Reliability of mantle tomography models assessed by spectral element simulation

Yilong Qin,[1] Yann Capdeville,[1] Jean-Paul Montagner,[1] Lapo Boschi[2]
and Thorsten W. Becker[3]

[1]*Département de Sismologie, Institut de Physique du globe de Paris,* 4, *Place Jussieu,* 75252 *Paris, France. E-mail: qyl@ipgp.jussieu.fr*
[2]*Eidgenossische Technische Hochschule, Zurich, Switzerland*
[3]*Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA*

SUMMARY
Global tomographic models collected in the Seismic wave Propagation and Imaging in Complex (SPICE media: a European network) model library (http://www.spice-rtn.org/research/planetaryscale/tomography/) share a similar pattern of long, spatial wavelength heterogeneity, but are not consistent at shorter spatial wavelengths. Here, we assess the performance of global tomographic models by comparing how well they fit seismic waveform observations, in particular Love and Rayleigh wave overtones and fundamental modes. We first used the coupled spectral element method (CSEM) to calculate long-period ($>$100 s) synthetic seismograms for different global tomography models. The CSEM can incorporate the effect of three-dimensional (3-D) variations in velocity, anisotropy, density and attenuation with very little numerical dispersion. We then compared quantitatively synthetic seismograms and real data. To restrict ourselves to high-quality overtone data, and to minimize the effects of the finite extent of seismic sources and of crustal heterogeneity, we favour deep ($>$500 km) earthquakes of intermediate magnitude ($M_w \sim 7$). Our comparisons reveal that: (1) The 3-D global tomographic models explain the data much better than the one-dimensional (1-D) anisotropic Preliminary Reference Earth Model (PREM). The current 3-D tomographic models have captured the large-scale features of upper-mantle heterogeneities, but there is still some room for the improvement of large-scale features of global tomographic models. (2) The average correlation coefficients for deep events are higher than those for shallow events, because crustal structure is too complex to be completely incorporated into CSEM simulations. (3) The average correlation coefficient (or the time lag) for the major-arc wave trains is lower (or higher) than that for the minor-arc wave trains. Therefore, the current tomographic models could be much improved by including the major-arc wave trains in the inversion. (4) The shallow-layer crustal correction has more effects on the fundamental surface waves than on the overtones.

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

GJI Seismology

1 INTRODUCTION

Over the last two decades, numerous three-dimensional (3-D) models of seismic velocity in the earth's interior have been derived from different seismological measurements (body wave traveltime, phase velocity or surface wave full waveforms) and with different inversion methods (e.g. Becker & Boschi 2002). Depending on the choice of inversion algorithm (direct or iterative, linear or non-linear), regularization scheme (norm or roughness minimization), geographic parameterization (spherical harmonics, continuous function, splines or blocks) and data types, tomographic models are generally different from each other. The assessment of tomographic model quality

and robustness has mainly relied, so far, on a model's correlation with other existing models, on the calculation of variance reduction, on the agreement of a model's dominant features with surface tectonics and on checkerboard resolution tests. However, none of these assessment techniques are entirely satisfactory: the correlation between existing models has nothing to do with their relation to the true Earth; variance reduction depends strongly on the set-up of a specific inversion problem and varies from study to study; comparing tomography with *a priori* expectations of surface tectonics gives, at best, a qualitative judgment; the checkerboard test would only be valid if the theoretical framework used to build the synthetic data was perfect, but in practice, this is never the case.

An objective way of assessing tomographic models is to compare their predictions of seismic observables directly against recorded data, that is, to test how well tomographic models explain observed seismograms (whether or not used to derive the models themselves). The higher the predictive power of a 3-D model is, the better the model captures the true Earth's structures. For example, Ritzwoller & Lavely (1995) tested 3-D tomographic models against structure coefficients derived from observed normal mode splitting. Trampert & Woodhouse (2001) tested phase velocity models against a set of raw seismograms. Bozdag (2007) (http://www.spice-rtn.org/events/workshops/cargese2007/) assessed the reliability of tomographic mantle models derived with different regularization schemes by comparing the spectral element method (SEM) seismograms with raw observations, limited to the minor-arc and intermediate periods (>40 s), which are much affected by crustal structures.

Becker & Boschi (2002) conducted a comprehensive quantitative analysis of the similarities and differences between a suite of tomographic models. To visualize correlation as a function of depth and scale length, Becker & Boschi (2002) expanded each model into spherical harmonic coefficients and calculated correlation as a function of harmonic degree (http://geodynamics.usc.edu/~becker/tomography/). After analysing the degree-dependent correlation between all possible pairs of models, Becker & Boschi (2002) found that most models have good correlations only at relatively long wavelengths. We attempt here to assess how well the consistent large-scale features identified by Becker & Boschi (2002) represent true Earth's structures; with this goal, we measure the correlation between observations and predictions for long-period (>100 s), three-component records, including minor- and major-arc fundamental mode/overtone waves. This exercise requires that synthetic seismograms from 3-D earth models be calculated as accurately as possible. Fortunately, after the introduction of SEM (Komatitsch & Vilotte 1998) and coupled spectral element method (CSEM) (Capdeville *et al.* 2003) in seismology, synthetic seismograms can be calculated for complex 3-D anisotropic models with very little intrinsic numerical dispersion. For example, Tsuboi *et al.* (2004) used the SEM to assess the quality of tomographic model S20RTS with short-period data (>5 s). In this study, we use the CSEM to simulate 3-D global seismic wave propagations generated by deep and intermediate earthquakes in different tomographic models.

Because we are limiting ourselves to long-period data, we can use point sources as a good approximation of real, finite-extent ones. This is convenient, also because the reliability and quality of finite-extent rupture models is still an open question. In this regime, it is also legitimate to partly neglect the influences of uncertainties in our knowledge of crustal structure, which has more impact on short- and intermediate-period surface waves. In this paper, we first validate CSEM simulations and then introduce the numerical implementations and comparison results for four deep events and two shallow events and a suite of tomographic models of the earth's mantle, with a focus on recent models that include information from surface wave propagation: S20RTS (Ritsema & van Heijst 2000), SAW24B16 (Mégnin & Romanowicz 2000), SB4L18 (Masters *et al.* 2002) and Smean (Becker & Boschi 2002).

2 VALIDATION OF CSEM SIMULATIONS

We first check whether the effects of gravitation and oceans have been properly incorporated into the CSEM codes. We then evaluate the effects of the finite extent of sources on long-period surface waves. Last, we verify that numerical dispersion is negligible, by

comparing CSEM and normal-mode solutions for the spherically symmetric model Preliminary Reference Earth Model (PREM) (Dziewonski & Anderson 1981).

2.1 Overview of CSEM

The CSEM (Capdeville *et al.* 2003) is based on the partition of the Earth into a heterogeneous outer shell and an inner spherically symmetric sphere. The two domains are coupled through a spherical interface via a Dirichlet-to-Neumann (DtN) operator. For the inner spherically symmetric sphere, corresponding in this study to the earth's core, the solution is found by means of normal-mode theory so as to decrease the total computational cost. In the heterogeneous outer shell, the solution is sought in terms of the SEM. The SEM is a high-order polynomial version of the finite-element method and has very small intrinsic numerical dispersion (Chaljub *et al.* 2003). Therefore, the CSEM can also incorporate 3-D variations in seismic wave velocity, attenuation, anisotropy, ellipticity, topography, bathymetry and crustal thickness (Komatitsch & Vilotte 1998; Komatitsch & Tromp 2002b; Capdeville *et al.* 2003). Like SEM, the CSEM also uses approximations to the full physics of seismic wave propagation, such as neglect of self-gravitation, and a simplified treatment of attenuation and oceans. As the implementation of attenuation and gravitation has been tested in Komatitsch & Tromp (2002a, 2002b), their validation will be omitted here.

2.2 Effects of self-gravitation on surface wave propagation

As self-gravitation (Chaljub & Valette 2004) is not taken into account in the CSEM, it is important to test its effect on long-period ($T >100$ s) seismograms such as those considered here. We have compared normal-mode synthetic seismograms derived in self-gravitating versus non-self-gravitating regimes; see, for example, in Fig. 1 seismograms that were derived from anisotropic PREM for an event in New Britain at 68.2 km depth, recorded at station APEZ, 122° away. The seismograms are low-pass filtered with a corner period of 100 s. Fig. 1 confirms that the contribution of self-gravitation is negligible for periods considered here.

2.3 Effects of the finite extent of sources on long-period surface waves

Generally, earthquake faults rupture along a certain direction with a velocity equal to or lower than that of *S* waves propagating in the same materials. Consequently, for intermediate-to-large earthquakes, seismic waves generated at a breaking segment of the fault will arrive at each station before those generated at a segment that ruptures later, resulting in the azimuthal dependence of source time functions for finite faults (directivity). For stations located along the direction of rupture propagation, the source time function is very narrow and has higher amplitude. On the contrary, for stations located along the direction opposite to that of rupture propagation, the source time function is spread out and has small amplitude. The main issue with the simulation of seismic wave propagation from finite-extent sources is that, for most earthquakes, finite-extent fault models are either unavailable or known only with high uncertainty. Here, we model earthquakes as point sources; we validate this approximation below.

The effects of the finite extent of sources become weaker with increasing period of seismic waves. We therefore limited our simulations to a minimum period of 100 s. Although modelling preferably large events, corresponding to a higher signal-to-noise ratio, we
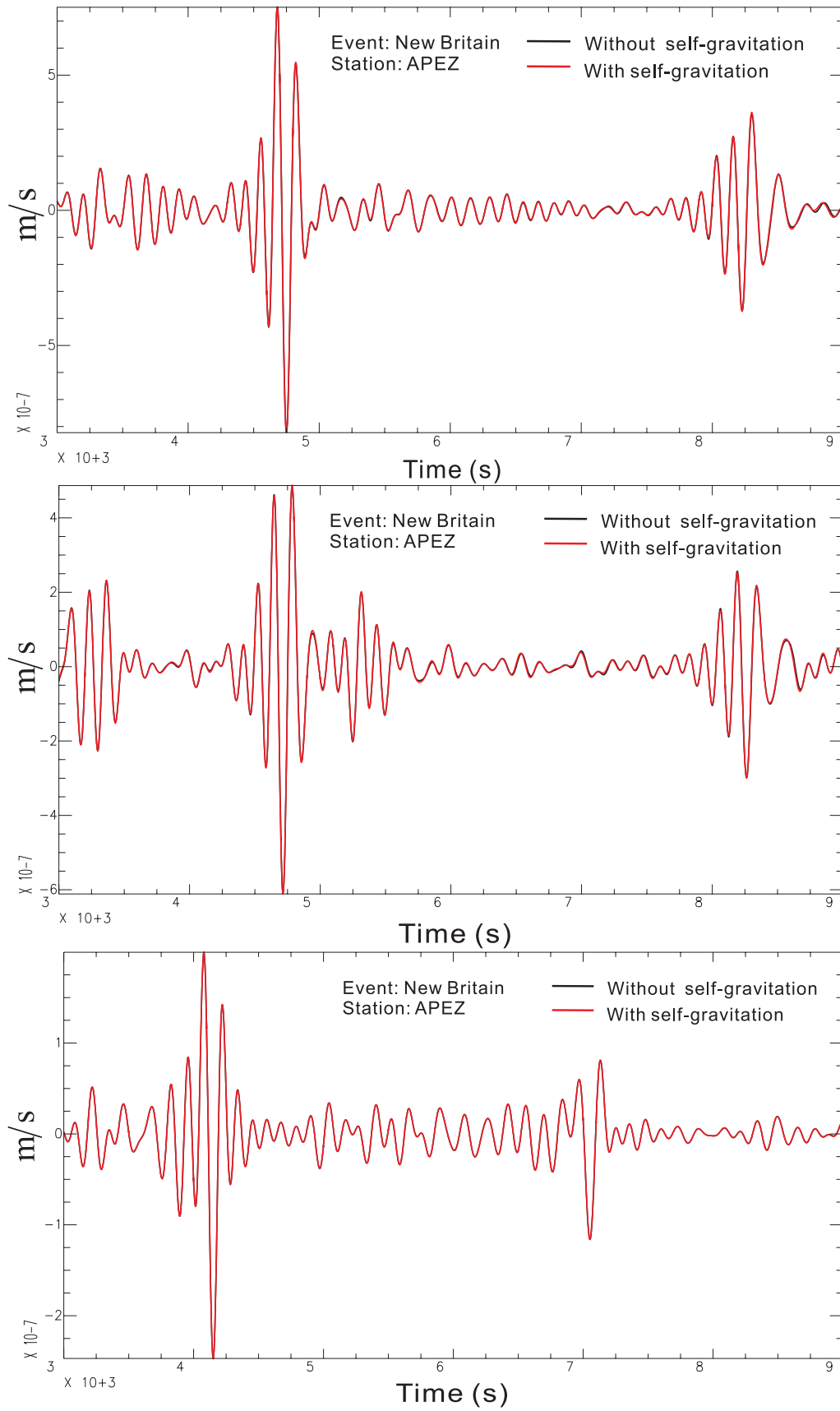
Figure 1. Comparison of vertical- (top), longitudinal- (middle) and transverse- (bottom) component displacement seismogram with full implementation of self-gravitation (red line) and without self-gravitation (black line) using normal-mode summation. The source is a shallow earthquake in New Britain recorded at station APEZ.
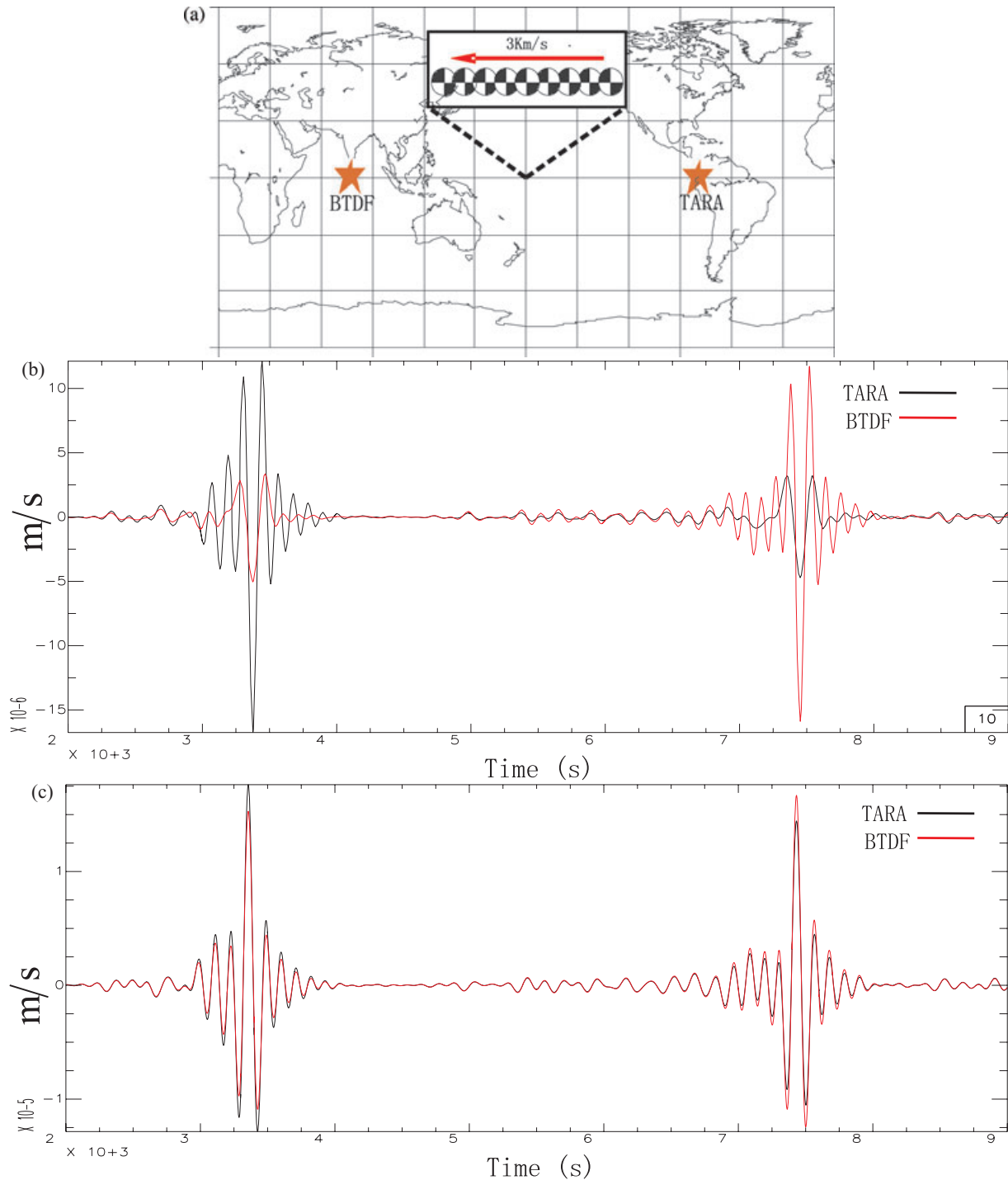
Figure 2. Synthetic test of the effects of finite-extent sources. (a) Acquisition geometry. Nine sub-sources are used to approximate a finite-extent, unilateral strike-slip source with a rupture velocity of 3 km s^{-1} from east to west. Stations BTDF and TARA, which have the same epicentral distance of 80°, are located at the east and the west of the source, respectively. (b) Comparison of records between station BTDF (red line) and station TARA (black line) for a finite-extent source with a duration of 64 s. The variation of amplitudes is consistent with the directivity of finite fault model. It can be deduced that, for long-period (>100 s) records, if the duration of fault rupture is larger than 64 s, the point source approximation is not appropriate. (c) Comparison of records between stations BTDF and TARA for a finite-extent source with a duration of 24 s. The records have been low-pass filtered with a corner frequency of 100 s. The small difference in synthetic seismograms between stations BTDF and TARA demonstrates that, for long-period surface waves (>100 s), if the duration of the finite-extent sources is less than 24 s, the finite-extent sources can be well approximated by a point source.

also try to pick them small enough so that the duration of fault rupture (proportional to magnitude) be short. To identify a good compromise value for magnitude, we computed synthetic seismograms for a finite-extent source with a series of different durations.

Following Tsuboi *et al.* (2004), we approximate one finite-extent source with nine subevents and with a rupture velocity of 3 km s^{-1}. As depicted in Fig. 2(a), the rupture propagates from east to west, and two stations are deployed to record seismograms, along the

direction of rupture propagation (BTDF), and along the opposite direction (TARA), at the same epicentral distance.

We compare in Fig. 2(b) synthetics predicted at BTDF versus TARA from the same finite-extent source with a duration of 64 s. The records have a lower corner frequency of 100 s. As the rupture propagates along the east–west direction, the minor-arc fundamental mode at station BTDF has significantly higher amplitude. At station TARA, conversely, the major-arc amplitude is much enhanced. We infer that, for rupture duration >64 s, the finiteness of fault cannot be ignored, or in other words, that the point source approximation is not valid, even if the minimum period of seismic records is >100 s. We repeat the comparison after reducing rupture duration to 24 s only (the rupture duration is reduced by decreasing the fault length). For both minor-arc and major-arc surface waves, the difference between seismograms recorded at BTDF and TARA is now very small (Fig. 2c). We infer that, if the duration of fault rupture is less than 24 s, it is reasonable to use a point source approximation to simulate synthetic seismograms with a period of 100 s and longer. Theoretically, the wavelength (about 400 km) is about five times the rupture length. According to the earthquake scale law (e.g. Lay & Wallace 1995), the 24-s long duration of rupture generally corresponds to an earthquake with a magnitude (M_w) of ∼7. In the following section, we shall simulate long-period seismic waves generated by deep events with an M_w of about 7.

2.4 Numerical dispersion

To simulate seismic waves with a minimum period of 100 s, we discretize 3-D earth models with an average grid distance of ∼30 km in the uppermost 80 km of the Earth, and ∼60 km elsewhere. The CSEM time step then equals 0.4 s. To validate the mesh and time step, the CSEM solution is compared with the normal-mode one for a one-dimensional (1-D) anisotropic PREM model, as shown in Fig. 3, in the neglect of attenuation. The good agreement between the normal-mode- and CSEM-predicted waveforms demonstrates that, with the designed meshes, the CSEM synthetics are accurate for a lower corner frequency of 100 s. Komatitsch & Tromp (2002a) verified that the effect of attenuation is reproduced well by spectral element calculations.

3 TESTED *S*-WAVE MODELS

In this study, we considered models S20RTS (Ritsema & van Heijst 2000), SAW24B16 (Mégnin & Romanowicz 2000), SB4L18 (Masters *et al.* 2000) and Smean (Becker & Boschi 2002) (Fig. 4). All these models incorporate information from surface waves, which is crucial for upper mantle structure.

All 3-D models are originally defined in terms of perturbations with respect to different background models. PREM (Dziewonski & Anderson 1981), a spherically symmetric global model, was derived from observations of normal-mode eigenfrequencies and Q values, traveltimes of P and S waves, mass and moment of inertia. PREM has vertical discontinuities at 80 km, 220 km, 400 km, 600 km and 670 km depth. Radial anisotropy is restricted to the uppermost 220 km, where the horizontal P-wave (horizontally polarized S wave) velocity (V_{PH} and V_{SH}, respectively) is 2–4% faster than the vertical P-wave (vertically polarized S wave) velocity (V_{PV} and V_{SV}, respectively).

3.1 S20RTS

S20RTS (Ritsema & van Heijst 2000) is a V_{SV} model of the earth's mantle. It is parameterized horizontally in spherical harmonics up

to degree 20 and radially with 21 vertical spline functions. Ritsema & van Heijst (2000) derived this model from the linear, ray theoretical inversion of observations of fundamental and higher-mode Rayleigh wave dispersion, teleseismic body wave traveltimes and normal-mode eigenfrequency splitting with damped least-squares (LSQR) method. They assumed that V_P heterogeneity be identical to V_S heterogeneity, except for a depth-dependent scaling factor, which increases linearly from 1.3 at the surface to 3.0 at the core–mantle boundary. The perturbation to the depth of discontinuities is not taken into account. A non-linear crustal correction was implemented based on model CRUST5.1. In the following section, we treat S20RTS as an isotropic V_S model, assuming V_{SH} heterogeneities to coincide with V_{SV} ones.

3.2 SAW24B16

SAW24B16 (Mégnin & Romanowicz 2000) is a V_{SH} model parameterized laterally in spherical harmonics up to degree 24 and radially with 16 unevenly spaced cubic B splines. This model was derived from handpicked transverse-component waveform data, including horizontally polarized body wave traveltimes, and minor- and major-arc Love waves with a stochastic LSQR inversion algorithm (Tarantola & Valette 1982). Moho depth and source parameters were also inverted, jointly with the tomographic inversion and starting from an *a priori* smooth (maximum harmonic degree = 12) crustal model. Mégnin & Romanowicz (2000) used broadband kernels for body wave waveforms and surface waves, derived via non-asymptotic coupling theory (Li & Romanowicz 1995; Li & Romanowicz 1996). As our modelling algorithm requires knowledge of both V_{SH} and V_{SV}, for this purely V_{SH} model, we assume $\delta \ln V_{SH} = \delta \ln V_{SV}$.

Compared with S20RTS, model SAW24B16 is characterized by much higher velocities at a depth of about 150 km under the Pacific Ocean. This is a result of SAW24B16 being a V_{SH} model based on Love wave data, whereas S20RTS is a V_{SV} model based on Rayleigh wave data, and of important radial anisotropy in the uppermost mantle (Montagner & Tanimoto 1991; Ekström & Dziewonski 1998; Boschi & Ekström 2002; Gung *et al.* 2003).
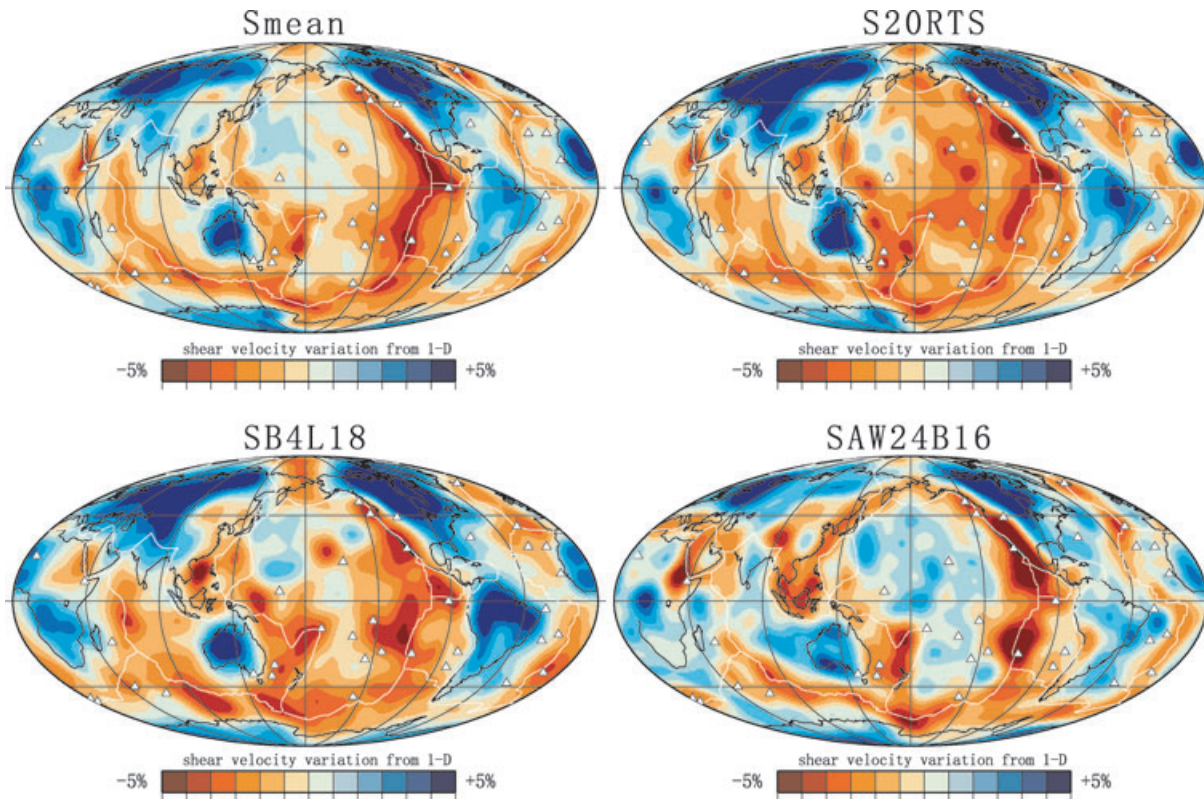
3.3 SB4L18

SB4L18 (Masters *et al.* 2000) is a V_S model. The model consists of 18 layers, with a thickness of ∼100 km in the upper mantle and transition zone and 200 km in the lower mantle. Laterally, each layer is divided into equal-area pixels, with a horizontal extent of $4° × 4°$ at the equator. SB4L18 was derived from observations of Rayleigh and Love wave dispersion, splitting coefficients of fundamental and higher-order spheroidal and toroidal modes and S-wave absolute/differential traveltimes with the LSQR algorithm. In the inversion process, perturbations in V_P and ρ were fixed to S-wave variations via the relationships $\delta \ln V_S = 1.7 × \delta \ln V_P$ and $\delta \ln V_S = 2.5 × \delta \ln \rho$. Crustal structure was fixed to model CRUST5.1 (Mooney *et al.* 1998).

3.4 Smean

Smean (Becker & Boschi 2002) is a composite V_S model that was derived in order to isolate robust, long-wavelength features by taking a weighted average of models NGRAND (Grand 1994; Grand *et al.* 1997), S20RTS and SB4L18. The weighting coefficients were depth-averaged root mean square powers (Becker & Boschi 2002). The derivation of Smean did not itself involve the direct inversion
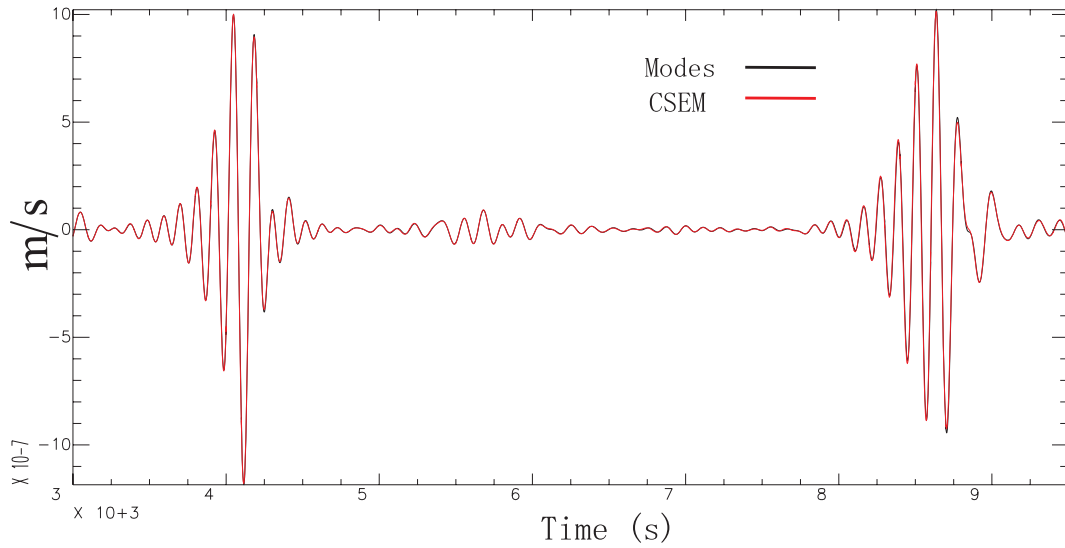
Figure 3. Comparisons of vertical components between CSEM solution and normal-mode solution with a minimum period of 100 s. The source is located in Tibet, and station ACSO is located in Alum Creek State Park, USA.

Figure 4. The velocity variations at a depth of 150 km for *S*-wave model Smean, *SV*-wave model S20RTS, *SV*-wave model SB4L18 and *SH*-wave model SAW24B16.

of seismic data; yet, compared with other published tomographic models of the mantle, Smean has been shown to provide the best fit of geodynamical observations (Steinberger & Calderwood 2006).

4 MODEL FORMAT

3-D tomographic models are parameterized differently (e.g. voxels, splines and spherical harmonics) and defined with respect to different 1-D reference profiles (e.g. ak135 and PREM). To unify

these formats, Becker & Boschi (2002) first converted all models to relative deviations from anisotropic PREM. Then, at a discrete set of depths, the model variations were expanded into spherical harmonic coefficients up to degree 31 as:

$$\delta V(\theta, \phi) = \sum_{l=0}^{31} \left[a_l^0 P_l^0(\theta) + \sqrt{2} \sum_{m=-l}^{m=l} P_l^m(\theta) \right.$$
$$\left. \times (a_l^m \cos m\phi + b_l^m \sin ml) \right],$$
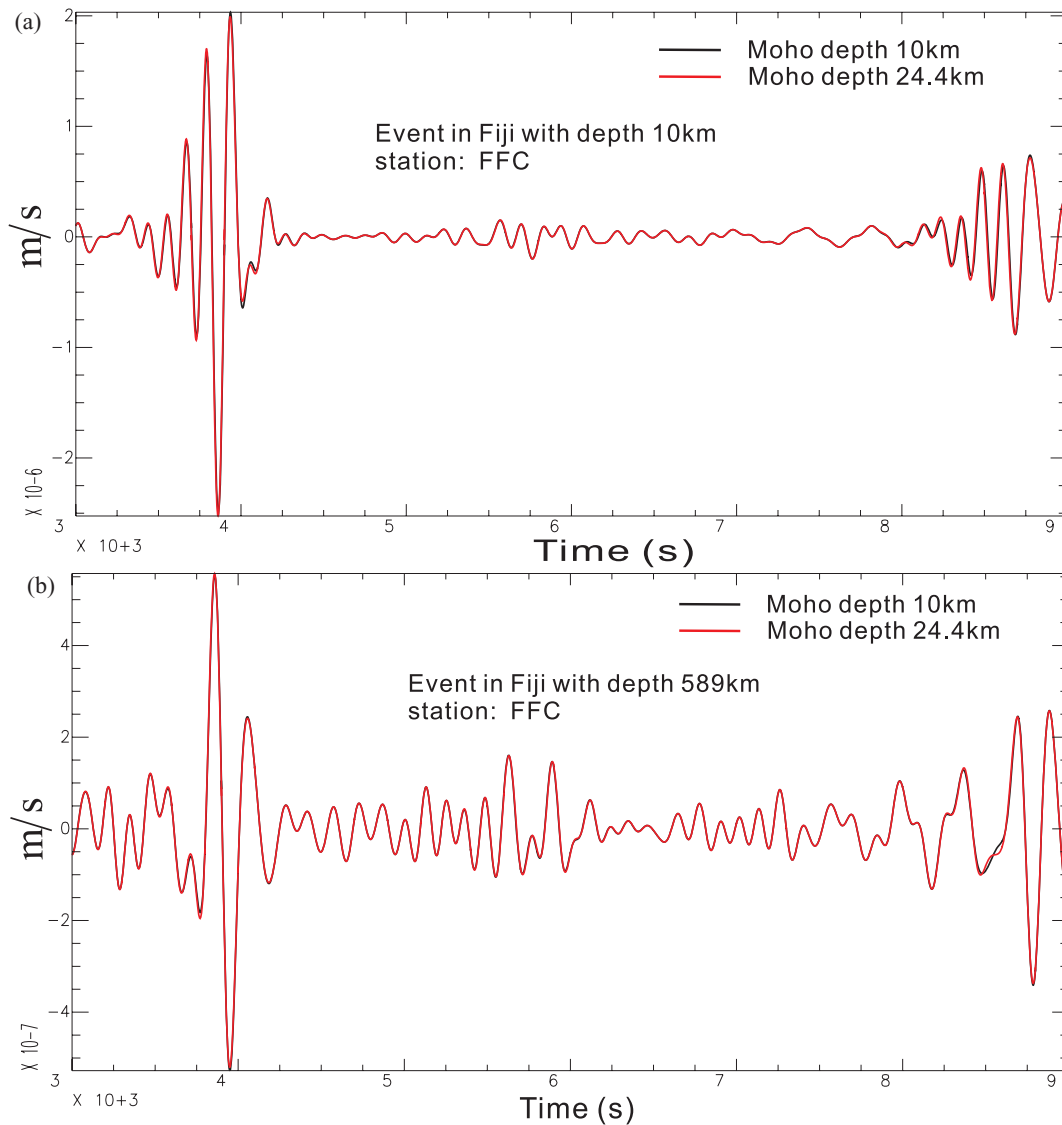
Figure 5. Test of the effect of replacing CRUST2.0 Moho depth under ocean with PREM Moho depth on the vertical component. Black line: the Moho depth is 10 km; red line: the Moho depth is 24.4 km. (a) The event is located underneath Fiji and has a depth of 10 km. (b) The event is also located underneath Fiji but has a depth of 589 km. In this test, we assume that the ocean is 3 km thick, and the crust has a thickness of 10 km with a $Vs = 3.8$ km s^{-1}.

where $\delta V(\theta, \phi)$ denotes relative velocity variation, $P_l^m(\theta)$ is the normalized associated Legendre function, and a_l^m and b_l^m are the spherical harmonic coefficients.

For discrete-layer parameterizations, Becker & Boschi (2002) found harmonic coefficients at the mean depth of each parameterization layer; for vertical spline parameterizations, they computed the coefficients at the original spline knots; for the Chebyshev polynomials, they computed the coefficients at equally spaced nodes, oversampling the resolution of the polynomials. With these unified spherical harmonic expansion model files, we can calculate velocity values at any point within the Earth by linear interpolations. Model coefficients in Becker & Boschi's (2002) format and a software to interpret them are available via the Seismic wave Propagation and Imaging in Complex (SPICE) tomographic model library (http://www.spice-rtn.org/research/planetaryscale/tomography/) and software repository (http://www.spice-rtn.org/library/software), respectively.

5 NUMERICAL IMPLEMENTATIONS

5.1 Incorporation of crustal models in CSEM

It is well recognized that surface wave dispersion is very sensitive to shallow structures, such as Moho depth, crustal velocity, bathymetry and topography, in non-linear and non-intuitive ways, even at long periods (Montagner & Jobert 1988; Boschi & Ekström 2002; Boschi *et al.* 2004). To simulate seismic wave propagation in a realistic 3-D earth model, it is very important to adequately incorporate a 3-D model of the earth's crust, which is challenging. For example, models CRUST2.0 (Bassin *et al.* 2000) and 3SMAC (Nataf & Ricard 1996) are both characterized by several very thin homogeneous layers with sharp contrasts in their elastic properties. To correctly approximate those velocity contrasts, each layer should be matched by element boundaries in the CSEM mesh. This dramatically reduces the minimum size of elements in the vertical direction, and therefore puts strong constraints on the time step required for the time integration schemes. As an alternative approach,

Fichtner & Igel (2008) proposed to smooth the thin near-surface layers via dispersion curve matching for a given period range. McManus (2007) proposed to use smoothed versions of crustal models and is now modifying a spectral element algorithm accordingly. Capdeville & Marigo (2007) presented a homogenization technique that effectively honours crustal effects and are now implementing this method in the CSEM. Unfortunately, neither algorithm is yet available.

Stutzmann & Montagner (1994) showed that, in the period range considered here, the effect of Moho depth is much larger than that of heterogeneity within the crust. In this study, we therefore model only the Moho discontinuity, ignoring other crustal velocity interfaces. At each location, we describe the crust with its thickness-weighted average velocity and density, derived from the global $2° \times 2°$ crustal model CRUST2.0 (Bassin *et al.* 2000). The Moho depth in CRUST2.0 varies by a factor of 4 between oceans and continents; to keep the mesh relatively simple and limit computational costs, under oceanic regions, we replace the CRUST2.0 value of Moho depth with that of PREM. Numerical experiments (Fig. 5) showed that the resulting error is negligible for periods considered here (100 s or more). In the horizontal direction, we smooth CRUST2.0 to suppress sharp transitions between neighbouring pixels (Fig. 6). Trying to mesh the crust more realistically would result, at this stage, in prohibitive computational costs.

5.2. Numerical computations

We combine the effects of self-gravitation, oceans, rotation, ellipticity, topography and bathymetry with a 3-D model to simulate true seismic data. We only consider *S*-wave velocity (V_S) models and find *P*-wave velocity (V_P) and density (ρ) variations from *S*-wave ones via the depth-independent scaling factors 0.588 and 0.4, respectively, that is, $\delta\ln V_P = 0.588\delta\ln V_S$, $\delta\ln\rho = 0.4 V_S$ (Karato 1993). Velocity and density perturbations are always defined with respect to anisotropic PREM. For V_{SV} or V_{SH} models, it was assumed that the V_{SV} anomalies are the same as the V_{SH} anomalies, that is, $\delta\ln V_{SH} = \delta\ln V_{SV}$. This assumption is consistent, for example, with Boschi & Ekström's (2002) approach in surface wave tomography. For V_S model, it was assumed that $\delta\ln V_{SH} = \delta\ln V_{SV} = \delta\ln V_S$. The quality factors are assumed to have no 3-D variations. Model ETOPO5 (from the US National Oceanic and Atmospheric Administration) for topography and bathymetry was also used to define the depth of oceans.

The simulations were implemented at the Institut de Physique du Globe de Paris (IPGP) on a PC cluster. To make sure that the second wave train of surface waves be properly modelled, all generated seismograms are 3-hr long.

6 SELECTIONS OF EVENTS AND TRACES

To test how different *S*-wave models explain not only fundamental mode data but also overtone data, we mostly modelled deep events, which excite higher modes and provide more information on the deeper parts of the Earth. We modelled only two shallow events, for comparison. As discussed above, to maximize the signal-to-noise ratio of available real data while minimizing event duration (<24 s), the M_w of events is fixed to ~7. As we want to compare waveform fitting for first higher modes (X1/G1), first fundamental modes (R1/L1), major-arc higher modes (X2/G2) and major-arc fundamental modes (R2/L2), only traces with epicentral distance
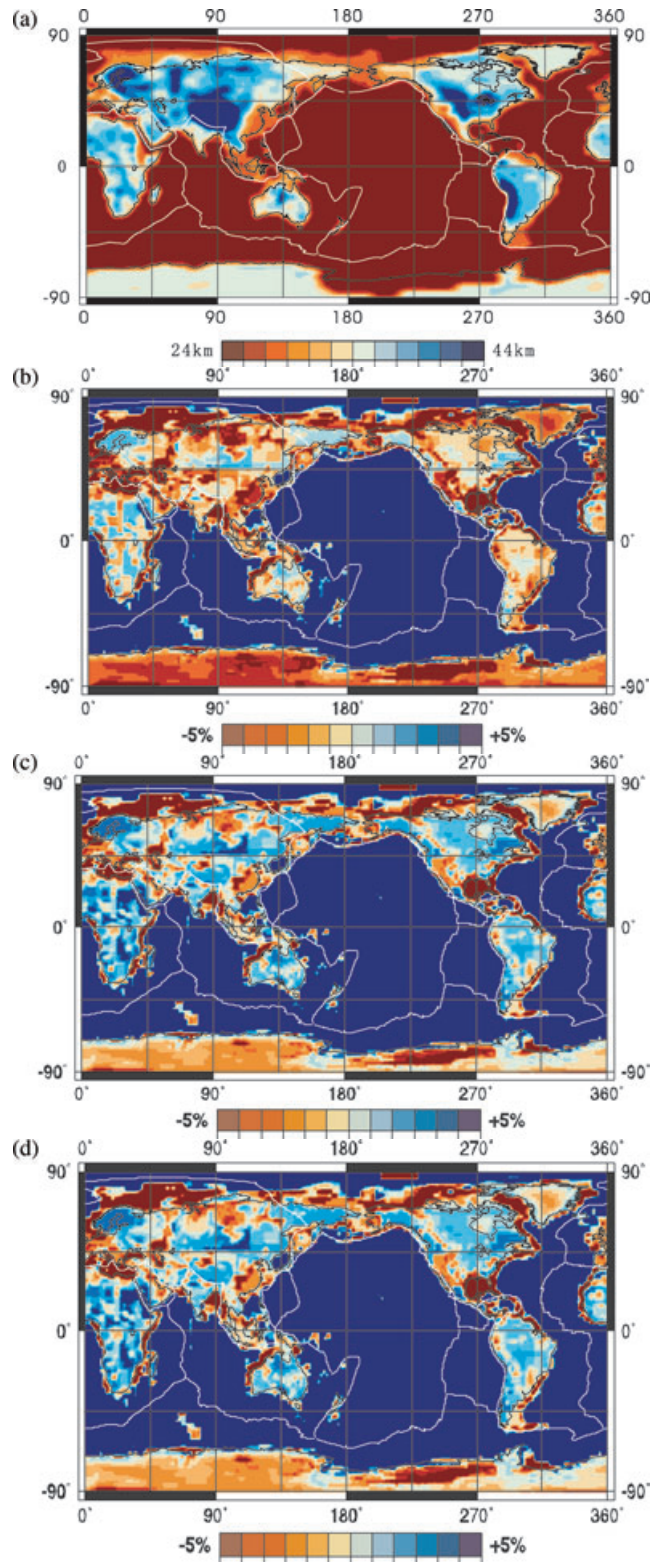
Figure 6. The crustal structures used in this study. (a) Crustal depth (with respect to sea level), (b) average *P*-wave velocity percentage variations in the crust (with respect to the average velocity of 6.5 km s^{-1}), (c) average *S*-wave velocity percentage variations in the crust (with respect to the average velocity of 3.6 km s^{-1}) and (d) average density percentage variations in the crust (with respect to the average density is 2.85 g cm^{-3}). Plate boundaries are drawn with white lines.
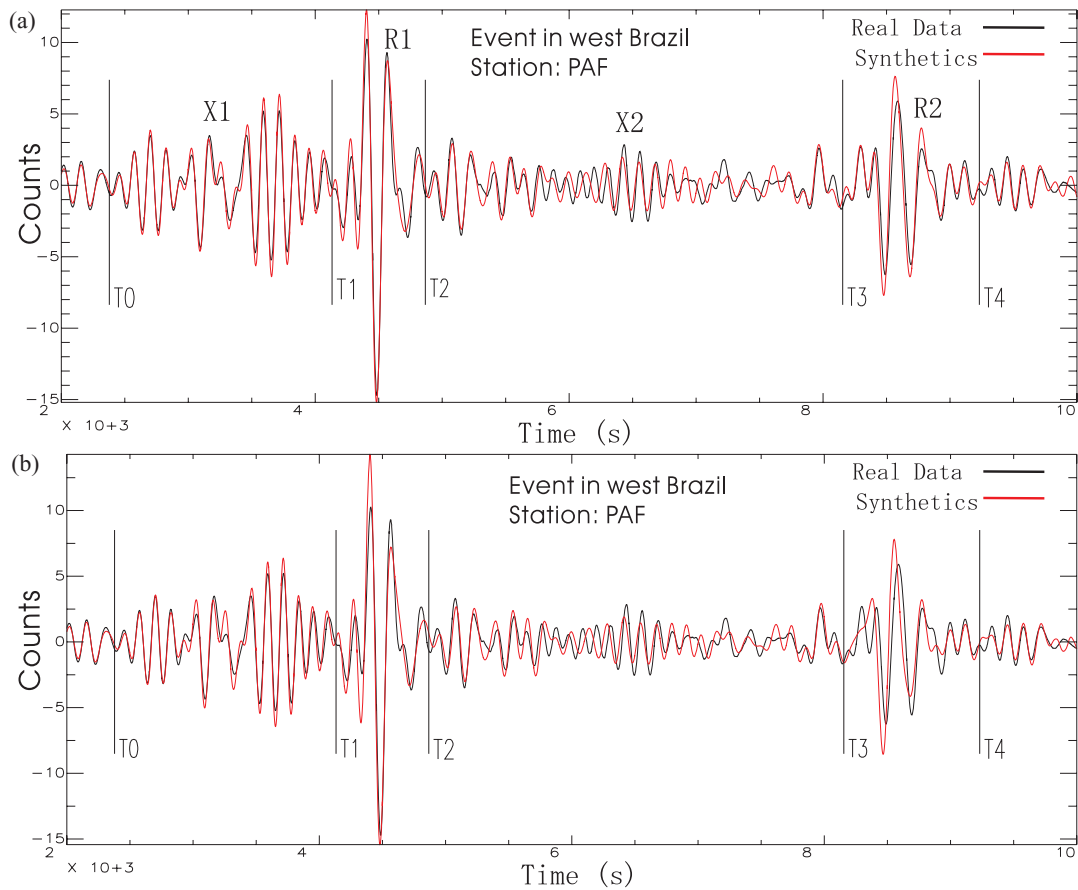
Figure 7. The waveform comparison between data (black line) and model prediction (red line). The handpicked windows include minor-arc overtones (T0-T1), minor-arc fundamental modes (T1-T2), major-arc overtones (T2-T3) and major-arc fundamental modes (T3-T4). (a) For model S20RTS, the correlation coefficient is 0.9817 for X1, 0.9791 for R1, 0.7715 for X2 and 0.9075 for R2. (b) For model SB4L18, the correlation coefficient is 0.9644 for X1, 0.9442 for R1, 0.7061 for X2 and 0.6558 for R2. The earthquake underlying this waveform occurred in Brazil on 2003 June 20 (depth 658 km) and was recorded by GEOSCOPE station PAF. Instrument response has been added to the synthetics.

between 50° and 140° are modelled. In this range of epicentral distances, fundamental mode surface waves are well separated in time domain from higher-mode ones, and the minor and major-arc surface waves do not overlap.

7 MANUAL 'PICKING' AND COMPUTATION OF THE CORRELATION AND TIME LAG

To identify the four different time windows corresponding to minor- and major-arc fundamental modes and overtones, we conducted five manual pickings per seismogram, as illustrated in Fig. 7. At that stage, we also controlled data quality via a visual comparison of the amplitude and phase of data and corresponding synthetics. We discarded real data that were largely and obviously inconsistent with the predictions of 3-D models. Before calculating the correlation between data and synthetics, it is essential to correct the former for instrument response. We convolved the instrument response with synthetic seismograms rather than deconvolving it from the real data, as deconvolution is sometimes an unstable procedure.

In view of the difficulties inherent in matching of amplitude, we employ the normalized correlation coefficient (Press *et al.*

1993)

$$\rho_{xy} = \frac{\sum\limits_{n=1}^{N_s} x_n y_n}{\sqrt{\sum\limits_{n=1}^{N_s} x_n^2 \sum\limits_{n=1}^{N_s} y_n^2}}$$

(where x_n and y_n are the nth samples on seismograms x and y, respectively) to measure the similarity between two waveforms.

An example of waveform comparison and corresponding ρ_{xy} is shown in Fig. 7. Note the high sensitivity of ρ_{xy} to phase shift (e.g. large loss of R2 correlation for model SB4L18 with respect to model S20RTS) and, in general, the consistency between ρ_{xy} and an intuitive visual analysis. An additional measure of consistency between data and synthetics consistency is 'time lag', defined as the time shift between a real and a predicted seismogram, which maximizes correlation ρ_{xy} between the two.

We systematically measure correlation and time lag with real data for all computed synthetic seismograms. In this study, the average cross-correlation coefficients and time lag are used for the comparisons of different models, because the averaging values is helpful to mitigate the error due to manual picking, visual choice of traces and low signal-to-noise ratio for some traces. Thus, the averaging value will be a less biased metric for the assessment

Table 1. Four deep earthquakes used for the comparisons between observations and synthetics.

No.	Event date	Time	Depth (km)	Latitude	Longitude	M_w	Half-duration (s)	Region name
1	2002.06.28	17:19:40	581.5	43.74	130.45	7.3	10.9	Eastern Russia–northeastern China border
2	2003.06.20	06:19:47	556.2	−7.37	−71.89	7.0	7.9	Western Brazil
3	2004.07.15	04:27:19	577.2	−17.68	−178.52	7.1	8.2	Fiji island region
4	2006.01.02	22:13:44	589.5	−19.80	−177.72	7.2	9.3	Fiji island region

of tomographic models. But the averaging values reduce the full usefulness of the measurements.

8 ASSESSMENT OF MODEL QUALITY FROM THE SIMULATION OF FOUR DEEP EARTHQUAKES

We simulate four different deep earthquakes, listed in Table 1, in each of the four V_S models described in section 6. As these events occurred after the publications of the tomographic models used in this study, the corresponding seismic recordings cannot have been used in the models' derivation. In the following section, the correlation between observed and predicted fundamental modes and overtones are discussed.

8.1 Eastern Russia–northeastern China border event

The eastern Russia–northeastern China border event in 2002 has an $M_w = 7.3$. Its geographic location gives us the opportunity to study both pure continental paths through Eurasia and pure oceanic paths through the Pacific. The source location and focal mechanisms are available from the Harvard Centroid Moment Tensor (CMT) catalogue. Fig. 8 shows the distribution of ray paths associated with seismic recordings of high signal-to-noise ratio. We note that: (1) The path coverage is extremely inhomogeneous, with more paths in the northern than in the southern hemisphere, and more paths in Europe and North America than over the rest of the globe. (2) There are more paths for the vertical components than for the horizontal ones, because the former have a systematically higher signal-to-noise ratio. (3) In oceanic regions, particularly few horizontal-component traces are available, as noise in oceanic regions is higher than that in continental regions.

The average ρ_{xy} for the different models and different wave trains (R1, R2, X1 and X2 for vertical and longitudinal components; L1, L2, G1 and G2 for transverse component) are shown in the left column of Fig. 9. It is immediately apparent that the differences in the average correlation achieved by different tomographic models are only minor. As expected, the average ρ_{xy} between PREM synthetics and data is systematically lower than between 3-D tomography-based synthetics and data. Correspondingly, the time lag (middle column of Fig. 9) between PREM synthetics and data is systematically larger than that associated with any tomographic model. Correlation (time lag) also systematically decreases (or grows) with increasing propagation distance (compare major-arc results with minor-arc results), suggesting that the cumulative discrepancy between tomography and real Earth structure grows with increasing travelled distance. The excellent fit to minor-arc seismograms suggests that augmenting the number of inverted minor-arc data would not significantly improve model quality; the poorer fit to major-arc seismograms indicates, on the other hand, that inversion of larger set of such data would be highly beneficial.

To assess how well seismogram amplitudes are predicted by tomography, the ratio between synthetic and observed amplitude was also calculated for different components and models, as shown in the right column of Fig. 9. Interestingly, we find that 3-D tomographic models do not fit amplitudes better than PREM, which may be ascribed to limitations in the simple 1-D attenuation model used in our simulation. Another possible factor is the significant drift in the instrument response for some stations, deteriorating the quality of the data that we are trying to fit.

8.2 Western Brazil event

We next model an $M_w = 7.1$ event in western Brazil, which occurred on 2003 June 20 at a depth of 556 km and show in Fig. 10 the corresponding measures of data-to-synthetics agreement, in the same fashion as in Fig. 9.

In agreement with Fig. 9, these results confirm that the agreement between observations and synthetics is improved for 3-D versus 1-D models, especially for major-arc surface wave trains. Here, and in Fig. 9, the performance of the purely V_{SH} model SAW26B16 in predicting transverse-component seismograms is surprisingly low; this can probably be explained in terms of the crustal structure associated with this model, importantly different from that implemented here. (This problem emerged during the preparation of this manuscript, and we have found it technically too difficult to reconstruct the crustal component of SAW26B16, which is not provided by the authors.) SAW24B16 nonetheless achieves an acceptable fit of vertical- and radial-component data.

8.3 2004 Fiji island region event

We next simulate a deep earthquake that occurred in the Fiji island region on 2004 July 15, at a depth of 577 km and with an $M_w = 7.1$ and a half-duration of 8.2 s. The corresponding average ρ_{xy} (Fig. 11) are particularly low for the vertical and radial components in SAW24B16; this is not surprising as, again, SAW24B16 is a purely V_{SH} model, and the simulated earthquake is located in the middle of a highly radially anisotropic region (e.g. Ekström & Dziewonski 1998; Boschi & Ekström 2002).

8.4 2006 Fiji island region event

To try and explain the mentioned low values of correlation associated with model SAW24B16, we simulated another deep event located in the Fiji region, very close to the one discussed in section 8.3, on 2006 January 2. Fig. 12 shows the corresponding source mechanism and path coverage. The recordings made at stations in California and Alaska are associated with paths that sample only the Pacific Ocean.

Fig. 13(a) shows the resulting average ρ_{xy} for different models and different components. Because SAW24B16 is a V_{SH} model, and because the upper mantle is more radially anisotropic under the Pacific Ocean than elsewhere (e.g. Ekström & Dziewonski 1998; Boschi & Ekström 2002), the average data-to-synthetics ρ_{xy} of minor-arc fundamental modes for this model is lower than that for the other
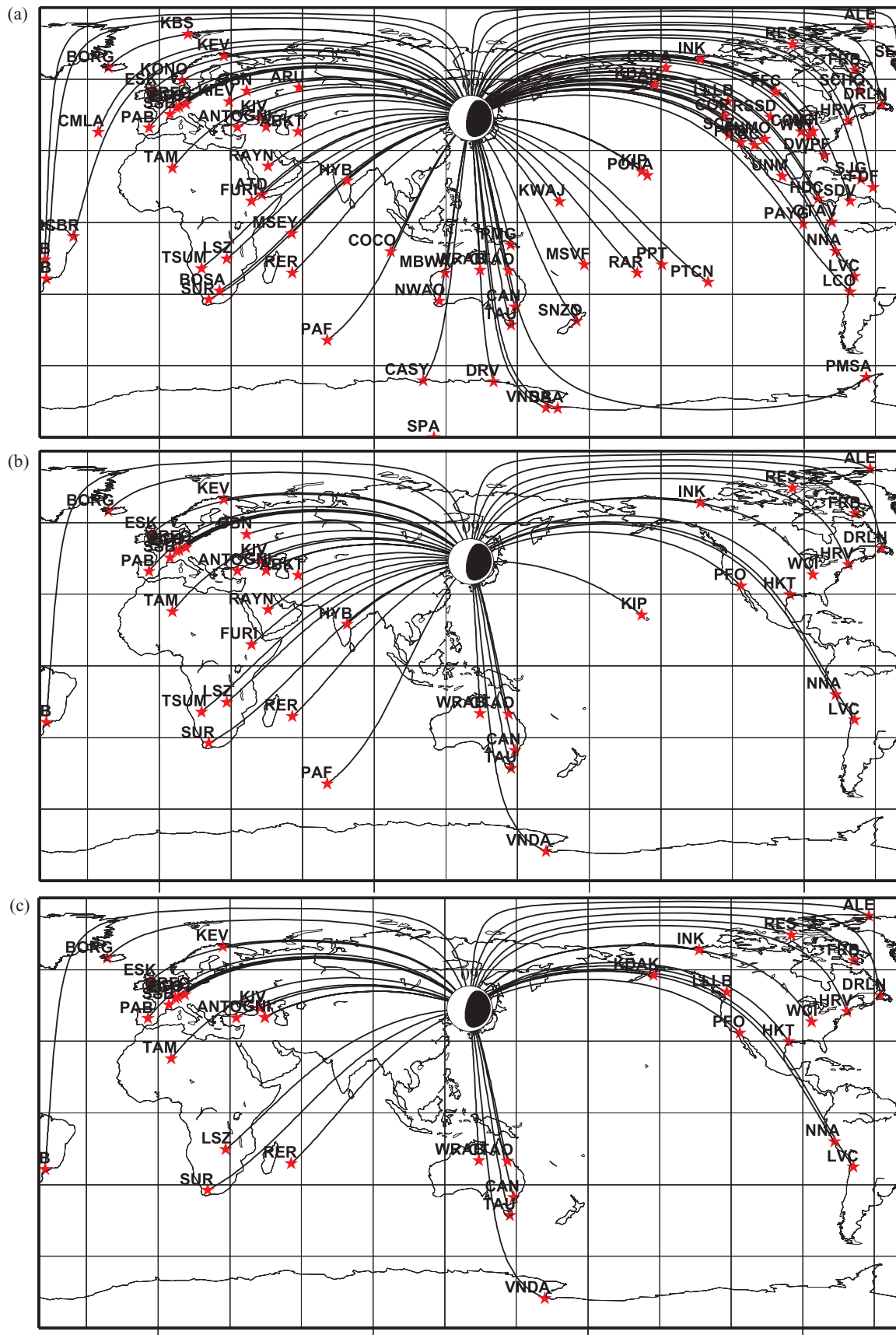
Figure 8. Path coverages for eastern Russia–northeastern China border event that occurred on 2002 June 28. (a) Vertical component, (b) longitudinal component and (c) transverse component. The stations are indicated by red stars. Solid black lines represent the minor-arc great-circle path. The beach ball shows the focal mechanism and is plotted at the epicenter.
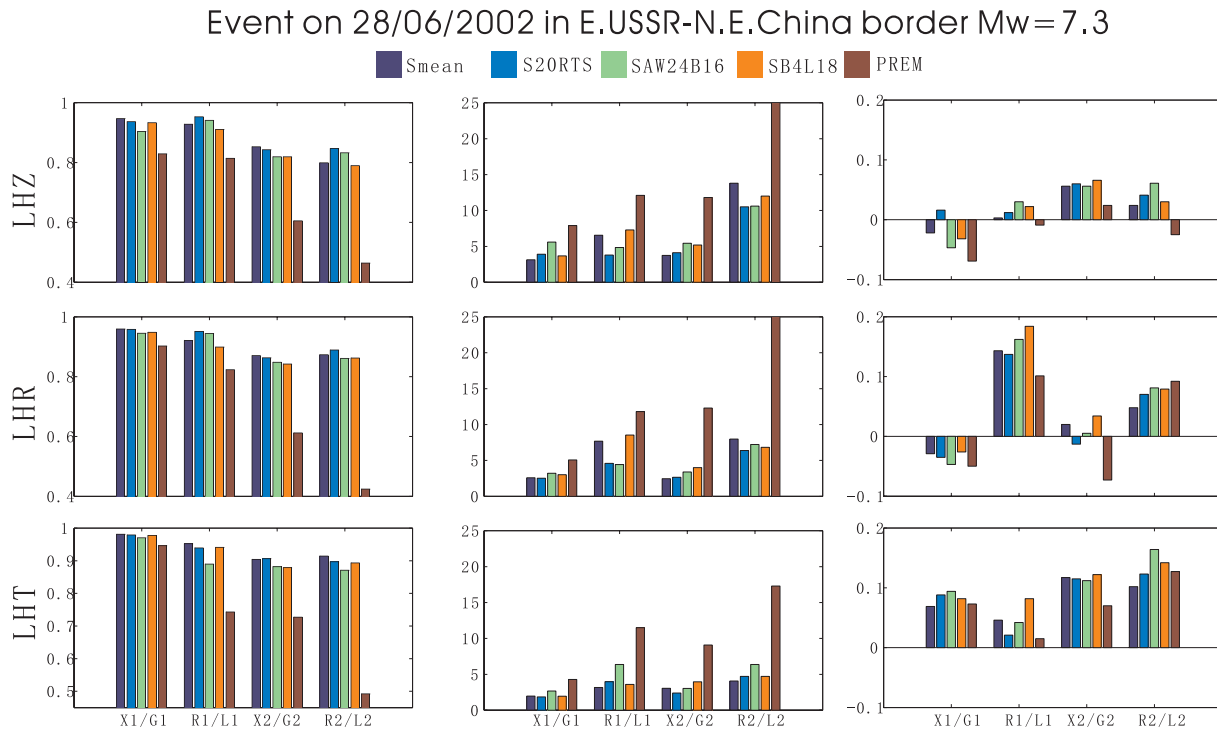
Figure 9. The average correlation coefficients (left column), time lag (middle column) and amplitude ratio (right column) for different components (R1, R2, X1 and X2 for vertical and longitudinal components; L1, L2, G1 and G2 for transverse component), different tomographic models and different time windows for a deep event (645 km deep) located in eastern Russian–northeastern China border which occurred on 2002 June 28. L1, minor-arc fundamental Love train; L2, major-arc fundamental Love train; R1, minor-arc fundamental Rayleigh train; R2, major-arc fundamental Rayleigh train; G1, minor-arc higher-mode Love train; G2, major-arc higher-mode Love train; X1, minor-arc higher-mode Rayleigh train; X2, major-arc higher-mode Rayleigh train. The results clearly demonstrated that the agreement in the phase between observations and synthetics is significantly improved by 3-D tomographic models by comparing with PREM model, but not in the amplitude. For the amplitude ratio in the right column, the value 1.0 has been removed.

3-D models, both on the vertical and on the radial components. Accordingly, modelled events with no or little path coverage of the Pacific had not given rise to a similar effect (sections 8.1 and 8.2).

In Fig. 14, we compare data (black line) and synthetics (red line), based on model S20RTS for vertical-, radial- and transverse-component recordings of the 2006 Fiji event, made at station LLLB in western North America. In Fig. 15, the comparison is repeated, with synthetics now derived from model SAW24B16. Comparing Figs 14 and 15, it is clear that S20RTS achieves a much better fit of minor-arc fundamental mode Rayleigh waves (R1) than SAW24B16. SAW24B16 explains well only minor-arc fundamental mode Love waves (L1), which however are explained well also by S20RTS. It is not surprising that SAW24B16 predicts Love waves, as it is a V_{SH} model derived from Love wave data. This problem is particularly severe in the Pacific Ocean, characterized by strong radial anisotropy (Ekström & Dziewonski 1998). Our findings from station LLLB are confirmed by an analogous analysis, conducted on synthetics and data recorded at station KDAK in Alaska. Again, a possible explanation for the relatively low performance of SAW24B16 resides in its associated crustal model, which is unavailable to us and is possibly inconsistent with the crustal model assumed here.

We specifically explore the implications of different crustal correction by recomputing synthetics for the Russia–China border and the 2006 January 2 events for model S20RTS combined with the 1-D crustal structure of PREM. The resulting time lags between data and synthetics are compared in Fig. 13(b) with those obtained from PREM and the previously implemented combination of S20RTS and CRUST2.0. Because PREM crust is a better approximation of

oceanic rather than continental crust, seismic phases associated with the Fiji event (right panels of Fig. 13b), largely sampling the Pacific region, are modelled almost equally well, regardless of the employed crustal model. On the other hand, phases modelled from the Russia–China border event, more sensitive to continental structure, match the data much less well if PREM crust rather than CRUST2.0 is employed. The above considerations on the relevance of crustal correction (model and implementation) are confirmed.

9 ASSESSMENT OF MODEL QUALITY FROM THE SIMULATION OF TWO SHALLOW EARTHQUAKES

In addition to the four deep events discussed above, we tested the same tomographic models on two shallow events. One occurred on 1997 November 8 in Tibet, with an $M_w = 7.5$ and a depth of 16.4 km. The same event has been used by Ferreira & Woodhouse (2007) in a validation of full ray theory for surface waves. The other occurred on 2005 June 15 in Northern California, with an $M_w = 7.2$ and a depth of 20.4 km.

9.1 Tibet event

The average ρ_{xy} and time lags between data and 3-D synthetics for the shallow Tibet event are shown in Fig. 16. ρ_{xy} is lower, and time lag is accordingly larger, than those obtained from deeper events and that are discussed above (compare Fig. 16 with Figs 9–11 and with Fig. 13). This effect is particularly large for major-arc wave

# Event on 20/06/2003 in west Brazil Mw = 7.0
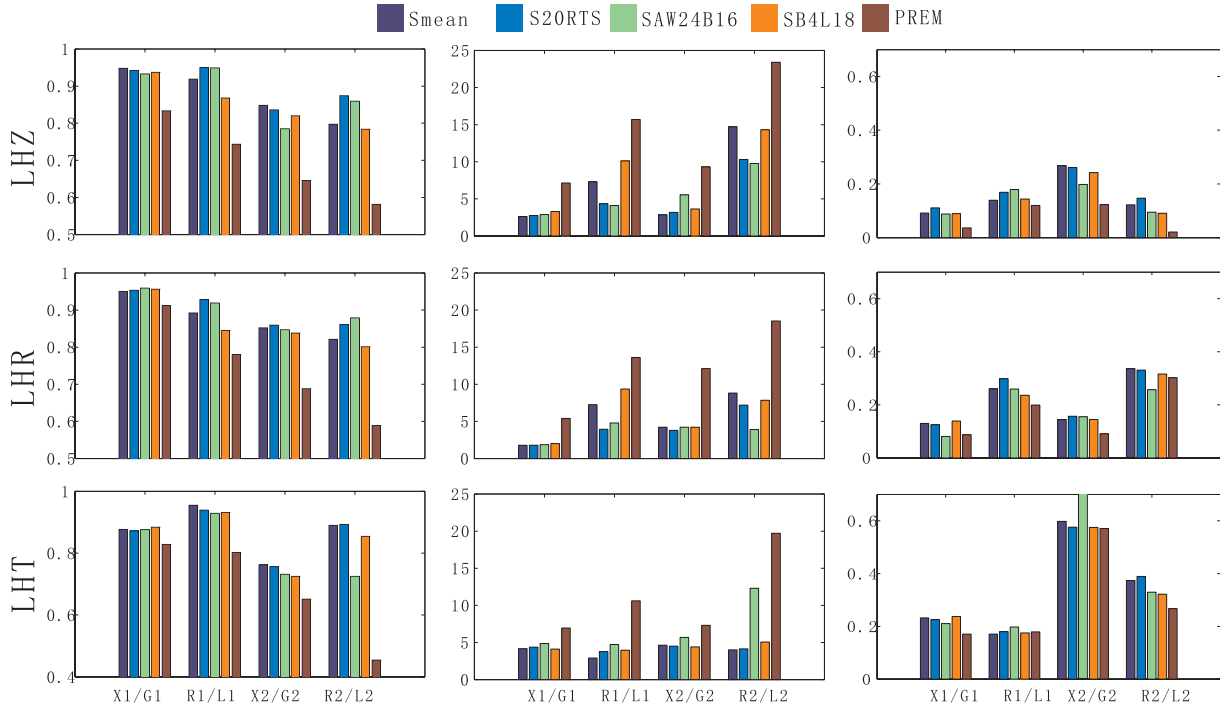
Figure 10. The average correlation coefficients (left column), time lag (middle column) and amplitude ratio (right column) for different component, different tomographic models and different time windows for the event located in western Brazil on 2003 June 20. The X1, X2, R1, R2, L1, L2, G1 and G2 have the same meaning as indicated in Fig. 9.
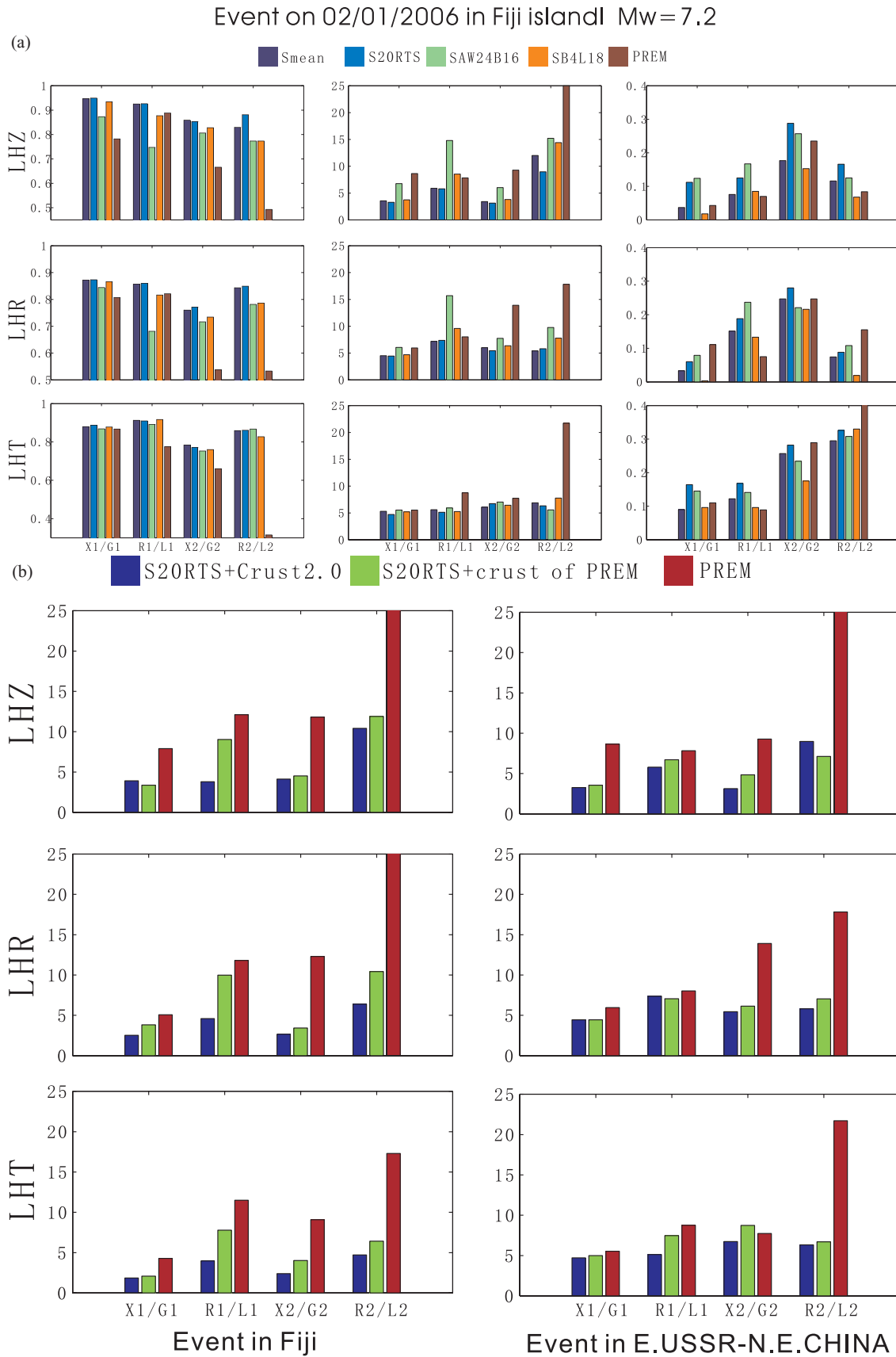
Event on 15/07/2004 in Fiji islandl Mw = 7.1
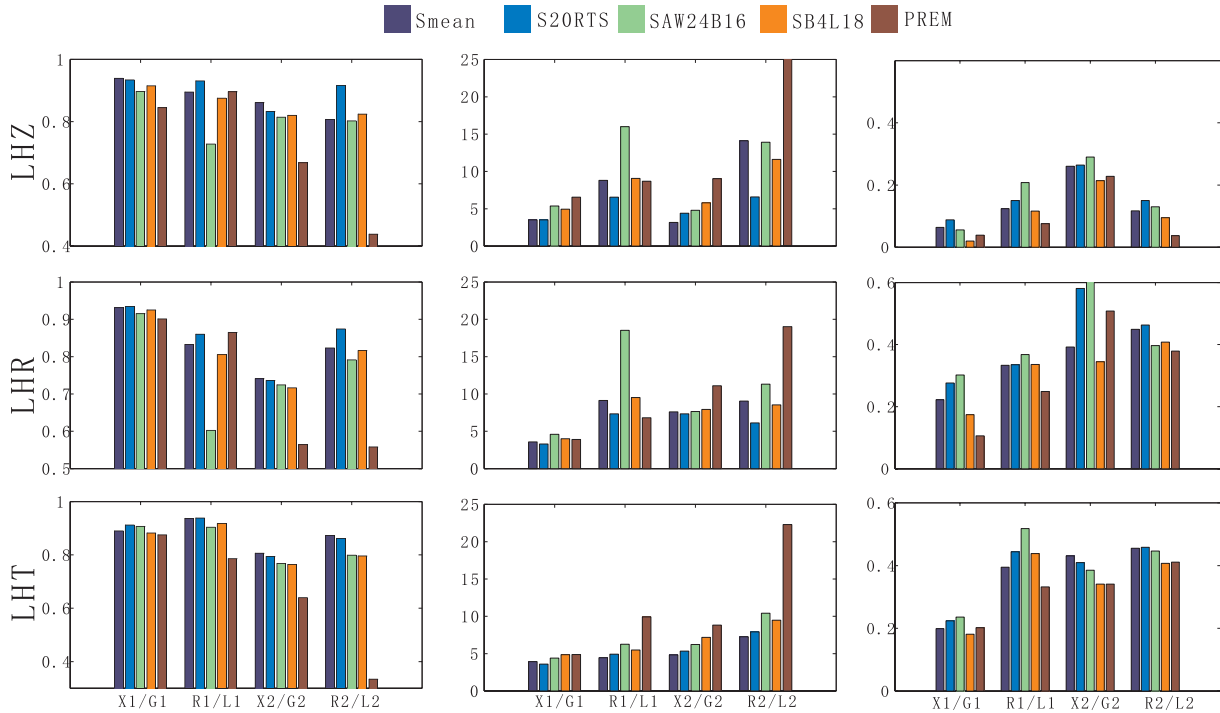
Figure 11. Same as Fig. 10, but for the event located in Fiji island region on 2004 July 15.
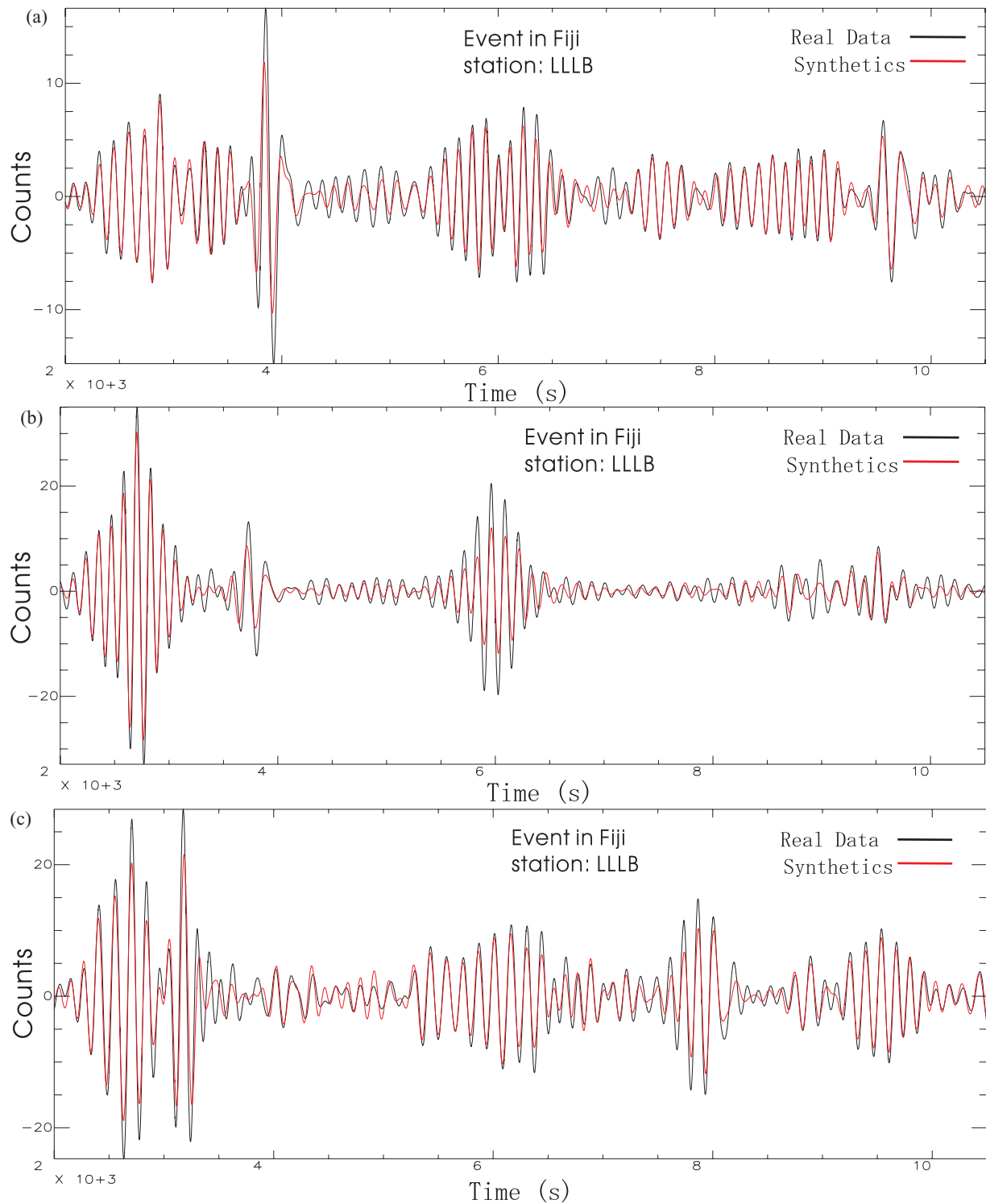
(a)

(b)

(c)
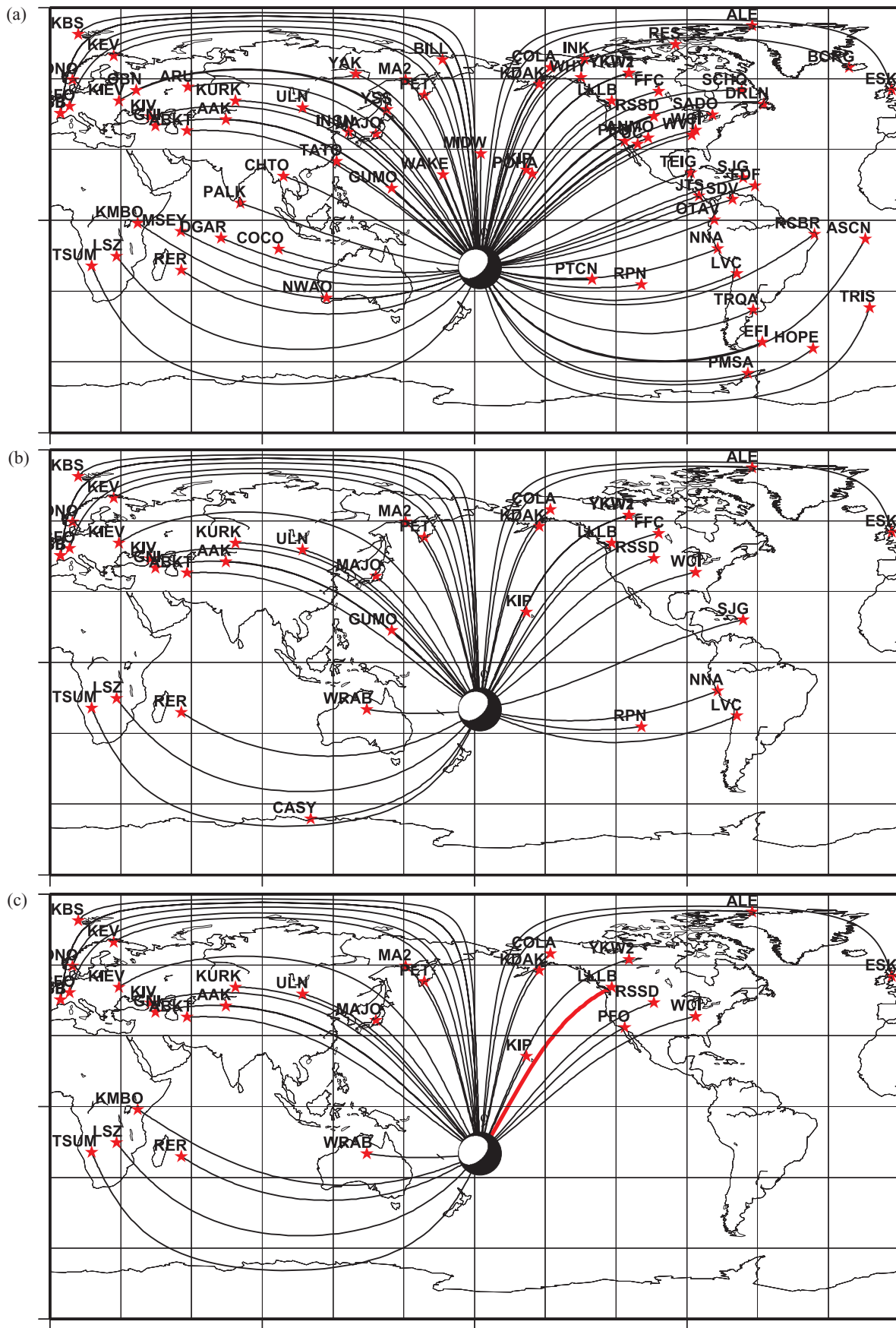
Figure 12. The path coverage for the event located in Fiji region in 2006 January 2. (a) Vertical component, (b) longitudinal component and (c) transverse component.

Figure 13. (a) Same as Fig. 9, but for the event located in Fiji island region on 2006 January 2. (b) Comparison of the average time lag between three models: S20RTS with average CRUST2.0 in the crust, S20RTS with PREM in the crust, and PREM. The left panel is for the event located in eastern Russian–northeastern China border on 2002 June 28, and the right is for the event located in Fiji island region on 2006 January 2.

Figure 14. Comparison of observations (black line) and synthetics (red line) with S20RTS model for station LLLB (western USA). The event is located in Fiji region as indicated in Fig. 12(c). The minor-arc great-circle path crosses the Pacific ocean. (a) Vertical component, (b) longitudinal component and (c) transverse component.

trains. As surface waves generated by shallow events are more sensitive to shallow (crustal) structure, the most likely explanation is that, even at the relatively long periods that we considered, crustal structure plays a very important role in governing surface wave propagation. The smooth crustal model (section 3.1) that we incorporated in the CSEM is adequate to simulate deep earthquakes,

but not shallow ones. In the future, computational costs permitting, more accurate CSEM synthetics will have to be generated after a finer meshing of the shallowest regions of the Earth. For higher-mode seismograms, the comparison is complicated by the lack of energy in these modes, which are not much excited by shallow events.
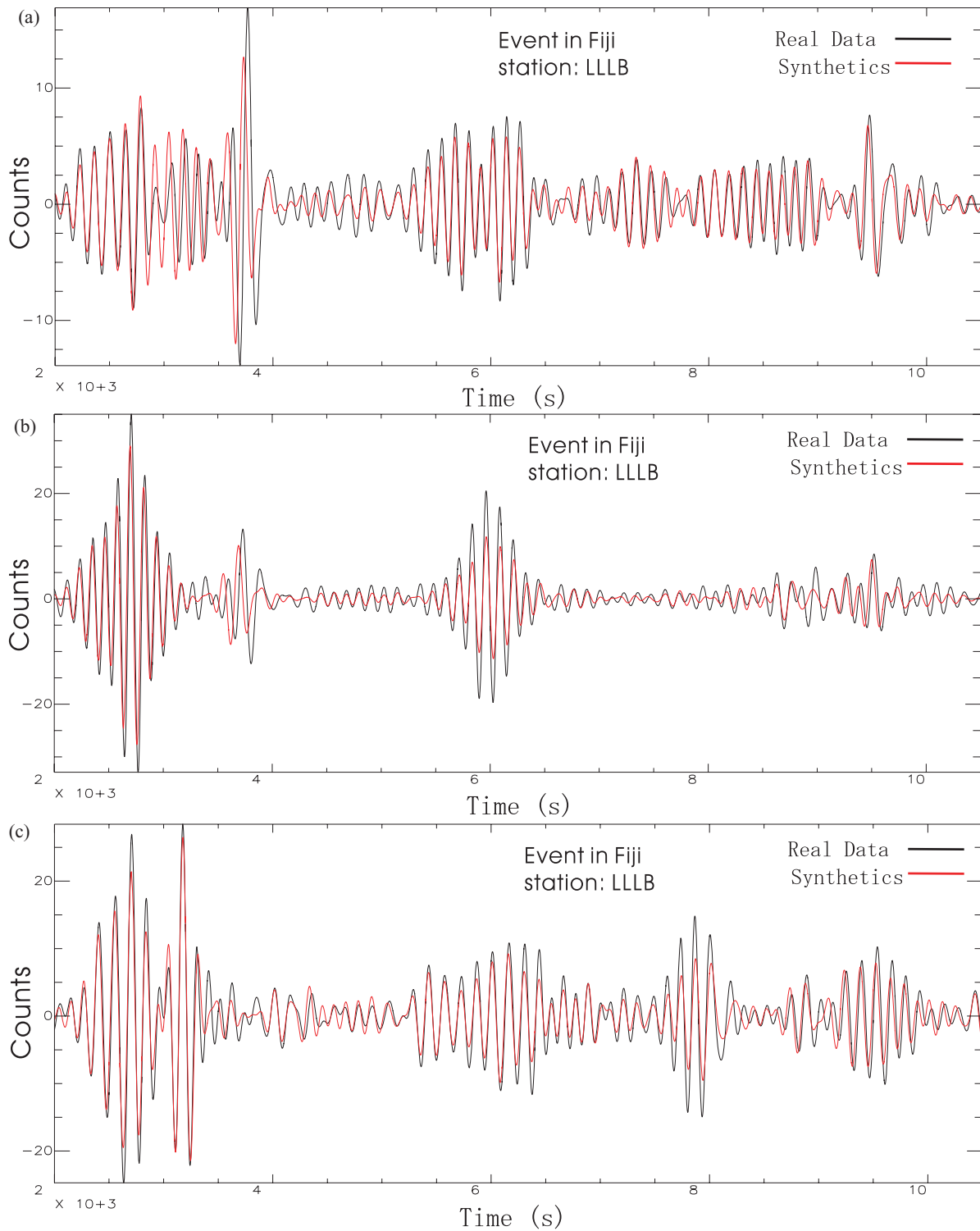
Figure 15. The same kind of comparison as in Fig. 14, but for model SAW24B16.
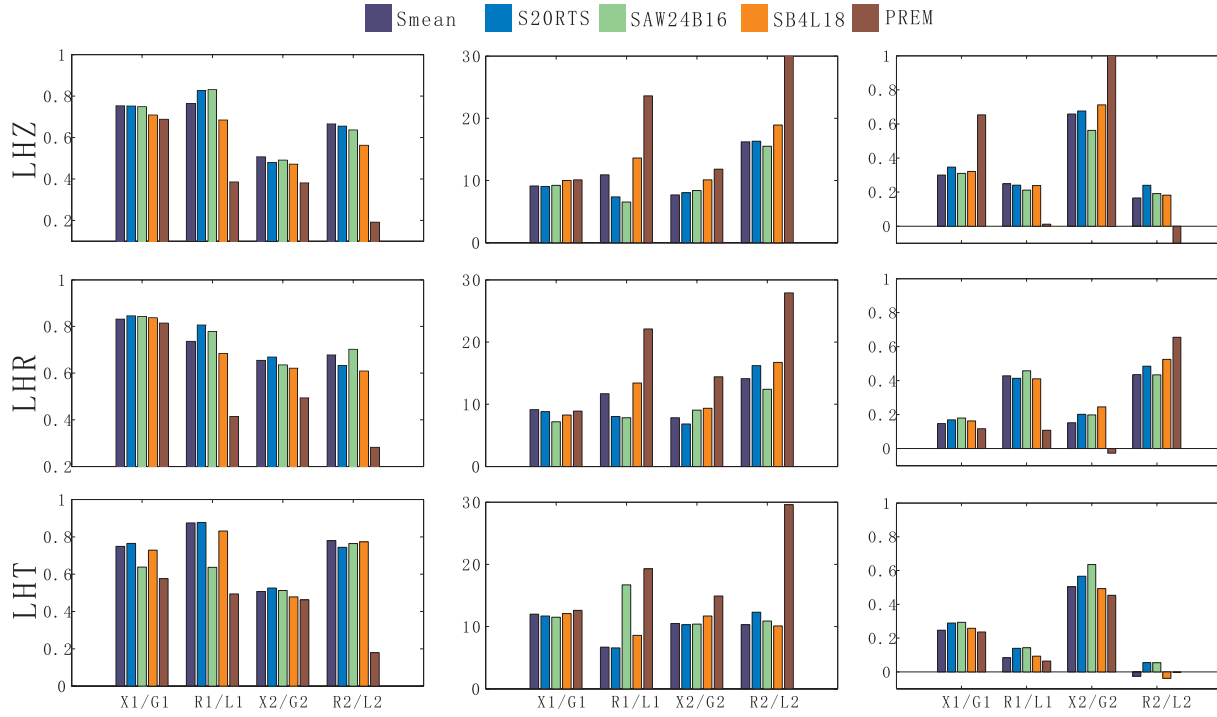
9.2 California event

The purpose of modelling this event is to investigate, again, the implication of having predominantly oceanic source station paths. The results of this experiment are illustrated in Fig. 17, which confirms the inferences that we made on the basis of the Fiji events discussed above, that is, SAW24B16 performs less well than S20RTS on the vertical and radial components along purely oceanic Pacific paths.
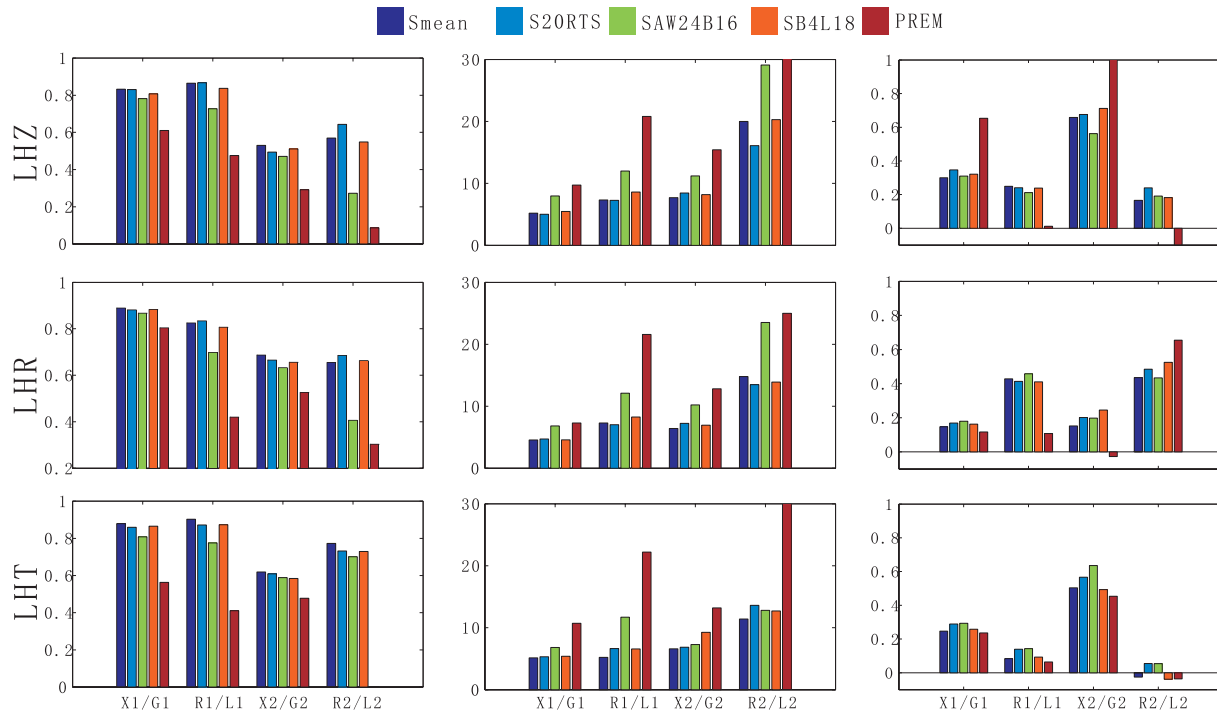
10 DISCUSSION AND CONCLUSION

We have simulated long-period (>100 s) synthetic seismograms for four deep and two shallow events on a suite of recently published

Event on 08/11/1997 in Tibet Mw=7.5

Figure 16. Same as Fig. 10, but for the event located in Tibet on 1997 November 8.

Event on 15/06/2005 in California Mw=7.2

Figure 17. Same as Fig. 10, but for the event located in California on 2005 June 15.

tomographic models and compared quantitatively the resulting synthetics with real seismic data. Our experiment represents the first attempt at a comparative, quantitative assessment of the quality of tomographic models that is not biased by the application of approx-

imate theories (ray theory, 'finite-frequency' methods) but instead takes advantage of a powerful numerical algorithm for the realistic simulation of wave propagation. Generally, the average correlation coefficients for different events and models demonstrate that the

tested 3-D seismic tomographic models can explain the observed surface waves, at periods >100 s, better than PREM. But our results also demonstrate that there is room for the improvement of the large-scale features of global tomographic models, not only for the phase but also for the amplitude. It can be inferred that the overall large-scale picture of upper-mantle heterogeneities has been properly captured by current 3-D tomography. Different 3-D models are, in general, similarly effective in fitting observed seismograms, with a somewhat better performance of S20RTS and Smean over SB4L18 and SAW24B16. A problem that emerged with model SAW24B16 suggests that the crustal structure used in this study did not sufficiently approximate the one associated with that model.

Our comparisons must be interpreted while keeping in mind that they are limited by several factors: (1) the extreme sparsity and inhomogeneity of seismic ray path coverage; (2) the lack of a definitive crustal model, and the difficulty of re-parameterizing the crust at each simulation, to account for different crustal models assumed by different tomographers; (3) the restriction of our analysis to long periods (>100 s) to limit computational costs; (4) uncertainties in CMT solutions, and the inherent inaccuracy of the point source approximation; (5) the employment of simple, constant scaling factors to derive V_P and ρ from V_s tomography and (6) the radical simplification inherent in our spherically symmetric model of the quality factor. For these reasons, the model quality assessment provided by our exercise should not be understood as an ultimate rating of the models but rather as a first stab at the problem, which can guide future mapping efforts.

ACKNOWLEDGMENTS

REFERENCES

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

Boschi, L. & Ekström, G., 2002. New images of the Earth's upper mantle from measurements of surface-wave phase velocity anomalies, *J. Geophys. Res.,* **107,** 2059, doi:10.1029/2000JB000059.

Boschi, L., Ekström, G. & Kustowski, B., 2004. Multiple resolution surface wave tomography, the Mediterranean Basin, *Geophys. J. Int.,* **157,** 293–304.

Bozdag, E., 2007. Assessment of the tomographic mantle models using SEM seismogram. In *the 4th SPICE Research and Training Workshop, at the Institut d'Etudes Scientific de Cargèse in Cargèse, Corsica.*

Capdeville, Y. & Marigo, J.J., 2007. Second order homogenization of the elastic wave equation for non-periodic layered media, *Geophys. J. Int.,* **170**(2), 823–836.

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. & Valette B., 2004. Spectral element modelling of three-dimensional wave propagation in a self-gravitating Earth with an arbitrarily stratified outer core, *Geophys. J. Int.* **158**(1), 131–141.

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. Comput. Phys.,* **187,** 457–491.

Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary Reference Earth Model (PREM), *Phys. Earth Planet. Inter.,* **25,** 297–356.

Ekström, G.A. & Dziewonski, A.M., 1998. The unique anisotropy of the Pacific upper mantle, *Nature,* **394,** 168–172.

Ferreira, A.M.G. & Woodhouse, J.H., 2007. Source, path and receiver effects on seismic surface waves, *Geophys. J. Int.* **168**(1), 109–132.

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), *Geophys. J. Int.,* **173**(2), 519–533.

Grand, S.P., 1994. Mantle shear structure beneath the Americas and surrounding oceans, *J. Geophys. Res.,* **99,** 11591–11621.

Grand, S.P., Van Der Hilst, R.D. & Widiyantoro, S., 1997. Global seismic tomography: a snapshot of convection in the Earth, *GSA Today,* **7**(4), 1–7.

Gung, Y.C., Panning, M. & Romanowicz, B., 2003. Anisotropy and thickness of the lithosphere, *Nature,* **422,** 707–711.

Karato, S., 1993. Importance of inelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.* **20,** 1623–1626.

Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation, Part I: validation, *Geophys. J. Int.,* **149,** 390–412.

Komatitsch, D. & Tromp, J., 2002b. 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.

Lay, T. & Wallace, T.C., 1995. Earthquake Kinematics and Dynamics, in *Modern Global Seismology,* Academic Press, San Diego, p. 390.

Li, X.B. & Romanowicz, B., 1995. Comparison of global waveform inversions with and without considering cross-branch modal coupling, *Geophys. J. Int.* **121,** 695–709.

Li, X.B. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. Geophys. Res.,* **101,** 11245–22271.

Masters, G., Laske, G., Bolton, H. & Dziewonski, A., 2000. The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: implications for chemical and thermal structure, in *Earth's Deep Interior,* pp. 63–87, eds. Karato S., Forte A.M., Liebermann R.C., Masters G. & Stixrude L., AGU Monograph 117, AGU, Washington, DC.

McManus, E. 2007. A method for simplifying crustal structure for the purpose of three dimensional wavefield modelling, in *the SPICE Workshop,* Corsica.

Mégnin, C. & Romanowicz, B., 2000. The shear velocity structure of the mantle from the inversion of body, surface, and higher modes waveforms, *Geophys. J. Int.,* **143,** 709–728.

Montagner, J.P. & Jobert, N., 1988. Vectorial tomography II: application to the Indian ocean, *Geophys. J. Roy. Astr. Soc.,* **94,** 309–344.

Montagner, J.P. & Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies, *J. Geophys. Res.,* **96**(B12), 20337–20351.

Mooney, W.D., Laske, G. & Masters, G., 1998. Crust 5.1: a global crustal model at 5 × 5 degrees, *J. Geophys. Res.,* **103,** 727–747.

Nataf, H. & Ricard, Y., 1996. 3SMAC: an a priori tomographic model of the upper mantle based on geophysical modeling, *Phys. Earth Planet. Int.,* **95,** 101–122.

Press, W.H., Flannery, B.P., Teukolsky S.A. & Vetterling, W.T., 1993. *Numerical Recipes in FORTRAN 77: The Art of Scientific Computing,* 2nd edn, Cambridge University Press, Cambridge.

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.

Ritzwoller, M.H. & Lavely, E.M., 1995. Three-dimensional seismic models of the Earth's mantle, *Rev. Geophys.,* **33,** 1–66.

Steinberger, B. & Calderwood, A.R., 2006. Models of large-scale viscous flow in the Earth's mantle with constraints from mineral physics and surface observations, *Geophys. J. Int.,* **167,** 1461–1481.

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.

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

Trampert, J. & Woodhouse, J.H., 2001. Assessment of global phase velocity models, *Geophys. J. Int.,* **144,** 165–174.

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.