Geophysical Journal International

Geophys. J. Int. (2010) **182,** 1025–1042 doi: 10.1111/j.1365-246X.2010.04667.x

Frequency-dependent effects on global *S*-wave traveltimes: wavefront-healing, scattering and attenuation

Christophe Zaroli,[1] Eric Debayle[1]* and Malcolm Sambridge[2]

[1]*Institut de Physique du Globe de Strasbourg (UMR 7516 CNRS, Université de Strasbourg/EOST), 5 rue René Descartes, 67084 Strasbourg Cedex, France. E-mail: Christophe.Zaroli@eost.u-strasbg.fr*
[2]*Research School of Earth Sciences, Australian National University, Canberra ACT 0200, Australia*

Accepted 2010 May 17. Received 2010 April 2; in original form 2009 August 4

SUMMARY

We present a globally distributed data set of ∼400 000 frequency-dependent *SH*-wave traveltimes. An automated technique is used to measure teleseismic *S, ScS* and *SS* traveltimes at several periods ranging from 10 to 51 s. The targeted seismic phases are first extracted from the observed and synthetic seismograms using an automated time window algorithm. Traveltimes are then measured at several periods, by cross-correlation between the selected observed and synthetic filtered waveforms. Frequency-dependent effects due to crustal reverberations beneath each receiver are handled by incorporating crustal phases into WKBJ synthetic waveforms.

After correction for physical dispersion due to intrinsic anelastic processes, we observe a residual traveltime dispersion on the order of 1–2 s in the period range of analysis. This dispersion occurs differently for *S, ScS* and *SS*, which is presumably related to their differing paths through the Earth. We find that: (1) Wavefront-healing phenomenon is observed for *S* and to a lesser extent *SS* waves having passed through very low velocity anomalies. (2) A preferred sampling of high velocity scatterers located at the CMB may explain our observation that *ScS* waves travel faster at low-frequency than at high-frequency. (3) A frequency-dependent attenuation $q(\omega) \propto q_0 \times \omega^{-\alpha}$, with $\alpha \sim 0.2$, is compatible with the globally averaged dispersion observed for *S* waves.

Key words: Body waves; Seismic attenuation; Seismic tomography; Wave scattering and diffraction.

GJI Seismology

1 INTRODUCTION

Seismic tomography is a standard tool for constraining the structure of the Earth's interior. The resolution of global body wave seismic tomographic models has significantly improved in the last 25 years because of the growth in the number of seismic stations, increase in computational power and development of new analysis tools which extract more information from seismograms. Until recently, ray theory (RT) formed the backbone of all body wave tomographic studies. The long-wavelength structure of the Earth is similar across recent RT tomographic models, even though there is some disagreement on the amplitudes of even the most prominent structures (Romanowicz 2003). Global body wave tomography, based on RT, has revealed a variety of subducting slabs, some remain stagnant around the 660 km discontinuity, whereas others penetrate into the lower mantle (e.g. Grand *et al.* 1997; Albaréde & Van der Hilst 1999; Fukao *et al.* 2001), which is inconsistent with the hypothesis of a two-layered convection in the Earth. The detection of hypothesized thin thermal plumes (Morgan 1971) in the mantle has remained elusive in these RT-based tomographic images (Romanowicz 2003), but would be of considerable value in understanding the Earth's mantle dynamics.

To be valid, ray theory requires that wavelengths are short and Fresnel zones narrow. Short-period (∼1 s) *P*-wave traveltimes have so far been extensively used for global RT tomography of the Earth, but provide poor sampling of the upper mantle. Surface wave data, that are generally analysed at very long periods (∼40–300 s), may be combined with *S*-wave traveltimes measured at long periods (∼10–51 s). This provides a way to image the entire mantle, because surface waves are more sensitive to the upper mantle, and *S*-waves to the lower mantle. We have therefore chosen to focus on *S*-waves in this study. Long-period *S*-wave data have so far mostly been used in tomographic imaging of the very long-wavelength heterogeneity (>1000 km horizontally) in the Earth, for which a ray theoretical approach is acceptable. However, 1000 km represents a third of the Earth's mantle thickness, and RT breaks down when used for

*Now at: Laboratoire de Science de la Terre, Centre National de la Recherche Scientifique, Université Lyon 1 et Ecole Normale Supérieure de Lyon, France.

imaging smaller heterogeneities that are of considerable interest in geophysics such as parts of slabs sinking in the mantle, or hot rising plumes. These objects are likely to have dimensions that are rather limited in size (~200 km horizontally). They are very difficult to constrain with RT because wave scattering and wavefront-healing effects are ignored. The effect of wave diffraction phenomena is to make traveltime anomalies dependent on Earth structure in the entire 3-D region around the geometrical ray path, rather than only on the infinitesimally narrow ray path itself. Because this is not taken into account in RT, it seems progress toward obtaining higher-resolution images of small heterogeneities in the mantle requires a movement away from RT.

In an effort to improve upon the infinite-frequency approximations of RT, that are only applicable to the time of the wave *onset*, finite-frequency (FF) approaches have recently emerged in seismic tomography (e.g. Dahlen *et al.* 2000; Tromp *et al.* 2005). For instance, the FF theory developed by Dahlen *et al.* (2000) takes the effects of wave diffraction into account (single scattering only), which makes the imaging of smaller objects or anomalies possible. The RT ray paths are replaced by volumetric sensitivity (Fréchet) kernels. Delay-times (or time residuals) observed in different frequency bands contain information on the size of the heterogeneity. For instance, the healing of a wavefront depends on the ratio between wavelength and size of heterogeneity. In FF tomography, traveltime (and amplitude) anomalies are therefore frequency-dependent. In principle one can exploit this dependence, by performing inversions with data from different frequency bands simultaneously. This may lead to an increase in resolution of the tomographic imaging. In FF tomography, the general form of the linear inverse problem is

$$dt_i(T) = \int_{V_i(T)} K_i(\mathbf{r}; T) \times m(\mathbf{r}) \mathrm{d}^3\mathbf{r}, \tag{1}$$

where $dt_i(T)$ is a *frequency-dependent* delay-time measured between the observed and synthetic waveforms of the target seismic phase i, both filtered around the period T. The volume integral $V_i(T)$ is theoretically over the entire Earth, but in practice limited to the region where the Fréchet kernel $K_i(\mathbf{r}\,; T)$ has a significant amplitude. The model parameter $m(\mathbf{r})$ represents a velocity perturbation ($\delta c/c$). By measuring the traveltime of a seismic phase at several periods, there is a potential for increasing the amount of independent informations in the inverse problem, as at each period the waveform is influenced by a different weighted average of the structure, through the corresponding 3-D sensitivity kernel. Recently, Montelli *et al.* (2004a,b, 2006b) published *P*- and *S*-wave FF global tomography models claiming to confirm the existence of deep mantle plumes below a large number of postulated hotspots. They mainly attributed this to an improvement in the resolving power of their FF approach (Dahlen *et al.* 2000).

The potential benefit of using an FF theory for tomography, as well as its significance in mapping mantle plumes, is somewhat controversial. A number of recent studies (e.g. Sieminski *et al.* 2004; de Hoop & Van der Hilst 2005a; Dahlen & Nolet 2005; de Hoop & Van der Hilst 2005b; Julian 2005; Trampert & Spetzler 2006; Montelli *et al.* 2006a; Van der Hilst & de Hoop 2006; Boschi *et al.* 2006) suggest that the effect of such an FF theory could be smaller than that of practical considerations such as the level of damping, the weighting of different data sets and the choice of data fit. Another important factor, which could limit the benefits of the new theory, is the number of FF data used, which has until now remained relatively small compared to the large number of traveltimes analysed with RT. For instance, FF tomographic models of Montelli *et al.* (2004a,b) are constrained by about 90 000 long-period (~20 s) *P, P–PP* and *pP–P* traveltimes, compared to about 1 500 000 short-period (~1s) *P* and *pP* traveltimes, commonly analysed using RT (Van der Hilst & de Hoop 2005). Moreover, Montelli *et al.* (2004a,b, 2006b) use traveltimes measured by matching the 'first swing' of a long period (~20 s) observed waveform with a synthetic (Bolton & Masters 2001). As noticed by Montelli *et al.* (2004b), such a measurement scheme presents a possible bias in dominant frequency caused by the correlation operator emphasizing the early part of the waveform rather than the full period. Only analysing the early part of the waveform, which is closely related to the wave's onset, may prevent tomographers from taking full advantage of such FF approach. For instance, Dahlen *et al.* (2000) show that if scatterers located off the ray path do not affect the onset of the wave, they can still advance or delay the full waveform. The global *S*-wave FF tomographic model obtained by Montelli *et al.* (2006b) is also based on (*S, ScS–S* and *SS–S*) traveltimes measured in a single-frequency band (~20 s), and hence does not benefit from the increased spatial resolution afforded by sensitivity kernels for a range of frequencies. Recently, Sigloch & Nolet (2006) presented an approach for measuring FF body wave amplitudes and traveltimes of teleseismic *P*-waves, between periods of 2 and 24 s. They model the first 25 s of a seismogram after the direct *P* wave arrival, including the depth phases *pP* and *sP*. The best source parameters (source time function, moment tensor, depth) are determined for each earthquake, with a cluster analysis that needs many stations having recorded the same event. This approach is, however, better suited for local or regional tomographic studies, rather than for global ones.

To take full advantage of using an FF approach, it is necessary to use traveltimes measured in a way which is fully consistent with the kernels. It is our view that for significant progress to be made, a new global data set of multiple-frequency body wave traveltimes is needed. One measured by cross-correlation over a broad frequency range. In this study, we focus on *S*-wave traveltimes, because they may be readily combined with surface wave data, to obtain a high-resolution tomographic image of the entire mantle. To our knowledge, there is no global database of *S*-wave traveltimes measured at different frequencies. We aim to measure traveltimes of single (*S, ScS, SS*) or groups of (*S+sS, ScS+sScS, SS+sSS*) phases, within the 10–51 s period range. We use 30 years of broadband seismograms recorded at the Global Seismological Networks (GSN) and distributed by the IRIS and GEOSCOPE data centres.

In Section 2, we present how we obtained a global data set of ~400 000 frequency-dependent *S*-wave traveltimes. An automated scheme for measuring long period *S*-wave traveltimes in different frequency bands has been developed. Automation was necessary to process the type of massive data set needed for global seismic tomography applications. Traveltime measurements have been corrected for elliptical, topographic, crustal and attenuation effects (Tian *et al.* 2007a). Frequency-dependent effects due to crustal reverberations beneath each receiver have been handled by incorporating crustal phases into WKBJ (Chapman 1978) synthetic waveforms. A good control of the frequency content of the waveforms, associated with a given traveltime, enables us to associate each measurement with a kernel carrying the same frequency information. The resulting multiple-frequency traveltimes are fully compatible to be inverted with volumetric sensitivity (Fréchet) kernels, irrespective of whether these kernels are computed with adjoint, mode-coupling or paraxial methods (e.g. Tromp *et al.* 2005); traveltimes measured at a single period can also be inverted using ray theory.

In Section 3, we focus on frequency-dependent effects occurring on global *S*-wave traveltimes in the mantle. If a residual structural traveltime dispersion is indeed observable, we would have a new constraint on the nature of seismic heterogeneity and attenuation in the Earth's interior. We are then especially interested in pointing out in our global data set frequency-dependent effects associated to wavefront-healing, scattering and attenuation.

2 A GLOBAL DATA SET OF FREQUENCY-DEPENDENT *S*-WAVE TRAVELTIMES

In this section, we describe our method for building a global data set of frequency-dependent body-wave traveltimes. Our automated scheme consists of two main stages. The first involves an automated selection of time windows around a set of target phases, which are present on both the observed and synthetic seismograms. The second stage involves measurements of multiple-frequency traveltimes by cross-correlating the observed and synthetic waveforms, filtered within the 10–51 s period range, which are contained in the previously selected time windows.

Readers mainly interested in our observations of frequency-dependent effects in our global data set, such as wavefront-healing, scattering and attenuation, may skip to Section 3.

2.1 Time windows selection and seismic phases isolation

Our time windows selection algorithm follows several of the ideas developed by Maggi *et al.* (2009). These authors present an approach for automated window selection designed for adjoint tomography studies. This class of studies involves 3-D numerical simulations of the seismic wavefield and 3-D sensitivity (adjoint) kernels (e.g. Komatitsch & Tromp 1995; Komatitsch *et al.* 2002; Tromp *et al.* 2005). The advantage of this approach is that the adjoint kernel is obtained from a numerical calculation, with no need to identify specific seismic phases. That is, the kernel can be computed for any part of the seismogram and takes care of the relevant sensitivities. For this reason, Maggi *et al.* (2009) define time windows covering as much as possible of a given seismogram, while avoiding portions of the waveforms that are dominated by noise. They select time windows on the synthetic waveform only, without identifying specific seismic phases. Each time window on the synthetic seismogram is then associated with the same time window on the observed seismogram, assuming that they contain the same patterns of interference between seismic phases. This assumption will be valid for accurate synthetic seismograms, calculated by 3-D propagation through a good 3-D Earth model. However, it may not be fulfilled with more approximate spherical-Earth synthetics, computed in a 1-D Earth model like IASP91 (Kennett & Engdahl 1991), as used in this study. Because strong 3-D heterogeneities, present at the top and bottom of the mantle, can produce large delay-times between observed and 1-D synthetic waveforms, we chose to focus our time windows selection on both the observed *and* 1-D synthetic waveforms, to isolate well-identified seismic phases. In Appendix A, we describe our time windows selection and seismic phase isolation methodology, which largely makes use of the ideas of Maggi *et al.* (2009), tuned for our particular application (*cf.* Table A1).

Finally, before entering into the measurement process (Section 2.2), the selected observed and synthetic waveforms are tapered and extrapolated outside their isolation time windows with an amplitude set to zero. This is possible because body waves are finite duration pulses. A similar approach was followed by Pollitz (2007), who also used cross-correlation measurements based on narrow-window tapers.

2.2 Frequency-dependent traveltime measurements

Our time windows selection scheme has allowed us to isolate, in an automated way, a pair of observed and synthetic waveforms, associated with each target seismic phase. We now aim to measure multiple-frequency traveltimes by cross-correlating the observed and synthetic waveforms, filtered at different periods, which are contained in the previously selected time windows. We chose to build a global data set of multiple-frequency traveltimes within the 10–51 s period range. It is our experience that *S* waves are generally prominent compared to seismic noise in this period range. However, this is not always the case for the entire 10–51 s period range. In the following section, we determine the appropriate frequency range of analysis for each target phase to be measured.

2.2.1 Frequency range of analysis

Seismic body waves, associated with long paths through the Earth, have their high-frequency content more severely attenuated than those associated with shorter paths. For instance, *S* and *ScS* phases are generally associated with shorter ray paths and higher frequency content than *SS*. The long period nature of *SS* is also related to its longer journey in the shallow mantle, compared to *S* and *ScS*, which is strongly attenuating for high-frequencies. Moreover, seismic noise has a peak in amplitude at short periods ($\sim$6 s), mainly caused by the oceanic swell, that may significantly pollute the high-frequency content of broadband seismograms recorded at oceanic stations.

In our time windows selection and seismic phases isolation scheme (Section 2.1 and Appendix A), broadband seismograms recorded at the GSN are first bandpass filtered between 7 and 85 s with a non-causal Butterworth filter, whose short- and long-period corners are denoted by T_1 (7 s) and T_2 (85 s), respectively. In this study, we cover the period range between 7 and 85 s with five overlapping Gaussian filters, whose centre periods T are 10, 15, 22.5, 34 and 51 s (*cf.* Table 1). Our aim is to determine the largest frequency range, associated to each targeted waveform, for which traveltimes can be measured.

We first determine the minimum short period corner of the Butterworth filter, denoted by T'_1, for which the target phase can be isolated on the observed seismogram with the approach described in Section 2.1 and Appendix A. T'_1 is chosen among three trial values: 7, 11 and 16 s. In the following, we will only consider those of the five Gaussian filters whose centre periods T are greater than T'_1.

A second selection is then performed by computing, for each selected Gaussian filter, the signal-to-noise ratio $SNR(T)$ between the absolute amplitude maxima of the isolated target observed waveform and of the seismic noise. Seismic noise is evaluated from a 100 s time window (Fig. A1a), taken on the observed seismogram, before the first arrival time among the *S, ScS* and *SS* phases. Among the selected Gaussian filters, we only keep those for which a ratio $SNRT(T)$ greater than 3 is found. In the following Section 2.2.2, we describe how we measure time residuals at several periods. For each couple of observed and synthetic waveforms, measurements are made at the periods for which Gaussian filters have been selected.

2.2.2 Measuring time residuals

We now aim to measure multiple-frequency time residuals for each optimal pair of observed and synthetic waveforms, within the frequency range of analysis corresponding to each target phase

Table 1. Passband Gaussian filters for which frequency-dependent traveltimes were measured.

Band index	1	2	3	4	5
Centre period (s) T	10.0	15.0	22.5	34.0	51.0
Centre frequency (mHz) ν	100.0	66.7	44.5	29.4	19.6
1σ–corner frequency (mHz) $\nu \pm \sigma$	81.0–119.0	54.0–79.4	36.0–53.0	23.8–35.0	15.9–23.3
2σ–corner frequency (mHz) $\nu \pm 2\sigma$	62.0–138.0	41.4–92.0	27.6–61.4	18.2–40.6	12.2–27.0

Where ν is the central frequency, $T = 1/\nu$ is the central period and σ the standard deviation. These filters are Gaussian in the frequency domain.

(*cf.* Sections 2.1 and 2.2.1). These multiple-frequency time residuals are designed to be fully compatible with a FF approach for tomography, such as the one developed by Dahlen *et al.* (2000). That is, the FF time residual τ_m associated with the period T, is defined as the time maximizing the cross-correlation function, $\gamma_{d,s}(\tau)$, between the observed, $d(t)$, and synthetic, $s(t)$, waveforms, both filtered around the period T. The cross-correlation function is defined as

$$\gamma_{d,s}(\tau) = \int_{-\infty}^{+\infty} d(t) \times s(t-\tau)\mathrm{d}t. \tag{2}$$

Picking the maximum of the cross-correlation is usually accurate. However, in some cases the highest absolute value of the cross-correlation corresponds to a cycle-skip and leads to large error. By eye, such cycle-skips can often be recognized because the rest of the signal, away from the maximum, has a mismatch. We have noticed that cycle-skipping problems are more likely to occur when the two filtered signals, $d(t)$ and $s(t)$, are strongly dominated by one particular period.

To minimize biases due to cycle-skipping problems, our multiple-frequency time residuals are measured by determining the maximum position, by curve-fitting, of the new function: $F_3(\tau)$. This function $F_3(\tau)$ is designed to provide an accurate estimation of τ_m, while minimizing cycle-skipping. We show, in Appendix B, that the function $F_3(\tau)$ and the cross-correlation function $\gamma_{d,s}(\tau)$ are maximized for the same time residual τ_m. Therefore, Fréchet kernels designed to be used with time residuals measured by maximizing $\gamma_{d,s}(\tau)$ (e.g. Dahlen *et al.* 2000) can also be used with residuals obtained by maximizing $F_3(\tau)$. In the following, we define the function $F_3(\tau)$ and show how it assists in minimizing cycle-skips.

Following Ritsema & van Heijst (2002), we define two quantities, function of the time shift τ, that describe the similarity between the filtered observed, $d(t)$, and time-shifted synthetic, $s(t-\tau)$, waveforms. These two quantities are noted $F_1(\tau)$ and $F_2(\tau)$. They are defined such as

$$F_i(\tau) = \begin{cases} F_i(\tau) & \text{if } 0 \le F_i(\tau) \le 1, \quad i = 1, 2 \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

The first quantity $F_1(\tau)$ is the least-squares misfit between $d(t)$ and $s(t-\tau)$

$$F_1(\tau) = 1 - \frac{\int_{-\infty}^{+\infty}[d(t) - s(t-\tau)]^2 dt}{\int_{-\infty}^{+\infty} d^2(t) dt}. \tag{4}$$

Note that $F_1(\tau)$ is close to 1 when τ minimizes the misfit between $d(t)$ and $s(t-\tau)$. The second quantity $F_2(\tau)$ measures the wave shape similarity between $d(t)$ and $s(t-\tau)$

$$F_2(\tau) = \frac{\min[A_1(\tau), A_2(\tau)]}{\max[A_1(\tau), A_2(\tau)]}, \tag{5}$$

where $A_1(\tau)$ and $A_2(\tau)$ minimize the functions f and g, respectively. The functions f and g are defined as

$$\begin{cases} f(x) = \int_{-\infty}^{+\infty}[(d(t) - x \times s(t-\tau)]^2 \mathrm{d}t \\ g(x) = \int_{-\infty}^{+\infty}[x^{-1} \times d(t) - s(t-\tau)]^2 \mathrm{d}t. \end{cases} \tag{6}$$

Which leads to

$$\begin{cases} A_1(\tau) = \gamma_{d,s}(\tau)/\gamma_{s,s}(0), \\ A_2(\tau) = \gamma_{d,d}(0)/\gamma_{d,s}(\tau), \end{cases} \tag{7}$$

where $\gamma_{d,d}(0)$ and $\gamma_{s,s}(0)$ are the autocorrelation values, at zero lag-time, of the observed and synthetic waveforms, respectively. Note that $F_2(\tau)$ is close to 1 when τ maximizes the wave shape similarity between $d(t)$ and $s(t-\tau)$. Finally, the function $F_3(\tau)$ is defined as

$$F_3(\tau) = \frac{F_1(\tau) + F_2(\tau)}{2}. \tag{8}$$

The function $F_3(\tau)$ includes a more sophisticated information on the misfit, through $F_1(\tau)$, and on the wave shape similarity, through $F_2(\tau)$, than the cross-correlation function, $\gamma_{d,s}(\tau)$. When cycle-skips occur, it is statistically easier to find, in an automated way, the appropriate residual time by using the function $F_3(\tau)$, because its global maximum is enhanced and its secondary maxima (corresponding to cycle-skips) are minimized, compared to the ones of the cross-correlation function. An example of comparison between the two functions $F_3(\tau)$ and $\gamma_{d,s}(\tau)$ is shown in Fig. B1. By experimentation, we only retained the time residuals corresponding to a function $F_3(\tau)$ with a unique maximum greater than 80 per cent, and with no secondary maximum greater than 70 per cent. The use of the function $F_3(\tau)$, rather than a simple cross-correlation, has proved to be very useful for the automation of our measurement process (*cf.* Figs 1 and B1). The function $F_3(\tau)$ mimics very well the seismologist's, often visual, decision in choosing the appropriate time residual.

2.2.3 Measurement errors

Errors on our time residual estimates can result from waveform distortion, owing to the effects of both the noise and the approximations made in the synthetics computation. We aim here to approximate the standard deviation (σ) related to each measured time residual (τ_m). Following Chevrot (2002), we first compute the correlation coefficient, $\gamma_{d,s}(\tau_m)$, between the observed and time-shifted synthetic waveforms. This coefficient is then compared with the autocorrelation function, $\gamma_{s,s}(\tau)$, of the synthetic waveform. Finally, we approximate the error (σ) by the time lag at which the correlation coefficient, $\gamma_{d,s}(\tau_m)$, is observed in the autocorrelation function, $\gamma_{s,s}(\tau)$. That is

$$\sigma = \left\{\tau \mid \gamma_{d,s}(\tau_m) = \gamma_{s,s}(\tau)\right\}. \tag{9}$$

Hence, observed waveforms exhibiting a strong correlation with the synthetic waveform will be attributed low errors. On the other

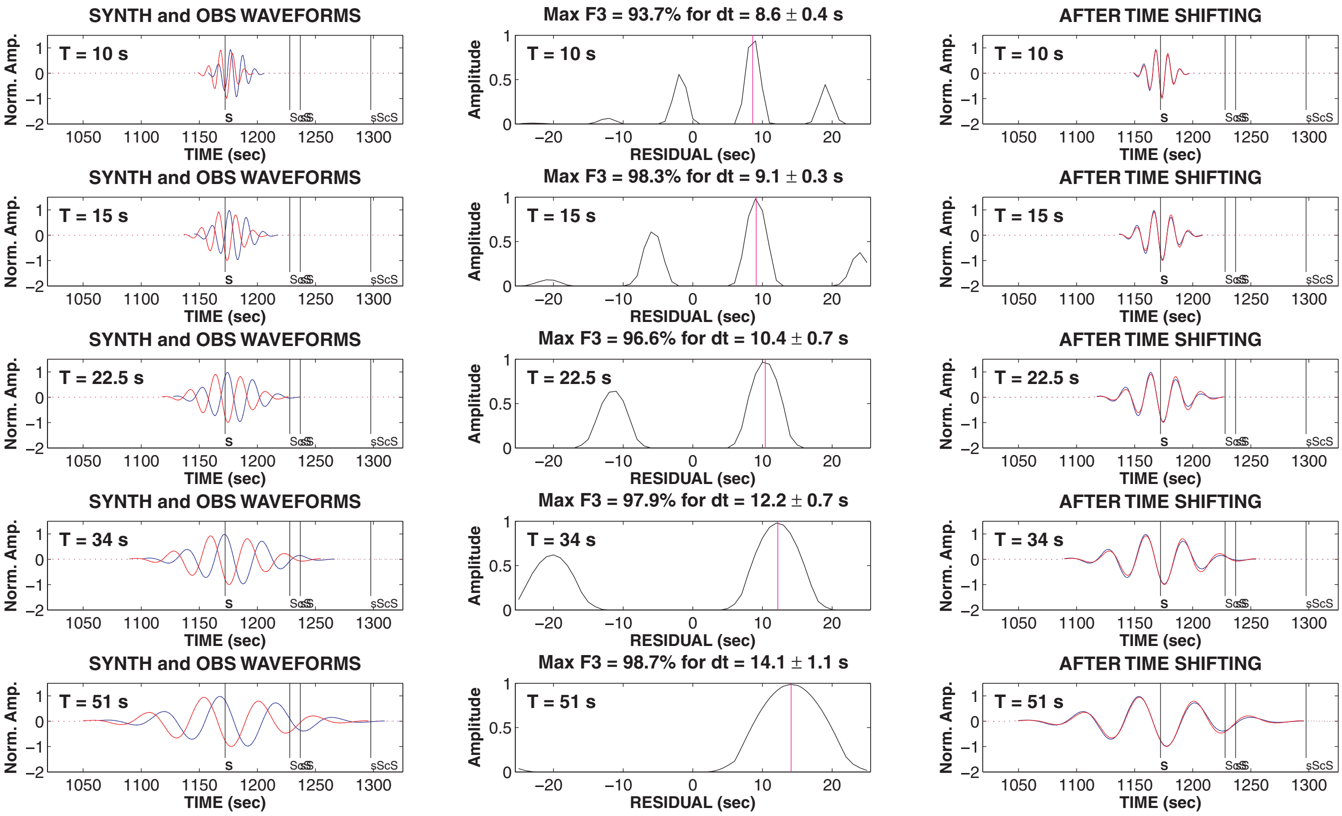

Figure 1. Measuring multiple-frequency time residuals. Left column: observed (blue) and synthetic (red) isolated and tapered *S*-phase waveforms (corresponding to Fig. A1a). Middle column: for each filtering period (T), $F_3(\tau)$ is maximized for a delay-time $\tau_m(T)$. Right column: observed and synthetic waveforms after time shifting of $\tau_m(T)$. Note that each raw corresponds to a different filtering period T.

hand, signals strongly contaminated by noise will produce larger traveltime residual errors.

2.2.4 Traveltime corrections for global seismic tomography

In this study, we aim to build a global data set of multiple-frequency time residuals, suitable for imaging the Earth's mantle structure using inversion schemes based on eq. (1). Time residuals are determined by cross-correlating observed and synthetic waveforms (*cf.* Section 2.2.2). The synthetic waveforms used in this study are computed in a spherical Earth with the WKBJ method for the IASP91 velocity model extended with the Q model from PREM (Dziewonski & Anderson 1981).

We need to apply corrections to the predicted traveltimes, computed in the radial IASP91 reference velocity model, to account for known deviations from spherical symmetry in the Earth. We use the software by Tian *et al.* (2007a), to compute the ellipticity, dt_{ell}, crustal, dt_{cru}, and topographic, dt_{top}, traveltime corrections, for each seismic phase (*S, sS, ScS, sScS, SS, sSS*) present in the WKBJ synthetic. The traveltime after correction, for each seismic phase, is

$$T_{\mathrm{corr}} = T_{\mathrm{BG}} + dt_{\mathrm{ell}} + dt_{\mathrm{cru}} + dt_{\mathrm{top}}, \tag{10}$$

where T_{BG} is the predicted traveltime for the spherically reference (background) model (IASP91 in this study). We use the 3-D global crustal model CRUST2.0 ($2^\circ \times 2^\circ$), by Gabi Laske, which is an updated version of an earlier model CRUSTS5.1 ($5^\circ \times 5^\circ$), by Mooney *et al.* (1998).

As they propagate through the Earth, seismic waves experience attenuation and dispersion resulting from microscopic dissipative processes, operating at a variety of relaxation times. Intrinsic attenuation causes dispersion of seismic velocities, decreasing the velocities of longer period waves, compared to shorter period ones. Properly correcting for the dispersion effect is crucial as we aim to use our multiple-frequency delay-times, determined in different frequency bands, to constrain velocities in the Earth. Frequency dependence of attenuation q can be represented by a power law

$$q \propto q_0 \times \omega^{-\alpha}. \tag{11}$$

Seismic studies routinely assume that, within the seismic band, α cannot be resolved and thus implicitly rely on the frequency-independent attenuation model, that is $\alpha = 0$, of Kanamori & Anderson (1977). Usually, the difference in wave-speeds due to an attenuation value q at two frequencies $\omega_{1,2}$ is calculated using the expression

$$\frac{V(\omega_2)}{V(\omega_1)} = 1 + \frac{q}{\pi} \times \ln(\omega_2/\omega_1), \tag{12}$$

which is only valid when $\alpha = 0$ (Kanamori & Anderson 1977). However, non-zero values of α (Section 3.3) require the use of a different expression (Anderson & Minster 1979):

$$\frac{V(\omega_2)}{V(\omega_1)} = 1 + \frac{q(\omega_1)}{2} \times \cot(\alpha\pi/2) \times [1 - (\omega_2/\omega_1)^{\alpha}]. \tag{13}$$

The values of α and $q(\omega_1)$ may significantly affect the magnitude of the dispersion correction. If one relies on a frequency-independent attenuation model, that is $\alpha = 0$, one should correct the multiple-frequency time residuals, measured by cross-correlation, by adding the physical dispersion correction, $dt_{\rm disp}^{\alpha=0}$, to eq. (10), with

$$dt_{\rm disp}^{\alpha=0}(T) = -\frac{t^*}{\pi} \times \ln\left(\frac{T_0}{T}\right). \tag{14}$$

T_0 is the reference period of the velocity model ($T_0 = 1$ s for IASP91), and T is the centre period of the Gaussian filter used to analyse the target phase. The parameter t^* is determined by kinematic ray tracing (Tian *et al.* 2007a)

$$t^* = \int_0^L \frac{dl}{c \times Q}. \tag{15}$$

The integration is along the ray path and Q is the quality factor ($Q = 1/q$) from the PREM model. On the other hand, if one relies on a frequency-dependent attenuation model, that is $\alpha \neq 0$, one should correct the multiple-frequency time-residuals by adding a different physical dispersion correction, $dt_{\rm disp}^{\alpha\neq0}$, to eq. (10), with

$$dt_{\rm disp}^{\alpha\neq0}(T) = -\frac{t^*}{2} \times \cot(\alpha\pi/2) \times [1 - (T/T_0)^\alpha]. \tag{16}$$

2.2.5 *Frequency-dependent crustal effects*

Removing the crustal signature from teleseismic traveltimes is very important to reduce the trade-off between crustal and mantle velocity heterogeneities in seismic tomography.

Yang & Shen (2006) discussed frequency-dependent effects due to continental crustal reverberations on teleseismic *P*-wave traveltimes, if strong reverberations arrive early enough to influence the cross-correlation. They observed a difference of traveltime up to 0.6 s between *P* waves filtered at short (i.e. 0.5–2 s) and long (i.e. 10–33 s) periods. Ritsema *et al.* (2009) also show that at relatively low frequencies, when the wavelengths of *P*- and *SH*-waves are similar to the thicknesses of the crust, crustal traveltimes are frequency-dependent, either for continental or ocean crustal structures. They use a global crustal model to calculate maps of frequency-dependent crustal effects for distinct *SH*-waves (e.g. *S, SS*). However, such crustal corrections should only be applied to 'single' seismic phases, free of interference with other seismic phases. When two phases interfere, such as *S* and *sS* in the case of a shallow earthquake, we cannot add linearly the frequency-dependent crustal traveltime corrections of the two seismic phases. Indeed, the interference pattern between two phases has a very complex and non-linear frequency dependence.

In modern global tomographic studies aimed at improving image resolution, we wish to include the large amount of seismograms corresponding to shallow earthquakes. As a consequence, if one aims to correct for frequency-dependent crustal effects, we cannot ignore the non-linear problem of interference pattern between direct and depth phases. One way to account for this non-linearity is to incorporate crustal reverberations into the synthetic waveform used to cross-correlate, but this requires prior knowledge of crustal structure, not just its thickness (e.g. Nolet 2008).

In this study, we directly incorporate into the WKBJ synthetics crustal phases reverberated on the receiver side, using the global crustal model CRUST2.0. We incorporate *S, sS, ScS, sScS, SS* and *sSS* crustal reverberations, when they are present in our time windows with a significant amplitude. The WKBJ approach does not allow us to simultaneously model crustal phases reverberated in different crustal models. We therefore ignore crustal phases associated to surface reflection points at the source side (e.g. *sS*), or at the bounce point in the case of *SS* waves.

Devilee *et al.* (2003) observed that interference between direct (e.g. *S*) and depth (e.g. *sS*) phases can also give apparent traveltime dispersion. Our modelling of the dispersion associated to crustal effects, and to direct and depth phases interference, may not be perfect, but we believe that our measured multiple-frequency traveltimes are accurate enough to assess frequency-dependent effects occurring on global *S*-wave traveltimes in the Earth's mantle. The validity of this assumption can be checked by using deep earthquakes, for which crustal phases reverberated at the source side arrive too late to influence the cross-correlation, so that there is no interference between direct and depth phases. Therefore, *S* and *ScS* waves are only affected by crustal phases reverberated at the receiver side, which are incorporated in our WKBJ synthetics. We checked that all the results of this study about wavefront healing (Section 3.1), scattering (Section 3.2) and attenuation (Section 3.3), remain the same if only deep earthquakes are used. Even if *SS* waves are affected by a surface reflection at the bounce point, whose associated crustal phases are not modelled here, this leads to the same result.

2.3 Data selection

A total of 28 810 earthquakes, with a body wave magnitude $m_b \geq 5.5$, were pre-selected from the Harvard centroid moment tensor (CMT) catalog, between 1976 January 1 and 2008 March 31. We obtained broadband seismograms (LH channel), associated with the selected events, from the IRIS and GEOSCOPE data centres, at almost 270 stations of the International Federation of Digital Seismograph Networks (FDSN). Only earthquakes with a magnitude such as $5.5 \leq m_b \leq 6.5$, and with a source half duration time $h_{\rm dur} < 6$ s, were used in this study. These criteria reject waveforms strongly complicated by the earthquake rupture process (Ritsema & van Heijst 2002). Moreover, Devilee *et al.* (2003) show that an asymmetric source time function may cause significant dispersion at periods shorter than the source duration time, but that this dispersion is small at greater periods. As we aim to measure multiple-frequency delay-times within the long period range 10–51 s, our measurements are not expected to be biased by this kind of dispersion. Therefore, we assume the source time function to be Gaussian, and use the expression given by Komatitsch *et al.* (2002). This assumption is appropriate for global teleseismic seismograms, but for local or regional studies, one should instead try to determine the exact source time function (Sigloch & Nolet 2006).

2.4 Data set robustness

Our multiple-frequency data set includes single-phase traveltimes (*S, ScS* and *SS*), completed with traveltime measurements for which the target phase interferes with its depth phase (*S*+*sS, ScS*+*sScS* and *SS*+*sSS*). This kind of interference is often associated with shallow earthquakes, whereas single-phase traveltimes generally correspond to deep events. We have specifically rejected measurements associated with waveforms that could be contaminated by other kind of interference. This is important for the tomographic inversion, as we aim to associate our multiple-frequency traveltimes with the appropriate sensitivity kernels.

Fig. 2 summarizes our traveltime observations for each target phase, superimposed on the theoretical traveltime curves, as a function of epicentral distance. Although we measure traveltimes of *SS* phases up to distances reaching $\sim 170°$, those *SS* measurements near antipodal epicentral distances ($\geq 140°$) should not be used with a kernel based upon the paraxial approximation (Tian *et al.* 2007b).

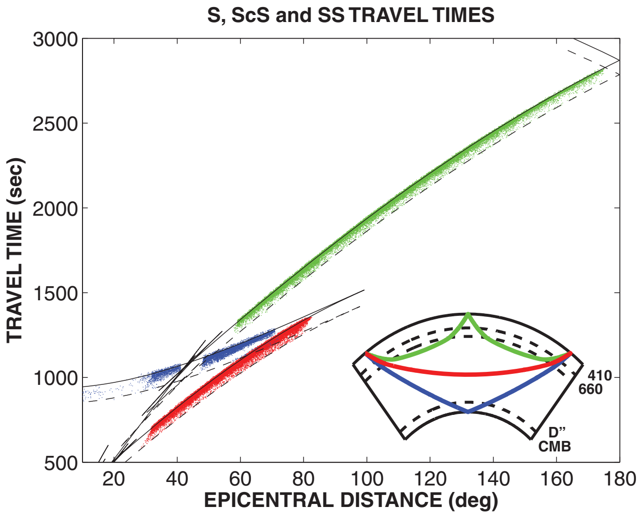

Figure 2. Traveltime versus epicentral distance plot showing our distribution of traveltimes for *S* (red), *ScS* (blue) and *SS* (green) seismic phases, superimposed to their corresponding theoretical traveltime curves. Traveltimes are given for the lowest filtering period *T* for which each target phase has been measured (*cf.* Section 2.2). Theoretical traveltime curves are shown with solid lines for 0 km source depth and dashed lines for 410 km source depth. We also show an example of *S* (red), *ScS* (blue) and *SS* (green) ray paths into the Earth.

Nevertheless, these near antipodal *SS* measurements could be used with more sophisticated kernels (Calvet & Chevrot 2005).

As in previous studies (e.g. Bolton & Masters 2001), the most subtle problem that we face is the accidental measurement of a depth phase (e.g. *sS*) when the direct phase (e.g. *S*) is poorly excited. Engdahl *et al.* (1998) show how the problem can be reduced using statistical methods. Bolton & Masters (2001) measure the arrival polarity to identify depth phase problems. Their measurements are based on the cross-correlation between the first swing of the observed and synthetic direct phases. Our analysis relies on the entire waveform(s) of the target phase(s). We impose (see Appendix A, Section A6) a high correlation coefficient ($CC_{\max} \geq 80$ per cent) between the observed and synthetic waveforms. Therefore, if a pattern of interference between two phases is present on the synthetic waveform, a similar pattern must also be met on the observed data. If this pattern is not found, the traveltimes are not measured. Uncertainties on Harvard centroid moment tensor (CMT) solutions are likely to affect the relative amplitudes of the direct and depth phases, especially when one of the take-off azimuth is near a nodal plane. In this case, the observed and synthetic two-phase waveforms are expected to differ, that is $CC_{\max}$ is low, and the data are rejected. The cross-correlation criteria, that is $CC_{\max} \geq 80$ per cent, associated with a waveform search in a specific time window (Appendix A and Section A5), enable us to reject a large number of data for which the CMT source mechanism is not reliable. This is especially true for near nodal measurements for which inaccuracies in the CMT solution often imply significant differences between the observed and synthetic waveforms. This allows us to discard most accidental measurements of a depth phase when the direct phase is poorly excited.

Although our database has been built for the transverse component (*SH*), our approach can also be easily extended to *P* or *SV* components, provided that (1) new crustal phases, as for instance from *P* to *S* conversions in the crust, are modelled and added in the synthetics; (2) new depth phases, such as *pS*, are added in the synthetics and (3) new seismic phases interference patterns, such as *S* with *SKS*, are taken into account.

2.5 Global patterns in the data

Fig. 3 displays the ray coverage achieved with our database, for different depth ranges covering the entire lower mantle. The current coverage of seismic stations allows us to achieve a good sampling of the Northern Hemisphere for all but the rays with the shallowest lower mantle turning depths. Coverage in the Southern Hemisphere

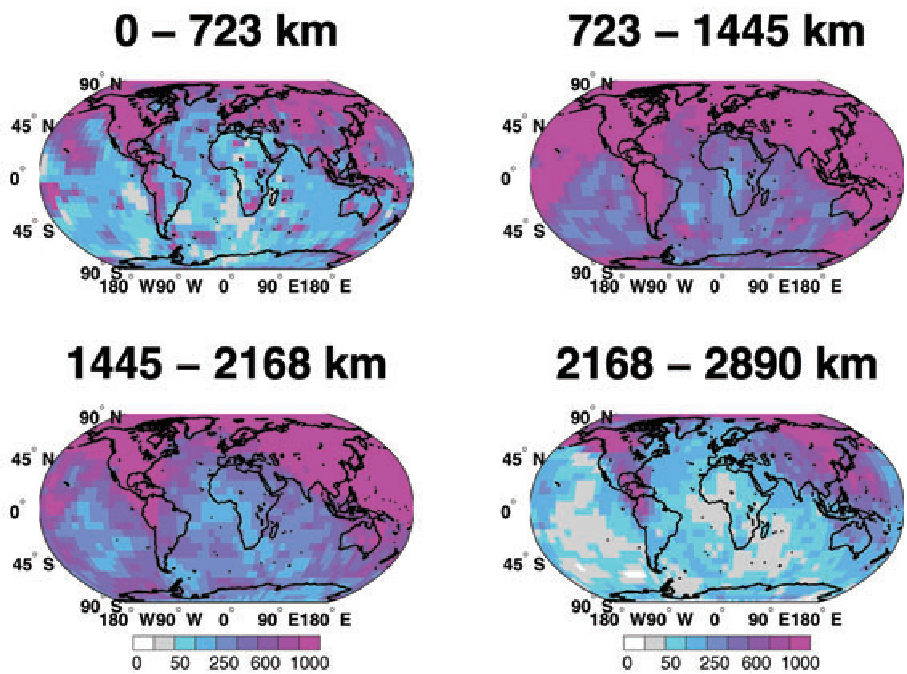

Figure 3. Ray density maps per 6° × 6° cells, shown at four different depth slices. The colourscale represents the number of rays in each cell, normalized by the cell size. Cells with less than 10 rays are shown in white.

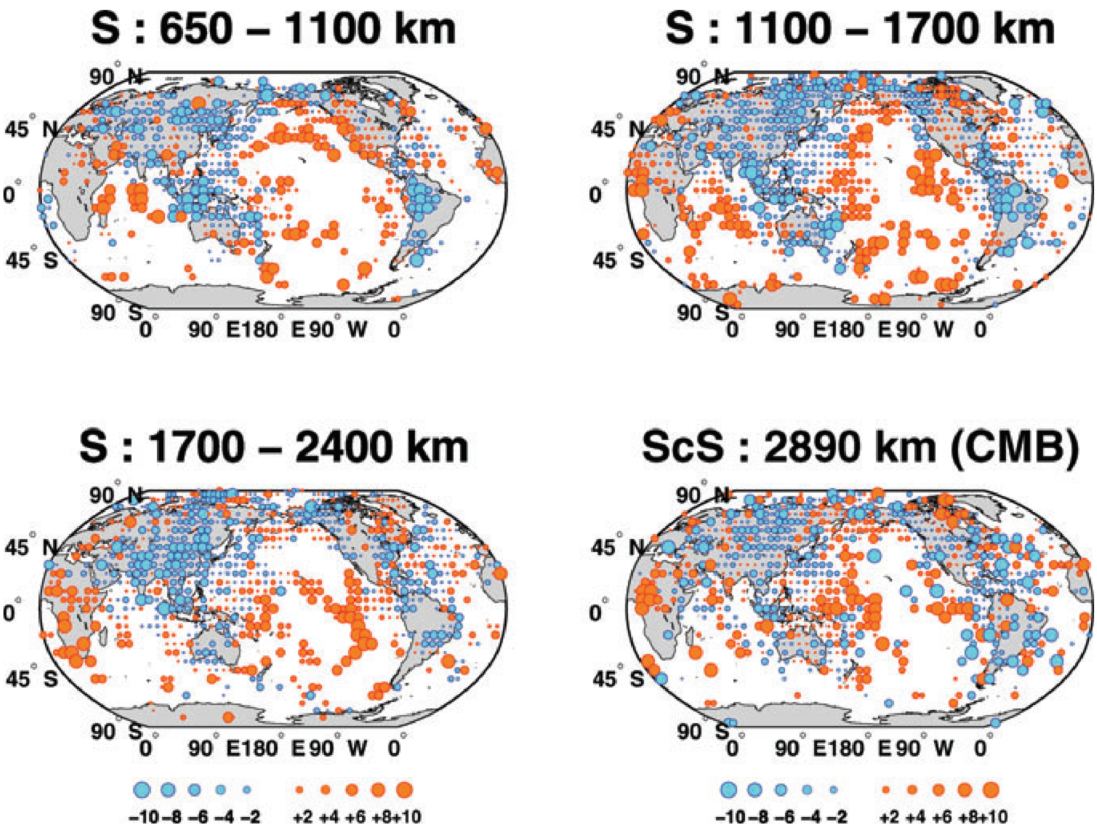

Figure 4. Data patterns: S and ScS (highest-frequency) time residuals (in s) are plotted at their turning points in $6^\circ \times 6^\circ$ cells (with three turning points at least in each cell), and shown at different depth slices. The mean (μ) from each depth slice has been removed: μ_S (650–1100 km) = 2 s, μ_S (1100–1700 km) = 1.7 s, μ_S(1700–2400 km) = 0.7 s and μ_{ScS}(2890 km) = −3.9 s. Because most cells contain residuals from many azimuths, we can infer that a majority of the signal shown here must be accumulated near the ray turning point.

still remains a problem, but many areas appear to be sampled well enough to reveal consistent patterns.

Fig. 4 shows the geographic distributions of the S and ScS residuals, plotted at the surface projection of the ray turning points. Residuals are averaged in $6^\circ \times 6^\circ$ cells and shown over four ray turning depth ranges. As in previous studies (Bolton & Masters 2001; Houser *et al.* 2008), we observe large-scale patterns in both sign and amplitude. These patterns are clearly associated with the long wavelength structure seen in global tomography. Our observations suggest fast regions beneath Asia, Arctic, North and South America in the depth range between 650 and 1700 km. For example, strong fast residuals observed at turning points between 650 and 1700 km, below the northern part of South America, correspond to the subduction of the Nazca Plate (Van der Hilst *et al.* 1997). These time residuals are consistent with the high-velocity ring around the Pacific seen in most S-wave tomographic models (e.g. Dziewonski 1984; Masters *et al.* 1996, 2000). The deep turning rays, deeper than 1700 km are delayed by the slow areas seen in global tomography (e.g. Ritsema *et al.* 1999) at the base of the mantle over much of the central Pacific Ocean and beneath South Africa. The agreement between our observations and global tomography suggests that mantle structure in the region of the ray turning point is responsible for most of the observed patterns.

3 FREQUENCY-DEPENDENT EFFECTS ON GLOBAL *S*-WAVE TRAVELTIMES

In this section, we focus on frequency-dependent effects occurring on global S-wave traveltimes in the mantle. If a residual traveltime

Table 2. Summarized global multiple-frequency time residuals.

Period (s)	10	15	22.5	34	51
S waves					
N	19008	36708	49089	46000	38238
μ (s)	0.6	0.9	1.4	1.7	2.6
σ (s)	±5.5	±5.6	±5.8	±5.8	±6.4
ScS waves					
N	4939	10094	13069	11480	8801
μ (s)	−2.8	−3.8	−4.7	−6.1	−8.3
σ (s)	±8.1	±8.4	±8.3	±8.1	±8.5
SS waves					
N	4189	21777	49963	50041	40882
μ (s)	0.2	0.7	1.2	1.2	1.5
σ (s)	±8.5	±8.1	±7.7	±7.5	±7.9

Where N is the number of measurements, μ is the mean, and σ is the standard deviation of the best fitting Gaussian function of each histogram of our S, ScS and SS data sets. Both single-phase (e.g. S) and two-phase (e.g. S+sS) time residuals are considered.

dispersion is indeed structural and observable, we would have a new constraint on the nature of seismic heterogeneity and attenuation in the Earth's interior.

Table 2 summarizes the mean and standard deviation of our S, ScS and SS multiple-frequency time residual measurements, in our period range of analysis. After correction for physical dispersion (*cf.* Section 3.3 and Fig. 9) due to intrinsic anelastic processes, under the hypothesis of a frequency-independent attenuation (i.e. $\alpha = 0$), we observe a clear frequency-dependency in our measurements. For instance, when the period increases, the mean delay-time decreases for *ScS* phases but increases for *S* and *SS* (*cf.* Table 2). At first glance, the frequency-dependency observed in our global measurements is not directly related to specific seismic heterogeneity or attenuation in the mantle.

So far, global tomographers have only relied on the inversion process (i.e. solving eq. 1) to unravel all the complex frequency-dependency information contained in their global multiple-frequency traveltime measurements (e.g. Montelli *et al.* 2004a). In the following, we aim to give evidence that a residual structural dispersion is contained in our data. We first point out, in Section 3.1, that wavefront-healing produced by very low velocity anomalies is clearly observed in our *S*-wave data set and may contribute to the observed *SS* dispersion. We also report on our observation that *ScS* waves of our global data set travel faster at low-frequency than at high-frequency. We suggest, in Section 3.2, that a preferred sampling of high-velocity scatterers located at the CMB, may explain the peculiar *ScS* dispersion pattern. Finally, we argue, in Section 3.3, that the globally averaged dispersion observed for *S* and *SS* traveltimes is compatible with a frequency-dependent attenuation model for the average mantle.

3.1 Evidence for wavefront-healing from local to global scale

An important effect caused by the wave's frequency being finite is wavefront-healing (Nolet & Dahlen 2000; Hung *et al.* 2001; Nolet *et al.* 2005). Wavefront-healing is a ubiquitous diffraction phenomenon, which depends upon the wave's frequency and the anomaly size. It occurs whenever the scale of any geometrical irregularities in a wavefront are comparable to the wavelength of the wave (Gudmundsson 1996), and affects cross-correlation traveltime measurements (Hung *et al.* 2001). That is, a low-velocity anomaly creates a delayed wavefront with an unperturbed zone that may be filled in (i.e. healed) by energy radiating from the sides, using Huygens' Principle (Nolet *et al.* 2005). Wavefronts of longer waves heal more quickly as a function of distance from the perturbation (e.g. Nolet 2008). Therefore, if a seismic wave passes through a low-velocity anomaly, the longer the wave period is, the more important the healing will be, and therefore the less the wave will be apparently delayed at the receiver. The corresponding time residuals, *dt*, measured by cross-correlation at different filtering periods, *T*, will then lead to a decreasing dispersion curve $dt(T)$.

3.1.1 Wavefront-healing at local scale

Here we focus on traveltime dispersion of *S* waves recorded at the LKWY broadband seismic station (Fig. 5), which belongs to the US network. This station has the particularity of being located above the Yellowstone hotspot, whose seismic signature is a very low velocity anomaly (Fig. 5c). When the wave's period increases, such as its wavelength grows to a length comparable to the dimension of the anomaly, wavefront-healing becomes significant even at short distance from the anomaly. A seismic wave travelling through the Yellowstone low-speed anomaly is then expected to be significantly affected by wavefront-healing when recorded at the LKWY receiver. The corresponding dispersion curve, $dt(T)$, is therefore expected to decrease. For comparison, we also analyse traveltime dispersion of *S* waves recorded at five other seismic stations located in the close vicinity of the LKWY station (Figs 5a and b). All these *S* waves are associated with earthquakes located in similar regions along the AA' profile (Fig. 5a), so that we can attribute the observed traveltime differences to the receiver side.

Fig. 6 shows the associated dispersion curves, measured within the 10–51 s period range. We plot $dt(T) - dt(T = 10\text{ s})$ so that increasing/decreasing dispersion curves are above/below zero of the *y*-axis. $dt(T = 10\,\text{s})$ provides an information on the average

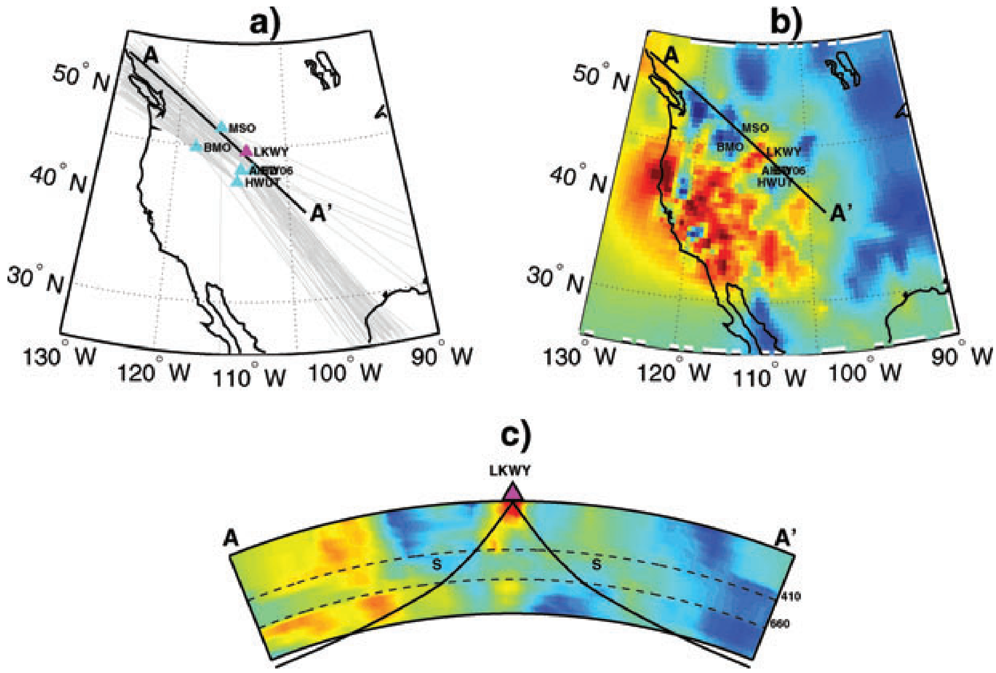

Figure 5. (a) Map showing all the ray paths used in Fig. 6 and six seismic stations located in the vicinity of the Yellowstone hostpot. (b) *P*-wave velocity anomaly model at 200 km depth around the Yellowstone hotspot. (c) Cross-section showing the Yellowstone hotspot (low velocity anomaly) beneath the seismic station LKWY. The tomographic model (MITP–USA–2007NOV) is from Burdick *et al.* (2008). The colourscale shows in red/blue the low/high velocity anomalies, respectively.

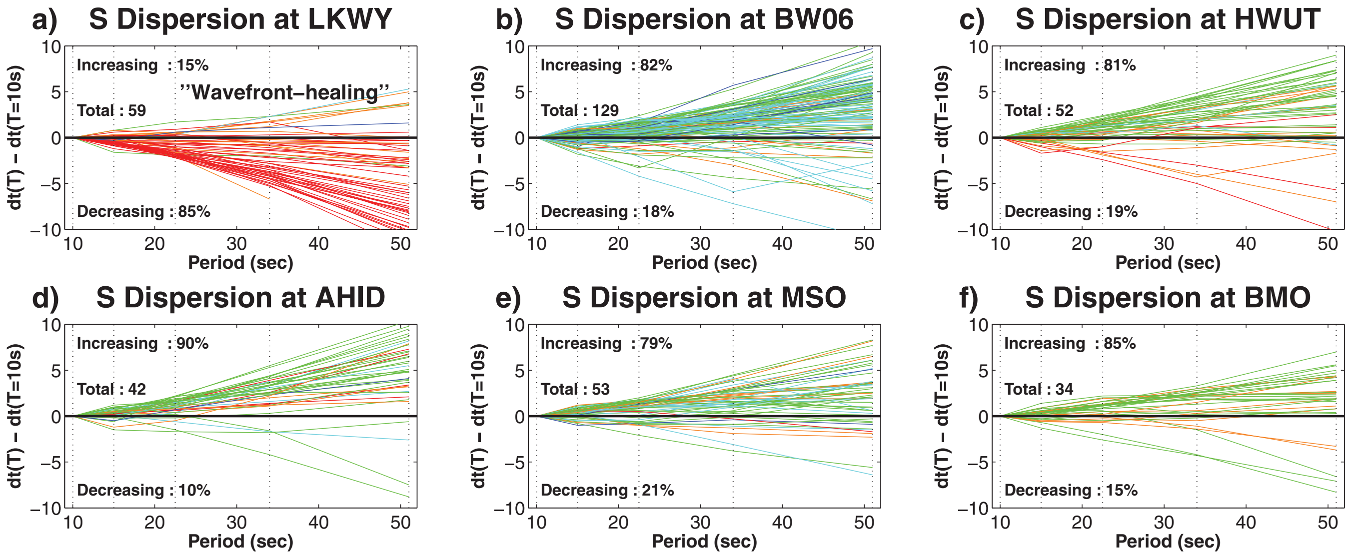

Figure 6. Dispersion curves $dt(T) - dt(T = 10s)$ of S waves recorded at six stations located in the vicinity of the Yellowstone hotspot. We plot dispersion curves with $dt(T = 10s) < 4s$ in cool colours (blue, cyan and green) and dispersion curves with $dt(T = 10s) \geq 4s$ in warm colours (orange and red), where blue/red are for the lowest/highest values of $dt(T = 10s)$. We see that (a) ∼85 per cent of the dispersion curves recorded at station LKWY, which is located above the Yellowstone hotspot, are decreasing and displayed in warm colours; (b–f) at the other stations, ∼83 per cent of the dispersion curves are increasing and mainly displayed in cool colours. This observation suggests that the particular dispersion pattern recorded at LKWY is due to wavefront-healing and related to the crossing of the Yellowstone low-speed anomaly.

velocity anomaly encountered by S waves between the source and the receiver. We therefore plot dispersion curves with $dt(T = 10\,\text{s}) < 4\,\text{s}$ in cool colours (blue, cyan and green) and dispersion curves with $dt(T = 10\,\text{s}) \geq 4\,\text{s}$ in warm colours (orange and red), where blue/red are for the lowest/highest values of $dt(T = 10\,\text{s})$. Fig. 6(a) shows that 85 per cent of the 59 dispersion curves recorded at LKWY are decreasing and associated with $dt(T = 10\,\text{s}) \geq 4\,\text{s}$ (warm colours). Figs 6(b–f) also show that at the other stations, ∼83 per cent of the dispersion curves are increasing and mainly associated with $dt(T = 10\,\text{s}) < 4\,\text{s}$ (cool colours). The large positive time residuals observed at station LKWY are likely to be due to the low-speed anomaly observed below Yellowstone (Fig. 5c). Our observations suggest that the particular dispersion pattern recorded at LKWY is due to wavefront-healing and related to the crossing of the Yellowstone low-speed anomaly.

3.1.2 *Wavefront-healing at global scale*

The case of Yellowstone hotspot (*cf.* Section 3.1.1) suggests that, at least at local scale, our frequency-dependent S-wave travel times contain structural dispersion. In this section, we show that wavefront-healing effect is also present at global scale.

We first consider ∼32 000 S dispersion curves for which S wave traveltimes have been successfully measured at 15, 22.5 and 34 s periods. Measurements at 10 s period were omitted because the number of measurements was not important enough (*cf.* Table 2), mainly because of the oceanic noise and mantle attenuation. Those at 51 s periods were also not used because of the often too large associated errors. With these two restrictions, we were able to extract a large subset of high quality S data (Fig. 7c). Fig. 7(a) shows the percentage of decreasing S dispersion curves as a function of the time residual at 15 s period, $dt(T = 15\,\text{s})$. That is, among all the S dispersion curves $dt(T)$ sharing the same value of $dt(T = 15\,\text{s})$, we plot the relative number of them that are decreasing. For S waves having encountered velocity anomalies producing $-10s \leq dt(T = 15\,\text{s}) < 5\,\text{s}$, the percentage of decreasing dispersion curves is almost constant and equal to ∼25 per cent. However, the percentage of decreasing dispersion curves linearly increases by a factor of 2.5 between $dt(T = 15s) = 5\,\text{s}$, where it is equal to ∼25 per cent, and $dt(T = 15\,\text{s}) = 12\,\text{s}$, where it is equal to ∼65 per cent. This observation suggests that, at global scale, S waves travelling across very low velocity anomalies experience wavefront-healing, a frequency-dependent effect which produces decreasing dispersion curves.

We then consider ∼17 500 SS dispersion curves for which traveltimes have been successfully measured at 15, 22.5 and 34 s periods (Fig. 7d). On Fig. 7(b), the percentage of decreasing dispersion curves associated with $dt(T = 15\,\text{s}) \leq -2\,\text{s}$, corresponding to SS waves having encountered high velocity anomalies, is almost constant and equal to ∼45 per cent. Then, it increases linearly by a factor of 1.5 from ∼45 per cent, at $dt(T = 15\,\text{s}) = -2\,\text{s}$, to ∼65 per cent, at $dt(T = 15\,\text{s}) = 13\,\text{s}$. This behaviour is more difficult to interpret than in the case of S waves. The fact that a smaller increase in the percentage of decreasing dispersion curves is observed over a broader interval of $dt(T = 15\,\text{s})$ values, not always indicating very low velocity anomalies, is at first glance more difficult to associate with wavefront healing. However, it is important to keep in mind that SS waves have a surface reflection at their bounce points, whose associated frequency-dependent crustal effects are not modelled in our WKBJ synthetics (Section 2.2.5). The associated traveltime sensitivity kernel is also more complex than for S waves, especially as SS waves encounter two caustics along their paths (e.g. Hung *et al.* 2000). Their longer journey into the lithosphere also makes them more likely to be affected by strong scattering effects (scattering effects will be discussed in Section 3.2). One part of the signal seen on Fig. 7(b) may be due to wavefront-healing effect. It is however likely that other effects compete and contribute to the SS dispersion. The behaviour of SS waves would reflect their more complex sensitivity to the 3-D structure.

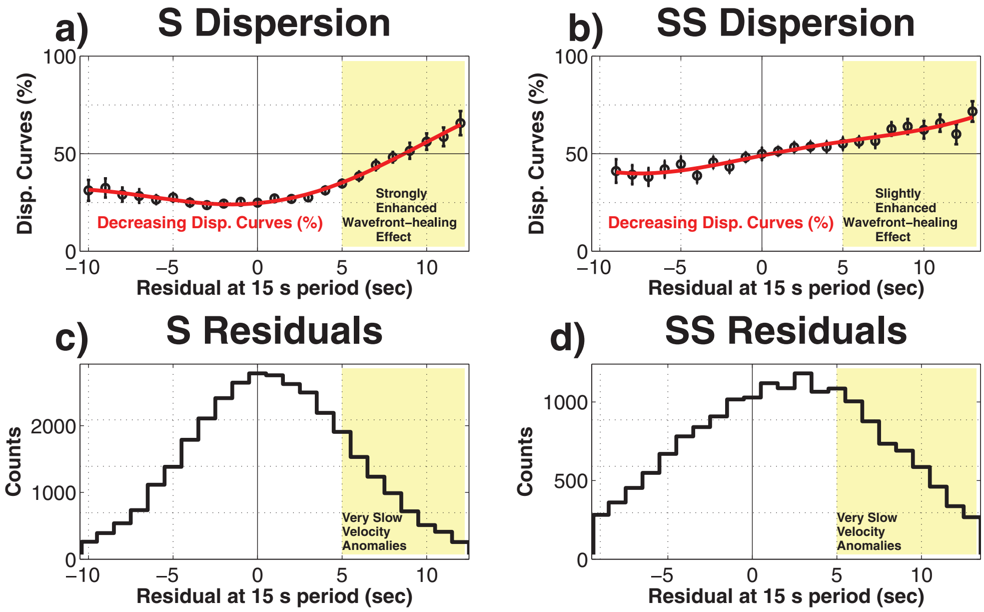

Figure 7. We consider ∼32 000 *S* and ∼17 500 *SS* dispersion curves for which time residuals have been successfully measured at 15, 22.5 and 34 s periods. (a,b) A drastic (smooth) increase in the percentage of *decreasing* dispersion curves is observed for *S* (*SS*) waves having travelled across very low velocity anomalies, associated to highly positive time residuals at 15 s period. This observation suggests that wavefront-healing effect is present at global scale. 2σ–error bars are determined by bootstrap technique. (c,d) Histograms of *S* and *SS* time residuals at 15 s period, showing the very low velocity anomalies producing enhanced wavefront-healing effect.

Finally, we consider ∼7500 *ScS* dispersion curves for which traveltimes have been successfully measured at 15, 22.5 and 34 s periods. We find that ∼85 per cent of these dispersion curves are decreasing. This tendency is also observed from the time residual of our entire *ScS* data set averaged at each period (Table 2). However, the large majority of decreasing dispersion curves cannot be due to wavefront healing, as this would require a preferential sampling of low velocity anomalies. We will see in the next section that, although our *ScS* data set provides a non-uniform sampling of the mantle, there are clear evidences for a preferential sampling of high-velocity anomalies near the CMB.

3.2 Scattering on *ScS* waves at CMB

Our *ScS* data set shows a peculiar behaviour with a large majority of decreasing dispersion curves associated with negative time residuals. In addition to wavefront-healing, we can reject intrinsic attenuation as a possible cause of this peculiar pattern. We show in Sections 2.2.4 and 3.3 that although intrinsic attenuation causes dispersion of seismic velocities, its effect is to produce increasing dispersion curves, by decreasing the velocity of long period waves compared to shorter period ones. In the following, we explore the possibility of explaining the dispersion pattern of our *ScS* data by scattering effect, related to high velocity scatterers located at the CMB.

We consider here a seismogram $s(t)$ as a succession of pulse-like arrivals $u_i(t)$, each with an amplitude A_i and a traveltime τ_i, plus some noise $n(t)$. In the framework of Born theory, we add the contribution δu_i of waves scattered from the wavefront around ray i. If we consider a *S*-wave striking a seismic heterogeneity, because the *S*-wave itself travels the path of minimum time, the scattered signal cannot arrive earlier than the direct wave. However, this does not mean that it always has a delaying influence on the *measured* traveltime (Nolet 2008). The addition of δu_i to u_i deforms the waveshape and therefore may have a delaying or an advancing effect, depending on the sign of the scattered wave. The sign of the scattered wave is determined by the sign of the velocity anomaly that causes the scattered wave. High- and low-velocity scatterers generate scattered waves with negative and positive polarities, respectively (Nolet *et al.* 2005). The effect of adding δu_i is to re-distribute the energy within the cross-correlation window. Under the paraxial approximation, the sensitivity kernel of traveltime with respect to velocity perturbation (Dahlen *et al.* 2000) may be written as

$$K_T^c(\mathbf{r}_x) = -\frac{1}{2\pi c(\mathbf{r}_x)} \times \frac{R_{rs}}{c_r R_{xr} R_{xs}} \times \xi \tag{17}$$

with

$$\xi = \frac{\int_0^\infty \omega^3 |\dot{m}(\omega)|^2 \times \sin[\omega \Delta T(\mathbf{r}_x) - \Delta\Phi(\mathbf{r}_x)]\mathrm{d}\omega}{\int_0^\infty \omega^2 |\dot{m}(\omega)|^2 \mathrm{d}\omega}. \tag{18}$$

$\Delta\Phi$ is the phase shift due to passage through caustics or super critical reflection, R_{rs}, R_{xr} and R_{xs} are the geometrical spreading factors, and ΔT is the detour time of the scattered wave. Unless the wave is supercritically reflected with an angle-dependent phase shift, $\Delta\Phi$ takes three possible values: 0, $-\pi/2$ and $-\pi$ (Dahlen *et al.* 2000; Hung *et al.* 2000). If we only consider *S* and *ScS* phases, we have $\Delta\Phi = 0$. The numerator of eq. (18) then consists of the term sin $(\omega\Delta T)$ modulated by the power spectrum $|\dot{m}(\omega)|^2$ and a factor ω^3. One may expect that the kernel has a maximum near $\omega_0 \Delta T = \pi/2$, or for $\Delta T = T_0/4$, if T_0 is the dominant period of the signal (Nolet 2008). If there is no phase shift, one may assume (e.g. Nolet *et al.* 2005) that δu_i preserves the shape of $u_i(t)$ (they will only differ by their amplitudes). Let the polarity of the scattered wave be negative, as for a high-velocity anomaly. The measurement process consists of cross-correlating the observed and synthetic waveforms, for instance filtered around the period $T_0 = 10$ s. The time residual *dt* corresponds to the maximum of the cross-correlation of the perturbed wave $u_i(t) + \delta u_i(t)$ (i.e. the observed waveform) with the unperturbed wave $u_i(t)$ (i.e. the synthetic waveform). The observed waveform is expected to be dominated by arrivals of scattered waves with detour times close to $\Delta T_i(T_0 = 10\,\mathrm{s}) = T_0/4 = 2.5\,\mathrm{s}$, corresponding to the maximum sensitivity of the associated kernel. The contribution of these scattered waves is to decrease the amplitude of the observed waveform, around the time $t \simeq \tau_i + \Delta T_i(T_0)$, compared to the synthetic waveform, such as this will have an advancing effect on the time residual *dt*. We have checked that this advancing effect may be expected to increase with the period T_0. For instance, at $T_0 = 34\,\mathrm{s}$, the observed waveform should be dominated by scattered waves with detour times close to $\Delta T_i(T_0 = 34s) = 8.5\,\mathrm{s}$. This will then decrease the amplitude of a latter part of the observed waveform, which means a greater advancing effect on *dt*. Therefore, in regions where high velocity scatterers dominate, we expect an apparent dispersion with $dt(T_0 = 10\,\mathrm{s}) > dt(T_0 = 34\,\mathrm{s})$, corresponding to a decreasing dispersion curve $dt(T)$. In regions where low-velocity scatterers dominate, we have checked that we may expect an increasing dispersion curve. In such low-velocity regions, we also expect that wavefront-healing (*cf.* Section 3.1) and scattering effects are competing.

A significant difference between *ScS* waves and the remaining part of our data set is that *ScS* waves cross the D″ discontinuity, which is located ∼300 km above the CMB. This D″ discontinuity is associated with a sharp increase in *S*-wave velocity and marks the top of a very heterogeneous zone at the bottom of the mantle. This region is not sampled by our deepest *S* and *SS* waves, which bottom near 2400 km depth. Deeper *S* waves interfere with the *ScS* waveforms and have been rejected by our selection process (*cf.* Appendix A). Using Sdiff waves would help to better understand frequency-dependent effects on global *S* waves in the D″ layer (e.g. To & Romanowicz 2009). However, Sdiff are not used in this study, as they cannot be properly synthesized with WKBJ synthetics.

We consider ∼3300 earthquake-station couples in the epicentral distance range 55–70 degrees, with both *S* and *ScS* dispersion curves successfully measured at 15, 22.5 and 34 s periods. At these distances, *S* and *ScS* waves have very similar traveltime sensitivity kernels except near the bottom of the mantle (Figs 8b and c), so that we can attribute their traveltime differences to velocity anomalies located above the CMB. Figs 8(a) and (d) show that the high-velocity ring around the Pacific and in eastern Asia at the CMB is preferentially sampled by our restricted *ScS* data set. The fast anomalies at the CMB are thought to be a collection of slab material (Van der Hilst *et al.* 1997), although this interpretation is still difficult to prove or disprove. Houser *et al.* (2008) also find fast anomalies at the CMB surrounding the entire Pacific Plate and attribute them to the cold remnants of past subduction. Very few of our *ScS* waves cross the low-velocity anomalies present at the base of the mantle over much of the central Pacific Ocean and beneath South Africa (e.g. Ritsema *et al.* 1999). For this restricted data set, we find that ∼85 per cent of the dispersion curves are decreasing for *ScS* waves, compared to ∼25 per cent for *S* waves (Figs 8e–f). This suggests that scattering effect, related to a preferential sampling of high-velocity scatterers located at the base of the mantle, is a plausible explanation for the peculiar dispersion observed for *ScS* waves.

3.3 Frequency-dependent attenuation

The mantle acts as an absorption band for seismic waves (e.g. Anderson 1976) and attenuation *q* depends on the frequency of oscillation. Within the absorption band, attenuation is relatively high and its frequency-dependent effects are expected to be weak for long period body waves (e.g. Sipkin & Jordan 1979), that is within the 10–51 s period range of analysis of this study. The frequency dependence of the attenuation *q* can be described by a power law $q \propto \omega^{-\alpha}$ (eq. 11), with a model-dependent α, usually thought to be smaller than 0.5 (Anderson & Minster 1979). Constraining the frequency dependence of intrinsic seismic attenuation in the Earth's mantle is crucial to properly correct for velocity dispersion due to attenuation. Global tomographic models usually rely on a frequency-independent attenuation model (Kanamori & Anderson 1977), corresponding to the case $\alpha = 0$. A non-zero α implies that seismic waves of different frequencies are differently

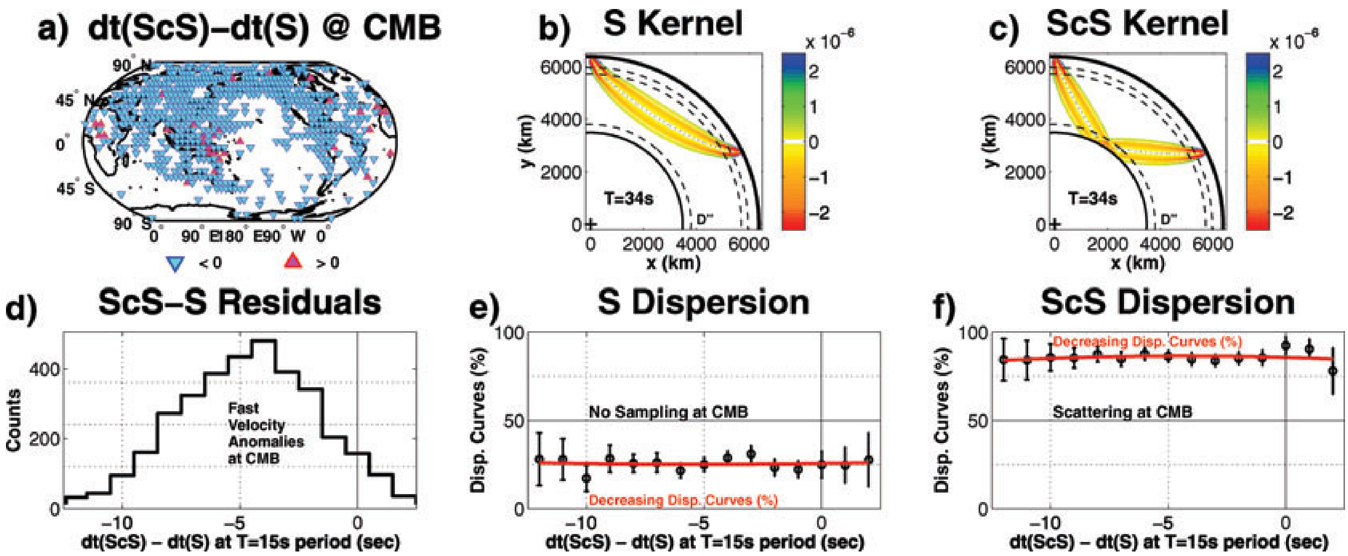

Figure 8. We selected a set of ∼3300 epicentre-station couples in the distance range 55–70 degrees, with both *S* and *ScS* dispersion curves successfully measured at 15, 22.5 and 34 s periods. (a) Difference of time residuals at 15 s period between *ScS* and *S* waves, that is $dt_{ScS}(T = 15\,\mathrm{s}) - dt_S(T = 15\,\mathrm{s})$, averaged in $6^\circ \times 6^\circ$ cells and geographically plotted at their corresponding CMB locations. (b) Traveltime sensitivity Fréchet kernel (in $\mathrm{s\,km^{-3}}$) for *S* wave, computed using the software by Tian *et al.* (2007b). (c) Fréchet kernel for *ScS* wave. (d) Histogram of *ScS*–*S* residuals at 15 s period. (e,f) Our results show that ∼85 per cent of the dispersion curves are decreasing for *ScS* waves, compared to ∼25 per cent for *S* waves. This argues in favour of strong scattering effect occurring on *ScS* waves, owing to preferential sampling of high velocity scatterers at CMB. 2σ–error bars are determined by bootstrap technique.

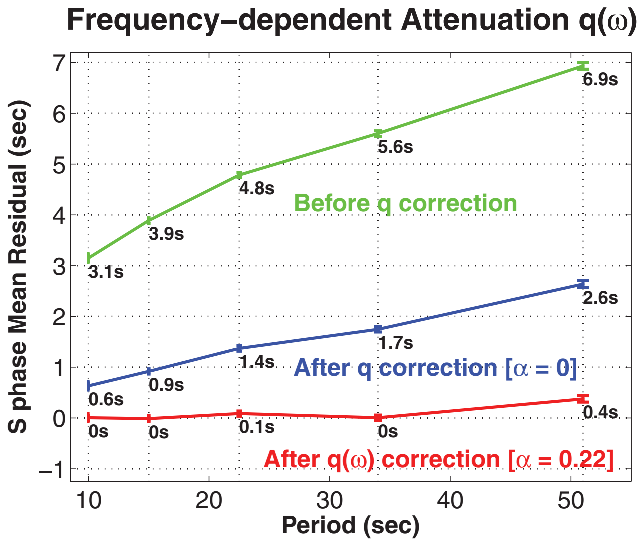

Figure 9. The green curve represents the globally averaged S-wave time residual, $\mu_S(T)$, at each period T between 10 and 51 s, with no attenuation correction applied to S traveltimes. The blue curve represents $\mu_S(T)$ corrected with a frequency-independent attenuation model, corresponding to $\alpha = 0$. The red curve represents $\mu_S(T)$ corrected with a 'frequency-dependent' attenuation model, $q(\omega) \propto q_0 \times \omega^{-\alpha}$, corresponding to a non-zero value of α. Our observations show that $\alpha \sim 0.2$ better accounts for our S observations, as it predicts $\mu_S(T) \sim 0$ in the full 10– 51 s period range. 2σ–error bars are determined by bootstrap technique.

attenuated, and accordingly modifies the velocity dispersion relation (Section 2.2.4).

Despite observational and experimental advances, no clear consensus concerning the value of α for the Earth's mantle has emerged over the past 25 yr. Nevertheless, theoretical predictions of $\alpha > 0$ have been systematically confirmed in various laboratory studies. A recent review by Romanowicz & Mitchell (2007) identifies a number of studies that collectively constrain α to the 0.1–0.4 range. Using normal mode and surface wave attenuation measurements, Lekic *et al.* (2009) find that $\alpha = 0.3$ should better approximate the α representative of the average mantle, at periods between 1 and 200 s. Their preferred model of frequency dependence of attenuation is also consistent with other studies that have relied upon body waves and have focused on higher frequencies. Looking at S/P ratios at periods lower than 25 s, several studies (Ulug & Berckhemer 1984; Cheng & Kennett 2002) have argued for α values in the 0.1–0.6 range. Shito *et al.* (2004) used continuous P-wave spectra to constrain α between 0.2 and 0.4 at periods shorter than 12 s. Flanagan & Wiens (1998) found an α value of 0.1–0.3 was needed to reconcile attenuation measurements on sS/S and pP/P phase pairs in the Lau basin.

In this study, we have measured globally distributed multiple-frequency time residuals of thousands of S waves, within the 10–51 s period range (Table 2). These measurements have been corrected from physical dispersion relying on a frequency-independent attenuation model (Kanamori & Anderson 1977). Sampling of the Earth's (lower) mantle corresponding to our S data set is mostly global (*cf.* Fig. 3). Table 2 shows the globally averaged time residual of S waves at each period T between 10 and 51 s, denoted by $\mu_S(T)$ in the following. We observe that, when the period T increases, the globally averaged time residual $\mu_S(T)$ slightly increases (*cf.* the blue curve on Fig. 9). At first glance, it is very difficult to explain with scattering effect only that $\mu_S(T)$ is positive and increases within our period range. Indeed, this would require a preferred sampling of low-velocity scatterers (Section 3.2) in the mantle, above 2400 km depth, for which there is no evidence at global scale. Wavefront-healing cannot explain such a positive and increasing averaged dispersion curve $\mu_S(T)$ (Section 3.1).

By only considering attenuation effect, long period seismic waves should arrive later than short period ones (*cf.* Section 2.2.4 and Fig. 9). An underestimation of this effect in our attenuation correction may therefore account for a major part of the observed increasing behaviour of $\mu_S(T)$. Here, we propose to explain the averaged S residual dispersion, remaining after the common correction of physical dispersion with $\alpha = 0$, by taking into account the possible frequency-dependency of attenuation with a non-zero α. We find that a frequency-dependent attenuation with $\alpha \sim 0.2$ better accounts for our frequency-dependent S traveltimes, as it predicts a globally averaged time residual $\mu_S(T)$ very close to zero at each period (*cf.* the red curve on Fig. 9). This value of $\alpha \sim 0.2$ is close to the value of 0.3 found by Lekic *et al.* (2009) for the average mantle, at periods lower than 200 s (and longer than 1 s). An α value of 0.2 is also compatible with other studies (e.g. Romanowicz & Mitchell 2007).

We need however to consider that there is a trade-off between Q (i.e. t^*) and α, as shown by eq. (16). That is, when considering a single S wave propagating in the mantle, we might also explain its residual dispersion by varying both Q and α. In this study, we use the radial PREM Q model, because 3-D variations of Q are not well constrained in the Earth's mantle. We believe that, considering a radial (1-D) Q model to interpret the observed globally averaged S residual dispersion (*cf.* Fig. 9), is reasonable because our thousands of S waves average the 3-D variations of Q sufficiently well in the average mantle. Our results suggest that applying a frequency-dependent attenuation correction with $\alpha \sim 0.2$, is a plausible explanation for the averaged residual dispersion of S waves observed in the entire 10–51 s period range.

Table 2 also suggests a slight increase of the globally averaged time residual $\mu_{SS}(T)$ for SS waves in the 10–51 s period range. In this case, we find that a frequency-dependent attenuation, with $\alpha \sim 0.1$, better accounts for our frequency-dependent SS traveltimes, as it predicts a globally averaged time residual $\mu_{SS}(T)$ very close to zero at each period. Compared with S waves, SS waves experience a longer journey into the lithosphere and upper mantle. It is therefore possible that the different α values obtained with SS and S waves reflect their different sampling of the Earth's mantle.

We observe a decrease of the averaged time residual $\mu_{ScS}(T)$ for ScS waves in the 10–51 s period range (Table 2). In this case, a frequency-dependent attenuation, with $\alpha > 0$, would reinforce the decreasing trend of the ScS residual dispersion. Our ScS dispersion pattern can therefore not be explained by a frequency-dependent attenuation with $\alpha > 0$. This favours scattering, instead of attenuation, to explain the particular ScS dispersion pattern (Section 3.2).

Our observation that frequency-dependent effects of Q might explain the averaged residual dispersion of our global S data set is compatible with the idea that other diffraction phenomena (e.g. wavefront-healing and scattering) can be predominant on individual data. As far as physical dispersion remains weak compared to the observed residual dispersion, the error that we make in the evaluation of this physical dispersion correction is unlikely to change the residual dispersion patterns we have observed and related to structural effects (*cf.* Sections 3.1 and 3.2). We have checked that wavefront-healing is similarly observed in our S and SS data sets with a new attenuation correction corresponding to a non-zero α.

This conclusion supports other previous studies which suggest that incorporating anelastic dispersion cannot completely account for the observed *S*-wave discrepancy (e.g. Liu *et al.* 1976; Baig & Dahlen 2004).

4 CONCLUSION

We have built a global database of ∼400 000 *S, ScS* and *SS* traveltimes measured at five different periods (10, 15, 22.5, 34 and 51 s). An automated scheme for measuring long period body wave traveltimes in different frequency bands has been presented. The scheme comprises of two main parts. The first involves an automated selection of time windows around the target phases present on both the observed and synthetic seismograms. The second stage involves measurements of multiple-frequency traveltimes by cross-correlating the selected observed and synthetic filtered waveforms. Frequency-dependent effects due to crustal reverberations beneath each receiver are handled by incorporating crustal phases into WKBJ synthetic waveforms. The obtained multiple-frequency *S*-wave traveltimes are well suited for global multiple-frequency tomographic imaging of the Earth's mantle.

After correction for physical dispersion due to intrinsic anelastic processes, we observe a residual dispersion on the order of 1–2 s in the period range of analysis. This dispersion occurs differently for *S, ScS* and *SS*, which is presumably related to their differing paths through the Earth. Our results show that: (1) Wavefront-healing phenomenon produced by very low velocity anomalies is observed in our *S* and, to a lesser extent, *SS* traveltimes. (2) A preferred sampling of high velocity scatterers located at the CMB may explain our observation that *ScS* waves travel faster at low frequency than at high frequency. (3) The globally averaged dispersion observed for *S* and *SS* traveltimes favour a frequency-dependent attenuation model $q(\omega) \propto q_0 \times \omega^{-\alpha}$, with an α value of ∼0.2 for *S* waves and ∼0.1 for *SS* waves.

Our results therefore suggest that the residual dispersion observed in our data is, at least partly, related to seismic heterogeneity and attenuation in the Earth's interior. With this, we feel that tomographic reconstruction schemes, that explicitly take account of frequency dependency, should help to build a more accurate picture of the Earth's mantle. Our expectations are that, with the newly processed observations, one may be able to shed light on some key small-scale features present in the mantle, and in doing so, better constrain mantle dynamics.

ACKNOWLEDGMENTS

This work was supported by the young researcher ANR TOMOGLOB no ANR-06-JCJC-0060. The authors thank the Iris and Geoscope data centres for providing seismological data. Discussions with M. Cara, J-J. Leveque, A. Maggi, L. Rivera and B. Tauzin have been stimulating at various stages of this research. The authors thank J. Ritsema, J. Trampert and an anonymous reviewer for helpful reviews that improved the original paper.

REFERENCES

Albarede, F. & Van der Hilst, R.D., 1999. New mantle convection model may reconcile conflicting evidence, *EOS, Trans. Am. geophys. Un.,* **45,** 535–539.

Anderson, D., 1976. The Earth as a seismic absorption band, *Science,* **196,** 1104–1106.

Anderson, D. & Minster, J., 1979. The frequency dependence of *Q* in the Earth and implications for mantle rheology and Chandler wobble, *Geophys. J. R. astr. Soc.,* **58,** 431–440.

Bai, C.-Y. & Kennett, B.L.N., 2001. Phase identification and attribute analysis of broadband seismograms at far regional distances, *J. Seismol.,* **5,** 217–231.

Baig, A.M. & Dahlen, F.A., 2004. Traveltimes biases in random media and the S-wave discrepancy, *Geophys. J. Int.,* **158,** 922–938.

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

Boschi, L., 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.

Burdick, S. et al., 2008. Upper Mantle Heterogeneity beneath North America from Travel Time Tomography with Global and USArray Transportable Array Data, *Seism. Res. Lett.,* **79,** 384–392.

Calvet, M. & Chevrot, S., 2005. Traveltime sensitivity kernels for PKP phases in the mantle, *Phys. Earth planet. Inter.,* **153,** 21–31.

Chapman, C., 1978. A new method for computing synthetic seismograms, *Geophys. J. R. astr. Soc.,* **54,** 481–518.

Cheng, H.X. & Kennett, B., 2002. Frequency dependence of seismic wave attenuation in the upper mantle beneath the Australian region, *Geophys. J. Int.,* **150,** 45–57.

Chevrot, S., 2002. Optimal measurement of relative and absolute delay-times by simulated annealing, *Geophys. J. Int.,* **151,** 164–171.

Dahlen, F.A. & Nolet, G., 2005. Comment on 'On sensitivity kernels for 'wave-equation' transmission tomography' by de Hoop and Van der Hilst, *Geophys. J. Int.,* **163,** 949–951.

Dahlen, F.A., Hung, S.-H. & Nolet, G., 2000. Fréchet kernels for finite-frequency travel times – I. Theory, *Geophys. J. Int.,* **141,** 157–174.

de Hoop, M.V. & Van der Hilst, R.D., 2005a. On sensitivity kernels for 'wave equation' transmission tomography, *Geophys. J. Int.,* **160,** 621–633.

de Hoop, M.V. & Van der Hilst, R.D., 2005b. Reply to comment by F.A. Dahlen and G. Nolet on 'On sensitivity kernels for 'wave equation' transmission tomography', *Geophys. J. Int.,* **163,** 952–955.

Devilee, R.J.R., Trampert, J. & Paulssen, H., 2003. Dispersion measurements of P waves and their implications for mantle Qp, *Pure appl. geophys.,* **160,** 2223–2238.

Dziewonski, A.M., 1984. Mapping the lower mantle: determination of lateral heterogeneity in P velocity up to degree and order 6, *J. geophys. Res.,* **89,** 5929–5952.

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

Engdahl, E.R., Van der Hilst, R.D. & Bulland, R.P., 1998. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. seism. Soc. Am.,* **88,** 722–743.

Flanagan, M.P. & Wiens, D.A., 1998. Attenuation of broadband P and S waves in Tonga: observations of frequency dependent Q, *Pure appl. geophys.,* **153,** 345–375.

Fukao, Y., Widiyantoro, S. & Obayashi, M., 2001. Stagnant slabs in the upper and lower mantle transition region, *Rev. Geophys.,* **39,** 291–323.

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

Gudmundson, O., 1996. On the effect of diffraction on traveltime measurements, *Geophys. J. Int.,* **124,** 304–314.

Houser, C., Masters, G., Shearer, P.M. & Laske, G., 2008. Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms, *Geophys. J. Int.,* **174,** 195–212.

Hung, S.-H., Dahlen, F.A. & Nolet, G., 2000. Fréchet kernels for finite-frequency travel times – II. Examples, *Geophys. J. Int.,* **141,** 175–203.

Hung, S.-H., Dahlen, F.A. & Nolet, G., 2001. Wavefront-healing: a banana–doughnut perspective, *Geophys. J. Int.,* **146,** 289–312.

Julian, B.R., 2005. What can seismology say about hot spots?, in *Plates, Plumes, and Paradigms,*Vol. 388, pp. 155–170, eds Foulger, G.R., Natland, J.H., Presnall, D.C. & Anderson, D.L., Geological Society of America.

Kanamori, H. & Anderson, D.L., 1977. Importance of physical dispersion in surface wave and free oscillation problems, *Rev. Geophys. Space Phys.,* **15,** 105–12.

Kennett, B. & Engdahl, E., 1991. Travel times for global earthquake location and phase identification, *Geophys. J. Int.,* **105,** 429–465.

Komatitsch, D. & Tromp, J., 1995. Spectral-element simulations of global seismic wave propagation – I. Validation, *Geophys. J. Int.,* **122,** 108–124.

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

Lekic, V., Matas, J., Panning, M. & Romanowicz, B., 2009. Measurement and implications of frequency dependence of attenuation, *Earth planet. Sci. Lett.,* **282,** 285–293.

Liu, H.-P, Anderson, D.L. & Kanamori, H., 1976. Velocity dispersion due to anelasticity; implications for seismology and mantle composition, *Geophys. J. R. astr. Soc.,* **47,** 41–58.

Marquering, H., Dahlen, F.A. & Nolet, G., 1999. Three-dimensional sensitivity kernals for finite-frequency traveltimes: the banana–doughnut paradox, *Geophys. J. Int.,* **137,** 805–815.

Masters, G., Johnson, S., Laske, G. & Bolton, H., 1996. A shear velocity model of the mantle, *Phil. Trans. R. Soc. Lond.,* **A354,** 1385–1410.

Masters, G., Laske, G., Bolton, H. & Dziewonski, A.M., 2000. The relative behaviour of shear velocity, bulk sound speed and compressionnal velocity in the mantle: implications for chemical and thermal structure, in *Earth's Deep Interior,* pp. 63–88, Karato, S., Forte, A., Liebermann, R.C., Masters, G. & Stixrude, L., eds, AGU, Washington, D.C.

Maggi, A., Tape, C., Chen, M., Chao, D. & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.,* **178,** 257–281.

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

Montelli, R., Nolet, G., Masters, G., Dahlen, F.A. & Hung, S.-H., 2004b. Global P and PP traveltime tomography: rays versus waves, *Geophys. J. Int.,* **158,** 637–654.

Montelli, R., Nolet, G., Dahlen, F.A. & Masters, G., 2006a. Comment on 'Banana–doughnut kernels and mantle tomography' by van der Hilst and de Hoop, *Geophys. J. Int.,* **167,** 1204–1210.

Montelli, R., Nolet, G., Dahlen, F.A. & Masters, G., 2006b. A catalogue of deep mantle plumes. New results from finite-frequency tomography, *Geochem. Geophys. Geosyst.,* **7,** Q11007, doi:10.1029/2006GC001248.

Mooney, W.D., Laske, G. & Masters, G., 1998. CRUST5.0: a global crustal model at $5^\circ \times 5^\circ$, *J. geophys. Res.,* **103,** 727–747.

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

Nolet, G., 2008. *A Breviary of Seismic Tomography,* Cambridge University Press, Cambridge, UK.

Nolet, G. & Dahlen, F.A., 2000. Wave front healing and the evolution of seismic delay times, *J. geophys. Res.,* **105,** 19 043–19 054.

Nolet, G., Dahlen, F.A. & Montelli, R., 2005. Traveltimes and amplitudes of seismic waves: a re-assessment, in *Array Analysis of Broadband Seismograms,* pp. 37–48, eds Levander, A. & Nolet, G., AGU Monograph Series.

Pollitz, F.F., 2007. Finite-frequency travel time tomography of San Francisco Bay region crustal velocity structure, *Geophys. J. Int.,* **171,** 630–656.

Ritsema, J. & van Heijst, H.-J., 2002. Constraints on the correlation of P and S wave velocity heterogeneity in the mantle from P, PP, PPP and PKPab travel times, *Geophys. J. Int.,* **149,** 482–489.

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

Ritsema, J., van Heijst, H.-J., Woodhouse, J.H. & Deuss, A., 2009. Long-period body wave traveltimes through the crust: implications for crustal corrections and seismic tomography, *Geophys. J. Int.,* **179,** 1255–1261.

Romanowicz, B., 2003. Global mantle tomography: progress status in the past 10 years, *Ann. Rev. Earth planet. Sci.,* **31,** 303–328.

Romanowicz, B. & Mitchell, B., 2007. Deep Earth structure: *Q* of the Earth from crust to core, *Treatise on Geophysics,* **1,** pp. 731–774, ed. Schubert, G., Elsevier, Amsterdam.

Shito, A., Karato, S.I. & Park, J., 2004. Frequency dependence of Q in Earth's upper mantle inferred from continuous spectra of body waves, *Geophys. Res. Lett.,* **31,** L24614, doi:10.1029/2004GL019582.

Sieminski, A., Leveque, J.J. & Debayle, E., 2004. Can finite-frequency effects be accounted for in ray theory surface wave tomography?, *Geophys. Res. Lett.,* **31,** L24614, doi:10.1029/2004GL021402.

Sigloch, K. & Nolet, G., 2006. Measuring finite-frequency body wave amplitudes and travel times, *Geophys. J. Int.,* **167,** 271–287.

Sipkin, S.A. & Jordan, T.H., 1979. Frequency dependence of Q_{ScS}, *Bull. seism. Soc. Am.,* **69**. 1055–1079, doi:10.1785/0120060038.

Tian, Y., Nolet, G., Hung, S.-H., Montelli, R. & Dahlen, F.A., 2007a. Dynamic ray tracing and traveltime corrections for global seismic tomography, *J. Comput. Phys.,* **226,** 672–687.

Tian, Y., Montelli, R., Nolet, G. & Dahlen, F, A., 2007b. Computing traveltime and amplitude sensitivity kernels in finite-frequency tomography, *J. Comput. Phys.,* **226,** 2271–2288.

To, A. & Romanowicz, B., 2009. Finite frequency effects on global S diffracted traveltimes, *Geophys. J. Int.,* **179,** 1645–1657.

Trampert, J. & Spetzler, J., 2006. Surface wave tomography: finite frequency effects lost in the null space, *Geophys. J. Int.,* **164,** 394–400.

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana–doughnut kernels, *Geophys. J. Int.,* **160,** 195–216.

Ulug, A. & Berckhemer, H., 1984. Frequency dependence of Q for seismic body waves in the Earth's mantle, *J. Geophys.,* **56,** 9–19.

Van der Hilst, R.D. & de Hoop, M.V., 2005. Banana–doughnut kernels and mantle tomography, *Geophys. J. Int.,* **163,** 956–961.

Van der Hilst, R.D. & de Hoop, M.V., 2006. Reply to comment by R. Montelli, G. Nolet and F. A. Dahlen on 'Banana–doughnut kernels and mantle tomography', *Geophys. J. Int.,* **167,** 1211–1214.

Van der Hilst, R.D., Widiyantoro, S. & Engdahl, E.R., 1997. Evidence for deep mantle circulation from global tomography, *Nature,* **386,** 578–584.

Yang, T. & Shen, Y., 2006. Frequency-dependent crustal corrections for FF seismic tomography, *Bull. seism. Soc. Am.,* **96**(6), 2441–2448, doi:10.1785/0120060038.

APPENDIX A: TIME WINDOWS SELECTION

We describe here our time windows selection and seismic phase isolation methodology, which largely makes use of the ideas of Maggi *et al.* (2009). The main differences are the following: (1) in step 1 we work on the rotated *SH* component; (2) in steps 2, 3, 4 and 5, we perform all the operations on the observed and synthetic seismograms (i.e. not on the synthetic only); (3) in step 6, we test all possible combinations of time windows before deciding on an optimal pair of observed and synthetic waveforms, corresponding to the target seismic phase. For this study, our codes have been tuned to measure traveltimes of *SH* waves, which have the advantage of being free of *P* energy. There is no difficulty in applying the same approach to *SV* and *P* waves, although it is likely that interference between *S* and *P* energy would result in fewer windows surviving the selection criteria.

A1 Step 1: pre-selection

The purpose of this step is to pre-process input seismograms and reject noisy records. Three components seismograms are first rotated along the *SH* component. The observed seismograms are then bandpass filtered with a non-causal Butterworth filter, whose short- and long-period corners are denoted by T_1 and T_2, respectively.

Table A1. Overview of standard and fine tuning parameters, as defined by Maggi *et al.* (2009), with the values used in this study.

Standard tuning parameters	
$r_P = 1$	Power signal-to-noise ratio threshold
$r_A = 2.5$	Amplitude signal-to-noise ratio threshold
$r_0 = 2.5$	Signal-to-noise ratio single windows
$WL = 0.1$	Water level
Fine tuning parameters	
$C_0 = 0.7$	For rejection of internal minima
$C_1 = 3$	For rejection of short windows
$C_2 = 0.1$	For rejection of un-prominent windows
$C_{3a} = 1, C_{3b} = 2$	For rejection of multiple distinct arrivals
$C_{4a} = 3, C_{4b} = 12$	For curtailing of windows

Where the tuning parameters C_{3a}, C_{3b}, C_{4a} and C_{4b} are used in this study as described by Maggi *et al.* (2009).

Following Maggi *et al.* (2009), we define the time-normalized power in the signal and noise portions of the data by

$$P_{\text{signal}} = \frac{1}{t_E - t_A} \int_{t_A}^{t_E} d^2(t)\, \mathrm{d}t \tag{A1}$$

and

$$P_{\text{noise}} = \frac{1}{t_A - t_0} \int_{t_0}^{t_A} d^2(t)\, \mathrm{d}t, \tag{A2}$$

where $d(t)$ is the observed seismogram, t_0 is the start time, t_A is set to be slightly before the time of the first arrival and t_E is the end of the main signal. We compute the power signal-to-noise ratio $SNR_P = P_{\text{signal}}/P_{\text{noise}}$ and the amplitude signal-to-noise ratio $SNR_A = A_{\text{signal}}/A_{\text{noise}}$, where A_{signal} and A_{noise} are the maximum values of $|d(t)|$ in the signal and noise time-spans, respectively. We reject records for which $SNR_P < r_P$ or $SNR_A < r_A$ (Table A1).

A2 Step 2: short-term/long-term average ratios

Seismic phase arrivals are detected using standard short-term/long-term (STA:LTA) average ratios. In contrast to Maggi *et al.* (2009), this STA:LTA analysis is applied to both the observed and synthetic seismograms. We first compute the envelopes of the observed and synthetic seismograms. The envelope $e(t)$ of a seismogram $s(t)$, whose Hilbert transform is noted $H[s(t)]$, is given by:

$$e(t) = |s(t) + i \cdot H[s(t)]|. \tag{A3}$$

Assuming that both observed and synthetic waveforms are discretized with time-step δt, we compute the short-term $S(i)$ and long-term $L(i)$ averages for each time sample i of the envelope, using:

$$S(i) = C_S \cdot S(i-1) + e(i) \tag{A4}$$

$$L(i) = C_L L(i-1) + e(i) \tag{A5}$$

and evaluate their ratios

$$E(i) = S(i)/L(i). \tag{A6}$$

The constants C_S and C_L determine the decay of the relative weighting of earlier parts of the signal in the calculation of the current average. Following Bai & Kennett (2001), we use $C_S = 10^{-\delta t/T_1}$ and $C_L = 10^{-\delta t/12T_1}$. Fig. A1(a) shows an example of observed and synthetic waveforms. The corresponding envelopes $e(t)$ and STA:LTA waveforms $E(i)$ are shown in Figs A1(b) and (c), respectively.

A3 Step 3: time windows isolation

At this stage, the intention is to list all possible time windows present on the observed and synthetic STA:LTA waveforms $E(i)$. As underlined by Maggi *et al.* (2009), the agreement between local maxima in $E(i)$ and the position of seismic phases on the observed and synthetic seismograms, as well as the correspondence between local minima and the transitions between successive phases, suggest that time windows should start and end at local minima surrounding a maximum in $E(i)$ (Fig. A1). We first select all maxima in $E(i)$ lying above a given water level *WL* (Table A1) on the observed and synthetic waveforms. The water level is identical for the observed and synthetic waveforms $E(i)$. Each maximum is then taken as a 'seed' maximum about which all possible candidate time windows can be created around it. The time windows start and end at local minima of the STA:LTA waveforms $E(i)$. We consider all local minima before the seed maximum as a potential start time for the window, and all local minima after the seed maximum as a potential end time. Therefore, each candidate time window is defined by three times: its start time, its end time and the time of its seed maximum.

A4 Step 4: shape-based time windows rejection

At this stage, we are left with a list of possible time windows surrounding the target phases present on the observed and synthetic seismograms. We first reject windows based on the shape of the STA:LTA waveforms $E(i)$. The aim of this shape-based window rejection is to extract observed and synthetic time windows with well-developed single phases (e.g. *S*) or groups of phases (e.g. *S*+*sS*). We use the same criteria as in Maggi *et al.* (2009), except that we apply them on both the synthetics and observed STA:LTA waveforms. First, we reject all time windows that contain internal local minima in $E(i)$ whose amplitude is less than $C_0 WL$ (Table A1). This choice forces a partition of unequivocally distinct seismic phases into separate time windows. Secondly, we reject short windows whose length is smaller than $C_1 T_1$ (Table A1). This criteria allows us to reject windows which are too short to contain useful information. Thirdly, we reject time windows whose seed maximum rises by less than $C_2\ WL$ above either of its adjacent minima (Table A1). Finally, we reject time windows containing at least one strong phase arrival that is well separated in time from the seed maximum time. This allows us to distinguish inseparable phase groups from distinct seismic phases that arrive close in time.

A5 Step 5: SNR and time interval based windows rejection

At this stage, we are left with several pairs of observed and synthetic time windows containing well-developed single (or groups of) seismic phase(s). We wish to extract the optimal pair of observed and synthetic time windows for each target phase. This task is not trivial to implement in an automated way. One of the main difficulties is that, in most cases, the observed and synthetic time windows corresponding to the same target phase have different start and end times. This is especially true when the time residual, between the observed and 1-D synthetic seismic phases, becomes large. In addition, when the target phase interferes with other phases, our automated scheme should ideally ensure that the observed and synthetic waveforms, present in the retained time windows, do carry the same pattern of interference.

First, we compute for each (observed and synthetic) candidate time window a signal-to-noise ratio (SNR): $SNR_w = A_{\text{window}}/A_{\text{noise}}$, where A_{window} and A_{noise} are the maximum absolute amplitude

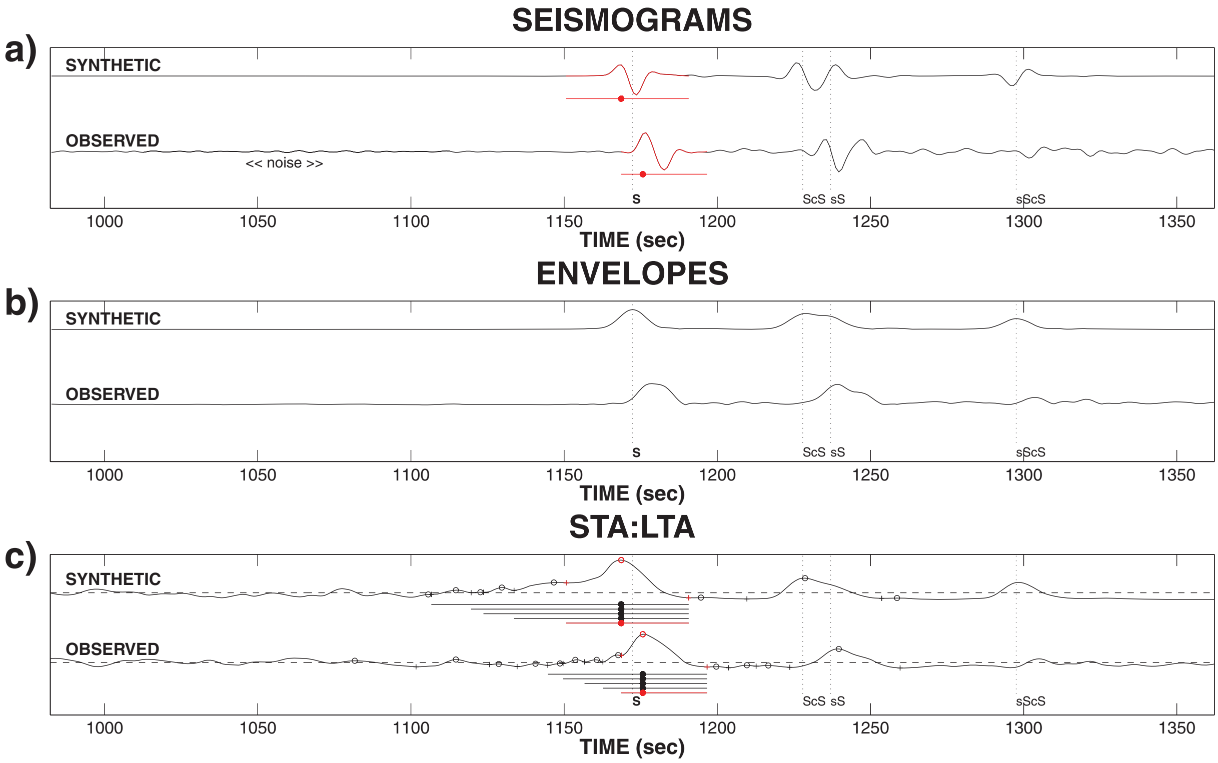

Figure A1. Time windows selection process, illustrated for the observed and synthetic S waveforms isolation. (a) Observed and synthetic seismograms for a deep earthquake located at Kyushu, Japan (date: 2005/11/21, centroid depth: 155 km, body wave magnitude: 5.9) and recorded at CAN station (Geoscope network). Seismograms are discretized with a time step δt of 1 s, and filtered with a Butterworth filter (corner frequencies: $T_1 = 7$ s and $T_2 = 85$ s). The optimal observed and synthetic time windows are shown in red. (b) Observed and synthetic envelopes. (c) Corresponding STA:LTA waveforms. In black dashed line overlaid on the STA:LTA waveforms is the water level (WL). Local maxima and minima of STA:LTA waveforms are denoted by circles and crosses, respectively. All the observed and synthetic time windows represented in black (and red for the optimal pair), correspond to the candidate windows having survived several selection criteria (i.e. steps 4 and 5 of Sections A4 and A5). The selection of the optimal pair (represented in red) among these candidate time windows is explained in Section A6. Note that the amplitude of the observed and synthetic seismograms have been normalized to 1.

values of the seismic signal contained in the candidate time window and in the noise time-span, respectively. We reject each (observed and synthetic) candidate time window if $SNR_w < r_0$ (Table A1). Second, on the synthetic seismogram, we retain time windows around the predicted arrival time (t_p) of the target phase. Third, on the observed seismogram, we retain time windows whose seed maxima are contained in the time interval $w_{\text{obs}} = [t_p - 25s - T, t_p + 25s + T]$, where T is the dominant period of the target phase. This choice is based on the fact that, for global S-wave tomography, delay-times have been observed to vary in the interval $[-25s, +25s]$ (Bolton & Masters 2001). Therefore, for a target phase with a dominant period $T \sim 10$ s, we span the time interval $w_{\text{obs}} = [t_p - 35s, t_p + 35s]$, which must only contain the target single phase (e.g. S), or group of phases (e.g. S+sS), for avoiding unwanted phases interference (e.g. ScS with SS). The remaining (observed and synthetic) candidate time windows, at the end of step 5, are shown in Fig. A1(c).

A6 Step 6: selection of the optimal pair of time windows

At this stage, we may still be left with several candidate time windows, around a given target phase (Fig. A1c). To select the optimal pair among them, we first test all combinations of cross-correlation between all the remaining pairs of observed and synthetic waveforms. The aim of this cross-correlation step is to help with the association of a synthetic time window with its best equivalent on the observed seismogram. For each pair of observed and synthetic waveforms, we obtain a cross-correlation maximum ($CC_{\max}$) and a corresponding delay-time ($dt_{\max}$). We keep those pairs of candidate time windows whose $CC_{\max}$ is greater than 80 per cent. Although this choice ensures a strong similarity between observed and synthetic waveforms, it does not always guarantee that they include the same portion of signal. We use the delay-time $dt_{\max}$ for discriminating wrong pairs of candidate time windows among those with $CC_{\max} > 80$ per cent. We then compute the ratio

$$P = \frac{CC_{\max}}{\max(\epsilon, |\mathrm{dt}_{\max}|)} \tag{A7}$$

and select as our optimal pair of observed and synthetic time windows the one with the highest parameter P. We use $\epsilon = 0.1$ s for avoiding to divide by zero, and because our delay times are determined with a precision down to ± 0.1 s. If several observed waveforms present a high degree of similarity with several synthetic waveforms, this choice is a compromise that favours small delay-times, because they are closer to the reference model.

APPENDIX B: TIME RESIDUAL

We aim to prove that the function $F_3(\tau)$ and the cross-correlation function $\gamma_{d,s}(\tau)$ are maximized for the same time residual. We call τ_m^{CC} and $\tau_m^{F_3}$ the time residuals maximizing $\gamma_{d,s}(\tau)$ and $F_3(\tau)$,

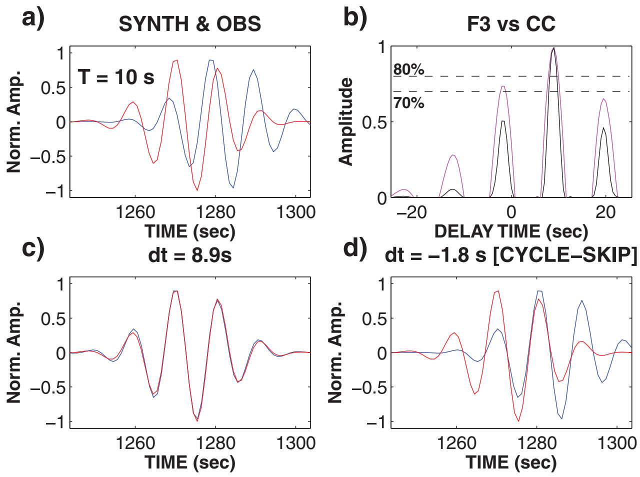

Figure B1. $F_3(\tau)$ versus $\gamma_{d,s}(\tau)$ for a *S*-wave recorded at station ATD (Geoscope network), event 2004 October 15. (a) Observed and synthetic waveforms filtered at 10 s period. (b) $F_3(\tau)$, in black, versus $\gamma_{d,s}(\tau)$, in magenta. Note that $F_3(\tau)$ has secondary maximums lower than for $\gamma_{d,s}(\tau)$. (c) Waveforms after appropriate time-shifting. (d) Waveforms after wrong time-shifting, corresponding to a cycle-skip.

respectively. The recorded signal at the receiver consists of a direct wave arrival $u(t)$ and a scattered wave arrival $\delta u(t)$. Therefore, the observed and synthetic waveforms are, respectively

$$\begin{cases} d(t) = u(t) + \delta u(t) \\ s(t) = u(t). \end{cases} \tag{B1}$$

The autocorrelation of the unperturbed wave u is given by

$$\gamma_{s,s}(\tau) = \int_{-\infty}^{+\infty} u(t)u(t-\tau)\,\mathrm{d}t. \tag{B2}$$

The time residual τ_m^{CC} is defined as maximizing the following cross-correlation function, between the observed signal $(u + \delta u)$ and the unperturbed wave (u)

$$\gamma_{d,s}(\tau) = \int_{-\infty}^{+\infty} [u(t) + \delta u(t)]u(t-\tau)\,\mathrm{d}t \tag{B3}$$

which leads to

$$\gamma_{d,s}(\tau) = \gamma_{s,s}(\tau) + \delta\gamma(\tau) \tag{B4}$$

with

$$\delta\gamma(\tau) = \int_{-\infty}^{+\infty} \delta u(t)u(t-\tau)\,\mathrm{d}t. \tag{B5}$$

For the unperturbed wave, the cross-correlation reaches its maximum at zero lag-time, so

$$\dot{\gamma}_{s,s}(0) = 0, \tag{B6}$$

and for the perturbed wave the maximum is reached for τ_m^{CC}, so

$$\dot{\gamma}_{d,s}\left(\tau_m^{CC}\right) = \dot{\gamma}_{s,s}\left(\tau_m^{CC}\right) + \delta\dot{\gamma}\left(\tau_m^{CC}\right) = 0, \tag{B7}$$

where the dot denotes the time differentiation. Developing $\dot{\gamma}$ to first order, we find (e.g. Marquering *et al.* 1999):

$$\dot{\gamma}_{d,s}\left(\tau_m^{CC}\right) = \dot{\gamma}_{s,s}(0) + \ddot{\gamma}_{s,s}(0)\tau_m^{CC} + \delta\dot{\gamma}(0) + O(\delta^2) = 0 \tag{B8}$$

which leads to

$$\tau_m^{CC} = -\frac{\delta\dot{\gamma}(0)}{\ddot{\gamma}_{s,s}(0)}. \tag{B9}$$

In Section 2.2.2, we have defined the function $F_3(\tau)$ as

$$F_3(\tau) = \frac{F_1(\tau) + F_2(\tau)}{2}. \tag{B10}$$

The first quantity $F_1(\tau)$ is given by

$$F_1(\tau) = 1 - \frac{\int_{-\infty}^{+\infty}[d(t) - s(t-\tau)]^2\mathrm{d}t}{\int_{-\infty}^{+\infty} d^2(t)\mathrm{d}t} \tag{B11}$$

which leads to

$$F_1(\tau) = \frac{2\gamma_{d,s}(\tau) - \gamma_{s,s}(0)}{\gamma_{d,d}(0)}. \tag{B12}$$

The second quantity $F_2(\tau)$ is given by

$$F_2(\tau) = \frac{\gamma_{d,s}^2(\tau)}{\gamma_{d,d}(0)\gamma_{s,s}(0)} \quad \text{if } A_1(\tau) < A_2(\tau). \tag{B13}$$

The maximum of the functions $F_1(\tau)$, $F_2(\tau)$ and $F_3(\tau)$ are reached for $\tau_m^{F_1}$, $\tau_m^{F_2}$ and $\tau_m^{F_3}$, respectively, such as

$$\dot{F}_1\left(\tau_m^{F_1}\right) = 0, \quad \dot{F}_2\left(\tau_m^{F_2}\right) = 0, \quad \dot{F}_3\left(\tau_m^{F_3}\right) = 0. \tag{B14}$$

Note that if $A_1(\tau) > A_2(\tau)$ (Section 2.2.2), we should analyse the maximum of $1/F_2(\tau)$. This maximum will be reached for the same time residual $\tau_m^{F_2}$, as

$$\frac{d}{d\tau}\frac{1}{F_2(\tau)} = -\frac{1}{F_2(\tau)^2}\dot{F}_2(\tau) = 0 \Rightarrow \tau = \tau_m^{F_2}. \tag{B15}$$

We then have

$$\begin{cases} \dot{F}_1\left(\tau_m^{F_1}\right) = \frac{2\dot{\gamma}_{d,s}(\tau_m^{F_1})}{\gamma_{d,d}(0)} = 0, \\ \dot{F}_2\left(\tau_m^{F_2}\right) = \frac{2\gamma_{d,s}(\tau_m^{F_2})\dot{\gamma}_{d,s}(\tau_m^{F_2})}{\gamma_{d,d}(0)\gamma_{s,s}(0)} = 0, \end{cases} \tag{B16}$$

which leads to

$$\begin{cases} \dot{F}_1\left(\tau_m^{F_1}\right) = 0 \Rightarrow \dot{\gamma}_{d,s}\left(\tau_m^{F_1}\right) = 0, \\ \{\dot{F}_2\left(\tau_m^{F_2}\right) = 0 \text{ and } \gamma_{d,s}\left(\tau_m^{F_2}\right) > 0\} \Rightarrow \dot{\gamma}_{d,s}\left(\tau_m^{F_2}\right) = 0. \end{cases} \tag{B17}$$

As previously, by developing $\dot{\gamma}$ to first order, we find

$$\tau_m^{F_1} = \tau_m^{F_2} = \tau_m^{F_3} = -\frac{\delta\dot{\gamma}(0)}{\ddot{\gamma}_{s,s}(0)}. \tag{B18}$$

Finally, the equality $\tau_m^{CC} = \tau_m^{F_3}$ is verified.