Structural, petrophysical and geological constraints in potential field inversion using the Tomofast-x v1.0 open-source code

Jérémie Giraud[1,2], Vitaliy Ogarko[1,2,3], Roland Martin[4], Mark Jessell[1,2], Mark Lindsay[1,2]

5

[1] Centre for Exploration Targeting (School of Earth Sciences), University of Western Australia, 35 Stirling Highway, 6009 Crawley Australia.

[2] Mineral Exploration Cooperative Research Centre, School of Earth Sciences, University of Western Australia, 35 Stirling Highway, WA Crawley 6009 Australia.

10 [3] The International Centre for Radio Astronomy Research, University of Western Australia, 7 Fairway, WA Crawley 6009 Australia.

[4] Laboratoire de Géosciences Environnement Toulouse GET, CNRS UMR 5563, Observatoire Midi-Pyrénées, Université Paul Sabatier, 14 avenue Edouard Belin, 31400 Toulouse France.

15

Abstract. The quantitative integration of geophysical measurements with data and information from other disciplines is becoming increasingly important in answering the challenges of undercover imaging and of the modelling of complex areas. We propose a review of the different techniques for the utilisation of structural, petrophysical and geological information in single physics and joint inversion as implemented in the Tomofast-x open-source inversion platform. We detail the range of constraints that can be applied to the inversion of potential field data. The inversion examples we show illustrate a selection of scenarios using a realistic synthetic dataset inspired by real-world geological measurements and petrophysical data from the Hamersley region (Western Australia). Using Tomofast-x's flexibility, we investigate inversions combining the utilisation of petrophysical, structural and/or geological constraints while illustrating the utilisation of the L-curve principle to determine regularisation weights. Our results suggest that the utilisation of geological information to derive disjoint interval bound constraints is the most effective method to recover the true model. It is followed by model smoothness and smallness conditioned by geological uncertainty, and cross-gradient minimisation.

1 Introduction

Geophysical data provide detailed information about the structure and composition of the Earth's interior otherwise not accessible by direct observation methods, and thus plays a central role in every major Earth imaging initiative. Applications of geophysical modelling range from deep Earth imaging to study the crust and the mantle to shallow investigations of the subsurface for the exploration of natural resources. Recent integration of different geophysical methods has been recognised as a means to reduce interpretation ambiguity and uncertainty. Further developments introduce uncertainty estimates from other geoscientific disciplines such as geology and petrophysics to produce more reliable and plausible models. Various techniques integrating different geophysical techniques have been developed with the aim to produce more geologically meaningful models, as reviewed by Parsekian et al. (2015), Lelièvre and Farquharson (2016), Moorkamp et al. (2016), and Ren and Kalscheuer (2019), (Meju and Gallardo, 2016), and several kinds of optimization for such problems exist (Bijani et al., 2017). In the natural resource exploration sector, the calls of Wegener (1923), Eckhardt, (1940) and Nettleton (1949) for the development of comprehensive, thorough multi-disciplinary and multi-physical integrated modelling have been acknowledged by the scientific community, and data integration is now an area of active research: quoting André Revil's preface of the compilation of reviews proposed by (Moorkamp et al., 2016a): "The joint inversion of geophysical data with different sensitivities […] is also a new frontier". The integration of multiple physical fields (both geophysical and geological) is particularly relevant for techniques relying on potential field gravity and magnetic data, as these constitute the most commonly acquired and widely available geophysical datatypes worldwide. The needs for integrated techniques is partly due to the interpretation ambiguity of geophysical data and resulting effects of non-uniqueness on inversion. Therefore, effective inversion of potential field data necessitates the utilisation of constraints derived from prior information extracted from geological and petrophysical measurements or other geophysical techniques whenever available.

A number of methods for the introduction of geological and petrophysical prior information into potential-field inversion have been developed. For example, when limited geological information is available, the assumption is that spatial variation of density and magnetic susceptibility are collocated. This can be enforced through simple structural constraints encouraging structural correlation between the two models using Gramian constraints (Zhdanov et al., 2012) or the cross-gradient technique introduced in Gallardo and Meju, (2003). When petrophysical information is available, petrophysical constraints can be applied during inversion to obtain inverted properties that match certain statistics (see techniques introduced by Paasche and Tronicke (2007), De Stefano et al. (2011), Sun and Li (2016, 2011, 2015), Lelièvre et al. (2012), Carter-McAuslan et al. (2015), Zhang and Revil (2015), Giraud et al. (2016, 2017, 2019c), Heincke et al. (2017). Furthermore, when geological data are available, geological models can be derived and their statistics can be used to derive a candidate model for forward modelling (Guillen et al., 2008, Lindsay et al., 2013, de La Varga et al., 2019), to derive statistical petrophysical constraints for inversion (Giraud et al., 2017, 2019d, 2019c), and to restrict the range of accepted values using spatially varying disjoint bound constraints (Ogarko et al., 2021a) or multinary transformation (Zhdanov and Lin, 2017).

In this paper, we present a versatile inversion platform designed to integrate geological and petrophysical constraints to the inversion of gravity and magnetic data at different scales. We present Tomofast-x ('x' for 'extendable') as an open-source inversion platform capable of dealing with varying amounts and quality of input data. Tomofast-x is designed to conduct constrained single-physics and joint-physics inversion. The need for

reproducible research (Peng 2011) is facilitated by open-source codes (Gil et al., 2016), thus we introduce and detail the different constraints implemented in Tomofast-x before providing a realistic synthetic application example using selected functionalities. We illustrate the use of Tomofast-x by performing a realistic synthetic study investigating several modelling scenarios typically encountered by practitioners, and provide information to get free access to the source-code and to run it using the synthetic data shown in this paper. We perform single physics inversion of gravity data and study the influence of prior information using several amounts and types of constraints, and run joint inversion of gravity and magnetic data. The flexibility of Tomofast-x is exploited to test the effect of structural constraints combined with petrophysical and geological prior information that are yet to be demonstrated in the published literature. A challenging geological setting is used to examine the capability of cross-gradient constraints within the joint inversion method. The mathematical formulation of geophysical problems and solutions are detailed throughout the paper and sufficient information is provided to allow the reproducibility of this work using Tomofast-x.

The remainder of the contributions revolves around two main aspects. We first review the theory behind the inversion algorithm and the different techniques used, with an emphasis on the mathematical formulation of the problem. We then present a synthetic example inspired from a geological model in the Hamersley province (Western Australia), where we investigate two case scenarios. In the first case, we apply structural constraints to an area where geology contradicts the assumption of collocated and correlated density and magnetic susceptibility variations. In the second case, we investigate a novel combination of petrophysical and structural information to constrain single physics inversion. Finally, we place Tomofast-x in the general context of research in geophysical inverse modelling and conclude this article.

2 Inverse modelling platform Tomofast-x

2.1 Purpose of Tomofast-x

Tomofast-x can be used in a wide range of geoscientific scenarios as it can integrate multiple forms of prior information to constrain inversion and follow appropriate inversion strategies. Constraints can be applied through Tikhonov-style regularisation of the inverse problem (Tikhonov and Arsenin 1977, 1978). In single-physics inversion, these comprise model smallness (also called 'model damping', minimizing the norm of the model, see Hoerl and Kennard 1970) and model smoothness (also called 'gradient damping', minimizing the norm of the spatial gradient of the model, see Li and Oldenburg 1996). For more detailed imaging, petrophysical constraints using Gaussian mixture models (Giraud et al., 2019c) as well as structural constraints (Giraud et al., 2019d, Martin et al., 2020), multiple interval bound constraints (Ogarko et al., 2021a), can be used depending on the requirements of the study and the information available. In the case of single-physics inversion with structural constraints, structural similarity between a selected reference model and the inverted models can be maximised using structural constraints based on cross-gradients (Gallardo and Meju 2003), locally weighted gradients in the same philosophy as Brown et al. (2012), Wiik et al. (2015), Yan et al. (2017), Giraud et al. (2019d). Generally speaking, in the joint inversion case, the two models inverted for are linked using the structural constraints just mentioned or petrophysical clustering constraints in the same spirit as Carter-McAuslan et al. (2015), Kamm et al. (2015), Sun and Li (2015, 2017), Zhang and Revil (2015), Bijani et al. (2017). In addition to the underlying assumptions defining the relationship between properties jointly inverted for, prior information from previous modelling or
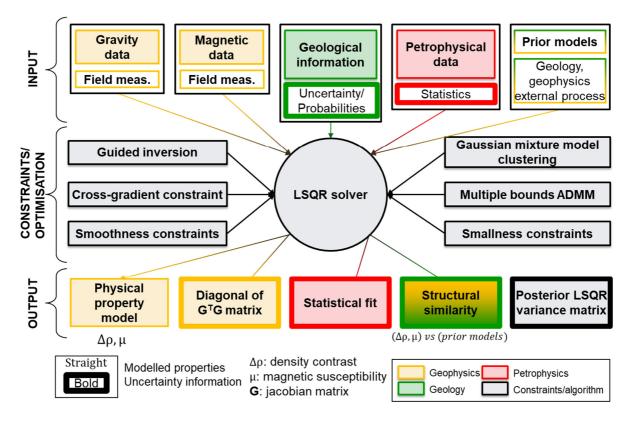
geological information can be incorporated in inversion using model and structural covariance matrices by assigning weights that vary spatially. In such case, Tomofast-x allows utilising prior information extensively. Furthermore, Tomofast-x allows the use of an arbitrary number of prior and starting models enabling the investigation of the subsurface in a detailed and stochastic-oriented fashion. Tomofast-x was initially developed for application to regional or crustal studies (areas covering hundreds of kilometres), and retains this capability. The current version of Tomofast-x is now more versatile as development is now directed toward use for exploration targeting and the monitoring of natural resources (kilometric scale).

Lastly, in addition to inversion, Tomofast-x offers the possibility to assess uncertainty in the recovered models. The uncertainty assessments include: statistical measures gathered from the petrophysical constraints; posterior least-squares variance matrix of the recovered model (in the Least Squares with QR-factorization algorithm – LSQR – sense of Paige and Saunders 1982, see 2.5), and the degree of structural similarity between the models (for joint inversion or structurally constrained inversion). From a practical point of view, associated with the inversion algorithm is a user manual covering most functionalities and a reduced 2D Python notebook illustrating concepts (see Sect. 7 for more information) that can be used for testing or educational purposes. A summary of the inverse modelling workflow of Tomofast-x is shown in Figure 1.

2.2 General design

The implementation we present extends the original inversion platform "Tomofast" (Martin et al., 2013, 2017). Tomofast-x is an extended implementation proposed and modified by Martin et al., 2018, Giraud et al. (2019d, 2019c), Martin et al. (2020), Ogarko et al. (2021a). Tomofast-x follows the object-oriented Fortran 2008 standard and utilizes classes designed to account for the mathematics of the problem. This introduces enhanced modularity based on the implementation of specific modules than can be called depending on the type of inversion required. The utilisation of classes in Tomofast-x also eases the addition of new functionalities and permits to reduce software complexity while maintaining flexibility. Our implementation uses an indexed hexahedral solid body mesh, giving the possibility to adapt the problem geometry, allowing to regularize the problem in the same fashion as Wiik et al. (2015) or to perform overburden stripping. By default, the sensitivity matrix to geophysical measurements is stored in a sparse format (using the Compressed Sparse Row format) to reduce memory consumption and for fast matrix-vector multiplications.

Attention has also been given to computational aspects. The only dependency of Tomofast-x is the Message Passing Interface (MPI) libraries, which eases installation and usage. This allows optimal usage of multi-CPU systems regardless of the number of CPUs. Parallelization is made on the model cells using a domain decomposition approach in space. That is, the model is divided into nearly equal, non-overlapping contiguous parts distributed among the CPUs, hence enforcing minimum load imbalance. Consequently, the code is fully scalable as the maximum number of CPUs is not limited by the number of receivers or measured data points. For large 3D models, Tomofast-x can run on hundreds of CPUs for a typical problem with 10^5-10^6 model cells and 10^3-10^4 data points. Parallel efficiency tests reveal excellent scalability and speed performance provided that the portions of the model sent to the CPUs are of sufficient size. In the current implementation, the optimum number of elements per CPU is 512. Interested readers can refer to 0.

Figure 1. Modelling workflow summary (modified from personal communication by J. Giraud).

2.3 Cost function

2.3.1 General formulation

Tomofast-x inversions rely on optimization of a least-squares cost function and optimized iteratively. The choice of a least-squares framework was motivated by flexibility in the number of constraints and forms of prior information used in the optimization process.

The objective function θ is derived from the log-likelihood of a probabilistic density function Θ (see Tarantola 2005, Chapters 1 and 3, for details). In the case of geophysical inversion, it is representative of the 'degree of knowledge that we have about the values of the parameters of our system' (Tarantola and Valette 1982), as summarized below. Let us first define Θ as follows:

$$\Theta(\boldsymbol{d}, \boldsymbol{m}) = \Theta_d(\boldsymbol{d}, \boldsymbol{m}) \prod_{i \, \in \, constraints} \Theta_i(\boldsymbol{m}), \tag{1}$$

where $\Theta_d(\boldsymbol{d}, \boldsymbol{m})$ is the density function over the geophysical data $\boldsymbol{d}$ that model $\boldsymbol{m}$ represents, and $\Theta_i(\boldsymbol{m})$ is the density function for the ith type of prior information available (the $constraints$ set).

On the premise that Gaussian probability densities approximate the problem appropriately, $\Theta_d(\boldsymbol{d}, \boldsymbol{m})$ can be expressed as:

$$\Theta_d(d, m) = \mathrm{A} \exp\left(-\left\|\boldsymbol{W_d}\big(\boldsymbol{d} - \boldsymbol{g}(\boldsymbol{m})\big)\right\|_2^2\right), \mathrm{A} \, \in \, \mathbb{R}^+ \backslash \{0\}, \tag{2}$$

where $\boldsymbol{g}(\boldsymbol{m})$ is the forward data set calculated by the forward operator $\boldsymbol{g}$; the matrix $\boldsymbol{W_d}$ weights the data points. Similarly, we formulate the different $\Theta_{i \, \in \, constraints}$ as:

$$\Theta_i(m) = C_i \exp\left(-\alpha_i \|\boldsymbol{W}_i \mathrm{f}(\boldsymbol{m})\|_2^2\right), C_i \in \mathbb{R}^+ \backslash \{0\}, \tag{3}$$

where f is a function of the model and prior information. $\boldsymbol{W}_i$ is a covariance matrix weighting of $\mathrm{f}(\boldsymbol{m})$ and α_i contains positive scalars that are introduced to adjust the relative importance of the ith constraint term. $\boldsymbol{W}_i$ and α_i are derived from prior information, or set according to study objectives.

From equation 3, it is clear that maximizing $\Theta(\boldsymbol{d}, \boldsymbol{m})$ is equivalent to minimizing its negative logarithm, $\theta(d, m)$, defined as:

$$\theta(d, m) = -\log \Theta(\boldsymbol{d}, \boldsymbol{m}) = \alpha_d \|\boldsymbol{W}_d(\boldsymbol{d} - \boldsymbol{g}(\boldsymbol{m}))\|_{\mathbf{L_2}}^2 + \sum_{i \,\in\, constraints} \alpha_i \|\boldsymbol{W}_i \mathrm{f}(\boldsymbol{m})\|_{\mathbf{L_2}}^2, \tag{4}$$

which corresponds to the general formulation of a cost function as formalized in the least-squares framework; α_d is a weight controlling the importance of the corresponding data term (i.e., gravity or magnetic) in the overall cost function.

2.3.2 Definition of regularization constraints

Adapting the formulation of the second term of equation (4) to the different types of prior information that we can accommodate leads to the following aggregate cost function:

$$\theta(d, m) = \|W_d(d - g(m))\|_2^2 + \alpha_m \|W_m(m - m_{pr})\|_p^2 + \alpha_g \|W_g \nabla m\|_2^2$$
$$+ \alpha_x \|W_x(\nabla m_1 \times \nabla m_2)\|_2^2 + \alpha_{pe} \|W_{pe} P(m)\|_2^2 + \alpha_a \|W_a(m - v + u)\|_2^2, \tag{5}$$

where the different terms following the data misfit term $\|W_d(d - g(m))\|_2^2$ constitute constraints for the inversion of geophysical data acting as regularization in the fashion of Tikhonov regularisation (Tikhonov and Arsenin 1977). In equations (2-5), $\boldsymbol{W}_d$ represents geophysical data weighting. Generally, $\boldsymbol{W}_d$ should be the data covariance. It is calculated by Tomofast-x as follows:

$$\boldsymbol{W}_d = \left(\sum_{i=1}^{n_{data}} (d_i)^2\right)^{-1} \boldsymbol{I}_w, \tag{6}$$

where $\boldsymbol{I}_w$ is a diagonal matrix equal to the identity matrix in single domain inversion, or giving the weight of one data misfit term (i.e, gravity data) against the other (i.e. magnetic data) in joint inversion; d_i is the ith datum. By convention, we fix $\boldsymbol{I}_w$ to the identity matrix $\boldsymbol{I}$ for gravity inversion, and use $\boldsymbol{I}_w = \boldsymbol{I}\alpha_{mag}$ in joint inversion. In such cases, α_{mag} is a strictly positive scalar.

The main terms of the cost function are defined below. The other individual terms are defined in the next subsection and summarized in Appendix B:

- $\boldsymbol{m}_{pr}$ refers to prior model,
- $\|\boldsymbol{W}_m(\boldsymbol{m} - \boldsymbol{m}_{pr})\|_p^2$ represents the smallness term (detailed in Sect. 2.4.1); subscript p refers to the L-p norm (here taken such that $1 < p \leq 2$) ;
- ∇ is the operator calculating the spatial gradient of the model;
- $\|\boldsymbol{W}_g \nabla \boldsymbol{m}\|_2^2$ represents the smoothness constraint on the model (detailed in Sect. 2.4.2);

- $\|W_x \nabla m^1 \times \nabla m^2\|_2^2$ represents cross-gradient constraints between the models m^1 and m^2 (detailed in Sect. 2.4.3).

- $\|W_{pe} P(m)\|_2^2$ represents petrophysics term (clustering constraint), into which $P(m)$ represents petrophysical distributions used to impose petrophysical constraints (detailed in Sect. 2.4.4);

- $\|W_a(m - v + u)\|_2^2$ is the formulation of the multiple-bounds constraint using the alternating direction of multipliers method (ADMM, detailed in Sect. 2.4.5);

In the case of joint inversion, the vectors defined above are concatenated and the matrices expanded as follows:

$$g(m) = \begin{bmatrix} g_G(m^G) \\ g_M(m^M) \end{bmatrix}, m = \begin{bmatrix} m^1 \\ m^2 \end{bmatrix} = \begin{bmatrix} m^G \\ m^M \end{bmatrix}, d = \begin{bmatrix} d_G \\ d_M \end{bmatrix},$$

$$\alpha_{i \in [d,m,g,x,pe,a]} = \begin{bmatrix} \alpha^G_{i \in [d,m,g,xgpe,a]} & 0 \\ 0 & \alpha^M_{i \in [d,m,g,x,pe,a]} \end{bmatrix},$$

$$W_{i \in [d,m,g,x,pe,a]} = \begin{bmatrix} W^G_{i \in [d,m,g,x,pe,a]} & 0 \\ 0 & W^M_{i \in [d,m,g,x,pe,a]} \end{bmatrix},$$

(7)

where T denotes the transpose operator; for more illustrative purposes of the joint inversion, we take here the gravity and magnetic joint inversion example; G and M refer to gravity and magnetic problems, respectively.

In the case of single domain inversion, m^1 is the model inverted for, and is equal to m^G or m^M depending on the type of geophysical data inverted, and m^1 is a reference model that can be used to constrain inversion from a structural point of view (see Sect. 2.4.2 and 2.4.3, and 4.4 and 4.5 for theory and example of utilisation, respectively).

In equation 5-7, subscript m, g, x, pe and a refer respectively to model, gradient, cross-gradient, petrophysics, and ADMM bound constraints, respectively. The different α terms are trade-off parameters that control the importance given to the different terms during the inversion. These terms therefore play an important role in inversion and need to be determined carefully (see Sect. 4.1 and 4.2 for more details).

As mentioned above, the cross-gradient constraints can be applied either to joint or single domain inversion. In the case of ADMM constraints, single or multiple bounds can be applied to define bounds for inverted model values. Such bound constraints can vary in space, and be made of an arbitrary number of intervals, be they disjoint or not (see Sect. 2.4.5 and 4.5). Qualitatively, the case with multiple disjoint intervals can be interpreted as applying a dynamic smallness constraint term.

In Tomofast-x, we introduce prior information in the diagonal variance matrices W_i such that they are no longer homogenous and can vary in space. Note that in the implementation of gravity and magnetic inversion, $g(m) = Sm$, with S the sensitivity matrix relating to measured geophysical data and corresponding recovered physical property (see Appendix C for details about their calculation). Introducing the sensitivity matrix S_G and S_M for gravity and magnetic data, respectively, we obtain:

$$g(m) = Sm = \begin{bmatrix} S_G & 0 \\ 0 & S_M \end{bmatrix} \begin{bmatrix} m^G \\ m^M \end{bmatrix} = \begin{bmatrix} g_G(m^G) \\ g_M(m^M) \end{bmatrix}$$

(8)

8

For reference, the terms defined or used here are summarized in Appendix B. Tomofast-x uses the Least Squares with QR-factorization (LSQR) algorithm (Paige and Saunders, 1982) to solve the least-squares problem. The full matrix formulation of the problem and the related system of equations are provided in Appendix D.

Generally, not all of the terms in equation 5 are used during a single inversion. The activation of selected terms from the cost function (setting $\alpha_i > 0$ and non-null $\boldsymbol{W_i}$) depends on the information available or on the requirements of the modelling to be performed. For example, a term not used during inversion has the corresponding weighting simply set to 0 (the corresponding matrix $\boldsymbol{W_i}$ is set to 0). Conversely, setting a specific weight to a relatively large value leads to the corresponding constraint to dominate the other terms. Such practice is typically used in sensitivity analysis to examine the effect of incorrectly assigned extreme weighting values on the inversion by providing an example to aid detection of this unintended situation.

In the following subsection, we detail the implementation of the different terms. The terms are introduced and detailed following the order they appear in equation 5.

2.4 Detailed formulation of constraints for inversion

In this Sect., we introduce the mathematical formulation of constraints applied during inversion. Throughout this paper, *geological information* relates to information extracted from probabilistic geological structural modelling. Petrophysical information relates to the statistics of the values inverted for (density contrast and magnetic susceptibility).

2.4.1 Smallness term

We repeat the smallness term:

$$\left\|\boldsymbol{W_m}(\boldsymbol{m} - \boldsymbol{m_{pr}})\right\|_p^2. \tag{9}$$

The smallness term corresponds to the ridge regression constraint, or smallness term of (Hoerl and Kennard, 1970). To simplify the problem, the covariance matrix $\boldsymbol{W_m}$ is assumed to be a diagonal matrix. In Tomofast-x, it is used to adjust the strength of the constraint either globally (i.e, $\boldsymbol{W_m} = I$) or locally (i.e., the elements of $\boldsymbol{W_m}$ may vary from one cell to another). In the second case, $\boldsymbol{W_m}$ can be determined using prior information such as uncertainty from geological modelling, or models and structural or statistical information derived from other geophysical techniques (e.g., seismic attributes, probabilistic results from magnetotellurics).

2.4.2 Smoothness term

The smoothness model term (Li and Oldenburg 1996), is a total variation (TV)-like regularisation term based on an original idea of Rudin et al. (1992). It constrains the degree of structural complexity allowed in the inverted model. We repeat the term:

$$\left\|\boldsymbol{W_g}\nabla\boldsymbol{m}\right\|_2^2. \tag{10}$$

The covariance matrix $\boldsymbol{W_g}$ modulates the importance of the term by assigning the weights to each cell. For the sake of simplicity, the matrix $\boldsymbol{W_g}$ is commonly assumed to be a diagonal matrix. It is commonly set as the identity

matrix ($W_g = I$), but several works vary the values in space accordingly with prior information. For instance, Brown et al. (2012), Yan et al. (2017) use seismic models to calculate such weights for the inversion of electromagnetic data, and Giraud et al. (2019a), who present an application case using Tomofast-x, invert gravity data using geological uncertainty information to calculate W_g. In Tomofast-x, it can be set either globally (i.e, $W_g = I$) or locally (i.e., the elements of W_g may vary from one cell to another).

2.4.3 Cross-gradient

The cross-gradient constraints were introduced as a means to link two models that are inverted jointly by encouraging structural correlation between them (Gallardo and Meju, 2004). We refer the reader to Meju and Gallardo (2016) for a review of applications using this technique. We repeat the term below:

$$\left\| W_{xg}(\nabla m^1 \times \nabla m^2) \right\|_2^2. \tag{11}$$

The matrix W_{xg} modulates the importance of the term by assigning the weights to each cell. In previous works, it is always (to the best our knowledge) set as the identity matrix ($W_{xg} = I$), to the exception of Rashidifard et al. (2020), who define such weights using seismic reflectivity and apply this approach to single physics inversion of gravity data constrained by fixed seismic velocity. In Tomofast-x, three finite difference numerical schemes can be chosen to calculate the cross-gradient derivatives: forward, centered, and mixed. In what follows, we use a 'mixed' finite difference scheme, where inversion iteration with an odd number use a forward scheme and even numbers backward scheme (e.g., iteration 3 will use a forward scheme and iteration 4 a backward scheme). This scheme was chosen as it reduces the influence of the border effects of both the forward and backward schemes onto the inverted model.

2.4.4 Statistical petrophysical constraints

One strategy to enforce the petrophysical constraints using statistics from petrophysical measurements is performed by encouraging the statistics of the recovered model to match that of a statistical model derived either from measurements made from the study area or literature values. In the current implementation of Tomofast-x, a mixture model representing the expected statistics of the modelled rock units is used. We use a Gaussian mixture model to approximate the petrophysical properties of the lithologies in the studied area. In the mixture model, the weight of each Gaussian can be set in the input. We suggest to use the probability of the corresponding rock unit when this information is available. The mismatch between the statistics of the recovered model and the mixture model is minimized in the optimization framework following the same procedure described in Giraud et al., 2018, 2019b.

In the ith model-cell, the likelihood term $P(m_i)$ of model-cell m_i is calculated as, for the kth lithology:

$$\mathcal{N}_k = \omega_k \mathcal{N}(m_i | \boldsymbol{\mu_k}, \boldsymbol{\sigma_k}) \tag{12}$$

$$P(m_i) = -\log\left(\sum_{k=1}^{n_f} \mathcal{N}_k \right) + \log\left(\max \mathcal{N}_{k=1..n_f} \right), \tag{13}$$

where:

$$\begin{cases} \omega_k = \dfrac{1}{n_f} \, everywhere \ in \ the \ absence \ of \ spatially - varying \ prior \ information \ (a) \\ \qquad \omega_k = \psi_{k,i} \ in \ the \ i^{th} \ cell \ using \ prior \ information \ (b) \end{cases} \tag{14}$$

and $\mathcal{N}$ symbolises the normal distribution and n_f is the total number of rock formations observed in the modelled area.

In practice, an expectation maximisation algorithm (McLachlan and Peel, 2000) can be used to estimate the mean $\boldsymbol{\mu}_k$ and standard deviation $\boldsymbol{\sigma}_k$ of the petrophysical measurements.

In equation 12-14, ω_k is the weight assigned to the Gaussian distribution representative of the petrophysics of the $\boldsymbol{k^{th}}$ lithology in the mixture. In equation 14a, the weight ω_k assigned to each lithology is constant across the model, while in equation 14b, the weight $\psi_{k,i}$ is derived from information derived from another modelling technique (geology, seismic, electromagnetic methods, etc.) and varies spatially.

We note that a small number of Gaussian distributions might not be suitable to approximate certain types of distributions like bimodal (magnetic susceptibility) or lognormal (electrical resistivity) distributions. However, we point out that increasing the quality of such approximation depends on the number of Gaussian distributions used for approximation (McLachlan and Peel, 2000).

2.4.5 Dynamic bound constraints using the ADMM algorithm

The objective of the dynamic bound constraints is to optimize eq. 5 while ensuring that, in every model cell $m_{1 \leq i \leq n_m}$, the inverted value lies within the prescribed bounds such that $m_i \in B_i$, defined as (Ogarko et al., 2021a):

$$B_i = \bigcup_{l=1}^{L_i} [a_{i,l}, b_{i,l}], \text{with } b_{i,l} > a_{i,l}, \forall \, l \in [1, L_i], \tag{15}$$

where $a_{i,l}$ and $b_{i,l}$ are the lower and upper bounds for the ith model-cell, and l is the lithology index; $L_i \leq n_f$ is the total number of bounds allowed for the considered cell, corresponding to the number of distinct rock units allowed by such constraints. During inversion, such multiple bound constraints on physical property values inverted for are gradually enforced using the ADMM algorithm. Implementation details are beyond the scope of this paper, but more information is provided in Appendix D and we refer the reader to Ogarko et al., 2021a, for details. Details about the general mathematical formulation of the ADMM algorithm can be found in Dykstra (1983), chapter 7 in Bertsekas (2016) and Boyd (2011).

Note that the application of the ADMM bound constraints can be interpreted as are analogous to clustering constraints where (taking the example of the kth model-cell):

1. the centre values depend on both the current model $\boldsymbol{m}$ at any given integration and petrophysical information defining a and b;
2. the weight assigned to each centre values changes from one iteration to the next as a function of the distance between m_k and the closest bound and the number of iterations m_k has remained outside B_k.

2.4.6 Depth weighting and data weighting

To balance the decreasing sensitivity of potential field data with the depth of the considered model cell, Tomofast-x offers the possibility for the calculation of the depth weighting operator. The first one, which we use in this paper utilizes the integrated sensitivities technique following Portniaguine and Zhdanov (2002). For each model cell i, a weight D_{ii} is introduced:

$$D_{ii} = \left(\sum_{k=1}^{n_{data}} (S_{ki})^2 \right)^{\frac{1}{4}} \tag{16}$$

The second option relies on the application of an inverse depth-weighting power law function following Li and Oldenburg (1998) and Li and Chouteau (1998) for gravity, and Li and Oldenburg (1996) for magnetic data:

$$D_{ii} = \left(\frac{1}{z_i + \varepsilon} \right)^{\beta} \tag{17}$$

Where z_i is the depth of the ith model cell and ε is introduced to ensure numerical stability, such that $\mathbf{z} \gg \boldsymbol{\varepsilon}$; the value of β depends on the type of data considered (gravity or magnetic). For more details about the use of depth weighting and selection of values of β, the reader is referred to the references provided in this subsection. The application of depth weighting as a preconditioner to the matrix system of equation solved for during inversion is shown in Appendix D.

2.5 Posterior uncertainty metrics

Uncertainty information is an important building block of modelling and a critical aspect of decision making (Scheidt et al., 2018). When available, uncertainty information can be communicated and used in subsequent modelling or for decision making (see examples of Ogarko et al., 2021a, who use uncertainty information in the model recovered by another method as input to their modelling using Tomofast-x) . Tomofast-x allows the calculation of metrics reflecting the degree of uncertainty in the models before and after inversion. It allows monitoring the evolution of the different terms of the cost function during inversion. Tomofast-x also calculates uncertainty metrics that are specific to the kind of constraints used in inversion:

- the posterior covariance matrix of model *m* as estimated by the LSQR algorithm (Paige and Saunders, 1982 p. 53, for details and A.1.1 for a brief introduction);
- the value of the cross-gradient in each cell;
- the individual $\mathcal{N}_k$ values (eq. 12) of the different Gaussians making up the Gaussian mixture used to define the petrophysical constraints.

The implementation of this series of indicators was performed with the intent to provide metrics for posterior analysis in detailed case studies. More information about these posterior uncertainty indicators is provided in Appendix A, which details functionalities of Tomofast-x not explored here.

3 Synthetic model and data

In this Sect., we introduce how the data used for synthetic modelling were derived and we present examples of using prior information derived from geological modelling. The process of simulating a realistic field case study is described with the design of the numerical experiment.

3.1 Geological framework

The original geological model is based on a region in the Hamersley province (Western Australia). It was built using the map2loop algorithm (Jessell et al., 2021b) to parse the raw data and the Geomodeller® implicit modelling engine for geological interpolation (Calcagno et al., 2008, Guillen et al., 2008) to model the contacts, stratigraphy and orientation data measurement in the area (see geographical location in Figure 2). Data used to generate the model comprise the 2016 1:500 000 Interpreted Bedrock Geology map of Western Australia (https://catalogue.data.wa.gov.au/dataset/1-500-000-state-interpreted-bedrock-geology-dmirs-016, last accessed on 02/12/2020) and the WAROX outcrop database (https://catalogue.nla.gov.au/Record/7429427, last accessed on 02/12/2020). Geological modelling was assisted by interpretation of the magnetic anomaly grid compilations at 80 meters and the 400 meters gravity anomaly grid from the Geological Survey of Western Australia (https://www.dmp.wa.gov.au/Geological-Survey/Regional-geophysical-survey-data-1392.aspx, last accessed on 02/12/2020). More information about data availability is provided in Sect. 7.

In what follows, we use an adapted version of the original structural geological framework of the selected region by increasing the vertical dimensions of the model cells and assuming a flat topography. The resulting reference geological model used to generate the physical properties for geophysical modelling measurements is shown in Figure 2 in terms of its geological units.
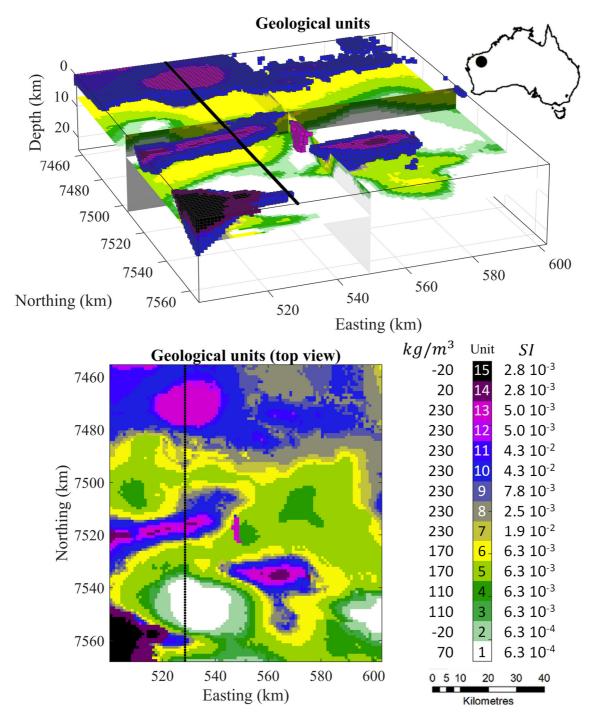
13

Figure 2. True model used for geologic modelling and geophysical inversion. The top part shows the map view. The black line represents the surface coordinates of the 2D profile considered here for illustration of Tomofast-x's utilisation. The values given on either side of the colour bar indicate the density contrast and magnetic susceptibility attached to each rock unit. Note that several rock units present similar density contrast or magnetic susceptibilities, making them undistinguishable using either gravity or magnetic inversion.

In addition to the modification of the structural model, we make adjustments on the original density values derived from field petrophysical measurements by reducing the differences between the density contrasts of different rock units. Doing so, we increase the interpretation ambiguity of inversion results and decrease the differentiability of the different rock units. The same procedure is applied to magnetic susceptibility to make accurate imaging using

14

inversion more challenging. The three-dimensional (3D) density contrast and magnetic susceptibility models used to generate the gravity and magnetic data are shown in Figure 3.
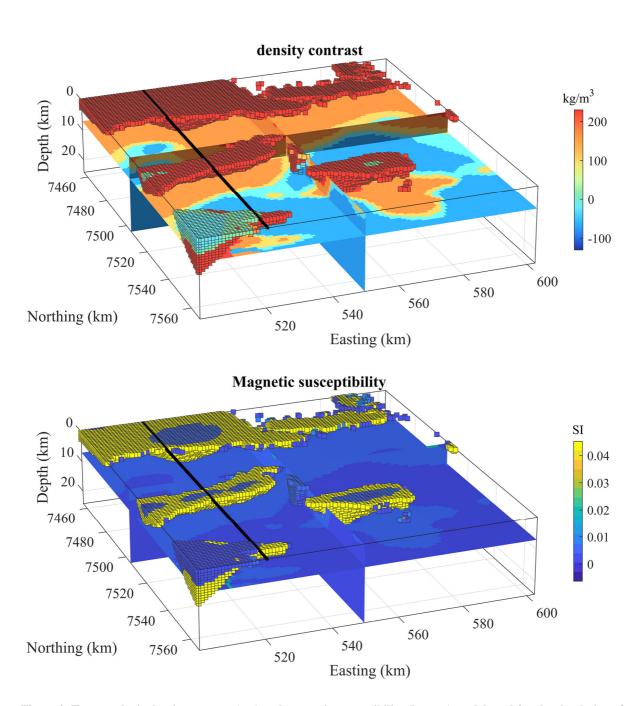
370

Figure 3. True synthetic density contrast (top) and magnetic susceptibility (bottom) model used for the simulation of geophysical data. The black line represents the surface coordinates of the 2D profile considered here. The voxels represent lithologies 10 through 15, color-coded with their respective density contrast and magnetic susceptibility values.

375

3.2 Geophysical simulations and model discretization

The core 3D model is discretised into $N_x \times N_y \times N_z = 103 \times 113 \times 33$ cells of dimensions equal to $999 \times 996 \times 745$ m^3. For both gravity and magnetics, we generate one geophysical measurement per cell along the

horizontal axis, leading to $N_x \times N_y = 11{,}639$ data points for each method, and add 10 padding cells in each

380   horizontal direction to limit numerical border effects in the forward calculation, leading to a total of $123 \times 133 \times 33 = 539{,}847$ model-cells. We simulate a ground gravity survey by locating the measurements 1 m above ground level, and aeromagnetic data acquired by a fixed-wing aircraft flying 100 meters above surface. To test the robustness of our inversion code to noise content in the data, the geophysical data inverted here are contaminated by noise.

385   The noise component was generated as follows. For each gravity measurement, we first add a perturbation value randomly sampled from the standard normal distribution of the whole dataset multiplied by 9% of the measurement's amplitude. We then add a second perturbation value randomly sampled from the standard normal distribution with an amplitude of 3 mGal (2% of the dynamical range). These values were derived manually to obtain a realistic noise contamination. To simulate small-scale spatial coherence in the noise generated in this

390   fashion, we then apply a two-dimensional Gaussian filter to the 2D noise map obtained from the previous step. We then apply a two-dimensional median filter to the resulting noise-contaminated gravity data to simulate de-noising. For magnetic data, we apply the same procedure, using, 12.5% of the measurement's amplitude for the first step, and 15 nT (1% of the dynamical range) for the second step. Similarly to gravity data, these values were derived manually; no levelling noise was simulated. For comparison, the noise-free and contaminated synthetic

395   measurements are shown in Figure 4. The resulting noise standard deviation σ_{noise} for gravity and magnetic data are equal to 1.2 mGal and 8.5 nT, respectively.

The gravity data modelled here correspond to the complete Bouguer anomaly. Magnetic data are simulated using the magnetic strength of the Hamersley province (53011 nT, which approximates the International Geomagnetic Reference Field in the area) reduced to the pole.
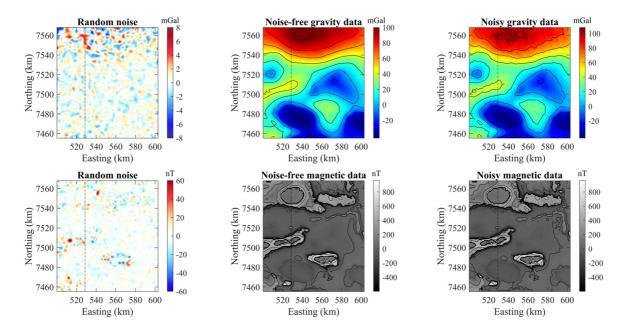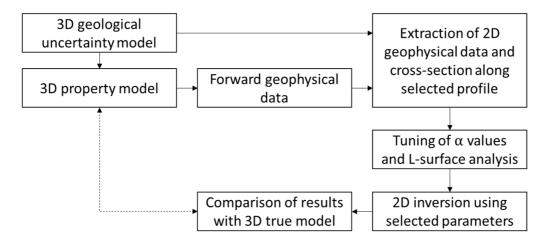

400

Figure 4. Noise (left hand side column) added to the data calculated from the true model (central column) and resulting noisy data (right hand side) for the gravity (top) and magnetic (bottom) datasets. The contour lines shown corresponding to the ticks shown in the palette's colour bar. The black line represents the location profile we use for the inversions performed here.

16

405 To complete the 3D modelling procedure, a series of 100 structural geological models are generated using Monte-Carlo perturbations of the geological measurements (foliation and contact points between geological units) constraining the geological structures. This was performed using the Monte Carlo Uncertainty Estimator (MCUE) technique of Pakyuz-Charrier et al. (2018, 2019). The result is an ensemble of models that all fit the geological measurements within a given set of prior uncertainty parameters. The ensemble is thus assumed to represent the

410 geological model space, rather than just a single 'best-guess' model. Probabilities for the occurrence of different rock units can be calculated from the ensemble and used to constrain geophysical inversion (see examples of Giraud et al., 2017, 2019a; Ogarko et al., 2021a). More specifically, MCUE is useful to obtain the probability $\psi_{i,l}$ of occurrence of the different lithologies l for every i^{th} model cell, and to calculate the related uncertainty indicators (Sect. 3.3). Detailing the probabilistic geological modelling procedure and analysing the results in 3D

415 is beyond the scope of this paper and interested readers are referred to Lindsay et al. (2012), Pakyuz-Charrier et al. (2018), Wellmann and Caumon (2018) and references therein.

In this contribution, we simulate a case study where modelling is carried out along the 2D profile materialized by the black line in Figure 3 and Figure 4, extracted from the 3D modelling framework as detailed in the next subsection.

420 **3.3 2D model simulation in a 3D world**

As mentioned above, we perform the inversion of geophysical data along a 2D profile for simplicity and to simulate the challenging case of 2D data acquired in a 3D geological setting, in a part of the model where subhorizontal or gently dipping features can be observed. The philosophy of the numerical study presented here is summarized in Figure 5.
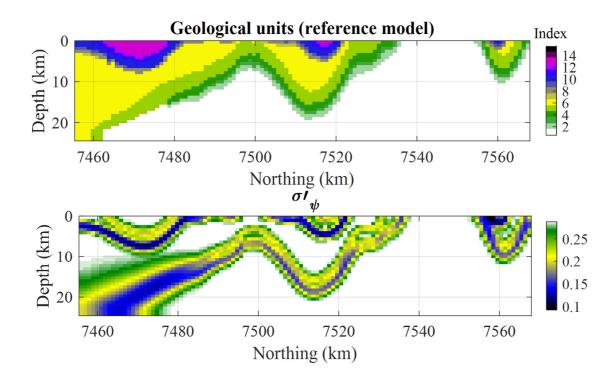

425

Figure 5. Summary of experimental protocol for synthetic study and testing of different functionalities of Tomofast-x.

The 2D geological Sect. considered is shown in Figure 6 (the black line marked in Figure 3 and Figure 4). Geological certainty is estimated using a measure of the dispersion away from the perfectly uninformed case where the all rock units are equiprobable. In each model-cell, this measure, which we write σ'_ψ, is calculated as a

430 function of the standard deviation σ_ψ of the probability ψ of observing the different rock units, as follows:

17

$$\sigma'_{\psi} = \sqrt{\frac{1}{card}} - \sigma_{\psi}, \tag{18}$$

where **card** is the geological cardinality of the model and is equal to the number of possible rock units observed in one location across the entire ensemble. From equation 18, σ'_{ψ} is maximum where rock units are well constrained, and minimum where the model is the most uncertain. This geological certainty metric is shown in Figure 6 for the 2D Sect. considered in this example.

Figure 6. Two-dimensional slice extracted from the 3D model along the profile: geological reference model (top), the σ'_{ψ} metric (bottom). Here, the geological uncertainty metric considered is the standard deviation of the probability of the different lithologies as per equation 18.

The probabilities of observation of the different lithologies are shown in Figure 7. Note that for the purpose of the tests we perform on gravity data inversion, we reduce the set of probabilities by grouping rock units into fictitious units with the same density contrasts as single rock units. This reduces the number of rock units to six units that can be distinguished by gravity inversion as several units may be assigned the same density contrast.
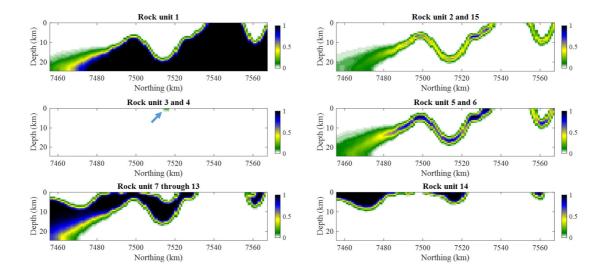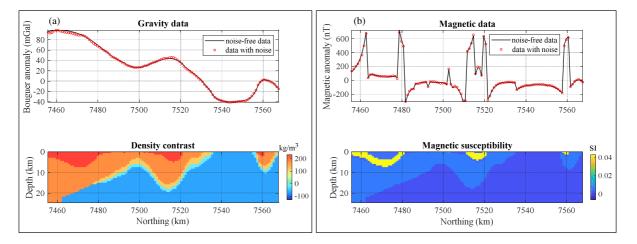
Figure 7. Observation probabilities for the rock units that present differencing density contrasts. Units 3 and 4 are nearly absent from the Sect (see maximum probability area marked by the arrow).

The geophysical data we use for inversion are extracted along the line marked in Figure 3 and Figure 4. The geophysical data and reference (true) petrophysical model extracted in this fashion are shown in Figure 8. Care was taken not to use the same mesh for both generating the synthetic dataset and its inverse modelling.

Figure 8. Two-dimensional slice extracted from the 3D model: gravity data and density contrast (a), magnetic data and magnetic susceptibility (b).

To inverse model the data shown in Figure 8, we generate a mesh centred on the profile (oriented along the y-direction) and add padding to either side and the northern and southern extremities. The resulting model comprises $n_x \times n_y \times n_z = 13 \times 133 \times 33$ cells of dimensions equal to $2998 \times 996 \times 745$ m. Note that we increased model-cells' size in the direction perpendicular to the profile. The gravity and magnetic datasets along the profile each comprise 113 data points evenly distributed along the line.

While we focus on a 2D section extracted from a 3D model presented here (see location in Figure 3 and Figure 4), the 3D model and the associated gravity and magnetic datasets shown here are made publicly available (see Sect. 7).

19

4 Application example: sensitivity analysis to constraints and prior information

In the examples shown below, we first perform single domain and joint (multiple domain) inversion (using the cross-gradient constraint) assuming identity matrices for $\boldsymbol{W_m}$, $\boldsymbol{W_g}$, and $\boldsymbol{W_x}$. We then investigate the influence of prior information on single domain inversion by combining structural and petrophysical information in the case of gravity inversion. The combination of petrophysical and structural constraints derived from geology is tested. The intention is to address knowledge gaps in the literature that describes the effects of parameterization of such constraints.

4.1 Experimental protocol

It is necessary to determine the appropriate weights α assigned to the terms defining the constraints applied during inversion to optimize the cost function in equation 5. The α values that define the weights of the different terms in the cost function constitute hyperparameters of the inverse problem. Appropriate estimation of these hyperparameters is necessary to approximate the optimum value of the global misfit function. To this end, we use the L-curve principle (Hansen and O'Leary 1993, Hansen and Johnston 2001, Santos and Bassrei 2007) for each of the cases presented below. We perform series of inversions sampling α values spanning the plausible range of potential choices using a heuristic approach.

When two constraint terms are used in inversion (i.e., with α > 0), we extend the L-curve approach to the two-parameter cases. In such case, the optimum values for the α weights are determined applying the L-curve criterion using L-surfaces (or elbow surface) instead of L-curves (Belge et al., 2002) (we note that the L-curves as plotted here can also be referred to as 'Tikhonov curves' in the case where data misfit is plotted as a function of regularisation value). The optimum value for the α weight of the two constraint terms is therefore obtained by identification of the inflection point of the surface made up of the variations of the data misfit as a function of the weights under consideration. We chose this approach for its simplicity and note that there exist other techniques that use an automated process, such as the generalized cross-validation technique (Craven and Wahba, 1978). We refer the reader to Farquharson and Oldenburg (2004) for a general introduction and Giraud et al. (2019b), Martin et al. (2020, for application of this principle to inversions using Tomofast-x. The role the L-surface analysis plays in the synthetic case presented here is reminded in the workflow shown in Figure 5. In our analysis, we set the objective value for the data misfit $\left\| \boldsymbol{W_d}\big(\boldsymbol{d} - \boldsymbol{g(m)}\big) \right\|_2^2$ to be equal to the objective data misfit $\Theta_d^{obj}(\boldsymbol{d}, \boldsymbol{m})$, defined as:

$$\Theta_d^{obj}(\boldsymbol{d}, \boldsymbol{m}) \geq \frac{n_{data}\sigma_{noise}}{\sum_{i=1}^{n_{data}}(d_i)^2}, \tag{19}$$

so that the data is reproduced with a level of error superior or equal to the estimated noise level of the data. Here, this leads to $\Theta_d^{obj} = 5.01 \times 10^{-4}$ for gravity inversion, and to $\Theta_d^{obj} = 1.55 \times 10^{-4}$ for magnetic data inversion, respectively.
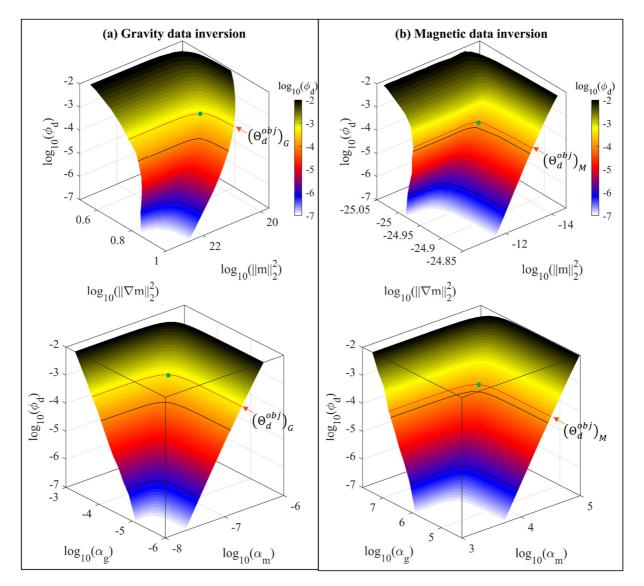
For the sake of consistency in our study of the influence of constraints onto inversion, we set $\boldsymbol{m_{pr}} = 0$ kg/m³ for gravity data inversion and $\boldsymbol{m_{pr}} = 0$ SI for magnetic data inversion.

4.2 Homogenously constrained potential field inversions

We first perform single physics inversion following the common strategy of constraining the model using smallness and smoothness constraints. Obtaining a good approximation of the optimum values of these parameters gives insights into the numerical structure of the problem. It constitutes valuable knowledge when using other kinds of constraints and we consider it a good practice to run such inversion prior to using more advanced constraints. Here, the first α parameters to determine are α_m and α_g, for both gravity and magnetic data inversion, assuming identity $\boldsymbol{W}_m$ and $\boldsymbol{W}_g$ matrices so that the constraints are applied homogenously over the entire model.

We generate grids in the (α_m, α_g) plane using $\alpha_m \in [10^{-8}, 10^{-6}]$ and $\alpha_g \in [10^{-6}, 10^{-3}]$ for gravity inversion, and $\alpha_m \in [10^3, 10^5]$ and $\alpha_g \in [10^3, 10^8]$ for magnetic data inversion, respectively. These ranges were determined empirically and assumed to comprise the optimums. In this subsection, all matrices $\boldsymbol{W}$ in equation (5) are set as the identity matrix.

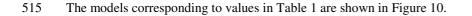**Figure 9. Elbow surfaces for gravity (a) and magnetic (b) inversions (top) and plot of the data misfit term as a function of the α weights (bottom). Each plot uses a total of 1260 points sampling the (α_m, α_g) plane. The black lines show the contour values corresponding to the ticks shown in the palette's colour bar, which shows the value of the data misfit**

21

510   For accurate estimation, the (α_m, α_g) values are sampled more finely closer to the estimated optimum values. The resulting L-surfaces are shown in Figure 9, where the vicinity of the optimum value of (α_m, α_g) is shown with a green dot. From these values, we estimate the optimum values of (α_m, α_g) reported in Table 1.

Table 1. Optimum values of (α_m, α_g) estimated from L-surface analysis.

	α_m	α_g
Gravity inversion	2.1×10^{-7}	1.8×10^{-4}
Magnetic inversion	3.22×10^{4}	4.52×10^{6}

515   The models corresponding to values in Table 1 are shown in Figure 10.
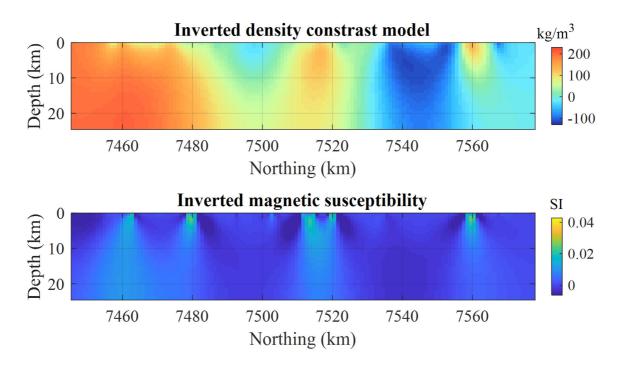

Figure 10. Results from separate inversions using smallness and smoothness constraints. The starting and prior models model are equal to zero everywhere and the smoothness constraint is applied homogeneously.

520   The values of α_m and α_g obtained for such constraints can be used as a starting point in subsequent inversions to understand the influence of prior information when varying amounts and types of prior information are available about the structure of the subsurface or its petrophysics. For instance, in what follows we will investigate the utilisation of geological information to define $\boldsymbol{W}_m$ and $\boldsymbol{W}_g$ (eq. 9 and 10, respectively) and see how it can combined with petrophysical data to define B (eq. 15) (see following subsection where we use global and

525   structural and/or petrophysical information). In the case of structural constraints relying on the spatial derivatives of model values (cross-gradient or local smoothness), the value of α_m may be kept constant and while the other α parameters (α_g or α_x) are adjusted. Conversely, α_g may be kept and α_m set to 0 for the utilisation of petrophysical

constraints acting on the model values themselves instead of the spatial derivatives (ADMM or statistical petrophysical constraints). Here, we restrict our analysis to two α values being strictly superior to zero, thereby 530   accounting for prior information in up to two constraints terms in the definition of the regularisation term in eq. 5.
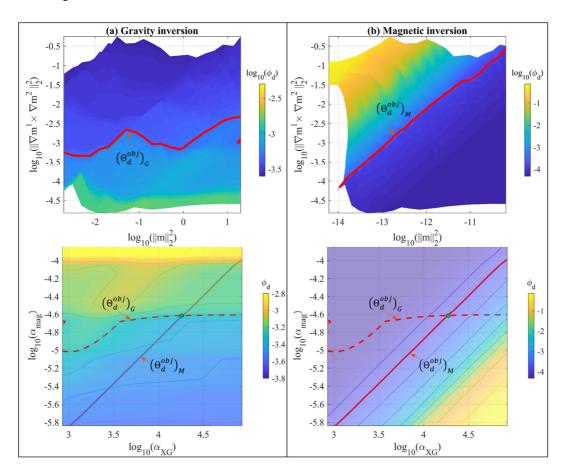
4.3 Joint inversion using the cross-gradient constraint

We start from the previous step to perform joint inversion using the cross-gradient constraint. Keeping the α_m weight constant and equal to the values determined from single domain inversion, it remains necessary to estimate 535   the optimum values of the cross-gradient constraint weight, α_x, and the relative importance given to the gravity and magnetic data misfit terms (setting $\alpha^G = 1$, it remains to determine α^M). We therefore investigate values in the (α_x, α^M) plane, which we sample in the same fashion as in the previous sub-section. The resulting surfaces are shown in Figure 11.
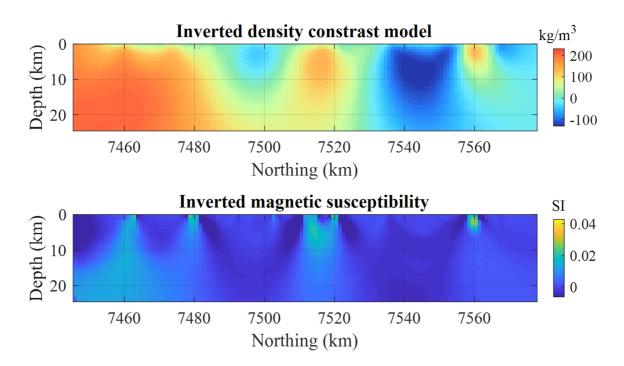

540   **Figure 11. Determination of the optimum (α_{xg}, α^M) parameters in the case of the joint inversion using the cross-gradient constraint. Top view of the elbow surfaces for gravity (a) and magnetic (b) inversions (top) and plot of the data misfit term as a function the α weights (bottom). The solid lines show the contour values of the data misfit, which values are given by the color bar on the side. The bold red solid line shows the contour level of Θ_d^{obj} for the corresponding dataset (gravity or magnetic) while the dashed line shows the same quantity for the other dataset. The green dot marks 545   their interSect., indicating the optimum (α_{xg}, α^M) values.**

In contrast to the single-physics inversion shown in Figure 9, it appears from Figure 11 that the two hyperparameters to be determined here, α_{xg} and α^M, influence the inversion differently. While the contour levels of the magnetic data misfit show a linear trend in the (α_{xg}, α^M) plane, it is clearly non-linear in the case of gravity

data misfit. This difference might be explained by the fact that the cross-gradient is a second order regularisation (product of two spatial derivatives of model values) linking two models that are otherwise decoupled. In addition, this suggests that in cases differing from this one, the hyperparameter selection may be non-unique. Nevertheless, the value of the optimum value is unambiguous in our case and can be determined easily. From Figure 11 we obtain $(\alpha_{xg}, \alpha_{mag}) = (1.995 \times 10^4, 2.57 \times 10^{-5})$. The corresponding inversion results are shown in Figure 12.

Figure 12. Joint inversion results obtained from utilisation of the cross-gradient constraint.

Compared to Figure 12, we observe that the application of the cross-gradient constraint leads to adjustments of the model ensuring more structural consistency between density contrast and magnetic susceptibility, illustrating the applicability of the approach presented here. Also note that the model is also visually closer to the true model from approximately 7520 km Northing and more. However, despite the increased structural consistency between the density contrast and magnetic susceptibility models, some of the structures of the model are not recovered accurately. For instance, the basin-shape structure around 7500 km Northing mirrors the actual geological structure (see Figure 8) and is an effect of non-uniqueness onto inversion. In this case, this illustrates the need for prior information in our inversion. While joint inversion of gravity and magnetic data using the cross-gradient constraint improves imaging comparatively with an inversion constrained only using smallness and smoothness constraints, prior geological information or petrophysical information may be necessary to alleviate the remaining uncertainty.

4.4 Smallness and smoothness constraints using geological information

In this subsection, a sensitivity analysis to prior information in inversion is performed through a series of scenarios where geological structural information is introduced to adjust the smallness and smoothness constraints through W_m and W_g, respectively. In what follows, we apply this approach to gravity inversion.

The influence of geological information in defining the smallness and smoothness terms (detailed in 2.4.1 and 2.4.2) is analysed by investigating three additional scenarios allowed by the utilisation of either homogenous or geologically-derived W_m and W_g matrices. In each case, we start from the (α_m, α_g) weights estimated in Sect. 4.2 from the analysis of the L-surface, which we adjust to obtain the geophysical misfit sufficiently close to objective values. We remind that α_m and α_g weight the overall contribution of the model smallness and smoothness, respectively, in the cost function (eq. 5).

In the first scenario we investigate (scenario b in Table 2), geological uncertainty information is used to define W_m while keeping W_g homogenous. This allows us to test the influence of geological prior information onto the smallness term. The values of the diagonal variance matrix W_m are calculated using the geological certainty metric σ'_ψ (eq. 18, shown in Figure 6b for the 2D section modelled here), and keep W_g homogenous.

Contrarily to the previous tests (see 4.2) where $W_m = I_{n_m}$, we have, for the k$^{\text{th}}$ model cell:

$$(W_m)_{kk} = \left(\sigma'_\psi\right)_k. \tag{20}$$

Because $0 \leq \left(\sigma'_\psi\right)_k \leq 1 \; \forall \, k$, we have:

$$tr(W_m) = \sum_{i=1}^{n_m} \left(\sigma'_\psi\right)_i \leq tr\left(I_{n_m}\right) = n_m. \tag{21}$$

Consequently, setting W_m in this fashion and keeping α_m constant translates in a lower overall relative importance of the smoothness term in the least-squares cost function (eq. 5), thereby moving away from the trade-off inferred from the L-curve principle (Sect. 4.2). To mitigate this, we adjust α_m to a value α'_m such that:

$$\alpha'_m = \alpha_m \sqrt{\frac{n_m}{\sum_{i=1}^{n_m}\left(\sigma'_\psi\right)_i}}, \tag{22}$$

which equates $(\alpha'_m)^2 tr(W_m)$ with $(\alpha_m)^2 n_m$ so that the overall weight assigned to the smallness term remains the same with and without geological structural information (left-hand side and right-hand side of inequality in equation 20, respectively). Because the values of $\|W_m m\|$ depend on both W_m and m, which vary in space and also depends on the other terms of the cost function, minor adjustments of the value of α'_m are necessary to reach the objective value of data misfit. In this example, this leads to tune the suggested $\alpha'_m = 8.4 \times 10^{-7}$ to $\alpha'_m = 8.85 \times 10^{-7}$ (keeping α_g constant). The corresponding inverted model is shown in Figure 13b. The corresponding α weights are repeated in Table 2.

In the second scenario we test (scenario c in Table 2), geological uncertainty information is then used to define W_g while keeping W_m homogenous. This allows us to test the influence of geological prior information onto the smoothness term. Following the same procedure as for the smallness term, we adjust the suggested $\alpha'_g = 3.6 \times 10^{-4}$ to $\alpha'_g = 4.1 \times 10^{-4}$. The corresponding inverted model is shown in Figure 13c. Finally, we test the case where both W_m and W_g are defined using geological information in the form of σ'_ψ. Starting from values of α'_m and α'_g used in the previous tests, minor tuning is performed, leading to $\alpha'_m = 6.1 \times 10^{-7}$ and $\alpha'_g = 6.0 \times 10^{-4}$. Inversion results in the model shown in Figure 13d.

600 **Table 2. α values derived for simultaneous usage of local and global smallness and smoothness constraints. Scenario (a) is a reminder of values obtained in 4.2 when only global constraints are used.**

	α_m	α_g
Global smoothness, global smallness constraints (a)	2.1×10^{-7}	1.8×10^{-4}
Global smoothness, local smallness constraints (b)	8.85×10^{-7}	1.8×10^{-4}
Local smoothness, global smallness constraints (c)	2.1×10^{-7}	4.1×10^{-4}
Local smoothness, local smallness constraints (d)	6.1×10^{-7}	6.0×10^{-4}

Figure 13. Results from gravity inversion constrained by: (a) homogenous smallness and homogenous smoothness constraints, (b) geologically-derived smallness and homogenous smoothness constraints (c) homogenous smallness and geologically-derived smoothness constraints, (d) geologically-derived smallness and geologically-derived smoothness constraints.

As can be seen in Figure 13 by comparing Figure 13a-b with Figure 13c-d, the utilisation of geological structural

information to adjust the smoothness regularization strength spatially has more impact on inversion than adjusting

the smallness term. While incorporating prior geological information in $\boldsymbol{W}_m$ constrains the model to a certain extent, using $\boldsymbol{C}_v$ to derive $\boldsymbol{W}_g$ has more influence on the inverted model than for $\boldsymbol{W}_m$, with resulting models that closer to the reference model.

Comparing Figure 13c and Figure 13d indicates that the use of geological uncertainty information to adjust the smallness regularization strength spatially (through $\boldsymbol{W}_m$) in addition to the model smoothness term (through $\boldsymbol{W}_g$) modifies inversion results further towards the reference model. Figure 13d, which results from inversion using prior information in both constraint terms, provides the model closest to the reference. While most interfaces are well-recovered when using geological information to define both $\boldsymbol{W}_g$ and $\boldsymbol{W}_m$, the recovered density contrasts remain affected by the ambiguity inherent to gravity data in the presence of subhorizontal geological units (around 7460 km northing). This suggests that in this example, more prior information might be useful in recovering the causative model more truthfully, especially in cases where potential field inversion is ambiguous (e.g., subhorizontal interfaces for gravity inversion). This is the object of the next subsection, which describes a new single-physics inversion scenario where petrophysical constraints are combined with structural constraints and geological information.

4.5 Structural and petrophysical constraints

In addition to the definition of matrices $\boldsymbol{W}_m$ and $\boldsymbol{W}_g$, geological information can be combined with petrophysical knowledge to define the range of density values allowed in inversion. This is achieved with spatially varying bound constraints on the property inverted for – density contrast in this case (see Sect. 2.4.5). Here, such bounds are defined using multiple intervals, each one corresponding to the range of density contrast values expected for a geological unit. Such bounds can be defined globally (homogenously) where all intervals are allowed everywhere in the model, or locally when prior information about the presence of the rock units is available. In this work, we use the probability of occurrence of the different rock units to derive bounds that vary in space accordingly with the probability of observation of each of the rock units. In a given cell, only the bound values corresponding to rock units with a probability $\Psi > 0$ are considered. Starting from eq. 15, such spatially-varying bounds B_k of the kth model-cell are obtained as follows:

$$B_k = \bigcup_{\substack{l=1 \\ \Psi_{k,l}>0}}^{n_f} [a_{k,l}, b_{k,l}], \tag{23}$$

where a and b correspond to lower and upper bounds. We consider narrow bounds such that $b_{k,l} = a_{k,l} + \varepsilon$, with $\varepsilon \ll a_{k,l}$, to encourage inversion to use density contrasts that closely resemble values defined a priori. Equation 23 corresponds to the application of a Boolean operator to the probabilities $\Psi_{k=1..n_f}$ in every cell to divide the studied area into domains defined by rock units with a probability $\Psi > \Psi_{th}$. In such case, the ADMM bound constraints act as a proxy for a prior model dynamically constrained by petrophysical information.

Four additional scenarios are tested to determine the influence of prior information onto inversion to accommodate the addition of both the damping gradient and ADMM bound constraint term. The use of prior information is illustrated in Figure 14.

Figure 14. Tested combinations for the utilisation of prior information into inversion. Bold frames indicate the utilisation of geological information to define the constraints.

At a given iteration, the ADMM bounded constraint encourages inverted values to evolve inside one of the prescribed intervals depending on the current model m. As mentioned above, we can then make the analogy with a smallness term that is dynamically updated. For this reason, we treat the ADMM bounded constraints in the same fashion as the smallness term, which we apply simultaneously to the model smoothness term.

Following the same protocol as Sect. 4.1 to determine α_g and α_a, we first perform inversion without the use of geological information in the form of the probabilities for the occurrence of different rock units or metrics that can be derived from them (Figure 14a, i.e., with $\boldsymbol{W}_g$ and $\boldsymbol{W}_a$ equal to the identity matrix and the corresponding regularization term weighted by α_g and α_a, respectively). It is therefore necessary to determine the value of α_g and α_a (eq. 5). We perform an L-surface analysis and sample values in the (α_g, α_a) plane to estimate the optimum values for these hyperparameters (see Figure 15). Values of α_g varying from 1.585×10^{-7} and 1.585×10^{-5}, and values of α_a vary from 2.484×10^{-5} and 2.484×10^{-7}. The resulting L-surface is shown in Figure 15.
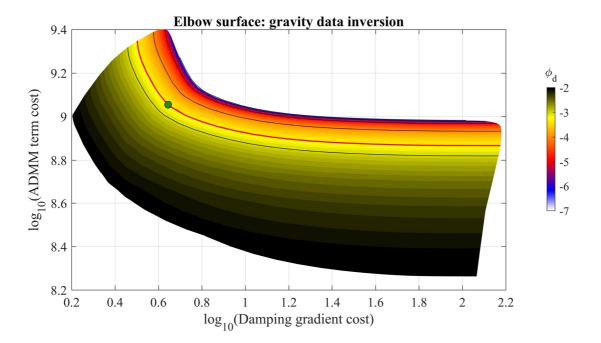
Figure 15. Elbow surfaces for gravity inversions. A total of 840 points sampling the (α_g, α_a) plane were used. The black lines show the contour values corresponding to the ticks shown in the palette's colour bar, which shows the value of the data misfit term. The red line indicates the contour value of $\Theta_d^{obj} = 5.008 \times 10^{-4}$, guiding the selection of the optimum (α_g, α_g) values, and the green point indicates the curve inflection point.

From Figure 15, we estimate the hyper-parameters (α_g, α_a) to be $(\alpha_g, \alpha_a) = (2.2 \times 10^{-5}, 1.3 \times 10^{-7})$ in the case no geological information is used, meaning that both constraints are applied homogenously across the model.

From there, we follow the same procedure as described above (Sect. 4.1 and 4.2) to obtain an estimate for the values of α_a and α_g in the different configurations shown in Figure 14b-d. The resulting inverted models are shown in Figure 16, and the estimates of (α_g, α_a) are provided in Table 3.

Table 3. α values derived for simultaneous usage of global and local smoothness and ADMM bound constraints. Cases (a) through (d) correspond to cases (a) through (d) in Figure 14.

	α_g	α_a
Global constraints (a)	2.2×10^{-5}	1.3×10^{-7}
Global ADMM, local gradients (b)	3.3×10^{-4}	2.6×10^{-7}
local ADMM, global gradients (c)	1.1×10^{-4}	3.6×10^{-7}
local ADMM, local gradients (d)	3.1×10^{-4}	3.25×10^{-7}

Figure 16. Results from gravity inversion using: (a) global ADMM clustering and homogenous smoothness constraints, (b) global ADMM clustering and geologically-derived smoothness constraints, (c), ADMM clustering and geologically-derived smoothness constraints, and (d) geologically-derived ADMM clustering and geologically-derived smoothness constraints. For visual comparison, the true model is reminded at the bottom of the Figure.

Figure 16 shows that the use of ADMM facilitates recovery of better-defined interfaces between rock units than in previous inversions (Figure 10, Figure 12, and Figure 13), and decreases the misfit with the causative model (shown in Figure 6). Unsurprisingly, without the use of geological information (Figure 16a) inversion results remain inconsistent with geology in several parts of the model, especially around the position 7500 km north. The

31

680   inconsistent results can be partly mitigated by using geologically-derived smoothness constraints (Figure 16b). In comparison, however, Figure 16c shows that use of geological information to determine the bounds recovers features much closer to the causative model.

While Figure 16d shows the more robust results overall, Figure 16c and Figure 16d present generally similar features. This indicates that, in this case, geological uncertainty information in structural constraints only allows

685   refining features largely controlled by the utilisation of the ADMM constraints. This statement is supported by Figure 16 where the comparison cases (a,b) and (c,d) reveal that the effect of using geological information to define bounds dominates over the effect of using uncertainty to define structural constraints.

The comparison of cases (a,b) and (c,d) in Figure 16 can be extrapolated to Figure 13 and Figure 16, to compare constraints more broadly. This is discussed in 5.1, which presents a short comparative analysis of all gravity

690   inversion results.

5 Discussion

5.1 Sensitivity analysis summary: comparison of constrained inversions

Tomofast-x was developed with the intent of providing practioners with an inversion platform accounting for various forms of prior information and geophysical datasets. We have tested a series of constraints involving joint

695   inversion, geological and petrophysical information. The inverted density contrast models for inversion using global smallness and smoothness constraints, joint inversion using the cross-gradient technique, geologically-derived smallness and smoothness constraints, and ADMM bound constraints (both global and using geological information) are shown in Figure 17. We remind that all models shown here produce a similar data misfit Θ_d^{obj} accordingly with equation 19.

(a) Global smallness and global smoothness constraints

α_m α_g

Model misfit: 96.4 kg/m^3

(d) Global smallness and local smoothness constraints

α_m α_g

Model misfit: 94.5 kg/m^3

(g) Global bound and local smoothness constraints

α_g α_a

Model misfit: 73.4 kg/m^3

(b) Local smallness and global smoothness constraints

α_m α_g

Model misfit: 94.0 kg/m^3

(e) Local smallness and local smoothness constraints

α_m α_g

Model misfit: 94.0 kg/m^3

(h) Local bound and global smoothness constraints

α_g α_a

Model misfit: 40.8 kg/m^3

(c) Global smallness and cross-gradient constraints

α_m α_x

Model misfit: 98.4 kg/m^3

(f) Global bound and global smoothness constraints

α_g α_a

Model misfit: 79.4 kg/m^3

(i) Local bound and local smoothness constraints

α_g α_a

Model misfit: 31.6 kg/m^3

α_m Model damping term α_x Cross-gradient α_g Gradient damping term α_a ADMM term No geological information Using geological information

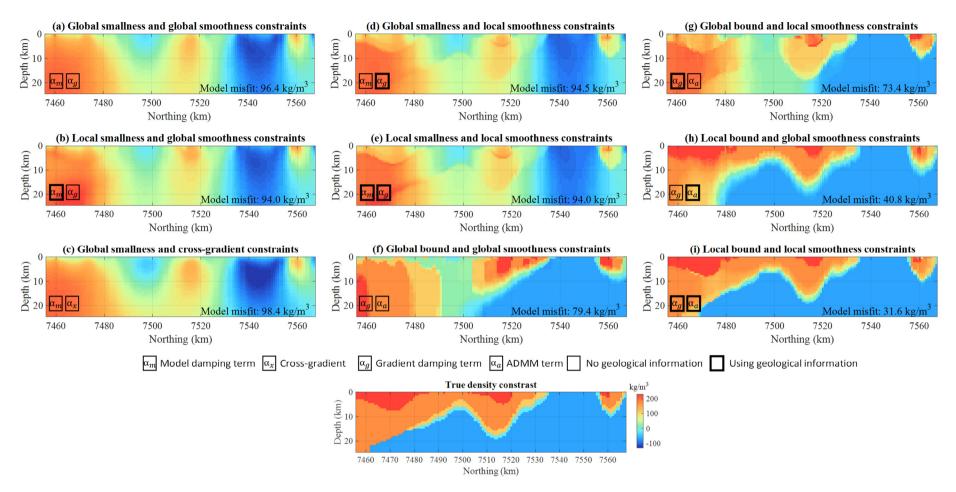**True density constrast**

kg/m^3

700

Figure 17. General comparison of all inversion results obtained from gravity inversion. The legend on the Figure identifies the different types of inversion shown. We remind the true model at the bottom of the Figure. The model misfit indicated on each panel is calculated as the root-mean-square of the difference between the inverted and true models.

Firstly, it appears from Figure 17 that regardless of the type of constraints considered, the utilisation of geological information (cases b, d, e, g, h, i) to derive spatially-varying constraints for the W matrix of both terms used

705 provides the models that are visually closest to the true model. In this category, the utilisation of petrophysical information in the ADMM bound constraints provides (cases h-i) models that are closest to the true model (lowest model misfit values indicated in the titles of the panels in the Figure). Secondly, the comparison of cases (a) and (d), (b) and (e), (f) and (g), and (h) and (i) indicates that while it has a less significant influence on the results, incorporating geological information in the definition of the smoothness term also influences inversion results

710 significantly. Lastly, comparison of cases (a) and (b), and (d) and (e) suggests that the utilisation of geological information to adjust the smallness strength spatially has an effect on inversion that is, with the cross-gradient constraints (where structural information is passed on from another geophysical dataset), the lowest.

From the results shown in Sect. 4.2 through 4.5 and compared in Figure 17, it is possible to make a qualitative ranking of the constraints accordingly with their influence on the resulting model (from the most important

715 influence to the least important influence):

ADMM bound constraints > smoothness constraints > smallness constraints > cross-gradient constraints. This observation is also corroborated by the values of the root-mean-square misfit between the true and inverted model. We note that this ranking remains speculative as it might apply only to models sharing similarities with the case we investigated

720 From these observations we also deduce that when geological uncertainty information is added to the definition of constraints (i.e., σ'_ψ for defining W_m and W_g and probabilities for defining B), the term of the cost function with the highest influence on the process will determine the main features of the model, which will be adjusted by the other term.

Tomofast-x was developed, with the intent of providing practioners with an inversion platform allowing various

725 forms of prior information and geophysical data. Constraints that represent uncertainty and our level of epistemic knowledge provide useful constraint to inversion. This is encouraging as the Tomofast-x platform addresses a gap in inversion schemes that rely on a single model, and that model being as similar as possible to the target region, an often impossible requirement to meet. Thus Tomofast-x opens additional research avenues to the community that are widely acknowledged, but remain largely unaddressed. Conceptual uncertainty relating the prior

730 assumptions made about tectonic event history of the region, and thus the structure under study can be analysed. Different event histories and topologies can be considered, giving a wider scope to the model space, and allowing the geophysics to invalidate implausible histories, but giving us pause to consider other that may be less likely, but nonetheless possible.

5.2 Outlook for future developments

735 Another research avenue under consideration is the integration of results from probabilistic modelling of seismic and electrical data into Tomofast-x. As stated in introduction, one of the goals born in mind during the design of

Tomofast-x is interoperability. Current work involves the integration of Tomofast-x into the Loop[1] open source 3D probabilistic geological and geophysical modelling platform (Ailleres et al., 2019), in an effort to unify geological and geophysical modelling at a more fundamental level than the more common cooperative approaches. Ongoing developments include the possibility to adjust weight assigned to the bound constraints accordingly with uncertainty levels in prior information used to defined spatially varying intervals.

Future research includes the utilisation of implicit geological modelling (in the sense of Calcagno et al., 2008) with Tomofast-x to define geological structures and rules that inversion will be encouraged to follow. It also comprises the incorporation of topological laws previously used a posteriori (Giraud et al., 2019b) directly into inversion. The electrical capacitance tomography component of Tomofast-x (Martin et al., 2018), which we have not detailed here, can be extended to acoustic/seismic or electromagnetic data inversions that rely on the resolution of similar non-linear inversion problems. It opens the door to more versatility in the code and can be applied to joint inversion in similar ways but on more than two physical domains.

In addition, future developments comprise the collaborative and joint inversion of seismic and potential field data. It is planned to develop an interface between Tomofast-x and Unisolver (not yet open-source released by its authors), which is an extension of Seimic_Cpml codes (Komatitsch and Martin 2007, Martin et al., 2010, 2019) where integrated seismic imaging solvers are implemented. Unisolver is a multi-purpose 2D/3D seismic imaging platform based on high order finite difference and finite volume discretization as well as nonlinear seismic data inversion procedures. Such interface would allow performing collaborative or joint inversion of seismic and gravity or magnetic data and obtain the resulting models on the same mesh while benefitting from Tomofast-x's various functionalities. This will be an easy way to provide Tomofast-x with separate seismic information like sensitivity kernels on the fly as another physical domain.

In the implementation presented here, only the truncation of the matrix system based on maximum distance thresholding was discussed. It is planned to reduce memory requirements using the wavelet compression of the matrix system of the inverse problem in the same fashion as (Martin et al., 2013).

We have shown a number of tests using a selected set of functionalities of Tomofast-x. However, more, or different, tests could be done. For instance, an interesting research avenue is to exploit Tomofast-x's capability to read an arbitrary number of prior and starting models to test the geological archetypes that can be identified by clustering of the set of geological models probabilistic geological modelling can produce (Pakyuz-Charrier et al., 2019). Additional features of Tomofast-x which testing lies beyond the scope of this paper are Jacobian matrix truncating, the different kinds of depth weighting, and their effects on the different types of inversion. Last, we have not used posterior uncertainty indicators listed in Sect. 2.5 and A.1 as the paper focusses on the inversion capabilities of Tomofast. The output results of Tomofast-x allow, however, to study uncertainty in the same fashion as Giraud et al. (2017, 2019c) where some of them are used.

---

[1] https://www.loop3d.org/

770   Results obtained using the cross-gradient technique for joint inversion of gravity and magnetic data showed that it can improve imaging of geological structures. However, our study also revealed some of the limitations of this method. In the synthetic example, structurally coherent features of the resulting model contradict the geology of the true model. In addition, our L-curve (or L-surface) analysis suggests that the determination of the optimum α weights of the cost function using the cross-gradient technique may be affected by non-uniqueness and that

775   multiple sets of weights could equally satisfy the L-curve criterion. One interpretation is that this method remains affected by uncertainty and could be producing several families of models fitting geophysical data equally well. This observation differs from similar analysis performed in the case of joint inversion using petrophysical constraints, where such potential ambiguity was not suggested by the L-surfaces (Giraud et al., 2019c). These impressions, however, require a more detailed investigation and constitutes a new research avenue.

780   In our sensitivity analysis, we have produced a series of models that can be considered geophysically equivalent because they fit the geophysical data equally well. These models are the result of deterministic inversion, where prior information guides inversion towards one of the modes of the probability density function describing the problem (equation 1), or modifies them. It is therefore safe to assume that each mode is representative of an archetype of models from the geophysical data's null-space. This highlights the interest of using 'null-space

785   shuttles' allowing navigation of the null-space (Deal and Nolet 1996, Muñoz and Rath 2006, Vasco 2007, Fichtner and Zunino 2019) to explore the space of possible models without extensive sampling and to assess the robustness of the result. In addition, the plots of the L-curves corresponding to the problem we presented suggest the presence of multiple optima in the hyperparameter space (weights α), which it might be interesting to investigate in future research, especially in the joint inversion case.

790   **6 Conclusion**

We have introduced the open source joint inversion platform Tomofast-x and demonstrated its capabilities with a realistic dataset taken from the Hamersley region in central Western Australia. The geophysical theoretical background of Tomofast-x was explained in depth to guide users in understanding and using the modelling approach implemented in the source code.

795   We leveraged the modularity of Tomofast-x to study the sensitivity of inversion to prior structural, geological and petrophysical information, joint inversion, and the code's scalability. We tested a new combination of constraints incorporating geological structural information in the smoothness term and dynamic prior model definition using petrophysical knowledge (ADMM bound constraints), a feature usually not available to most inversion software. Our sensitivity analyses on prior information and different constraints reveal that constraints using petrophysics

800   (ADMM clustering bound constraints) dominate over gradient-based constraints (smoothness and cross-gradient constraints), which in turn exert more influence onto inversion than smallness constraints. This shows the importance of prior information in inversion and illustrates the need to study the space of geophysically equivalent models.

The examples described here were designed to replicate a typical, rigorous approach to the development of a

805   geoscientific model and be relevant to real-world application. The aim to ensure rigour and reproducibility of the

result presented is facilitated by the release of the source code, datasets, and a reduced, modified python version of the algorithm that accompany this paper.

7 Source code, documentation and data availability

The source code for Tomofast-x as used in this manuscript can be found in Ogarko et al. (2021b) (https://zenodo.org/record/4454220#.YFwpEK8zZaQ). The latest version of Tomofast-x is available at: https://github.com/TOMOFAST/Tomofast-x

The geological model, a description of the input data and the geophysical models are given in Jessell et al. (2021). It also contains a dataset using the same model projected onto a finer mesh of approximately 4,2M cells and 80,000 geophysical data. The datasets are licensed under the Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license (see https://creativecommons.org/licenses/by-sa/4.0/legalcode for details). Tomofast-x's source code is licensed under the MIT License (https://opensource.org/licenses/MIT).

8 Competing interests

The authors declare that they have no conflict of interest.

9 Acknowledgements

10 Authors' contribution

JG designed the geophysical study and ran all inversions shown in the paper. He adjusted the geological model presented here, which was initially built by MJ and ML. JG performed posterior analysis and interpretation of results. JG is the main contributor to the writing of this article and preparation of the assets. JG carried out the scaling tests shown in Appendix with the collaboration of VO.

VO and JG worked together on the implementation of the gravity and magnetic inversion methodologies in Tomofast-x. VO is the main developer of this version of Tomofast-x, the development of which was carried in collaboration with JG and RM. The main contributors to the code are VO, JG, and RM. JG performed the initial testing of the different data integration techniques presented to the exception of the cross-gradient technique, which was implemented independently by VO.

RM participated in the redaction of this paper. RM initiated the Tomofast project and implemented the initial LSQR solver for gravity inversion, which was subsequently used as a starting point for the current version of the code. RM and VO added the possibility of performing data inversion using an Lp norm ($1 < p \leq 2$) smallness term. RM and VO developed and tested the electrical capacitance tomography component of Tomofast.

JG and RM explored the functionalities of Tomofast-x and performed extensive testing for robustness and validation of techniques implemented, especially the cross-gradient technique.

ML and MJ produced the reference geological model from field measurements and carried out probabilistic geological modelling. ML and MJ were involved in the redaction of the manuscript and participated in the supervision of the project. MJ and ML have been involved in the validation of the methodology at the initial development stage and supervised the overall progress of the presented work.

All authors participated in the revisions of the manuscript.

Appendix A. Other functionalities of Tomofast-x

A.1   Posterior uncertainty indicators

A.1.1   Posterior LSQR variance matrix

At the first and last iteration of the inversion, the diagonal elements of the posterior covariance matrix of the recovered model is calculated in Tomofast-x. This calculation is performed following (Kostina et al., 2009) who introduce an extension of the LSQR algorithm where such matrix is calculated at each iteration. The variances are part of the outputs of Tomofast-x for further analysis by the user, such as the estimation of uncertainty in the recovered property model.

A.1.2   Jacobian of the cost function

Tomofast-x offers the possibility to examine the Jacobian matrix of the misfit function (eq. 5), which encapsulates the contribution of several constraint terms (see for example eq. 5), by calculating its derivative with respect to the model m, $\frac{\partial \theta(d,m)}{\partial m}$. This feature takes advantage of the LSQR solver. In the LSQR algorithm, $\frac{\partial \theta(d,m)}{\partial m}$ is calculated at the beginning of each iteration when approximating a solution to the system of equations (Appendix D). The value of $\frac{\partial \theta(d,m)}{\partial m}$ can then be calculated before or after application of the depth weighting operator. It is computed as the product of the transpose of the matrix representing the left-hand side of the system of equations to be solved by the LSQR solver, by the vector of the corresponding quantities to minimize (data misfit, cross-gradients, damping terms, etc., constituting the right hand side of the corresponding equation, as shown in Appendix D). Importantly, its dimension is equal to the number of model parameters. It is therefore possible to

store it on disk to provide a measure of the sensitivity of the data and the different terms of the misfit function to model variations at depth or in any part of the computational domain. By computing $\left\|\frac{\partial \theta(d,m)}{\partial m}\right\|$, it is possible to study the convergence of the algorithm, small values indicating convergence. In addition, it is a metric that measures the stability of the algorithm and which is useful to determine whether the system of equations is well conditioned.

A.1.3 Identification of rock units

Membership analysis of the inverted model can be performed when statistical petrophysical constraints were applied to inversion from the values of $\mathcal{N}_k$ reached after inversion converged. Membership values can be used to assess inverted models by reconstructing a rock unit model from the recovered inverted physical properties (Doetsch et al., 2010, Sun et al., 2012, Giraud et al., 2019c). Rock units labels can also be assigned to model-cells when the ADMM bound constraints have converged. It allows attaching a petrophysical property interval to each model-cell, allowing direct identification of rock types.

A.1.4 Cross-product of gradients

The cross-product of model gradients in 3D can be stored after inversion and its L_2 norm is given after each inversion cycle. It allows to assess the degree of structural similarity between the models, and to delineate areas showing specific structural similarities or dissimilarities.

A.2 Jacobian matrix truncation

Tomofast-x offers the possibility to use a moving sensitivity domain approach (Čuma et al., 2012, Čuma and Zhdanov 2014) limiting the sensitivity domain to a cylinder which radius is chosen by the user to reduce computational requirements (the option for a sphere is also present in the source code, but commented in the current version). The underlying assumption is that cells beyond a given distance exert a negligible influence on the measurement. Generally speaking, this radius is provided by the users and should be chosen carefully.

A.3 L_p norm

Tomofast-x also offers the possibility of performing data inversion using a L_p norm ($1 < p \leq 2$) to define the smallness term, as it has been proposed for electrical capacitance tomography (ECT) in (Martin et al., 2018) in the framework of Tomofast-x. The L_p norm inverse problem is non-linear and can be solved iteratively using L_2 minimization. In the L_p norm case, the regularization parameter can be approximated by a p-power law of the model at each point of the computational domain and must also be recomputed at each new inversion cycle. When the L_p norm is introduced for $p < 2$, this procedure allows to obtain sharper models with better interface definition, and determine stronger contrasts for the specific cases under study. If $p = 2$, the smallness term is reduced to L_2 norm minimization (the commonly-used Tykhonov-like regularization) as used in this work. The choice of other values such that $1 < p \leq 2$ is at the discretion of the user or depending on prior information.

A.4 Electrical capacitance tomography

Detailing electrical capacitance tomography (ECT) is beyond the scope of this paper, but we can apply to joint gravity and magnetic inversion the functionalities that have been introduced to solve the ECT inverse problem

based on L₂ data misfit norm and L_p (1<p≤2) damping term minimization (see Sect. A.3). In Tomofast-x, ECT is based on the finite-volume method for the forward problem and on a non-linear and iterative LSQR method to solve the inverse problem. As in propagative and diffusive geophysical inverse problems in frequency domain, the sensitivity matrix and the damping term depend on the current model and must be recomputed at each new iteration. We refer the reader to Martin et al. (2018) for more details on this technique. Note that algorithms developed in relation to this method can easily be extended to propagative and diffusive geophysical inverse problems.

Appendix B. Summary of the notation and terms used in the paper

915

Symbol	Definition
Subscripts and superscripts	
d	Refers to 'data': 'mag' or 'grav'.
m	model
pr	Refers to 'prior'
g	gradient
x	Cross-gradient
pe	Refers to 'petrophysics'
a	ADMM
G	Gravity
M	Magnetics
Model and physical quantities	
$\boldsymbol{m}$	Property model inverted for
$\boldsymbol{v}$	ADMM variable
$\boldsymbol{u}$	ADMM variable
ω	Membership value in Gaussian mixture
ψ	Membership value in Gaussian mixture (from geology)
$\boldsymbol{\mu}$	Mean value of petrophysical properties
$\boldsymbol{\sigma}$	Standard deviation of petrophysical properties
$\boldsymbol{z}$	ADMM variable
ε	Positive threshold real number such that z>> ε
n_{data}	Number of geophysical data points
n_f	Number of rock formations
$\boldsymbol{d}$	Geophysical data
β	Exponent for depth weight power law
Mathematical operators or notations	
$g(.)$	Geophysical forward operator
$P(\cdot)$	Petrophysical distribution operator

$\nabla \cdot$	Gradient operator
$\mathbf{L_p}$	L-p norm
$\mathbf{L_2}$	L-2 norm (sum-of-squares)
S	Geophysical data sensitivity matrix
Diagonal matrices $\boldsymbol{W}$	
$\boldsymbol{W_d}$	diagonal matrix; all elements equal to the sum-of-squares of the data
$\boldsymbol{W_m}$	Smallness term covariance matrix
$\boldsymbol{W_g}$	Smoothness term covariance matrix
$\boldsymbol{W_x}$	Cross-gradient term covariance matrix
$\boldsymbol{W_{pe}}$	Petrophysical term covariance matrix
$\boldsymbol{W_a}$	ADMM term covariance matrix
$\boldsymbol{D}$	Depth weighting operator
Weighting terms α	
α_m	model
α_g	Gradient
α_x	Cross-gradient
α_{pe}	petrophysics
α_a	Multiple bounds constraints
α^G	Weight assigned to the gravity inverse problem (used only in joint inversion)
α^M	Weight assigned to the magnetic inverse problem (used only in joint inversion)

Appendix C. Forward gravity and magnetic data calculation

In this Appendix we summarize the forward calculation of gravity and magnetic data as performed in Tomofast-x. In practice, Tomofast-x calculates forward data using input data expressed in units from *Système International* (SI), expressed in kilogram, metre, and second. The gravity field $\boldsymbol{f}$ of a distribution of density anomalies $\Delta\rho$ over a volume of rock V at a location $\boldsymbol{r'} = [x', y', z']$ can be expressed as follows:

$$\boldsymbol{f(r)} = G \iiint_V \Delta\rho(\boldsymbol{r}) \frac{\boldsymbol{r} - \boldsymbol{r'}}{|\boldsymbol{r} - \boldsymbol{r'}|^3} dV \tag{24}$$

Where $\boldsymbol{r} = [x, y, z]$ defines the location of mass density anonaly $\Delta\rho(\boldsymbol{r})$ and G is the universal gravity constant. While Tomofast-x is implemented in such a way that the three spatial components of $\boldsymbol{f}$ can be obtained, we consider only the vertical direction here, which we simply write f for sake of clarity (note that here, when taken for the whole model, $f = \boldsymbol{g_G}$). In our implementation, the volume V is descretized in N_m rectangular prisms (model cells) of constant density. Discretised, equation 26 then rewrites as:

$$f(x,y,z) = G \sum_{k=1}^{N_m} \Delta\rho_k \frac{z_k - z'}{((x - x_k)^2 + (y - y_k)^2 + (z - z_k)^2)^{\frac{3}{2}}} \Delta x_k \Delta y_k \Delta z_k, \tag{25}$$

where Δx_k, Δy_k and Δz_k define the dimensions of the kth rectangular prism. In our discretization, we assume a model constituted of $n_x \times n_y \times n_z$ cells, with n_x, n_y and n_z representing the number of cells in each direction. This leads to the computation of f using the following formulation of equation 26:

$$f(x,y,z) = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} \Delta\rho_{i,j,k} S_{i,j,k} \tag{26}$$

Where, using the formulation of (Blakely, 1995), the elements of the sensitivity matrix S are given as:

$$S_{i,j,k} = G \sum_{m=1}^{2} \sum_{q=1}^{2} \sum_{t=1}^{2} (-1)^{m+q+t} \left[\zeta_t \text{atan}\left(\frac{\xi_m \eta_q}{\zeta_t R_{m,q,t}}\right) - \xi_m \log(R_{m,q,t} + \eta_q) - \eta_q \log(R_{m,q,t} + \xi_m) \right], \tag{27}$$

where, ξ_m, η_q and ζ_t are the coordinates of the vertices of the prism, and

$$R_{m,q,t} = \left[(x - \xi_m)^2 + (y - \eta_q)^2 + (z - \zeta_t)^2 \right]^{\frac{1}{2}} \tag{28}$$

In Tomofast-x, the total magnetic field anomaly is calculated by summing the responses of the prisms making up the model, following Bhattacharyya (1964, 1980). The Regional magnetic field is denoted $F = (F_x, F_y, F_z)$, and the Magnetization $\boldsymbol{M} = (M_x, M_y, M_z)$. We write $F = \|\boldsymbol{F}\|$ and $M = \|\boldsymbol{M}\|$. Note that remnant magnetization is not accounted for.

Using the formalism of (Blakely, 1995), we denote ΔT the magnitude of the total magnetic field anomaly generated by a prism oriented parallel to the x, y, and z axes of the mesh similarly to the gravity case. We have:

$$\Delta T(x,y,z) = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} \chi_{i,j,k} S_{i,j,k} \tag{29}$$

where χ is the magnetic susceptibility. The sensitivity S is given as:

$$
\begin{aligned}
S_{i,j,k} = \mu_0 F \sum_{m=1}^{2} \sum_{q=1}^{2} \sum_{t=1}^{2} (-1)^{t+1} & \left[\frac{\alpha_{yz}}{2} \log\left(\frac{R_{m,q,t} - \xi_m}{R_{m,q,t} + \xi_m}\right) + \frac{\alpha_{xz}}{2} \log\left(\frac{R_{m,q,t} - \eta_q}{R_{m,q,t} + \eta_{q_m}}\right) \right. \\
& - \alpha_{xy} \log(R_{m,q,t} + \zeta_t) - M_x F_x \, \text{atan}\left(\frac{\xi_m \eta_q}{\xi_m^2 + R_{m,q,t}\zeta_t + \zeta_t^2}\right) \\
& \left. - M_y F_y \, \text{atan}\left(\frac{\xi_m \eta_q}{R_{m,q,t}^2 + R_{m,q,t}\zeta_t - \xi_m^2}\right) + M_z F_z \, \text{atan}\left(\frac{\xi_m \eta_q}{R_{m,q,t}\zeta_t}\right) \right],
\end{aligned}
\tag{30}
$$

where, $\xi_{m=1,2}$, $\eta_{q=1,2}$ and $\zeta_{t=1,2}$ are the coordinates of the vertices of the prism along the x, y, and z directions, respectively. The other terms of equation 30 are defined below:

$$\alpha_{xz} = F_x M_z + F_z M_z, \tag{31}$$

$$\alpha_{xy} = F_x M_y + F_y M_x, \tag{32}$$

$$\alpha_{yz} = F_y M_z + F_z M_y. \tag{33}$$

945   Scaling tests

Although Tomofast-x can run on personal computers in a few seconds or minutes for 2D inversions and small 3D volumes (typically a few minutes on a laptop for models smaller than approx. 100,000 model cells), it necessitates a supercomputer for realistic size 3D case studies (e.g., models exceeding 500,000 model cells and 10,000 geophysical data points).

950   We assess Tomofast-x's parallel efficiency using the strong scaling as an indicator. The strong scaling curve is given by plotting the number of CPUs as a function of user time. It is complemented by the relative speedup curve $t_1/t_{n_{cpu}}$, where t_1 and $t_{n_{cpu}}$ are respectively the user times to complete inversion using the number of CPUs $n_{cpu} = 1$ and a given number of CPU n_{cpu}, respectively. We performed the scaling tests on the EOS machine from the CALMIP supercomputing centre (https://www.top500.org/site/50539/, https://www.calmip.univ-

955   toulouse.fr/spip.php?article388 – the latter being in French only, both last accessed on 10/11/2020).

The full-size model we used is made of $N_x N_y N_z = 128 * 128 * 32 = 524,288$ cells (i.e., 2^{19} cells), which we reduce by a factor of 2 by reducing the physical domain incrementally to $N_x N_y N_z = 32 * 32 * 32 = 32,768$ cells (i.e., 2^{15} cells) to be able to use it on a single CPU, for the purpose of parallelization efficiency analysis. In the configuration we use, the number of data points modelled is equal to: $N_{data} = N_x N_y$.
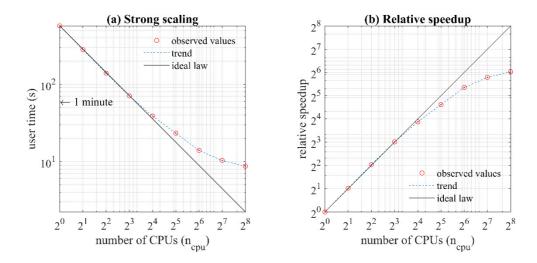

960

Figure A 1 - Strong scaling (a), relative speedup (b) and number of elements per CPU (c) plots for a number of cpus equal to 1, 2, 4, 8, 16, 32, 64, 128, 256, and 512. in (c), the intersection. between the two curves symbolises the number of elements below which computational resources usage is suboptimal. The line marked 'ideal law' indicates perfect scalability for the tests that were performed.

965   Figure A 1a shows the parallelization efficiency. It reveals that the scaling is nearly perfect for up to 16 CPUs, very good for 32 CPUs and that it deteriorates above 64 CPUs. This corresponds to relative speedups (ratio between run time for one CPU and a given number of CPUs) of about 14.5, 25 and 40 (Figure A 1b), respectively. For the cases using 64, 128 and 256 CPUs, speedup increases from 40 to 65, indicating that overhead inter-processor communication time for $n_{cpu} \geq 64$ increasingly impacts the total computation time, for this (small)

970   problem size. This is noticeable in Figure A 1a and Figure A 1b as both curves seem to adopt an asymptotic behaviour for the largest numbers of CPUs. This illustrates the deterioration of performances due to inter-processor communications (Kumar et al., 1994). The deterioration of performances due to inter-processor communications is due to the number of elements (or model cells) processed by each CPU becoming smaller, while the number of elements involved in communications increases; ultimately, the time spent in pure

975   computations in each core becomes smaller than the time spent in inter-processor communications.

The efficiency curves (e.g., Figure A 1a and Figure A 1b) allow us to determine the minimum number of elements per CPU that run efficiently (Hammond and Lichtner 2010). The case of $n_{cpu} = 64$ marks an inflection point in Figure A 1b, corresponding to point of diminishing return equal to a number of element per CPUs of 512. This indicates that for this particular configuration, it is preferable to run inversions with $n_{cpu} \leq 64$ to maximize

980   parallel efficiency. For better understanding and interpretation of scaling and speedup, we remind the number of elements n_{el} per CPU as a function of n_{cpu}:

$$n_{el} = N_x N_y N_z / n_{cpu}. \tag{34}$$

As a general rule, we recommend to respect the condition of $n_{el} \geq 512$. For smaller number of elements, the allocated resources are used in a suboptimum manner. Note that the memory requirements vary proportionally with $N_x N_y N_z N_{data} = n_m n_d$.

985   **Appendix D. Matrix formulation of least-squares problem and resolution of the inverse problem**

This appendix introduces the matrix formulation of equation 5 and its resolution.

We can write the system of equations to be solved in the least-squares sense as:

$$\begin{bmatrix} \boldsymbol{Sm} \\ \alpha_m \boldsymbol{W}_m \boldsymbol{m} \\ \alpha_g \boldsymbol{W}_g \nabla \boldsymbol{m} \\ \alpha_{pe} \boldsymbol{W}_{pe} \boldsymbol{P}(\boldsymbol{m}) \\ \alpha_x \boldsymbol{W}_x (\nabla \boldsymbol{m_1} \times \nabla \boldsymbol{m_2}) \\ \alpha_a \boldsymbol{W}_a \boldsymbol{m} \end{bmatrix} = \begin{bmatrix} \boldsymbol{d} \\ \alpha_m \boldsymbol{W}_m \boldsymbol{m}_{pr} \\ 0 \\ \alpha_{pe} \boldsymbol{W}_{pe} \boldsymbol{P}_{max} \\ 0 \\ \alpha_a \boldsymbol{W}_a (\boldsymbol{v} - \boldsymbol{u}) \end{bmatrix} \tag{35}$$

At iteration k, the system of equation is linearized around the current model. It is solved for the optimum update

990   of the current model $\boldsymbol{m^k}$ model update as described below. Models $\boldsymbol{m_1}$ and $\boldsymbol{m_2}$ are set accordingly with the type of inversion considered.

In Tomofast-x, depth weighting $\boldsymbol{D}$ is applied as a sensitivity matrix preconditioner. The resulting system is solved using the LSQR algorithm (Paige and Saunders 1982) as:

44

$$\begin{bmatrix} D^{-1}S \\ \alpha_m D^{-1}W_m \\ \alpha_g D^{-1}W_g \dfrac{\partial \nabla m_k}{\partial m} \\ \alpha_{pe} D^{-1}W_{pe}P'(m_k) \\ \alpha_x D^{-1}W_x \dfrac{\partial}{\partial m}(\nabla m_1 \times \nabla m_2) \\ \alpha_a D^{-1}W_a \end{bmatrix} \Delta \overline{m}_{k+1} = - \begin{bmatrix} (g(m^k) - d) \\ \alpha_m W_m(m_k - m_{pr}) \\ \alpha_m W_g \nabla m_k \\ \alpha_{pe} W_{pe}(P(m_k) - P_{max}) \\ \alpha_x W_x(\nabla m_k \times \nabla m_x) \\ \alpha_a W_a(m_k - z^k + u^k) \end{bmatrix} \qquad (36)$$

At each kth inversion cycle, we solve this system of equations and calculate the model update the model as follows:

$$\Delta m_{k+1} = D^{-1}\Delta \overline{m}_{k+1}. \qquad (37)$$

The model m^k can then be updated to obtain m^{k+1}.

$$m_{k+1} = m_k + \Delta m_{k+1}, \qquad (38)$$

Following Ozgarko et al. (2021a), $u^0 = 0$, $z^0 = 0$. The updated ADMM variables z^{k+1} and u^{k+1} are calculated using the ADMM algorithm introduced by Boyd (2010):

$$z^{k+1} = \pi_B(m_{k+1} + u^k), \qquad (39)$$

$$u^{k+1} = u^k + m_{k+1} - z^{k+1}, \qquad (40)$$

where π_B is a projection onto the bounds B such that:

$$\pi_B(x) = \left[\pi_{B_1}(x_1), \pi_{B_2}(x_2), \dots, \pi_{B_n}(x_n)\right], \text{ with} \qquad (41)$$

$$\pi_{B_i}(x_i) = \arg\min_{y \in B_i}\|x_i - y\|_2 \qquad (42)$$

The value of the starting model m^0 is provided by the user.

11 References

Ailleres, L., Jessell, M., de Kemp, E., Caumon, G., Wellmann, F., Grose, L., Armit, R., Lindsay, M., Giraud, J., Brodaric, B., Harrison, M. and Courrioux, G.: Loop - Enabling 3D stochastic geological modelling, ASEG Ext. Abstr., 2019(1), 1–3, doi:10.1080/22020586.2019.12072955, 2019.

Belge, M., Kilmer, M. E. and Miller, E. L.: Efficient determination of multiple regularization parameters in a generalized L-curve framework, Inverse Probl., 18(4), 314, doi:10.1088/0266-5611/18/4/314, 2002.

Bertsekas, D. P.: Nonlinear programming, Second Edition., 1999.

Bhattacharyya, B. K.: Magnetic anomalies due to prism-shaped bodies with arbitrary polarization, Geophysics, 29(04), 517–531, 1964.

Bhattacharyya, B. K.: A generalized multibody model for inversion of magnetic anomalies, GEOPHYSICS, 45(2), 255–270, doi:10.1190/1.1441081, 1980.

Bijani, R., Lelièvre, P. G., Ponte-Neto, C. F. and Farquharson, C. G.: Physical-property-, lithology- and surface-geometry-based joint inversion using Pareto Multi-Objective Global Optimization, Geophys. J. Int., 209(2), 730–748, doi:10.1093/gji/ggx046, 2017.

Blakely, R. J.: Potential Theory in Gravity and Magnetic Applications, Cambridge University Press., 1995.

Boyd, S.: Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, Found. Trends® Mach. Learn., 3(1), 1–122, doi:10.1561/2200000016, 2010.

Brown, V., Key, K. and Singh, S.: Seismically regularized controlled-source electromagnetic inversion, GEOPHYSICS, 77(1), E57–E65, doi:10.1190/geo2011-0081.1, 2012.

Calcagno, P., Chilès, J. P., Courrioux, G. and Guillen, A.: Geological modelling from field data and geological knowledge. Part I. Modelling method coupling 3D potential-field interpolation and geological rules, Phys. Earth Planet. Inter., 171(1–4), 147–157, doi:10.1016/j.pepi.2008.06.013, 2008.

Carter-McAuslan, A., Lelièvre, P. G. and Farquharson, C. G.: A study of fuzzy c -means coupling for joint inversion, using seismic tomography and gravity data test scenarios, GEOPHYSICS, 80(1), W1–W15, doi:10.1190/geo2014-0056.1, 2015.

Craven, P. and Wahba, G.: Smoothing noisy data with spline functions, Numer. Math., 31(4), 377–403, doi:10.1007/BF01404567, 1978.

Čuma, M. and Zhdanov, M. S.: Massively parallel regularized 3D inversion of potential fields on CPUs and GPUs, Comput. Geosci., 62, 80–87, doi:10.1016/j.cageo.2013.10.004, 2014.

Čuma, M., Wilson, G. A. and Zhdanov, M. S.: Large-scale 3D inversion of potential field data, Geophys. Prospect., 60(6), 1186–1199, doi:10.1111/j.1365-2478.2011.01052.x, 2012.

Deal, M. M. and Nolet, G.: Nullspace shuttles, Geophys. J. Int., 124(2), 372–380, doi:10.1111/j.1365-246X.1996.tb07027.x, 1996.

Doetsch, J., Linde, N., Coscia, I., Greenhalgh, S. a. and Green, A. G.: Zonation for 3D aquifer characterization based on joint inversions of multimethod crosshole geophysical data, GEOPHYSICS, 75(6), G53–G64, doi:10.1190/1.3496476, 2010.

Dykstra, R. L.: An Algorithm for Restricted Least Squares Regression, J. Am. Stat. Assoc., 78(384), 837–842, doi:10.1080/01621459.1983.10477029, 1983.

Eckhardt, E. A.: Partnership between geology and geophysics in prospecting for oil, GEOPHYSICS, 5(3), 209–214, doi:10.1190/1.1441804, 1940.

Farquharson, C. G. and Oldenburg, D. W.: A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems, Geophys. J. Int., 156(3), 411–425, doi:10.1111/j.1365-246X.2004.02190.x, 2004.

Fichtner, A. and Zunino, A.: Hamiltonian Nullspace Shuttles, Geophys. Res. Lett., 46(2), 644–651, doi:10.1029/2018GL080931, 2019.

Gallardo, L. a. and Meju, M. a.: Characterization of heterogeneous near-surface materials by joint 2D inversion of dc resistivity and seismic data, Geophys. Res. Lett., 30(13), doi:10.1029/2003GL017370, 2003.

1050 Gallardo, L. A. and Meju, M. A.: Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints, J. Geophys. Res. Solid Earth, 109(B3), B03311, doi:10.1029/2003JB002716, 2004.

Gil, Y., David, C. H., Demir, I., Essawy, B. T., Fulweiler, R. W., Goodall, J. L., Karlstrom, L., Lee, H., Mills, H. J., Oh, J.-H., Pierce, S. A., Pope, A., Tzeng, M. W., Villamizar, S. R. and Yu, X.: Toward the Geoscience Paper of the Future: Best practices for documenting and sharing research from data to software to provenance, Earth Sp.

1055 Sci., 3(10), 388–415, doi:10.1002/2015EA000136, 2016.

Giraud, J., Jessell, M., Lindsay, M., Parkyuz-Charrier, E. and Martin, R.: Integrated geophysical joint inversion using petrophysical constraints and geological modelling, in SEG Technical Program Expanded Abstracts 2016, pp. 1597–1601, Society of Exploration Geophysicists., 2016.

Giraud, J., Pakyuz-Charrier, E., Jessell, M., Lindsay, M., Martin, R. and Ogarko, V.: Uncertainty reduction

1060 through geologically conditioned petrophysical constraints in joint inversion, GEOPHYSICS, 82(6), ID19–ID34, doi:10.1190/geo2016-0615.1, 2017.

Giraud, J., Pakyuz-Charrier, E., Ogarko, V., Jessell, M., Lindsay, M. and Martin, R.: Impact of uncertain geology in constrained geophysical inversion, ASEG Ext. Abstr., 2018(1), 1, doi:10.1071/ASEG2018abM1_2F, 2018.

Giraud, J., Lindsay, M., Ogarko, V., Jessell, M., Martin, R. and Pakyuz-Charrier, E.: Integration of geoscientific

1065 uncertainty into geophysical inversion by means of local gradient regularization, Solid Earth, 10(1), 193–210, doi:10.5194/se-10-193-2019, 2019a.

Giraud, J., Ogarko, V., Lindsay, M., Pakyuz-Charrier, E., Jessell, M. and Martin, R.: Sensitivity of constrained joint inversions to geological and petrophysical input data uncertainties with posterior geological analysis, Geophys. J. Int., 218(1), 666–688, doi:10.1093/gji/ggz152, 2019b.

1070 Giraud, J., Lindsay, M., Jessell, M. and Ogarko, V.: Towards plausible lithological classification from geophysical inversion: honouring geological principles in subsurface imaging, Solid Earth, 11(2), 419–436, doi:10.5194/se-11-419-2020, 2020.

Guillen, A., Calcagno, P., Courrioux, G., Joly, A. and Ledru, P.: Geological modelling from field data and geological knowledge. Part II. Modelling validation using gravity and magnetic data inversion, Phys. Earth Planet.

1075 Inter., 171(1–4), 158–169, doi:10.1016/j.pepi.2008.06.014, 2008.

Hammond, G. E. and Lichtner, P. C.: Field-scale model for the natural attenuation of uranium at the Hanford 300 Area using high-performance computing, Water Resour. Res., 46(9), doi:10.1029/2009WR008819, 2010.

Hansen, P. C. and Johnston, P. R.: The L-Curve and its Use in the Numerical Treatment of Inverse Problems, in Computational Inverse Problems in Electrocardiography, pp. 119–142. [online] Available from:

1080 file://g/Publications/PDFs/hansen2001_Lcurve.pdf, 2001.

Hansen, P. C. and O'Leary, D. P.: The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems, SIAM J. Sci. Comput., 14(6), 1487–1503, doi:10.1137/0914086, 1993.

Heincke, B., Jegen, M., Moorkamp, M., Hobbs, R. W. and Chen, J.: An adaptive coupling strategy for joint inversions that use petrophysical information as constraints, J. Appl. Geophys., 136, 279–297, doi:10.1016/j.jappgeo.2016.10.028, 2017.

Hoerl, A. E. and Kennard, R. W.: Ridge Regression: Application to nonorthogonal problems, Technometrics, 12(1), 69–82, doi:10.1080/00401706.1970.10488634, 1970.

Jessell, M., Giraud, J. and Lindsay, M.: 3-D geological and petrophysical models with synthetic geophysics based on data from the Hamersley region (Western Australia), , doi:http://doi.org/10.5281/zenodo.4431796, 2021a.

Jessell, M. W., Ogarko, V., Lindsay, M. D., Joshi, R., Piechocka, G., Grose, L., Varga, M. de la, Ailleres, L. and Pirot, G.: Automated geological map deconstruction for 3D model construction, Geosci. Model Dev. Discuss. [preprint], doi:https://doi.org/10.5194/gmd-2020-400, 2021b.

Kamm, J., Lundin, I. A., Bastani, M., Sadeghi, M. and Pedersen, L. B.: Joint inversion of gravity, magnetic, and petrophysical data — A case study from a gabbro intrusion in Boden, Sweden, GEOPHYSICS, 80(5), B131–B152, doi:10.1190/geo2014-0122.1, 2015.

Komatitsch, D. and Martin, R.: An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation, Geophysics, doi:10.1190/1.2757586, 2007.

Kostina, E., Saunders, M. A. and Schierle, I.: Computation of Covariance Matrices for Constrainted Parameter Estimation Problems using LSQR, Stanford Optim. Lab Rep., 4026(BM2009-01), 1–11 [online] Available from: http://www.uni-marburg.de/fb12/forschung/berichte/berichtemathe/pdfbfm/bfm09-01, 2009.

Kumar, V., Grama, A., Gupta, A. and Karypis, G.: Introduction to Parallel Computing: Design and Analysis of Algorithms, Benjamin-Cummings Publishing Co., Inc., Redwood City, CA, USA., 1994.

De La Varga, M., Schaaf, A. and Wellmann, F.: GemPy 1.0: Open-source stochastic geological modeling and inversion, Geosci. Model Dev., 12(1), 1–32, doi:10.5194/gmd-12-1-2019, 2019.

Lelièvre, P., Farquharson, C. and Hurich, C.: Joint inversion of seismic traveltimes and gravity data on unstructured grids with application to mineral exploration, Geophysics, 77(1), K1–K15, doi:10.1190/geo2011-0154.1, 2012.

Lelièvre, P. G. and Farquharson, C. G.: Integrated Imaging for Mineral Exploration, in Integrated Imaging of the Earth: Theory and Applications, pp. 137–166., 2016.

Li, X. and Chouteau, M.: Three Dimensional Modeling in All Space, Surv. Geophys., 19(4), 339–368, doi:10.1023/A:1006554408567, 1998.

Li, Y. and Oldenburg, D. W.: 3-{D} inversion of magnetic data, Geophysics, 61(02), 394–408, 1996.

Li, Y. and Oldenburg, D. W.: 3-D inversion of gravity data, GEOPHYSICS, 63(1), 109–119, doi:10.1190/1.1444302, 1998.

Lindsay, M. D., Aillères, L., Jessell, M. W., de Kemp, E. A. and Betts, P. G.: Locating and quantifying geological uncertainty in three-dimensional models: Analysis of the Gippsland Basin, southeastern Australia,

Tectonophysics, 546–547, 10–27, doi:10.1016/j.tecto.2012.04.007, 2012.

Lindsay, M. D., Perrouty, S., Jessell, M. W. and Aillères, L.: Making the link between geological and geophysical uncertainty: geodiversity in the Ashanti Greenstone Belt, Geophys. J. Int., 195(2), 903–922, doi:10.1093/gji/ggt311, 2013.

Martin, R., Komatitsch, D., Gedney, S. D. and Bruthiaux, E.: A high-order time and space formulation of the unsplit perfectly matched layer for the seismicwave equation using auxiliary differential equations (ADE-PML), C. - Comput. Model. Eng. Sci., doi:10.3970/cmes.2010.056.017, 2010.

Martin, R., Monteiller, V., Komatitsch, D., Perrouty, S., Jessell, M., Bonvalot, S. and Lindsay, M. D.: Gravity inversion using wavelet-based compression on parallel hybrid CPU/GPU systems: application to southwest Ghana, Geophys. J. Int., 195(3), 1594–1619, doi:10.1093/gji/ggt334, 2013.

Martin, R., Ogarko, V., Komatitsch, D. and Jessell, M.: Parallel three-dimensional electrical capacitance data imaging using a nonlinear inversion algorithm and Lp norm-based model, Measurement, 128(May), 428–445, doi:10.1016/j.measurement.2018.05.099, 2018.

Martin, R., Bodet, L., Tournat, V. and Rejiba, F.: Seismic wave propagation in nonlinear viscoelastic media using the auxiliary differential equation method, Geophys. J. Int., doi:10.1093/gji/ggy441, 2019.

Martin, R., Giraud, J., Ogarko, V., Chevrot, S., Beller, S., Gégout, P. and Jessell, M.: Three-dimensional gravity anomaly inversion in the Pyrenees using compressional seismic velocity model as structural similarity constraints, Geophys. J. Int., doi:10.1093/gji/ggaa414, 2020.

McLachlan, G. and Peel, D.: Finite Mixture Models, Wiley Ser. Probab. Stat. New York Wiley, 419, doi:10.1198/tech.2002.s651, 2000.

Meju, M. A. and Gallardo, L. A.: Structural Coupling Approaches in Integrated Geophysical Imaging, pp. 49–67., 2016.

Moorkamp, M., Lelièvre, P. G., Linde, N. and Khan, A., Eds.: Integrated Imaging of the Earth, John Wiley & Sons, Inc, Hoboken, NJ., 2016a.

Moorkamp, M., Heincke, B., Jegen, M., Hobbs, R. W. and Roberts, A. W.: Joint Inversion in Hydrocarbon Exploration, in Integrated Imaging of the Earth: Theory and Applications, pp. 167–189., 2016b.

Muñoz, G. and Rath, V.: Beyond smooth inversion: the use of nullspace projection for the exploration of non-uniqueness in MT, Geophys. J. Int., 164(2), 301–311, doi:10.1111/j.1365-246X.2005.02825.x, 2006.

Nettleton, L. L.: Geophysics, geology and oil finding, GEOPHYSICS, 14(3), 273–289, doi:10.1190/1.1437535, 1949.

Ogarko, V., Giraud, J., Martin, R. and Jessell, M.: Disjoint interval bound constraints using the alternating direction method of multipliers for geologically constrained inversion: Application to gravity data, GEOPHYSICS, 86(2), G1–G11, doi:10.1190/geo2019-0633.1, 2021a.

Ogarko, V., Giraud, J. and Martin, R.: Tomofast-x v1.0 source code, , doi:10.5281/zenodo.4454220, 2021b.

Paasche, H. and Tronicke, J.: Cooperative inversion of 2D geophysical data sets: A zonal approach based on fuzzy c -means cluster analysis, GEOPHYSICS, 72(3), A35–A39, doi:10.1190/1.2670341, 2007.

Paige, C. C. and Saunders, M. A.: LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares, ACM Trans. Math. Softw., 8(1), 43–71, doi:10.1145/355984.355989, 1982.

1155   Pakyuz-Charrier, E., Lindsay, M., Ogarko, V., Giraud, J. and Jessell, M.: Monte Carlo simulation for uncertainty estimation on structural data in implicit 3-D geological modeling, a guide for disturbance distribution selection and parameterization, Solid Earth, 9(2), 385–402, doi:10.5194/se-9-385-2018, 2018.

Pakyuz-Charrier, E., Jessell, M., Giraud, J., Lindsay, M. and Ogarko, V.: Topological analysis in Monte Carlo simulation for uncertainty propagation, Solid Earth, 10(5), 1663–1684, doi:10.5194/se-10-1663-2019, 2019.

1160   Parsekian, A. D., Singha, K., Minsley, B. J., Holbrook, W. S. and Slater, L.: Multiscale geophysical imaging of the critical zone, Rev. Geophys., 53(1), 1–26, doi:10.1002/2014RG000465, 2015.

Peng, R. D.: Reproducible Research in Computational Science, Science (80-.)., 334(6060), 1226–1227, doi:10.1126/science.1213847, 2011.

Portniaguine, O. and Zhdanov, M. S.: 3-D magnetic inversion with data compression and image focusing, 1165   GEOPHYSICS, 67(5), 1532–1541, doi:10.1190/1.1512749, 2002.

Rashidifard, M., Giraud, J., Ogarko, V., Jessell, M. and Lindsay, M.: Cooperative Inversion of Seismic and Gravity Data Using Weighted Structure-Based Constraints, in NSG2020 3rd Conference on Geophysics for Mineral Exploration and Mining, pp. 1–5, European Association of Geoscientists & Engineers., 2020.

Ren, Z. and Kalscheuer, T.: Uncertainty and Resolution Analysis of 2D and 3D Inversion Models Computed from 1170   Geophysical Electromagnetic Data, Surv. Geophys., (0123456789), doi:10.1007/s10712-019-09567-3, 2019.

Rudin, L. I., Osher, S. and Fatemi, E.: Nonlinear total variation based noise removal algorithms, Phys. D Nonlinear Phenom., 60(1–4), 259–268, doi:10.1016/0167-2789(92)90242-F, 1992.

Santos, E. T. F. and Bassrei, A.: L- and Θ-curve approaches for the selection of regularization parameter in geophysical diffraction tomography, Comput. Geosci., 33(5), 618–629, doi:10.1016/j.cageo.2006.08.013, 2007.

1175   Scheidt, C., Li, L. and Caers, J.: Quantifying Uncertainty in Subsurface Systems, John Wiley & Sons, Inc., Hoboken, NJ, USA., 2018.

De Stefano, M., Golfré Andreasi, F., Re, S., Virgilio, M. and Snyder, F. F.: Multiple-domain, simultaneous joint inversion of geophysical data with application to subsalt imaging, Geophysics, 76(3), R69–R80, doi:10.1190/1.3554652, 2011.

1180   Sun, J. and Li, Y.: Geophysical inversion using petrophysical constraints with application to lithology differentiation SEG San Antonio 2011 Annual Meeting SEG San Antonio 2011 Annual Meeting, pp. 2644–2648., 2011.

Sun, J. and Li, Y.: Multidomain petrophysically constrained inversion and geology differentiation using guided fuzzy c-means clustering, Geophysics, 80(4), ID1–ID18, doi:10.1190/geo2014-0049.1, 2015.

1185   Sun, J. and Li, Y.: Joint inversion of multiple geophysical data using guided fuzzy c -means clustering, GEOPHYSICS, 81(3), ID37–ID57, doi:10.1190/geo2015-0457.1, 2016.

Sun, J. and Li, Y.: Joint inversion of multiple geophysical and petrophysical data using generalized fuzzy clustering algorithms, Geophys. J. Int., 208(2), 1201–1216, doi:10.1093/gji/ggw442, 2017.

Sun, J., Li, Y. and Studies, M.: Joint inversion of multiple geophysical data : A petrophysical approach using
1190   guided fuzzy c-means clustering SEG Las Vegas 2012 Annual Meeting SEG Las Vegas 2012 Annual Meeting, , (4), 1–5, 2012.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics., 2005.

Tarantola, A. and Valette, B.: Inverse Problems = Quest for Information, J. Geophys., 50(3), 159–170,
1195   doi:10.1038/nrn1011, 1982.

Tikhonov, A. N. and Arsenin, V. Y.: Solutions of Ill-Posed Problems, John Wiley, New York., 1977.

Tikhonov, A. N. and Arsenin, V. Y.: Solution of Ill-Posed Problems, Math. Comput., 32(144), 491, doi:10.2307/2006360, 1978.

Vasco, D. W.: Invariance, groups, and non-uniqueness: The discrete case, Geophys. J. Int., 168(2), 473–490,
1200   doi:10.1111/j.1365-246X.2006.03161.x, 2007.

Wegener, A.: Die Entstehung der Kontinente und Ozeane, Geogr. J., 61(4), 304, doi:10.2307/1781265, 1923.

Wellmann, F. and Caumon, G.: 3-D Structural geological models: Concepts, methods, and uncertainties, edited by C. Schmelzback, pp. 1–121, Cambridge, Massachusetts., 2018.

Wiik, T., Nordskag, J. I., Dischler, E. Ø. and Nguyen, A. K.: Inversion of inline and broadside marine controlled-
1205   source electromagnetic data with constraints derived from seismic data, Geophys. Prospect., 63(6), 1371–1382, doi:10.1111/1365-2478.12294, 2015.

Yan, P., Kalscheuer, T., Hedin, P. and Garcia Juanatey, M. A.: Two-dimensional magnetotelluric inversion using reflection seismic data as constraints and application in the COSC project, Geophys. Res. Lett., 44(8), 3554–3563, doi:10.1002/2017GL072953, 2017.

1210   Zhang, J. and Revil, A.: 2D joint inversion of geophysical data using petrophysical clustering and facies deformation, Geophysics, 80(5), M69–M88, 2015.

Zhdanov, M. S. and Lin, W.: Adaptive multinary inversion of gravity and gravity gradiometry data, GEOPHYSICS, 82(6), G101–G114, doi:10.1190/geo2016-0451.1, 2017.

Zhdanov, M. S., Gribenko, A. and Wilson, G.: Generalized joint inversion of multimodal geophysical data using
1215   Gramian constraints, Geophys. Res. Lett., 39(9), L09301, doi:10.1029/2012GL051233, 2012.