

# JGR Solid Earth

## RESEARCH ARTICLE

10.1029/2021JB021962

### Key Points:

- We present a novel strategy to invert 3D magnetotelluric (MT) data together with other data sets in a fully probabilistic manner
- Perform the first joint probabilistic inversion of 3D MT and surface-wave data for imaging the conductivity distribution in the lithosphere
- Demonstrate the applicability of our approach to include 3D MT data into joint probabilistic inversions for imaging the Earth's interior

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

M. C. Manassero and J. C. Afonso,  
[maria-constanza.manassero@mq.edu.au](mailto:maria-constanza.manassero@mq.edu.au);  
[juan.afonso@mq.edu.au](mailto:juan.afonso@mq.edu.au)

### Citation:

Manassero, M. C., Afonso, J. C., Zyserman, F. I., Jones, A. G., Zlotnik, S., & Fomin, I. (2021). A reduced order approach for probabilistic inversions of 3D magnetotelluric data II: Joint inversion of MT and surface-wave data. *Journal of Geophysical Research: Solid Earth*, 126, e2021JB021962. <https://doi.org/10.1029/2021JB021962>

Received 26 FEB 2021







Accepted 8 NOV 2021

### Author Contributions:

**Conceptualization:** M. C. Manassero, J. C. Afonso, F. I. Zyserman, S. Zlotnik  
**Formal analysis:** M. C. Manassero  
**Funding acquisition:** J. C. Afonso  
**Investigation:** M. C. Manassero  
**Methodology:** M. C. Manassero  
**Project Administration:** J. C. Afonso  
**Software:** M. C. Manassero, F. I. Zyserman, I. Fomin  
**Supervision:** J. C. Afonso  
**Validation:** M. C. Manassero, J. C. Afonso, F. I. Zyserman, A. G. Jones  
**Visualization:** M. C. Manassero

© 2021. American Geophysical Union.  
 All Rights Reserved.

## A Reduced Order Approach for Probabilistic Inversions of 3D Magnetotelluric Data II: Joint Inversion of MT and Surface-Wave Data

M. C. Manassero<sup>1</sup> , J. C. Afonso<sup>1</sup> , F. I. Zyserman<sup>2</sup> , A. G. Jones<sup>3,4,5</sup> , S. Zlotnik<sup>6</sup> , and I. Fomin<sup>1</sup> 

<sup>1</sup>Department of Earth and Environmental Sciences, Australian Research Council Centre of Excellence for Core to Crust Fluid Systems/GEMOC, Macquarie University, Sydney, NSW, Australia, <sup>2</sup>CONICET, Facultad de Ciencias Astronómicas y Geofísicas, Universidad de La Plata, La Plata, Argentina, <sup>3</sup>Formerly Dublin Institute for Advanced Studies, Dublin, Ireland, <sup>4</sup>University of Western Australia, Centre for Exploration Targeting, Perth, WA, Australia, <sup>5</sup>Now at Complete MT Solutions Inc., Ottawa, ON, Canada, <sup>6</sup>Laboratori de Càlcul Numèric, Escola Tècnica Superior d'Enginyers de Camins, Canals i Ports, Universitat Politècnica de Catalunya, Barcelona, Spain

**Abstract** Joint probabilistic inversions of magnetotelluric (MT) and seismic data have great potential for imaging the thermochemical structure of the lithosphere as well as mapping fluid/melt pathways and regions of mantle metasomatism. In this contribution, we present a novel probabilistic (Bayesian) joint inversion scheme for 3D MT and surface-wave dispersion data particularly designed for large-scale lithospheric studies. The approach makes use of a recently developed strategy for fast solutions of the 3D MT forward problem (Manassero et al., 2020, <https://doi.org/10.1093/gji/ggaa415>) and combines it with adaptive Markov chain Monte Carlo (MCMC) algorithms and parallel-in-parallel strategies to achieve extremely efficient simulations. To demonstrate the feasibility, benefits and performance of our joint inversion method for imaging the temperature and conductivity structures of the lithosphere, we apply it to two numerical examples of increasing complexity. The inversion approach presented here is timely and will be useful in the joint analysis of MT and surface wave data that are being collected in many parts of the world. This approach also opens up new avenues for the study of trans-lithospheric and trans-crustal magmatic systems, the detection of metasomatized mantle, and the incorporation of MT into multi-observable inversions for the physical state of the Earth's interior.

**Plain Language Summary** The joint analysis of two or more geophysical data sets is becoming common practice for imaging the Earth's interior. This approach has great potential when the different data sets offer complementary sensitivities to the properties of interest. Such is the case of combining MT, that is, magnetotelluric (a technique based on observed electric and magnetic fields) and seismic data. In this context, probabilistic methods offer a powerful and general platform for the joint analysis of these data sets and their uncertainties. However, the inclusion of MT data into 3D probabilistic schemes has so far not been possible due to the significant computational cost of the MT problem. In this contribution we present a novel probabilistic joint analysis scheme for MT and surface-wave data which is based on a recently developed strategy to achieve fast and efficient solutions to the MT problem (Manassero et al., 2020, <https://doi.org/10.1093/gji/ggaa415>). We demonstrate the feasibility, benefits and performance of our approach for whole-lithosphere and upper mantle structures up to ~400 km depth. The approach presented here will be useful in the joint analysis of MT and seismic data that are being collected in many parts of the world.

## 1. Introduction

Joint inversions of two or more geophysical data sets are becoming common practice for imaging the Earth's interior and elucidating the physical state of the planet. When the inverted data sets have complementary sensitivities to the properties of interest, joint inversions significantly reduce the ambiguity inherent in single-data set inversions, achieve more stable solutions, increase identifiability of features and enhance model resolution. Perhaps more importantly, certain properties of the Earth's interior can only be revealed by combining observations from different techniques. An example is the bulk composition of the lithospheric mantle, which requires independent constraints on the bulk density (e.g., from gravity data sets) and shear-wave velocity (e.g., from surface-wave data). Recent discussions on the benefits and limitations of joint approaches for imaging the structure of the lithosphere and upper mantle can be found in, for example, Khan et al. (2006); Afonso, Fullea, Griffin,

Writing – original draft: M. C. Manassero  
 Writing – review & editing: M. C. Manassero, J. C. Afonso, F. I. Zyserman, A. G. Jones

et al. (2013); Afonso, Moorkamp, & Fullea (2016), and Moorkamp (2017). The joint inversion of magnetotelluric (MT) with seismic data (e.g., Afonso, Rawlinson, et al., 2016; Bennington et al., 2015; Gallardo & Meju, 2007; Jegen et al., 2009; Jones et al., 2017; Khan et al., 2006; Moorkamp et al., 2007, 2010; Vozar et al., 2014) is of particular interest as they offer complementary sensitivities to temperature, composition, and fluid/melt content that are impossible to obtain with other data sets (e.g., Afonso, Moorkamp, & Fullea, 2016; Afonso, Rawlinson, et al., 2016; Gallardo & Meju, 2007; Jones et al., 2009; Moorkamp et al., 2007, 2010; Selway et al., 2019).

In the context of whole-lithosphere structure, both seismic (or seismic plus gravity) and MT data can be used to put constraints on the background (or regional) thermal, and mineralogical structure (e.g., Afonso, Moorkamp, & Fullea, 2016; Afonso, Rawlinson, et al., 2016; Jones et al., 2009; Karato & Wang, 2013), but only MT is exponentially sensitive to temperature, and strongly sensitive to hydrogen content, minor conductive phases, and/or small volumes of fluid or melt (Evans, 2012; Karato, 1990, 2006; Khan, 2016; Selway, 2014; Yoshino, 2010). Therefore, while both data sets should converge toward a consistent view of the background thermochemical structure, they will diverge in regions where the electrical conductivity of rocks is affected by factors other than temperature or bulk composition. This makes MT + seismic joint inversions a powerful means to detect fluid pathways in the lithosphere (e.g., Evans et al., 2019; Selway & O'Donnell, 2019), including the locus of partial melting, ore deposits, and hydrated (or metasomatized) lithologies. This unique potential of joint MT + seismic inversions has also given impetus to the acquisition of collocated MT and seismic data over large regions. Concrete examples are the MAGIC and EarthScope USArray in USA ([www.usarray.org](http://www.usarray.org)), the AusLAMP program and AusArray in Australia ([www.ga.gov.au/efft/minerals/nawa](http://www.ga.gov.au/efft/minerals/nawa)), the IberArray ([www.iberarray.ictja.csic.es/](http://www.iberarray.ictja.csic.es/)) in Europe, and the Sinoprobe in China ([www.sinoprobe.org](http://www.sinoprobe.org)). These programs are providing high-quality seismic and MT data with unprecedented resolution and coverage, allowing the pursuit of large-scale 3D joint inversions for the physical state of the whole lithosphere and upper mantle.

The actual approach to the joint inversion of MT and seismic data is still a matter of much debate. While traditional deterministic methods are computationally efficient, they are not well suited to deal with the inherent non-uniqueness of geophysical data sets, and MT data in particular (e.g., Mallick & Verma, 1979; Oldenburg, 1979; Parker, 1971, 1980; Wait, 1962). This is because resolution and/or global uncertainty analyses are difficult, if not impossible, to obtain (e.g., Afonso, Moorkamp, & Fullea, 2016; Ledo & Jones, 2005; Moorkamp, 2017). Probabilistic inversion methods represent an attractive alternative (Gregory, 2005; Mosegaard & Hansen, 2016; Tarantola, 2005) as they are less susceptible to the above-mentioned limitations and provide substantially more information on the parameters of interest via full probability distributions. In probabilistic or Bayesian approaches, the solution to the inverse problem is given by the so-called posterior probability density function (PDF) over the model parameter space. This PDF summarizes all the information about the unknown parameters and their uncertainties conditioned on the data and modeling assumptions. As such, it represents the most general solution to the inverse problem. For nonlinear problems and/or complex priors, the posterior PDF cannot be represented analytically and needs to be sampled point-wise using, for example, Markov chain Monte Carlo (MCMC) algorithms (Gilks et al., 1995; Gregory, 2005; Mosegaard & Tarantola, 1995; Tarantola, 2005). This particular sampling-based approach to probabilistic inversions makes them less efficient than deterministic approaches, as they typically require the numerical solution of millions of forward problems. When the forward problems are computationally expensive, probabilistic approaches can be rendered impractical.

Joint probabilistic inversions of MT and seismic data have been successfully implemented by, for example, Khan et al. (2006, 2008); Afonso, Fullea, Griffin, et al. (2013); Afonso, Fullea, Yang, et al. (2013); Vozar et al. (2014); and Jones et al. (2017) in the context of 1D MT data only. For the cases of 2D and 3D MT data, however, the large computational cost of the MT forward problem has been the main impediment for pursuing probabilistic inversions, as the number of forward solutions required are typically on the order of  $10^5$ – $10^7$ .

In recent years, various methods and strategies for reducing the cost of full forward solutions have been proposed (see reviews in Frangos et al., 2011; Peherstorfer et al., 2018). The general idea behind these methods is the construction of an approximation, called the *low-fidelity* or *surrogate* model, which can be used instead of, or combined with, the costly full forward or *high-fidelity* solution. Having a faster surrogate of the forward problem is beneficial in a number of contexts, but it is particularly attractive in the context of MCMC schemes used to estimate the posterior PDF in a probabilistic inversion (Christen & Fox, 2005; Conrad et al., 2016; Cui et al., 2015; Florentin & Díez, 2012; Manassero et al., 2020; Ortega-Gelabert et al., 2020; Zhang & Taflanidis, 2019). In traditional implementations, the surrogates are computed in an offline stage (prior to the probabilistic inversion)

at specific locations within the parameter space called “snapshots”. However, it has been shown recently (Cui et al., 2015; Manassero et al., 2020; Ortega-Gelabert et al., 2020; Yan & Zhou, 2019; Zhang & Taflanidis, 2019) that in the context of high- and ultra-high-dimensional probabilistic inversions, it is practically impossible to pre-explore the parameter space in an offline stage to create surrogates that will guarantee accurate solutions within the so far unknown high-probability regions. In these situations, an adaptive MCMC approach where the surrogate is refined *online* during the MCMC simulation is a more effective and efficient approach.

A strategy to reduce the computational cost of the 3D MT forward solver and perform full probabilistic 3D MT inversions has recently been presented by Manassero et al. (2020). This novel strategy, called Reduced Basis (RB) + MCMC, combines (a) an efficient parallel-in-parallel structure to solve the 3D MT forward problem, (b) a Reduced Basis Method to create fast and accurate surrogate models of the *high-fidelity* solution, and (c) adaptive strategies for both the MCMC algorithm and the surrogate model.

This paper builds on our previous work (Manassero et al., 2020) and presents the first joint inversion of 3D magnetotelluric and surface-wave data within the context of MCMC-driven, fully probabilistic inversions. Specifically, we focus on a realistic 3D mapping of the electrical conductivity structure of the lithosphere including the locus of deep thermochemical anomalies and fluid pathways. We adopt the RB + MCMC strategy to compute 3D MT surrogate models and propose complementary parameterizations to couple both data sets. Using realistic, whole-lithosphere synthetic models, we demonstrate the benefits and general capabilities of our method for 3D joint probabilistic inversions of MT with surface-wave data in particular, and with other data sets in general.

## 2. Bayesian Inversion

Within the context of Bayesian inference, the most general solution to the inverse problem is represented by a multidimensional probability density function (PDF) over the combined parameter-data space (cf. Gilks et al., 1995; Gregory, 2005; Kaipio & Somersalo, 2006; Mosegaard & Hansen, 2016; Mosegaard & Tarantola, 2002; Tarantola & Valette, 1982). This distribution is known as the *posterior* PDF and can be thought of as an objective measure of our best state of knowledge on the problem at hand. It is obtained as a conjunction of the available information on the model parameters ( $\mathbf{m}$ ), the data ( $\mathbf{d}$ ), and their uncertainties. In particular, the conditional probability density for the model parameters given the observed data,  $P(\mathbf{m}|\mathbf{d})$ , is formally given by

$$P(\mathbf{m}|\mathbf{d}) \propto \mathcal{L}(\mathbf{m})P(\mathbf{m}). \quad (1)$$

where  $P(\mathbf{m})$  is a PDF encoding *a priori* information on the parameter space (what we know or believe about the unknown model parameters prior to considering the actual data) and  $\mathcal{L}(\mathbf{m})$  is the so-called *likelihood* function, which describes the probability of obtaining the observed data  $\mathbf{d}$  given  $\mathbf{m}$ . In general,  $P(\mathbf{m}|\mathbf{d})$  will be nonlinear and high-dimensional (and possibly multi-peaked), with no simple analytical description. When this is the case, unbiased approximations of  $P(\mathbf{m}|\mathbf{d})$  are commonly obtained via Markov chain Monte Carlo (MCMC) methods (Gilks et al., 1995; Gregory, 2005; Mosegaard & Tarantola, 1995; Tarantola, 2005). These types of algorithms are designed to output a large population of models (more precisely, a Markov chain) that has  $P(\mathbf{m}|\mathbf{d})$  as its equilibrium distribution. They do so by randomly and repeatedly drawing models  $\mathbf{m}$ , and evaluating their posterior probability  $P(\mathbf{m}|\mathbf{d})$  (via solution of the forward problem). A large number of MCMC methods have been proposed in the literature, all with relative merits and drawbacks. We refer the reader to the excellent monographs by, for example, Tarantola & Valette (1982); Gilks et al. (1995); Gregory (2005); Calvetti & Somersalo (2007); and Mosegaard & Hansen (2016) for in-depth treatments of Bayesian and MCMC methods applied to inverse problems. In the following, we restrict ourselves to describing only the most relevant theoretical and computational aspects for our purposes.

### 2.1. The Likelihood Function

The construction of an appropriate likelihood function  $\mathcal{L}(\mathbf{m})$  is a critical part of any Bayesian inference problem.  $\mathcal{L}(\mathbf{m})$  is typically specified by the distribution of the data uncertainty, which includes both observational and modeling errors. In most cases, observational errors are relatively straightforward to model. Modeling errors, on the other hand, are more complex (and commonly ignored in most geophysical studies) to describe and typically involves exploratory assessments of both numerical errors—for example, convergence analyses - and Monte Carlo estimates of the correlations between different data sets (see discussions and approaches in Afonso, Fulla,

Griffin, et al., 2013; Gouveia & Scales, 1998). In the convenient (and most popular) case where both observational and modeling errors can be assumed to be approximately Gaussian, the likelihood function takes the form:

$$\mathcal{L}(\mathbf{m}) \propto \exp\left(-\frac{1}{2}(\mathbf{g}(\mathbf{m}) - \mathbf{d})^t \mathbf{C}^{-1}(\mathbf{g}(\mathbf{m}) - \mathbf{d})^t\right) = \exp(\phi), \quad (2)$$

where  $\mathbf{C}$  is the data covariance matrix and  $\mathbf{g}(\mathbf{m})$  denotes the data predicted by the *forward problem* for model  $\mathbf{m}$ . The term  $\phi$  within the parenthesis in Equation 2 is commonly referred to as the *misfit* of model  $\mathbf{m}$ .

If the data errors are uncorrelated,  $\mathbf{C}$  is a diagonal matrix and the misfit can be written as

$$\phi = -\frac{1}{2} \sum_{i=1}^N \left( \frac{g_i(\mathbf{m}) - d_i(\mathbf{m})}{s_i} \right)^2 \quad (3)$$

where  $N$  is the total number of data and  $s_i$  denotes the standard deviation for the  $i$ th data error.

A more robust and often more realistic assumption (e.g., in the case of MT data Farquharson & Oldenburg (1998); Rosas-Carbajal et al. (2013)) is that data errors follow a Laplace (double exponential) distribution. In this situation, and considering uncorrelated data errors, the data misfit is given by Tarantola (2005)

$$\phi = - \sum_{i=1}^N \frac{|g_i(\mathbf{m}) - d_i(\mathbf{m})|}{s_i}. \quad (4)$$

In the case of joint inversions of independent observational data sets, the likelihood function can be written as the product of partial likelihoods:

$$\mathcal{L}(\mathbf{m}) = \prod \mathcal{L}_j(\mathbf{m}), \quad (5)$$

where  $\mathcal{L}_j$  refers to the likelihood associated with the data set  $\mathbf{d}^j$ . The assumption of independent observational data is well justified in most practical situations, and in particular in the MT + seismic case discussed in this paper, as different data sets are commonly gathered in separate surveys using different instrumentation. An important practical advantage of the factorization of the likelihood into partial likelihoods (Equation 5) is that it makes it possible to adopt a Cascaded Metropolis (CM) approach (Hassani & Renaudin, 2013; Tarantola, 2005), which is typically more efficient than a standard Metropolis-Hastings algorithm applied to the total likelihood.

## 2.2. Cascaded-Metropolis Algorithm

The CM algorithm is particularly useful when the different data sets jointly inverted are independent, have complementary sensitivities to different aspects of the problem, and at least one of the forward solvers is more computationally demanding than the others. The basic idea is to apply a Metropolis criterion sequentially to each partial posterior (prior  $\times$  partial likelihood), which becomes an updated prior in the evaluation of the subsequent partial posterior (e.g., Hassani & Renaudin, 2013, 2018). The practical benefits of the above procedure are significant when the partial likelihoods are arranged in order of computational complexity or cost, as there is no need to compute expensive forward solutions for models that are rejected early in the sequence (see e.g., Tarantola, 2005, for further details).

The basic procedure for the case of two forward operators is as follows: For a new sample  $\mathbf{m}_r$ , the first partial posterior  $P_1(\mathbf{m}_r|\mathbf{d}) = \mathcal{L}_1(\mathbf{m}_r)P(\mathbf{m}_r)$  is always computed using the computationally inexpensive forward solution. If  $P_1(\mathbf{m}_r|\mathbf{d}) > P_1(\mathbf{m}_{r-1}|\mathbf{d})$ , this first posterior becomes a prior in the evaluation of the second partial posterior, which is now obtained from the expensive forward:

$$P_2(\mathbf{m}_r|\mathbf{d}) = \mathcal{L}_2(\mathbf{m}_r)P_1(\mathbf{m}_r|\mathbf{d}). \quad (6)$$

If  $P_1(\mathbf{m}_r|\mathbf{d}) < P_1(\mathbf{m}_{r-1}|\mathbf{d})$ , the algorithm randomly decides to evaluate  $P_2(\mathbf{m}_r|\mathbf{d})$  or to reject the proposed move with a probability  $P = P_1(\mathbf{m}_r|\mathbf{d})/P_1(\mathbf{m}_{r-1}|\mathbf{d})$  of going to the second step. At the second step, the acceptance of the proposed move is computed as in the standard Metropolis-Hastings algorithm. In this work,  $P_1(\mathbf{m}_r|\mathbf{d})$  and  $P_2(\mathbf{m}_r|\mathbf{d})$  correspond to the surface-wave dispersion solver and the 3D MT solver, respectively (see details in Section 3).

We will also make use of the Adaptive Metropolis (AM) approach of Haario et al. (2001) to ameliorate the problem of choosing an optimal proposal before the start of the MCMC simulation and to obtain a more efficient sampling strategy of the parameter space that exploits correlations in the model parameters. We leave the presentation of this method to Section 5, where the general sampling strategy is discussed in detail.

### 3. Forward Problems

#### 3.1. The Magnetotelluric Forward Problem

In this section, we introduce the 3D magnetotelluric (MT) forward problem, the finite element high-fidelity solver and the RB + MCMC approach to compute surrogate solutions. The reader is referred to Douglas et al. (1999, 2000) and Zyserman and Santos (2000) for an in-depth treatment of the theory behind the formulation of the 3D MT problem and to Manassero et al. (2020) for a detailed description of the surrogate approach.

##### 3.1.1. High-Fidelity Solver for the MT Forward Problem in 3D

Using the secondary field formulation of Douglas et al. (1999, 2000) and the absorbing boundary conditions defined by Sheen (1997), the MT forward problem in 3D is defined as follows:

Find  $\mathbf{E}$  and  $\mathbf{H}$  such that

$$\sigma \mathbf{E} - \nabla \times \mathbf{H} = -\mathbf{F} \quad \text{in } \Omega, \quad (7a)$$

$$i\omega\mu_0\mathbf{H} + \nabla \times \mathbf{E} = 0 \quad \text{in } \Omega, \quad (7b)$$

$$(1-i)P_\tau a\mathbf{E} + \nu \times \mathbf{H} = 0 \quad \text{on } \partial\Omega \equiv \Gamma, \quad (7c)$$

where  $\mathbf{E}$  is the electric field [V/m];  $\mathbf{H}$  is the magnetic field (A/m);  $\mu_0$  is the magnetic permeability of free space (Vs/Am);  $\sigma$  is the electrical conductivity (S/m) of the medium  $\Omega \in \mathbb{R}^3$ , and  $\Gamma \equiv \partial\Omega$  is the boundary of the domain  $\Omega$ .  $a$  is defined as  $a = (\sigma/2\omega\mu_0)^{1/2}$  and  $P_\tau \boldsymbol{\varphi} = \boldsymbol{\varphi} - \nu(\nu \cdot \boldsymbol{\varphi})$  is the projection of the trace of any vector  $\boldsymbol{\varphi}$  on  $\Gamma$  where  $\nu$  is the unit outer normal to  $\Gamma$ .

High-fidelity numerical solutions to Equation 7 are sought via an optimized version of the finite element (FE) code developed by Zyserman & Santos (2000). In this optimized version, once the variational formulation of Equation 7 is discretized in terms of the FE shape functions, Equation 7 are converted into the following linear system of equations:

$$\mathbb{K}\mathbf{U} = \mathbf{F}, \quad (8)$$

where  $\mathbb{K}^{N_{FE} \times N_{FE}}$  is a sparse and symmetric matrix (the so-called FE *stiffness matrix*) and  $N_{FE}$  is the number of degrees of freedom (usually very large).  $\mathbf{F}^{N_{FE} \times 1}$  is the force vector and  $\mathbf{U}^{N_{FE} \times 1}$  is a vector containing the unknown coefficients for the electric field in the whole domain. In Manassero (2019) and in the Supporting Information S1, we present typical benchmarks (using models with strong resistivity contrasts) aiming at comparing the solution and the computational efficiency of this optimized version with other available codes. Additional benchmarks of the original code can be found in Zyserman and Santos (2000).

In MT, the numerical forward solution for a conductivity model requires the computation of two (typically orthogonal) components of the electromagnetic (EM) fields per frequency. Here, these components are referred to as  $\mathbf{U}^{S^i}$  and  $\mathbf{U}^{S^\perp}$ , for a frequency  $i$ . Once these solutions are computed, their coefficients and the FE shape functions are used to derive the electric and magnetic fields in the whole domain and at the surface of the Earth (for comparison with the observed data). It is worth noting that although the EM fields that satisfy Equation 7 are the actual solution to the forward problem, we will refer to the vector  $\mathbf{U}$  (either  $\mathbf{U}^{S^i}$  or  $\mathbf{U}^{S^\perp}$ ) as the *high-fidelity* solution to the forward problem.

As previously mentioned, the overall cost of computing the high-fidelity solution has been the main limitation preventing probabilistic inversions of 3D MT data. In the following section, we briefly describe the RB + MCMC strategy introduced in our previous paper (Manassero et al., 2020) to obtain fast and accurate approximations of the high-fidelity solutions.

### 3.1.2. Surrogate Solutions: A Reduced Basis + MCMC Approach

The RB + MCMC approach combines three main elements (a) a RB method to obtain fast approximations of the high-fidelity solution; (b) an MCMC algorithm that drives the sampling of the parameter space, and (c) an efficient parallel-in-parallel structure to solve the 3D MT forward problem (for both the surrogate and high-fidelity solvers). The first level of parallelization is defined by frequency, that is, different processors are in charge of computing the forward solution for different frequencies. The second level of parallelization includes a group of processors linked to each frequency, which compute (when needed) the costly high-fidelity solutions using the parallel solver *MUMPS* (Amestoy et al., 2001, 2006).

The general idea behind RB approaches is to seek for surrogate solutions as projections onto a space of small dimensionality, referred to as the *reduced basis*. We generate a reduced basis space  $\mathcal{V}_{RB}$  per frequency and field orientation, with dimension  $N_{RB} \ll N_{FE}$  and basis vectors  $\mathbf{V}_j$ . These *bases* are high-fidelity solutions of Equation 8 for specific realizations  $\theta$  of the conductivity model,  $\boldsymbol{\sigma}(\mathbf{x}, \theta)$ . In contrast to traditional RB approaches, these bases are not sampled in a preinversion stage, but rather during the MCMC inversion. In this way, each  $\mathcal{V}_{RB}$  is automatically updated (enriched) by adding new bases as needed during the evolution of the MCMC chain. This *online* enrichment approach circumvents the need of costly *offline* stages to build the reduced basis and increases the overall efficiency of the method (e.g., Manassero et al., 2020).

In the following, we summarize the main steps of the RB + MCM procedure. Note that items (1) – (4) are implemented per frequency  $i$  and field orientation ( $S^i$  and  $S^i_{\perp}$ ):

1. If there are bases available from an *offline* stage or from a preliminary probabilistic inversion, we load these bases as the initial basis matrix  $\mathbb{V}_{RB}$ . Otherwise, we compute the high-fidelity solution of the starting model of the Markov chain and add it as a column vector in the initial  $\mathbb{V}_{RB}$ .
2. For a new sample  $\mathbf{m}_i = \boldsymbol{\sigma}(\mathbf{x}, \theta)$ , we first seek for a surrogate solution to the forward problem by solving

$$\mathbb{K}_{RB}(\theta)\mathbf{a} = \mathbf{F}_{RB}(\theta) \quad (9)$$

for the coefficients  $\mathbf{a}(\theta)$ ; where  $\mathbb{K}_{RB}(\theta)^{N_{RB} \times N_{RB}} = \mathbb{V}_{RB}^T \mathbb{K}(\theta) \mathbb{V}_{RB}$  is the RB matrix,  $\mathbf{F}_{RB}(\theta)^{N_{RB} \times 1} = \mathbb{V}_{RB}^T \mathbf{F}(\theta)$  is the RB force vector, and  $\mathbb{V}_{RB}^{N_{FE} \times N_{RB}} = [\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_{N_{RB}}]$  is the matrix of basis vectors of  $\mathcal{V}_{RB}$ . The surrogate solution,  $\mathbf{U}_{RB}(\theta)$ , is then found as a linear combination of the basis vectors in  $\mathcal{V}_{RB}$  by substituting the coefficients  $\mathbf{a}(\theta)$  into the following equation:

$$\mathbf{U}_{RB}(\mathbf{x}, \theta) = \sum_{j=1}^{N_{RB}} a_j(\theta) \mathbf{V}_j = \mathbb{V}_{RB} \mathbf{a}(\theta). \quad (10)$$

Since the linear system of equations in Equation 9 is of size  $N_{RB} \ll N_{FE}$ , its computational cost is only a small fraction of the time consumed in solving Equation 8.

3. The following relative error is computed to assess the accuracy of the surrogate (Hesthaven et al., 2016; Quarteroni et al., 2015):

$$\mathbf{R}_{RB} := \frac{\|\mathbb{K} \mathbf{U}_{RB} - \mathbf{F}\|}{\|\mathbf{F}\|}, \quad (11)$$

where  $\|\cdot\|$  is the  $L_2$  norm.

4. The surrogate solution is considered admissible if  $\mathbf{R}_{RB} \leq \beta$  for a prescribed tolerance  $\beta$ .
5. If all the errors  $\mathbf{R}_{RB}$  are smaller than  $\beta$ , we accept  $\mathbf{U}_{RB}^{S^i}$  and  $\mathbf{U}_{RB}^{S^i_{\perp}}$  as acceptable approximations of the high-fidelity solution for all frequencies. In this case, the corresponding approximate likelihood,  $\tilde{\mathcal{L}}_2(\mathbf{m}_i)$ , is computed and the sample is either accepted or rejected according to the Metropolis-Hastings (MH) criterion.
6. In the case of any  $\mathbf{R}_{RB} \gg \beta$ , the high-fidelity FE solution for that frequency and component of the EM field is computed for  $\mathbf{m}_i$  and added as a new basis vector to enrich the corresponding space  $\mathcal{V}_{RB}$ . Since the posterior probabilities of the proposed sample  $\mathbf{m}_i$  and that of the current sample  $\mathbf{m}_{i-1}$  are no longer comparable (i.e., they were computed with different solvers, FE and RB, respectively), we recompute the surrogate solution (and the associated likelihood) at sample  $\mathbf{m}_{i-1}$  using the newly enriched RB space. If  $\mathbf{m}_i$  is rejected by the MH criterion, a new trial  $\mathbf{m}_i^*$  is proposed in the vicinity of  $\mathbf{m}_i$  and its likelihood is computed with the newly

enriched RB space. This new trial  $\mathbf{m}_i^*$  is accepted/rejected according to a modified Metropolis ratio to account for the delayed rejection (i.e., two proposals) step (see e.g., Haario et al., 2006; Mira, 2001).

As explained in Manassero et al. (2020), the last step above is required to preserve the ergodicity of the algorithm, but it is not the only possible option. We refer the reader to our previous work (Manassero et al., 2020) for further details on the combined RB + MCMC approach and additional functionalities to improve the efficiency of the method (e.g., use of variable tolerances and Singular Value Decomposition of the basis).

### 3.2. The Surface-Wave Forward Problem

Surface waves (SW) provide one of the most valuable data sets to study the lithospheric structure (e.g., Afonso, Fullea, Griffin, et al., 2013; Huang et al., 2009; Yang et al., 2008). One of the most common approaches involves (a) the generation of dispersion curves or 2D phase velocity maps at a number of periods via seismic tomography and (b) the subsequent 1D inversion of local dispersion curves for the shear velocity structure at depth (e.g., Afonso, Rawlinson, et al., 2016; Bensen et al., 2009; Ritzwoller et al., 2002; Shen et al., 2013; Yang et al., 2008). Here, we do not deal with the tomography part, for which many approaches are possible and covered in detail elsewhere, and consider only the inversion of dispersion curves. The relevant forward problem is therefore the computation of dispersion curves as functions of 1D vertical velocity structures, for which we use a modified version of the forward code *disp96* (Afonso, Fullea, Yang, et al., 2013; Afonso, Rawlinson, et al., 2016; Herrmann & Ammon, 2002). We compute anelastic wave velocities ( $V_s$  and  $V_p$ ) of mantle rocks as (Afonso et al., 2005, 2008, 2010):

$$V_s = V_{s0}(T, P)[1 - (1/2)\cot(\alpha\pi/2)Q_s^{-1}(T_o, T, P, d)], \quad (12)$$

$$V_p = V_{p0}(T, P)[1 - (2/9)\cot(\alpha\pi/2)Q_s^{-1}(T_o, T, P, d)], \quad (13)$$

where  $V_{s0}$  and  $V_{p0}$  are the unrelaxed, high-frequency (anharmonic) wave velocities at a given temperature ( $T$ ) and pressure ( $P$ ) (cf. Afonso et al., 2010). Without loss of generality, here we compute them as

$$V_{p0} = V_p^{ref} + \frac{\partial V_p}{\partial T} \Delta T + \frac{\partial V_p}{\partial P} \Delta P, \quad (14)$$

$$V_{s0} = V_s^{ref} + \frac{\partial V_s}{\partial T} \Delta T + \frac{\partial V_s}{\partial P} \Delta P, \quad (15)$$

where  $V_p^{ref}$  and  $V_s^{ref}$  are reference velocities at  $T_{ref}$  and  $P_{ref}$ ;  $\Delta T = T - T_{ref}$  and  $\Delta P = P - P_{ref}$ . The factor  $Q_s^{-1}$  is obtained as (Jackson & Faul, 2010; Jackson et al., 2002)

$$Q_s^{-1} = A \left[ \frac{T_o}{d} \exp\left(\frac{-E + VP}{RT}\right) \right]^\alpha, \quad (16)$$

where  $T_o$  is the oscillation period,  $d$  is grain size,  $E$  is the activation energy,  $V$  is the activation volume,  $\alpha$  is an empirical exponent,  $A$  is a pre-exponential constant and  $R$  is the universal gas constant. Although more sophisticated/realistic approaches for computing anelastic seismic velocities are possible (e.g., Afonso, Fullea, Griffin, et al., 2013; Afonso, Fullea, Yang, et al., 2013; Khan et al., 2008; Matas & Bukowinski, 2007; Vozar et al., 2014), the set represented by Equations 12–16 is sufficient for the goals of this paper.

## 4. Model Parameterization and Discretization

A key difficulty in the joint inversion of two or more disparate geophysical data sets is how to define the interdependence between model parameters in an internally consistent manner. For instance, if our goal was to jointly invert first arrivals of compressional waves ( $V_p$ ) and gravity anomalies (a common approach in geophysics), we would need to answer the following question: how is  $V_p$  related to bulk density in our medium? A typical assumption in this case is considering a linear correlation between  $V_p$  and density (e.g., Birch, 1961, 1964; Feng et al., 1986; Yasar & Erdogan, 2004). While this is a popular and practical assumption, the actual relationship between  $V_p$  and density also depends on temperature, pressure, and bulk composition (see e.g., Afonso, Fullea, Griffin, et al., 2013; Guerri et al., 2016). Several authors therefore distinguish between primary and secondary

parameters (e.g., Afonso, Fullea, Griffin, et al., 2013; Bosch, 1999; Khan et al., 2006). The latter are the most commonly used in geophysical inversions and refer to those that enter the governing equations of the forward problems (e.g.,  $V_p$ , density, electrical conductivity); the former are more fundamental in their nature and thus control the values of the secondary ones (e.g., temperature, porosity, and pressure).

In the case of joint inversions of SW and MT data, the primary parameters controlling both the seismic velocities and electrical conductivity ( $\sigma$ ) in the mantle are temperature ( $T$ ), bulk major-element composition ( $C$ ), and pressure  $P$  (e.g., Evans, 2012; Fullea et al., 2011; Jones et al., 2009; Selway, 2014). Using empirically calibrated equations of state of the type  $V_p(T, P, C)$ ,  $V_s(T, P, C)$ , and  $\sigma(T, P, C)$ , and thermodynamic constraints, we can establish direct relationships between the primary and secondary parameters (Bosch, 1999; Fullea et al., 2011; Jones et al., 2009; Khan et al., 2006; Xu et al., 2000; Yoshino, 2010). Since the electrical conductivity is also highly sensitive to hydrogen content, minor conductive constituents and localized melt/fluid pathways, we can explicitly write  $\sigma(T, P, C, X)$ , where  $X$  stands for any factor other than the bulk major-element composition of the rock. This distinction emphasizes the fact that although both seismic velocities and electrical conductivity can constrain the background  $T$ - $P$ - $C$  field, the electrical conductivity offers sensitivity to additional factors. The chosen model parameterization should thus be able to accommodate representative variations in both primary parameters (that simultaneously control  $V_p$ ,  $V_s$ , and  $\sigma$ ) and those responsible for conductivity anomalies above the background values. At the same time, as in any other inverse geophysical problem, the choice of model parameterization needs to be based on the principles of (a) flexibility, (b) parsimony, (c) parameter identifiability, and (d) suitability for the intended use.

With all of these in mind, and given our particular interest in lithospheric-scale imaging, we focus on a mixed parameterization of the conductivity distribution as the superposition of two contributions: a *background* conductivity related to the long-wavelength thermo-physical state of the lithosphere and an *anomalous* conductivity distribution associated with the presence of features such as fluid pathways, melt-rich regions, hydrogen-rich domains, anomalous mineral assemblages, etc. Following Afonso, Fullea, Griffin, et al. (2013); Afonso, Fullea, Yang, et al. (2013), we choose the depth to the lithosphere-asthenosphere boundary (LAB) and the bulk mantle composition as the main model parameters to constrain the background velocity and conductivity structures. We discuss this parameterization in more detail in Section 4.1. In order to account for smaller-scale conductivity anomalies superimposed on the background, we use a more standard parameterization based on conductivity nodes. This parameterization is only relevant to the MT forward problem and it is described in detail in Section 4.2. As shown in the numerical examples of Section 6, the advantage of using this combined parameterization is that a rapid convergence is achieved by using the LAB depths to constrain the first-order conductivity background at the beginning of the inversion. Once this first-order convergence has been achieved, the nodal values are used to locally modify the background to fit the smaller-scale features of the data.

#### 4.1. Background Parameterization

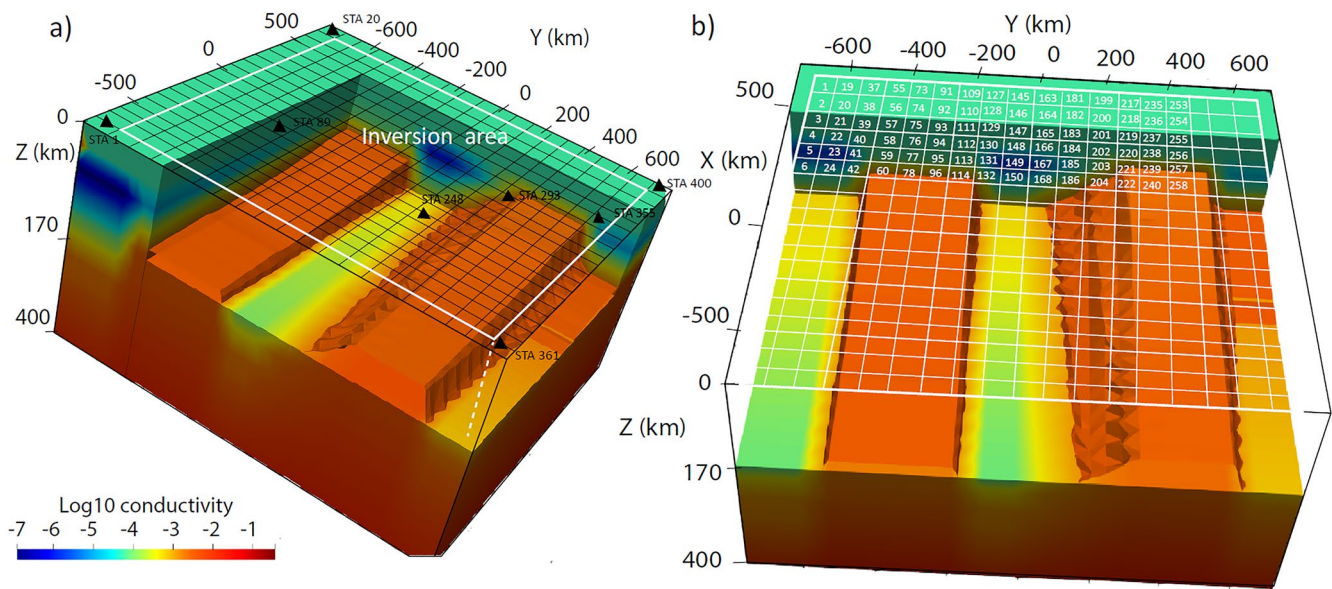
The 3D numerical model is made up of a collection of  $M_{\text{col}}$  columns (see Figure 1b). Each individual column is characterized by its own LAB depth. Here, we identify the LAB with the depth to the 1250°C isotherm (cf. Afonso, Moorkamp, & Fullea, 2016). In order to obtain the background conductivity structure from the LAB structure, we first compute the thermal profile of each column by solving the steady-state heat transfer problem with Dirichlet boundary conditions at the surface ( $T_0 = 10^\circ\text{C}$ ) and bottom of the lithosphere ( $T_{\text{LAB}} = 1250^\circ\text{C}$ ). For simplicity, we assume a linear temperature gradient between the LAB and 410 km depth, where the temperature is fixed at  $T_{410} = 1550^\circ\text{C}$ . This gradient is extrapolated to the bottom of the numerical domain (460 km). A pressure profile is also computed in each column using the following quadratic lithostatic-type approximation:

$$P(z) = 0.99 \times (4.4773 \times 10^{-3} z^2 + 3.2206 \times 10^4 z - 1.284278 \times 10^8), \quad (17)$$

where  $P$  is pressure in Pa and  $z$  is depth in meters.

As a further simplification, we assume a dry and homogeneous mantle composition with the following mineral modes: 56, 18.2, 10.8, and 15 vol% for olivine, orthopyroxene, clinopyroxene, and garnet, respectively. This is a standard mantle composition and it can be changed depending on the nature of the study area. While more realistic/sophisticated approaches to map major-element composition into mineral phases should be used when working with real data (e.g., Khan et al., 2006; Afonso, Fullea, Griffin, et al., 2013; Afonso, Fullea, Yang, et al.,





**Figure 1.** 3D views of the true conductivity structure where the iso-surface of  $-2.8 \log_{10} \text{ S/m}$  is plotted as a reference. The white rectangle in (a) indicates the region used for the inversion. Panel (a) illustrates the  $20 \times 20$  station-grid in black and eight of the 400 stations (black triangles). The model parameters are the depths to the LAB of 324 columns. Panel (b) displays the location of these columns (white small squares) and 96 column-parameters as a reference. The reader is referred to Section 4.1 for details on the parameterization.

2013; Afonso, Rawlinson, et al., 2016; Jones et al., 2017), this simplification does not affect the main results and conclusions of this paper. The electrical conductivity for each mineral phase is obtained using Equations A3a, A3b, and A3c with parameters specified in Table A1 and the bulk electrical conductivity (i.e., that of the mineral aggregate or rock) of each FE cell in the mantle is computed using the Hashin-Shtrikman averaging scheme (Hashin & Shtrikman, 1962, 1963).

For the surface-wave dispersion problem, each 1D column is further subdivided into 60 layers, each with constant density and wave velocities. The density of each layer is computed as a function of  $T$  and  $P$  values at the depth of its midpoint as follows:

$$\rho(P, T) = \rho_0 + 1 - \alpha(T - T_0) + \eta(P - P_0), \quad (18)$$

with  $\rho_0 = 3,355 \text{ kg/m}^3$ ,  $T_0 = 10^\circ\text{C}$ ,  $P_0 = 0 \text{ Pa}$ ,  $\alpha = 3.6 \times 10^{-5} \text{ 1/}^\circ\text{C}$ , and  $\eta = 1.1 \times 10^{-11} \text{ 1/Pa}$ . For a particular layer, the  $V_p$  and  $V_s$  are obtained using Equations 12 and 13 with the following values:  $A_v = 750 \text{ s}^{-\alpha} \mu\text{m}^\alpha$ ,  $\alpha = 0.26$ ,  $E = 424 \text{ kJ mol}^{-1}$ ,  $V = 1.3 \times 10^{-5} \text{ m}^3 \text{ mol}^{-1}$ , and grain size  $d = 5.0 \mu\text{m}$ . Given the periods of interest for surface waves, we adopt  $T_o = 50 \text{ s}$  in Equation 16 (Lebedev & Van Der Hilst, 2008; Liu et al., 1976; Moorkamp et al., 2020). The values for the parameters used in Equations 14 and 15 are listed in Table 1 (after Afonso et al., 2010).

#### 4.2. Node-Based Parameterization

Any conductivity anomaly that departs from the background is described with  $N_{\text{nodes}}$  nodes located within the numerical domain. In order to define the nodal locations (Figure 8c), the domain is first subdivided into horizontal layers of variable thickness. The midpoints of these layers correspond to the nodal depths. The horizontal distance between different locations within each layer is chosen relative to the mesh size and skin depth for the range of periods and apparent resistivities shown in the observed data (see for example Figure 4). The parameters of interest to be retrieved by the inversion are the conductivity values of these nodes. During the probabilistic inversion, the nodal values are interpolated to each FE cell of the numerical domain via

**Table 1**  
Parameters Used in the Computation of  $V_{s0}$  and  $V_{p0}$

$T_{\text{ref}}$	800.0°C
$P_{\text{ref}}$	0 Gpa
$\partial V_p / \partial T$	$-5.1 \times 10^{-4} \text{ (km/s}^\circ\text{C)}$
$\partial V_p / \partial P$	0.110 (km/sGPa)
$\partial V_s / \partial T$	$-3.3 \times 10^{-4} \text{ (km/s}^\circ\text{C)}$
$\partial V_s / \partial P$	0.03 (km/sGPa)

kriging interpolation (see e.g., Cressie, 1993; Omre, 1987; Williams & Rasmussen, 1996) using spatially varying correlation lengths (Section B1). Details about the implementation of the interpolation are given in Appendix B.

Intuitively, the range of anomalous conductivity values for the nodes should allow for positive and negative perturbations with respect to the background. However, as the electrical conductivity values can span several orders of magnitude, nodal values are typically obtained from proposal distributions defined in logarithmic scale (e.g., Jeffreys and log-normal distributions). Since the domain of the logarithmic function is the set of all positive real values, the sampled anomalous conductivity values (in linear scale) are always positive. Alternatively, one could consider the sign of the anomaly at each node as an additional parameter to be recovered by the inversion, but this option would double the number of model parameters. In practice, the use of positive anomalies is not a limitation, as resistive structures (i.e., negative deviations from the background) are generally determined solely by changes in the thermo-physical state (e.g., temperature and/or composition changes) whereas anomalous features of interest, such as presence of melt and/or fluid, hydrogen content, grain-boundary graphite films, and interconnected sulfides produce positive conductivity anomalies (e.g., Hu et al., 2017; Selway, 2014). Considering positive anomalous values over the background is therefore sufficient to represent any feature of interest while keeping the number of parameters small.

#### 4.3. A Note on the Combined Background + Nodes Parameterization

The current combined parameterization is specifically tailored to constrain the first-order conductivity background and to locally accommodate smaller-scale anomalies. This parameterization also allows for considerable model variance/flexibility, as it is capable of approximating any conductivity structure, and it favors a rapid convergence at the beginning of the inversion. However, it is difficult to know *a priori* the optimal number of parameters necessary to retrieve the true model. An over-parameterization of the model can seriously compromise the convergence of the MCMC algorithm, whereas an under-parameterization can introduce spurious features in regions where the conductivity nodes are far from each other (since the kriging-like interpolation produces spurious values where poor or no information from the surrounding nodes is available; see Appendix B and B1).

In practice, these issues are addressed by running preliminary inversions (similar to what is done in deterministic inversions with the variance-resolution trade-off diagram; Menke, 2018). A more efficient approach would be to implement transdimensional algorithms (e.g., Brodie & Jiang, 2018; Bodin & Sambridge, 2009; Ray & Myer, 2019), where the optimal dimensionality of the problem is identified as required by the data. In particular, the combination of the kriging interpolation (also known as Gaussian process regression) with a transdimensional algorithm is a promising approach (e.g., Ray & Myer, 2019) that warrants further investigation.

## 5. Sampling Strategy

The sampling strategy is specifically tailored to take advantage of the differential sensitivities of the SW and MT data sets to the conductivity structure of the lithosphere. With this in mind, we subdivide the MCMC inversion into four main stages. The first stage aims to constrain the background conductivity associated with the first-order temperature structure defined by the LAB depths (if we were interested in inverting for bulk chemical composition, we would also sample this parameter). In the second stage, conductivity anomalies over the background start to be sampled. During these first two stages, we sample both the LAB depths and the conductivity nodes using a *metropolized-independent* sampler (Tierney, 1994) where the proposal does not depend on the current state. Once enough information (i.e., enough samples) has been acquired for both sets of parameters, we incorporate ergodic adaptive strategies (Haario et al., 2001, 2006) to efficiently sample the full parameter space during the third and fourth stages. We briefly describe each of these stages below.

### 5.1. First Stage: Focus on Background Fields

1. Randomly select a column in the 3D domain using a *metropolized-independent sampler*.
2. Randomly propose an LAB depth for that column from its proposal distribution.
3. Recompute the temperature and pressure profiles and update the conductivity and wave velocities ( $\mathbf{m}_i$ ), as explained in Section 4.1.
4. Evaluate the first partial likelihood  $P_1(\mathbf{m}_i|\mathbf{d})$  with the SW solver.

5. Evaluate  $P_2(\mathbf{m}, \mathbf{d})$  with probability  $P = P_1(\mathbf{m}_t, \mathbf{d})/P_1(\mathbf{m}_{t-1}, \mathbf{d})$  using the MT forward solution:
  - (a) Seek for a surrogate RB solution to the 3D MT forward problem (Section 3.1.2).
  - (b) If  $R_{RB} < \beta$  for all frequencies,  $\mathbf{m}_t$  is accepted or rejected according to the Metropolis-Hastings criterion.
  - (c) If any  $R_{RB} > \beta$ , the high-fidelity FE solution is computed at  $\mathbf{m}_t$ . The RB surrogate is recomputed at  $\mathbf{m}_{t-1}$  and the algorithm proposes a new move in the vicinity of  $\mathbf{m}_t$  whose acceptance is evaluated with a Delayed Rejection criterion (Section 3.1.2).

### 5.2. Second Stage: Conductivity Nodes Begin to Be Sampled

When the number of MCMC steps reaches a predefined number of simulations (*LAB-stage*):

1. Randomly chose a type of parameter to sample (i.e., LAB depths or nodes) at each MCMC step.
2. If chosen parameter = LAB, the algorithm follows the **first stage**.
3. If chosen parameter = conductivity nodes:
  - (a) Randomly select  $n_1$  nodes at a time, with all nodes having the same probability of being chosen.
  - (b) Assign a random conductivity value to each node from their individual proposal distributions.
  - (c) Update the 3D conductivity model via kriging interpolation.
  - (d)  $P_1(\mathbf{m}, \mathbf{d})$  remains unchanged, that is, it only changes when a new LAB value is proposed.
  - (e) Evaluate  $P_2(\mathbf{m}, \mathbf{d})$  with the MT solver following items (a–c) of the **first stage**.

### 5.3. Third Stage: Adaptive Strategy for the LAB Depths

When the number of MCMC steps reaches a predefined number of simulations (*LAB-adapt*):

1. Compute a new multivariate Gaussian proposal distribution (via the Adaptive Metropolis algorithm of Haario et al. (2001)) using the history of the MCMC chains. This proposal now has information about spatial correlations in the LAB.
2. Randomly chose a type of parameter to sample (i.e., LAB depths or nodes) at each MCMC step.
3. If chosen parameter = LAB:
  - (a) Randomly select  $m$  columns at a time, with all columns having the same probability of being chosen.
  - (b) Propose a new sample for the selected LAB depths using the global multivariate Gaussian proposal.
  - (c) Follow items (3–5) of the **first stage**.
4. If chosen parameter = conductivity nodes, the algorithm follows items (a–f) of the **second stage**.

### 5.4. Fourth Stage: Adaptive Strategy for the Conductivity Nodes

When the number of MCMC steps reaches a predefined number of simulations (*nodes-adapt*):

1. Compute a multivariate log-normal proposal distribution via the Adaptive Metropolis algorithm using the MCMC chains of all nodes.
2. Randomly chose a type of parameter to sample (i.e., LAB depths or nodes) at each MCMC step.
3. If chosen parameter = LAB, follow item (3) of the **third stage**.
4. If chosen parameter = conductivity nodes:
  - (a) Randomly select  $n_2$  nodes with a *metropolized-independent sampler*.
  - (b) Use the multivariate log-normal distribution to propose new conductivity values for the  $n_2$  random nodes with probability  $q(\cdot|\cdot)$  defined in Eq. C2.
  - (c) Follow items (c–f) of the **second stage**.

The first stage only needs a moderate number of models to significantly reduce the original range of possible LAB values. This rapid convergence is due to the strong combined sensitivity of SW and MT to the background field; it also allows the MCMC inversion to focus on the last three stages (i.e., on conductivity anomalies not related to the background  $T$ - $P$ - $C$  conditions) while still allowing a continuous improvement of the background field. Additional gain in convergence efficiency is obtained with adaptive sampling strategies applied to both LAB and

conductivity nodes. The implementation of these strategies is almost imperative given the high-dimensionality of the problem.

Note that the *burn-in* period needs to be larger than the total number of steps in the first stage (*LAB-stage*) to ensure the overall ergodicity (e.g., Meyn & Tweedie, 2012) and correct convergence of the sampler (Adaptive Metropolis and the RB + MCMC procedure of stages 3 and 4 maintain ergodicity, Haario et al., 2001, 2006; Manassero et al., 2020, see also Section 3.1.2). We also note that while more advanced sampling strategies (e.g., parallel tempering, differential evolution, and auto-regressive chains) can be implemented to further improve efficiency, we deliberately use this practical (and basic) four-step adaptive strategy to test our joint inversion algorithm under adverse circumstances.

## 6. Numerical Examples

### 6.1. Example 1: Large-Scale Thermal Lithospheric Structure

The aim of this example is to demonstrate the improved resolution and efficiency of the joint MT + SW inversion to recover the background conductivity structure compared to the probabilistic inversion of MT data only presented in Manassero et al. (2020). Accordingly, we only use the LAB parameterization in the first and third stages (Sections 5.1 and 5.3), simple noise statistics for the data and the same model discretization as in Manassero et al. (2020).

#### 6.1.1. Synthetic Data

The synthetic data correspond to a large-scale lithospheric model with dimensions  $1600 \times 1600 \times 460$  km (Figure 1). The MT synthetic data are the off-diagonal apparent resistivities and phases computed for 12 periods between 3.2 s and  $10^4$  s at 400 stations. The stations are located on a grid of  $20 \times 20$  (Figure 1a) with an interstation distance of 80 km. The data uncertainties are assumed to be uncorrelated and normally distributed. We use a standard deviation of 8% for the apparent resistivities and  $2.5^\circ$  for the phases. For the case of the SW, the synthetic data are the fundamental mode Rayleigh wave phase velocities for periods between 15 and 175 s, computed at the locations of the MT stations. We assume normally distributed and uncorrelated data errors with a standard deviation (std) of 20% of the period (e.g., std = 5 m/s at 25 s and std = 35 m/s at 175 s). For both data sets, the misfit function is given by Equation 3. To minimize the so-called “inversion crime,” where the forward model used during the inversion is the same as that used to create the synthetic data (Kaipio & Somersalo, 2006), we compute the synthetic MT and SW data with a finer FE mesh than that used in the inversion.

#### 6.1.2. Model Setup

The inversion area is subdivided into  $18 \times 18$  columns (white squares in Figure 1b) of size  $80 \times 80 \times 460$  km. Each column is comprised of  $4 \times 4 \times 20$  FE cells (i.e., the computational domain is discretized with  $40 \times 40 \times 20$  finite elements). The model includes two padding cells at the borders of the domain and four air-layers (total thickness of 106 km) with resistivity  $\rho = 10^{15} \Omega\text{m}$ . As shown in Manassero (2019) and in the Supporting Information S1, the absorbing boundary conditions allow for a small number of padding cells (and, therefore, considerable size reduction of the computational domain) without compromising the solution. The model parameters are the depths to the LAB of the 324 columns within the inversion area, that is, there is one model parameter per column. The true conductivity model is shown in Figure 1 and it is controlled by the subsurface thermal structure. The resistivity in the crust (Moho at 49 km depth) is held constant and equal to  $20,000 \Omega\text{m}$  (see Manassero et al., 2020).

#### 6.1.3. Prior and Proposal Distributions

The priors for the LAB depths are uniform distributions defined in a range of  $\pm 70$  km, centered on the true value of each column. The proposals used in the first stage of the inversion are Gaussian distributions centered on the current sample with a standard deviation of 20 km. The proposal is adapted in the third stage (starting at 80,000 steps), and therefore it becomes a multivariate Gaussian distribution that reflects the spatial correlations between LAB values of all columns (see Section 5.3). The initial model (i.e., starting point of the MCMC inversion) has a flat LAB located at 180 km depth.

#### 6.1.4. Inversion Results

We ran a total of 600,000 MCMC simulations using two processors (Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50 GHz processors) per frequency and variable RB tolerances of  $\beta = 0.07$  for the first 50,000 MCMC steps and  $\beta = 0.05$  for the rest of the simulation. After computing Geweke's convergence diagnostics (Geweke, 1992) for all parameters, the *burn-in* period was set to 100,000 steps. This *burn-in* and run length of the simulation is also consistent with the Raftery-Lewis diagnostics (Raftery & Lewis, 1992). Despite the small number of processors used, the joint inversion took only 25 hr, with a staggering average of 0.15 s per simulation. This means >99.5% gain in computational efficiency compared to the high-fidelity solution (~30 s). For the same model, and using the same number and type of processors, the RB + MCMC inversion of MT data only (see Manassero et al., 2020) took ~30 days (an average of 1.03 s per MCMC iteration) and convergence was achieved after 2,000,000 MCMC simulations (based on visual inspection of the chain's evolution and basis size). This dramatic gain in efficiency of the joint inversion is due mainly to (a) the implementation of the CM algorithm, (b) the use of adaptive MCMC strategies, and (c) the high sensitivity of SW data to the background thermal structure.

The posterior PDFs of 60 of the 324 parameters are shown in Figure 2. The data PDFs for the dispersion curves at two illustrative stations and the data PDFs for MT at one station are shown in Figure 3 and Figure 4, respectively. Additional results can be found in the Supporting Information S1. The results clearly show that the posterior PDFs for all parameters are well-behaved (i.e., single valued and approximately Gaussian) and include the true solution, which is always close to the peaks of the PDFs. The resulting uncertainties affecting the LAB values are comparable to those obtained in real inversions (e.g., Afonso, Moorkamp, & Fulla, 2016; Zhang et al., 2019). The data fit is excellent for both data sets (see Figures 3 and 4).

The true, Maximum Likelihood Estimation (MLE) and mean models are shown in Figure 5, together with the 95% credibility intervals of the posterior PDFs. The root-mean-square (rms) values of the MLE and mean conductivity models, as well as the rms for the LAB structure, are included in Table 2. As a comparison, we have also included the rms values obtained for the same model after the RB + MCMC inversion of 3D MT data only (see Manassero et al., 2020), which are considerable higher than those obtained with the joint inversion.

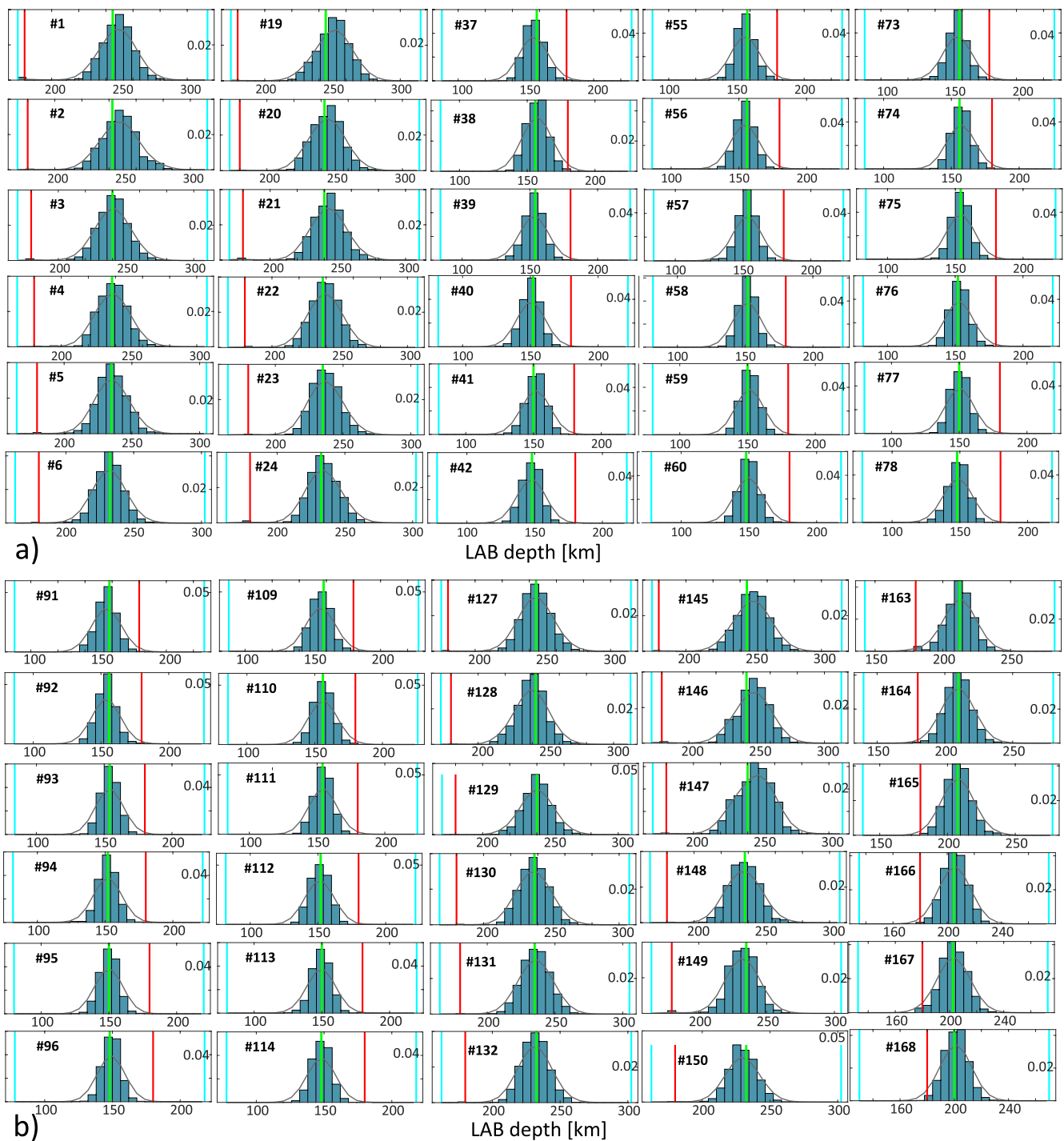
The evolution of the misfits for MT and SW data is shown in Figure 6. The number of bases computed per frequency and field orientations are shown in Figure 7. In all cases, a rapid increment in the basis size is observed during the first 100,000 simulations, which correlates with a rapid decrease in the overall misfits (Figure 6). This rapid increment in the number of basis is the combination of two factors: (a) the starting point of the inversion is far from the high probability region and (b) the initial proposal distribution is not optimal and of large variance. The MCMC algorithm thus samples a wide spectrum of models in its attempt to locate the best paths to the high probability regions. During this exhaustive exploration, the moves or "jumps" through the parameter space are large. Consequently, the resulting conductivity models are significantly different from each other and the surrogate needs to be constantly enriched in order to produce accurate solutions for all possible models.

After ~100,000 MCMC steps, the basis size reaches a *plateau* (i.e., saturation of the surrogate) for all frequencies and orientations. This means that (a) the chain has reached the high probability regions and (b) the RB surrogate is "rich enough" to be able to deliver accurate solutions within these regions (as only a small number of new bases are subsequently required). At this point, we could stop the adaptation or enrichment of the surrogate without compromising the accuracy of the final solution.

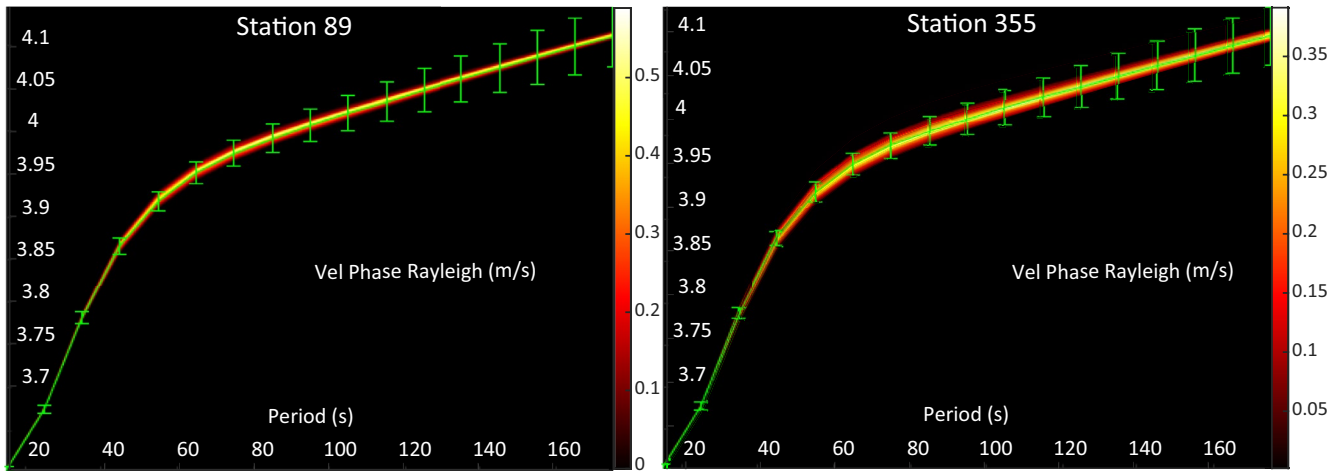
**Table 2**  
*Root-Mean-Square (RMS) Values of the Mean and Maximum Likelihood Estimation (MLE) Conductivity and Lithosphere-Asthenosphere Boundary (LAB) Models With Respect to the True model*

	RMS conductivity ( $\log_{10} \Omega\text{m}$ )		RMS LAB depth (km)	
	MLE	Mean model	MLE	Mean model
Joint RB + MCMC	0.08	0.02	6.89	2.21
RB + MCMC	0.19	0.15	21.20	17.01

*Note.* The rms values obtained after the Reduced Basis + Markov chain Monte Carlo (RB + MCMC) inversion of 3D magnetotelluric (MT) data only are also included (extracted from Manassero et al., 2020).

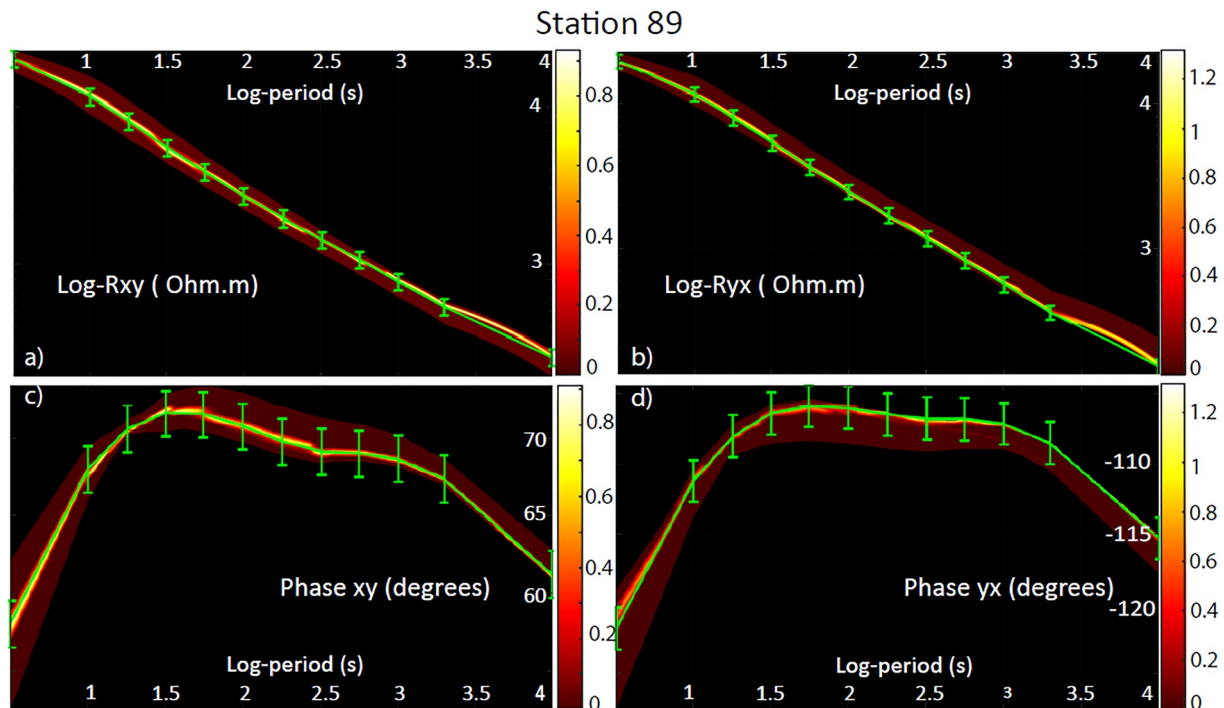


**Figure 2.** Marginal posterior probability density functions (PDFs) (blue bars) of 60 model parameters obtained after 600,000 Reduced Basis + Markov chain Monte Carlo (RB + MCMC) simulations. The real value, starting value, and prior bounds of each parameter are shown in green, red, and light blue vertical lines, respectively. The best Gaussian fits to the real PDFs given by the histograms are shown in black lines. The numbers within each panel correspond to the columns highlighted in Figure 1b.

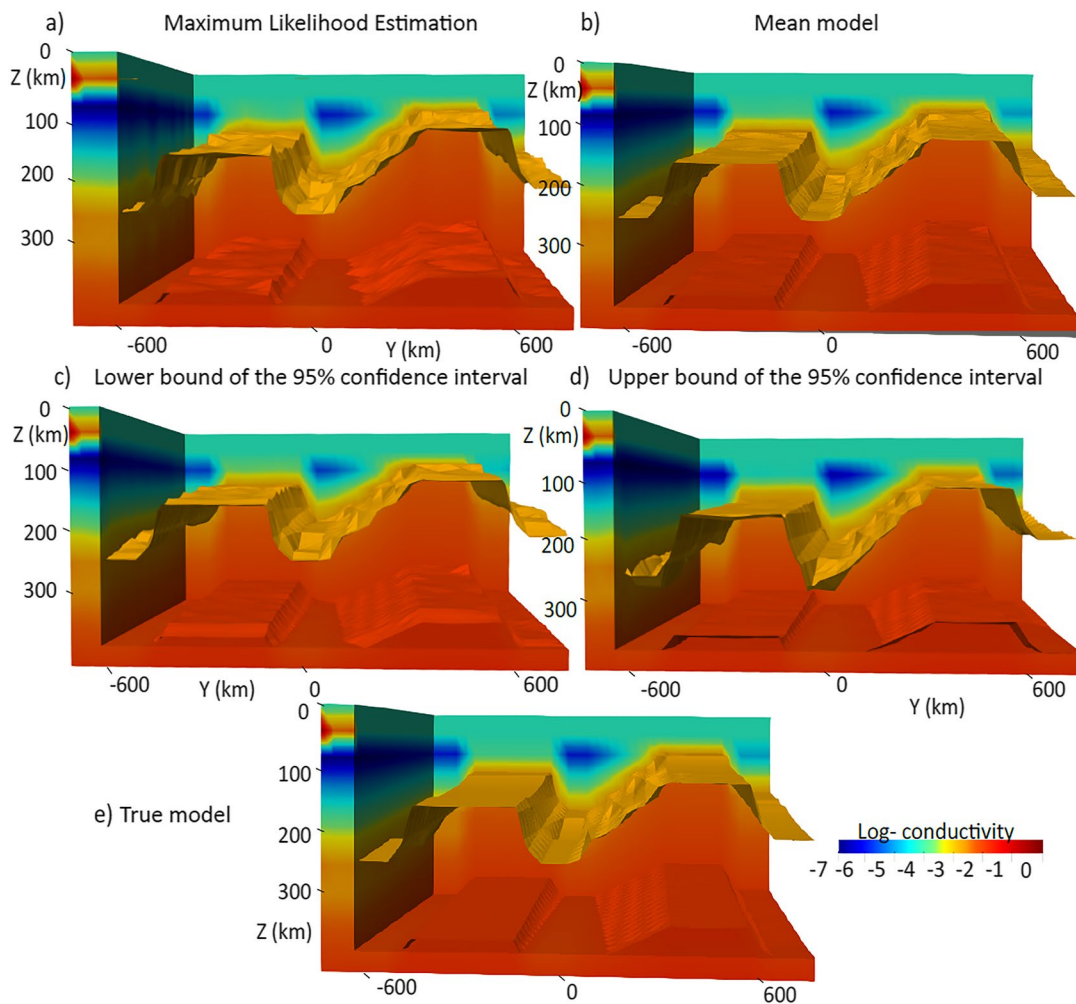


**Figure 3.** Posterior probability density functions (PDFs) of Rayleigh dispersion curves for stations (a) 89 and (b) 355. Synthetic data and error bars are plotted in green. The location of the stations are shown in Figure 1a.

These results demonstrate that our RB + MCMC approach successfully solves the joint probabilistic inversion problem and retrieves the first order conductivity structure (and associated uncertainties) from noise-free MT and SW data. Moreover, we demonstrate that the addition of the SW data increases the overall efficiency of the algorithm and significantly reduces the range of acceptable conductivity models compared to those obtained from the inversion of MT data only.



**Figure 4.** Posterior probability density functions (PDFs) of magnetotelluric (MT) data for station 89. Synthetic data and error bars are plotted in green. (a and b) Posterior PDFs of the off-diagonal apparent resistivities. Both axes are in log scale (c and d) Posterior PDFs of the off-diagonal apparent phases. The location of the stations is shown in Figure 1a.

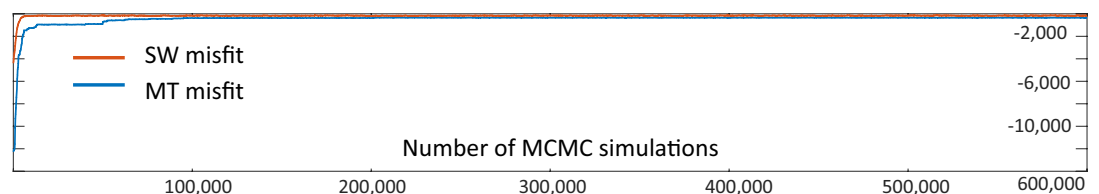


**Figure 5.** Conductivity structures corresponding to the (a) maximum likelihood estimation (best-fitting model); (b) mean model; conductivity models corresponding to the lower (c) and upper bound (d) of the 95% credibility interval of the posterior probability density function (PDF) obtained after 600,000 MCMC steps; and (e) true model. The iso-surfaces of  $-2.8$  and  $-2 \log_{10} S/m$  are plotted as a reference.

## 6.2. Example 2: Large-Scale Lithospheric Structure With Conductivity Anomalies

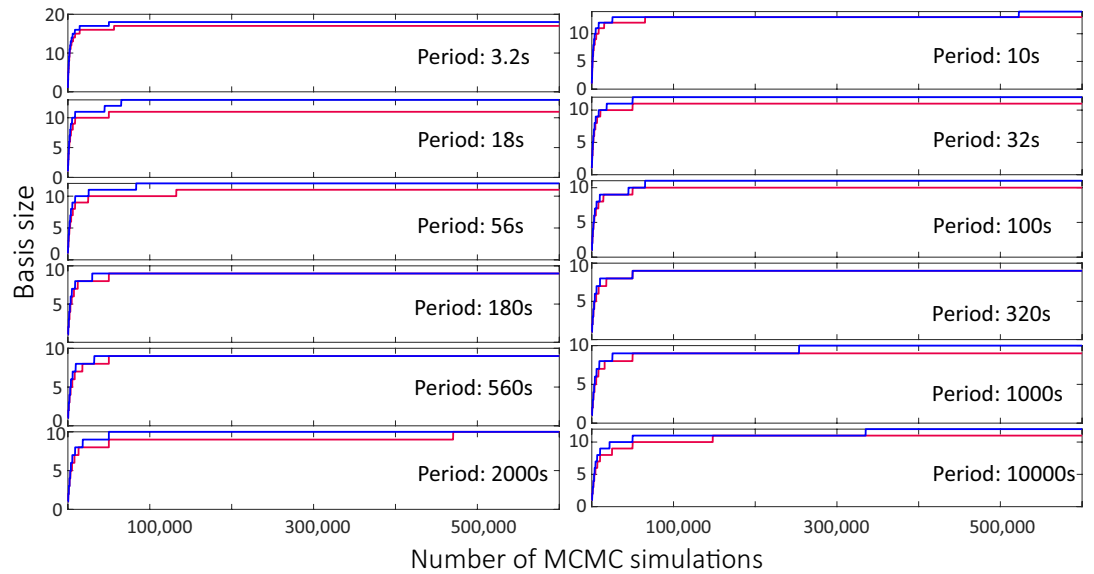
### 6.2.1. Model Setup

In order to assess the applicability of our method to more realistic scenarios, we have created a synthetic model using the crustal conductivity structure (Moho  $\sim 40$  km) reported for southeast Australia using data from the dense AusLAMP Array of MT stations, with a spacing of  $\sim 55$  km (Kirkby et al., 2020). In this model, the ocean has been replaced by crustal structures with resistivity values of  $100 \Omega m$  at the southwest and southeast of the



**Figure 6.** Data misfits for the dispersion curves (red line) and magnetotelluric (MT) (blue line) for each one of the 600,000 Reduced Basis + Markov chain Monte Carlo (RB + MCMC) steps.





**Figure 7.** Basis size as a function of the Markov chain Monte Carlo (MCMC) simulations for different frequencies and field orientations ( $S_{\perp}$  mode in blue and  $S$  mode in red).

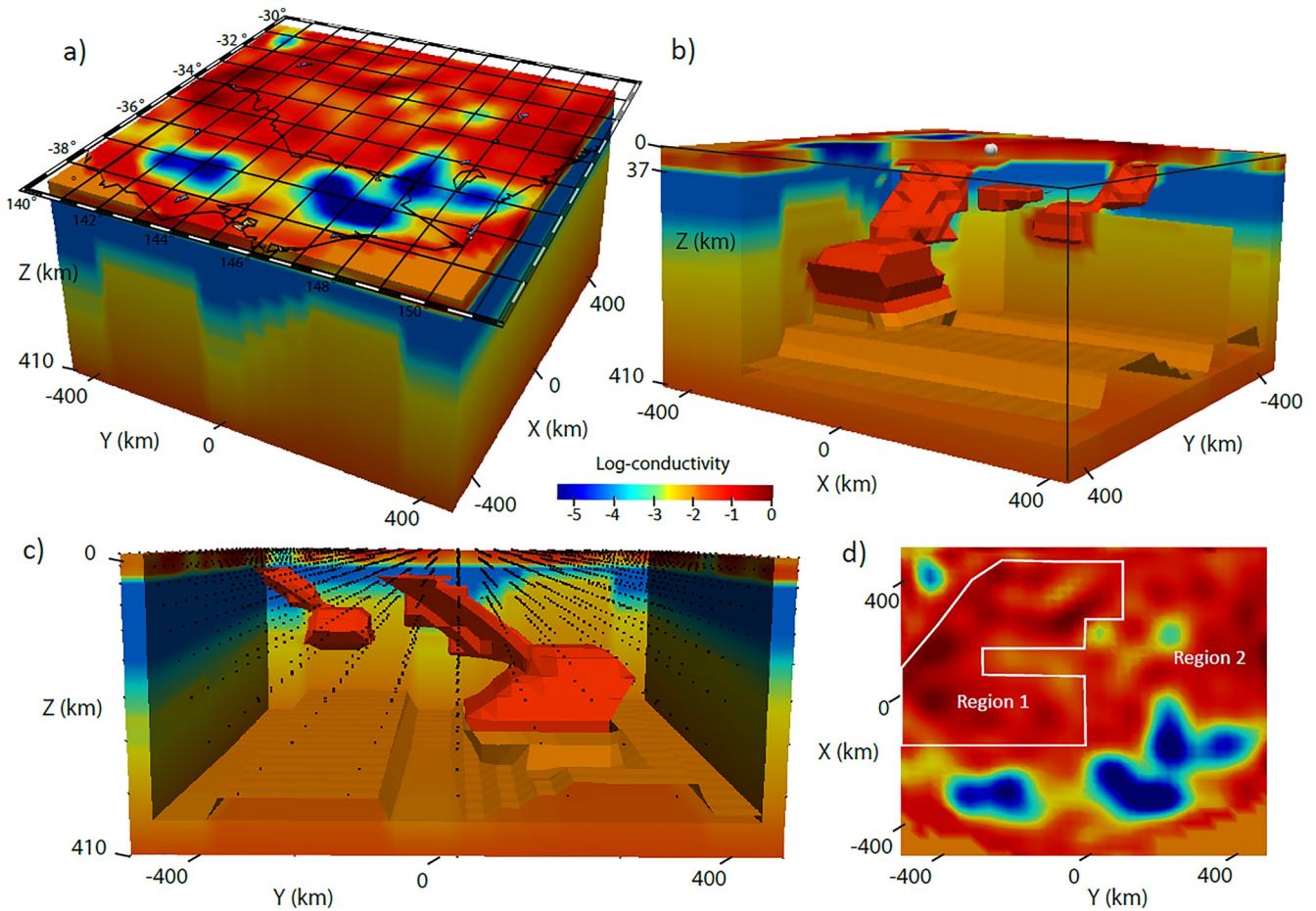
model and  $10 \Omega\text{m}$  between those regions. These values were chosen to add another strong resistivity contrast to those already present in the crust. The area of interest is shown in Figures 8a and 9. We include the possibility of both correlated and uncorrelated velocity-conductivity structures in the crust. In Region 1 (white region in Figure 8d), we assign a constant ratio  $V_p/V_s = 1.78$  and assume  $V_p$  to be correlated with electrical conductivity (Meju et al., 2003) as  $\log_{10}(1/\sigma) = m \log_{10}(V_p) + c$ , where  $m = 3.88$  and  $c = 13$  for consolidated rocks. For the crustal velocity structure of Region 2 (see Figure 8d), we assign constant velocities  $V_p = 6.8 \text{ km/s}$  and  $V_s = 3.9 \text{ km/s}$  and assume them independent of electrical conductivity.

The mantle includes the lithospheric model of Section 6.1 as a background (with an additional cut-off for resistivity values higher than  $20,000 \Omega\text{m}$ ) plus several multi-scale (and of variable geometry) conductivity anomalies that simulate realistic geological features. In particular, we have included two elongated anomalies resembling trans-lithospheric and trans-crustal magmatic systems (Figure 8b and 8c). Region 1 represents the continuity of these translithospheric structures into the crust. The goal here is to assess the identifiability of the true conductivity structure, including background and conductivity anomalies within both the crust and the mantle, from noisy 3D MT and SW measurements.

There are 2290 conductivity nodes (black dots in Figure 8c) sparsely located within the inversion volume ( $900 \times 900 \times 410 \text{ km}$ ), which is discretized into 361 columns. The vector of model parameters therefore contains 361 LAB values and 2290 nodal conductivity values. The inversion volume is discretized with  $38 \times 38 \times 20$  finite elements (of  $25 \times 25 \text{ km}$  in the  $x$ - $y$  direction, and variable vertical thickness). As before, the computational domain includes four extra layers of air with resistivity  $\rho = 10^{15} \Omega\text{m}$  and total thickness of  $106 \text{ km}$ , and two padding cells at the lateral boundaries of the domain. The conductivity value of each numerical cell is obtained by adding the background conductivity derived from the LAB structure (Section 4.1) and the anomalous conductivity obtained after interpolation of the nodal values (Section 4.2).

### 6.2.2. Synthetic Data

The MT synthetic data were computed at 298 stations located according to the real AusLAMP deployment (black dots in Figure 9) in New South Wales and Victoria, Australia. The data are the full impedance tensor computed for 18 periods between  $15.80 \text{ s}$  and  $39,800 \text{ s}$ , which correspond to the frequencies found in the AusLAMP data. All data were contaminated with random noise (see below). The error floors are set to  $5\%$  of  $\max(|Z_{xx}|, |Z_{xy}|)$  for the components  $Z_{xx}$  and  $Z_{xy}$  of the impedance tensor, and  $5\%$  of  $\max(|Z_{yy}|, |Z_{yx}|)$  for the components  $Z_{yy}$  and  $Z_{yx}$ . The data errors are assumed to be uncorrelated and exponentially distributed, that is, we assume that



**Figure 8.** 3D views of the true conductivity structure. Panel (a) shows the coordinates of the inversion volume. Conductivity anomalies and background structure are highlighted in both (b) and (c) panels using the iso-surfaces of  $-1.5$  and  $-2.0 \log_{10}(\text{S/m})$ . Black dots in (c) indicate the position of the conductivity node-parameters within the inversion volume. A horizontal slice at 1 km depth is shown in panel (d). The region where seismic velocity is assumed correlated with electrical conductivity is highlighted in white (Region 1).

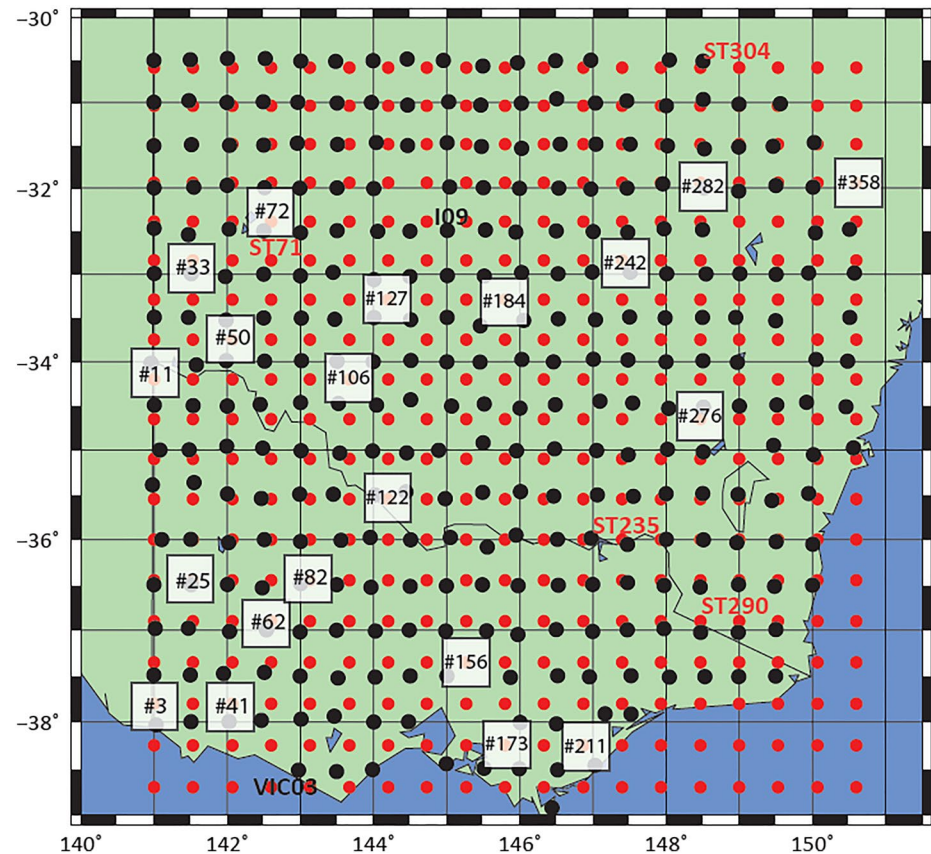
the data misfit follows Equation 4. The MT data are generated with the true conductivity value for each FE cell, whereas the models used during the inversion are obtained via interpolation of the nodes' values (Section 4.2). This reduces the inversion crime and simulates a more realistic scenario where the (unknown) true structures are approximated via a chosen parameterization in the inversion. This also implies, however, that a perfect data fit may not be achievable.

For the case of SW, the synthetic data are the fundamental mode Rayleigh wave phase velocities for 34 periods between 15 and 180 s. The stations are located on a grid of  $19 \times 19$  (red dots in Figure 9) with an interstation distance of 50 km (comparable to those in the WOMBAT Array; Rawlinson et al., 2008). We have added random noise to all the data (see example in Figure 13) assuming normally distributed data errors (i.e., the misfit function is given by Equation 3) with a representative standard deviation of 1% of the velocity in meters per second, comparable to the data errors expected for real SW data in dense arrays (Moorkamp et al., 2010; Wang et al., 2020; Yang & Forsyth, 2006).

### 6.2.3. Prior and Proposal Distributions

The prior and proposal distributions for the LAB parameters are the ones defined in Section 6.1.3. For the conductivity nodes, we use Gaussian prior distributions centered on the background conductivity value (in log-scale) with a standard deviation of  $1.5 \log_{10}(\text{S/m})$ . This prior information behaves as a regularization term, that is, it penalizes the introduction of anomalies that are not required by the data. The initial proposal distributions are

## Eastern Australia



**Figure 9.** Location of the AusLAMP MT stations (black dots) in southeast Australia. Red dots denote the locations of the computed SW dispersion data used in Example 2.

log-normal (Equation C1) centered on the current node value  $m_{l-1}^i$  with standard deviations of  $0.8 \log_{10}(S/m)$ . During the fourth stage, we use an adapted multivariate log-normal distribution centered on the current sample (see Section 5.4). The starting conductivity model is obtained by setting all the LAB depths to 180 km and a value of  $\log_{10}(\text{true val}) - 2.0$  (i.e., two order of magnitude more resistive than the real value) for all the conductivity nodes.

The first stage was set to 10,000 steps, where we sample LAB depths one column at a time. During the second stage, the algorithm randomly decides to sample either the LAB depth of one column or the conductivity values of  $n_1 = 2$  nodes. The multivariate proposal for the LAB (third stage) is computed when the chains reach 250,000 samples and it is adapted every 100,000 LAB-samples during the rest of the inversion. During this third stage, we propose conductivity values of  $n_1 = 2$  random nodes or LAB depths of  $m = 2$  random columns (from the adapted multivariate proposal distributions; see Section 5.3). The multivariate log-normal proposal distribution for the nodes is computed when their chains reach 500,000 samples (fourth stage) and it is subsequently adapted every 100,000 steps. During this stage, we randomly select  $n_2 = 10$  nodes or  $m = 2$  columns at a time (see Section 5.4).

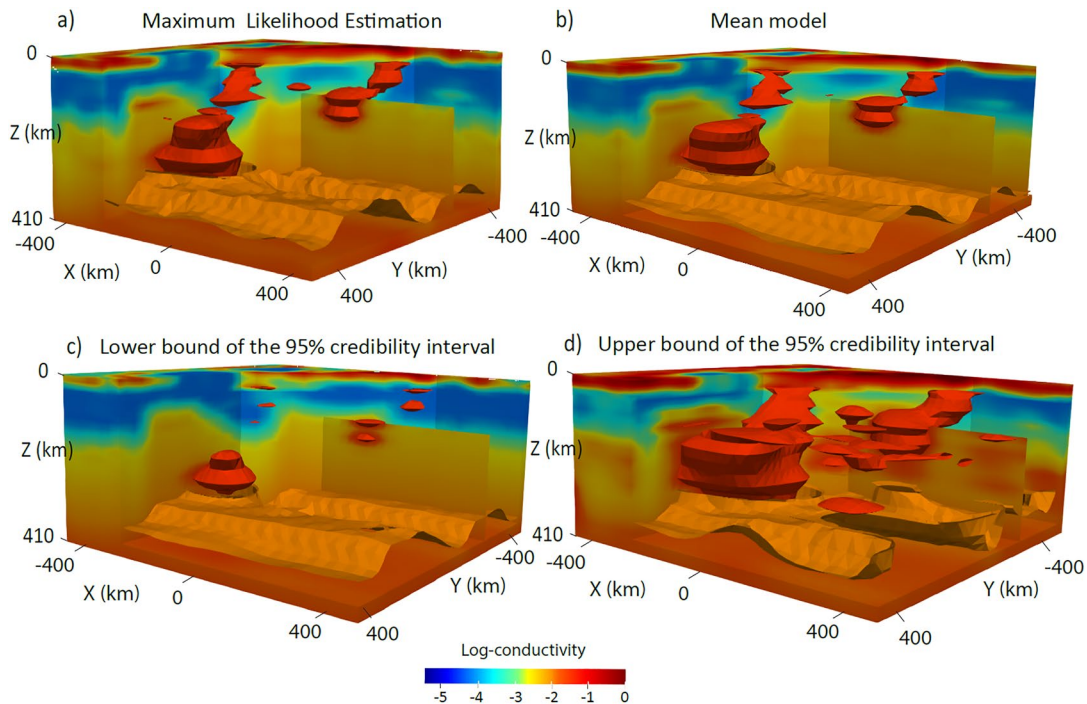
### 6.2.4. Inversion Results

We ran a total of 1,250,000 MCMC simulations for 18 frequencies using two processors (Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50 GHz) per frequency. The tolerances used were  $\beta = 0.08$  for the first 150,000 steps and  $\beta = 0.068$  for the remaining of the simulation.

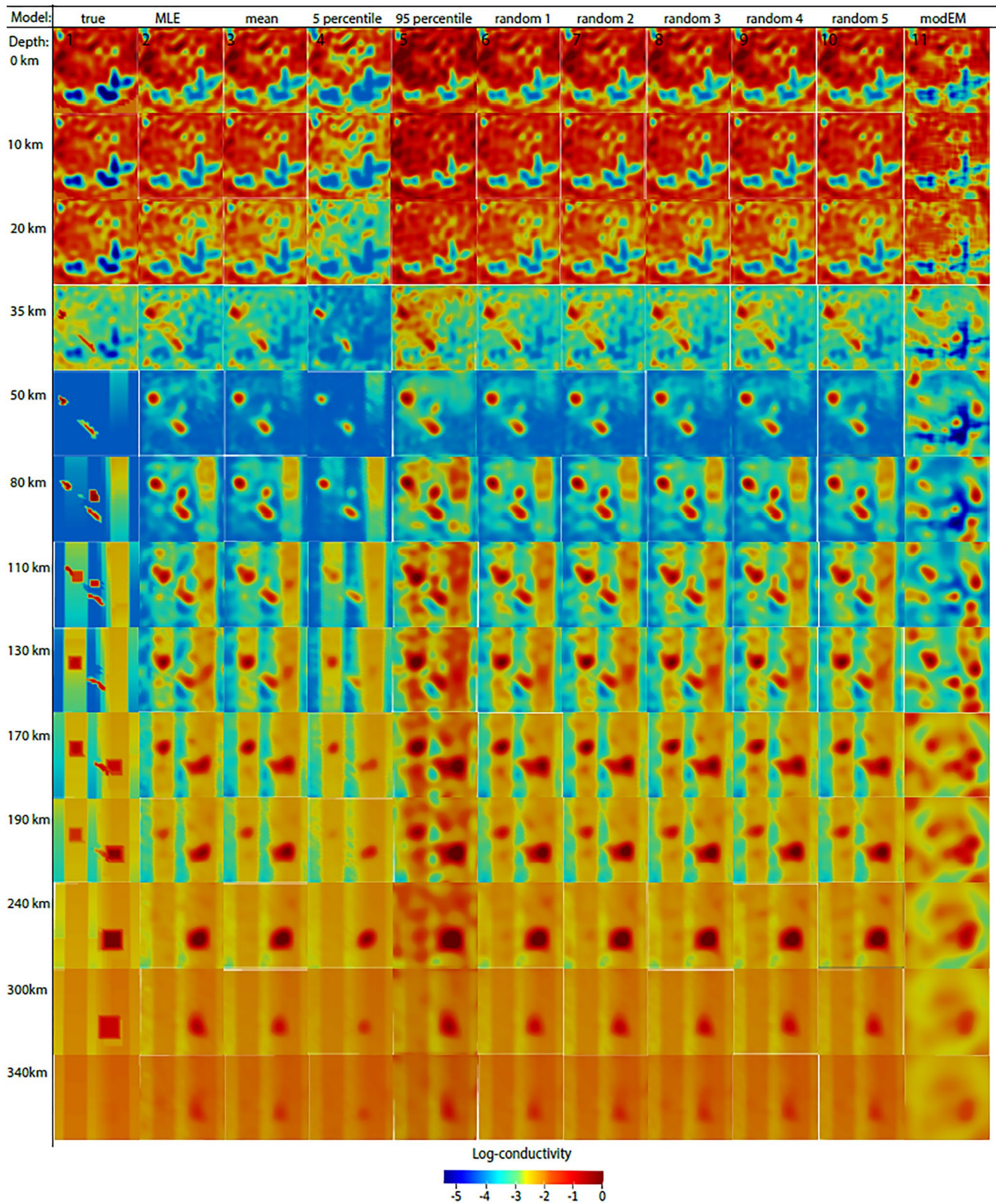
The *burn-in* period was set to 200,000 steps. This is larger than the length of the first stage (condition to maintain ergodicity of the chain) and the *burn-in* suggested by Geweke's convergence diagnostics (Geweke, 1992) for all parameters. Again, even with modest computational resources, the inversion took 27.4 days with an average of 1.9 s for each simulation. This represents a time reduction of  $\sim 95\%$  for each forward computation. We note that the average time spent in each simulation is higher than those in Example 1. This is mainly due to the large number of bases ( $\sim 190$ ) required in order to explain the complexity of this 3D model (see Figure 16).

The Maximum Likelihood Estimation (MLE) and mean models are shown in Figure 10 together with the 95% credibility intervals of the posterior PDF. The crustal and background conductivity structure and the location and volume of the conductivity anomalies are well resolved. Depth slices from the 95% credibility intervals, MLE, true and mean models are shown in Figure 11. In this figure we also include depth slices from five random subsets from the posterior, each obtained as the mean of 10 randomly chosen models from the entire ensemble of conductivity models making up the posterior PDF. By design, features that are well resolved by the inversion are persistent in all subsets, whereas poorly resolved features show higher variability among subsets (Tarantola, 2005). The identifiability of the background structure is also illustrated in Figure 12 where we show that the true LAB depths are close to the mean value of the marginal posterior PDFs. The sizes of the basis per frequency and the SW-MT data misfits for each of the 1,250,000 steps are shown in Figures 16 and 17, respectively, and show a similar pattern to those in Figure 6. It is worth noting that contrary to what we would expect from an inversion of MT data alone (see results in, e.g., Manassero et al., 2020; Rosas-Carbajal et al., 2013), model variability decreases with depth. The reason for this is the tighter constraints that the SW data puts on the background thermochemical structure.

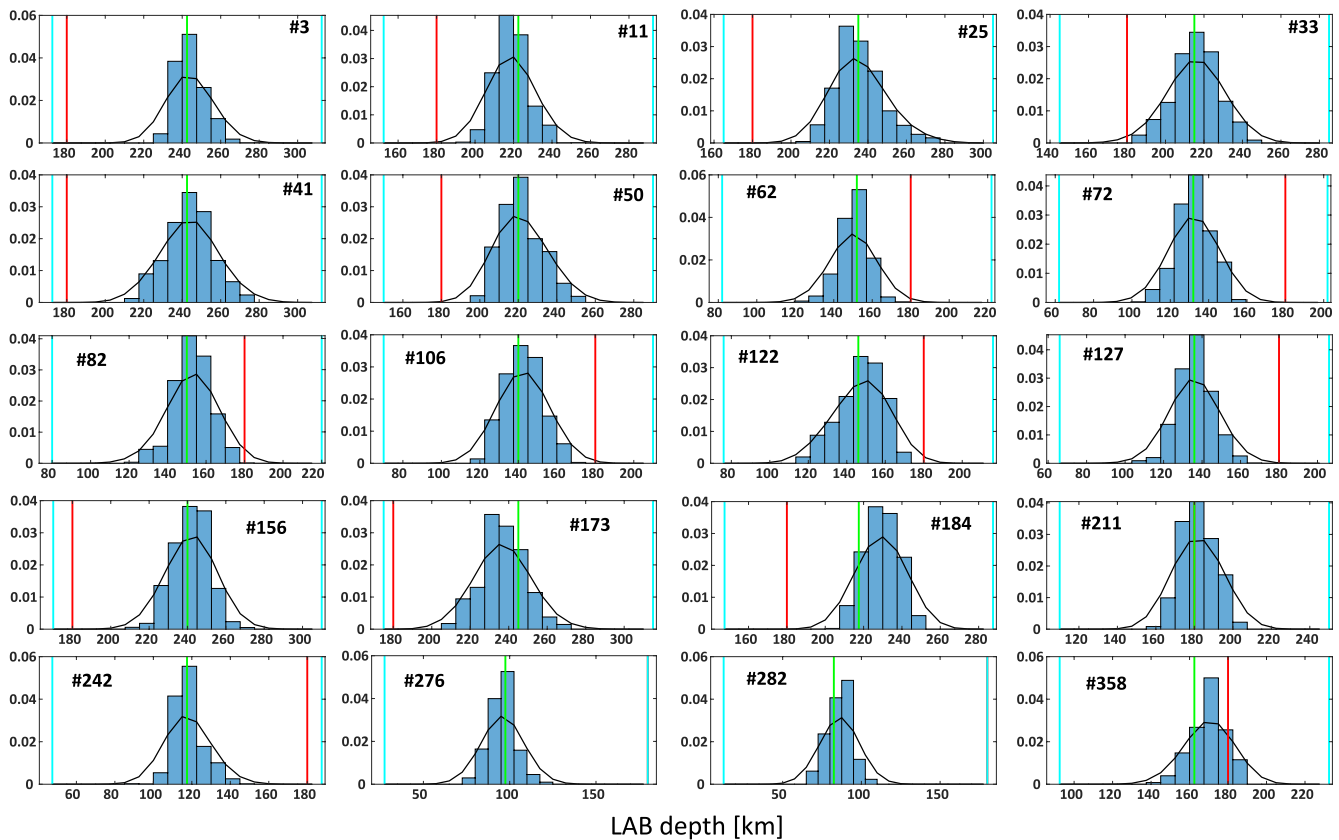
For comparison, Figure 11 includes the results obtained from a deterministic inversion using the software ModEM3DMT (Egbert & Kelbert, 2012; Kelbert et al., 2014) with the same initial model (used also as prior model), numerical discretization, and MT data as in the joint probabilistic inversion. Multiple inversions were run using different damping factors ( $\lambda$ ) and model covariance. Column (11) in Figure 11 shows depth-slices of the best



**Figure 10.** 3D views of the conductivity structure corresponding to the (a) *maximum likelihood estimation* model; (b) mean model; (c) lower, and (d) upper bound of the 95% credibility interval of the posterior PDF obtained after 1,250,000 Reduced Basis + Markov chain Monte Carlo (RB + MCMC) steps. The iso-surfaces of  $-1.5$  and  $-2.0 \log_{10} \text{S/m}$  are plotted in all panels to highlight the background structure and the conductivity anomalies in the mantle.



**Figure 11.** Columns (1–5): depth slices from the (1) true model; (2) MLE model, (3) mean and conductivity models corresponding to (4) the 5% percentile, and (5) the 95% percentile of the posterior PDF. Columns (6–10): depth slices for five mean models computed with 10 random samples of the posterior PDF. Columns (11): best model from a ModEM deterministic inversion. Selected depths are shown on the left of the figure.

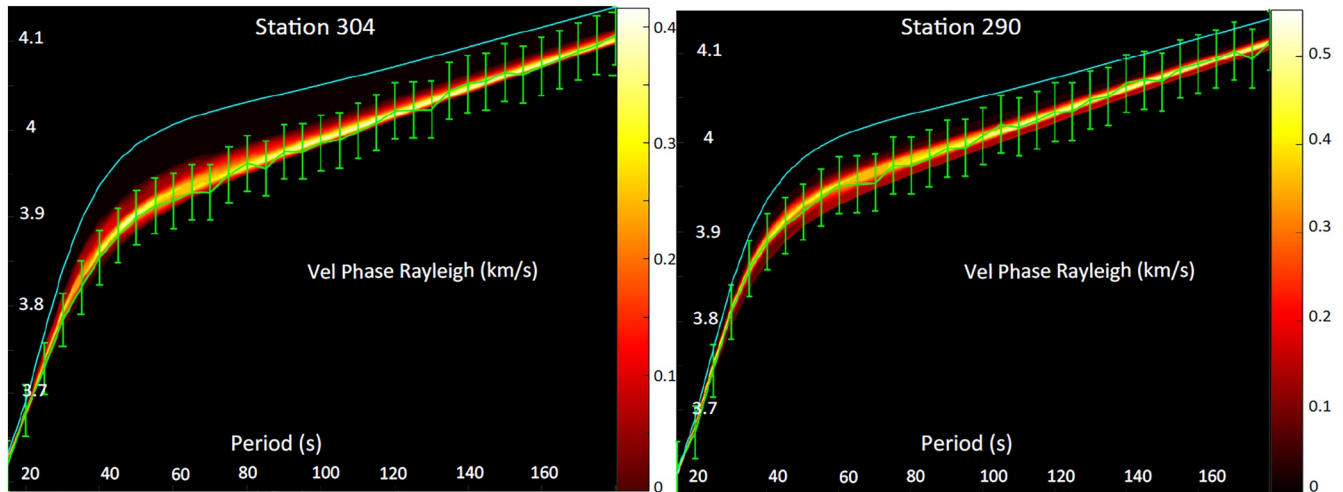


**Figure 12.** Marginal posterior probability density functions (PDFs) (blue bars) of 20 LAB depths obtained after 1,250,000 Reduced Basis + Markov chain Monte Carlo (RB + MCMC) steps. The real value, starting value, and prior bounds of each parameter are indicated by the green, red, and blue lines, respectively. The best Gaussian fits to the real PDFs given by the histograms are shown in black lines. The numbers within each panel correspond to the columns highlighted in Figure 9 (white squares).

model obtained after 82 iterations using  $\lambda = 1$  and covariance of 0.2. The inversion took 3.62 hs using 40 processors and the final data rms was 2.9. The main reason for this relatively large rms is the coarse mesh used in the inversion (same size as in the RB + MCMC inversion for a valid comparison); the effect of cell size on the rms is explored in Robertson et al. (2020) and Meqbel et al. (2014). We also observe that the recovered conductivity structure in the mantle is not satisfactory and overall more conductive than the true conductivity value. As shown in Robertson et al. (2020), this effect can also be minimized by decreasing the cell size.

Examples of the posterior PDFs of SW and MT data are shown in Figures 13, 14 and 15; additional posterior PDFs can be found in the Supporting Information S1. All of the dispersion data points are contained within one standard deviation of the posterior PDFs (e.g., Figure 13). This is also the case for the great majority of the MT data, although a poor data fit (or bias) is observed in some stations. As mentioned in Section 6.1.1, the noisy synthetic MT data is computed with the true conductivity model (Figure 8), whereas the conductivity models used in the actual inversion are derived from the interpolation of nodal values. This discrepancy or inadequacy between models and the considerable random noise that was added to the data are the main reasons of the poorer data fit seen at some stations (e.g., Smith, 2013).

The results from this example demonstrate that the joint probabilistic inversion of surface wave dispersion and MT data (a) is a practical option with modest computational resources, (b) succeeded in identifying the true LAB and conductivity structures in the crust and mantle (background plus anomalies), and (c) produced well behaved posterior distributions and global measures of uncertainty and correlations between model parameters.



**Figure 13.** Posterior probability density functions (PDFs) of Rayleigh wave dispersion curves for stations (a) ST304 and (b) ST290. Synthetic data and error bars are plotted in green and the computed data for the initial model is plotted in blue. The location of the stations are shown in Figure 9.

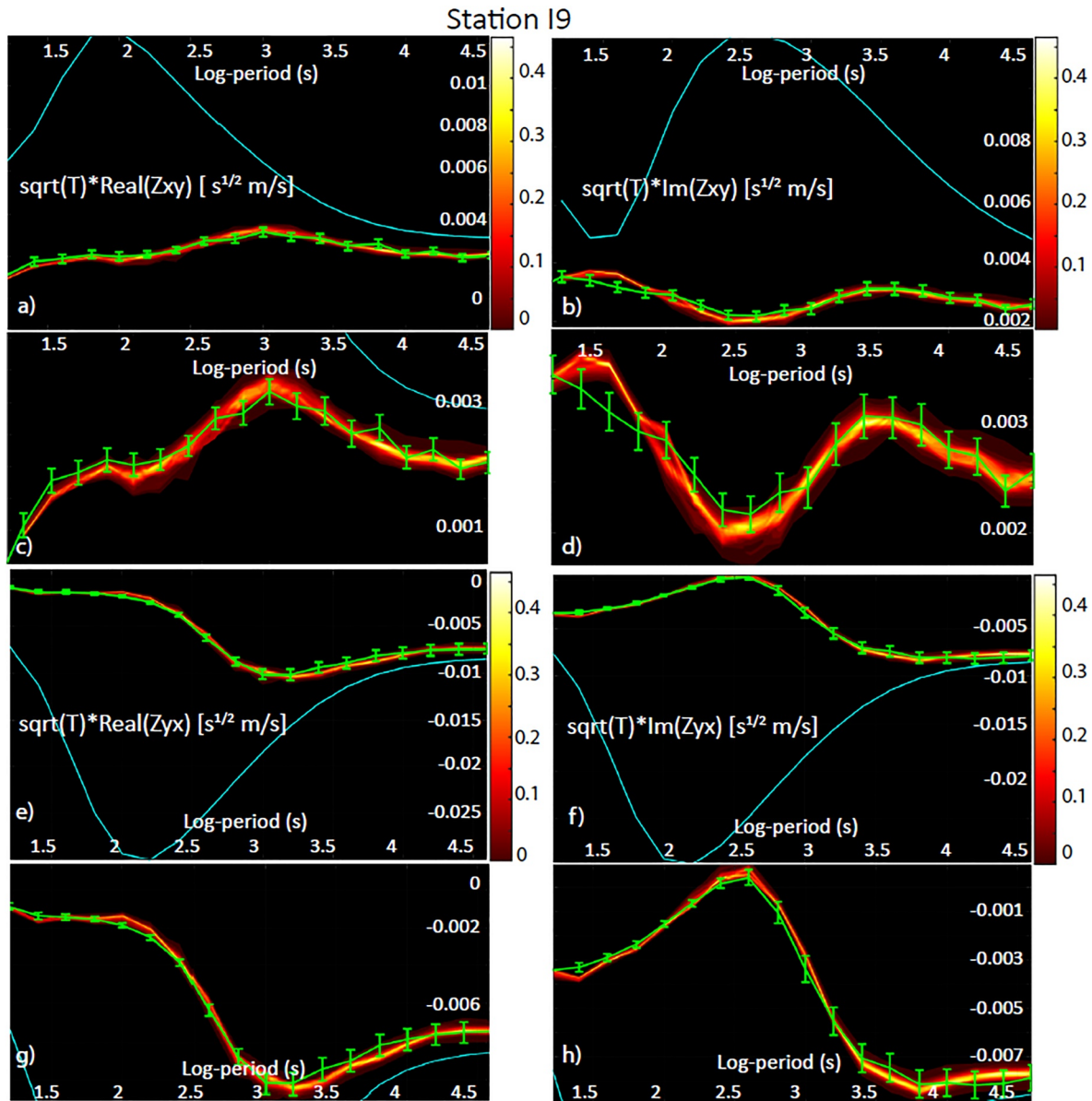
### 6.3. A Note on Crustal Velocity Structure and Bulk Mantle Composition

In all the numerical examples discussed so far, we considered a constant major-element composition for the mantle. This simplification seems appropriate in applications with emphasis on the general structure of the lithosphere, as the sensitivity of SW and MT to bulk major-element composition is of second-order compared to factors such as temperature and fluid content. If the mantle's major-element composition is of interest, other data sets such as gravity and/or geoid anomalies would be required (Afonso, Fullea, Griffin, et al., 2013; Afonso, Fullea, Yang, et al., 2013; Afonso, Rawlinson, et al., 2016).

In agreement with the main goal of this work -the deep lithospheric structure-, we have also assumed a fixed seismic structure for the crust (see e.g., Section 6.2.1). This would correspond to the case in which reliable prior information is available from previous studies such as ambient noise tomography and/or receiver function studies (e.g., Bello et al., 2021; Kennett & Salmon, 2012; Kennett et al., 2011; Young et al., 2013, in southeast Australia). A similar idea was applied to a real joint inversion by Jones et al. (2017) in Southern Africa. However, if the seismic structure of the crust needs to be inverted for, we can readily expand the vector of model parameters to include, e.g., the bulk density,  $V_s$  and  $V_p$  of the  $n$  layers used to discretize the crust in each 1D column. A more efficient alternative is to use the existing crustal conductivity nodes (e.g., Figure 8c) to invert for crustal seismic parameters as well. The bulk density,  $V_s$  and  $V_p$  values of each numerical cell within the crust are then obtained by interpolation using the same scheme as for  $\sigma$ . This option is particularly useful when the interstation spacings of both MT and seismic arrays are comparable. We are currently assessing these schemes and results will be presented in a forthcoming publication.

### 6.4. A Note on Model Discretization

The numerical examples discussed above have been discretized using finite elements with a horizontal dimension of  $25 \times 25$  km. While this is likely to be too coarse for real regional applications, Section 3 of the Supporting Information S1 includes tests that show that the current discretization is sufficient to represent the computed fields and the deep lithospheric structures of interest, while keeping the computational cost low. However, we emphasize that in order to ensure numerical accuracy of the forward computation, special care is needed when selecting the model discretization in real studies with more focus on shallow, small-scale structures.

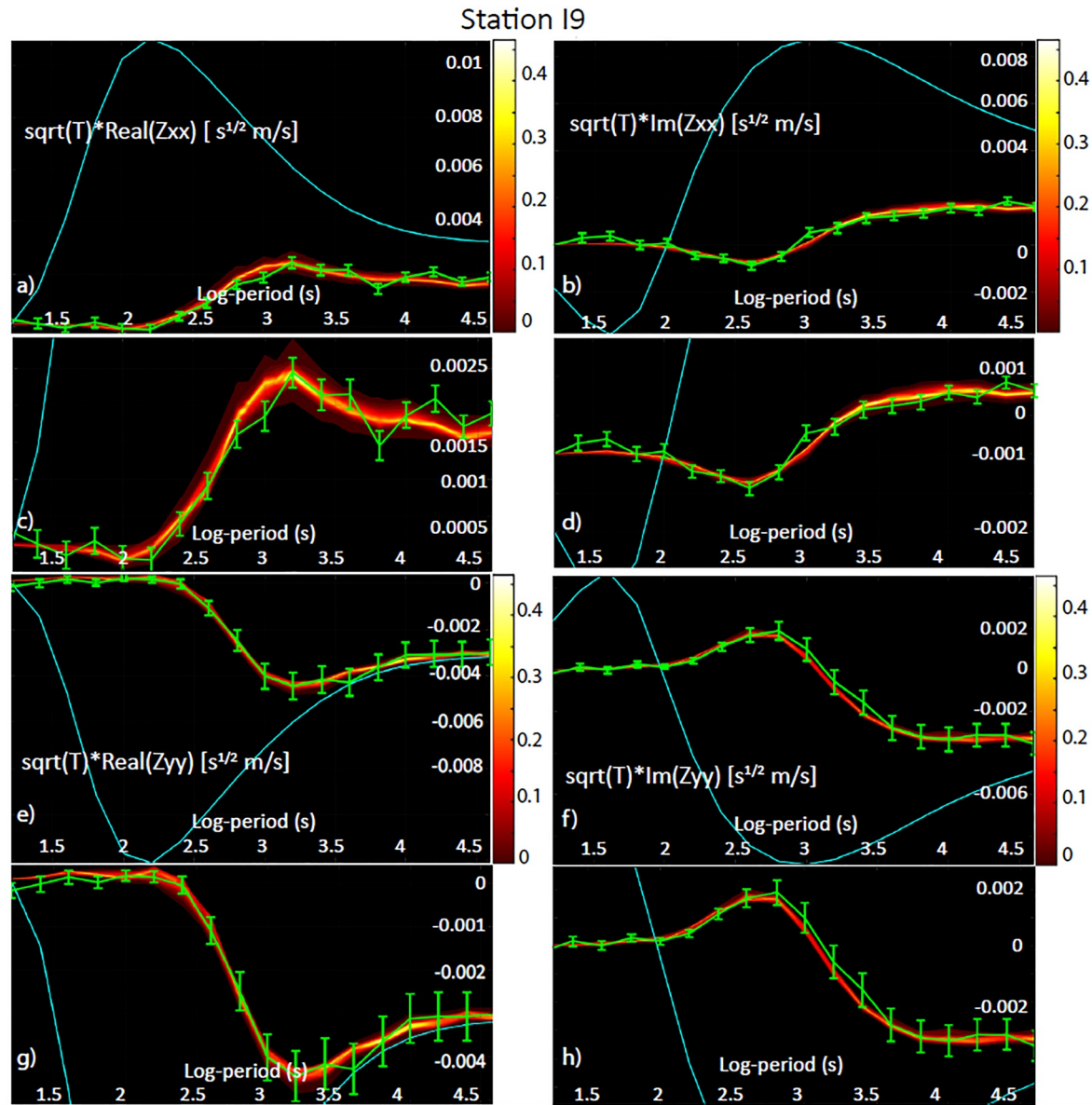


**Figure 14.** Posterior probability density functions (PDFs) of the off-diagonal components of the MT impedance tensor for station I9. Synthetic data and error bars are plotted in green and the computed data for the initial model is plotted in blue. Panels (a), (b), (e), and (f): Posterior PDFs of the real and imaginary parts of the off-diagonal components. A zoom of the PDFs and input data is shown in panels (c), (d), (g), and (h). In all panels, the data has been scaled by the square-root of the period. The location of the station is shown in Figure 9.

## 7. Conclusions

We presented a novel, MCMC-driven probabilistic joint inversion of 3D magnetotelluric (MT) and surface-wave (SW) dispersion data for imaging the electrical conductivity and temperature structures of the whole lithosphere and sublithospheric upper mantle. The method is based on (a) an efficient parallel-in-parallel structure to solve the 3D MT forward problem, (b) the combination of a reduced order, MCMC-driven strategy to compute fast and accurate surrogate solutions to the 3D MT forward problem, (c) adaptive strategies for both the MCMC algorithm and the surrogate, and (d) an efficient dual parameterization to couple both data sets.

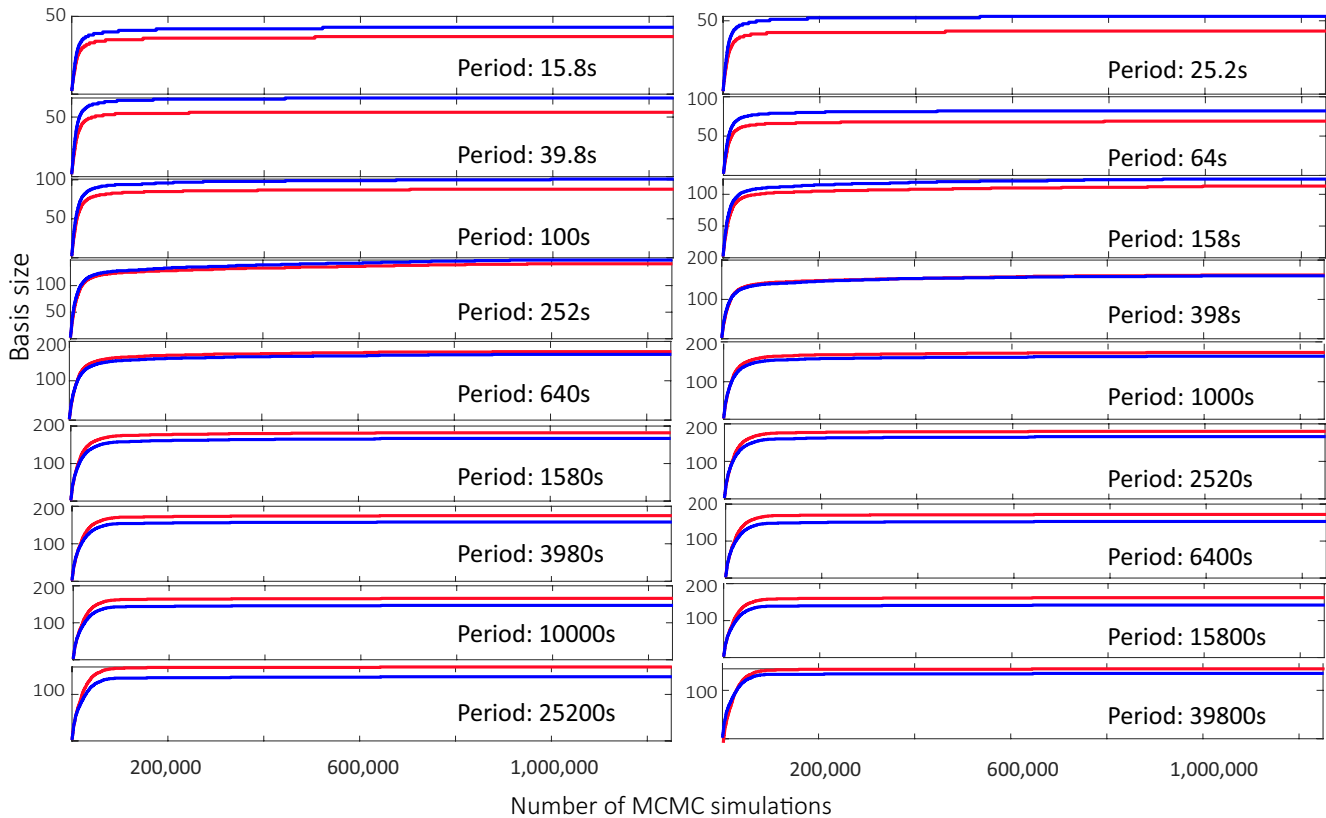




**Figure 15.** Posterior probability density functions (PDFs) of the diagonal components of the MT impedance tensor for station I9. Synthetic data and error bars are plotted in green and the computed data for the initial model is plotted in blue. Panels (a), (b), (e), and (f): Posterior PDFs of the real and imaginary parts of the diagonal components of the impedance tensor. A zoom of the PDFs and input data is shown in panels (c), (d), (g), and (h). In all panels, the data have been scaled by the square-root of the period. The location of the station is shown in Figure 9.

The feasibility, potential, and efficiency of our algorithm to solve the joint inverse problem are demonstrated with two realistic whole-lithosphere examples of increasing complexity. In both cases, we obtain staggering gains in computational efficiency (>96%) compared to a traditional MCMC implementation. Average times per MCMC step are of the order of 1 s, even when using modest computational resources. We also show that the inclusion of SW data and a simple Cascade-Metropolis algorithm resulted in drastic improvements in computational efficiency and quality of the recovered models compared to the RB + MCMC inversion of MT data only (Manassero et al., 2020).

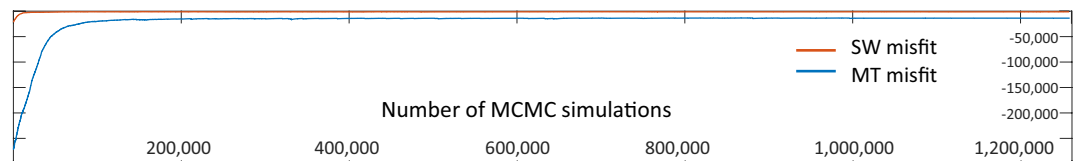
The model parameterization takes advantage of the differential sensitivities of MT and SW dispersion data to different aspects of the problem by using two sets of parameters. The first set corresponds to LAB depths, which control the large-scale (background) conductivity/velocity structure. The second set corresponds to



**Figure 16.** Basis size as a function of Reduced Basis + Markov chain Monte Carlo (RB + MCMC) steps for the 18 frequencies and field orientations used to compute the MT forward solution ( $S_1$  mode in blue, and  $S$  mode in red).

conductivity nodes inside the model, which control the small-scale conductivity anomalies. An additional advantage of using this parameterization is that a rapid convergence is achieved by using the LAB depths to constrain the first-order conductivity/velocity background at the beginning of the inversion. Once this first-order convergence has been achieved, the nodes are used to locally modify the background to fit the smaller-scale features of the data.

Finally, we note that proposed method is general enough to incorporate more advanced MCMC algorithms (e.g., trans-dimensional schemes, parallel tempering, and differential evolution), additional model parameters (e.g., bulk mantle composition) and other forward operators (e.g., gravity anomalies). An application to real MT and seismic data will be presented in a forthcoming publication (Part III).



**Figure 17.** Data misfits for dispersion curves (red line) and MT (blue line) for each of the 1,250,000 Reduced Basis + Markov chain Monte Carlo (RB + MCMC) steps.

## Appendix A: Mapping Thermochemical Parameters to Electrical Conductivity

The temperature dependence of electrical conductivity can be described with an Arrhenius-type Equation:

$$\sigma = \sigma_0 \exp\left(\frac{-\Delta H}{k_B T}\right), \quad (\text{A1})$$

where  $\sigma_0$  is the so-called pre-exponential factor,  $T$  [K] is absolute temperature, and  $k_B$  [eV/K] the Boltzmann's constant.  $\Delta H$  [eV] is the pressure-dependent activation enthalpy, defined as

$$\Delta H = \Delta E + P\Delta V, \quad (\text{A2})$$

where  $P$  is the pressure (GPa),  $\Delta E$ , and  $\Delta V$  are the activation energy and activation volume, respectively.

The main bulk conduction mechanisms in mantle minerals are ionic conduction, small polaron (hopping) conduction and proton conduction (e.g., Yoshino, 2010). Each mechanism follows an Arrhenius-type equation with particular activation energies depending on their charge mobility. These three conduction mechanisms can be integrated in a model for the electrical conductivity of mantle rocks as a function of pressure, temperature, water content, and composition (via Fe content) for each mineral phase (see also Fullea et al., 2011; Yoshino et al., 2009):

$$\sigma = \sigma_0 \exp\left(\frac{-\Delta H(X_{Fe}, P)}{k_B T}\right) + \sigma_{0i} \exp\left(\frac{-\Delta H_i}{k_B T}\right) + \sigma_p \quad (\text{A3a})$$

$$\sigma_p = f(C_w) \exp\left(\frac{-\Delta H_{wet}(C_w)}{k_B T}\right), \quad (\text{A3b})$$

$$-\Delta H(X_{Fe}, P) = a + bX_{Fe} + cX_{Fe}^2 + dX_{Fe}^3 + eX_{Fe}^4 + fX_{Fe}^5 + P\Delta V, \quad (\text{A3c})$$

where  $\sigma_0$ ,  $\sigma_{0i}$  (S/m), and  $f(C_w)$  are the small polaron, ionic, and proton pre-exponential factors, respectively,  $\Delta V$  (cm<sup>3</sup>/mol) is activation volume,  $\Delta H$ ,  $\Delta H_i$  (eV), and  $\Delta H_{wet}$  are activation enthalpies and  $X_{Fe}$  is the bulk Fe content in wt%.

The first term in the right-hand side of Equation A3a describes the contribution from small polaron conduction. As mentioned above, the activation enthalpy for this process depends on the iron content and pressure. This dependence is represented by a polynomial on  $X_{Fe}$  (Equation A3c) plus a term that depends on pressure (the coefficients  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$ ,  $f$  are determined experimentally). The second term of Equation A3a represents ionic conduction at high temperature and the third term ( $\sigma_p$ ) represents the proton conduction due to the presence of “water” (hydrogen diffusion).  $f(C_w)$  and  $\Delta H_{wet}$  are functions of the water content  $C_w$  (wt%) and they are obtained from laboratory experiments. The reader is referred to Fullea et al. (2011); Pommier (2014); Jones (2014); and Jones (2016) for a summary on the results from different laboratories Table A1.

**Table A1**  
Parameters Used to Compute Mantle Conductivity

Phase	$\sigma_0$	$\sigma_{0i}$	$a$	$b$	$c$	$d$	$e$	$F$	$\Delta V$	$\Delta H_i$	$X_{Fe}$
Olivine	2.70	4.73	1.64	0.246	-4.85	3.26			0.68	2.31	0.10
Opx	3.0		1.90	-2.77	2.61	-1.09					0.107
Cpx	3.25		2.07	-2.77	2.61	-1.09					$5.84 \times 10^{-2}$
Garnet		4.96	2.60	-15.33	80.40	-194.6	202.6	-75.0			0.168

## Appendix B: Kriging Interpolation

Kriging, or Gaussian process regression, is one of the most common methods for spatial interpolation (see e.g., Cressie, 1993; Gibbs, 1998; Gibbs & MacKay, 1997; Omre, 1987; Rasmussen, 1997; Williams & Rasmussen, 1996). The main idea is to predict (or interpolate) the value of a function  $Z$  at  $m$  locations from  $n$  observations by computing average spatial weights ( $W$ ). In simple kriging, these weights are derived using a known covariance function  $c$  between observations (given by the matrix  $K_{\text{obs}}$ ) and between the observations and the  $m$  estimation locations (given by the covariance matrix  $K_{\text{loc}}$ ):

$$W = K_{\text{obs}}^{-1} \cdot K_{\text{loc}}, \quad (\text{B1})$$

$$\text{where } K_{\text{obs}} = \begin{pmatrix} c(x_1^{\text{obs}}, x_1^{\text{obs}}) & \dots & c(x_1^{\text{obs}}, x_n^{\text{obs}}) \\ \dots & \dots & \dots \\ c(x_n^{\text{obs}}, x_1^{\text{obs}}) & \dots & c(x_n^{\text{obs}}, x_n^{\text{obs}}) \end{pmatrix} \text{ and } K_{\text{loc}} = \begin{pmatrix} c(x_1^{\text{obs}}, x_1^{\text{loc}}) & \dots & c(x_1^{\text{obs}}, x_m^{\text{loc}}) \\ \dots & \dots & \dots \\ c(x_n^{\text{obs}}, x_1^{\text{loc}}) & \dots & c(x_n^{\text{obs}}, x_m^{\text{loc}}) \end{pmatrix}.$$

The interpolation (or estimated value) at the  $m$  locations is then given by  $Z^{\text{loc}} = W \cdot Z^{\text{obs}}$ , where  $Z^{\text{obs}}$  is the vector containing the  $n$  observations.

The covariance function  $c$  can take any form with the only constrain that it must generate a nonnegative definite covariance matrix. A common form is given by (e.g., Gibbs & MacKay, 1997):

$$c(\mathbf{x}_m, \mathbf{x}_n) = \theta_1 \exp\left(-\frac{1}{2} \sum_l \frac{(x_m^l - x_n^l)^2}{r_l^2}\right) + \theta_2, \quad (\text{B2})$$

where  $x_n^l$  is the  $l$  component of  $\mathbf{x}_n$ .  $\theta_1$  and  $\theta_2$  are hyperparameters, where  $\theta_1$  represents the overall vertical scale relative to the mean field and  $\theta_2$  gives the vertical uncertainty.  $r_l$  is the correlation or scale length and it characterizes the distance in the direction  $l$  over which the value of  $Z$  varies significantly. It should be noted that since the spatial weights ( $W$ ) depends on the covariance function  $c$ , the interpolated values at the  $m$  locations also depends on the chosen form for  $c$ .

### B1. Spatially Varying Length Scales

The covariance function of Equation B2 assumes that the correlation length ( $r_l$ ) is fixed in each direction ( $l$ ) and location ( $\mathbf{x}$ ). In the most general case, however, assuming a fixed  $r_l$  might lead to a simplistic and poor representation of the conductivity model. We, therefore, use a positive definite covariance function with spatially variable correlation lengths (Gibbs, 1998; Gibbs & MacKay, 1997):

$$c(\mathbf{x}_m, \mathbf{x}_n) = \theta_1 \prod_l \left( \frac{2r_l(\mathbf{x}_m)r_l(\mathbf{x}_n)}{r_l^2(\mathbf{x}_m) + r_l^2(\mathbf{x}_n)} \right)^{1/2} \exp\left(-\sum_l \frac{(x_m^l - x_n^l)^2}{r_l^2(\mathbf{x}_m) + r_l^2(\mathbf{x}_n)}\right) \quad (\text{B3})$$

where  $r_l(\mathbf{x})$  is an arbitrary parameterized function of position  $\mathbf{x}$  defined in  $[-1, 1]^2 \times [0, 1]$ . The form of  $r_l(\mathbf{x})$  as a function of the scaled coordinates ( $x, y, z$ ) used in Examples 1 and 2 in the main text is shown in Procedure 1. This covariance function has the property that the variance is independent of  $\mathbf{x}$  and equal to  $\theta_1$ . Since a change in  $\theta_1$  will produce changes in the vertical scale in the whole domain (see previous section), the inclusion of  $\theta_1$  as an additional parameter of the inversion can (potentially) benefit the efficiency and convergence of the algorithm. The implementation of  $\theta_1$  as a hyper-parameter of the inversion is left for future work.

## Appendix C: Log-Normal Proposal Distributions

The log-normal distribution (Gaussian in log-scale) used in the second stage is defined as:

$$y(m_i^j) = \frac{1}{\sqrt{2\pi m_i^j s}} \exp\left(-\frac{\ln(m_i^j) - \mu^2}{2s^2}\right), \quad (\text{C1})$$

where  $y(m_i^j)$  is the proposed value for node  $i$ , and  $\mu$  and  $s$  are the mean and standard deviation in log-scale.

In Section 5, we have chosen to define a multivariate Gaussian proposal of dimension  $N_{\text{nodes}} \times N_{\text{nodes}}$ , where  $N_{\text{nodes}}$  is the number of conductivity nodes in the model. Since the nodes' conductivity values can span several orders of magnitude, the Gaussian proposal is defined in log-scale but we evaluate its probability  $q(\cdot)$  in linear scale, that is, a multivariate log-normal PDF centered at the current state  $\mathbf{m}_{t-1}$  with covariance  $\Sigma$ :

$$q(\mathbf{m}_t | \mathbf{m}_{t-1}) = \frac{1}{(2\pi)^{\frac{N_{\text{nodes}}}{2}} (\det \Sigma)^{\frac{1}{2}} \prod_{j=1}^{N_{\text{nodes}}} m_t^j} \exp \left[ -\frac{1}{2} (\ln(\mathbf{m}_t) - \ln(\mathbf{m}_{t-1}))^T \Sigma^{-1} (\ln(\mathbf{m}_t) - \ln(\mathbf{m}_{t-1})) \right], \quad (\text{C2})$$

where  $\mathbf{m}_t$  is the proposed value for all nodes and  $\mathbf{m}_{t-1}$  is the current sample.

**Algorithm 1. Definition of  $r_l(x)$  as a function of the scaled coordinates  $(x, y, z)$**

```

procedure  $r_l(x)$ 
  if  $z >= 0.9$  then
     $r_3 = 0.5$ 
     $r_2 = r_1 = 0.48$ 
  else if  $z < 0.9$  and  $z >= 0.8$  then
     $r_3 = 0.45$ 
     $r_2 = r_1 = 0.43$ 
  else if  $z < 0.8$  and  $z >= 0.7$  then
     $r_3 = 0.4$ 
     $r_2 = r_1 = 0.4$ 
  else if  $z < 0.7$  and  $z >= 0.6$  then
     $r_3 = 0.4$ 
     $r_2 = r_1 = 0.38$ 
  else if  $z < 0.6$  and  $z >= 0.5$  then
     $r_3 = 0.35$ 
     $r_2 = r_1 = 0.33$ 
  else if  $z < 0.5$  and  $z >= 0.4$  then
     $r_3 = 0.33$ 
     $r_2 = r_1 = 0.3$ 
  else if  $z < 0.4$  and  $z >= 0.3$  then
     $r_3 = 0.3$ 
     $r_2 = r_1 = 0.28$ 
  else if  $z < 0.3$  and  $z >= 0.2$  then
     $r_3 = 0.28$ 
     $r_2 = r_1 = 0.24$ 
  else if  $z < 0.2$  and  $z >= 0.1$  then
     $r_3 = 0.25$ 
     $r_2 = r_1 = 0.23$ 
  else if  $z < 0.1$  then
     $r_3 = 0.2$ 
     $r_2 = r_1 = 0.15$ 
  end if
end procedure

```

## Data Availability Statement

The authors acknowledge that data were not used, nor created for this research. The 3D rendering views were created using ParaView (Ahrens et al., 2005).

## Acknowledgments

They thank editor Michael Bostock for handling this manuscript, Max Moorkamp and two anonymous reviewers for their thorough comments and suggestions. They thank Farshad Salajegheh for providing his Matlab codes for plotting results. Special thanks to Sinan Özyaydin, Kate Selway, and Anandaroop Ray for their suggestions at different stages of this work. M. C. Manassero thanks support from an International Macquarie Research Excellence Scholarship (iMQRES). M. C. Manassero and J. C. Afonso acknowledge support from ARC Grant DP160103502, ARC Linkage Grant LP170100233, the ARC Centre of Excellence Core to Crust Fluids Systems (<http://www.ccfsc.mq.edu.au>), and the Centre for Earth Evolution and Dynamics, Geoscience Australia and the European Space Agency via the “3D Earth - A Dynamic Living Planet.” F. I. Zyserman acknowledges support from CONICET through grant PIP 112-201501-00192. S. Zlotnik has been funded by the Spanish Ministry through grant DPI2017-85139-C2-2-R, by the Catalan government through grant 2017-SGR-1278 and by the EU’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 777778.

## References

- Afonso, J. C., Fernandez, M., Ranalli, G., Griffin, W., & Connolly, J. (2008). Integrated geophysical-petrological modeling of the lithosphere and sublithospheric upper mantle: Methodology and applications. *Geochemistry, Geophysics, Geosystems*, 9(5). <https://doi.org/10.1029/2007gc001834>
- Afonso, J. C., Fullea, J., Griffin, W., Yang, Y., Jones, A., Connolly, J., & O’Reilly, S. (2013). 3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle. I: A priori petrological information and geophysical observables. *Journal of Geophysical Research*, 118(5), 2586–2617. <https://doi.org/10.1002/jgrb.50124>
- Afonso, J. C., Fullea, J., Yang, Y., Connolly, J., & Jones, A. (2013). 3-D multi-observable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle. II: General methodology and resolution analysis. *Journal of Geophysical Research*, 118(4), 1650–1676. <https://doi.org/10.1002/jgrb.50123>
- Afonso, J. C., Moorkamp, M., & Fullea, J. (2016). Imaging the lithosphere and upper mantle: Where we are at and where we are going. In N. L. M. Moorkamp, P. Lelievre, & A. Khan (Eds.), *Integrated imaging of the earth: Theory and applications* (pp. 191–218). John Wiley & Sons. <https://doi.org/10.1002/9781118929063.ch10>
- Afonso, J. C., Ranalli, G., & Fernández, M. (2005). Thermal expansivity and elastic properties of the lithospheric mantle: Results from mineral physics of composites. *Physics of the Earth and Planetary Interiors*, 149(3–4), 279–306. <https://doi.org/10.1016/j.pepi.2004.10.003>
- Afonso, J. C., Ranalli, G., Fernández, M., Griffin, W. L., O’Reilly, S. Y., & Faul, U. (2010). On the Vp/Vs–Mg# correlation in mantle peridotites: Implications for the identification of thermal and compositional anomalies in the upper mantle. *Earth and Planetary Science Letters*, 289(3–4), 606–618. <https://doi.org/10.1016/j.epsl.2009.12.005>
- Afonso, J. C., Rawlinson, N., Yang, Y., Schutt, D. L., Jones, A. G., Fullea, J., & Griffin, W. L. (2016). 3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle: III. Thermochemical tomography in the Western-Central US. *Journal of Geophysical Research: Solid Earth*, 121(10), 7337–7370. <https://doi.org/10.1002/2016jb013049>
- Ahrens, J., Geveci, B., & Law, C. (2005). Paraview: An end-user tool for large data visualization. *The visualization handbook* (pp. 717–731). <https://doi.org/10.1016/b978-012387582-2/50038-1>
- Amestoy, P. R., Duff, I. S., L’Excellent, J.-Y., & Koster, J. (2001). A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM Journal on Matrix Analysis and Applications*, 23(1), 15–41. <https://doi.org/10.1137/s0895479899358194>
- Amestoy, P. R., Guermouche, A., L’Excellent, J.-Y., & Pralet, S. (2006). Hybrid scheduling for the parallel solution of linear systems. *Parallel Computing*, 32(2), 136–156. <https://doi.org/10.1016/j.parco.2005.07.004>
- Bello, M., Cornwell, D. G., Rawlinson, N., Reading, A. M., & Likkason, O. K. (2021). Crustal structure of southeast Australia from teleseismic receiver functions. *Solid Earth*, 12, 463–481. <https://doi.org/10.5194/se-12-463-2021>
- Bennington, N. L., Zhang, H., Thurber, C. H., & Bedrosian, P. A. (2015). Joint inversion of seismic and magnetotelluric data in the Parkfield Region of California using the normalized cross-gradient constraint. *Pure and Applied Geophysics*, 172(5), 1033–1052. <https://doi.org/10.1007/s00024-014-1002-9>
- Bensen, G., Ritzwoller, M., & Yang, Y. (2009). A 3-D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise. *Geophysical Journal International*, 177(3), 1177–1196. <https://doi.org/10.1111/j.1365-246x.2009.04125.x>
- Birch, F. (1961). Composition of the Earth’s mantle. *Geophysical Journal International*, 4(1), 295–311. <https://doi.org/10.1111/j.1365-246x.1961.tb06821.x>
- Birch, F. (1964). Density and composition of mantle and core. *Journal of Geophysical Research*, 69(20), 4377–4388. <https://doi.org/10.1029/jz069i020p04377>
- Bodin, T., & Sambridge, M. (2009). Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, 178(3), 1411–1436. <https://doi.org/10.1111/j.1365-246x.2009.04226.x>
- Bosch, M. (1999). Lithologic tomography: From plural geophysical data to lithology estimation. *Journal of Geophysical Research*, 104(B1), 749–766. <https://doi.org/10.1029/1998jb900014>
- Brodie, R., & Jiang, W. (2018). Trans-dimensional Monte Carlo inversion of short period magnetotelluric data for cover thickness estimation. *ASEG Extended Abstracts*, 2018(1), 1–7. [https://doi.org/10.1071/aseg2018abt5\\_1f](https://doi.org/10.1071/aseg2018abt5_1f)
- Calvetti, D., & Somersalo, E. (2007). *An introduction to bayesian scientific computing: Ten lectures on subjective computing*. Springer Science & Business Media.
- Christen, J. A., & Fox, C. (2005). Markov chain Monte Carlo using an approximation. *Journal of Computational & Graphical Statistics*, 14(4), 795–810. <https://doi.org/10.1198/106186005x76983>
- Conrad, P. R., Marzouk, Y. M., Pillai, N. S., & Smith, A. (2016). Accelerating asymptotically exact MCMC for computationally intensive models via local approximations. *Journal of the American Statistical Association*, 111(516), 1591–1607. <https://doi.org/10.1080/01621459.2015.1096787>
- Cressie, N. (1993). *Statistics for spatial data* (Revised Edition ed.). Wiley-Interscience.
- Cui, T., Marzouk, Y. M., & Willcox, K. E. (2015). Data-driven model reduction for the Bayesian solution of inverse problems. *International Journal for Numerical Methods in Engineering*, 102(5), 966–990. <https://doi.org/10.1002/nme.4748>
- Douglas, J., Jr., Santos, J. E., & Sheen, D. (2000). A nonconforming mixed finite element method for Maxwell’s Equations. *Mathematical Models and Methods in Applied Sciences*, 10(04), 593–613. <https://doi.org/10.1142/s02182050000032x>
- Douglas, J., Jr., Santos, J. E., Sheen, D., & Ye, X. (1999). Nonconforming galerkin methods based on quadrilateral elements for second order elliptic problems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 33(4), 747–770. <https://doi.org/10.1051/m2an:1999161>
- Egbert, G. D., & Kelbert, A. (2012). Computational recipes for electromagnetic inverse problems. *Geophysical Journal International*, 189(1), 251–267. <https://doi.org/10.1111/j.1365-246x.2011.05347.x>
- Evans, R. (2012). Conductivity of Earth materials. In J. A. Chave A (Ed.), *The magnetotelluric method, theory and practice* (pp. 50–95). Cambridge University Press.

- Evans, R., Benoit, M. H., Long, M. D., Elsenbeck, J., Ford, H. A., Zhu, J., & Garcia, X. (2019). Thin lithosphere beneath the central Appalachian Mountains: A combined seismic and magnetotelluric study. *Earth and Planetary Science Letters*, 519, 308–316. <https://doi.org/10.1016/j.epsl.2019.04.046>
- Farquharson, C. G., & Oldenburg, D. W. (1998). Non-linear inversion using general measures of data misfit and model structure. *Geophysical Journal International*, 134(1), 213–227. <https://doi.org/10.1046/j.1365-246x.1998.00555.x>
- Feng, R., Yan, H., & Zhang, R. (1986). Fast inversion method and corresponding programming for 3D potential field. *Acta Geologica Sinica*, 4(3), 390–402.
- Florentin, E., & Diez, P. (2012). Adaptive reduced basis strategy based on goal oriented error assessment for stochastic problems. *Computer Methods in Applied Mechanics and Engineering*, 225, 116–127. <https://doi.org/10.1016/j.cma.2012.03.016>
- Frangos, M., Marzouk, Y., Willcox, K., & van Bloemen Waanders, B. (2011). Surrogate and reduced-order modeling: A comparison of approaches for large-scale statistical inverse problems. In O. G. M. H. D. K. B. M. Y. M. L. T. B. v. B. W. L. G. BieglerBiros, & K. Willcox (Eds.), *Computational methods for large-scale inverse problems and quantification of uncertainty* (pp. 266–290). John Wiley & Sons.
- Fullea, J., Muller, M., & Jones, A. (2011). Electrical conductivity of continental lithospheric mantle from integrated geophysical and petrological modeling: Application to the Kaapvaal Craton and Rehoboth Terrane, southern Africa. *Journal of Geophysical Research*, 116(B10), 94–105. <https://doi.org/10.1029/2011JB008544>
- Gallardo, L. A., & Meju, M. A. (2007). Joint two-dimensional cross-gradient imaging of magnetotelluric and seismic traveltime data for structural and lithological classification. *Geophysical Journal International*, 169(3), 1261–1272. <https://doi.org/10.1111/j.1365-246x.2007.03366.x>
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian statistics* (pp. 169–193). Oxford University Press.
- Gibbs, M. (1998). *Bayesian Gaussian processes for regression and classification*. (Unpublished doctoral dissertation). Citeseer.
- Gibbs, M., & MacKay, D. J. (1997). *Efficient implementation of Gaussian processes*. Neural Computation.
- Gilks, W. R., Richardson, S., & Spiegelhalter, D. (1995). *Markov chain Monte Carlo in practice*. Chapman and Hall/CRC.
- Gouveia, W. P., & Scales, J. A. (1998). Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis. *Journal of Geophysical Research*, 103(B2), 2759–2779. <https://doi.org/10.1029/97jb02933>
- Gregory, P. (2005). *Bayesian logical data analysis for the physical sciences: A comparative approach with Mathematica® support*. Cambridge University Press.
- Guerri, M., Cammarano, F., & Tackley, P. J. (2016). Modelling Earth's surface topography: Decomposition of the static and dynamic components. *Physics of the Earth and Planetary Interiors*, 261, 172–186. <https://doi.org/10.1016/j.pepi.2016.10.009>
- Haario, H., Laine, M., Mira, A., & Saksman, E. (2006). DRAM: Efficient adaptive MCMC. *Statistics and Computing*, 16(4), 339–354. <https://doi.org/10.1007/s11222-006-9438-0>
- Haario, H., Saksman, E., Tamminen, J., et al. (2001). An adaptive Metropolis algorithm. *Bernoulli*, 7(2), 223–242. <https://doi.org/10.2307/3318737>
- Hashin, Z., & Shtrikman, S. (1962). A variational approach to the theory of the effective magnetic permeability of multiphase materials. *Journal of Applied Physics*, 33(10), 3125–3131. <https://doi.org/10.1063/1.1728579>
- Hashin, Z., & Shtrikman, S. (1963). A variational approach to the theory of the elastic behaviour of multiphase materials. *Journal of the Mechanics and Physics of Solids*, 11(2), 127–140. [https://doi.org/10.1016/0022-5096\(63\)90060-7](https://doi.org/10.1016/0022-5096(63)90060-7)
- Hassani, B., & Renaudin, A. (2013). *The cascade Bayesian approach for a controlled integration of internal data, external data and scenarios*.
- Hassani, B., & Renaudin, A. (2018). The cascade Bayesian approach: Prior transformation for a controlled integration of internal data, external data and scenarios. *Risks*, 6(2), 47. <https://doi.org/10.3390/risks6020047>
- Herrmann, R. B., & Ammon, C. J. (2002). *Computer programs in seismology: Surface waves, receiver functions and crustal structure*. St Louis University.
- Hesthaven, J. S., Rozza, G., & Stamm, B. (2016). *Certified reduced basis methods for parametrized partial differential Equations*. Springer.
- Hu, H., Dai, L., Li, H., Hui, K., & Sun, W. (2017). Influence of dehydration on the electrical conductivity of epidote and implications for high-conductivity anomalies in subduction zones. *Journal of Geophysical Research: Solid Earth*, 122(4), 2751–2762. <https://doi.org/10.1002/2016jb013767>
- Huang, Z., Li, H., Zheng, Y., & Peng, Y. (2009). The lithosphere of North China craton from surface wave tomography. *Earth and Planetary Science Letters*, 288(1–2), 164–173. <https://doi.org/10.1016/j.epsl.2009.09.019>
- Jackson, I., & Faul, U. H. (2010). Grain-size-sensitive viscoelastic relaxation in olivine: Towards a robust laboratory-based model for seismological application. *Physics of the Earth and Planetary Interiors*, 183(1–2), 151–163. <https://doi.org/10.1016/j.pepi.2010.09.005>
- Jackson, I., Fitz Gerald, J. D., Faul, U. H., & Tan, B. H. (2002). Grain-size-sensitive seismic wave attenuation in polycrystalline olivine. *Journal of Geophysical Research*, 107(B12). <https://doi.org/10.1029/2001jb001225>
- Jegen, M. D., Hobbs, R. W., Tarits, P., & Chave, A. (2009). Joint inversion of marine magnetotelluric and gravity data incorporating seismic constraints: Preliminary results of sub-basalt imaging off the Faroe Shelf. *Earth and Planetary Science Letters*, 282(1–4), 47–55. <https://doi.org/10.1016/j.epsl.2009.02.018>
- Jones, A. G. (2014). Reconciling different equations for proton conduction using the Meyer-Neldel compensation rule. *Geochemistry, Geophysics, Geosystems*, 15(2), 337–349. <https://doi.org/10.1002/2013gc004911>
- Jones, A. G. (2016). Proton conduction and hydrogen diffusion in olivine: An attempt to reconcile laboratory and field observations and implications for the role of grain boundary diffusion in enhancing conductivity. *Physics and Chemistry of Minerals*, 43(4), 237–265. <https://doi.org/10.1007/s00269-015-0790-5>
- Jones, A. G., Afonso, J. C., & Fullea, J. (2017). Geochemical and geophysical constraints on the dynamic topography of the Southern African Plateau. *Geochemistry, Geophysics, Geosystems*, 18(10), 3556–3575. <https://doi.org/10.1002/2017gc006908>
- Jones, A. G., Evans, R. L., & Eaton, D. W. (2009). Velocity–conductivity relationships for mantle mineral assemblages in Archean cratonic lithosphere based on a review of laboratory data and Hashin–Shtrikman extremal bounds. *Lithos*, 109(1–2), 131–143. <https://doi.org/10.1016/j.lithos.2008.10.014>
- Kaipio, J., & Somersalo, E. (2006). *Statistical and computational inverse problems*. Springer Science & Business Media.
- Karato, S.-i. (1990). The role of hydrogen in the electrical conductivity of the upper mantle. *Nature*, 347(6290), 272–273. <https://doi.org/10.1038/347272a0>
- Karato, S.-i. (2006). Remote sensing of hydrogen in Earth's mantle. *Reviews in Mineralogy and Geochemistry*, 62(1), 343–375. <https://doi.org/10.2138/rmg.2006.62.15>
- Karato, S.-i., & Wang, D. (2013). Electrical conductivity of minerals and rocks. *Physics and Chemistry of the Deep Earth*, 5, 145–182. <https://doi.org/10.1002/9781118529492.ch5>
- Kelbert, A., Meqbel, N., Egbert, G. D., & Tandon, K. (2014). ModEM: A modular system for inversion of electromagnetic geophysical data. *Computers & Geosciences*, 66, 40–53. <https://doi.org/10.1016/j.cageo.2014.01.010>

- Kennett, B., & Salmon, M. (2012). Ausrem: Australian seismological reference model. *Australian Journal of Earth Sciences*, 59(8), 1091–1103. <https://doi.org/10.1080/08120099.2012.736406>
- Kennett, B., Salmon, M., Saygin, E., & Group, A. W. (2011). AusMoho: The variation of Moho depth in Australia. *Geophysical Journal International*, 187(2), 946–958. <https://doi.org/10.1111/j.1365-246x.2011.05194.x>
- Khan, A. (2016). On Earth's mantle constitution and structure from joint analysis of geophysical and laboratory-based data: An example. *Surveys in Geophysics*, 37(1), 149–189. <https://doi.org/10.1007/s10712-015-9353-z>
- Khan, A., Connolly, J., & Olsen, N. (2006). Constraining the composition and thermal state of the mantle beneath Europe from inversion of long-period electromagnetic sounding data. *Journal of Geophysical Research*, 111(B10). <https://doi.org/10.1029/2006jb004270>
- Khan, A., Connolly, J., & Taylor, S. (2008). Inversion of seismic and geodetic data for the major element chemistry and temperature of the Earth's mantle. *Journal of Geophysical Research*, 113(B9), B09308. <https://doi.org/10.1029/2007JB005239>
- Kirkby, A. L., Musgrave, R. J., Czarnota, K., Doublier, M. P., Duan, J., Cayley, R. A., & Kyi, D. (2020). Lithospheric architecture of a Phanerozoic orogen from magnetotellurics: AusLAMP in the Tasmanides, southeast Australia. *Tectonophysics*, 793, 228560. <https://doi.org/10.1016/j.tecto.2020.228560>
- Lebedev, S., & Van Der Hilst, R. D. (2008). Global upper-mantle tomography with the automated multimode inversion of surface and S-wave forms. *Geophysical Journal International*, 173(2), 505–518. <https://doi.org/10.1111/j.1365-246x.2008.03721.x>
- Ledo, J., & Jones, A. G. (2005). Upper mantle temperature determined from combining mineral composition, electrical conductivity laboratory studies and magnetotelluric field observations: Application to the intermontane belt, northern Canadian cordillera. *Earth and Planetary Science Letters*, 236(1–2), 258–268. <https://doi.org/10.1016/j.epsl.2005.01.044>
- Liu, H.-P., Anderson, D. L., & Kanamori, H. (1976). Velocity dispersion due to anelasticity; implications for seismology and mantle composition. *Geophysical Journal International*, 47(1), 41–58. <https://doi.org/10.1111/j.1365-246x.1976.tb01261.x>
- Mallick, K., & Verma, R. (1979). Time-domain electromagnetic sounding—Computation of multi-layer response and the problem of equivalence in interpretation. *Geophysical Prospecting*, 27(1), 137–155. <https://doi.org/10.1111/j.1365-2478.1979.tb00962.x>
- Manassero, M. C. (2019). *A reduced order approach for probabilistic inversions of 3d magnetotelluric data (Unpublished doctoral dissertation)*. Macquarie University.
- Manassero, M. C., Afonso, J. C., Zyserman, F., Zlotnik, S., & Fomin, I. (2020). A Reduced Order Approach for Probabilistic Inversions of 3D Magnetotelluric Data I: General Formulation. *Geophysical Journal International*, 223(3), 1837–1863. <https://doi.org/10.1093/gji/ggaa415>
- Matas, J., & Bukowski, M. S. (2007). On the anelastic contribution to the temperature dependence of lower mantle seismic velocities. *Earth and Planetary Science Letters*, 259(1–2), 51–65. <https://doi.org/10.1016/j.epsl.2007.04.028>
- Meju, M. A., Gallardo, L. A., & Mohamed, A. K. (2003). Evidence for correlation of electrical resistivity and seismic velocity in heterogeneous near-surface materials. *Geophysical Research Letters*, 30(7). <https://doi.org/10.1029/2002gl016048>
- Menke, W. (2018). *Geophysical data analysis: Discrete inverse theory*. Academic press.
- Meqbel, N. M., Egbert, G. D., Wannamaker, P. E., Kelbert, A., & Schultz, A. (2014). Deep electrical resistivity structure of the northwestern US derived from 3-D inversion of USArray magnetotelluric data. *Earth and Planetary Science Letters*, 402, 290–304. <https://doi.org/10.1016/j.epsl.2013.12.026>
- Meyn, S. P., & Tweedie, R. L. (2012). *Markov chains and stochastic stability*. Springer Science & Business Media.
- Mira, A. (2001). On Metropolis-Hastings algorithms with delayed rejection. *Metron*, 59(3–4), 231–241.
- Moorkamp, M. (2017). Integrating electromagnetic data with other geophysical observations for enhanced imaging of the earth: A tutorial and review. *Surveys in Geophysics*, 38(5), 935–962. <https://doi.org/10.1007/s10712-017-9413-7>
- Moorkamp, M., Fulla, J., Aster, R., & Weise, B. (2020). Inverse methods, resolution and implications for the interpretation of lithospheric structure in geophysical inversions. *Earth and Space Science Open Archive ESSOAr*.
- Moorkamp, M., Jones, A., & Eaton, D. (2007). Joint inversion of teleseismic receiver functions and magnetotelluric data using a genetic algorithm: Are seismic velocities and electrical conductivities compatible? *Geophysical Research Letters*, 34(16). <https://doi.org/10.1029/2007gl030519>
- Moorkamp, M., Jones, A., & Fishwick, S. (2010). Joint inversion of receiver functions, surface wave dispersion, and magnetotelluric data. *Journal of Geophysical Research*, 115(B4). <https://doi.org/10.1029/2009jb006369>
- Mosegaard, K., & Hansen, T. M. (2016). Inverse methods: Problem formulation and probabilistic solutions. *Integrated imaging of the Earth: Theory and applications. Geophysical Monograph* (pp. 9–27).
- Mosegaard, K., & Tarantola, A. (1995). Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research*, 100(B7), 12431–12447. <https://doi.org/10.1029/94jb03097>
- Mosegaard, K., & Tarantola, A. (2002). Probabilistic approach to inverse problems. *International Geophysics Series*, 81(A), 237–265. [https://doi.org/10.1016/s0074-6142\(02\)80219-4](https://doi.org/10.1016/s0074-6142(02)80219-4)
- Oldenburg, D. W. (1979). One-dimensional inversion of natural source magnetotelluric observations. *Geophysics*, 44(7), 1218–1244. <https://doi.org/10.1190/1.1441004>
- Omre, H. (1987). Bayesian kriging—Merging observations and qualified guesses in kriging. *Mathematical Geology*, 19(1), 25–39. <https://doi.org/10.1007/bf01275432>
- Ortega-Gelabert, O., Zlotnik, S., Afonso, J. C., & Diez, P. (2020). Fast Stokes flow simulations for geophysical-geodynamic inverse problems and sensitivity analyses based on reduced order modeling. *Journal of Geophysical Research: Solid Earth*, 125. <https://doi.org/10.1029/2019JB018314>
- Parker, R. L. (1971). The inverse problem of electrical conductivity in the mantle. *Geophysical Journal International*, 22(2), 121–138. <https://doi.org/10.1111/j.1365-246x.1971.tb03587.x>
- Parker, R. L. (1980). The inverse problem of electromagnetic induction: Existence and construction of solutions based on incomplete data. *Journal of Geophysical Research*, 85(B8), 4421–4428. <https://doi.org/10.1029/jb085ib08p04421>
- Peherstorfer, B., Willcox, K., & Gunzburger, M. (2018). Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 60(3), 550–591. <https://doi.org/10.1137/16m1082469>
- Pommier, A. (2014). Interpretation of magnetotelluric results using laboratory measurements. *Surveys in Geophysics*, 35(1), 41–84. <https://doi.org/10.1007/s10712-013-9226-2>
- Quarteroni, A., Manzoni, A., & Negri, F. (2015). *Reduced basis methods for partial differential equations: An introduction*. Springer.
- Raftery, A. E., & Lewis, S. (1992). How many iterations in the Gibbs Sampler? In *Bayesian statistics 4* (pp. 763–773). Oxford University Press.
- Rasmussen, C. E. (1997). *Evaluation of Gaussian processes and other methods for non-linear regression (Unpublished doctoral dissertation)*. University of Toronto.
- Rawlinson, N., Tkalcic, H., & Kennett, B. (2008). New results from WOMBAT: An ongoing program of passive seismic array deployment in Australia. In *AgU fall meeting abstracts*.



- Ray, A., & Myer, D. (2019). Bayesian geophysical inversion with trans-dimensional Gaussian process machine learning. *Geophysical Journal International*, 217(3), 1706–1726. <https://doi.org/10.1093/gji/ggz111>
- Ritzwoller, M. H., Shapiro, N. M., Barmin, M. P., & Levshin, A. L. (2002). Global surface wave diffraction tomography. *Journal of Geophysical Research*, 107(B12). <https://doi.org/10.1029/2002jb001777>
- Robertson, K., Thiel, S., & Meqbel, N. (2020). Quality over quantity: On workflow and model space exploration of 3D inversion of MT data. *Earth Planets and Space*, 72(1), 1–22. <https://doi.org/10.1186/s40623-019-1125-4>
- Rosas-Carbajal, M., Linde, N., Kalscheuer, T., & Vrugt, J. A. (2013). Two-dimensional probabilistic inversion of plane-wave electromagnetic data: Methodology, model constraints and joint inversion with electrical resistivity data. *Geophysical Journal International*, 196(3), 1508–1524. <https://doi.org/10.1093/gji/ggt482>
- Selway, K. (2014). On the causes of electrical conductivity anomalies in tectonically stable lithosphere. *Surveys in Geophysics*, 35(1), 219–257. <https://doi.org/10.1007/s10712-013-9235-1>
- Selway, K., & O'Donnell, J. (2019). A small, unextractable melt fraction as the cause for the low velocity zone. *Earth and Planetary Science Letters*, 517, 117–124. <https://doi.org/10.1016/j.epsl.2019.04.012>
- Selway, K., O'Donnell, J., & Özaydin, S. (2019). Upper mantle melt distribution from petrologically constrained magnetotellurics. *Geochemistry, Geophysics, Geosystems*, 20(7), 3328–3346. <https://doi.org/10.1029/2019gc008227>
- Sheen, D. (1997). Approximation of electromagnetic fields: Part I. Continuous problems. *SIAM Journal on Applied Mathematics*, 57(6), 1716–1736. <https://doi.org/10.1137/s0036139995289234>
- Shen, W., Ritzwoller, M. H., Schulte-Pelkum, V., & Lin, F.-C. (2013). Joint inversion of surface wave dispersion and receiver functions: A Bayesian Monte-Carlo approach. *Geophysical Journal International*, 192(2), 807–836. <https://doi.org/10.1093/gji/ggs050>
- Smith, R. C. (2013). *Uncertainty quantification: Theory, implementation, and applications*. Siam.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. siam.
- Tarantola, A., & Valette, B. (1982). Inverse problems= quest for information. *Journal of Geophysics*, 50(1), 159–170.
- Tierney, L. (1994). Markov chains for exploring posterior distributions. *Annals of Statistics*, 22, 1701–1728. <https://doi.org/10.1214/aos/1176325750>
- Vozar, J., Jones, A. G., Fullea, J., Agius, M. R., Lebedev, S., Le Pape, F., & Wei, W. (2014). Integrated geophysical-petrological modeling of lithosphere-asthenosphere boundary in central Tibet using electromagnetic and seismic data. *Geochemistry, Geophysics, Geosystems*, 15(10), 3965–3988. <https://doi.org/10.1002/2014gc005365>
- Wait, J. R. (1962). Theory of magnetotelluric fields. *Journal of Research of the National Bureau of Standards, Section D*, 66(5), 509–541. <https://doi.org/10.6028/jres.066d.052>
- Wang, K., Lu, L., Maupin, V., Ding, Z., Zheng, C., & Zhong, S. (2020). Surface wave tomography of northeastern Tibetan Plateau using beamforming of seismic noise at a dense array. *Journal of Geophysical Research: Solid Earth*, 125(4), e2019JB018416. <https://doi.org/10.1029/2019jb018416>
- Williams, C. K., & Rasmussen, C. E. (1996). Gaussian processes for regression. In *Advances in neural information processing systems* (pp. 514–520).
- Xu, Y., Shankland, T. J., & Poe, B. T. (2000). Laboratory-based electrical conductivity in the Earth's mantle. *Journal of Geophysical Research*, 105(B12), 27865–27875. <https://doi.org/10.1029/2000jb900299>
- Yan, L., & Zhou, T. (2019). Adaptive multi-fidelity polynomial chaos approach to Bayesian inference in inverse problems. *Journal of Computational Physics*, 381, 110–128. <https://doi.org/10.1016/j.jcp.2018.12.025>
- Yang, Y., & Forsyth, D. W. (2006). Rayleigh wave phase velocities, small-scale convection, and azimuthal anisotropy beneath southern California. *Journal of Geophysical Research*, 111(B7). <https://doi.org/10.1029/2005jb004180>
- Yang, Y., Ritzwoller, M. H., Lin, F.-C., Moschetti, M., & Shapiro, N. M. (2008). Structure of the crust and uppermost mantle beneath the western United States revealed by ambient noise and earthquake tomography. *Journal of Geophysical Research*, 113(B12). <https://doi.org/10.1029/2008jb005833>
- Yasar, E., & Erdogan, Y. (2004). Correlating sound velocity with the density, compressive strength and young's modulus of carbonate rocks. *International Journal of Rock Mechanics and Mining Sciences*, 41(5), 871–875. <https://doi.org/10.1016/j.ijrmms.2004.01.012>
- Yoshino, T. (2010). Laboratory electrical conductivity measurement of mantle minerals. *Surveys in Geophysics*, 31(2), 163–206. <https://doi.org/10.1007/s10712-009-9084-0>
- Yoshino, T., Matsuzaki, T., Shatskiy, A., & Katsura, T. (2009). The effect of water on the electrical conductivity of olivine aggregates and its implications for the electrical structure of the upper mantle. *Earth and Planetary Science Letters*, 288(1–2), 291–300. <https://doi.org/10.1016/j.epsl.2009.09.032>
- Young, M., Cayley, R., McLean, M., Rawlinson, N., Arroucou, P., & Salmon, M. (2013). Crustal structure of the east gondwana margin in south-east Australia revealed by transdimensional ambient seismic noise tomography. *Geophysical Research Letters*, 40(16), 4266–4271. <https://doi.org/10.1002/grl.50878>
- Zhang, A., Afonso, J. C., Xu, Y., Wu, S., Yang, Y., & Yang, B. (2019). The deep lithospheric structure of the Junggar Terrane, NW China: Implications for its origin and tectonic evolution. *Journal of Geophysical Research: Solid Earth*, 124(11), 11615–11638. <https://doi.org/10.1029/2019jb018302>
- Zhang, J., & Taflanidis, A. A. (2019). Accelerating MCMC via Kriging-based adaptive independent proposals and delayed rejection. *Computer Methods in Applied Mechanics and Engineering*, 355, 1124–1147. <https://doi.org/10.1016/j.cma.2019.07.016>
- Zyserman, F. I., & Santos, J. E. (2000). Parallel finite element algorithm with domain decomposition for three-dimensional magnetotelluric modelling. *Journal of Applied Geophysics*, 44(4), 337–351. [https://doi.org/10.1016/s0926-9851\(00\)00012-4](https://doi.org/10.1016/s0926-9851(00)00012-4)