

# Full waveform tomography of the upper mantle in the South Atlantic region: Imaging a westward fluxing shallow asthenosphere?



Lorenzo Colli<sup>a,\*</sup>, Andreas Fichtner<sup>b</sup>, Hans-Peter Bunge<sup>a</sup>

<sup>a</sup> Department of Earth and Environmental Sciences, Ludwig Maximilian University, Theresienstrasse 41, 80333 Munich, Germany

<sup>b</sup> Department of Earth Sciences, ETH Zurich, Sonneggstrasse 5, 8092 Zurich, Switzerland

## ARTICLE INFO

### Article history:

Received 23 July 2012

Received in revised form 18 June 2013

Accepted 19 June 2013

Available online 28 June 2013

### Keywords:

Full waveform inversion

Adjoint method

Seismic tomography

Pressure-driven asthenospheric flow

Upper mantle convective planform

## ABSTRACT

A prominent feature of the South Atlantic region is its strongly asymmetric residual bathymetry across the ocean basin. It has been suggested that the residual bathymetry is dynamic in nature, arising from the large slow velocity seismic anomaly located in the lower mantle beneath the African plate. Unfortunately, the pattern of mantle heterogeneity particularly in the upper mantle is not well known owing to the sparsity of seismic stations and the existence of large aseismic regions on the African and South American plates. Here we present a new seismic tomographic study of the South Atlantic upper mantle. Our model is based on a full seismic waveform inversion of  $\approx 4000$  high-quality seismograms for isotropic 3-D seismic structure using a powerful adjoint methodology capable of extracting maximum information from each seismogram. The theory requires simulation of seismic wave propagation in 3-D heterogeneous earth models computed with a spectral-element method where the differences between observed and synthetic seismograms are quantified using phase misfits obtained through a time–frequency transform. The model images a continuous channel of pronounced slow seismic velocity in the shallow sublithospheric mantle between  $\approx 150$  and  $\approx 300$  km depth that branches in between the cratonic roots under the African and South American continents. At greater depth, below 300–350 km, the slow anomalies are less pronounced and a change in the convective planform is indicated by isolated, round shaped patches in an overall faster mantle. It is possible that the depthwise change of the convective planform from vertical to horizontal advection of hot buoyant material in a low viscosity asthenosphere can reconcile the anomalous residual bathymetry in the region, and we present a simple fluid dynamic model of pressure driven flow to assess the feasibility of this scenario.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

The South Atlantic region is characterized by major structural elements that include the conjugate margins of South America and Africa with their adjacent cratonic cores, the Parana and Etendeka continental flood basalts and their hotspot tracks along the Walvis Ridge and Rio Grande Rise, as well as the Tristan da Cunha, St. Helena and Ascension hotspots (Fig. 1). The presence of well preserved magnetic isochrons (Fig. 2a), moreover, allows one to constrain the opening history of the basin, so that residual ocean floor topography can be assessed after removing thermal cooling effects within the oceanic lithosphere. An outstanding observation from the residual topography maps (Fig. 2b) is the existence of strongly anomalous bathymetry cutting across the structural elements. Elevated topography,

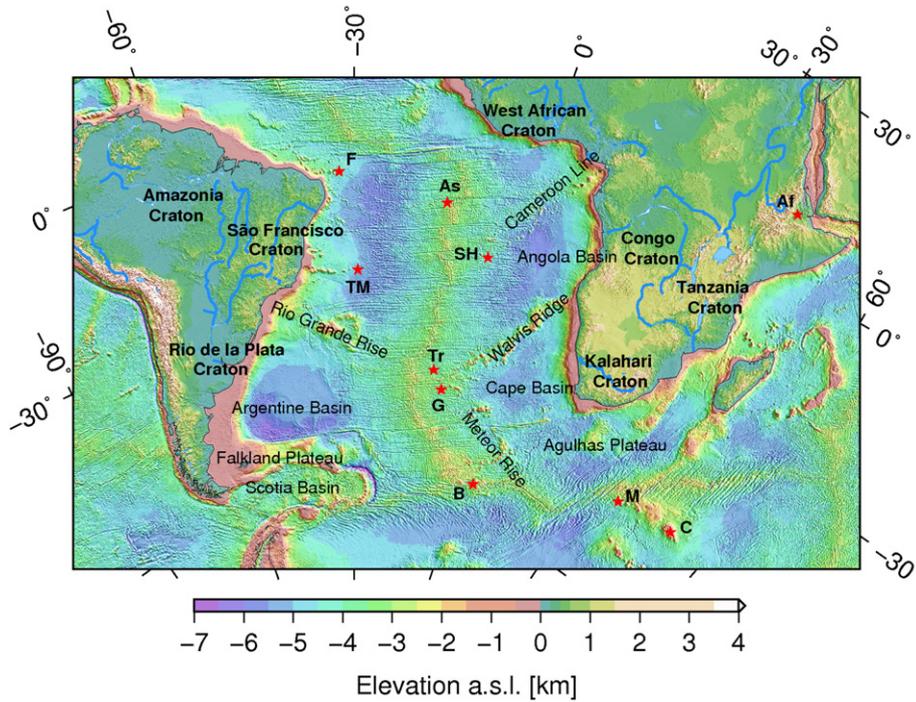
termed the African superswell by Nyblade and Robinson (1994), consists of uplifted portions of the African continent and areas of abnormally high bathymetry in the south-eastern Atlantic, while much of the south-western Atlantic especially in the Argentine Basin is abnormally deep.

Pronounced residual bathymetry in the South Atlantic implies that significant topography must be supported by heterogeneity beneath the tectonic plates. Whether this topography originates from upper mantle flow directly beneath the lithosphere, or whether it reflects dynamic support of deeply seated buoyancy in the lower mantle, remains unclear. Global tomography models (e.g., Grand, 2002; Grand et al., 1997; Ritsema et al., 2011; Simmons et al., 2007) persistently image slow seismic velocities in the lower mantle beneath the African plate, and a substantial portion of this wave speed reduction is probably due to highly elevated temperature (Schuberth et al., 2009a, 2009b) so that the region may act as a source of considerable thermal instabilities.

Anderson (1982) noted that the sub-African mantle had long been shielded from subduction by the former supercontinent Pangea.

\* Corresponding author. Tel.: +49 89 2180 4215; fax: +49 89 2180 4205.

E-mail addresses: [lorenzo.colli@geophysik.uni-muenchen.de](mailto:lorenzo.colli@geophysik.uni-muenchen.de) (L. Colli), [andreas.fichtner@erdw.ethz.ch](mailto:andreas.fichtner@erdw.ethz.ch) (A. Fichtner), [hans-peter.bunge@geophysik.uni-muenchen.de](mailto:hans-peter.bunge@geophysik.uni-muenchen.de) (H.-P. Bunge).



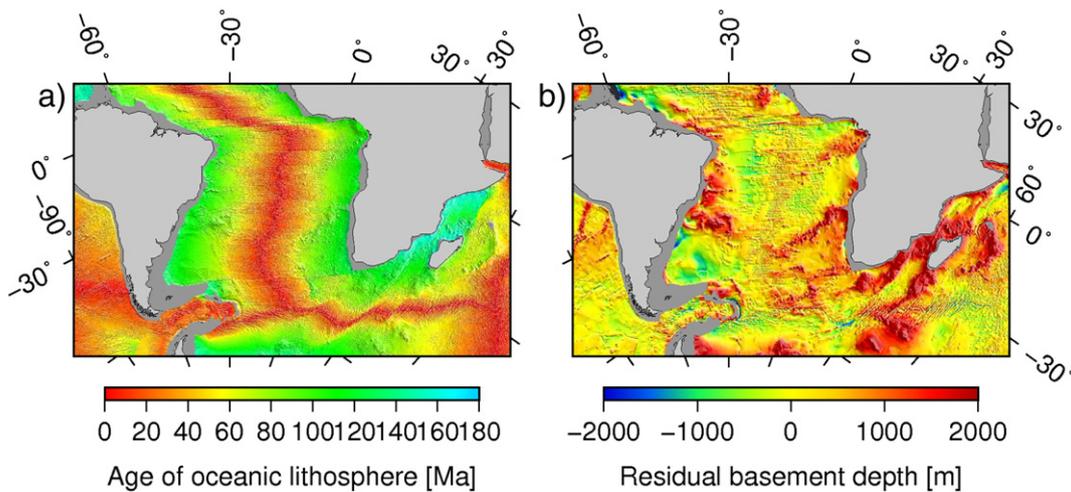
**Fig. 1.** Topographic map of the South Atlantic and adjacent continents from EOTPO1 (Amante and Eakins, 2009), annotated with major structural elements cited in the text. Oblique Mercator projection: central parallel passes through the points (30°S, 30°W) and (0°S, 60°E), projection centred on (28.024°S, 7.204°W). Craton names are boldface, while stars denote prominent hotspots (F: Fernando de Noronha; As: Ascension; Af: Afar; SH: Saint Helena; TM: Trinidad and Martim Vaz; Tr: Tristan da Cunha; G: Gough Island; B: Bouvet Island; M: Marion; C: Crozet Islands).

Significant hot thermal upwellings are thus expected in the region, in agreement with inferences that Africa experienced greater uplift in the Tertiary than other continents (e.g., Burke and Gunnell, 2008). A variety of geodynamic models (e.g., Forte et al., 2010; Gurnis et al., 2000; Lithgow-Bertelloni and Silver, 1998; Moucha and Forte, 2011) suggest deep mantle heterogeneity as a plausible cause for high topography in Africa and the South Atlantic, and that the entire region is influenced by a major mantle convection cell (Husson et al., 2012).

The African plate contains many volcanic centres that may be interpreted as the surface expression of mantle plumes rising from the lower mantle, but the path of deep material taken on its passage

from the lower into the upper mantle, and its advective redistribution within the asthenosphere and toward the Mid-Atlantic ridge system is not well known. It must be inferred from seismic studies.

Upper mantle structure in the South Atlantic region remains poorly studied owing to the sparse distribution of seismic stations and the existence of large aseismic areas on the African and South American plates. Temporary deployments of seismometers were carried out at some localities, for example across the shields in southern Africa, to provide detailed seismic images of the crust and upper mantle (e.g., Chevrot and Zhao, 2007; Freybourger et al., 2001). Tomographic studies have revealed thick, seismically fast



**Fig. 2.** Left: age-area distribution of the ocean floor from Müller et al. (2008); right: residual basement depth grid computed by calculating the difference between the predicted basement depth and the sediment unloaded basement depth. Predicted basement depth is obtained by applying Crosby et al.'s (2006) North Pacific thermal boundary layer model to the age-area distribution from Müller et al. (2008). Oblique Mercator projection. Note the anomalous bathymetry cutting across the regional structural elements, with excess topography in correspondence with the African superswell and anomalously high bathymetry in the south-eastern Atlantic.

“keels” under the cratonic regions of Africa, with slow structure prevailing elsewhere (e.g., Priestley et al., 2006; Ritsema et al., 1999). Many seismic studies relied on surface wave data (Pasyanos and Nyblade, 2007; Ritsema and van Heijst, 2000) for spatial coverage, mapping significant shear velocity features down to depth levels of 200 km and corresponding mostly to cratonic structures. Maps of upper mantle anisotropy (Debayle et al., 2005; Priestley et al., 2008; Sebai et al., 2006) have yielded additional constraints on material flow directions in the region and have concluded that there is a strong and spatially complex azimuthal anisotropy in the African lithosphere, while the signal is weaker and simpler in the asthenosphere below.

While raypath tomography is progressing, a refined physical modelling of seismic wave propagation has become possible: finite-frequency effects, such as wave scattering and wavefront healing, bear significantly on the behaviour of seismic waves, and should be taken into account to improve the seismic images. To yield synthetic seismograms, one must compute the propagation of the full wavefield in a three-dimensional heterogeneous Earth model. Finite-frequency effects must also be accounted for in the inversion procedure and in the construction of effective misfit functions.

Full waveform tomography, made possible by the dramatic gain of computational resources in recent years, takes advantage of a numerical solution of the wave equation to account for the effects of 3-D heterogeneous seismic structures on wave propagation. The numerical solution allows one to treat simultaneously all direct, reflected and scattered body and surface waves, thus considerably increasing the exploitable information carried by each seismogram. Here we employ this technique to invert for seismic structure in the South Atlantic region, which is characterized by scarcity of high quality seismograms. In order to best extract information and invert for 3-D seismic structure, a suitable misfit function needs to be used. To this end, we apply a time–frequency misfit (Fichtner et al., 2008) for comparison of observed and synthetic data, together with an efficient adjoint method (Fichtner et al., 2006) for the iterative correction of model parameters.

## 2. Method

### 2.1. The adjoint method

The aim of our present work is to obtain the speed of propagation of seismic waves in the South Atlantic upper mantle region. As direct, in situ measurements of the seismic wave speed are impossible, we have to resort to an indirect one. By exploiting the data recorded at seismic stations across the region of interest, we can obtain a synthetic model of the upper mantle seismic structure, defined by a set of parameters  $\mathbf{p}$ , that can reproduce the observed ground motion. We thus want to minimize the misfit between real data recorded at seismic stations and synthetic data generated with our model.

To quantify the misfit between synthetic and observed seismograms, one must establish a misfit function  $\chi(\mathbf{u}, \mathbf{p})$  that depends on the model parameters  $\mathbf{p}$  and the synthetic seismogram  $\mathbf{u}$  generated using those parameters. It is possible to improve the synthetic model starting from an initial guess for the model parameters  $\mathbf{p}_0$  that produced the mismatched initial solution  $\mathbf{u}_0$ . One then calculates a parameter correction  $\mathbf{d}$  which minimizes the misfit function by gradient descent, that is by computing the total derivative (Fréchet derivative) of  $\chi$  with respect to the model parameters  $\mathbf{p}$ :

$$D_{\mathbf{p}}\chi(\mathbf{u}, \mathbf{p})(\mathbf{d}) = \partial_{\mathbf{u}}\chi(\mathbf{u}, \mathbf{p})D_{\mathbf{p}}\mathbf{u}(\mathbf{d}) + \partial_{\mathbf{p}}\chi(\mathbf{u}, \mathbf{p})(\mathbf{d}). \quad (1)$$

While the partial derivatives of  $\chi$  can be computed quite easily, the computation of  $D_{\mathbf{p}}\mathbf{u}(\mathbf{d})$ , that is the variation of the synthetic seismogram  $\mathbf{u}$  due to a change  $\mathbf{d}$  in the model parameters, is involved.

Classical finite differencing techniques can be employed, but they become infeasible in the case of a large parameter space, like the one needed to properly describe the seismic structure of the Earth. It is here that the adjoint method provides an elegant way to compute the derivative with respect to the parameters by combining the synthetic forward solution  $\mathbf{u}$  with an ancillary function  $\boldsymbol{\psi}$ , the solution of the adjoint equation.

The adjoint technique is a general and versatile method to compute partial derivatives with respect to model parameters, and has been implemented for a range of geophysical applications that include, in addition to seismic tomography, lithosphere, mantle and core dynamics (Bunge et al., 2003; Chen et al., 2007; Fournier et al., 2010; Iaffaldano et al., 2007; Tape et al., 2009; Tarantola, 1988; Tromp et al., 2005). The adjoint method exploits the fundamental fact that the change in the observed quantity generated by a change in the model parameters is governed by the physics of the system. In the case of seismology, a change in the elastic properties of some part of the model causes the wavefield passing through that region to be reflected, refracted and diffracted differently, leading to a variation in the synthetic seismograms. This perturbation of the wavefield is still a wavefield itself, and propagates in accordance to the constraints of the wave equation. It is possible to backtrack the place of origin of the perturbation by solving for a subsidiary wavefield subject to a set of equations, called adjoint equations, that are derived from the original wave equation. The adjoint equations are still of the wave type, and the source term in the adjoint equation is derived from the misfit between observed and synthetic seismograms (further details are available for the interested reader in Appendix A). The choice of the misfit function  $\chi$  is thus particularly important: since we are using a first-order algorithm, a misfit function which relates linearly to perturbations in the parameters will give a faster and smoother convergence.

We next present the adjoint problem for continental-scale full waveform tomography and describe our choice for the misfit function.

### 2.2. Numerical method and initial model

As we mentioned before, it is important to compute the propagation of the full wavefield through the 3-D heterogeneous seismic model, in order to avoid artefacts generated by simplifying approximations. Since this is beyond an analytic solution, simulation of elastic wave propagation and computation of Fréchet kernels is done numerically with the code SES3D (Fichtner et al., 2009), implemented on dedicated capacity computing infrastructure (Oeser et al., 2006). The similarity between the original wave equation and its adjoint equation implies that the same code used to solve for the forward wavefield can be used, with only minor changes, to solve for the adjoint field. SES3D is based on the spectral-element method and operates in a spherical section. The unphysical boundaries of the spherical section cause spurious reflections that prevent a meaningful comparison between synthetic and observed data. In order to avoid these reflections we absorb waves close to the boundaries by using the anisotropic perfectly matched layer technique proposed by Teixeira and Chew (1997) and Zheng and Huang (1997). We employ a spherical section of  $140^\circ \times 80^\circ \times 1440$  km going from the surface to the mid-mantle, divided into nearly 165,000 elements that are  $1.5^\circ \times 1.6^\circ \times 40$  km. Inside each element the dynamic fields are approximated by degree 6 Lagrange polynomials, for a total of nearly  $30 \times 10^6$  gridpoints. The minimum theoretical period for this grid is around 55 s. We implement Ritsema's s20rts (Ritsema and van Heijst, 2004) as our initial model, to speed up convergence, so that the longest wavelength features of the South Atlantic upper mantle are already present in the initial model. This ensures an acceptable initial fit between synthetic seismograms and data and allows us to effectively correct the model.

### 2.3. Crustal structure

The long-period (120 s and longer) waves used in this study are influenced by the crust without being able to resolve its fine layering. The only way to represent the crustal structure accurately is to honour all its thin layers within the spectral-element mesh, which results in an extremely fine mesh. However, as the dominant wavelength is roughly 400 km in our study, such fine mesh would waste computational time, both because the wavefield is well represented on a coarser grid, and because the waves are not as sensitive to each thin layer of the model as they are to its long-wavelength features. We therefore implement a smooth long-wavelength equivalent (Fichtner and Igel, 2008) of the MDN crustal model by Meier et al. (2007) constructed explicitly to provide a crustal correction for long-period tomography. This allows us to use larger elements without introducing artefacts due to the absence of a proper crust. Together with the rest of the model, we then improve the crustal model iteratively even if it remains a long-wavelength equivalent. The upper 50 km of our model thus cannot be interpreted in terms of real-earth seismic structure.

### 2.4. Misfit function

As stated before, it is important to choose a suitable misfit function for the quantification of the differences between synthetic and observed seismograms, in order to extract the maximum amount of meaningful information from each seismogram and to improve the convergence rate. A direct comparison of the seismograms' amplitudes (e.g., the  $L_p$  norm of the difference  $\mathbf{u}^0 - \mathbf{u}$ ) is unsatisfactory, since amplitude is influenced non-linearly by seismic anomalies, in particular by local anomalies near the receiver. A small variation in seismic velocity, producing a minor time shift of the waveform, can therefore induce a large instantaneous amplitude misfit even if the total energy arriving at the receiver remains almost constant; moreover, source inversions are often not accurate enough for the needs of an amplitude tomography.

A better choice is the use of the phase difference between synthetic and observed seismograms. This measure, akin to a measure of instantaneous time delay but performed in a time–frequency domain, is more linearly related to seismic anomalies and less degraded than amplitude by interference of different waves. Hence we disregard amplitudes and use phase differences to quantify the misfit between synthetic and observed waveforms: given an observed seismogram  $u_i^0(t)$ , where the index  $i$  runs over the three components, and the corresponding synthetic seismogram  $u_i(t)$ , we map the seismograms into the time–frequency domain using the windowed Fourier transform

$$\tilde{v}(t, \omega) = F_h[v](t, \omega) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} v(\tau) h(\tau - t) e^{-i\omega\tau} d\tau, \quad (2)$$

with  $h(t) = (\pi\sigma)^{-1/4} e^{-t^2/2\sigma^2}$  and  $\sigma$  set approximately equal to the dominant period of the considered waveform – in our present study, we used a value of 100 s. We then write the transformed observed and synthetic seismograms in exponential form, making explicit their phase and magnitude

$$\tilde{u}_i^0(t, \omega) = |\tilde{u}_i^0(t, \omega)| e^{i\phi_i^0(t, \omega)}, \quad \tilde{u}_i(t, \omega) = |\tilde{u}_i(t, \omega)| e^{i\phi_i(t, \omega)}, \quad (3)$$

and define a phase misfit  $E$  as the weighted  $L_2$  norm of the phase difference  $\phi_i^0(t, \omega) - \phi_i(t, \omega)$ :

$$E^2(u_i^0, u_i) := \int_{\mathbb{R}^2} W^2(t, \omega) [\phi_i^0(t, \omega) - \phi_i(t, \omega)]^2 dt d\omega. \quad (4)$$

$W(t, \omega)$  is a positively defined weighting function that allows us to select and weigh the various waveforms both time- and frequency-wise, excluding phase jumps and reducing the influence of seismic noise, and thus enhancing convergence. In practice it consists of the sum of various time windows that select the usable portion of each waveform and decay to zero outside this region. Each time window is multiplied by a constant between 0 and 100 that upweighs waveforms with a better signal-to-noise ratio and less frequent station–event combinations (e.g., north–south trending over east–west trending).  $W(t, \omega)$  is constant in frequency between 1/300 and 1/120 Hz and decays exponentially outside this interval.

A complication of the phase misfit is given by the possibly discontinuous nature of the phases  $\phi_i^0(t, \omega)$  and  $\phi_i(t, \omega)$ , which implies discontinuities up to  $\pm 2\pi$  for the phase difference  $\Delta\phi_i = \phi_i^0(t, \omega) - \phi_i(t, \omega)$ . Some of these phase jumps are contingent and can be removed by calculating the phase difference via the Fourier transform of the correlation function (Fichtner et al., 2008). In other cases, when data and synthetics are too dissimilar and out of phase, discontinuities are fundamental in nature and cannot be removed: they are symptomatic of cycle skips between synthetics and real data, and they signal a point where the misfit function is no longer physically meaningful. These jumps must be excluded by tapering in the time–frequency domain via a proper choice of  $W(t, \omega)$ .

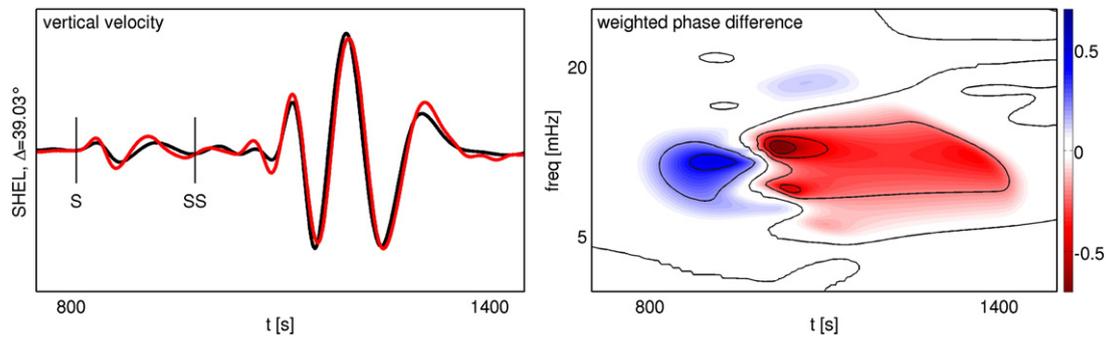
The total misfit  $\chi$  is computed as a sum of each individual phase misfit over all the seismograms and all the events.

### 2.5. Sensitivity kernels

Applying the above methodology to each individual event we obtain the total derivative of the misfit function with respect to the model parameters (see Eqs. (1) and (A.11)) for that event, also called misfit sensitivity kernel, an example of which is given in Fig. 4. The complex kernel structure arises from the sum of different concomitant factors: first and foremost the complexity of each single-frequency source–receiver kernel, then their interfering superposition both frequency- and receiver-wise. Due to finite-frequency effects a kernel obtained using a single frequency of a single seismogram already has a fully developed 3-D structure: the region with erroneous elastic parameters that cause the discrepancy between the observed and the synthetic seismogram may lie well outside the theoretical 1-D raypath. Moreover, waves travelling with different frequencies are characterized by Fresnel zones of different widths, and thus grant different spatial resolutions. In addition it is well known that higher frequency surface waves are more sensitive to shallower anomalies, resulting in a natural overall variation with depth of the kernel's resolution. When adding the contribution of all the receivers, the single kernels will interfere constructively in the regions that need to be corrected, and they will interfere destructively in those regions that are already well characterized. This trend is reinforced and enhanced when the kernels obtained from all events are summed, giving the direction  $\mathbf{d}$  of the total derivative of our chosen misfit function with respect to the model parameters.

### 2.6. Inversion procedure

The inversion procedure consists thus of the following steps: (1) the propagation of the full wavefield in the current  $n$ -th model, defined by the set of parameters  $\mathbf{p}_n$ , is computed; (2) the difference between synthetic and observed seismograms is quantified using the phase misfit, obtaining the source term for the adjoint equation; (3) the adjoint equation is solved and the total derivative of the misfit function is calculated; (4) an optimal step length is determined using a parabolic line search: assuming the misfit function  $\chi(\mathbf{u}, \mathbf{p})$  to be quadratic in  $\mathbf{p}$  near the best model  $\mathbf{p}_{min}$ , we fit a quadratic polynomial in the trial step length  $s$  through the values of misfit for the current  $n$ -th model  $\mathbf{p}_n$  and the two trial models  $\mathbf{p}_n + s_1\mathbf{d}$  and  $\mathbf{p}_n + s_2\mathbf{d}$ . The values



**Fig. 3.** Left-hand panel: vertical-component velocity seismograms from an event in Tanzania recorded at the station SHEL on Saint Helena. The recorded seismogram is plotted as a thick black line while the synthetic is plotted as a thick red line. Right-hand panel: weighted phase difference  $W(t,\omega)\Delta\phi(t,\omega)$  corresponding to the seismogram on the left-hand side. Contour lines are plotted at multiples of 20% of the maximum value. Both S and SS modelled body waves arrive early relative to the observed phases and show a positive phase difference, while the Rayleigh wave is slightly late and shows an overall negative phase difference. Notice however the frequency dependence of the phase difference: the time of maximum phase delay at medium frequencies is a time of maximum phase advance at higher frequencies.

of the trial step lengths  $s_1$  and  $s_2$  are assigned arbitrarily on the base of physical intuition, direct experience with the inversion and rough quantitative analysis of the misfits. In practice they are such that the maximum resulting correction is between a fraction of a percent and a few percent in wavespeed. Their values tend to decrease with further iterations, as the model gets better and smaller corrections are needed. Finally, (5) a new  $n + 1$ -th model is obtained by applying the computed correction to the current  $n$ -th model. These steps are repeated iteratively until the differences between synthetic and observed seismograms become sufficiently small.

### 3. Data

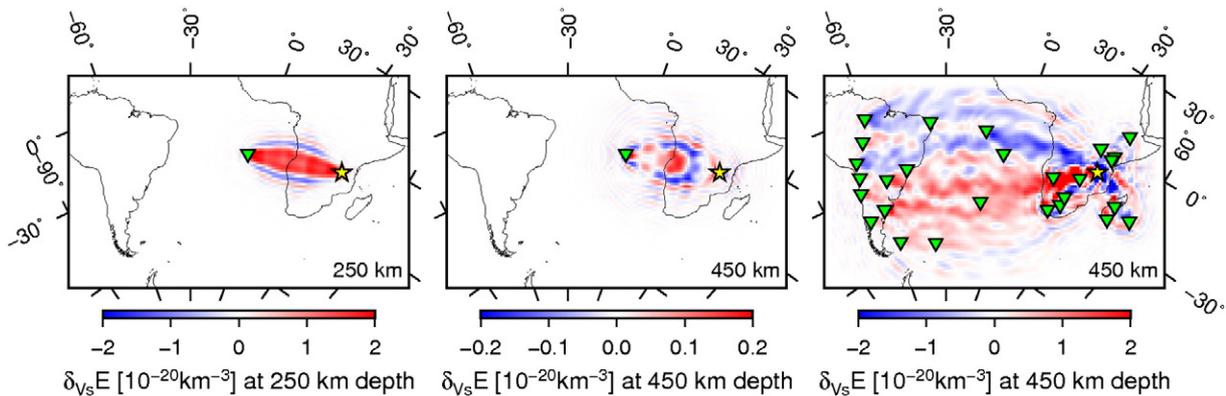
We selected 68 events occurring between 2007 and 2011 for a total of 3916 seismograms, mainly vertical-component (Fig. 5). About 95% of the data comes from permanent networks operated by IRIS, GEOSCOPE, USGS, AfricaArray, University of Sao Paulo and Universidad de Chile, while the remaining 5% comes from the Chile RAMP and Sierras Pampeanas temporary networks.

In order to neglect finite-source effects we accepted only events whose magnitude is smaller than  $M_s$  6.9. Moreover, we discarded events with less than 20 high-quality recordings in our region of

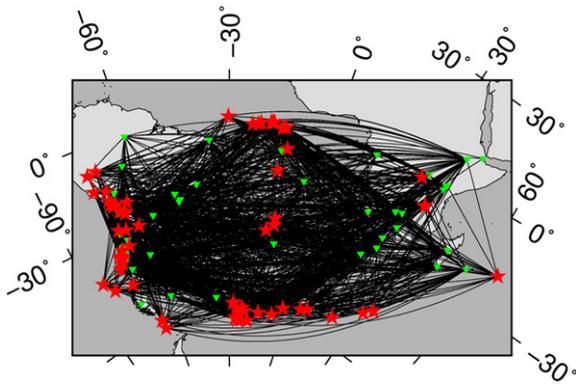
interest: the adjoint method is maximally efficient when there is a large seismogram-to-event ratio, and our computational resources are limited.

The majority of the waveforms consist of fundamental-mode surface waves, which have the best signal-to-noise ratio. The remaining waveforms are higher-mode surface waves and S body waves. We individually upweigh less frequent waveforms in order to keep the model unbiased toward a particular seismic phase, and to improve the resolution at depth. We used surface waves recorded at epicentral distances between  $5^\circ$  and  $125^\circ$ , while we used only direct body waves recorded at distances of no more than  $40^\circ$  because beyond that distance the synthetic body wave is influenced by the unphysical domain boundary of our mesh at 1440 km depth. For this reason we don't have a good resolution in the lower mantle.

Hypocentre coordinates and moment tensor components for all 68 events were taken from the Centroid Moment Tensor catalogue (<http://www.globalcmt.org>). We decided against inverting for source parameters for two main reasons: on the one hand, we use rather long-period waves, so our inversion is not affected much by errors in them; on the other hand for our inversion we would be able to use waves coming only from limited values of azimuths, since the overwhelming majority of the events is close to the border of our



**Fig. 4.** Left-hand panel: horizontal slice at 250 km depth through the sensitivity kernel calculated with the adjoint source derived from the waveforms shown in Fig. 3. The position of the epicentre is marked by a yellow star while the station SHEL is denoted by a green triangle. The regions in blue are those that can possibly be generating the mismatch between the synthetic and the observed seismogram because of a too high  $s$ -wave speed, while those in red have a possibly too slow  $s$ -wave speed. This relatively shallow depth is mainly illuminated by the surface wave, that was characterized by a phase delay. The kernel thus points to an increase of the  $s$ -wave speed in the region between the station and the receiver that is crossed by the main wave, with alternating rings of slow and fast scatterers surrounding this region. Central panel: horizontal slice at 450 km depth through the same sensitivity kernel of the left-hand panel. This portion of the model is illuminated by both the surface and the body waves, causing a more complex pattern of slow and fast regions. Right-hand panel: horizontal slice at 450 km depth through the sensitivity kernel calculated for the same event but using all available waveforms. Oblique Mercator projection.



**Fig. 5.** Ray coverage for the 3916 seismograms used in this study. Epicentres are marked by red stars while stations are denoted by green triangles. Oblique Mercator projection.

computational domain. Inverting for hypocentre coordinates and moment tensor components would result in a misfit reduction without implying a real improvement of the model.

## 4. Results

### 4.1. 3-D S-wave speed model

Using the method described in Section 2 we inverted the data presented in the last section for isotropic upper mantle seismic structure in the South Atlantic region. The resulting isotropic model after five iterations is presented in Figs. 6–10.

At a depth of 100 km (Fig. 6) the model is mainly influenced by the lithosphere's compositional and structural variations, both under the continents and under the ocean. In continental regions cratons stand out as fast anomalies surrounded by relatively slower regions. Moving into the oceanic regions, the thick and relatively fast oceanic lithosphere near the continents gives way to slower mantle material towards the mid-oceanic ridges. The main exception is given by the two slow anomalies extending from the Mid-Atlantic Ridge towards South America on one side and towards Africa on the other.

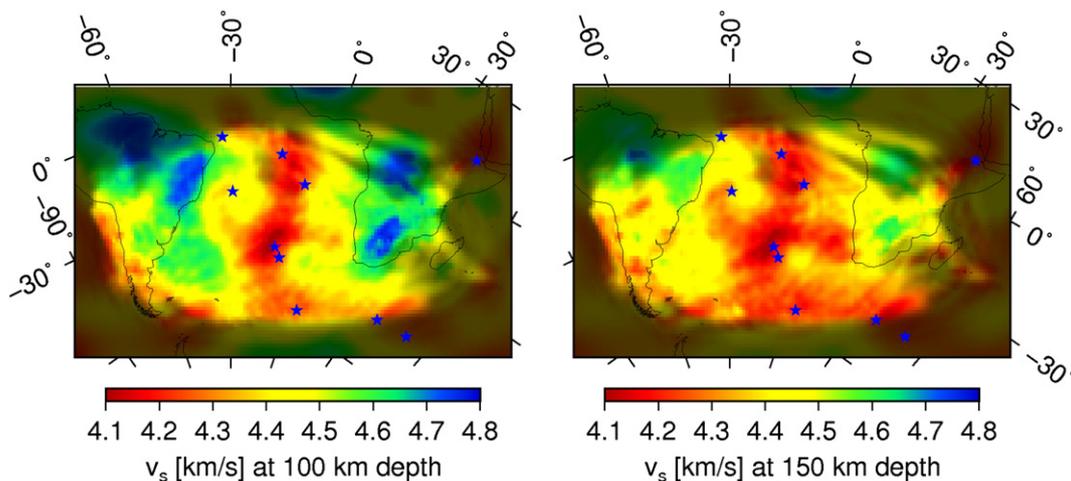
Deeper down in the mantle (Figs. 6, 7 and 10) the fast anomalies related to the continental cratons become smaller and fainter, while

slow velocities are overall dominant, in particular in the sub-oceanic mantle where the slow velocity anomalies are mainly elongated at an angle with the strike of the Mid-Atlantic Ridge. Between 200 and 300 km depths some slow regions protrude under the continents, prolongations of the slow anomalies under the Rio Grande Rise and the Walvis Ridge. Under Africa a slow anomaly weaves its way between the fast velocity anomalies of the Kalahari and the Congo cratons, connecting the slow region around Afar with the slow anomalies under the Atlantic Ocean. Under South America the slow velocity anomaly splits into two branches, one protruding northwards under central Brazil while the other extending westward under central Argentina and Chile. Below 300 km depth (Fig. 8) slow anomalies are no longer dominant, with also a distinct change in their shape from elongated to rounded.

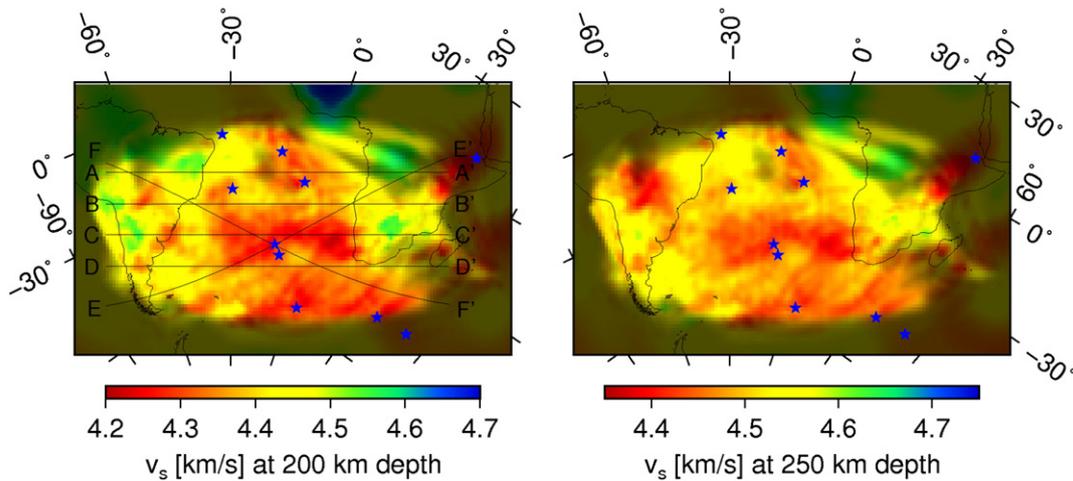
### 4.2. South America and Africa

Two very fast velocity anomalies in the South American lithosphere between 100 and 150 km (Fig. 6) depths are located under the Amazonia and São Francisco cratons in the northern part of the continent (see profile A–A' in Fig. 10), while a moderately fast and shallower anomaly is found under the Rio de la Plata craton (see profiles B–B' and C–C'). The southern part of the continent including the continental platform is instead underlain by moderately fast seismic anomalies and is divided from the northern faster regions by a slower east–west anomaly around 40° S (see profile D–D'). Deeper in the mantle the fast cratonic anomalies weaken and reduce in size, giving space to an anomaly starting from the coast off southern Brazil and extending more and more northwards under the continent with increasing depth. This slow anomaly branches off westward under Bolivia, reaching towards the border between Chile and Peru. A slow anomaly appears with depth also under the coast of Uruguay and Argentina, connecting the slow anomaly off the coast of southern Brazil with the slow anomaly around 40°S under Argentina. An isolated slow anomaly appears also near the eastern tip of Brazil, possibly due to the Fernando de Noronha hotspot. The Nazca slab appears as a set of discontinuous moderately fast regions under Peru, Chile and Argentina.

The lithosphere under Africa is also characterized by two large and fast anomalies corresponding to the Congo (see profiles A–A' and E–E') and Kalahari (see profiles C–C') cratons, while the Tanzania craton is underlain by a moderately fast anomaly only at shallow depths (see profiles B–B'). The centre of the fast anomaly corresponding to the



**Fig. 6.** Horizontal slices through the final tomographic model at 100 km depth (left) and 150 km depth (right). Oblique Mercator projection. At these depths continental cratons stand out as fast anomalies surrounded by relatively slower regions, while variations of seismic velocity in the oceanic realm are mainly due to the thermal cooling of the lithosphere, with fast anomalies near the continents varying smoothly into slower ones towards the mid-oceanic ridges. The main exception is given by the two slow anomalies extending from the Mid-Atlantic Ridge towards South America on one side and towards Africa on the other.



**Fig. 7.** Horizontal slices through the final tomographic model at 200 km depth (left) and 250 km depth (right). Black lines on the 200 km depth slice indicate the locations for the cross sections shown in Fig. 9. Oblique Mercator projection. The fast anomalies corresponding to continental cratons are less pronounced than at shallower depths, while slow velocities are overall dominant, in particular in the sub-oceanic mantle but protruding also significantly under the continents, winding between the cratons.

Kalahari craton shifts with depth towards the north-east, becoming less pronounced and disappearing completely below 250 km depth. The fast anomaly corresponding to the Congo craton instead is more stable and remains clearly visible even at 250 km and deeper. Below 300 km depth two other prominent slow anomalies are particularly noticeable, one under the Angola basin and the other under the gulf of Guinea extending all the way north under the West African craton.

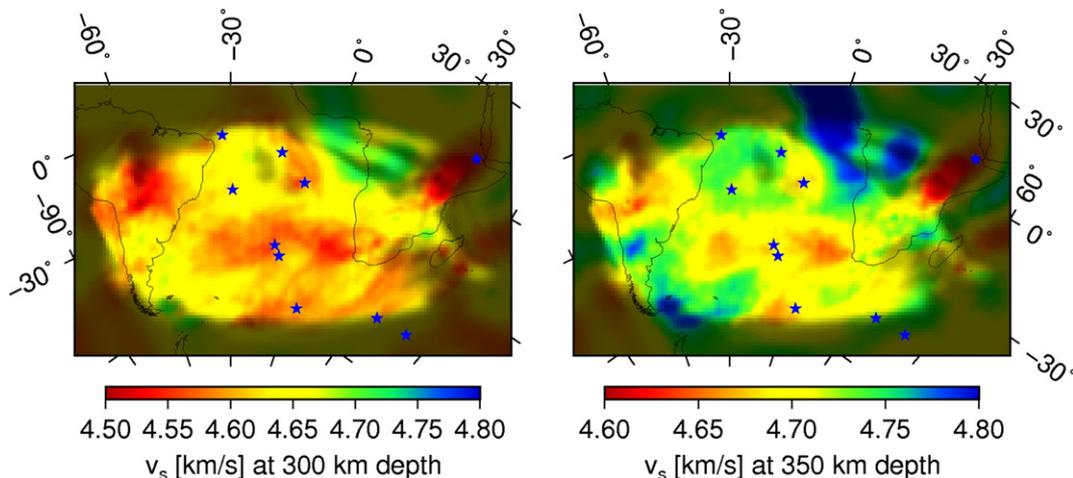
#### 4.3. South Atlantic Ocean

The seismic anomalies in the oceanic regions at 100 and 150 km depths are mainly correlated with the thermal cooling of the ageing lithosphere: the slowest material is at the mid-oceanic ridge and gets faster towards the continents. Two anomalously slow regions branch out from the Mid-Atlantic ridge heading north-east towards southern Africa and north-west towards Brazil. Interestingly, these anomalies are spatially quite well correlated with the Rio Grande Rise and the Walvis Ridge. The slow anomaly on the African side then bends then southwards all along Namibia and South Africa, connecting with a slow anomaly corresponding to the Cape Rise and the paleo-location of the Mid-Atlantic Ridge off the coast of South Africa at the time of the initial South Atlantic opening.

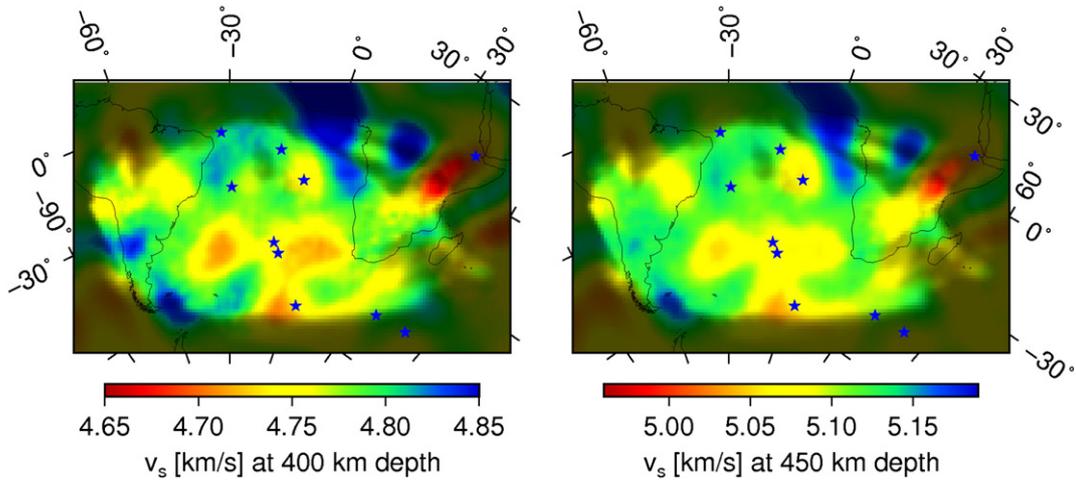
Below 150 km depth the lithospheric signature disappears and the north-south trending slow anomaly corresponding to the mid-oceanic ridge is replaced by an overall slow mantle streaked with an east-west slow anomaly extending from the Rio Grande Rise to the coast of Namibia. The slow velocity anomaly off the coast of Namibia becomes more pronounced, together with the slow anomaly under Bouvet island and the Atlantic-Antarctica ridge triple junction. The Argentine basin (see profiles E-E') and some parts of the Brazil basin (see profiles A-A') are instead characterized by relatively faster velocities throughout the upper 400 km of the mantle.

#### 4.4. Waveform fit

In Figs. 11–14 we show some examples of the increased waveform fit for some representative source-receiver configurations. During the first iteration we were able to use only a limited number of waveforms: the initial model was too fast under the ocean, and the synthetic waveforms for paths crossing the Atlantic were frequently too out of phase with the observed data, causing phase jumps in the misfit function and making the resulting adjoint source unusable. As the model was getting better, with further iterations, and the synthetics were becoming more similar to the observed data, we were able to exploit more seismograms and



**Fig. 8.** Horizontal slices through the final tomographic model at 300 km depth (left) and 350 km depth (right). Oblique Mercator projection. At these depths the pattern of slow velocity anomalies starts to change into round shaped slow regions in an overall fast mantle.

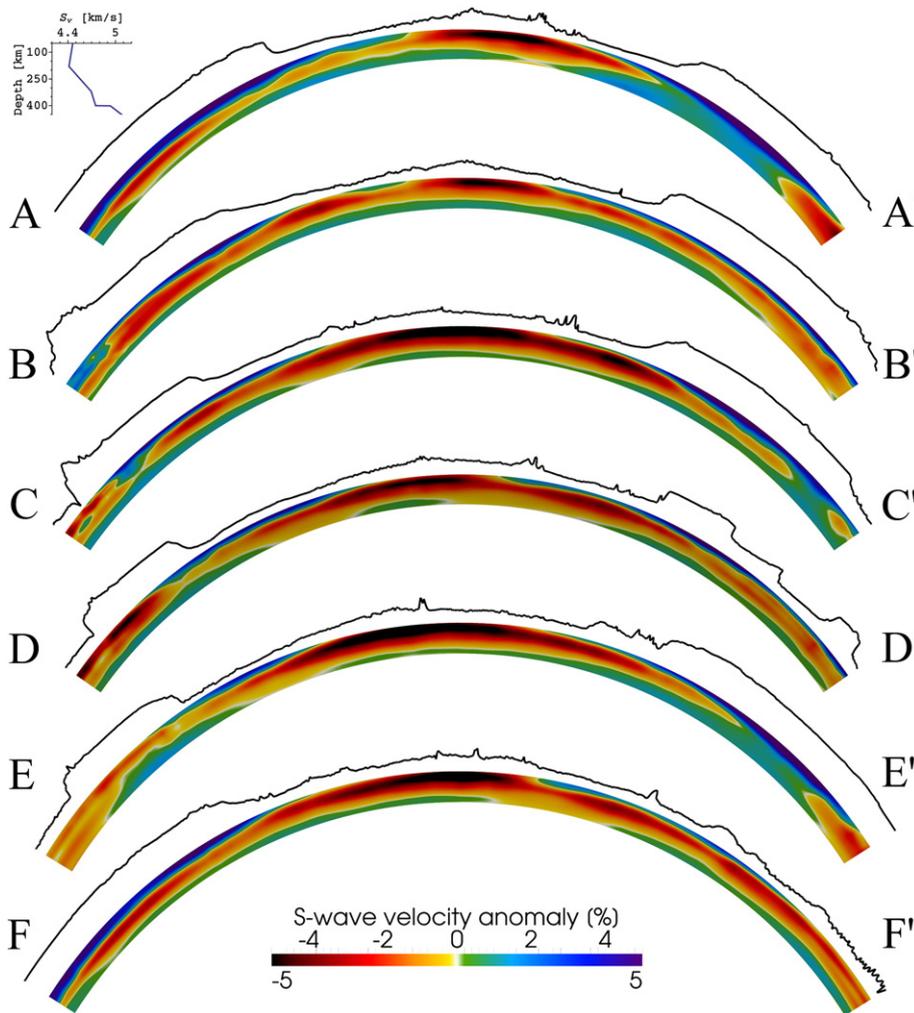


**Fig. 9.** Horizontal slices through the final tomographic model at 400 km depth (left) and 450 km depth (right). Oblique Mercator projection. Slow velocity anomalies are more rounded and embedded in an overall faster mantle.

to expand the time window, that is the function  $W(t,\omega)$  in Eq. (4), to include longer portions of the waveforms. As a result, the total time window length of the fitted portions has increased by a factor of 2, with a 50% misfit reduction over the original time windows.

4.5. Point spread functions

A point spread function (PSF) (Fichtner and Trampert, 2011) represents the sensitivity of a single adjoint iteration to a small localized



**Fig. 10.** Vertical profiles through the final tomographic model between 50 and 450 km depth. The upper 50 km have been removed since we are using a long wavelength equivalent crustal model. The thick black line on top of each profile is the (50× exaggerated) surface topography. The locations of the profiles are indicated in Fig. 7. Velocity variations are expressed as percent anomalies with respect to the 1-D S-wave velocity profile shown in the inset in the upper-left corner.

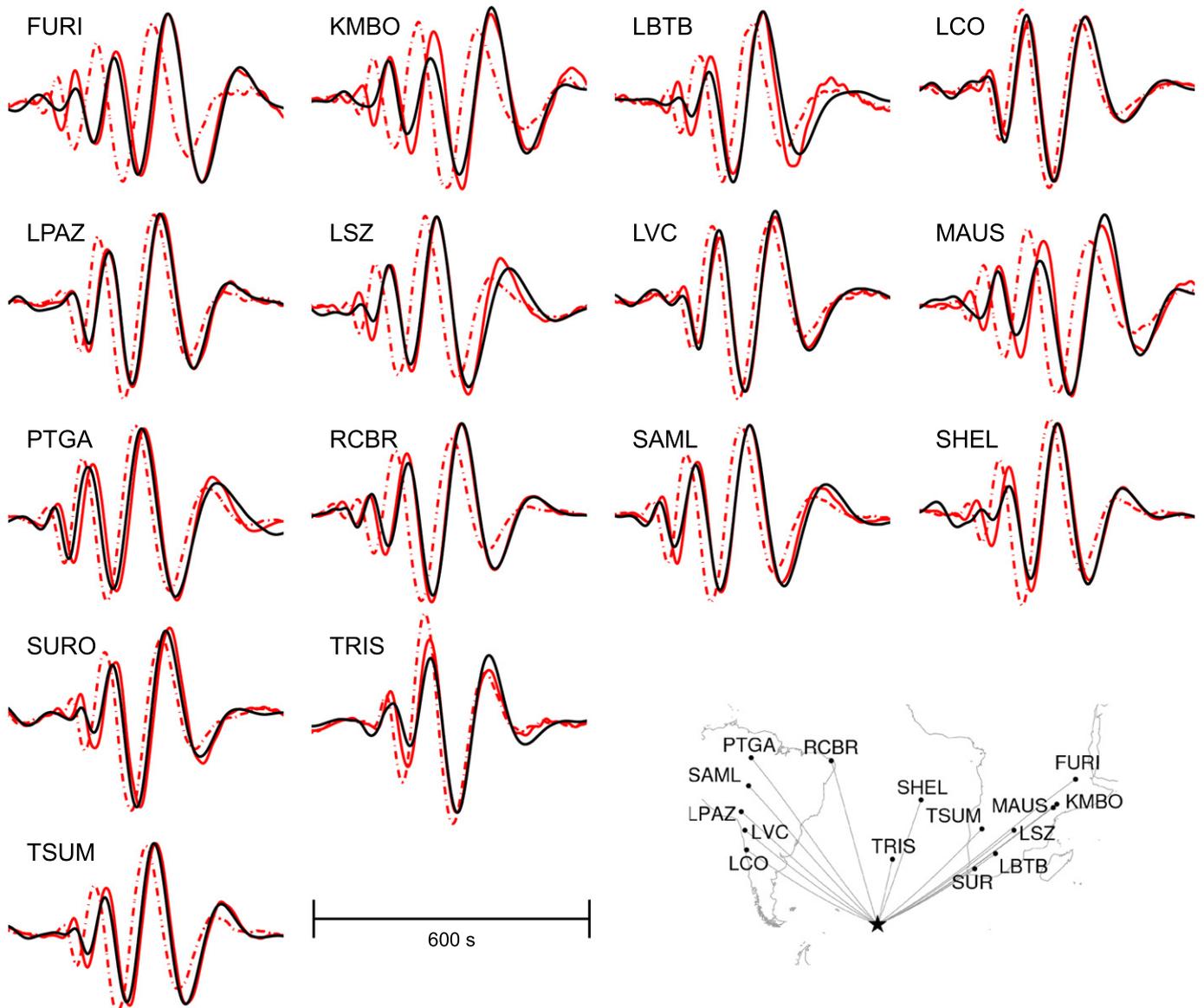


Fig. 11. Waveform comparison. Data are plotted as thick black lines, synthetics calculated with the final model as thick red lines, synthetics calculated with the initial model as dash-dotted red lines. Seismograms for an  $M$ , 6.8 event occurred on January 5th, 2010 east of the South Sandwich islands ( $58^{\circ}\text{S}$ ,  $15^{\circ}\text{W}$ ).

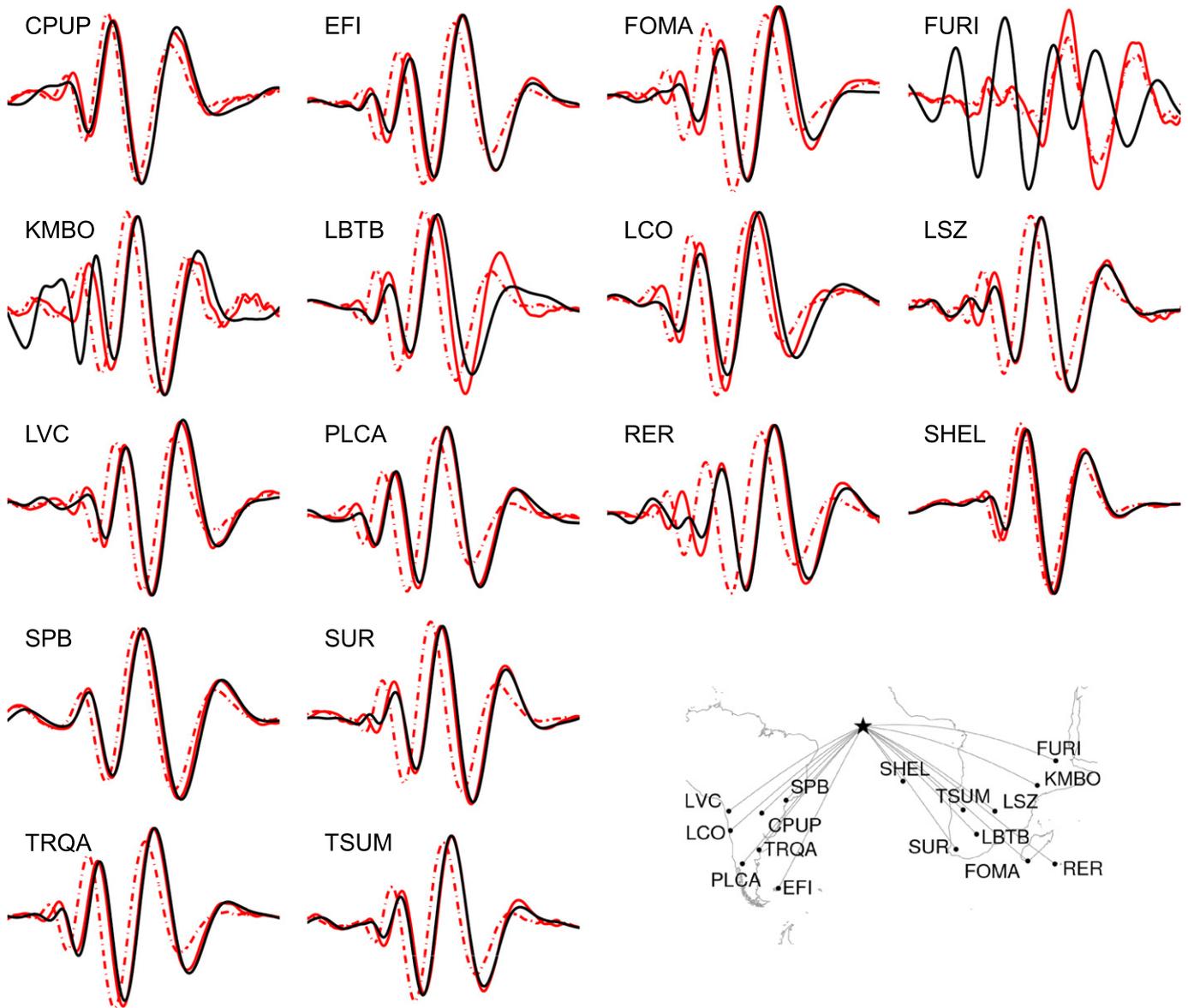
anomaly, and it can thus be used to assess the local resolution of the model and the eventual trade-offs with structures elsewhere in the model. In order to compute a PSF, a small anomaly is added to the current tomographic model and a single adjoint iteration is performed, obtaining the total derivative of the misfit function for this perturbed model. The PSF is then obtained by subtracting the total derivative of the unperturbed model from the total derivative of the perturbed model. Fig. 15 shows the three localized slow velocity anomalies used as input and the resulting PSF. The geometry of each anomaly is reconstructed reasonably well, although there is a consistent trade-off between negative sensitivity at the anomaly and positive sensitivity around it. The sensitivity decreases from 150 km depth to 305 km depth by about a factor of five.

## 5. Discussion

We have inverted a wide collection of broadband seismograms for isotropic seismic structure in the South Atlantic region. The full solution of the equations of motion for a realistic three dimensional

heterogeneous model through a numerical spectral-element technique ensures that differences between synthetic and recorded seismograms arise only from yet undiscovered Earth structure or errors in source parameters. The advantages of the full waveform method are particularly relevant in strongly heterogeneous portions of the Earth, such as the prominent ocean–continent–transitions of the South Atlantic region.

The number of events used in our inversion was limited by the available computational resources. The comparatively small number of seismograms is in part compensated by extracting as much waveform information as possible, e.g., through the application of the time frequency (TF) misfits. TF misfits as defined by Fichtner et al. (2008) and reviewed in Fichtner et al. (2009) offer several advantages: (1) they separate phase and envelope information, (2) apply to any type of seismic wave, (3) yield a quasi-linear relation to Earth structure and, most importantly, (4) exploit the complete waveform information. However, the phase misfit is meaningful only when observed and synthetic waveforms are sufficiently close to avoid phase jumps. We have picked and weighted the time windows manually to ensure that no phase jumps occur. An alternative



**Fig. 12.** Waveform comparison. Data are plotted as thick black lines, synthetics calculated with the final model as thick red lines, synthetics calculated with the initial model as dash-dotted red lines. Seismograms for an  $M_s$  6.4 event occurred on December 9th, 2009 at the Central Mid-Atlantic Ridge ( $1^\circ\text{S}$ ,  $21^\circ\text{W}$ ). Notice the complete difference between the observed and the modelled waveform for station FURI, that prevented the computation of a meaningful misfit and the improvement of the fit. Around 3% of the seismograms are not fitted well.

to manual selection has recently been proposed by Maggi et al. (2009).

An expected result is that the low velocity anomalies mapped in our inversion are spatially more extended and significantly slower than in the starting model, since their influence on wave propagation relative to raypath tomography is more pronounced when finite frequency effects, like waveform healing, are taken into account (Zhou et al., 2005).

The fast anomaly corresponding to the Congo craton is imaged down to great depth (see profiles A–A' and E–E' in Fig. 10). It is likely that amplitude and spatial extent of this feature are overestimated: the paucity of waves crossing the area (see Fig. 5) and the strong difference between synthetic and observed waveforms (see Fig. 12) prevented us from extracting robust structural information for this region, so that the initial model remains uncorrected in the inversion. Neighbouring stations are too few and too distant to build up significant sensitivity in this region. This is true both at shallower depth,

above 150 km, where the sensitivity of surface waves is restricted to a narrower Fresnel zone, and deeper, below 300 km, where increased width of the Fresnel zone is mitigated by lower overall sensitivity of the wavefield. A comparison with other seismic studies of the region (Begg et al., 2009; Reusch et al., 2010; Sebai et al., 2006) shows that the anomaly is probably due to a depthwise smearing towards the ocean of the overlying fast regions. Support for this interpretation comes from the fact that the waveform fit for seismograms crossing two unfavourable, outlying regions – north Central Africa and the Gulf of Guinea – improved only marginally (see e.g., Fig. 12, station FURI).

### 5.1. Comparison with previous tomographic studies

In the well-resolved parts of our computational domain, our model shows excellent agreement with previous regional tomographic studies for the South American (Feng et al., 2007; Heintz et al., 2005) and the

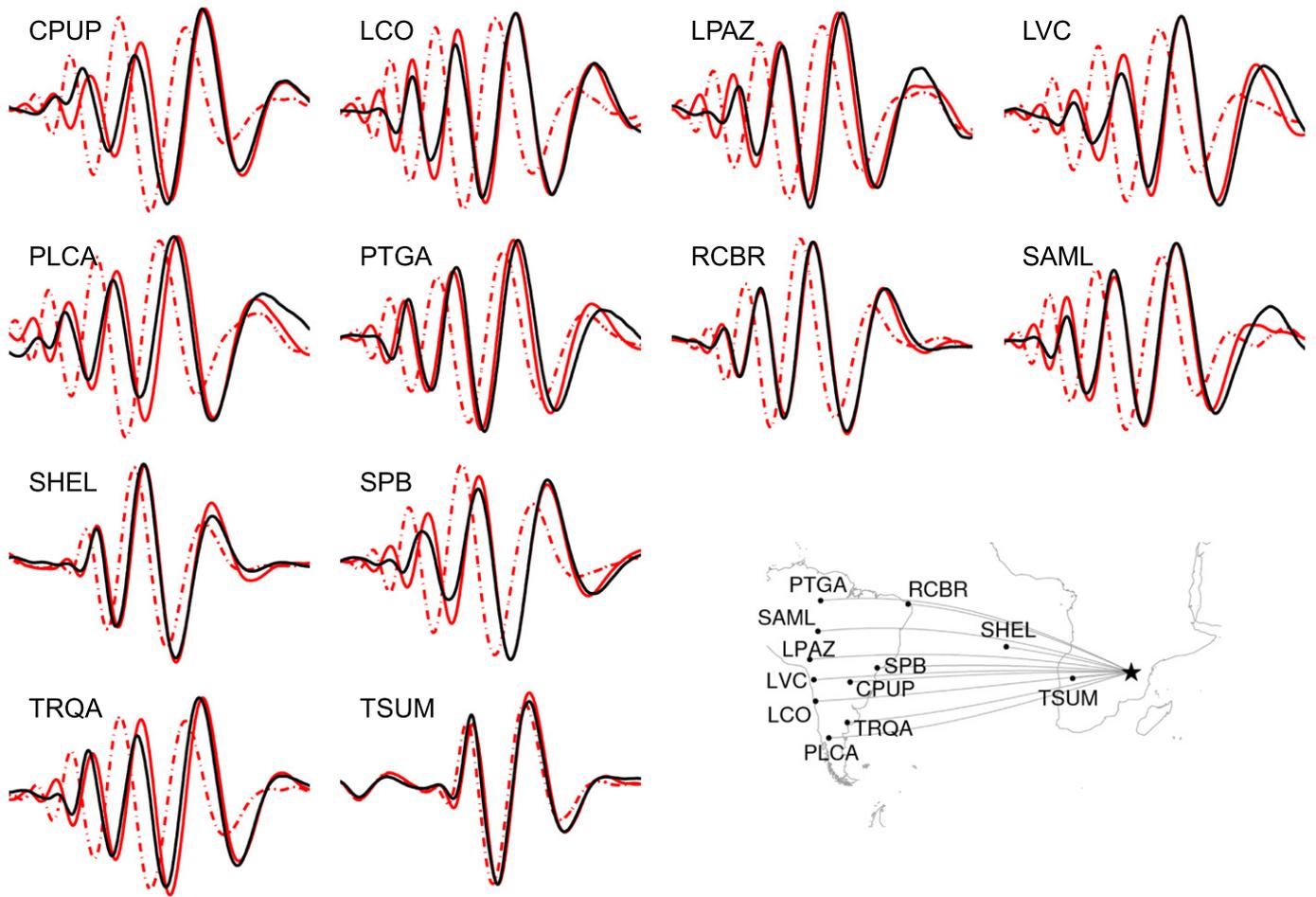


Fig. 13. Waveform comparison. Data are plotted as thick black lines, synthetics calculated with the final model as thick red lines, synthetics calculated with the initial model as dash-dotted red lines. Seismograms for an  $M_s$  5.9 event occurred on December 8th, 2009 in Tanzania ( $10^\circ\text{S}$ ,  $34^\circ\text{E}$ ).

African region (Begg et al., 2009; Chevrot and Zhao, 2007; Fishwick, 2010; Sebai et al., 2006). Discrepancies at smaller length scales are probably caused by differences in methodology and number and types of seismic waveform used. Since there are no regional studies of the oceanic region, we can compare our model only to global S wave tomography models (Grand et al., 1997; Lekić and Romanowicz, 2011; Masters et al., 1999; Mégnin and Romanowicz, 1999). We find a good agreement with our results in the upper 250 km, while deeper down only the model SEMum by Lekić and Romanowicz (2011) presents features similar to those of our model. In particular, model SAW24B16 by Mégnin and Romanowicz (1999) is characterized by a fast velocity anomaly halfway between South America and Africa at 350 km depth that is absent in our model.

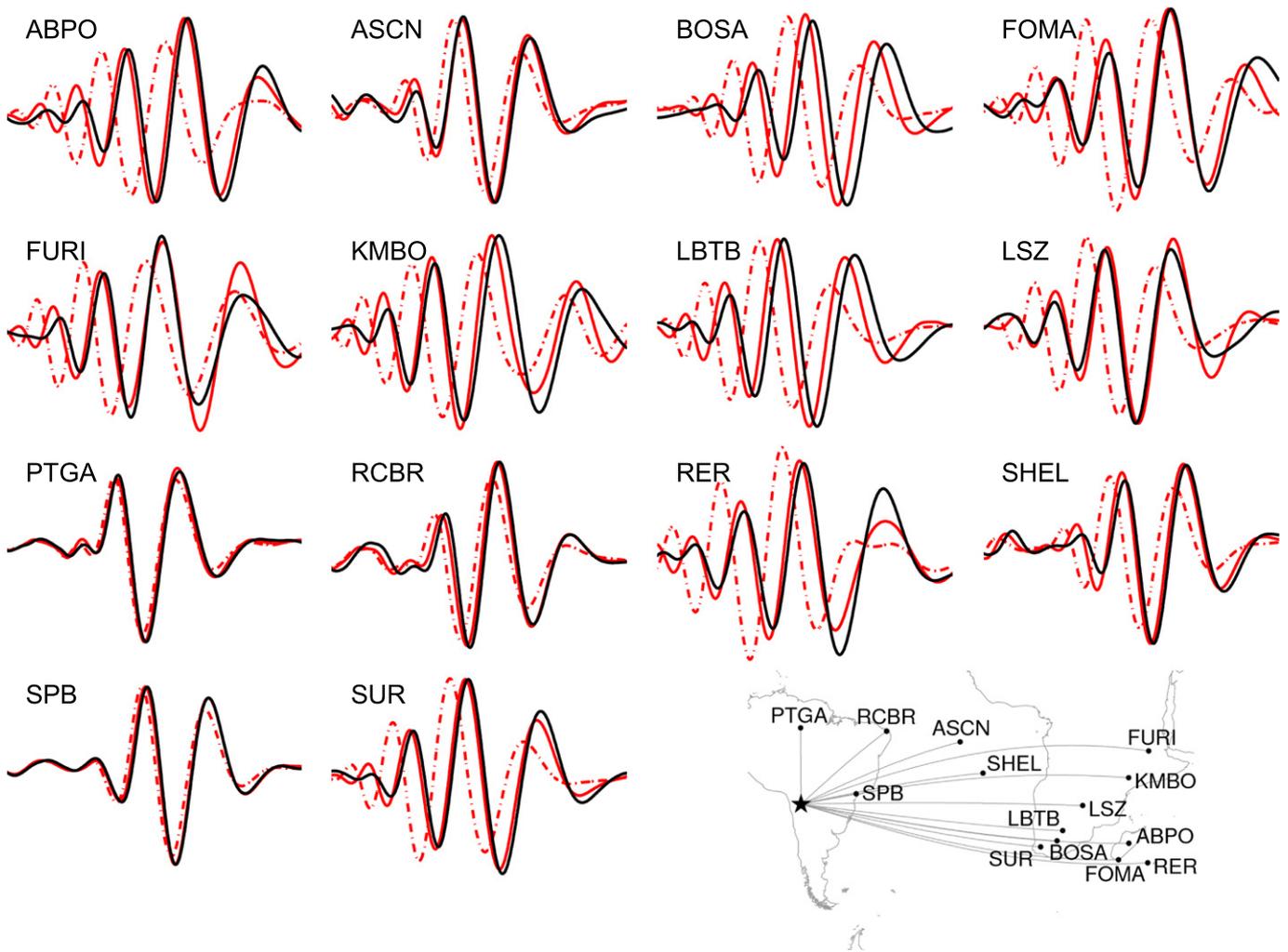
The initial starting model for our iterative inversion is taken from Ritsema's s20rts (Ritsema and van Heijst, 2004) global seismic study, as mentioned before. The same working group recently published a new global tomography model, s40rts (Ritsema et al., 2011). While based on data types and inversion method similar to s20rts, it takes advantage of an extended data set to obtain a higher resolution. There is a good correlation in the upper 250 km, although our model presents broader slow velocity anomalies.

## 5.2. Geodynamic implications

Since this is a preliminary model obtained using only long period waves the model resolution is low, and the predominance of east-

west trending paths is likely to cause some artefacts in the shape of the anomalies. For this reason we give here a possible interpretation for its broadest and best constrained features only.

The 3-D heterogeneity model for the South Atlantic region is characterized by a strong radial change in the seismic anomaly pattern (see Figs. 6–10): down to 150 km depth the imaged anomalies reflect the continental and oceanic lithospheric structure; between  $\approx 150$  and  $\approx 300$  km depth the only lithospheric features are given by the fast anomalies associated with cratonic roots, while the mantle is occupied by seismically slower material assuming horizontally elongated forms; deeper than  $\approx 350$  km the mantle is characterized by vertically oriented, round-shaped slow anomalies embedded in a faster mantle. This variation in heterogeneity may reflect a depthwise change in the convective planform, with hot buoyant mantle material supplied from a greater depth dominating horizontal upper mantle advection in a narrow channel reminiscent of the asthenosphere. A variety of evidence from seismology (Dziewonski and Anderson, 1981), rock mechanics (Karato and Wu, 1993), the geoid (Hager and Richards, 1989), convective planform studies (Bunge et al., 1996; Busse et al., 2006) and post-glacial rebound (Mitrovica, 1996) strongly argues for the existence of a mechanically weak upper mantle layer. But inherent trade-offs between layer thickness and viscosity reduction make it difficult to constrain the depth extent of such a layer on the ground of geodynamic observations alone (Paulson et al., 2007; Schaber et al., 2009). Although our model doesn't have enough resolution at depth to pinpoint the thickness



**Fig. 14.** Waveform comparison. Data are plotted as thick black lines, synthetics calculated with the final model as thick red lines, synthetics calculated with the initial model as dash-dotted red lines. Seismograms for an  $M_s$  6.2 event occurred on November 14th, 2009 in Jujuy, Argentina ( $23^\circ\text{S}$ ,  $67^\circ\text{W}$ ).

of the layer of low seismic speed nor how sharply it ends, it favours a relatively small (i.e., between 150 and 300 km) rather than a large (i.e., 400 km or more) thickness. Our results are compatible with a limited depth of the asthenosphere for the South Atlantic region.

A thin asthenosphere requires a stronger viscosity reduction with respect to the lower mantle to fit post-glacial rebound data (Mitrovica, 1996). This is bound to have a strong impact on upper mantle dynamics. In particular, it implies a greater mobility for material in the asthenosphere: hot buoyant material rising from the lower mantle can spread horizontally at a much faster rate than the overlying plate, keeping the asthenospheric layer relatively hot and homogeneous, and thus generating the type of seismic structure that we see in our model.

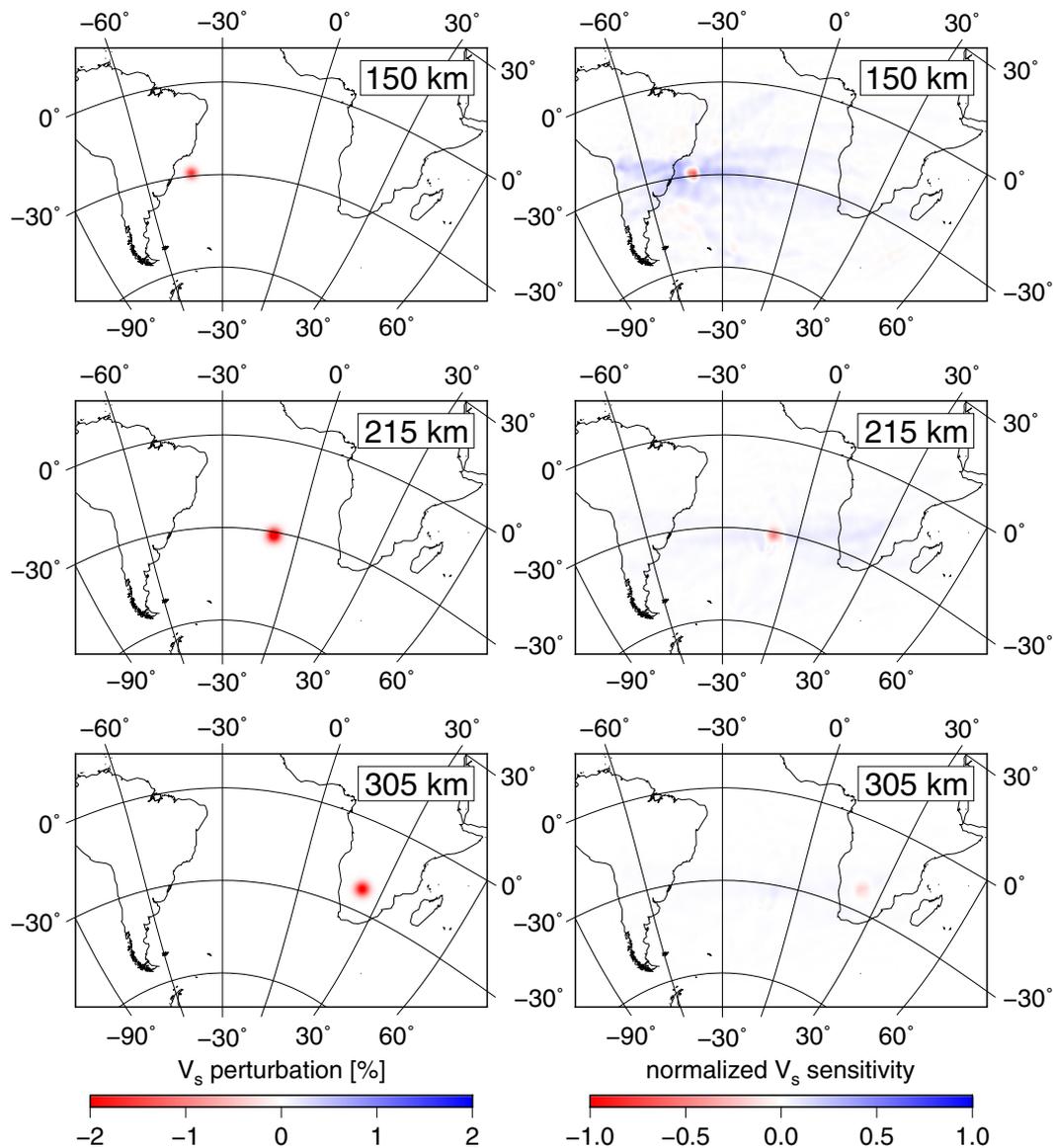
Likely sources of hot buoyant material to supply the South Atlantic upper mantle are a number of primary hotspots in the region (Courtilot et al., 2003). These hotspots are located within an area of elevated topography known as the African superswell (Nyblade and Robinson, 1994). The superswell is characterized by a dynamically supported topography of up to  $\approx 1$  km, which contrasts sharply with a significant negative dynamic topography of up to  $\approx 1$  km inferred for the conjugate South American margin (e.g., Winterbourne et al., 2009). The pronounced dynamic topography gradient across

the South Atlantic region implies substantial westward directed channel flow as suggested by Behn et al. (2004) and advocated by Phipps Morgan et al. (1995) from a variety of geodynamic and geochemical considerations. We note that significant basal shear tractions arising from westward fluxing sub-Atlantic asthenosphere would help to balance the budget of driving and resisting forces acting on the South American plate by balancing the gravitational load associated with the Andes (Husson et al., 2012; Iaffaldano and Bunge, 2009).

We assess the dynamical feasibility of westward fluxing asthenosphere across the South Atlantic region from a simple fluid dynamic model for pressure-driven flow in a low viscosity channel, where the asthenosphere channel is approximated as a viscous fluid confined between two fixed, horizontal infinite planes separated by distance  $h$ . The excess velocity due to the horizontal pressure gradient  $d p/d x$  that arises in the channel varies parabolically with depth (Landau and Lifshitz, 1987) such that the mean fluid velocity is:

$$\bar{v} = -\frac{h^2}{12\eta} \frac{dp}{dx}, \quad (5)$$

where  $\eta$  is the viscosity. An estimate of the pressure gradient across the South Atlantic can be obtained from the oceanic residual topography



**Fig. 15.** Horizontal slices showing the input S-wave velocity anomaly (left column) and the computed PSF sensitivity (right column) for three small anomalies localized at 150 km depth off the coast of Brazil (top row), at 215 km depth below Tristan da Cunha (middle row) and at 305 km depth below southern Africa (bottom row).

(Fig. 1) difference of up to  $\approx 2$  km between the Argentine Basin and that of the Cape Basin, from which we infer a dynamically maintained pressure difference of  $\approx 60$  MPa over a distance of  $\approx 5000$  km. This yields a velocity of  $\approx 12.5$  cm/yr for an asthenosphere of a viscosity of  $10^{19}$  Pa  $\cdot$  s and a channel thickness  $h$  of 200 km. A thicker layer ( $h = 500$  km) would instead be characterized by a higher viscosity ( $\eta = 3 \cdot 10^{20}$  Pa  $\cdot$  s) and the mean fluid velocity would thus be much smaller ( $\approx 2.6$  cm/yr).

The pressure gradient across the South Atlantic Ocean can thus produce a significant excess velocity in a thin and low-viscosity asthenosphere with respect to both the overlying plate and the underlying mantle. Hot buoyant material rising to the upper mantle under southern Africa can thus spread westward at a fast rate, keeping the asthenosphere relatively homogeneous and considerably warmer than the ambient mantle. This would in turn result in a layer characterized by widespread slow seismic velocities. It is thus possible for the layer of slow seismic velocity imaged by our seismic tomography to be primarily thermal in nature and due to fast horizontal advection of hot material rising from deeper in the mantle.

## 6. Conclusions

We have successfully obtained a full waveform tomography for isotropic upper-mantle structure of a large and sparsely sampled region. This result has been achieved thanks to an accurate numerical solution of the elastic wave equation for 3-D heterogeneous media, together with a careful choice of the misfit function and the use of the adjoint method. We have thus been able to avoid the artefacts generated by common simplifying approximations and to extract the maximum amount of information from each seismogram, thus compensating for the paucity of data.

The resulting tomographic image shows a broad region in the shallow upper mantle characterized by particularly low velocities that may reflect the dynamics of a thin and very mobile asthenosphere. In order to better assess the spatial extension of this region and its geodynamical significance we plan to extend the frequency window up to 33 mHz (30 s period). This will allow us not only to use seismic waves characterized by a different spatial sensitivity, but also to exploit a larger number of seismograms, thanks to the more favourable signal-to-noise ratio at these higher frequencies.

## Acknowledgments

This study is funded by the German Research Foundation (DFG) within the priority program SPP 1375 - SAMPLE. We would like to thank Brian Kennett for many useful discussions and insightful advices.

## Appendix A. The adjoint equation

We present here only a brief summary of the theory of the adjoint method applied to seismology; those interested in more details are referred to Fichtner et al. (2006, 2008, 2009).

Let us assume we are confident that the recorded elastic displacement  $\mathbf{u}^0$  can be described as the solution of the wave equation, which we will denote via the shorthand  $\mathbf{L}$ , in a model Earth described by the set of parameters  $\mathbf{p}$  in the presence of an external forcing  $\mathbf{g}_0$ :

$$\mathbf{L}(\mathbf{u}, \mathbf{p}) = \rho(\mathbf{x}) \partial_t^2 \mathbf{u}(\mathbf{x}, t) - \nabla \cdot \int_{-\infty}^t \mathbf{C}(\mathbf{x}, t - \tau) : \nabla \mathbf{u}(\mathbf{x}, \tau) \, d\tau = \mathbf{g}_0(\mathbf{x}, t). \quad (\text{A.1})$$

The left hand side contains the physics of the system, while the right hand side  $\mathbf{g}_0$  is the particular earthquake that gave rise to that particular seismogram,  $\mathbf{u}^0$ . The model parameters  $\mathbf{p}$  comprise the mass density  $\rho$  and the elastic parameters contained in the relaxation tensor  $\mathbf{C}$ . Solving the inverse problem means to find the particular value of the parameters  $\mathbf{p}_{\min}$  for which the solution  $\mathbf{u}_{\min}$  of the differential equation is closest to the observable  $\mathbf{u}^0$ .

We start from an initial guess for the model parameters that produce an initial mismatched solution, and we then iteratively correct the model parameters by descending the gradient of the misfit function  $\chi$ . The adjoint method provides an elegant way of computing the gradient and is particularly efficient when the parameter space is very large. Assuming that the misfit function  $\chi$  can be expressed as a bilinear form, i.e.,

$$\chi(\mathbf{u}, \mathbf{p}) = \langle \mathbf{1}, f(\mathbf{u}, \mathbf{p}) \rangle, \quad (\text{A.2})$$

we can rewrite Eq. (1) as

$$D_p \chi(\mathbf{u}, \mathbf{p})(\mathbf{d}) = \langle \mathbf{d}, \partial_p f(\mathbf{u}, \mathbf{p}) \rangle + \langle \mathbf{v}, \partial_u f(\mathbf{u}, \mathbf{p}) \rangle, \quad (\text{A.3})$$

where  $\mathbf{v} = D_p \mathbf{u}(\mathbf{d})$ . Differentiating (A.1) with respect to  $\mathbf{p}$  yields

$$D_p \mathbf{L}(\mathbf{u}; \mathbf{p})(\mathbf{d}) = \partial_u \mathbf{L}(\mathbf{u}; \mathbf{p})(\mathbf{v}) + \partial_p \mathbf{L}(\mathbf{u}; \mathbf{p})(\mathbf{d}) = 0, \quad (\text{A.4})$$

since the r.h.s. of Eq. (A.1) doesn't depend on the model parameters. Taking the scalar product with an arbitrary but sufficiently regular test function  $\psi$ , we find that

$$\langle \partial_u \mathbf{L}(\mathbf{u}; \mathbf{p})(\mathbf{v}), \psi \rangle + \langle \partial_p \mathbf{L}(\mathbf{u}; \mathbf{p})(\mathbf{d}), \psi \rangle = 0. \quad (\text{A.5})$$

Assuming that there exist two transposed operators  $\partial_u \mathbf{L}^*(\mathbf{u}; \mathbf{p})$  and  $\partial_p \mathbf{L}^*(\mathbf{u}; \mathbf{p})$  such that

$$\langle \partial_u \mathbf{L}(\mathbf{u}; \mathbf{p})(\mathbf{v}), \psi \rangle = \langle \mathbf{v}, \partial_u \mathbf{L}^*(\mathbf{u}; \mathbf{p})(\psi) \rangle, \quad (\text{A.6})$$

$$\langle \partial_p \mathbf{L}(\mathbf{u}; \mathbf{p})(\mathbf{d}), \psi \rangle = \langle \mathbf{d}, \partial_p \mathbf{L}^*(\mathbf{u}; \mathbf{p})(\psi) \rangle, \quad (\text{A.7})$$

then Eq. (A.5) can be rewritten as

$$\langle \mathbf{v}, \partial_u \mathbf{L}^*(\mathbf{u}; \mathbf{p})(\psi) \rangle + \langle \mathbf{d}, \partial_p \mathbf{L}^*(\mathbf{u}; \mathbf{p})(\psi) \rangle = 0. \quad (\text{A.8})$$

Finding an explicit expression for the transpose operators can be difficult, and may require some conditions on the test function  $\psi$ . In case of the wave equation, its adjoint equation is again of the wave

type but reversed in time, and the test function  $\psi$  is subject to boundary conditions that are dual with respect to the boundary conditions on  $\mathbf{u}$ , and to terminal conditions instead of initial conditions.

Adding together Eqs. (A.3) and (A.8) yields

$$D_p \chi(\mathbf{u}, \mathbf{p})(\mathbf{d}) = \langle \mathbf{d}, \partial_p f(\mathbf{u}, \mathbf{p}) + \partial_p \mathbf{L}^*(\mathbf{u}; \mathbf{p})(\psi) \rangle + \langle \mathbf{v}, \partial_u f(\mathbf{u}, \mathbf{p}) + \partial_u \mathbf{L}^*(\mathbf{u}; \mathbf{p})(\psi) \rangle. \quad (\text{A.9})$$

The term  $\mathbf{v} = D_p \mathbf{u}(\mathbf{d})$  may now be eliminated by choosing that particular test function  $\psi$  that satisfies

$$\partial_u f(\mathbf{u}, \mathbf{p}) + \partial_u \mathbf{L}^*(\mathbf{u}; \mathbf{p})(\psi) = 0, \quad (\text{A.10})$$

which is called the adjoint equation; therefore the total derivative of the objective function with respect to the model parameters takes the form

$$D_p \chi(\mathbf{u}, \mathbf{p})(\mathbf{d}) = \langle \mathbf{d}, \partial_p f(\mathbf{u}, \mathbf{p}) + \partial_p \mathbf{L}^*(\mathbf{u}; \mathbf{p})(\psi) \rangle. \quad (\text{A.11})$$

The adjoint method consists thus in obtaining the transpose operators for the physical problem at hand. One then finds a solution of the adjoint problem, that is a solution of the adjoint equation plus the corollary conditions on the test function, and finally calculates the total derivative of the objective function relative to the model parameters.

## Appendix B. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.tecto.2013.06.015>.

## References

- Amante, C., Eakins, B.W., 2009. ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24.
- Anderson, D.L., 1982. Hotspots, polar wander, Mesozoic convection and the geoid. *Nature* 297, 391–393.
- Begg, G.C., Griffin, W.L., Natapov, L.M., O'Reilly, S.Y., Grand, S.P., O'Neill, C.J., Hronsky, J.M.A., Djomani, Y.P., Swain, C.J., Deen, T., Bowden, P., 2009. The lithospheric architecture of Africa: seismic tomography, mantle petrology, and tectonic evolution. *Geosphere* 5, 23–50.
- Behn, M.D., Conrad, C.P., Silver, P.G., 2004. Detection of upper mantle flow associated with the African Superplume. *Earth and Planetary Science Letters* 224, 259–274.
- Bunge, H.P., Richards, M.A., Baumgardner, J.R., 1996. Effect of depth-dependent viscosity on the planform of mantle convection. *Nature* 379, 436–438.
- Bunge, H.P., Hagelberg, C.R., Travis, B.J., 2003. Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography. *Geophysical Journal International* 152, 280–301.
- Burke, K., Gunnell, Y., 2008. The African erosion surface: a continental-scale synthesis of geomorphology, tectonics, and environmental change over the past 180 million years. *Memoir (Geological Society of America)*. Geological Society of America 201, 1–66.
- Busse, F.H., Richards, M.A., Lenardic, A., 2006. A simple model of high Prandtl and high Rayleigh number convection bounded by thin low-viscosity layers. *Geophysical Journal International* 164, 160–167.
- Chen, P., Zhao, L., Jordan, T.H., 2007. Full 3D tomography for the crustal structure of the Los Angeles region. *Bulletin of the Seismological Society of America* 97, 1094–1120.
- Chevrot, S., Zhao, L., 2007. Multiscale finite-frequency Rayleigh wave tomography of the Kaapvaal craton. *Geophysical Journal International* 169, 201–215.
- Courtillot, V., Davaille, A., Besse, J., Stock, J., 2003. Three distinct types of hotspots in the Earth's mantle. *Earth and Planetary Science Letters* 205, 295–308.
- Crosby, A.G., McKenzie, D., Sclater, J.G., 2006. The relationship between depth, age and gravity in the oceans. *Geophysical Journal International* 166, 553–573.
- Debayle, E., Kennett, B., Priestley, K., 2005. Global azimuthal seismic anisotropy and the unique plate-motion deformation of Australia. *Nature* 433, 509–512.
- Dziewonski, A.M., Anderson, D.L., 1981. Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors* 25, 297–356.
- Feng, M., van der Lee, S., Assumpção, M., 2007. Upper mantle structure of South America from joint inversion of waveforms and fundamental mode group velocities of Rayleigh waves. *Journal of Geophysical Research* 112.
- Fichtner, A., Igel, H., 2008. Efficient numerical surface wave propagation through the optimization of discrete crustal models — a technique based on non-linear dispersion curve matching (DCM). *Geophysical Journal International* 173, 519–533.
- Fichtner, A., Trampert, J., 2011. Resolution analysis in full waveform inversion. *Geophysical Journal International* 187, 1604–1624.

- Fichtner, A., Bunge, H.P., Igel, H., 2006. The adjoint method in seismology I. Theory. *Physics of the Earth and Planetary Interiors* 157, 86–104.
- Fichtner, A., Kennett, B.L.N., Igel, H., Bunge, H.P., 2008. Theoretical background for continental- and global-scale full-waveform inversion in the time–frequency domain. *Geophysical Journal International* 175, 665–685.
- Fichtner, A., Kennett, B.L.N., Igel, H., Bunge, H.P., 2009. Full seismic waveform tomography for upper-mantle structure in the australasian region using adjoint methods. *Geophysical Journal International* 179, 1703–1725.
- Fishwick, S., 2010. Surface wave tomography: Imaging of the lithosphere–asthenosphere boundary beneath central and southern Africa? *Lithos* 120, 63–73.
- Forté, A.M., Quéré, S., Moucha, R., Simmons, N.A., Grand, S.P., Mitrovića, J.X., Rowley, D.B., 2010. Joint seismic–geodynamic–mineral physical modelling of African geodynamics: a reconciliation of deep-mantle convection with surface geophysical constraints. *Earth and Planetary Science Letters* 295, 329–341.
- Fournier, A., Hulot, G., Jault, D., Kuang, W., Tangborn, A., Gillet, N., Canet, E., Aubert, J., Lhuillier, F., 2010. An introduction to data assimilation and predictability in geomagnetism. *Space Science Reviews* 155, 247–291.
- Freybourger, M., Gaherty, J.B., Jordan, T.H., Kaapvaal Seismic Grp, 2001. Structure of the Kaapvaal craton from surface waves. *Geophysical Research Letters* 28, 2489–2492.
- Grand, S.P., 2002. Mantle shear-wave tomography and the fate of subducted slabs. *Philosophical Transactions of the Royal Society A* 360, 2475–2491.
- Grand, S.P., van der Hilst, R.D., Widiyantoro, S., 1997. Global seismic tomography: a snapshot of convection in the earth. *GSA Today* 7, 1–7.
- Gurnis, M., Mitrovića, J.X., Ritsema, J., van Heijst, H.J., 2000. Constraining mantle density structure using geological evidence of surface uplift rates: the case of the African superplume. *Geochemistry, Geophysics, Geosystems* 1.
- Hager, B.H., Richards, M.A., 1989. Long-wavelength variations in Earth's geoid: physical models and dynamical implications. *Philosophical Transactions of the Royal Society A* 328, 309–327.
- Heintz, M., Debayle, E., Vauchez, A., 2005. Upper mantle structure of the South American continent and neighboring oceans from surface wave tomography. *Tectonophysics* 406, 115–139.
- Husson, L., Conrad, C.P., Faccenna, C., 2012. Plate motions, Andean orogeny, and volcanism above the South Atlantic convection cell. *Earth and Planetary Science Letters* 317–318, 126–135.
- Iaffaldano, G., Bunge, H.P., 2009. Relating rapid plate-motion variations to plate-boundary forces in global coupled models of the mantle/lithosphere system: effects of topography and friction. *Tectonophysics* 474, 393–404.
- Iaffaldano, G., Bunge, H.P., Buckler, M., 2007. Mountain belt growth inferred from histories of past plate convergence: a new tectonic inverse problem. *Earth and Planetary Science Letters* 260, 516–523.
- Karato, S.I., Wu, P., 1993. Rheology of the upper mantle: a synthesis. *Science* 260, 771–778.
- Landau, L.D., Lifshitz, E.M., 1987. *Fluid Mechanics*, Second edition. Course of Theoretical Physics, 6. Butterworth-Heinemann.
- Lekić, V., Romanowicz, B., 2011. Inferring upper-mantle structure by full waveform tomography with the spectral element method. *Geophysical Journal International* 185, 799–831.
- Lithgow-Bertelloni, C., Silver, P.G., 1998. Dynamic topography, plate driving forces and the African superswell. *Nature* 395, 269–272.
- Maggi, A., Tape, C., Chen, M., Chao, D., Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography. *Geophysical Journal International* 178, 257–281.
- Masters, G., Bolton, H., Laske, G., 1999. Joint seismic tomography for p and s velocities: how pervasive are chemical anomalies in the mantle? *Eos, Transactions, American Geophysical Union* 80.
- Mégnin, C., Romanowicz, B., 1999. The effects of the theoretical formalism and data selection on mantle models derived from waveform tomography. *Geophysical Journal International* 138, 366–380.
- Meier, U., Curtis, A., Trampert, J., 2007. Fully nonlinear inversion of fundamental mode surface waves for a global crustal model. *Geophysical Research Letters* 34.
- Mitrovića, J.X., 1996. Haskell [1935] revisited. *Journal of Geophysical Research* 101, 555–569.
- Moucha, R., Forté, A.M., 2011. Changes in African topography driven by mantle convection. *Nature Geoscience* 4, 707–712.
- Müller, R.D., Sdrolias, M., Gaina, C., Roest, W.R., 2008. Age, spreading rates, and spreading asymmetry of the world's ocean crust. *Geochemistry, Geophysics, Geosystems* 9.
- Nyblade, A.A., Robinson, S.W., 1994. The African superswell. *Geophysical Research Letters* 21, 765–768.
- Oeser, J., Bunge, H.P., Mohr, M., 2006. Cluster design in the Earth sciences: TETHYS. In: Gerndt, M., Kranzlmüller, D. (Eds.), *High Performance Computing and Communications – Second International Conference, HPCC 2006*. Springer, Munich, Germany, pp. 31–40.
- Pasyanos, M.E., Nyblade, A.A., 2007. A top to bottom lithospheric study of Africa and Arabia. *Tectonophysics* 444, 27–44.
- Paulson, A., Zhong, S., Wahr, J., 2007. Limitations on the inversion for mantle viscosity from postglacial rebound. *Geophysical Journal International* 168, 1195–1209.
- Phipps Morgan, J., Morgan, W.J., Zhang, Y.S., Smith, W.H.F., 1995. Observational hints for a plume-fed, suboceanic asthenosphere and its role in mantle convection. *Journal of Geophysical Research* 100, 12753–12767.
- Priestley, K., McKenzie, D., Debayle, E., 2006. The state of the upper mantle beneath southern Africa. *Tectonophysics* 416, 101–112.
- Priestley, K., McKenzie, D., Debayle, E., Piliidou, S., 2008. The African upper mantle and its relationship to tectonics and surface geology. *Geophysical Journal International* 175, 1108–1126.
- Reusch, A.M., Nyblade, A.A., Wiens, D.A., Shore, P.J., Ateba, B., Tabod, C.T., Nnange, J.M., 2010. Upper mantle structure beneath Cameroon from body wave tomography and the origin of the Cameroon volcanic line. *Geochemistry, Geophysics, Geosystems* 11.
- Ritsema, J., van Heijst, H.J., 2000. New seismic model of the upper mantle beneath Africa. *Geology* 28, 63–66.
- Ritsema, J., van Heijst, H.J., 2004. Global transition zone tomography. *Journal of Geophysical Research* 109.
- Ritsema, J., van Heijst, H.J., Woodhouse, J.H., 1999. Complex shear wave velocity structure imaged beneath Africa and Iceland. *Science* 286, 1925–1928.
- Ritsema, J., Deuss, A., van Heijst, H.J., Woodhouse, J.H., 2011. S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltimes and normal-mode splitting function measurements. *Geophysical Journal International* 3, 1223–1236.
- Schaber, K., Bunge, H.P., Schubert, B., Malservisi, R., Horbach, A., 2009. Stability of the rotation axis in high-resolution mantle circulation models: weak polar wander despite strong core heating. *Geochemistry, Geophysics, Geosystems* 10.
- Schubert, B.S.A., Bunge, H.P., Ritsema, J., 2009a. Tomographic filtering of high-resolution mantle circulation models: can seismic heterogeneity be explained by temperature alone? *Geochemistry, Geophysics, Geosystems* 10.
- Schubert, B.S.A., Bunge, H.P., Steinle-Neumann, G., Moder, C., Oeser, J., 2009b. Thermal versus elastic heterogeneity in high-resolution mantle circulation models with pyrolytic composition: high plume excess temperatures in the lowermost mantle. *Geochemistry, Geophysics, Geosystems* 10.
- Sebai, A., Stutzmann, E., Montagner, J.P., Sicilia, D., Beucler, E., 2006. Anisotropic structure of the African upper mantle from Rayleigh and Love wave tomography. *Physics of the Earth and Planetary Interiors* 155, 48–62.
- Simmons, N.A., Forté, A.M., Grand, S.P., 2007. Thermochemical structure and dynamics of the African superplume. *Geophysical Research Letters* 34.
- Tape, C., Liu, Q., Maggi, A., Tromp, J., 2009. Adjoint tomography of the Southern California crust. *Science* 325, 988–992.
- Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation. *Pure and Applied Geophysics* 128, 365–399.
- Teixeira, F.L., Chew, W.C., 1997. Systematic derivation of anisotropic PML absorbing media in cylindrical and spherical coordinates. *IEEE Microwave and Guided Wave Letters* 7, 371–373.
- Tromp, J., Tape, C., Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana doughnut kernels. *Geophysical Journal International* 160, 195–216.
- Winterbourne, J., Crosby, A., White, N., 2009. Depth, age and dynamic topography of oceanic lithosphere beneath heavily sedimented Atlantic margins. *Earth and Planetary Science Letters* 287, 137–151.
- Zheng, Y., Huang, X., 1997. Anisotropic perfectly matched layers for elastic waves in Cartesian and curvilinear coordinates. Consortium Report. MIT Earth Resources Laboratory.
- Zhou, Y., Dahlen, F.A., Nolet, G., Laske, G., 2005. Finite-frequency effects in global surface-wave tomography. *Geophysical Journal International* 163, 1087–1111.