Geophysical Journal International

Geophys. J. Int. (2017) **211,** 1341–1353
Advance Access publication 2017 August 22
GJI Marine geosciences and applied geophysics

doi: 10.1093/gji/ggx355

Time-domain induced polarization – an analysis of Cole–Cole parameter resolution and correlation using Markov Chain Monte Carlo inversion

Line Meldgaard Madsen, Gianluca Fiandaca, Esben Auken and Anders Vest Christiansen

HydroGeophysics Group, Department of Geoscience, Aarhus University, C.F. Møllers Allé 4, *DK*-8000 *Aarhus C, Denmark.*
E-mail: linemeldgaard@geo.au.dk

Accepted 2017 August 18. Received 2017 June 20; in original form 2016 December 8

SUMMARY

The application of time-domain induced polarization (TDIP) is increasing with advances in acquisition techniques, data processing and spectral inversion schemes. An inversion of TDIP data for the spectral Cole–Cole parameters is a non-linear problem, but by applying a 1-D Markov Chain Monte Carlo (MCMC) inversion algorithm, a full non-linear uncertainty analysis of the parameters and the parameter correlations can be accessed. This is essential to understand to what degree the spectral Cole–Cole parameters can be resolved from TDIP data. MCMC inversions of synthetic TDIP data, which show bell-shaped probability distributions with a single maximum, show that the Cole–Cole parameters can be resolved from TDIP data if an acquisition range above two decades in time is applied. Linear correlations between the Cole–Cole parameters are observed and by decreasing the acquisitions ranges, the correlations increase and become non-linear. It is further investigated how waveform and parameter values influence the resolution of the Cole–Cole parameters. A limiting factor is the value of the frequency exponent, C. As C decreases, the resolution of all the Cole–Cole parameters decreases and the results become increasingly non-linear. While the values of the time constant, τ, must be in the acquisition range to resolve the parameters well, the choice between a 50 per cent and a 100 per cent duty cycle for the current injection does not have an influence on the parameter resolution. The limits of resolution and linearity are also studied in a comparison between the MCMC and a linearized gradient-based inversion approach. The two methods are consistent for resolved models, but the linearized approach tends to underestimate the uncertainties for poorly resolved parameters due to the corresponding non-linear features. Finally, an MCMC inversion of 1-D field data verifies that spectral Cole–Cole parameters can also be resolved from TD field measurements.

Key words: Electrical properties; Inverse theory; Probability distributions.

1 INTRODUCTION

Time-domain (TD) induced polarization (IP) is a geophysical method that is applied increasingly in environmental investigations and mineral exploration. Today, field applications include landfill characterization (Gazoty *et al.* 2012a, 2013; Wemegah *et al.* 2014), lithology discrimination (Gazoty *et al.* 2012b; Chongo *et al.* 2015; Johansson *et al.* 2015, 2017; Maurya *et al.* 2016), prospecting of quartz-sulphide-gold mineralizations (Tarasov & Gurin 2016), time-lapse monitoring of CO_2 injections (Fiandaca *et al.* 2015; Doetsch *et al.* 2015a) and investigation of permafrost freezing processes (Doetsch *et al.* 2015b). Recent laboratory applications include studies of the relationship between IP and the hydraulic properties of the subsurface (e.g. Titov *et al.* 2010).

The TDIP phenomenon is manifested by a transient potential decay following termination of an applied electric direct current (DC) conducting through the subsurface. Previously, only the integral chargeability of the transient IP decay curve was considered when inverting and interpreting TDIP data (Oldenburg & Li 1994). This method, however, disregards the spectral information given by the shape of the IP decay. More recently, modelling methods have been introduced where the entire IP decay is used to resolve, for example, the Cole–Cole parameters (Cole & Cole 1941; Pelton *et al.* 1978), which parametrize the spectral information. One of these methods divide the IP decay into time gates and invert each time gate independently with DC algorithms (Yuval & Oldenburg 1997; Hönig & Tezkan 2007). Such method works well for a wide range of parameter contrasts (Hördt *et al.* 2006), however, errors may arise

from neglecting the actual transmitter waveform. Fiandaca *et al.* (2012) presented a 1-D laterally constrained inversion algorithm, which takes into account the full IP potential decay as well as the transmitter waveform and the receiver transfer function when inverting for the spectral content of the IP signal. The algorithm has been extended to handle 2-D spectral inversions as well (Fiandaca *et al.* 2013).

Access to full-waveform recordings is a necessity in order to retrieve spectral information from TDIP data. This is today possible given the recent improvements in acquisition equipment. With the full-waveform recording, it is possible to facilitate comprehensive processing of the recorded IP signals. This is especially important in order to access the IP data at early times after current alternation. Olsson *et al.* (2016) presented a method to remove harmonic noise from IP data, thus improving the signal-to-noise ratio at early times significantly, moving the first usable time gate up to a few milliseconds. They also introduced a processing scheme with spike-and-drift removal. The combined effect of these processing schemes doubles the usable range of the TDIP signal to almost four decades in time (Olsson *et al.* 2016). Furthermore, Olsson *et al.* (2015) suggested the use of a 100 per cent duty cycle waveform for data acquisition. Here, the injected current is switched directly from positive to negative and the data measurements are performed in the current on time. This improves the signal-to-noise ratio further due to the superposition of the IP decay signal and the ongoing current injection.

It is a non-linear problem to retrieve the spectral Cole–Cole parameters from TDIP data and many different deterministic inversion methods can be applied (e.g. Yuval & Oldenburg 1997; Kemna 2000; Xiang *et al.* 2003). Often, the problem is linearized and solved using gradient-based methods. The gradient-based spectral inversion scheme presented in Fiandaca *et al.* (2012) applies a traditional Taylor expansion for linearization and iterates to minimize a given objective function. A linearized uncertainty analysis of the spectral Cole–Cole parameters can then be computed from the posterior covariance matrix of the forward problem. However, it is in general understood that the linearized methods tend to underestimate the parameter uncertainties (Sambridge & Mosegaard 2002).

To investigate the linearity of the spectral TDIP problem and to evaluate how well a linearized inversion approach preforms, a full non-linear inversion and uncertainty analysis must be carried out. By exploring the space of models, which are consistent with an observed data set, instead of just finding a single best model as given by gradient-based methods, it is possible to make a complete interpretation of the problem. This can be done by applying a probabilistic formulation.

The probabilistic formulation of non-linear inversion problems is well known within geophysical methods, where the aim is to compute probability distributions for data and model parameters (Tarantola & Valette 1982). Ghorbani *et al.* (2007) applied a Bayesian model to invert synthetic TD and frequency-domain (FD) IP data and used numerical integration techniques to obtain marginal posterior probability distribution of each Cole–Cole parameter. The Markov Chain Monte Carlo (MCMC) method is another way to compute probability densities and it is more frequently applied in geophysical inversions today (Mosegaard & Tarantola 1995; Sambridge & Mosegaard 2002). Chen *et al.* (2008) compared the MCMC inversion method of synthetic and laboratory FDIP data to a Gauss–Newton deterministic method and points out the advances of the MCMC approach to quantify uncertainties from probability distributions.

We have developed a 1-D MCMC algorithm to invert TDIP data, which uses the forward algorithm described by Fiandaca *et al.* (2012) taking into account the full IP decay. We use the advantages of the MCMC method to study non-linearized parameter uncertainties and parameter correlations in order to understand to what degree the spectral Cole–Cole parameters can be retrieved from TDIP data. The following work investigates the influence of different factors on the model sensitivity: current duty cycle, acquisition range and parameter values. Several examples with both synthetic data and field data are investigated, and the results are also compared to Ghorbani *et al.* (2007) findings. Additionally, the MCMC inversion results are compared to results gained from a gradient-based linearized inversion approach.

2 INVERSION METHODOLOGY

2.1 Data space and noise model

The data space, $\boldsymbol{d}_{\mathbf{obs}}$, of the MCMC analysis of this study consists of apparent resistivity and full-decay chargeability values, which are log-transformed to enhance the linearity of the forward mapping:

$$\boldsymbol{d}_{\mathbf{obs}} = \left\{\log\left(\rho_i\right), \log(M_{i,j})\right\};$$

$$i = 1 : N_{\text{quads}};\ j = 1 : N_{\text{gates}}, \tag{1}$$

where the apparent resistivity, ρ_i, is given for each quadrupole measurement, i, and the chargeability, $M_{i,j}$, is computed in each time gate, j, of the TDIP signal. For layered models, the synthetic data are generated from a vertical sounding with 20 quadrupoles, with $|\text{AB}| = 7.5 - 500$ m and $|\text{MN}| = 2.5 - 65$ m (Table 1), which limits the geometric factor to 3000 m. For homogenous half-spaces, a single quadrupole with electrode spacing $|\text{AB}| = 37.5$ m and $|\text{MN}| = 2.5$ m is used (quadrupole 7 in Table 1). Based on experience, each IP decay is divided into seven log-distributed time gates per time decade. If fewer gates are used, parts of the signal will not be modelled, and adding more gates does not change the uncertainty, because no more information is retrieved from the signal.

Both 50 per cent and the 100 per cent duty cycle waveform have been used together with the noise model described in Olsson *et al.* (2015). This noise model models a uniform noise and a random noise that depends on the signal level. A total standard deviation, $\text{STD}_{\text{total}}$, is computed for each gate, j, by summing the quadrature of the uniform term, STD_{uni}, and a voltage-dependent term, STD_{vth}:

$$\text{STD}^2_{\text{total}}(j) = \text{STD}^2_{\text{uni}}(j) + \text{STD}^2_{\text{vth}}(j), \tag{2}$$

where, STD_{vth}, is controlled by a voltage noise threshold V_{TH} and is given as

$$\text{STD}_{\text{vth}}(j) = \frac{V_{\text{TH}}}{V_{\text{IP}}(j)} \cdot \frac{\sqrt{D_{\text{norm}}}}{\sqrt{D\ (j)}} \cdot \frac{1}{\sqrt{N_{\text{stacks}}}}, \tag{3}$$

where $V_{\text{IP}}(j)$, $D(j)$, and N_{stacks} represent the voltage level, the integration time and the stack size of the jth gate. D_{norm} is the nominal integration time of the voltage noise threshold V_{TH}. For all the tests in this study, $\text{STD}_{\text{uni}} = 5$ per cent and $\text{STD}_{\text{uni}} = 2$ per cent have been used for IP and DC data, respectively, with $V_{\text{TH}} = 0.1$ mV (for $D_{\text{norm}} = 0.01$ s) and $N_{\text{stacks}} = 3$, which resemble the noise expected in the field when neglecting background drift as justified by Gazoty *et al.* (2013). Fig. 1 shows the standard deviations derived from the noise model on two representative IP decays.

Table 1. Electrode spacing of the 20 quadrupoles used for generation of synthetic vertical soundings.

Quadrupole	1	2	3	4	5	6	7	8	9	10
|AB| (m)	7.5	12.5	17.5	22.5	27.5	32.5	37.5	42.5	47.5	52.5
|MN| (m)	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
Quadrupole	11	12	13	14	15	16	17	18	19	20
|AB| (m)	57.5	72.5	92.5	117.5	147.5	192.5	240	305	390	500
|MN| (m)	2.5	2.5	2.5	15	15	15	15	65	65	65

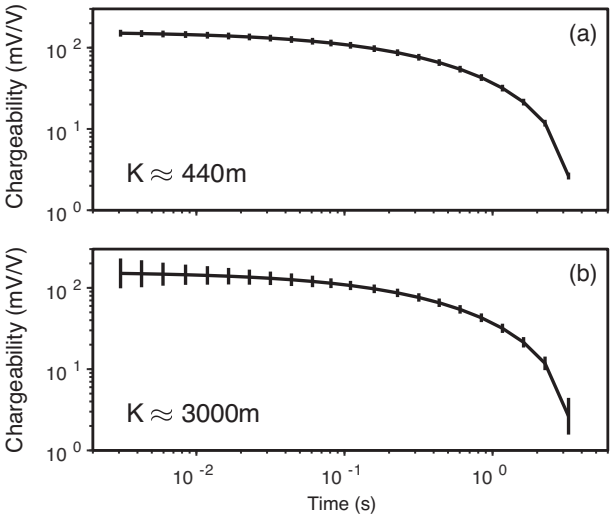

Figure 1. Forward response and data uncertainty (two standard deviations) of a Cole–Cole model with the parameters: $\rho = 100$ Ωm, $m_0 = 200$ mV, $\tau = 1$ s and $C = 0.6$. The data uncertainty is shown for two values of the geometrical factor K, that is, (a) $K = 440$ m and (b) $K = 3000$ m, which corresponds to quadrupoles 7 and 20 in Table 1, respectively.

2.2 Cole–Cole parametrization

The spectral content of the IP data is parametrized in terms of the Cole–Cole model (Cole & Cole 1941; Pelton *et al.* 1978). Despite being phenomenological, the model is often applied to field and laboratory data measured in both TD (e.g. Yuval & Oldenburg 1997; Hönig & Tezkan 2007; Fiandaca *et al.* 2012) and FD (e.g. Yoshioka & Zhdanov 2005; Loke *et al.* 2006). The complex resistivity is given as

$$\zeta(\omega) = \rho\left[1 - m_0\left(1 - \frac{1}{1 + (i\omega\tau)^C}\right)\right], \tag{4}$$

where ρ is the DC resistivity, m_0 is the (model space) chargeability as defined by Seigel (1959), τ is the relaxation time (or time constant), C is the frequency exponent, ω is the angular frequency and i is the imaginary unit. The 1-D model space can thus be defined as

$$\boldsymbol{m} = \left\{\log(\rho_i), \log(m_{0i}), \log(\tau_i), \log(C_i), \log\left(\mathrm{thk}_j\right)\right\};$$

$$i = 1 : N_{\text{layers}};\ j = 1 : N_{\text{layers}} - 1, \tag{5}$$

where thk represents the layer thickness.

The chargeability, which is given in mV V^{-1}, describes the magnitude of the polarization of the subsurface and thus depends on the quantity of polarizable elements (Pelton *et al.* 1978). In this way, knowledge of the chargeability helps to distinguish between lithotypes, which otherwise have identical resistivity responses. The dimensionless frequency exponent depends on the size distribution of the polarizable elements with small values expected for inhomogeneous grain sizes (Luo & Zhang 1998). The range of the frequency exponent is between 0.1 and 0.6 (e.g. Pelton *et al.* 1978; Luo & Zhang 1998). The relaxation time is strongly controlled by pore sizes and short relaxation times can be expected in materials with small pore throats (e.g. Binley *et al.* 2005; Titov *et al.* 2010; Revil *et al.* 2014). The relaxation time varies over a large interval with typical values in the range 10^{-3}–10^{3} (e.g. Revil *et al.* 2014).

Due to the logarithmic transform applied on data and model parameters (eqs 1 and 5), the uncertainties on the model parameters are given as standard deviation factors (STDFs), where the 68 per cent confidence interval for the parameter p lies between

$$\frac{p}{\text{STDF}} < p < p \cdot \text{STDF}. \tag{6}$$

A perfect resolution will give a STDF = 1. Using the terminology of Auken *et al.* (2005), a STDF < 1.2 is a well-resolved parameter, 1.2 < STDF < 1.5 is a moderately resolved parameter, 1.5 < STDF < 2 is a poorly resolved parameter and an STDF > 2 is an unresolved parameter.

2.3 Forward modelling

Given a layered model of the subsurface, the TDIP forward response is computed using the algorithm described by Fiandaca *et al.* (2012). The algorithm computes the TD forward response via a Hankel transform of the FD response. Once a complex resistivity, σ^*, is defined in each layer in terms of the Cole–Cole parameters, it is possible to compute the system transfer function, $K^*(\sigma^*(\omega, \boldsymbol{m}), q)$, were q is an arbitrary quadrupole. The TD step response can then be computed from the Fourier transform of the kernel, $K^*(\sigma^*(\omega, \boldsymbol{m}), q)/i\omega$, i being the imaginary unit. The transform is implemented in terms of the Hankel transform described by Johansen & Sørensen (1979). Fiandaca *et al.* (2012) also address the issue of modelling the effects of the transmitted current waveform, stacking of measurements and the effect of low-pass filters in the receiver instrument, which are all accounted for in the algorithm. However, the code does not take into account inductive effects from the ground nor from inductance among cables.

2.4 MCMC inversion

The probabilistic formulation of inversion problems is widely treated in connection with geophysical problems (e.g. Tarantola & Valette 1982; Mosegaard & Tarantola 1995; Malinverno 2002; Tarantola 2005). The posterior probability distribution, P_{post}, of the model $\boldsymbol{m}$ can be computed as (Tarantola 2005)

$$P_{\text{post}}(\boldsymbol{m}) = k\, P_{\text{prior}}(\boldsymbol{m})\ P_{\text{like}}(\boldsymbol{d}_{\text{obs}}|\boldsymbol{m}), \tag{7}$$

where $P_{\text{prior}}(\boldsymbol{m})$ is the prior probability distribution given by the prior information and $P_{\text{like}}(\boldsymbol{m})$ is the likelihood function describing the degree of fit between the observed data, $\boldsymbol{d}_{\text{obs}}$, and the forward

response, g($\boldsymbol{m}$). k is a normalization constant. The objective of the inversion is to describe the posterior probability distribution. With a MCMC inversion scheme, it is possible to sample the posterior probability distribution without knowing the exact prior probability distribution.

We apply a Metropolis–Hastings sampling algorithm (Metropolis *et al.* 1953; Hastings 1970), where a random walk samples models from the model space based on likelihood. The algorithm takes two steps. (1) A model is proposed. (2) The model is either accepted or rejected based on the likelihood of the model compared to the likelihood of the last accepted model. In this way, the model $\boldsymbol{m}_i$ only depends on $\boldsymbol{m}_{i-1}$ and none of the previous accepted models. The steps are repeated 5000–5000 000 times depending on the complexity of the model and a sequence of models is sampled (a Markov Chain of models). As models with high probability will be sampled more frequently, this chain will converge toward the posterior probability distribution and thus map the probability density for each Cole–Cole parameter.

The implementation of the 1-D TDIP MCMC inversion algorithm is described in the Appendix.

3 RESULTS

In the following, the MCMC inversion results of synthetic and field data are shown. All the synthetic tests have been carried out with both 50 per cent and 100 per cent duty cycles, always with equivalent results. For this reason, only the results obtained with the 100 per cent duty cycle are explicitly presented.

3.1 Effect of acquisition range

We show here how crucial the acquisition range is for resolving the spectral Cole–Cole parameters from TDIP data. A homogenous half-space model is inverted with the MCMC algorithm using different acquisition ranges. The model is given with the following parameters: $\rho = 100$ Ωm, $m_0 = 200$ mV V^{-1}, $\tau = 1$ s and $C = 0.6$.

Fig. 2 shows the resulting marginal posterior probability distributions of the four Cole–Cole parameters for three different acquisition ranges. In Fig. 2(a), an acquisition range of one decade running from 100 to 1000 ms (seven time gates) is applied. This acquisition range does not resolve the spectral Cole–Cole parameters. The small amount of data points results in a non-uniqueness, where the marginal distribution of ρ has more than one maximum. Furthermore, the distribution of m_0 and τ are strongly skewed and the correlations between the parameters are found to be strongly non-linear, which also gives rise to the poor resolution where STDF > 1.5 for the IP parameters. The STDFs are all listed together with the distributions in the figure.

By increasing the acquisition range to two decades (10–1000 ms, 14 time gates), the parameters become better resolved (Fig. 2b). However, the marginal posterior probability distributions are still wide and slightly skewed and the IP parameters are moderately resolved or unresolved.

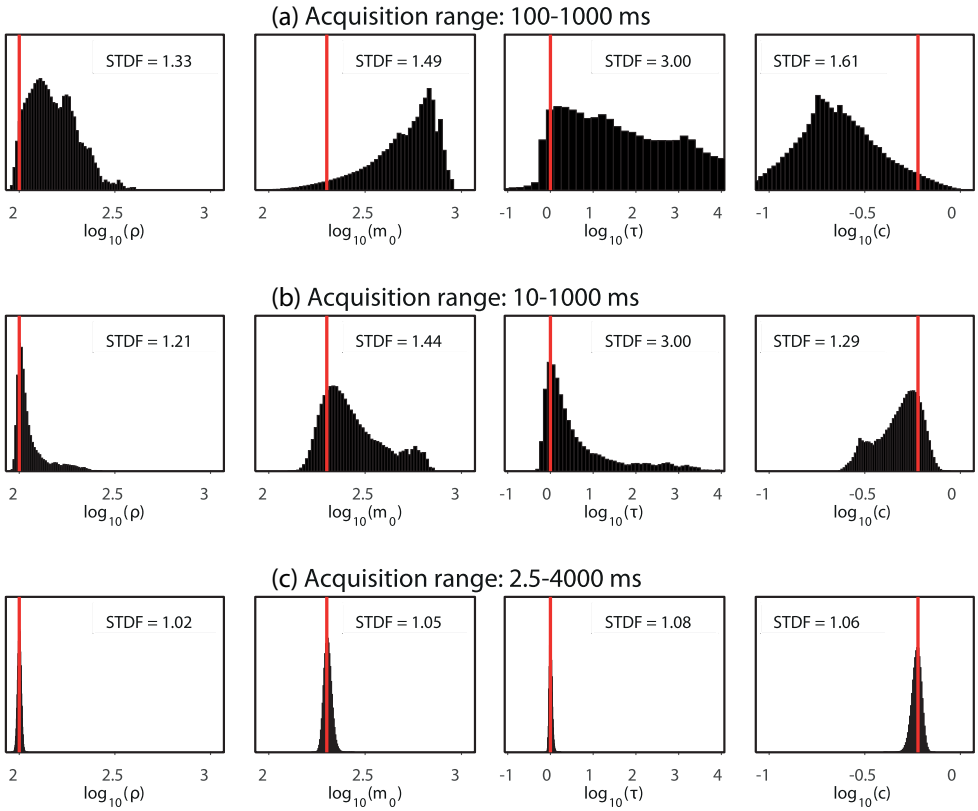

Figure 2. Marginal posterior probability distributions for the four Cole–Cole parameters determined from inversion with an MCMC inversion algorithm. Three different acquisition ranges have been applied for computation of synthetic data: (a) 100–1000 ms, (b) 10–1000 ms and (c) 2.5–4000 ms. The decays have been divided into seven time gates per decade. The true model values ($\rho = 100$ Ωm, $m_0 = 200$ mV V^{-1}, $\tau = 1$ s and $C = 0.6$) are shown in red. The standard deviation factors (STDFs) of each marginal distribution are shown together with each distribution. If a distribution has not converged, the STDF is put to 3 in the figure. Note that the plots of τ are scaled wider than the other Cole–Cole parameters.

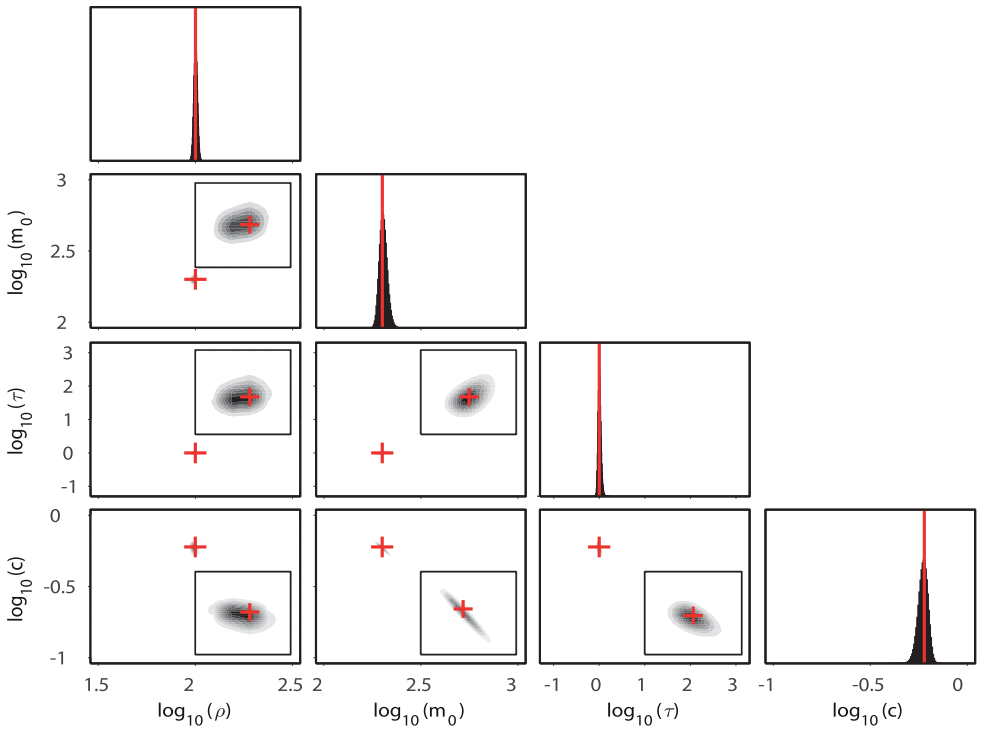

Figure 3. MCMC inversion results for the model with the Cole–Cole parameters: $\rho = 100$ Ωm, $m_0 = 200$ mV V^{-1}, $\tau = 1$ s and $C = 0.6$. The diagonal plots show the marginal posterior probability distributions of the four Cole–Cole parameters and the off-diagonal plots show the cross-correlations of the parameters (with zoom-in in the insets). The red colour indicates the true model value.

Acquisition ranges of three decades (1–1000 ms, 21 time gates), four decades (1–10 000 ms, 28 time gates) and combinations in between have been examined as well. We see that as the acquisition range increases, the resolution of the Cole–Cole parameters increases as well. The skewness almost disappears and the marginal posterior probability distributions become bell-shaped, when the acquisition range gets above three decades.

In the following, the acquisition range 2.5–4000 ms (22 time gates) is applied. A delay of 2.5 ms is sufficient for retrieving good IP data in the field when harmonic de-noising is performed (Olsson *et al.* 2016), and by measuring up to 4000 ms a bell-shaped probability distribution is obtained. With the model shown in Fig. 2(c), all the Cole–Cole parameters are well resolved (STDF$<$ 1.2) for this acquisition range.

3.2 Parameter values, resolution and correlations

Fig. 3 shows MCMC inversion results of the model described previously ($\rho = 100$ Ωm, $m_0 = 200$ mV V^{-1}, $\tau = 1$ s and $C = 0.6$), which was inverted with the 2.5–4000 ms acquisition range. Here, the diagonal plots show the histograms of the marginal posterior probability distributions of the Cole–Cole parameters and the off-diagonal plots show the cross-correlations between the parameters.

The distributions are all approximately bell-shaped with a maximum at the true model value. The cross-plots show that the DC resistivity is roughly uncorrelated to the IP parameters. There is a strong negative correlation between m_0 and C. Minor correlations are also visible between m_0 and τ and between τ and C. The observed correlations are all linear in log-space. However, if the resolution of the model is decreased, e.g. if the signal-to-noise level is decreased, the linearity slowly breaks down.

The negative $m_0 - C$ correlation can be explained by studying the influence of the IP parameter values on the shape of the TD and FD IP forward response. Fig. 4(a) shows that increasing/decreasing m_0 shifts the TD response up/down relative to the chargeability. The same is the case when C is increasing/decreasing. In FD, the maximum phase is increased/decreased when m_0 or C is increased/decreased (Fig. 4b).

In Fig. 5, the value of the frequency exponent, C, is decreased from $C = 0.6$ to 0.3. The FD responses in Fig 4(a) show that when C is decreased, the phase peak will become less pronounce and it will therefore be more difficult to determine the peak frequency and thereby also more difficult to resolve the value of τ. In Fig. 5, we see that the reduction in C decreases the resolution of all the IP parameters. A secondary maximum appears in ρ, m_0, and C. A secondary maximum can 'trap' a linearized inversion. However, in the linearized inversion approach applied for this study (see Appendix), boundary values are given to all the parameters and for C the minimum value is 0.1. In this way, the iterative procedure of the linearized inversion does not reach this secondary maximum. Secondary maxima have also been observed in other poorly resolved models in association with very low C values.

The strong linear m_0–C correlation observed in Fig. 3 is also present in Fig. 5 and has been observed in all tested models. The correlations between the Cole–Cole parameters are no longer linear when $C = 0.3$ (Fig. 5). The correlations, except the m_0–C, are found to change sign and magnitude from model to model.

To test the limit of parameter resolution, a number of models were inverted with different values for C. The STDFs were computed for each Cole–Cole parameter for each inversion and plotted in Fig. 6(a). The figure illustrates that the parameter resolution decreases as C is decreased. Equivalent sensitivity analyses have been

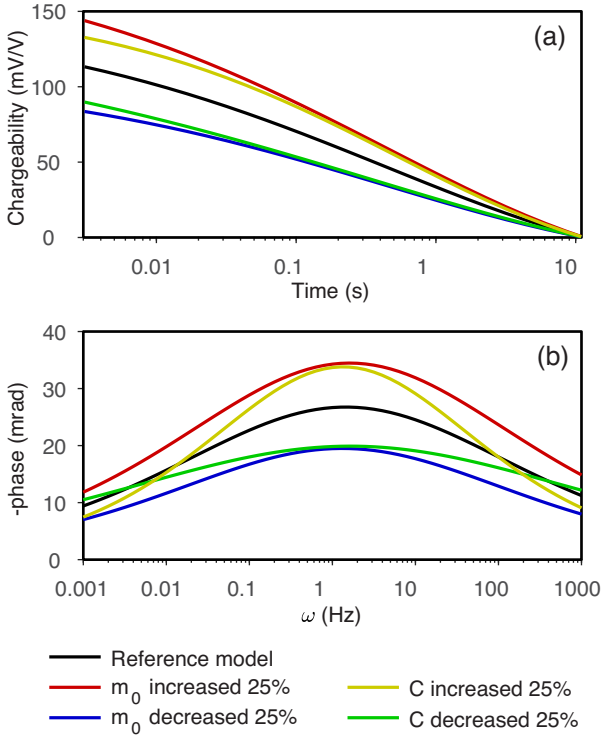

Figure 4. Variations of (a) time-domain and (b) frequency-domain forward responses with variations of the model parameters m_0 and C. The reference model (black line) has the model parameters: $\rho = 100$ Ωm, $m_0 = 200$ mV V^{-1}, $\tau = 1$ s and $C = 0.3$.

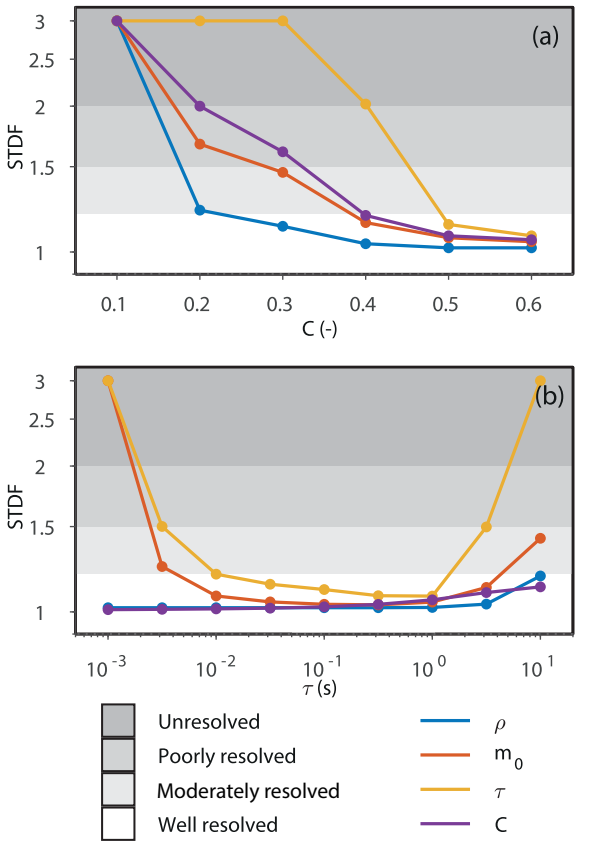

Figure 6. (a) Standard deviation factors (STDFs) of the four Cole–Cole parameters of the model: $\rho = 100$ Ωm, $m_0 = 200$ mV V^{-1}, $\tau = 1$ s and $C = [0.1–0.6]$. (b) STDFs of the four Cole–Cole parameters of the model: $\rho = 100$ Ωm, $m_0 = 200$ mV V^{-1}, $\tau = [10^{-3}–10^{1}]$ s and $C = 0.6$. If an STDF is above 3, it is put to 3 in the figure.

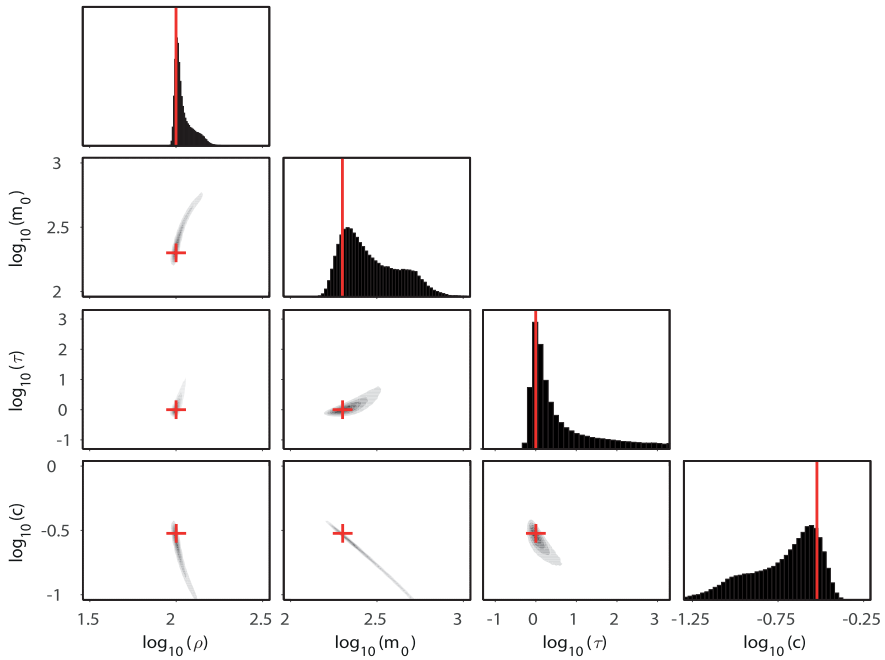

Figure 5. MCMC inversion results for the model with the Cole–Cole parameters: $\rho = 100$ Ωm, $m_0 = 200$ mV V^{-1}, $\tau = 1$ s and $C = 0.3$. The diagonal plots show the marginal posterior probability distribution of the four Cole–Cole parameters and the off-diagonal plots are the cross-correlations of the parameters (with zoom-in in the insets). The red colour indicates the true model value.

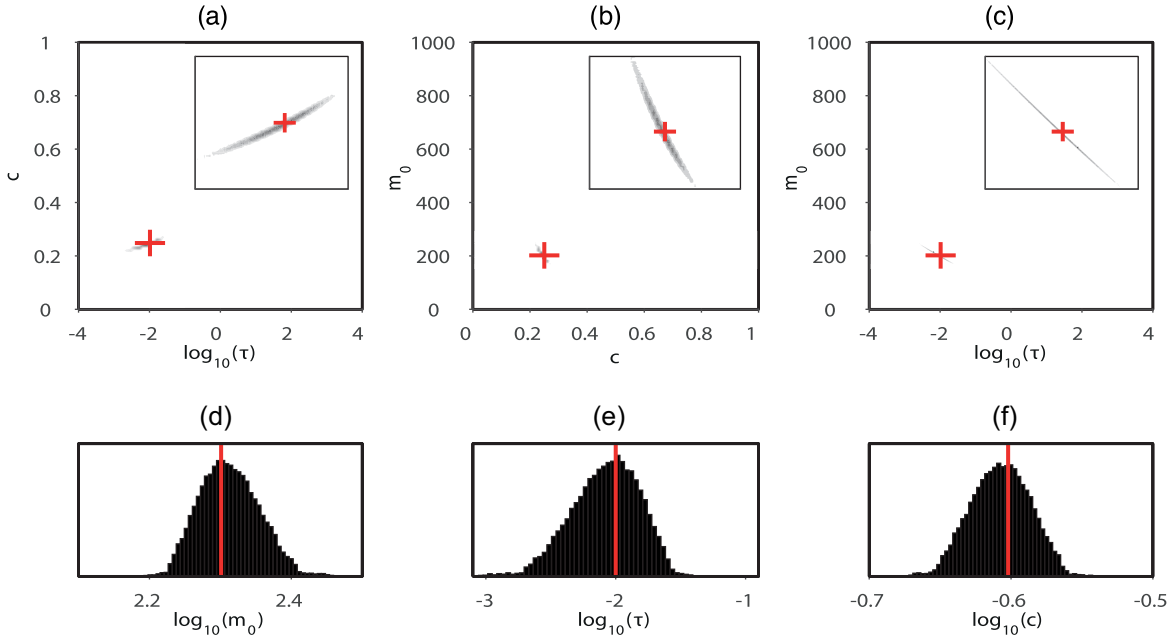

Figure 7. MCMC inversion results of a model previous studied by Ghorbani *et al.* (2007): $m_0 = 200$ mV V^{-1}, $\tau = 0.01$ s and $C = 0.25$. Subfigures (a)–(c) show cross-plots of the IP parameters accepted by the MCMC algorithm (with zoom-in in the insets) and the true model (red crosses). The layout of the plots is comparable with the one used by Ghorbani *et al.* (2007). Subfigures (d)–(f) show the marginal posterior probability distributions.

carried out for changing values of ρ, m_0 and τ. When ρ is decreased, the resolution of all the parameters decrease due to the decreased signal-to-noise ratio. Similarly, when m_0 is decreased, the IP decay is parallel shifted downwards, which again means that the signal-to-noise ratio is decreased. Furthermore, it was found that the parameters are best resolved when τ is in the range 0.01–1 s, which means that the best resolution is gained when τ is inside the acquisition range (Fig. 6b).

3.3 Comparison with previous publication

Ghorbani *et al.* (2007) have previously applied a Bayesian approach to invert TDIP data. From studies of homogenous half-spaces, they found it difficult to resolve the spectral Cole–Cole parameters from synthetic data. Their resulting posterior probability distributions showed multiple maxima and strong non-linear correlations. Based on the observed non-linearity, they concluded that it is almost impossible to resolve the spectral Cole–Cole parameters from TDIP data with a linearized inversion approach.

In this study, we have tried to replicate the results from Ghorbani *et al.* (2007), but without success. The same settings have been applied for the forward computation: 50 per cent duty cycle, two-decade acquisition range (20–2000 ms) with a total of 20 gates and DC resistivity independent on the IP parameters.

The model shown in Fig. 7 has the following parameters: $m_0 = 200$ mV V^{-1}, $\tau = 0.01$ s and $C = 0.25$. In Figs 7(a)–(c), the Cole–Cole parameters are plotted against each other with the same layout as in Ghorbani *et al.* (2007), allowing for a direct comparison. In Figs 7(d)–(f), the marginal posterior probability distributions are shown. Contrary to results of Ghorbani *et al.* (2007), the posterior probability has only one maximum and the marginal distributions are approximately bell-shaped in the logarithmic space, which means that the forward problem is linearizable (Mosegaard & Tarantola 2002). The same can be concluded for all the remaining models studied in Ghorbani *et al.* (2007) except for models were τ is far from the acquisition range. In these models, the distributions can show a skewness.

3.4 Three-layer model

The resolution and the correlation of the Cole–Cole parameters have also been studied for layered models. Here, we show the results of a three-layer model, which is characterized by a highly chargeable middle layer. The true model parameters are: $\rho = [20, 20, 20]$ Ωm, $m_0 = [5, 300, 5]$ mV V^{-1}, $\tau = [0.08, 8, 1]$ s, $C = [0.6, 0.6, 0.6]$ and thk $= [7, 7, -]$ m.

The MCMC inversion results are shown in Fig. 8 and the STDFs are listed in Table 2. In the first two layers, all the Cole–Cole parameters, including the layer thicknesses, are well to moderately resolved, and the maxima of the MCMC posterior probability distributions (yellow dashed line) agree very well with the true model (red line). In the bottom layer, m_0 and τ are unresolved. In general, it is observed that the resolution decreases significantly below chargeable layers.

The correlations between the Cole–Cole parameters in the three-layer model (Fig. 8) are shown in the correlation matrix in Fig. 9. The degree of correlation is given as Persons's correlation coefficient (Upton & Cook 2008):

$$\mathrm{PCC}_{x,y} = \frac{\mathrm{cov}\,(x, y)}{\mathrm{STD}_x\ \mathrm{STD}_y}. \tag{8}$$

This is a simple normalization of the covariance of the model parameters x and y. PCC $= 1$ is a total positive correlation, PCC $= 0$ is no correlation and PCC $= -1$ is a total negative correlation

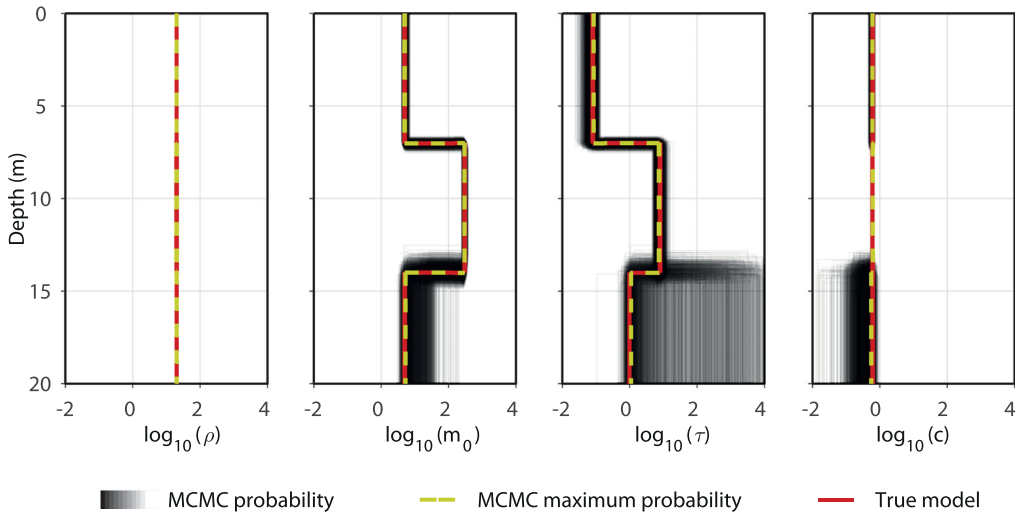

Figure 8. MCMC inversion results of a three-layer model: $\rho = [20, 20, 20]$ Ωm, $m_0 = [5, 300, 5]$ mV V^{-1}, $\tau = [0.08, 8, 1]$ s, $C = [0.6, 0.6, 0.6]$ and thk = [7, 7, –] m. The density of the black lines illustrates the probability of the models. The true model is shown with the red lines and the most probable model is shown with the dashed yellow line.

Table 2. Standard deviation factors (STDFs) describing the uncertainty of the Cole–Cole parameters of the three-layer model presented in Fig. 8.

Layer	STDF (ρ)		STDF (m_o)		STDF (τ)		STDF (C)		STDF (thk)	
	MCMC	Lin.	MCMC	Lin.	MCMC	Lin.	MCMC	Lin.	MCMC	Lin.
1	1.01	1.01	1.08	1.07	1.26	1.21	1.07	1.06	1.02	1.02
2	1.03	1.03	1.07	1.06	1.19	1.18	1.03	1.03	1.06	1.06
3	1.01	1.01	2.00	1.20	3.00	1.25	1.81	1.19		

Notes: The STDFs are computed from the mean of the marginal posterior probability distribution found with the MCMC inversion approach (MCMC). The STDFs determined from a linearized analysis of the true model (Lin.) are listed as well.

(anticorrelation). The correlation matrix shows the same strong negative correlation between m_0 and C in all three layers as was observed for homogenous half-spaces (Figs 3 and 5). Other strong correlations can be observed within each individual layer.

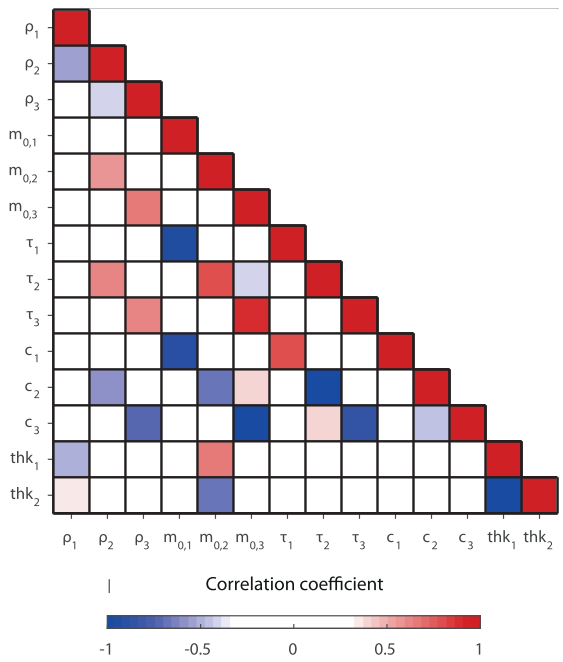

Figure 9. Persons's correlation coefficients for the model presented in Fig. 8.

The results of a linearized uncertainty analysis of the true model are listed together with the uncertainties computed from the MCMC in Table 2. The two methods agree very well in the first two layers. In the bottom layer, the MCMC method finds the IP parameters to be poorly resolved or unresolved, and the linearized approach finds the parameters well resolved. The explanation of this inconsistency is the presence of a secondary maximum found for m_0 in layer three as shown clearly by the marginal posterior probability in Fig. 10. The MCMC approach sees both maxima and therefore computes a higher STDF than the linearized approach, which only sees one of the maxima and computes a local uncertainty.

3.5 Field example

Today, TDIP field data are often acquired in 2-D or 3-D. However, due to the time required to carry out a MCMC inversion of full-decay 2-D/3-D IP data, it is not realistic to apply the MCMC approach as a standard inversion tool for such surveys. Here, we instead present MCMC inversion results of a 1-D TDIP sounding in order to investigate the uncertainty and the correlation of the Cole–Cole parameters retrieved from field measurements. These results will also be valid for 2-D and 3-D inversions in general.

The 1-D sounding is extracted from a larger 3-D data set, which was acquired in Grindsted, Denmark, in connection with an environmental survey carried out by the project GEOCON (www.geocon.env.dtu.dk). TD DCIP data were collected with a 100 per cent duty cycle and a current on time of 4000 ms. A total of 21 quadrupoles were used with electrode spacing in the range: |AB| = 18–119 m and |MN| = 2–16 m. In the data processing, one restively data point was silenced along with the

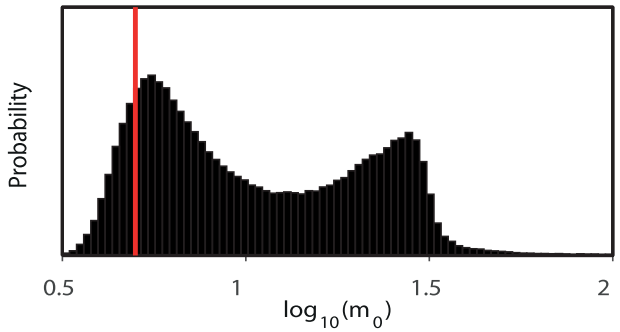

Figure 10. Marginal posterior probability distribution of the chargeability, m_0, in layer three of the model shown in Fig. 8. The distribution has two maxima: a global maximum close to the true model (red line) and a secondary maximum at a higher chargeability.

first gates of the IP decays due to coupling effects. The field data are shown in Fig. 11. The data were inverted with a three-layer model and the inversion was started from a homogenous half-space model ($\rho = 200\ \Omega$m, $m_0 = 10\,\text{mV V}^{-1}$, $\tau = 0.1$ s and $C = 0.5$).

The MCMC inversion results are shown in Fig. 12 together with a log of lithotypes determined from a borehole (DGU 114.2507) drilled close to the sounding. The uncertainties of the model parameters are listed in Table 3 and the correlation matrix is shown in Fig. 13.

The posterior probability distribution of the Cole–Cole parameters shows a well-resolved top layer with a thickness and resistivity of approximately 9 m and 90 Ωm, which fit well with the sand layer from the borehole. In the second layer, the restively is low (63 Ωm) and the chargeability is high (501 mV V^{-1}), which is expected for lignite and for silty sand layers. The thickness and the restively of the second layer are strongly correlated (Fig. 13), which indicates a minor equivalence problem (Fitterman *et al.* 1988). However, the Cole–Cole parameters of the layer are all resolved except for the time constant (Table 3). The low value of C in the second layer can be explained by the strong negative $m_0 - C$ correlation (Fig. 13). The third layer is very poorly resolved, because it is below the focus depth of the survey (Fig. 11).

We have compared the MCMC inversion results to results of the gradient-based linearized inversion approach described in the Appendix. The linearized results (Fig. 11, blue line) agree well with the MCMC probability maximum (yellow dashed line) in the top layer (Table 3). In the second layer, the two inversions find different values for m_0 and C, which could be due to the correlation between the two parameters, but the values for ρ and for τ agree very well.

4 DISCUSSION

In this study, we have used the advantages of the MCMC inversion approach to obtain a non-linearized sensitive analysis of spectral Cole–Cole parameters retrieved from TDIP data. The method has been applied to show to what degree the posterior probability distributions of the Cole–Cole parameters are bell-shaped and unimodal, that is, to what degree the inversion problem is linearizeable as discussed by Mosegaard & Tarantola (2002) and Tarantola (2005). Compared to gradient-based inversion methods, the MCMC approach has an advantage when it comes to quantifying parameter uncertainties and correlation as shown with the results. Furthermore, it was found that the MCMC approach is largely insensitivity to the start model of the inversion as long as a satisfying number of iterations are used. This is in contrast to gradient-based inversions where the stating model often has a large influence on the model result.

However, this study does not suggest MCMC as a general scheme for inversion of 2-D/3-D IP field or laboratory data. Inversion of a 2-D data set can be done in a few hours using classic gradient-based inversion, but will takes days using the MCMC approach. Furthermore, the results of this study indicate that it is justified to apply a linearization of the TDIP inversion problem and that the spectral Cole–Cole parameters can be retrieved using a gradient-based inversion approach as long as a proper acquisition time has been used in the data acquisition.

5 CONCLUSIONS

A 1-D MCMC inversion algorithm has been implemented to evaluate the spectral Cole–Cole parameters retrieved from TDIP data. Furthermore, the effect of acquisition time and parameter values on model resolution has been studied. The MCMC algorithm was applied to both synthetic data and field data and the results show that it is possible to resolve the Cole–Cole parameters from TDIP data.

It was found that the acquisition range of the IP signal is a very important factor controlling the resolution of the Cole–Cole parameters. Between two and three decades in time are necessary for resolving the IP parameters. With only two decades or less,

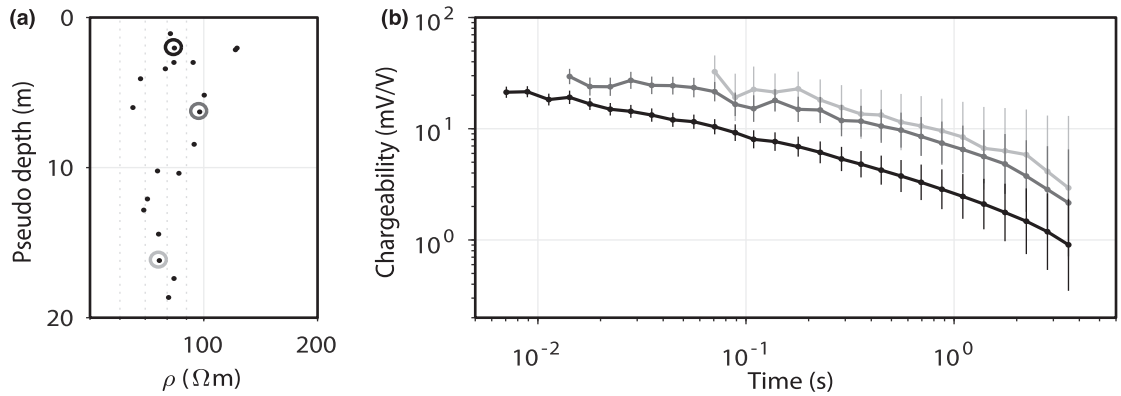

Figure 11. (a) Apparent resistivity data from a 1-D sounding acquired in Grindsted, Denmark. (b) Examples of measured IP decays corresponding to the resistivity measurements marked with circles in (a).

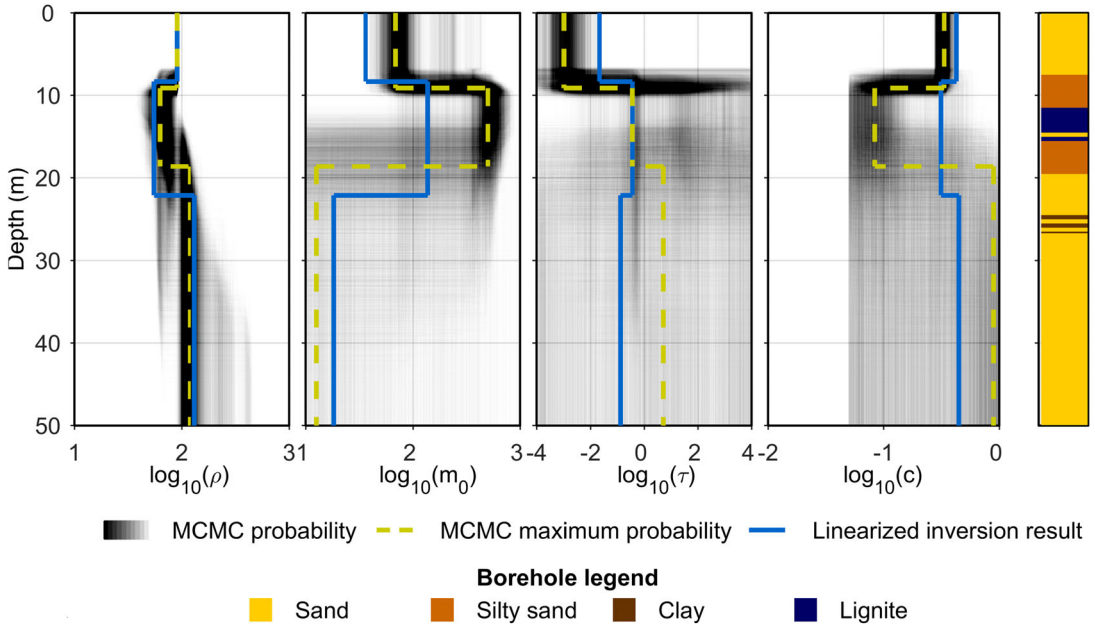

Figure 12. Results of an MCMC inversion with three layers of a 1-D sounding from Grindsted in Denmark (data shown in Fig. 11). The density of the black lines shows the probability distribution of the Cole–Cole parameters and the yellow dashed lines show their maximum. The blue lines are the result of a linearized inversion of the data. A log of lithotypes determined from a borehole drilling close to the sounding is show to the right. Note that the plot of τ is scaled wider to capture the entire parameter range.

Table 3. The parameter values and the standard deviation factors (STDFs) of the three-layer inversion result of field data from Grindsted, Denmark.

(a) Resulting Cole–Cole parameter values

Layer	ρ (Ωm)		m_0 (mV V)		τ (s)		C (−)		thk (m)	
	MCMC	Lin.	MCMC	Lin.	MCMC	Lin.	MCMC	Lin.	MCMC	Lin.
1	90.8	90.8	69.0	36.5	1.0×10^{-3}	2.1×10^{-2}	0.33	0.43	9.1	8.3
2	63.1	55.2	501.2	137.5	3.6×10^{-1}	3.6×10^{-1}	0.83	0.31	9.5	13.8
3	114.8	130.6	–	18.3	–	1.3×10^{-1}	–	0.45		

(b) Uncertainty analysis

Layer	STDF (ρ)		STDF (m_0)		STDF (τ)		STDF (C)		STDF (thk)	
	MCMC	Lin.	MCMC	Lin.	MCMC	Lin.	MCMC	Lin.	MCMC	Lin.
1	1.00	1.01	1.15	1.19	1.28	1.67	1.14	1.12	1.07	1.15
2	1.05	1.22	1.11	1.94	–	–	1.37	1.98	1.30	1.75
3	1.06	1.30	–	–	–	–	–	–		

Notes: The results are listed for an MCMC inversion (MCMC) and a linearized gradient-based inversion (Lin.) of the data. Non-values indicate unresolved parameters.

the parameters are poorly resolved and the linearity of the problem breaks down. This is seen as non-linear correlations between the model parameters and as posterior probability distributions with more than one maximum.

The resolution of the Cole–Cole parameters also depends strongly on the value of the parameters themselves. The inversion linearity breaks down also when the frequency exponent, C, approaches zero. As the chargeability, m_0, and/or the resistivity, ρ, decrease, the noise increases making it more difficult to resolve the parameters. Furthermore, the time constant, τ, must be within the acquisition range for being well resolved.

The posterior probability distributions and the parameter correlations were studied for homogenous half-spaces and layered models. For resolved models, the distributions are approximately bell-shaped with a single maximum at the true model value. In general, this means that the problem can be linearized. Results from a linearized inversion approach were compared to the results from the MCMC inversion. The comparison showed that, for resolved models, the linearized inversion finds the model with the highest probability. It also showed that generally the linearized uncertainty analysis approximates the MCMC uncertainty well, as long as no multiple maxima are present in the MCMC results.

ACKNOWLEDGEMENTS

The work was cofounded by the project GEOCON (Advancing GEOlogical, geophysical and CONtaminant monitoring technologies for contaminated site investigation, www.geocon.env.dtu.dk, last accessed 1 September 2017). The project was also supported by Formas (The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning), BeFo (Swedish Rock Engineering

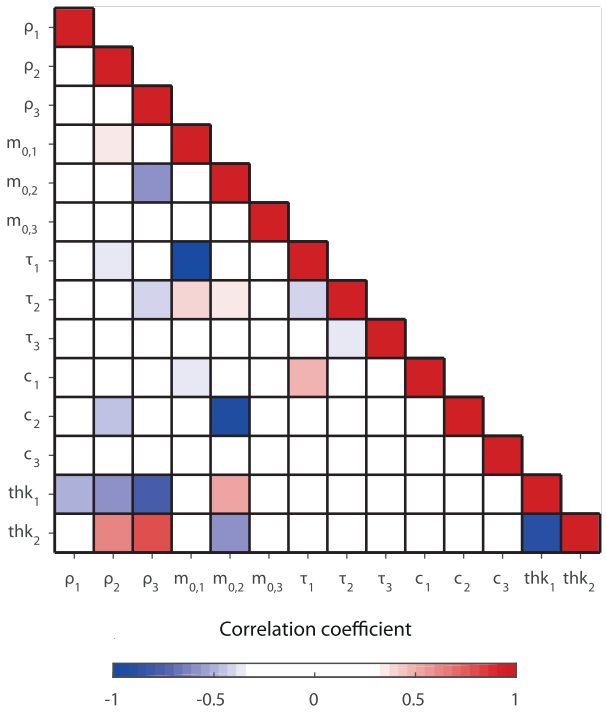

Figure 13. Persons's correlation coefficients of the Cole–Cole model parameters determined from an MCMC inversion of TDIP field data (data shown in Fig. 11).

Research Foundation) and SBUF (The Development Fund of the Swedish Construction Industry) within the Geoinfra-TRUST framework (Transparent Underground Structure, www.trust-geoinfra.se, last accessed 1 September 2017).

REFERENCES

Auken, E., Christiansen, A.V., Jacobsen, B.H., Foged, N. & Sørensen, K.I., 2005. Piecewise 1D laterally constrained inversion of resistivity data, *Geophys. Prospect.,* **53,** 497–506.

Binley, A., Slater, L.D., Fukes, M. & Cassiani, G., 2005. Relationship between spectral induced polarization and hydraulic properties of saturated and unsaturated sandstone, *Water Resour. Res.,* **41,** 1–13.

Chen, J., Kemna, A. & Hubbard, S.S., 2008. A comparison between Gauss-Newton and Markov-chain Monte Carlo-based methods for inverting spectral induced-polarization data for Cole-Cole parameters, *Geophysics,* **73,** F247–F259.

Chongo, M., Christiansen, A.V., Fiandaca, G., Nyambe, I.A., Larsen, F. & Bauer-Gottwein, P., 2015. Mapping localised freshwater anomalies in the brackish paleo-lake sediments of the Machile-Zambezi Basin with transient electromagnetic sounding, geoelectrical imaging and induced polarisation, *J. appl. Geophys.,* **2015,** 81–92.

Cole, K.S. & Cole, R.H., 1941. Dispersion and absorption in dielectrics, *J. Chem. Phys.,* **9,** 341–351.

Doetsch, J., Fiandaca, G., Auken, E., Christiansen, A.V., Cahill, A.G. & Jacobsen, J.D., 2015a. Field scale time-domain spectral induced polarization monitoring of geochemical changes induced by injected CO_2 in a shallow aquifer, *Geophysics,* **80,** WA113–WA126.

Doetsch, J., Ingemann-Nielsen, T., Christiansen, A.V., Fiandaca, G., Auken, E. & Elberling, B., 2015b. Direct current (DC) resistivity and induced polarization (IP) monitoring of active layer dynamics at high temporal resolution, *Cold Reg. Sci. Technol.,* **119,** 16–28.

Fiandaca, G., Auken, E., Gazoty, A. & Christiansen, A.V., 2012. Time-domain induced polarization: full-decay forward modeling and 1D laterally constrained inversion of Cole-Cole parameters, *Geophysics,* **77,** E213–E225.

Fiandaca, G., Ramm, J., Binley, A., Gazoty, A., Christiansen, A.V. & Auken, E., 2013. Resolving spectral information from time domain induced polarization data through 2-D inversion, *Geophys. J. Int.,* **192,** 631–646.

Fiandaca, G., Doetsch, J., Vignoli, G. & Auken, E., 2015. Generalized focusing of time-lapse changes with applications to direct current and time-domain induced polarization inversions, *Geophys. J. Int.,* **203,** 1101–1112.

Fitterman, D.V., Meekes, J.A.C. & Ritsema, I.L., 1988. Equivalence behavior of three electrical sounding methods as applied to hydrogeological problems, in *50th Annual meeting of EAEG,* The Hague, The Netherlands.

Gazoty, A., Fiandaca, G., Pedersen, J., Auken, E. & Christiansen, A.V., 2012a. Mapping of landfills using time-domain spectral induced polarization data: the Eskelund case study, *Near Surf. Geophys.,* **10,** 575–586.

Gazoty, A., Fiandaca, G., Pedersen, J., Auken, E., Christiansen, A.V. & Pedersen, J.K., 2012b. Application of time domain induced polarization to the mapping of lithotypes in a landfill site, *Hydrol. Earth Syst. Sci.,* **16,** 1793–1804.

Gazoty, A., Fiandaca, G., Pedersen, J., Auken, E. & Christiansen, A.V., 2013. Data repeatability and acquisition techniques for Time-Domain spectral Induced Polarization, *Near Surf. Geophys.,* **11,** 391–406.

Gelman, A., Roberts, G.O. & Gilks, W.R., 1996. Efficient Metropolis jumping rules, Bayesian statistics, 5, 599-608, 42.

Ghorbani, A., Camerlynck, C., Florsch, N., Cosenza, P. & Revil, A., 2007. Bayesian inference of the Cole-Cole parameters from time- and frequency-domain induced polarization, *Geophys. Prospect.,* **55,** 589–605.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications, *Biometrika,* **57,** 97–109.

Hönig, M. & Tezkan, B., 2007. 1D and 2D Cole-Cole-inversion of time-domain induced-polarization data, *Geophys. Prospect.,* **55,** 117–133.

Hördt, A., Hanstein, T., Hönig, M. & Neubauer, F.M., 2006. Efficient spectral IP-modelling in the time domain, *J. appl. Geophys.,* **59,** 152–161.

Johansen, H.K. & Sørensen, K.I., 1979. Fast Hankel transforms, *Geophys. Prospect.,* **27,** 876–901.

Johansson, S., Fiandaca, G. & Dahlin, T., 2015. Influence of non-aqueous phase liquid configuration on induced polarization parameters: conceptual models applied to a time-domain field case study, *J. appl. Geophys.,* **123,** 295–309.

Johansson, S., Sparrenbom, C., Fiandaca, G., Lindskog, A., Olsson, P.-I., Dahlin, T. & Rosqvist, H., 2017. Investigations of a Cretaceous limestone with spectral induced polarization and scanning electron microscopy, *Geophys. J. Int.,* **208**(2), 954–972.

Kemna, A., 2000. Tomographic inversion of complex resistivity: theory and application, *PhD thesis*, Ruhr-Universität Bochum, Der Andere Verlag Osnabrück, Germany.

Loke, M.H., Chambers, J.E. & Ogilvy, R.D., 2006. Inversion of 2D spectral induced polarization imaging data, *Geophys. Prospect.,* **54,** 287–301.

Luo, Y. & Zhang, G., 1998. *Theory and Application of Spectral Induced Polarization*, Geophysical Monograph Series, Vol. 8, pp. 1–171, Society of Exploration Geophysicists.

Malinverno, A., 2002. Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem, *Geophys. J. Int.,* **151,** 675–688.

Maurya, P.K., Fiandaca, G., Auken, E. & Christiansen, A.V., 2016. Lithological characterization of a comtaminated site using direct current resistivity and time domain induced polarization, in *IP2016/4th International Workshop on Induced Polarization,* Aarhus, Denmark, Available at: http://hgg.au.dk/fileadmin/www.gfs.au.dk/DIV/Abstracts_from_Session_B.pdf, pp. 21–23.

Metropolis, N., Rosenbluth, M.N., Rosenbluth, A.W. & Teller, A.H., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.,* **21,** 1987–1092.

Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.,* **100,** 12 431–12 447.

Mosegaard, K. & Tarantola, A., 2002. Probabilistic approach to inverse problems, in *International Handbook of Earthquake and Engineering Seismology*, pp. 237–265, eds. Lee, W., Jennings, P., Kisslingers, C. & Kanamori, H., Academic Press.

Oldenburg, D.W. & Li, Y., 1994. Inversion of induced polarization data, *Geophysics,* **59,** 1327–1341.

Olsson, P.-I., Fiandaca, G., Larsen, J.J., Dahlin, T. & Auken, E., 2016. Doubling the spectrum of time-domain induced polarization by harmonic de-noising, drift correction, spike removal, tapered gating, and data uncertainty estimation, *Geophys. J. Int.,* **207**(2), 774–784.

Olsson, P.I., Dahlin, T., Fiandaca, G. & Auken, E., 2015. Measuring time-domain spectral induced polarization in the on-time: decreasing acquisition time and increasing signal-to-noise ratio, *J. appl. Geophys.,* **123,** 316–321.

Pelton, W.H., Ward, S.H., Hallof, P.G., Sill, W.R. & Nelson, P.H., 1978. Mineral discrimination and removal of inductive coupling with multifrequency IP, *Geophysics,* **43,** 588–609.

Revil, A., Florsch, N. & Camerlynck, C., 2014. Spectral induced polarization porosimetry, *Geophys. J. Int.,* **198,** 1016–1033.

Roberts, G.O., Gelman, A. & Gilks, W.R., 1997. Weak convergence and optimal scaling of random walk metropolis algorithms, *Ann. Appl. Probab.,* **7,** 110–120.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.,* **40**(3), 1009, doi:10.1029/2000RG00089.

Seigel, H.O., 1959. Mathematical formulation and type curves for induced polarization, *Geophysics,* **24,** 547–565.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation,* SIAM.

Tarantola, A. & Valette, B., 1982. Inverse problems = quest for information, *J. Geophys.,* **50,** 159–170.

Tarasov, A. & Gurin, G., 2016. Spectral induced polarization of the ore zone of the gold deposit Sukhoi Log, in *IP2016/4th International Workshop on Induced Polarization,* Aarhus, Denmark, Available at: http://hgg.au.dk/fileadmin/www.gfs.au.dk/DIV/Abstracts_from_Poster_Session_B.pdf.

Titov, K., Tarasov, A., Ilyin, Y., Seleznev, N. & Boyd, A., 2010. Relationships between induced polarization relaxation time and hydraulic properties of sandstone, *Geophys. J. Int.,* **180,** 1095–1106.

Upton, G. & Cook, I., 2008. *A Dictionary of Statistics,* 2nd edn, Oxford Univ. Press.

Wemegah, D., Fiandaca, G., Auken, E., Menyeh, A. & Danuor, S., 2014. Time-domain spectral induced polarization and magnetics for mapping municipal solid waste deposits in Ghana, in *Near Surface Geoscience 2014–20th European Meeting of Environmental and Engineering Geophysics*, Available at: http://www.earthdoc.org/publication/publicationdetails/?publication=77751.

Xiang, J., Cheng, D., Schlindwein, F. & Jones, N., 2003. On the adequacy of identified Cole–Cole models, *Comput. Geosci.,* **29,** 647–654.

Yoshioka, K. & Zhdanov, M.S., 2005. Three-dimensional nonlinear regularized inversion of the induced polarization data based on the Cole-Cole model, *Phys. Earth planet. Inter.,* **150,** 29–43.

Yuval & Oldenburg, D.W., 1997. Computation of Cole-Cole parameters from IP data, *Geophysics,* **62,** 436–448.

APPENDIX: IMPLEMENTATIONS

A1 Gradient-based inversion and uncertainties

In this study, we have compared the results of a gradient-based linearized inversion approach to results of a MCMC inversion. Furthermore, the linearized approach has been integrated in the MCMC algorithm to determine a start model and scale the model perturbation as described below. The gradient-based inversion scheme follows Fiandaca *et al.* (2012), where a first term Taylor expansion is used for linearization of the non-linear forward problem. A linearized uncertainty analyses can then be computed based on the posterior covariance matrix (Tarantola & Valette 1982), which is given by

$$C_{\text{est}} = \left(G^T C_{\text{obs}} G\right)^{-1}, \tag{A1}$$

where G is the Jacobi matrix and C_{obs} is the data covariance matrix holding the data uncertainties. Under the assumption that the model parameters are normally distributed in the logarithmic space, the standard deviation can be obtained by taking the square root of the diagonal of C_{est}. Due to the logarithmic transform applied on the parameter, a STDF of the ith parameter can then be computed as

$$\text{STDF}\,(m_i) = \exp\left(\sqrt{C_{\text{est}(i,i)}}\right). \tag{A2}$$

Table A1. Implementation of 1-D MCMC inversion algorithm.

1:	Set N_{ite}, k_{step}
2:	Run linearized inversion to determine starting model, $\boldsymbol{m}_{\text{start}}$, and the covariance matrix, $\boldsymbol{C}_{\text{est}}(\boldsymbol{m}_{\text{start}})$
3:	$\boldsymbol{m}_{\text{cur}} = \boldsymbol{m}_{\text{start}}$
4:	**for** $i = 1, N_{\text{ite}}$ **do**
5:	Compute a new model proposal: $\boldsymbol{m}_{\text{new}} = \boldsymbol{m}_{\text{cur}} + \boldsymbol{L}\boldsymbol{n}k_{\text{step}}$
6:	Compute acceptance probability: $P_{\text{acc}} = \frac{P_{\text{like}}(\boldsymbol{m}_{\text{new}})}{P_{\text{like}}(\boldsymbol{m}_{\text{cur}})}$
7:	Draw random number (α) from a uniform distribution [0:1]
8:	**if** $P_{\text{acc}} > \alpha$
9:	$\boldsymbol{m}_{\text{cur}} = \boldsymbol{m}_{\text{new}}$
10:	**else**
11:	$\boldsymbol{m}_{\text{cur}} = \boldsymbol{m}_{\text{cur}}$
12:	**end if**
13:	**end for**

A2 MCMC inversion

The algorithm presented in Table A1 describes the implementation of the MCMC scheme used for inversion of TDIP data. The algorithm is based on a Metropolis–Hastings sampling algorithm (Metropolis *et al.* 1953; Hastings 1970), which applies a random walk to sample models to a Markov Chain, which converge toward the posterior probability distribution of the model space.

Starting model

With the MCMC sampling method, it is difficult to determine when a sampling has converged and when all local probability maxima have been sampled. To avoid that a walker is trapped in a local maximum, multiple walkers can be started independently from different location in the model space. For the models studied here, we found that all walkers found the same maxima as long as a sufficient number of iterations (N_{ite}) was used. Furthermore, the step length (k_{step}) constant should be chosen so the acceptance rate of models is approximately 30 per cent for homogenous half-space models, which ensures the most effective sampling (Gelman *et al.* 1996; Roberts *et al.* 1997). For layered models, the acceptance rate should be 40 per cent–50 per cent. We apply walkers, which are started from the resulting model of a gradient-based inversion (Table A1, steps 2 and 3). In this way, the walker will start in a local maximum and the burn-in phase is minimized.

Model perturbation

For each iteration, a new model ($\boldsymbol{m}_{\mathbf{new}}$) is proposed by the algorithm to be the next model in the Markov Chain (Table A1, step 5). The new model is computed by adding a model perturbation to the current model ($\boldsymbol{m}_{\mathbf{cur}}$). The perturbation is given by $\boldsymbol{L}$, which is a lower triangular matrix defined as the Cholesky decomposition of $\boldsymbol{C}_{\mathbf{est}}$ (determined in step 2 in Table A1) multiplied by a vector of random numbers drawn from a Gaussian distribution ($\boldsymbol{n}$) and k_{step}. By applying this realization of $\boldsymbol{C}_{\mathbf{est}}$, the length of the model perturbation is adapted for each model parameter individually according to their interval range and uncertainty. The algorithm then takes larger steps for poorly defined parameters and can thus search the subspace faster. It was found that the application of this scaling lets the algorithm converge over 200 times faster than a standard Gaussian proposer.

Acceptance probability

An acceptance probability, P_{acc}, of the new model determines whether the model is accepted to the Markov Chain or rejected (Table A1, steps 6–12). Because we apply a symmetric proposal distribution, where the probability of going from $\boldsymbol{m}_{\mathbf{cur}}$ to $\boldsymbol{m}_{\mathbf{new}}$ is the same as going from $\boldsymbol{m}_{\mathbf{new}}$ to $\boldsymbol{m}_{\mathbf{cur}}$, the acceptance probability can be computed as given by Mosegaard & Tarantola (1995):

$$P_{\text{acc}} = \min\left[1,\ \frac{P_{\text{like}}(m_{\text{new}})}{P_{\text{like}}(m_{\text{cur}})}\right], \tag{A3}$$

where, if the uncertainties of the model parameters are assumed to be Gaussian, the likelihood can be computed as given by Tarantola (2005):

$$P_{\text{like}}(\boldsymbol{d}|\boldsymbol{m}, I) = K \cdot \exp\left[-\frac{1}{2}(\boldsymbol{g}(\boldsymbol{m}) - \boldsymbol{d}_{\mathbf{obs}})^T \boldsymbol{C}_{\mathbf{obs}}{}^{-1} (\boldsymbol{g}(\boldsymbol{m}) - \boldsymbol{d}_{\mathbf{obs}})\right], \tag{A4}$$

where $g(\boldsymbol{m})$ is the forward response of the model $\boldsymbol{m}$ and K is a normalization constant. How this is implemented can be seen in steps 6–12 in Table A1.

No priors have been applied in the inversion except for hard boundaries set on the model space geometry to confine the random walk. This can be considered as a uniform distribution prior, that is, the prior probability is the same for all models within the boundaries, why these can be neglected in the probability calculations. The boundaries applied are: $\rho = 0.1 - 20000\,\Omega\text{m}$, $m_0 = 0.1\text{–}1000\,\text{mV V}^{-1}$, $\tau = 10^{-5}\text{–}10^{5}\,\text{s}$, $C = 0.05\text{–}1$ and $\text{thk} = 0.1\text{–}100\,\text{m}$.