Dissertation Methodology

Notes:

* Mention that this method is pretty much the same as Uieda only difference is the cross validation.
* Brief overview on parametrization, forward problem, inverse problem, regularisation, and Bott’s method.
* Cross validation – split up have training size ranging between 2/3, 3/4, and 4/5 to see the mean error between gravity and seismic estimates. Want to see how many points are needed to still get relatively accurate values (small MSE’s) so we know that when seismic point estimates are not present the gravity moho calculated is still accurate (at around 2.3-2.5km error).
* Statement about using python software.

Methodology:

To accurately obtain a model of the Moho depth one must first remove all other effects that contribute overall gravitational values recorded over an area. This is achieved through correcting for things such as removing the scalar gravity of an ellipsoidal reference Earth (the Normal Earth) and then the removal of all other effects via a Bouguer correction otherwise known as the correction for topography. Initially though the effect of the Normal Earth needs to be removed from the same point as where the gravity observation was made and is calculated from the closed-form solution in Li & Götze (2001). The value obtained here is called the gravity disturbance and can be seen in the equation below (insert gravity disturbance equation once on .tex file). (Also include Figure 1 in Uieda & Barbosa paper, stages of gravity correction).

The gravity disturbance is still not a direct result of the change in density associated with the Moho discontinuity but also an amalgamation of topography with reference to the normal ellipsoid, variations of density in the crust (e.g. sedimentary basins and igneous intrusions), anomalies below the upper mantle, and mass deficiency due to the oceans. These effects are removed through a topography correction, (insert topography correction equation). This is the method used in Uieda & Barbosa (2017) and assumes that the effects of other crustal and mantle sources are negligible, and after this correction the gravitational values attained are purely a result of the density variations either side of the Moho discontinuity. Seeing as the data is obtained for a sufficiently large area (South America) this correction for topography amongst other things is calculated using tesseroids as part of a spherical Earth approximation. Effects of the tesseroids are calculated using a GLQ integration presented in Asgharzadeh et al. 2007) and improved upon in Uieda et al. 2016 through the adaptive discretization scheme. (insert Figure 2 in Uieda & Barbosa, tesseroids).

Overview of the methodology of calculating a gravity derived Moho model from Uieda (2017)

Note: for each section don’t write more than 200 words, around 1000 overall for this bit.

Parametrization, the forward problem, inverse problem, regularization, Bott’s method, estimating the parameters:

As the implementation of cross validation to estimate uncertainty in the model is implemented as part of the code used in the Uieda & Barbosa (2017) paper the method in calculating the model is largely similar. For context, an overview of this method will be given but for more detail see Uieda & Barbosa (2017).

Upon the calculation of the Bouguer disturbance from the removal of topography, sediments etc. the forward model is parameterized by discretizing the anomalous Moho onto tesseroids. The forward model aims to calculate the difference between the Normal Earth Moho and the true Moho depth, and depending on which is shallower will result in either a positive (red) or negative (blue) density contrast displayed by the colour of the tesseroids. The overall absolute value of the density contrast is predetermined, this produces a nonlinear problem with equation, (insert equation 3 from Uieda, don’t include i values just generic equation). Where d is the data vector, p is the parameter vector containing Moho depths, and f is the non-linear function.

Leading on to the inverse problem the parameter vector is estimated using least squares that reduces the misfit to the data, (insert least-squares equation). Where d0 is the observed gravity data, the equation means that this is a non-linear inverse problem, but we can calculate the minimum error using optimization, where a perturbation vector Δp0 is iterated until a minimum is reach which leads to the minimum value of φ(p).

The optimization of the least squares estimate however is not enough for estimating the relief associated with the Moho and needs regularization in the form of a first-order Tikhonov regularization (Tikhonov & Arsenin 1977) to ensure smoothness in the model and no sharp jumps in Moho depth, to provide a realistic model. (inset regularization equation). R a matrix composed of first order differences between tesseroid depths. This along with the least squares estimate leads to an inverse problem that is solved by minimising the goal function, (insert goal function), µ is the regularization parameter that helps control the fit to the observed data and the smoothness.

After the rearrangement and substitution of equations we arrive at a linear equation system which can calculate the real Earth Moho depths with reference to the Normal Earth Moho. (insert linear equation system equation), where Ak is the Jacobian matrix, and Δpk is the parameter perturbation vector.

Bott’s method from Bott, 1960, is a method that calculates the thickness of a sedimentary basin based of gravitational data, the method is iterative so recalculates a new vector of basement depths from the previous calculation until a value where the residuals (equation numerator) fall below the noise level. (insert Bott’s method equation) Where Δp is the density contrast between the sediment and the reference density, and G is the gravitational constant (6.67x10-11 m-3 kg-1 s-2). However, in 2014 Silva et al showed that Bott’s method can be written as (insert special case Bott’s method) and the main advantage to this is that this method does not need the solution of equation system.

For calculating the depth to the Moho Uieda & Barbosa, 2017 uses Bott’s method in the inversion process and adapts it onto a spherical coordinate system using tesseroids. Stating in the paper that this method retains the efficiency of Bott’s method while accounting for the stability problem previously present in the method. Following this step, the final part of the method involves calculating the hyperparameters (regularization parameter µ, Moho density-contrast Δp, and depth of the Normal Earth Moho zref) which will be used in the inversion process. The calculation of the regularization parameter is through a method of hold out cross validation from Hansen, 1992 and from this optimal regularization value the other two hyperparameters can be calculated as these depend on the value of the regularization but not vice versa. The main way these parameters are calculated is by finding the smallest Mean Square Error (MSE) through a cross validation method which compares known Moho depths to calculated ones from the hyperparameters and picks the values with the smallest associated MSE.

Implementation of cross validation in error estimation

In order to test the quality of the gravitationally derived Moho model a method of Cross Validation (CV) will be used to compare the model to seismic point estimates to calculate the Mean Square Error (MSE) and subsequently the Mean Error. This value gives the uncertainty of the Moho model so in areas where no point data is available we can predict the accuracy of the model with an appropriate error estimate.

Cross validation as mentioned will be used to determine how well this model will perform with the absence of seismic points. In particular k-fold random subsampling will be utilized, by splitting the full data randomly into a training and testing set, the training set is compared to the model whilst the testing set is held back and used to attain an error estimate between the data and the model. (find and insert figure here about random subsample cross validation). This process will be repeated k times for a number of different training set sizes ranging between 2/3, 3/4, and 4/5 of the full data these sizes are selected as usually the training set contains 70% to 90% of the full data to be considered useful (Berrar, 2018). The point of varying the ratio of the training to testing (validating) set is to see how many seismic point estimates are needed for a model, in this case a model of South America, to produce a reasonable estimate of the uncertainty or error in the depth to the Moho discontinuity. The estimate reached will come from a histogram plot of all the MSE values for a select training size from which the average MSE value can be calculated which will give the error on the model.

(Need to finish this section)

Possible rewrite?

Steps:

* summary of cross validation method in short steps
* Explanation in detail of the steps.
* What do the results show without specifics.

Software Implementation

This inversion and error estimation method put forward in the methodology is executed in the Python programming language. Software is available under the BSD 3 clause open-source software license. The code in this project depend on open-source libraries scipy and numpy (Harris et al., 2020) for computational number exploitation, matplotlib (Hunter, 2007, <http://matplotlib.org>) and seaborn (Waskom et al. 2015, <http://stanford.edu/∼mwaskom/software/seaborn>) for plots and maps, Fatiando a Terra (Uieda et al. 2013, <http://www.fatiando.org>) for geophysics tasks. scipy.sparse package is implemented for use on sparse matrix arithmetic and linear algebra and solves the linear equation system equation.

The use of Jupyter notebooks (Perez & Granger 2007, <http://jupyter.org/>), which merge the source code, results, and figures of the project.

All source code, Jupyter notebooks, data, and error estimate results are available through an online repository (<https://github.com/compgeolab/moho-uncertainty>).