Geologically constrained geometry inversion and null-space navigation to explore alternative geological scenarios: a case study in the Western Pyrenees

Jérémie Giraud, Mary Ford, Guillaume Caumon, Vitaliy Ogarko, Lachlan Grose, Roland Martin, Paul Cupillard

Geophysical Journal International

Geophys. J. Int. (2024) **239,** 1359–1379
Advance Access publication 2024 June 07
GJI Geodynamics and Tectonics

https://doi.org/10.1093/gji/ggae192

Geologically constrained geometry inversion and null-space navigation to explore alternative geological scenarios: a case study in the Western Pyrenees

Jérémie Giraud [1,2], Mary Ford [3], Guillaume Caumon [1,4], Vitaliy Ogarko [2,5], Lachlan Grose [6], Roland Martin [7] and Paul Cupillard [1]

[1] *GeoRessources, Université de Lorraine, CNRS,* 54500 *Vandoeuvre-lès-Nancy, France. E-mail: jeremie.giraud@uwa.edu.au*
[2] *Centre of Exploration Targeting (School of Earth Sciences), University of Western Australia, Crawley, WA* 6009, *Australia*
[3] *Université de Lorraine, CNRS, CRPG,* 54000 *Nancy, France*
[4] *Institut Universitaire de France (IUF), 75000 Paris, France*
[5] *Mineral Exploration Cooperative Research Centre, The University of Western Australia, Crawley, WA 6009, Australia*
[6] *School of Earth, Atmosphere and Environment, Monash University, VIC* 3800, *Australia*
[7] *Géoscience Environnement Toulouse, Observatoire Midi-Pyrénées, Université de Toulouse, 31400 Toulouse, France*

Accepted 2024 May 1. Received 2024 May 1; in original form 2023 December 4

SUMMARY

Reducing the gap between geophysical inversion and geological interpretation can be achieved by integrating geological modelling into geophysical inversion. For this, we use a generalized, iterative level-set gravity inversion scheme in which geological units are deformed automatically. During the inversion process, a regularization term is defined using automated geological modelling to account for geological data and principles. This provides model-dependent geological constraints and encourages geological realism throughout inversion. To alleviate the dependence on the starting model and consider the possibility of features unseen by direct observations, an automated geophysical data-driven method is proposed to insert new rock units in the model. Uncertainty quantification is achieved through the null-space shuttle algorithm, which is used to generate a series of alternative models that are consistent with geophysical data. This methodology is applied to assess the uncertainties of a pre-existing 3-D crustal-scale geological model of the Western Pyrenean orogeny (France, Spain). The area is characterized by a positive gravity anomaly generally attributed to the presence of a shallow mantle body. The impact of variations in shape and density of key crustal and mantle features is investigated. Different scenarios are explored in 3-D space to produce a range of viable, relatively simple crustal-scale models of the area. This application demonstrates the capability and potential of this approach to evaluate alternative interpretations of geophysical data. The results show the plausibility of scenarios with a shorter subducted Iberian lower crust and a denser Axial Zone than in the pre-existing model.

Key words: Composition and structure of the continental crust; Gravity anomalies and Earth structure; Inverse theory; Numerical modelling.

1 INTRODUCTION

Geological interpretation of geophysical data is central to geoscientific discovery but is prone to uncertainty owing to the ill-posedness of geophysical inverse problems. To better quantify interpretation uncertainties, several modelling techniques have recently been introduced to incorporate user interpretation into the inversion process, be it after the completion of the inverse modelling procedure or through the introduction of constraints during inversion. In potential field inversion, post-inversion methods rely on assigning the different parts of the model to different rock units (see Giraud *et al.* 2020; Li *et al.* 2019; Melo *et al.* 2017, and references therein). Alternatively, the use of clustering algorithms during inversion (e.g. Carter-McAuslan *et al.* 2015; Giraud *et al.* 2017, 2019; Maag & Li 2018; Paasche & Tronicke 2007; Sun & Li 2016), where the inverted property values are encouraged to cluster around discrete values known *a priori*, or to get as near to said values as desired (Ogarko *et al.* 2021), allows direct labelling of rock units in the model. These techniques still suffer from uncertainty as the assignment of a volume to a given cluster, and hence to a given rock type,

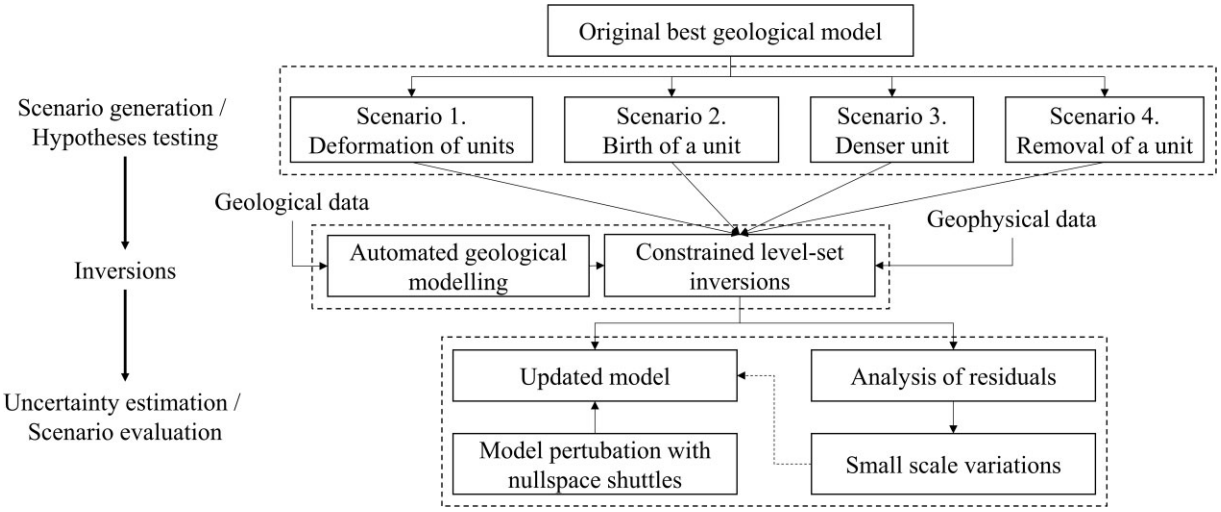

Figure 1. Workflow summary and scenario evaluation.

may remain ambiguous. In addition, they can yield geologically implausible juxtapositions between different rock units, and the use of clustering constraints leads to additional computational cost.

An alternative approach that is free of some of these limitations is presented by discrete-valued inversion, through the use of explicit surfaces (Galley *et al.* 2020, 2021; Lu *et al*. 2024) or level-set functions to define the location of interfaces between rock units of known physical properties during the inversion process (Giraud *et al.* 2021a; Kadu *et al.* 2016; Li *et al.* 2014, 2017, 2020; Liu *et al.* 2018; Muir & Tsai 2019; Rashidifard *et al.* 2021, 2024; Zheglova *et al.* 2018, 2013; Zheglova & Farquharson 2016). The application of level-set functions in geophysical inversion is relatively recent. In exploration settings with more than three rock units, multibody geometrical inversion (e.g. Giraud *et al.* 2021a) may be an alternative to the tedious process of forward modelling and manual adjustment of rock unit interfaces to fit geophysical data. It may also alleviate the ambiguity arising when relying solely on geological data to model undersampled subsurface features. However, like all other geophysical level set inversions, this approach may compromise the geological validity of the proposed models. A strategy to mitigate this limitation was proposed by Giraud *et al.* (2022, 2024), who calculate the geological projection of the inversion's current model to use it as constraint during inversion: at each iteration of the inversion, an implicit geological modelling procedure calculates a geologically acceptable model using the geophysical inversion's current model together with available geological measurements and principles. In this work, we use this geologically plausible model as a dynamic prior model to constrain level set inversion, thereby reducing the risk of producing geologically unrealistic outcomes. In what follows, we introduce essential elements of the proposed inversion algorithm (Section 2.1) and present an application example using crustal-scale gravity data (Section 5).

Alternative scenarios can be generated by perturbing the density of rock units for a given structural geometry or by adding new rock units into, or remove units from, the model, changing the dimensionality of the level-set inverse problem. Using a geophysical data-guided process, new rock units can be automatically inserted in specific regions of the geological model to test scenarios using geophysical inversion (Section 3). To complement this, Section 4 investigates the possibility of removing features or rock units from the model a posteriori while maintaining geophysical data misfit to a predefined level using the null-space shuttle algorithm (Fichtner & Zunino 2019), which allows generating intermediate models fitting the data.

To show the applicability of the proposed methods, we select an area of the Western Pyrenean orogen centred on the Mauléon Basin (Section 5). While the area has been well studied and the related body of literature is significant (see for instance the recent special issue in BSGF—Earth Science Bulletin, https://www.bsgf.fr/component/toc/?task=topic&id=1264, last accessed on 27/05/2024), there remain unanswered questions and debate regarding the orogen's deep structure in this area. Some geological bodies that have been interpreted from indirect measurements therefore merit further investigation. For instance, to date, there is no consensus on the 3-D morphology of the Moho discontinuity in the Western Pyrenees around and in relation to the strong Labourd-Mauléon gravity anomaly (Casas *et al.* 1997). From the review of Chevrot *et al.* (2022), 'Constraining density [. . .] should allow us to further constrain the composition of deep crustal materials'. Building on previous studies, we propose an application case focusing on the modelling of the Labourd gravity anomaly using the proposed integrated inversion workflow and methods. The purpose of this paper being to present and apply new techniques, this data set provides an opportunity to critically appraise existing geological models based on geophysical data and, where appropriate, to explore the impact of additional elements pertaining to the structure and evolution of the crust in the Western Pyrenees. The input crustal-scale 3-D model of the Western Pyrenees of Wehr (2017) (which slightly differs from that published in Wehr *et al.* 2018) is used here to test the techniques mentioned above and as a starting point to generate other plausible scenarios. We investigate and evaluate:

(i) The deformation of the different rock units to fit the gravity data with geometrical inversion.
(ii) The addition (or birth) of a rock unit in the model prior to inversion.
(iii) The removal (or death) of a rock unit in the model prior to inversion.
(iv) The increase in density of selected units prior to inversion.
(v) Perturbations of the model after inversion to evaluate uncertainty.

The proposed workflow applied to the Pyrenees case study is summarized in Fig. 1.

2 GEOMETRICAL INVERSION

We hypothesize that, at the first order, only a small number of geological units are required to explain most of the signal contained in geophysical data. For this reason, we use a level set inversion to optimize the shape of the rock units populated with discrete values. However, this approach might fail in some cases to capture all the information about rock unit properties present in the data. Therefore, we retain the possibility of performing continuous inversion to invert the residuals in each rock unit to obtain low amplitude petrophysical property variations.

2.1 Generalized level set inversion

The method we propose is an extension of Giraud *et al.* (2022, 2024), who use implicit geological modelling to ensure the geological consistency of inverted models. Fundamental to level set inversion is the consideration of signed distances to interfaces as the quantity linking the outline of rock units to geophysical data. We consider a set of signed distances $\boldsymbol{\phi} = \{\boldsymbol{\phi}_k,\ k = 1, \ldots, n_r\}$ corresponding to the boundaries of n_r rock units, such that:

$$\boldsymbol{\phi}_k \begin{cases} > 0 \textit{ inside rock } k \\ = 0 \textit{ at interfaces} \\ < 0 \textit{ outside rock } k, \end{cases} \tag{1}$$

which we calculate using the Fast Marching Method (FMM) of Sethian (1996) as implemented in the Python extension module scikit-fmm (Furtney 2015). From the signed-distance values $\boldsymbol{\phi}$, a transform is applied to obtain a physical property model (e.g. density values):

$$\boldsymbol{m}\left(\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{n_r}\right) = \sum_{i=1}^{n_r} v_i H\left(\boldsymbol{\phi}_i\right) \left[\prod_{j=1,\ j\neq i}^{n_r} \left(1 - H\left(\boldsymbol{\phi}_j\right)\right)\right], \tag{2}$$

where $\boldsymbol{v} \in \mathbb{R}^{n_r}$ is a vector containing the physical property value assigned to each of the n_r rock units modelled; H is the smeared-out Heaviside function, which we calculate following Osher & Fedkiw (2003, p. 15).

We formulate the inverse problem using a least-squares framework, solving for the update of the signed-distance, $\delta\boldsymbol{\phi} = \boldsymbol{\phi}^n - \boldsymbol{\phi}^{n-1}$, between two iterations. This inversion combines a geophysical data term, a term minimizing the difference between the geophysical level set solution and the closest implicit geological model, and a regularization term. Formally, the cost function for the nth iteration reads:

$$\begin{aligned} \Psi\left(\boldsymbol{\phi}^n,\ \boldsymbol{d}^{geophy}, \boldsymbol{d}^{geol}\right) &= \| \boldsymbol{W}_d^{-1}\left(\boldsymbol{d}^{geophy}_{obs.} - \boldsymbol{d}^{geophy}_{calc.}\left(\boldsymbol{\phi}^n\right)\right)\|_2^2 \\ &+ \| \boldsymbol{W}_p^{-1}\left(\boldsymbol{\phi}^n - \boldsymbol{\phi}^{geol}\left(\boldsymbol{d}^{geol}_{obs.}, \boldsymbol{d}^{geol}_{calc.}\left(\boldsymbol{\phi}^{n-1}\right)\right)\right)\|_2^2 \\ &+ \| \boldsymbol{B}_I\left(\boldsymbol{\phi}^n - \boldsymbol{\phi}^{n-1}\right)\|_2^2, \end{aligned} \tag{3}$$

where $\boldsymbol{d}^{geophy}_{obs.}$ and $\boldsymbol{d}^{geophy}_{calc.}$, are the observed and calculated gravity data, respectively; $\boldsymbol{\phi}^{geol}$ contains signed distances to the geological interfaces corresponding to the implicit geological model closest to $\boldsymbol{\phi}^{n-1}$, given the geological field data $\boldsymbol{d}^{geol}_{obs.}$ and geological information extracted from $\boldsymbol{\phi}^{n-1}$ contained in $\boldsymbol{d}^{geol}_{calc.}(\boldsymbol{\phi}^{n-1})$; $\boldsymbol{W}_d$ is a diagonal data covariance matrix. Here, it is set as the identity matrix for simplicity; and $\boldsymbol{W}_p$ is a diagonal covariance matrix controlling the strength of the constraints assigned to each model-cell. It can be set to account for prior geological or geophysical information. Here, $\boldsymbol{W}_p$ is set with values that are higher for the shallower model cells and decrease with depth. The reader is referred to Rashidifard *et al.* (2021, 2024) and Giraud *et al.* (2024), for more details about the use of prior information such as seismic interpretations in $\boldsymbol{W}_p$. $\boldsymbol{B}_I$ is a block identity matrix of dimensions $n_m \times (n_m n_r)$. The third term in eq. (3) is a regularization term introduced here. During inversion, it encourages minimizing $\| \sum_{i=1}^{n_r} \delta\boldsymbol{\phi}_i \|_2^2$ at interfaces between two iterations. This corresponds to imposing the conservation of the sum of signed distances at the interface between units and contributes to stabilizing the inversion.

Note that while the distance to interfaces between rock units is inverted, eq. (3) is similar to the 'traditional' least-squares formulation of inverse problems considering continuous property inversion. To solve eq. (3), we use the LSQR algorithm (Paige & Saunders 1982) as implemented in SciPy (Virtanen *et al.* 2020). Inversion stops when the data misfit reaches a prescribed value or when model changes between two iterations become negligible.

In what follows, the geophysical data misfit ψ^d is $\psi^d = \| \boldsymbol{d}^{geophy}_{obs.} - \boldsymbol{d}^{geophy}_{calc.} \|_2^2$. The forward calculation is performed using a forward operator g applied to model $\boldsymbol{m} = \boldsymbol{m}(\boldsymbol{\phi})$ obtained from $\boldsymbol{\phi}$ using eq. (2):

$$\boldsymbol{d}^{geophy}_{calc.} = \boldsymbol{d}^{geophy}_{calc.}(\boldsymbol{m}) = g(\boldsymbol{m}). \tag{4}$$

In the case of gravity data, $\boldsymbol{S}^m$ is introduced as the sensitivity matrix of the gravity data $\boldsymbol{d}$ to changes in densities following the relation

$$\boldsymbol{d}^{geophy}_{calc.}(\boldsymbol{m}) = g(\boldsymbol{m}) = \boldsymbol{S}^m \boldsymbol{m}, \tag{5}$$

and define a data misfit term:

$$\psi^d = \psi^d\left(\boldsymbol{d}^{geophy}_{obs.},\ \boldsymbol{m}\right) = \| \boldsymbol{d}^{geophy}_{obs.} - \boldsymbol{S}^m \boldsymbol{m} \|_2^2. \tag{6}$$

In eq. (3), at any given iteration n of the inversion, $\boldsymbol{\phi}^{geol}$ is calculated as the geological projection of the signed distance model $\boldsymbol{\phi}^{n-1}$. For this, contact location and orientation data are extracted and stored in $\boldsymbol{d}^{geol}_{calc.}$. Geological data in $\boldsymbol{d}^{geol}_{obs.}$ and $\boldsymbol{d}^{geol}_{calc.}$ are then fed to an implicit geological modelling code. From this, the closest geological model honouring these contacts together with geological data and knowledge encapsulated in $\boldsymbol{d}^{geol}$ (which include orientation data, foliations, etc.), is calculated. The resulting implicit geological model $\boldsymbol{\phi}^{geol}$ accounts for relationships between units (e.g. intrusion, erosion), known locations between contacts and orientation data from both the geological field measurements and information recovered from the geophysical level set model $\boldsymbol{\phi}^{n-1}$. The calculation of the geological model is performed using the implicit geological modelling code LoopStructural (Grose *et al.* 2021). For more information about this process, the reader is referred to Giraud *et al.* (2022, 2024).

To accelerate convergence and to mitigate possible trade-offs in the objective function (eq. 3), we use a line search algorithm (see Nocedal & Wright 2006, for details) modulating the value of the update $\delta\boldsymbol{\phi}$, with a scalar β. At each iteration of the inversion, the value of $\boldsymbol{m}(\boldsymbol{\phi} + \beta_i \delta\boldsymbol{\phi})$ is calculated for $i = 1, ., n_\beta$ (for β ranging between 0.1 and 3 and $n_\beta = 30$ in our experiments) and the update is scaled by the value β_{best} minimizing the geophysical misfit term:

$$\beta_{best} = \arg\min_\beta \psi^d\left(\boldsymbol{d}^{geophy}_{obs.},\ \boldsymbol{m}(\boldsymbol{\phi} + \beta_i \delta\boldsymbol{\phi})\right). \tag{7}$$

2.2 Small-scale variations from the residuals

In cases where the deformation of constant density rock units may be insufficient to explain the geophysical data, we propose to complement the above level-set inversions with continuous value property inversion.

Consider the residuals $\boldsymbol{r} = \boldsymbol{d}_{obs.}^{geophy} - \boldsymbol{d}_{calc.}^{geophy}(\boldsymbol{m}(\boldsymbol{\phi}))$ as the data to invert. This is similar to removing a known background model approximating the subsurface to invert for departures from this model. Here, the background model is the model recovered from level set inversion.

We obtain continuous, small-scale variations $\boldsymbol{m}_{cont}$ after level set inversion assuming that densities fitting the gravity data can be modelled using:

$$\tilde{\boldsymbol{m}} = \boldsymbol{m}(\boldsymbol{\phi}) + \boldsymbol{m}_{cont}, \tag{8}$$

such that the data misfit $\psi^d(\boldsymbol{d}_{obs.}^{geophy}, \tilde{\boldsymbol{m}})$ is reduced as compared to that of the piecewise constant model (eq. 2). In this work, the higher-order density updates $\boldsymbol{m}_{cont}$ are obtained by continuous property value inversion using the Tomofast-x open-source inversion code (Giraud *et al.* 2021b; Ogarko *et al.* 2024). To make sure these local density updates are petrophysically and geologically consistent, bound constraints and smoothing regularization are applied to limit density variations in $\boldsymbol{m}_{cont}$ using the disjoint interval bound constraints as detailed in Ogarko *et al.* (2021).

3 INSERTING AND REMOVING ROCK UNITS IN A PRE-EXISTING MODEL

3.1 Automated calculation of optimum rock unit densities

As an alternative to the above gradual density update, another strategy to minimize geophysical data misfit involves changing the constant density values $v_{i=1,..,n_r}$ inside the n_r rock units. Whereas these values are often chosen by manual tuning and with expert knowledge, this process can be automated by reparametrizing the model by the number of rock units n_r (and not by the number of model cells n_m as in Section 2.2), and then solving for the optimum density vector $\boldsymbol{v} = \{v_{i=1,..,n_r}\}$. Considering a piecewise constant density model populated with values from $\boldsymbol{v}$, we have

$$\boldsymbol{d}_{calc.}^{geophy} = \boldsymbol{S}^v \boldsymbol{v}, \tag{9}$$

with

$$\boldsymbol{S}^v = \frac{\partial \boldsymbol{d}}{\partial \boldsymbol{v}} = \frac{\partial \boldsymbol{d}}{\partial \boldsymbol{m}} \frac{\partial \boldsymbol{m}}{\partial \boldsymbol{v}} = \boldsymbol{S}^m \frac{\partial \boldsymbol{m}}{\partial \boldsymbol{v}}. \tag{10}$$

$\boldsymbol{S}^v$ is a matrix of dimensions $n_d \times n_r$ (n_d the number of data, and n_r the number of rock units). In the general case, $n_r \ll n_m$ and the corresponding system of equations is small. Injecting eq. (9) into eq. (3), we obtain:

$$\psi^d\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{v}\right) = \|\boldsymbol{d}_{obs.}^{geophy} - \boldsymbol{d}_{calc.}^{geophy}\|_2^2 = \|\boldsymbol{d}_{obs.}^{geophy} - \boldsymbol{S}^v \boldsymbol{v}\|_2^2. \tag{11}$$

Minimizing $\psi^d(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{v})$ in a constrained least-squares framework, prior petrophysical information can again be infused through bound constraints to maintain the densities $\boldsymbol{v}$ within realistic ranges (e.g. with values $\boldsymbol{v}_{min}$ and $\boldsymbol{v}_{max}$ to allow ±5 per cent density variations). We obtain $\boldsymbol{v}$ solving:

$$\min \psi^d\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{v}\right) \text{ subject to } \boldsymbol{v}_{min} \leq \boldsymbol{v} \leq \boldsymbol{v}_{max} \tag{12}$$

using the least-squares solver ('lsq_linear') implemented in SciPy Python library (Branch *et al.* 1999). Note that this algorithm can be applied to the optimization of all or only selected units by restricting variations of $\boldsymbol{v}$ by $\boldsymbol{v}_{min}$ and $\boldsymbol{v}_{max}$.

3.2 Adding or removing rock units

The above methodologies optimize the geometry and the density values of geological units present in the reference geological model either at the grid cell level (Section 2.2) or the rock unit level (Section 3.1). However, the interpretation ambiguities and the general quest for parsimonious models may lead to underestimating the number of rock units actually present in the subsurface (Wellmann & Caumon 2018). For instance, an intrusive unit not observed in geological data may be unravelled by geophysical data. To account for this, we designed a geophysical data-driven framework to automatically insert new geological units in an existing 3-D geological model.

The birth of rock units is based on the following approach (Fig. 2). We first identify regions of the model that present the most potential for geophysical data misfit reduction. For this, we calculate the sensitivity $\nabla\psi^d$ of the data misfit term ψ^d to changes in density:

$$\nabla\psi^d\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m}\right) = \frac{\partial \psi^d\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m}\right)}{\partial \boldsymbol{m}} = -2\boldsymbol{S}^{m\,T}\left(\boldsymbol{d}_{obs.}^{geophy} - \boldsymbol{S}^m \boldsymbol{m}\right). \tag{13}$$

The model cells whose sensitivity norms $|\nabla\psi^d|$ are superior to a selected threshold $\tau_{\nabla\psi^d}$ then define a subdomain

$$\Omega = \mathbf{1}_{|\nabla\psi^d| \geq \tau_{\nabla\psi^d}} \tag{14}$$

with $\mathbf{1}$ representing the indicator function. Because the subdomain Ω may consist of several parts, we detect face-adjacent connected components using the *cc3d* Python library of Silversmith (2021). This way, Ω is divided into n_{cc} subdomains of connected model cells such that $\Omega = \{\Omega_1, \ldots, \Omega_{n_{cc}}\}$, as exemplified in Fig. 2(a). Note that the smallest or least significant of these connected components may be filtered out.

Finally, we calculate n_{cc} models $\boldsymbol{m}_{birth}^{i=1,..,\,n_{cc}}$ where the density assigned to each subdomain $\Omega_{i=1,..,\,n_{cc}}$ is obtained solving eq. (12) while the rest of the model remains unchanged.

The misfit reduction incurred by adjusting the density within the selected clusters (e.g. $\Omega_{1,..,4}$ in Fig. 2a) is used to rank the new units' individual contribution to improving the data fit. Here, we select the case allowing the largest reduction of ψ^d as the optimum subdomain for the birth of a rock unit. Formally, the model resulting from the birth process of a rock unit is obtained from:

$$\boldsymbol{m} \leftarrow \boldsymbol{m}_{birth}^{i_{best}} \text{ with } i_{best} = \arg\min_i \psi^d\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m}_{birth}^{i=1,..,\,n_{cc}}\right). \tag{15}$$

In detail, the above process can generate units which cross-cut existing geological units. This may make sense for geological intrusions for example (Fig. 2a). However, it can also be interesting to further restrict each connected component to the initial rock types (Fig. 2b). For instance, this may make sense for accounting for hydrothermal alteration which depends on the host rock. Therefore, in our implementation, we can decide whether the considered component Ω_i is a connected component of Ω or is formed by the intersection of a connected component and a pre-existing rock unit. In this paper, we chose to insert rock units that conform to a pre-existing geological model so that (some) geological rules remain unviolated. In this particular case, the birth of rock units conforming to pre-existing geometries can be used to simulate intraunit density variations or inclusions (Fig. 2b). To simulate several scenarios, multiple threshold values $\tau_{\nabla\psi^d}$ can be used to select the most favourable case.

The flowchart of the process introduced here is illustrated in Fig. 3 and its pseudo-code is given in the Appendix.

Rock units are removed from the model through a 'death' process. In contrast to the automated birth process, which uses an optimization algorithm (eq. 12), the death process simply involves assigning

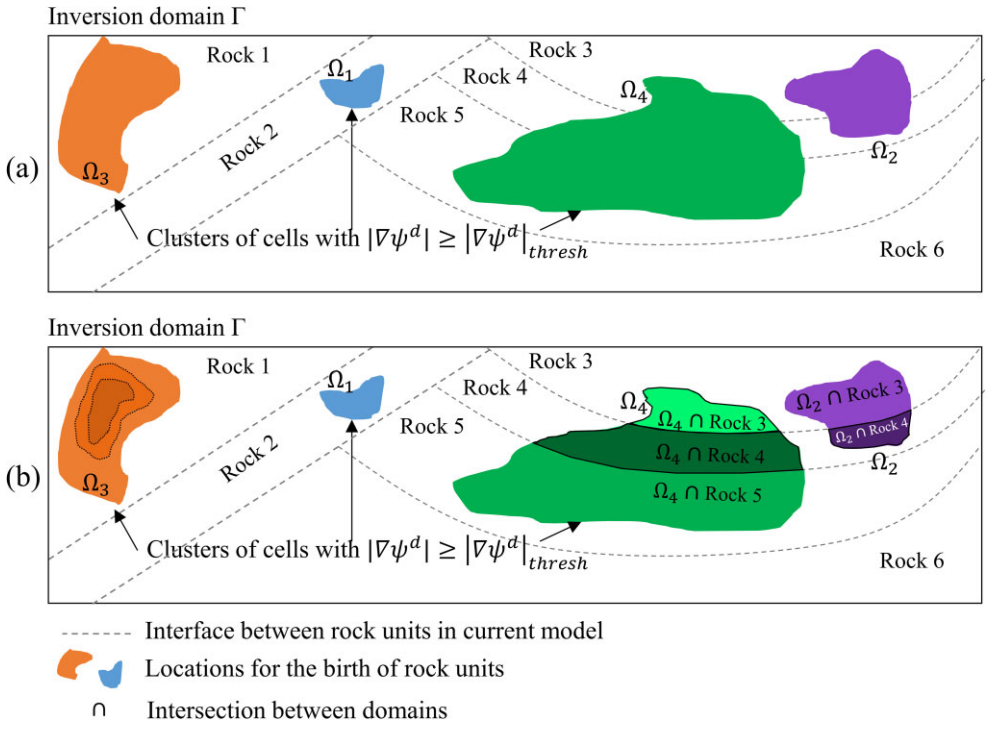

Figure 2. Conceptual example of the birth of rock units: rock units inserted without regard pre-existing rock units in the model (a) or conforming to them (b).

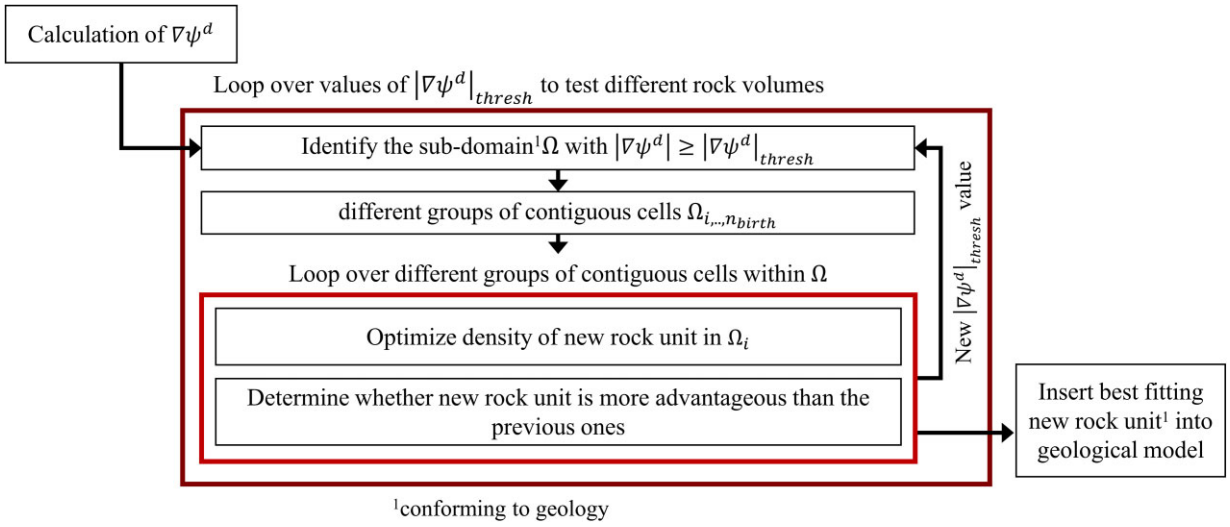

Figure 3. Proposed strategy to automate the insertion of rock units in a geological model using geophysical data.

a rock unit the density of another. Consequently, the resulting model utilizes one less density and can be described with one less signed distance field in $\boldsymbol{\phi}$.

4 UNCERTAINTY ANALYSIS USING NULL-SPACE SHUTTLES

4.1 Null-space analysis: shuttles

After inversion, we wish to study the importance of specific geological features in fitting the geophysical data and the necessity of their presence in the model using the null space of the inverse problem. For this, we approximate the null space of an inversion around a given inverse model $\boldsymbol{m}$ as the ensemble of models $\boldsymbol{m} + \Delta\boldsymbol{m}$ living in the vicinity of $\boldsymbol{m}$ that fit the data within a given tolerance ε_χ, such that for a given cost function χ, $|\chi(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m} + \delta\boldsymbol{m}) - \chi(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m})| \leq \varepsilon_\chi$. Starting from inversion results, it is possible to generate a series of models with nearly equivalent geophysical data misfits but presenting different characteristics by iteratively modifying the model. This is achieved using the algorithm of the so-called null 'space shuttles' (Deal & Nolet 1996; Fichtner & Zunino 2019; De Wit *et al.* 2012). Selecting a particular feature of interest from the model, null-space shuttles can be used to tune some of its characteristics, such as its size, location and physical property value, and analyse how the other parts of the model are adjusted to maintain geophysical data misfit within tolerance. By analysing the models recovered in this fashion, it is possible to estimate the sensitivity of the model to selected features

and to start quantifying uncertainty related to, for example, a lack of knowledge.

In this work, we apply the Hamiltonian null-space shuttles (Fichtner & Zunino 2019) to generate a series of alternative scenarios that are compatible with gravity measurements. We proceed as described below, based on the algorithm detailed in Fichtner & Zunino (2019), of which we provide only the elements that are relevant to this paper. The Hamiltonian null-space shuttle algorithm relies on maintaining the value of an 'artificial Hamiltonian', H, constant. It involves solving Hamilton's equations iteratively to keep the Hamiltonian constant while concurrently modifying its constitutive terms. In this formulation, H is taken as the sum of a given geophysical cost function χ that is treated as analogous to a potential energy, and a model perturbation term k, analogous to a kinetic energy:

$$H\left(\boldsymbol{d}_{obs.}^{geophy},\ \boldsymbol{m},\ \boldsymbol{p}\right) = \chi\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m}\right) + k(\boldsymbol{p}). \tag{16}$$

$\chi(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m})$ and $k(\boldsymbol{p})$ are expressed as:

$$\begin{cases} \chi\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m}\right) = \frac{1}{2}\left(\psi^{d}\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m}\right) + \boldsymbol{m}^{t}\boldsymbol{W}_{m}\boldsymbol{m}\right), \\ k(\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^{T}\boldsymbol{M}^{-1}\boldsymbol{p}, \end{cases} \tag{17}$$

where $\boldsymbol{W}_m$ is a weighting matrix (diagonal in our current implementation) and $\boldsymbol{M}$ is the so-called 'mass matrix'. Here, $\boldsymbol{M}$ is set as a diagonal matrix for simplicity. The variable $\boldsymbol{p}$ is called 'momentum' and varies during the null-space shuttle navigation. Using an initial value equal to

$$\boldsymbol{p} = \gamma \boldsymbol{M}\Delta\boldsymbol{m}, \tag{18}$$

the null-space shuttle algorithm can be used to impose perturbations $\Delta\boldsymbol{m}$ to $\boldsymbol{m}$ (see section 3 in Fichtner & Zunino 2019, for details) designed to test an interpretation hypothesis. The positive real number γ scales the magnitude of the perturbation and is obtained as follows. Allowing $\chi(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m})$ to vary up to a selected threshold, we set

$$k(\boldsymbol{p}) = \varepsilon. \tag{19}$$

Manipulating eqs (17)–(19), we obtain γ using the initial value of $\Delta\boldsymbol{m}$ (i.e. a perturbation direction imposed as input) at the first epoch as

$$\gamma = (2\varepsilon)^{\frac{1}{2}}\left(\Delta\boldsymbol{m}^{T}\boldsymbol{M}^{T}\Delta\boldsymbol{m}\right)^{-\frac{1}{2}}, \tag{20}$$

and can use Hamiltonian null-space shuttles using H, $\boldsymbol{p}$, $\boldsymbol{m}$ expressed as in eqs (16)–(18). Following Fichtner & Zunino (2019), we use the leapfrog method to calculate the update of the model. Using eqs (17)–(20), at a given step $t-1$ (or epoch of artificial time t), the update with step Δt is given by:

$$\boldsymbol{p}_{t} = \boldsymbol{p}_{t-1} - \frac{\Delta t}{2}\nabla\chi\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m}\right)\Big|_{t-1} \tag{21}$$

$$\boldsymbol{m}_{t} = \boldsymbol{m}_{t-1} + \Delta t\,\boldsymbol{M}^{-1}\boldsymbol{p}_{t} \tag{22}$$

where

$$\nabla\chi\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m}\right)\Big|_{t-1} = -\boldsymbol{S}^{m\,T}\left(\boldsymbol{d}_{obs.}^{geophy} - \boldsymbol{S}^{m}\boldsymbol{m}_{t-1}\right) + \boldsymbol{W}_{m}\boldsymbol{m}_{t-1}. \tag{23}$$

Beyond the work presented here, the implementation used in this work is interoperable with the Tomofast-x open source platform and supports wavelet compression of the sensitivity matrix (https://github.com/TOMOFAST/Tomofast-x).

4.2 Gravity data misfit

To monitor the evolution of the null-space shuttle algorithm and to measure the consistency of inverted models with gravity data, we calculate the geophysical data misfit as the difference between the observed and the calculated geophysical data as the following a root mean square measure of misfit:

$$ERR_{d}\left(\boldsymbol{d}_{obs.}^{geophy},\ \boldsymbol{m}\right) = \sqrt{\frac{\psi^{d}\left(\boldsymbol{d}_{obs.}^{geophy}, \boldsymbol{m}\right)}{n_{d}}}, \tag{24}$$

where n_d is the number of data points.

5 FIELD APPLICATION

5.1 Geological and geophysical context

The Pyrenean mountain range, which lies between France and Spain (Figs 4a and b), formed during the Late Cretaceous to early Miocene by convergence between the Iberian and European plates resulting in the northward subduction of the Iberian Plate (Choukroune 1989). On the northwestern flank of the Pyrenees, the vicinity of the plate boundary is characterized by high compressional velocity at depth (Chevrot *et al.* 2022; Lehujeur *et al.* 2021; Wehr *et al.* 2018), and strong positive Bouguer anomalies (see gravity data map in Fig. 4c). Several studies have focused on this area and its anomalies (see for instance Chevrot *et al.* 2018; Lescoutre 2019; Martin *et al.* 2021; Wehr *et al.* 2018, and references therein). Passive seismic imaging using data acquired during the PYROPE and IBERRARRAY campaigns (Chevrot *et al.* 2015, 2018) suggest the presence of dense mantle material lying with subvertical boundaries between the continental plates and reaching to within approximately 12 km of the surface (Wang *et al.* 2016, Wehr *et al.* 2018, Lescoutre 2019). However, in this context and area, the imaging of steeply dipping boundaries using passive seismic imaging is challenging and shows only limited sensitivity to density (Wang *et al.* 2016). In addition, the gravity data misfit of models integrating geological knowledge and seismic modelling (Wehr *et al.* 2018) remains slightly above gravity data uncertainty locally, possibly in part due to the uncertainty affecting the constraining seismic models in the presence of steep or subvertical boundaries and to the presence of small-scale heterogeneities. In contrast to seismic data, the inversion of gravity measurements offers good lateral resolution in terms of density contrasts. Here, we leverage this by integrating the model of Wehr (2017) within a geometrical inversion of gravity data. In this model, the density of the European upper crust differs from that of the Iberian upper crust, which differs from the orogen-scale model of Wehr *et al.* (2018) where these two units have the same density. Wehr (2017) explains that this density difference improved the fit of the geophysical data in the Pyrenees as a whole.

We first project the structural model of Wehr (2017) for the Mauléon area onto a different mesh that is coarser in the vertical direction and finer in the horizontal directions compared with the mesh used in Wehr (2017). This allows us to reduce memory usage by decreasing the number of model cells while increasing the density of gravity data used to model the area.

While Wehr (2017) modelled the whole Pyrenean orogen, our model is focused around a smaller area of approximately 65 × 100 km^2 in the northwestern part of the Pyrenean mountain range in and around the Mauléon Basin (Figs 4a–c). During the inversion, constraints using automated geological modelling as summarized in Section 2.1 are applied only to the two uppermost sedimentary units

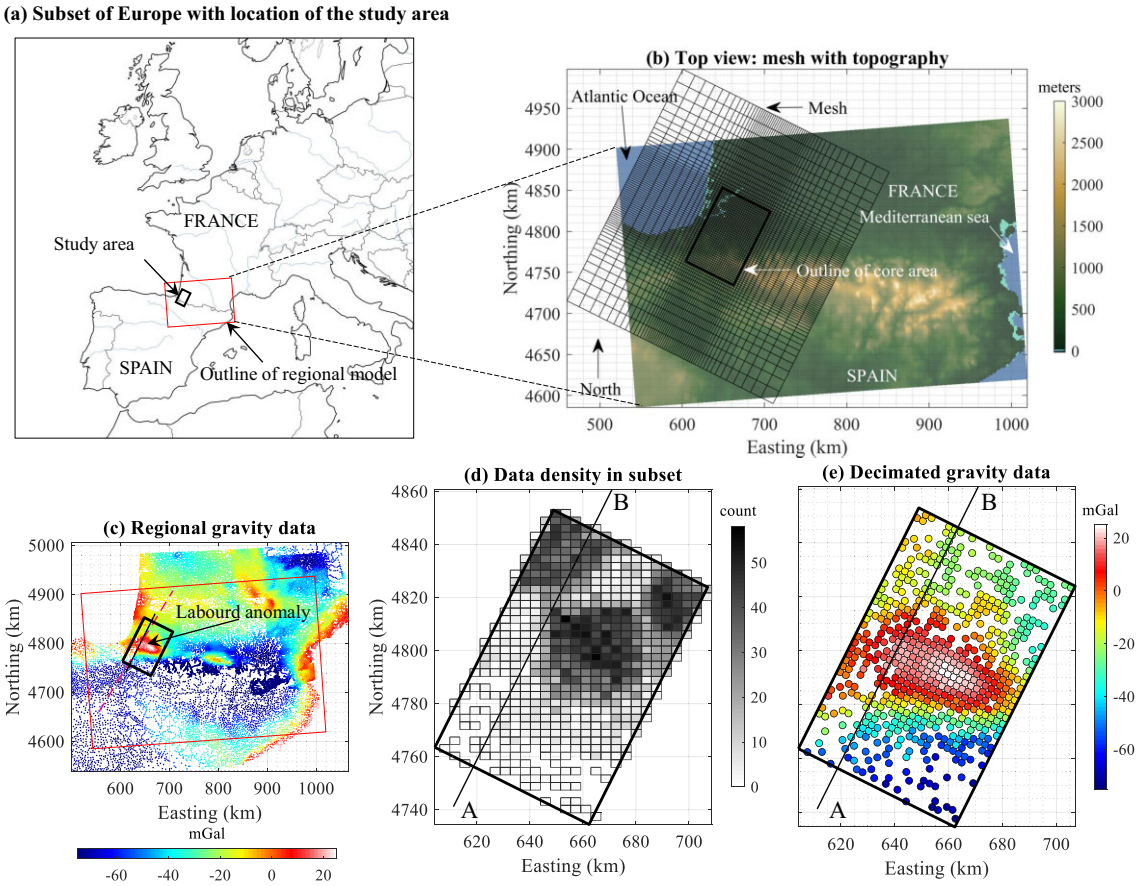

Figure 4. (a) Geographical context and gravity data: location of the study area in France and Spain, (b) top view of mesh used for modelling, (c) regional gravity data from the BGI (approximately 65 000 data points), (d) data density in subset selected here and (e) data used for geophysical inversion. Map from: https://www.freeworldmaps.net/europe/outline.html (lass accessed on 27/05/2024). Topography plot using ETOPO1 (https://catalog.data.gov/dataset/etopo1-1-arc-minute-global-relief-model, last accessed on 27/05/2024). In panel (c), the pink dashed line shows the location of profile 'A_03″ from Wehr *et al.* (2018) and can be used as a reference for visualization. We note that it almost coincides with the line A-B, shown in black in panels (c), (d) and (e). The red box represents the area covered by the full model Wehr (2017) and Wehr *et al.* (2018). In panel (d), the bins used for the calculation of the data density correspond to a regular grid with 27 × 41 elements.

(Mesozoic and Cenozoic stratigraphic units). Except at the location of contacts with either of these two units, the rest of the model can evolve more freely as it is not directly constrained by automated geological modelling (performed using LoopStructural). The considered field geological data is extracted from maps and seismic interpretations of the Mesozoic and Cenozoic stratigraphic units (the geological data locations are shown in Fig. 5c). This is motivated by the assumption that the deeper units are poorly constrained by borehole and reflection seismic data, hence that interpretations are more uncertain at depth (Wehr 2017). Consequently, while other units can evolve more freely during inversion to fit the gravity data during inversion, the geometry of the Mesozoic and Cenozoic stratigraphic units remains close to that of Wehr (2017) as projected onto the mesh used here.

5.2 Data and model preparation

5.2.1 Geophysical data and mesh preparation

We use data from the Bouguer anomaly from the region as provided by the Bureau Gravimétrique International/International Gravimetric Bureau (https://bgi.obs-mip.fr/, last accessed on 27/05/2024). We isolate the data located in the study area (black rectangle in Fig. 4, covering approximately 65 × 100 km^2), which we refer to as the core area. We discretize the horizontal plane within this area using a regular grid. For each grid cell containing more than one data point, we calculate the average Bouguer anomaly for all the values in that cell. The calculated data density calculated shows a >50-fold difference in data density in some areas (Fig. 4d). Because of this imbalance, there are many more data points in the northeast than in the southwest of the core area where some cells have no data. To mitigate this, we apply a random decimation algorithm to balance the number of data points across the area, reducing the number of data points from 10 124 to 593 by reducing the number of data points in the most data rich regions. This not only homogenizes data density, but also reduces the computational time required to run the inversion.

A higher data density is maintained in the central core area to have maximum sensitivity to the region that we assume to be the most geologically uncertain due to the potential presence of abnormally high densities. The grid defining gravity data coverage in the core area is made of a 2.5 km resolution 2-D grid and comprises 27 × 41 squared cells. Overall, this process leads to an inversion using 593 Bouguer anomaly estimations (see Fig. 4e for the resulting data considered for inversion), with expected noise levels in excess of 1.5 mGal, which is consistent with legacy data (Barnes *et al.* 2011). These data are projected at elevations provided by topography interpolated from the 1 mn arc resolution ETOPO1 topography model (https://catalog.data.gov/dataset/etopo1-1-arc-minute-global-relief-model, last accessed on 27/05/2024).

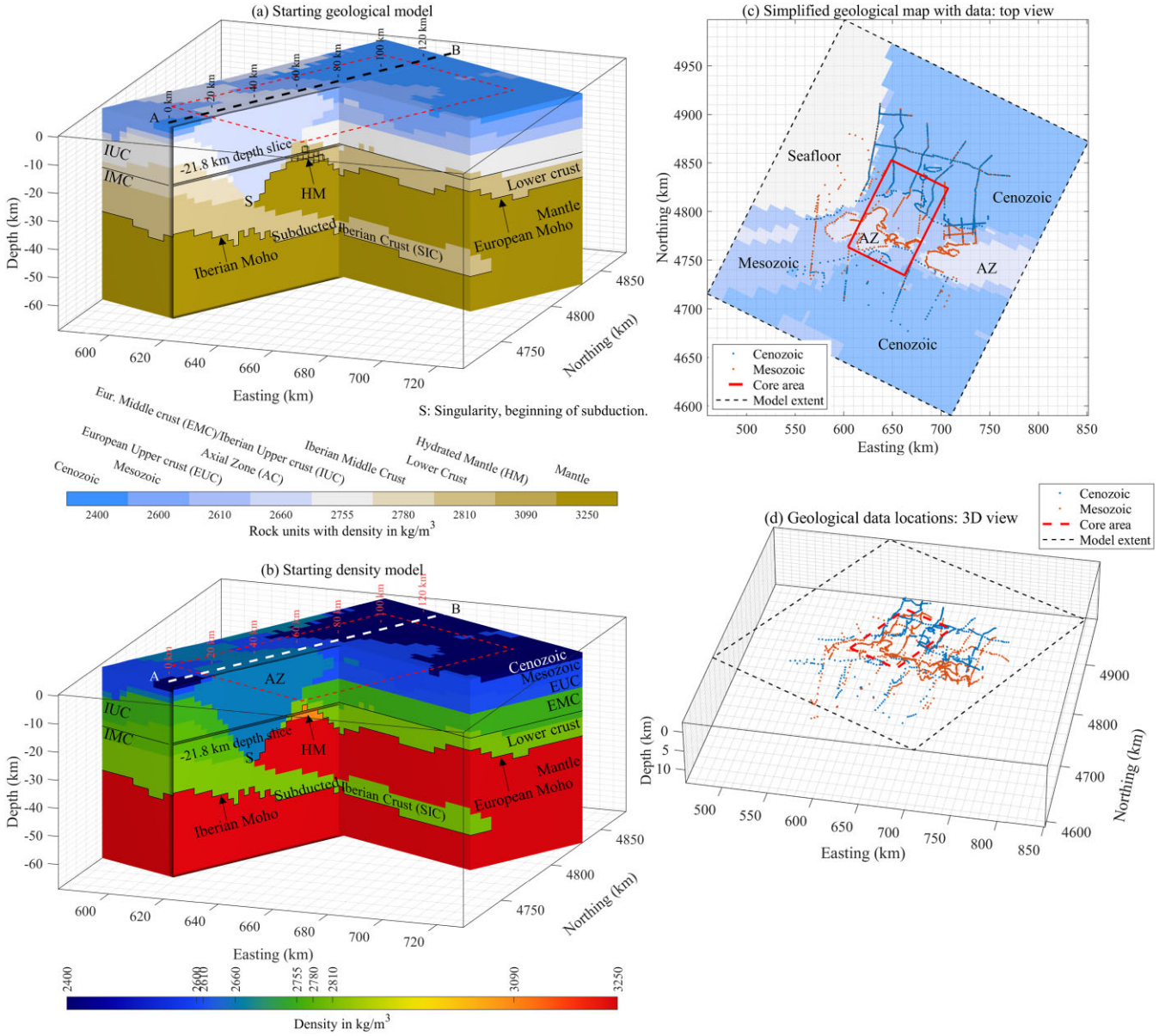

Figure 5. Starting geological model based on Wehr (2017). (a) The starting geological model and (b) the corresponding starting density model used for modelling. The core area is shown by the horizontal red square. The modelling of this volume is represented in Fig. 7 through to Fig. 11 by the cross-section A-B and the depth slice at 21.8 km. (c) Geological data points used by Wehr *et al.* (2018) and (Wehr 2017) to model Cenozoic and Mesozoic units in the area covered by the model, derived from interpretations of Cenozoic and Mesozoic stratigraphic units on reflection seismic profiles and in boreholes (Serrano *et al.* 2006) combined with seismic interpretations from the central ECORS deep seismic line (Choukroune 1989), the PYROPE (Chevrot *et al.* 2014; Wang *et al.* 2016), and OROGEN passive seismic acquisition campaigns. In (c) the top image is a top view of the geological data with topography as shown in Fig. 4(b), while the bottom image is a 3-D representation of only the geological data.

A 3-D mesh of rectangular prisms to be used for inversion is generated from the 2-D grid used for gravity data. The top of the upper layer of the model cells corresponds to the altitudes obtained from ETOPO1. Below this, we use a regular mesh from a depth of 2 km. Starting with a depth increment of 2 km, a growth factor of 1.008 is applied to reach approximately 2.5 km at 70 km with a total of 31 model cells in the vertical direction.

To reduce edge effects, the mesh is extended in both horizontal directions (Fig. 4b) with largers cells outside the core area. Overall, the grid comprises 54 × 68 × 31 model cells. Using eq. (3), the consistency with the implicit geological modelling of Mesozoic and Cenozoic units is encouraged by assigning higher values of $\boldsymbol{W}_p^{-1}$ in model cells corresponding to these two units.

In contrast to Wehr (2017), we do not interpolate the gravity measurements in areas not covered by field data during projection on the inversion grid. A consequence of this is the absence of smoothing in the data. We note that portions of the core area have a low data coverage as illustrated by the presence of bins with zero data count in Fig. 4d. Although data processing and mesh differ, Wehr *et al.* (2018) is used as a reference for gravity modelling in the area and level-set inversion stop when ERR_d reaches 3 mGal.

5.2.2 *Geological data and information*

In this study, geological information comes primarily in the form of the geological model derived by Wehr (2017) and Wehr *et al.*

(2018), which is based on a compilation and interpretation of seismic lines, outcrops and boreholes in the area. We use the original implicit model of Wehr (2017) and Wehr *et al.* (2018) calculated using GeoModeller (Guillen *et al.* 2008, Calcagno *et al.* 2008). In particular, we start from a projection onto a regular mesh similar to Martin *et al.* (2021) at regional scale in the area outlined in red in Fig. 4(a) incorporating the whole Pyrenean mountain belt, which also contains the region studied in this application (black box in Fig. 4a). We interpolate this rock unit model in the region of interest onto the mesh designed for this study (see Section 5.2.1) and use it as a starting model. This model is shown in Figs 5(a) and (b), with the original densities of Wehr (2017) reproduced in the colour bar. In this model we note that (a) the European Upper and Middle Crustal layers (EUC, EMC) are less dense than the equivalent Iberian layers (IUC, IMC) and (b) the Axial Zone (AZ), consisting of upper crustal imbricates is less dense than the IUC and denser than the EUC.

Rock units distinguished in this model are the Mesozoic and Cenozoic sedimentary cover units, upper, middle and lower crust for each of the two plates (Iberia, Europe), the Axial Zone, the upper mantle and a hydrated mantle unit (Fig. 5). Crustal structure represents a doubly vergent collisional orogen with the lower continental crust of the Iberian plate subducted north below the European Plate. Equivalent Iberian upper crustal units were thrust southward to create the thickened Axial Zone. This crustal geometry for the Pyrenees has been described on the two ECORS deep seismic lines traversing the central and western mountain range (Choukroune 1989; Choukroune *et al.* 1990; Roure *et al.* 1989) and more recently imaged using passive seismic tomography (Chevrot *et al.* 2014, 2018, 2022). The northern boundary of the Axial Zone is generally identified as the plate boundary. In this model, this steep contact juxtaposes the lower Axial Zone against a dome-like area of uplifted mantle below a European Moho at around 15–20 km depth. This is an unusual feature to find in the core of an orogenic belt. The dense mantle material relates to the Labourd positive Bouguer anomaly and its presence at shallow levels has been explained as inherited from pre-orogenic hyperextensional phase (Lehujeur *et al.* 2021; Lescoutre & Manatschal 2020; Saspiturry *et al.* 2020; Wang *et al.* 2016).

No oceanic crust was created during the Apto-Cenomanian rifting phase that preceded convergence. Instead, it is widely accepted that hyperextension exhumed mantle rocks in the centre of a narrow, segmented rift system that linked the Bay of Biscay to the Tethys ocean to the east (e.g. García-Senz *et al.* 2020; Jammes *et al.* 2009; Lehujeur *et al.* 2021; Lescoutre & Manatschal 2020; Manatschal *et al.* 2021; Tugend *et al*. 2014). Field evidence suggests that at least some of the exhumed mantle was hydrated (e.g. Lagabrielle & Bodinier 2008). Hence, the model of Wehr (2017) integrates a hydrated mantle unit (HM, Figs 5a and b) between thinned European crust and the underlying mantle. Thus, this model integrates both features inherited from the rifting phase (thinned European margin, domed mantle, hydrated mantle) and from the orogenic phase (thickened Iberian crust in the Axial Zone, subducted Iberian lower crust). The preservation of earlier rift-related features is supported by the very low overall convergence in this orogen (see Ford *et al.* 2022, and references therein), in particular in this western area (e.g. Saspiturry *et al.* 2020; Teixell 1996; Teixell *et al.* 2018).

The methodology presented in this paper offers the opportunity to interrogate several aspects of this model, to explore the range of geometries and density values that could be used while still respecting constraining geophysical and geological data. First, we note that the Axial Zone has a lower density than of the Iberian upper crust. As the Axial Zone consists at outcrop level of imbricates of Iberian upper crust, we investigate the impact of changing the density of the Axial Zone rock unit to correspond to that assigned to the Iberian upper crust. Secondly, we ask if the presence of a hydrated mantle unit is necessary. Given the variation in shape and length of the subducted Iberian crust (SIC) below the Western Pyrenees proposed by various authors based on geophysical data (e.g. Chevrot *et al.* 2014, 2015; Saspiturry *et al.* 2020; Teixell 1996; Teixell *et al.* 2018; Wang *et al.* 2016), we consider that there is high uncertainty associated with the thickness and length of the subducted Iberian lower crust, which will therefore be interrogated.

5.2.3 *Geological constraints*

As mentioned above, directly accessible geological data are mainly restricted to the two shallowest units (Cenozoic and Mesozoic sedimentary units, Fig. 5) modelled in the original model of Wehr (2017) and Wehr *et al.* (2018), who consider them to be the best constrained in the model. Our working assumption is that these units are well constrained by outcrop, borehole and shallow seismic data, which constitute the observed geological measurements $\boldsymbol{d}^{geol}_{calc.}$ used in the implicit geological modelling component of eq. (3). For this reason, we model only Cenozoic and Mesozoic sedimentary units to derive the constraining $\boldsymbol{\phi}^{geol}(\boldsymbol{d}^{geol}_{obs.}, \boldsymbol{d}^{geol}_{calc.})$ in eq. (3) at each iteration of the inversion using the method introduced in Section 3.1. The corresponding data considered as field observations $\boldsymbol{d}^{geol}_{obs.}$ used in the modelling are shown in Fig. 5(c). In contrast, we assume that the geometry of deeper units is more affected by the uncertainty inherent in the interpretations used to build the geological model and group them together under a third unit assigned to locations not belonging to either Cenozoic or Mesozoic. It follows that to constrain the geophysical inversion using automated geological modelling as introduced in Section 2, we reduce the constraining geological model to three units and consider only the geological data used to model these units in $\boldsymbol{d}^{geol}_{calc.}$ (second term in eq. 3). In such a case, the constraining geological model consists of Cenozoic and Mesozoic units, and another fictitious unit comprising all other rock types. As a consequence, the geometry of the former two units is predominantly constrained by geological data and knowledge, while variations during level set inversion within the rest of the model remain mostly driven by geophysical data. In our analysis of three-dimensional modelling results, we will use the NNE–SSW section A-B located in Fig. 5 along with a horizontal depth slice at −21.8 km.

To derive the geological constraints $\boldsymbol{\phi}^{geol}(\boldsymbol{d}^{geol}_{obs.}, \boldsymbol{d}^{geol}_{calc.})$ that are applied at each iteration of the inversion (see eq. 3), we obtain $\boldsymbol{d}^{geol}_{calc.}$ from the surfaces making up the base of the Cenozoic and Mesozoic stratigraphic units. From this surface, we extract the contact and orientation data used in the modelling. In eq. (3), we assign the surface layer a higher value of $\boldsymbol{W}_p^{-1}$ to reduce the possibility of inversion modifying the shallowest part of the model. $\boldsymbol{W}_p^{-1}$ then decreases with depth to account for the increase in uncertainty with depth. This relaxes the constraints and leaves more freedom for inversion to modify the geometry of deeper units. Overall, the geological constraints from automated geological modelling extend down to a depth of 10 km. As another constraint, we apply morphological closing to prevent the presence of isolated cells presenting a rock unit differing from all neighbouring cells.

5.3 Geophysical investigations

5.3.1 *Simulated scenarios*

Using the methods introduced in Section 2 (geologically constrained geometrical inversion) and Sect. 3 (birth of units) we design a series of 5 scenarios with hypotheses different from Wehr *et al.* (2018) and evaluate their plausibility:

Table 1. Scenarios investigated: perturbations of the original geological model investigated through gravity inversion.

Scenario	Deformation of geometries during inversion	Type of modification prior to geometrical inversion.		
		Birth of units	Death of a unit	Change in density
1	Yes	No	No	No
2	Yes	Birth of two units	No	No
3	Yes	No	No	Axial Zone changed from 2660 kg m^{-3} to the value of the Iberian Upper Crust: 2755 kg m^{-3}
4	Yes	No	Removal of the hydrated mantle unit (3090 kg m^{-3}). Replaced by lower crust (2810 kg m^{-3})	No
5	Yes	No	Removal of the hydrated mantle unit (3090 kg m^{-3}). Replaced by lower crust (2810 kg m^{-3})	Axial Zone changed from 2660 kg m^{-3} to the value of the Iberian Upper Crust: 2755 kg m^{-3}

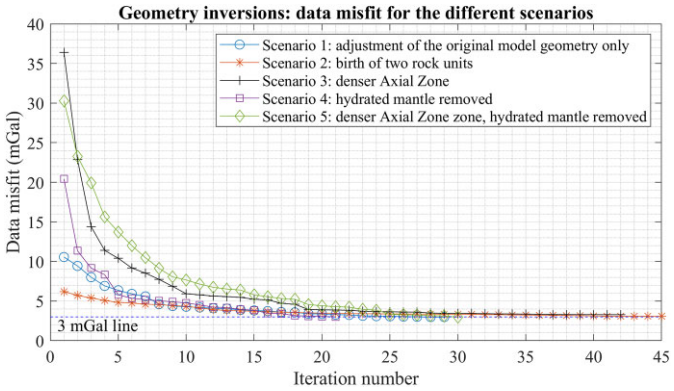

Figure 6. Evolution of the data misfit ERR_d during inversion for scenarios 1 through 5.

(i) *Scenario 1*: adjustment of the original geological model using geophysical inversion assisted by automated geological modelling to maintain consistency with geological field observations and interpretations of the Mesozoic and Cenozoic units; the objective is to improve the data fit of the original model given the considered data and to evaluate the required amount of change to the original model.

(ii) *Scenario 2*: birth of two rock units (Section 2.2) prior to inversion to generate a new starting model with 2 more units; the objective is to evaluate the impact of the birthed units on the updates of the model from inversion and the geological validity of the outcome. Note that scenario 2 mainly aims at testing the algorithm with the birth of two rock units, so it has limited relevance in terms of geodynamics.

(iii) *Scenario 3*: higher density in the Axial Zone, from the original value (2660 kg m^{-3}) to that of the Iberian upper crust (2755 kg m^{-3}); the objective is to evaluate the geological configuration where the Axial Zone is composed of material Iberian crust material that accumulated during the subduction of the Iberian Plate.

(iv) *Scenario 4*: Scenario 3 with the removal of the hydrated mantle wedge (3090 kg m^{-3}), replaced by lower crust (2810 kg m^{-3}); the objective is to evaluate the geological scenario where the hydrated mantle is not present in the model by analysing how inversion adjusts the model to fit the data without it.

(v) *Scenario 5*: higher density in the Axial Zone (as Scenario 3) with the removal of the hydrated mantle unit (similar to Scenario 4); the objective is to test yet another geological scenario combining scenarios 3 and 4.

The tested scenarios are summarized in Table 1.

5.3.2 *Geometrical inversions*

This subsection presents the series of inversions considering the scenarios in Table 1. The evolution of the corresponding data misfit ERR_d (Fig. 6) shows that all scenarios converge to a stable solution with ERR_d falling within the prescribed range. This indicates that inversion results can be intrepreted and compared. In Fig. 6, it is evident that the starting data misfit of Scenarios 3–5 is significantly higher than for Scenarios 1 and 2. Scenarios 3 and 5 exhibit the strongest starting misfit. This indicates that setting the density of the Axial Zone to that of the upper crust deviates significantly from the hypotheses made during the construction of the original model. In fact, the increase of the Axial Zone density in Scenarios 3 and 5 has a greater impact on the data misfit than the other perturbations because the Axial Zone, being voluminous and shallow, is the most pronounced deviation investigated. Although the difference with the original models is stronger in Scenario 5 than in Scenario 3, the corresponding starting misfit is lower, indicating that the removal of the hydrated mantle partly mitigates the increase in the density of the Axial Zone. Furthermore, the data misfit for scenario 4 is significantly higher than for scenario 1, which indicates that the hydrated mantle is crucial to explain the gravity anomaly in the original model. For a more detailed analysis of the results, we use the starting density models and inversion results obtained for these five scenarios as shown in Figs 7–11.

In Figs 7 and 8, we observe that in both Scenarios 1 and 2, the addition of relatively shallow mantle material at locations originally set as Axial Zone is needed by inversion to improve the fit of gravity data. A comparison of depth slices in Fig. 7 also indicates an increase in the lateral extension of the uplifted mantle material

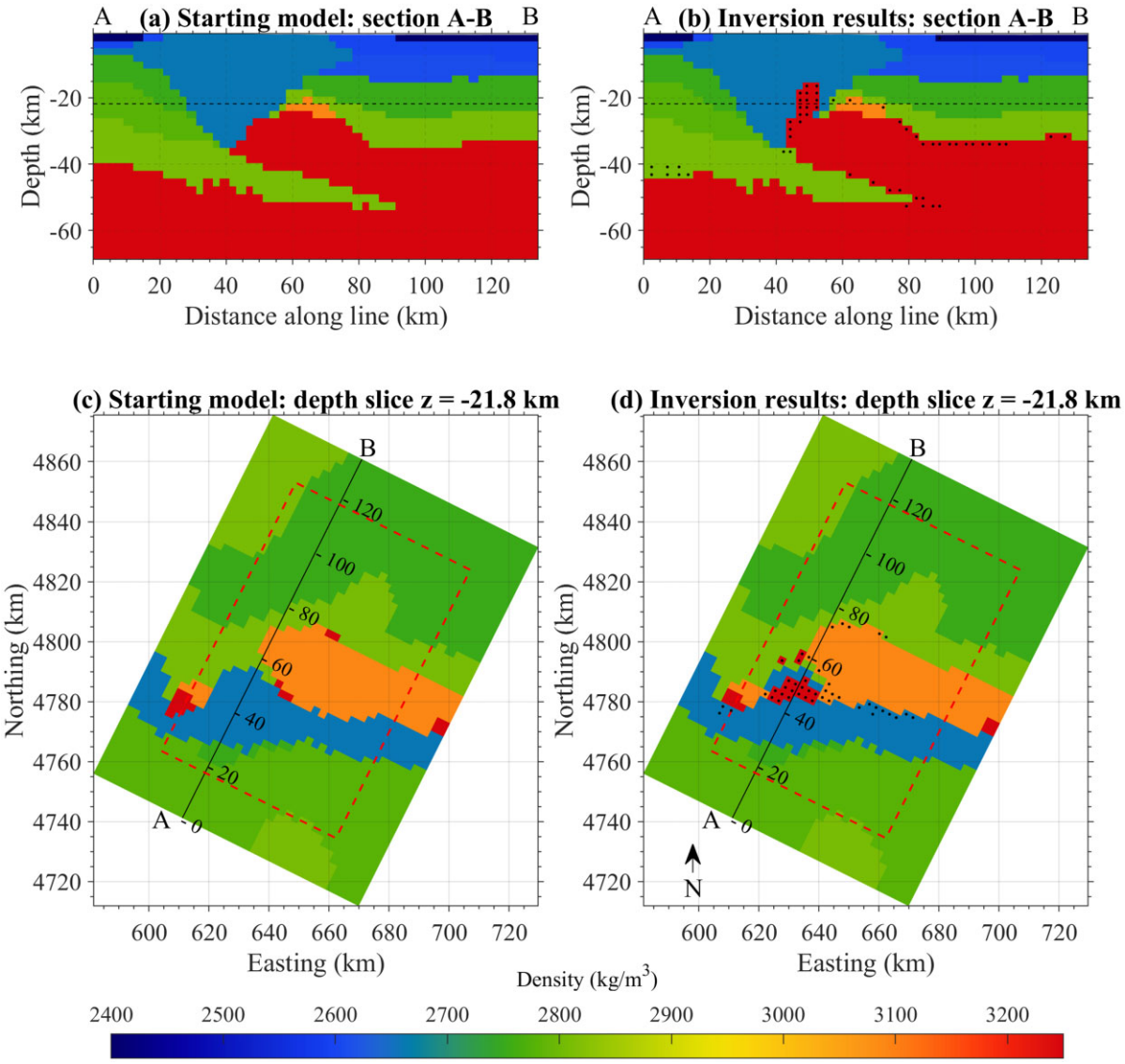

Figure 7. Inversion results for scenario 1 (original number of rock units): starting model (left-hand column) and inversion results (right-hand column). In (b), the black dots show the changes that occurred during inversion.

after inversion. Comparison of depth slices for Scenarios 1 and 2 reveals that the rock units that were added to Scenario 2 reduce the modifications to the mantle unit required to fit the gravity data mostly to the contact between the mantle and the Axial Zone. The main qualitative interpretations that we make are summarized in Table 2.

From the analysis of Scenarios 1–5, three geologically and geophysically plausible changes to the original model emerge: a denser Axial Zone, a shorter subducted crust, and the removal of the hydrated mantle unit. Prior to further examination using the null-space shuttles, we note that the data misfit in the models discussed above correspond to an RMS misfit of 3 mGal. This relatively large error represents the lower limit of what geometrical inversions can accommodate at the considered scale and with the current model discretization. Instead of adding smaller units to fit small-scale anomalies from the gravity data using the birth process presented above (which is not easily justified from a geodynamic standpoint), we consider continuous valued inversion to reduce data misfit further and to obtain intra-unit density variations.

5.3.3 *Heterogeneous density variations*

For the recovery of heterogeneous density variations, we invert for the residuals obtained after geometry inversion and impose bound constraints to density variations using the inversion approach of Ogarko *et al.* (2021), restricting the changes to ± 15 kg m^{-3} in Tomofast-x. In each of the scenarios investigated above, the data misfit ERR_d decreases to approximately 1.70–1.80 mGal. In each case, the distribution of residuals is close to a normal distribution centred on zero. Except for a minor trend showing larger residuals in the South, the residuals map reveals no clear pattern (see Fig. 12 for the case of scenario 1). The residual map shown in Fig. 12(f) also indicates that using non-interpolated data generates local, strong residual values in the Iberian side characterized by rapid lateral variations.

In the original model, the closest density values are those for shallow units consisting of the European upper crust (2610 kg m^{-3}) and Mesozoic units (2600 kg m^{-3}) in the northern part of the model and are difficult to differentiate using gravity inversion. Density

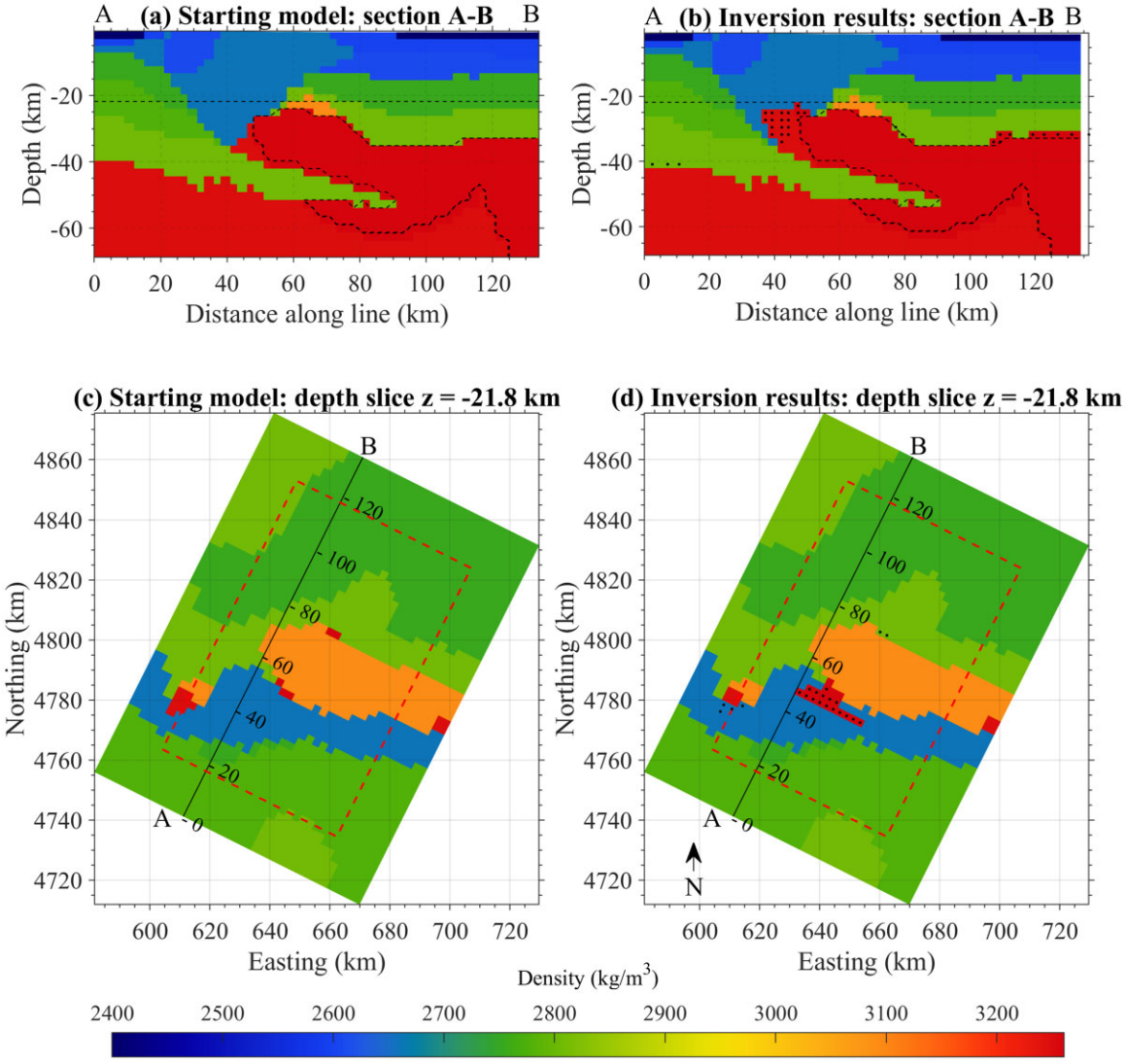

Figure 8. Inversion results for scenario 2 (birth of a rock unit, shown by the white arrow): starting model (a, c, left-hand column) and inversion results (b, d, right-hand column). The unit within the curved black dashed line has a slightly larger density (+11 kg m^{-3} compared to the original mantle density) than the surrounding mantle (3250 kg m^{-3}). In panels (a) and (b), the white arrow and the black dashed shape indicate the two units added to the model using the birth process. In panel (b), the black dots shown with the inversion results identify the changes in the geometry of the mantle.

changes that may warrant swapping rock units between the European upper crust and Mesozoic units may occur for a few model cells in some places. However, assuming that densities within a rock unit can generally be described using a normal distribution with a standard deviation greater than 10 kg m^{-3} due to possible variations of density within the rock, it is safe to consider that small-scale variations as recovered here may occur naturally within each rock unit. In addition, variations in density picked up by continuous inversion with bound constraints may be related to lateral intracrustal variations that require a more detailed analysis, but which lies beyond the scope of this paper.

5.4 Additional scenario evaluation using null-space shuttles: reducing the length of the subducted Iberian crust

We perform null-space exploration to evaluate alternative mantle structures starting from the inversion results of Scenario 1.

The proposed perturbation $\Delta \boldsymbol{m}$ used in eqs (18)–(20) is added to the inverted model with the intention of reducing the length of the subducted crust. For this, $\Delta \boldsymbol{m}$ is set to +440 kg m^{-3} (difference between Iberian lower crust and mantle) in the selected model cells (see envelope in red on the 3-D view of the inverted model in Fig. 13), and zero elsewhere.

With this, we start the null-space navigation using the algorithm in eqs (18)–(23), which stops after 262 epochs, when the density values in the cells with $\Delta \boldsymbol{m} \neq \boldsymbol{0}$ are sufficiently close to that of the mantle. To reduce the geographical range of possible model perturbations, we prevent model changes that are shallower than 10 km above the Moho during null-space navigation. The model resulting from replacement of crustal material by mantle is shown in Fig. 14 along the A–B section.

Fig. 14(e) shows that at the end of the null-space navigation, the data misfit is nearly equal to that at the beginning of the process. An animation showing the shortening of the crust on a profile west of A–B is available (Giraud *et al.* 2023b).

The forward data $\boldsymbol{S}^m \Delta \boldsymbol{m}$ produced by the original perturbation $\Delta \boldsymbol{m}$ produces values in excess of 20 mGal. With a root mean square value $n_d{}^{-1/2} \| \boldsymbol{S}^m \Delta \boldsymbol{m} \|_2$ of approximately 19.2 mGal, it is necessary to fit the gravity data, and the application of $\Delta \boldsymbol{m}$ needs to be compensated by significant model changes (Fig. 14d). Overall, this

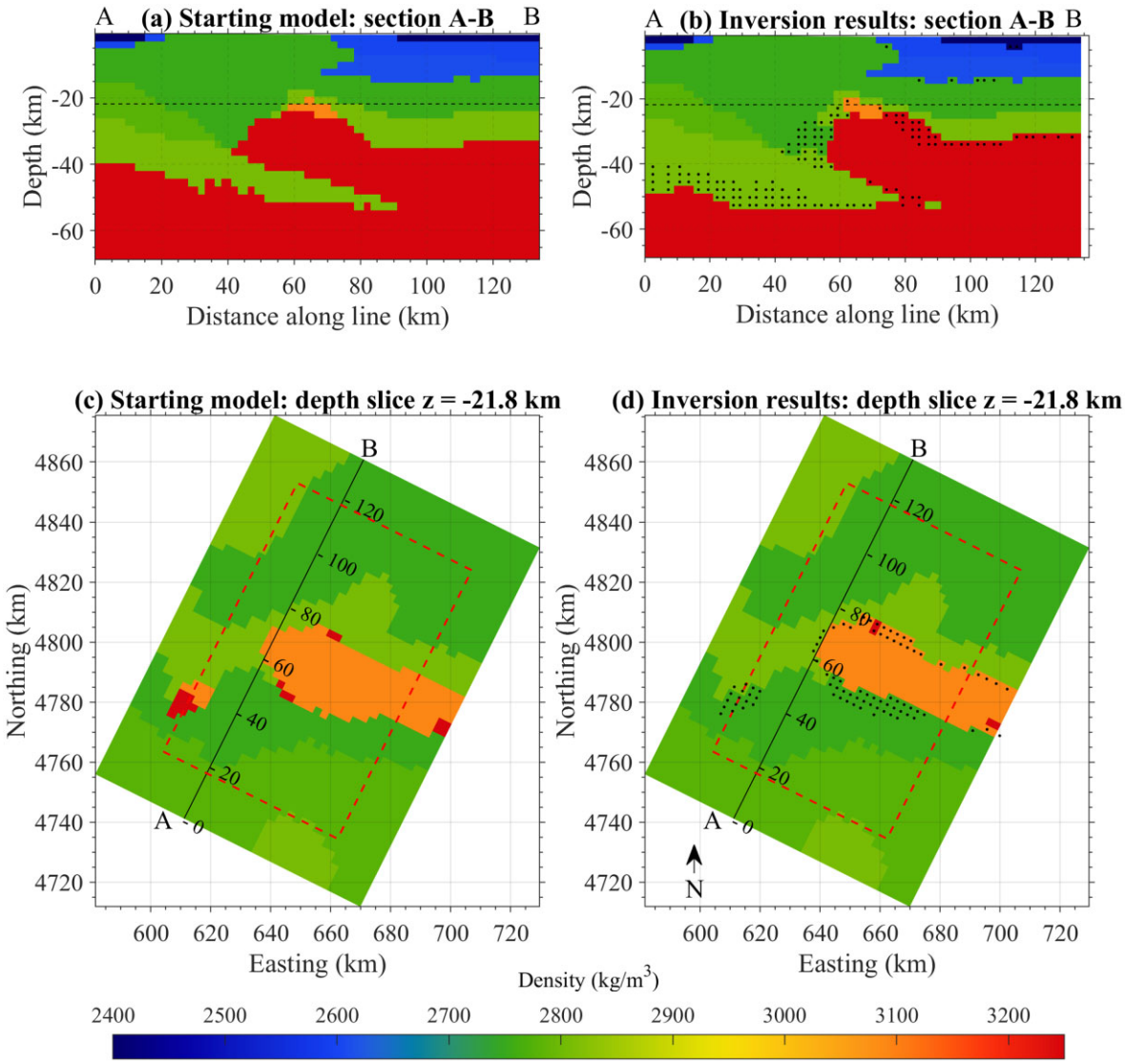

Figure 9. Inversion results for scenario 3 (denser Axial Zone): starting model (a, c, left-hand column) and inversion results (b, d, right-hand column). In panel (b), the black dots show the changes that occurred during inversion.

shows that, under the conditions of our experiments, reducing the length of the subducted Iberian crust as proposed here is consistent with the gravity data. In addition, as indicated by geometry inversion (Section 5.3.2), this hypothesis is consistent with scenarios 3–5. The changes brought by the null-space shuttles to the model indicate that a shorter subducted crust may be compensated by the addition of denser material in the Axial Zone and by a reduction in density in the vicinity of the hydrated mantle unit (see Fig. 13 for its outline). On seismic tomography images in the same area, the interpreted position of the Iberian Moho can vary but overall appears to correspond most consistently with the shorter subducted crust (Chevrot *et al.* 2015, 2018).

5.5 Geological and geodynamic implications

It could be argued that the alternatives proposed here (Scenarios 3–5) are extreme cases that constitute end-members of plausible scenarios. From the differences shown in Fig. 14, we observe that, in the conditions of the experiment, two main changes are required to accommodate the shorter subducted Iberian Crust in the gravity model:

(i) The addition of dense material in the Axial Zone (shallower cluster of black dots in Fig. 14d), in its central-northern part to reach the same values as in the Iberian middle crust.

(ii) The reduction in density within, and slightly north of, the hydrated mantle (deeper cluster of black dots in Fig. 14d).

The shortening tested with the null-space shuttles is significant (approximately 25 km) and may be exaggerated. To investigate intermediate scenarios, we performed another experiment consisting of a sequence of six smaller, successive perturbations using the null-space shuttles. In this way, we obtain a series of six models with progressively shorter subducted crust. An animation showing the evolution of the density model during the transition from one of these models to the next is available in the Supporting Information (Giraud *et al.* 2023b).

Generally, starting from Wehr (2017) our modelling suggests alternatives of the model, including:

(i) Crustal thickening in the southwest region of the model area above the Iberian Moho, and thickening of the European mantle below the Axial Zone.

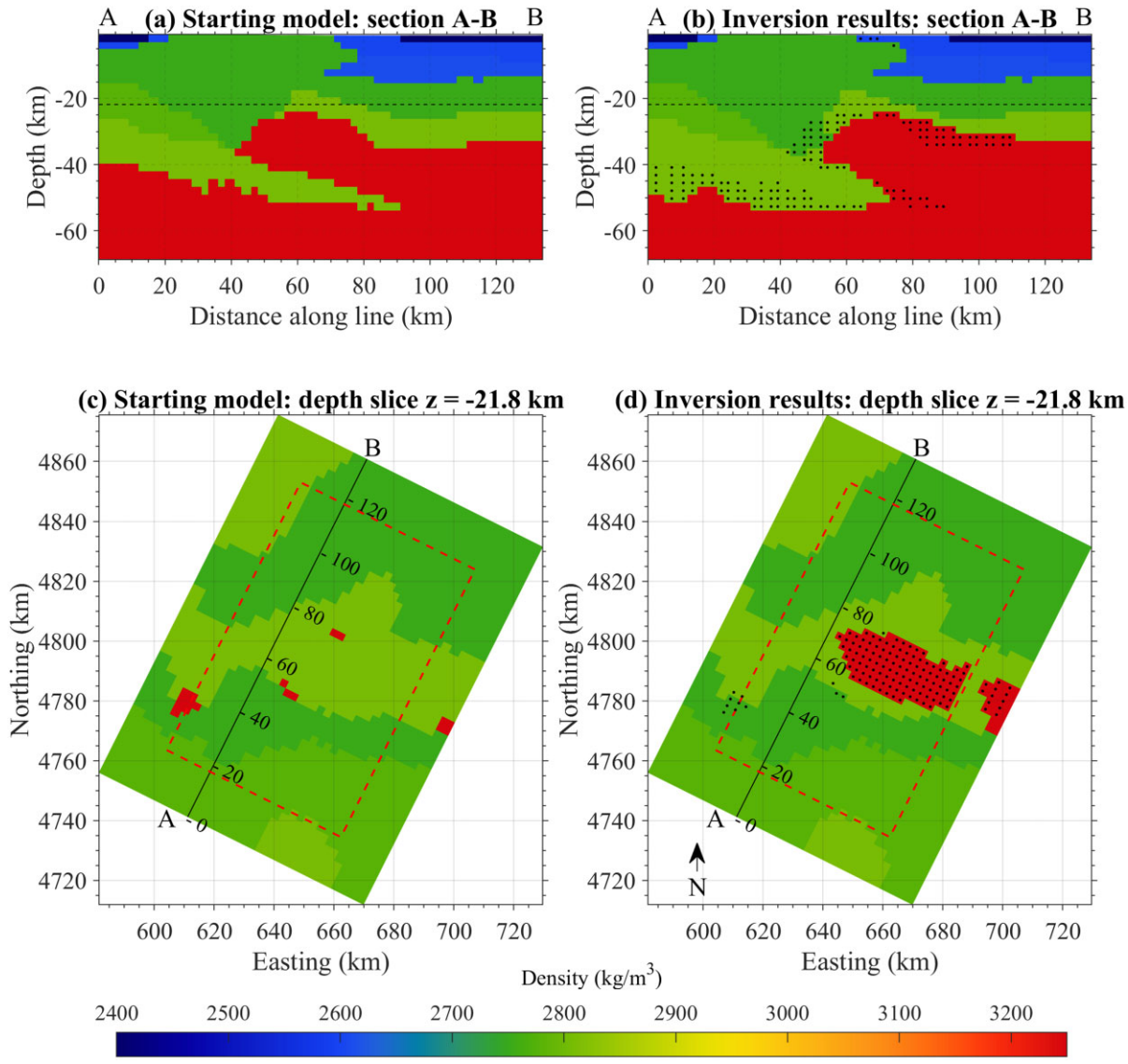

Figure 10. Inversion results for scenario 4 (removal of the hydrated mantle unit): starting model (a, c, left-hand column) and inversion results (b, d, right-hand column). In panel (b), the black dots show the changes that occurred during inversion.

(ii) A two-part deep density zone at the base of the Axial Zone which comprises a less dense southwestern end and denser north-eastern end.

(iii) some changes in the geology-constrained part, but these remain minor and may be explained by the relatively coarse model resolution used in this study.

(iv) tuned geometries improving the fit to the gravity data but which remain largely consistent with the original structural model and proposed units.

6 DISCUSSION

6.1 Methodological implications and limitations

The method proposed here to estimate the best fitting densities within each of the considered units (Section 3.1) is straightforward to implement as all required packages are easy to use and open source. Such a density adjustment can be useful to compensate for the uncertainties inherent to choosing the densities of rock unit from sparse and incomplete borehole data (especially at depth). Beyond the work presented here, this approach might remove the need for manual tuning of rock unit density (or magnetic susceptibility) in some potential field inversion exercises, as it simultaneously provides a unique optimum density for each considered rock unit. This may be useful to refine the density values when they are uncertain or to assess the capability of a given structural model to explain the gravity data.

The workflow we propose relies mostly on a deterministic approach, wherein the model space is explored by altering the parametrization of the problem based on a few expert-designed scenarios. This is useful for rapidly investigating alternative geological scenarios without the complexity of, e.g., methods like trans-dimensional Monte Carlo sampling; however, our approach lacks exhaustiveness.

The work presented here is applied to gravity data, but it can easily be transposed to magnetic data as the only change required is the calculation of the corresponding sensitivity matrix.

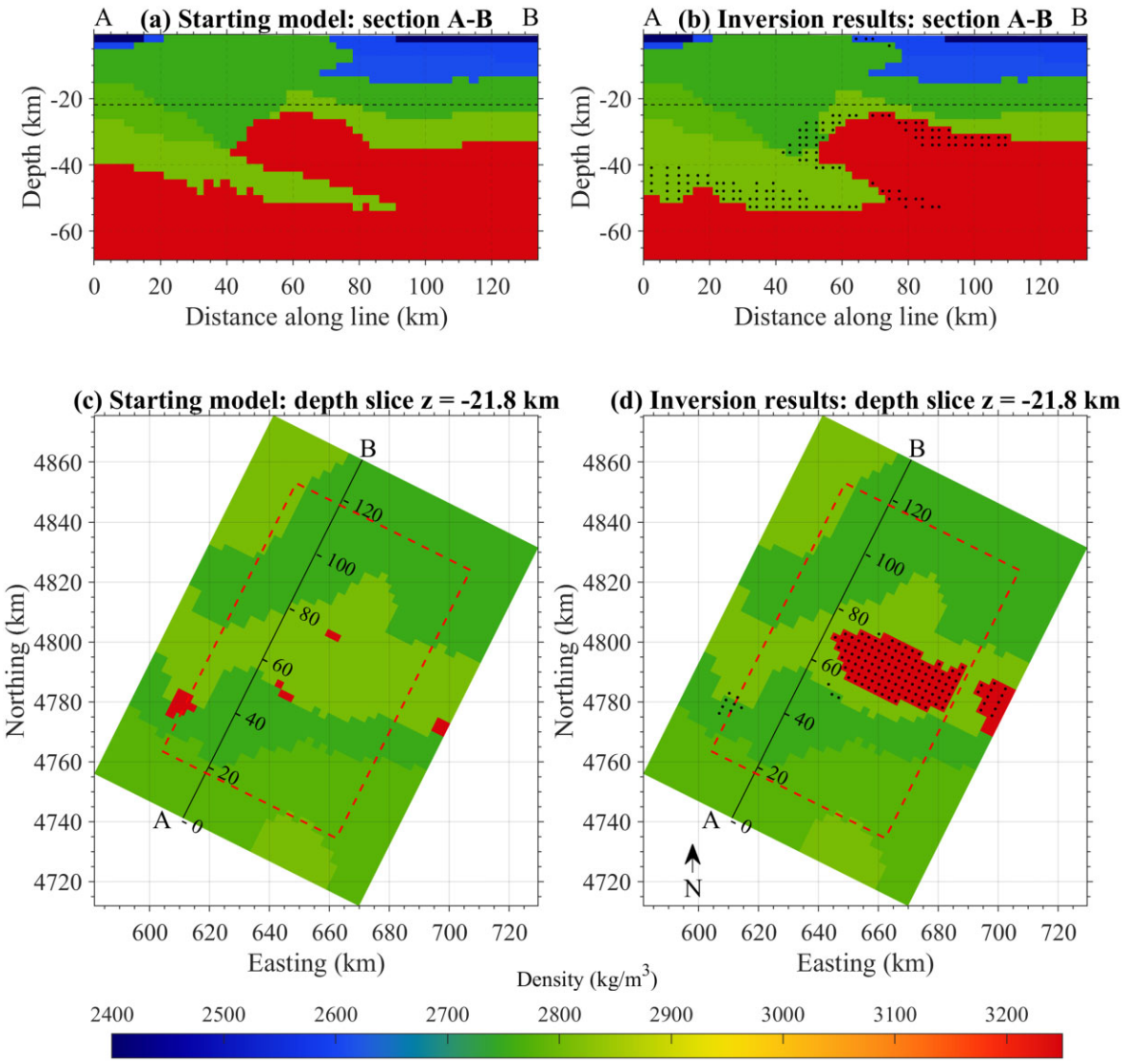

Figure 11. Inversion results for scenario 5 (denser Axial Zone and removal of the mantle wedge): starting model (a, c, left-hand column) and inversion results (b, d, right-hand column). In panel (b), the black dots show the changes that occurred during inversion.

Table 2. Summary of observations and preliminary interpretations, with comparisons made with the original model of Wehr (2017).

Scenario number, figure number	Observation on affected units, compared to the original model		
	Axial Zone (AZ)	Subducted Iberian Crust (SIC)	Mantle, Moho
Scenario 1, Fig. 7	Addition of mantle material in deep AZ	Minor reduction in length	Deeper Iberian Moho, shallower European Moho
Scenario 2, Fig. 8	Birth of one unit: lower density in the upper Southern AZ (-51 kg m^{-3}; white arrow Fig. 8). Mantle material to replace nearly all the deep AZ.	No significant change in shape.	Birth of one unit: higher density in the European side of the mantle ($+11$ kg m^{-3}). Mantle material replaces nearly all the deep AZ. Slight shallowing of the northern European Moho
Scenario 3, Fig. 9	No significant change in shape.	SIC is shorter and thicker	Deeper Iberian Moho. Shallower European Moho in the North.
Scenario 4, Fig. 10	Dense mantle rocks replace part of the deep Axial Zone	Slightly shorter SIC	Slightly deeper Iberian Moho. Slight shallowing of the European Moho both to the north and above the mantle dome
Scenario 5, Fig. 11	No significant change in shape.	Shorter and thicker SIC	Deeper Iberian Moho, shallower European Moho and a mantle dome

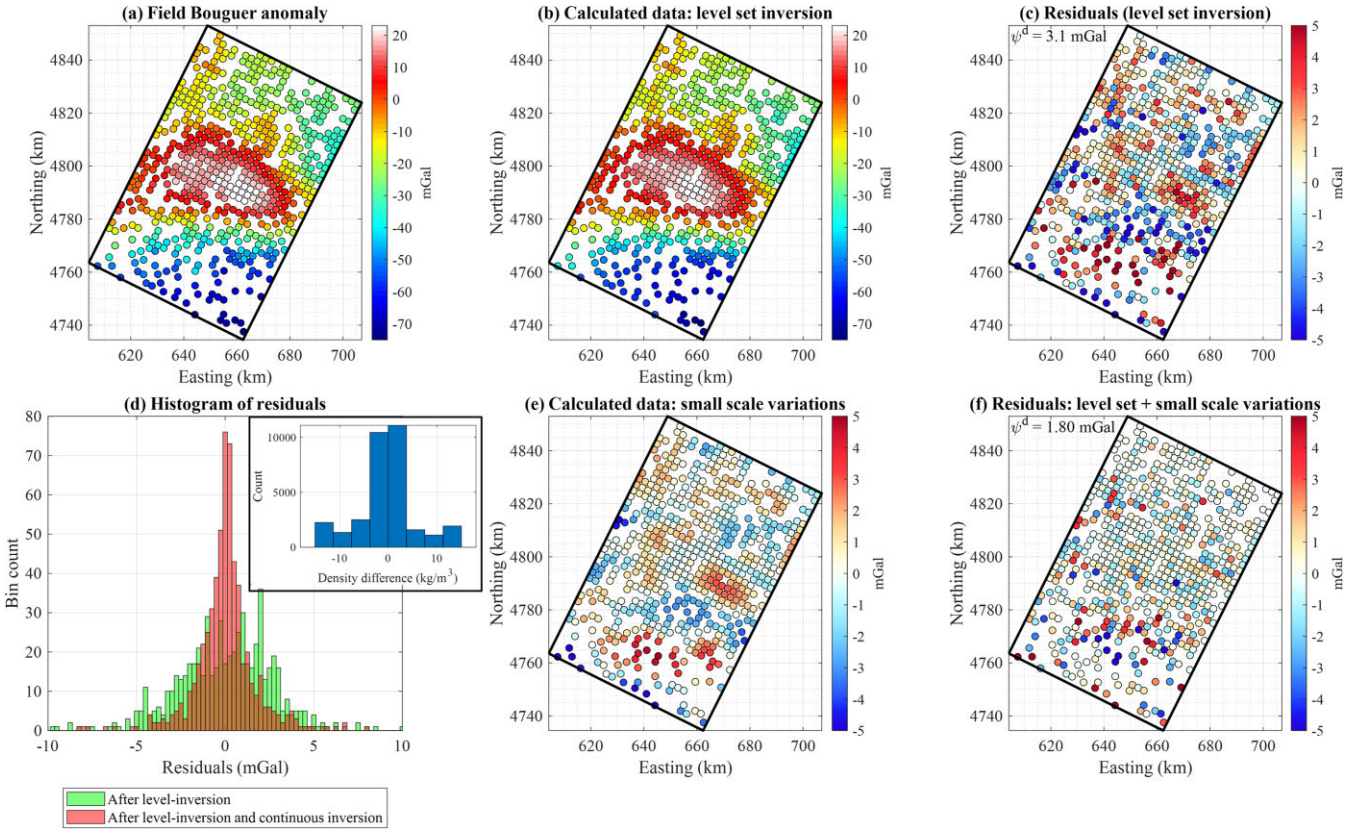

Figure 12. Gravity data analysis with inverted data (a), calculated data after level set inversion (b) and the associated residuals (c); comparison of residuals between after level set inversion and level set inversion followed by continuous value inversion (d), calculated data for the corresponding variations (e) and the associated residuals (f). The inset in (d) shows the histogram of density variations with respect to the model obtained from geometry inversion.

6.2 Geological implications and limits of the investigations

Previous studies have discussed the potential presence of native hydrogen at depth in the studied area associated with hydrated mantle lithologies at high levels in the subsurface. This potential is partly inferred from measurements that indicate the existence of chemical elements associated with native hydrogen, as well as a favourable geological context (Lefeuvre *et al*. 2021; Tichadou *et al*. 2021). The geological framework these studies use in their assessment has been influenced, in part, by the body of literature used in Wehr (2017), and Wehr *et al.* (2018) to build the geological model used as a starting point for our analysis. The plausibility of alternative scenarios investigated in this paper suggests a broader range of feasible models. Furthermore, these scenarios imply that the origin or possible migration pathways (such as shallow mantle bodies) might extend deeper than previously speculated.

We investigated five scenarios. More could be explored in the future, such as increasing the thickness of the European upper crust at the expense of Mesozoic units, setting similar values for the crustal densities of the European and Iberian crusts, or decreasing the density of the Iberian mantle.

We have proposed variations of the model of Wehr (2017) and discussed the limitations of certain hypotheses used to construct the original model by, for example, deviating from 'the hypothesis that positive Bouguer anomalies are produced by exhumed mantle bodies' (Wehr *et al.* 2018). However, the alternative models we propose are also affected by uncertainty due to approximations made during the design of the geophysical survey. Interpretations and comparisons must therefore be approached with caution. Besides, the lack of an automated geological consistency check outside of the Cenezoic and Mezosoic units means that geologically unrealistic contacts between rock units may be present. Unlike Wehr (2017) and Wehr *et al.* (2018), our modelling focuses on a subset of the Pyrenees. Moreover, the gravity data we invert exhibit slightly different features due to variations in the gravity data preparation workflow, and our inversions and subsequent analyses were performed using different algorithms.

6.3 Future work

We here summarize seven aspects of the presented work that we have identified as avenues for future research to improve the proposed modelling technique and workflow.

(1) The method for the 'birth' of rock units is designed to test the impact of adding initially non-identified rock units to the geological model. Starting from a pre-defined number of rock units, the process of inserting new units in a given model modifies the dimensionality of the level-set problem. This opens up possibilities for the development of trans-dimensional level set inversions, which we may investigate in the future.

(2) The rock units generated from the geophysical data-driven process are only crudely constrained in the geological model. One way to improve their definition may be to automatically incorporate prior geological knowledge or principles when inserting rock units in the model.

(3) An alternative to the use of automated geological modelling to recover the geological image of a given set of signed distance

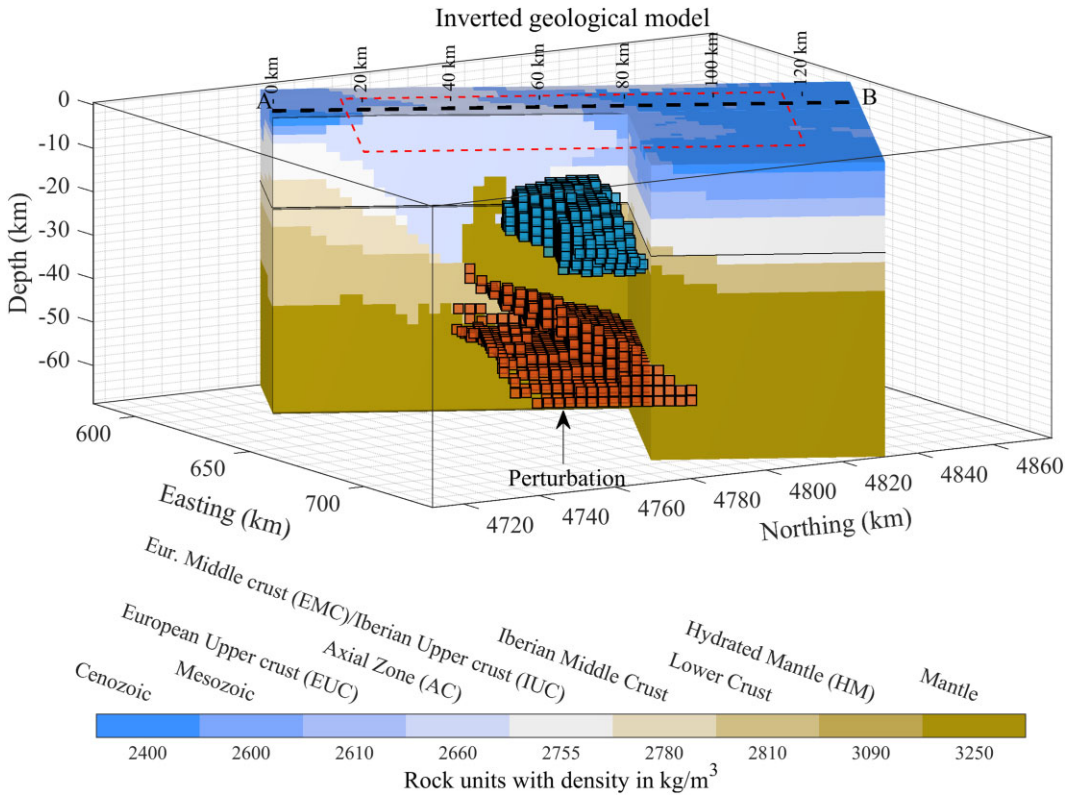

Figure 13. 3-D view of the inverted model for scenario 1 with the envelope of the perturbation for the shortening of the subducted crust for null-space exploration shown by the brown cubes. The blue cubes are a reminder of the location of the hydrated mantle in the original model.

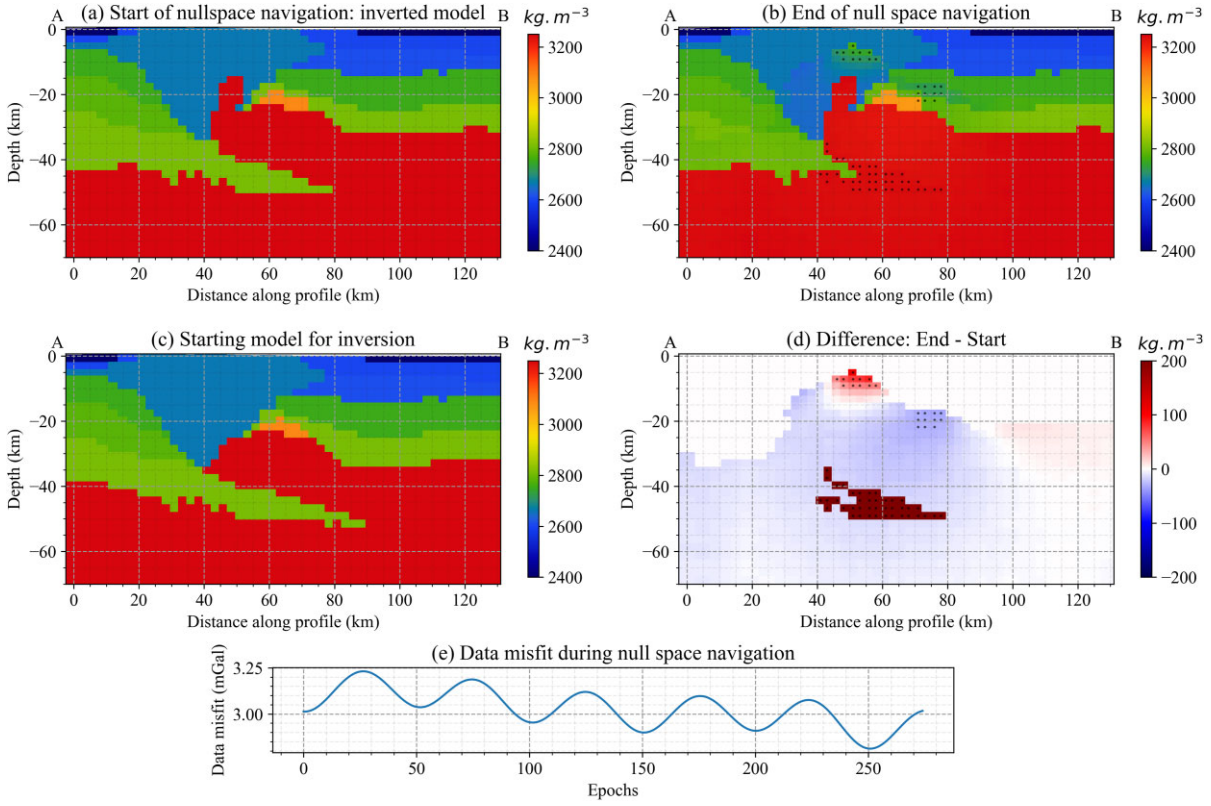

Figure 14. Null-space navigation, A–B section: (a) starting point of null-space navigation, (b) last model of the navigation which can be compared to the starting model used for inversion in (c); (d) shows the differences in density between models in (a) and (b). The black dots shown in (d) indicate the location of density changes with values greater than 30 kg m^{-3} and (e) evolution of gravity data misfit during null-space navigation.

values proposed by geophysical data may be to use deep learning techniques. In particular, pre-trained multiple-view convolutional networks could be used in our workflow as a starting point to identify the closest type of geological scenario that could fit the area. This might come as an extension of the work of Guo *et al.* (2021) in which deep learning is used as a means to obtain a prior model to constrain inversion in a similar spirit as Colombo *et al.* (2021).

(4) It is reasonable to expect that results obtained from the integration of geological and geophysical data as presented in this paper be less uncertain than when such integration is not performed. However, while we have presented elements for the analysis of alternative scenarios, the rigorous analysis that is required to properly estimate the uncertainty associated with our inversion results has not been carried out. For this reason, we acknowledge that the proposed workflow should integrate a thorough uncertainty analysis using sampling techniques. Such a study would ideally need to also consider the seismological data in a quantitative way.

(5) In our level set inversion approach, we recalculate the signed distances to interfaces to maintain the signed distance properties of $\boldsymbol{\phi}$ at each iteration. This is done using the FMM, which can be costly for large models and limit the application to small models. In our algorithm, we have the possibility to use the Narrow Band FMM (Chopp 1993), which reduces computation time by up to an order of magnitude. To reduce the computational cost further, an avenue is to use the update of signed distances from the LSQR algorithm to calculate extension velocities. The operation of updating signed distances with extension velocities is simple and does compromise the signed distance properties (Adalsteinsson & Sethian 1999), thereby alleviating the need to re-initialize the signed distances with the FMM. In addition, it could unlock the potential for uncertainty analysis (Yang *et al.* 2019).

(6) The proposed techniques were applied to a regional study focussing on deep crustal-scale structures. However, the underlying principles have a broad applicability in nature so that the workflow can be readily transferred to scenarios such as undercover imaging in mineral exploration or basin studies. Likewise, the methods presented here can readily be transposed to modelling magnetic data as the forward solver for both gravity and magnetic data is linear. In particular, the open-source Python implementation for null-space exploration is capable of loading sensitivity matrices (compressed or not) calculated with Tomofast-x. Because Tomofast-x can also be used to generate the sensitivity matrix for magnetic data, the null-space code requires only minor tuning to be used with such data.

(7) Our manual investigation of a trans-dimensional problem with the birth of units (in scenario 2) and the death of one (scenario 4) highlights the usefulness of the automated process of inserting and removing rock units to better sample the model space. A natural extension of this work is therefore to address trans-dimensional sampling in a stochastic inversion context.

7 CONCLUSIONS

We have presented an extension of existing inversion workflows focussing on:

(i) A geometrical inversion approach relying on the inversion of signed distances between units constrained by geological data.

(ii) The introduction of a geophysical data-driven technique for the birth of rock units within a pre-existing geological model.

(iii) The use of null-space shuttles for uncertainty analysis and the addition or removal of selected features while maintaining the geophysical data fit of the model.

We applied these techniques to the northwestern Pyrenees, starting from a pre-existing geological model accounting for previous studies of the Pyrenees. This case study serves as an example of the direct integration of automated implicit geological modelling, level-set inversions and exploration of equivalent geophysical models. In this study, we utilized the proposed workflow to modify the geometry of geological units, adjust their density and add or remove units while fitting gravity data. Our investigations show the plausibility of alternative models to the initial input model presenting:

(i) A shorter subducted Iberian crust.

(ii) Densities in the Axial Zone equal to those of the Iberian upper crust.

(iii) The absence of hydrated mantle at shallow depths, together with some intermediate cases, demonstrating that the deep structure of the northwestern Pyrenees remains ambiguous and requires further investigation.

FUNDING INFORMATION

JG was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions grant agreement No. 101032994. JG, GC and PC acknowledge support from the academic and industrial sponsors of the RING consortium (https://www.ring-team.org/consortium) managed by ASGA (Association Scientifique pour la Géologie et ses Applications), last accessed 10/06/2024). LG was supported, in part, by the Australia Research Council Linkage grant LP17010985 – Enabling 3-D Stochastic Geological Modelling, supporting the development of the Loop platform, and by the Loop consortium (https://loop3d.org/, last accessed 10/06/2024). Contribution of VO to this research was supported by the Mineral Exploration Cooperative Research Centre (MinEx CRC), whose activities are funded by the Australian Government's Cooperative Research Centre Programme.

ACKNOWLEDGMENTS

We thank the Bureau Gravimétrique International/International Gravimetric Bureau (Geosciences Environment Toulouse, France) for providing the gravity database. We also thank Simon Lopez and Christelle Loiselet from the BRGM (French geological survey) for their support in accessing the original geological model developed by the BRGM RGF-Pyrenees programme. The authors express consideration to Sébastien Chevrot and Lucia Seoane from CNRS for interesting discussions about the modelling and history of the Pyrenees and for the compilation of the gravity data set the data used here was extracted from, respectively. JG also acknowledges Gloria Arienti from the University of Milano Bicocca, for discussions about the structural geology of mountain ranges, Ben McCarthy from Epiroc for his suggestions of algorithms to use in future developments and Mark Lindsay, Mark Jessell and Mahtab Rashidifard from the University of Western Australia for interesting discussion about geoscientific integration.

AUTHOR CONTRIBUTIONS

Jérémie Giraud (Conceptualization [lead], Data curation [supporting], Formal analysis [lead], Funding acquisition [lead], Investigation [lead], Methodology [lead], Software [lead], Visualization [lead], Writing – original draft [lead], Writing – review & editing [lead]), Mary Ford (Conceptualization [supporting], Formal analysis [equal], Investigation [supporting], Methodology [supporting], Validation [supporting], Visualization [supporting], Writing – original draft [supporting], Writing – review & editing [supporting]), Guillaume Caumon (Conceptualization [supporting], Formal analysis [supporting], Funding acquisition [supporting], Investigation [supporting], Methodology [supporting], Project administration [equal], Software [supporting], Supervision [supporting], Validation [supporting], Visualization [supporting], Writing – original draft [supporting], Writing – review & editing [supporting]), Vitaliy Ogarko (Conceptualization [supporting], Methodology [supporting], Software [supporting], Validation [supporting], Writing – review & editing [supporting]), Lachlan Grose (Conceptualization [supporting], Data curation [supporting], Investigation [supporting], Methodology [supporting], Software [supporting], Validation [supporting], Writing – review & editing [supporting]), Roland Martin (Conceptualization [supporting], Data curation [supporting], Investigation [supporting], Validation [supporting], Writing – review & editing [supporting]), and Paul Cupillard (Conceptualization [supporting], Funding acquisition [supporting], Investigation [supporting], Methodology [supporting], Supervision [supporting], Validation [supporting], Writing – review & editing [supporting])

SUPPORTING INFORMATION

Supplementary data are available at *GJIRAS* online.

ggae192.zip

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

COMPETING INTERESTS

The authors declare that they have no competing interests.

DATA AND CODE AVAILABILITY

The original model of the Pyrenees (Wehr 2017) as provided to JG by the Bureau de Recherches Géologiques et Minières (BRGM, French geological survey) can be obtained upon request. The gravity and geological data used in the inversions presented here can be accessed using (Giraud *et al.* 2023a).

The Python script used for nullspace navigation is made available by Giraud & Ogarko (2024).

Open Source Python libraries used in this work that are mentioned in the text are those of Branch *et al.* (1999) and Silversmith (2021).

Tomofast-x: https://github.com/TOMOFAST/Tomofast-x.

LoopStructural: https://github.com/Loop3D/LoopStructural.

Null-space shuttles: https://github.com/jgeomos/GeoMos-nullspace.

The regional gravity dataset can be obtained from the Bureau Gravimétrique International/International Gravimetry Bureau (https://bgi.obs-mip.fr/).

The ETOPO1 topography model can be accessed at https://catalog.data.gov/dataset/etopo1-1-arc-minute-global-relief-model.

All links or data sets mentioned in this section were last accessed on 22/08/2024.

REFERENCES

Adalsteinsson, D. & Sethian, J.A., 1999. The fast construction of extension velocities in level set methods, *J. Comput. Phys.,* **148,** 2–22.

Barnes, G.J., Lumley, J.M., Houghton, P.I. & Gleave, R.J., 2011. Comparing gravity and gravity gradient surveys, *Geophys. Prospect.,* **59,** 176–187.

Branch, M.A., Coleman, T.F. & Li, Y., 1999. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems, *SIAM J. Sci. Comput.,* **21,** 1–23.

Calcagno, P., Chilès, J.P., Courrioux, G. & Guillen, A., 2008. 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,** 147–157.

Carter-McAuslan, A., Lelièvre, P.G. & Farquharson, C.G., 2015. A study of fuzzy C-means coupling for joint inversion, using seismic tomography and gravity data test scenarios, *Geophysics,* **80,** W1–W15.

Casas, A., Kearey, P., Rivero, L. & Adam, C.R., 1997. Gravity anomaly map of the Pyrenean region and a comparison of the deep geological structure of the western and eastern Pyrenees, *Earth planet. Sci. Lett.,* **150,** 65–78.

Chevrot, S. *et al.*, 2014. High-resolution imaging of the Pyrenees and Massif Central from the data of the PYROPE and IBERARRAY portable array deployments, *J. geophys. Res.,* **119,** 6399–6420.

Chevrot, S. *et al.*, 2018. The non-cylindrical crustal architecture of the Pyrenees, *Sci. Rep.,* **8,** doi:10.1038/s41598-018-27889-x.

Chevrot, S. *et al.*, 2022. Passive imaging of collisional orogens: a review of a decade of geophysical studies in the Pyrénées. *Bulletin de la Société Géologique de France,* **193,** 1–18.

Chevrot, S., Sylvander, M., Diaz, J., Ruiz, M. & Paul, A., 2015. The Pyrenean architecture as revealed by teleseismic P-to-S converted waves recorded along two dense transects, *Geophys. J. Int.,* **200,** 1094–1105.

Chopp, D.L., 1993. Computing minimal surfaces via level set curvature flow, *J. Comput. Phys.,* **106,** 77–91.

Choukroune, P., 1989. The Ecors Pyrenean deep seismic profile reflection data and the overall structure of an orogenic belt, *Tectonics,* **8,** 23–39.

Choukroune, P., Roure, F. & Pinet, B., 1990. Main results of the ECORS Pyrenees profile, *Tectonophysics,* **173,** 411–423.

Colombo, D., Turkoglu, E., Li, W., Sandoval-Curiel, E. & Rovetta, D., 2021. Physics-driven deep-learning inversion with application to transient electromagnetics, *Geophysics,* **86,** E209–E224.

De Wit, R.W.L., Trampert, J. & Van Der Hilst, R.D., 2012. Toward quantifying uncertainty in travel time tomography using the null-space shuttle, *J. geophys. Res.,* **117,** 1–20.

Deal, M.M. & Nolet, G., 1996. Nullspace shuttles, *Geophys. J. Int.,* **124,** 372–380.

Fichtner, A. & Zunino, A. 2019. Hamiltonian nullspace shuttles. *Geophys. Res. Lett.,* **46,** 644–651.

Ford, M. *et al.* 2022. Evolution of a low convergence collisional orogen: a review of Pyrenean orogenesis, *Bulletin de la Société Géologique de France,* **193**(1). doi:10.1051/bsgf/2022018.

Furtney, J. 2015. scikit-fmm: the fast marching method for Python. [online]. Available from: https://github.com/scikit-fmm/scikit-fmm (accessed June 2024).

Galley, C., Lelièvre, P., Haroon, A., Graber, S., Jamieson, J., Szitkar, F., Yeo, I. *et al.*, 2021. Magnetic and Gravity Surface Geometry Inverse Modeling of the TAG Active Mound. *J. Geophys. Res. Solid Earth,* 126. doi:10.1029/2021JB022228.

Galley, C.G., Lelièvre, P.G. & Farquharson, C.G., 2020. Geophysical inversion for 3D contact surface geometry, *Geophysics,* **85,** K27–K45.

García-Senz, J., Pedrera, A., Ayala, C., Ruiz-Constán, A., Robador, A. & Rodríguez-Fernández, L.R., 2020. Inversion of the north Iberian hyperextended margin: the role of exhumed mantle indentation during continental collision, *Geol. Soc. Lond., Spec. Publ.,* **490,** 177–198.

Giraud, J. & Ogarko, V, 2024. Scripts for null space exploration using gravity data: application to field data from the Western Pyrenees and synthetic example. (v1.0), *Zenodo,* doi:10.5281/zenodo.12784672.

Giraud, J., Caumon, G., Grose, L. & Cupillard, P., 2022. Geometrical inversion coupled with automated geological modelling, in *Proceedings of the 83rd EAGE Annual Conference & Exhibition,* pp. 1–5, European Association of Geoscientists & Engineers.

Giraud, J., Caumon, G., Grose, L., Ogarko, V. & Cupillard, P. 2024. Integration of automatic implicit geological modelling in deterministic geophysical inversion. *Solid Earth,* **15,** 63–89.

Giraud, J., Grose, L. & Martin, R., 2023a. Input geophysical and geological data for "Geologically constrained geometry inversion and null-space navigation to explore alternative geological scenarios: a case study in the Western Pyrenees" [Data set], *Zenodo,* doi: 10.5281/zenodo.10255782.

Giraud, J., Ford, M., Caumon, G. & Ogarko, V., 2023b. Supplementary material: animated GIF of null space shuttle with shortening of the subducted Iberian crust [image], https://zenodo.org/record/8263575.

Giraud, J., Lindsay, M. & Jessell, M., 2021a. Generalization of level-set inversion to an arbitrary number of geologic units in a regularized least-squares framework, *Geophysics,* **86,** R623–R637.

Giraud, J., Ogarko, V., Martin, R., Jessell, M. & Lindsay, M., 2021b. Structural, petrophysical, and geological constraints in potential field inversion using the Tomofast-x v1.0 open-source code, *Geosci. Model Dev.,* **14,** 6681–6709.

Giraud, J., Lindsay, M., Jessell, M. & Ogarko, V., 2020. Towards plausible lithological classification from geophysical inversion: honouring geological principles in subsurface imaging, *Solid Earth,* **11,** 419–436.

Giraud, J., Ogarko, V., Lindsay, M., Pakyuz-Charrier, E., Jessell, M. & Martin, R., 2019. Sensitivity of constrained joint inversions to geological and petrophysical input data uncertainties with posterior geological analysis, *Geophys. J. Int.,* **218,** 666–688.

Giraud, J., Pakyuz-Charrier, E., Jessell, M., Lindsay, M., Martin, R. & Ogarko, V., 2017. Uncertainty reduction through geologically conditioned petrophysical constraints in joint inversion, *Geophysics,* **82,** ID19–ID34.

Grose, L., Ailleres, L., Laurent, G. & Jessell, M., 2021. LoopStructural 1.0: time-aware geological modelling, *Geosci. Model Dev.,* **14,** 3915–3937.

Guillen, A., Calcagno, P., Courrioux, G., Joly, A. & Ledru, P., 2008. Geological modelling from field data and geological knowledge. Part II. Modelling validation using gravity and magnetic data inversion, *Phys. Earth planet. Inter.,* **171,** 158–169.

Guo, J., Li, Y., Jessell, M.W., Giraud, J., Li, C., Wu, L., Li, F. & Liu, S., 2021. 3D geological structure inversion from Noddy-generated magnetic data using deep learning methods, *Comput. Geosci.,* **149,** doi:10.1016/j.cageo.2021.104701.

Jammes, S., Manatschal, G., Lavier, L. & Masini, E., 2009. Tectonosedimentary evolution related to extreme crustal thinning ahead of a propagating ocean: example of the western Pyrenees, *Tectonics,* **28**(4), doi:10.1029/2008TC002406.

Kadu, A., van Leeuwen, T. & Mulder, W.A., 2016. Salt reconstruction in full-waveform inversion with a parametric level-set method, *IEEE Trans. Comput. Imaging,* **3,** 305–315.

Lagabrielle, Y. & Bodinier, J.-L., 2008. Submarine reworking of exhumed subcontinental mantle rocks: field evidence from the Lherz peridotites, French Pyrenees, *Terra Nov,* **20,** 11–21.

Lefeuvre, N., Truche, L., Donzé, F.V., Ducoux, M., Barré, G., Fakoury, R.A., Calassou, S. *et al.*, 2021. Native H2 Exploration in the Western Pyrenean Foothills. *Geochemistry, Geophys. Geosystems,* **22,** 1–20.

Lehujeur, M., Chevrot, S., Villaseñor, A., Masini, E., Saspiturry, N., Lescoutre, R. & Sylvander, M., 2021. Three-dimensional shear velocity structure of the Mauléon and Arzacq Basins (Western Pyrenees), *Bulletin de la Société Géologique de France,* **192**(1), doi:10.1051/bsgf/2021039.

Lescoutre, R. & Manatschal, G., 2020. Role of rift-inheritance and segmentation for orogenic evolution: example from the Pyrenean-Cantabrian system, *BSGF - Earth Sci. Bull.,* **191,** doi:10.1051/bsgf/2020021.

Lescoutre, R., 2019. Formation et réactivation du système de rift, nstitut de Physique du Globe de Strasbourg, *Doctoral thesis*, University of Strasbourg.

Li, W., Leung, S. & Qian, J., 2014. A level-set adjoint-state method for crosswell transmission-reflection traveltime tomography, *Geophys. J. Int.,* **199,** 348–367.

Li, W., Lu, W., Qian, J. & Li, Y., 2017. A multiple level-set method for 3D inversion of magnetic data, *Geophysics,* **82,** J61–J81.

Li, W., Qian, J. & Li, Y., 2020. Joint inversion of surface and borehole magnetic data: a level-set approach, *Geophysics,* **85,** J15–J32.

Li, Y., Melo, A., Martinez, C. & Sun, J., 2019. Geology differentiation: a new frontier in quantitative geophysical interpretation in mineral exploration, *Leading Edge,* **38,** 60–66.

Liu, D., Khambampati, A.K. & Du, J., 2018. A parametric level set method for electrical impedance tomography, *IEEE Trans. Med. Imaging,* **37,** 451–460.

Lu, X., Farquharson, C.G. & Lelièvre, P., 2024. Surface geometry inversion of transient electromagnetic data. *Geophysics ,* **89**(4), 1–64.

Maag, E. & Li, Y., 2018. Discrete-valued gravity inversion using the guided fuzzy c -means clustering technique, *Geophysics,* **83,** G59–G77.

Manatschal, G. *et al.* 2021. The role of inheritance in forming rifts and rifted margins and building collisional orogens: a Biscay-Pyrenean perspective, *BSGF—Earth Sci. Bull.,* **192**. doi:10.1051/bsgf/2021042.

Martin, R., Giraud, J., Ogarko, V., Chevrot, S., Beller, S., Gégout, P. & Jessell, M., 2021. Three-dimensional gravity anomaly data inversion in the Pyrenees using compressional seismic velocity model as structural similarity constraints, *Geophys. J. Int.,* **225,** 1063–1085.

Melo, A.T., Sun, J. & Li, Y., 2017. Geophysical inversions applied to 3D geology characterization of an iron oxide copper-gold deposit in Brazil, *GeophysicS,* **82,** K1–K13.

Muir, J.B. & Tsai, V.C., 2019. Geometric and level set tomography using ensemble Kalman inversion, *Geophys. J. Int.,* **220,** 967–980.

Nocedal, J. & Wright, J., 2006. *Numerical Optimization,* Springer New York.

Ogarko, V., Frankcombe, K., Liu, T., Giraud, J., Martin, R. & Jessell, M., 2024. Tomofast-x 2.0: an open-source parallel code for inversion of potential field data with topography using wavelet compression, *Geosci. Model Dev.,* **17,** 2325–2345.

Ogarko, V., Giraud, J., Martin, R. & Jessell, M., 2021. Disjoint interval bound constraints using the alternating direction method of multipliers for geologically constrained inversion: application to gravity data, *Geophysics,* **86,** G1–G11.

Osher, S. & Fedkiw, R., 2003. *Level Set Methods and Dynamic Implicit Surfaces,* eds Antman, S. S., Marsden, J. E. & Sirovitch, L., Springer New York.

Paasche, H. & Tronicke, J., 2007. Cooperative inversion of 2D geophysical data sets: a zonal approach based on fuzzy c -means cluster analysis, *Geophysics,* **72,** A35–A39.

Paige, C.C. & Saunders, M.A., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Softw.,* **8,** 43–71.

Rashidifard, M., Giraud, J., Lindsay, M. & Jessell, M., 2024. Cooperative geophysical inversion integrated with 3-D geological modelling in the Boulia region, QLD, Geophysical Journal International, *Geophys. J. Int.,* **238**(2), 860–880. https://doi.org/10.1093/gji/ggae179.

Rashidifard, M., Giraud, J., Lindsay, M., Jessell, M. & Ogarko, V., 2021. Constraining 3D geometric gravity inversion with a 2D reflection seismic profile using a generalized level set approach: application to the eastern Yilgarn Craton, *Solid Earth,* **12,** 2387–2406.

Roure, F. *et al.*, 1989. Ecors deep seismic data and balanced cross sections: geometric constraints on the evolution of the Pyrenees, *Tectonics,* **8,** 41–50.

Saspiturry, N., Allanic, C., Razin, P., Issautier, B., Baudin, T., Lasseur, E., Serrano, O. & Leleu, S., 2020. Closure of a hyperextended system in an orogenic lithospheric pop-up, Western Pyrenees: the role of mantle buttressing and rift structural inheritance, *Terra Nov,* **32,** 253–260.

Serrano, O., Delmas, J., Hanot, F., Vially, R., Herbin, J., Houel, P. & Tourlière, B., 2006. Le Bassin d'Aquitaine: valorisation des données sismiques, cartographie structurale et potentiel pétrolier, Orleans. [online] Available from: http://infoterre.brgm.fr/rapports/RP-57867-FR.pdf (accessed June 2024).

Sethian, J.A., 1996. A fast marching level set method for monotonically advancing fronts, *Proc. Natl. Acad. Sci.,* **93,** 1591–1595.

Silversmith, W., 2021. cc3d: connected components on multilabel 3D & 2D images, Version 3.2.1, Zenodo, doi:10.5281/zenodo.5719536.

Sun, J. & Li, Y., 2016. Joint inversion of multiple geophysical data using guided fuzzy c -means clustering, *Geophysics,* **81,** ID37–ID57.

Teixell, A., 1996. The Ansó transect of the southern Pyrenees: basement and cover thrust geometries, *J. Geol. Soc. Lond.,* **153,** 301–310.

Teixell, A., Labaume, P., Ayarza, P., Espurt, N., de Saint Blanquat, M. & Lagabrielle, Y., 2018. Crustal structure and evolution of the Pyrenean-Cantabrian belt: a review and new interpretations from recent concepts and data, *Tectonophysics,* **724-725,** 146–170.

Tichadou, C., Godard, M., Muñoz, M., Labaume, P., Vauchez, A., Gaucher, E.C. & Calassou, S., 2021. Mineralogical and geochemical study of serpentinized peridotites from the North-Western Pyrenees: New insights on serpentinization along magma-poor continental passive margins, *Lithos,* **406–407,** 106521, doi: 10.1016/j.lithos.2021.106521.

Tugend, J., Manatschal, G., Kusznir, N.J., Masini, E., Mohn, G. & Thinon, I., 2014. Formation and deformation of hyperextended rift systems: insights from rift domain mapping in the Bay of Biscay-Pyrenees, *Tectonics,* **33,** 1239–1276.

Virtanen, P. *et al.*, 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python, *Nat. Methods,* **17,** 261–272.

Wang, Y. *et al.*, 2016. The deep roots of the western Pyrenees revealed by full waveform inversion of teleseismic P waves, *Geology,* **44,** 475–478.

Wehr, H., 2017. *Modélisation 3D des Pyrénées à partir des données géologiques, gravimétriques et sismiques,* Université Paul Sabatier (Toulouse III). [online] Available from: http://thesesups.ups-tlse.fr/3862/ (accessed June 2024).

Wehr, H., Chevrot, S., Courrioux, G. & Guillen, A., 2018. A three-dimensional model of the Pyrenees and their foreland basins from geological and gravimetric data, *Tectonophysics,* **734-735,** 16–32.

Wellmann, F. & Caumon, G., 2018. 3-D Structural geological models: concepts, methods, and uncertainties, in *Advances in Geophysics,* Vol. **59,** pp. 1–121, ed. Schmelzbach, C., Elsevier Ltd.

Yang, L., Hyde, D., Grujic, O., Scheidt, C. & Caers, J., 2019. Assessing and visualizing uncertainty of 3D geological surfaces using level sets with stochastic motion, *Comput. Geosci.,* **122,** 54–67.

Zheglova, P. & Farquharson, C., 2016. Joint level set inversion of gravity and travel time data: application to mineral exploration, in *SEG Technical Program Expanded Abstracts 2016,* pp. 2165–2169, Society of Exploration Geophysicists.

Zheglova, P., Farquharson, C.G. & Hurich, C.A., 2013. 2-D reconstruction of boundaries with level set inversion of traveltimes, *Geophys. J. Int.,* **192,** 688–698.

Zheglova, P., Lelièvre, P.G. & Farquharson, C.G., 2018. Multiple level-set joint inversion of traveltime and gravity data with application to ore delineation: a synthetic study, *Geophysics,* **83,** R13–R30.

APPENDIX

Table A1. Pseudo-code for the birth of rock unit.

1. Define the range of n_{thresh} thresholds values of $|\nabla\psi^d|$ to test:
$|\nabla\psi^d|_{thresh} = (|\nabla\psi^d|_{thresh})_{min} \dots (|\nabla\psi^d|_{thresh})_{max}$
2. Calculate gravity data misfit and initialize the model
$\boldsymbol{m}_{current} = \boldsymbol{m}$
$\psi^d_{current} = \psi^d(\boldsymbol{d}^{geophy}_{obs.}, \boldsymbol{m}_{current})$
3. Loop over values of $|\nabla\psi^d|$ to test different birth possibilities
for $j = 1 \dots n_{thresh}$:
Define set of model-cells with sufficiently high value of $|\nabla\psi^d|$
$\Omega = 1_{|\nabla\psi^d| \geq \tau_{\nabla\psi^d}(j)}$
Identify subsets of Ω: groups of connected cells from Ω
$\{\Omega_1, \dots, \Omega_{n_{cc}}\} = cc3d(\Omega)$ # *cc3d*: Multilabel Connected Components Labelling in 3-D
loop over the of n_{birth} groups of cells
for $k = 1 \dots n_{cc}$:
conform to geology: identify rock unit with largest overlap with Ω_k
calculate optimum constant density in selected cluster of cells using eq. [above]
$\boldsymbol{v}_{(k)} = solve(\boldsymbol{J}^{\boldsymbol{v}_{(k)}}, \boldsymbol{d}^{geophy}_{obs.}, inequality\ constraints)$
calculate new model $\boldsymbol{m}^k_{birth}(\boldsymbol{m}, \Omega_k, \boldsymbol{v}_{(k)})$ and misfit ψ^d_{new}
$\psi^d_{birth} = \psi^d(\boldsymbol{d}^{geophy}_{obs.}, \boldsymbol{m}^k_{birth})$
Compare misfit with previous trials: keep model with minimum misfit.
If $\psi^d_{clust} \leq \psi^d_{current}$
$\psi^d_{current} \leftarrow \psi^d_{birth}$
$\boldsymbol{m}_{current} \leftarrow \boldsymbol{m}^k_{birth}$
4. Keep the model with the optimum birth