submitted to *Geophys. J. Int.*

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

Leonardo Uieda[1,2], 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.*

SUMMARY

Estimating the relief of the Moho from gravity data is a computationally intensive non-linear inverse problem. What is more, the modeling must take the Earths curvature into account when the study area is of regional scale or greater. We present a regularized non-linear 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 hyper-parameters: 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 Guiana shield, the central Solimões and Amazon 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; Gravity anomalies and Earth structure; Satellite gravity; 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 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 seismologic and terrestrial gravity data are traditionally concentrated around large urban centers and coastal areas.

Estimating Moho depth from gravity data is a non-linear inverse problem. One can generalize this problem to estimating the relief of an interface, such as the basement of a sedimentary basin or the relief of the anomalous Moho. Several methods have been developed over the years to solve this inverse problem, for example Bott (1960); Barbosa et al. (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 modeling. Oldenburg (1974) showed that the fast FFT-based forward modeling 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.

Most non-linear gravity inversion methods discretize the relief of the interface into juxtaposed right-rectangular prisms with a known density contrast. The inverse problem is then to estimate the thickness of each prism from the gravity data. The use of rectangular prisms implies a planar Earth approximation and may not be adequate for continental and global scale studies. In such cases, a spherical Earth approximation is preferred. 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 approximation to estimate the global Moho relief from GOCE satellite gravity data. Conversely, one could adapt one of the methods developed for right-rectangular prisms to use tesseroids (spherical prisms) instead. 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 the relief of an interface, two hyper-parameters control the inversion results: the density-contrast between the two mediums and the reference level around which the interface undulates. The reference level is the depth of the Normal Earth Moho in the case of the anomalous Moho. For regularized inversions, an additional hyper-parameter is the regularization parameter that balances data-misfit and regularization. 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 non-linear 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 modeling. 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 cross-validation. Similarly to Silva et al. (2006) and Martins et al. (2010), this cross-validation procedure uses knowledge of the Moho depth at certain points. 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).

2 METHODOLOGY

In potential field methods, we must isolate the target anomalous density distribution prior to modeling 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 (Fig 1a-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,

$$\delta(P) = g(P) - \gamma(P). \tag{1}$$

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

In order to invert for the anomalous Moho relief, we must first isolate its gravitational attraction. Thus, all other effects must be either removed or assumed negligible. Here, we will remove the effect of the topography and oceans in order to obtain the full Bouguer disturbance (Fig 1d),

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

We will remove the effect of 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 modeling using tesseroids (Fig. 2). The tesseroid effects are calculated

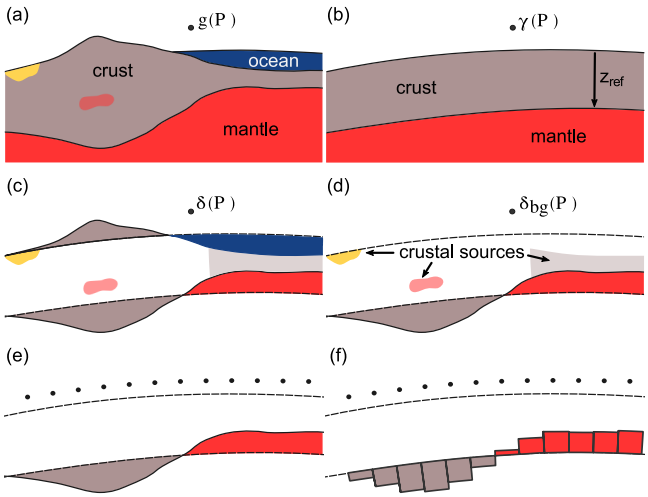

Figure 1. Sketch of the stages in gravity data correction and the discretization of the anomalous Moho relief using tesseroids. (a) The Earth and the measured gravity at point P ($g(P)$). (b) The Normal Earth and the calculated normal gravity at point P ($\gamma(P)$). z_{ref} is the depth of the Normal Earth Moho. (c) The gravity disturbance ($\delta(P)$) and the corresponding density anomalies after removal of the normal gravity: topography, oceans, crustal heterogeneities, and the anomalous Moho. (d) The Bouguer disturbance ($\delta_{bg}(P)$) 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.

numerically using Gauss-Legendre Quadrature (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

We parameterize the forward problem by discretizing the anomalous Moho into a grid of $M_{lon} \times M_{lat} = M$ 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). In cases where 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 non-linear function of the parameters z_k, $k = 1, \ldots, M$,

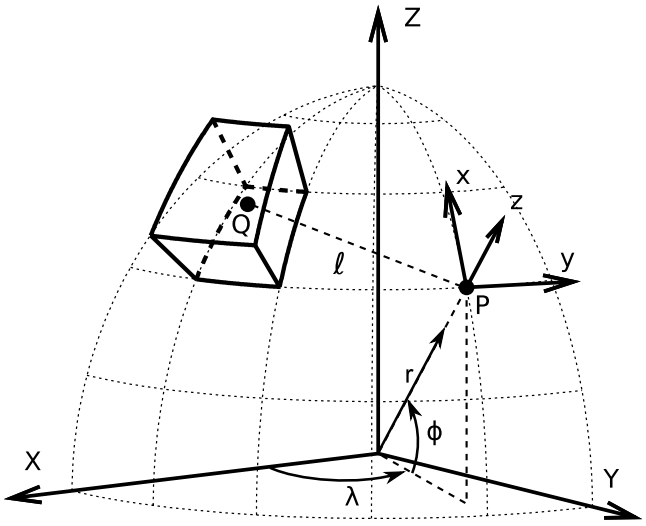

Figure 2. Sketch of a tesseroid (spherical prism) in a geocentric coordinate system (X, Y, Z). Observations are made at point P with respect to it's local North-oriented coordinate system (x, y, z). After Uieda (2015).

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

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 non-linear 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 anomaly data $\mathbf{d}^o$. The least-squares estimate is the one that minimizes the data-misfit function

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

Function $\phi(\mathbf{p})$ is non-linear 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 estimate $\mathbf{p}^0$ and iteratively update the estimate until a minimum is reached.

For the Gauss-Newton method, the update at the kth iteration, $\mathbf{\Delta p} = \mathbf{p}^{k+1} - \mathbf{p}^k$, is the solution of the linear system

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

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

The Steepest Descent method uses only the gradient direction to update the initial estimate (Kelley 1987). The update at the kth iteration is achieved by equating the Hessian in Eq. 5 to the identity matrix,

$$\mathbf{\Delta p} = -\nabla\phi^k. \tag{6}$$

Thus, it does not require the computation and storage of the Hessian matrix nor the solution of linear systems. However, the Steepest Descent method has poor convergence when the current solution is close to the minimum of the goal function (Kelley 1987).

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

$$\nabla\phi^k = -2\mathbf{A}^T[\mathbf{d}^o - \mathbf{d}(\mathbf{p}^k)], \tag{7}$$

and

$$\mathbf{H}^k \approx 2\mathbf{A}^T\mathbf{A}, \tag{8}$$

in which $\mathbf{A}$ is the Jacobian or sensitivity matrix,

$$A_{ij} = \frac{\partial f_i}{\partial p_j}(\mathbf{p^k}). \tag{9}$$

2.3 Regularization

Non-linear inversions for the relief of an interface (like the Moho) are ill-posed and require additional constraints in the form of regularization (**?**). A common approach is to use the first-order Tikhonov regularization to impose smoothness on the solution. The cost function for smoothness regularization is given by

$$\theta(\mathbf{p}) = \mathbf{p}^T\mathbf{R}^T\mathbf{R}\mathbf{p}, \tag{10}$$

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

The solution $\hat{\mathbf{p}}$ to the regularized inverse problem is the one that minimizes the goal function

$$\Gamma(\mathbf{p}) = \phi(\mathbf{p}) + \mu\theta(\mathbf{p}), \tag{11}$$

in which μ is the regularization parameter that controls the balance between fitting the observed data and obeying the smoothness constraint.

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

$$\nabla\boldsymbol{\Gamma}^k = -2\mathbf{A}^T[\mathbf{d}^o - \mathbf{d}(\mathbf{p}^k)] + 2\mu\mathbf{R}^T\mathbf{R}\mathbf{p}^k, \tag{12}$$

and

$$\mathbf{H}^k = 2\mathbf{A}^T\mathbf{A} + 2\mu\mathbf{R}^T\mathbf{R}. \tag{13}$$

The parameter updates for the regularized Gauss-Newton and Steepest Descent methods, respectively, then become

$$\left[\mathbf{A}^T\mathbf{A} + \mu\mathbf{R}^T\mathbf{R}\right]\boldsymbol{\Delta}\mathbf{p} = \mathbf{A}^T[\mathbf{d}^o - \mathbf{d}(\mathbf{p}^k)] - \mu\mathbf{R}^T\mathbf{R}\mathbf{p}^k, \tag{14}$$

and

$$\boldsymbol{\Delta}\mathbf{p} = \mathbf{A}^T[\mathbf{d}^o - \mathbf{d}(\mathbf{p}^k)] - \mu\mathbf{R}^T\mathbf{R}\mathbf{p}^k, \tag{15}$$

Producing the regularized solution 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 $\mathbf{A}$ and (2) the solution of the resulting $M \times M$ equation system (not required for Steepest Descent). In practice, the derivatives in the Jacobian (Eq. 9) are often calculated through a first-order finite-difference approximation. Thus, evaluating $\mathbf{A}$ requires $2 \times M \times N$ forward modeling operations for each iteration of the gradient descent algorithm. These computations are performed for each iteration of the optimization.

2.4 Bott's method

Bott (1960) developed an efficient method to determined the basement relief 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 estimate of the basement relief $\mathbf{p}^0$ equal to the null vector and updates the estimate using the formula

$$\boldsymbol{\Delta}\mathbf{p} = \frac{\mathbf{d}^o - \mathbf{d}(\mathbf{p}^k)}{2\pi G\Delta\rho}, \tag{16}$$

in which G is the gravitational constant and $\Delta\rho$ is the basin density contrast. The iterative process stops when the inversion residuals $\mathbf{r} = \mathbf{d}^o - \mathbf{d}(\mathbf{p}^k)$ 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. 9) to

$$\mathbf{A} = 2\pi G \Delta\rho \mathbf{I}, \tag{17}$$

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

We propose that Bott's method can also be formulated as a special case of the Steepest Descent method (Eq. 6) by setting the Jacobian matrix to

$$\mathbf{A} = \frac{1}{4\pi G \Delta\rho}\mathbf{I}. \tag{18}$$

In practice, both formulations lead to Eq. 16. One of the advantages of Bott's method over the traditional Gauss-Newton or Steepest Descent is eliminating the computation and storage of the dense Jacobian matrix $\mathbf{A}$. 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 Regularized Bott's method in spherical coordinates

We propose a regularized version of Bott's method to invert for the relief of the anomalous Moho in spherical coordinates. To adapt Bott's method to spherical coordinates, we replace the right-rectangular prisms in the forward modeling ($\mathbf{d}(\mathbf{p}^k)$ in Eq. 16) with tesseroids. The tesseroid forward modeling 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. 14) but replaces the full Jacobian matrix with the Bouguer plate approximation (Eq. 17). This linearizes the Jacobian matrix and reduces it to a sparse diagonal matrix, thus eliminating the cost of computing and storing $\mathbf{A}$. Matrix arithmetic operations can be performed efficiently by taking advantage of the sparse nature of matrices $\mathbf{A}$ and $\mathbf{R}$ (respectively, Eq. 17 and 10). The same is true for solving the equation system in the Gauss-Newton method (Eq. 14). However, the computational cost of forward modeling is still present. Particularly, forward modeling using tesseroids is more computationally intensive than using

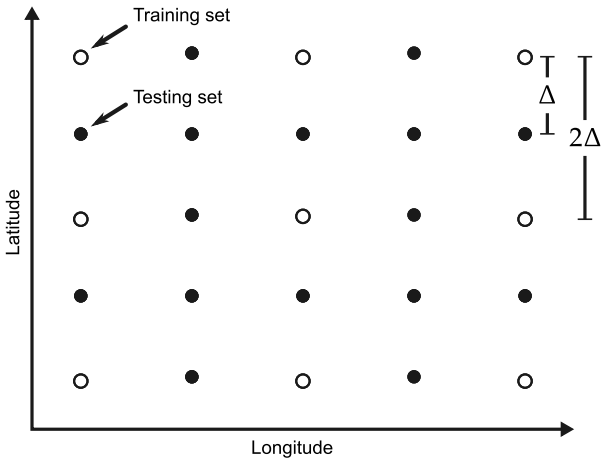

Figure 3. Sketch of a data grid separated into the training (white dots with black outlines) and testing (black dots) data sets. The training data set is still displayed on a regular grid but with twice the grid spacing of the original data grid.

right-rectangular prisms because of the numerical integration and adaptive discretization (Uieda et al. 2016). Our benchmarks suggest that sparse matrix multiplications and solving the sparse linear system in Eq. 14 account for less than 0.1% of the computation time of a single inversion (see section 3.1 and Table 1). Hence, by employing the use of sparse matrices, our formulation retains the efficiency of Bott's method while stabilizing the solution through the well established formalism of Tikhonov regularization.

2.6 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 $\mathbf{d}^o_{inv}$ and a testing set $\mathbf{d}^o_{test}$. 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_k, the training set is inverted using μ_k to obtain an estimate $\hat{\mathbf{p}}^k$. This estimate is used to calculate predicted data on the same points as the testing set via forward modeling ($\mathbf{d}^k_{test} = \mathbf{f}(\hat{\mathbf{p}}^k)$). The metric chosen to evaluate μ_k is the mean square error (MSE) of the misfit between the observed and predicted testing data sets,

$$MSE_k = \frac{\|\mathbf{d}^o_{test} - \mathbf{d}^k_{test}\|^2}{N_{test}}, \tag{19}$$

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, i.e. the one that best predicts the testing data. We emphasize that the inversion is performed only on the training data set.

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

(i) Divide the observed data into the training ($\mathbf{d}^o_{inv}$) and testing ($\mathbf{d}^o_{test}$) sets.

(ii) For each $\mu_k \in [\mu_1, \mu_2, \ldots, \mu_{N_\mu}]$:

(a) Estimate $\hat{\mathbf{p}}^k$ by inverting the training set $\mathbf{d}^o_{inv}$.

(b) Use $\hat{\mathbf{p}}^k$ to calculate the predicted testing set $\mathbf{d}^k_{test}$.

(c) Calculate the mean square error MSE_k using Eq. 19.

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

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 (white dots with black outlines 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.

2.7 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 hyper-parameters of the inversion. That is, their value influences the final solution but they are not estimated during the inversion. Both hyper-parameters 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 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 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 mean square error (MSE) as a measure of how well a given inversion output $\hat{\mathbf{p}}^k$ 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 $\hat{\mathbf{p}}^k$ on the known points to obtain the predicted depths $\mathbf{z}_s^k$. The MSE is defined as

$$MSE = \frac{\|\mathbf{z}_s^o - \mathbf{z}_s^k\|^2}{N_s}. \tag{20}$$

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

(i) For every combination of $z_{ref,l} \in [z_{ref,1}, z_{ref,2}, \ldots, z_{ref,N_z}]$ 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 $\mathbf{d}_{inv}^o$ using $z_{ref,l}$, $\Delta\rho_m$, 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 $\mathbf{z}_s^{l,m}$.

(c) Calculate the MSE between $\mathbf{z}_s^o$ and $\mathbf{z}_s^{l,m}$ using Eq. 20.

(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 hyper-parameters, though in a less formalized way.

2.8 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 modeling using tesseroids. We use the scipy.sparse package for sparse matrix arithmetic and linear algebra. The sparse linear system in Eq. 14 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 results can be found at the online repository `https://github.com/pinga-lab/paper-moho-inversion-tesseroids`. The repository also contains

instructions for replicating all results presented here. An archived version of this repository is also available at http://dx.doi.org/... (**Note to reviewers: the archived version will be uploaded upon publication**).

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 data set 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 application uses cross-validation to estimate all three hyper-parameters: μ, $\Delta\rho$, 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 $M_{lat} \times M_{lon} = 40 \times 50$ grid of juxtaposed tesseroids (a total of $M = 2000$ model elements). The anomalous Moho density-contrast is $\Delta\rho = 400\ kg/m^3$ and the Normal Earth Moho depth is $z_{ref} = 30\ km$. Fig. 4a shows the model Moho depths, where each pixel in the pseudo-color image corresponds to a tesseroid of the model.

The synthetic data were forward modeled on a regular grid of $N_{lat} \times N_{lon} = 79 \times 99$ 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. 4b shows the noise-corrupted full synthetic data set. 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.

We separated the synthetic data into training and testing data sets following Fig. 3. The training data set is a regular grid of $N_{lat} \times N_{lon} = 40 \times 50$ points (a total of $N_{train} = 2000$). The testing data set is composed of $N_{test} = 5821$ 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,

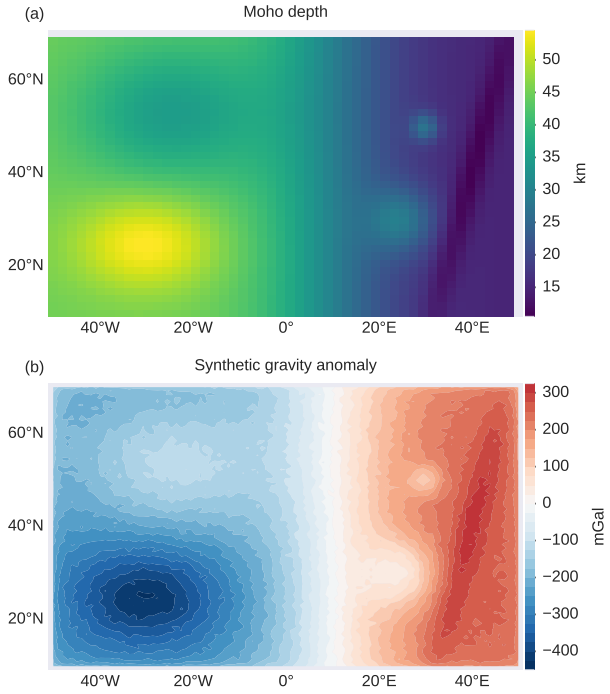

Figure 4. A simple Moho model made of tesseroids for synthetic data application. (a) The Moho depth of the model in kilometers. The model transitions from a deep Moho in the right to a shallow Moho in left, simulating the transition between a continental and an oceanic Moho. Each pixel in the pseudo-color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

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 mean square error (MSE, Eq. 19) for each estimate and chose as the final estimated Moho model the one that minimizes the MSE.

Fig. 5 summarizes the inversion results. Fig. 5a shows the final estimated Moho depth after cross-validation. The recovered model is smooth, indicating that the cross-validation procedure was effective in estimating an optimal regularization parameter. Fig. 5b 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. 5c shows inversion residuals (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. 5d, with 0.02 mGal mean and a standard deviation of 3.63 mGal. The cross-validation curve in Fig. 5e shows a clear minimum MSE at $\mu = 0.00046$

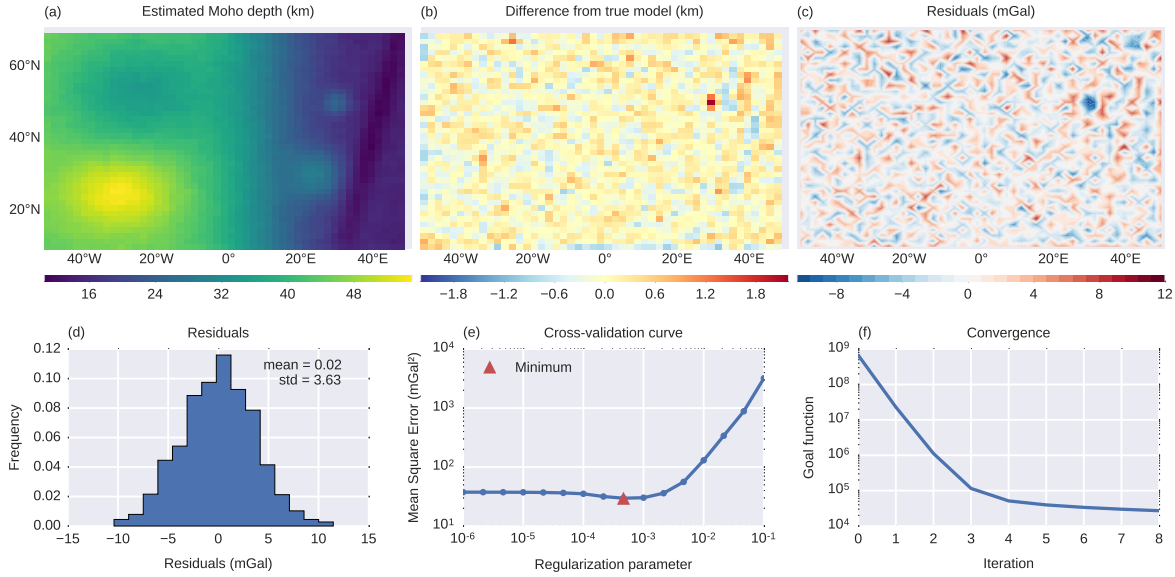

Figure 5. Results from the inversion of the simple synthetic data. (a) The estimated Moho depth. (b) The difference between the true model depths and the estimated depths. (c) The inversion residuals (observed data minus the data predicted by the estimate). (d) Histogram of the residuals. Also shown are the calculated mean and standard deviation (std) of the residuals. Note that the data were contaminated with normally distributed pseudo-random noise with zero mean and 5 mGal standard deviation. (e) Cross-validation curve used to determine the optimal regularization parameter (Eq. 11). Both axis are in logarithmic scale. The minimum Mean Square Error (Eq. 19) is found at $\mu = 0.00046$ (red triangle). (f) Goal function value (Eq. 11) per Gauss-Newton iteration showing the convergence of the gradient descent. The y-axis is in logarithmic scale.

(indicated by the red triangle). Fig. 5f 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 largest computational bottlenecks of the inversion: solving the linear system in Eq. 14 using the conjugate gradient method, performing the dot products required to compute the Hessian matrix (Eq. 13) and the gradient vector (Eq. 12), and forward modeling to calculate the predicted data (Eq. 3). The profiling results presented in Table 1 show that the time spent on forward modeling accounts for approximately 99.8% of the total computation time.

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

Function description	Time	Percentage of total time
Sparse conjugate gradient	0.021 s	0.050%
Sparse dot product	0.007 s	0.017%
Tesseroid forward modeling	42.059 s	99.824%

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 $M_{lat} \times M_{lon} = 80 \times 60$ juxtaposed elements, 4800 in total, using the Moho depths shown in Fig. 6a. In our model, the Normal Earth Moho is $z_{ref} = 30\ km$ 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 $N_{lat} \times N_{lon} = 159 \times 119$ points (a total of 18921 observations). We contaminate the synthetic data with normally distributed pseudo-random noise with zero mean and 5 mGal standard deviation (Fig. 6b).

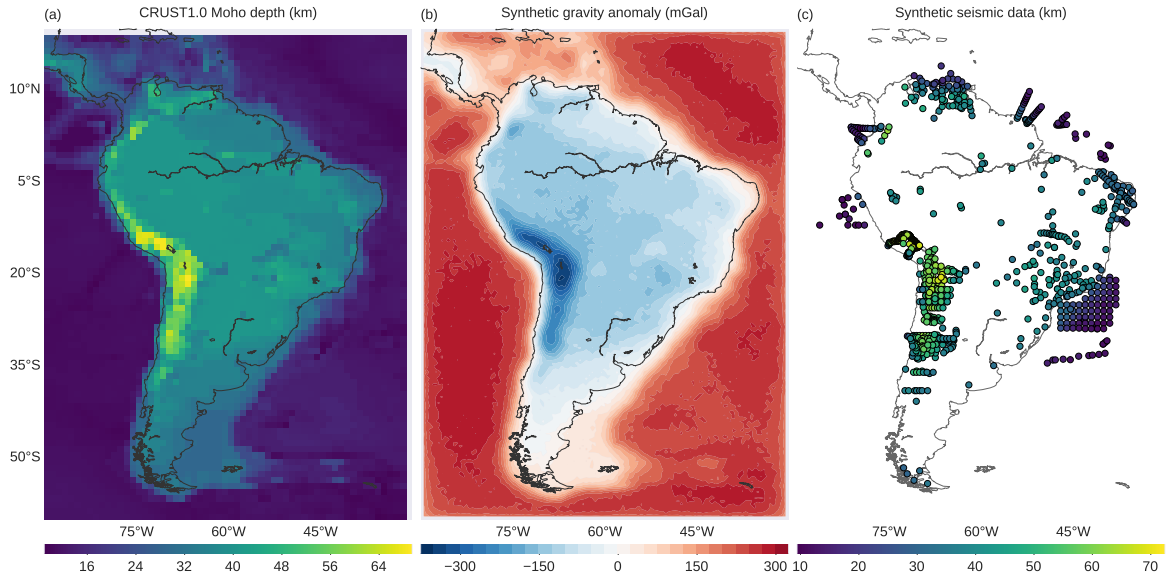

Figure 6. Synthetic data of a model derived from CRUST1.0. The model is made of tesseroids with an constant density-contrast of $\Delta\rho = 350\ kg/m^3$ and assuming a reference level of $z_{ref} = 30\ km$. (a) The Moho depth of the model in kilometers. Each pixel in the pseudo-color image corresponds to a tesseroid of the model. (b) Noise-corrupted synthetic gravity data generated from the model. (c) Synthetic seismic data simulating point estimates of Moho depth. The point estimates were obtained by interpolating the Moho depth in (a).

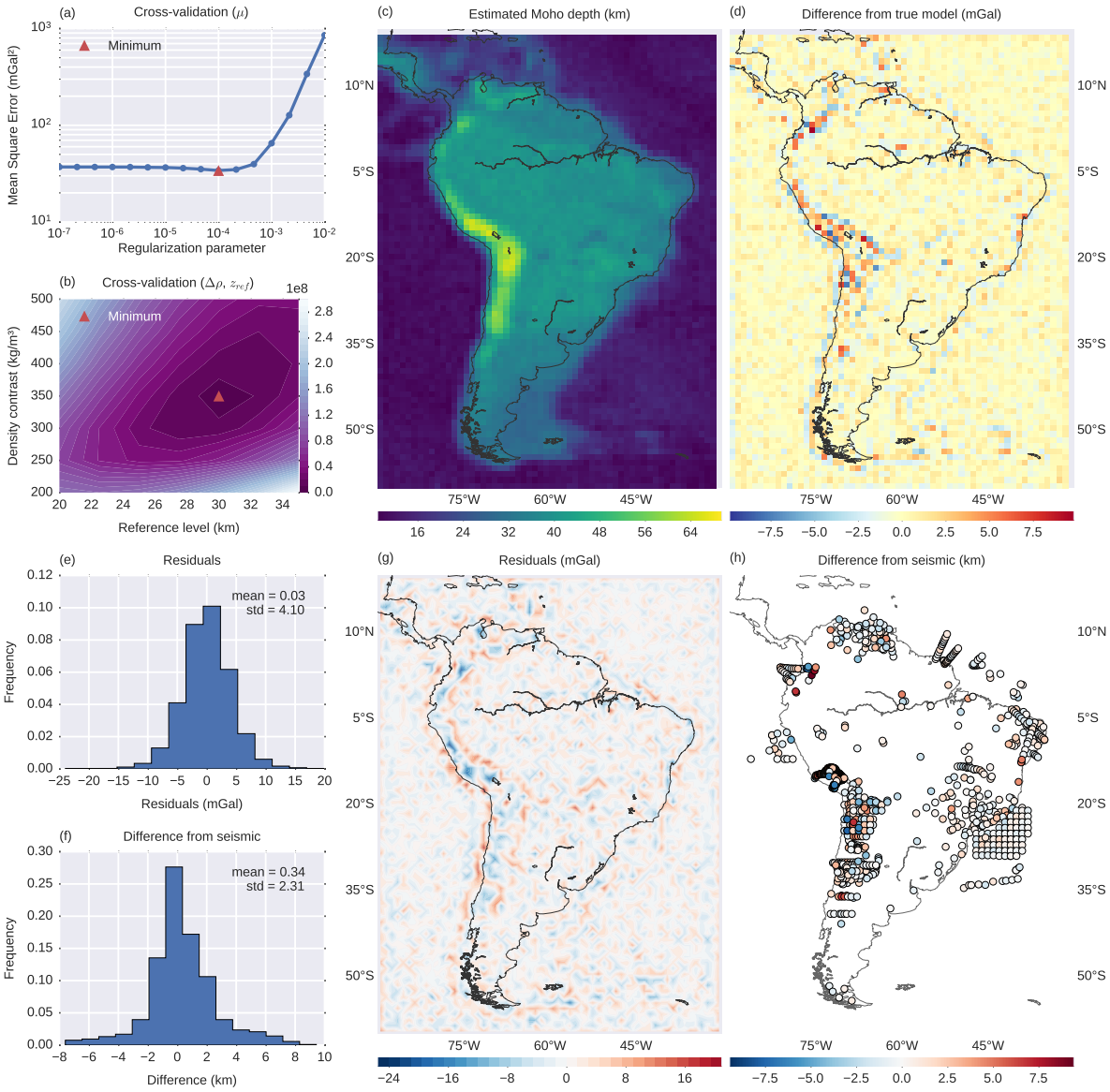

Figure 7. Inversion results from the CRUST1.0 synthetic data. (a) Cross-validation curve used to determine the regularization parameter (Eq. 11). The minimum Mean Square Error (Eq. 19) is found at $\mu = 0.0001$ (red triangle). (b) Cross-validation results used to determine the reference level (z_{ref}) and the density-contrast ($\Delta\rho$). The colored contours represent the Mean Square Error (Eq. 20) in km^2. The minimum (red triangle) is found at $z_{ref} = 30\ km$ 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.

The cross-validation procedure to determine $\Delta\rho$ and z_{ref} requires knowledge of the Moho depth at certain points ($\mathbf{z}_s^o$ in Eq. 20), 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. 6a on the same geographic coordinates as the 937 points from the Assumpção et al. (2013) data set. The resulting synthetic seismic data is shown in Fig. 6c.

We perform the cross-validation procedures 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. Our investigations suggest that the outcome of this round of cross-validation does not depend on the particular values of z_{ref} and $\Delta\rho$ used. Second, we use the estimated μ to run the cross-validation to estimate z_{ref} and $\Delta\rho$, thus obtaining the final estimated Moho depths. Fig. 7 summarizes the results from both cross-validation runs and the final inversion results.

For the first cross-validation, we separate the synthetic data (Fig. 3) into a training set with twice the grid spacing of the original data ($N_{lat} \times N_{lon} = 80 \times 60$) 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. 19), shown in Fig. 7a as function of μ. The optimal regularization parameter that minimizes the MSE is $\mu = 10^{-4}$ (indicated by the red triangle).

In the second cross-validation, 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 inversion for every combination of z_{ref} and $\Delta\rho$, totaling 49 inversions. Finally, we calculate the Mean Square Error (Eq. 20) for each of the 49 estimates and choose the values of z_{ref} and $\Delta\rho$ that minimize the MSE. Fig. 7b shows a colored-contour map of the MSE with a minimum (marked by the red triangle) at $z_{ref} = 30$ km and $\Delta\rho = 350$ kg/m^3.

Fig. 7c shows the final solution after both cross-validation procedures. The recovered model is smooth, indicating that the cross-validation procedure was effective in estimating an optimal regularization parameter. Fig. 7d shows the difference between the true Moho depths (Fig. 6a) and the estimated depths. 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. 7e and g show the inversion residuals (difference between the observed and predicted data). The inversion residuals appear normally distributed, with 0.03 mGal mean and a standard deviation of 4.10 mGal. The residuals follow a similar,

though reversed, pattern to the differences shown in Fig. 7d. 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. 7f 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 modeling 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.

4.1 Gravity and seismic 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. 8a show the calculated gravity disturbance of South America.

We remove the gravitational effect of the topography from the gravity disturbance by modeling the ETOPO1 digital terrain model (Amante & Eakins 2009, `http://dx.doi.org/10.7289/V5C8276M`) 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. 8c shows the calculated gravitational attraction of the to-

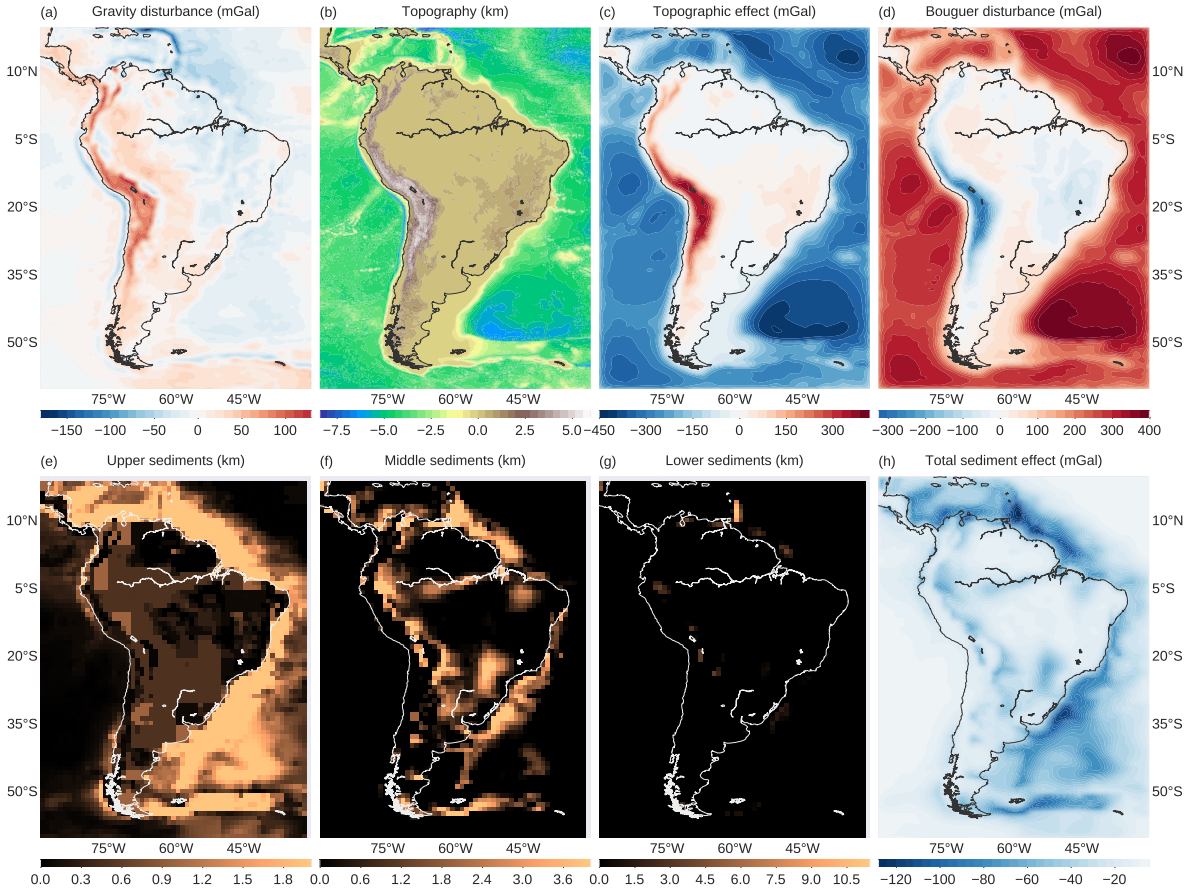

Figure 8. Gravity data for South America and the models used in the data corrections. (a) The gravity disturbance (Eq. 1) calculated from the raw gravity data. (b) Topography from ETOPO1. (c) Gravitational attraction of the topography calculated at the observation height using tesseroids. (d) The Bouguer 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.

pographic masses at 50 km height. Fig. 8d 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. 8e-g show the thickness of the upper, middle, and lower sedimentary layers, respectively. The density-contrasts of the tesseroid model is obtained by subtracting 2670 kg/m^3 from the density of each model element. Fig. 8h shows the combined gravitational attraction of the sedimentary basin tesseroid

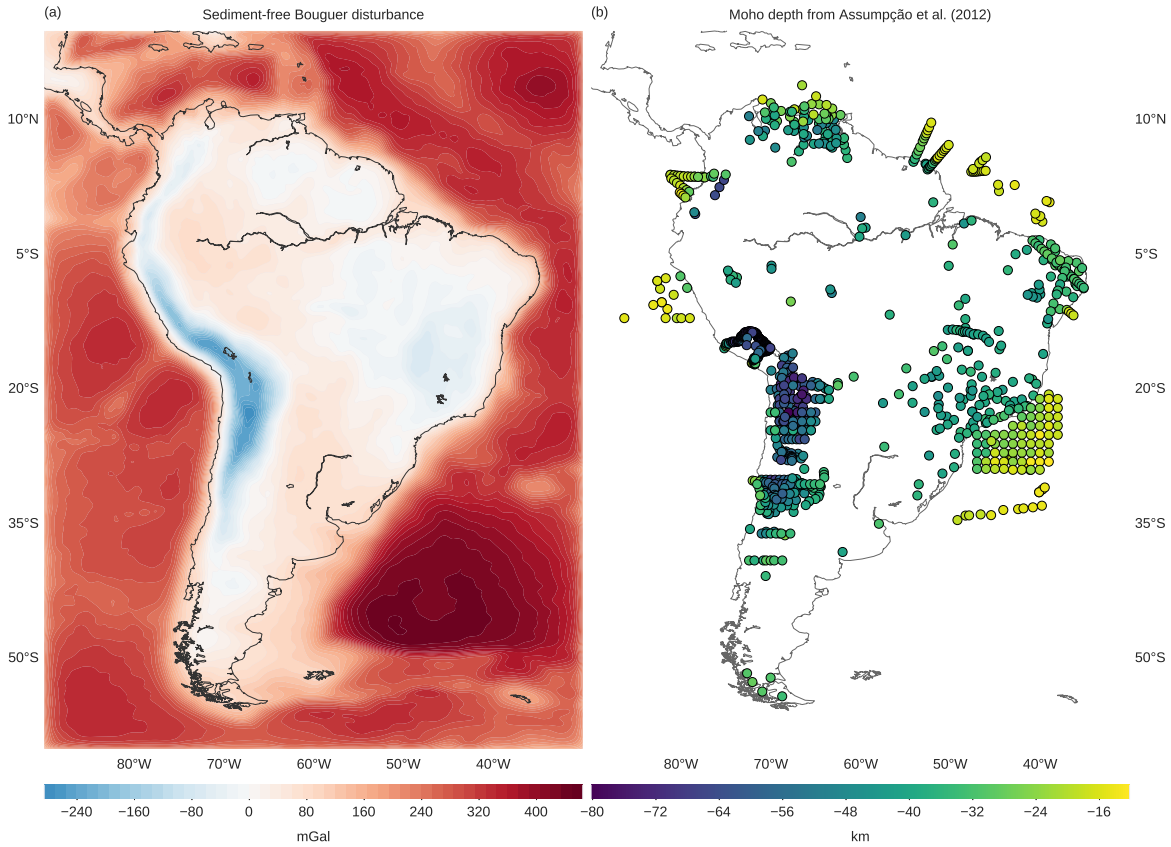

Figure 9. Input data for the South American Moho inversion. (a) Sediment-free Bouguer disturbance for South America. Obtained by subtracting the total sediment gravitational effect (Fig. 8h) from the Bouguer disturbance (Fig. 8d). (b) Seismological Moho depth estimates from Assumpção et al. (2013).

model. We subtract the total effect of sediments from the Bouguer disturbance in Fig. 8d to obtain the sediment-free Bouguer disturbance (Fig. 9a), which will be used as input for the inversion.

The seismic point estimates of Moho depth used in the cross-validation procedure are from the data set of Assumpção et al. (2013). The 937 data points in this data set are shown in Fig. 9b.

4.2 Inversion and cross-validation

As in the CRUST1.0 synthetic data test (section 3.2), we perform the cross-validation 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 cross-validation to estimate z_{ref} and $\Delta\rho$, thus obtaining the final estimated Moho depth model.

We split the sediment-free gravity data 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 $N_{lat} \times N_{lon} =$

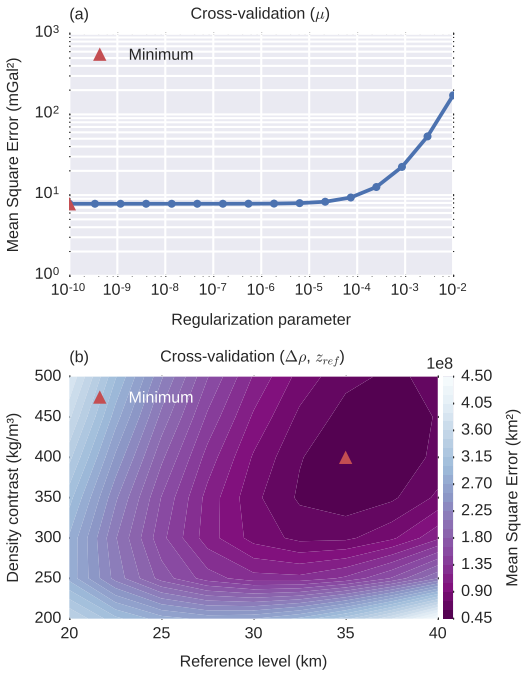

Figure 10. Cross-validation results for the South American Moho inversion. (a) Cross-validation to determine the regularization parameter μ (Eq. 11). The minimum Mean Square Error (Eq. 19), shown as a red triangle, corresponds to $\mu = 10^{-10}$. (b) Cross-validation to determine the reference level (z_{ref}) and the density-contrast ($\Delta\rho$). The colored contours represent the Mean Square Error (Eq. 20). The minimum (red triangle) is found at $z_{ref} = 35\ km$ and $\Delta\rho = 400\ kg/m^3$.

201×151 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. 10a 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 second cross-validation using $\mu = 10^{-10}$ in all inversions. We test all combinations of 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. Fig. 10b shows a map of the MSE with respect to the Assumpção et al. (2013) data set. The MSE has a well defined minimum, indicated by the red triangle, at $z_{ref} = 35\ km$ and $\Delta\rho = 400\ kg/m^3$.

4.3 Moho model for South America

The final Moho depth model for South America is shown as a pseudo-color map in Fig. 11a. The model is available in the online repository that accompanies this contribution (see section 2.8). Each model element is a $0.4° \times 0.4°$ tesseroid, represented by the pixels in the pseudo-color map.

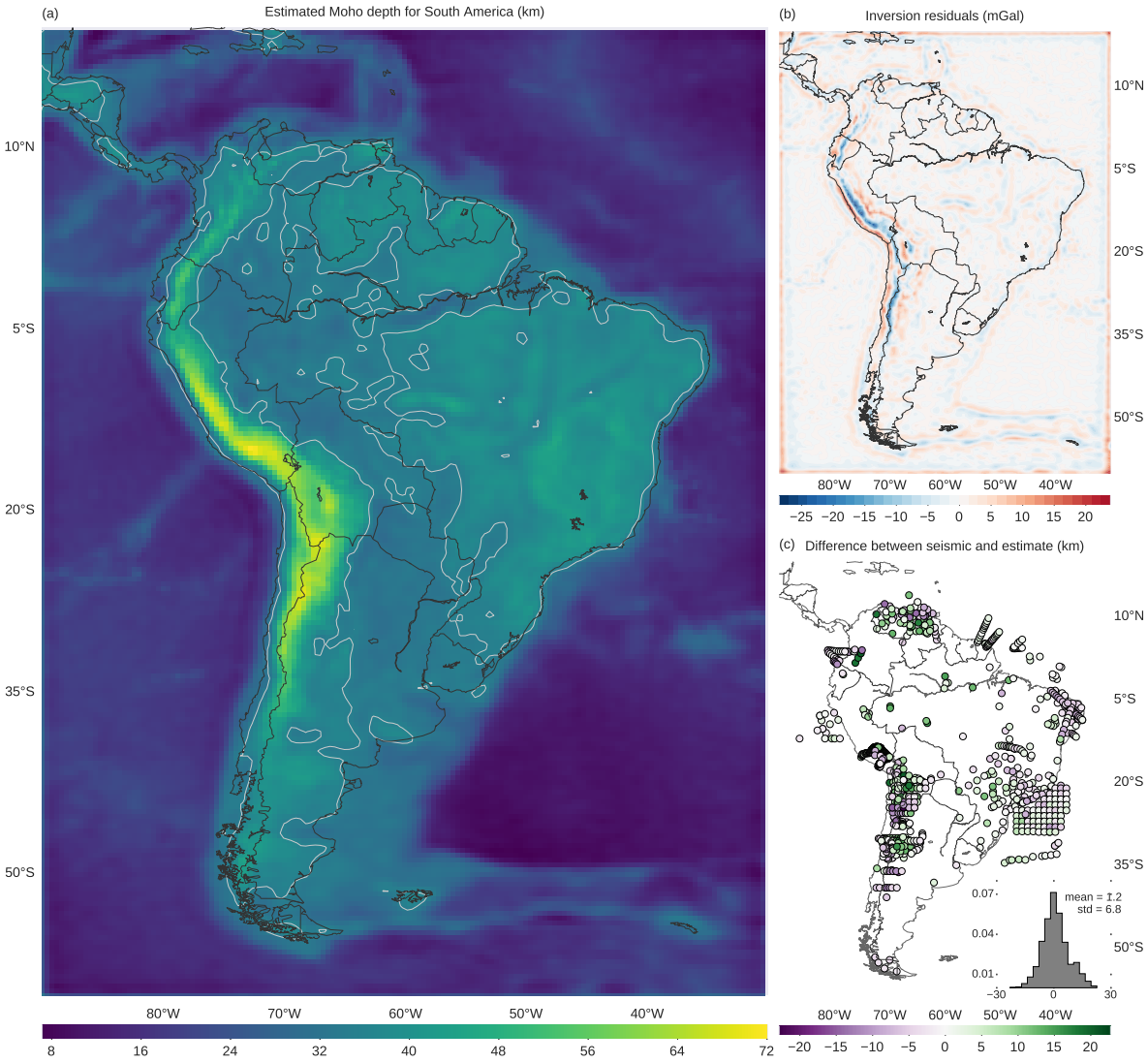

Figure 11. Inversion results for the South American Moho. (a) The estimated Moho depth of South America. The solid light grey line is the 35 km Moho depth contour. (b) Inversion residuals (observed data in Fig. 9a minus the data predicted by the estimate (a)). (c) Differences between the seismological depths of Assumpção et al. (2013) and our gravity-derived estimate shown in (a). The inset in (c) shows a histogram of the differences along with their calculated mean and standard deviation (std).

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 Guiana shields have a deeper Moho (greater than 35 km), with the deepest portion in the area of the São Francisco craton. The Moho is shallower than 35 km along the western Amazon and Andean foreland regions, as well as along the Amazon river.

Fig. 11b shows the inversion residuals (observed minus predicted data) and Fig. 11c shows the differences between the seismic-derived depths of Assumpção et al. (2013) (Fig. 9b) and the depths in our model. The differences range from approximately -23 to 23 km and have a mean of 1.18 km and a standard deviation of 6.84 km. The residuals and differences from seismic are smallest in the oceanic areas, southern Patagonia, and the eastern coast of the continent. The largest 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 center flanked by positive values to the East and West. This same pattern is observed in the CRUST1.0 synthetic test results (Fig. 7), suggesting that this is a byproduct of the inversion method, not the data. Likewise, larger residuals also appear to be associated with sharp variations in the estimated Moho depth. Along the Andes, large differences with seismic data are correlated with the large inversion residuals. Conversely, this correlation is absent from the large differences seen in points around Venezuela. In the Borborema province, northeastern Brazil, our model slightly overestimates the Moho depth. On the other hand, our model underestimates the depths in the Amazon region and the Paraná basin. Particularly in the Amazon basin, where our model predicts a Moho depth of approximately 30 km, the residuals and the differences with the seismic data are larger than in the Paraná basin.

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 (Silva et al. 2014) to use tesseroids as model elements and smoothness regularization. We retain the computational efficiency of Bott's method by taking advantage of the sparse nature of all matrices involved. We employ two cross-validation techniques to estimate the hyper-parameters of the inversion: the regularization parameter, the Moho density-contrast, and the Normal Earth Moho depth.

The test on simple synthetic data shows that our inversion method is able to recover a smooth Moho relief with a homogeneous density-contrast. The inversion was not able to fully recover the shortest wavelength feature in the model, possibly due to the smoothness constraints which tends to soften high-frequency (sharp) variations. The cross-validation Mean Square Error curve in Fig. 5e 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 smooth Moho relief and acceptable data misfit.

The source code profiling results presented in Table 1 confirm the efficiency of the proposed method. When using sparse matrices, solving linear systems and performing matrix multiplications together account for a mere 0.067% of the total computation time required for a single inversion.

The majority of the computation time (99.824%) is spent on forward modeling. Thus, we are able to retain the high computational efficiency of Bott's method while using 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. 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).

The more complex synthetic data test based on CRUST1.0 (Fig. 7) shows that the cross-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 Total Variation (Martins et al. 2011), Cauchy norm regularization (Sacchi & Ulrych 1996; Pilkington 2008), or an adaptive mixed smoothness-sharpness regularization (Sun & Li 2014).

We applied the method proposed here to estimate the Moho depth for South America. Our Moho depth model is in accordance with previous results by van der Meijde et al. (2013). 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. This might indicate the improper use of smoothness regularization, as suggested by the CRUST1.0 synthetic data test, or the presence of crustal or mantle density anomalies that were unaccounted for during the data corrections. In the coastal region of Venezuela, along the central Amazon and Solimões basins, and in the Paraná basin, the model is able to fit the gravity data but differs significantly from the seismic data. Mariani et al. (2013) and Nunn & Aires (1988) explain these discrepancies in the Paraná and Amazon basins, respectively, 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.

6 ACKNOWLEDGMENTS

We are indebted to the developers and maintainers of the open-source software without which this work would not have been possible. 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).

Coloco o projeto da FAPERJ também?

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, *Geophysical Journal International*, **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., 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), *Journal of Geodesy*, **86**(10), 932–934.

Bott, M. H. P., 1960. The use of Rapid Digital Computing Methods for Direct Gravity Interpretation of Sedimentary Basins, *Geophysical Journal International*, **3**(1), 63–67.

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

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

Hansen, P., 1992. Analysis of Discrete Ill-Posed Problems by Means of the L-Curve, *SIAM Review*, **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, *Journal of Geodesy*, **81**(2), 121–136.

Hunter, J. D., 2007. Matplotlib: A 2D graphics environment, *Computing in Science & Engineering*, **9**(3), 90–95.

Jones, E., Oliphant, T., Peterson, P., & others, 2001. SciPy: Open source scientific tools for Python, http://www.scipy.org/, Accessed 22-08-2015.

Kelley, C. T., 1987. *Iterative Methods for Optimization*, Society for Industrial and Applied Mathematics, Philadelphia, 1st edn.

Kim, J.-H., 2009. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap, *Computational Statistics & Data Analysis*, **53**(11), 3735–3745.

Laske, G., Masters, G., Ma, Z., & Pasyanos, M., 2013. Update on CRUST1.0 - A 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, *Journal of Applied Geophysics*, **73**(3), 221–231.

Mariani, P., Braitenberg, C., & Ussami, N., 2013. Explaining the thick crust in Paraná basin, Brazil, with satellite GOCE gravity observations, *Journal of South American Earth Sciences*, **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., Pail, R., Gruber, T., Fecher, T., Rexer, M., Schuh, W. D., Kusche, J., Brockmann, J. M., Rieser, D., Zehentner, N., Kvas, A., Klinger, B., Baur, O., Höck, E., Krauss, S., & Jäggi, A., 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, *Journal of Geophysical Research*, **93**(B1), 415.

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, *Geophysical Journal International*, **31**(4), 447–455.

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

Pilkington, M., 2008. 3D magnetic data-space inversion with sparseness constraints, *GEOPHYSICS*, **74**(1), L7–L15.

Reguzzoni, M., Sampietro, D., & Sanso, F., 2013. Global Moho from the combination of the CRUST2.0 model and GOCE data, *Geophysical Journal International*.

Sacchi, M. D. & Ulrych, T. J., 1996. Estimation of the discrete Fourier transform, a linear inversion approach, *Geophysics*, **61**(4), 1128–1136.

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., 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, *Geophysical Journal International*, p. ggu067.

Uieda, L., 2015. A tesserioid (spherical prism) in a geocentric coordinate system with a local-North-oriented coordinate system, http://dx.doi.org/10.6084/m9.figshare.1495525, Accessed 02-11-2015.

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

Uieda, L., Barbosa, V. C. F., & Braitenberg, C., 2016. Tesseroids: forward modeling gravitational fields in spherical coordinates, *Geophysics*, In press (accepted).

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, *Journal of Geodynamics*, **84**, 1–18.

Waskom, M., Botvinnik, O., Hobson, P., Warmenhoven, J., Cole, J. B., Halchenko, Y., Vanderplas, J., Hoyer, S., Villalba, S., Quintero, E., Miles, A., Augspurger, T., Yarkoni, T., Evans, C., Wehner, D., Rocher, L., Megies, T., Coelho, L. P., Ziegler, E., Hoppe, T., Seabold, S., Pascual, S., Cloud, P., Koskinen, M., Hausler, C., kjemmett, Milajevs, D., Qalieh, A., Allan, D., & Meyer, K., 2015. seaborn: v0.6.0, http://dx.doi.org/10.5281/zenodo.19108, Accessed 02-11-2015.

Wieczorek, M. A. & Phillips, R. J., 1998. Potential anomalies on a sphere: Applications to the thickness of the lunar crust, *Journal of Geophysical Research: Planets*, **103**(E1), 1715–1724.

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