Investigating uncertainties in travel time tomography using the null space shuttle

Final draft

Ralph de Wit

November 26, 2010

Abstract

Due to the underdetermined nature of large tomographic inverse problems, a sizable null space exists. It is therefore important to investigate the uncertainties in tomographic models produced by inverse problems with multiple solutions. Conventional methods to analyze model quality all have shortcomings, which we discuss in this report. As an alternative, we use the the null space shuttle technique (Deal and Nolet, 1996). This method has been designed to exploit components of the model null space, along with a priori information or a physical model, in order to improve or enhance the original solution. We generalize the null space shuttle technique to analyze the robustness of a global P wave speed perturbation mantle model PMIT08 produced by a classical study of travel time tomography (Li et al., 2008) and examine a range of models that is consistent with the travel time data. We find that the RMS amplitude of velocity perturbations within this set of solutions ranges from 0.2 to 0.6 % in the lowermost mantle and 0.3 to 1.3 % in the upper mantle and deduce that the travel time data provide little constraint on the amplitudes of the solution. Furthermore, solutions exist that contain structures different in geometry than those imaged in PMIT08. Upper mantle portions of slab-like anomalies are altered or removed from the tomogram in some regions, while the vertical extent of low-velocity anomalies in PMIT08 appears to be poorly constrained. On the contrary, most slab-like anomalies in the lower mantle seem to be robust. We strongly advise against physical interpretations of a single solution of an underdetermined tomographic problem. The null space shuttle technique is straightforward to implement and can be utilized in an efficient way to analyze the solution robustness. We suggest that this technique should be routinely applied before physical interpretation of tomographic images are made.

Contents

1 Introduction

The tomographic imaging of the deep interior of the Earth has since its advent in the 1970's evolved into a branch of seismology comprising a wide range of techniques, including tomographic modelling of body waves (e.g. Dziewonski et al. (1977); Ritsema et al. (1999); Kárason and van der Hilst (2001)), surface waves (e.g. Trampert and van der Hilst (2005); Ekström et al. (1997)), normal modes (e.g. Li et al. (1991); Ishii and Tromp (2004)) and more recently full waveforms (e.g. Fichtner et al. (2009); Tape et al. (2009)). Much work has been done on travel time tomography, a method that uses arrival times of seismic phases generated by earthquakes to infer the velocity structure within the Earth. One of the advantages of travel time tomography is the abundance of arrival time data and the relatively straightforward utilization of these data. Over the past four decades, thousands of receivers around the globe have registered millions of seismic phase arrivals generated by thousands of earthquakes. These data have been faithfully catalogized into large arrival time databases, such as the Bulletin of the International Seismological Centre (ISC), making them available to every seismologist willing to use travel time data to image the Earth. However, despite the large amount of data available, there are severe shortcomings to the data set. The most prominent problem with arrival time data sets is the uneven sampling of the Earth's interior by body waves as a result of the biased distribution of both sources and receivers. Almost all earthquakes occur along plate boundaries, while seismometers mainly reside on land. This uneven data coverage is most problematic in the upper mantle below oceanic regions and in the Southern hemisphere, where far less stations are installed than on the Northern hemisphere (Li et al. (2008), Figure 3 therein). Furthermore, the travel time data include both random and systematic errors that are difficult or even impossible to accurately estimate, although attempts have been made (e.g. Gudmundsson et al. (1990); Röhm et al. (1999)). In other words, in travel time tomography (and most other seismic tomographic problems) one has contradictory information on some model parameters, while other parameters can not be determined due to lack of information and the inverse problem is ill-posed. In general, the quality of a tomographic image is limited by the quality of the data, the data coverage, i.e. the uniformity of the spatial distribution of the data, and approximations made in the formulation of both forward and inverse problem.

Until now, much time in seismic tomographic studies has been devoted to the improvement of tomographic images by reducing the minimum resolvable length scales of structures. Reliable analyses of solution quality on the other hand are far from well developed, although solution verification is a critical aspect of any study. To analyze the quality of a tomographic image, studies of seismic tomography often employ sensitivity tests (e.g. van der Hilst et al. (1997); Bijwaard et al. (1998); Kárason and van der Hilst (2000)) or attempt to calculate the resolution (Aki et al., 1977; Boschi, 2003; Soldati and Boschi,

2005) and covariance matrices (Vasco and Johnson, 2003; Soldati et al., 2006). However, sensitivity tests have significant shortcomings (Lévêque et al., 1993). We will discuss the shortcomings of such sensitivity tests in the current report. The resolution matrix R describes the fictitious coupling between solution coefficients (model parameters), while the covariance matrix of the solution coefficients describes the uncertainties associated with the model parameter estimates and their correlations. However, the knowledge of both resolution and model covariance matrices does not yet define the quality of the solution. At best, these matrices could serve as an indication how true Earth has been blurred in the tomographic image by the tomographic inversion.

Clearly, a more robust and accurate analysis of the quality of tomographic models and associated uncertainties is needed. Deal and Nolet (1996) designed the null space shuttle to exploit components of the model null space of a certain tomographic problem, in combination with a priori information, to enhance the corresponding tomographic image (Deal et al., 1999). We generalize this technique and use the null space shuttle to investigate the robustness of a model obtained from a classical travel time tomography study (Li et al., 2008). We use theoretical models based on the original tomographic image to provide some quantitive bounds on the uncertainties in the model. In addition, we illustrate what is and what is not constrained by the data in the original image.

2 Data

We use the same data set as Li et al. (2008) and therefore only briefly summarize the data here. For further details, we refer the reader to this study. Li et al. (2008) use travel time residuals with respect to travel times computed from *ak135* (Kennett et al., 1995). The data set comprises millions of travel times and consists of three types of data: (1) routinely picked and processed travel times from global and regional networks; (2) differential times measured by waveform cross correlation; and (3) phase arrivals from temporary arrays. The first category contains the EHB data set (Engdahl et al., 1990) and data from the Annual Bulletin of Chinese Earthquakes (ABCE) and Chinese Seismograph Network (e.g. Li et al. (2006)). The EHB database consists of ISC data that were reprocessed by relocating all events therein that are teleseismically well constrained. The EHB data used here contains more than 10 ten million travel time residuals associated with more than 450,000 well constrained teleseismic and regional earthquakes and includes P, pP, pwP, Pg and Pn phases. Waveform-based differential times include high-frequency PKP differential travel times (McSweeney, 1995), P_{diff} phases that are diffracted along the core-mantle boundary, measured at ∼0.05 Hz (Wysession, 1996) and long-period $PP-P$ data, measured at a frequency ∼0.04 Hz (Bolton and Masters, 2001). The processing of these differential travel time data and their integration with the routinely processed travel time data is described by Kárason and van der Hilst (2001). The third part of the data set comes from temporary arrays in Australia (van der Hilst and Mann, 1994), the Tibetan plateau and SW China (Li et al., 2006), Africa (e.g. Benoit et al. (2006)) and the United States (*USArray*, Burdick et al. (2008)). The distribution of seismograph stations and a schematic overview of the ray paths of the phases used in this study can be found in Li et al. (2008). Despite the addition of regional data, data coverage remains uneven across the Earth's surface. A second data-related problem comes from data errors and we therefore summarize knowledge on these errors from other studies here.

2.1 Data errors

In travel time tomography, the data fit, or data residual, is an important criterion to analyze the agreement between real data and data predicted by an acquired solution. One can readily state that the relevance of a certain data fit, and therefore the choice of a corresponding solution, is questionable when errors in the data are large. Accurate knowledge of both random and systematic errors in the data is therefore desirable. Over the years, the vast quantities of seismic data in the various existing databases have been recorded and processed with varying precision. It is therefore difficult to obtain a reliable estimation of the errors in the millions of data available for travel time tomography. A few studies have attempted to estimate observational errors in the ISC database. Gudmundsson et al. (1990) estimate random errors in teleseismic ISC delay times by performing statistical tests and find a value of 0.5 s. Amaru et al. (2008) approximate picking errors in the ISC data as 0.75 s and claim that this value is in agreement with the estimate of Gudmundsson et al. (1990). The millions of travel time data can be used to average over closely spaced ray paths and thereby reduce observational errors by the concept of summary or composite rays (Spakman and Nolet, 1988). A composite ray is a ray bundle forged from all rays travelling from a event cluster volume to a single station and intersects all cells traversed by the rays in the ray bundle. If the bundle is narrow, rays are expected to have approximately the same delay and differences in delay time are due to observational errors. Using this concept, Bijwaard et al. (1998) find by extrapolation to an infinitely narrow ray bundle that the limit value for the standard deviation is approximately 0.3 s and assume this value to be representative for observational errors. However, this estimate is not directly applicable to the composite rays used in the tomography by Li et al. (2008). The reprocessing of the ISC data by Engdahl et al. (1990) may have reduced original observational errors somewhat, but no estimates of such errors are currently available. While reading or picking errors are presumably random (Nolet, 2008) and can therefore be reduced by using composite rays, this is not the case for systematic errors in the dataset. Röhm et al. (1999) test delay times of the EHB database for systematic variations in time. They find temporal variations of the median delay time of 0.5-1.0 s, in agreement with Grand (1990) who reports a late arrival bias in the ISC data of 0.5 s related to the gain of stations. Because of their systematic nature, Röhm et al. (1999) expect that these errors will not necessarily cancel out by using the large number of travel times in the ISC Bulletin. Error estimates for data from regional networks used in this study are scarce or non-existent. Schaff and Richards (2004) find an average of 2 s for P phases as the relative picking error in the ABCE data. For a temporary array in Ethiopia, Benoit et al. (2006) estimate uncertainties in delay times to be no greater than 0.15 s.

3 Methodology

The relationship between the data and model parameters, $d = g(m)$ forms the basis of all tomographic methods. For an observed data set d_{obs} and an initial model m_0, the difference $(g(m_0) - d_{obs})$ presents an indication of how well this initial model satisfies the observed data. The goal of the inverse problem in tomography is to minimize this difference between observed and predicted data, the data residual, by adjusting the model parameters, provided some regularization that may be imposed. In particular, in travel time tomography the goal is to translate these differences between observed and theoretical travel times into aspherical variations in wave speed in the Earth's interior. The final model will be a (mathematical) representation of true Earth structure. The quality of this representation, i.e. the accuracy of the model, depends on various factors, including:

- The data residual, i.e. how well the predicted model satisfies the observed data
- Errors in the observed data
- The quality of the predicted (theoretical) data obtained from $g(m)$
- The extent to which the data constrain the model parameters
- The parameterization of the model and assumptions made therein

In summary, the steps required to obtain a tomographic image from observed seismic data can be defined by the following:

1. Model parameterization: The seismic structure of the studied region is represented by a set of unknown model parameters. The choice of parameterization restricts the range of possible solutions and can therefore be seen as a form of ad hoc regularization. The structure may be parameterized by a set of blocks of constant value or nodes with a specified interpolation function (e.g. splines), a set of interfaces of which the geometry can be varied to fit the data or a combination of such parameters. The most straightforward approach involves the use of a regular static parameterization, but other

possibilities include irregular parameterizations, in which the minimum scale length of structure is variable, and adaptive parameterizations, where the inversion process is involved to adjust the location and/or number of parameters to better suit the resolving power of the data. An initial estimate of model parameters is often required by the tomographic method used.

2. Forward calculation: The method used to produce predicted data from a certain set of model parameter values ($d = g(m)$). In seismic tomography, many techniques exist to solve the forward problem, most of which make use of the high-frequency assumption of geometric optics (ray approximation). Techniques include both ray-based (shooting and bending) and grid-based methods (e.g. eikonal solvers), which are designed to solve the data prediction problem by finding the path taken by seismic energy from source to receiver.

3. Inversion: the manipulation of model parameters in order to reduce the data residual, i.e. difference between observed and predicted data, while simultaneously abiding by imposed regularization. This process is usually performed iteratively in combination with the preceding step. One of the main problems in all practical applications of seismic tomography is that of solution non-uniqueness. This non-uniqueness is commonly dealt with by imposing regularization that reduces the amount of data-fitting solutions. A drawback of this approach is that often subjective regularization, i.e. regularization not based on somewhat objective physical constraints, is imposed. Another important issue in most studies of seismic tomography concerns the non-linearity of the inverse problem. Often, the inverse problem is linearized around a certain reference model.

4. Analysis of solution quality: the robustness of the solution is often investigated by either calculating models from synthetic data sets or estimating resolution and covariance operators associated with the performed tomographic inversion.

The latter category is of interest here. Analysis of the quality of a result is vital to any study. As mentioned in the *Introduction* section of this report, traditional measures for analysis of solution robustness in travel time tomography have some major shortcomings. As an alternative for analyzing solution quality, we use the null space shuttle technique as introduced by Deal and Nolet (1996). Although the method of travel time tomography used here is quite standard nowadays, we first summarize the general methodology for travel time tomography here. Subsequently, the details of the global tomographic inversion used in the current study (following Li et al. (2008)) are briefly described. Furthermore, the usage of the null space shuttle technique is outlined.

3.1 Seismic travel time tomography

In seismic travel time tomography the observed data are travel times and the model parameters represent velocity variations. In a medium of continous velocity $v(\bar{x})$ the travel time of a ray can be written as:

$$t = \int_{L(v)} \frac{1}{v(\bar{x})} \mathrm{d}l \tag{3.1}$$

where L is the ray path and $v(\bar{x})$ is the velocity field. Since the ray path depends on the velocity, (3.1) is non-linear. This non-linearity makes the inverse problem much harder to solve. Several approaches exist to solve tomographic inverse problems, which can be loosely separated into three categories: linear tomography, iterative non-linear tomography and fully non-linear tomography. However, the latter approach is not used in travel time tomography in practice. In the first approach, ray paths are calculated only once for a certain initial or reference model. Consequently, the relationship between velocity anomalies and travel times is linearized about this reference model and modifications of the velocity field are made relative to the reference model. Iterative non-linear tomography also ignores the dependence of the velocity correction to the source-receiver path, but does account for the non-linear aspect of the problem by iteratively manipulating the model parameters and subsequently re-tracing rays. This can be viewed as repeating steps 2 and 3 in the procedure described above. In this study we adopt the framework of linearized travel time tomography, which is commonly used in other studies (e.g. van der Hilst et al. (1997); Kárason and van der Hilst (2000); Li et al. (2008)). These studies involve up to millions of data and a number of model parameters on the order of 10^5, making it computationally expensive to use the iterative non-linear approach, although some efforts have been made using this method (e.g. Bijwaard and Spakman (2000)). The performed linearization is acceptable as long as the ray path is not significantly perturbed by the modifications made to the velocity field during the inversion step. Consider a perturbation $\delta v(\bar{x})$ to a reference velocity field $v_0(\bar{x})$, so that $v(\bar{x}) = v_0(\bar{x}) + \delta v(\bar{x})$. Due to this deviation from the original velocity field $v_0(\bar{x})$, both the ray path and the travel time will also be perturbed. The new ray path $L(v) = L_0 + \delta L$ where L_0 is the ray path in $v_0(\bar{x})$ and $t = t_0 + \delta t$ where t_0 is the travel time along L_0, the ray path calculated for the reference velocity field $v_0(\bar{x})$. The travel time in $v(\bar{x})$ can then be written as:

$$t = \int_{L_0+\delta L} \frac{1}{v_0(\bar{x}) + \delta v(\bar{x})} \mathrm{d}l \tag{3.2}$$

Without showing a complete derivation, we adopt Fermat's Principle to prove the validity of the linearization. Fermat's Principle states that between two points the travel time of a ray is stationary with respect to perturbations in the path ($\partial t/\partial L = 0$). As a result, for a velocity perturbation along the original ray path the corresponding travel time perturbation is to first order given by:

$$\delta t = -\int_{L_0} \frac{\delta v}{v_0^2} \mathrm{d}l + O(\delta v^2) \tag{3.3}$$

Finally, if we substitute the velocity field in Equation (3.1) by the corresponding slowness field

$$s(\bar{x}) = \frac{1}{v(\bar{x})} \tag{3.4}$$

then the expression for a travel time perturbation relative to the reference model is given by:

$$\delta t = \int_{L_0} \delta s \mathrm{d}l + O(\delta s^2) \tag{3.5}$$

The travel time perturbation is now linearly dependent on a perturbation in slowness.

The forward problem of finding source-receiver paths for a given reference model can be dealt with by ray tracing techniques such as shooting and bending methods, first-arrival wavefront tracking on a grid (e.g. eikonal solvers) and network methods, known as Shortest Path Raytracing (SPR). For an explanation of these methods, see for example Rawlinson and Sambridge (2003). For the representation of the slowness perturbations, two main classes of model parameterization exist in global seismic tomography: superposition of global basis functions, such as spherical harmonic functions (e.g. Woodhouse and Dziewonski (1984)) and local discrete cell functions (e.g. van der Hilst et al. (1997)). Here we focus on the latter by parameterizing the model vector as non-overlapping constant slowness cells and parameters associated with hypocenter mislocation (Spakman and Nolet, 1988). For this type of parameterization, the linearized inverse problem without regularization can be expressed in matrix notation as the system of equations

$$Gm = d + \epsilon \tag{3.6}$$

where G_{ij} represents the distance the ith ray travels in cell j and the partial derivatives for the simultaneous source relocation, m_j the magnitude of the slowness perturbations from the reference model in cell j and parameters associated with the source relocation (origin time, longitude, latitude, depth), $(d_i + \epsilon_i)$ the travel time delay for ray i and ϵ the noise present in the data. The least-squares solution is derived from the normal equations

$$G^T G m = G^T d \tag{3.7}$$

where d now is the vector representing both travel time residuals and data noise. Solving the inverse problem for m in (3.7) gives the solution

$$m = Ld \tag{3.8}$$

with the inverse operator

$$L = (G^T G)^{-1} G^T \tag{3.9}$$

However, solving for m in (3.7) is practically never possible in travel time tomography (and seismic tomography in general) without the introduction of some form of regularization. Due to an uneven distribution of sources and receivers and thus ray paths, the model parameters are unevenly sampled. Furthermore, due to errors in the data the system of equations is inconsistent. Consequently, no unique solution exists that fits the data exactly and the problem is ill-posed. To resolve this issue, regularization is imposed. As a result, solving the inverse problem in travel time tomography can be formulated as minimizing some cost function containing a data residual term and one or more regularization terms. In general, a trade-off exists between obtaining the best data fit and the minimum-norm solution and we cannot minimize both simultaneously. For example, in a standard least-squares problem often the 'optimal' solution chosen is the solution that both provides a good fit to the data, i.e. has a low data residual, and whose Euclidean norm is not too large. While regularization is needed to deal with the ill-posedness and ill-conditioned characteristics of the inverse problem, it introduces a large ambiguity and hence subjectivity in the choice of the 'optimal' solution. Therefore, solution robustness and uncertainties should always be analyzed carefully to deduce the quality and value of the obtained solution.

In this study, we follow the inversion procedure as described in (Li et al., 2008), who use an iterative least squares method (LSQR) (Paige and Saunders, 1982; Nolet, 1985) to minimize the following cost function

$$\xi = \parallel Gm - d \parallel^2 + \lambda_1 \parallel Sm \parallel^2 + \lambda_2 \parallel m \parallel^2 + \lambda_3 \parallel C - m_C \parallel^2 \tag{3.10}$$

Here, G is the sensitivity matrix, m is the solution vector, d is the data vector, L is a smoothing operator, C is the a priori crustal model, m_C the crustal part of the model space m and λ_i control the weights of the three regularization terms relative to the first term on the right-hand side (r.h.s.) in (3.10). The second term on r.h.s. represents the gradient damping that smoothes the model both laterally and radially, while norm damping (third term on r.h.s.) aims to find the best model with small variations from the reference model. The solution m consists of $\sim$ 650,000 parameters and includes both P wave speed perturbations (in non-overlapping and constant-slowness blocks) relative to *ak135* (Kennett et al., 1995) and parameters associated with hypocenter mislocation (origin time, longitude, latitude and depth) (Kárason and van der Hilst, 2001). Ray tracing is carried out by using *ak135.* The part of the sensitivity matrix A associated with short-period data consists of the total length of the rays traversing the constant-slowness cells, whereas the part of A representing the low-frequency data is obtained through projection of 3-D kernels onto such cells. To reduce the size of the sensitivity matrix and better balance the sampling of cells, weighted composite rays are used for the EHB data (Spakman and

Nolet, 1988). (Li et al., 2008) construct an adaptive parameterization scheme on the basis of sampling density, i.e. hit count, of the short-period data used, thereby developing a grid with an irregular cell size. The cells in the regular base grid have a size of approximately 0.7°in latitude and longitude and a depth of 45 km throughout the mantle. With increasing depth, the minimum cell size increases in accordance with the increasing width of the Fresnel zones of short-period P waves. Due to the 3D crust correction incorporated by (Li et al., 2008), the solution to the inversion of synthetic data always contains a bias to the a priori defined crustal model. For our purpose, this bias is unwanted and we therefore remove the crustal correction from the inverse problem. Removing this crust correction does not influence the stability of the inversion. Minor differences exist in some locations between the solution obtained with and without the 3D crust correction, but we are confident that excluding the crust correction does not significantly influences our results and interpretations made thereon. For further details regarding the adaptive parameterization, crust correction or 3-D sensitivity kernels, we again refer the reader to (Li et al., 2008). All solutions presented in the current report were obtained after 100 iterations, after which most of the solution has converged.

3.2 Null space shuttle

Consider a solution m_t that might violate the data. The part of m_t that lies in the null space of G can be found using the null space shuttle technique as defined by Deal and Nolet (1996). The solution m_t can be seen as the sum of the components lying in the range and in the null space of G, so that

$$m_t = m_t^{range} + m_t^{null} \tag{3.11}$$

Defining

$$h = Gm_t \tag{3.12}$$

gives

$$h = Gm_t = Gm_t^{range} + Gm_t^{null} \tag{3.13}$$

Since $Gm_t^{null} = \bar{0}$ by definition, solving the inverse problem (3.13) would provide one with m_t^{range} and it would be trivial to obtain m_t^{null} via (3.11). We employ the same method for the evaluation of m_t^{range} as is used to solve the original inverse problem (LSQR). However, we want to emphasize that we would only obtain m_t^{range} exactly if the algorithm used to solve the inverse problem yields the minimum-norm solution. A mininum-norm solution would have no components in the model null space and would fit the data exactly. Consequently, the associated null space component m_t^{null} derived from (3.11) would have no components in the range of the model space and would therefore not affect the data fit. However, the solution to our inverse problem is a compromise between the minimum-norm and least-squares solution. Therefore, we do not obtain m_t^{range} exactly, but instead we get an estimation:

$$\tilde{m}_t^{range} = Lh \tag{3.14}$$

with L the inverse operator corresponding to our original inverse problem and $\tilde{m}_t^{range} \neq m_t^{range}$. Since we do not obtain the minimum-norm solution, $\tilde{m}_t^{range}$ includes a small component from the model null space. Consequently, we only have an estimation of the null space component of m_t:

$$\tilde{m}_t^{null} = m_t - \tilde{m}_t^{range} \tag{3.15}$$

and $\tilde{m}_t^{null}$ contains elements lying in the range of G:

$$G\tilde{m}_t^{null} \neq \bar{0} \tag{3.16}$$

Since $G\tilde{m}_t^{null}$ is not exactly zero, the data fit is affected. However, as will be shown later, effects on the data fit are minor compared to data errors and utilization of the null space shuttle is valid. Once the null space component of the theoretical solution m_t is obtained, it is straightforward to calculate a new conservative solution

$$m_{new} = \tilde{m} + \tilde{m}_t^{null} \tag{3.17}$$

with $\tilde{m}$ the original solution. The term conservative refers to the fact that the new solution has a data fit comparable to that of the original solution. The null space shuttle provides us with a powerful tool to investigate the robustness of a tomographic image in a straightforward fashion.

4 Sensitivity tests

The quality assessment of tomographic models is as important as the imaging of Earth's interior itself. Despite this, while the detail in tomographic images has greatly improved over the last decades, the former has been given relatively little attention over the years and existing methods for quality analysis are far from perfect. The reliability of tomographic images is often assessed by executing sensitivity tests that use the ability to recover a known input model as a measure for image robustness. To this end, the so-called resolution operator is often employed in seismic tomography. Consider again (3.6), where the model m represents true Earth: $m = m_{true}$. By inserting (3.6) in (3.8) we obtain

$$\tilde{m} = LGm_{true} = Rm_{true} \tag{4.1}$$

with R the linear resolution operator relating the estimated solution to the true solution. When R = I, according to (4.1) the estimated solution is equal to the true solution and the true solution is said to be perfectly resolved. If R $\neq$ I, the estimated model parameters can be interpreted as weighted averages of the true model parameters. Instead of $m = m_{true}$, any theoretical model can be inserted in (4.1). However, this type of assessment is rather qualitative and can not be used to obtain a quantitative interpretation of model uncertainties and thus model quality. In a sensitivity test, a theoretical model is used to generate synthetic data that is subsequently inverted for to derive a tomographic image (the output of the test). Various schemes have been developed to investigate the reliability or 'resolution' of global seismic tomographic images. Many studies of seismic tomography use the term resolution to qualify the outcome of synthetic tests and to judge the quality of their tomographic inversions (e.g. Spakman and Nolet (1988); Spakman et al. (1989); Bijwaard et al. (1998); Li et al. (2008)). By this definition, resolution does not imply anything on the resolved spatial length scales in the tomographic images. The most common types of synthetic tests include so-called checkerboard tests (e.g. Spakman and Nolet (1988)), which use harmonic input patterns, and hypothesis tests (e.g. van der Hilst et al. (1997)), for which input models of velocity anomalies (e.g. subducting slabs) are constructed. We will discuss both types and interpretations made based on the outcome of these tests.

4.1 Checkerboard tests

The shortcomings of checkerboard tests to assess quality or resolution of the tomographic method were shown by (Lévêque et al., 1993). The main shortcoming of the method lies in the fact that the measure of resolution depends on the input model used for the checkerboard test. In mathematical terms, for an input model that corresponds to eigenvalues of the resolution operator close to 1, the outcome of a sensitivity test would show a very good recovery of the input model. The opposite would occur for input models containing components in the subspace associated with eigenvalues close to 0 and in both cases a wrong conclusion on the resolving power or resolution of the inverse problem could be drawn. While for a particular linear tomographic inversion the resolution operator R in (4.1) is uniquely defined, inferences on R and model quality vary for checkerboards differing in dominant wavelength of the input pattern. Obviously, this is an undesirable aspect for any quality analysis. However, despite these notions, the outcome of these checkerboards tests is still used in many studies as being indicative of resolution or resolving power of the performed seismic tomography. To aid our discussion of the matter, we here perform a checkerboard test similar to the test done by Li et al. (2008). The outcome of a checkerboard test shows the part of the input model that is recovered by the inversion of synthetic data (in accordance with (4.1)). This implicitly means, following (3.11), that the part of the input model that is not recovered lies in the null space of G:

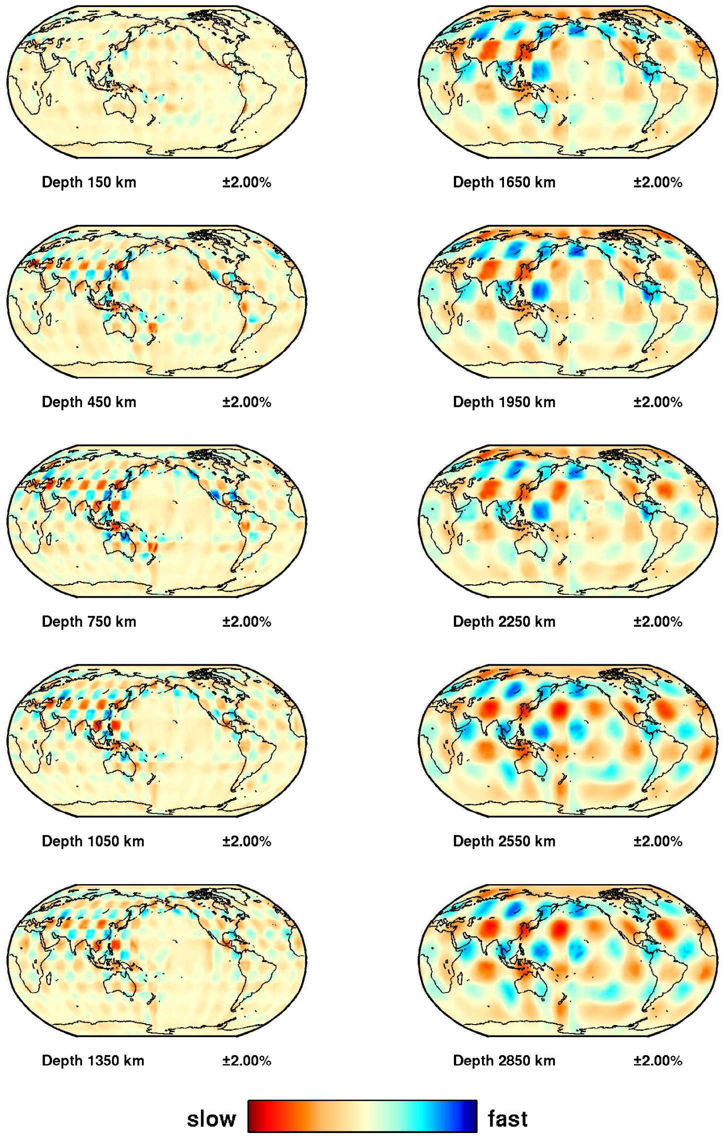
$$m_{in} = m_{in}^{range} + m_{in}^{null} \tag{4.2}$$

To illustrate this we plot both terms on the r.h.s. of (4.2) in Figures 1a and 1b, respectively. We use harmonic noiseless input patterns with a fixed amplitude of 2 % that were put in for one depth at a time. Similar to Figure 5 in Li et al. (2008), we take an input pattern with a half wavelength of $\sim 5°$ in the upper half of the mantle and a half wavelength of $\sim 10°$ in the lower half. The synthetic travel times were inverted following the same method as used for the real data in this study. The colour scale in Figures 1a and 1b corresponds to the amplitude of the input pattern. It is apparent that amplitudes of the null space component are much larger than the recovered component. The amplitude recovery is variable but on the order of $\sim$30% or less, which implicitly leads to an unrecovered component of 70% or higher in terms of amplitude. Of course, this can not be an argument to state that resolution is poor, for the same reason that interpreting structures in the output as well-resolved is too simple. By including these figures here we simply aim to illustrate the large model null space of this particular inverse problem. Most studies using these checkerboard tests (e.g. Bijwaard et al. (1998); van der Hilst et al. (1997); Li et al. (2008) acknowledge that the loss of amplitude is partly due to the regularization used in the inversion and that therefore the amplitudes are underestimated. However, it might be unwise to take this underestimation of amplitudes in sensitivity tests for granted. A too strong regularization may dominate the solution instead of observed data. As a consequence, the robustness of both amplitudes and geometry of imaged structures might not be guaranteed. In our view, loosing such a large portion of the input pattern can not be simply interpreted as 'amplitude underestimation' and should warn one that resolving power may be low and that parts of the solution may be non-robust.

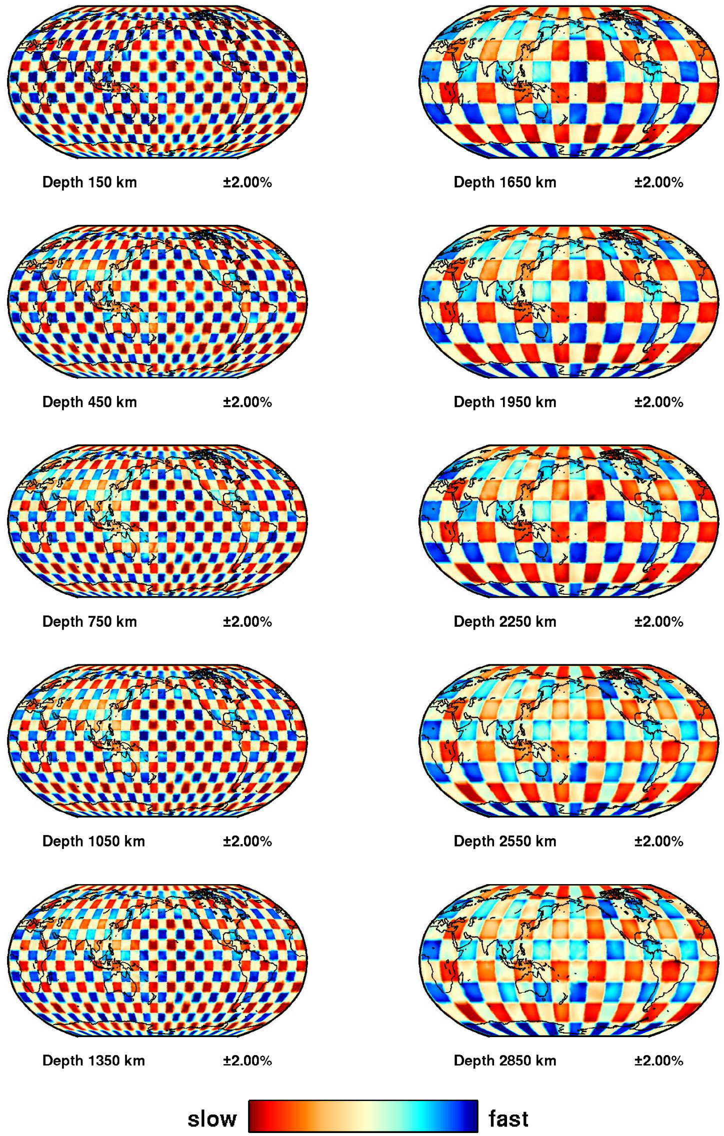
As an alternative to checkerboard tests, (Spakman and Nolet, 1988) introduced so-called 'spike tests', where the synthetic model contains a 3D network of spatially well-separated spikes, i.e. cells with a alternately negative and positive anomaly spaced a few cells apart in every direction. This test aims to investigate the smearing of the spikes by the inversion. As long as the smeared features in the solution do not overlap, the test should provide one with independent estimates of the effects of regularization (Nolet, 2008). However, the regularity of the pattern introduces a dominant wavelength into the input model and the same problems as for the checkerboard tests could arise when making interpretations of resolving power.

4.2 Hypothesis tests

Furthermore, we want to discuss the use of so-called hypothesis tests which focus on inverting synthetic anomalies similar in shape and amplitude to those imaged by the inversion of real data (e.g. van der Hilst et al. (1997); Bijwaard et al. (1998)). In these tests, (part of) a solution to an inversion of real data $\tilde{m}$ is used to create synthetic data and compute a new solution $\tilde{m}_{new}$. The goal of these tests is to analyze whether the tomography would be able to image certain structures for a given model parameterization. The well-resolved structures in the output solution are commonly thought of as indications that the structures imaged in the original tomogram are required by the data and are not an artifact of processes such as smearing or leaking (Bijwaard et al., 1998; van der Hilst et al., 1997; Li et al., 2008). However, inverting for synthetic data based on $\tilde{m}$ will naturally produce a solution closely resembling $\tilde{m}$, since the major part of this solution already lies in the range of G. The differences between test input and output result from the fact that the solution $\tilde{m}$ is not a minimum-norm solution and contains components of the model null space. Although these hypothesis tests can give some general information on the seismic tomography performed, we again emphasize that the danger lies in the type of interpretations made based on test results and that one should be careful when conclusions are drawn on model quality.



(a) Recovered components.



(b) Null space components.

Figure 1: Recovery fields (left figure) and corresponding null space components (right figure) of a global checkerboard test (similar to Figure 5 in (Li et al., 2008). We use harmonic input patterns with constant amplitude $\pm 2\%$ in wave speed throught the mantle. (Left column) Half wavelength of $\sim 5^\circ$ (spatial wavelength of $\sim$550 km at the surface). (Right column) Half wavelength of $\sim 10^\circ$ ($\sim$600 km at the CMB).

5 Results

We compute a set of conservative solutions of the form (3.17) by designing several theoretical models m_t based on the original solution (Figure 2). Conducting a full statistical analysis by computing thousands of new conservative solutions would give more complete estimates on uncertainties, but is time-consuming. Instead, we are confident that a relative small set of conservative solutions can give sufficient information on the robustness of velocity perturbations in the original model. We interpret an anomaly as robust when the anomaly is unequivocally required by the data. If this is the case, the majority of the components of the anomaly lie in the range of G and the anomaly retains its geometry and amplitude after application of the null space shuttle. To estimate uncertainties, we first have to fix ideas on the concept of uncertainty in the tomographic image. It is clear that introducing one model parameter with a very large velocity perturbation would not affect the global data fit measure. Does this mean that the uncertainty for this particular model parameter is very large? Not necessarily. Meaningful estimates of model parameters can only be obtained over a volume large enough to affect the data. It is simply a disadvantage of global travel time tomography that we are not able to recognize local data misfits, since ray paths may traverse many model cells. However, it does mean that one has to be extremely careful when influencing the data fit as long as one does not have regional information on the fit to the data. To avoid this problem, when applying the null space shuttle we only use global theoretical models containing many non-zero model parameters. This should ensure that unwarranted modifications to the original solution are recognizable in the global data fit. We work with the root mean square (RMS) measure, which is defined as:

$$x_{rms} = \sqrt{\frac{\sum_{k=1}^{N} x_k^2}{N}} \tag{5.1}$$

with N the number of delay time data in the data set. The RMS of the whole data set prior to inversion is 1.92 s, while the RMS data residual is 1.46 s for the original tomographic image shown in Figure 2, corresponding to a variance reduction of 42 %.

5.1 Uncertainties in solution norm

One of the most well known issues in travel time tomography is the classical trade-off between the fit of the data, represented by the forward problem RMS residual (Gm - d), and the size of the model, i.e. the norm of the solution. Due to this trade-off it ultimately takes the subjective choice of a researcher to produce a certain final model. We use the same two quantities as a basis to establish some quantification of the uncertainties present in the original tomographic solution. To this end, we expand (3.17) with a

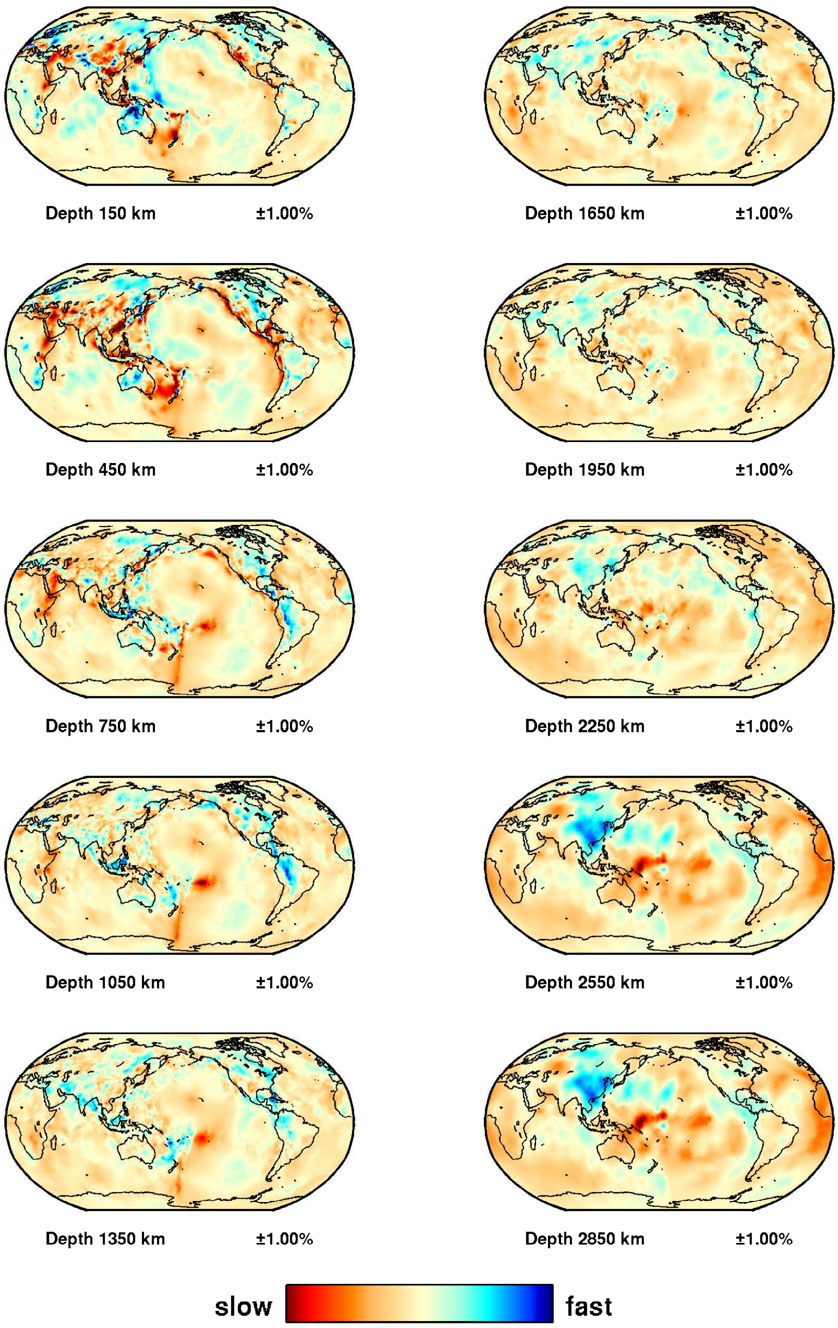


Figure 2: The original P wave speed perturbation model PMIT08 without crust correction. The colour scale ranges from -1 to +1 %.

simple multiplication factor α:

$$m_c = \tilde{m} + \alpha \tilde{m}_t^{null} \tag{5.2}$$

First, we analyze the range of acceptable solutions by taking the theoretical model m_t in (5.2) equal to the original model $\tilde{m}$. We computed a set of conservative solutions by letting α range from -10 to 10 with a stepsize of 0.1. Figure 3 shows the resulting RMS data residual and RMS solution norm (the total of slowness cells and relocation parameters) for the obtained solution set. The red asterik denotes the original solution $\tilde{m}$ as shown in Figure 2. It is clear that we would be able to either minimize the solution norm or to improve the data fit. The original solution is minimized in RMS norm by 24 % for α = -1.7. The reduction of solution norm occurs in both velocity perturbation and hypocenter relocation parameters. We find that modifying these mislocation parameters with the null space shuttle has a relatively small influence on the RMS data fit compared to the effect of changing the slowness parameters (not shown here). Furthermore, a large range of solutions seems to exist that produces an acceptable data fit in view of the errors present in the data. The term acceptable is subjective in this sense, and we therefore leave it to the reader to decide what change in the data fit is acceptable. For the moment, let us choose the lowermost horizontal line representing a 0.1 s change in RMS data residual in Figure 3, in our view a conservative choice given the errors in the data mentioned earlier. For α = -1.4 and α = 6.6, we find two solutions that give a RMS data misfit 0.1 s larger than the original data misfit, while the solution providing the best fit to the data is found for α = 2.6 (Figure 3).

We translate the range of possible solutions (for the 0.1 s change in data residual) to a plot of the corresponding RMS velocity perturbations (so excluding relocation parameters) versus depth (Figure 4). For α = -1.4, we observe that the solution norm is reduced over the whole depth extent of the mantle. The solution associated with α = 6.6 is on average about three times as large as the original solution $\tilde{m}$ in terms of RMS norm, while the solution with the best data fit found for this exercise is roughly 1.5 to 2 times as large. It is clear that the amplitudes of the velocity perturbations are not well constrained. For the range of possible solutions norms presented in Figure 4, the maximum difference in average P wave velocity perturbation is about 0.4 % in the lower mantle and approximately 1 % in the upper mantle around 350 km depth. Large differences thus exist in average global amplitudes of velocity perturbations. It is intelligible to compare this range of possible average amplitudes to the amplitudes of the structures in the original solution. For every solution parameter we compute the average for the solutions for α = -1.4 and α = 6.6. We define an uncertainty range centred around these model parameter averages and the uncertainty in a model parameter is defined as the absolute difference between the average and either extreme of the range, i.e. the new solutions for α = -1.4 and α = 6.6. We show the original solution for the South American region in Figure 5. The uncertainty as defined here for this region is presented in Figure 6. It is clear that uncertainties in amplitudes are of similar size as the amplitudes of anomalies in the original solution, particularly in the upper mantle. However, this does not mean that these anomalies are not required by the data. We will investigate the robustness of individual anomalies imaged in the original model in the next section. Whether or not anomalies are robust is interesting from a physical point of view, since it is the presence of these anomalies in the tomographic image that spark hypotheses on physical processes such as downwelling of slabs, upwelling mantle plumes and ultimately

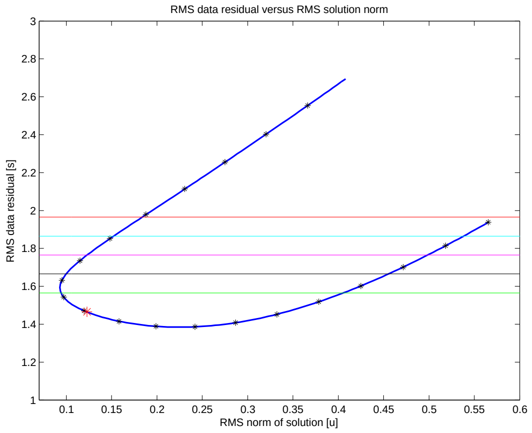


Figure 3: RMS data residual versus RMS norm of the tomographic solution. Red asterisk denotes the original solution $\tilde{m}$. Black asterisks represent the multiplication factor α from -10 to 10 with an interval of 1 as a reference. Coloured lines represent changes in the data residual from 0.1 to 0.5 s (upwards) with 0.1 s intervals.

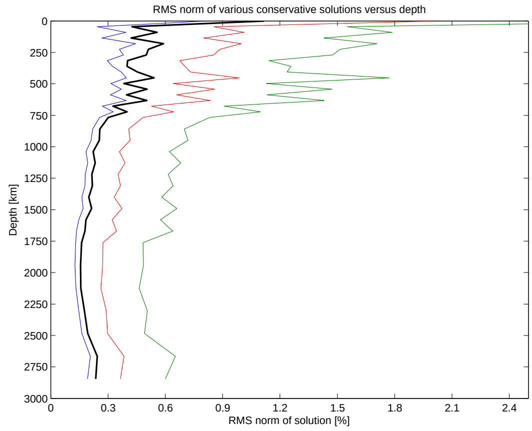


Figure 4: RMS norms of various solution versus depth. Bold black line represents the original solution $\tilde{m}$. Blue and green line represent solutions for a 0.1 s data residual change for α equal to -1.4 and 6.6, respectively. Red line shows the solution with the best data fit (α = 2.6).

mantle convection.

5.2 Implications for PMIT08

We investigate the difference between the four solutions presented in Figure 4 by visualizing tomograms for the South American region (cf. Figure 9 in (Li et al., 2008)). The result is shown in Figures 5 and 7 to 9. Structures recognizable in the original solution in Figure 5 are slab-like high velocity anomalies that extent into the lower mantle and low velocity anomalies present in the upper mantle around the slab-like structures. Figure 7 shows a reduction of solution norm and hence anomalies by using the null space shuttle and α = -1.4 in (5.2). The effect on the visible structures is apparent. The high-velocity structure in slice 10 has vanished, while structure in slices 7 and 8 are largely diminished, in accord with the notion by (Li et al., 2008) that resolving power in the South American region becomes less towards the south. Other conspicuous changes with respect to the original solution are the reduction of upper mantle low velocity anomalies (e.g. slice 3) and the mostly absent upper mantle portion of slabs (slices 5 and 9), which creates the impression that a large gap exists between the lithospheric and lower mantle part of the slab-like structures. Naturally, for α positive and $m_t = \tilde{m}$ the structures in the new conservative solutions (Figures 8 and 9) resemble the structures in the original solution closely, apart from the obvious increase in amplitude. Note that the colour scale in Figures 8 and 9 has been modified to prevent complete saturation of the tomographic slices and to show the size of the amplitudes in these models. It is interesting to see the range of possible amplitudes of anomalies such as the slab-like structures. For instance, one could question the relevance of the amplitude of $\sim$ 0.8 % for the slab in slice 9 in Figure 5 when other acceptable solutions show amplitudes ranging from 0-0.3 % (Figure 7) to approximately 3 % (Figure 9), amounting to a broad range of possible velocity perturbation of 3 % for this particular slab. Recall that this amplitude range will increase further for solutions deviating more than 0.1 s from the original RMS data residual (Figure 3).

As a second example we analyze the African region, underneath which a low velocity anomaly is commonly observed in tomographic studies (e.g. Ritsema et al. (1999); Grand (2002); Li et al. (2008). The low velocity anomaly is believed to vertically extent from above the CMB to the upper mantle, with the deeper portions shifted southwards relative to the shallower low velocity structure (Figure 12. The low velocity underlying Eastern Africa in the upper mantle is pronounced in this solution. Figure 13 again shows the new solution for α = -1.4. While the upper mantle component of the low velocity anomaly is still clearly visible, the lower mantle portion has been partly removed, causing it to barely stand out from the background.

We can explore the model null space further by choosing a more specific theoretical model, i.e. a model not equal to the original solution. We use theoretical models based on the upper or lower mantle part of the original solution. We feel that this provides us with enough non-zero model parameters to use the global misfit criterion to verify the new solutions. We take the original solution as a basis and set the model parameters representing the lower mantle (>660 km) to zero, thus leaving the original upper mantle parameters and the relocation parameters untouched. The null space shuttle in (5.2) is applied for the same range of values for α as earlier. Here we focus on the solutions obtained for negative α values for which the structures in the upper mantle are minimized, i.e. for which the structures in the solution are biased towards the lower mantle. For α = -1.8, the RMS norm of the new solution is minimized. The corresponding new solution $\tilde{m}_u$ has a RMS data residual of 1.568 s, a difference of $\sim$ 0.1 s with the data misfit of the original solution. Figure 10 shows the solution for the South American region. A strong reduction of both the size and amplitudes of the anomalies in the upper mantle is visible. High-velocity anomalies have been partly or completely removed. Some lower mantle structures just below the 660 km discontinuity have slightly increased in amplitude (e.g. slice 5), an example of the bias towards the lower mantle imposed by our choice of theoretical model and α. Figures 17 and 18 show $\tilde{m}$ and $\tilde{m_u}$ for a region centred on Indonesia, respectively. The main structures in the original solution are the slab-like anomalies below the Sunda arc (slices 19, 20), the Celebes Sea (21) and the Philippine trench (22). Similar differences are observed, the most apparent feature being the disappearance of slab-like structures in the upper mantle (e.g. slices 19, 20, 22). By comparing $\tilde{m}_u$ to $\tilde{m}$, variable resolving power for the upper mantle is observed within the Mediterranean region (Figures 15 and 16). Slices 1-3 are oriented perpendicular to the Hellenic arc south of Greece, while slice 4 transects the Calabria arc located at the southern tip of Italy. While most anomalies below the Hellenic arc reappear in the new solution $\tilde{m}_u$, a large part of the original upper mantle structure below the Calabria arc is absent, including the low-velocity anomaly at the western edge of the profile and most of the slab-like structure.

Now let us reverse the previous exercise, i.e. set the upper mantle model parameters in the theoretical model to zero. As before, here we only analyze the solutions found for negative α, thus solutions with structure biased towards the upper mantle. For α = -1.9, the RMS norm is minimized and the data fit of the new solution $\tilde{m_l}$ is 1.519 s, an increase of $\sim$ 0.05 s compared to the original RMS data fit. Figure 11 shows the South American region for $\tilde{m_l}$. As opposed to $\tilde{m}_u$, most of the original high-velocity anomalies in the lower mantle are still present in the new solution, although amplitudes of the anomalies have decreased (e.g. slices 3-5). However, in some locations anomalies are removed to a large extent (e.g. slices 3, 7, 8). For the latter two, this might be expected given the aforementioned southwards decreasing resolving power. The high-velocity anomalies in the lower mantle also appear in $\tilde{m_l}$ below Indonesia (Figure 19). These examples of the new solution $\tilde{m_l}$ might indicate that high-velocity slab-like anomalies protruding into the lower mantle are required and therefore resolved by the data. The contrary is observed for the African low-velocity anomaly. Whereas in Figure 13 the low-velocity anomaly is still visible to a small extent, the anomaly is virtually completely absent in Figure 14, indicating that the llow velocity anomaly commonly observed in the lower mantle below Africa is not required by the data.

Figures 20b to 20d show $\tilde{m}$, $\tilde{m_u}$ and $\tilde{m_l}$ for the Hawaiian region, respectively. The solution in Figure 20c lacks the upper mantle low velocity anomaly, while the lower mantle component of this anomaly has stronger amplitudes than in the original solution. The solution in Figure 20d has a strong shallow anomaly, but the supposed continuation of the Hawaiian low velocity anomaly into deeper parts of the mantle is now completely missing. Apparently, the P wave data require a low velocity anomaly to be present, but a constraint on the depth of this anomaly seems to be absent for the larger part. An explanation lies in the insufficient ray path coverage in the region due to the lack of stations in the Pacific Ocean except for a few stations on the Hawaiian island chain (Nolet et al., 2007). Near vertical arriving rays below the Hawaiian stations induce vertical smearing and the imaged anomaly follows ray geometry (Figure 20b), resulting in a poor vertical resolution and little constraint on the vertical extent of the anomaly.

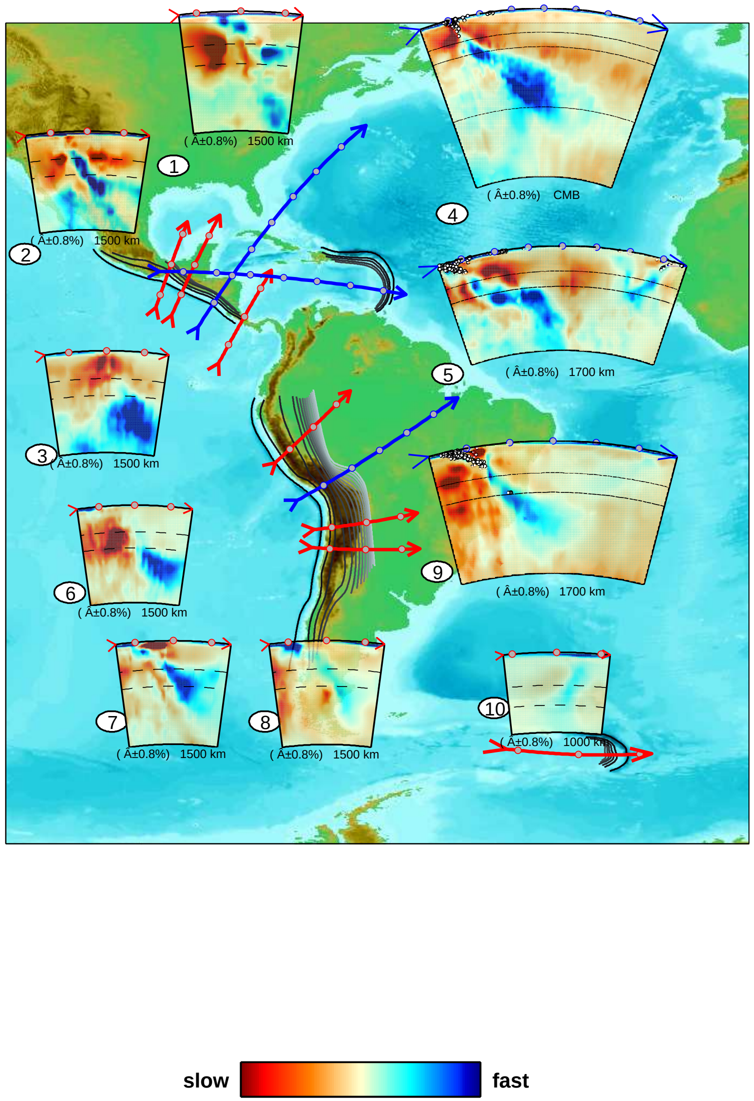


Figure 5: Original model $\tilde{m}$ for South American region. The colour scale ranges from -0.8 to +0.8 %.

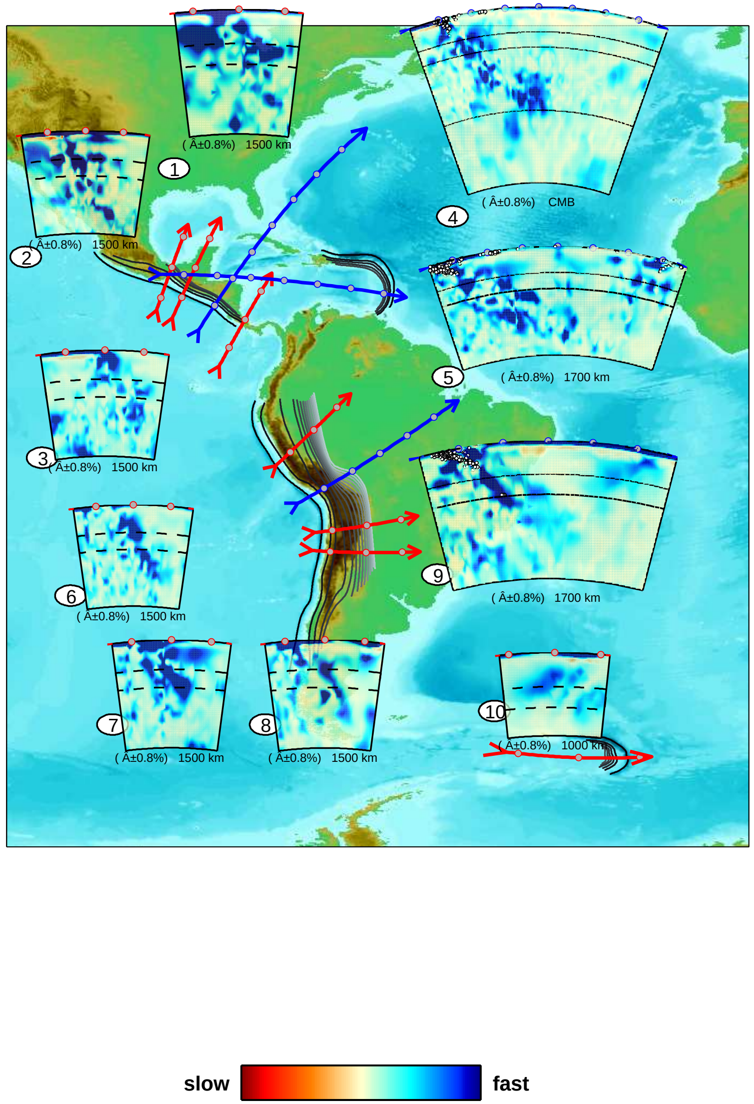


Figure 6: Representation of uncertainty as defined in text. The colour scale ranges from -0.8 to +0.8 %.

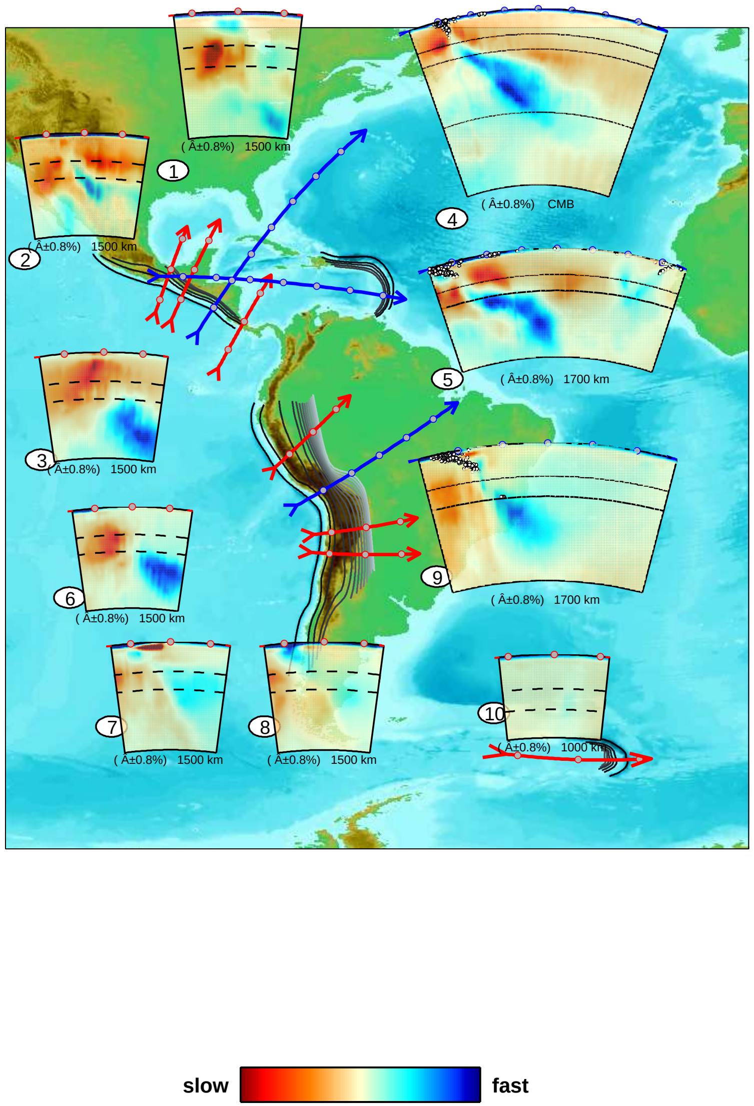


Figure 7: Conservative solution for α = -1.4 and m$_t$ = $\tilde{m}$, corresponding to blue curve in left frame of Figure 4. The colour scale ranges from -0.8 to +0.8 %.

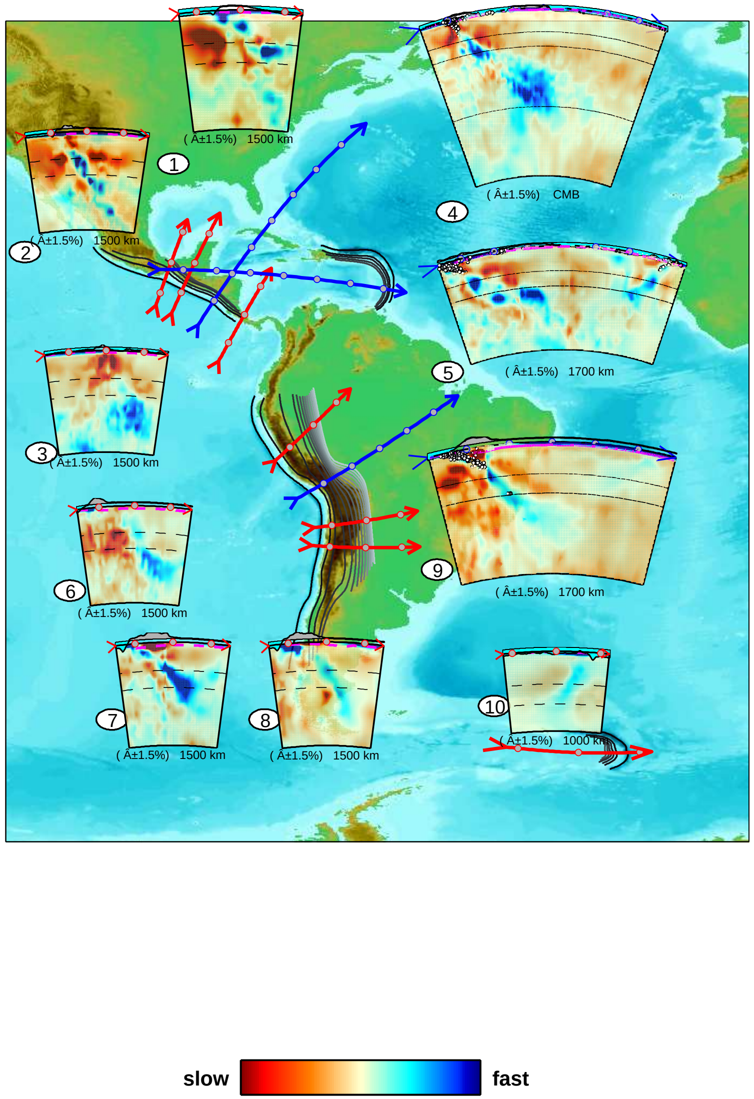


Figure 8: Conservative solution for $\alpha = +2.6$ and $m_t = \tilde{m}$, corresponding to red curve in left frame of Figure 4. The colour scale ranges from -1.5 to +1.5 %.

Figure 9: Conservative solution for $\alpha = +6.6$ and $m_t = \tilde{m}$, corresponding to green curve in left frame of Figure 4. The colour scale ranges from -3 to +3 %.

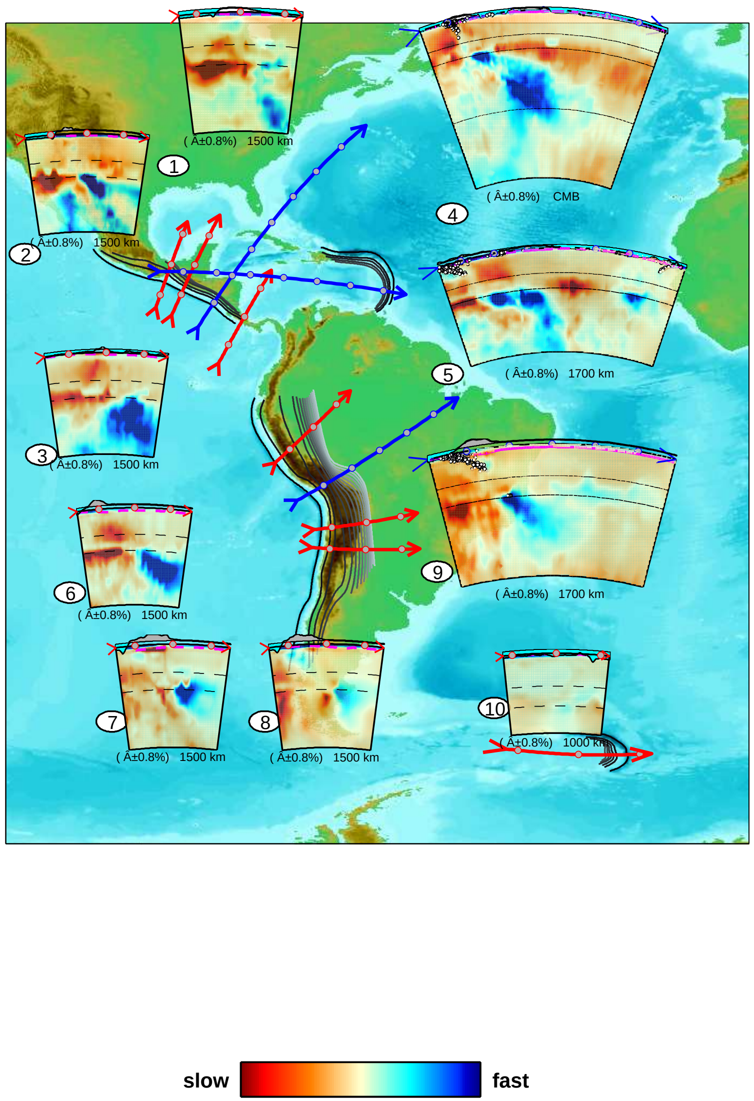


Figure 10: Conservative solution $\tilde{m}_u$ for α = -1.8. The colour scale ranges from -0.8 to +0.8 %.

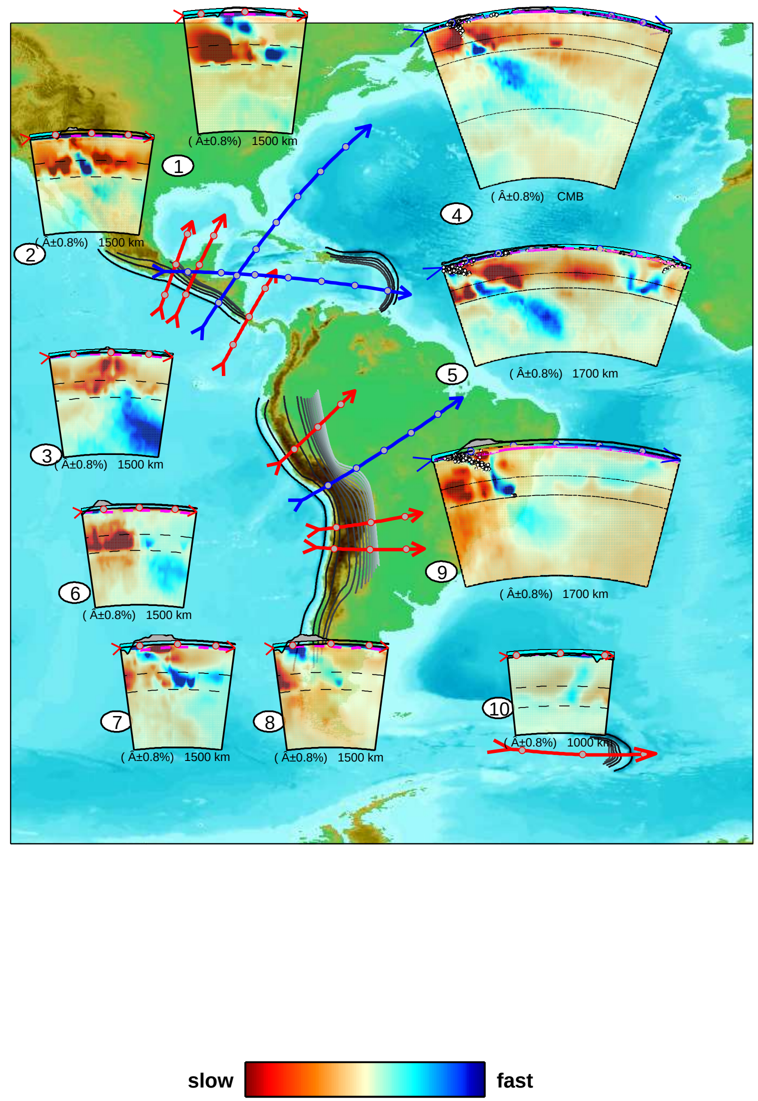


Figure 11: Conservative solution $\tilde{m}_l$ for α = -1.9. The colour scale ranges from -0.8 to +0.8 %.

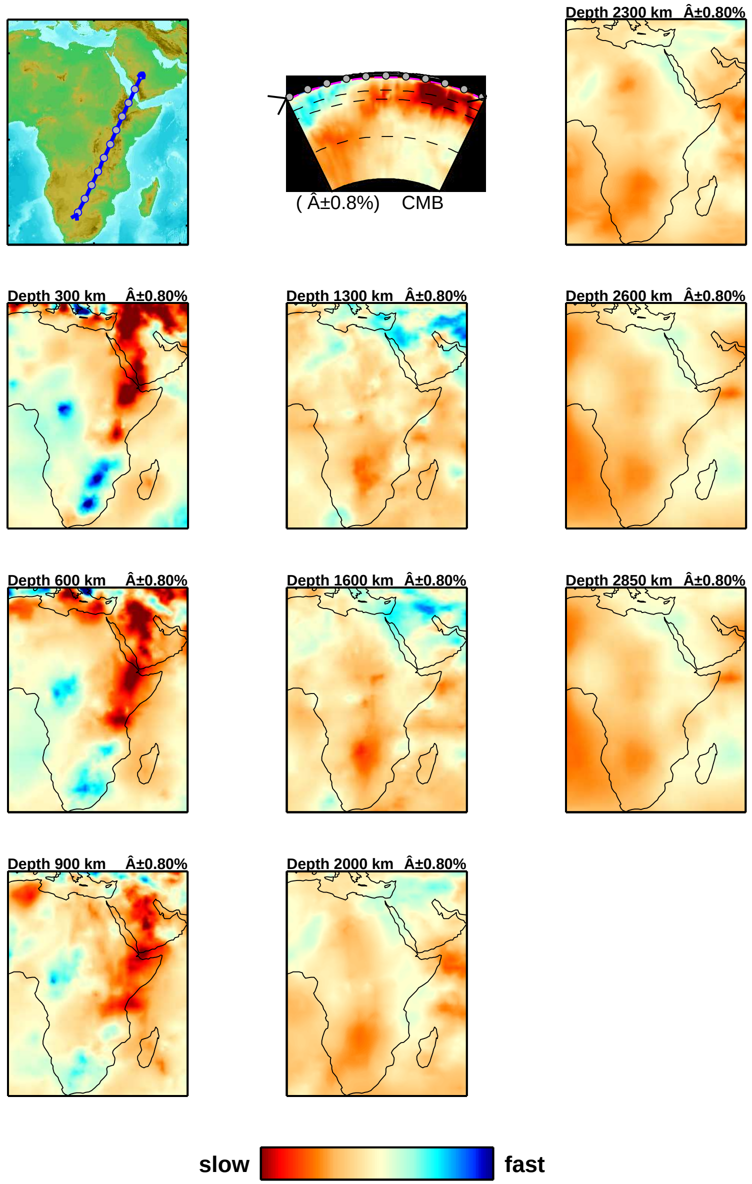


Figure 12: Original model $\tilde{m}$ for African region. The colour scale ranges from -0.8 to +0.8 %.

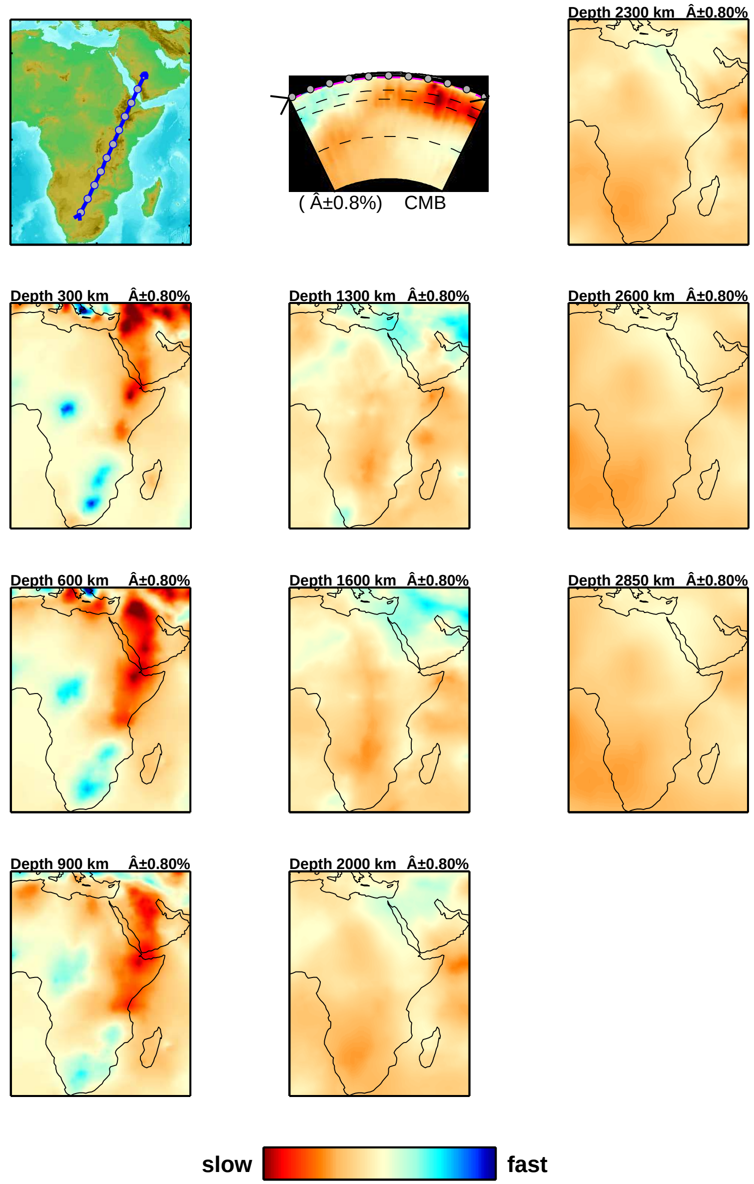


Figure 13: Conservative solution for α = -1.4 and $m_t = \tilde{m}$, corresponding to blue curve in left frame of Figure 4. The colour scale ranges from -0.8 to +0.8 %.

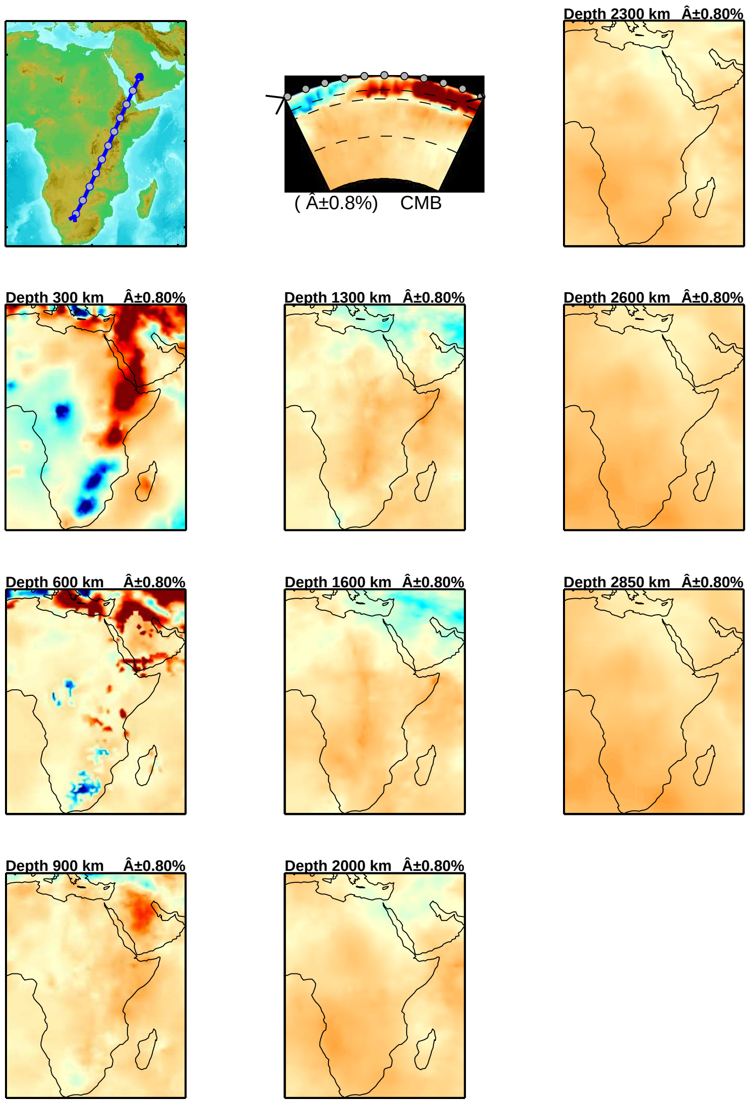


Figure 14: Conservative solution $\tilde{m}_l$ for α = -1.9. The colour scale ranges from -0.8 to +0.8 %.

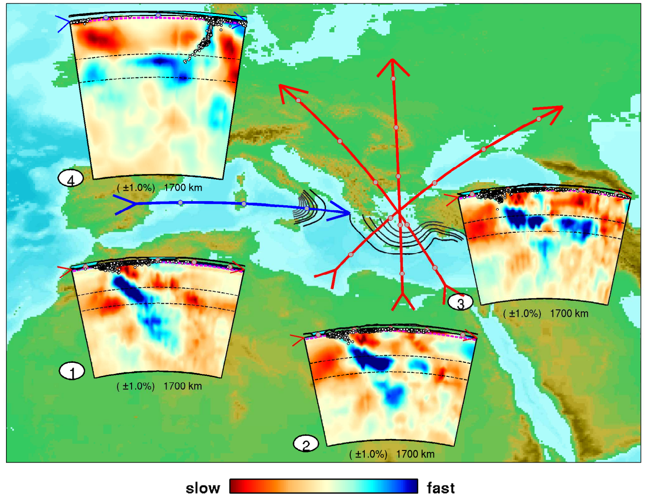


Figure 15: Original model $\tilde{m}$ for Mediterranean region. The colour scale ranges from -0.8 to +0.8 %.

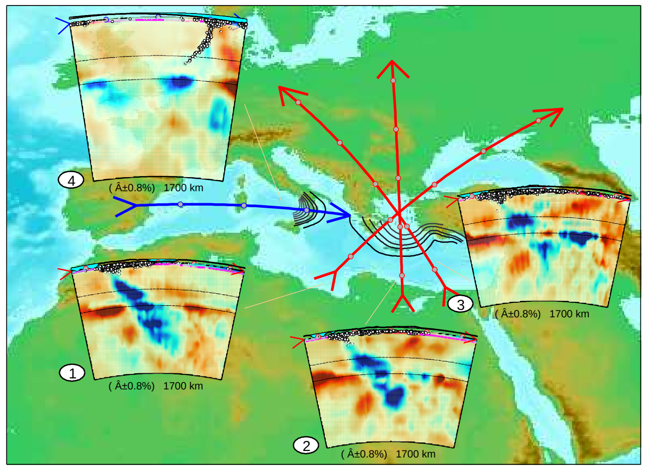


Figure 16: Conservative solution $\tilde{m}_u$ for α = -1.8. The colour scale ranges from -0.8 to +0.8 %.

5.3 Incorporation of regional data sets

The new solutions presented in the previous tomographic figures give an indication of the information in the model null space and whether anomalies observed are resolved by the data or not. It is apparent that in some regions the solution is significantly altered due to the addition of a model null space component. Most changes occur in the upper mantle, while slab-like anomalies in the lower mantle remain (partly) intact in most regions. However, the low-velocity anomaly underneath Africa is largely reduced in size and amplitude in the lower mantle. Clearly, resolving power of the data should ideally be higher for most regions. To achieve this, some global tomographic studies, such as (Li et al., 2008), incorporate seismic data from regional arrays in the tomographic inversion, as described earlier. These regional data should image more robust velocity perturbations compared to an inversion done without these data. We analyze the success of this incorporation of regional data by comparing various solutions for the North American and Chinese region.

5.3.1 China

We exclude the data from the Chinese Seismological Network (CSN) and the temporary array in Tibet. After inverting for the new data set, we aim to minimize the RMS norm of the resulting solution $\tilde{m_{chn}}$ by applying 5.2, where in this case both $\tilde{m}$ and $\tilde{m}_t$ are equal to $\tilde{m}_{chn}$. For α = -1.2 the conservative solution based on $\tilde{m}_{chn}$ is minimized in RMS norm. In Figures 22 and 23 we compare the solutions found with and without the Chinese regional data to their respective solutions with minimized RMS norms. A comparison between the original solution and the solution without the regional data in Figures 22a and 23a shows that in the latter solution less anomalies are imaged. For example, the high-velocity anomaly at locations 6 and 7 is not detected without the regional data. This could indicate that a higher resolution is obtained, i.e. that smaller wavelength structures can be detected better by incorporating the CSN and Tibet array data, which obviously is the desired result. We observe that anomalies in the original solution are only reduced in amplitude and not so much in size or position when minimizing the RMS norm (Figure 22). Possible physical interpretations based on these tomograms would not differ much between the two solutions and one might conclude that the regional data provide sufficient resolving power to keep the uncertainties in this part of the model adequately small. On the contrary, most low-velocity anomalies in $\tilde{m}_{chn}$ are absent after application of the null space shuttle (e.g. around location 3, Figure 23b). High-velocity anomalies at locations 1, 2 and 6 are not modified significantly by the null space shuttle and seem to be required by the global data. Overall, the incorporation of regional Chinese travel time data seems to lead to a higher resolution and slightly more robust imaging of velocity perturbations.

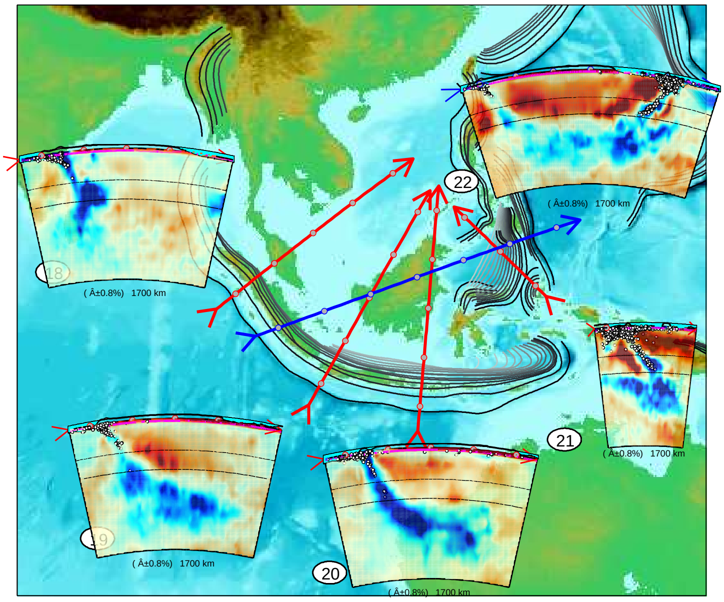


Figure 17: Original model $\tilde{m}$ for Indonesian region. The colour scale ranges from -0.8 to +0.8 %.

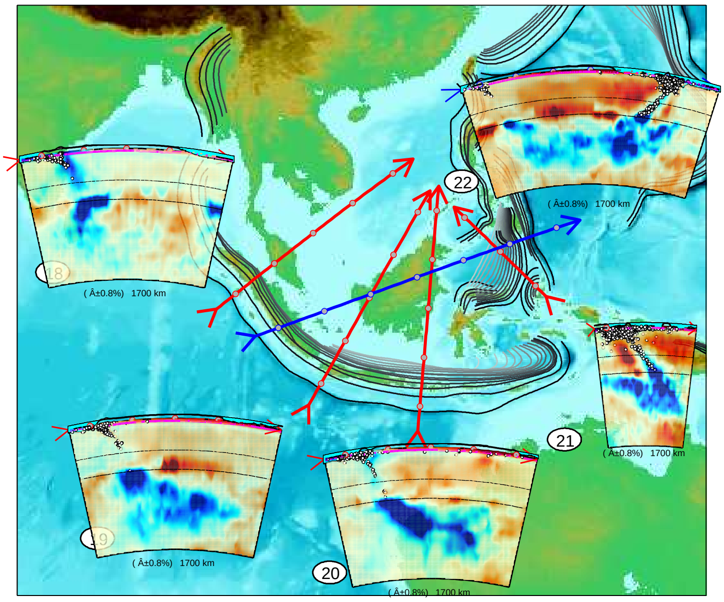


Figure 18: Conservative solution $\tilde{m}_u$ for α = -1.8. The colour scale ranges from -0.8 to +0.8 %.

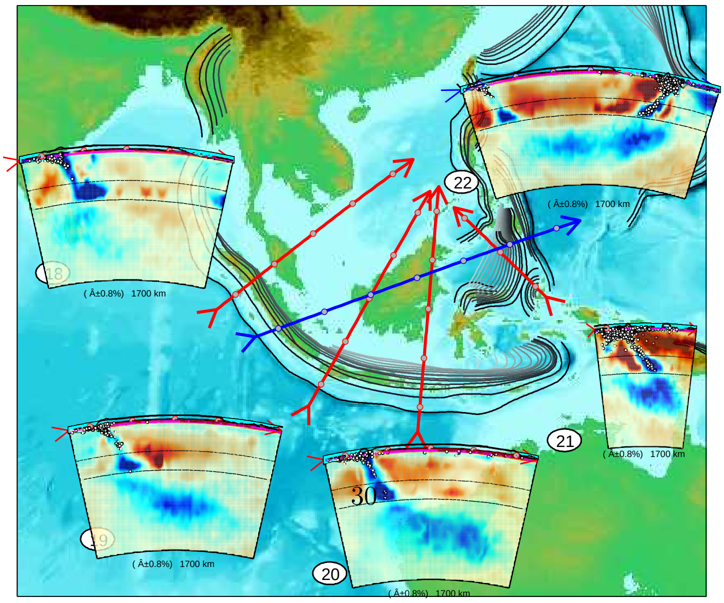
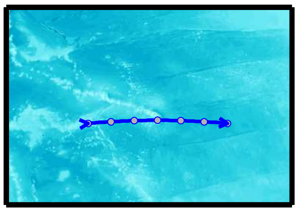
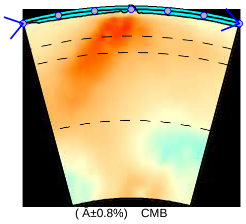
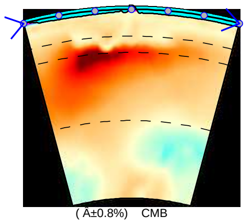


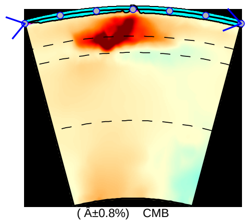
Figure 19: Conservative solution $\tilde{m}_l$ for α = -1.9. The colour scale ranges from -0.8 to +0.8 %.



(a) Map with location of slices for Hawaiian region.







(b) Original model $\tilde{m}$ for Hawaiian region.

(c) Conservative solution $\tilde{m_u}$ for α = -1.8.

(d) Conservative solution $\tilde{m_l}$ for α = -1.9.

Figure 20: Slices for the Hawaiian region for three different solutions. The colour scale ranges from -0.8 to +0.8 %.

5.3.2 North America

We repeat the above exercise by excluding the P wave data from the USArray. An equivalent comparison is made in Figures 25 and 26. The new solution excluding the regional USArray data is minimized in RMS norm for α = -1.6. As opposed to the Chinese example above, the difference between the solutions with and without the regional data is small (left column in Figures 25 and 26). In addition, no relevant difference in the robustness of anomalies is observed, as similar changes to the tomograms occur due to the null space shuttle (right column in Figures 25 and 26). The high-velocity anomalies in the first and third slices are partly removed from the image, while low-velocity anomalies are reduced in amplitude and shifted deeper into the mantle. The conspicuous low-velocity anomaly associated with the Yellowstone hotspot track in the bottom slice is almost completely eliminated from the solution. To adequately increase the resolving power, more regional data is needed than just the P wave data used in this study. (Li et al., 2008) admit that resolving power beneath North America is not as high as beneath east Asia, but that the incorporation of more data from the USArray will overcome this in the near future (Burdick et al., 2008, 2010).

6 Discussion

In this paper we utilize the null space shuttle to investigate uncertainties in the global P wave speed model for the Earth's mantle PMIT08 (Li et al., 2008). We acknowledge that the analyses performed in

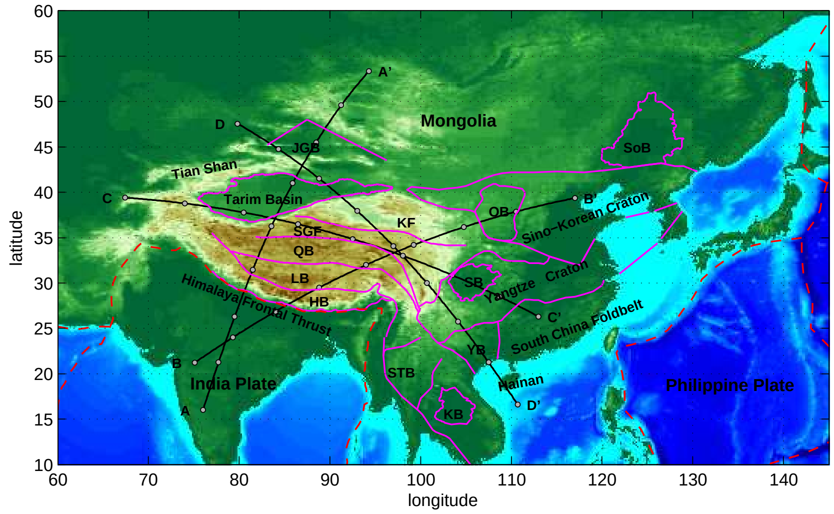


Figure 21: Map of China showing locations of the tomographic slices in Figures 22 and 23.

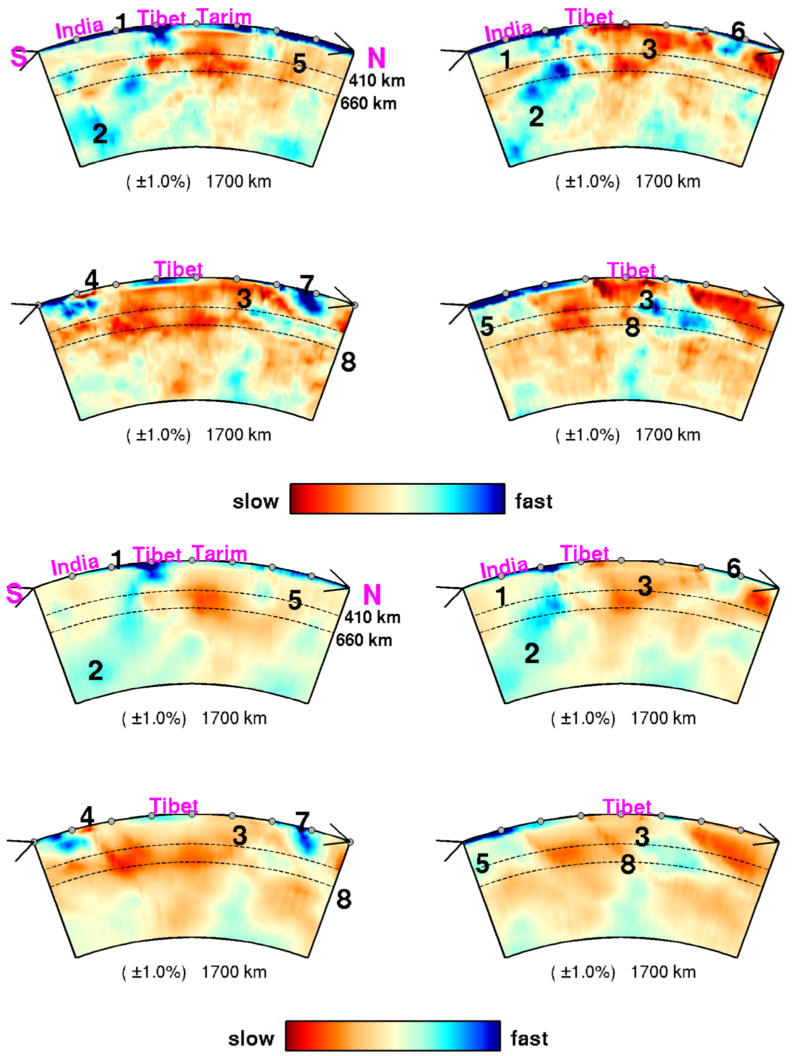


Figure 22: Original solution (top) and for the conservative solution for α = -1.7 and $m_t = \tilde{m}$ (bottom). Profiles AA' to DD' shown in Figure 21a are displayed in four slices with the topleft slice corresponding to profile AA', topright to BB', bottomright to CC' and bottomleft to DD'. The colour scale ranges from -1 to +1 %.

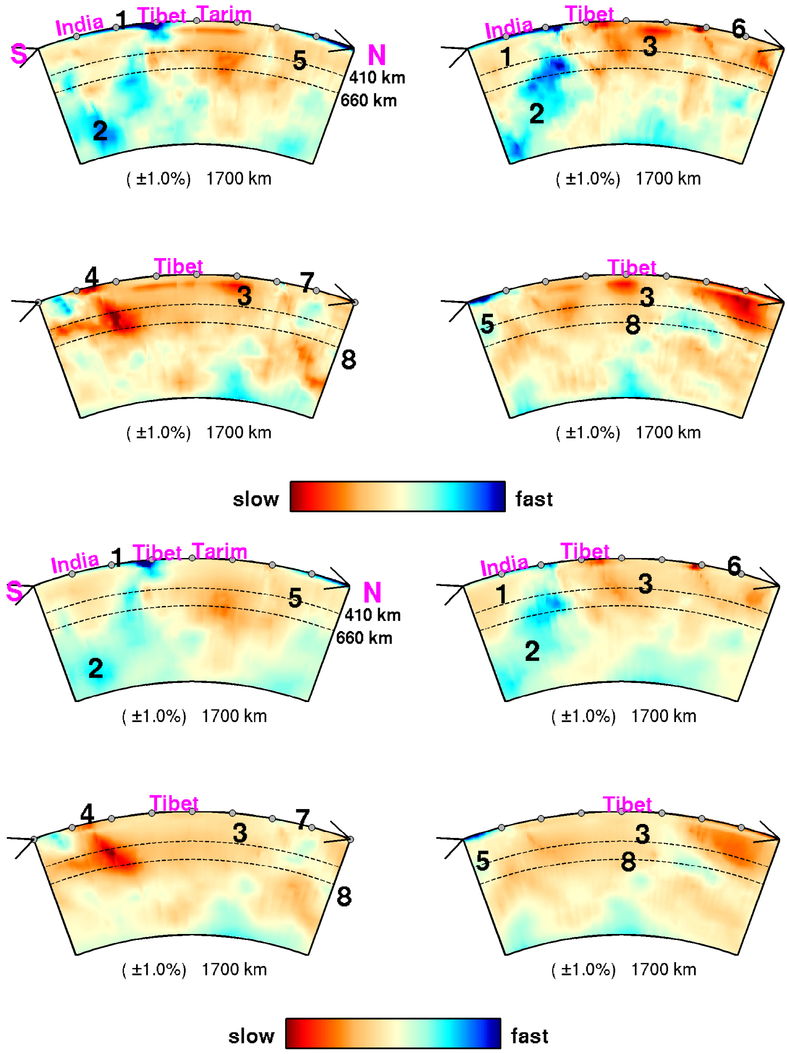


Figure 23: Solution obtained after exclusion of data from CSN and temporary array in Tibet (top) and correspoding new solution for α = -1.2 (bottom). Location of slices is identical to that of the slices in Figure 22. The colour scale ranges from -1 to +1 %.

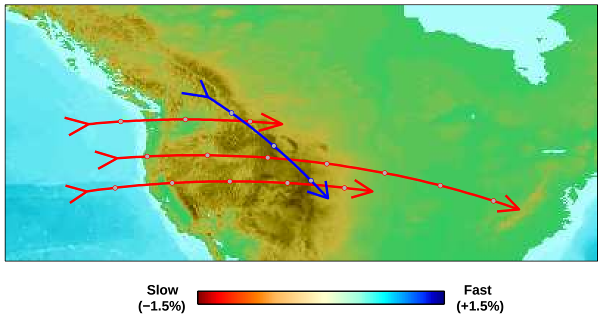


Figure 24: Map of North America showing locations of the tomographic slices in Figures 25 and 26. The slice displayed by the blue arrow crosses the Yellowstone hotspot track. The colour scale applies to all slices in Figures 25 and 26.

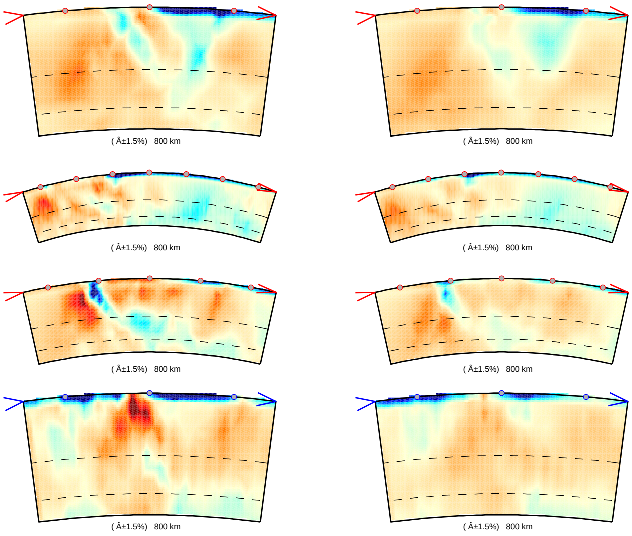


Figure 25: Original solution (left) and for the conservative solution for α = -1.7 and $m_t = \tilde{m}$ (right). Top slice corresponds to top red arrow in Figure 24, second slice to the middle red arrow, etc. Location of bottom slice is shown by the blue arrow.

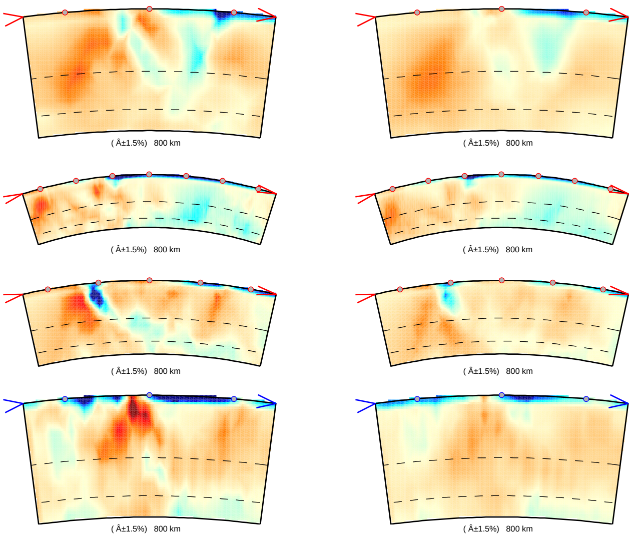


Figure 26: Solution obtained after exclusion of *P* phase data from *USArray* (left) and correspoding new solution for α = -1.6 (right).

this study can not be regarded as a complete uncertainty analysis for this model, since an infinite number of solutions exists that provide an acceptable fit to the data. However, by using the null space shuttle we are able to show the variability in amplitudes within the range of possible solutions and investigate to what extent individual anomalies in PMIT08 are constrained by the data. A detailed description of the robustness of individual mantle structures in the original model is beyond the scope of this paper and we therefore have given some illustrative examples of changes that can be made to PMIT08 by exploiting components of the model null space.

The non-uniqueness of the original solution (Figure 2) is apparent in Figure 3. While part of the regularization aims to reduce the non-uniqueness by imposing *a priori* constraints, it is clear that a broad range of viable solutions still exists. Many solutions could have been chosen as a final solution, with the two solutions minimizing the model norm or fitting the data best being the most notable. In Figure 4 the solution with maximum average amplitudes has roughly three (lower mantle) to four (upper mantle) times as large a norm as the solution with smallest norm in this figure. We interpret this as a lack of constraints on amplitudes of anomalies and see this as an example of the solution being dominated by the imposed regularization. The data have little bearing on amplitudes and the minimization of these amplitudes in PMIT08 by regularization reflects the subjectivity in the choice of the final model.

Furthermore, we want to point out that it is virtually impossible to use velocity perturbations in PMIT08 alone to infer robust information on thermochemical variations in the Earth. For instance, the RMS velocity perturbation in the lowermost mantle ranges from 0.2 to 0.6 % in Figure 4. Deschamps and Trampert (2003) find that the sensitivity of P wave velocity to temperature in the lowermost mantle is $\sim$ $-1.5 \cdot 10^{-5}$. To explain the velocity perturbations in the lowermost mantle in PMIT08 by a purely thermal origin would thus require temperature anomalies of 130 and 400 K for average velocity perturbations of 0.2 and 0.6 %, respectively. While such temperature anomalies may be possible in the lowermost mantle, the contrary is true for the upper mantle. The RMS velocity perturbation in the upper mantle around 350 km ranges from 0.3 to 1.3 % for our set of acceptable solutions (Figure 4). Cammarano et al. (2003) calculate a sensitivity of P wave velocity to temperature of $\sim -0.5 \cdot 10^{-5}$ for a 1300 °C adiabat around a depth of 350 km. To explain the velocity perturbations at this depth in PMIT08 by a purely thermal origin would thus require temperature anomalies of 600 and 2600 K for average velocity perturbations of 0.3 and 1.3 %, respectively. The upper limit of 2600 K is unrealistically large given temperatures in the Earth and one would conclude from these values that velocity perturbations may for the largest part be caused by chemical variations in the Earth. On the contrary, one could attribute the lower limit of 0.3 % to a more thermal origin. We thus obtain two different interpretations of thermochemical variations in the upper mantle for the range of RMS velocity perturbations inferred from the set of acceptable solutions. Consequently, one may question the value of such an interpretation of thermochemical variations based on a single 'optimal' solution, knowing that many other solutions fit the data that could lead to different interpretations.

The travel time data used in this study do not adequately resolve all anomalies present in the model PMIT08 to make these velocity perturbations robust. Parts of this model can be modified in such a way that physical interpretations of velocity perturbations in the original tomogram would require revision. As an example, let us discuss some low-velocity anomalies in PMIT08 extending vertically throughout the whole mantle, e.g. underneath Africa (Figure 12) and Hawaii (Figure 20b). Low-velocity anomalies are often interpreted as mantle plumes, i.e. upwellings of mantle material of anomalously high temperatures (e.g. Morgan (1971); Zhao (2001); Ritsema and Allen (2003)), although the existence of such mantle plumes is still heavily debated (see Foulger et al. (2005) for an overview). Furthermore, the origin depth of mantle plumes is a fundamental aspect in the discussion on the sources of hotspots. Existing hypotheses on these sources range from mantle plumes rising from the CMB into the upper mantle to the opposing theory that hotspots could be caused by shallow, plate-related stresses that fracture the lithosphere and induce volcanism along these fractures (Anderson, 2000; Foulger and Natland, 2003). These hypotheses heavily rely on the imaging of low-velocity perturbations and their associated depth extent imaged in high-resolution models, i.e. models containing short wavelength structures, produced by travel time tomography (e.g. Montelli et al. (2004); Li et al. (2008); Wolfe et al. (2009)). Naturally, this should urge one to question to what degree these low-velocity anomalies and their depths are actually constrained by the travel time data. In Figures 14 and 20d we presented an in our view perfectly acceptable solution to the tomographic problem at hand. From both figures, one might conclude that the low-velocity anomalies underlying the African continent and Hawaii are only present in the upper mantle. From this new solution one would never hypothesize that upper mantle plumes originate from upwellings near the CMB. So where does this leave us? Neither solution is false or true, but physical interpretations of these low-velocity anomalies oppose each other. Therefore, we strongly suggest that one should refrain from making physical interpretations based on one solution to the highly non-unique inverse problem when information on the robustness of this solution and structures therein is not available.

Furthermore, we would like to point out that using only delay times of body waves might not lead to robust images of velocity anomalies when the data used covers a small range of incidence angles, as is the case for the mantle below Hawaii in our study. The solution in Figure 20c has no low-velocity anomaly in the upper mantle underneath Hawaii. Nolet et al. (2007) points out that the vertical resolution in tomographic studies is dependent on the extent to which ray paths cross each other. Clearly, for the poor data coverage of the Hawaiian region, it is not possible to constrain the depth of the low-velocity anomaly in PMIT08 using P waves only. We encourage other workers on the tomography of the region such as Wolfe et al. (2009) to use the null space shuttle to verify the robustness of inferred continuous low-velocity anomalies.

A nowadays less controversial aspect of tomographic models regards the imaging of slabs,represented by high-velocity anomalies, and the variability in depth to which these slabs appear to sink in the mantle. More than a decade ago, (van der Hilst et al., 1997) showed evidence of some slabs sinking deep into the lower mantle and used this observation as an argument advocating whole mantle convection, observations that were confirmed by later studies (Kárason and van der Hilst, 2000; Li et al., 2008). Using the null space shuttle, we find that most slab segments in the lower mantle in PMIT08 are resolved well and are required by the data (cf. Figures 5 and 11). On the contrary, we find that resolving power in the upper mantle and the transition zone (between the 410 and 660 km discontinuities) in particular is highly variable and can be poor locally. Examples presented in this report showed that slabs in the upper mantle underneath Italy and Indonesia can be filtered out of the original solution with the null space shuttle (Figures 15 and 19). An attempt to minimize the total amplitude of anomalies in the transition zone (not shown in this report) even reduced the RMS norm of this region to 40 % of the original transition zone RMS norm.

In summary, we find that uncertainties in the original solution PMIT08 are largest in the upper mantle and the transition zone in particular. Most slab-like anomalies in the lower mantle appear to be robust wich may validate physical interpretations based on these structures. The opposite is true for low-velocity anomalies commonly interpreted as mantle plumes. In general, this type of velocity perturbation is not constrained well by the data throughout the whole depth extent of the mantle, the main lack of constraint being in the radial direction due to inadequate data coverage. Therefore, despite the fact that a model as PMIT08 has a large *potential resolution* as it is parameterized by many small blocks, the actual resolution may in some parts of the mantle not necessarily be higher than that inferred from long-wavelength models (Boschi and Dziewonski, 1999). One should bear in mind that our ability to image velocity anomalies in the real Earth does not merely depend on the type of parameterization, but that uniformity of data distribution and data quality are of paramount importance to obtain robust tomographic images. Boschi and Dziewonski (1999) verify the consistency of anomalies imaged in high- and low-resolution tomographic models and find that the long-wavelength component of high-resolution models is consistent with structures in models of lower resolution. A similar comparison between different types of tomographic modelling would be desirable to investigate the consistency of small-scale structures (under 100 km for slabs) imaged in PMIT08. However, all other models we know that aim to resolve structures at such short wavelengths are all constructed by travel time tomography (e.g. Bijwaard et al. (1998)) and may therefore suffer from similar shortcomings in resolving power and robustness as we showed for PMIT08 here. Assessing the quality of the model in terms of small-scale structure thus resides in methods such as the null space shuttle.

Inconsistencies or errors in the data and their size are a crucial aspect in the inversion. Available estimates of the random and systematic errors range from 0.3 s to 1.0 s for the ISC data. Given the difference of only 0.5 s between the RMS data residual of the reference model and the model in Figure 2, one may question the robustness of latter model given the high magnitude of the ISC data errors. These errors are hopefully not representative for the complete data set used in this study, but reliable estimates of errors in the data from regional networks and in the waveform-based differential times unfortunately remain largely unknown. Since reliable error estimations of both global and regional data sets are lacking, in this study we restricted ourselves to finding new solutions that have an average deviation of ~ 0.1 s from the original global data fit. We believe that this restriction is conservative given the estimates of data errors and that therefore the range of possible solutions presented in this study provides a first indication of the actual uncertainties in this particular travel time tomography problem. We leave it up to the reader to extrapolate the changes made to the original solution as shown in Figures 5 to 20d for larger deviations from the original solution's data fit.

The incorporation of phase arrival times from dense seismograph networks (e.g. Li et al. (2006); Burdick et al. (2008)) increases the amount of detail in the tomographic image, as shown for the Chinese region (cf. Figures 22a and 23a). More importantly, it seems that the incorporation of travel time data from Chinese seismograph networks increases the robustness of imaged anomalies, if we follow our simple comparison before and after minimizing the solution norm. As noted before, the difference between Figures 22a and 22b is not so dramatic that physical interpretations of the velocity anomalies would change, whereas the opposite could be concluded for the low-velocity anomalies in Figures 23a and 23b. We want to emphasize that, although arrival time data from dense networks increases resolving power, still more data is required to create tomographic images in which all structures in a region are adequately constrained by the data and are not the result of a priori constraints. Another example was given by the velocity structure beneath North America in Figures 25 and 26, from which no relevant differences in resolving power were observed between data sets with or without P wave travel times from the USArray. Again we note that we used only one phase and that incorporation of more data from this array as done by Burdick et al. (2010) will probably give more robust results for this region.

We commented on the use of several types of sensitivity tests to assess the reliability of tomographic images and features therein. We can not stress enough the major shortcoming of these tests that do not provide an analysis of true resolution independent of the choice of input model. A workaround commonly employed is to perform synthetic tests for input models of varying dominant wavelength content (e.g. Bijwaard et al. (1998)) but such tactis will never represent a complete and reliable analysis of resolving power. Nolet (2008) notes that by regularizing the solution beyond what is imposed by Bayesian constraints, i.e. somewhat objective physical constraints, the results of sensitivity tests will always reflect the subjectivity inherently imposed by a tomographer by including regularization that is not based on any physical information and aims at obtaining a smooth model. As an alternative to sensitivity tests, we advocate the use of the null space shuttle technique in linear tomographic problems to analyze the quality of a tomographic image and the robustness of its individual components. Besides being straightforward to implement, the method produces a new solution that can be easily visualized and interpreted in a physical sense, similar to the representation of the original solution. All that is required to correctly use the null space shuttle is a linear inverse problem and a corresponding linear resolution operator.

7 Conclusions

The global P wave speed mantle model PMIT08 is one of many solutions to the highly non-unique inverse problem. For a conservative increase in RMS data misfit of 0.1 s, a broad range of acceptable solutions can be found by using the null space shuttle. For a theoretical model equal to the original solution (Equation (5.2)), the RMS velocity perturbations of these solutions vary from 0.2 to 0.6 % in the lowermost mantle and from 0.3 to 1.3 % in the upper mantle (Figure 4). Such large variations in average amplitudes could indicate that constraints from the travel time data on amplitude are lacking and inhibit robust inferences on thermochemical variations in the Earth from PMIT08 only.

Furthermore, not all structures in PMIT08 are robust in terms of geometry of the anomalies. We find that a subset of the acceptable solutions contains structures in the upper mantle that strongly differ from those imaged in PMIT08, as demonstrated by the alternate shape or even absence of continous high-velocity anomalies in the upper mantle in these solutions. Velocity perturbations in the transition zone between 410 and 660 km depth are least constrained. On the contrary, the lower mantle component of high-velocity anomalies in PMIT08 seems to be robust. The radial extent of low-velocity anomalies (e.g. below Hawaii, Africa), associated with mantle plumes in the literature, is poorly constrained by the travel time data used. To acquire robust recovery of such anomalies, a more evenly spaced distribution of data (receivers) is needed.

We conclude that high-resolution models such as PMIT08 do not necessarily provide robust images of small-scale structures, such as slabs or narrow plume conduits. While the *potential resolution* of such models can be very high due to large amount of model parameters, the actual resolution might not be representative for all invididual anomalies imaged in the solution. The uneven distribution of both earthquakes and seismic stations inhibits an adequate sampling of the entire model space. The use of strong regularization might introduce anomalies in the model that are not required by the data. Therefore, physical interpretations of velocity perturbations in the tomographic image may be unreliable.

We strongly advise against inferring physical interpretations from a single solution to an underdetermined tomographic problem. In this study, we illustrated how the straightforward and efficient null space shuttle technique can be utilized to analyze solution robustness and we recommend the use of this method in future tomographic studies.

Acknowledgements

I am grateful for the assistance and advise from Jeannot Trampert, my supervisor here at Utrecht University, who was always available to answer my questions. Further, I thank the people at MIT: Rob van der Hilst for helping me on my way in this project and for providing me with his model PMIT08, for which I also thank Chang Li. I wish to thank Scott Burdick, who was always willing to lend me a hand during my time at MIT. My gratitude goes out to Samantha Hansen at Pennsylvania State University for providing me with a 64-bit version of the source code, which brought computation time down from hours to minutes. Computational resources for this work were provided by the Netherlands Research Center for Integrated Solid Earth Science (ISES 3.2.5 High End Scientific Computation Resources).

References

Aki, K., A. Christoffersson, and E. S. Husebye (1977). Determination of the three-dimensional seismic structure of the lithosphere. *J. Geophys. Res. 82*, 277–296.

Amaru, M. L., W. Spakman, A. A. Villaseñor, S. Sandoval, and E. Kissling (2008). A new absolute arrival time data set for Europe. *Geophys. J. Int. 173*, 465–472.

Anderson, D. L. (2000). The thermal state of the Upper Mantle; no role for mantle plumes. *Geophys. Res. Lett. 27*, 3623–3626.

Benoit, M. H., A. A. Nyblade, and J. C. VanDecar (2006). Upper mantle P wave speed variations beneath Ethiopia and the origin of the Afar Hotspot. *Geology 34*, 329–332.

Bijwaard, H. and W. Spakman (2000). Non-linear global P-wave tomography by iterated linearized inversion. *Geophys. J. Int. 141*, 71–82.

Bijwaard, H., W. Spakman, and E. R. Engdahl (1998). Closing the gap between regional and global travel time tomography. *J. Geophys. Res. 103*, 30,055–30,078.

Bolton, H. and G. Masters (2001). Travel times of P and S from global digital seismic networks: Implication for the relative variation of P and S velocity in the mantle. *J. Geophys. Res. 106*, 13,527–13,540.

Boschi, L. (2003). Measures of resolution in global body wave tomography. *Geophys. Res. Lett. 30*.

Boschi, L. and A. Dziewonski (1999). High and low resolution images of the Earth's mantle - Implications of different approaches to tomographic modeling. *J. Geophys. Res. 104*, 25,567–25,594.

Burdick, S., C. Li, V. Martynov, T. Cox, J. Eakins, L. Astiz, F. Vernon, G. Pavlis, and R. van der Hilst (2008). Upper mantle heterogeneity beneath North America from travel time tomography with global and USArray transportable array data. *Seismological Research Letters 79*, 384–392.

Burdick, S., R. van der Hilst, F. Vernon, V. Martynov, T. Cox, J. Eakins, L. Astiz, and G. Pavlis (2010). Model Update January 2010: Upper mantle heterogeneity beneath North America from travel time tomography with global and USArray Transportable Array data. *Seismological Research Letters 81*, 689–693.

Cammarano, F., S. Goes, P. Vacher, and D. Giardini (2003). Inferring upper-mantle temperatures from seismic velocities. *Phys. Earth Planet. Inter. 138*, 197–2222.

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

Deal, M. M., G. Nolet, and R. D. van der Hilst (1999). Slab temperature and thickness from seismic tomography, 1. Method and application to Tonga. *J. Geophys. Res. 124*, 372–380.

Deschamps, F. and J. Trampert (2003). Mantle tomography and its relation to temperature and composition. *Phys. Earth Planet. Inter. 140*, 277–291.

Dziewonski, A. M., B. H. Hager, and R. J. O'Connell (1977). Large-scale heterogeneities in the lower mantle. *J. Geophys. Res. 82*, 239–255.

Ekström, G., J. Tromp, and E. W. F. Larson (1997). Measurements and global models of surface wave propagation. *J. Geophys. Res. 102*, 8137–8157.

Engdahl, E. R., R. D. van der Hilst, and R. Buland (1990). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Geophys. J. Int. 102*, 24–43.

Fichtner, A., B. L. N. Kennett, H. Igel, and H.-P. Bunge (2009). Full waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophys. J. Int. 179*(3), 1703–1725.

Foulger, G. R. and J. H. Natland (2003). Is ḦotspotV̈olcanism a Consequence of Plate Tectonics? *Science 300*, 921–922.

Foulger, G. R., J. H. Natland, P. D. C., and D. L. E. Anderson (2005). Plates, plumes, and paradigms. *Geol. Soc. Am. Special Volume 288*, pp881+xi.

Grand, S. P. (1990). A possible station bias in travel time measurements reported to ISC. *Geophys. Res. Lett. 17(1)*, 17–20.

Grand, S. P. (2002). Mantle shear-wave tomography and the fate of subducted slabs. *Phil. Trans. R. Soc. Lond. 360*, 2475–2491.

Gudmundsson, O., J. H. Davies, and R. W. Clayton (1990). Stochastic analysis of global traveltime data: mantle heterogeneity and random errors in the ISC data. *Geophys. J. Int. 102*, 24–43.

Ishii, M. and J. Tromp (2004). Constraining large-scale mantle heterogeneity using mantle and inner-core sensitive normal modes. *Phys. Earth Planet. Inter. 146*, 113–124.

Kárason, H. and R. D. van der Hilst (2000). Constraints on mantle convection from seismic tomography. *History and Dynamics of Plate Motion, Geophys. Monogr. Ser. 121*, 277–288. edited by M. A. Richards, R. Gordon and R. D. van der Hilst.

Kárason, H. and R. D. van der Hilst (2001). Tomographic imaging of the lowermost mantle with differential times of refracted and diffracted core phases (PKP, P_{diff}). *J. Geophys. Res. 106*, 6569–6588.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from travel times. *Geophys. J. Int. 122*, 108–124.

Lévêque, J. J., L. Rivera, and G. Wittlinger (1993). On the use of the checker-board test to assess the resolution of tomographic inversions. *Geophys. J. Int. 115*, 313–318.

Li, C., R. D. van der Hilst, E. R. Engdahl, and S. Burdick (2008). A new global model for P wave speed variations in Earth's mantle. *Geochem. Geophys. Geosyst 9*(5).

Li, C., R. D. van der Hilst, and N. M. Toksöz (2006). Constraining spatial variations in P-wave velocity in the upper mantle beneath SE Asia. *Phys. Earth Planet. Inter. 154*, 180–195.

Li, X.-D., D. Giardini, and J. H. Woodhouse (1991). Large-scale three-dimensional even-degree structure of the Earth from splitting of long-period normal modes. *J. Geophys. Res. 96*, 551–577.

McSweeney, T. J. (1995). Seismic constraints on core structure and dynamics. Ph.D. thesis, Univ. of Wash., Seattle.

Montelli, R., G. Nolet, F. A. Dahlen, G. Masters, E. R. Engdahl, and S.-H. Hung (2004). Finite-frequency tomography reveals a variety of plumes in the mantle. *Science 303*, 338–343.

Morgan, W. (1971). Convection plumes in the lower mantle. *Nature 230*, 42–43.

Nolet, G. (1985). Solving or resolving inadequate and noisy tomographic systems. *J. Geophys. Res. 112*.

Nolet, G. (2008). A breviary of seismic tomography: imaging the interior of the earth and sun. *Cambridge University Press*, 344pp.

Nolet, G., R. Allen, and D. Zhao (2007). Mantle plume tomography. *Chemical Geology 241*, 248–263.

Paige, C. C. and M. A. Saunders (1982). LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Software 8*, 43–71.

Rawlinson, N. and M. Sambridge (2003). Seismic traveltime tomography of the crust and lithosphere. *Advances in Geophysics 46*, 81–198.

Ritsema, J. and R. M. Allen (2003). The elusive mantle plume. *Earth and Planetary Science Letters 207*, 1–12.

Ritsema, J., H. J. Heijst, and J. H. Woodhouse (1999). Complex shear wave velocity structure imaged beneath Africa and Iceland. *Science 286*, 1925–1928.

Röhm, A., J. Trampert, H. Paulssen, and R. Snieder (1999). Bias in reported seismic arrival times deduced from the ISC catalogue. *Geophys. J. Int. 137*, 163–174.

Schaff, D. P. and P. G. Richards (2004). Repeating Seismic Events in China. *Science 303*, 1176–1178.

Soldati, G. and L. Boschi (2005). The resolution of whole Earth seismic tomographic models. *Geophys. J. Int. 161*, 143–153.

Soldati, G., L. Boschi, and A. Piersanti (2006). Global seismic tomography and modern parallel computers. *Annals of Geophysics 49*, 977–986.

Spakman, W. and G. Nolet (1988). Imaging algorithms, accuracy and resolution in delay time tomography. *Mathematical Geophysics: A Survey of Recent Developments in Seismology and Geodynamics*, 155–188. edited by N. J. Vlaar.

Spakman, W., S. Stein, R. van der Hilst, and R. Wortel (1989). Resolution experiments for NW Pacific subduction zone tomography. *Geophys. Res. Lett. 16*, 1097–1100.

Tape, C., Q. Liu, A. Maggi, and J. Tromp (2009). Adjoint tomography of the southern California crust. *Science 325*, 988–992.

Trampert, J. and R. van der Hilst (2005). Towards a quantitative interpretation of global seismic tomography. *Geophysical Monograph (American Geophysical Union, Washington, D.C.) 160*, 47–62.

van der Hilst, R. D. and P. Mann (1994). Tectonic implications of tomographic images of subducted lithosphere beneath northwestern South America. *Geology 106*, 451–454.

van der Hilst, R. D., S. Widyantoro, and E. Engdahl (1997). Evidence for deep mantle circulation from global tomography. *Nature 386*, 578–584.

Vasco, D. W. and L. R. Johnson (2003). Resolution, uncertainty, and whole Earth tomography. *J. Geophys. Res. 108 (B1)*, 2022.

Wolfe, C. J., S. C. Solomon, G. Laske, J. A. Collins, R. S. Detrick, J. A. Orcutt, D. Bercovici, and E. H. Hauri (2009). Mantle Shear-Wave Velocity Structure Beneath the Hawaiian Hot Spot. *Science 326*, 1388–1390.

Woodhouse, J. H. and A. M. Dziewonski (1984). Mapping the Upper Mantle: Three-Dimensional Modeling of Earth Structure by Inversion of Seismic Waveforms. *J. Geophys. Res. 89*, 5953–5986.

Wysession, M. E. (1996). Large-scale structure at the core-mantle boundary from diffracted waves. *Nature 382 (6588)*, 244–248.

Zhao, D. (2001). Seismic structure and origin of hotspots and mantle plumes. *Earth Planet. Sci.Lett. 192*, 251–265.