

FULL PAPER

Open Access



# Effect of data error correlations on trans-dimensional MT Bayesian inversions

Rongwen Guo<sup>1,2,3</sup>, Liming Liu<sup>1,2</sup>, Jianxin Liu<sup>1,2</sup>, Ya Sun<sup>1,2</sup> and Rong Liu<sup>1\*</sup> 

## Abstract

Real magnetotelluric (MT) data errors are commonly correlated, but MT inversions routinely neglect such correlations without an investigation on the impact of this simplification. This paper applies a hierarchical trans-dimensional (trans-D) Bayesian inversion to examine the effect of correlated MT data errors on the inversion for subsurface geoelectrical structures, and the model parameterization (the number of conductivity interfaces) is treated as an unknown. In the inversion considering error correlations, the data errors are parameterized by the first-order autoregressive (AR(1)) process, which is included as an unknown in the inversion. The data information itself determines the AR(1) parameter. The trans-D inversion applies the reversible-jump Markov chain Monte Carlo algorithm to sample the trans-D posterior probability density (PPD) for the model parameters, model parameterization and AR(1) parameters, accounting for the uncertainties of the model dimension and data error correlation in the uncertainty estimates of the conductivity profile. In the inversion ignoring the correlation, we neglect the correlation effect by turning off the AR(1) parameter. Then the correlation effect on the MT inversion can be examined upon comparing the posterior marginal conductivity profiles from the two inversions. Further investigation is then carried out for a synthetic case and a real MT data example. The results indicate that for strong correlation cases, neglecting error correlations can significantly affect the inversion results.

**Keywords:** Trans-D Bayesian inversion, Autoregressive model, Parameterization, Magnetotelluric method

## Introduction

Magnetotelluric (MT) data inversions are widely carried out to understand the subsurface geoelectrical structure with applications such as geothermal investigations (Heise et al. 2008; He et al. 2016), exploration for ore deposits (Zhang and Chouteau 1992) and hydrocarbon reservoirs (He et al. 2010), and tectonic imaging (Becken and Ritter 2012; Espurt et al. 2014; Goto et al. 2005). MT inversions are commonly conducted based on the assumption that the data errors are uncorrelated either over frequencies or over space with the consideration of variances (not covariances) in the inversion (e.g., Yoshimura et al. 2018; Grayver 2015; Guo et al. 2011; Usui et al. 2016; Wheelock 2012). This simplification makes the calculation of the data misfit function easier

and more efficient. For example, many studies utilize only diagonal elements (variances) of the data covariance matrix to construct the weighting matrix for data misfit calculations (e.g., Grayver 2015; Guo et al. 2011), while other studies treat all errors as providing equal contributions to the inversion by employing identical data variances (Lee et al. 2009). The validity of these simplifications depends on how strongly the correction of the data errors affects those errors.

It is well known that MT transfer functions estimated from the variations of horizontal electromagnetic fields can include frequency- and spatially correlated noise. Frequency-correlated noise can arise from the data measurement process (Egbert 1997; Eisel and Egbert 2001) and the simplified error statistics used in the inversion (e.g., due to the idealized physics of source and forward models and the simplified parameterizations of the Earth models). Relative to frequency-correlated errors, spatially correlated errors have been studied more extensively (Goubau et al. 1978; Gamble et al. 1979; Goubau et al. 1984). Even

\*Correspondence: liurongkaoyan@126.com

<sup>1</sup> School of Geosciences and Info-physics, Central South University, Lushan Road South 932, Changsha 410083, China

Full list of author information is available at the end of the article

with the most advanced data processing techniques, it is impossible to completely eliminate such errors from estimated MT transfer functions (Ritter et al. 1998; Weckmann et al. 2005). Due to these ubiquitous correlations, it is worthwhile to closely investigate their effect on MT inversions, which is the major goal of this paper.

The importance of properly handling correlated errors in inversions has been demonstrated in the literature (Sambridge 1999; Agostinetti and Malinverno 2010; Dosso and Dettmer 2011) based on point estimate inversions. While quantitatively evaluating the effect of error correlations on inversion results can be nontrivial in single best-fit model inversions (Backus 1970; Oldenburg 1983), Bayesian inversions can easily consider the error correlation effect in the estimates of nonlinear parameter uncertainties, providing a natural scheme to assess such effects on the inversion results. For example, the effects of error correlations in the inversion of ocean acoustic data for seabed geoacoustic parameters with a known model parameterization were examined by Dosso et al. (2006) and Huang et al. (2007), who found that the inversion results are strongly affected by neglecting those correlations, producing significantly underestimated parameter uncertainties. In fixed-dimensional MT Bayesian inversions, the impacts of both frequency- and spatially correlated noise on the inversion results (both parameter estimates and parameter uncertainties) were extensively investigated (Guo et al. 2014). Their work indicates that ignoring the correlated noise in an inversion can result in a potentially serious bias on the parameter estimation and an unrealistic underestimation of the parameter uncertainties; thus, the results should not be used routinely without a close examination.

Due to its ability to account for uncertainties in model parameterizations, the trans-dimensional (trans-D) Bayesian inversion technique (Mandolesi et al. 2018; Blatter et al. 2019), in which the model dimension is considered unknown and determined in the inversion based on the data and prior information, has recently become a common approach for Bayesian inversion applications. In the more general Bayesian inversion scheme, while the literature (Bodin et al. 2012; Gehrman et al. 2015) focuses on estimating the data error correlation based on data residuals in the inversion, some authors (Dettmer et al. 2012; Ray et al. 2013b) reexamined the data error correlation effect on the inversion results for geophysical applications. Neglecting correlated errors leads to inversion results with underestimated uncertainties, multimodal marginals and abundant dubious structures (Ray et al. 2013a). Fortunately, these phenomena can be handled properly by considering data error correlations in trans-D inversions using parametric AR models (Warner, et al. 2015; Steininger et al. 2013, Xiang et al.

2018), described by relatively few numbers of low-order AR parameters, which are also treated as unknown and are sampled trans-dimensionally as model parameters in the inversion. However, due to the common practice of neglecting error correlations in MT inversions, an examination of the effect of ignoring the correlation on the inversion results for a more general trans-D MT Bayesian inversion (treating the data error as an unknown) is of practical importance, as the topic has not yet been investigated in the literature and is the focus of this paper.

In this paper, a quantitative examination of the effect of neglecting error correlations over a range of frequencies is performed based on a trans-D Bayesian inversion, in which the uncertainty of the model parameterization is accounted for in the posterior and is thus included in the uncertainties of the inversion results. This is accomplished by comparing the marginal probability profiles of the conductivity for inversions based on sampling over the trans-D AR model space and sampling of the data variance (i.e., with and without accounting for data error correlations). An advanced reversible-jump Markov chain Monte Carlo (rjMCMC) sampling algorithm is applied to sample the multiple model subspaces of different dimensions, by employing techniques such as parallel tempering (Jasra et al. 2007; Dosso et al. 2012; Dettmer and Dosso 2012; Sambridge 2013), principal component (PC) space parameter perturbations, and birth proposals from the prior (Dosso et al. 2014; Xiang et al. 2018). Data errors are parameterized by the first-order AR process (denoted AR(1) in this paper) and sampled trans-dimensionally in the inversion considering the error correlation while fixing AR(1) at 0 in the cases where the error correlation is neglected.

## Inverse theory and algorithms

### Trans-D Bayesian formulation

In the trans-D Bayesian approach, the model of unknown parameters  $\mathbf{m}_k$  with  $M_k$  dimensions and data  $\mathbf{d}$  with  $N$  dimensions are related by Bayes' rule as follows (Green 1995; Sambridge et al. 2006; Dosso et al. 2014; Xiang et al. 2018)

$$P(k, \mathbf{m}_k | \mathbf{d}) = \frac{P(k)P(\mathbf{m}_k | k)P(\mathbf{d} | \mathbf{m}_k, k)}{\sum_{k' \in K} \int_{\mathcal{M}_{k'}} P(k')P(\mathbf{m}'_{k'} | k')P(\mathbf{d} | k', \mathbf{m}'_{k'}) d\mathbf{m}'_{k'}}, \quad (1)$$

where the denominator is the total evidence, that is the integral over all  $k$ -dimensional models  $\mathbf{m}_k$  and over all discrete values  $k$ . In Eq. (1),  $P(k)P(\mathbf{m}_k | k)$  denotes the prior information of the state  $(k, \mathbf{m}_k)$ , and  $P(\mathbf{d} | \mathbf{m}_k, k)$  denotes the data information, which can be interpreted as the likelihood of  $(k, \mathbf{m}_k)$  given (fixed) observed data  $\mathbf{d}$ . The denominator on the right side of Eq. (1) represents the total evidence of the set of models. On the left side,  $P(k, \mathbf{m}_k | \mathbf{d})$  is the posterior probability density (PPD), defined over the trans-D model space  $(k, \mathbf{m}_k)$ ,

providing the most general solution to the Bayesian inverse problem.

Due to the variable parameterization in a trans-D inversion, the results are typically considered in terms of profiles of the parameter of interest. For instance, for a depth-dependent MT inversion, the results can be considered in terms of the marginal probability profile of the conductivity computed at a given depth  $z$  by finding the histogram of all sampled conductivities at depth  $z$  by a trans-D Markov chain method (see “Reversible-jump MCMC sampling” section).

### Reversible-jump MCMC sampling

By generalizing the Metropolis–Hastings criterion, transitions from the current state  $(k, \mathbf{m}_k)$  of the chain to a proposed state  $(k', \mathbf{m}_{k'})$  are accepted with the probability given by the Metropolis–Hastings–Green criterion (Green 1995)

$$A(k', \mathbf{m}_{k'} | k, \mathbf{m}_k) = \min \left[ 1, \frac{Q(k, \mathbf{m}_k | k', \mathbf{m}_{k'})}{Q(k', \mathbf{m}_{k'} | k, \mathbf{m}_k)} \frac{P(k')P(\mathbf{m}_{k'} | k')}{P(k)P(\mathbf{m}_k | k)} \frac{L(k', \mathbf{m}_{k'})}{L(k, \mathbf{m}_k)} |\mathbf{J}| \right], \quad (2)$$

where  $|\mathbf{J}|$  is the determinant of the Jacobian matrix for a diffeomorphism from  $(k, \mathbf{m}_k)$  to  $(k', \mathbf{m}_{k'})$ . The reversible-jump MCMC algorithm is commonly implemented by the birth–death scheme in practical applications (Green 1995; Malinverno 2002; Sambridge et al. 2006; Dosso et al. 2014). The birth–death scheme applied here is based on a layered geoelectrical partition model (Xiang, et al. 2018), the parameters of which consist of  $k$  interface depths  $z_k$  above a maximum depth  $z_b$  (such that the data are insensitive to the conductivity below this depth), and the conductivity  $\sigma$  of each of the  $k + 1$  layers. The conductivity of a layer is associated with its lower interface, and the deepest layer represents a uniform half-space. The prior for  $\sigma$  is taken to be (logarithmically) uniform over a bounded interval  $[\sigma_{\min}; \sigma_{\max}]$ . The prior for the number of interfaces  $k$  is also uniform over a bounded interval  $[k_{\min}; k_{\max}]$ . In the birth–death scheme, one of three types of steps (perturbation, birth and death) is applied at each iteration. The perturbation step applies perturbations to the current parameters, without a change in the model dimension ( $k$  is unchanged). In the birth step, a new layer interface is generated randomly over  $[0, z_b]$ , increasing  $k$  by 1, and the conductivity parameter of the new layer is drawn from some distributions (e.g., a Gaussian distribution). A death move randomly selects and deletes an existing layer, decreasing  $k$  by 1; the conductivity parameter of the new, thicker layer is set to that below the removed interface. The above birth–death scheme leads to a unit determinant of the Jacobian matrix  $\mathbf{J}$  (Green 1995).

If the parameters for the birth steps are proposed from the prior, the Metropolis–Hastings–Green criterion for perturbation, birth and death reduces to (Dosso et al. 2014)

$$A_{P,B,D}(k', \mathbf{m}_{k'} | k, \mathbf{m}_k) = \min \left[ 1, \frac{L(k', \mathbf{m}_{k'})}{L(k, \mathbf{m}_k)} \right]. \quad (3)$$

In this paper, an advanced sampling algorithm based on the application of a parallel tempering technique, and efficient proposals for perturbation and birth steps is used to examine the effect of ignoring data error correlations in inversions. The parallel tempering technique (Jasra et al. 2007; Dosso et al. 2012) employs a series of parallel interacting Markov chains with successively relaxed likelihood functions to achieve a reasonable acceptance rate for the birth/death steps and to transit between multiple modes, which may exist for nonlinear problems (particularly in high-dimensional parameter subspaces). The parameter perturbation is carried out in the principal component (PC) space (Dosso et al. 2014), defined by the eigenvectors of the unit-lag model covariance to overcome the inefficient sampling problem caused by parameter correlations. The proposal for the birth step originates from the parameter uniform prior, as proposed by previous study (Dosso et al. 2014); this approach has been verified to be more efficient in MT Bayesian inversion (Xiang et al. 2018) than in other localized distributions (e.g., the Gaussian distribution).

### Trans-D autoregressive data error model

To carry out trans-D Bayesian inversions, the data error distribution,  $P(\mathbf{d} | \mathbf{m}_k, k)$ , must be specified. However, due to the complicated sources of errors (including errors in the measurement and theory), the specific form of this distribution is typically unknown. In practice, a simple multivariate (complex) Gaussian distribution can generally be applied based on the central-limit theorem (Dettmer et al. 2012; Steininger et al. 2013; Xiang, et al. 2018), given by

$$P(\mathbf{d} | \mathbf{m}_k, k) = \frac{1}{\pi^N |\mathbf{C}_d|} \exp \left( -(\mathbf{d} - \mathbf{d}(k, \mathbf{m}_k))^* \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{d}(k, \mathbf{m}_k)) \right), \quad (4)$$

where  $\mathbf{C}_d$  is the data covariance matrix,  $\mathbf{d}(k, \mathbf{m}_k)$  presents the modeled data, and  $*$  is the conjugate transpose. For uncorrelated data errors, the data covariance  $\mathbf{C}_d$  is diagonal with elements representing the data variances. For data contaminated with correlated noise,  $\mathbf{C}_d$  is generally not diagonal, and the off-diagonal elements quantify the correlations between errors (e.g., at different frequencies). In most geophysical inversions, the data error correlation (i.e., the off-diagonal terms of  $\mathbf{C}_d$ ) is routinely ignored without justification.

In this section, we apply the AR(1) process to model the error correlations between frequencies. Note that while it is possible to consider higher-order AR processes to model errors, our experience is that great care is required beyond the first order as higher-order processes can be sufficiently general to model the informative data structure rather than just errors, degrading the inversion results. In the AR(1) model, the data can be written as

$$d_i = d_i(k, \mathbf{m}_k) + d_i(a) + \varepsilon_i, \quad i = 1, \dots, N, \quad (5)$$

where  $\varepsilon_i$  is the uncorrelated Gaussian random error for data point  $i$ , and

$$d_i(a) = a(d_{i-1} - d_{i-1}(k, \mathbf{m}_k)), \quad (6)$$

is the correlated error component which depends on the AR(1) parameter  $a$  and  $d_1(a) = 0$  (assuming the correlation process begins with the first datum). Mathematically,  $a$  is limited to  $[-1, 1]$ , with positive values indicating errors with serial correlation and negative values indicating anticorrelation. Assuming that we have removed the correlation part of the error  $d_i(a)$ ,  $\mathbf{C}_d$  is the diagonal data covariance matrix with only elements composed of variances. For generality, the covariance matrix is further treated as  $\mathbf{C}_d = s^2 \mathbf{C}'_d$ , where  $\mathbf{C}'_d$  includes the preliminary relative data variance estimates and  $s$  is an unknown scale factor for all data (Xiang et al. 2018). To include the AR(1) model in an inversion, the likelihood in Eq. (4) can be rewritten in the form

$$P(\mathbf{d}|\mathbf{m}_k, k) = \frac{1}{\pi^N s^{2N} |\mathbf{C}'_d|} \exp \left( -(\mathbf{d} - \mathbf{d}(k, \mathbf{m}_k) - \mathbf{d}(a))^* (s^2 \mathbf{C}'_d)^{-1} (\mathbf{d} - \mathbf{d}(k, \mathbf{m}_k) - \mathbf{d}(a)) \right) \quad (7)$$

where  $s$  and  $a$  are hyperparameters included in the inversion. An implicit sampling method is applied here for the scale factor  $s$  (see Xiang et al. 2018 for more details). During sampling, the AR(1) process is applied (turned on) if needed (e.g., by the data), but this process is not applied (turned off) if it is unnecessary; this is performed to avoid overparameterizing the error model. For the parallel tempering technique (Xiang et al. 2018) used in this paper, an arbitrary sampling temperature  $T \geq 1$  can be used, and the likelihood is relaxed by  $1/T$ .

### Synthetic test case

In this section, an eight-layer model is considered to examine the correlation effect on MT Bayesian inversions with both the model parameterization and the data error parameterization unknown (the true model is listed in Table 1 and the considered AR(1) parameters are 0.3 and 0.8). Synthetic impedance data are generated at 40 logarithmically spaced periods from 0.0025 to 250 s with the addition of 2% correlated Gaussian-distributed noise. The priors applied in the inversion are uniform distributions in the range  $[1, 30]$  for the parameterization index number  $k$ , and the range  $[-5, 1]$   $\text{S m}^{-1}$  for the log conductivity; the maximum penetrating depth is set to  $z_b = 100$  km. The prior information for AR(1) is set to be uniform on  $[-0.5, 1]$ , and the state space of the AR error process is given to be  $[0, 1]$  (i.e., AR(0) and AR(1)). A total of 5 chains are applied in the parallel tempering scheme with temperatures of  $T_i = 1.5^{(i-1)}$ ,  $i = 1, \dots, 5$ . The sampling starts at a best-fit model solution from a fixed-D MT inversion using a nonlinear optimization method. Then a burn-in process (typically consisting of several tens of thousands of samples) is applied, after which samples of untempered ( $T = 1$ ) MCMC chains are used to calculate the PPD properties. Convergence of the chain is verified by comparing the conductivity marginal probability profiles computed from two-thirds of the samples with those computed for all samples (no significant difference indicates a convergence, typically reached at several tens of millions of samples). The acceptance rates in the sampling for each step are adjusted to be approximately 30%.

As a general MT inversion, we treat both the model parameterization and the data error model as unknowns supported by the data. The effects of data error correlations on trans-D inversions are examined by comparing the marginal profiles of the conductivity for an inversion considering the AR(1) model for the data errors with those considering a model ignoring data error correlations. The inversion results in terms of the marginal profiles for the interface depth and conductivity are shown in Fig. 1.

In general, when correlated errors are considered in the inversion, the conductivity marginal profiles capture the features of the true model very well, whereas ignoring the correlation in the inversion leads to

**Table 1** The true parameters of the eight-layer model

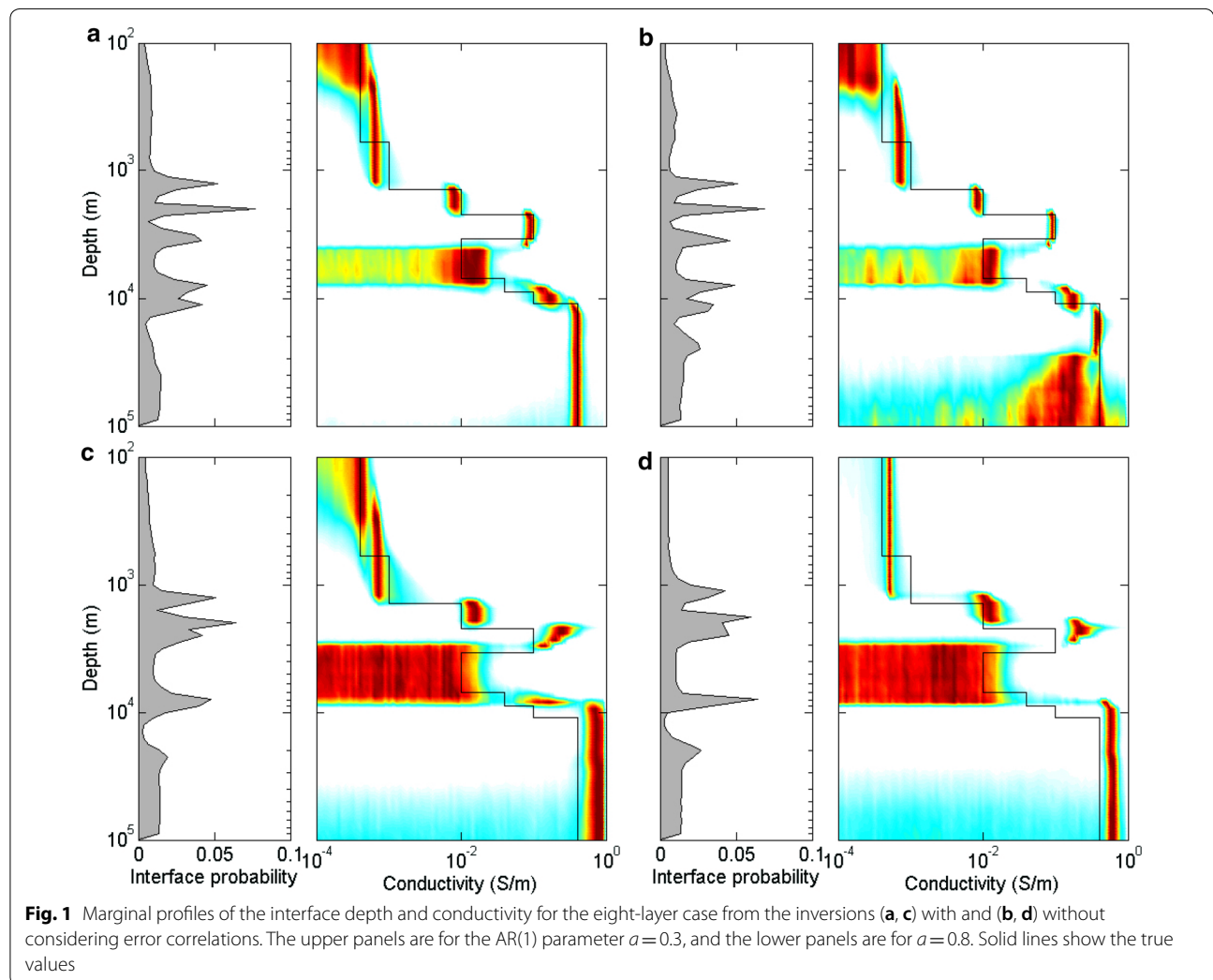
Layer	1	2	3	4	5	6	7	8
Conductivity ( $\text{S m}^{-1}$ )	0.0004	0.001	0.01	0.1	0.01	0.04	0.1	0.4
Thickness (m)	600	800	800	1200	3600	2000	2000	Half-space

structures with a very spurious resolution. In the case with a correlation of 0.3, ignoring the error correlation (Fig. 1b) tends to produce unjustified and slightly varying conductivity structures at great depth. In the case with a correlation of 0.8, ignoring the correlation in the inversion (Fig. 1d) leads to an incorrectly fine resolution for structures in the second layer and in the sixth and seventh layers. When the correlated errors are considered, the conductivity uncertainties are reasonably resolved. Ignoring the correlation in the inversion produces a conductivity marginal profile with slightly smaller uncertainties for most layers, and thus, the profile pretends to exhibit data that contain more information about the model structure than are actually contained therein. The large uncertainty for the fifth layer can be understood as the insensitivity of a relatively low conductivity structure to the data, as the layer sits within a conductor. When the correlated errors are considered, the marginal profile for the

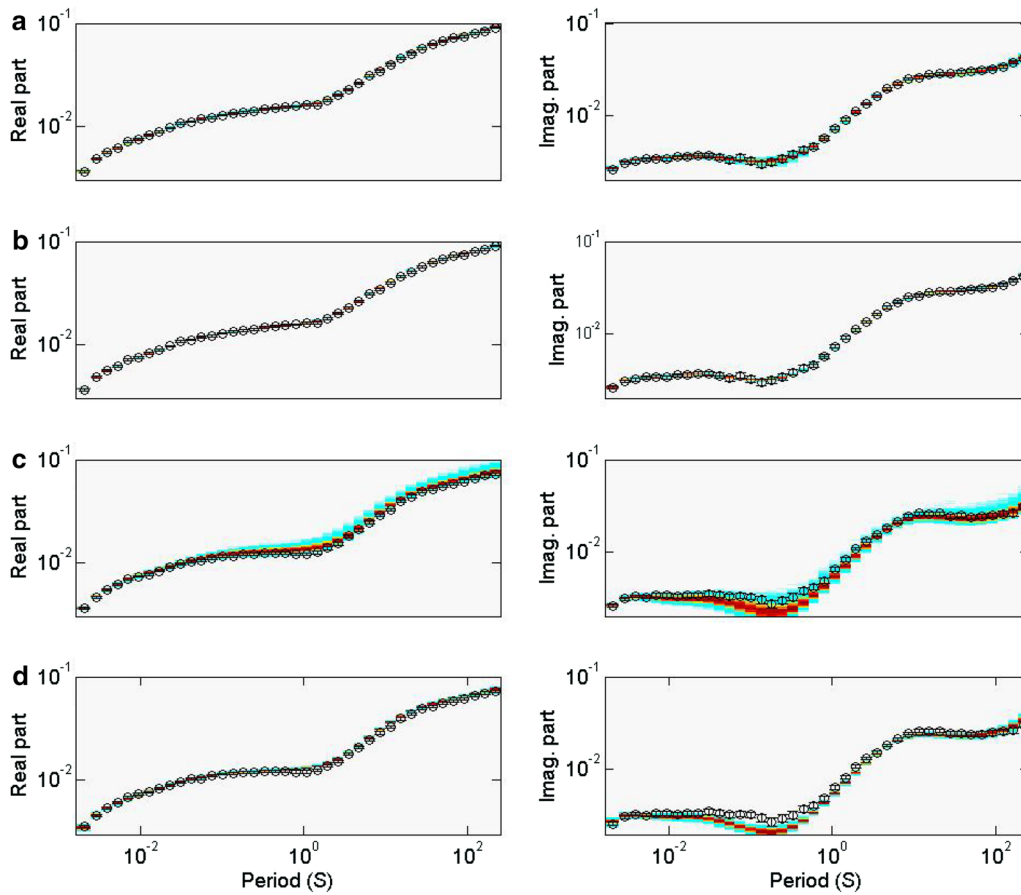
interface depth shows that most interfaces are resolved with reasonable uncertainties and there is some sensitivity to the second layer interface and eighth layer interface, although the general form of the conductivity profile is well resolved. For the inversion that does not consider error correlations, the probabilities of most interfaces present results similar to the conductivity marginal profiles.

The predicted data density plots calculated using the model ensemble from the trans-D PPD are shown in Fig. 2. When correlated errors are ignored, the predicted data uncertainties for both the real and the imaginary parts of the impedance data are smaller than those consider error correlation in the inversion.

The effect of including an AR model in the inversion is examined by posterior analysis of the normalized total residual, defined by



**Fig. 1** Marginal profiles of the interface depth and conductivity for the eight-layer case from the inversions (a, c) with and (b, d) without considering error correlations. The upper panels are for the AR(1) parameter  $a=0.3$ , and the lower panels are for  $a=0.8$ . Solid lines show the true values



**Fig. 2** Densities of the predicted data for the eight-layer model from the posterior samples for the inversions (a, c) with and (b, d) without considering error correlations. The upper panels are for the AR(1) parameter  $a=0.3$ , and the lower panels are for  $a=0.8$ . Circles with error bars represent the noisy synthetic data

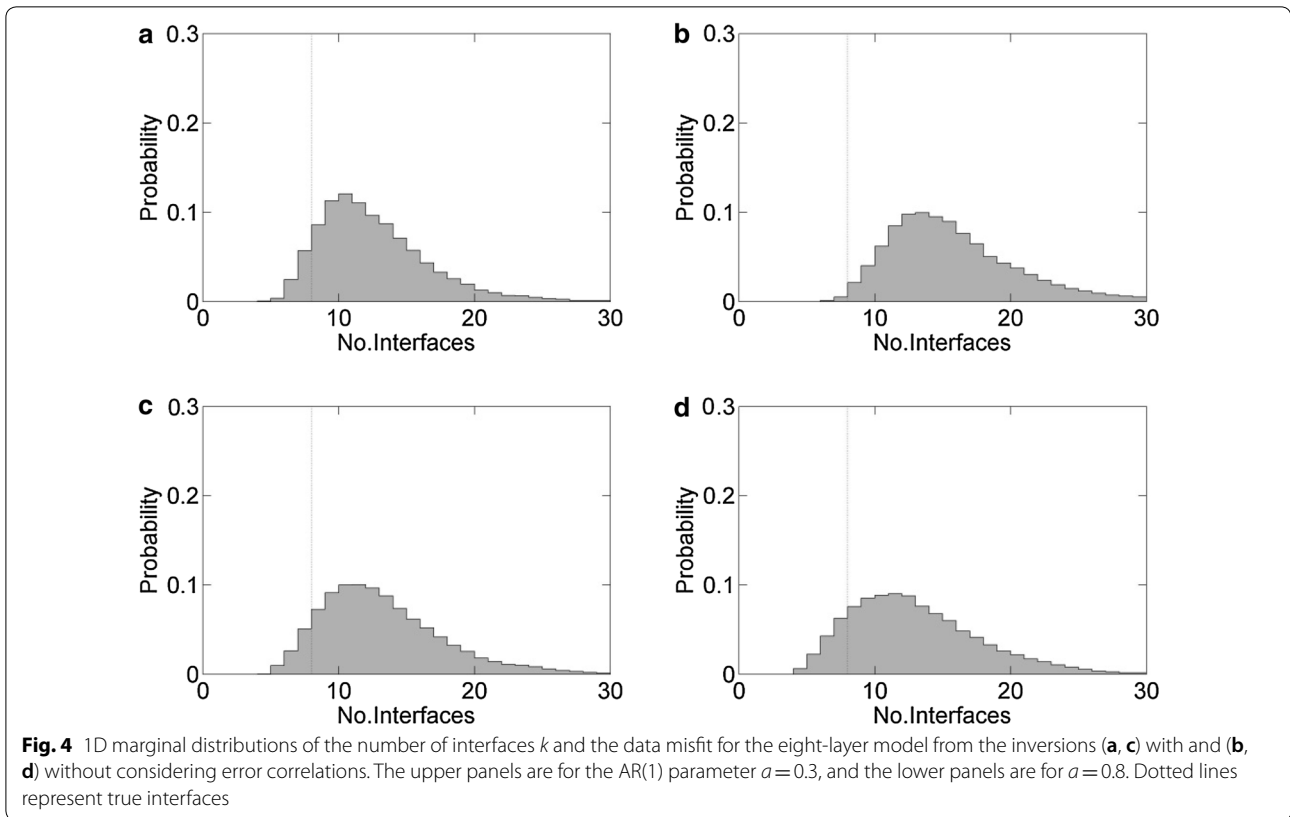
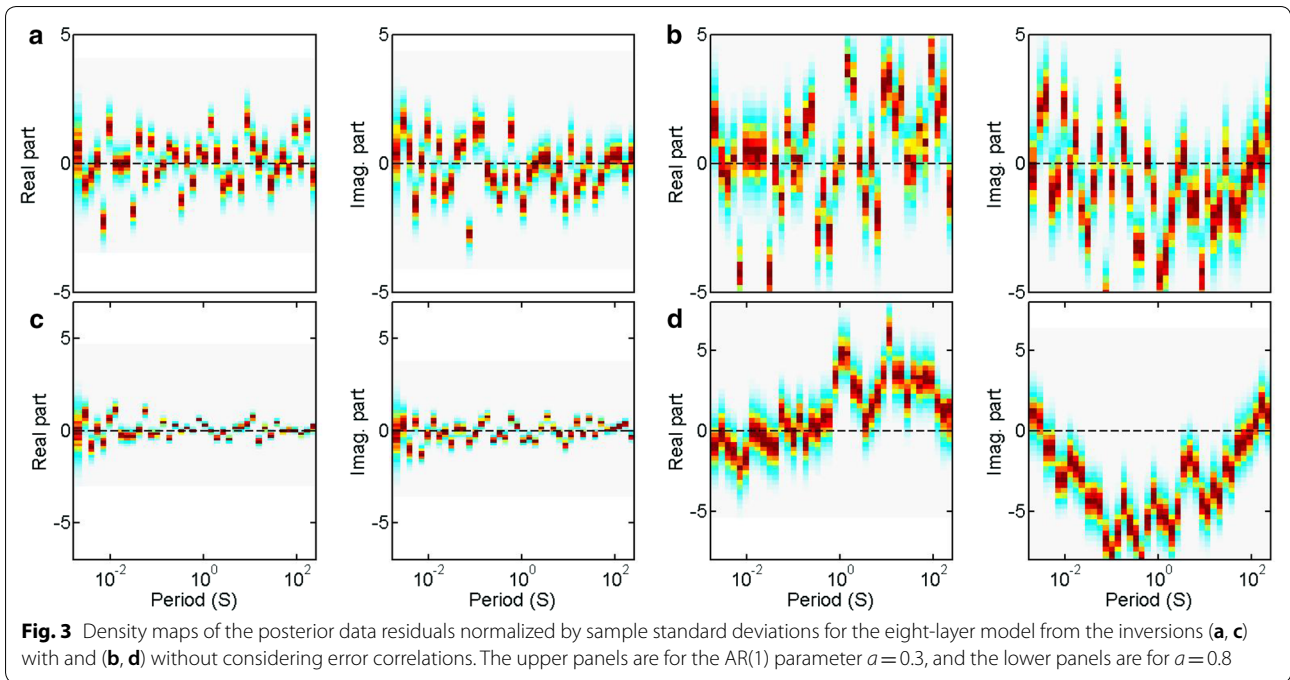
$$r_i = \frac{d_i - d_i(k, \mathbf{m}_k) - d_i(a)}{s_i} \tag{8}$$

Density maps of the data residuals normalized by the sample standard deviations based on the posterior are shown in Fig. 3. In the cases with correlations of 0.3 and 0.8, when correlated errors are ignored, the data residuals for both the real and the imaginary parts of the impedance data are much larger than those considering error correlations in the inversion. The test runs for the randomness of the residuals indicates that, in the case with a correlation of 0.3, 93% of the real residuals and 99% of the imaginary residuals support the null hypothesis (by applying the AR(1) model in the inversion), compared to 49% and 78%, respectively, for the inversion without considering the effect of error correlations; in the case with a correlation of 0.8, 98% and 99% of the real and imaginary residuals, respectively, support the null hypothesis for the inversions considering error correlations compared to 0%

and 0%, respectively, for the inversions without including the correlation.

The marginal distributions for the number of interfaces  $k$  in the cases with correlations of 0.3 and 0.8 are shown in Fig. 4, which shows that the probability distribution for the index  $k$  is not well constrained by the data with a heavy right tail. For both cases, the minimum number of interfaces is 4, and the probability peaks at approximately 10, which is close to the number of interfaces of the true model. When correlated errors are ignored, the uncertainties in the number of interfaces of the model (8 layers) are larger than those in inversions considering error correlations for the cases with correlations of both 0.3 and 0.8.

A summary of the inversion results for the data errors at  $T=1$  is shown in Fig. 5. The inversion results for the AR(1) parameters peak at approximately 0.5 and 0.75 for the cases with correlations of 0.3 and 0.8, respectively, which both slightly deviate from the true values. Using the AR(1) parameters to represent the noise



in the inversion leads to a high percentage of AR(1)-on rjMCMC samples for both cases, and the  $t$  test for the AR(1) parameter produces  $t_{\text{amax}} = 4.21$  in the case with

a 0.3 correlation and  $t_{\text{amax}} = 20.93$  in the case with a 0.8 correlation, indicating that the AR(1) parameter is significant. Hence, by including the AR(1) parameters in

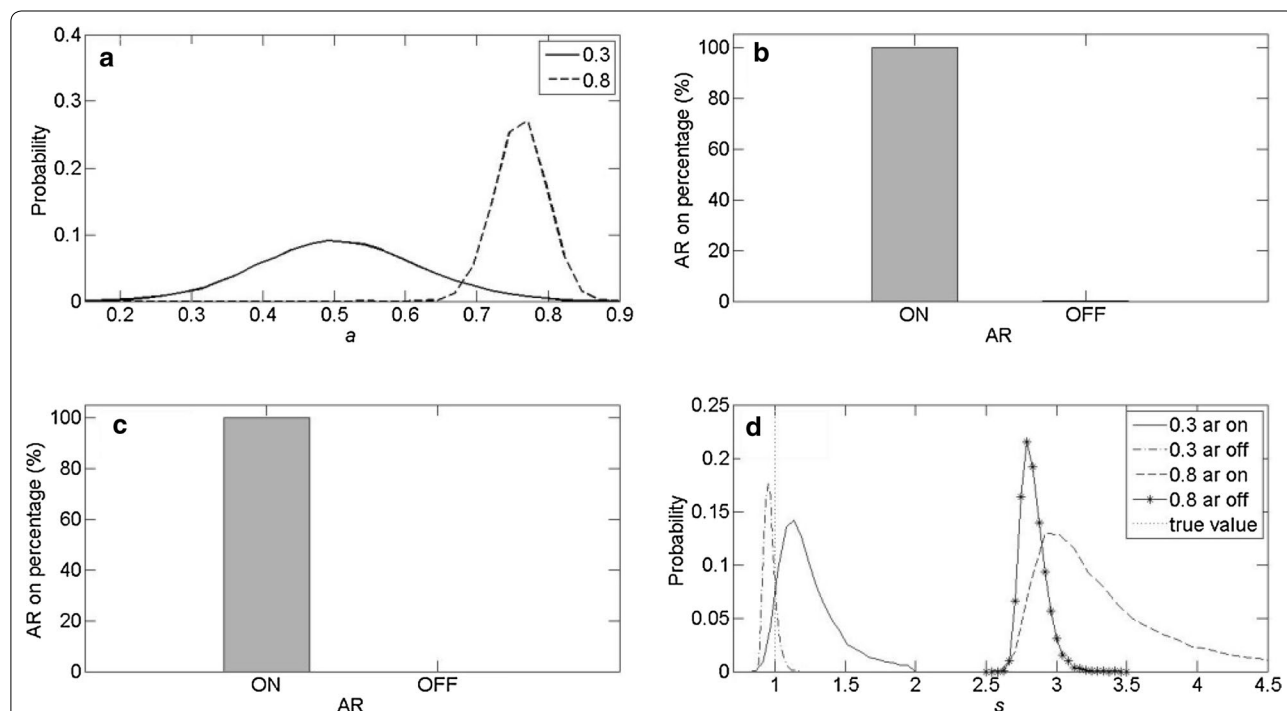
the inversion, the recovered error scale hyperparameter  $s$  tends to be larger than both the true value and the values obtained by ignoring the error correlation.

### Real application

This section considers the measured COPROD1 MT data (Jones and Hutton 1979), which have been used widely as a 1D inversion example (e.g., Constable et al. 1987; Guo et al. 2011; Guo 2011). These data are composed of MT impedance responses at 15 periods from 28.5 to 1960.7 s. The standard data errors given by Constable et al. (1987) are used here. The model is taken to represent the log conductivity to determine the structure over several orders of magnitude. Bounded uniform priors are applied over  $[10^{-5}, 1] \text{ S m}^{-1}$  for the conductivity, and the maximum penetrating depth is set to  $z_b=1000 \text{ km}$ . The number of interfaces  $k$  is taken to be a uniform distribution over the range  $[1, 30]$ , the prior information for AR(1) is set to be uniform on  $[-0.5, 1]$ , and the state space of the AR error process is given to be  $[0, 1]$  (i.e., AR(0) and AR(1)). A total of 5 chains are applied in the parallel tempering scheme, where the temperatures  $T_i = 1.5^{(i-1)}$ ,  $i = 1, \dots, 5$  are used in the synthetic case.

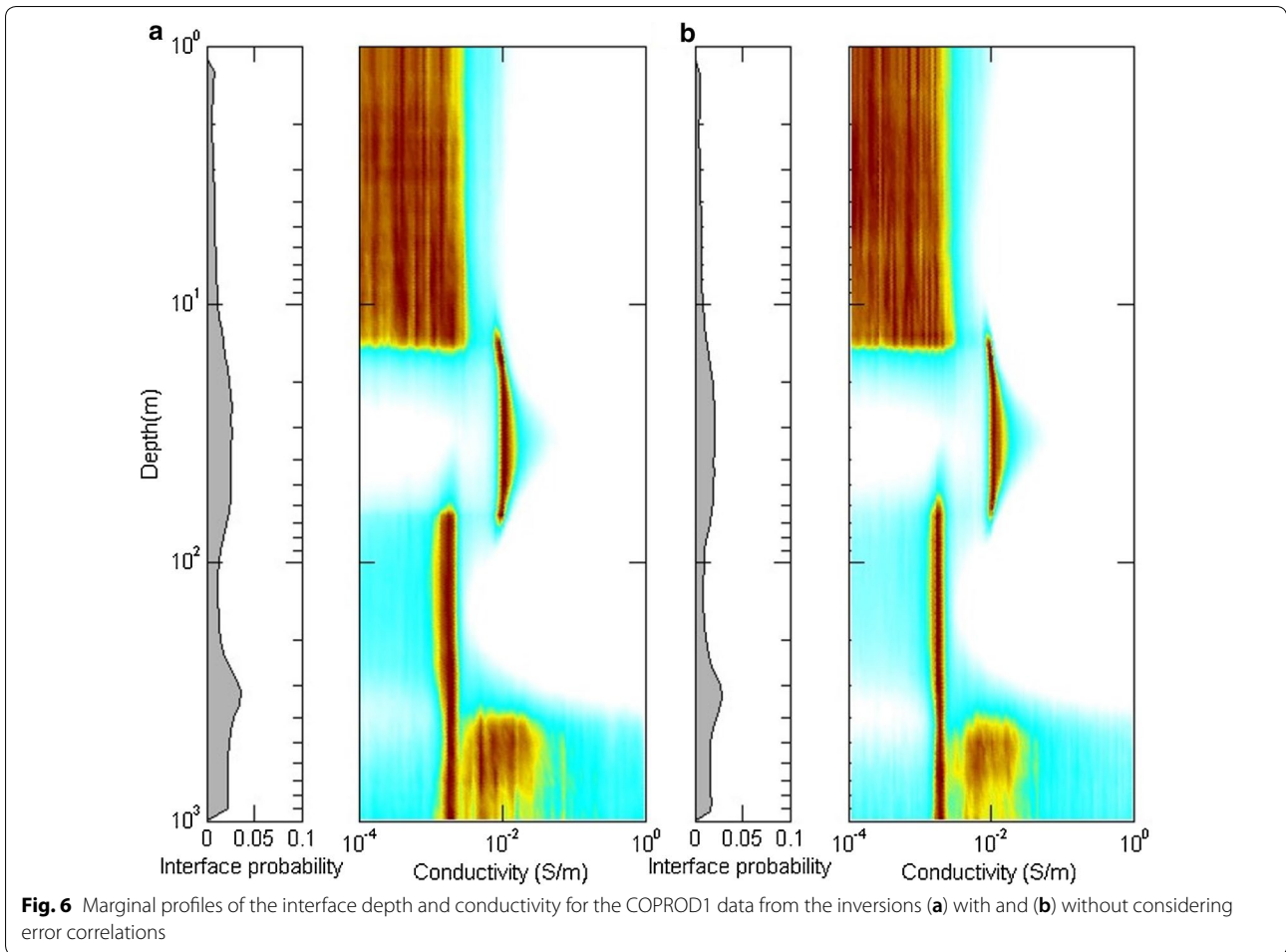
First, the results of the trans-D inversion are considered in terms of the interface probability densities and

marginal profiles for the conductivity, as shown in Fig. 6. The interface marginal profiles show poor resolutions at the layer interfaces, probably due to the diffusive nature of electromagnetic waves. However, the marginal profile of the conductivity illustrates a general four layer model; large interface-depth uncertainties are observed for the first and second recovered layers, and thus, the interface profile density is indistinguishable between these layers. The model indicated in the conductivity marginal profile consists of an upper resistive layer approximately 10 km thick above a relatively high conductivity layer (approximately 10–60 km) and a resistive layer. Below approximately 400 km, the probability density becomes multimodal with two possibilities: either the resistive layer extends to 1000 km or there is a transition to a somewhat more conductive half-space, the latter is consistent with the results inferred by Guo et al. (2011). The conductivity uncertainties are quite large over the uppermost ~10 km, probably caused by the poor high-frequency coverage for the COPROD1 data. The data resolve the structure at an intermediate depth with relatively small uncertainties between ~10 km and ~400 km but poorly constrain the deeper structure. The results for the inversions with and without considering error correlations are almost identical.



**Fig. 5** Posterior error results for the eight-layer model. **a, d** are the 1D marginal distributions for the first-order autoregressive parameter  $a$  and error scale hyperparameter  $s$ , respectively. Solid line and dashed lines show the results for  $a=0.3$  with and without considering the error correlation, respectively, and the dashed-dotted lines and lines with asterisks are for  $a=0.8$ . **b, c** indicate portions of AR(1)-on and AR(1)-off rjMCMC samples for  $a=0.3$  and  $a=0.8$ , respectively. Dotted lines represent true values





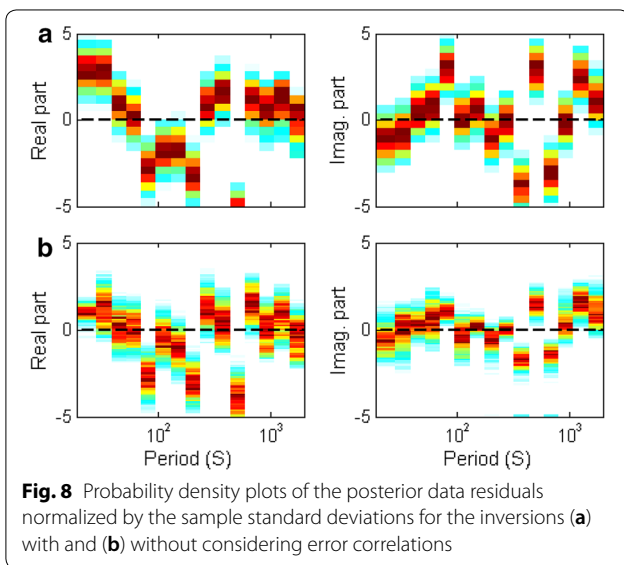
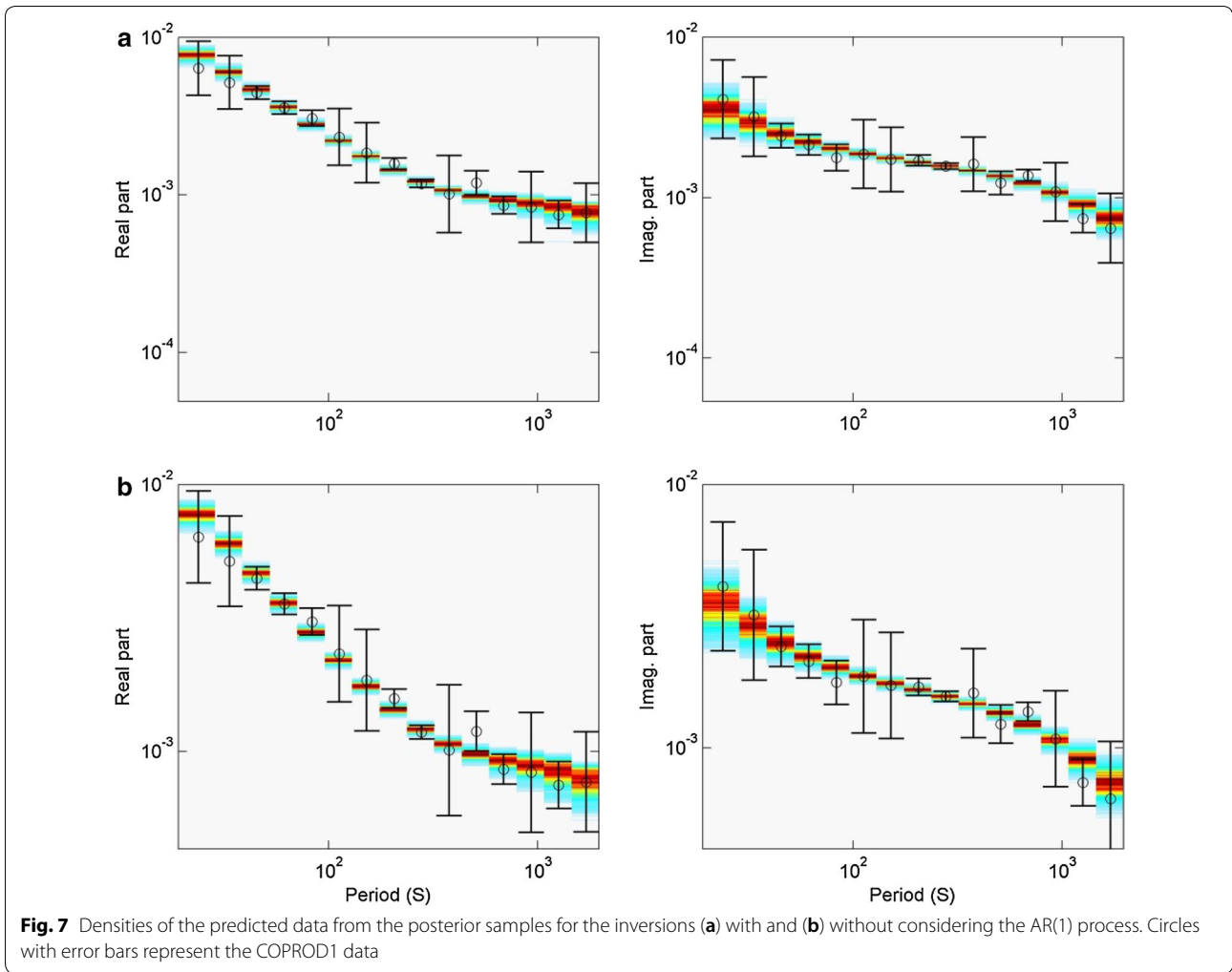
The predicted data density plots calculated using the model ensemble from the trans-D PPD are shown in Fig. 7. When correlated errors are ignored, the predicted data uncertainties for both the real and the imaginary parts of the impedance data are almost the same as those of the inversions that consider error correlations.

The density plots of the data residuals normalized by the sample standard deviations for the model ensemble from the posterior are shown in Fig. 8. When correlated errors are ignored, the data residuals for both the real and the imaginary parts of the impedance data are slightly larger than those of the inversions that consider error correlations. The test runs for the randomness of the residuals indicate that 97% of the real residuals and 99% of the imaginary residuals support the null hypothesis (after applying the AR(1) model in the inversion), compared to 98% and 99%, respectively, for the inversion without considering the effect of error correlation. This indicates that the data noise follows a Gaussian

distribution without significant correlations between each data point, which may explain why the results in Figs. 6, 7 for the inversions with and without considering error correlations are almost the same.

Figure 9 shows the marginal density distribution for the number of interfaces  $k$  and data misfit. The peak of the marginal density for index  $k$  is at approximately 6, and the uncertainty is large due to the insensitivity of the MT data to the interface, as observed in the literature for MT problems (Xiang et al. 2018).

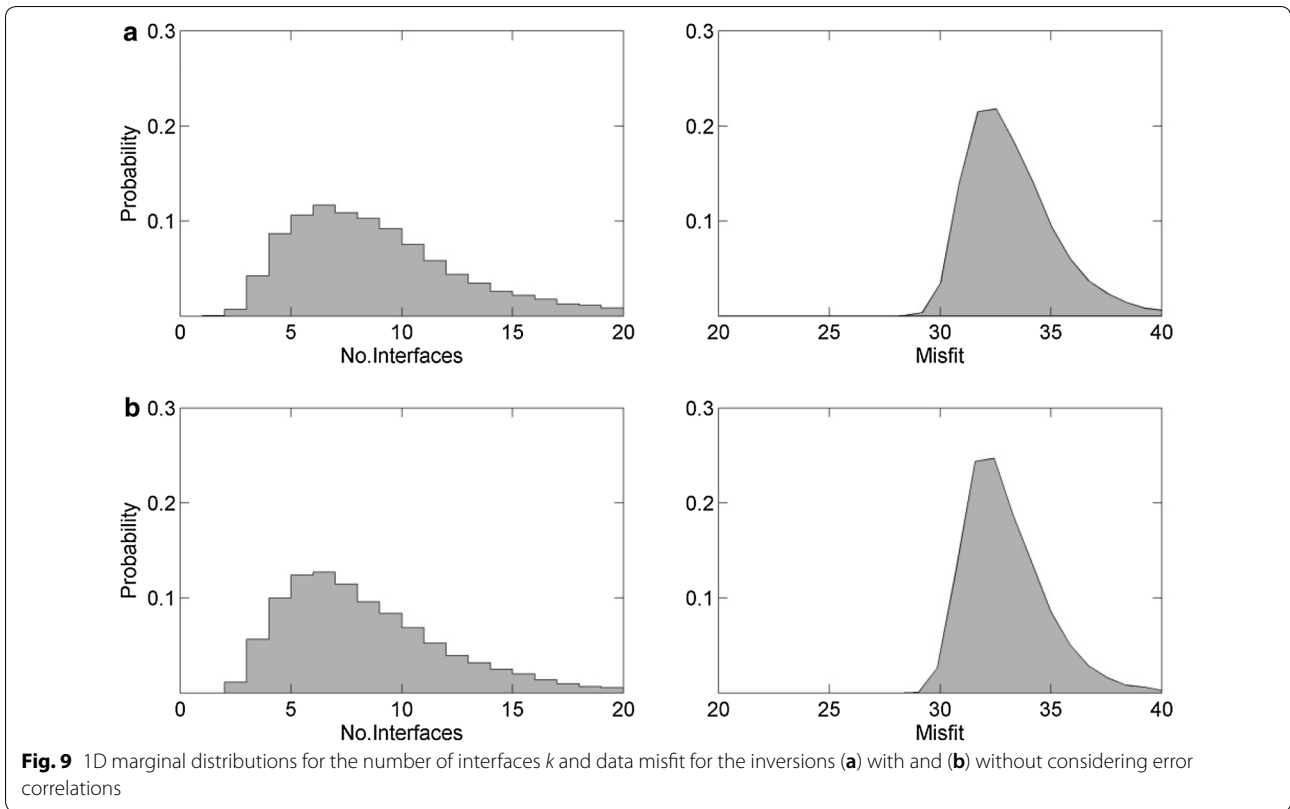
A summary of the inversion results for the data errors at  $T=1$  is shown in Fig. 10. The inversion results for the AR(1) parameter peak at approximately 0, again indicating that the data errors are mostly independent. The  $t$  test for the AR(1) parameter produces  $t_{\text{amax}} = -0.19$  indicating that the AR(1) parameter is insignificant. Using the AR(1) process to represent the noise in the inversion leads to a relatively low percentage of AR(1)-on rjMCMC samples, probably due to the insignificant



error correlation present in the data. The error scale hyperparameter  $s$  is narrowly constrained, indicating that the data are sensitive to  $s$  for both the inversions with and the inversions without considering error correlations. Their marginal distributions are almost the same with peaks at approximately 0.76.

### Conclusions

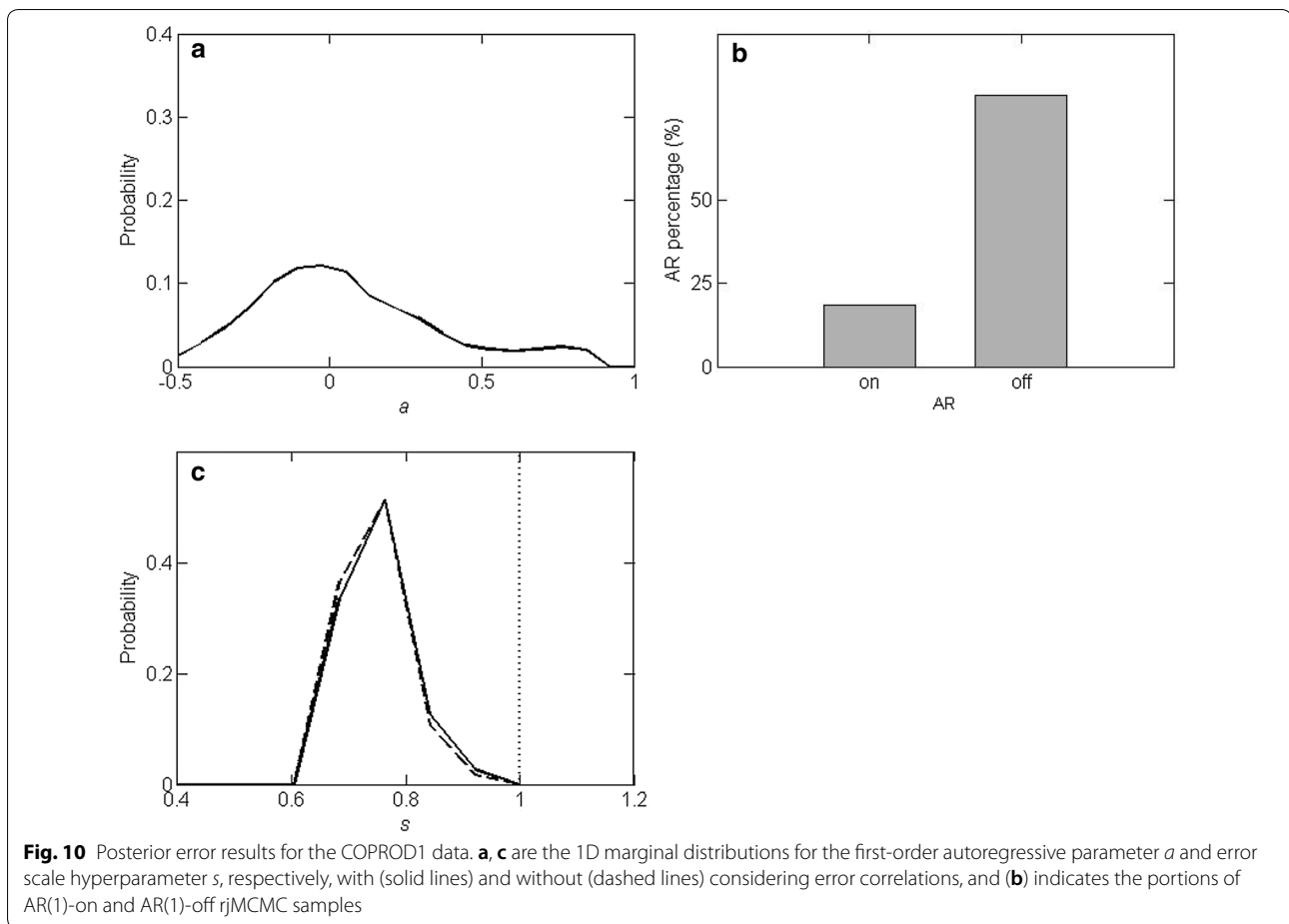
In this paper, a trans-D Bayesian method is applied to investigate the data error correlation effect on the inversion of MT data for subsurface conductivity structures. To avoid unnecessary complexity in the error model, the error process is parameterized by the first-order AR process. In hierarchic Bayesian inversions, the AR parameters are treated as unknowns and are included as hyperparameters that are sampled by the trans-D algorithm over the prior subspace (e.g., AR(0) and AR(1) are used in this paper). The limited knowledge on the data



error parameterization can be quantified in the posterior model uncertainty estimates. Ignoring the serial correlation certainly affects the model uncertainty estimation and model estimates in an unknown way since additional model structures are required to fit the correlated part of the data. Thus, inversions with AR(1) models and without considering error correlations are compared to examine the effect of neglecting the correlation on the inversion results.

The above inversion approaches are applied to an eight-layer synthetic model and to the measured COPROD1 MT dataset. In both cases, the data can resolve both the model parameterization and the data error parameterization with reasonable uncertainties, indicating that the data contain enough information

about the model and data error parameterizations as well as the model information. For the synthetic cases, ignoring the data error correlations leads to a general underestimation of the model uncertainties and overfitting of the data. When ignoring the serial correlations in data errors, the inversion for the case with a correlation of 0.3 tends to produce deeper small-scale structures to account for error correlations; for the synthetic data with a correlation of 0.8, the inversion results from the case neglecting the correlation lose resolution for the second layer, and the sixth and seventh layers of the true model. For the COPROD1 MT data, the recovered AR(1) parameter peaks at 0, indicating that the data error correlation is essentially weak. The inversion results for the data with and without considering error



correlations are almost identical except for a slight underestimation of the model uncertainty when simply ignoring the correlation effect.

#### Abbreviations

MT: magnetotelluric; trans-D: trans-dimensional; AR(1): first-order autoregressive; PPD: posterior probability density; rjMCMC: reversible-jump Markov chain Monte Carlo; AR: autoregressive; PC: principal component.

#### Acknowledgements

Great thanks to Professor Stan Dosso from University of Victoria, Canada, who reviewed the manuscript and improved this paper significantly.

#### Authors' contributions

RG developed the ideas and contributed to the final version of the manuscript. JL and YS supervised the project. LL and RL ran the examples, plotted the results and wrote the draft. All authors read and approved the final manuscript.

#### Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 41674079, 41674080, 41404061 and 41404042).

#### Availability of data and materials

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup> School of Geosciences and Info-physics, Central South University, Lushan Road South 932, Changsha 410083, China. <sup>2</sup> Human Key Laboratory of Nonferrous Resources and Geological Hazards Exploration, Central South University, Lushan Road South 932, Changsha 410083, China. <sup>3</sup> Key Laboratory of Metallogenic Prediction of Nonferrous Metals, Ministry of Education, Central South University, Lushan Road South 932, Changsha 410083, China.

Received: 23 April 2019 Accepted: 25 November 2019

Published online: 05 December 2019

#### References

- Agostinetti NP, Malinverno A (2010) Receiver function inversion by trans-dimensional Monte Carlo sampling. *Geophys J Int* 181:858–872
- Backus GE (1970) Inference from inadequate and inaccurate data. I. *Proc Natl Acad Sci USA* 65:1–7
- Becken M, Ritter O (2012) Magnetotelluric studies at the San Andreas Fault Zone: implications for the role of fluids. *Surv Geophys* 33(1):65–105
- Blatter D, Key K, Ray A, Gustafson C, Evans R (2019) Bayesian joint inversion of controlled source electromagnetic and magnetotelluric data to image freshwater aquifer offshore New Jersey. *Geophys J Int* 218(3):1822–1837

- 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* 117:B02301
- Constable SC, Parker RL, Constable CG (1987) Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics* 52:289–300
- Dettmer J, Dosso SE (2012) Trans-dimensional matched-field geoaoustic inversion with hierarchical error models and interacting Markov chains. *J Acoust Soc Am* 132:2239–2250
- Dettmer J, Molnar S, Steininger G, Dosso SE, Cassidy JF (2012) Trans-dimensional inversion of microtremor array dispersion data with hierarchical autoregressive error models. *Geophys J Int* 188:719–734
- Dosso SE, Dettmer J (2011) Bayesian matched-field geoaoustic inversion. *Inverse Probl* 27:055009
- Dosso SE, Nielsen PL, Wilmut MJ (2006) Data error covariance in matched-field geoaoustic inversion. *J Acoust Soc Am* 119:208–219
- Dosso SE, Holland CW, Sambridge M (2012) Parallel tempering in strongly nonlinear geoaoustic inversion. *J Acoust Soc Am* 132:330–340
- Dosso SE, Dettmer J, Steininger G, Holland CW (2014) Efficient trans-dimensional Bayesian inversion for geoaoustic profile estimation. *Inverse Probl* 30:114018
- Egbert G (1997) Robust multiple-station magnetotelluric data processing. *Geophys J Int* 130:475–496
- Eisel M, Egbert G (2001) On the stability of magnetotelluric transfer function estimates and the reliability of their variances. *Geophys J Int* 144:65–82
- Espurt N, Hippolyte JC, Kaymakci N, Sangu E (2014) Lithospheric structural control on inversion of the southern margin of the Black Sea Basin, Central Pontides, Turkey. *Lithosphere* 6(1):26–34
- Gamble T, Goubau W, Clarke J (1979) Magnetotellurics with a remote reference. *Geophysics* 44:53–68
- Gehrmann RA, Dettmer J, Schwalenberg K, Engels M, Dosso SE, Özmaral A (2015) Trans-dimensional Bayesian inversion of controlled-source electromagnetic data in the German North Sea. *Geophys Prospect* 63:1314–1333
- Goto TN, Wada Y, Oshiman N, Sumitomo N (2005) Resistivity structure of a seismic gap along the Atotsugawa Fault, Japan. *Phys Earth Planet Inter* 148(1):55–72
- Goubau W, Gamble T, Clarke J (1978) Magnetotelluric data analysis: removal of bias. *Geophysics* 43:1157–1166
- Goubau WM, Maxton PM, Koch RH, Clarke J (1984) Noise correlation lengths in remote reference magnetotellurics. *Geophysics* 49:433–438
- Grayver AV (2015) Parallel three-dimensional magnetotelluric inversion using adaptive finite-element method. Part I: theory and synthetic study. *Geophys J Int* 202(1):584–603
- Green PJ (1995) Reversible jump MCMC computation and Bayesian model determination. *Biometrika* 82:711–732
- Guo R (2011) Nonlinear analysis and uncertainty estimate based on Bayesian MT inversion. Ph.D. thesis at Central South University
- Guo R, Dosso SE, Liu J, Dettmer J, Tong X (2011) Non-linearity in Bayesian 1-D magnetotelluric inversion. *Geophys J Int* 185:663–675
- Guo R, Dosso SE, Liu J, Liu Z, Tong X (2014) Frequency-and spatial-correlated noise on layered magnetotelluric inversion. *Geophys J Int* 199:1205–1213
- He Z, Hu Z, Luo W, Wang C (2010) Case history, mapping reservoirs based on resistivity and induced polarization derived from continuous 3D magnetotelluric profiling: case study from Qaidam Basin, China. *Geophysics* 75:B25–B33
- He L, Chen L, Xi X, Zhao X, Chen R, Yao H (2016) Mapping the geothermal system using AMT and MT in the Mapamyum (QP) field, Lake Manasarovar, Southwestern Tibet. *Energies* 9:85
- Heise W, Caldwell TG, Bibby HM, Bannister SC (2008) Three-dimensional modelling of magnetotelluric data from the Rotokawa geothermal field, Taupo Volcanic Zone, New Zealand. *Geophys J Int* 173:740–750
- Huang CF, Gerstoft P, Hodgkiss WS (2007) On the effect of error correlation on matched-field geoaoustic inversion. *JASA Express Lett* 121:64–70
- Jasra A, Stephens DA, Holmes CC (2007) Population-based reversible jump Markov chain Monte Carlo. *Biometrika* 94:787–807
- Jones AG, Hutton R (1979) A multi-station magnetotelluric study in southern Scotland, I., Fieldwork, data analysis and results. *Geophys J Int* 56:329–349
- Lee SK, Kim HJ, Song Y, Lee CK (2009) MT2DInvMatlab—A program in MATLAB and FORTRAN for two-dimensional magnetotelluric inversion. *Comput Geosci* 35:1722–1734
- Malinverno A (2002) Parsimonious Bayesian Markov chain Monte Carlo inversion in a non linear geophysical problem. *Geophys J Int* 151:675–688
- Mandolesi E, Ogaya X, Campaña J, Agostinetti NP (2018) A reversible-jump Markov chain Monte Carlo algorithm for 1D inversion of magnetotelluric data. *Comput Geosci* 113:94–105
- Oldenburg DW (1983) Funnel functions in linear and nonlinear appraisal. *J Geophys Res Solid Earth* 88(B9):7387–7398
- Ray A, Alumbaugh DL, Hoversten GM, Key K (2013a) Robust and accelerated Bayesian inversion of marine controlled-source electromagnetic data using parallel tempering. *Geophysics* 78:E271–E280
- Ray A, Key K, Bodin T (2013b) Hierarchical Bayesian inversion of marine CSEM data over the Scarborough gas field—a lesson in correlated noise. In: 2013 SEG annual meeting, society of exploration geophysicists. pp 723–727
- Ritter O, Junge A, Dawes GJ (1998) New equipment and processing for magnetotelluric remote reference observations. *Geophys J Int* 132:535–548
- Sambridge M (1999) Geophysical inversion with a neighbourhood algorithm—i. Searching a parameter space. *Geophys J Roy Astron Soc* 139(2):479–494
- Sambridge M (2013) A parallel tempering algorithm for probabilistic sampling and multimodal optimization. *Geophys J Int* 196:357–374
- Sambridge M, Gallagher K, Jackson A, Rickwood P (2006) Trans-dimensional inverse problems, model comparison and the evidence. *Geophys J Int* 167:528–542
- Steininger G, Dettmer J, Dosso SE, Holland CW (2013) Trans-dimensional joint inversion of seabed scattering and reflection data. *J Acoust Soc Am* 133:1347–1357
- Usui Y, Ogawa Y, Aizawa K, Kanda W, Hashimoto T, Koyama T, Kagiya T (2016) Three-dimensional resistivity structure of Asama Volcano revealed by data-space magnetotelluric inversion using unstructured tetrahedral elements. *Geophys J Int* 208(3):1359–1372
- Warner GA, Dosso SE, Dettmer J, Hannay DE (2015) Bayesian environmental inversion of airgun modal dispersion using a single hydrophone in the Chukchi Sea. *J Acoust Soc Am* 137(6):3009–3023
- Weckmann U, Magunia A, Ritter O (2005) Effective noise separation for magnetotelluric single site data processing using a frequency domain selection scheme. *Geophys J Int* 169:635–652
- Wheelock BD (2012) Electromagnetic imaging of the crust and upper mantle across the continental margin in Central California. University of California, San Diego
- Xiang E, Guo R, Dosso SE, Liu J, Dong H, Ren Z (2018) Efficient hierarchical trans-dimensional Bayesian inversion of magnetotelluric data. *Geophys J Int* 213(3):1751–1767
- Yoshimura R, Ogawa Y, Yukutake Y, Kanda W, Komori S, Hase H, Kamo M (2018) Resistivity characterisation of Hakone volcano, Central Japan, by three-dimensional magnetotelluric inversion. *Earth Planets Space* 70:1–10. <https://doi.org/10.1186/s40623-018-0848-y>
- Zhang P, Chouteau M (1992) The use of magnetotellurics for mineral exploration: an experiment in the Chibougamau region of Quebec. *Can J Earth Sci* 29:621–635

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.