Fast non-linear gravity inversion in spherical coordinates with application to the South American Moho

by Leonardo Uieda and Valéria C. F. Barbosa (2017)

Version of record: doi.org/10.1093/gji/ggw390

Source code: The Python code that produces the results presented here is available under a BSD 3-clause open-source license at github.com/pinga-lab/paper-moho-inversion-tesseroids

Citation:
Uieda, L., and V. C. F. Barbosa (2017), Fast nonlinear gravity inversion in spherical coordinates with application to the South American Moho, *Geophysical Journal International*, 208(1), 162-176, doi:10.1093/gji/ggw390

Geophysical Journal International

Geophys. J. Int. (2017) **208,** 162–176
Advance Access publication 2016 October 17
GJI Gravity, geodesy and tides

doi: 10.1093/gji/ggw390

Fast nonlinear gravity inversion in spherical coordinates with application to the South American Moho

Leonardo Uieda[1,2] and Valéria C.F. Barbosa[2]

[1]*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil. E-mail: leouieda@gmail.com*
[2]*Observatório Nacional, Rio de Janeiro, Brazil*

Accepted 2016 October 13. Received 2016 October 7; in original form 2016 March 24

SUMMARY

Estimating the relief of the Moho from gravity data is a computationally intensive nonlinear inverse problem. What is more, the modelling must take the Earths curvature into account when the study area is of regional scale or greater. We present a regularized nonlinear gravity inversion method that has a low computational footprint and employs a spherical Earth approximation. To achieve this, we combine the highly efficient Bott's method with smoothness regularization and a discretization of the anomalous Moho into tesseroids (spherical prisms). The computational efficiency of our method is attained by harnessing the fact that all matrices involved are sparse. The inversion results are controlled by three hyperparameters: the regularization parameter, the anomalous Moho density-contrast, and the reference Moho depth. We estimate the regularization parameter using the method of hold-out cross-validation. Additionally, we estimate the density-contrast and the reference depth using knowledge of the Moho depth at certain points. We apply the proposed method to estimate the Moho depth for the South American continent using satellite gravity data and seismological data. The final Moho model is in accordance with previous gravity-derived models and seismological data. The misfit to the gravity and seismological data is worse in the Andes and best in oceanic areas, central Brazil and Patagonia, and along the Atlantic coast. Similarly to previous results, the model suggests a thinner crust of 30–35 km under the Andean foreland basins. Discrepancies with the seismological data are greatest in the Guyana Shield, the central Solimões and Amazonas Basins, the Paraná Basin, and the Borborema province. These differences suggest the existence of crustal or mantle density anomalies that were unaccounted for during gravity data processing.

Key words: Inverse theory; Satellite gravity; Gravity anomalies and Earth structure; South America.

1 INTRODUCTION

The Mohorovičić discontinuity (or Moho) that marks the transition from the crust to the mantle, is studied almost exclusively through indirect geophysical methods. The two main geophysical methods used to estimate the depth of the Moho are seismology, with both natural and controlled sources, and gravimetry. With the advent of satellite gravimetry missions like GRACE and GOCE, gravity-derived crustal models can be produced in regional or global scales (e.g. Reguzzoni *et al.* 2013; van der Meijde *et al.* 2013, 2015). New spherical harmonic gravity models that use these satellite observation, like GOCO5S (Mayer-Guerr *et al.* 2015), provide almost homogeneous data coverage in difficult to access regions traditionally poor in terrestrial data. An example is South America, where seismological and terrestrial gravity data are traditionally concentrated around urban centres and coastal areas, resulting in large areas (e.g. forests and mountains) devoid of data.

Estimating Moho depth from gravity data is a nonlinear inverse problem. One can generalize this problem of estimating the depths of an interface separating two media, such as the sediment-basement interface of a sedimentary basin or the crust-mantle interface (Moho). Several methods have been developed over the years to solve this inverse problem, for example, Bott (1960); Barbosa *et al.* (1997, 1999a,b); Barnes & Barraud (2012); Leão *et al.* (1996); Martins *et al.* (2010, 2011); Oldenburg (1974); Reguzzoni *et al.* (2013); Santos *et al.* (2015); Silva *et al.* (2006, 2014), to name a few. Solving the inverse problem is computationally demanding because it requires the construction of large dense matrices and the solution of large linear systems. As a result, some authors search for ways to increase the computational efficiency of this class of inverse problem. Bott (1960) proposed a method based on iteratively applying

corrections to a starting estimate based on the inversion residuals. The algorithm is fast because it bypasses the construction and solution of linear systems and only involves forward modelling. Oldenburg (1974) showed that the fast FFT-based forward modelling of Parker (1973) could be rearranged to estimate the relief. Barnes & Barraud (2012) use a form of adaptive discretization to compute the Jacobian, or sensitivity, matrix. For each data point, the discretization will be progressively coarser the further way from the point. This reduces the matrix and, consequently, the linear systems to a sparse form that can be solved efficiently. Recently, Silva *et al.* (2014) extended and generalized the original method of Bott (1960) and Santos *et al.* (2015) used this extension to estimate a basement relief with sharp boundaries.

A spherical Earth approximation is preferred when estimating the Moho depth from gravity data in continental and global scale studies. Wieczorek & Phillips (1998) developed a spherical harmonic equivalent of the Parker-Oldenburg FFT algorithm and applied it to estimate the crustal structure of the Moon. Reguzzoni *et al.* (2013) use a spherical Earth approximation to estimate the global Moho relief using data from the GOCE satellite mission. Another approach is to use non-spectral (space domain) gravity inversion methods. Many such methods were developed for estimating the basement relief of a sedimentary basin (e.g. Barbosa *et al.* 1997, 1999a,b; Martins *et al.* 2010, 2011; Sun & Li 2014). These methods approximate the sedimentary pack by a set of juxtaposed right-rectangular prisms. The top of the prisms coincide with the Earth's surface and the prisms' thicknesses represent the depths to the basement and are the parameters to be estimated in the inversion. The use of rectangular prisms implies a planar Earth approximation and may not be adequate for depth-to-Moho estimates in continental- or global-scale study. A straightforward way to circumvent this hindrance is to adapt one of the methods developed for rectangular prisms to use tesseroids (spherical prisms). One of the difficulties of this approach is that the forward problem for a tesseroid must be solved numerically. Two alternatives proposed in the literature to the numerical solution are Taylor series expansion (Heck & Seitz 2007; Grombein *et al.* 2013) and the Gauss-Legendre Quadrature (Asgharzadeh *et al.* 2007). Numerical experiments by Wild-Pfeiffer (2008) suggest that the Gauss-Legendre Quadrature (GLQ) offers superior results. However, the GLQ suffers from numerical instability when the computation point is close to the tesseroid (Asgharzadeh *et al.* 2007). To overcome the numerical instability, Li *et al.* (2011) proposed an adaptive discretization algorithm which was later improved upon by Uieda *et al.* (2016).

In any gravity inversion for estimating the relief of an interface, two hyperparameters control the inversion results: the density-contrast between the two media and the reference level around which the interface undulates. The reference level is the constant depth of the Normal Earth Moho in the case of the anomalous Moho. For regularized inversions, an additional hyperparameter is the regularization parameter that balances the relative importance between the data-misfit measure and the regularizing function. The two most commonly used methods for estimating the regularization parameter are the L-curve criterion and Generalized Cross Validation (GCV). Farquharson & Oldenburg (2004) provide for a thorough comparison of both methods. Estimating the density-contrast in a sedimentary basin context has been tackled by Silva *et al.* (2006) and Martins *et al.* (2010) when the basement depth is known at a few points. To the authors' knowledge no attempt has been made to estimate the reference level.

We present a nonlinear gravity inversion to estimate the Moho depth in a spherical Earth approximation. Our method is based on the Silva *et al.* (2014) Gauss-Newton formulation of the method of Bott (1960). We use tesseroids to discretize the anomalous Moho and the adaptive discretization algorithm of Uieda *et al.* (2016) for the forward modelling. The stability of the inversion is achieved through smoothness regularization. In order to maintain the computational efficiency of Bott's method, we exploit the sparse nature of all matrices involved in the computations. We employ a variant of GCV known as hold-out cross-validation (Kim 2009) to estimate the regularization parameter. Additionally, we estimate the density-contrast and reference level simultaneously in a second validation step using knowledge of the Moho depth at a few points, similarly to Silva *et al.* (2006) and Martins *et al.* (2010). Finally, we apply the proposed method to estimate the Moho depth for South America using gravity data from the GOCO5S model (Mayer-Guerr *et al.* 2015) and the seismological data of Assumpção *et al.* (2013).

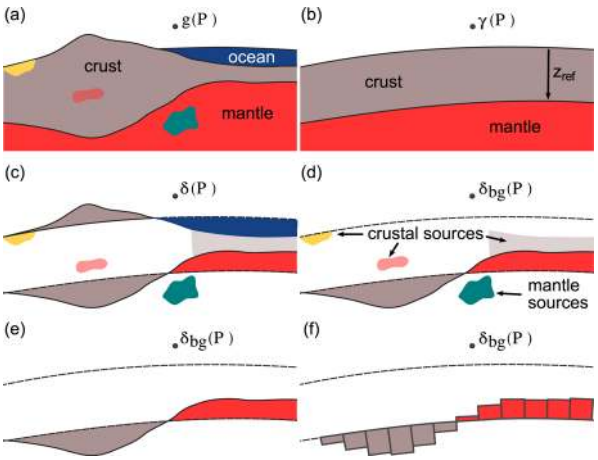

Figure 1. Sketch of the stages in gravity data correction and the discretization of the anomalous Moho relief using tesseroids. (a) The Earth and the measured gravity at point P ($g(P)$). (b) The Normal Earth and the calculated normal gravity at point P ($\gamma(P)$). $z_{\rm ref}$ is the depth of the Normal Earth Moho. (c) The gravity disturbance ($\delta(P)$) and the corresponding density anomalies after removal of the normal gravity: topography, oceans, crustal and mantle heterogeneities, and the anomalous Moho. (d) The Bouguer disturbance ($\delta_{\rm bg}(P)$) after topographic correction and the remaining density anomalies. (e) All density anomalies save the anomalous Moho are assumed to have been removed before inversion. (f) The discretization of the anomalous Moho in tesseroids. Grey tesseroids will have a negative density contrast while red tesseroids will have a positive one.

2 METHODOLOGY

In potential field methods, we must isolate the target anomalous density distribution before modelling and inversion. In our case, the target is the relief of the real Moho undulating around a reference Moho. We do this by removing all other effects from the gravity observations. The first correction is to remove the scalar gravity of an ellipsoidal reference Earth (the Normal Earth), hereafter denoted as γ. This effect is calculated on the same point P where the gravity observation was made (Figs 1a and b). $\gamma(P)$ is calculated using the closed-form solution presented by Li & Götze (2001). The difference between the observed gravity at point P ($g(P)$) and Normal gravity at the same point is known as the gravity disturbance,

$$\delta(P) = g(P) - \gamma(P). \quad (1)$$

The disturbance contains only the gravitational effects of density distributions that are anomalous with respect to the Normal Earth

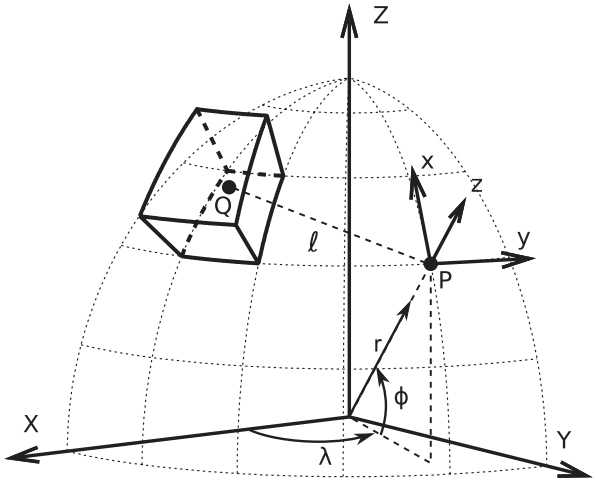

Figure 2. Sketch of a tesseroid (spherical prism) in a geocentric coordinate system (X, Y, Z). Observations are made at point P with respect to its local north-oriented coordinate system (x, y, z). After Uieda (2015).

(see Fig. 1c). This includes all masses above the surface of the ellipsoid (the topography), the mass deficiency of the oceans, the mass deficiency of sedimentary basins, crustal sources (e.g. igneous intrusions, lateral density changes, etc.), heterogeneities below the upper mantle, and the effect of the difference between the real Moho topography and the Moho of the Normal Earth.

To estimate the anomalous Moho relief from gravity data, we must first isolate its gravitational attraction. Thus, all other gravitational effects must be either removed or assumed negligible. Here, we will remove the gravitational effect produced by the known topography and ocean masses to obtain the full Bouguer disturbance (Fig. 1d),

$$\delta_{\text{bg}}(P) = \delta(P) - g_{\text{topo}}(P). \tag{2}$$

We will also remove the gravitational effect of know sedimentary basins but assume that the effects of other crustal and mantle sources are negligible. Thus, the only effect left will be that of the anomalous Moho relief (Fig. 1e). The gravitational attraction of the topography, oceans, and basins are calculated in a spherical Earth approximation by forward modelling using tesseroids (Fig. 2). The tesseroid effects are calculated numerically using GLQ integration (Asgharzadeh *et al.* 2007). The accuracy of the GLQ integration is improved by the adaptive discretization scheme of Uieda *et al.* (2016).

2.1 Parametrization and the forward problem

We parameterize the forward problem by discretizing the anomalous Moho into a grid of $M_{\text{lon}} \times M_{\text{lat}} = M$ juxtaposed tesseroids (Fig. 1f). The true (real Earth) Moho varies in depth with respect to the Moho of the Normal Earth. Hereafter we will refer to the depth of the Normal Earth Moho as z_{ref} (see Fig. 1b). If the true Moho is above z_{ref}, the top of the kth tesseroid is the Moho depth z_k, the bottom is z_{ref}, and the density-contrast ($\Delta\rho$) is positive (red tesseroids in Fig. 1f). If the Moho is below z_{ref}, the top of the tesseroid is z_{ref}, the bottom is z_k, and $\Delta\rho$ is negative (grey tesseroids in Fig. 1f).

Considering that the absolute value of the density-contrasts of the tesseroids is a fixed parameter, the predicted gravity anomaly of the Moho is a nonlinear function of the parameters z_k, $k = 1, \ldots, M$,

$$d_i = f_i(\mathbf{p}), \tag{3}$$

in which d_i is the ith element of the N-dimensional predicted data vector $\mathbf{d}$, $\mathbf{p}$ is the M-dimensional parameter vector containing the M Moho depths (z_k), and f_i is the ith nonlinear function that maps the parameters onto the data. The functions f_i are the radial component of the gravitational attraction of the tesseroid Moho model.

2.2 Inverse problem

We wish to estimate the parameter vector $\mathbf{p}$ from a set of observed gravity data $\mathbf{d}^{\text{o}}$. The least-squares estimate is the one that minimizes the data-misfit function

$$\phi(\mathbf{p}) = [\mathbf{d}^{\text{o}} - \mathbf{d}(\mathbf{p})]^T [\mathbf{d}^{\text{o}} - \mathbf{d}(\mathbf{p})]. \tag{4}$$

Function $\phi(\mathbf{p})$ is nonlinear with respect to $\mathbf{p}$. Thus, we can determine its minimum using gradient-based iterative optimization methods like Gauss–Newton or Steepest Descent. Such methods start from an initial approximation to the model parameter vector $\mathbf{p}^0$ and estimate a parameter perturbation vector $\Delta\mathbf{p}^0$. The perturbation vector is used to update $\mathbf{p}^0$ to $\mathbf{p}^1 = \mathbf{p}^0 + \Delta\mathbf{p}^0$. This procedure is repeated until a minimum of function $\phi(\mathbf{p})$ (eq. 4) is reached.

For the Gauss-Newton method, the parameter perturbation vector at the kth iteration $\Delta\mathbf{p}^k$ is obtained by solving the linear system

$$\mathbf{H}^k \Delta\mathbf{p}^k = -\nabla\phi^k, \tag{5}$$

in which $\nabla\phi^k$ and $\mathbf{H}^k$ are, respectively, the gradient vector and the Hessian matrix of $\phi(\mathbf{p})$.

The gradient vector and the Gauss-Newton approximation of the Hessian matrix of $\phi(\mathbf{p})$ are, respectively,

$$\nabla\phi^k = -2\mathbf{A}^{k\,T}[\mathbf{d}^{\text{o}} - \mathbf{d}(\mathbf{p}^k)], \tag{6}$$

and

$$\mathbf{H}^k \approx 2\mathbf{A}^{k\,T}\mathbf{A}^k, \tag{7}$$

in which $\mathbf{A}^k$ is the $N \times M$ Jacobian or sensitivity matrix whose elements are

$$A^k_{ij} = \frac{\partial f_i}{\partial p_j}(\mathbf{p}^k). \tag{8}$$

2.3 Regularization

Nonlinear gravity inversions for estimating the relief of an interface separating two media (like the Moho) are ill-posed and require additional constraints in the form of regularization (Silva *et al.* 2001). A common approach is to use the first-order Tikhonov regularization (Tikhonov & Arsenin 1977) to impose smoothness on the solution. The cost function for smoothness regularization is given by

$$\theta(\mathbf{p}) = \mathbf{p}^T\mathbf{R}^T\mathbf{R}\mathbf{p}, \tag{9}$$

where $\mathbf{R}$ is an $L \times M$ finite-difference matrix representing L first-order differences between the depths of adjacent tesseroids.

To transform the ill-posed inverse problem into a well-posed one via Tikhonov regularization, we adopted the well-established procedure of formulating a constrained inverse problem that is solved by minimizing an unconstrained goal function

$$\Gamma(\mathbf{p}) = \phi(\mathbf{p}) + \mu\theta(\mathbf{p}), \tag{10}$$

in which μ is the regularization parameter that controls the balance between fitting the observed data and obeying the smoothness constraint imposed by the regularizing function $\theta(\mathbf{p})$ (eq. 9).

The goal function $\Gamma(\mathbf{p})$ is also nonlinear with respect to $\mathbf{p}$ and can be minimized using the Gauss-Newton method. The gradient vector and Hessian matrix of the goal function are, respectively,

$$\nabla\Gamma^k = -2\mathbf{A}^{k\,T}[\mathbf{d}^{\mathrm{o}} - \mathbf{d}(\mathbf{p}^k)] + 2\mu\mathbf{R}^T\mathbf{R}\mathbf{p}^k, \tag{11}$$

and

$$\mathbf{H}^k = 2\mathbf{A}^{k\,T}\mathbf{A}^k + 2\mu\mathbf{R}^T\mathbf{R}. \tag{12}$$

At the kth iteration, the parameter perturbation vector $\Delta\mathbf{p}^k$ is obtained by solving the linear equation system

$$\left[\mathbf{A}^{k\,T}\mathbf{A}^k + \mu\mathbf{R}^T\mathbf{R}\right]\Delta\mathbf{p}^k = \mathbf{A}^{k\,T}[\mathbf{d}^{\mathrm{o}} - \mathbf{d}(\mathbf{p}^k)] - \mu\mathbf{R}^T\mathbf{R}\mathbf{p}^k. \tag{13}$$

Estimating the Moho depths using the above equations is computationally costly because of two main factors: (1) the evaluation and storage of the dense $N \times M$ Jacobian matrix $\mathbf{A}^{\mathrm{k}}$ and (2) the solution of the resulting $M \times M$ equation system. In practice, the derivatives in the Jacobian (eq. 8) are often calculated through a first-order finite-difference approximation. Thus, evaluating $\mathbf{A}^{\mathrm{k}}$ requires $2 \times M \times N$ forward modelling operations for each iteration of the gradient descent algorithm. These computations are performed for each iteration of the optimization of the goal function $\Gamma(\mathbf{p})$.

2.4 Bott's method

Bott (1960) developed an efficient method to determine the depth of the basement of a sedimentary basin from gravity observations. The method requires data on a regular grid of $N_x \times N_y = N$ observations. The basement relief is then discretized into an equal grid of $M_x \times M_y = M$ elements with $M_x = N_x$ and $M_y = N_y$. Bott's iterative method starts with an initial approximation of the basement depths $\mathbf{p}^0$ equal to the null vector. The method updates the approximation by calculating a parameter perturbation vector $\Delta\mathbf{p}^k$ using the formula

$$\Delta\mathbf{p}^k = \frac{\mathbf{d}^{\mathrm{o}} - \mathbf{d}(\mathbf{p}^k)}{2\pi G\Delta\rho}, \tag{14}$$

in which G is the gravitational constant and $\Delta\rho$ is the contrast between the density of the sediments and the reference density. The iterative process stops when the inversion residuals $\mathbf{r}^k = \mathbf{d}^{\mathrm{o}} - \mathbf{d}(\mathbf{p}^k)$ fall below the assumed noise level of the data.

Silva *et al.* (2014) showed that Bott's method can be formulated as a special case of the Gauss–Newton method (eq. 5) by setting the Jacobian matrix (eq. 8) to

$$\mathbf{A} = 2\pi G\Delta\rho\mathbf{I}, \tag{15}$$

where $\mathbf{I}$ is the identity matrix. In this framework, Bott's method uses a Bouguer plate approximation of the gravitational effect of the relief, $d_i = 2\pi G\Delta\rho\, p_i$. The derivative of d_i with respect to the parameter p_i is $2\pi G\Delta\rho$, thus linearizing the Jacobian matrix. However, the nonlinearity of the predicted data $\mathbf{d}(\mathbf{p}^k)$ is preserved.

One of the advantages of Bott's method over the traditional Gauss–Newton or Steepest Descent is the elimination of the computation and storage of the dense Jacobian matrix $\mathbf{A}^k$. Furthermore, Bott's method also does not require the solution of equation systems. However, a disadvantage of Bott's method is that it suffers from instability (Silva *et al.* 2014). A common approach to counter this issue is to apply a smoothing filter after the inversion to the unstable estimate, as in Silva *et al.* (2014).

2.5 Combining Bott's method, regularization and tesseroids

We propose a regularized version of Bott's method to invert gravity data for estimating the depth of the Moho in spherical coordinates. To adapt Bott's method to spherical coordinates, we replace the right-rectangular prisms in the forward modelling ($\mathbf{d}(\mathbf{p}^k)$ in eq. 14) with tesseroids. The tesseroid forward modelling uses the adaptive discretization algorithm of Uieda *et al.* (2016) to achieve accurate results. Furthermore, our formulation maintains the regularized solution for the Gauss-Newton method (eq. 13) but replaces the full Jacobian matrix with the Bouguer plate approximation. Here, the Jacobian matrix is replaced by a diagonal matrix (eq. 15) whose elements are invariant along successive iterations. Using this approximation eliminates the cost of computing and storing the full $N \times M$-dimensional Jacobian matrix $\mathbf{A}^k$ at each iteration (eq. 8). Traditionally, the full Jacobian matrix is computed using a first-order finite difference scheme, which requires $2 \times N \times M$ forward modelling operations per iteration. Using eq. (15) requires N multiplications that need only be performed once. This provides a considerable speed gain.

Matrix arithmetic operations can be performed efficiently by taking advantage of the sparse nature of matrices $\mathbf{A}$ and $\mathbf{R}$ (respectively, eqs 15 and 9). The same is true for solving the equation system in the Gauss–Newton method (eq. 13). However, the computational cost of forward modelling is still present. Particularly, forward modelling using tesseroids is more computationally intensive than using right-rectangular prisms because of the numerical integration and adaptive discretization (Uieda *et al.* 2016). We show later in this article that sparse matrix multiplications and solving the sparse linear system in eq. (13) account for less than 0.1 per cent of the computation time required for a single inversion.

The main advantage of our formulation is that it retains the efficiency of Bott's method while stabilizing the solution through the well-established formalism of Tikhonov regularization. For example, the total variation approach used by Martins *et al.* (2011) could potentially be implemented in a more straight forward manner than done by Santos *et al.* (2015).

2.6 Estimating the inversion hyperparameters

Parameters that influence the inversion result but are not estimated directly in the inversion are known as hyperparameters. In the case of our regularized Moho depth inversion, the hyperparameters are the regularization parameter μ (eq. 10), the Moho density-contrast $\Delta\rho$ (eq. 15), and the depth of the Normal Earth Moho, or reference level, z_{ref} (Fig. 1b).

We estimate these hyperparameters in two steps. First, we assume fixed values for z_{ref} and $\Delta\rho$ and perform a hold-out cross-validation procedure (Hansen 1992) to estimate an optimal value for μ. Our investigations suggest that the optimal value of μ does not depend on the particular values of z_{ref} and $\Delta\rho$ used. Second, we use the estimated μ to perform a validation procedure to estimate z_{ref} and $\Delta\rho$. The final outcomes of both steps are the values of the three hyperparameters and the final estimated Moho depths.

2.6.1 Estimating the regularization parameter

The regularization parameter μ controls how much smoothness is applied to the inversion result. An optimal value of μ will

stabilize and smooth the solution while not compromising the fit to the observed data. Two widely used methods to estimate an optimal μ are the L-curve criterion and cross-validation (Hansen 1992). Here, we will adopt the hold-out method of cross-validation (Kim 2009). The hold-out method consists of splitting the observed data set into two independent parts: a training set $\mathbf{d}^{\mathrm{o}}_{\mathrm{inv}}$ and a testing set $\mathbf{d}^{\mathrm{o}}_{\mathrm{test}}$. The training set is used in the inversion while the testing set is kept back and used to judge the quality of the chosen value of μ. For a value of the regularization parameter μ_n, the training set is inverted using μ_n to obtain an estimate $\hat{\mathbf{p}}^n$. This estimate is used to calculate predicted data on the same points as the testing set via forward modelling

$$\mathbf{d}^n_{\mathrm{test}} = \mathbf{f}(\hat{\mathbf{p}}^n). \tag{16}$$

The metric chosen to evaluate μ_n is the mean square error (MSE) of the misfit between the observed and predicted testing data sets,

$$\mathrm{MSE}_n = \frac{\|\mathbf{d}^{\mathrm{o}}_{\mathrm{test}} - \mathbf{d}^n_{\mathrm{test}}\|^2}{N_{\mathrm{test}}}, \tag{17}$$

in which N_{test} is the number of data in the testing set. The optimal value of μ will be the one that minimizes the MSE, that is, the one that best predicts the testing data. We emphasize that the inversion is performed on the training data set only.

The algorithm for the hold-out cross-validation is summarized as follows:

(i) Divide the observed data into the training ($\mathbf{d}^o_{\mathrm{inv}}$) and testing ($\mathbf{d}^o_{\mathrm{test}}$) sets.
(ii) For each $\mu_n \in [\mu_1, \mu_2, \ldots, \mu_{N_\mu}]$:

(a) Estimate $\hat{\mathbf{p}}^n$ by inverting the training set $\mathbf{d}^o_{\mathrm{inv}}$.
(b) Use $\hat{\mathbf{p}}^n$ to calculate the predicted testing set $\mathbf{d}^n_{\mathrm{test}}$ using eq. (16).
(c) Calculate the mean square error MSE_n using eq. (17).

(iii) The final solution is the $\hat{\mathbf{p}}^n$ corresponding to the smallest MSE_n.

The separation of the training and testing data sets is commonly done by taking random samples from the full data set. However, we cannot perform the separation in this way because Bott's method requires data on a regular grid as well as having model elements directly below each data point. Thus, we take as our training set the points from the observed data grid that fall on a similar grid but with twice the grid spacing (open circles in Fig. 3). All other points from the original data grid make up the testing data set (black dots in Fig. 3). This separation will lead to a testing data set with more points than the training data set. A way to balance this loss of data in the inversion is to generate a data grid with half of the desired grid spacing, either through interpolation or from a spherical harmonic model.

2.6.2 *Estimating z_{ref} and $\Delta\rho$*

The depth of the Normal Earth Moho (z_{ref}) and the density-contrast of the anomalous Moho ($\Delta\rho$) are other hyperparameters of the inversion. That is, their value influences the final solution but they are not estimated during the inversion. Both hyperparameters cannot be determined from the gravity data alone. Estimating z_{ref} and $\Delta\rho$ requires information that is independent of the gravity data, such as knowledge of the parameters (Moho depths) at certain points. This information can be used in a manner similar to the cross-validation described in the previous section. In this study, we use

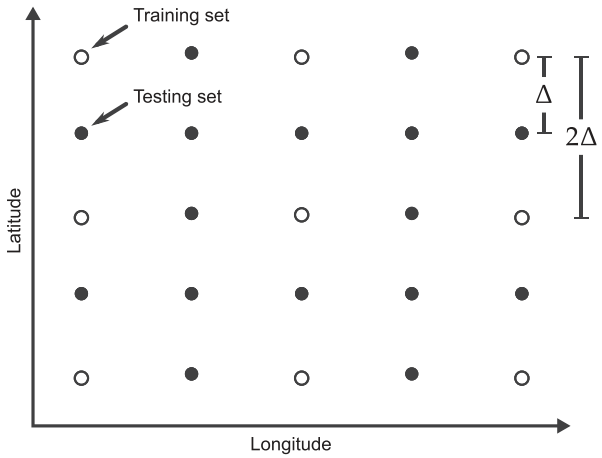

Figure 3. Sketch of a data grid separated into the training (open circles) and testing (black dots) data sets. The training data set is still displayed on a regular grid but with twice the grid spacing of the original data grid.

point estimates of the Moho depth to determine the optimal values of z_{ref} and $\Delta\rho$. These points will generally come from seismologic studies, like receiver functions, surface wave dispersion, and deep refraction experiments.

Let $\mathbf{z}^o_s$ be a vector of N_s known Moho depths. We use the MSE as a measure of how well a given inversion output $\hat{\mathbf{p}}^{l,m}$ fits the know depths. The optimal values of z_{ref} and $\Delta\rho$ are the ones that best fit the independent known Moho depths (i.e. produce the smallest MSE). However, the points do not necessarily coincide with the model elements of the inversion. Before computing the MSE, we interpolate $\hat{\mathbf{p}}^{l,m}$ on the known points to obtain the predicted depths $\mathbf{z}^{l,m}_s$. The MSE is defined as

$$\mathrm{MSE} = \frac{\|\mathbf{z}^{\mathrm{o}}_s - \mathbf{z}^{l,m}_s\|^2}{N_s}. \tag{18}$$

The algorithm for estimating z_{ref} and $\Delta\rho$ is:

(i) For every combination of $z_{\mathrm{ref},l} \in [z_{\mathrm{ref},1}, z_{\mathrm{ref},2}, \ldots, z_{\mathrm{ref},N_z}]$ and $\Delta\rho_m \in [\Delta\rho_1, \Delta\rho_2, \ldots, \Delta\rho_{N_\rho}]$:

(a) Perform the inversion on the training data set $\mathbf{d}^{\mathrm{o}}_{\mathrm{inv}}$ using $z_{\mathrm{ref},l}$, $\Delta\rho_m$, and the previously estimated value of μ. The inversion output is the vector $\hat{\mathbf{p}}^{l,m}$.
(b) Interpolate $\hat{\mathbf{p}}^{l,m}$ on the known points to obtain the predicted depths $\mathbf{z}^{l,m}_s$.
(c) Calculate the MSE between $\mathbf{z}^o_s$ and $\mathbf{z}^{l,m}_s$ using eq. (18).

(ii) The final solution is the $\hat{\mathbf{p}}^{l,m}$ corresponding to the smallest MSE.

A similar approach was used by Silva *et al.* (2006) and Martins *et al.* (2010) to estimate the parameters defining the density-contrast variation with depth of a sedimentary basin. van der Meijde *et al.* (2013) also had a similar methodology for dealing with the hyperparameters, though in a less formalized way.

2.7 Software implementation

The inversion method proposed here is implemented in the Python programming language. The software is freely available under the terms of the BSD 3-clause open-source software license. Our implementation relies on the open-source libraries scipy and numpy (Jones *et al.* 2001, http://scipy.org)

for array-based computations, matplotlib (Hunter 2007, http://matplotlib.org) and seaborn (Waskom *et al.* 2015, http://stanford.edu/~mwaskom/software/seaborn) for plots and maps, and Fatiando a Terra (Uieda *et al.* 2013, http://www.fatiando.org) for geophysics specific tasks, particularly for forward modelling using tesseroids. We use the scipy.sparse package for sparse matrix arithmetic and linear algebra. The sparse linear system in eq. (13) is solved using the conjugate gradient method implemented in scipy.sparse.

The computational experiments (e.g. data processing, synthetic tests, real data application) were performed in Jupyter (formerly IPython) notebooks (Pérez & Granger 2007, http://jupyter.org/). The notebook files combine the source code used to run the experiments, the results and figures generated by the code, and rich text to explain and document the analysis.

All source code, Jupyter notebooks, data, and model results are made available through an online repository (Uieda & Barbosa 2016, http://dx.doi.org/10.6084/m9.figshare.3987267 or https://github.com/pinga-lab/paper-moho-inversion-tesseroids). The repository also contains instructions for replicating all results presented here.

3 APPLICATION TO SYNTHETIC DATA

We test and illustrate the proposed inversion method by applying it to two noise-corrupted synthetic data sets. The first one is generated by a simple Moho model simulating the transition from a thicker continental crust to a thinner oceanic crust. This application uses cross-validation to estimate the regularizing parameter (μ) while assuming that the anomalous Moho density-contrast ($\Delta\rho$) and the Normal Earth Moho depth (z_{ref}) are known quantities. This first test is simplified in order to investigate solely the efficiency of the inversion and the cross-validation procedure to estimate μ. The second data set is generated by a more complex model derived from the South American portion of the global CRUST1.0 model (Laske *et al.* 2013). This second application uses cross-validation to estimate μ and the validation procedure using synthetic seismological data to estimate $\Delta\rho$ and z_{ref}. The model and corresponding synthetic data are meant to simulate with more fidelity the real data application.

3.1 Simple model

We simulate the transition from a continental-type Moho to an oceanic-type Moho using a model composed of $M_{\text{lat}} \times M_{\text{lon}} = 40 \times 50$ grid of juxtaposed tesseroids (a total of $M = 2000$ model elements). The anomalous Moho density-contrast is $\Delta\rho = 400\,\text{kg m}^{-3}$ and the Normal Earth Moho depth is $z_{\text{ref}} = 30\,\text{km}$. Fig. 4(a) shows the model Moho depths where we can clearly see an eastward crustal thinning. In Fig. 4(a), each pixel in the pseudo-colour image corresponds to a tesseroid of the model.

The synthetic data were forward modelled on a regular grid of $N_{\text{lat}} \times N_{\text{lon}} = 79 \times 99$ points (a total of $N = 7821$ observations) at a constant height of 50 km. The data were contaminated with pseudo-random noise sampled from a normal distribution with zero mean and 5 mGal standard deviation. Fig. 4(b) shows the noise-corrupted full synthetic data set exhibiting an eastward increase due to the simulated eastward crustal thinning shown in Fig. 4(a). The data grid spacing is half the grid spacing of the tesseroid model so that, when separating the training and testing data sets (Fig. 3), the training data set points will fall directly above each model element.

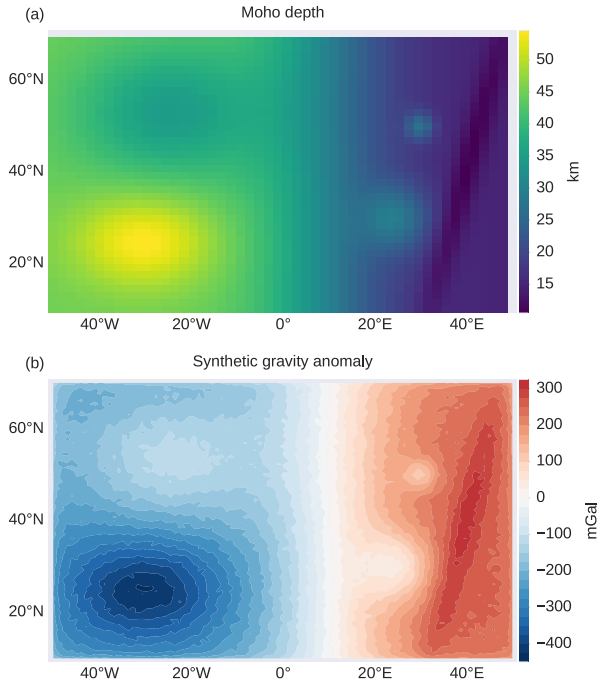

Figure 4. A simple Moho model made of tesseroids for synthetic data application. (a) The Moho depth of the model in kilometres. The model transitions from a deep Moho in the right to a shallow Moho in left, simulating the transition between a continental and an oceanic Moho. Each pixel in the pseudo-colour image corresponds to a tesseroid of the model. (b) Noise-corrupted synthetic gravity data generated from the model shown in (a).

We separated the synthetic data into training and testing data sets following Fig. 3. The training data set is a regular grid of $N_{\text{lat}} \times N_{\text{lon}} = 40 \times 50$ points (a total of $N_{\text{train}} = 2000$). The testing data set is composed of $N_{\text{test}} = 5821$ observations. We used cross-validation to estimate an optimal regularization parameter (μ) from a set of $N_{\mu} = 16$ values equally spaced on a logarithmic scale between 10^{-6} and 10^{-1}. We ran our regularized inversion on the training data set for each value of μ, obtaining 16 Moho depth estimates. For all inversions, the initial Moho depth estimate used to start the Gauss-Newton optimization was set to 60 km depth for all inversion parameters. Furthermore, z_{ref} and $\Delta\rho$ are set to their respective true values. Finally, we computed the MSE (eq. 17) for each estimate and chose as the final estimated Moho model the one that minimizes the MSE.

Fig. 5(a) shows the final estimated Moho depth after the cross-validation. The recovered model is smooth, indicating that the cross-validation procedure was effective in estimating an optimal regularization parameter. Fig. 5(b) shows difference between the true Moho depth (Fig. 4a) and the estimated Moho depth. The differences appear to be semi-randomly distributed with a maximum coinciding with a short-wavelength feature in the true model. The maximum and minimum differences are approximately 2.19 and −2.13 km, respectively. Fig. 5(c) shows inversion residuals, defined as the difference between the observed and predicted data (in mGal). The largest residual (in absolute value) coincides with the largest difference between the true model and the estimate. The inversion residuals are normally distributed, as shown in Fig. 5(d), with 0.02 mGal mean and a standard deviation of 3.63 mGal. The cross-validation curve in Fig. 5(e) shows a clear minimum MSE

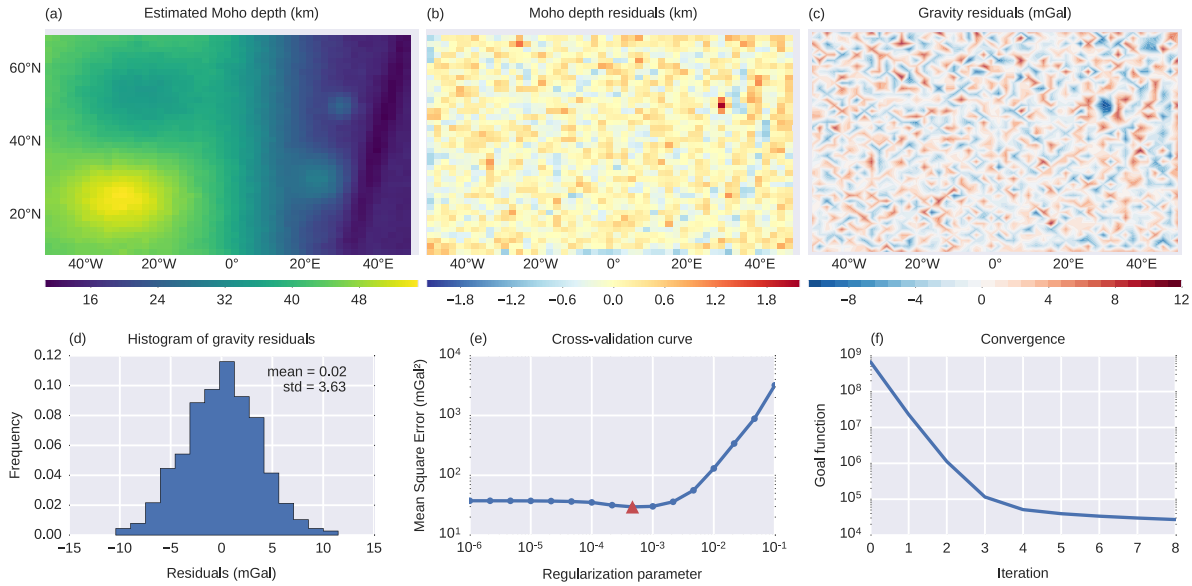

Figure 5. Results from the inversion of the simple synthetic data. (a) The estimated Moho depth. (b) The Moho depth residuals (difference between the true and estimated Moho depths). (c) The gravity residuals (difference between the observed and predicted gravity data). (d) Histogram of the gravity residuals shown in panel (c), with the calculated mean and standard deviation (std) of the residuals in mGal. (e) Cross-validation curve used to determine the optimal regularization parameter (eq. 10). Both axes are in logarithmic scale. The minimum mean square error (eq. 17) is found at $\mu = 0.00046$ (red triangle). (f) Goal function value (eq. 10) per Gauss–Newton iteration showing the convergence of the gradient descent. The y-axis is in logarithmic scale.

Table 1. Time spent on each function during a single inversion of simple synthetic data. The inversion was performed on a laptop computer with an Intel(R) Core(TM) i7-3612QM CPU @ 2.10 GHz processor. The total time for the inversion was 42.133 s.

Function description	Time (s)	Percentage of total time (per cent)
Sparse conjugate gradient	0.021	0.050
Sparse dot product	0.007	0.017
Tesseroid forward modelling	42.059	99.824

at $\mu = 0.00046$ (indicated by the red triangle). Fig. 5(f) shows the convergence of the Gauss-Newton optimization in eight iterations.

We also investigated the computation time spent in each section of the inversion process using a source code profiler. The profiler measures how much time is spent inside each function during the execution of a program. We ran the profiler on a single inversion of the training data set using the estimated regularization parameter. We tracked the total time spent inside each of the three functions that represent the potential bottlenecks of the inversion: solving the linear system in eq. (13) using the conjugate gradient method, performing the dot products required to compute the Hessian matrix (eq. 12) and the gradient vector (eq. 11), and forward modelling to calculate the predicted data (eq. 3). The profiling results presented in Table 1 show that the time spent on forward modelling accounts for approximately 99.8 per cent of the total computation time.

3.2 Model based on CRUST1.0

In this test, we simulate the anomalous Moho of South America using Moho depth information extracted from the CRUST1.0 model (Laske *et al.* 2013). We construct a tesseroid model with $M_{\text{lat}} \times M_{\text{lon}} = 80 \times 60$ juxtaposed elements, 4800 in total, using the Moho depths shown in Fig. 6(a). In our model, the Normal Earth Moho is $z_{\text{ref}} = 30$ km and the density-contrast is $\Delta\rho = 350$ kg m^{-3}. We produce the synthetic data at a constant height of 50 km and on a regular grid of $N_{\text{lat}} \times N_{\text{lon}} = 159 \times 119$ points (a total of 18 921 observations). We contaminate the synthetic data with normally distributed pseudo-random noise with zero mean and 5 mGal standard deviation (Fig. 6b).

The validation procedure to determine $\Delta\rho$ and z_{ref} requires knowledge of the Moho depth at certain points ($\mathbf{z}_s^o$ in eq. 18), usually from seismic experiments. Thus, we must also generate synthetic seismic data about the Moho depth. We produce such data by interpolating the Moho depth shown in Fig. 6(a) on the same 937 geographic coordinates pinpointed in the data set of Assumpção *et al.* (2013). The resulting synthetic seismic data is shown in Fig. 6(c).

We estimate the three hyperparameters in two parts. First, we run the cross-validation to estimate an optimal regularization parameter (μ). The starting estimate for all inversions is 60 km depth for all model parameters. For this cross-validation, we keep z_{ref} and $\Delta\rho$ fixed to 20 km and 500 kg m^{-3}, respectively. Second, we use the estimated μ to run the validation procedure with the synthetic seismological data to estimate z_{ref} and $\Delta\rho$, thus obtaining the final estimated Moho depths. Fig. 7 summarizes the results.

For the cross-validation, we separate the synthetic data (Fig. 3) into a training set with twice the grid spacing of the original data ($N_{\text{lat}} \times N_{\text{lon}} = 80 \times 60$) and a testing set with 14 121 observations. We run the inversion for 16 different values of μ equally spaced in a logarithmic scale between 10^{-7} and 10^{-2}. For each of the 16 estimates we compute the MSE (eq. 17), shown in Fig. 7(a) as function of μ. The optimal regularization parameter that minimizes the MSE is $\mu = 10^{-4}$ (red triangle in Fig. 7a).

In the validation using seismological data, we use the estimated value of μ in all inversions. We test seven values of z_{ref} from 20 to 35 km with 2.5 km intervals and seven values of $\Delta\rho$ from 200 to 500 kg m^{-3} with 50 kg m^{-3} intervals. We run the

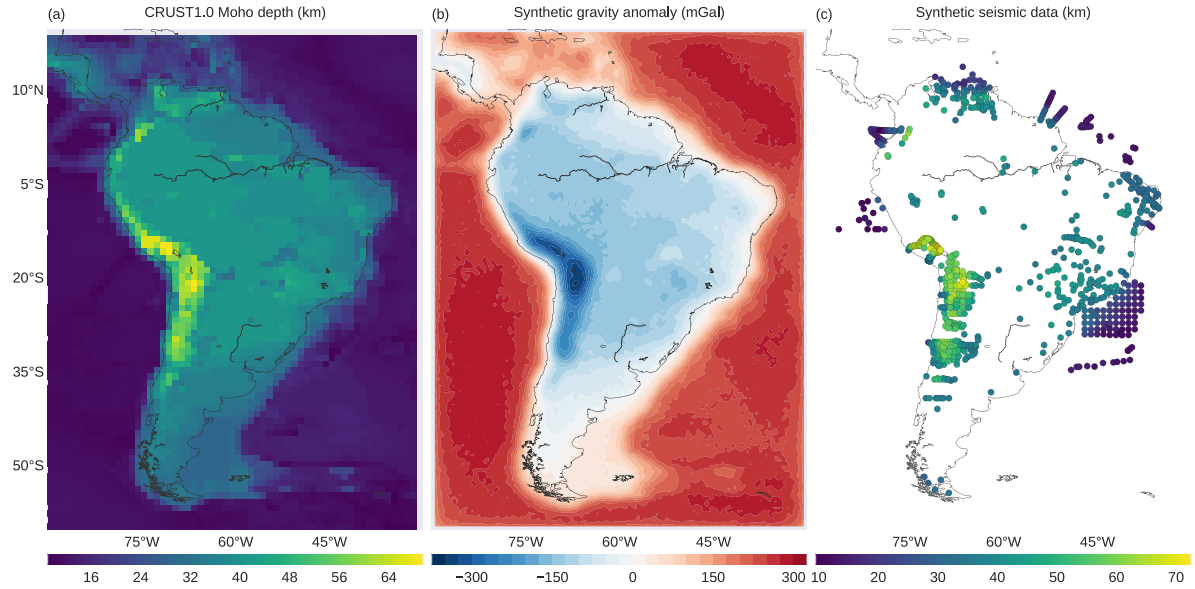

Figure 6. Synthetic data of a model derived from CRUST1.0. The model is made of tesseroids with an constant density-contrast of $\Delta\rho = 350\,\mathrm{kg\,m^{-3}}$ and assuming a reference level of $z_{\mathrm{ref}} = 30\,\mathrm{km}$. (a) The Moho depth of the model in kilometres. Each pixel in the pseudo-colour image corresponds to a tesseroid of the model. (b) Noise-corrupted synthetic gravity data generated from the model. (c) Simulated points where the Moho depths are known from seismological estimates (colour dots). Here, these point were obtained by interpolating the Moho depth in (a).

inversion for every combination of z_{ref} and $\Delta\rho$, totalling 49 inversions. Finally, we calculate the MSE (eq. 18) for each of the 49 estimates and choose the values of z_{ref} and $\Delta\rho$ that minimize the MSE. Fig. 7(b) shows a pseudo-colour map of the MSE with a minimum (marked by the red triangle) at $z_{\mathrm{ref}} = 30\,\mathrm{km}$ and $\Delta\rho = 350\,\mathrm{kg\,m^{-3}}$.

Fig. 7(c) shows the final solution after cross-validation and validation using seismological data. The recovered model is smooth, indicating that the cross-validation procedure was effective in estimating an optimal regularization parameter. Fig. 7(d) shows the difference between the true Moho depths (Fig. 6a) and the estimated depths (Fig. 7c). The maximum and minimum differences are, respectively, 9.8 and −8.2 km. The largest absolute differences are located along the central and northern Andes, where there is a sharp increase in the true Moho depth (Fig. 6a). Positive differences (indicating a too shallow estimate) appear along the central portion of the Andes, flanked by regions of negative differences (indicating a too deep estimate) on the continental and Pacific sides. Figs 7(e) and (g) show the gravity residuals, defined as the difference between the observed and predicted gravity data. The residuals appear normally distributed, with 0.03 mGal mean and a standard deviation of 4.10 mGal. The gravity residuals follow a similar, though reversed, pattern to the differences shown in Fig. 7(d). The largest residuals (in absolute value) are along the Andes, with the central portion being dominated by negative residuals and flanked by positive residuals on both sides. Figs 7(f) and (h) show the differences between the synthetic seismic data (Fig. 6c) and the estimated Moho depths. Once more, the largest differences are concentrated along the Andes, particularly in the central Andes and near Ecuador and Colombia. The differences are smaller along the Atlantic coast of South America, with notable larger differences in a few points of northeastern Brazil and along the Amazon river. In general, large residuals are associated with sharp increases in Moho depth.

4 APPLICATION TO THE SOUTH AMERICAN MOHO

We apply the inversion method proposed here to invert for the Moho depth of the South American continent. We follow the application of van der Meijde *et al.* (2013) but with some differences, mainly using a different data set and performing all modelling in spherical coordinates using tesseroids. The data are corrected of the effects of topography and sedimentary basins. Crust and mantle heterogeneities cannot be properly accounted for in regions where information coverage is sparse and readily accessible models are not available, like in South America and Africa. Hence, for the purposes of this study, we will assume to be negligible all other crustal and mantle sources, including lateral variations in density along the Moho. All tesseroid models are defined with respect to a spherical Earth of radius 6378.137 km.

4.1 Gravity and seismological data

The raw gravity data are generated from the satellite only spherical harmonic model GOCO5S (Mayer-Guerr *et al.* 2015). The GOCO5S model combines data from 15 satellites, including the complete mission data from the GOCE satellite. The data were downloaded from the International Centre for Global Earth Models (ICGEM) web-service (Barthelmes & Köhler 2012, http://icgem.gfz-potsdam.de/ICGEM/) in the form of the complete gravity field on a regular grid with 0.2° grid spacing at ellipsoidal height 50 km. We calculate the gravity disturbance ($\delta(P)$ in eq. 1) by subtracting from the raw data the normal gravity of the WGS84 reference ellipsoid ($\gamma(P)$) using the formula of Li & Götze (2001). Fig. 8(a) show the calculated gravity disturbance of South America.

We remove the gravitational effect of the topography from the gravity disturbance by modelling the ETOPO1 digital terrain model (Amante & Eakins 2009, http://dx.doi.org/10.7289/V5C8276M)

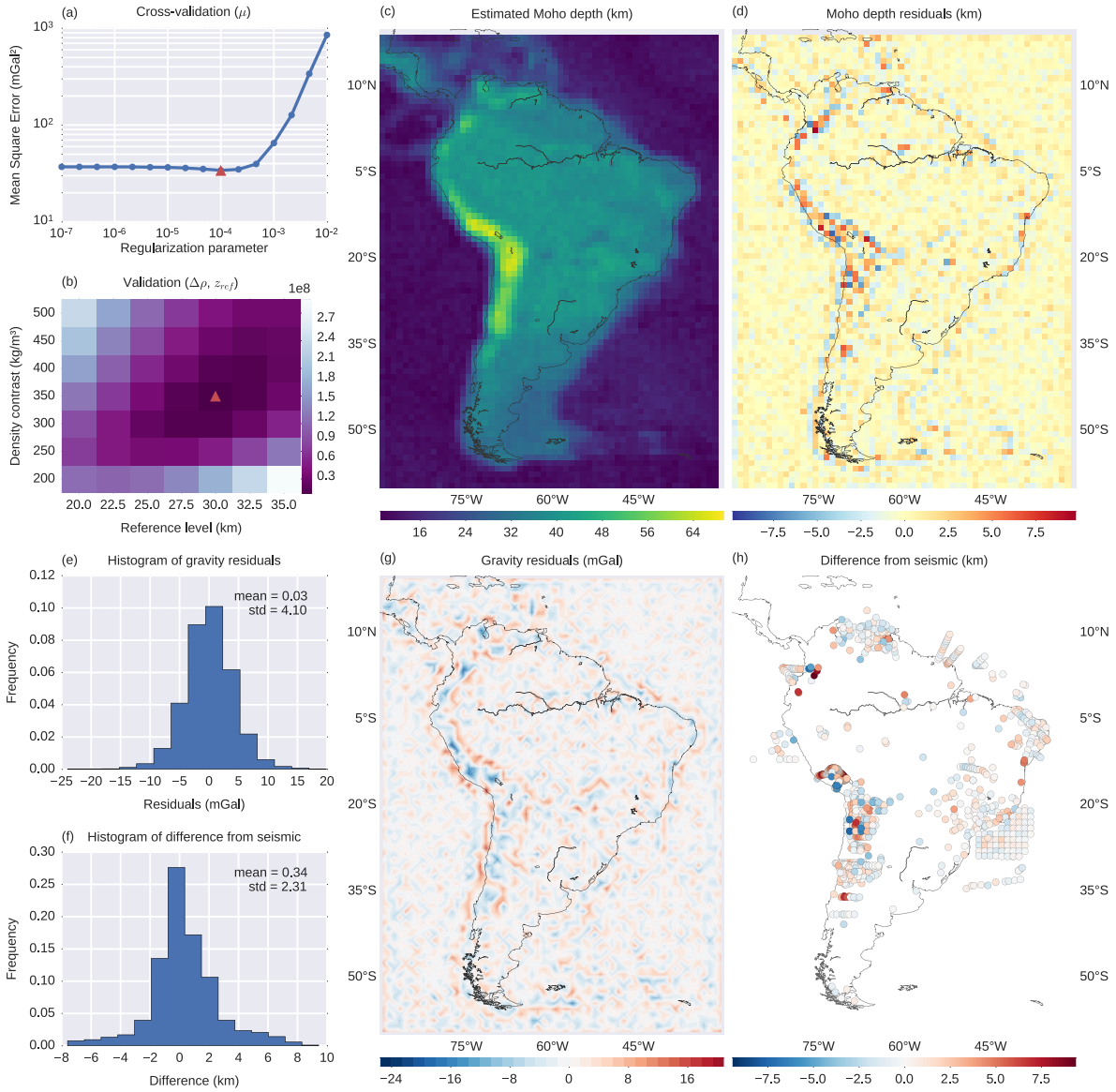

Figure 7. Inversion results from the CRUST1.0 synthetic data. (a) Cross-validation curve used to determine the regularization parameter (eq. 10). The minimum MSE (eq. 17) is found at $\mu = 0.0001$ (red triangle). (b) Validation results used to determine the reference level (z_{ref}) and the density-contrast ($\Delta\rho$). The colours represent the MSE (eq. 18) in km^2. The minimum (red triangle) is found at $z_{\text{ref}} = 30$ km and $\Delta\rho = 350$ kg m^{-3}. (c) The estimated Moho depth. (d) Difference between the CRUST1.0 model depths (Fig. 6a) and the estimated depths. (e) Histogram of the inversion residuals (observed minus predicted data). (f) Histogram of the differences between the synthetic seismic observations (Fig. 6c) and the estimated depths. (g) The inversion residuals. (h) Difference between the seismic and the estimated depths.

using tesseroids (Fig. 8b). We used the standard densities of 2670 kg m^{-3} for continents and -1630 kg m^{-3} for the oceans. Fig. 8(c) shows the calculated gravitational attraction of the topographic masses at 50 km height. Fig. 8(d) shows the Bouguer disturbance (eq. 2) obtained after subtracting the topographic effect from the gravity disturbance.

The effect of sedimentary basins is removed using tesseroid models of the three sedimentary layers present in the CRUST1.0 model (Laske *et al.* 2013, http://igppweb.ucsd.edu/~gabi/rem.html). Each sedimentary layer model includes the density of each $1^\circ \times 1^\circ$ model cell. Figs 8(e)–(g) show the thickness of the upper, middle, and lower sedimentary layers, respectively. The density-contrasts of the tesseroid model are obtained by subtracting 2670 kg m^{-3} from the density of each model element. Fig. 8(h) shows the combined gravitational attraction of the sedimentary basin tesseroid model. We subtract the total effect of sediments from the Bouguer disturbance in Fig. 8(d) to obtain the sediment-free Bouguer disturbance (Fig. 9a), which will be used as input for the inversion.

Fig. 9(b) shows the 937 known Moho depths (coloured dots) which were estimated from seismological data by Assumpção *et al.* (2013). This data set is used in the validation procedure.

4.2 Inversion, cross-validation, and validation using seismological data

As in the CRUST1.0 synthetic data test (Section 3.2), we estimate the hyperparameters in two steps. First, we run the cross-validation to estimate an optimal regularization parameter (μ). The starting

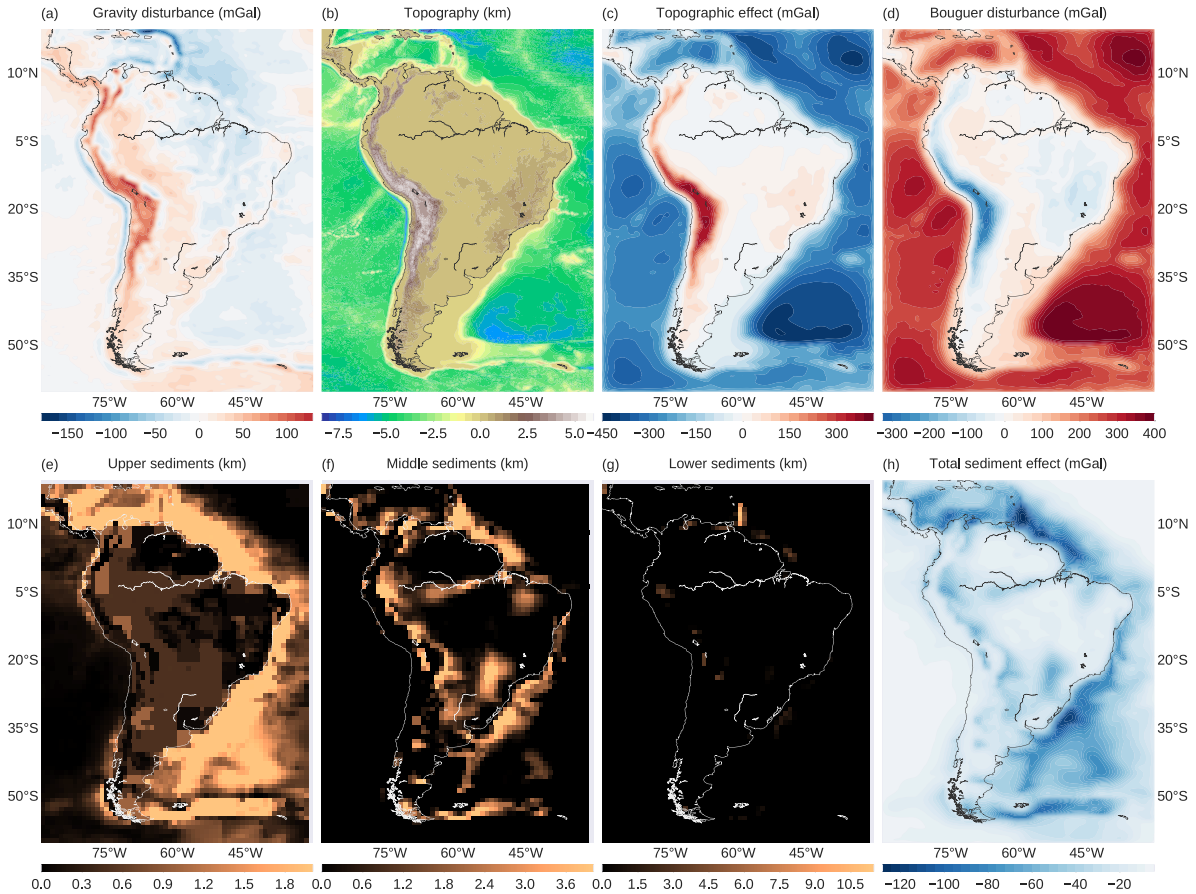

Figure 8. Gravity data for South America and the models used in the data corrections. (a) The gravity disturbance (eq. 1) calculated from the raw gravity data. (b) Topography from ETOPO1. (c) Gravitational attraction of the topography calculated at the observation height using tesseroids. (d) The Bouguer disturbance (eq. 2) obtained by subtracting (c) from (a). The upper (e), middle (f) and lower (g) sediment layer thicknesses from the CRUST1.0 model. (h) The total gravitational attraction of the sediment layers shown in (e), (f) and (g), calculated using tesseroids.

estimate for all inversions is 60 km depth for all model parameters. For this cross-validation, we keep z_{ref} and $\Delta\rho$ fixed to 20 km and 500 kg m^{-3}, respectively. Second, we use the estimated μ to run the validation using the seismological data of Assumpção *et al.* (2013) to estimate z_{ref} and $\Delta\rho$, thus obtaining the final estimated Moho depth model.

We split the sediment-free gravity disturbance (Fig. 9a) into the training and testing data sets. The training data set is a regular grid with 0.4° grid spacing (twice the spacing of the original data grid) and $N_{\text{lat}} \times N_{\text{lon}} = 201 \times 151$ grid points, a total of 30 351 observations. The remaining 90 350 points compose the testing data set. We test 16 values of the regularization parameter (μ) equally spaced on a logarithmic scale between 10^{-10} and 10^{-2}. Fig. 10(a) shows the Mean Square Error (MSE) as a function of μ. The minimum MSE is found at $\mu = 10^{-10}$, the lowest value of μ tested, suggesting that little or no regularization is required.

We proceed with the validation using seismological data using $\mu = 10^{-10}$ in all inversions. We test all combinations of nine values of z_{ref}, from 20 to 40 km with 2.5 km intervals, and seven values of $\Delta\rho$, from 200 to 500 kg m^{-3} with 50 kg m^{-3} intervals. Fig. 10(b) shows a pseudo-colour map of the MSE with respect to the Assumpção *et al.* (2013) data set. The MSE has a minimum, indicated by the red triangle, at $z_{\text{ref}} = 35$ km and $\Delta\rho = 400$ kg m^{-3}. The minimum is not as well-defined as for the CRUST1.0 synthetic (Fig. 7b), which is expected because in reality $\Delta\rho$ is not homogeneous across all of South America and the surrounding oceans.

4.3 Moho model for South America

The final Moho depth model for South America is shown as a pseudo-colour map in Fig. 11. The model is available in the online repository that accompanies this contribution (see Section 2.7). Each model element is a 0.4° × 0.4° tesseroid, represented by the pixels in the pseudo-colour map.

Our model differs significantly from CRUST1.0 (Fig. 6a) but contains most of the large-scale features present in the GMSA12 gravity-derived model of van der Meijde *et al.* (2013). The deepest Moho is along the central Andes, reaching depths upward of 70 km. The oceanic areas present the shallowest Moho, ranging approximately from 7.5 to 20 km. The Brazilian and Guyana Shields have a deeper Moho (greater than 35 km), with the deepest portions in the area around the São Francisco Craton and the northern border of the Parecis Basin. The Moho is shallower than 35 km along the Guyana Basin, the Andean foreland basins, the Chaco Basin, and along the centres of the Solimões, Amazonas and Paraná Basins.

Fig. 12(a) shows the gravity residuals, defined as the difference between the observed and predicted gravity data. Fig. 12(b) shows the differences between the seismic-derived Moho depths of

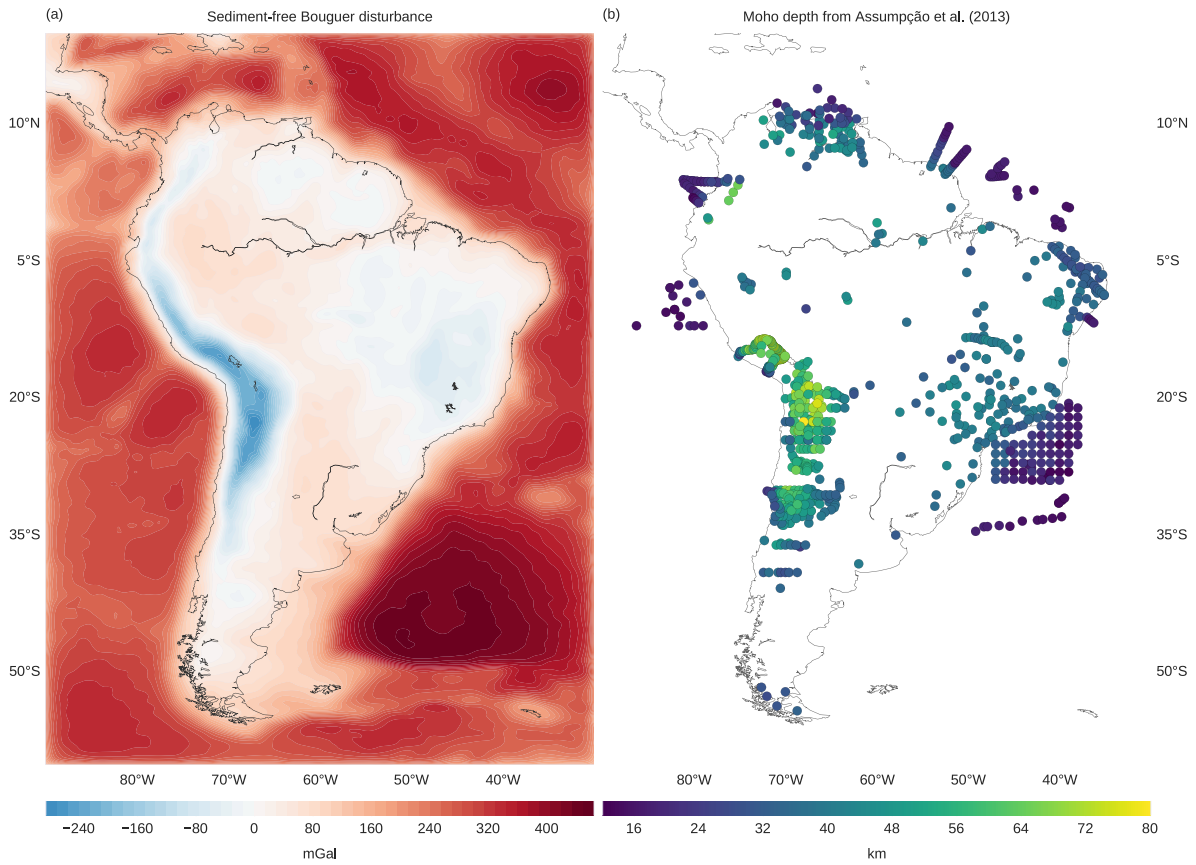

Figure 9. Input data for the South American Moho inversion. (a) Sediment-free Bouguer disturbance for South America. Obtained by subtracting the total sediment gravitational effect (Fig. 8h) from the Bouguer disturbance (Fig. 8d). (b) Seismological Moho depth estimates from Assumpção *et al.* (2013).

Assumpção *et al.* (2013; Fig. 9b) and the depths of our gravity-derived model (Fig. 11). The differences shown in Fig. 12(b) range from approximately −23 to 23 km and have a mean of 1.18 km and a standard deviation of 6.84 km. The gravity residuals and Moho depth differences from seismic are smallest in the oceanic areas, southern Patagonia, and the eastern coast of the continent. The largest gravity residuals are located along the Andes and correlate with the deepest Moho depths. These large residuals follow a pattern of a negative value in the centre flanked by positive values to the east and west. This same pattern is observed in the CRUST1.0 synthetic test results (Fig. 7). In general, larger gravity residuals appear to be associated with sharp variations in the estimated Moho depth. Along the Andes, large differences with seismic data are correlated with the larger gravity residuals. Conversely, this correlation is absent from the large differences seen in the Guyana, Paraná, and the Solimões Basins. In the Borborema province, northeastern Brazil, our model slightly overestimates the Moho depth. On the other hand, our model underestimates the Moho depths in the Amazonas, Solimões, and Paraná Basins. Particularly in the Amazonas and Solimões Basins, where our model predicts a Moho depth of approximately 30 km, the differences with the seismological estimates can reach 10 km or more.

4.4 Discussion

Differences between our Moho depth model and the seismological data (Fig. 12b) may indicate regions where our initial assumptions (summarized in Fig. 1) are inadequate or where we have failed to correct for all crustal and mantle sources. The largest differences are seen along the Andean Province and are likely caused by the fact that our model does not include the subducting Nazca plate. Furthermore, the CRUST1.0 synthetic data test (Fig. 7) suggests that our inversion method is not able to fully recover deep Moho depths in the Andes, even without the effect of the subducting plate. In the Guyana Basin, our model is able to fit the gravity data but differs from the seismological data by up to ±10 km with no clear pattern for the distribution of the differences. A possible explanation is an inaccuracy in the CRUST1.0 sediment model (Laske *et al.* 2013) used to correct our gravity data. The inversion results will be biased if the input data includes effects other than the anomalous Moho.

In the Amazon and Paraná Basins, our model fits the gravity data but underestimates the seismological data by up to 15 km. This indicates that a mass excess may be present in the crust or in the upper mantle. A body with positive density contrast whose gravitational effect was not removed from the data during processing will make the observed gravity disturbance greater than it would be otherwise. This will cause the inversion to produce a shallower Moho estimate. These discrepancies between gravity and seismological estimates have been noted before by Nunn & Aires (1988) for the Amazon Basin and Mariani *et al.* (2013) for the Paraná Basin. Both studies propose high density rocks in the lower crust as probable causes for the discrepancies. Another possible cause for the observed discrepancy in our model could be our failure to fully remove from the data the effects of the igneous intrusions present in both basins. Using a sediment model for South America more

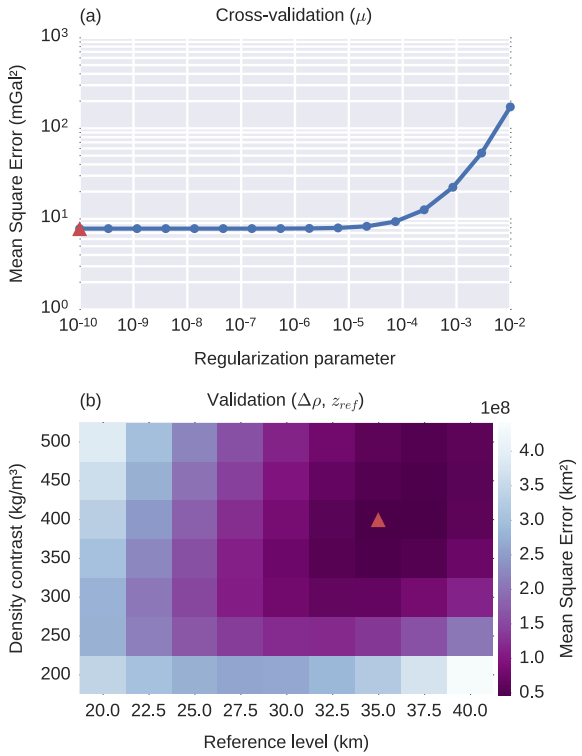

Figure 10. Cross-validation results for the South American Moho inversion. (a) Cross-validation to determine the regularization parameter μ (eq. 10). The minimum Mean Square Error (eq. 17), shown as a red triangle, corresponds to $\mu = 10^{-10}$. (b) Validation to determine the reference level (z_{ref}) and the density-contrast ($\Delta\rho$). The colours represent the Mean Square Error (eq. 18). The minimum (red triangle) is found at $z_{\text{ref}} = 35$ km and $\Delta\rho = 400\,\text{kg m}^{-3}$.

detailed than CRUST1.0 might lead to more accurate results for these basins.

Greater confidence in our Moho model can be had in areas where it is able to fit both the gravity and the seismological data. In such places, a gravity-derived model can serve as an interpolator for the seismological point estimates. In general, our model fits both data sets in the oceans, the Atlantic coast of the continent, and central Brazil. In the Parnaíba Basin, our model has small differences with the few seismological data points that fall inside and on the borders of the basin. The Moho surrounding the basins continental borders is deep, following inward with a shallower Moho, then a deeper step, and finally a shallower part in the middle of the basin. Recent deep seismic reflection studies by Daly *et al.* (2014) that cross the basin from west to east agree with our model. The Northern border of the Parecis Basin has a deep Moho in our model that is corroborated by a single seismological data point.

5 CONCLUSIONS

We have developed a computationally efficient gravity inversion method in spherical coordinates. Our method extends the Gauss-Newton formulation of Bott's method to use tesseroids as model elements and Tikhonov regularization. The computational efficiency of our method is due to using Bott's approximation for the Jacobian matrix and using sparse matrix algorithms for arithmetic operations and the solution of linear systems. This approximation for the Jacobian matrix is adequate and the method converges even when the data are at higher altitudes and the model is not outcropping, as shown by the applications to synthetic data. We employ hold-out cross-validation to estimate the regularization parameter and a validation procedure using seismological data to estimate the Moho density-contrast and the Normal Earth Moho depth.

There are two main advantages of the proposed method over previous works. First, unlike the Parker-Oldenburg method or methods using rectangular prisms, our inversion method does not require the data to be projected onto a plane. Second, the Parker-Oldenburg method and methods derived from Bott's method cannot apply the traditional methodology for constraining inverse problems. Our method has no such restriction because we use the formalism of traditional Tikhonov inversion. However, the proposed method is not without limitations. It requires data on a regular grid and restricts the model to be a regular mesh tied to the data grid. Like Bott's method, our method only works for gravity disturbances and not for the gravity gradients.

The test on simple synthetic data shows that our inversion method is able to recover a smooth Moho relief with a homogeneous density-contrast distribution. The inversion was not able to fully recover the shortest wavelength feature in the model, possibly due to the smoothness constraints which tend to soften high-frequency (sharp) variations. The cross-validation Mean Square Error curve has a well-defined minimum, indicating a value of the regularization parameter (μ) whose corresponding estimate best predicts data that were not included in the inversion. Using this value of μ in the inversion leads to a stable solution characterized by a smooth Moho relief with an acceptable data misfit.

The efficiency of the proposed method is because solving linear systems and performing matrix multiplications together account for a mere 0.067 per cent of the total computation time required for a single inversion. The majority of the computation time (99.824 per cent) is spent on forward modelling. Thus, we are able to retain the high computational efficiency of Bott's method and use a classic Tikhonov regularization formulation. This approach could, in theory, be extended to other types of regularization (e.g. total variation) and misfit functions (e.g. re-weighted least squares) already available in the literature.

The more complex synthetic data test based on CRUST1.0 shows that the validation using pointwise Moho depth information is able to correctly estimate the density-contrast ($\Delta\rho$) and Normal Earth Moho depth (z_{ref}). This test indicates that the inversion neither correctly estimates Moho depth nor adequately fits the gravity and pointwise data when sharp variations in Moho depth occur. This phenomenon is particularly strong in the region below the Andes. A likely explanation is that the smoothness regularization is intrinsically unable to produce sharp variations in Moho depth. These effects might be mitigated with the use of sharpness-inducing regularization, like the weighted smoothness inversion, Cauchy norm regularization, entropic regularization, total variation regularization, or an adaptive mixed smoothness-sharpness regularization.

We applied the method proposed here to estimate the Moho depth for South America. Our estimated Moho depth model is in accordance with previous results. The model fits well the gravity and seismic data in all oceanic regions, the central portion of the Andean foreland, Patagonia, and coastal and central parts of Brazil. However, the model is unable to fit the gravity and seismic data in places with sharp variations in Moho depth, particularly below the Andes and in the boundaries of the main geotectonic provinces of the South American Plate, like the Borborema province, the Parnaiba Basin, and the São Francisco Craton. This might indicate that

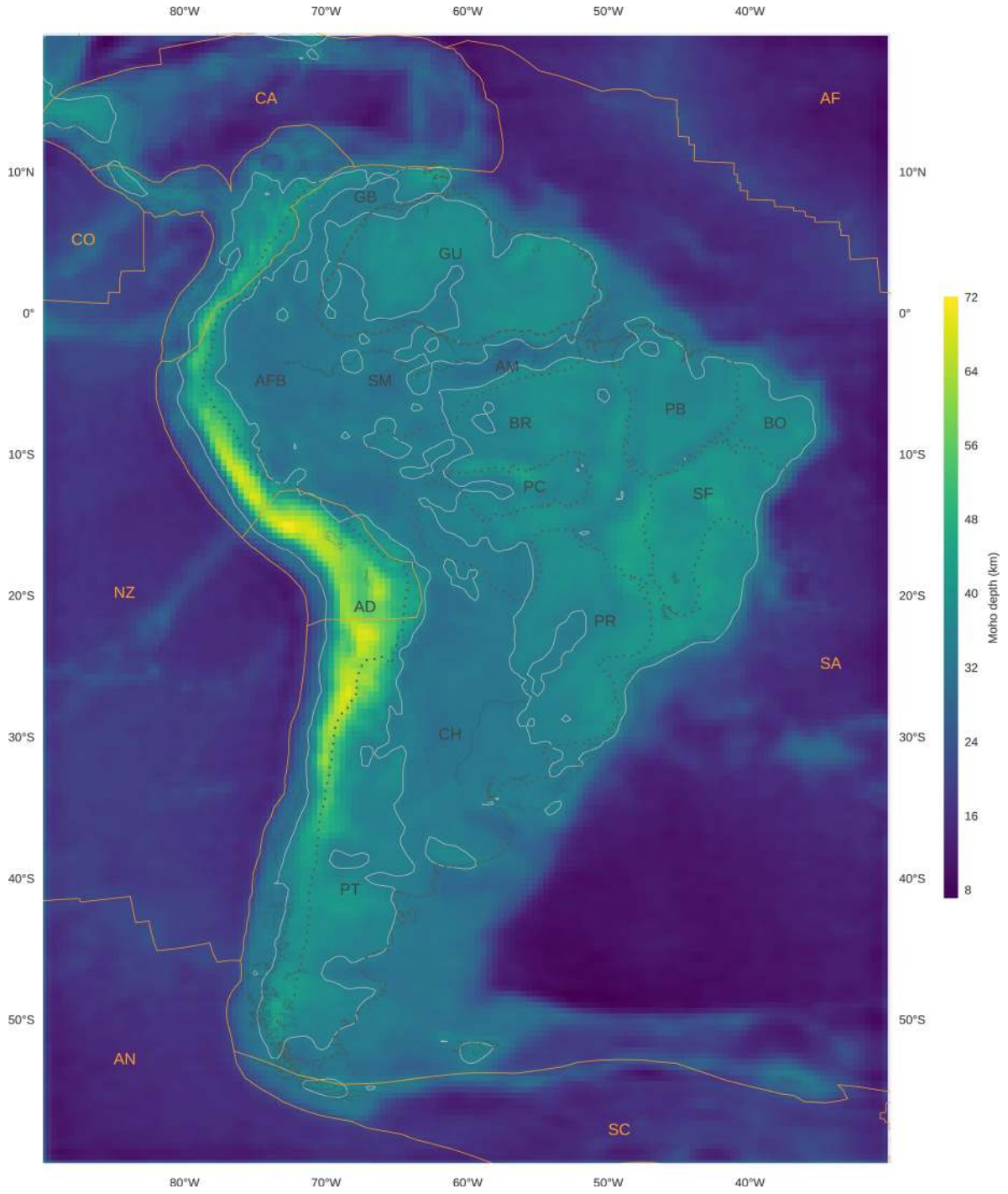

Figure 11. The estimated Moho depth of South America. Dotted lines represent the boundaries between major geologic provinces (after Assumpção *et al.* 2013; Goutorbe *et al.* 2015); AD: Andean Province, AFB: Andean foreland basins, AM: Amazonas Basin, BR: Brazilian Shield, BO: Borborema province, CH: Chaco Basin, GB: Guyana Basin, GU: Guyana Shield, PB: Parnaíba Basin, PC: Parecis Basin, PR: Paraná Basin, PT: Patagonia province, SF: São Francisco Craton, SM: Solimões Basin. Solid orange lines mark the limits of the main lithospheric plates (Bird 2003); AF: Africa Plate, AN: Antarctica Plate, CA: Caribbean Plate, CO: Cocos Plate, SA: South America Plate, SC: Scotia Plate, NZ: Nazca Plate. The solid light grey line is the 35 km Moho depth contour.

smoothness regularization should not be applied indiscriminately to the whole model, as suggested by the CRUST1.0 synthetic data test. Another reason for the observed misfit might be the presence of crustal or mantle density anomalies whose gravitational effects were not removed during the data corrections. In the Guyana Basin on the coastal region of Venezuela, along the central Amazonas and Solimões Basins, and in the Paraná Basin, our Moho depth model is able to fit the gravity data but differs significantly from the seismic data. These discrepancies in the Paraná and Amazonas Basins are interpreted in the literature as high density rocks in the lower crust. In general, differences between a gravity and a seismically derived Moho model may indicate the presence of crustal or mantle density anomalies that were unaccounted for in the data processing. Such locations warrant further detailed investigation.

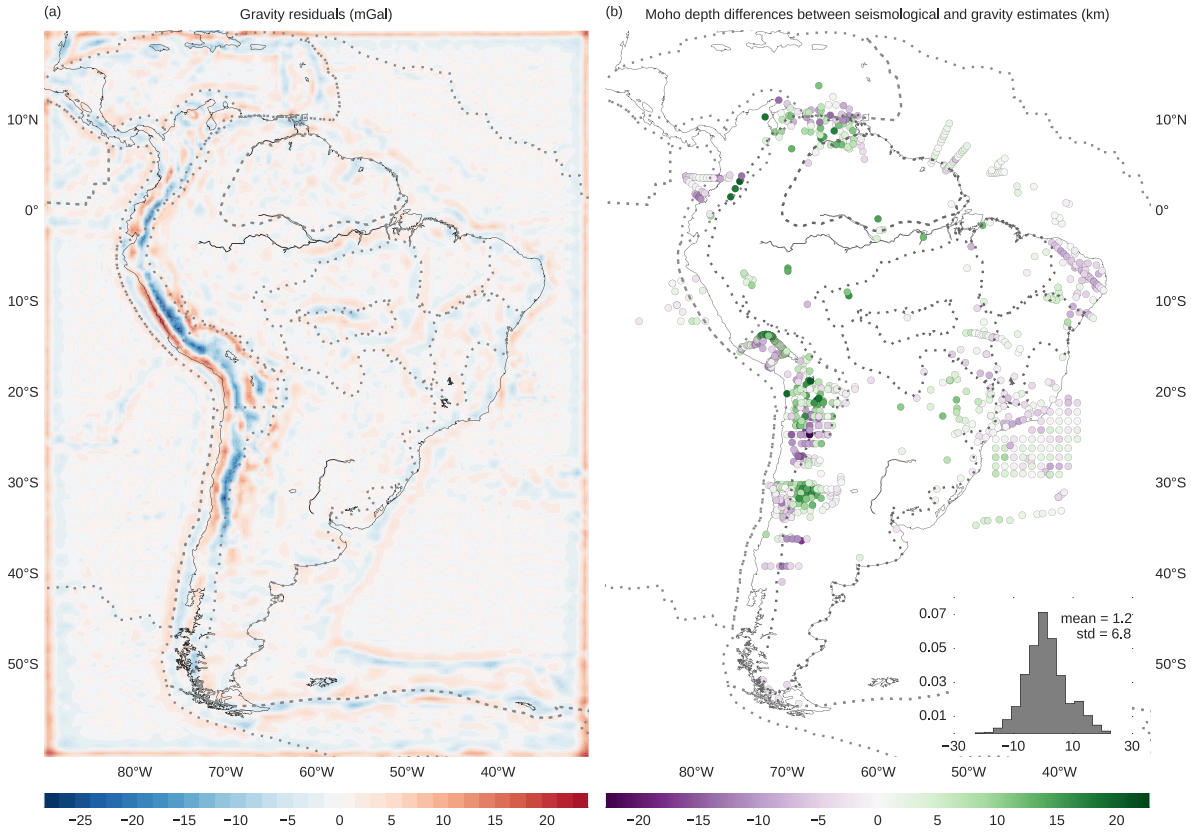

Figure 12. Residuals for the estimated South American Moho depth in Fig. 11. (a) Gravity residuals, defined as the difference between the observed data in Fig. 9(a) and the data predicted by the estimate in Fig. 11. (b) Differences between the seismological depth estimates of Assumpção *et al.* (2013) and our gravity-derived Moho depth estimate. The inset in b shows a histogram of the differences along with their calculated mean and standard deviation (std). Dotted lines mark the limits of major geologic provinces and lithospheric plates.

Our gravity-derived Moho model for South America can be downloaded from the online repository http://dx.doi.org/10.6084/m9.figshare.3987267.

ACKNOWLEDGEMENTS

We are indebted to the developers and maintainers of the open-source software without which this work would not have been possible. We thank Marcelo Assumpção for providing the seismological Moho depth estimates, Naomi Ussami, Julio C.S.O. Lyrio, Cosme F.P. Neto, and Daniel R. Franco for their constructive feedback, and Editor Dr Gary Egbert and two anonymous reviewers for their thoughtful comments that helped improve this work. L. Uieda was supported by a scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). V.C.F. Barbosa was supported by a fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

REFERENCES

Amante, C. & Eakins, B.W., 2009. *ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, NOAA Technical Memorandum NESDIS NGDC-24*. National Geophysical Data Center, NOAA.

Asgharzadeh, M.F., von Frese, R.R.B., Kim, H.R., Leftwich, T.E. & Kim, J.W., 2007. Spherical prism gravity effects by Gauss-Legendre quadrature integration, *Geophys. J. Int.,* **169**(1), 1–11.

Assumpção, M., Feng, M., Tassara, A. & Julià, J., 2013. Models of crustal thickness for South America from seismic refraction, receiver functions and surface wave tomography, *Tectonophysics,* **609,** 82–96.

Barbosa, V., Silva, J. & Medeiros, W., 1997. Gravity inversion of basement relief using approximate equality constraints on depths, *Geophysics,* **62**(6), 1745–1757.

Barbosa, V., Silva, J. & Medeiros, W., 1999a. Stable inversion of gravity anomalies of sedimentary basins with nonsmooth basement reliefs and arbitrary density contrast variations, *Geophysics,* **64**(3), 754–764.

Barbosa, V.C.F., Silva, J.B.C. & Medeiros, W.E., 1999b. Gravity inversion of a discontinuous relief stabilized by weighted smoothness constraints on depth, *Geophysics,* **64**(5), 1429–1437.

Barnes, G. & Barraud, J., 2012. Imaging geologic surfaces by inverting gravity gradient data with depth horizons, *Geophysics,* **77**(1), G1–G11.

Barthelmes, F. & Köhler, W., 2012. International Centre for Global Earth Models (ICGEM), *J. Geod.,* **86**(10), 932–934.

Bird, P., 2003. An updated digital model of plate boundaries, *Geochem. Geophys. Geosyst.,* **4**(3), 1027, doi:10.1029/2001GC000252.

Bott, M.H.P., 1960. The use of rapid digital computing methods for direct gravity interpretation of sedimentary basins, *Geophys. J. Int.,* **3**(1), 63–67.

Daly, M.C., Andrade, V., Barousse, C.A., Costa, R., McDowell, K., Piggott, N. & Poole, A.J., 2014. Brasiliano crustal structure and the tectonic setting of the Parnaíba basin of NE Brazil: results of a deep seismic reflection profile, *Tectonics,* **33**(11), 2102–2120.

Farquharson, C.G. & Oldenburg, D.W., 2004. A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems, *Geophys. J. Int.,* **156**(3), 411–425.

Goutorbe, B., Coelho, D.L.d.O. & Drouet, S., 2015. Rayleigh wave group velocities at periods of 6–23 s across Brazil from ambient noise tomography, *Geophys. J. Int.,* **203**(2), 869–882.

Grombein, T., Seitz, K. & Heck, B., 2013. Optimized formulas for the gravitational field of a tesseroid, *J. Geod.,* **87**(7), 645–660.

Hansen, P., 1992. Analysis of Discrete Ill-Posed Problems by Means of the L-Curve, *SIAM Rev.,* **34**(4), 561–580.

Heck, B. & Seitz, K., 2007. A comparison of the tesseroid, prism and point-mass approaches for mass reductions in gravity field modelling, *J. Geod.,* **81**(2), 121–136.

Hunter, J.D., 2007. Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.,* **9**(3), 90–95.

Jones, E., Oliphant, T., Peterson, P. & others, 2001. SciPy: Open source scientific tools for Python, accessed August 2015.

Kim, J.H., 2009. Estimating classification error rate: repeated cross-validation, repeated hold-out and bootstrap, *Comput. Stat. Data Anal.,* **53**(11), 3735–3745.

Laske, G., Masters, G., Ma, Z. & Pasyanos, M., 2013. Update on CRUST1.0—1-degree Global Model of Earth's Crust, in *EGU General Assembly Conference Abstracts,* vol. 15, pp. EGU2013–2658.

Leão, J., Menezes, P., Beltrão, J. & Silva, J., 1996. Gravity inversion of basement relief constrained by the knowledge of depth at isolated points, *Geophysics,* **61**(6), 1702–1714.

Li, X. & Götze, H., 2001. Ellipsoid, geoid, gravity, geodesy, and geophysics, *Geophysics,* **66**(6), 1660–1668.

Li, Z., Hao, T., Xu, Y. & Xu, Y., 2011. An efficient and adaptive approach for modeling gravity effects in spherical coordinates, *J. appl. Geophys.,* **73**(3), 221–231.

Mariani, P., Braitenberg, C. & Ussami, N., 2013. Explaining the thick crust in Paraná basin, Brazil, with satellite GOCE gravity observations, *J. South Am. Earth Sci.,* **45,** 209–223.

Martins, C., Barbosa, V. & Silva, J., 2010. Simultaneous 3D depth-to-basement and density-contrast estimates using gravity data and depth control at few points, *Geophysics,* **75**(3), I21–I28.

Martins, C., Lima, W., Barbosa, V. & Silva, J., 2011. Total variation regularization for depth-to-basement estimate: Part 1 — Mathematical details and applications, *Geophysics,* **76**(1), I1–I12.

Mayer-Guerr, T. *et al.*, 2015. The combined satellite gravity field model GOCO05s, in *EGU General Assembly Conference Abstracts,* vol. 17, pp. EGU2015–12364.

Nunn, J.A. & Aires, J.R., 1988. Gravity anomalies and flexure of the lithosphere at the Middle Amazon Basin, Brazil, *J. geophys. Res.,* **93**(B1), 415–428.

Oldenburg, D., 1974. The inversion and interpretation of gravity anomalies, *Geophysics,* **39**(4), 526–536.

Parker, R.L., 1973. The Rapid Calculation of Potential Anomalies, *Geophys. J. Int.,* **31**(4), 447–455.

Pérez, F. & Granger, B.E., 2007. IPython: A System for Interactive Scientific Computing, *Comput. Sci. Eng.,* **9**(3), 21–29.

Reguzzoni, M., Sampietro, D. & Sansò, F., 2013. Global Moho from the combination of the CRUST2.0 model and GOCE data, *Geophys. J. Int.,* **195**(1), 222–237.

Santos, D., Silva, J., Martins, C., dos Santos, R., Ramos, L. & de Araújo, A., 2015. Efficient gravity inversion of discontinuous basement relief, *Geophysics,* pp. G95–G106.

Silva, J., Medeiros, W. & Barbosa, V., 2001. Potential-field inversion: choosing the appropriate technique to solve a geologic problem, *Geophysics,* **66**(2), 511–520.

Silva, J., Costa, D. & Barbosa, V., 2006. Gravity inversion of basement relief and estimation of density contrast variation with depth, *Geophysics,* **71**(5), J51–J58.

Silva, J., Santos, D. & Gomes, K., 2014. Fast gravity inversion of basement relief, *Geophysics,* **79**(5), G79–G91.

Sun, J. & Li, Y., 2014. Adaptive Lp inversion for simultaneous recovery of both blocky and smooth features in a geophysical model, *Geophys. J. Int.,* **197,** 882–899.

Tikhonov, A.N. & Arsenin, V.Y., 1977. *Solutions of Ill-posed Problems,* Scripta Series in Mathematics, Winston.

Uieda, L., 2015. A tesserioid (spherical prism) in a geocentric coordinate system with a local-North-oriented coordinate system, *figshare*, Available at: https://dx.doi.org/10.6084/m9.figshare.1495525.v1, last accessed November 2015.

Uieda, L. & Barbosa, C.F.V., 2016. A gravity-derived moho model for south america: source code, data, and model results from 'Fast non-linear gravity inversion in spherical coordinates with application to the South American Moho', *figshare*, Available at: http://dx.doi.org/10.6084/m9.figshare.3987267, accessed October 2016.

Uieda, L., Oliveira, V.C. Jr. & Barbosa, V.C.F., 2013. Modeling the Earth with Fatiando a Terra, in *Proceedings of the 12th Python in Science Conference,* Austin, Texas, USA, pp. 91–98.

Uieda, L., Barbosa, V. & Braitenberg, C., 2016. Tesseroids: forward-modeling gravitational fields in spherical coordinates, *Geophysics,* **81,** F41–F48.

van der Meijde, M., Julià, J. & Assumpção, M., 2013. Gravity derived Moho for South America, *Tectonophysics,* **609,** 456–467.

van der Meijde, M., Fadel, I., Ditmar, P. & Hamayun, M., 2015. Uncertainties in crustal thickness models for data sparse environments: a review for South America and Africa, *J. Geodyn.,* **84,** 1–18.

Waskom, M. *et al.*, 2015. *seaborn: v0.6.0*, accessed November 2015.

Wieczorek, M.A. & Phillips, R.J., 1998. Potential anomalies on a sphere: applications to the thickness of the lunar crust, *J. geophys. Res.,* **103**(E1), 1715–1724.

Wild-Pfeiffer, F., 2008. A comparison of different mass elements for use in gravity gradiometry, *J. Geod.,* **82**(10), 637–653.