Geophys. J. Int. (2008) **175,** 598–616 doi: 10.1111/j.1365-246X.2008.03904.x

SPICE benchmark for global tomographic methods

Yilong Qin,[1] Yann Capdeville,[1] Valerie Maupin,[2] Jean-Paul Montagner,[1] Sergei Lebedev[3,4] and Eric Beucler[5]

[1] *Département de Sismologie, Institut de Physique du globe de Paris, 4, Place Jussieu, 75252 Paris, France. E-mail: qyl@ipgp.jussieu.fr*
[2] *Department of Geosciences, University of Oslo,* PO Box 1047 *Blindern,* 0316 *Oslo, Norway*
[3] *Department of Earth Sciences, Utrecht University, Budapestlaan* 4*, Utrecht,* 3584 *CD, the Netherlands*
[4] *Dublin Institute for Advanced Studies, Geophysics Section,* 5 *Merrion Square, Dublin* 2*, Ireland*
[5] *Université de Nantes, Nantes Atlantique Universités, CNRS-UMR* 6112*, Laboratoire de Planétologie et Géodynamique,* BP 92205*, Nantes, France*

Accepted 2008 June 27. Received 2008 June 23; in original form 2008 March 13

SUMMARY

The existing global tomographic methods result in different models due to different parametrization, scale resolution and theoretical approach. To test how current imaging techniques are limited by approximations in theory and by the inadequacy of data quality and coverage, it is necessary to perform a global-scale benchmark to understand the resolving properties of each specific imaging algorithm.

In the framework of the Seismic wave Propagation and Imaging in Complex media: a European network (SPICE) project, it was decided to perform a benchmark experiment of global inversion algorithms. First, a preliminary benchmark with a simple isotropic model is carried out to check the feasibility in terms of acquisition geometry and numerical accuracy. Then, to fully validate tomographic schemes with a challenging synthetic data set, we constructed one complex anisotropic global model, which is characterized by 21 elastic constants and includes 3-D heterogeneities in velocity, anisotropy (radial and azimuthal anisotropy), attenuation, density, as well as surface topography and bathymetry. The intermediate-period ($>$32 s), high fidelity anisotropic modelling was performed by using state-of-the-art anisotropic anelastic modelling code, that is, coupled spectral element method (CSEM), on modern massively parallel computing resources. The benchmark data set consists of 29 events and three-component seismograms are recorded by 256 stations. Because of the limitation of the available computing power, synthetic seismograms have a minimum period of 32s and a length of 10500s.

The inversion of the benchmark data set demonstrates several well-known problems of classical surface wave tomography, such as the importance of crustal correction to recover the shallow structures, the loss of resolution with depth, the smearing effect, both horizontal and vertical, the inaccuracy of amplitude of isotropic *S*-wave velocity variation, the difficulty of retrieving the magnitude of azimuthal anisotropy and so on.

The synthetic data set can be used to validate and calibrate new processing methodologies and has been made available to the scientific community at the Institut de Physique du Globe de Paris (IPGP) website (www.ipgp.jussieu.fr/∼qyl). Any group wishing to test their tomographic algorithm is encouraged to download the synthetic data.

Key words: Surface waves and oscillations; Seismic anisotropy; Seismic tomography; Computational seismology.

1 INTRODUCTION

As a result of the deployment of many digital broad-band networks over the past two decades, there has been a tremendous growth in the volume of acquired seismic data. With this wealth of data, the different existing global tomographic methods (Romanowicz 2003) result in different models of velocity, seismic anisotropy and attenuation (see, for example, Becker & Boschi 2002 for a compilation of isotropic mantle models). In fact, depending on the inversion strategy (regularization, correlation length), theory approximation (ray theory, Born approximation or full wavefield), data processing, *a priori* assumptions on the Earth properties, data selection (fundamental surface wave, overtones, body waves or normal modes), model parametrization, data spatial and azimuthal coverage, the final tomographic images are generally different in details, sometimes even incompatible with each other. Becker & Boschi (2002)

© 2008 The Authors
Journal compilation © 2008 RAS

and Boschi *et al.* (2007, 2008) made quantitative comparison between existing global models and showed that their correlation is generally high at long wavelength, but lower at shorter wavelength. To interpret the fine details of structure images in terms of the associated physical processes (e.g. mantle convection, plate motion or fluid flow in a reservoir), it is essential to understand in a quantitative way the resolving properties of each specific imaging algorithm. But, since the Earth structure still remains ultimately unknown, it is hard to properly evaluate the efficiency of imaging algorithms or to know the resolution limit of each tomographic method.

Various methods have been used to measure tomographic resolution of model parameters (e.g. isotropic velocity, radial anisotropy, azimuthal anisotropy and anelasticity), such as the checker-box test (Zhou *et al.* 2005), L-curve (Boschi *et al.* 2006) and Backus–Gilbert resolution kernel (Backus & Gilbert 1968; Montagner & Jobert 1981; Ritsema & van Heijst 2000), which are insightful in understanding the effects of incomplete data coverage and inversion regularization. But for checkerboard tests or the resolution kernel, the same theoretical formulation is used to calculate synthetic data and to solve the inverse problem. In that way, the limits in the theoretical formulation itself could not be assessed. This becomes possible if the synthetic data are based on an accurate numerical simulation. Therefore, the current methodologies of model assessment should be supplemented by a high-accuracy benchmark data set. Examples of benchmark models that have gained widespread popularity for various types of testing are: the 2-D Marmousi model (Versteeg 1994), elastic Marmousi model (Martin *et al.* 2006), SEG/EAGE 3-D salt and overthrust models (Aminzadeh *et al.* 1995, 1996, 1997; House *et al.* 1996), 2-D wide-angle reflection/refraction benchmark data set (Brenders & Pratt 2007; Hole *et al.* 2005) and BP velocity inversion benchmark (Billette & Brandsberg-Dhal 2005). These benchmark models and synthetic data sets have demonstrated their great value of advancing imaging and tomography technique in the oil and gas industries. To meet the need for innovative imaging techniques for 3-D and anisotropic imaging, Society of Exploration Geophysicist (SEG) is currently producing the next generation of 3-D elastic model for calibrating multicomponent processing. At the same time, the European Association of Geoscientists and Engineers (EAGE) research committee is conducting 3-D anisotropic elastic data set for calibration of anisotropic inversion and imaging (Mulder *et al.* 2006). So far, however, there has been no benchmark data set for the global-scale seismological community.

To generate one global-scale benchmark, there are two mandatory prerequisite conditions: (1) the chosen numerical modelling method must accurately incorporate the effect of velocity, density, attenuation, rotation and gravity on the seismic wave propagation and (2) available computer resources must afford the vast computational cost of fully numerical global-scale simulation. Fortunately, in the last 10 yr, the field of numerical modelling of global-scale seismic wave propagation in 3-D anisotropic structures has made significant progress due to the introduction of the spectral-element method (SEM; Komatitsch & Vilotte 1998; Chaljub *et al.* 2003) and the coupled spectral-element method (CSEM; Capdeville *et al.* 2003). The SEM has very little intrinsic numerical dispersion and can accurately incorporate the effects of 3-D variations of seismic wave velocity, density, anelasticity, discontinuity, ellipticity, topography, ocean effect, rotation and self-gravity. Computational speed has also been improved dramatically, thanks to the introduction of PC clusters, which has made it possible to perform global-scale fully anisotropic modelling. For example, on a PC Beowulf Cluster with 150 processors, Komatitsch *et al.* (2002) reached a minimum seismic period of 18s. Tsuboi *et al.* (2004) simulated seismic waves accurately down to a minimum period of 5 s by using 1944 processors and down to a minimum period of 3.5 s by using 4056 processors of the earth simulator in Japan.

In the framework of the European training network Seismic Wave Propagation and Imaging in Complex media: a European Network (SPICE), it has been decided to carry out a global-scale benchmark, testing global inversion algorithms through the inversion of the same high-accuracy synthetic data set. The goal of this benchmark data set is not only to compare different inversion methods in terms of resolution to assess their limits and advantages, but also to enable seismologists to analyse how current imaging techniques are limited or not by approximations in theory, the inadequacy of data quality and coverage (Qin *et al.* 2006). For example, how do an inappropriate *a priori* model of shallow layers, the ignorance of 3-D anisotropy and attenuation and finite-frequency effects at long period bias the inverted model? How do the great-circle approximation of propagation, the linearization of sensitivity of phase velocity to mantle structure and the scaling of P and S velocity affect the inversion results? How do the model parametrization and inversion strategy affect the tomographic models? Furthermore, this benchmark data set will also contribute to pinpoint the important aspect that should be addressed by future methodological development and verify the signal analysis at the pre-processing phase (such as the picks of the overtones and fundamental modes, filtering of the waveform, selection and rotation of the components), which introduces a certain level of arbitrariness in the tomography inversion procedures.

Before constructing a benchmark model as sophisticated as possible and generating the synthetic data set, a very time-consuming endeavour, special care has been taken to test the feasibility of this exercise. To this end, the benchmark experiment has been carried out in two stages by increasing complexity in the model and by increasing the frequency bandwidth of the synthetic seismograms. The first step is the preliminary benchmark in which the input model is a very simple isotropic model and the synthetic seismograms have a minimum period of 50 s. Then we perform the second benchmark with complex anisotropic input model and wider frequency band.

In this paper, we first present the preliminary benchmark, then the second benchmark. For each of them, we discuss the construction of the input model, the generation of the synthetic data set and the corresponding inversion.

2 PRELIMINARY BENCHMARK

To confirm that the numerical dispersion, sampling rate and acquisition geometry are suitable for the benchmark test, we decided to first carry out a preliminary benchmark.

2.1 Acquisition geometry

Global seismic tomography is limited by the uneven distribution, in time and space, of the sources and by the incomplete coverage with recording stations. To test these effects on seismic tomography, the modelling process must emulate the real distribution of stations and events.

The primary seismic sources used for global tomography are earthquakes, which occur only at certain locations around the world, primarily plate boundaries. Therefore, in this experiment, we distributed 27 sources, mainly along the plate boundaries, as shown in Fig. 1(a). The total number of sources is determined according to the project schedule and computation cost of each event. To get a homogeneous path coverage with minimum computational time, these events are well separated, however, with a little concentration

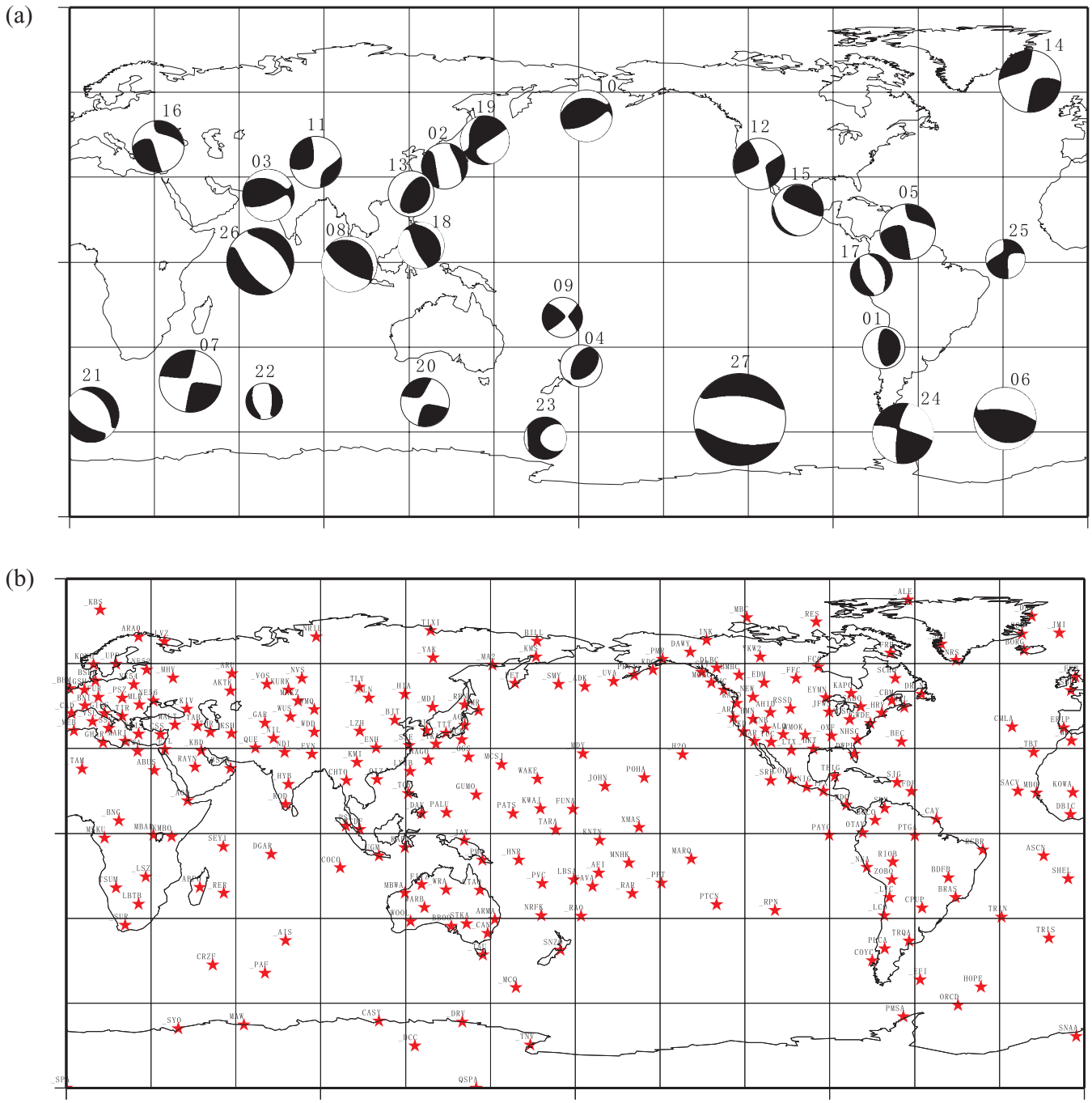

Figure 1. Acquisition geometry of the preliminary Benchmark. (a) Distribution of 27 events with inhomogeneous magnitude. All of the sources are deliberately distributed along the plate boundaries to mimic the real seismic activity. (b) Distribution of the 256 stations chosen from the Federation of Digital Broadband Seismograph Network (FDSN). In oceanic regions, there are fewer stations than in continental areas. In the southern hemisphere, there are fewer stations than in the northern hemisphere.

on West Pacific. Since only medium- and large-scale anomalies are inserted into the input model, 27 sources (each is recorded by 256 stations) can produce enough path coverage to recover these anomalies. The moments of these events vary from about 10^{26} dyn-cm to about 10^{24} dyn-cm. As for the depth range of all events, only one event has a depth of about 118 km and the others are shallow with depths between 15 and 50 km. In Fig. 1(a), the pattern of the beach balls shows the diversity of source mechanisms.

The Federation of Digital Broad-Band Seismograph Network (FDSN) station book contains information about 800 stations from all networks that contribute seismic data. Due to economic and geographical reasons, the global coverage of recording stations is rather non-homogeneous. Most stations are installed in developed countries and oceanic islands. Since in some regions such as Japan and Europe, the stations are very densely distributed, we extracted only 256 stations from FDSN stations, with the criteria that the spacing between a pair of station is larger than 5°, even though including more stations does not increase the computational time. The distribution of 256 stations is shown in Fig. 1(b). In Europe and America, station distributions are dense, whereas in oceanic areas

and Southern Hemisphere, the density of stations is very low. For all these stations, we assume that the gain and normalization are equal to one.

2.2 Construction of input model

To test the basic feasibility of global-scale benchmark, it is reasonable to start with a simple model easy to be recovered. For this goal, we design the input model as follows: (1) The model is an isotropic model described by only four parameters, namely, *S*-wave velocity, *P*-wave velocity, density and attenuation. But only the *S*-wave velocity is independent. The variation of *P*-wave velocity and density are scaled to perturbation of *S*-wave velocity by a factor of 0.5 and 0.25, respectively. (2) The quality factor has no 3-D lateral variation. (3) Exclude the effect of ocean and surface topography, ellipticity, Earth's rotation and 3-D crust, since these structures will increase the complexity of surface wave inversion and crustal correction. (4) The heterogeneities of shear wave velocity consist of several depth-independent cylinder-shaped anomalies, which are inserted smoothly into the spherically symmetric earth model. These anomalies extend from the surface down to the core–mantle boundary and can be thought of as large-scale plume-like anomalies (Capdeville *et al.* 2002). Fig. 2(b) (left-hand column) shows that the heterogeneities include three large-scale anomalies with radii larger than 2500 km and five medium-scale anomalies with radii of about 1000 km. The large-scale heterogeneities have a peak percentage perturbation of about 2.5 per cent, and the medium-scale heterogeneity have peak percentage perturbation of about 6 per cent. The long-wavelength character of this isotropic input model can insure that the linearized tomographic methods are sufficiently accurate; so, the preliminary benchmark will serve to test the data set, not the tomographic methods.

The reference spherically symmetric model consists of six layers. The boundaries of the six layers are at 3480.0, 3780.0, 4700.0, 5701.0, 5791.0, 6345.0 and 6371.0 km, respectively. The 1-D reference model is different from PREM model (Dziewonski & Anderson 1981). For example, the 1-D reference model has fewer discontinuities, which make it easy to design the mesh in CSEM.

2.3 Calculation of synthetic data set

The synthetic seismograms are calculated by using Coupling Spectral Elements and Modal Solution (Capdeville *et al.* 2003). For this preliminary benchmark, as described before, we did not account for the influence of ellipticity, ocean and rotation of real Earth. But we account for the influence of gravity and attenuation of the

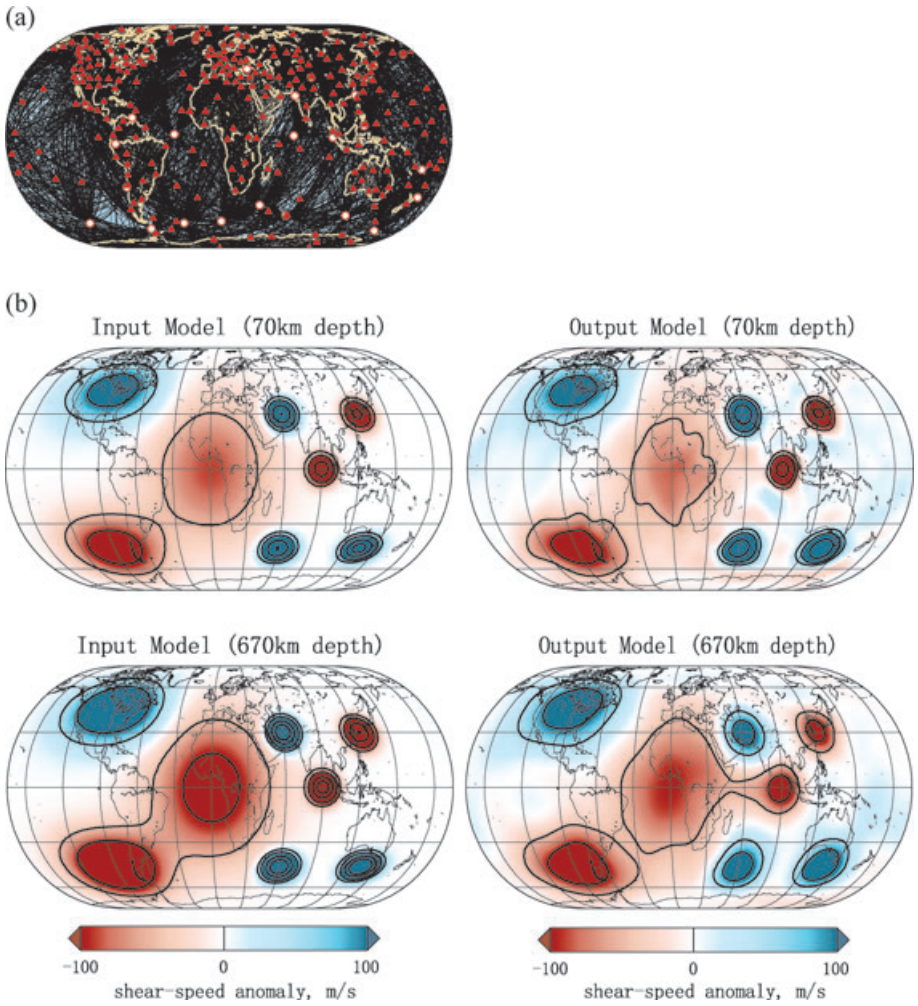

Figure 2. (a) 3134 minor-arc paths are used in this inversion. Red triangles indicate the station positions. Source locations are indicated by white circles. (b) Comparison of the inversion results (right-hand panel) using automated multimode inversion (AMI; Lebedev *et al.* 2005) with the input model (left-hand panel). The velocity variation is expressed in terms of the absolute perturbation with respect to the 1-D reference model.

© 2008 The Authors, *GJI*, **175,** 598–616
Journal compilation © 2008 RAS

1-D reference model. Because of limited computer resources and project time schedule, the minimum period to be modelled is set to 50 s, which is enough to recover the medium-scale anomalies. The source time function has a frequency band of 0.004–0.02 Hz. For each of 27 events, 256 stations are used to record the wavefield with vertical, south–north and east–west components. For each trace, to record the second surface wave train, the record length is 10500 s.

2.4 Inversion of preliminary benchmark

The automated multimode inversion (AMI) technique (Lebedev *et al.* 2005) has been applied to this preliminary data set. AMI method is highly automated and can invert this large data set within a few days.

The final inversion results are shown in Fig. 2. The path coverage used in this inversion is shown Fig. 2(a), where 3134 paths are selected from a total of 6912 paths. Due to the irregular distribution of sources and stations, the path coverage is uneven. For example, the coverage over the Pacific (especially near the southern tip of South America) and Indian Oceans is much sparser than in Eurasia and North America.

A comparison of the inversion results at 70 and 670 km with the input model, shown in Fig. 2(b), suggests that the resolution of the imaging decreases with depth. Although the shape of the anomalies is retrieved reasonably accurately at both depths, the amplitude is underestimated at 670 km. Structure in the mantle transition zone (410–670 km) is sampled by *S* and multiple *S* waves and by very-long-period fundamental-mode surface waves. Waves of both types are less abundant in a typical data set than the intermediate-period fundamental-mode surface waves. This may contribute to the decrease in the resolution, together with the regularization of the inversion and, more fundamentally, the assumption of 2-D instead of full 3-D sensitivity kernels for the *S* waves. Examining the inversion results, we note that close to the southern tip of South America, even though the path coverage has an obvious gap, the low-velocity anomaly is well recovered because this anomaly has a very large lateral extent. But when compared with the input model, this anomaly is somewhat smeared along the main path coverage direction. Similarly, the medium-scale high-velocity anomaly between Africa and Antarctic is also somewhat smeared along the north direction, affected by the azimuthal unevenness of the path coverage. At 70 km depth, there are some artificial low-amplitude anomalies around the medium-scale anomalies, which can also be attributed to insufficient path coverage. Overall, however, since the input model is very simple, the retrieved image matches the input model very well.

This preliminary benchmark not only demonstrates that numerical accuracy and acquisition geometry is suitable for global tomographic test, but also illustrates some interesting well-known problems of tomographic inversions: (1) generally, the tomographic images show smoother velocity variations compared with the input model; (2) the shape of heterogeneities can be deformed by the inhomogeneous path distribution; (3) the effects of horizontal smearing can result in merging of some separated anomalies into a single one and (4) the small-scale, weak anomalies are not robust.

3 BENCHMARK

3.1 Acquisition geometry

The station distribution is the same as that of the preliminary benchmark in Fig. 1(b), but the source distribution (Fig. 3) is modified and different from that of the preliminary benchmark in two aspects: (1) The magnitude of modelled seismic events is much more homogeneous. As a results, the noise level of synthetics is also relatively homogeneous since the same noise is added to all the records. (2) The total number of events is 29, and three of them are located close to each other, which will help to validate higher-mode phase-velocity measurement techniques that use clustering of events (Stutzmann & Montagner 1994; Beucler *et al.* 2003). The source location is the same as for the preliminary benchmark.

3.2 Constructions of input model

Our goal is not to build a tectonically or geodynamically realistic model, and our model should not necessarily be in agreement with any existing tomographic or geodynamic ones. Besides, no care is taken for having a realistic chemical composition. Nevertheless, it should be complex enough to fully test the resolving limit of tomographic methods. Since anisotropy inversion is of significant interest in seismic tomography and critical for understanding mantle convection and plate motion, adding anisotropic variation is

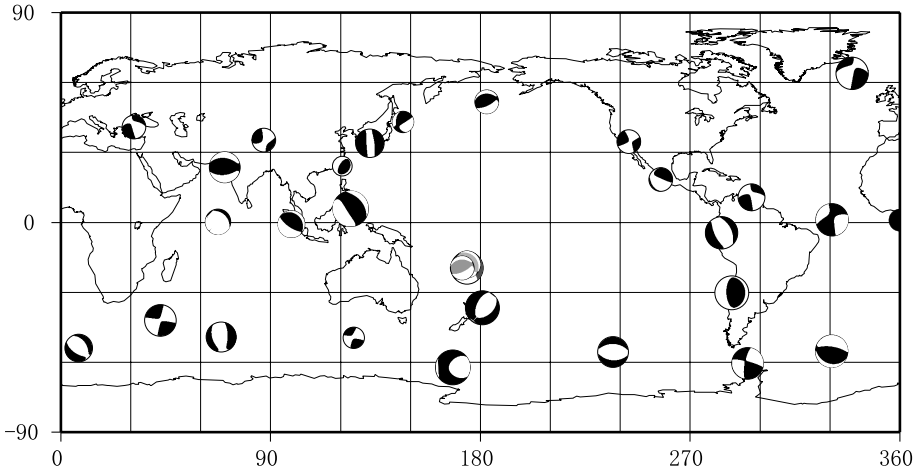

Figure 3. Distribution of 29 realistic events. These events are located along plate boundaries, representing the real seismicity of the world, with divergent, convergent and strike-slip plate margins represented in 'beach ball' diagrams. The size of the beach ball increases with the magnitude of earthquake represented. Near the location of the coloured beach ball, three events are deliberately positioned close to each other.

essential for calibrating radial and azimuthal anisotropic inversion algorithms. To this end, the model should contain various spatial scales and different types of heterogeneities in velocity, radial anisotropy, azimuth anisotropy, attenuation and density.

In addition, because crustal correction plays a key role in global seismic tomography, it is essential to incorporate the shallow, low-velocity crustal layers in the benchmark model so that the synthetic data set can be used for the validation of various crustal-correction methods and evaluating the effect of the corrections.

With these goals in mind, we constructed the input model by assembling the following parts: 1-D reference model, topography and bathymetry, crustal structures and 3-D perturbation in isotropic and anisotropic velocity and density in the whole mantle. Although these different parts are specified with different parametrizations, finally, they are all converted into elastic tensor (attenuation factor as one independent parameter) and incorporated into CSEM algorithm.

3.2.1 1-D reference model

The 1-D reference model consists of six layers, which correspond to the lowermost mantle, lower mantle, middle mantle, transition zone, upper mantle and crust. Their boundaries are at 3480.0, 3780.0, 4700.0, 5701.0, 5971.0, 6345.0 and 6371.0 km, respectively. Like for PREM, the radius-dependent structure within each layer is described by a third-order polynomial.

Parameters for each layer are *P*-wave vertical velocity, *P*-wave horizontal velocity, *S*-wave vertical velocity, *S*-wave horizontal velocity, *S*-wave attenuation coefficient, anisotropic parameter $\eta = \frac{F}{A-2L}$ (where A, L, F are the Love parameters). But the radial anisotropy is limited to 2 per cent in the transition zone and 5 per cent in the lowermost mantle (D$''$ layer). Fig. 4 shows a comparison of this 1-D reference model with PREM. Our reference model has 670 and 400 km discontinuities, but there are no 220 km (Lehmann) or 80 km (Gutenberg) discontinuities. Including fewer discontinuities in the 1-D reference model makes the mesh design easier and, so, improves the computation efficiency.

3.2.2 3-D variation of mantle model

3-D perturbations are superimposed on the reference model. They consist of four parts: (1) intermediate to large-scale stochastic features; (2) large-scale deterministic anomalies; (3) small-scale deterministic anomalies; (4) the azimuthally anisotropic perturbation.

Stochastic heterogeneities. Horizontally, the random part is defined in spherical harmonic domain with a maximum angular degree of 20. The variation of the spectrum amplitude with angular order is chosen to follow the derivative of a Gaussian function. The angular order with maximum value varies from 2 to 6, depending on layer and parameter. For example, for the isotropic and VTI parameters, the angular order with maximum value is 6 in the D$''$ layer, ensuring small scale heterogeneities in this layer, 3 in the lower mantle, 4 in the transition zone and 5 in the upper mantle. For the azimuthally anisotropic parameters, the angular order with maximum power varies from 2 to 5. Radially, the model is defined in wavenumber domain with a smallest wavelength of 200 km. The form of the power spectrum along radial direction is also a Gaussian function.

To mimic the real Earth, different elastic parameters are coupled in some layers. For example, in upper mantle and lower mantle, the *S*-wave velocity, density and quality factor variations are coupled with that of *P*-wave velocity by different scaling factors. But in the transition zone and lower mantle, the *S*-wave velocity and density variation are decoupled from *P*-wave variation, and quality factor Q_μ is coupled with *S*-wave perturbation. Moreover, the adjacent layers are mixed with a depth interval of 250 km.

The 3-D perturbations are defined in terms of *P*-wave isotropic average velocity, *S*-wave isotropic average velocity, *S*-wave anisotropic parameter$\xi = N/L$, *P*-wave anisotropic parameter $\varphi = C/A$, density and anisotropic parameter $\eta = F/(A-2L)$. Here A, C, N, L and F are the Love parameters, related to the elastic tensor $C_{ij}(i = 1 \ldots 6; j = 1 \ldots 6)$ in the coordinate system (θ, φ, r) (r is radius or depth, θ is latitude, φ is longitude) (Montagner & Nataf 1986), as follows,

$$A = \rho V_{PH}^2 = \frac{3}{8}(C_{11} + C_{22}) + \frac{1}{4}C_{12} + \frac{1}{4}C_{66},$$

$$C = \rho V_{PV}^2 = C_{33},$$

$$F = \frac{1}{2}(C_{13} + C_{23}),$$

$$L = \rho V_{SV}^2 = \frac{1}{2}(C_{44} + C_{55}),$$

$$N = \rho V_{SH}^2 = \frac{1}{8}(C_{11} + C_{22}) - \frac{1}{4}C_{12} + \frac{1}{2}C_{66},$$

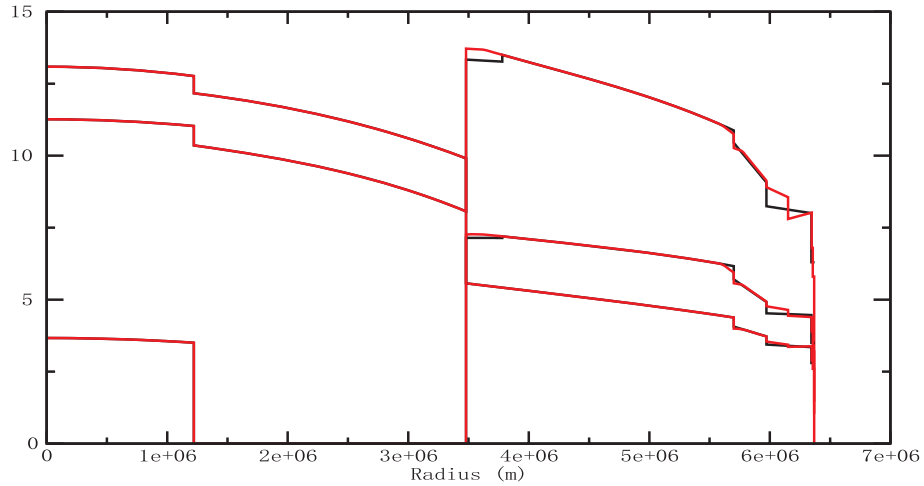

Figure 4. Comparison between 1-D reference model (black) model and PREM model (red) for the average V_S, V_P and density.

© 2008 The Authors, *GJI*, **175,** 598–616
Journal compilation © 2008 RAS

where the 6 × 6 elastic tensor with elements $C_{IJ}(I = 1 \ldots 6; J = 1 \ldots 6)$, is related to the elements of the elastic tensor c_{ijkl} in the coordinate system (θ, φ, r), by

$$\begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ & & C_{33} & C_{34} & C_{35} & C_{36} \\ & & & C_{44} & C_{45} & C_{46} \\ & & & & C_{55} & C_{56} \\ & & & & & C_{66} \end{pmatrix}$$

$$= \begin{pmatrix} c_{\theta\theta\theta\theta} & c_{\theta\theta\varphi\varphi} & c_{\theta\theta rr} & c_{\theta\theta\varphi r} & c_{\theta\theta\theta r} & c_{\theta\theta\theta\varphi} \\ & c_{\varphi\varphi\varphi\varphi} & c_{\varphi\varphi rr} & c_{\varphi\varphi\varphi r} & c_{\varphi\varphi\theta r} & c_{\varphi\varphi\theta\varphi} \\ & & c_{rrrr} & c_{rr\varphi r} & c_{rr\theta r} & c_{rr\theta\varphi} \\ & & & c_{\varphi r\varphi r} & c_{\varphi r\theta r} & c_{\varphi r\theta\varphi} \\ & & & & c_{\theta r\theta r} & c_{\theta r\theta\varphi} \\ & & & & & c_{\theta\varphi\theta\varphi} \end{pmatrix}$$

The variation pattern of amplitude of 3-D stochastic heterogeneities also looks like the typical feature of real Earth, that is, strong heterogeneities in the upper mantle, weak heterogeneities in the middle mantle, strong heterogeneities in the lowermost mantle. The range of shear wave velocity variations is between −7 and 7 per cent. To mimic the anelastic behaviour and dispersive properties of the real Earth, anomalies in the quality factor vary from −50 to 60 per cent.

Large-scale deterministic anomalies. To ensure that the input model can give basic validation of the applied tomography algorithm, one large-scale positive anomaly with a maximum perturbation of 5 per cent and one large-scale negative anomaly with a maximum perturbation of −5 per cent are inserted beneath Eastern Pacific and Western US (as in the Fig. 10a) and beneath Indian Ocean, respectively. For these two larger-scale deterministic anomalies, the perturbation of average *P*-wave velocity, density, quality factor and anisotropic parameters $\xi = N/L$, $\varphi = C/A$ and $\eta = F/(A - 2L)$ are all coupled to the variation of average *S*-wave velocity. We expect that these two larger anomalies are easy to be retrieved, and so, can serve as the basic calibration of each inversion technique.

Small-scale deterministic anomalies. To ensure that the complexity of the model is adequate to identify the resolution limit of current global surface-wave tomographic algorithms, a series of small-scale anomalies are added to the model. These deterministic heterogeneities consist of 16 depth-independent cylinder-shaped anomalies, which vary in size and amplitude. Their diameters decrease from 400 km in the north to 200 km in the south. The maximum *S*-wave velocity anomaly varies from 5 per cent for the larger ones to 3 per cent for the smaller ones. The velocity perturbation also changes the sign between two adjacent anomalies. Therefore, these small-scale heterogeneities are expected indicators for evaluating the horizontal resolution of each surface-wave tomography algorithm. Similarly, for these small-scale anomalies, the variations of the average *P*-wave, density, quality factor and anisotropic parameters $\xi = N/L$, $\varphi = C/A$ and $\eta = F/(A - 2L)$ are also scaled with respect to the average *S*-wave velocity. From Fig. 10 (uppermost line), it can be noted that the size of cylinder-shaped anomalies decreases from north to south, and that these are the most challenging elements for tomography inversion due to their small scale and velocity complexity.

Azimuthally anisotropic heterogeneities. To test how a conventional isotropic inversion might be biased by ignoring azimuthal anisotropy, we include azimuthally anisotropic heterogeneities in the model. The azimuthal anisotropy is generated by rotating anisotropic minerals around one of their symmetry axes. Since olivine is believed to be the primary component of the upper mantle and a main contributor to its anisotropy, we use the elastic coefficients of olivine crystals to define the anisotropy in the upper mantle. We assume that the olivine crystals orient with their *a*-axis in the direction of maximum deformation, and that their *b*- and *c*-axes orient randomly in the perpendicular plane. The *a*-axis defines, therefore, a direction of cylindrical symmetry, which is then oriented in different directions in the mantle. The direction of the symmetry axis and the amplitude of the anisotropy are determined by generating a vector field satisfying some statistical properties, in the same way as we generated the isotropic stochastic heterogeneity. The amplitude of the vector is used to scale the amount of anisotropy, and we choose parameters which ensure that we have more horizontally oriented crystals than vertically oriented. For the middle mantle and lower mantle, no azimuthal anisotropy is prescribed. In the D″ layer, we use the elastic coefficients of magnesiowustite since those of the newly discovered and more adequate post-perovskite were not yet known where this study was done. In this study, the parameters used for olivine are A = 232 Pa, C = 320 Pa, F = 70 Pa, L = 78 Pa and N = 67 Pa, and the properties of magnesiowustite are defined by A = 1000 Pa, C = 1070 Pa, F = 360 Pa, L = 210 Pa and N = 280 Pa. From Fig. 12 (left-hand panel), it can be seen that the azimuthal anisotropy in our model varies in amplitude, but the direction of the fast axis is constant in the south–north direction. Originally, we intended to design the azimuthal anisotropy with 3-D variation in the orientation of the fast axis, but programming error led to a constant north–south direction. The error was discovered after the synthetic seismograms were calculated, and limitations in our computational resources did not allow us to repeat the experiment with a corrected model. These added azimuthal anisotropies can, however, allow the data set to be used in examining the advantages and the drawbacks of various anisotropic tomography techniques.

3.2.3 Crustal model

Global long-period seismic data used in mantle tomography are sensitive to variations in crustal structure. But the crust is too thin to be resolved by these data. For this reason, crustal correction is an essential step for global-scale tomography to recover unbiased deeper structures. To validate a crustal correction method, it is important to incorporate crustal structure in the benchmark model.

The real crust is characterized by large velocity variations and the presence of very thin layers with sharp contrast in their elastic properties. Additionally, the depth of the Moho varies from 7 (oceanic crust) to 75 km (underneath Himalaya). In CSEM, to honour properly such a crustal structure, mesh size in the crust would be not only too small compared with that in mantle, but also excessively deformed, which will decrease the time step to fulfill the time-integral stability condition. In this benchmark experiment, we construct a crustal model by means of a Lagrange interpolation with four nodes at depths of 0, 8, 19 and 26 km. The values of seismic velocities at the three shallowest nodes are the same and are generated by smoothing horizontally the CRUST2.0 model (Bassin *et al.* 2000), suppressing the sharp transitions between neighbouring $2^\circ \times 2^\circ$ blocks. Velocities at the 26 km node are generated with the 1-D reference model plus the perturbations of uppermost mantle. So

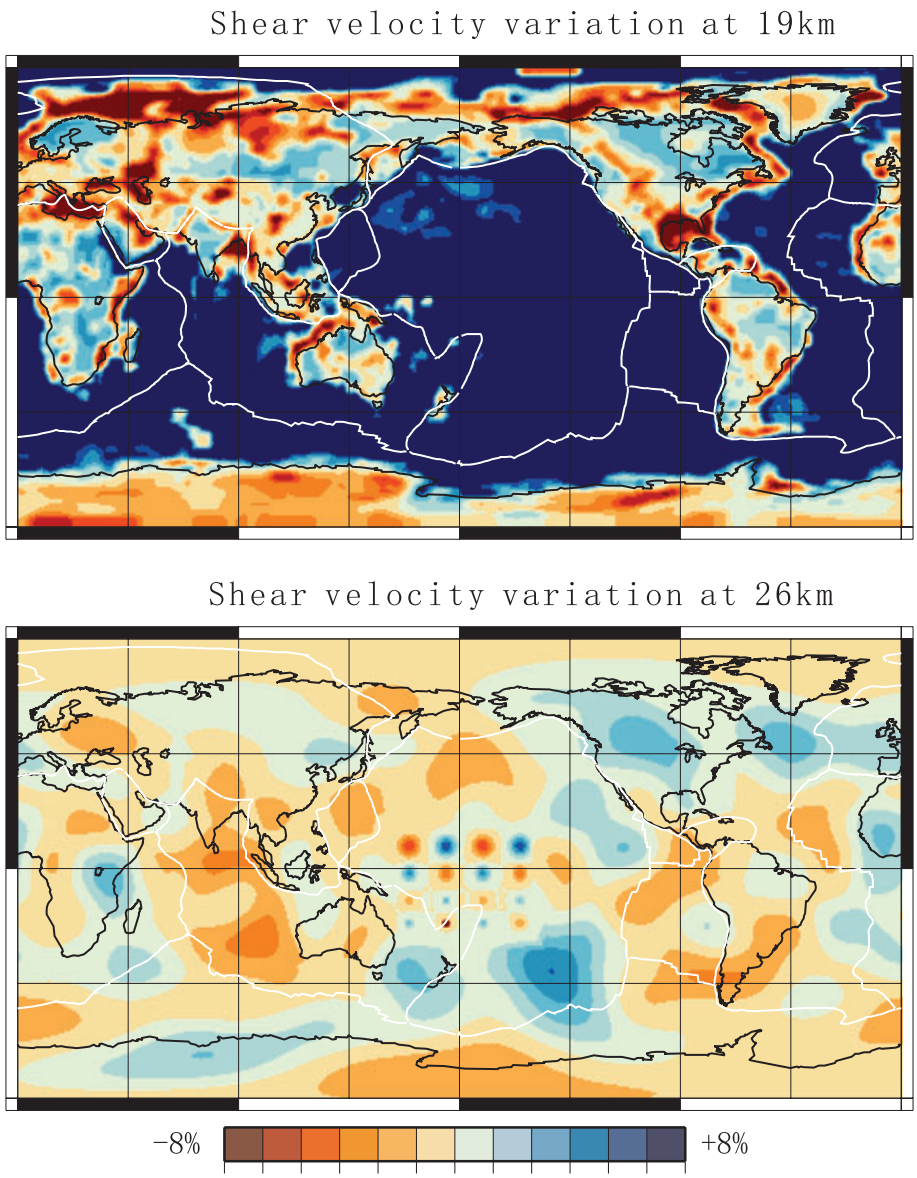

Figure 5. *S*-wave velocity variation in the crust. Upper panel: depth of 19 km; Lower panel: depth of 26 km. The perturbation is expressed with respect to the average value of 3.5 kms^{-1}.

Fig. 5 shows only *S*-wave velocity variations at 19 and 26 km depth. The crustal model is characterized by a strong contrast between higher velocity in oceanic areas and lower velocity in continental areas. The Moho depth is constant, that is, 26 km, which is consistent with the Moho depth of the 1-D reference model. The choice of a constant Moho depth makes the crustal correction easier, and more importantly, lowers the computational cost.

3.2.4 Topography and bathymetry

A smooth version of the global 5 × 5 min bathymetry and topography model ETOPO5 (NOAA 1988) has been smoothed (Fig. 6) and used to define the topography and the thickness of the oceans at the surface in this model. The CSEM honours the effect of ocean on the global wave propagation, by introduction of an equivalent load at the ocean floor, without having to explicitly mesh the water layer (Komatitsch & Tromp 2002). Besides, the Earth's rotation and ellipticity are also included in the input model.

3.3 Numerical simulations

The main objective of our benchmark exercise is to create a broadband synthetic data set over a complex model, but the extensive wave propagation computations throughout the whole earth are still too heavy for the available computation resources. To reduce the computation cost, the coupling spectral-element (CSEM) is used to calculate the synthetic data set (Capdeville *et al.* 2003). CSEM uses SEM to solve the wave propagation in regions with 3-D structures and use the modal solution where the model can be assumed to be spherically symmetric. To use CSEM, the first important parameter to be determined is the maximum frequency that one wants to reach, because maximum frequency partly determines mesh design and time step. In view of the limited computer resources available

© 2008 The Authors, *GJI*, **175,** 598–616
Journal compilation © 2008 RAS

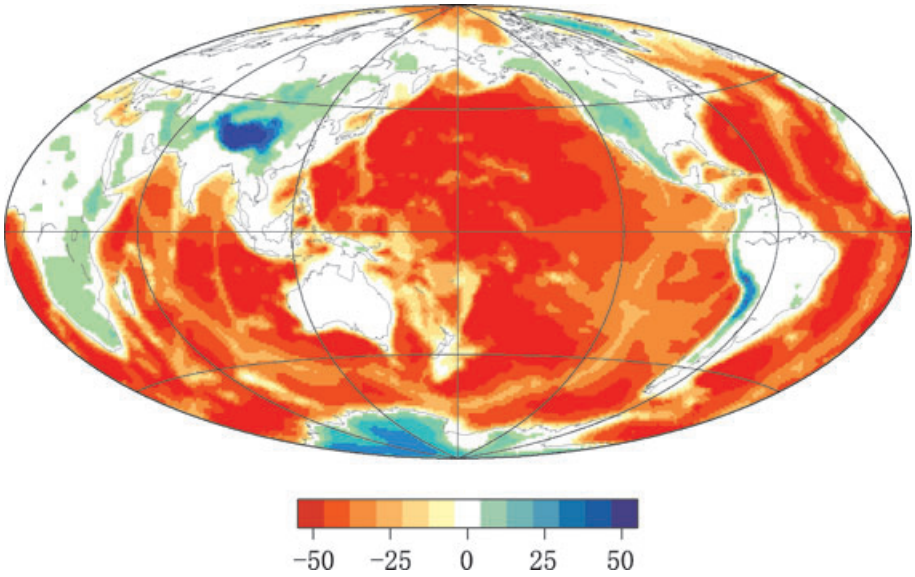

Figure 6. Topography and bathymetry of the Earth, obtained by smoothing the 5 × 5 min ETOPO5 model (NOAA 1998).

to us, we decided to compute the synthetics down to a minimum period down of 32s. Note that the requirements of memory and computation cost increase as a frequency, raised to the power 4. For instance, a factor of 16 multiplies the numerical cost each time one wants to increase the maximum frequency by a factor of 2, because the mesh size needs to be divided by 2 in three directions of space, and in addition, the time step needs to be divided by two in the fully-explicit, conditionally-stable time integration scheme.

3.3.1 Meshing the Earth

In this section, we briefly introduce the grid design in CSEM simulation of seismic wave propagation in this benchmark. The CSEM computes the wave propagation in global earth models, based on the coupling between the spectral method and a modal solution method (Capdeville *et al.* 2003). In this experiment, the Earth is decomposed into two parts: the whole mantle and crust with 3-D lateral heterogeneities and the core with only spherically symmetric heterogeneities. In the core, the solution is sought for in terms of a modal solution, which saves some of the computation cost; in the whole mantle and crust, the solution is sought for in terms of the SEM. The coupling is introduced within the spectral element method via a Dirichlet-to-Neumann (DtN) operator.

In the whole mantle and crust, the sphere is meshed using hexahedra only, based on an analytical mapping from the six sides of a unit cube to a six-block decomposition of the surface of the sphere, which is called the 'cube sphere' (Ronchi *et al.* 1996). Then each of the six sides of the cubed sphere is divided into 8 × 8 slices. And we allocate total of 8 × 8 processor for all slices; so, each processor is charged of computations in six slices.

Without considering the discontinuities and clustering of Gauss–Lobatto–Legendre (GLL) points, the grid size can be estimated using the equation below

$$\mathrm{d}x \leq \frac{V_{\min}}{f_{\max} * n_{\lambda}},$$

where dx is the possible minimum grid size, $V_{\min}$ is the minimum velocity in the model and n_{λ} is the number of samples per wavelength. To have a well-sampled wavefield, at least four or five points per minimum wavelength are needed. So, each slice is further subdivided into 3 × 3 spectral elements at the surface. And within each surface element, we use 9 × 9 = 81 GLL gridpoints for horizontal directions. This translates to an average grid spacing of about 26 km at the surface. To maintain a relatively constant number of gridpoints per wavelength and reduce the computation cost, a non-conforming interface is set for the discontinuity at 670 km depth. Below the 670 km interface, the mesh has been de-refined by a factor 2, along the horizontal directions. Vertically, the mesh used in this benchmark has been built to match the six discontinuities that are included in the 1-D reference model. Vertically, the element size is about 8.6 km in the crust, 31 km in the upper mantle, 33.7 km in the transition zone. With this mesh, we can calculate accurate synthetic seismograms at periods of 32 s and longer. The mesh file is created once and for all and stored on a large capacity disk to avoid having to recreate the mesh every time we model a new source.

After creating this kind of mesh, the possible maximum time step is determined by the Courant–Friedrichs–Levy (CFL) stability condition (Courant *et al.* 1928):

$$\mathrm{d}t = C\frac{\mathrm{d}x_{\min}}{V_{\max}},$$

where d$x_{\min}$ is the smallest distance between two mesh points, $V_{\max}$ is the maximum wave speed at this point and C is the Courant number (for 3-D Earth C is about 0.4). Although a smaller time step might decrease the numerical dispersion, when taking account of the constraint of computer resources, the time step is finally fixed at 0.35 s.

The meshes are validated on a 1-D reference model with minimum S-wave velocity of 3.5 kms^{-1}. Only at few nodes of the crustal model, velocities are lower than 3.5 kms^{-1}. And we consider that a more reasonable average minimum shear wave velocity is around 3.5 kms^{-1}. The final grid spacing of 26 km gives about 4.3 samples per wavelength for the highest frequency and absolute minimum velocity, which results in some dispersion in the wavefield. However, for all velocity over 4 kms^{-1} and all periods larger than 32 s, there should be negligible dispersion.

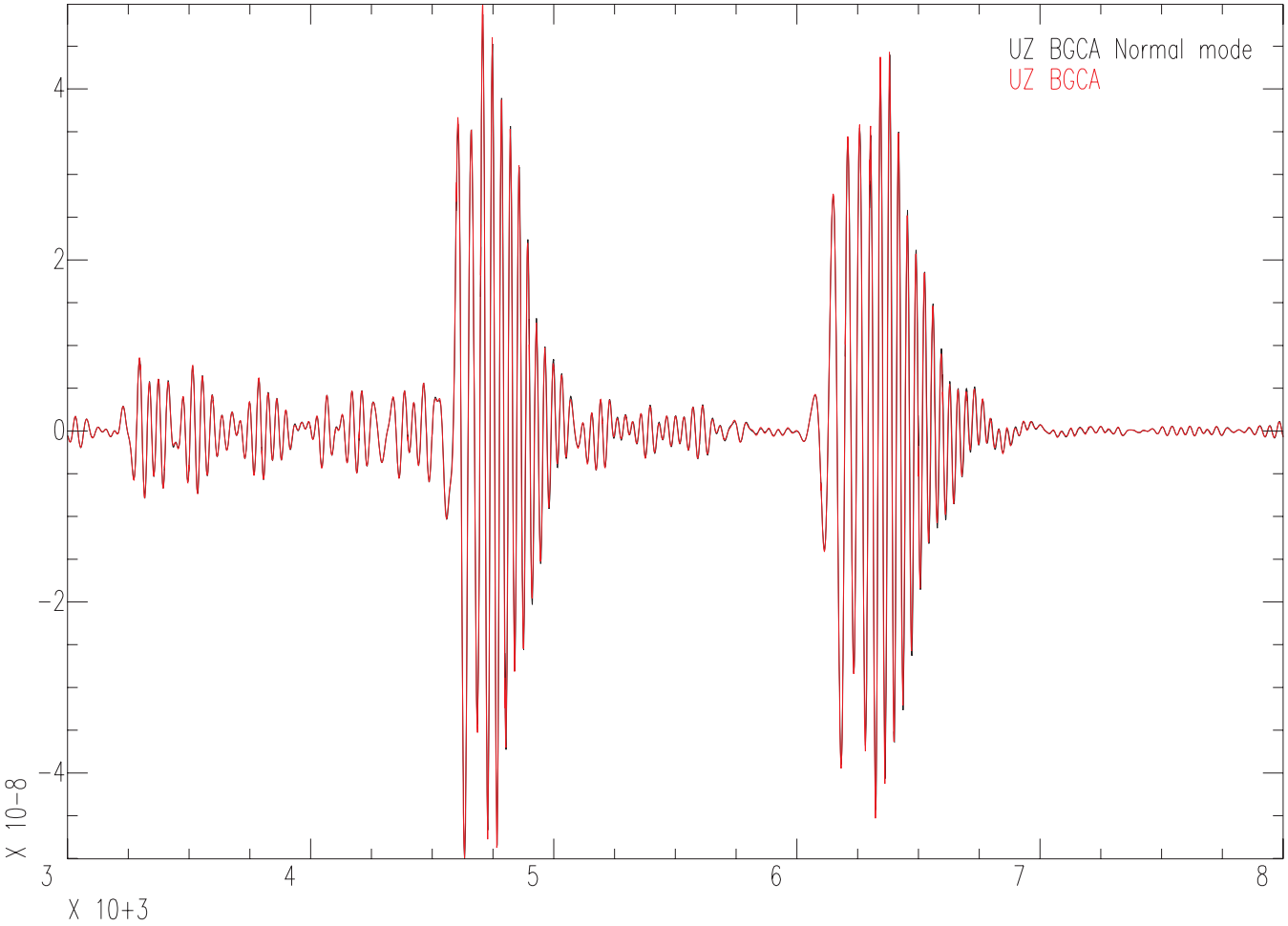

Figure 7. Numerical dispersion with a minimum period of 32s. Comparison between normal-mode solution (black) and CSEM solution (red; without attenuation and gravity).

3.3.2 Numerical dispersion

To validate the meshes and the time step used in the CSEM, synthetics seismograms for a 1-D spherically symmetric model are computed and compared with the solution obtained by quasi-analytical normal-mode solution. Indeed, in a spherically symmetric Earth model, normal modes are known with a very good precision, and the normal-mode solution is a very accurate reference solution. So, this comparison can illustrate the numerical accuracy of our simulations. From Fig. 7, it can be seen that the surface waves are very accurately modelled by the CSEM (here, attenuation and gravity are not included in the modelling; so, the amplitude of surface waves decays more slowly). Note that the mode solution is quasi-analytical but contains some numerical errors (e.g. related to the numerical integration and root-finding results); so, the difference between the CSEM and normal-mode solution is just an approximate measurement of the accuracy of CSEM. We conclude that, at a period of 32 s, the current meshes are satisfactory.

3.3.3 Source wavelet

For long-period study, the source fault can be well approximated by a point source with a set of force couples. And the derivative of time history of force couples can be approximated with one impulse signal. In this study, since the designed mesh enables the seismic waves with minimum period of 32 s to be modelled accurately, the source signature, as shown in Fig. 8, is a band-limited wavelet with frequency range of 800–500–35.7–32 s.

3.3.4 Implementation and synthetic data set

To correctly record the second orbit of surface waves, the computed record length is of 10 500 s. The 3 hr length of records, accurate for periods larger than 32 s, costs about 30 hr of CPU times on 64 processors of a modern PC cluster. The synthetic data were generated between 2006 March and July, using total of about 65 000 CPU hours.

The final set consists of 7424 three-component (south–north, east–west and vertical component) velocity wavefields with a minimum period of 32 s, a sampling rate of 1 s and a duration of 10 500 s.

Fig. 9(a) shows the comparison of a noise-free vertical component calculated in the 3-D model versus one calculated in the 1-D reference model. The synthetic seismogram exhibits clearly the main propagation features of surface waves such as well-defined higher modes and fundamental modes and the Airy phase, which corresponds to the minimum group velocity at ∼250 s period. The fundamental (R1 and R2) and higher-mode (X1 and X2) Rayleigh waves can be clearly identified. The spectrum of the traces in

© 2008 The Authors, *GJI*, **175,** 598–616
Journal compilation © 2008 RAS

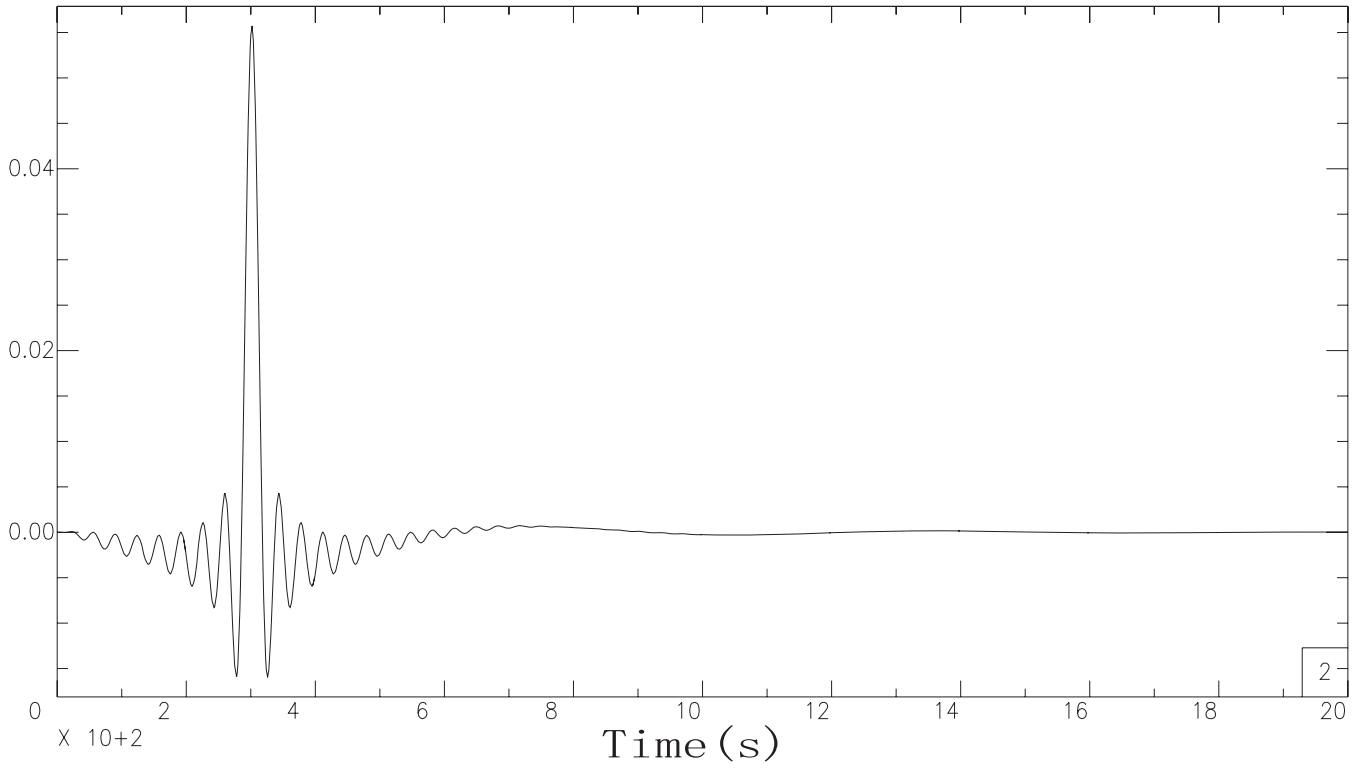

Figure 8. The source time function for the second benchmark.

Fig. 9(b) for the 3-D model has lower amplitude than the spectrum of the 1-D synthetic seismogram, especially for the higher frequency parts. This is because of the strong wave scattering attenuation in structurally complex media.

At last, to make the data set more realistic, not only the real noise from GEOSCOPE station TAM is added to the synthetics, but also CMT solution are perturbed with 10 per cent random error. The noise in the E–W and N–S components is larger than that in the *Z*-component, and most energy is concentrated at low frequency (Stutzmann *et al.* 2000). Therefore, the added noise has a significant effect on the inversion of short-period phase velocity, and the higher-mode Love wave phase velocity will be more difficult to invert because the higher-modes have much smaller amplitude than fundamental modes and are greatly affected by strong horizontal noises.

The data set is released in two versions: a noise-free one and a noise-contaminated one. Each of them consists of 29 gzipped-compressed AH files with total size of 900 megabytes and can be downloaded from the Institut de Physique du Globe de Paris (IPGP) website (http://www.ipgp.jussieu.fr/∼qyl/). Along with the data set, a document explaining the experiment, the source centroid moment tensor (CMT), source time function, crustal model, topography and bathymetry model and 1-D reference model are also provided.

4 INVERSION OF THE BENCHMARK DATA SET

The synthetic data set has been processed and inverted by two institutes with state-of-the-art surface-wave tomography procedures. At the University of Utretcht, Sergei Lebedev processed this data set with his AMI method. In Institut de Physique du Globe de Paris (IPGP), Yilong Qin inverted this data set with the more traditional and classic three-step method, that is, phase-velocity measurement, regionalization and depth inversion.

4.1 Inversion of benchmark data set with AMI method

This benchmark data set was first inverted with the AMI of surface and *S* waveforms (Lebedev *et al.* 2005). AMI used, as observables, complete vertical-component seismic waveforms, including fundamental-mode surface waves, *S* and multiple *S* waves. In this method, an elaborate scheme of filtering, windowing and weighting enables a highly complete and balanced use of structural information from seismic recordings. To ensure the validity of the assumed JWKB approximation, structural constraints are extracted only from waveforms within specially selected time-frequency windows, such that observed waveforms in all of them can be matched with JWKB synthetics simultaneously and nearly exactly. Waveform information is related to shear- and compressional-speed structure within approximate waveform sensitivity areas (Lebedev & van der Hilst 2008). The sensitivity areas are similar to the 'influence zones' computed by Yoshizawa & Kennett (2002) and to the traveltime sensitivity kernels calculated by Zhou *et al.* (2005); they are computed for the fundamental mode at a frequency in the middle of the frequency band of the inversion of the seismogram. The 3-D model is parametrized on a global triangular grid of knots, with an approximately equal interknot spacing of around 4°. The tomographic inversion is performed with LSQR with horizontal and vertical smoothing and slight norm damping. The recent application of the method to a large global data set of 50 000 seismograms (Lebedev & van der Hilst 2008) has yielded a model with a high lateral resolution (a few hundred kilometres, varying with data sampling); accurate mapping of the structure known to be in the mantle (e.g.

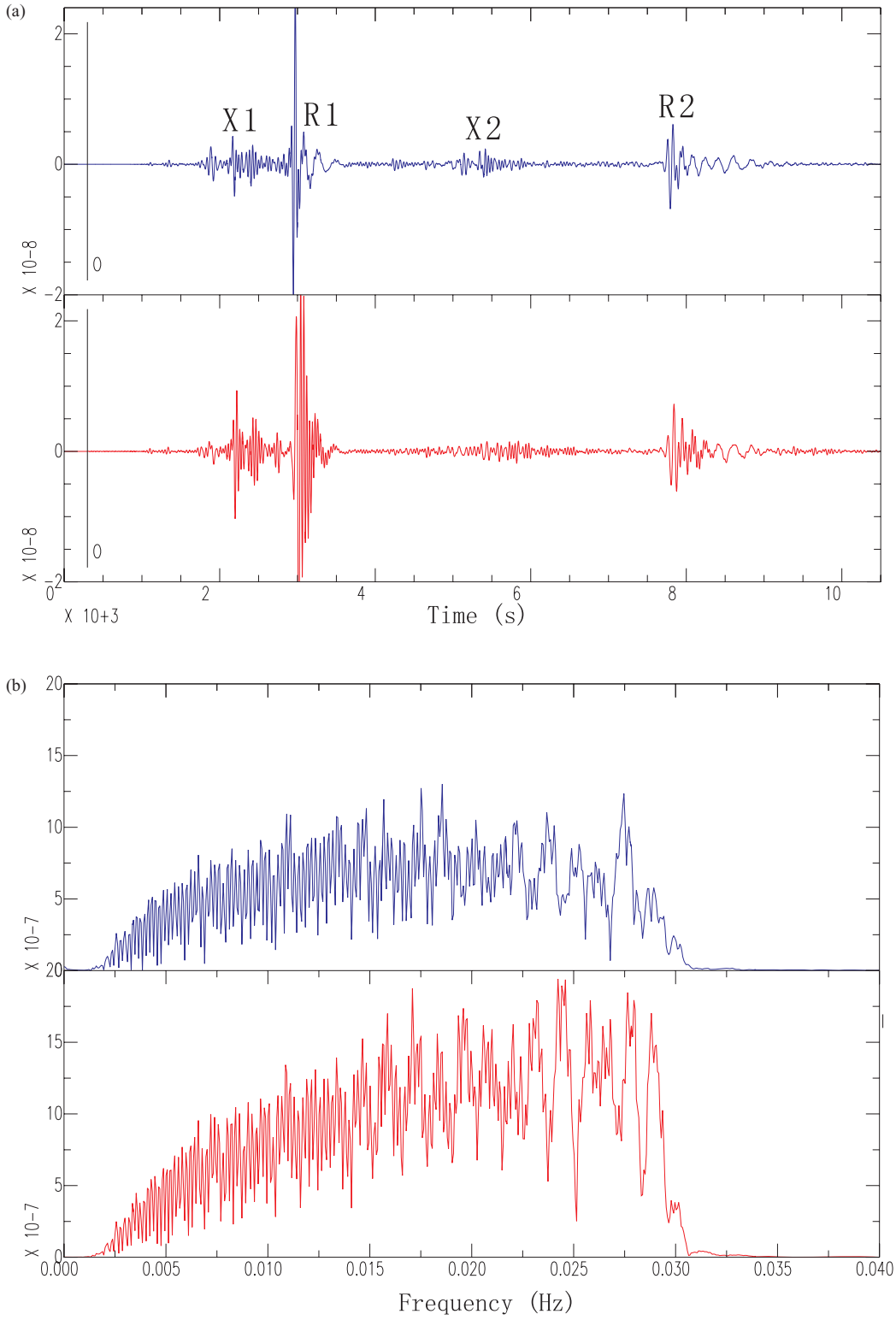

Figure 9. Comparison of vertical component seismograms for the 3-D model (blue) and the 1-D model (red). (a) The seismogram is filtered between 32 and 500s. Synthetic seismograms exhibit clearly the higher modes and fundamental modes and Airy phases; (b) the corresponding spectrum of (a). For the 3-D model, the higher-frequency energy is much more attenuated because of scattering attenuation.

the seismic expression of subduction zones or craton–lithosphere boundaries) has attested to the validity of the methods.

The results of the application of AMI to this benchmark data set are shown in Fig. 10(a) (second line from top). Signal at near-nodal azimuths has not been considered. If the criteria of Lebedev *et al.* (2005) are assumed, the average proportion of recordings from a large regional or global data set that are near-nodal at all of the frequencies of interest, is around 28 per cent. For the remaining

© 2008 The Authors, *GJI*, **175,** 598–616
Journal compilation © 2008 RAS

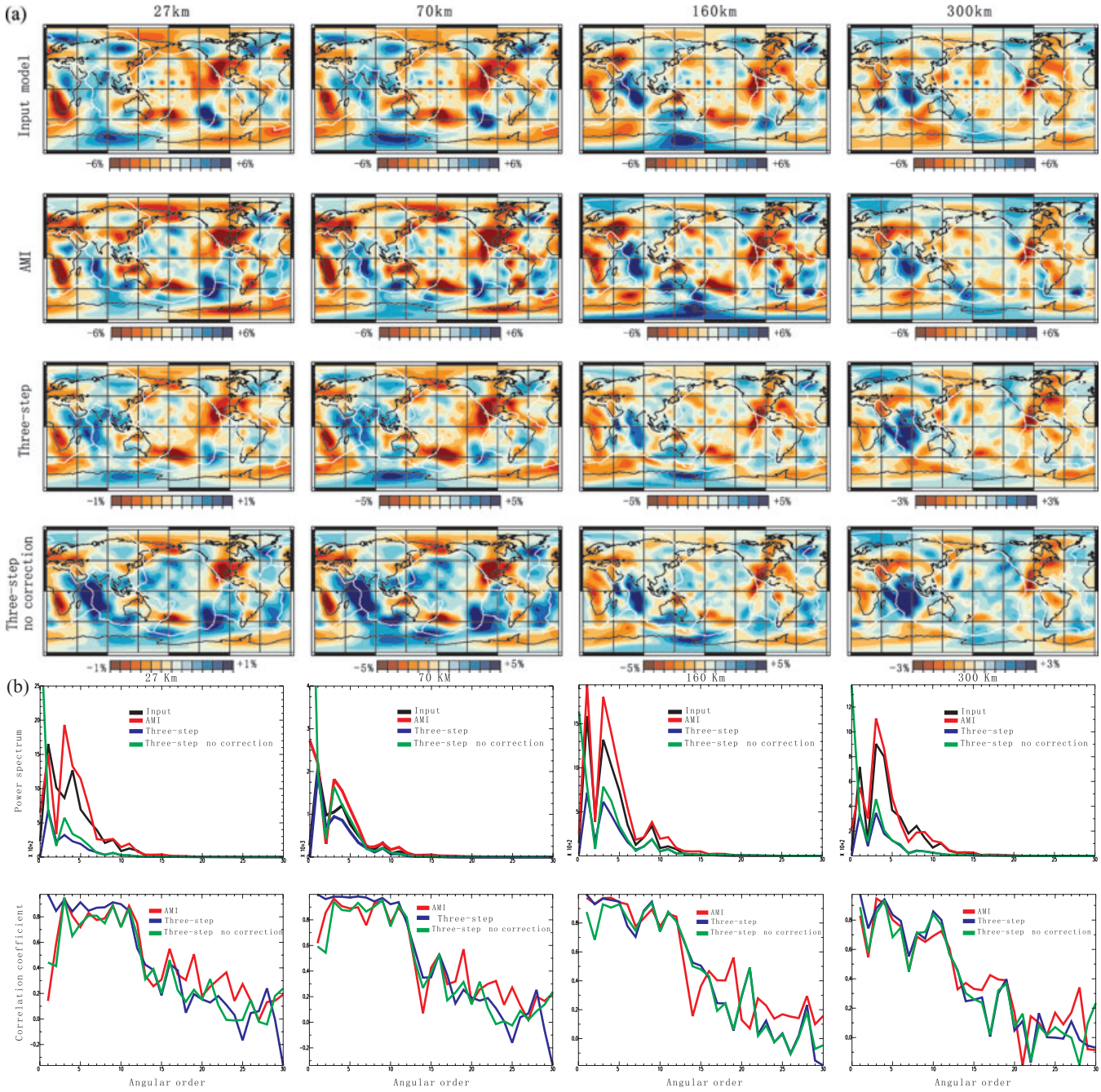

Figure 10. (a) Comparison of *SV* velocity variation of the input model and the retrieved model at the depths of 27, 70, 160 and 300 km. Upper row: input model; second row from top: inversion results of AMI; third row from top: inversion results of three-step method; lower row: inversion results of three-step method without crustal correction. (b) Upper panel: comparison of power spectra. Note that for comparison at the 27 km depth, the power of three-step results has been multiplied by 10 to make the amplitude comparable. Lower panel: comparison of correlation coefficients between the input model and different inversion results.

seismograms, waveform fitting is performed in series of time-frequency windows selected such that both the far-field and point-source approximations are valid. If a nearly exact fit is obtained in a set of windows (Lebedev *et al.* 2005), structural information from the seismogram is accepted and put in the form of linear equations with uncorrelated uncertainties. The equations obtained from the 3134 seismograms of the synthetic data set combine to make up the 3-D tomographic problem. Structural constraints are extracted from the fundamental and higher modes in frequency windows with central frequencies in the range from 3.17 to 18.48 mHz (central periods 54 to 315 s). The frequency band of the inversion varies from seismogram to seismogram. Taking into account the finite width of the Gaussian filters, the full period range of the data that constrain the model is 32 to 400 s. Source–station distances are in the 1958–18 508 km range.

Previously, AMI has been implemented and used with a 3-D, laterally heterogeneous crustal model (Lebedev & van der Hilst 2008). In the inversion in this study, however, the crust was assumed

to be laterally homogeneous, as represented by the 1-D reference model. This will enable us to evaluate the effect that erroneous assumptions regarding crustal structure may have on mantle tomographic images.

4.2 Inversion of benchmark data set with the three-step method

We invert the fundamental modes of vertical-component Rayleigh waves of this benchmark data set with the following three-step procedure: (1) use the roller coaster technique to measure phase velocity along each source-receiver path (Beucler *et al.* 2003); (2) use the computation of large anisotropic seismic heterogeneities (CLASH) method (Beucler & Montagner 2006) to regionalize phase velocity and (3) conduct depth inversion at each geographic location to obtain a 1-D depth profile, as described by Montagner & Nataf (1986). As a last step, all 1-D depth profiles are combined to give 3-D images.

4.2.1 Phase velocity measurement with roller coaster methods

The roller coaster technique developed by Beucler *et al.* (2003) is an improved version of the Fourier-domain waveform-inversion techniques of Stutzmann & Montagner (1993). The roller coaster method relies on Fourier-domain decomposition of one seismogram:

$$A_{\mathrm{R}}(\Delta,\omega)\exp\left[\mathrm{i}\phi_{\mathrm{R}}(\Delta,\omega)\right] = A_{\mathrm{S}}(\omega)\exp\left\{\mathrm{i}\left[\phi_{\mathrm{S}}(\omega)-\omega\Delta n_s\left(1+\frac{\delta n}{n_s}\right)\right]\right\},$$

where A_{R} and ϕ_{R} are the amplitude and phase of the real data spectrum, Δ is the epicentral distance in km, n_{s} is the phase slowness computed in the reference model and $\frac{\delta n}{n_{\mathrm{s}}}$ is the relative slowness perturbation to be retrieved. A_{S} and $\phi_{\mathrm{S}}$are, respectively, the amplitude and phase terms that include source term, geometrical spreading and receiver response.

The phase velocity is non-unique due to the 2π ambiguity. To deal with this, the roller coaster method finds $\frac{\delta n}{n_{\mathrm{s}}}$in two steps: (1) global detection of all possible large-scale solutions in the whole frequency range and (2) each of these smooth solutions is used as the starting model in the linearized least-squares solution to match the shorter-wavelength variations of the model space. Here, since only fundamental modes are inverted to make the fundamental parts and overtones well separated in the time domain (the first and the second surface wave trains not overlapped); only the traces with epicenter distance larger than 55° and less than 135° are selected. Moreover, to ensure the waveform quality and the fundamental modes not to be contaminated by overtones, the time windows of fundamental modes, for each trace, are checked and picked manually. These constrains considerably decrease the number of traces used in the inversion. Only 4318 paths were selected from a total of 6912 traces (256 stations record the seismograms for each of 27 events).

4.2.2 Regionalization of path-average phase velocity

To convert the phase velocities integrated along the ray paths, for each period of a given mode branch, into local phase velocity, the CLASH method (Beucler & Montagner 2006) relies on a classical azimuthal decomposition of phase velocity for a weak anisotropic earth (Smith & Dahlen 1973). For each angular frequency ω, the local phase velocity C at a given point (located at latitude θ and longitude ϕ) can be expressed as

$$\begin{aligned} C(\theta,\phi,\omega,\psi) = C_0(\theta,\phi,\omega) + \frac{1}{2C_0(\theta,\phi,\omega)}[&A_1(\theta,\phi,\omega) \\ &+A_2(\theta,\phi,\omega)\cos 2\psi + A_3(\theta,\phi,\omega)\sin 2\psi \\ &+A_4(\theta,\phi,\omega)\cos 4\psi + A_5(\theta,\phi,\omega)\sin 4\psi] \end{aligned} \quad (1)$$

Where ψ denotes the local azimuth of the horizontal wave vector, measured clockwise from the north, C_0 is the azimuthal average phase velocity (called isotropic phase velocity) and the coefficients $A_i(i = 1 \ldots 5)$ are the phase velocity azimuthal anisotropy terms. For weak anisotropic medium, A_i are depth-integral functions of 13 parameters (A, C, F, L, N, B_c, B_s, H_c, H_s, G_c, G_s, E_c and E_s) involving Rayleigh or Love displacement eigenfunctions, which are linear combinations of elastic tensor components. Using this phase velocity series expansion, Beucler & Montagner (2006) have developed a regionalization method to obtain the distribution of each of the five coefficients$A_i(i = 1 \ldots 5)$.

Phase velocity has been inverted to obtain lateral variations of azimuthal terms $A_i(i = 1 \ldots 5)$ in the period range 45–273 s, for 11 discrete periods: 45, 55, 68, 84, 103, 127, 156, 192, 220, 240 and 273 s. The grid size used for regionalization for all periods is the same, that is, 500 km. It maps the whole Earth surface into 2036 cells. The results of regionalization are controlled by three parameters: *a priori* phase velocity measurement error, *a priori* error on parameters and the correlation length, which can be thought of *a priori* errors on parameters defining the variation range of the model parameters, that is, velocity and anisotropy. The phase-velocity measurement error is the posterior error of the roller coaster algorithm. *A priori* variances on parameter are set to be 5 per cent. The correlation length should be larger than the grid size and should increase with period—we set it to 600 km for 45 s period, 670 km for 127 s, 900 km for 192 s and 1200 km for 240 s.

From the results of phase-velocity regionalization shown in Fig. 11(b), it can be noted that the resolution decreases with increasing period. For example, at periods of 45 and 127 s, the small-scale cylinder-shaped anomalies can be imaged well, but at period of 240 s, only parts of these small-scale anomalies are reproduced at all. At all these periods, the two large-scale anomalies can be well resolved, but the smallest cylinder-shaped anomalies can not be resolved, perhaps because the grid size is too coarse and the path coverage is not dense enough. Compared to the input model (Fig. 10a), it can be noted that isotropic phase-velocity distributions after regionalization can robustly show the location of anomalies.

From Fig. 11(c), it can be observed that the fast-axis direction is almost south–north, which is consistent with the fast-axis direction of input model (Fig. 12). But the magnitude of azimuthal anisotropy is not consistent with input model. Therefore, it can be inferred that, for the anisotropic phase-velocity images, the variation of fast-axis direction is relatively reliable for interpretation, but not its amplitude.

4.2.3 Inversion of local phase velocities to map 3-D structure at depth

Local phase perturbations are next converted into velocity and anisotropic perturbations at depth. The phase velocity perturbation can be expressed by 13 parameters: $A, C, F, L, N, B_c, B_s, H_c, H_s, G_c, G_s, E_c$ and E_s (Montagner & Nataf 1986). The first five parameters

© 2008 The Authors, *GJI*, **175,** 598–616
Journal compilation © 2008 RAS

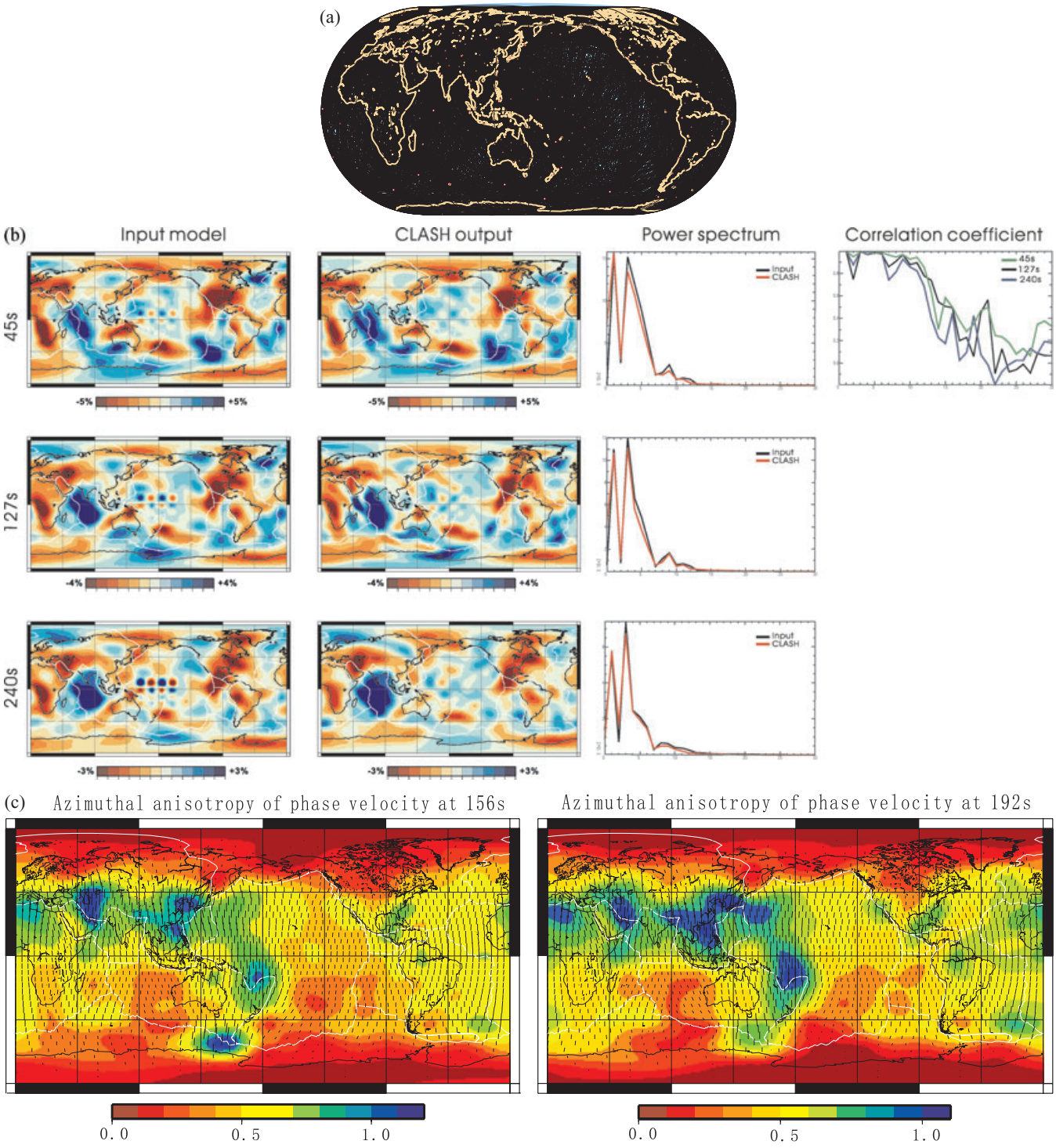

Figure 11. (a) Path coverage used for regionalization of fundamental Rayleigh mode phase velocity. There are about 4318 paths totally. The sources are indicated by red dot. (b) comparison of phase velocity between input model and CLASH results for three different periods 45, 127 and 240 s. Right-hand column: the correlation coefficients of spherical harmonics coefficients. Second column from right-hand side: the comparison of power spectrum. (c) Azimuthal anisotropy of fundamental Rayleigh phase velocity at different periods 156 and 192 s. The bar length corresponds the magnitude of azimuthal anisotropy. The colour scale in the background corresponds to the peak-to-peak amplitude of anisotropy, expressed with respect to the mean phase velocity.

(A, C, F, L, N) describe the equivalent transverse isotropic medium with vertical axis. They correspond to the isotropic phase-velocity term C_0. B_c, B_s, H_c, H_s, G_c and G_s describe the 2ψ-azimuthal variations, whereas E_c and E_s define the 4ψ-variation. The Rayleigh phase perturbation, at each location for each period T and azimuth ψ, is expressed as follows

$$\delta C_{\mathrm{R}}(T,\psi) = \int_0^a \frac{\partial C_{\mathrm{R}}}{\partial A}(\delta A + B_{\mathrm{c}}\cos 2\psi + B_{\mathrm{s}}\sin 2\psi + E_{\mathrm{c}}\cos 4\psi + E_{\mathrm{s}}\sin 4\psi)\frac{\mathrm{d}z}{\Delta h} + \int_0^a \frac{\partial C_{\mathrm{R}}}{\partial C}\delta C\frac{\mathrm{d}z}{\Delta h}$$

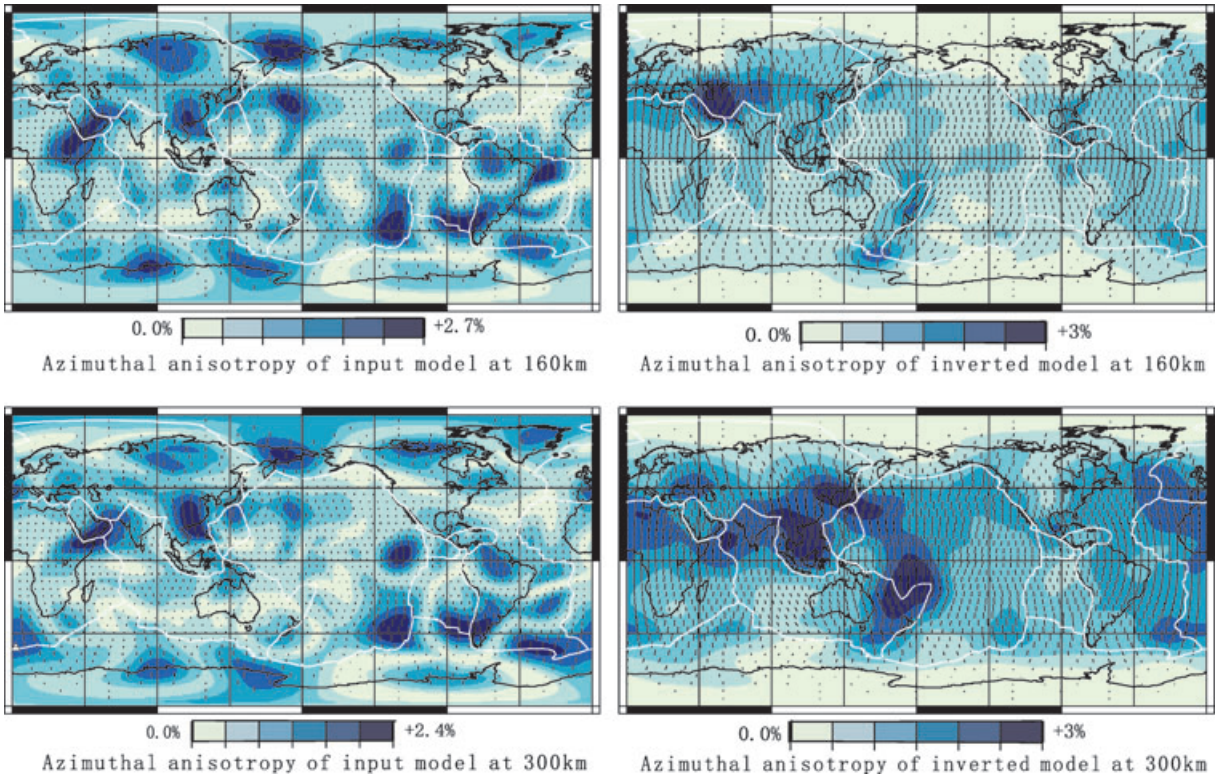

Figure 12. Left-hand panel: *SV*-wave azimuthal anisotropy of input model at two depths, 160 and 300 km. Right-hand panel: *SV*-wave velocity azimuthal anisotropy from three-step method. The directions of bars indicate the fast-axis direction of *SV*-wave, with magnitude proportional to the magnitude of azimuthal anisotropy. The map in the background corresponds to peak-to-peak amplitude of azimuth anisotropy.

$$+\int_0^a \frac{\partial C_{\mathrm{R}}}{\partial F}\left(\delta F + H_{\mathrm{c}}\cos 2\psi + H_{\mathrm{s}}\sin 2\psi\right)\frac{\mathrm{d}z}{\Delta h}$$
$$+\int_0^a \frac{\partial C_{\mathrm{R}}}{\partial L}\left(\delta L + G_{\mathrm{c}}\cos 2\psi + G_{\mathrm{s}}\sin 2\psi\right)\frac{\mathrm{d}z}{\Delta h}$$

Montagner & Nataf (1986) demonstrate that Rayleigh wave phase velocity is mainly sensitive the azimuthal anisotropy via the 2ψ-coefficient (A_2 and A_3) of eq. (1). Therefore, we invert the three parameters C_0, A_2 and A_3 only and neglect 4ψ-anisotropy.

In this inversion, vertical parametrization consists of about 20 km thick layers between Moho and 670 km discontinuity, and a Gaussian correlation with correlation length of 20 km is introduced between adjacent layers. The data error on phase velocity is obtained from the regionalization procedure as *a posteriori* error. Those errors depend on period and correspond to the resolution of local phase velocity, which is a function of the uncertainties along each path. The kernels $\partial C_{\mathrm{L}}/\partial p_i$ are calculated in a transverse isotropic reference medium, with vertical symmetry axis. A total of 11 periods are selected so that the interval between their logarithm periods is approximately constant.

13 parameters per layers are inverted for, via a least-square inversion algorithm (Tarantola & Valette 1982). The available data cannot resolve all these parameters, and Rayleigh waves are mainly sensitive to parameters L, G_c and G_s (Montagner & Nataf 1986). The 3-D model can be described by the following parameters:

$$V_{SV} = \sqrt{\frac{L + G_c\cos 2\psi + G_s\sin 2\psi}{\rho}}$$

Where V_{SV} is the *SV*-wave velocity.

Figs 10 and 12 show the results of our three-step inversion. It can be noted that: (1) the inverted model with crustal correction reflects the relative velocity variation, but the amplitude of velocity contrast is obviously smaller than in the input model and (2) for the 0° azimuthally anisotropic variation of *SV* waves, its direction is recovered but not its amplitude.

4.3 Qualitative Analysis the tomographic images

In this section, we illustrate the differences and similarities between our input model and tomographic results of AMI and the three-step method, as shown in Fig. 10.

27 km depth. In our application of the AMI method, we intentionally used no effective crustal corrections. In the output model of the test, the heterogeneities just underneath the Moho are not well recovered. Especially under Eurasia, the recovered model has lower velocity than the input model since the unaccounted-for low-velocity crustal structures are smeared to this depth. Under the Pacific, the retrieved velocity is higher than the input model, which is an effect of not correcting for crustal structure. Due, again, to the absence of a crustal correction, the phase of determined stronger anomalies, such as the large low-velocity anomaly under north America and high-velocity anomaly under the southeastern Pacific, are not reproduced entirely well. For the small-scale cylinder-shaped anomalies in eastern Pacific, only the relatively larger ones (upper two lines) are resolved but do not appear clearly.

With the three-step method, shallow heterogeneities, such as the high-velocity anomaly of Eurasia continent, are imaged relatively

© 2008 The Authors, *GJI*, **175,** 598–616
Journal compilation © 2008 RAS

well since the crustal correction has been taken into account. But the small-scale cylinder-shaped anomalies are imaged less well than by the AMI method, maybe because only data with period larger than 45 s were used or because relatively larger correlation lengths were used. For the three-step inversion without crustal correction, the retrieved velocity under Pacific is higher than input model.

70 km depth. AMI inversion and three-step method without crustal correction still do not recover well the high-velocity anomalies under Eurasia, for the same reason as above. On the other hand, AMI's resolution of small-scale cylinder-shaped anomalies is higher, compared with the 27-km-deep images. Around South America, the amplitude of heterogeneities is stronger than in the input model.

The three-step inversion maps a pattern of heterogeneities around Indian ocean that is more similar to the input model. But the cylinder-shaped anomalies are imaged less well than by AMI results.

160 km depth. Depth structure mapped by AMI under the Euro–Asia continent and the Indian Ocean match the input model very well, in spite of the absence of crustal correction. This implies that the effect of crustal correction is weak at this depth. The northernmost two lines of cylinder-shaped anomalies seem to be well resolved.

The three-step method splits the large high-velocity anomaly around Indian Ocean into one large anomaly and two smaller anomalies, which may be a result of poor path coverage. Similarly, the low-velocity anomaly in Africa is also split into several small anomalies. Only the northernmost line of cylinder-shaped anomalies is recovered, with resolution lower than that achieved by AMI at the same depth. The three-step result without crustal correction is very similar to three-step results since the effect of shallow-layer correction is small at this depth.

300 km depth. AMI achieves a resolution of tomographic images higher than that of three-step method because not only fundamental modes but also higher modes are employed by AMI. For example, the cylinder-shaped anomalies are better imaged than by the three-step method. The higher-velocity anomaly under the Indian Ocean is also better retrieved than by the three-step method. This illustrates to what extent inverting higher-mode waveforms is important to properly resolve deeper structures. The higher-velocity anomaly under Madagascar Island is poorly imaged by both AMI and three-step inversions; the explanation might be the limited vertical extent of the anomalous region, which can result in smearing from high-velocity anomalies above and below.

The azimuthal anisotropy. Similar to the variation of phase-velocity azimuthal anisotropy, the tomographic results of SV-wave azimuthal anisotropy (Fig. 12) show that the azimuth of the fast-axis of SV-waves is retrieved, whereas the amplitude of anisotropy is not. We infer that surface waves alone are not sufficient in this case—with relatively sparse azimuthal coverage—to resolve weak azimuthal anisotropy. This reminds us that caution should be taken when interpreting the strength of tomography-mapped azimuthal anisotropy.

In summary, the tomographic images are generally very similar to the input model in terms of large-scale heterogeneities. However, as expected, they do not entirely resolve the shape gradient and the boundaries of inverted heterogeneities are smoother than those in the 'true' model and smeared according to the dominant direction of path distributions. Besides, the depth range of a heterogeneous body is difficult to capture exactly, due to vertical smearing and the inherently finite vertical resolution of surface waves. The accuracy of model recovery decreases with increasing depth and with increasing input model complexity. Perhaps, this can be explained by the following facts: (1) the sensitivity of surface waves to Earth structure decreases with increasing depth and (2) the deeper structures can only be sensed by higher-mode surface waves. In maps derived with the three-step method, the amplitude of tomographic results is generally smaller than that of the input model. Tomographic maps resulting from AMI have stronger amplitude in some areas but weaker amplitude in other regions.

4.4 Quantitative comparison of input models and inversion results

Fig. 10(b) shows the spherical harmonic power spectra of different models and the correlation coefficient between input model and different inversion results. At the depth of 27 km, the three-step tomographic results with crustal correction are in good agreement with the input model in terms of pattern but in less good agreement in terms of amplitude. At the depth of 70 km, three-step tomographic results are still similar to the input model, and also the amplitude is now comparable to that of input model. At 27 and 70 km depths, AMI (applied assuming laterally homogeneous crust) and three-step results without crustal correction are less successful in reproducing the input model as a result of the choice not to apply any crustal correction in this test. This confirms that crustal correction is really critical to retrieve the structure, at least down to about 100 km depth. A comparison of the input and output models also reveals the magnitude of uncertainties in uppermost-mantle shear-speed structure, which one can expect to result from erroneous assumptions regarding crustal structure.

At a depth of 160 km, for angular extent of less than 10°, both AMI and three-step results are in agreement with the input model, but the power spectrum of AMI results shows larger amplitude than the input spectrum, whereas the power spectrum of three-step results has much lower amplitude. At the depth of 300 km, for angular extent less than 7°, the power spectra from AMI and three-step results are both similar to the input model spectrum. At the same depth, for angular extent less than 2°, the power spectra of results from AMI is lower than that of the input model; for angular extent between 2° and 6°, it is somewhat higher than that of the input model. It can also be seen that, at the depth of 160 km, for the lowest 2° angular extent, the three-step result without crustal corrections correlates with the input, with a somewhat smaller correlation coefficient.

5 DISCUSSIONS AND CONCLUSIONS

We have created a global-scale, 3-D, fully anisotropic, anelastic model and used it to generate a three-component synthetic data set for the seismological community. This data set is suitable for a wide variety of experiments, including calibration of phase-velocity measurements, radially anisotropic and azimuthally anisotropic inversion, density inversion, attenuation inversion, polarization inversion, three-component signal processing techniques and CMT inversions using surface waves. We hope the growing use of this data set over the coming years.

From the inversion of the preliminary benchmark and second benchmark data sets, we have contributed to a better quantitative understanding of the following (known) phenomena: (1) the resolution decreases with depth; (2) crustal corrections are essential to

recover shallow velocity structures; (3) the higher-mode waveform is indispensable to retrieve deeper structures; (4) the depth range of heterogeneities is difficult to determine accurately, especially for relatively thin anomalous layers, but their horizontal location is robustly constrained; (5) the horizontal boundary or shape of anomalies is not always reliable; (6) the location of larger-scale heterogeneities is reliable, but the weak anomalies of smaller scale are often artefacts; (7) the absolute amplitude of variation of isotropic *S*-wave velocity is difficult to determine accurately; (8) the amplitude of azimuthal anisotropy of *S*-wave velocity is not correctly mapped using only surface wave phase-velocity measurements; (9) smearing occurs, both horizontally and vertically, and must be taken into account for interpretation and (10) splitting of single anomalies into apparent multiple ones may also occur. Since only two tomography techniques are applied and only vertical components have been used, these preliminary conclusions are subject to modification, based on future results from other tomography methods.

The major limitation of our synthetic data set is that the minimum period is 32 s, which makes it not suitable for benchmarking body wave or crustal tomography. Other drawbacks include the fact that the receiver distribution is too sparse for the present data set to be appropriate for two-station surface-wave analysis techniques and the fact that all of the discontinuities, including the Moho, 400 km and 670 km discontinuities, are at constant depths. Due to the limited number of modelled seismic events, the path coverage is sparser than used in actual global tomographic inversions. We hope that in the near future, a new global-scale seismological benchmark will be conducted, with a much broader frequency range, more sources and more stations.

ACKNOWLEDGMENTS

This global-scale benchmark is a cooperative effort of IPGP, Oslo university, Utrecht university and Nantes university. We thank Jeannot Trampert as the editor, Lapo Boschi and an anonymous reviewer for their comments and suggestions on this paper in its initial form. This project is funded by the European Commission's Human Resources and Mobility Programme Marie Curie Research Training Network SPICE contract MRTN-CT-2003-504267. The computations were implemented at IPGP, the Institut du Développement et des Ressources en Informatique Scientifique (IDRIS) and the Centre Informatique National de l'Enseignement Supérieur (CINES).

REFERENCES

Aminzadeh, F., Burkhard, N., Kunz, T., Nicoletis, L. & Rocca, F., 1995. 3-D modeling project: 3rd report, *Leading Edge,* **14**(2), 125–128.

Aminzadeh, F., Burkhard, N., Long, J., Kunz, T. & Duclos, P., 1996. Three dimensional SEG/EAEG models - an update, *Leading Edge,* **15**(2), 131–134.

Aminzadeh, F., Brac, J. & Kunz, T., 1997. *3-D Salt and Overthrust Models,* 3-D Modeling Series, Vol. 1, *Soc. of Expl. Geophysicists.*.

Backus, G.E. & Gilbert, J.F., 1968. The resolving power of gross earth data, *Geophys. J. R. astr. Soc.,* **16,** 169–205.

Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *EOS, Trans. Am. geophys. Un.,* **81,** F897.

Becker, T.W. & Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosyst.,* **3**(1), 1003, doi:10.1029/2001GC000168.

Beucler, E. & Montagner, J.-P., 2006. Computation of large anisotropic seismic heterogeneities (CLASH), *Geophys. J. Int.,* **165,** 447–468.

Beucler, É., Stutzmann, É. & Montagner, J.P., 2003. Surface wave higher-mode phase velocity measurements using a roller-coaster-type algorithm, *Geophys. J. Int.,* **155**(1), 289–307.

Billette, F. & Brandsberg-Dhal, S., 2005. The 2004 BP Velocity Benchmark, in *Proceedings of the 67th Meeting,* EAGE Conference & Exhibition, Expanded Abstract, B035, Madrid, Spain, 13–16 June 2005.

Boschi, L., Becker, T.W., Soldati, G. & Dziewonski, A.M., 2006. On the relevance of Born theory in global seismic tomography, *Geophys. Res. Lett.,* **33,** L06302, doi:10.1029/2005GL025063.

Boschi, L., Becker, T.W. & Steinberger, B., 2007. Mantle plumes: dynamic models and seismic images, *Geochem. Geophys. Geosyst.,* **8,** Q10006, doi:10.1029/2007GC001733.

Boschi, L., Becker, T.W. & Steinberger, B., 2008. On the statistical significance of correlations between synthetic mantle plumes and tomographic models, *Phys. Earth planet. Inter.,* **167,** 230–238.

Brenders, A.J. & Pratt, R.G., 2007. Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model, *Geophys. J. Int.,* **168**(1), 133–151.

Capdeville, Y., Larmat, C., Vilotte, J.P. & Montagner, J.P., 2002. Direct numerical simulation of the scattering induced by a localized plume like using a coupled spectral element and modal solution, *Geophys. Res. Lett.,* **29**(9), doi:10.1029/2001GL013747.

Capdeville, Y., Chaljub, E., Vilotte, J.P. & Montagner, J.P., 2003. Coupling spectral elements and modal solution: a new efficient tool for numerical wave propagation in laterally heterogeneous earth models, *Geophys. J. Int.,* **152,** 34–66.

Chaljub, E., Capdeville, Y. & Vilotte, J.P., 2003. Solving elastodynamics in a fluid-solid heterogeneous sphere: a parallel spectral element approximation on non-conforming grids, *J. Comp. Phy.* **187,** 457–491.

Courant, R., Friedrichs, K. & Lewy, H., 1928. Über die partiellen Differenzengleichungen der mathematischen Physik, *Mathematische Annalen,* **100**(1), 32–74.

Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference earth model (PREM), *Phys. Earth planet. Inter.,* **25,** 297–356.

Hole, J.A., Zelt, C.A. & Pratt, R.G., 2005. Advances in controlled-source seismic imaging, *EOS, Trans. Am. geophys. Un.,* **86,** 177–188.

House, L., Fehler, M., Aminzadeh, F., Barhen, J. & Larson, S., 1996. A national laboratory-industry collaboration to use the SEG/EAEG model data sets, *Leading Edge,* **15,** 135–136.

Komatitsch, D. & Tromp, J., 2002. Spectral-element simulations of global seismic wave propagation -II. 3-D models, oceans, rotation, and gravity, *Geophys. J. Int.,* **150,** 303–318.

Komatitsch, D. & Vilotte, J.P., 1998. The spectral element method: an effective tool to simulate the seismic response of 2D and 3D geological structures, *Bull. seism. Soc. Am.,* **88,** 368–392.

Komatitsch, D., Ritsema, J. & Tromp, J., 2002. The spectral-element method, beowulf computing, and global seismology, *Science*, **298,** 1737–1742.

Lebedev, S. & Van Der Hilst, R.D., 2008. Global upper-mantle tomography with the automated multimode inversion of surface and S wave forms, *Geophys. J. Int.,* **173,** 505–518.

Lebedev, S., Nolet, G., Meier, T. & Van Der Hilst, R.D., 2005. Automated multi-mode inversion of surface and S waveforms, *Geophys. J. Int.,* **162**(3), 951–964.

Martin, G.S., Wiley, R. & Marfurt, K.J., 2006. Marmousi2: an elastic upgrade for Marmousi, *Leading Edge,* **25**(2), 156–166.

Montagner, J.P. & Jobert, N., 1981. Investigation of upper mantle structure under young regions of the South-East Pacific using long-period Rayleigh waves, *Phys. Earth planet. Inter.,* **27,** 206–222.

Montagner, J.P. & Nataf, H.C., 1986. A simple method for inverting the azimuthal anisotropy of surface waves, *J. geophys. Res.,* **91**(B1), 511–520.

Mulder, W.A., Nicoletis, L. & Alkhalifah, T., 2006. The EAGE 3D anisotropic elastic modelling project, in *Proceedings of the 68th EAGE Conference & Exhibition,* Vienna (2006).

NOAA, 1988. National Oceanic and Atmospheric Administration (NOAA) product information catalog—ETOPO5 earth topography 5-minute digital

model, Technical report, US Department of Commerce, Washington, DC, p. 171.

Qin, Y., Capdeville, Y., Maupin, V. & Montagner, J.P., 2006. A SPICE blind test to benchmark global tomographic methods, *EOS (Washington DC),* **87,** 512–513.

Ritsema, J. & van Heijst, H.J., 2000. Seismic imaging of structural heterogeneity in Earth's mantle: evidence for large-scale mantle flow, *Sci. Prog.,* **83,** 243–259.

Romanowicz, B., 2003. Global mantle tomography: progress status in the last 10 years, *An. Rev. Geophys. Space Phys.,* **31**(1), 303.

Ronchi, C., Ianoco, R. & Paolucci, P.S., 1996. The "Cubed Sphere": a new method for the solution of partial di.erential equations in spherical geometry, *J. Comput. Phys.,* **124,** 93–114.

Smith, M.L. & Dahlen, F.A., 1973. The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropicmedium, *J. geophys. Res.,* **78,** 3321–3333.

Stutzmann, E. & Montagner, J.-P., 1994. Tomography of the transition zone from the inversion of higher mode surface waves, *Phys. Earth planet. Int.,* **86,** 99–115.

Stutzmann, E. & Montagner, J.P., 1993. An inverse technique for retrieving higher mode phase velocity and mantle structure, *Geophys. I. Int.,* **113,** 669–683.

Stutzmann, E., Roult, G. & Astiz, L., 2000. GEOSCOPE station noise levels, *Bull. seismol. Soc. Am.,* **90,** 690–701.

Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. geophys. Space Phys.,* **20**(2), 219–232.

Tsuboi, S., Komatitsch, D., Chen, J. & Tromp, J., 2004. Modeling of global seismic wave propagation on the earth simulator, *J. Earth Simulator,* **1,** 57–66.

Versteeg, R., 1994. The Marmousi experience: velocity model determination on a synthetic complex data set, *Leading Edge,* **13,** 927–936.

Yoshizawa, K. & Kennett, B.L.N., 2002. Determination of the influence zone for surface wave paths, *Geophys. J. Int.,* **149,** 440–453.

Zhou, Y., Dahlen, F.A., Nolet, G. & Laske, G., 2005. Finite-frequency effects in global surface-wave tomography, *Geophys. J. Int.,* **163,** 1087–1111.