

FULL PAPER

Open Access



Simultaneous inversion for source field and mantle electrical conductivity using the variable projection approach

Jingtao Min¹ and Alexander Grayver^{2*}

Abstract

Time-varying electromagnetic field observed on the ground or at a spacecraft consists of contributions from (i) electric source currents, such as those in the ionosphere and magnetosphere, and (ii) corresponding fields induced by source currents within the conductive Earth's interior by virtue of electromagnetic induction. Knowledge about the spatio-temporal structure of inducing currents is a key component in ionospheric and magnetospheric studies, and is also needed in space weather hazard evaluation, whereas the induced currents depend on the Earth's subsurface electrical conductivity distribution and allow us to probe this physical property. In this study, we present an approach that reconstructs the inducing source and subsurface conductivity structures simultaneously, preserving consistency between the two models by exploiting the inherent physical link. To achieve this, we formulate the underlying inverse problem as a separable nonlinear least-squares (SNLS) problem, where inducing current and subsurface conductivity parameters enter as linear and nonlinear model unknowns, respectively. We solve the SNLS problem using the variable projection method and compare it with other conventional approaches. We study the properties of the method and demonstrate its feasibility by simultaneously reconstructing the ionospheric and magnetospheric currents along with a 1-D average mantle conductivity distribution from the ground magnetic observatory data.

Keywords Geomagnetism, Electromagnetic induction, Variable projection

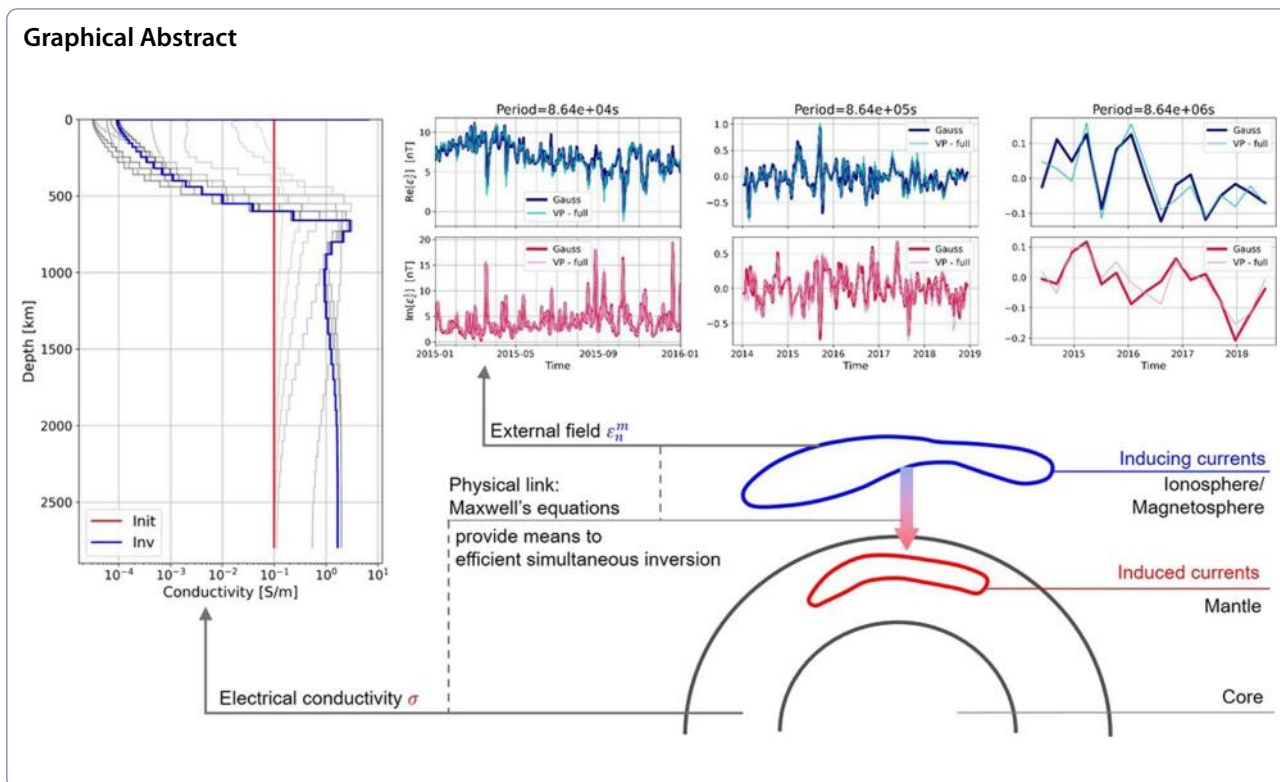
*Correspondence:

Alexander Grayver
agrayver@uni-koeln.de

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



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.



Introduction

Time variations of magnetic field that we observe on the ground or at a spacecraft represent a superposition of the inducing (primary) and induced components. There is substantial interest in knowing both inducing and induced components of the field as accurately as possible. On the one hand, knowledge about space-time variability of the inducing field constrains the state of source electric currents in the ionosphere and magnetosphere (Yamazaki and Maute 2017; Balasis and Egbert 2006; Tsyganenko 2019), which in turn represents a crucial input for accurate geomagnetic field modelling (Maus and Weidelt 2004; Finlay et al. 2017) and space weather hazard evaluation (Pulkkinen et al. 2003; Kelbert 2020; Juusola et al. 2020). On the other hand, relation between the inducing and the induced field variations, governed by Maxwell's equations, can be used to probe the electrical conductivity distribution in the Earth's subsurface (Olsen 1999; Kuvshinov 2012; Kelbert et al. 2009). However, separation of the magnetic field into inducing and induced components is often non-trivial owing to their nonlinear relationship that depends on the 3-D distribution of electrical conductivity in the Earth's interior. Our goal here is to elaborate on this problem further.

To keep the study concise and focused, we make several assumptions that are implied in the derivations and discussions below. First, we concentrate on time-variations

with periods longer than a few hours, which is beyond the band where a simple plane-wave source assumption is valid (this assumption can be used to model external source fields (Kelbert and Lucas 2020) and can be used in the magnetotelluric method (Chave and Jones 2012) for probing the electrical conductivity of subsurface). Second, we assume that field variations are due to the extraneous electric currents and the corresponding electromagnetic response from the conductive Earth's interior. In other words, the contributions from all other magnetic field sources, such as the crust or the core, were subtracted from the data (it is clear that some residual fields from these sources are always present, but these problems are beyond the scope of our study). Further, the extraneous electric currents are assumed to have their origin in the ionosphere and magnetosphere. By this, we exclude the ocean-induced electromagnetic fields, which require dedicated modelling and inversion approaches (e.g., Velínský et al. 2018).

In the most general form, the extraneous source structure needs to be parameterized with spatially heterogeneous functions and estimated from the data along with the subsurface electrical conductivity distribution by solving a corresponding inverse problem. However, joint estimation of conductivity and external field structures represents a notoriously difficult task. Conventionally, the Gauss method has been used to separate the magnetic

field into time series of Spherical Harmonic (SH) coefficients of internal and external origins (Backus et al. 1996). By relating internal and external SH coefficients, one can estimate a transfer function between them and perform the inversion in terms of subsurface electrical conductivity (Olsen 1999; Schmucker 1999; Kuvshinov 2012) or fit the time series of SH coefficients directly (Velínský and Knopp 2021). However, this approach is only applicable to potential fields where inducing field contribution is external to the observer. Moreover, due to sparse measurements, one is typically limited to using a small set of Spherical Harmonic functions to describe the inducing and induced parts of the field (Kuvshinov et al. 2021; Velínský and Knopp 2021).

Recognizing these limitations, a number of recent studies (Koch and Kuvshinov 2013; Sun et al. 2015; Guzavina et al. 2019; Egbert et al. 2021; Zhang et al. 2022) have adopted an alternative strategy where the source structure is estimated given some prior knowledge about the subsurface conductivity. With this estimated source structure, the inversion in terms of subsurface conductivity is subsequently performed and the updated conductivity model can again be used to re-estimate the source coefficients. This approach (i) allows for a more general ansatz to describe the source geometry (Zenhäusern et al. 2021; Egbert et al. 2021), (ii) enables derivation of alternative families of transfer functions (Püthe and Kuvshinov 2014; Guzavina et al. 2019), which are not limited to the potential field assumption, and (iii) facilitates incorporation of the prior knowledge on the induction effects due to the ocean and marine sediments (Grayver et al. 2021). Listed points make it possible to mitigate or completely overcome the limitations imposed by the conventional Gauss method. In these aforementioned studies, determination of the inducing source field and the mantle conductivity is performed in an alternating manner on the two separate model spaces (hereinafter, we term this procedure an "alternating approach"). Such separate estimation of the two model spaces is assumed to result in progressively refined knowledge of both the source and conductivity models.

In this study, we develop this idea further and pose a problem in a general form that allows us to simultaneously estimate the source and subsurface conductivity directly from the data. Since the model space consists of one part (i.e., inducing source currents), upon which the dependence of the observable is linear, and another part (i.e., subsurface electrical conductivity), which enters the objective in a nonlinear manner, the underlying inverse problem (under squared loss) belongs to a class of special optimization problems known as the Separable Nonlinear Least-Squares (SNLS) problem. We will show that the naive "alternating approach" described above is the

simplest way of solving the SNLS problem, although it may lack consistency and suffer from slow convergence. We will explore more efficient ways of solving the SNLS problem. In particular, the variable projection method (hereafter referred to as VP) has been proposed as an optimal method for solving SNLS problems that benefits from both computational efficiency and fast convergence (Golub and Pereyra 1973, 2003). In essence, VP exploits the linear dependency in one part of the model and estimates this part via linear least squares at each iteration, thus optimally (in least-squares sense) projecting the complete model space onto a reduced subspace for efficient nonlinear optimization.

The advantage of variable projection naturally appeals to a number of geophysical inverse problems where the unknown parameters intrinsically constitute separable least squares. Such behavior is typical of seismic wave propagation and electromagnetic induction, where source characterization is linearly filtered by a medium response that depends nonlinearly on medium properties. In the last decade, this algorithm has been recognized in seismology as an efficient way to invert for velocity structure while simultaneously characterizing the source (Rickett 2013; De Ridder and Madsen 2018), the source-related calibration parameters (Li et al. 2013), or both the source and the receiver factors (Hu et al. 2021). Despite an early conceptualization (Fainberg et al. 1990), this method, to our knowledge, has not yet been elaborated in the context of electromagnetic induction problems, where the merit of VP is potentially much more pronounced: the full model inversion including the source and conductivity, which is prohibitive due to aforementioned high dimension and nonlinearity, becomes tractable thanks to linear variable projection. Here we present the application of the VP method to a problem of electromagnetic induction sounding.

We demonstrate that not only does VP enable simultaneous estimation of the inducing field structure and the electrical conductivity using a natural physical link between them, but it also provides insights into the interplay between determination of inducing field and conductivity models.

Methods

Electromagnetic (EM) field variations are governed by Maxwell's equations, which in the frequency domain read

$$\begin{aligned}\nabla \times \mathbf{E} &= -i\omega\mathbf{B}, \\ \frac{1}{\mu_0}\nabla \times \mathbf{B} &= \sigma\mathbf{E} + \mathbf{j},\end{aligned}\tag{1}$$

where ω is the angular frequency, \mathbf{r} is the position vector, $\sigma(\mathbf{r}) \in \mathbb{R}$ denotes electrical conductivity of a medium,

and $\mathbf{B}(\mathbf{r}, \omega)$ and $\mathbf{E}(\mathbf{r}, \omega)$ are the magnetic and electric fields, respectively. $\mathbf{j}(\mathbf{r}, \omega)$ is the extraneous (impressed) current density. The extraneous currents are assumed to originate within the ionosphere and magnetosphere, separated from the solid Earth by a layer of insulating air. We neglected displacement currents and took $\mu = \mu_0$ for the magnetic permeability. Here, we adopted the following convention for the Fourier transform:

$$\begin{aligned} X(\omega) &= \mathcal{F}[x(t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x(t)e^{-i\omega t} dt, \\ x(t) &= \mathcal{F}^{-1}[X(\omega)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} X(\omega)e^{i\omega t} d\omega. \end{aligned} \tag{2}$$

The solution of Eq. (1) can be found when both the inducing source \mathbf{j} and the conductivity σ are given. In this case, the magnetic field due to an arbitrary distribution of current density can be formally expressed as

$$\mathbf{B}(\mathbf{r}, \omega; \sigma, \mathbf{j}) = \int_{\Omega} \mathbf{G}(\mathbf{r}, \mathbf{r}', \omega; \sigma) \cdot \mathbf{j}(\mathbf{r}', \omega) d\mathbf{r}', \tag{3}$$

where \mathbf{G} is the Green's tensor of the medium (Kuvshinov 2008) and Ω is the volume occupied by extraneous currents. A corresponding time-domain counterpart contains a temporal convolution, and has the form

$$\mathbf{B}(\mathbf{r}, t; \sigma, \mathbf{j}) = \int_{-\infty}^t \int_{\Omega} \mathbf{G}(\mathbf{r}, \mathbf{r}', t - t'; \sigma) \cdot \mathbf{j}(\mathbf{r}', t') d\mathbf{r}' dt'. \tag{4}$$

Independent of the domain where we operate, the modeling process can always be expressed in the operator form

$$\mathbf{B} = \mathbf{L}(\sigma) \mathbf{j}, \tag{5}$$

where $\mathbf{L}(\sigma) : (V_j(\Omega_j))^3 \mapsto (V_B(\Omega_B))^3$ is a linear operator, mapping an electric current field to a vector magnetic field, and $\mathbf{L}(\cdot) : V_\sigma(\Omega_\sigma) \mapsto \mathcal{L}((V_j(\Omega_j))^3, (V_B(\Omega_B))^3)$ is a nonlinear function that maps the conductivity distribution into a linear operator. Here, Ω_j , Ω_σ and Ω_B are the domains of inducing currents, induced currents, and observed magnetic fields, respectively. V_j , V_σ , and V_B are function spaces on the corresponding domains, and $\mathcal{L}(U, V)$ denotes the linear space of the linear maps from U to V .

Equation 5 shows that the magnetic field is related to the source by a linear operator, which is a nonlinear functional of the electrical conductivity. The equivalent for the electric field is straightforward, but we omit it because we only consider magnetic field observations in this study. We can thus express the forward modeling in a concise algebraic form as

$$\mathbf{d}^{\text{mod}}(\mathbf{m}, \mathbf{c}) = \mathbf{F}(\mathbf{m}) \mathbf{c}, \tag{6}$$

where \mathbf{d}^{mod} is the modelled data vector, \mathbf{m} is a parameterization of the conductivity model, \mathbf{c} is the inducing source vector, and $\mathbf{F}(\mathbf{m})$ is a functional of \mathbf{m} that links the field to the extraneous currents. The specific form of $\mathbf{F}(\mathbf{m})$ depends on the adopted discretization and parameterization of σ and \mathbf{j} , but the stated general algebraic form accommodates the full set of modelling approaches. Our goal is to estimate the unknown variables consisting of electrical conductivity model \mathbf{m} and extraneous currents \mathbf{c} from observations of the magnetic field taken at specified locations and times. To achieve this goal, we seek a combination of \mathbf{m} and \mathbf{c} that minimizes the data misfit

$$\chi^2 = d(\mathbf{d}^{\text{obs}}, \mathbf{d}^{\text{mod}}(\mathbf{m}, \mathbf{c})) = d(\mathbf{d}^{\text{obs}}, \mathbf{F}(\mathbf{m}) \mathbf{c}), \tag{7}$$

where \mathbf{d}^{obs} is the observational data vector, given by magnetic field observations in our case, and $d(\cdot, \cdot)$ denotes the distance metric induced by the corresponding Banach space. A popular choice for such metric in EM induction soundings is the distance induced by the vector norm weighted by the data covariance

$$\begin{aligned} \chi^2 &= \frac{1}{2} (\mathbf{d}^{\text{obs}} - \mathbf{d}^{\text{mod}})^H \mathbf{C}_d^{-1} (\mathbf{d}^{\text{obs}} - \mathbf{d}^{\text{mod}}) \\ &= \frac{1}{2} (\mathbf{d}^{\text{obs}} - \mathbf{F}(\mathbf{m}) \mathbf{c})^H \mathbf{C}_d^{-1} (\mathbf{d}^{\text{obs}} - \mathbf{F}(\mathbf{m}) \mathbf{c}), \end{aligned} \tag{8}$$

where \mathbf{C}_d is the data covariance matrix. Note we use the superscript H to denote the Hermitian transpose of the matrix or vector, as the data vector may be complex. In the absence of co-variances, the data samples are assumed to be mutually independent, in which case $\mathbf{C}_d = \text{diag}(s_i^2)$, where s_i^2 is the variance of the i -th datum. Introducing $\mathbf{W} = \mathbf{C}_d^{-1/2} = \text{diag}(s_i^{-1})$, the data misfit can be rewritten as the squared ℓ_2 norm of the weighted residual

$$\chi^2 = \frac{1}{2} \|\mathbf{r}_w\|_2^2 = \frac{1}{2} \|\mathbf{W}(\mathbf{d}^{\text{obs}} - \mathbf{F}(\mathbf{m}) \mathbf{c})\|_2^2 = \frac{1}{2} \|\mathbf{d}_w^{\text{obs}} - \mathbf{F}_w(\mathbf{m}) \mathbf{c}\|_2^2. \tag{9}$$

Here, $\|\cdot\|_2$ denotes the ℓ_2 norm, $\mathbf{r} = \mathbf{d}^{\text{obs}} - \mathbf{F}(\mathbf{m}) \mathbf{c}$ is the residual vector, and $\mathbf{r}_w = \mathbf{W} \mathbf{r}$, $\mathbf{d}_w = \mathbf{W} \mathbf{d}$, and $\mathbf{F}_w = \mathbf{W} \mathbf{F}$ are the weighted forms of the residual vector, the data vector, and the linear operator, respectively. To mitigate the inherent non-uniqueness of the problem, we add a regularization term $\lambda R(\mathbf{m})$, where $R(\cdot)$ is the penalty function, and λ is the regularization strength. Here, we consider the penalty function that penalizes the ℓ_2 norm of model structure, given by $R(\mathbf{m}) = \frac{1}{2} \|\mathbf{\Gamma} \mathbf{m}\|_2^2$, where $\mathbf{\Gamma}$ is known as the Tikhonov matrix. In the experiments that follow, we shall use a first-order difference operator as $\mathbf{\Gamma}$ to enforce the smoothness of the model. Similarly,

we can also regularize the linear parameters directly, i.e., $\frac{\lambda_c}{2} \|\Gamma_c \mathbf{c}\|_2^2$. The full optimization problem is then given by

$$\min_{\mathbf{m}, \mathbf{c}} \frac{1}{2} \|\mathbf{d}_w^{\text{obs}} - \mathbf{F}_w(\mathbf{m}) \mathbf{c}\|_2^2 + \frac{\lambda}{2} \|\Gamma \mathbf{m}\|_2^2 + \frac{\lambda_c}{2} \|\Gamma_c \mathbf{c}\|_2^2. \tag{10}$$

Regularization on the source parameters \mathbf{c} allows the use of prior knowledge on the source geometry (Sun et al. 2015; Laundal et al. 2021), e.g., while using a high-dimensional parameter space for the source. This would become beneficial or even necessary when an accurate complex source estimation is required. In this study, however, we shall proceed by defining $\lambda_c = 0$, as our experiments all assume a low-degree source structure. Generalization to include the regularization on \mathbf{c} is mathematically straightforward but further complicates the formulas, and hence is only documented for completeness in Appendix A.

Seeking a solution to the stated problem directly in the joint model space of \mathbf{m} and \mathbf{c} induces a fully nonlinear least-squares problem in a high-dimensional space. We note again that the magnetic field is a linear functional of the extraneous currents represented by \mathbf{c} , but a nonlinear functional of the subsurface electrical conductivity expressed through \mathbf{m} . This particular structure of the inverse problem with the data misfit defined in Eq. 9 makes it an example of the so-called separable nonlinear least squares (SNLS). While the problem in Eq. 10 can be linearized and solved in the full model space, this "naive" approach is inefficient and prohibitive for problems of interest. Fortunately, the particular structure of an SNLS problem allows us to adopt more efficient solution strategies.

Variable projection approach

Variable projection (VP) has been first proposed by Golub and Pereyra (1973) as an optimization method for solving SNLS problems. Exploiting the linear dependency on \mathbf{c} , at each given conductivity model \mathbf{m} , the best-fitting linear part can be obtained via a linear regression $\mathbf{c} = \hat{\mathbf{c}}(\mathbf{m}) = \mathbf{F}_w^+(\mathbf{m}) \mathbf{d}_w^{\text{obs}}$, where \mathbf{F}_w^+ denotes the Moore–Penrose pseudoinverse of \mathbf{F}_w . We use $\hat{\mathbf{c}}$ to explicitly denote the dependency of the least-squares solution of \mathbf{c} on \mathbf{m} . With the linear regression at each iteration, the optimization is then optimally (in statistical sense) constrained to the nonlinear part of the model space

$$\begin{aligned} \min_{\mathbf{m}} \frac{1}{2} \|\mathbf{d}_w - \mathbf{F}_w(\mathbf{m}) \hat{\mathbf{c}}(\mathbf{m})\|_2^2 + \frac{\lambda}{2} \|\Gamma \mathbf{m}\|_2^2 \\ = \min_{\mathbf{m}} \frac{1}{2} \|\mathbf{d}_w - \mathbf{F}_w(\mathbf{m}) \mathbf{F}_w^+(\mathbf{m}) \mathbf{d}_w\|_2^2 + \frac{\lambda}{2} \|\Gamma \mathbf{m}\|_2^2 \\ = \min_{\mathbf{m}} \frac{1}{2} \|\mathbf{P}_{\mathbf{F}_w}^\perp(\mathbf{m}) \mathbf{d}_w\|_2^2 + \frac{\lambda}{2} \|\Gamma \mathbf{m}\|_2^2, \end{aligned} \tag{11}$$

where $\mathbf{P}_{\mathbf{F}_w}^\perp = \mathbf{I} - \mathbf{F}_w \mathbf{F}_w^+$ is a projector onto the orthogonal complement of the range of $\mathbf{F}_w(\mathbf{m})$. Note we used $\mathbf{d} = \mathbf{d}^{\text{obs}}$ for brevity.

A minimum to the nonlinear least-squares problem in Eq. 11 can be found using either a gradient-based or a Newton-based optimization method. In both cases, the update on the nonlinear model involves evaluation of the Fréchet derivatives with respect to the nonlinear parameters. In turn, this requires us to incorporate the implicit dependency of \mathbf{c} on \mathbf{m} . In what follows, we will use \mathbf{DA} to denote the derivative of \mathbf{A} with respect to \mathbf{m} , where \mathbf{A} is a functional of \mathbf{m} . In its discrete form where $\mathbf{A} \in \mathbb{C}^{i_1 \times i_2 \times \dots \times i_l}$, the result \mathbf{DA} is a tensor of order $l + 1$, and the last dimension denotes the differentiation component. More explicitly

$$(\mathbf{DA}(\mathbf{m}))_{i_1 i_2 \dots i_{l+1}} = \frac{\partial A_{i_1 i_2 \dots i_l}}{\partial m_{i_{l+1}}}. \tag{12}$$

For $l \geq 2$, matrix multiplications involving \mathbf{DA} are always assumed to be performed on the leading 2 dimensions. Golub and Pereyra (1973) derives the expression for the gradient of the objective function and the Jacobian of the residual vector in terms of pseudoinverses and derivatives of the linear operator \mathbf{F}_w . We adopt the notations used in Hong et al. (2017), and introduce the following two partial Jacobians, as derivatives taken explicitly on the original data misfit without variable projection (Eq. 9)

$$\begin{aligned} \mathbf{J}_c &= \frac{\partial \mathbf{r}_w}{\partial \mathbf{c}} = -\mathbf{F}_w(\mathbf{m}), \\ \mathbf{J}_m &= \frac{\partial \mathbf{r}_w}{\partial \mathbf{m}} = -\mathbf{D}\mathbf{F}_w(\mathbf{m}) \mathbf{c} = \mathbf{D}\mathbf{J}_c \mathbf{c}. \end{aligned} \tag{13}$$

Now, the linear projection can also be stated as $\hat{\mathbf{c}} = -\mathbf{J}_c^+ \mathbf{d}_w$, and the orthogonal projector is given by $\mathbf{P}_{\mathbf{F}_w}^\perp = \mathbf{I} - \mathbf{J}_c \mathbf{J}_c^+ = \mathbf{P}_{\mathbf{J}_c}^\perp$. We note that the two explicit Jacobians are coupled in the model space (i.e., \mathbf{J}_m and \mathbf{J}_c are dependent upon \mathbf{c} and \mathbf{m} , respectively). This will be clearly seen in the case of VP, where the complete Jacobian of the variable-projected misfit term (Eq. 11) is given by

$$\mathbf{J}(\mathbf{m}, \hat{\mathbf{c}}(\mathbf{m})) = \mathbf{D}\mathbf{r}_w = \mathbf{J}_m + \mathbf{J}_c \mathbf{D}\hat{\mathbf{c}} = \mathbf{J}_m - \mathbf{J}_c \mathbf{D}\mathbf{J}_c^+ \mathbf{d}_w. \tag{14}$$

Invoking the derivative of pseudoinverse (see Golub and Pereyra 1973 for derivation details)

$$\begin{aligned} \mathbf{DA}^\dagger &= -\mathbf{A}^\dagger \mathbf{D}\mathbf{A} \mathbf{A}^\dagger + \mathbf{A}^\dagger (\mathbf{A}^\dagger)^H (\mathbf{D}\mathbf{A})^H \mathbf{P}_{\mathbf{A}}^\perp \\ &\quad + \left(\mathbf{P}_{\mathbf{A}^H}^\perp\right)^H (\mathbf{D}\mathbf{A})^H (\mathbf{A}^\dagger)^H \mathbf{A}^\dagger, \end{aligned} \tag{15}$$

the complete Jacobian of the variable-projected system can hence be reiterated and expressed solely in terms of $\mathbf{J}_m, \mathbf{J}_c$ together with its derivative and pseudoinverse

$$\left(\text{Re}[\mathbf{J}^H \mathbf{J}] + \lambda \mathbf{\Gamma}^T \mathbf{\Gamma} \right) \Delta \mathbf{m} = -D\chi^2 - \lambda \mathbf{\Gamma}^T \mathbf{\Gamma} \mathbf{m}. \quad (20)$$

$$\begin{aligned} \mathbf{J}(\mathbf{m}, \hat{\mathbf{c}}(\mathbf{m})) &= \mathbf{J}_m - \mathbf{J}_c \left(-\mathbf{J}_c^\dagger \text{DJ}_c \mathbf{J}_c^\dagger + \mathbf{J}_c^\dagger (\mathbf{J}_c^\dagger)^H (\text{DJ}_c)^H \mathbf{P}_{\mathbf{J}_c}^\perp + \left(\mathbf{P}_{\mathbf{J}_c^\dagger}^\perp \right)^H (\text{DJ}_c)^H (\mathbf{J}_c^\dagger)^H \mathbf{J}_c^\dagger \right) \mathbf{d}_w \\ &= \mathbf{J}_m - \mathbf{J}_c \mathbf{J}_c^\dagger \text{DJ}_c \hat{\mathbf{c}} - \mathbf{J}_c \mathbf{J}_c^\dagger (\mathbf{J}_c^\dagger)^H (\text{DJ}_c)^H \mathbf{P}_{\mathbf{J}_c}^\perp \mathbf{d}_w + \mathbf{J}_c \left(\mathbf{P}_{\mathbf{J}_c^\dagger}^\perp \right)^H (\text{DJ}_c)^H (\mathbf{J}_c^\dagger)^H \hat{\mathbf{c}} \\ &= \mathbf{J}_m - \mathbf{J}_c \mathbf{J}_c^\dagger \mathbf{J}_m - (\mathbf{J}_c^\dagger)^H (\text{DJ}_c)^H \mathbf{P}_{\mathbf{J}_c}^\perp \mathbf{d}_w. \end{aligned} \quad (16)$$

The last step uses the fact that $\mathbf{A}\mathbf{A}^\dagger(\mathbf{A}^\dagger)^H = (\mathbf{A}\mathbf{A}^\dagger)^H(\mathbf{A}^\dagger)^H = (\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger)^H = (\mathbf{A}^\dagger)^H$ and $\mathbf{A}(\mathbf{P}_{\mathbf{A}^\dagger}^\perp)^H \equiv \mathbf{0}$. Part of the dependency of \mathbf{c} upon \mathbf{m} , namely the 3rd term in Eq. 15, has no contribution to the complete Jacobian, since it is perpendicular to \mathbf{J}_c . The complete Jacobian reads

$$\begin{aligned} \mathbf{J}(\mathbf{m}, \hat{\mathbf{c}}(\mathbf{m})) &= \mathbf{J}_m - \mathbf{J}_c \mathbf{J}_c^\dagger \mathbf{J}_m - (\mathbf{J}_c^\dagger)^H (\text{DJ}_c)^H \mathbf{P}_{\mathbf{J}_c}^\perp \mathbf{d}_w \\ &= \mathbf{J}_m - \mathbf{J}_c \mathbf{J}_c^\dagger \mathbf{J}_m - (\mathbf{J}_c^\dagger)^H (\text{DJ}_c)^H \mathbf{r}_w. \end{aligned} \quad (17)$$

Note that if an inverse problem were posed solely in the space of conductivity model, then only the first term, namely \mathbf{J}_m , would be present. The trailing two terms involve the dependency of the source estimate on the change in the subsurface conductivity, confining the model updates of \mathbf{m} to the hyperplane defined by the regression of \mathbf{c} . Reintroducing linear operators via Eq. 13, we arrive at the expression for the Jacobian of the residual vector

$$\mathbf{J} = -\mathbf{P}_{\mathbf{F}_w}^\perp \text{DF}_w \mathbf{F}_w^\dagger \mathbf{d}_w - \left(\mathbf{P}_{\mathbf{F}_w}^\perp \text{DF}_w \mathbf{F}_w^\dagger \right)^H \mathbf{d}_w. \quad (18)$$

Accordingly, the gradient of the misfit function reads

$$\begin{aligned} \text{grad} \chi^2 &= D\chi^2 = \text{Re}[\mathbf{J}^H \mathbf{r}_w] = -\text{Re} \left[\left(\mathbf{P}_{\mathbf{F}_w}^\perp \text{DF}_w \mathbf{F}_w^\dagger \mathbf{d}_w \right)^H \mathbf{r}_w + \mathbf{d}_w^H \mathbf{P}_{\mathbf{F}_w}^\perp \text{DF}_w \mathbf{F}_w^\dagger \mathbf{r}_w \right] \\ &= -\text{Re} \left[(\text{DF}_w \mathbf{F}_w^\dagger \mathbf{d}_w)^H \mathbf{P}_{\mathbf{F}_w}^\perp (\mathbf{P}_{\mathbf{F}_w}^\perp \mathbf{d}_w) + \mathbf{d}_w^H \mathbf{P}_{\mathbf{F}_w}^\perp \text{DF}_w \mathbf{F}_w^\dagger (\mathbf{P}_{\mathbf{F}_w}^\perp \mathbf{d}_w) \right] \\ &= -\text{Re} \left[(\text{DF}_w \hat{\mathbf{c}})^H \mathbf{P}_{\mathbf{F}_w}^\perp \mathbf{d}_w + \mathbf{0} \right] = \text{Re}[\mathbf{J}_m^H \mathbf{r}_w]. \end{aligned} \quad (19)$$

The second line uses the fact that $(\mathbf{P}_{\mathbf{F}_w}^\perp)^2 = \mathbf{P}_{\mathbf{F}_w}^\perp$, and $\mathbf{F}_w^\dagger \mathbf{P}_{\mathbf{F}_w}^\perp = \mathbf{F}_w^\dagger (\mathbf{I} - \mathbf{F}_w \mathbf{F}_w^\dagger) = \mathbf{0}$, and the last equality uses $\mathbf{J}_m = -\text{DF}_w \hat{\mathbf{c}}$ and $\mathbf{r}_w = \mathbf{P}_{\mathbf{F}_w}^\perp \mathbf{d}_w$. Equations 18–19 define the first order Fréchet derivatives with the variable projection. To avoid higher order derivatives, we use the Gauss–Newton algorithm to update conductivity model, where the Hessian is approximated as $\mathbf{H} \approx \text{Re}[\mathbf{J}^H \mathbf{J}]$. The model update thus takes the form

We refer to the inversion scheme that calculates Jacobian via Eq. 18 as the full-VP scheme. In the case of a 1-D radial conductivity model, the calculation of DF_w is cheap and can often be obtained semi-analytically. For a general 3-D conductivity model, the explicit evaluation and storage of DF_w is often prohibitive. In this case, the adjoint method (Pankratov and Kuvshinov 2010; Egbert and Kelbert 2012) can be used to efficiently calculate the gradient and create a low-rank representation of the DF_w (Egbert 2012) or solve Eq. (20) for the model update using Krylov subspace methods, both avoiding storage and evaluation of large matrices (e.g., Jacobian). However, even without explicit evaluation of Jacobian, the full-VP algorithm entails additional calculations due to interactions between the linear and nonlinear parts of the model space. It is therefore desirable to explore approximations that allow for fewer evaluations of DF_w .

Two such approximations have been proposed by Ruhe and Wedin (1980). One option is to drop the last term in Eq. 17, effectively dropping the 2nd term in Eq. 18, yielding

$$\mathbf{J} = -\mathbf{P}_{\mathbf{F}_w}^\perp \text{DF}_w \mathbf{F}_w^\dagger \mathbf{d}_w = \mathbf{P}_{\mathbf{J}_c}^\perp \mathbf{J}_m. \quad (21)$$

We adopt the terminology used by Hong et al. (2017) and hereinafter refer to this as the VP-RW2 scheme. The dropped term is considered a higher order refinement. This scheme retains high convergence rate and accuracy, while outperforming the full-VP in terms of computational efficiency (Ruhe and Wedin 1980; O’Leary and Rust 2013). The second option is to drop both the 2nd and the 3rd terms in Eq. 17, leading to the very simple form

$$\mathbf{J} = -\mathbf{D}\mathbf{F}_w \mathbf{F}_w^+ \mathbf{d}_w = \mathbf{J}_m, \quad (22)$$

hereinafter referred to as the VP-RW3 scheme. This is equivalent to assuming fixed inducing currents (i.e., $\mathbf{D}\mathbf{c} = 0$) at each iteration while searching for updates on the conductivity structure. This is in contrast to both full-VP and VP-RW2 schemes, where Jacobian contains additional information on the implicit feedback of the source. These three variants are closely related in the scope of VP but have different levels of approximation, and will be compared in the context of EM induction sounding. Despite poor performance of the VP-RW3 scheme previously reported by Hong et al. (2017) for matrix factorization problems, we chose to consider this scheme here, particularly because of its resemblance to what we call the "alternating approach", which we will revisit later under the framework of variable projection.

As a final remark, we observe that the gradient $\mathbf{D}\chi^2$ always has the same expression as in Eq. 18, regardless of the approximation used for constructing the Jacobian. This is due to the fact that as $\hat{\mathbf{c}} = \mathbf{F}_w^+ \mathbf{d}_w$ guarantees that the source parameters minimize the least-squares misfit of the data, the residual inevitably lives in the orthogonal complement of the linear operator, and only manifest itself through \mathbf{J}_m . In other words, as long as the current source estimation minimizes the data misfit, gradients do not "sense" the implicit feedback of the source, but always view the source as truly fixed (as if it were the ground truth model of the source), as has been noticed by Aravkin and van Leeuwen (2012). Therefore, purely gradient-based optimization schemes are not affected by the choice of the variant of VP. Optimization schemes utilizing higher order information, such as Gauss–Newton method and Levenberg–Marquardt algorithms, are however different for different VP variants.

Alternating approach

Conventionally, models of magnetospheric/ionospheric current systems and models of the mantle electrical conductivity are estimated separately, using dedicated approaches. Combining these procedures, Koch and Kuvshinov (2013) proposed a scheme where, starting

from an initial model of subsurface conductivity, one first obtains a preliminary estimate of inducing currents, then re-calculates the conductivity model with the estimated source, and then goes back to refining the source with the "updated" mantle conductivity. This procedure can in principle be repeated several times, until model estimates or data misfits reach certain convergence criteria. The same alternating method has most recently been utilized by Zhang et al. (2022) to invert for the conductivity in the mantle transition zone (MTZ), in combination with their physics-based representation of the inducing currents.

Similar to variants of the variable projection, the alternating approach also offers a way to optimize on external currents and mantle conductivity simultaneously, without resorting to fully nonlinear inversion schemes. It can be viewed as a conglomeration of successive inversions, conventionally carried out independently with respect to external currents and mantle conductivity. The major difference from VP is that in the case of a naive alternating approach, once one part of the model is estimated, inversion on the other part is carried out in a complete standalone stage to minimize the objective. This behavior is especially pronounced during inversion of the electrical conductivity, where a significant number of iterations are usually needed to capture the nonlinear dependence of the predicted data on the conductivity model. In VP, estimate on the source is projected and updated at each iteration step and is only used for one update, whereas for alternating approaches, all iterations on the conductivity model in one inversion phase are conducted under a fixed source. Such approach may potentially lead to high redundancy in iterations and result in biased model estimates.

In this study, we revisit and generalize the idea of the alternating approach, by implementing a flexible version of the inversion scheme for our problem. Our implementation is based on nonlinear model updates: at each iteration, update on the nonlinear model is generated using the Gauss–Newton method, while the source is kept fixed. At iterations pre-defined by certain criteria (referred to as linear update criteria), the inducing source is updated. The scheme can be summarized by the following pseudo-code:

$$\begin{aligned}
 \text{Iteration } k = 0 : \quad & \mathbf{c}^{(0)} = \mathbf{F}_w^\dagger(\mathbf{m}^{(0)})\mathbf{d}_w \\
 \text{Iteration } k > 0 : \quad & \text{Calculate } \mathbf{J}_m^{(k-1)} = \mathbf{J}_m(\mathbf{m}^{(k-1)}, \mathbf{c}^{(k-1)}), \\
 & \mathbf{g}^{(k-1)} = \mathbf{D}\chi^2(\mathbf{m}^{(k-1)}, \mathbf{c}^{(k-1)}) + \lambda\mathbf{\Gamma}^T\mathbf{\Gamma}\mathbf{m}^{(k-1)}. \\
 & \text{Solve } \left(\text{Re} \left[\left(\mathbf{J}_m^{(k-1)} \right)^H \mathbf{J}_m^{(k-1)} \right] + \lambda\mathbf{\Gamma}^T\mathbf{\Gamma} \right) \Delta\mathbf{m}^{(k)} = -\mathbf{g}^{(k-1)} \\
 & \text{Update } \mathbf{m}^{(k)} = \mathbf{m}^{(k-1)} + \Delta\mathbf{m}^{(k)}. \\
 & \text{If } k \text{ satisfies linear update criteria, } \mathbf{c}^{(k)} = \mathbf{F}_w^\dagger(\mathbf{m}^{(k)})\mathbf{d}_w; \\
 & \text{Else, } \mathbf{c}^{(k)} = \mathbf{c}^{(k-1)}.
 \end{aligned} \tag{23}$$

By varying the linear update criterion, this implementation can potentially incorporate a spectrum of inversion schemes. For instance, by disabling update on the linear model until the inversion on the nonlinear part has converged (or stagnated), one obtains one end-member scenario, which is exactly the approach described in Koch and Kuvshinov (2013). This scenario contains the least frequent linear model updates. In contrast, by forcing linear model regression at each iteration, one obtains the other end-member, a scheme equivalent to the VP-RW3 (Eq. 22). A customized linear update criterion allows for intermediate solutions between these two end-members.

Both VP and alternating approaches provide alternative means to solve the joint model space inversion (Eq. 10). Although beyond the scope of this work, it can be further shown that the linear system for nonlinear updates resulting from VP/alternating approaches at each iteration is also closely related to that obtained in the joint model space inversion (see Appendix B). Therefore, these surrogate methods all sample subsets of the manifold describing the objective function in the higher dimensional joint model space.

As a final remark, we discuss in brief the computational aspects of different variants of VP and alternating approaches assuming a common scenario where evaluation of $\mathbf{D}\mathbf{F}_w(\mathbf{m})$ is the most resource-demanding part in computation of Fréchet derivatives. Specifically, each evaluation of the matrix–vector multiplication $\mathbf{u}^H(\mathbf{D}\mathbf{F}_w\mathbf{v})$ or $\mathbf{v}^H\mathbf{D}\mathbf{F}_w^H\mathbf{u}$ would incur a forward or an adjoint solution of the electromagnetic modelling problem, respectively, with \mathbf{u} and \mathbf{v} being arbitrary vectors of matching dimensions. Following Eqs. 18–22, evaluation of the Jacobians in VP-RW2 and VP-RW3 involves only one evaluation of $\mathbf{D}\mathbf{F}_w$, while full-VP incurs two evaluations. Therefore, the computational cost per iteration for VP-RW2 and VP-RW3 is roughly half of that for full-VP. Compared to VP-RW3, VP-RW2 involves an additional linear projection in calculation of the Jacobian. Since the cost for

linear regression is often marginal in comparison to evaluation of derivatives, the difference between the cost per iteration between VP-RW2 and VP-RW3 will not play a significant role. There is virtually no difference between the evaluation of Fréchet derivatives in VP-RW3 and alternating approaches, and hence, the two approaches should be considered equal in terms of cost per iteration, except for extra linear regressions required at each iteration in VP-RW3. As was already mentioned in the previous section, in cases where explicit (e.g., for a 3-D conductivity parameterization) storage of the Jacobian is prohibitive, it can be avoided for all variants of VP or alternating approaches by evaluating Jacobian-vector products on the fly. Obviously, this still preserves the relative cost of different methods discussed above.

Forward modelling

We remind the reader that the separation of the joint model space into linear and nonlinear parts is an innate property of EM induction sounding stemming from the governing Maxwell's equations. Therefore, the formulation provided above is general and will apply to any electromagnetic imaging problem where both source and physical properties are unknown. To test different inversion approaches, we need to choose a specific form of inducing source parameterization \mathbf{c} and a forward modelling operator $\mathbf{F}(\sigma)$. We limit the experiment in this study to a simple scenario satisfying the following two assumptions. First, we consider only observations made within a current-free space between the inducing source and the induced currents. In other words, the observed magnetic field is assumed to be potential ($\mathbf{B} = -\nabla V$), where the potential V can be expanded using Spherical Harmonic (SH) functions in the frequency domain as

$$V(\mathbf{r}, \omega) = a \sum_{n,m} \left[\varepsilon_n^m(\omega) \left(\frac{r}{a} \right)^n + i_n^m(\omega) \left(\frac{r}{a} \right)^{-(n+1)} \right] Y_n^m(\theta, \varphi), \tag{24}$$

where $\sum_{n,m} \equiv \sum_{n=1}^N \sum_{m=-n}^n$; $Y_n^m(\theta, \varphi) = P_n^{|m|}(\cos \theta)e^{im\varphi}$ is a complex SH function of degree n and order m , with $P_n^{|m|}$ being Schmidt quasi-normalized associated Legendre functions, $\mathbf{r} = (r, \theta, \varphi)$ is the position vector in spherical coordinates, and a is the Earth radius; ε_n^m and l_n^m are the external and internal SH coefficients, respectively. These assumptions will facilitate the comparison of our methods with conventional Gauss-based workflows.

Second, we assume a 1-D radial conductivity structure of the Earth (that is, $\sigma(\mathbf{r}) \equiv \sigma(r)$). This assumption allows us to use a Q-response to describe the induction in the

model (Olsen 1999). Q-response is a frequency-dependent global transfer function (TF) that is independent of the SH order m for a 1-D radially symmetric conductivity, and is formally defined as the ratio between the internal and the corresponding external Gauss coefficients

$$Q_n(\omega; \sigma) = \frac{l_n^m(\omega; \sigma)}{\varepsilon_n^m(\omega)}. \tag{25}$$

Then, the forward operator that links magnetic field (\mathbf{B}) with model parameters (external coefficients ε and conductivity σ) can be stated as follows:

$$\begin{aligned} B_r(\mathbf{r}, \omega) &= -\sum_{n,m} \left[n \left(\frac{r}{a}\right)^{n-1} - (n+1)Q_n(\omega; \sigma) \left(\frac{r}{a}\right)^{-(n+2)} \right] Y_n^m(\theta, \varphi) \varepsilon_n^m(\omega), \\ B_\theta(\mathbf{r}, \omega) &= -\sum_{n,m} \left[\left(\frac{r}{a}\right)^{n-1} + Q_n(\omega; \sigma) \left(\frac{r}{a}\right)^{-(n+2)} \right] \frac{\partial Y_n^m(\theta, \varphi)}{\partial \theta} \varepsilon_n^m(\omega), \\ B_\varphi(\mathbf{r}, \omega) &= -\sum_{n,m} \left[\left(\frac{r}{a}\right)^{n-1} + Q_n(\omega; \sigma) \left(\frac{r}{a}\right)^{-(n+2)} \right] \frac{1}{\sin \theta} \frac{\partial Y_n^m(\theta, \varphi)}{\partial \varphi} \varepsilon_n^m(\omega). \end{aligned} \tag{26}$$

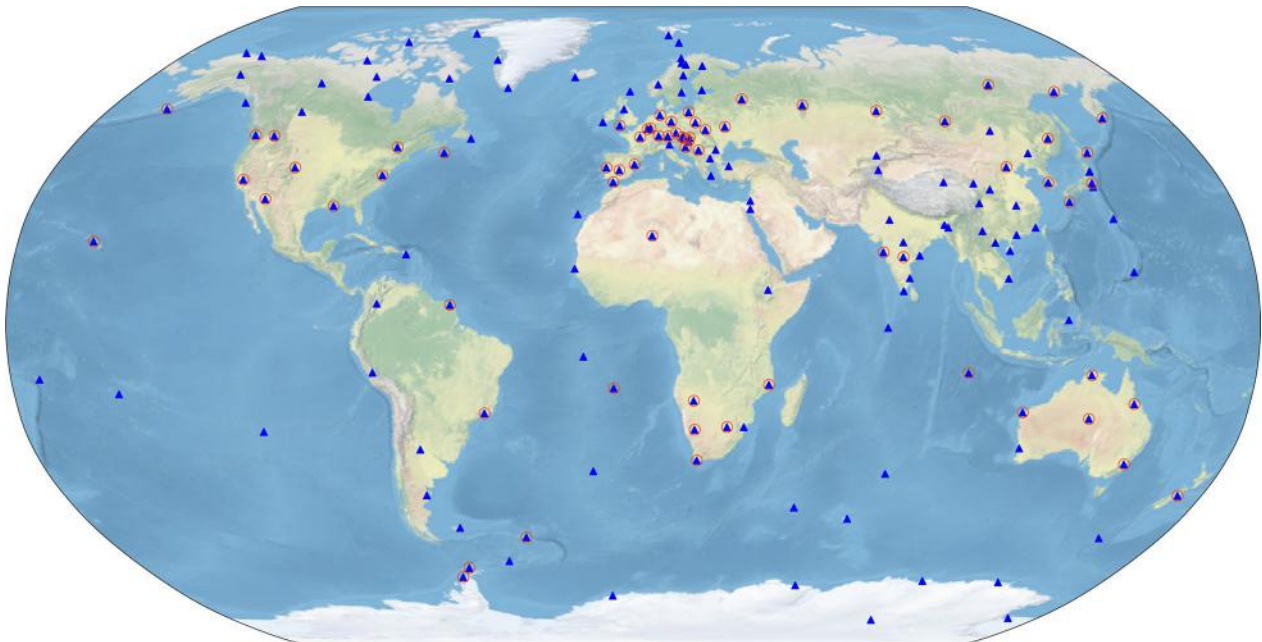


Fig. 1 Distribution of observatories. The blue triangles show all observatories in the dataset. Note the observatories whose absolute magnetic latitudes are greater than 56° or less than 5° are later excluded from experiments. The red circles mark the observatories used in the experiment from the Discussion section

Equation 26 gives the magnetic field at a position \mathbf{r} and at frequency ω in terms of unknown variables σ and ε_n^m , which can be written in the vector form as

$$\mathbf{B}(\mathbf{r}, \omega) = \sum_{n,m} \mathbf{B}_n^m(\mathbf{r}, \omega; \sigma) \varepsilon_n^m(\omega), \tag{27}$$

where $\mathbf{B}(\mathbf{r}, \omega) \in \mathbb{C}^3$ is the vector magnetic field in the frequency domain, and $\mathbf{B}_n^m(\mathbf{r}, \omega; \sigma) \in \mathbb{C}^3$ is the transfer function related to mode ε_n^m for a given \mathbf{r} and ω , whose detailed expression is given in Eq. 26.

While the SH coefficients ε_n^m appear as coefficients of spherical harmonic expansion for the potential magnetic field, they can also be used for representing the inducing current. To this end, consider an extraneous sheet current floating at an altitude h , then the sheet current density can be written as $\mathbf{j}(\mathbf{r}, \omega) = -\delta(r - b)\hat{\mathbf{e}}_r \times \nabla_H \Psi^{\text{ext}}(\theta, \phi)$, where $b = a + h$, and the external current stream function can be expanded in SH using ε_n^m as

$$\Psi^{\text{ext}}(\theta, \phi) = -\frac{a}{\mu_0} \sum_{n,m} \frac{2n+1}{n+1} \left(\frac{b}{a}\right)^n Y_n^m(\theta, \phi) \varepsilon_n^m(\omega). \tag{28}$$

It follows that the coefficients $\varepsilon_n^m(\omega)$ give the parameterization of the inducing currents, and constitute the aforementioned source vector \mathbf{c} .

We stress here that forward operators with other parameterizations of source currents, not limited to a potential representation (Egbert et al. 2021; Zenhäuser et al. 2021), and a general 3-D conductivity distribution (Grayver et al. 2021) are possible and can be incorporated in the formalism of Eq. 6, but this leads to a rather lengthy and technically cumbersome implementation. Choosing a simplified forward operator here allows us to concentrate on studying the properties of the SNLS problem and variable projection method, which we consider to be the main contribution of this study.

To capture the temporal behavior of the external field as well as its properties in the frequency domain, the forward operator and the inversion are both established in the windowed Fourier domain, where each window is considered a realization of the source. For a given frequency ω and a time window τ , the magnetic field is related to the source coefficients via

where \mathbf{d} denotes the data vector, \mathbf{c} denotes the vector of external spherical harmonic coefficients, and their subscripts τ and ω indicate the time window and frequency, respectively. As a common practice, we always inverted for logarithmic electrical conductivity directly in this study, in which case $\mathbf{m} = \log(\sigma)$ denotes the logarithmic conductivity (in S/m) of the subsurface. By concatenating time windows and periods, the forward operator with respect to the complete set of observations can be recast to the algebraic form given by Eq. 6.

Data

Our original dataset consists of hourly mean observations of the magnetic field at 163 observatories shown in Fig. 1. Both synthetic and real data experiments were carried out based on these observations. Synthetic data were generated using a realistic external field, and a two-layer mantle electrical conductivity model, following the procedures outlined below. We took observatory magnetic field hourly means for years 2014–2018 and subtracted the core, crust, and ionospheric contributions as given by the Comprehensive Inversion within the ESA Swarm data processing chain (Olsen et al. 2013). Then, hourly time series of the external and internal coefficients up to degree and order three were estimated using SHA and robust regression. Only mid-latitude observatories (observatories with geomagnetic latitudes from 5° to 56° north and south) were used in the process. The purpose of this step is to obtain time series of the external field that are representative of the real large-scale magnetospheric/ionospheric currents in terms of spatial and temporal characteristics. Using the estimated external field coefficients and a pre-defined 1-D Earth electrical conductivity model, synthetic magnetic field time series at real observatory locations were obtained using Eq. 26. In this synthetic test, we set up a simple but realistic two-layer mantle conductivity model, with an upper mantle (surface to 660km depth) electrical conductivity of 0.01 S/m, and a lower mantle (660km to 2900km depth) electrical conductivity of 1.0 S/m. Finally, we contaminated the synthetic time series with a realization of an independent and identically distributed Gaussian white noise with a standard deviation of 1nT.

$$\mathbf{d}_{\tau,\omega}^{\text{mod}} = \begin{bmatrix} \mathbf{B}(\mathbf{r}_1, \tau, \omega) \\ \vdots \\ \mathbf{B}(\mathbf{r}_{N_r}, \tau, \omega) \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1^0(\mathbf{r}_1, \omega; \sigma) & \cdots & \mathbf{B}_N^N(\mathbf{r}_1, \omega; \sigma) \\ \vdots & \ddots & \vdots \\ \mathbf{B}_1^0(\mathbf{r}_{N_r}, \omega; \sigma) & \cdots & \mathbf{B}_N^N(\mathbf{r}_{N_r}, \omega; \sigma) \end{bmatrix} \begin{bmatrix} \varepsilon_1^0(\tau, \omega) \\ \vdots \\ \varepsilon_N^N(\tau, \omega) \end{bmatrix} = \mathbf{B}_{\tau,\omega}(\mathbf{m}) \mathbf{c}_{\tau,\omega}, \tag{29}$$

Since our implementations of the VP methods and the alternating approach are posed in the frequency domain, the data vector \mathbf{d} is prepared by transforming the time series of magnetic field observations using the windowed spectral transform with tapering, which is defined in discrete form as

$$X(\tau, \omega) = \mathcal{F}_{\tau, \omega}[x(t)] = \frac{1}{\sum_n w_n} \sum_{n=1}^{N_\tau} w_n x(t_n) e^{-i\omega t_n}, \tag{30}$$

where t_n marks time points within a time window, N_τ is the number of points within a time window τ , and w_n is a weighting coefficient associated with the n th point for a tapering window function. Choice of normalization does not affect the inversion scheme, but adopting this specific convention preserves the amplitude during the transform (i.e., a monochromatic oscillating field of period T and amplitude A will be transformed to a peak of amplitude A at frequency $1/T$), yielding physical meaning to the recovered inducing field.

Finally, we comment on the uncertainty of the data. Recall that the data misfit in Eq. 9 is normalized by the standard deviation. Therefore, the uncertainty of the data in the Fourier domain affects weighting and alters the topography of the objective function. We identify two contributions of the uncertainty of the windowed spectrum. First, since measurements are made in the time domain, the noise in the time domain propagates to the Fourier domain, which we term the propagated spectral uncertainty, denoted as s_{prop} . For an arbitrary time series with a Gaussian white noise of variance s_0^2 in time domain, the corresponding uncertainty in windowed Fourier domain following transform defined in Eq. 30 is given by

$$s_{\text{prop}}^2(\tau, \omega) = \text{Var}[X(\tau, \omega)] = \mathbb{E}[|X(\tau, \omega)|^2] - |\mathbb{E}[X(\tau, \omega)]|^2 = \frac{\sum_n w_n^2}{(\sum_n w_n)^2} s_0^2. \tag{31}$$

Therefore, the propagated spectral uncertainty is not merely proportional to its temporal counterpart, but is also dependent on the length of the time window. The proof of this property is given in Appendix C. For this to hold, the noise is assumed to be independent identically distributed with zero mean. The noise we add to the synthetic data indeed satisfies this assumption, and thus, our estimation of this uncertainty is optimal.

In addition to the propagated spectral uncertainty, modelling in the frequency domain in short time windows and using tapering functions introduce extra errors to the problem. While such segmentation and windowing enhance robustness of frequency-domain inversion, it also introduces spectral leakage (see Appendix D for details). This error varies depending on the shape of the spectrum in the vicinity of the frequency, but we assume that it has similar magnitudes for different frequencies. Denoting this error as s_{spec} and assuming s_{spec} and s_{prop} are independent, we can write

$$s^2(\tau, \omega) = s_{\text{prop}}^2(\tau, \omega) + s_{\text{spec}}^2. \tag{32}$$

The windowed transformed data are then inversely weighted by the complete variance $s^2(\tau, \omega)$. In cases where s_{prop} is very small (e.g., at long periods), s_{spec} serves as the baseline (or error floor) for the overall uncertainty. We note that the choice of s_{spec} is generally problem dependent. By comparing the windowed spectrum of a time series modelled with a long time window and the spectrum modelled window-wise, we chose $s_{\text{spec}} \approx 0.05\text{nT}$ for our experiments. This would have only marginal effect for real data because the noise s_0 is large, but it can play a role for our synthetic experiments with the idealized white noise model.

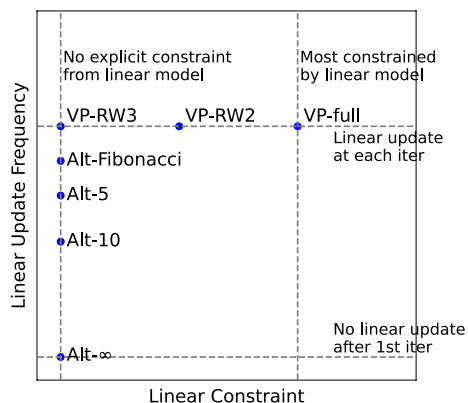


Fig. 2 Schematic regime plot of different inversion methods of VP and alternating approaches. These inversion schemes can be categorized by how strong the linear constraint is enforced, and how often the linear model is updated. Note VP-RW3 can be interpreted as an alternating approach which updates the linear model at each iteration

Results

We conducted both synthetic and real data experiments. In both cases, we implemented a conventional approach for field separation and subsurface conductivity inversion. The conventional inversion scheme includes separating the internal and external fields using the Gauss method, estimating an EM transfer function, and then inverting the transfer function for subsurface electrical conductivity. For consistency, the same transfer function as is used in our forward operator, i.e., the Q -response,

is chosen, and the nonlinear optimization problem for a subsurface conductivity distribution is stated as

$$\min_{\mathbf{m}} \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^K \left| \frac{Q_n^{\text{obs}}(\omega_k) - Q_n^{\text{mod}}(\omega_k; \mathbf{m})}{dQ_n(\omega_k)} \right|^2 + \frac{\lambda}{2} \|\Gamma \mathbf{m}\|_2^2, \tag{33}$$

where Q_n^{obs} gives the Q response estimated from ε_n^m and ι_n^m , and dQ_n is the formal uncertainty. This method will be referred to as Q -response inversion. The model vector \mathbf{m} is again the logarithmic electrical conductivity, same

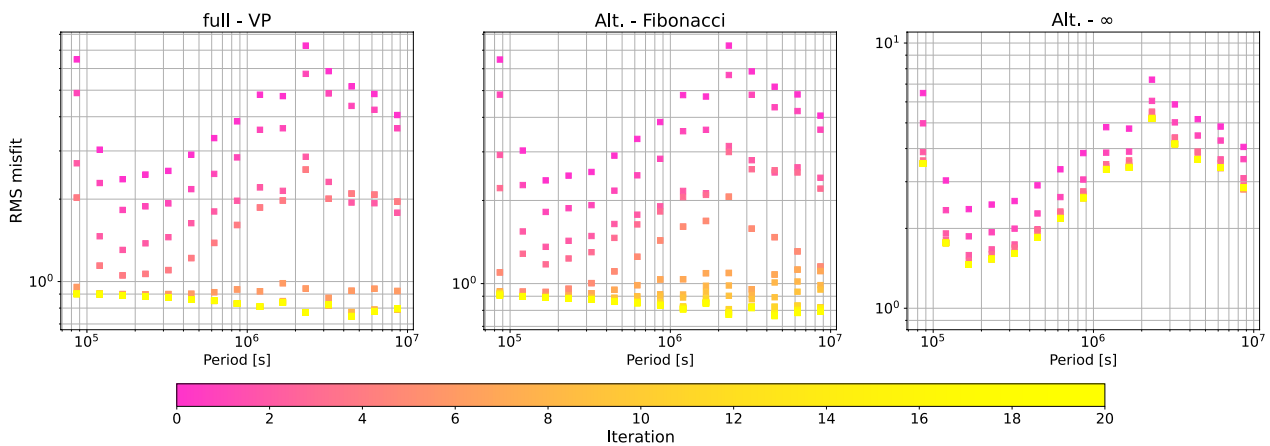


Fig. 3 RMS misfits by periods for full-VP (left), alt-Fibonacci (middle), and alt- ∞ (right). The RMS misfits are color-coded by the iteration number in each plot

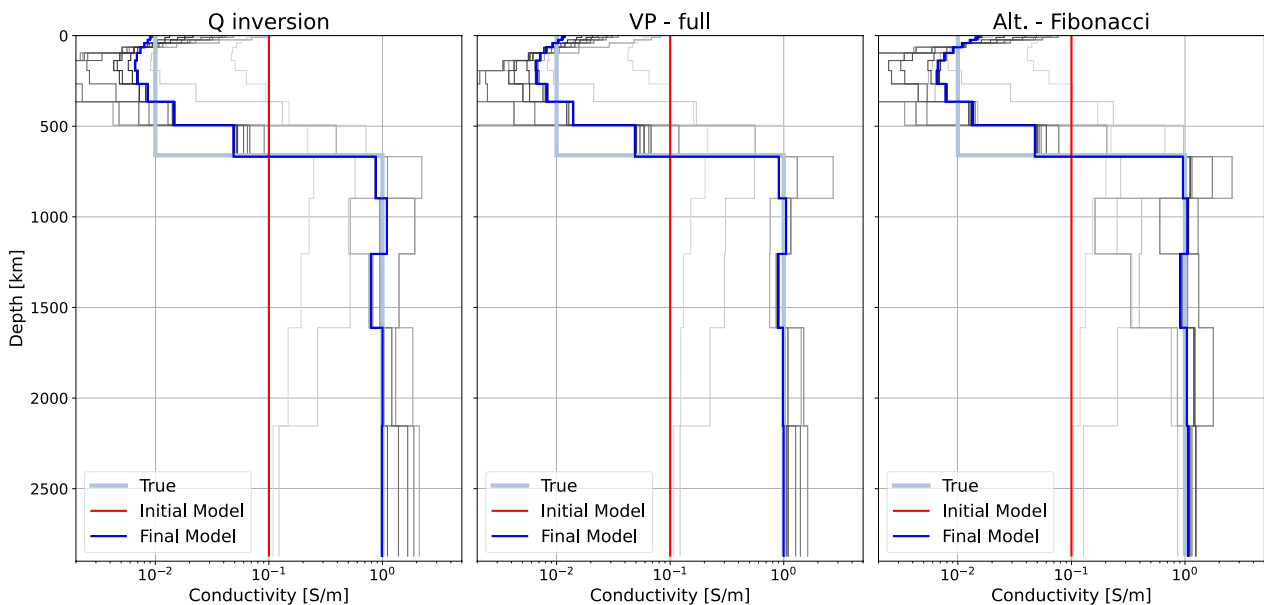


Fig. 4 Electrical conductivity profiles recovered for representative inversion schemes. The results come from Q -response inversion (left), full-VP (middle), and alternating approach using the Fibonacci linear update rule (right). In each case, the final inversion result (blue) is plotted together with the initial model (red) and the ground truth (light blue). The intermediate models along the iterations are also plotted in gray scale, color-coded by iteration numbers. Light gray corresponds to early iterations, while dark gray corresponds to later iterations

as our forward operator in VP/alternating approach. The Tikhonov matrix Γ in Eqs. 10 and 33 is equivalent. A solution that minimizes the objective function (33) is sought using a Newton-based algorithm. The regularization strength as a hyperparameter in inversion is identified using the L-curve analysis (Hansen and O’Leary 1993).

We note that the absolute values of the data misfit for Q -response inversion and VP/alternating methods are not directly comparable. Even when the data misfit is normalized by the number of data samples, Q -response inversion and VP/alternating methods perform inversion with respect to different data. The latter directly works on the magnetic field data in the frequency domain, the uncertainties of which are propagated from the time-domain estimates, while the former is conducted on estimates of transfer functions and their formal uncertainties obtained from regression in spectral domain. We therefore do not intend to directly compare the magnitude of the objective function between Gauss and VP/alternating methods. It naturally follows that the respective suitable regularization strengths are also not directly comparable in magnitude, as they would vary between these two types of scheme. It is, however, worth mentioning that VP and alternating methods, regardless of their variants, share the same data misfit evaluation, and are thus comparable in terms of data misfit as well as regularization strengths. We stick to the standard procedure of choosing regularization for each scheme, but such choices may be general across different variants of VP and alternating approaches, as is indeed the case in the experiments shown later.

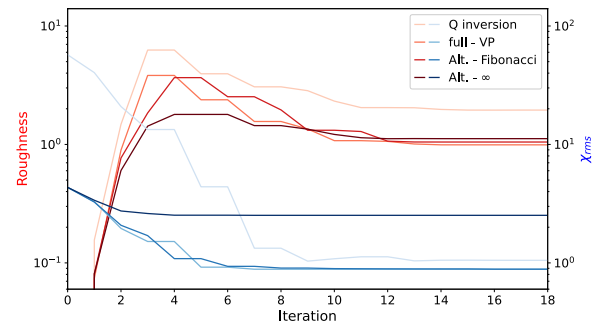


Fig. 6 Evolution of root-mean-square misfit and model roughness. The evolution of normalized RMS misfit (right axis) is plotted in blue, and the evolution of model roughness (left axis) is plotted in red. Roughness is evaluated by calculating $\|\Gamma \mathbf{m}\|_2^2$. From light to dark colors, the parameter evolution curves are plotted for Q -response inversion, full-VP, alt-Fibonacci scheme, and alt- ∞ scheme, respectively

Synthetic experiment

Since both Q -response inversion and the VP/alternating approaches are done in the frequency domain, we choose the same discrete frequencies for all inversions. For the synthetic experiment, we chose 15 periods log-spaced between 1 and 100 days. The inducing field to be determined is parameterized using SH functions up to SH degree and order 3, and the electrical conductivity model is parameterized as a 15-layer 1-D profile. Although the coherence of the Q -responses obtained by Gauss method is adequately high for all modes and frequencies due to the ideal synthetic data, we chose to invert only the degree one response Q_1 , as is done in practice. VP/alternating approaches, on the other hand, automatically try to fit all modes and frequencies simultaneously.

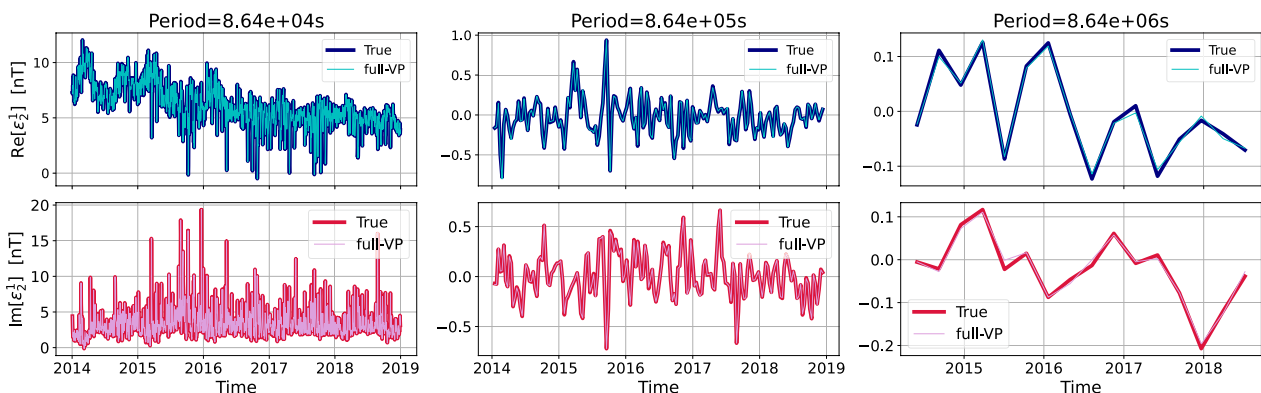


Fig. 5 Recovered inducing field coefficient e_2^1 in the frequency domain using full-VP. Three frequency bands (left: 1 day; middle: 10 days; right: 100 days) are shown. The windowed spectra are split into real parts (upper panel) and imaginary parts (lower panel). In each subplot, the inversion results (cyan for real parts and light pink for imaginary parts) are shown together with the windowed spectra of the ground truth (thick blue and red lines)

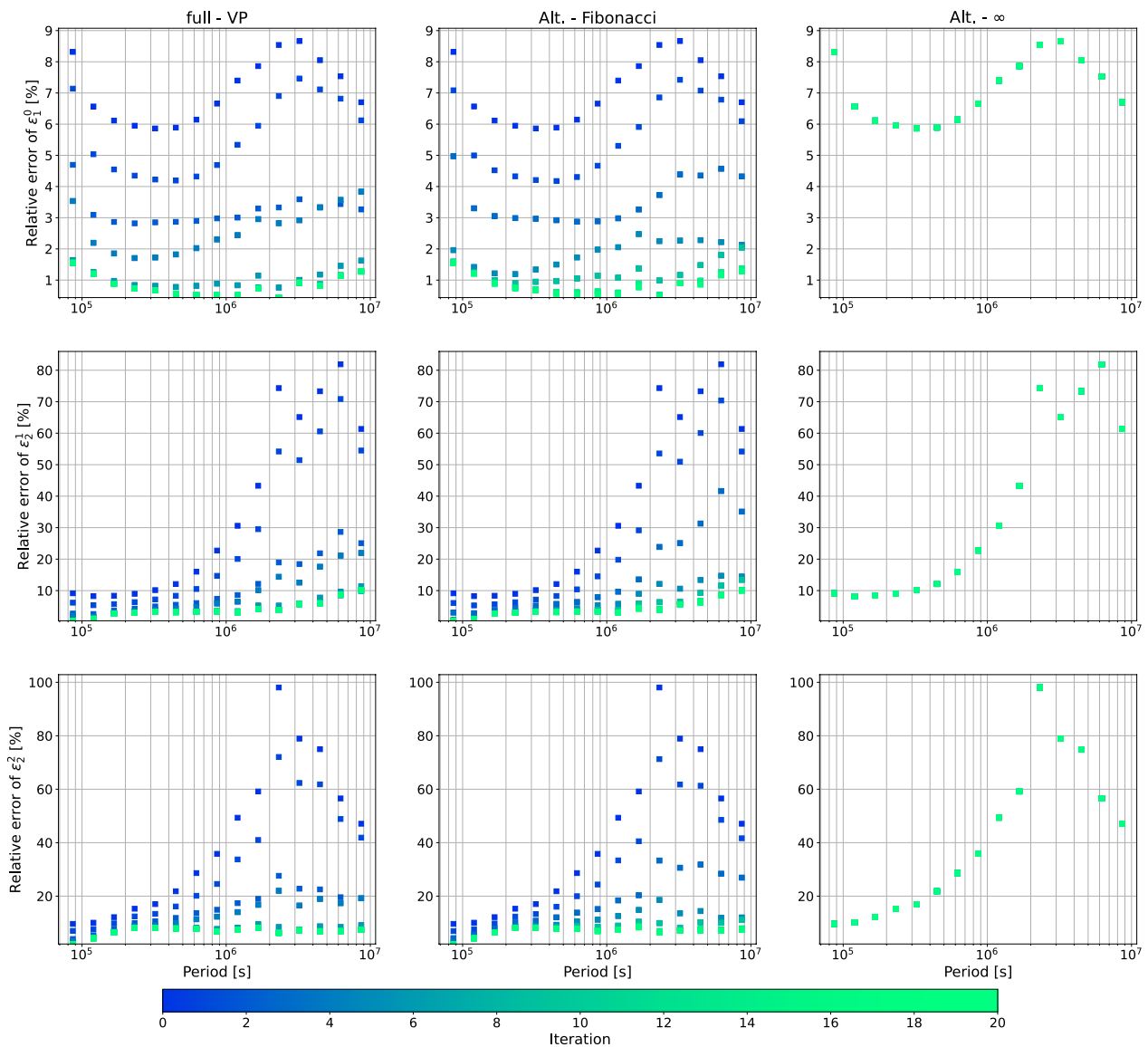


Fig. 7 Frequency-wise relative errors between the estimated and the true external SH coefficients for different modes. The columns show the RMS errors using full-VP (left), alt-Fibonacci (middle), and alt-∞ (right), respectively. Different rows are for different SH modes, i.e., (1, 0) (top), (2, 1) (middle), and (2, 2) (bottom). In each plot, the RMS errors are color-coded by the iteration number

For VP methods, all three variants were tested on the synthetic dataset. For alternating approaches, we tested four different linear update rules: (1) the external field is estimated once at the beginning using some initial conductivity model, and never updated afterwards; (2) the external field is re-estimated every ten iterations; (3) the external field is re-estimated every five iterations, and (4) the external field is updated following the Fibonacci sequence (that is, at iterations 1, 2, 3, 5, 8,...). The alternating methods with these four linear update rules will be abbreviated as alt-∞, alt-10, alt-5, and alt-Fibonacci,

respectively. These variants of VP and alternating approaches can be uniformly described by the two controlling parameters, i.e., the incorporation of linear constraints, which varies for different variants of VP, and the frequency of linear update, which varies for different variants of alternating approaches. The relation of these inversion schemes is schematically summarized in Fig. 2.

Most of the synthetic data inversions produce satisfactory results of the conductivity profile within ~ 20 iterations. In particular, we first examine two representative cases, namely full-VP and alt-Fibonacci, in addition to the

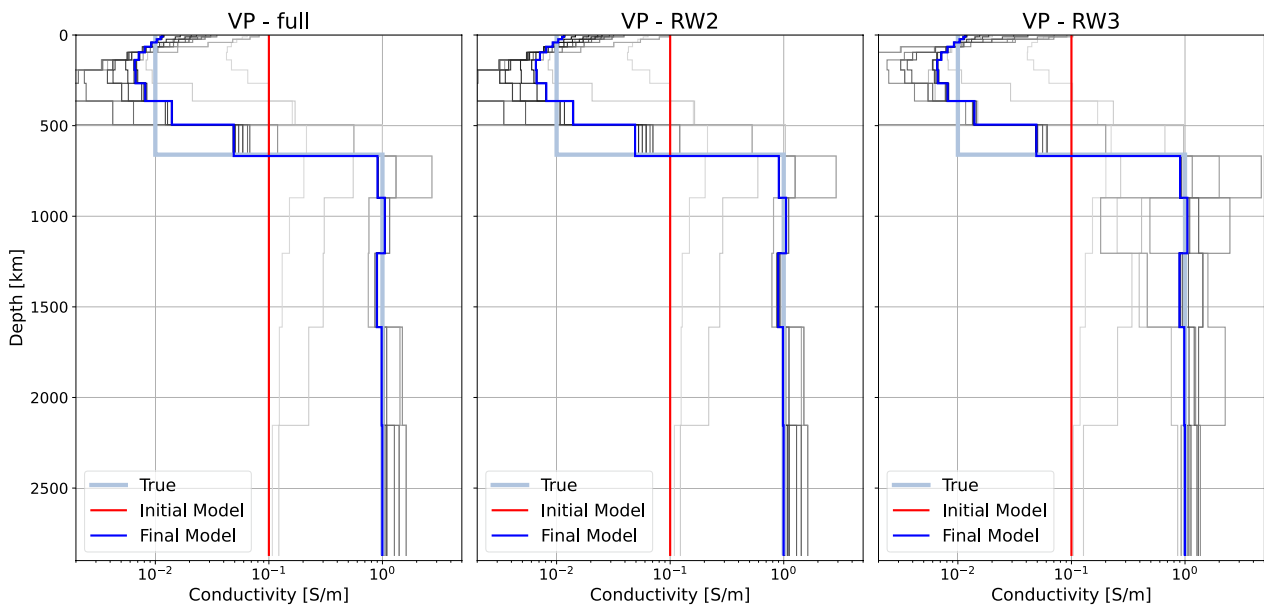


Fig. 8 Electrical conductivity profile recovery for variants of VP. The results come from full-VP method (left), VP-RW2 (middle), and VP-RW3 approximation schemes (right). In addition to the ground truth (light blue), the initial model (red), and the final inversion result (blue), the intermediate models along the iterations are also plotted in gray scale, color-coded by iteration numbers. Light gray corresponds to early iterations, while dark gray corresponds to later iterations

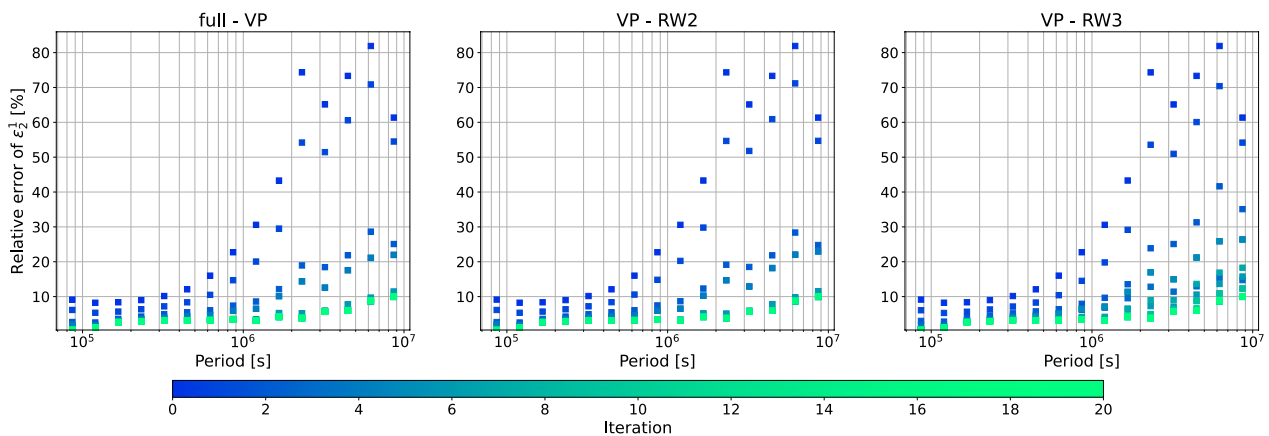


Fig. 9 Relative error between the estimated and the true $\epsilon_2^1(\omega)$ coefficients for different variants of the VP method at different inversion iterations (color-coded). Shown are full-VP (left), VP-RW2 (middle), and VP-RW3 (right)

reference Q-response inversion. These include the best-performing schemes within VP variants and alternating approach variants, as will become clear in the results and discussions that follow. In full-VP and alt-Fibonacci schemes, the converged models yield normalized RMS values (χ_{rms}) of ≈ 0.9 , indicating successful data fitting. In addition, the RMS misfits are roughly at a uniform level across the considered range of frequencies (Fig. 3). The recovered mantle conductivity profiles are shown in Fig. 4. The conductive lower mantle is recovered almost

perfectly, especially in the case of full-VP and alt-Fibonacci, while the inverted upper mantle conductivity follows a gradual decrease from 600km depth upwards, and, in these cases, experience a mild reverse trend at lithospheric depths, mostly appearing as a result of regularization and low sensitivity to these depths.

Variable projection inversions as well as alternating approaches simultaneously produce an estimate of the linear model, i.e., the inducing field in this context, along with an estimate of the mantle electrical

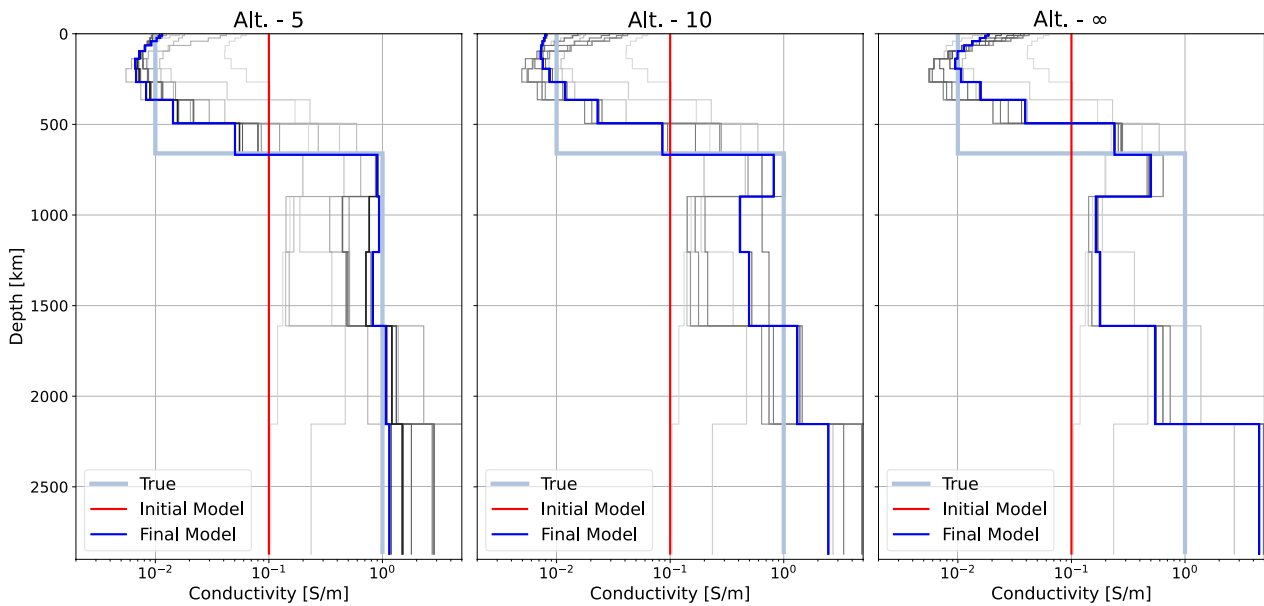


Fig. 10 Electrical conductivity models recovered using variants of alternating approaches. The variants shown update inducing source coefficients every five (left) or ten (middle) iterations, or never update the inducing source after the initial estimation (right). The color coding of the lines and legends is the same as in Fig. 8

conductivity. Since the inversions are carried out in the frequency domain, the linear model is estimated in the form of windowed spectrum $\varepsilon_H^m(\tau, \omega)$. In the synthetic test, the ground truth external coefficients are known, and hence can be used to validate the windowed spectrum of the inducing field SH coefficients inverted using these approaches. A comparison for three period bands of $\varepsilon_2^1(\tau, \omega_i)$ with $\omega_i = 2\pi/T_i$ and $T_i = 1, 10$ and 100 days obtained using full-VP is presented in Fig. 5. From visual inspection, our synthetic tests for VP yield almost perfect recovery of the windowed spectrum of the external

field. This is not limited to frequency bands or spherical harmonic components with strong external signals (e.g., the daily band ~ 1 day for ε_2^1 , Fig. 5 left panel), but also applies to less energetic frequency bands and SH modes (e.g., Fig. 5 right panel). Similar results are observed for alternating scheme with the Fibonacci linear update rule (Fig. 22).

We have hence demonstrated that with appropriate hyperparameters and specific variants, all types of inversion methods are able to yield satisfactory solution on the synthetic dataset. Furthermore, we observe

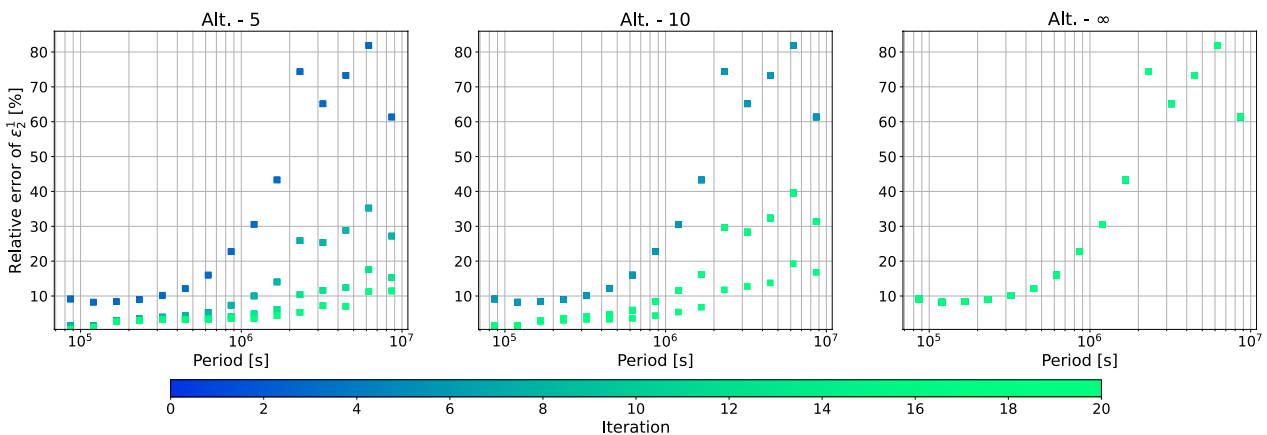


Fig. 11 Relative error between the estimated and the true coefficient $\varepsilon_2^1(\omega)$ for different variants of the alternating approach. The variants shown are alt-5, alt-10, and alt-∞, which encode the frequency of the linear model update

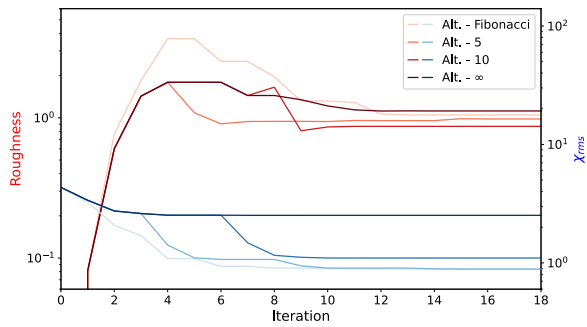


Fig. 12 Evolution of RMS misfit and model roughness for alternating approaches. The displayed quantities and the calculation of roughness are the same as in Fig. 6. The curves are color-coded by inversion schemes, the lightest ones to the darkest ones corresponding to alt-Fibonacci, alt-5, alt-10, and alt-∞, respectively

similar convergence behavior across these representative inversion schemes (Fig. 6). All of these inversion cases converged to the stable solution within 20 iterations. We stress that, at least for this experiment, the full-VP scheme and alt-Fibonacci scheme exhibit convergence rates that are at least as good as or slightly outperform the Q-response inversion, even though the SNLS problem has seemingly more "work" to do, because it also estimates the source structure for all 15 SH coefficients at all frequencies.

In all cases, we observe an initial upsurge in the model roughness, followed by a gradual decrease, accompanied by an almost monotonic decrease of the data misfit. This behavior is expected as the inversions start from a uniform model with zero roughness, thus they first attempt to fit the data at the cost of expanding model complexity, and then stabilize by settling for a smoother model after the data misfit reaches a certain level.

Since the inducing field model is known for the synthetic study, we also examine how the inducing field estimates converge towards the ground truth solution for VP and alternating methods. We introduce the frequency-wise relative error for the SH coefficients, defined as

$$\epsilon_{SH}^{nm} = \sqrt{\frac{\sum_i |\epsilon_n^{m,true}(\tau_i, \omega) - \epsilon_n^{m,est}(\tau_i, \omega)|^2}{\sum_i |\epsilon_n^{m,true}(\tau_i, \omega)|^2}}. \quad (34)$$

The evolutions of $\epsilon_{SH}^{nm}(\omega)$ as a function of iterations for full-VP and alt-Fibonacci are shown in Fig. 7. For VP, the inducing field monotonically approaches the ground truth except for marginal oscillations around the converged solution, and eventually converges at iteration 6, hence slightly earlier than the conductivity model (Fig. 4). For the alt-Fibonacci scheme, the inducing field does not converge until the 13th iteration. Since alternating

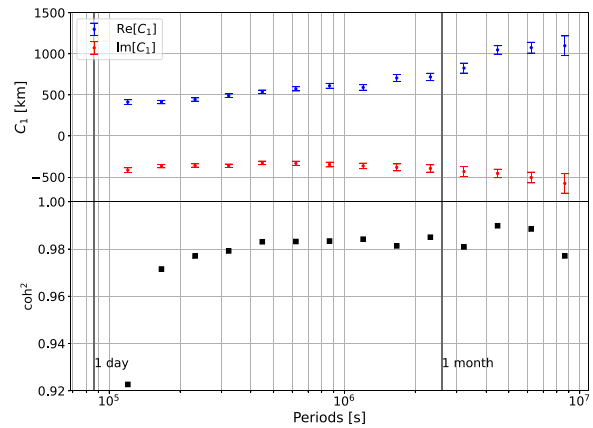


Fig. 13 C-responses and their corresponding coherence. The top panel shows C_1 -response estimated from (1, 0) spherical harmonic coefficients with their uncertainties, and the lower panel shows the corresponding squared coherence

approaches do not update the inducing currents at each iteration, only five linear projections have been made by iteration 13, following the Fibonacci linear update rule. Slow convergence of the inducing field solution is especially observed at long periods when using the alt-Fibonacci scheme. For instance, the estimated inducing field at iteration 8 has the same level of relative error with the field at iteration 5 for the 100-day period band (Fig. 7, middle column). This shows the importance of maintaining consistency between the source and conductivity models, which is most consistently done in the full-VP method.

Our tests using different variants of VP show that all variants of VP converge to almost the exact same electrical conductivity model that is satisfactorily close to the ground truth for the synthetic data (Fig. 8). However, we do observe that VP-RW3 exhibits slower convergence for both linear (Fig. 9) and nonlinear parts of the model space (Fig. 8). Most period bands of the external field take 8–10 iterations to converge. Deterioration of the solution, i.e., increased error in the external field at later iterations, is also observed for VP-RW3 for periods longer than 10^6 s, while such increase in error is absent in full-VP and VP-RW2. As is shown in Eqs. 18–22, implicit feedback of the inducing field estimate is utilized by both full-VP and VP-RW2, but absent in VP-RW3. The observed slower convergence as well as deterioration is hence the result of omitting the relevant terms in VP-RW3. We anticipate that the performance of VP-RW3 (and alternating approaches) will further deteriorate for more complex and high-dimensional models.

The conductivity model recovery for different variants of the alternating approach is shown in Fig. 10. As was shown above, updating the inducing field coefficients

at iterations using a Fibonacci sequence still allows the inversion to reach a solution that is fairly close to the VP solution (Fig. 4) within 20 nonlinear iterations with only six linear updates, but as soon as the frequency of linear updates is reduced to every five iterations, considerable deterioration of the electrical conductivity recovery occurs, particularly in the lower mantle (Fig. 10). Not only is the inversion taking longer to reach a stationary point, but the scheme fails to locate the best-fitting nonlinear solution within 20 iterations as well, proving to be at best only half as efficient as the VP or alt-Fibonacci.

Interestingly, despite the deteriorated recovery of the mantle conductivity, alt-5 produces an estimate of the external field that is almost the same in accuracy to VP or alt-Fibonacci scheme (Fig. 11), with slightly increased error only in some long period bands. However, as seen in Fig. 12, the misfits and roughness values for alt-5 have already stagnated at the final stage, and the convergence criterion is satisfied at iteration 18, indicating that the optimization has already converged. Therefore, the misfit between the ground truth and the inverted conductivity models can only be attributed to the marginal difference in external field, and the final inverted conductivity profile should be considered the optimizer of the manifold constrained by the slightly incorrect external field. Alt-5 provides such a clear example where a relatively small error in the inducing source field leads to significant artifacts in conductivity model. When the number of linear updates are cut even further, both the external field

estimation and the mantle conductivity recovery deteriorate further, as in the case of alt-10 and alt-∞ (Figs. 10 and 11).

It is worth mentioning that we observe two types of "stagnation" behavior in our inversions. In one scenario, the model estimates along with the diagnostic parameters, such as χ_{rms} and roughness, either fulfill the convergence criterion and terminate the inversion, or oscillate mildly in the vicinity of a stationary value. The model is considered converged in this scenario, whether to the vicinity of the ground truth (as in VP variants and alt-Fibonacci scheme) or to a local optimum (undoubtedly the case in alt-5) of the manifold in the joint model space. In another scenario (observed in the case of alt-∞), the objective is not improved for more than eight iterations. This behavior indicates that the trust-region Newton method we employed for optimization repeatedly rejected all update proposals, likely because of a poor local quadratic approximation and severe ill conditioning far from the optimum.

Real data inversion

We applied the VP method to the real ground geomagnetic observatory data measured between years 2014–2018 to simultaneously reconstruct mantle conductivity and external field spectrum. The data come from 120 geomagnetic observatories within the mid-low geomagnetic latitude range of 5°–56°. Several amendments were

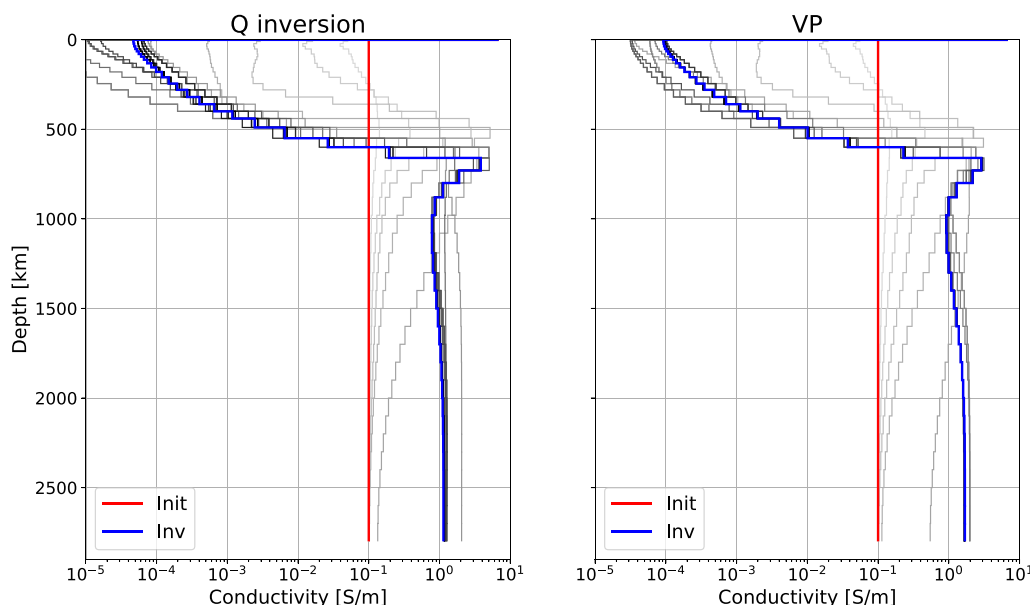


Fig. 14 Mantle conductivity obtained from ground magnetic field observations from years 2014–2018, using the variable projection inversion. The legends and the color coding of the intermediate inversion steps are the same as in Fig. 8. A thin surface layer with radial conductance of 6600S was fixed throughout the inversion

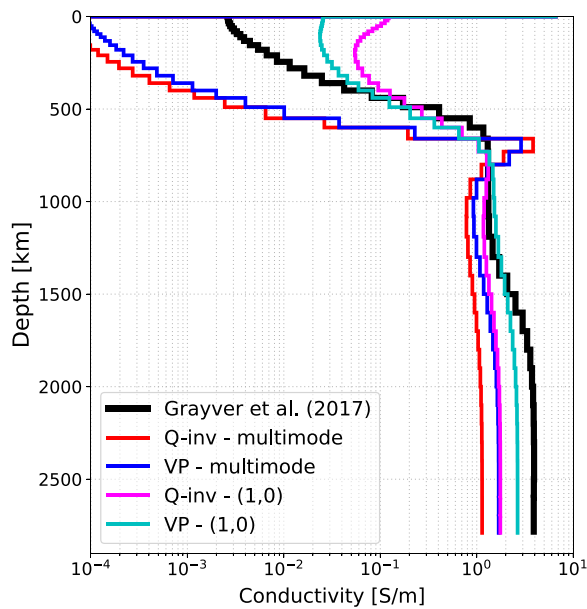


Fig. 15 Mantle conductivity profiles from this study and Grayver et al. (2017). The inversions using all three SH degrees are shown in blue and red lines, whereas the inversions using only the first zonal harmonic are shown in cyan and magenta lines

introduced to the workflow used in numerical experiments to adjust the method to real-Earth sounding. First, a new frequency band (12 hours) is added to the 15 frequency bands spanning 2 decades. This is used for improving constraints to the asthenospheric conductivity. Second, we expanded the parameterization of the mantle conductivity to a 45-layer 1-D profile. Following Grayver et al. (2017), we used a fixed surface layer with the conductance of 6600 S that represents an average ocean-sediments conductance over the globe.

First, we obtained an estimate of the Q -responses by applying conventional Gauss method and robust spectral stacking. The responses can be converted to and visualized as a global C -response (e.g., Olsen 1999), via the equation

$$C_n(\omega) = \frac{a}{n+1} \frac{1 - \frac{n+1}{n} Q_n(\omega)}{1 + Q_n(\omega)}. \quad (35)$$

Having the dimension of length, the real part of the C -response corresponds to the central depth of induced currents, and is hence indicative of the penetration depth of the EM field at a certain frequency (Weidelt 1972). The available data and frequency bands used in this study are most sensitive to a depth range of 500–1500 km, as can be seen from the C_1 responses (Fig. 13). Based on squared coherences, we use the following Q -responses as the input data for subsequent conductivity inversion: Q_1 estimated from the SH coefficient (1, 0) at periods longer than 1 day, Q_2 estimated from the SH coefficient (2, 1) at diurnal band (24 hours), and Q_3 estimated from the SH coefficient (3, 2) at semi-diurnal band (12 h). Variations in the daily band are mostly driven by ionospheric current systems.

For VP, we took only observatories with at least 99% of valid data in each time window and linearly interpolated the missing observations before conducting the windowed spectral transform (Eq. 30). For real time series of the magnetic field, no information on either spatio-temporal properties or magnitude of the noise is available. It is apparent that when assuming a Gaussian white noise, the assumed noise magnitude serves as a mere normalizing factor for the data misfit (Eq. 31), but would not alter the topography of the objective function. It then follows that the properties for inversions (convergence, results, etc.) are preserved except for corresponding

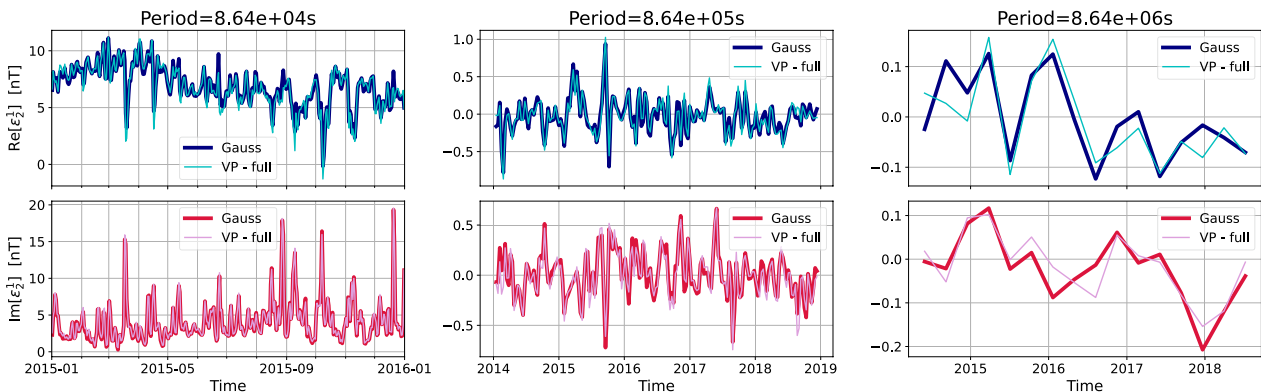


Fig. 16 Recovered inducing field coefficient ϵ_2^1 in the frequency domain. The frequency bands are the same as Fig. 5. In each subplot, we show the VP inversion results (thin cyan and pink lines) on top of the windowed spectra of the external field obtained by Gauss method (thick blue and red lines)

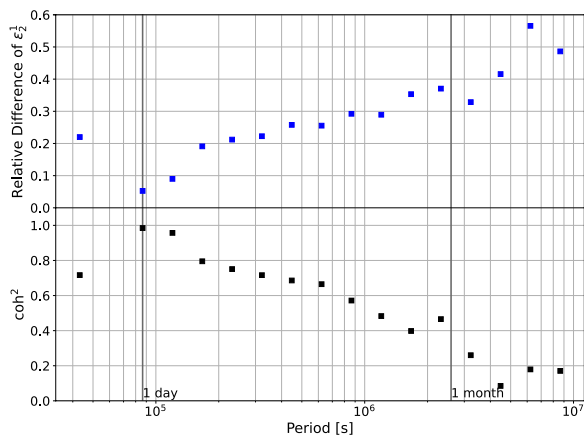


Fig. 17 Relative inducing field difference and coherence of the Q-response. The top panel shows the relative difference between ϵ_2^1 estimates from Gauss method and VP, and the bottom panel shows the coherence of Q_2 estimated from Gauss-method-derived ϵ_2^1 and ι_2^1

normalization factors for regularization terms and uncertainties. Meanwhile, the spectral and temporal characteristics of the unknown noise pose a much greater threat. Without the Gaussian white noise assumption, our estimation of the Fourier-domain uncertainty would be invalidated. Nevertheless, we proceed by assuming a Gaussian white noise with the standard deviation of $1nT$ for real observations.

A series of VP inversions were conducted with varying regularization strengths, and the desired hyperparameter was chosen based on the L-curve (Fig. 24). The inversion results of the mantle conductivity profile using VP

as well as conventional Q-response inversion are shown in Fig. 14. Both conductivity models show a resistive upper mantle with a conductivity monotonically increasing from $10^{-3}S/m$ at lithospheric depths to $1S/m$ at the bottom of the MTZ. Due to the limited sensitivity of our data to the upper mantle (Fig. 13), however, the magnitude and the detailed shape of conductivity in this region are less reliable. The mantle conductivity at 750km depth just beneath the MTZ is well constrained by our data at 2 to $3S/m$, characterized by a conductive peak, a feature that appears quite robust when using weaker regularizations. Beneath this conductive layer, the lower mantle is characterized by a resistive kink, followed by mildly increasing conductivity from $\sim 1S/m$ at 1200 km depth to $2S/m$ at 1600.

Despite resemblance between the conductivity profiles produced using Q-response inversion and VP inversion, these models show considerable discrepancies compared to previous 1-D conductivity models, e.g., the 1-D profile inverted from C_1 in Grayver et al. (2017), shown in Fig. 15. Compared to the previous models, the conductive peak and the resistive kink in our inverted models are much more pronounced, and the depth within the MTZ in our model is considerably more resistive. These discrepancies can be reconciled, however, if we used only the (1, 0) SH mode in the inversion. Instead of combining data with different SH degrees (i.e., $n \leq 3$), by inverting only the Q_1 responses estimated from (1, 0), or by parameterizing the source structure using only the first zonal harmonic (controlled by the magnetospheric ring current) in the VP inversion, the obtained conductivity profiles match well with the previous profiles within the depth range where

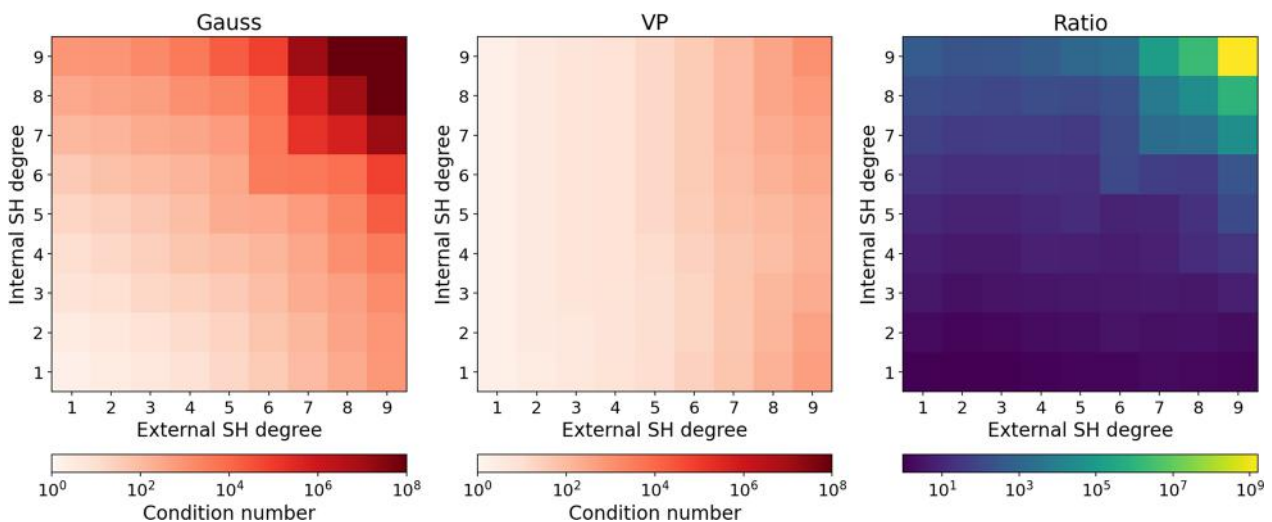


Fig. 18 Condition number of the linear system for field estimation in Gauss method (left) and VP (middle), with varying truncation SH degree for external and internal fields. The ratio between the two condition numbers (condition number of Gauss method divided by that of VP) is shown on the right

our data have adequate sensitivity (Fig. 15). It is not the focus of this study to discuss the differences between models in Figs. 14 and 15. Whether these differences are dictated by source effects or induced by subsurface differences that become more pronounced in other than P_1^0 terms, they imply that the adopted parameterization of the inducing field has a strong impact on the retrieved conductivity model.

As in the synthetic tests, we obtained the windowed spectra of the external field SH coefficients up to degree and order 3 from the VP method. The estimated external field is close to that obtained by the Gauss method (Fig. 16), but the misfit between the two is generally dependent on energy of the mode. For energetic spatial modes, e.g., ε_1^0 at period bands longer than 1 day (Fig. 23), ε_2^1 at 1 day period (Fig. 16 left column), the results are very close. Modes with low power also show correlated trends, but estimated magnitudes disagree. Similar to the relative error defined in Eq. 34, a relative measure of the discrepancy of the estimated SH coefficients can be introduced as

$$\epsilon_{VP-Gauss}^{nm}(\omega) = \sqrt{\frac{\sum_i |\varepsilon_n^{m,VP}(\tau_i, \omega) - \varepsilon_n^{m,Gauss}(\tau_i, \omega)|^2}{\sum_i |\varepsilon_n^{m,Gauss}(\tau_i, \omega)|^2}}. \tag{36}$$

We observe that the relative difference between the inducing fields derived from the Gauss and VP methods is strongly correlated with coherence in the corresponding Q -response estimation. For a spherical harmonic mode, frequency bands with higher Q -response coherence are associated with lower relative differences in the inducing field estimates from different methods, as is clearly shown in the case of mode (2, 1) (Fig. 17). The correlation between the transfer function coherence and the inducing source consistency does not come as a surprise, but is a natural consequence of the physical connection between the inducing field and its induced counterpart. The coherence in transfer function estimation is used to describe how much of the induced field can be causally explained by the transfer function, which in our case is the Q -response, whereas such physical link is explicitly incorporated in VP. Therefore, in frequency bands and SH modes where such physical connection is appropriate for explaining the data (i.e., high coherence), the inducing source estimate in the Gauss method should be more consistent with that obtained in VP.

Discussion

Separability of inducing source modes

We have emphasized that VP/alternating approaches are not limited to potential field and can accommodate

diverse measurements and sources in contrast to inversions based on TFs, such as the Q -response, estimated based on SH coefficients obtained from the Gauss method. In addition, VP and alternating approaches also benefit from more reliable source estimate when compared to the Gauss method. This is because the Gauss method needs to estimate both internal and external coefficients together, while the linear operator in VP or alternating approaches involves only variables for the inducing source.

To illustrate the effect, we explored the dependency of the condition number of linear regression operators upon different combinations of internal and external field parameterizations. High condition number implies nearly co-linear columns, which leads to poorly separable parameters and catastrophic amplification of data errors (Heath 2018). In the Gauss method, the internal and external fields are co-estimated; different truncation degrees of internal field give rise to linear systems of different dimensions. In VP/alternating approaches, the internal field is modelled. Varying the maximum degree of modelled internal field changes the columns in the matrix, but the column dimension of the linear system which corresponds to the external field parameters remains the same. We constructed linear systems for field estimation for both Gauss method and VP, assuming a real observatory layout with 68 observatory locations (red circles in Fig. 1), a configuration taken from the distribution of available observatories in the dataset on Nov 29, 2019. The result is shown in Fig. 18. While the condition number for the Gauss method increases substantially with both internal and external SH degrees, the condition number for VP depends mostly only on the maximum external SH degree, but hardly changes over one order of magnitude for different degrees of modelled internal field. As a result, source determination within the VP method remains well conditioned (e.g., condition number ≈ 20 for maximum external SH degree of 5), whereas the corresponding matrix for the Gauss method may already be very ill-conditioned (condition number $> 10^3$ for the same external field parameterization). This experiment reveals severe limitations of the conventional Gauss method for 3-D scenarios, where the intention is to go for higher degrees in internal field parameterization to capture lateral conductivity variations at desired length scales. In these settings, VP and alternating approaches offer a definite advantage.

Effect of linear update and derivative approximation

Both VP and alternating approaches provide efficient means to find a suitable solution in the joint model space by introducing point-dependent constraints. Whereas variable projection variants re-estimate the

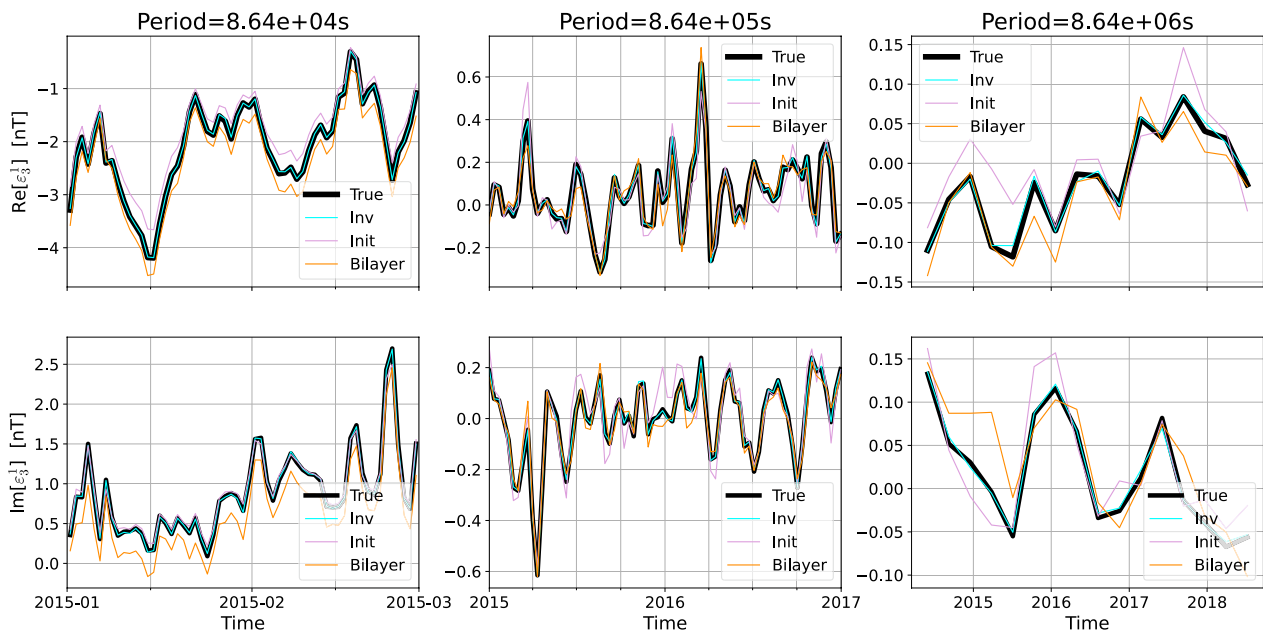


Fig. 19 Estimated windowed spectrum of the inducing field coefficient ϵ_3^T . Frequency bands are the same as in Fig. 5. For each frequency band, the inducing field coefficients are shown for the VP inversion result (cyan), estimation using a uniform mantle conductivity of 0.1S/m (magenta) and a simplistic perfect insulator-conductor bi-layer model (orange). The estimations are produced using the synthetic data, and are plotted on top of the ground truth (black)

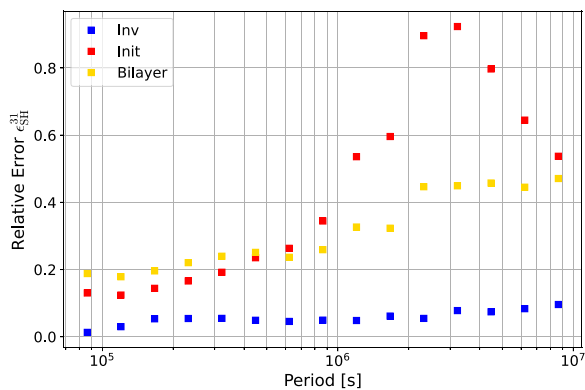


Fig. 20 Relative errors of the external field coefficients ϵ_3^T with different conductivity models. Errors are shown for the VP inversion result (blue), estimation using a uniform mantle conductivity of 0.1S/m (red), and a simplistic perfect insulator-conductor bi-layer model (orange)

local "optimal" constraint at each iteration, alternating approaches allow a user to delay the next re-projection of the linear model and re-estimation of the constraint, implicitly assuming that the constraint remains valid at each subsequent iteration that reuses the initial projection. Although this saves resources and time, we observe that such assumption may not be valid and can potentially cause considerable deterioration of the solution when utilizing the alternating concept. Excessive

iterations using a fixed external field model push a conductivity model away from an optimal solution, which in turn projects the inaccuracy back to the external field at the next iteration (Fig. 7, Alt. - Fibonacci). In real settings, without information about the ground truth, detecting such behavior is practically impossible. Therefore, as appealing as the alternating approach might be due to its simple nature, insufficiently frequent updates of the linear model create a risk of obtaining biased solutions, as was demonstrated in this study. On the other hand, we found that more elaborate update rules, such as alt-Fibonacci, succeed in locating the optimal solution. This process is facilitated by more frequent linear updates at early stages of the inversion. Therefore, alternating approaches should be used with care; in particular, for a given source model, the nonlinear inversion on the conductivity model should not be run until it stagnates, by which time the conductivity model (and with it, the estimate of the external field in the next stage) is probably already biased. In contrast, it is beneficial to alternate between conductivity inversion and inducing field estimation as often as possible initially, with more rare updates to be permitted at later stages.

For our rather simple synthetic tests, we observe no significant difference between the models obtained with different VP variants. However, we observe a slower convergence as well as deterioration of the models in intermediate iterations for the VP-RW3 variant (Fig. 9).

Different levels of approximation of the Fréchet derivatives seem to work almost equally well for the synthetic experiment, and in practice, it might be beneficial to adopt either VP-RW2 or VP-RW3 variants for the sake of computational efficiency, especially when evaluation of DF is expensive (to be the case once full 3-D forward operator is required).

Interplay between conductivity model and external field

The external field model and the mantle conductivity model are mutually dependent in the optimization problem defined in Eq. 10. In particular, the mantle conductivity model is sensitive to perturbations in the external field, as was evident in our experiment where alternations between linear and nonlinear models are done at varied frequencies (alt-Fibonacci, alt-5, alt-10, and alt- ∞). Note that the application of the VP method completely eliminates this problem, and preserves consistency between the linear and nonlinear model unknowns. It does not mean that the VP is less ambiguous than the alternating approach, but it allows one to attain the best (in least-squares sense) possible trade-off between source and conductivity models.

In turn, mantle conductivity has a non-negligible feedback on source reconstruction. To quantify and illustrate the effect, we compared the quality of inducing field reconstruction from synthetic data using simplistic subsurface conductivity models, such as a uniform mantle conductivity of 0.1S/m, which is used as the starting model for all our inversions, and a simplistic two-layer Earth model consisting of a 1200km-thick perfectly insulating mantle and a perfectly conducting core, hereinafter referred to as the bi-layer model. For the perfect insulator-conductor bi-layer model, the Q -response of degree n degenerates to a frequency-independent algebraic form

$$Q_n = \frac{n}{n+1} \left(1 - \frac{z}{a}\right)^{2n+1}, \quad (37)$$

where z is the thickness of the overlying perfectly insulating layer. Due to its simplicity, the bi-layer model is often used in space weather and geomagnetic field modelling to get a first-order estimate of the induction effect.

Figure 19 shows estimates of the external field windowed spectra obtained with different conductivity models. There are considerable discrepancies between the inducing field estimates and the ground truth field when using overly simplistic models. For instance, note the amplitude discrepancies of the field recovered using the bi-layer model, especially in short (e.g., 1 day) and long periods (e.g., 100 days). The discrepancies are also obvious from the relative field errors evaluated in energies (Fig. 20), calculated from Eq. 34. While the inversion

result from VP gives an external field ε_3^1 that is 1% to 5% different from the ground truth in most frequency bands, the aforementioned simplistic models yield external fields that typically exhibit over 15% error. The initial uniform conductivity model, for instance, gives relative inducing field errors of 15% in short periods, which increases to values of 90% at periods of about 1 month. For the simplistic two-layer model, the relative error of the external field increases from 20% in the diurnal band, to 45% in the period bands of 1–3 months. This large discrepancy is partially attributed to the low energies in these modes, but the patterns are the same in more energetic modes.

In short, source estimates calculated using overly simplistic or wrong conductivity models are prone to additional errors (up to 90% in our experiment), even in the scenario where the response is given by a simple 1-D model. For a realistic 3-D earth, the conductivity model might have a more pronounced effect, especially in the vertical component of the magnetic field, when strong lateral variations are present (Grayver et al. 2021).

Conclusions

We addressed the problem of inverting for the inducing source and subsurface conductivity through the solution of a Separable Nonlinear Least-Squares (SNLS) problem. By exploiting the inherent property whereby observations depend on source coefficients linearly, whereas the dependency on the subsurface electrical conductivity is nonlinear, we proposed a novel inversion scheme that solves the underlying SNLS problem using the variable projection to determine source and conductivity structures simultaneously and retain consistency between them. We applied this method to both synthetic tests and real observations. Although our experiments and inversions were limited to ground magnetic field observations and a rather simple 1-D conductivity model parameterization, the provided method derivations are general in both the observational data and the model parameterization aspects. Our derivations in the Methods section present a versatile generic framework for exploiting the VP and show how conventional inversion schemes that often already implement Jacobians for separate source and conductivity estimation can be reformulated into the SNLS form and solved using the VP or alternating approaches. To gain additional insight into the problem, we studied several variants of the VP and showed its relation to the full joint inversion as well alternating inversion approaches.

Alternating approach provides a simple alternative to the VP method that can be used to solve SNLS. However, one important aspect that was not identified in the previous studies is that alternating approach with too rare

source model updates can result in deteriorated model estimates along iterations, which eventually undermines the convergence and model recovery. To avoid this, alternating inversions need to perform frequent re-estimation of the inducing source, especially at early stages. We also observed a slower convergence of the alternating approaches compared to the full-VP method, although this can be compensated in practice by a lower computational cost per iteration.

We demonstrated that by introducing additional constraints to the joint model space, variable projection methods and alternating approaches are capable of recovering both external field and mantle conductivity simultaneously. They show comparable performance to the Gauss method and transfer function inversion on our (simple) test cases, where potential field assumption is applicable. However, unlike approaches that invoke the Gauss method, the SNLS problem solved by the VP method is not limited to the potential field scenarios. In particular, it can accommodate arbitrary source geometries at arbitrary locations [e.g., current loops, dipoles, spherical elementary current systems (SECS)]; incorporate electric field data as well as both ground and satellite observations. Importantly, VP methods make the explicit use of the physical link (through Maxwell's equations) between the source and conductivity, which ensures that consistency between both model spaces is preserved (at least to an extent for which data coverage and quality allow). This is in contrast to the conventional approaches where the source and conductivity are estimated independently and it is often (implicitly) assumed in subsequent transfer function estimation and inversion that the external source estimate is "noise-free". Our synthetic tests showed that even small inconsistencies in a source model can lead to significant artifacts in subsurface conductivity. We also showed that inadequate modelling of the induced field leads to a biased estimate of the external field structure.

Appendix A: VP with regularization on the linear part

For the sake of completeness, we present formulas for VP assuming a scenario where the source estimation requires regularization. This would be the case when a

rich parameter space is needed to represent a source with a complex geometry. In accordance with the metric used in the rest of this paper, we only consider the ℓ_2 regularization. Regularization on the linear parameters does not change the general pipeline of inversion, but would alter the form of the Fréchet derivatives, as will become clear in this following derivations.

We shall start from the general form of the joint optimization problem (Eq. 10). To exploit the pre-existing derivations for VP without the regularization term on linear parameters, we introduce the augmented residual vector, defined as

$$\begin{aligned}\tilde{\mathbf{r}}_w &= \begin{bmatrix} \mathbf{d}_w - \mathbf{F}_w \mathbf{c} \\ -\lambda^{1/2} \mathbf{\Gamma}_c \mathbf{c} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_w \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{F}_w \\ \lambda^{1/2} \mathbf{\Gamma}_c \end{bmatrix} \mathbf{c} \\ \mathbf{c} &= \begin{bmatrix} \mathbf{d}_w \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{F}_w \\ \tilde{\mathbf{\Gamma}}_c \end{bmatrix} \mathbf{c} = \tilde{\mathbf{d}}_w - \tilde{\mathbf{F}}_w \mathbf{c},\end{aligned}\quad (\text{A.1})$$

where we use the notation $\tilde{\mathbf{\Gamma}}_c = \lambda^{1/2} \mathbf{\Gamma}_c$. The explicit Jacobians on the augmented residual vector are linked to the original Jacobians (Eq. 13) through

$$\tilde{\mathbf{J}}_c = - \begin{bmatrix} \mathbf{F}_w \\ \tilde{\mathbf{\Gamma}}_c \end{bmatrix} = \begin{bmatrix} \mathbf{J}_c \\ -\mathbf{\Gamma}_c \end{bmatrix}, \quad \tilde{\mathbf{J}}_m = \begin{bmatrix} -\mathbf{D}\mathbf{F}_w \mathbf{c} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{D}\mathbf{J}_c \mathbf{c} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_m \\ \mathbf{0} \end{bmatrix}.\quad (\text{A.2})$$

The augmented formulation allows us to rewrite Eq. 10 as

$$\min_{\mathbf{m}, \mathbf{c}} \frac{1}{2} \left\| \tilde{\mathbf{d}}_w - \tilde{\mathbf{F}}_w(\mathbf{m}) \mathbf{c} \right\|_2^2 + \frac{\lambda}{2} \left\| \mathbf{\Gamma}_m \mathbf{m} \right\|_2^2.\quad (\text{A.3})$$

Since this is the same form as the VP formulation with $\lambda_c = 0$, all previous formulas apply, except for replacing the original derivatives with the derivatives associated with the augmented residual. Furthermore, with Eqs. A.1 and A.2, the estimated linear model, the Jacobian, and the gradient of the residual vector can be expressed explicitly in terms of the original quantities. The linear regression will be given by

$$\hat{\mathbf{c}}(\mathbf{m}) = -\tilde{\mathbf{J}}_c^\dagger \tilde{\mathbf{d}}_w = -(\mathbf{J}_c^H \mathbf{J}_c + \mathbf{\Lambda}_c)^{-1} \mathbf{J}_c^H \mathbf{d}_w = -\mathbf{J}_c^{-\Lambda_c} \mathbf{d}_w,\quad (\text{A.4})$$

where $\mathbf{\Lambda}_c = \lambda_c \mathbf{\Gamma}_c^H \mathbf{\Gamma}_c$ is the regularization matrix, and $\mathbf{J}_c^{-\Lambda_c} = (\mathbf{J}_c^H \mathbf{J}_c + \mathbf{\Lambda}_c)^{-1} \mathbf{J}_c^H$ is the regularized version of the pseudoinverse. If $\mathbf{J}_c^H \mathbf{J}_c$ is itself invertible, and $\mathbf{\Lambda}_c = \mathbf{0}$, the matrix $\mathbf{J}_c^{-\Lambda_c}$ defined in this manner would be exactly the Moore–Penrose pseudoinverse \mathbf{J}_c^\dagger . The complete Jacobian on the augmented residual vector is

$$\begin{aligned}\tilde{\mathbf{J}} &= \mathbf{D}\tilde{\mathbf{r}}_w(\mathbf{m}, \hat{\mathbf{c}}(\mathbf{m})) = \tilde{\mathbf{J}}_m - \tilde{\mathbf{J}}_c \tilde{\mathbf{J}}_c^\dagger \tilde{\mathbf{J}}_m - (\tilde{\mathbf{J}}_c^\dagger)^H (\mathbf{D}\tilde{\mathbf{J}}_c)^H \tilde{\mathbf{r}}_w \\ &= \begin{bmatrix} \mathbf{J}_m \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{J}_c \\ -\tilde{\mathbf{\Gamma}}_c \end{bmatrix} \mathbf{J}_c^{-\Lambda_c} \mathbf{J}_m - \begin{bmatrix} \mathbf{J}_c \\ -\tilde{\mathbf{\Gamma}}_c \end{bmatrix} (\mathbf{J}_c^H \mathbf{J}_c + \mathbf{\Lambda}_c)^{-1} (\mathbf{D}\mathbf{J}_c)^H (\mathbf{I} - \mathbf{J}_c \mathbf{J}_c^{-\Lambda_c}) \mathbf{d}_w \\ &= \begin{bmatrix} \mathbf{J}_m - \mathbf{J}_c \mathbf{J}_c^{-\Lambda_c} \mathbf{J}_m & -(\mathbf{J}_c^{-\Lambda_c})^H (\mathbf{D}\mathbf{J}_c)^H (\mathbf{I} - \mathbf{J}_c \mathbf{J}_c^{-\Lambda_c}) \mathbf{d}_w \\ + \tilde{\mathbf{\Gamma}}_c \mathbf{J}_c^{-\Lambda_c} \mathbf{J}_m & + \tilde{\mathbf{\Gamma}}_c (\mathbf{J}_c^H \mathbf{J}_c + \mathbf{\Lambda}_c)^{-1} (\mathbf{D}\mathbf{J}_c)^H (\mathbf{I} - \mathbf{J}_c \mathbf{J}_c^{-\Lambda_c}) \mathbf{d}_w \end{bmatrix}.\end{aligned}\quad (\text{A.5})$$

The equation above gives the Jacobian for the full-VP scheme. Different variants of VP can now be achieved by choosing different approximations for $\tilde{\mathbf{J}}$. For VP-RW2, only the first two terms on the second line of Eq. A.5 is kept, while for VP-RW3, only the first one is retained. The gradient of the augmented data misfit is once again independent of the variants of VP, and takes the form

$$\text{grad } \tilde{\phi}(\mathbf{m}) = \text{D}\tilde{\phi}(\mathbf{m}) = \text{Re} \left[\tilde{\mathbf{J}}_m^H \tilde{\mathbf{r}}_w \right] = \text{Re} \left[\mathbf{J}_m^H (\mathbf{I} - \mathbf{J}_c \mathbf{J}_c^{-\Lambda \epsilon}) \mathbf{d}_w \right], \quad (\text{A.6})$$

where $\tilde{\phi}(\mathbf{m}) = \frac{1}{2} \|\tilde{\mathbf{r}}_w(\mathbf{m}, \hat{\mathbf{c}}(\mathbf{m}))\|_2^2$. For the optimization problem (Eq. A.3), the update on the model parameter yielded by the Gauss–Newton algorithm satisfies

$$\left(\text{Re} \left[\tilde{\mathbf{J}}^H \tilde{\mathbf{J}} \right] + \lambda \Gamma^T \Gamma \right) \Delta \mathbf{m} = - \left(\text{D}\tilde{\phi} + \lambda \Gamma^T \Gamma \mathbf{m} \right). \quad (\text{A.7})$$

$$\begin{aligned} \left[\mathbf{J}_m^H (\mathbf{I} - \mathbf{J}_c \mathbf{J}_c^\dagger) \mathbf{J}_m + \mathbf{r}_w^H \text{D}\mathbf{J}_c (\mathbf{J}_c^H \mathbf{J}_c)^{-1} (\text{D}\mathbf{J}_c)^H \mathbf{r}_w + \lambda \Gamma^T \Gamma \right] \Delta \mathbf{m} &= -\mathbf{J}_m^H \mathbf{r}_w - \lambda \Gamma^T \Gamma \mathbf{m} \quad (\text{full - VP}), \\ \left[\mathbf{J}_m^H (\mathbf{I} - \mathbf{J}_c \mathbf{J}_c^\dagger) \mathbf{J}_m + \lambda \Gamma^T \Gamma \right] \Delta \mathbf{m} &= -\mathbf{J}_m^H \mathbf{r}_w - \lambda \Gamma^T \Gamma \mathbf{m} \quad (\text{VP - RW2}), \\ \left[\mathbf{J}_m^H \mathbf{J}_m + \lambda \Gamma^T \Gamma \right] \Delta \mathbf{m} &= -\mathbf{J}_m^H \mathbf{r}_w - \lambda \Gamma^T \Gamma \mathbf{m} \quad (\text{VP - RW3}). \end{aligned} \quad (\text{B.4})$$

Appendix B: Linking VP and joint model space optimization

We have alluded to the the role of VP and alternating approaches as surrogate methods to the joint model space inversion. Here, we shall more explicitly state how they are related. We consider optimization of the same problem (Eq. 10), but in joint model space. The model vector is then represented as a concatenated vector $\tilde{\mathbf{m}} = [\mathbf{m}^T, \mathbf{c}^T]^T$. The Jacobian and the Hessian of the misfit (under the same Gauss–Newton algorithm) in the full model space are given by

$$\mathbf{J}_{\tilde{\mathbf{m}}} = [\mathbf{J}_m, \mathbf{J}_c], \quad \mathbf{H}_{\tilde{\mathbf{m}}} \approx \mathbf{J}_{\tilde{\mathbf{m}}}^H \mathbf{J}_{\tilde{\mathbf{m}}} = \begin{bmatrix} \mathbf{J}_m^H \mathbf{J}_m & \mathbf{J}_m^H \mathbf{J}_c \\ \mathbf{J}_c^H \mathbf{J}_m & \mathbf{J}_c^H \mathbf{J}_c \end{bmatrix}, \quad (\text{B.1})$$

where \mathbf{J}_m and \mathbf{J}_c are defined in Eq. 13. Adding the regularization term yields the model update equation for the joint model space, given by the linear system

$$\begin{bmatrix} \mathbf{J}_m^H \mathbf{J}_m + \lambda \Gamma^T \Gamma & \mathbf{J}_m^H \mathbf{J}_c \\ \mathbf{J}_c^H \mathbf{J}_m & \mathbf{J}_c^H \mathbf{J}_c \end{bmatrix} \begin{bmatrix} \Delta \mathbf{m} \\ \Delta \mathbf{c} \end{bmatrix} = - \begin{bmatrix} \mathbf{J}_m^H \mathbf{r}_w + \lambda \Gamma^T \Gamma \mathbf{m} \\ \mathbf{J}_c^H \mathbf{r}_w \end{bmatrix}, \quad (\text{B.2})$$

where \mathbf{m} is the current estimate of the conductivity model. The linear system has a total dimension of $M_m + M_c$, where M_m and M_c are the dimension of the

models \mathbf{m} and \mathbf{c} , respectively. In geomagnetic deep sounding problems the inducing field parameterization usually occupies a much more higher dimensional subspace than the conductivity model, and the resulting system can be formidable to tackle. To compare the model update obtained using the joint model space optimization, we use Schur complement to extract the model update on \mathbf{m} , which is given by

$$\left[\mathbf{J}_m^H (\mathbf{I} - \mathbf{J}_c \mathbf{J}_c^\dagger) \mathbf{J}_m + \lambda \Gamma^T \Gamma \right] \Delta \mathbf{m} = -\mathbf{J}_m^H (\mathbf{I} - \mathbf{J}_c \mathbf{J}_c^\dagger) \mathbf{r}_w - \lambda \Gamma^T \Gamma \mathbf{m}. \quad (\text{B.3})$$

Let us now consider the model update induced by VP and its variants. Following Eq. 19, the gradient retains its form regardless of the adopted approximation. Using the Gauss–Newton algorithm with Jacobians given by Eqs. 18, 21 and 22, the model updates are expressed as

We can see that the system induced by the joint model space optimization is very similar to the model update within the VP-RW2 method. In particular, expressing the current model in the joint model space as $\tilde{\mathbf{m}} = [\mathbf{m}, \mathbf{c}] = [\mathbf{m}, -\mathbf{J}_c^\dagger \mathbf{d}_w]$, we can apply the orthogonal property of the residual vector with respect to \mathbf{J}_c

$$\mathbf{J}_c^H \mathbf{r}_w = \mathbf{0} \implies (\mathbf{I} - \mathbf{J}_c \mathbf{J}_c^\dagger) \mathbf{r}_w = \mathbf{r}_w. \quad (\text{B.5})$$

The linear system for joint model space optimization (Eq. B.3) is then effectively identical to VP-RW2 (Eq. B.4). We therefore conclude that given the same conductivity model \mathbf{m} and an optimized source model $\mathbf{c} = -\mathbf{J}_c^\dagger(\mathbf{m}) \mathbf{d}_w$, the conductivity model update proposed by VP-RW2 is exactly the same as that proposed by the joint model space inversion. However, we note that the proposed linear update in joint model space inversion is given by

$$\left[\mathbf{J}_c^H \mathbf{J}_c \right] \Delta \mathbf{c} = -\mathbf{J}_c^H (\mathbf{r}_w + \mathbf{J}_m \Delta \mathbf{m}), \quad (\text{B.6})$$

and does not yield optimized update on \mathbf{c} . In contrast, the VP methods conduct regression at every iteration, which guarantees that the choice of \mathbf{c} is optimal (in least-squares sense). Therefore, starting from the same \mathbf{m} , $\mathbf{c} = -\mathbf{J}_c^\dagger \mathbf{d}_w$ combinations, VP-RW2 is guaranteed to propose a linear model that yields smaller data misfit compared to joint model space inversion, without resorting to a more complex nonlinear model.

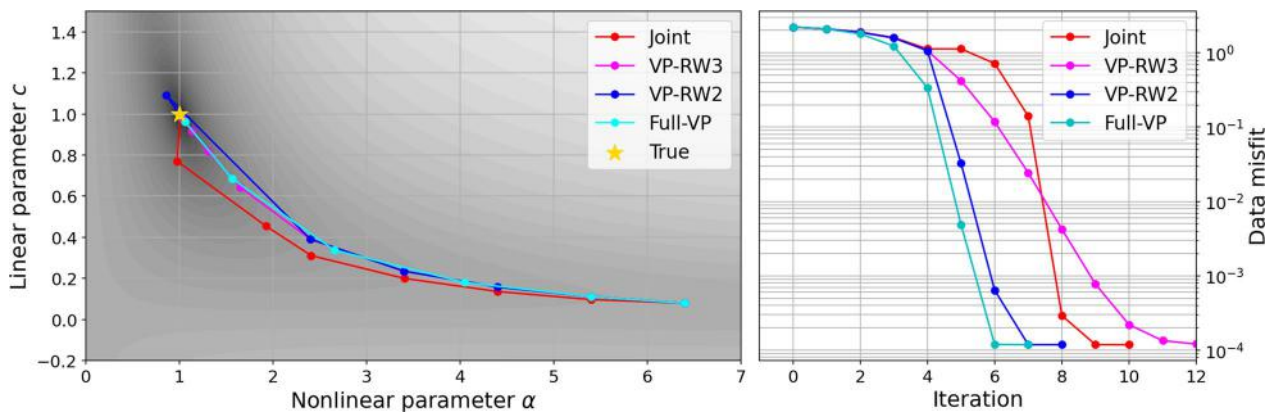


Fig. 21 Convergence of different inversion techniques on the wavelet fitting problem. Left: trajectories of inversion schemes in the parameter space for the wavelet fitting problem. The background color shows the data misfit in logarithmic scale. Right: data misfit as a function of iteration

As a simple illustration, we consider the wavelet fitting problem of a Ricker wavelet, which takes the form

$$\phi(t) = -\frac{d^2}{dt^2} (ce^{-\alpha t^2}) = 2c(\alpha - 2\alpha^2 t^2)e^{-\alpha t^2}, \quad (B.7)$$

where c controls the amplitude and enters the observation linearly, and α controls the width of the wavelet and enters the data nonlinearly. The wavelet fitting problem with squared misfit is an SNLS problem, and can be solved using the aforementioned methods. Different variants of VP were conducted, together with a joint model space inversion. The synthetic data were generated with $\alpha_* = 1$ and $c_* = 1$, and all inversion schemes started from an initial guess $\alpha_0 = 6$. For VP, no initial guess on the linear model is needed; for the joint model space inversion, we took the linear regression result \hat{c}_0 at $\alpha_0 = 6$.

The convergence of different inversion schemes in the 2-D model space is shown in Fig. 21. In this toy example, all inversion schemes converged to the ground truth, but we observe the convergence slows down as we go from full-VP, VP-RW2 to VP-RW3. Comparison between the

trajectories of VP-RW2 (blue line) and the joint model space inversion (red line) shows that within the first four iterations, the update on the nonlinear parameter α is very close, while the linear model c in the joint model space inversion slowly drifts away from the optimal linear least-squares solution, a phenomenon predicted by our derivations above. The non-optimal linear model for the joint model space inversion results in a detour of the trajectory after the fifth iteration, undermining the convergence of the inversion scheme.

Appendix C: Variance propagation for windowed Fourier transform

Here, we present a derivation for Eq. 31. To this end, we consider a time series $x(t_n)$, which is a sum of a deterministic signal $x^{(0)}(t_n)$, and a white noise $\epsilon(t_n) \sim \mathcal{N}(0, s)$. Thus, we have

$$\mathbb{E}[x(t_n)] = x^{(0)}(t_n), \quad s^2 \equiv \text{Var}[x(t_n)] = \text{Var}[\epsilon(t_n)]. \quad (C.1)$$

We shall use X , $X^{(0)}$ and \mathcal{E} to denote the windowed spectrum of x , $x^{(0)}$ and ϵ , respectively. Following the windowed Fourier transform (Eq. 30), we write the expectation of the windowed spectrum:

$$\begin{aligned} \mathbb{E}[X(\tau, \omega)] &= \mathbb{E}\left[\frac{1}{\sum_n w_n} \sum_{n=1}^{N_\tau} w_n x(t_n) e^{-i\omega t_n}\right] = \frac{1}{\sum_n w_n} \sum_{n=1}^{N_\tau} w_n \mathbb{E}[x(t_n)] e^{-i\omega t_n} \\ &= \frac{1}{\sum_n w_n} \sum_{n=1}^{N_\tau} w_n x^{(0)}(t_n) e^{-i\omega t_n} = X^{(0)}(\tau, \omega). \end{aligned} \quad (C.2)$$

Therefore, the windowed spectrum X is the unbiased estimate of $X^{(0)}$. On the other hand, the variance of the windowed spectrum X is determined by the variance of the windowed spectrum of the noise

$$\text{Var}[X(\tau, \omega)] = \mathbb{E}[(X(\tau, \omega) - \mathbb{E}[X(\tau, \omega)])^2] = \mathbb{E}[|\mathcal{E}(\tau, \omega)|^2]. \tag{C.3}$$

The expectation of $|\mathcal{E}|^2$ is in turn given by

$$\begin{aligned} \mathbb{E}[|\mathcal{E}(\tau, \omega)|^2] &= \mathbb{E}[\mathcal{E}^*(\tau, \omega)\mathcal{E}(\tau, \omega)] = \mathbb{E}\left[\frac{1}{(\sum_n w_n)^2} \sum_{n,m=1}^{N_\tau} w_n w_m \epsilon^*(t_n)\epsilon(t_m)e^{i\omega(t_n-t_m)}\right] \\ &= \frac{1}{(\sum_n w_n)^2} \sum_{n,m=1}^{N_\tau} w_n w_m \mathbb{E}[\epsilon^*(t_n)\epsilon(t_m)]e^{i\omega(t_n-t_m)} \\ &= \frac{1}{(\sum_n w_n)^2} \sum_{n,m=1}^{N_\tau} w_n w_m \text{Var}[\epsilon(t_n)]\delta_{nm}e^{i\omega(t_n-t_m)} = \frac{\sum_n w_n^2}{(\sum_n w_n)^2} \text{Var}[\epsilon(t)]. \end{aligned} \tag{C.4}$$

The last two steps use the fact that $\epsilon(t_n)$ and $\epsilon(t_m)$ are independent random variables with zero mean. This completes the proof for Eq. 31. For a boxcar window function, i.e. $w_n \equiv 1$, we have the following:

$$\text{Var}[X(\tau, \omega)] = \frac{\text{Var}[x(t)]}{N_\tau}. \tag{C.5}$$

Appendix D: Modelling in windowed Fourier domain

In our synthetic test, we observe that although $\chi_{\text{rms}} \approx 1$ can be obtained for the entire dataset, it is not the case with every frequency band (Fig. 3). This phenomenon should be attributed to the spectral leakage and, as a consequence, the inevitable imperfect nature of windowed-Fourier-domain modelling. In particular, we consider two time series $y(t)$ and $x(t)$, which are related in the frequency domain via

$$Y(\omega) = H(\omega)X(\omega), \tag{D.1}$$

where X, Y are the spectra of x and y , respectively, and $H(\omega)$ is the transfer function. In the general formulation of VP/alternating methods, X and Y correspond to the inducing current parameterization \mathbf{c} and the data vector

\mathbf{d} , and $H(\omega)$ corresponds to the forward operator $\mathbf{F}(\sigma)$ (Eq. 6); in the formulation of Q -response estimation, X and Y are ϵ_n^m and l_n^m , respectively, while $H(\omega)$ is nothing but the $Q_n(\omega)$ response (Eq. 25). Without loss of generality, we limit ourselves to scalars $X, Y, H \in \mathbb{C}$ in this appendix. The goal here is to show that the windowed spectra of x and y , given by

$$X(\tau, \omega) = \mathcal{F}_{\tau, \omega}[x(t)], \quad Y(\tau, \omega) = \mathcal{F}_{\tau, \omega}[y(t)] \tag{D.2}$$

with the transforms defined in Eq. 30 do not strictly satisfy the same relation in the frequency domain. In other words, in general, we have

$$Y(\tau, \omega) \neq H(\omega)X(\tau, \omega). \tag{D.3}$$

In this appendix, we strictly distinguish between the term windowed Fourier domain and the Fourier domain. The former is defined in Eq. 30, and the latter is defined in its continuous form in Eq. 2, and in its discrete form as the following convention for the discrete Fourier transform (DFT) and its inverse (iDFT):

$$\begin{aligned} X(\omega_q) &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x(t_k) e^{-i\omega_q t_k}, \\ x(t_k) &= \frac{1}{\sqrt{N}} \sum_{q=0}^{N-1} X(\omega_q) e^{i\omega_q t_k}, \end{aligned} \tag{D.4}$$

where $t_k = k\Delta t$ are the sampling time points, and $\omega_q = q/N\Delta t$ are the frequency points, $k, q = 0, \dots, N - 1$. The windowed spectral transform of x in time window τ at frequency ω can be written as

$$\begin{aligned}
 X(\tau, \omega) &= \frac{1}{\sum_k w_k} \sum_{k \in \{k_\tau\}} w_{k-k_{\tau 0}} x(t_k) e^{-i\omega(t_k-t_{k_{\tau 0}})} \\
 &= \frac{1}{\sum_k w_k} \sum_{k \in \{k_\tau\}} w_{k-k_{\tau 0}} \left[\frac{1}{\sqrt{N}} \sum_{q=0}^{N-1} X(\omega_q) e^{i\omega_q t_k} \right] e^{-i\omega(t_k-t_{k_{\tau 0}})} \\
 &= \frac{1}{\sqrt{N}} \sum_{q=0}^{N-1} X(\omega_q) \left[\frac{1}{\sum_k w_k} \sum_{k \in \{k_\tau\}} w_{k-k_{\tau 0}} e^{i(\omega_q-\omega)(t_k-t_{k_{\tau 0}})} \right] e^{i\omega_q t_{k_{\tau 0}}} \\
 &= \frac{1}{\sqrt{N}} \sum_{q=0}^{N-1} X(\omega_q) \left[\frac{1}{\sum_k w_k} \sum_{p=0}^{K_\tau-1} w_p e^{i(\omega_q-\omega)t_p} \right] e^{i\omega_q t_{k_{\tau 0}}}.
 \end{aligned} \tag{D.5}$$

Here, $k_{\tau 0}$ denotes the first time index of the time window τ , and K_τ denotes the total number of time points in the time window τ . Defining the normalized spectrum of the window function within the time window as

$$\widetilde{W}(\omega) = \frac{1}{\sum_{p=0}^{K_\tau-1} w_p} \sum_{p=0}^{K_\tau-1} w_p e^{-i\omega t_p}, \tag{D.6}$$

the windowed spectrum of x can be reiterated as

$$X(\tau, \omega) = \frac{1}{\sqrt{N}} \sum_{q=0}^{N-1} X(\omega_q) \widetilde{W}(\omega - \omega_q) e^{i\omega_q t_{k_{\tau 0}}}. \tag{D.7}$$

The trailing factor $e^{i\omega_q t_{k_{\tau 0}}}$ shifts the phases at respective frequencies to the beginning of the time window. The spectrum \widetilde{W} is normalized as such, so that $\widetilde{W}(0) \equiv 1$. In the extreme case of infinite length time series and time windows, we will have $\widetilde{W}(\omega) = \delta(\omega)$, leading to $X(\tau, \omega) \propto X(\omega)$. This is, however, never the case for finite length time series and time windows, where the windowed spectrum at frequency ω always contains the spectrum at adjacent DFT frequencies, i.e., $X(\omega_q)$, a phenomenon known as spectral leakage. Appropriate choices of the window function yield \widetilde{W} that can suppress the leakage, but this phenomenon still exists. The "imperfection" of the forward modelling in windowed Fourier domain becomes clear when we also write the windowed spectrum of Y in a similar form

$$\begin{aligned}
 Y(\tau, \omega) &= \frac{1}{\sqrt{N}} \sum_{q=0}^{N-1} Y(\omega_q) \widetilde{W}(\omega - \omega_q) e^{i\omega_q t_{k_{\tau 0}}} \\
 &= \frac{1}{\sqrt{N}} \sum_{q=0}^{N-1} H(\omega_q) X(\omega_q) \widetilde{W}(\omega - \omega_q) e^{i\omega_q t_{k_{\tau 0}}}.
 \end{aligned} \tag{D.8}$$

From Eqs. D.7 and D.8, we see that the $X(\tau, \omega)$ is not simply linked to $Y(\tau, \omega)$ through a product with the transfer function, i.e., Eq. D.3. Only under one specific condition, i.e., $H(\omega) \equiv H_0$, do the two quantities follow the same relation as their Fourier-domain counterparts, as

$$\begin{aligned}
 Y(\tau, \omega) &= \frac{1}{\sqrt{N}} \sum_{q=0}^{N-1} H_0 X(\omega_q) \widetilde{W}(\omega - \omega_q) e^{i\omega_q t_{k_{\tau 0}}} \\
 &= H_0 X(\tau, \omega).
 \end{aligned} \tag{D.9}$$

In other words, only when the transfer function has a flat spectrum (the impulse response is impulsive) is the modelling in windowed Fourier domain exactly the same as the modelling in the Fourier domain. Otherwise, the forward modelling as $Y(\tau, \omega) = H(\omega) X(\tau, \omega)$ cannot explain the scattering of the data $Y(\tau, \omega)$. This phenomenon should be perceived as an imperfection that affects both the TF estimation, as well as the VP/alternating approaches when using the specific form of forward modelling (Eq. 29). In Q -response estimation, this indicates that even for perfect synthetic data, there will be residuals in the fitting of $l(\tau, \omega)$ that cannot be explained by Eq. 25. In our implementation of VP/alternating approaches combined with the forward operators (Eq. 29), this imperfection is included in the uncertainty by the spectral transform error in Eq. 32.

Appendix E: Additional figures

See Figs. 22, 23, 24.

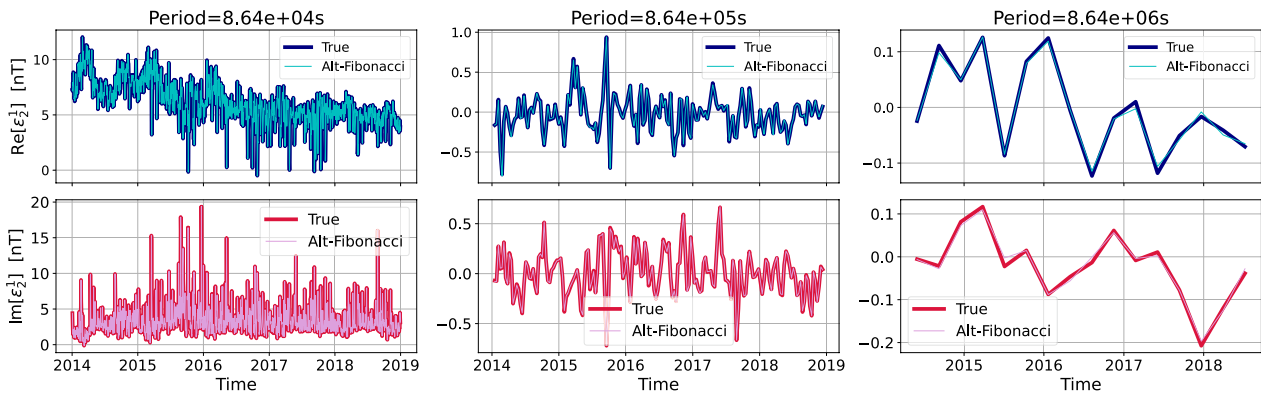


Fig. 22 Recovered windowed spectrum of the inducing field coefficient ϵ_2^1 using Alt-Fibonacci. The frequency bands and the legends are the same as in Fig. 5

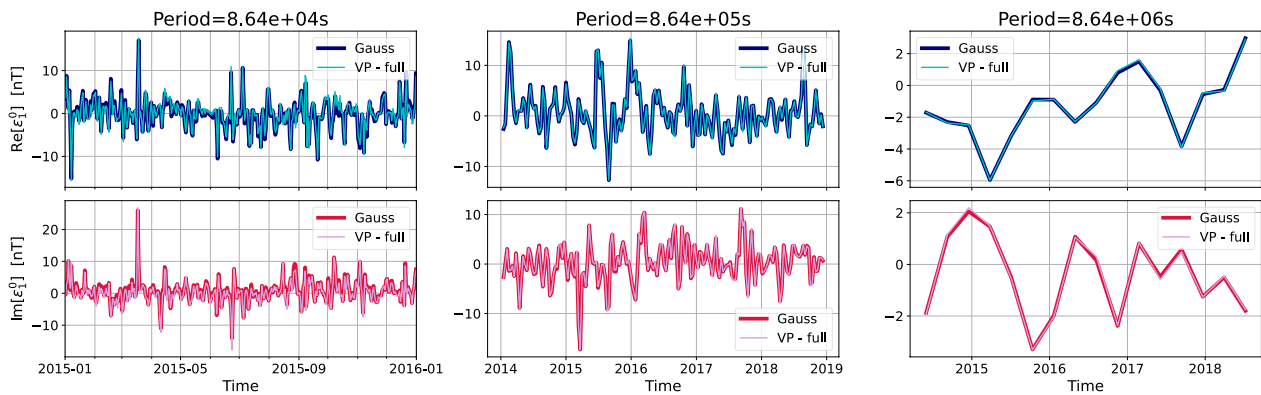


Fig. 23 Windowed spectrum of the inducing field coefficient ϵ_1^0 estimated from the real dataset estimated with VP and Gauss methods. The frequency bands and the legends are the same as Fig. 16

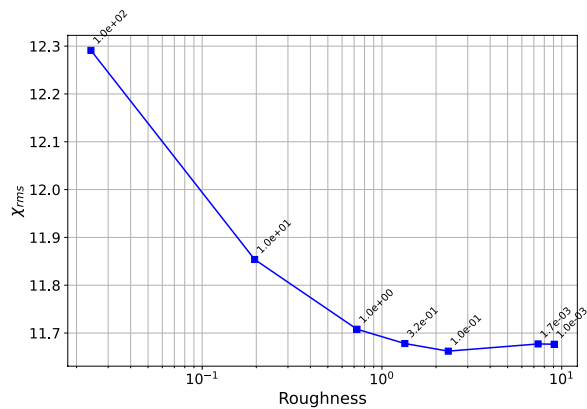


Fig. 24 L-curve of model roughness $\|\Gamma m\|_2^2$ and RMS misfit. All inversions eventually converged within around 20 iterations, but we notice that some inversions converged to sub-optimal solutions, probably local optima (regularization strengths 1.0×10^{-3} and 1.7×10^{-3}). Further, unlike the synthetic tests, the inversions of the real data converged to models with RMS misfits $\chi_{rms} \approx 11.7 \gg 1$. We however notice that the absolute reduction of RMS misfit ($\Delta \chi_{rms} \approx -2$) is comparable to that of synthetic tests. A regularization strength of 0.32 was chosen to produce the preferred model, which is located at the kink of the L-curve

Abbreviations

EM	Electromagnetic
GDS	Geomagnetic depth sounding
MT	Magnetotellurics
RMS	Root-mean-square
SH	Spherical harmonics
SHA	Spherical harmonic analysis
SNLS	Separable nonlinear least squares
VP	Variable projection method
VP-RW2	Second algorithm of variable projection, proposed by Ruhe and Wedin (1980)
VP-RW3	Third algorithm of variable projection, proposed by Ruhe and Wedin (1980)

Acknowledgements

AG is thankful to Malcolm Sambridge for fruitful discussions about SNLS problems. Constructive reviews by two anonymous reviewers helped improve the original draft substantially.

Author contributions

JM: methodology, software, data analysis, and writing—original draft; AG: conceptualization, methodology, data analysis, and writing—review and editing. Both authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. AG was supported by the ESA Swarm DISC (Contract No. 4000109587) and the Heisenberg Grant from the German Research Foundation, Deutsche

Forschungsgemeinschaft (Project No. 465486300). JM is grateful for funding from the European Research Council (Agreement No. 833848-UEMHP) under the Horizon 2020 programme.

Data availability

Time series of the hourly means at ground magnetic observatories were taken from the AUX_OBS ESA Swarm product <https://earth.esa.int/eogateway/missions/swarm/product-data-handbook/auxiliary-product-definitions>.

Declarations

Competing interests

The authors declare no competing interests.

Author details

¹Institute of Geophysics, ETH Zurich, Sonneggstrasse 5, Zurich 8092, Switzerland. ²Institute of Geophysics and Meteorology, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany.

Received: 17 October 2022 Accepted: 9 April 2023

Published online: 16 May 2023

References

- Aravkin AY, van Leeuwen T (2012) Estimating nuisance parameters in inverse problems. *Inverse Problems* 28(11):115016. <https://doi.org/10.1088/0266-5611/28/11/115016>
- Backus G, Parker RL, Constable C (1996) *Foundations of geomagnetism*. Cambridge University Press, Cambridge
- Balasis G, Egbert GD (2006) Empirical orthogonal function analysis of magnetic observatory data: further evidence for non-axisymmetric magnetospheric sources for satellite induction studies. *Geophys Res Lett* 33:2006GL025721. <https://doi.org/10.1029/2006GL025721>
- Chave AD, Jones AG (2012) *The magnetotelluric method: theory and practice*. Cambridge University Press, Cambridge
- De Ridder SA, Maddison JR (2018) Full wavefield inversion of ambient seismic noise. *Geophys J Int* 215:1215–1230. <https://doi.org/10.1093/GJI/GGY328>
- Egbert GD (2012) Hybrid conjugate gradient-Occam algorithms for inversion of multifrequency and multitransmitter EM data. *Geophys J Int* 190:255–266. <https://doi.org/10.1111/j.1365-246X.2012.05523.x>
- Egbert GD, Kelbert A (2012) Computational recipes for electromagnetic inverse problems. *Geophys J Int* 189(1):251–267. <https://doi.org/10.1111/j.1365-246X.2011.05347.x>
- Egbert GD, Alken P, Maute A, Zhang H (2021) Modelling diurnal variation magnetic fields due to ionospheric currents. *Geophys J Int* 225(2):1086–1109. <https://doi.org/10.1093/gji/ggaa533>
- Fainberg EB, Kuvshinov AV, Mishina LP, Singer BS (1990) The new approach to global deep sounding. *Pure Appl Geophys* 134:527–531. <https://doi.org/10.1007/BF00878016>
- Finlay CC, Lesur V, Thébaud E, Vervelidou F, Morschhauser A, Shore R (2017) Challenges handling magnetospheric and ionospheric signals in internal geomagnetic field modelling. *Space Sci Rev* 206(1):157–189. <https://doi.org/10.1007/s11214-016-0285-9>
- Golub G, Pereyra V (2003) Separable nonlinear least squares: the variable projection method and its applications. *Inverse Problems* 19(2):R1. <https://doi.org/10.1088/0266-5611/19/2/201>
- Golub GH, Pereyra V (1973) The differentiation of pseudo-inverses and nonlinear least square problems whose variables separate. *SIAM J Numer Anal* 10:413–432. <https://doi.org/10.1137/0710036>
- Grayver AV, Munch FD, Kuvshinov AV, Khan A, Sabaka TJ, Tøffner-Clausen L (2017) Joint inversion of satellite-detected tidal and magnetospheric signals constrains electrical conductivity and water content of the upper mantle and transition zone. *Geophys Res Lett* 44:6074–6081. <https://doi.org/10.1002/2017GL073446>
- Grayver AV, Kuvshinov A, Werthmüller D (2021) Time-domain modeling of three-dimensional earth's and planetary electromagnetic induction effect in ground and satellite observations. *J Geophys Res Space Phys* 126:e2020JA028672. <https://doi.org/10.1029/2020JA028672>
- Guzavina M, Grayver A, Kuvshinov A (2019) Probing upper mantle electrical conductivity with daily magnetic variations using global-to-local transfer functions. *Geophys J Int* 219:2125–2147. <https://doi.org/10.1093/gji/ggz412>
- Hansen PC, O'Leary DP (1993) The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J Sci Comput* 14:1487–1503. <https://doi.org/10.1137/0914086>
- Heath MT (2018) *Scientific Computing*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 10(1137/1):9781611975581
- Hong JH, Zach C, Fitzgibbon A (2017) Revisiting the variable projection method for separable nonlinear least squares problems. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 5939–5947. <https://doi.org/10.1109/CVPR.2017.629>
- Hu X, Wei G, Song J, Yang Z, Lu M, Gao L (2021) Full-waveform inversion with source and receiver coupling effects correction. *Geophysics* 86(4):R463–R470. <https://doi.org/10.1190/geo2020-0331.1>
- Juusola L, Vanhamäki H, Viljanen A, Smirnov M (2020) Induced currents due to 3D ground conductivity play a major role in the interpretation of geomagnetic variations. *Annales Geophysicae* 38:983–998. <https://doi.org/10.5194/angeo-38-983-2020>
- Kelbert A (2020) The role of global/regional earth conductivity models in natural geomagnetic hazard mitigation. *Surveys Geophys* 41:115–166. <https://doi.org/10.1007/s10712-019-09579-z>
- Kelbert A, Lucas GM (2020) Modified GIC estimation using 3-D earth conductivity. *Space Weather* 18(8):e2020SW002467. <https://doi.org/10.1029/2020SW002467>
- Kelbert A, Schultz A, Egbert G (2009) Global electromagnetic induction constraints on transition-zone water content variations. *Nature* 460:1003–1006. <https://doi.org/10.1038/nature08257>
- Koch S, Kuvshinov A (2013) Global 3-D EM inversion of Sq variations based on simultaneous source and conductivity determination: concept validation and resolution studies. *Geophys J Int* 195(1):98–116. <https://doi.org/10.1093/gji/ggt227>
- Kuvshinov A (2008) 3-D global induction in the oceans and solid earth: recent progress in modeling magnetic and electric fields from sources of magnetospheric, ionospheric and oceanic origin. *Surveys Geophys* 29:139–186. <https://doi.org/10.1007/s10712-008-9045-z>
- Kuvshinov A (2012) Deep electromagnetic studies from land, sea, and space: progress status in the past 10 years. *Surveys Geophys* 33:169–209. <https://doi.org/10.1007/s10712-011-9118-2>
- Kuvshinov A, Grayver A, Tøffner-Clausen L, Olsen N (2021) Probing 3-D electrical conductivity of the mantle using 6 years of Swarm, CryoSat-2 and observatory magnetic data and exploiting matrix Q-responses approach. *Earth Planets Space* 73:67. <https://doi.org/10.1186/s40623-020-01341-9>
- Laundal KM, Yee JH, Merkin VG, Gjerloev JW, Vanhamäki H, Reistad JP, Madelaire M, Sorathia K, Espy PJ (2021) Electrojet estimates from mesospheric magnetic field measurements. *J Geophys Res Space Phys* 126:e2020JA028644. <https://doi.org/10.1029/2020JA028644>
- Li M, Rickett J, Abubakar A (2013) Application of the variable projection scheme for frequency-domain full-waveform inversion. *Geophysics* 78:R249–R257. <https://doi.org/10.1190/geo2012-0351.1>
- Maus S, Weidelt P (2004) Separating the magnetospheric disturbance magnetic field into external and transient internal contributions using a 1D conductivity model of the earth. *Geophys Res Lett*. <https://doi.org/10.1029/2004GL020232>
- O'Leary DP, Rust BW (2013) Variable projection for nonlinear least squares problems. *Comput Optim Appl* 54:579–593. <https://doi.org/10.1007/s10589-012-9492-9>
- Olsen N (1999) Induction studies with satellite data. *Surveys Geophys* 20:309–340. <https://doi.org/10.1023/a:1006611303582>
- Olsen N, Friis-Christensen E, Foberghagen R, Alken P, Beggan CD, Chulliat A, Doornbos E, da Encarnação JT, Hamilton B, Hulot G, van den IJssel J, Kuvshinov A, Lesur V, Lühr H, Macmillan S, Maus S, Noja M, Olsen PEH, Park J, Plank G, Püthe C, Rauberg J, Ritter P, Rother M, Sabaka TJ, Schachtschneider R, Sirol O, Stolle C, Thébaud E, Thomson AWP, Tøffner-Clausen L, Velimský J, Vigneron P, Visser PN (2013) The Swarm satellite constellation application and research facility (SCARF) and Swarm data products. *Earth Planets Space* 65:1189–1200. <https://doi.org/10.5047/eps.2013.07.001>
- Pankratov OV, Kuvshinov AV (2010) Fast calculation of the sensitivity matrix for responses to the earth's conductivity: general strategy and examples,

- Izvestiya. Phys Solid Earth 46:788–804. <https://doi.org/10.1134/S1069351310090089>
- Pulkkinen A, Amm O, Viljanen A (2003) Ionospheric equivalent current distributions determined with the method of spherical elementary current systems. *J Geophys Res Space Phys* <https://doi.org/10.1029/2001JA005085>
- Püthe C, Kuvshinov A (2014) Mapping 3-D mantle electrical conductivity from space: a new 3-D inversion scheme based on analysis of matrix Q-responses. *Geophys J Int* 197(2):768–784. <https://doi.org/10.1093/gji/ggu027>
- Rickett J (2013) The variable projection method for waveform inversion with an unknown source function. *Geophys Prospect* 61:874–881. <https://doi.org/10.1111/1365-2478.12008>
- Ruhe A, Wedin PÅ (1980) Algorithms for separable nonlinear least squares problems. *SIAM Rev* 22:318–337. <https://doi.org/10.1137/1022057>
- Schmucker U (1999) A spherical harmonic analysis of solar daily variations in the years 1964–1965: Response estimates and source fields for global induction-i. methods. *Geophys J Int* 136:439–454. <https://doi.org/10.1046/j.1365-246X.1999.00742.x>
- Sun J, Kelbert A, Egbert GD (2015) Ionospheric current source modeling and global geomagnetic induction using ground geomagnetic observatory data. *J Geophys Res Solid Earth* 120(10):6771–6796. <https://doi.org/10.1002/2015JB012063>
- Tsyganenko NA (2019) Secular drift of the auroral ovals: How fast do they actually move? *Geophys Res Lett* 46:3017–3023. <https://doi.org/10.1029/2019GL082159>
- Velínský J, Knopp O (2021) Lateral variations of electrical conductivity in the lower mantle constrained by Swarm and CryoSat-2 missions. *Earth Planets Space* 73:4. <https://doi.org/10.1186/s40623-020-01334-8>
- Velínský J, Grayver A, Kuvshinov A, Šachl L (2018) On the modelling of M2 tidal magnetic signatures: effects of physical approximations and numerical resolution. *Earth Planets Space* 70:192. <https://doi.org/10.1186/s40623-018-0967-5>
- Weidelt P (1972) The inverse problem of geomagnetic induction. *J Geophys* 38:257–289
- Yamazaki Y, Maute A (2017) Sq and EEJ—a review on the daily variation of the geomagnetic field caused by ionospheric dynamo currents. *Space Sci Rev* 206:299–405. <https://doi.org/10.1007/s11214-016-0282-z>
- Zenhäusern G, Kuvshinov A, Guzavina M, Maute A (2021) Towards probing earth's upper mantle with daily magnetic field variations: exploring a physics-based parametrization of the source. *Earth Planets Space* 73:136. <https://doi.org/10.1186/s40623-021-01455-8>
- Zhang H, Egbert GD, Huang Q (2022) A relatively dry mantle transition zone revealed by geomagnetic diurnal variations. *Sci Adv* <https://doi.org/10.1126/sciadv.abo3293>

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
