method we use implicitly justifies the use of a certain number of parameters to model the environment, justification that in linear inversions takes efforts and is based more on ad hoc selection of parameters than on data evidences. The developed algorithm has been implemented as a parallel computer code, and has been successfully tested against two synthetic models and a field dataset.

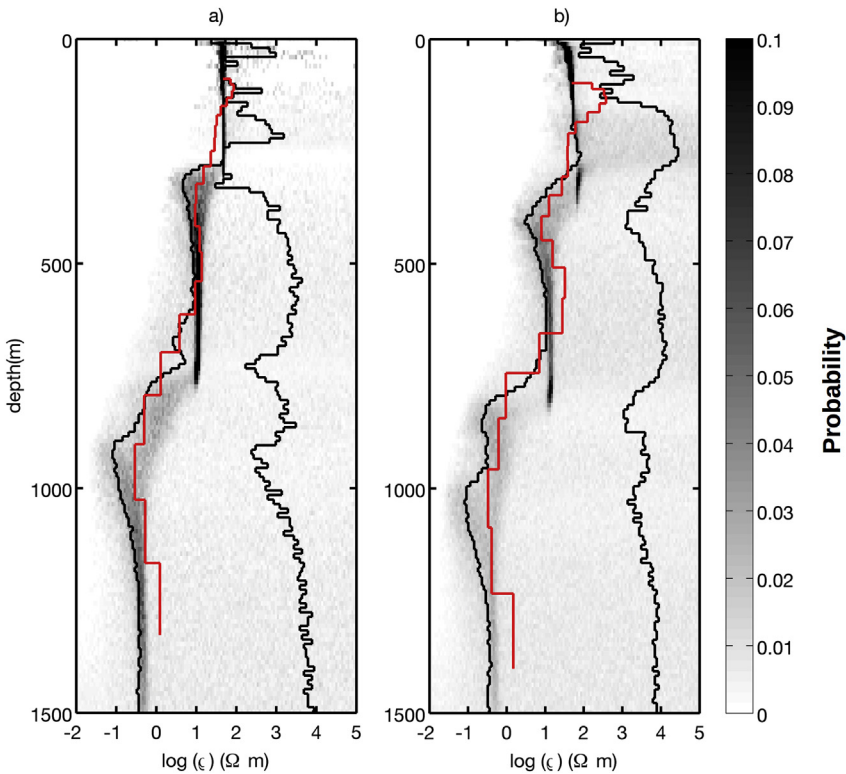


Fig. 16. Comparison of the inversion results obtained with the Occam's inversion (red solid line) and the bayesian inversion (background). The a) panel shows the XY polarisation, the b) panel shows the YX polarisation. The black solid line represents the boundaries of the 68% credibility intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Acknowledgements

NPA is deeply in debt with Daniele Melini for assistance on MPI codes. This research has been partially funded by Science Foundation Ireland projects 11/SIRG/E2174 and 14/IFB/2742. NPA research is conducted with the financial support of Austrian Science Fund (FWF), project number: M2218-N29. XO and JC would like to acknowledge the financial support from the IRECCSEM project funded by a Science Foundation of Ireland Investigator Project Award (SFI: 12/IP/1313). The authors would like to thank Petroleum Affairs Division of Ireland for providing data from Doonbeg borehole.

References

- Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., Rawlinson, N., 2012. Transdimensional inversion of receiver functions and surface wave dispersion. *J. Geophys. Res.: Solid Earth* 117.
- Cagniard, L., 1953. Basic theory of the magneto-telluric method of geophysical prospecting. *Geophysics* 18, 605–635.
- Caldwell, T.G., Bibby, H.M., Brown, C., 2004. The magnetotelluric phase tensor. *Geophys. J. Int.* 158, 457–469.
- Campanyà, J., Ledo, J., Queralt, P., Marcuello, A., Jones, A.G., 2014. A new methodology to estimate magnetotelluric (mt) tensor relationships: estimation of local transfer-functions by combining interstation transfer-functions (elicit). *Geophys. J. Int.* 198, 484–494.
- Cerv, V., Menvielle, M., Pek, J., 2007. Stochastic interpretation of magnetotelluric data, comparison of methods. *Ann. Geophys.* 50.
- Charvin, K., Gallagher, K., Hampson, G.L., Labourdette, R., 2009. A bayesian approach to inverse modelling of stratigraphy, part 1: Method. *Basin Res.* 21, 5–25.
- Chave, A.D., Jones, A.G., 2012. *The Magnetotelluric Method: Theory and Practice*. Cambridge University Press.
- Chave, A.D., Thomson, D.J., Ander, M.E., 1987. On the robust estimation of power spectra, coherences, and transfer functions. *J. Geophys. Res.: Solid Earth* 92, 633–648.
- Constable, S.C., Parker, R.L., Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics* 52, 289–300.
- Dettmer, J., Dosso, S.E., Holland, C.W., 2010. Trans-dimensional geoaoustic inversion. *J. Acoust. Soc. Am.* 128, 3393–3405.
- Dettmer, J., Molnar, S., Steininger, G., Dosso, S.E., Cassidy, J.F., 2012. Trans-dimensional inversion of microtremor array dispersion data with hierarchical autoregressive error models. *Geophys. J. Int.* 188, 719–734.
- Dosso, S.E., Wilmut, M.J., 2006. Data uncertainty estimation in matched-field geoaoustic inversion. *IEEE J. Ocean. Eng.* 31, 470–479.
- Dosso, S.E., Dettmer, J., Steininger, G., Holland, C.W., 2014. Efficient trans-dimensional bayesian inversion for geoaoustic profile estimation. *Inverse Probl.* 30, 114018.
- Egbert, G.D., 1997. Robust multiple-station magnetotelluric data processing. *Geophys. J. Int.* 130, 475–496.
- Egbert, G.D., Booker, J.R., 1986. Robust estimation of geomagnetic transfer functions. *Geophys. J. Int.* 87, 173–194.
- Farrelly, I., Loske, B., Neele, F., Holdstock, M., Assessment of the Potential for Geological Storage of Co₂ in Hypothetical Deep Saline Aquifers in the Vicinity of Moneypoint, Co. Clare, CCRP.
- Gallagher, K., Charvin, K., Nielsen, S., Sambridge, M., Stephenson, J., 2009. Markov chain Monte Carlo (mcmc) sampling methods to determine optimal models, model resolution and model choice for earth science problems. *Mar. Petrol. Geol.* 26, 525–535.
- Gamble, T., Goubau, W.M., Clarke, J., 1979a. Magnetotellurics with a remote magnetic reference. *Geophysics* 44, 53–68.
- Gamble, T., Goubau, W., Clarke, J., 1979b. Error analysis for remote reference magnetotellurics. *Geophysics* 44, 959–968.
- Geyer, C.J., 1991. Markov Chain Monte Carlo Maximum Likelihood.
- Geyer, C.J., Møller, J., 1994. Simulation procedures and likelihood inference for spatial point processes. *Scand. J. Stat.* 359–373.
- Grandis, H., Menvielle, M., Roussignol, M., 1999. Bayesian inversion with markov chains. the magnetotelluric one-dimensional case. *Geophys. J. Int.* 138, 757–768.
- Grandis, H., Menvielle, M., Roussignol, M., 2013. Thin-sheet inversion modeling of geomagnetic deep sounding data using mcmc algorithm. *Int. J. Geophys.* 2013.
- Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and bayesian model determination. *Biometrika* 82, 711–732.
- Guo, R., Dosso, S.E., Liu, J., Dettmer, J., Tong, X., 2011. Non-linearity in bayesian 1-d magnetotelluric inversion. *Geophys. J. Int.* 185, 663–675.
- Guo, R., Dosso, S.E., Liu, J., Liu, Z., Tong, X., 2014. Frequency-and spatial-correlated noise on layered magnetotelluric inversion. *Geophys. J. Int.* 199, 1205–1213.
- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 97–109.
- Hopcroft, P.O., Gallagher, K., Pain, C.C., 2007. Inference of past climate from borehole temperature data using bayesian reversible jump Markov chain Monte Carlo. *Geophys. J. Int.* 171, 1430–1439.
- Hopcroft, P.O., Gallagher, K., Pain, C.C., 2009. A bayesian partition modelling approach to resolve spatial variability in climate records from borehole temperature inversion. *Geophys. J. Int.* 178, 651–666.
- Jasra, A., Stephens, D.A., Gallagher, K., Holmes, C.C., 2006. Bayesian mixture modelling in geochronology via Markov chain Monte Carlo. *Math. Geol.* 38, 269–300.
- Kaufman, A.A., 1981. The magnetotelluric sounding method. *Meth. Geochem. Geophys.* 15.
- Kelbert, A., Meqbel, N., Egbert, G.D., Tandon, K., 2014. ModEM: a modular system for inversion of electromagnetic geophysical data. *Comput. Geosci.* 66, 40–53.
- Malinverno, A., 2002. Parsimonious bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophys. J. Int.* 151, 675–688.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21, 1087–1092.
- Minsley, B.J., 2011. A trans-dimensional bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data. *Geophys. J. Int.* 187, 252–272.
- Mosegaard, K., 2006. Monte Carlo Analysis of Inverse Problems. Ph.D. thesis.
- Mosegaard, K., Sambridge, M., 2002. Monte Carlo analysis of inverse problems. *Inverse Probl.* 18, R29.
- Mosegaard, K., Tarrantola, A., 1995. Monte Carlo sampling of solutions to inverse problems. *J. Geophys. Res.: Solid Earth* 100, 12431–12447 (1978–2012).
- Mosegaard, K., Singh, S., Snyder, D., Wagner, H., 1997. Monte Carlo analysis of seismic reflections from moho and the w reflector. *J. Geophys. Res.: Solid Earth* 102, 2969–2981 (1978–2012).
- Piana Agostinetti, N., Malinverno, A., 2010. Receiver function inversion by trans-dimensional Monte Carlo sampling. *Geophys. J. Int.* 181, 858–872.
- Piana Agostinetti, N., Giacomuzzi, G., Malinverno, A., 2015. Local 3D earthquake tomography by trans-dimensional Monte Carlo sampling. *Geophys. J. Int.* 201, 1598–1617. <https://doi.org/10.1093/gji/ggv084>.
- Ray, A., Key, K., 2012. Bayesian inversion of marine csem data with a trans-dimensional self parametrizing algorithm. *Geophys. J. Int.* 191, 1135–1151.
- Ray, A., Key, K., Bodin, T., Myer, D., Constable, S., 2014. Bayesian inversion of marine csem data from the scarborough gas field using a transdimensional 2-d parametrization. *Geophys. J. Int.* 199, 1847–1860.
- Rider, M., The namurian of west county clare, in: *Proceedings of the Royal Irish Academy. Section B: Biological, Geological, and Chemical Science*, JSTOR, pp. 125–142.
- Rodi, W., Mackie, R.L., 2001. Nonlinear conjugate gradients algorithm for 2-d magnetotelluric inversion. *Geophysics* 66, 174–187.
- Rudolph, M.L., Lekić, V., Lithgow-Bertelloni, C., 2015. Viscosity jump in earth's mid-mantle. *Science* 350, 1349–1352.
- Sambridge, M., 2014. A Parallel Tempering algorithm for probabilistic sampling and multimodal optimization. *Geophys. J. Int.* 196, 357–374. <https://doi.org/10.1093/gji/ggt342>.
- Sambridge, M., Gallagher, K., Jackson, A., Rickwood, P., 2006. Trans-dimensional inverse problems, model comparison and the evidence. *Geophys. J. Int.* 167, 528–542.
- Sambridge, M., Bodin, T., Gallagher, K., Tkalčić, H., 2013. Transdimensional inference in the geosciences. *Phil. Trans. R. Soc. A* 371, 20110547.
- Smirnov, M.Y., 2003. Magnetotelluric data processing with a robust statistical procedure having a high breakdown point. *Geophys. J. Int.* 152, 1–7.
- Steininger, G., Dettmer, J., Dosso, S.E., Holland, C.W., 2013. Trans-dimensional joint inversion of seabed scattering and reflection data. *J. Acoust. Soc. Am.* 133, 1347–1357.
- Stephenson, J., Gallagher, K., Holmes, C., 2006. Low temperature thermochronology and strategies for multiple samples: 2: partition modelling for 2d/3d distributions with discontinuities. *Earth Planet Sci. Lett.* 241, 557–570.
- Tikhonov, A., On determining electrical characteristics of the deep layers of the earths crust, in: *Sov. Math. Dokl.*, volume vol. 2, pp. 295–297.