



**HAL**  
open science

# Global seismic attenuation imaging using full-waveform inversion: a comparative assessment of different choices of misfit functionals

Haydar Karaoğlu, Barbara Romanowicz

## ► To cite this version:

Haydar Karaoğlu, Barbara Romanowicz. Global seismic attenuation imaging using full-waveform inversion: a comparative assessment of different choices of misfit functionals. *Geophysical Journal International*, 2018, 212, pp.807-826. 10.1093/gji/ggx442 . insu-03589378

**HAL Id: insu-03589378**

**<https://insu.hal.science/insu-03589378>**

Submitted on 25 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Global seismic attenuation imaging using full-waveform inversion: a comparative assessment of different choices of misfit functionals

Haydar Karaoğlu<sup>1,\*</sup> and Barbara Romanowicz<sup>1,2,3</sup>

<sup>1</sup>*Institut de Physique du Globe de Paris, 1 Rue Jussieu, F-752382 Paris Cedex 05, France. E-mail: [karaoglu@ipgp.fr](mailto:karaoglu@ipgp.fr)*

<sup>2</sup>*Berkeley Seismological Laboratory, McCone Hall, University of California, Berkeley, CA 94720, USA*

<sup>3</sup>*Collège de France, 11 Place Marcelin Berthelot, F-75005 Paris, France*

Accepted 2017 October 13. Received 2017 October 4; in original form 2017 March 17

## SUMMARY

We present the results of synthetic tests that aim at evaluating the relative performance of three different definitions of misfit functionals in the context of 3-D imaging of shear wave attenuation in the earth's upper mantle at the global scale, using long-period full-waveform data. The synthetic tests are conducted with simple hypothetical upper-mantle models that contain  $Q_\mu$  anomalies centred at different depths and locations, with or without additional seismic velocity anomalies. To build synthetic waveform data sets, we performed simulations of 50 events in the hypothetical (target) models, using the spectral element method, filtered in the period range 60–400 s. The selected events are chosen among 273 events used in the development of radially anisotropic model SEMUCB-WM1 and recorded at 495 stations worldwide. The synthetic Z-component waveforms correspond to paths and time intervals (fundamental mode and overtone Rayleigh waves) that exist in the real waveform data set. The inversions for shear attenuation structure are carried out using a Gauss–Newton optimization scheme in which the gradient and Hessian are computed using normal mode perturbation theory. The three different misfit functionals considered are based on time domain waveform (WF) and waveform envelope (E-WF) differences, as well as spectral amplitude ratios (SA), between observed and predicted waveforms. We evaluate the performance of the three misfit functional definitions in the presence of seismic noise and unresolved S-wave velocity heterogeneity and discuss the relative importance of physical dispersion effects due to 3-D  $Q_\mu$  structure. We observed that the performance of WF is poorer than the other two misfit functionals in recovering attenuation structure, unless anelastic dispersion effects are taken into account in the calculation of partial derivatives. WF also turns out to be more sensitive to seismic noise than E-WF and SA. Overall, SA performs best for attenuation imaging. Our tests show that it is important to account for 3-D elastic effects (focusing) before inverting for  $Q_\mu$ . Additionally, we show that including high signal-to-noise ratio overtone wave packets is necessary to resolve  $Q_\mu$  structure at depths greater than 250 km.

**Key words:** Seismic attenuation; Seismic tomography.

## 1 INTRODUCTION

Intrinsic attenuation is a material property that strongly depends on temperature, water content and partial melt (e.g. Karato 1993, 2003; Jackson *et al.* 2004; Dalton & Faul 2010), therefore, mapping it can advance our understanding of mantle structure and dynamics significantly. However, this task is a challenging one, as it requires

addressing uncertainties related to the seismic wave amplitude measurements.

The average (1-D) intrinsic attenuation structure as a function of depth is mostly constrained by normal mode data (e.g. Dziewonski & Anderson 1981; Widmer *et al.* 1991; Durek & Ekström 1996; Cammarano & Romanowicz 2008). While this structure is not known in as great detail as for elastic velocities, there is consensus on the presence of a high attenuation zone in the upper mantle, that more-or-less coincides with the well-established seismic low-velocity zone, and the average shear quality factors ( $Q_\mu$ ) in the upper- and lower-mantle are relatively well constrained.

\*Now at: Department of Earth Sciences, Universiteit Utrecht, Utrecht, the Netherlands. E-mail: [h.karaoglu@uu.nl](mailto:h.karaoglu@uu.nl)

Lateral variations in attenuation in the upper mantle have been mapped with increased resolution over the last 20 years, (e.g. Romanowicz 1990; Durek *et al.* 1993; Romanowicz 1995; Selby & Woodhouse 2000; Gung & Romanowicz 2004; Dalton & Ekström 2006; Ma *et al.* 2016). However, the proposed models lag behind the global seismic velocity models both in resolution and agreement. While some recent models, built by following the same procedure as proposed by Dalton & Ekström (2006) with varying data sets and approximate forward modelling techniques (Dalton & Ekström 2006; Bao *et al.* 2016; Ma *et al.* 2016), show improved agreement between each other (Dalton *et al.* 2017), the agreement with models built using different procedures is still poor even at the longest wavelengths (e.g. Adenis *et al.* 2017). This is primarily due to the difficulty of separating elastic and anelastic effects in the measurement of seismic wave amplitudes.

The anelastic nature of earth materials manifests itself through amplitude decay (attenuation) and velocity dispersion of seismic waves travelling through them. The main challenges with isolating the attenuation from seismic wave amplitudes are that: (i) Amplitudes are sensitive to second transverse gradients in the 3-D elastic structure (e.g. Woodhouse & Wong 1986; Romanowicz 1987), which causes de-focusing (local de-amplification), and, scattering at higher frequencies, (ii) the amplitude response of seismographs is not as accurately known as their phase response, (iii) the earthquake source mechanisms and seismic moments are not perfectly known, introducing possible biases in the amplitudes.

Attenuation imaging studies can be grouped into three categories based on the period range they consider. The first group uses short-period body waves (e.g. Bhattacharyya *et al.* 1993, 1996; Matheny & Nowack 1995; Bhattacharyya *et al.* 1996; Warren & Shearer 2002; Lawrence & Wyssession 2006; Hwang *et al.* 2011) and relies on time or frequency domain differential waveform analysis (e.g. Kanamori 1967; Teng 1968), which has the advantage of suppressing uncertainties related to the sources and receivers. They also apply stacking and averaging to address the challenges listed above.

The second and more common approach is to use intermediate- and long-period surface waves. Some of the corresponding studies try to minimize focusing effects by combining data from consecutive Rayleigh wave trains and overcome source and receiver uncertainties through careful data selection (e.g. Romanowicz 1990, 1994, 1995; Durek *et al.* 1993). Others ignore focusing after a careful data selection by reasoning that its effect is negligible for long wavelength structures (e.g. Selby & Woodhouse 2000; Gung & Romanowicz 2004). Some others try to account for focusing using asymptotic methods (e.g. Dalton & Ekström 2006; Dalton *et al.* 2008; Bao *et al.* 2016; Ma *et al.* 2016). The third group of attenuation imaging studies relies on the decay of amplitudes of free oscillations with time (e.g. Roullet 1975; Sailor & Dziewonski 1978; Masters & Gilbert 1983; Romanowicz *et al.* 1987; Roullet *et al.* 1990; Widmer *et al.* 1991). The advantage of these studies is that they are insensitive to source and receiver site uncertainties. However, normal mode studies require records from earthquakes large enough to excite Earth's free oscillations. This limits the available data and, combined with the long wavelength sensitivity of normal modes, renders normal-mode-based methods less suitable for high-resolution 3-D global attenuation imaging.

Using surface and body wave amplitude data, focusing effects need to be accounted for to attain higher resolution attenuation mapping. The focusing effects are usually calculated for an elastic 3-D model, which was built assuming a 1-D attenuation model. For these computations, methods based on linear approximations (e.g. Woodhouse & Wong 1986; Park 1987; Romanowicz 1987; Zhou *et al.* 2004) are generally used. For example, Dalton *et al.* (2014)

and Bao *et al.* (2016) show a comparison between the exact ray theory and finite frequency theory used for attenuation imaging.

The state-of-the-art for computing elastic effects at the global scale is to employ the spectral-element method (SEM) (Komatitsch & Vilotte 1998) for seismic wavefield computations. Although computationally more expensive than conventional methods, SEM provides higher accuracy in estimating focusing and scattering effects. Recently, Zhu *et al.* (2013) applied a waveform inversion methodology based on SEM and adjoint kernels to image upper mantle attenuation structure in Europe.

A full-waveform inversion methodology based on using normal mode perturbation theory both for the forward and inverse modelling steps was introduced by Li & Romanowicz (1995). The approach was first applied to elastic shear wave imaging (Li & Romanowicz 1996; Mégnin & Romanowicz 2000; Gung *et al.* 2003; Panning & Romanowicz 2006) and later extended to shear attenuation imaging of the upper mantle (Romanowicz & Gung 2002; Gung & Romanowicz 2004). More recently, as accurate numerical wavefield computations have become accessible, SEM was adopted for the forward modelling of the wavefield. Employing the SEM-based approach, long-period surface wave and overtone waveforms were used to build global shear velocity models for the upper mantle (Lekić & Romanowicz 2011; French *et al.* 2013). Later, French & Romanowicz (2014) added body waveforms to extend the model to the whole-mantle. In these studies, a smoothed version of the 1-D attenuation model QL6 (Durek & Ekström 1996) was fixed throughout the iterative inversion process. We are now in a position to extend this method to global attenuation imaging.

The selection of a suitable misfit functional, that quantifies the differences between the data and the synthetic seismograms, is a critical step in any inverse modelling problem. For attenuation imaging, it is desirable to work with the misfit functionals that prioritize the attenuation fingerprint in the seismic waves. Most global attenuation studies to date rely on the analysis of amplitudes in the spectral domain. These analyses are based on measuring the decay rate of normal mode peaks (e.g. Roullet & Clévéde 2000) or minimization of the misfit between synthetic and observed spectral amplitudes of individual surface wave trains (e.g. Dalton & Ekström 2006; Zhu *et al.* 2013; Ma *et al.* 2016) or body wave differential amplitudes (e.g. Bhattacharyya *et al.* 1996; Warren & Shearer 2002). In contrast, Gung & Romanowicz (2004) chose to minimize the individual waveform differences in the time domain, after careful selection of data to verify close enough phase alignment. In addition to these, misfit functionals that rely on instantaneous phase matching through Hilbert transform of seismograms have also been suggested for attenuation imaging and used for some regional studies (e.g. Tonn 1991; Matheny & Nowack 1995). More recently, Fichtner *et al.* (2008) suggested misfit functionals based on time- and frequency-dependent phase and envelope misfit of time-frequency transforms of seismograms allowing a complete analysis of the seismograms for imaging. Additionally, Bozdağ *et al.* (2011) suggested the use of misfit functionals that combine the instantaneous phase and envelope misfit in the time domain. These recent studies advocate the use of envelope misfit for attenuation imaging.

This study focuses on the selection of a proper misfit functional for attenuation imaging within the scope of a full-waveform inversion methodology that relies on the comparison of individual energy wave packets. In this paper, we carry out a comparative assessment of three misfit functionals—time-domain waveform comparison (as was employed by Gung & Romanowicz (2004)), time-domain envelope and spectral amplitude ratios, using our full-waveform inversion method for attenuation imaging. We have evaluated these approaches through synthetic tests, targeting to recover heterogeneous

hypothetical models. In these tests, our data set consists of SEM-computed seismograms in the hypothetical target models, which we will refer to as ‘synthetic data’ in what follows. For the seismograms computed in the starting models or in models obtained by inversion of the synthetic data, on the other hand, we will use the expressions ‘synthetic seismograms’ or ‘synthetic waveforms’.

In what follows, we first introduce the waveform inversion method and the misfit functionals to be tested. Second, we present synthetic tests, beginning with a test in a radially symmetric (1-D) model, in order to evaluate the accuracy of the methodology. Then, we present tests based on a synthetic data set computed in simple 3-D heterogeneous attenuation models with and without noise, and in the presence or absence of unresolved velocity heterogeneities. In all cases presented, we invert only for 3-D  $Q_\mu$  structure, whether or not the 3-D elastic model is known. In particular, we test the effect of unknown elastic structure on the retrieved 3-D  $Q_\mu$  structure.

## 2 METHODOLOGY

We utilize a probabilistic iterative inversion scheme, as originally proposed by Tarantola (1984) and apply it in the context of a hybrid full-waveform inversion approach, which combines the accuracy of seismic wavefield computations using SEM with the efficiency of Nonlinear Asymptotic Coupling Theory (NACT) used for partial derivative computations (Li & Romanowicz 1995).

At each iteration, we minimize the misfit functional given by:

$$\Phi(\mathbf{u}, \mathbf{m}) = \frac{1}{2} (\Psi(\mathbf{u}, \mathbf{d}) \mathbf{C}_D^{-1} \Psi(\mathbf{u}, \mathbf{d}) + \alpha (\mathbf{m}_k - \mathbf{m}_0) \mathbf{C}_M^{-1} (\mathbf{m}_k - \mathbf{m}_0)) \quad (1)$$

where  $\mathbf{C}_D$  is the data covariance matrix that weights the data to account for the non-uniform distribution of ray paths and balances contributions from high- and low-amplitude wave packets (Li & Romanowicz 1996).  $\mathbf{C}_M$  is the covariance matrix in model space that addresses the resolution level and is defined through the introduction of correlation lengths (Lekić & Romanowicz 2011),  $\alpha$  is a regularization parameter that is a scaling factor applied to the a priori variance of the model, and  $\mathbf{m}_k$  and  $\mathbf{m}_0$  are the  $k$ th iteration and reference models respectively.  $\Psi(\mathbf{u}, \mathbf{d})$  is the functional quantifying the difference between the synthetics ( $\mathbf{u}$ ) and data ( $\mathbf{d}$ ) that we will investigate further. During the inversion process, the model space is updated iteratively through small perturbations ( $\delta m_k$ ) as follows:

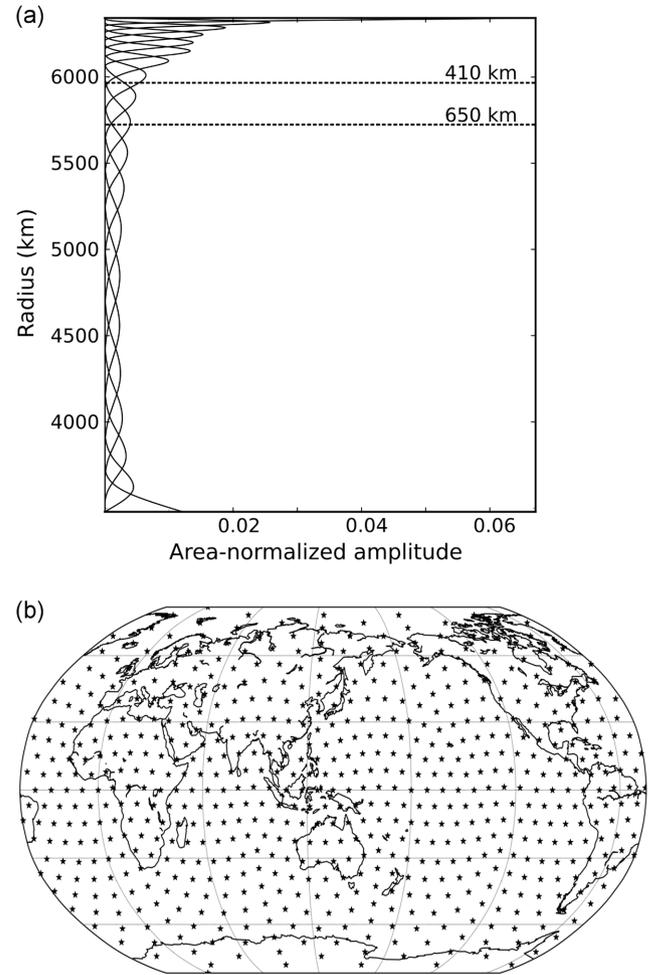
$$\delta m_k = -(\mathbf{G}_k^T \mathbf{C}_D^{-1} \mathbf{G}_k + \alpha \mathbf{C}_M^{-1})^{-1} \times (\mathbf{G}_k^T \Psi(\mathbf{u}_k, \mathbf{d}) + \alpha \mathbf{C}_M^{-1} (\mathbf{m}_k - \mathbf{m}_0)) \quad (2)$$

$$\mathbf{G}_k = \partial \Psi / \partial \mathbf{m}_k \quad (3)$$

where the subscript  $k$  denotes the iteration number. We update the partial derivative matrix  $\mathbf{G}_k$ , the model  $\mathbf{m}_k$  and the synthetics  $\mathbf{u}_k$  (waveforms or spectral amplitudes) computed for the current model iteratively.

### 2.1 Model parametrization

In recent transversely isotropic seismic velocity models built in our group (Lekić & Romanowicz 2011; French & Romanowicz 2014) and in some other studies (e.g. Auer *et al.* 2014; Moulik & Ekström 2014; Chang *et al.* 2015), the physical model is parametrized in terms of Voigt average isotropic  $S$ -wave velocity ( $V_{Siso}$ ) and radial anisotropy parameter ( $\xi = V_{SH}^2 / V_{SV}^2$ ). A transverse isotropic medium requires five independent elastic moduli. These are reduced to two by introducing scaling factors that relate the Voigt average



**Figure 1.** Model space discretization with splines. (a) Radially, we employ 20 b-splines distributed from CMB to the shallowest Moho depth of 30 km as in model SEMUCB-WM1 (French & Romanowicz 2014). (b) Laterally, spherical splines are defined over quasi-equidistant nodes. For the synthetic tests presented in this paper, we employed a level 4 spherical spline parametrization (Wang & Dahlen 1995), with node spacing of  $\sim 9^\circ$ .

$P$ -wave velocity and density perturbations to the  $S$ -wave velocity perturbations, and perturbations in anisotropy parameters  $\phi$  and  $\gamma$  to the perturbation in radial anisotropy parameter  $\xi$ . Further details on this approach and the underlying physical background can be found in appendix A of Panning & Romanowicz (2006).

In our inversion scheme, the model space is parametrized radially in b-splines ( $B_i(r)$ ) and laterally in spherical splines ( $H_j(\phi, \theta)$ ) (e.g. Wang & Dahlen 1995) as expressed by

$$m(\phi, \theta, r) = \sum_{i,j} B_i(r) H_j(\phi, \theta) m_{ij}. \quad (4)$$

Eq. (4) defines a continuous model represented by a set of spline coefficients ( $m_{ij}$ ).

As shown in Fig. 1, we use 20 b-splines with knots distributed from the core–mantle boundary to the shallowest Moho depth of 30 km, as in model SEMUCB-WM1 (French & Romanowicz 2014). The radii of the knots are 3480, 3600, 3775, 4000, 4275, 4550, 4850, 5150, 5375, 5575, 5750, 5900, 6050, 6100, 6150, 6200, 6250, 6300, 6346 and 6361 km. The knots become denser closer to the surface, as relatively higher radial resolution is expected there. Laterally, we use a grid with nodes located approximately  $9^\circ$  apart. This is fine enough to recover heterogeneities of size  $> \sim 2700$  km, which is

more-or-less equivalent to the distance covered by three neighbouring spherical spline nodes along a given direction.

## 2.2 Workflow

Our workflow consists of four major steps:

(1) *Forward modelling*. The synthetic seismograms are computed using the global earthquake simulator, developed by Capdeville *et al.* (2003) (hereinafter referred to as CSEM). CSEM couples computationally expensive SEM in the heterogeneous mantle with normal mode computations in the core through a Dirichlet-to-Neumann operator.

(2) *Wave-packet ‘picking’*. We select wave packets that correspond to time windows with significant seismic energy arrivals through a comparison of data and synthetic waveforms, where the latter are computed for an iteratively updated 3-D model. The wave packets are selected automatically through a procedure that ensures high signal-to-noise ratio, and avoidance of cycle-skipping (e.g. Li & Romanowicz 1996).

(3) *Sensitivity kernels*. The normal mode-based sensitivity kernels are computed in the framework of NACT (non-linear asymptotic coupling theory, (Li & Romanowicz 1995)). Improving on the path average approximation (PAVA; Woodhouse & Dziewonski 1984; Romanowicz 1987) that accounts for the coupling of mode multiplets along the same branch, NACT also includes across-branch coupling of multiplets. This leads to 2-D sensitivity kernels that bring out the sensitivity of body waves centred on the ray path, an important consideration also for surface wave overtones (e.g. Mégnin & Romanowicz 1999).

(4) *Inversion*. The predefined misfit functional is minimized through an iterative Gauss–Newton optimization scheme, in which the matrix  $(\mathbf{G}_k^T \mathbf{C}_D^{-1} \mathbf{G}_k + \alpha \mathbf{C}_M^{-1})$  (see eq. 2) is inverted using the ‘scalapack’ package.

The normal mode-based sensitivity kernels are updated at each iteration for the 3-D model. As argued in Lekić & Romanowicz (2011), and because our kernels already capture the effects of heterogeneities well to first order, the accuracy of the kernels is of second order importance compared to the accuracy of the misfit functional computation. Computing an approximate Hessian efficiently allows us to employ a fast converging Gauss–Newton inversion scheme. At the long wavelengths considered, the efficiency of the inversion methodology outweighs the errors in the kernels.

## 2.3 NACT extension to attenuation imaging

The sensitivity kernels are computed using NACT, which is described in detail by Li & Romanowicz (1995), for the elastic case. Here, we briefly explain how we extend it to the anelastic case assuming a frequency independent absorption band  $Q$  model as proposed by Liu *et al.* (1976) and Kanamori & Anderson (1977). While there is evidence for frequency dependence of attenuation in the earth (e.g. Anderson & Hough 1984; Lekić *et al.* 2009), the assumption of constant  $Q$  is sufficient in the frequency band considered here. In this case, the shear and bulk moduli of an anelastic medium can be represented as complex parameters:

$$\mu = \mu_0 \left( 1 + \frac{2}{\pi Q_\mu} \ln \left( \frac{f}{f_0} \right) + \frac{i}{Q_\mu} \right) \quad (5)$$

$$\kappa = \kappa_0 \left( 1 + \frac{2}{\pi Q_\kappa} \ln \left( \frac{f}{f_0} \right) + \frac{i}{Q_\kappa} \right) \quad (6)$$

where  $f_0$  is the reference frequency at which the shear and bulk moduli are denoted as  $\mu_0$  and  $\kappa_0$  respectively. Modelled as independent of the frequency,  $Q_\mu$  is the shear quality factor and  $Q_\kappa$  is the bulk quality factor.

The resulting expressions for synthetic seismograms and the partial derivative computations using normal mode perturbation theory are given in appendix A, where it is presented that  $Q_\mu$  and  $Q_\kappa$  heterogeneities lead to perturbations not only in the multiplet decay rate but also in the frequency. The latter gives rise to physical/anelastic dispersion, thus phase anomalies. With the absorption band model, the influence of the physical dispersion on the multiplets depends on the definition of the reference frequency ( $f_0$ ). Traditionally, this frequency is set to 1 Hz as originally recommended by Kanamori & Anderson (1977). The seismic velocity models such as PREM of Dziewonski & Anderson (1981), AK135 of Kennett *et al.* (1995) and SEMUCB-WM1 of French & Romanowicz (2014) were all built with reference frequency set to 1 Hz. Choosing different values of  $f_0$  requires significant redefinition of  $Q_\mu$  profiles used in previously built anelastic models (Oki *et al.* 2004). In accordance with the traditional approach, in this study, we set  $f_0$  to 1 Hz.

It is common practice to neglect the bulk attenuation and consider  $Q_\mu^{-1}$  as dominant in the mantle. In accordance with this, we neglect  $Q_\kappa^{-1}$  in our models and inversions, reducing our physical model space to three parameters:  $V_{Siso}$ ,  $\xi$  and  $Q_\mu^{-1}$ .

## 2.4 Misfit functionals

Anelastic tomography requires isolating the attenuation fingerprint in seismic records as mentioned earlier. To that end, the definition of the misfit functional, which defines the type of difference between the data ( $d(t)$ ) and synthetic waveforms ( $u(t)$ ), is critical. In this section, we present the three misfit functionals tested for attenuation imaging, defined as:

$$\Psi(u, d) = u(t) - d(t) \quad (7)$$

$$\begin{aligned} \Psi(u, d) &= env(u) - env(d) \\ &= \sqrt{u(t)^2 + \mathcal{H}(u(t))^2} - \sqrt{d(t)^2 + \mathcal{H}(d(t))^2} \end{aligned} \quad (8)$$

$$\Psi(u, d) = \log(|U(\omega)|/|D(\omega)|). \quad (9)$$

Eq. (7) defines what we call the ‘waveform misfit functional’ (WF). WF is a discrete function of time and computed at each time step used in sampling the waveforms within the selected time-windows. In this form, the misfit is very sensitive to the phase of waveforms, and therefore to elastic heterogeneity ( $V_S$ ,  $\xi$ ). Eq. (8) defines the envelope misfit functional (E-WF) which is also defined in the time domain, and requires the computation of the Hilbert transform ( $\mathcal{H}(t)$ ) and the amplitude of the analytical function for the original waveform. This prioritizes the amplitude and suppresses the phase differences at the expense of losing some of the high frequency content in the waveforms. Finally, eq. (9) defines the misfit functional based on the spectral amplitude ratio (SA). SA is defined in the frequency domain at a set of discrete frequencies. As the phase is naturally separated from the amplitude in the spectral domain, we expect this third misfit functional to be less sensitive to elastic effects.

For the three misfit functionals considered, the partial derivative matrix ( $\mathbf{G}_k$ ) is built iteratively for the model  $\mathbf{m}_k$ , in accordance with the definition of  $\Psi(u, d)$ . For WF, it simply consists in the partial derivatives of time-domain synthetic waveforms with respect to

model space parameters. For E-WF, the  $\mathbf{G}$  matrix (we drop the iteration number  $k$  for simplicity) takes the form

$$\mathbf{G} = \frac{1}{\text{env}(u(t))} \left( u(t) \frac{\partial u(t)}{\partial m} + \mathcal{H}(u(t)) \frac{\partial \mathcal{H}(u(t))}{\partial m} \right) \quad (10)$$

and for SA:

$$\mathbf{G} = \frac{1}{|U(\omega)|^2} \left( \Re(U(\omega)) \frac{\partial \Re(U(\omega))}{\partial m} + \Im(U(\omega)) \frac{\partial \Im(U(\omega))}{\partial m} \right) \quad (11)$$

where  $U(\omega)$  represents the Fourier transform of the original waveform, with  $\Re$  and  $\Im$  denoting the real and imaginary parts. In all these computations, we employ NACT for the computation of the partial derivatives ( $\partial u(t)/\partial m$ ,  $\partial \mathcal{H}(u(t))/\partial m$ ,  $\partial |U(\omega)|/\partial m$ ) and SEM for synthetic waveforms/spectral amplitudes ( $u(t)$ ,  $\mathcal{H}(u(t))$ ,  $|U(\omega)|$ ).

Among the misfit functionals we have tested for attenuation imaging, SA is computationally the most expensive one. This comes from working with selected seismic phase time windows (wave-packet picking). In the time domain, this procedure is equivalent to multiplying the full time history ( $f(t)$ ) by a boxcar function ( $w(t)$ ) that is non-zero from the beginning of wave packet ( $t_0$ ) to the end of it ( $t_0 + T_w$ ) as in eqs (12) and (13). In the spectral domain, this procedure corresponds to circular convolution as in eq. (14). The computation of the Fourier transforms and the circular convolution makes SA more expensive:

$$u(t) = f(t)w(t) \quad (12)$$

$$w(t) = \begin{cases} 1, & \text{if } t_0 + T_w \geq t \geq t_0 \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

$$U(\omega) = F(\omega) \otimes W(\omega) \quad (14)$$

$$W(\omega) = \frac{2\sin(\omega T_w/2)}{\omega} \exp(-i\omega t_0). \quad (15)$$

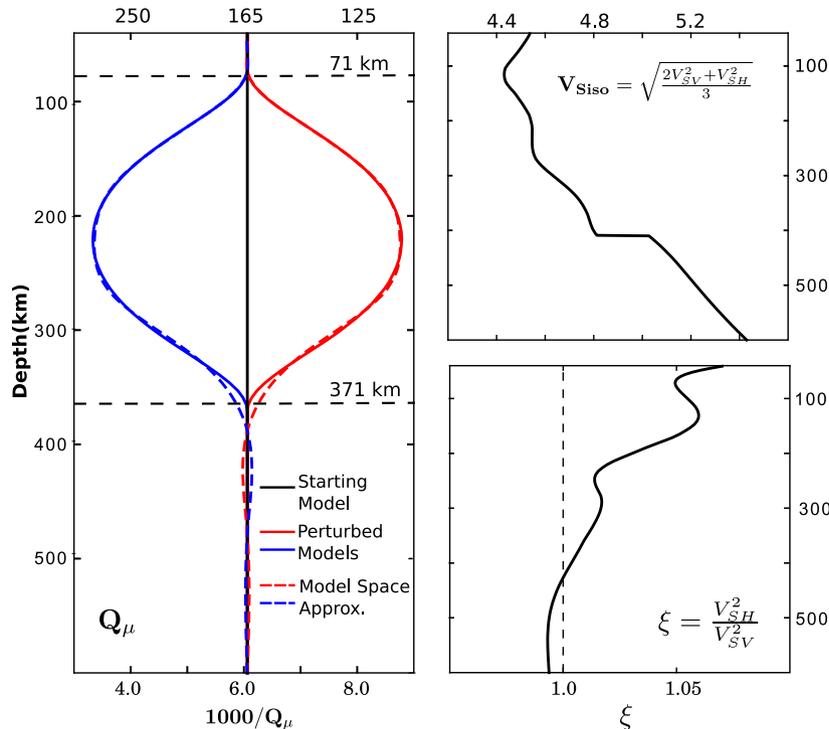
### 3 SYNTHETIC TESTS

In this section, we present the results of several synthetic tests based on simple hypothetical earth models, aimed at evaluating the performance of the three misfit functionals defined in the previous section, with the goal of guiding the development of a suitable methodology for application to real data.

#### 3.1 1-D synthetic tests

First, we test the accuracy of the multiplet frequency and attenuation perturbation computations introduced in Section 2.3, using our model parametrization, for two spherically symmetric (1-D) anelastic models. In these tests, for the elastic model ( $V_S, \xi$ ), we used the 1-D structure of SEMUCB-WM1 without any 3-D perturbations. To build the target shear attenuation models, we considered a 1-D, spherically symmetric  $Q_\mu$  model, where the values are set to 165 in the upper mantle (from surface to the 660 km discontinuity). The lower mantle  $Q_\mu$  model is as in Durek & Ekström (1996). This model is used as starting model in the following tests. To obtain the target model, we added smoothly varying 1-D (spherically symmetric) perturbations to the upper-mantle part of this model. For simplicity, in all the tests, we keep the crustal model and Moho depth fixed.

We consider two hypothetical 1-D models, one with positive and the other with negative  $Q_\mu^{-1}$  perturbations (Fig. 2). As we observed



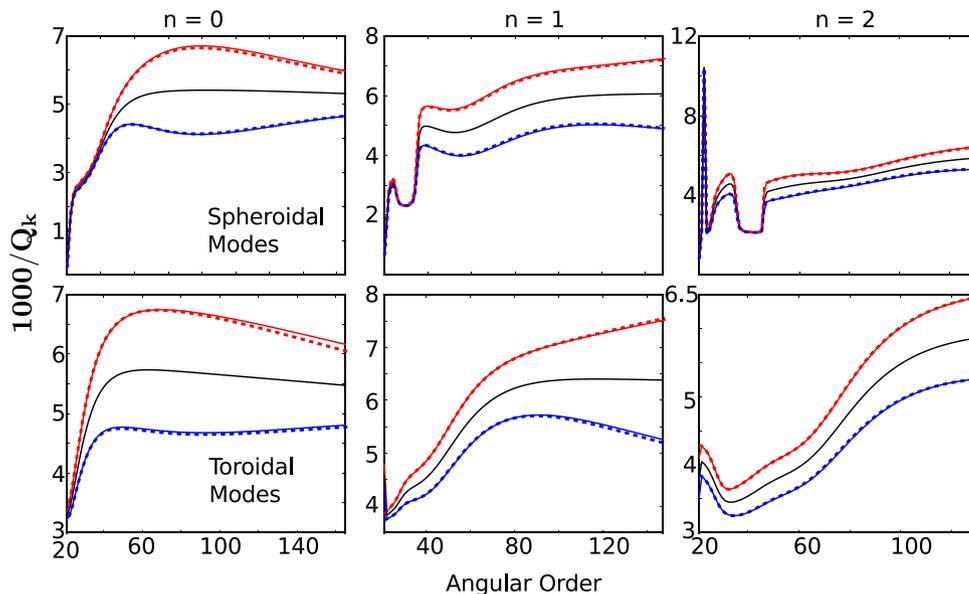
**Figure 2.** 1-D synthetic test reference models with positive (red) and negative (blue) attenuation perturbations introduced on top of a constant  $Q_\mu$  model down to 660 km. The hypothetical models are approximated by the radial b-splines and spherical splines that we use for model space discretization. The global averages of the approximated perturbations are shown by dashed lines. The elastic model ( $V_{Siso}, \xi$ ) is the same as the global-average (spherically symmetric) structure of the SEMUCB-WM1 model (French & Romanowicz 2014).

from our preliminary tests (not shown) that the inversions using our method cannot recover zero-degree  $Q_\mu^{-1}$  anomalies of amplitude larger than 50 per cent in one iteration, we set the peak anomaly amplitudes in this test to 45 per cent so that we can limit the number of our iterations to one.

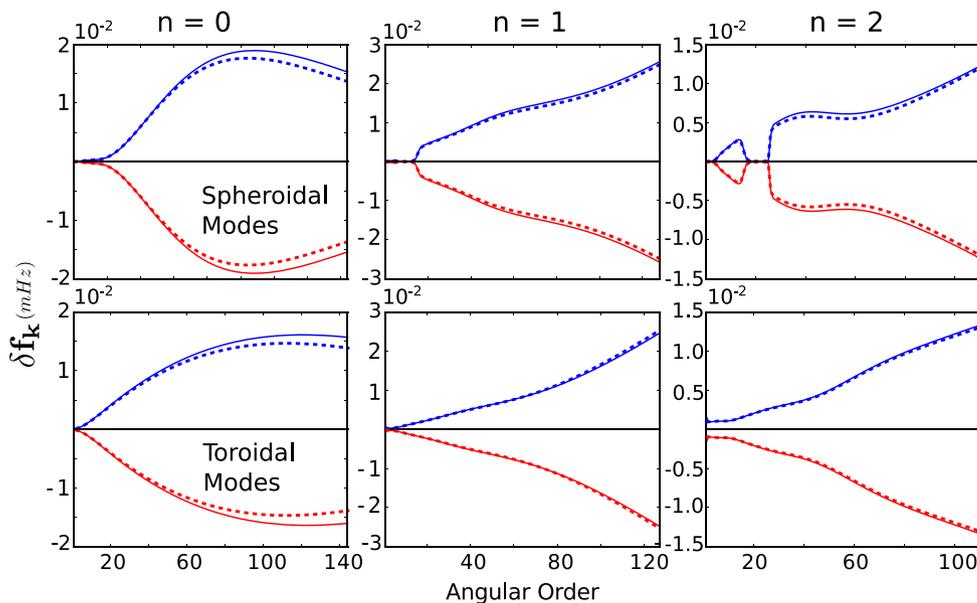
The quality factors ( $Q_k$ ) and frequencies ( $\omega_k$ ) of normal mode multiplet  $k = (l, n)$ , where  $l$  is the angular order and  $n$  the radial order of the multiplet, can be computed exactly for any 1-D model using normal mode theory (e.g. Woodhouse 1988). Having calculated them exactly for the starting and target ( $Q_\mu$  perturbed) models, we next calculated the  $\omega_k$ 's and  $Q_k$ 's in the target model approximately,

using first order perturbation theory applied to the starting model. For this calculation, we expanded the perturbation in the  $Q_\mu^{-1}$  models using our spline model parametrization (Section 2.1).

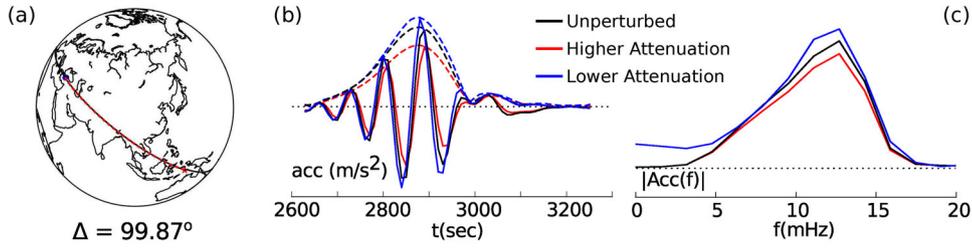
Next, we chose several different paths with lengths between  $60^\circ$ – $120^\circ$  and computed the corresponding  $Q_k$  and  $\omega_k$  perturbations. This is done to test the approximations introduced from using perturbation theory as well as errors stemming from the spline parametrization. The results are presented in Figs 3 and 4. The match between the two computations is almost perfect both for  $Q_k$  and  $\omega_k$  with slight differences due to imperfect  $Q_\mu$  model representation using our spline basis.



**Figure 3.** The quality factors of spheroidal and toroidal normal modes with periods longer than 60 s for the first three mode branches ( $n = 0, 1, 2$ ). The direct  $Q_k$  computations, where  $k$  is the normal mode index, for the unperturbed (black solid) and perturbed models (red and blue solid) are compared with those predicted by normal mode perturbation theory (dashed lines), computed for a low  $Q$  and a high  $Q$  path in the model, as a function of angular order. The colour coding is the same as in Fig. 2.



**Figure 4.** The eigenfrequency perturbations ( $\delta f_k$ ) of the spheroidal and toroidal normal modes with periods longer than 60 s with respect to the starting model for the first three mode branches ( $n = 0, 1, 2$ ). The direct  $f_k$  computations for the unperturbed (black solid) and perturbed models (red and blue solid lines) are compared with those predicted by the normal mode perturbation theory (dashed lines), computed for a low  $Q$  and a high  $Q$  path in the model, as a function of angular order. The colour coding is the same as in Fig. 2.



**Figure 5.** (a) The path of the  $R_1$  wave packet chosen to image the  $Q_\mu$  structure in the 1-D synthetic test. (b) The Z-component acceleration waveform (solid), envelope (dashed) time histories and (c) the corresponding amplitude spectra reflect the effects of the  $Q_\mu$  perturbations. The dotted lines in the time histories and in the spectra correspond to the zero value. The colour coding is the same as in Fig. 2.

We now consider a 1-D  $Q_\mu$  profile recovery test using the PAVA approximation for each of the perturbed models. For this simple problem, we expect the PAVA approximation to work well, even though the perturbation is large, because PAVA includes multiple forward scattering (see appendix A) and can accurately represent a constant perturbation over large distances (in contrast to the Born approximation, e.g. Romanowicz *et al.* 2008; Panning *et al.* 2009). For the inversion, we simulated a single event (2009/10/24, 14:40:43.7,  $M_w$ :7.0, Banda Sea) with a hypocentral depth of 154.6 km and chose a single Z-component record of the first arriving Rayleigh wave train ( $R_1$ ) recorded at station ANTO at a distance of  $\sim 100^\circ$  (Fig. 5). As expected, the model with higher attenuation leads to faster amplitude decay and positive time delay with respect to the reference model, whereas the model with lower attenuation leads to the opposite.

Using the three definitions of misfit functionals presented in Section 2.4, we carried out inversions for the  $Q_\mu$  structure along the chosen path, starting with the unperturbed model, with and without accounting for anelastic dispersion in the computation of the Fréchet derivatives (see appendix A). The inversions are carried out for the upper mantle and the transition zone, including only the top 10 b-splines (30–621 km depth range). For the synthetic waveforms used in the inversions, the sampling rate is 10 s, to ensure a dense enough sampling for the targeted period range of 60–400 s. It is 0.1 mHz for inversions in the frequency domain, as the corresponding frequency range is 2.5–16.6 mHz.

In our inversions, the regularization is defined with correlation lengths. Following French & Romanowicz (2014), in the followings tests, the radial correlation lengths have varying values from 50 km at the shallowest depths to 200 km around 600 km depth. The radial correlation lengths are defined in accordance with the distance between neighbouring radial nodes and following insights from preliminary tests. The list of chosen radial correlation lengths for the targeted 10 b-splines beginning with the bottom one is 200, 150, 100, 50, 50, 50, 50, 50, 50, 50 km. Laterally, we fixed the correlation lengths to 2400 km, which is approximately three times the mean nodal distance laterally (see Section 2.1). The inversions were carried out with decreasing norm damping parameter ( $\alpha$  in eq. 1) until the recovered model showed unrealistically large fluctuations that first appear in depth and next laterally, eventually destabilizing all the model space. The best models recovered in one iteration, using the smallest regularization parameter among the predefined range of values that preserve the stability of the inversions, are presented in Fig. 6 for each case.

WF is quite sensitive to anelastic dispersion, as expected from its sensitivity to the phase difference between data and predicted synthetic seismograms. We cannot recover the depth range of the attenuation perturbations using WF in the absence of anelastic dispersion. E-WF and SA work equally well whether we include anelastic dis-

persion or not. However, for the E-WF, if we ignore the anelastic dispersion, it is necessary to increase the regularization parameter value, illustrating the significance of the phase difference in the time domain, whether we consider waveform or envelope fitting. For the SA inversion, there is no such problem.

High sensitivity to anelastic dispersion of WF makes it the least favourable misfit functional to be employed in attenuation tomography, as expected. The other two approaches are more stable.

### 3.2 3-D synthetic tests

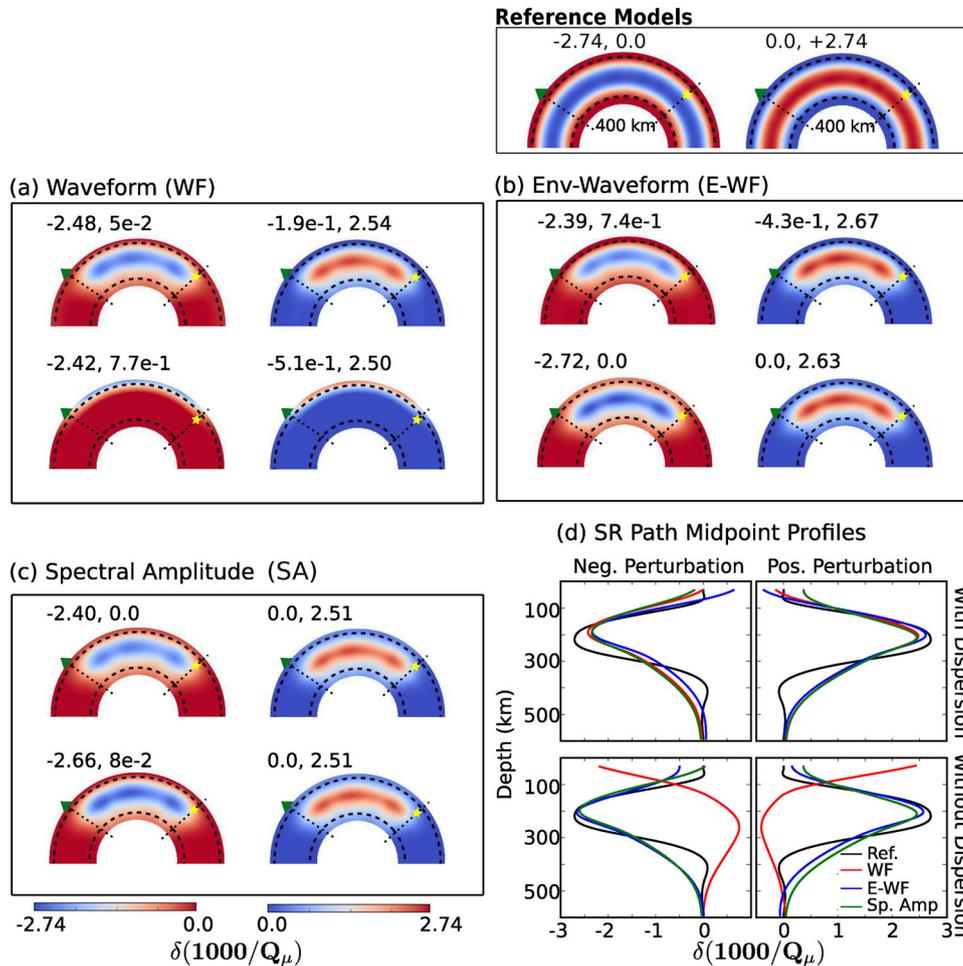
In this section, we carry out a comparative assessment of the misfit functionals in the case where 3-D heterogeneities are present. In all the tests presented in this section, we added 3-D heterogeneities to the unperturbed zero-degree structure used in the previous section. This radially symmetric model consists of the global average  $V_S$  and  $\xi$  profiles of SEMUCB-WM1 and the  $Q_\mu$  profile with values set to 300 from surface to fixed Moho depth of 30 km and 165 from Moho depth to 660 km discontinuity. Below 660 km, we use QL6 of Durek & Ekström (1996). In all the inversions presented below, we chose this radially symmetric model as our starting model. For simplicity, we ignored the Moho depth variations, crustal model and topography.

In the presented tests, we tried to recover shear attenuation anomalies in target models built with either only  $Q_\mu$  or both  $Q_\mu$  and  $V_S$  ellipsoidal heterogeneities added to the radially symmetric model described above. The perturbation amplitudes peak at the centre of the ellipsoids and fade away from it. For a model parameter  $m$ , these perturbations are defined as:

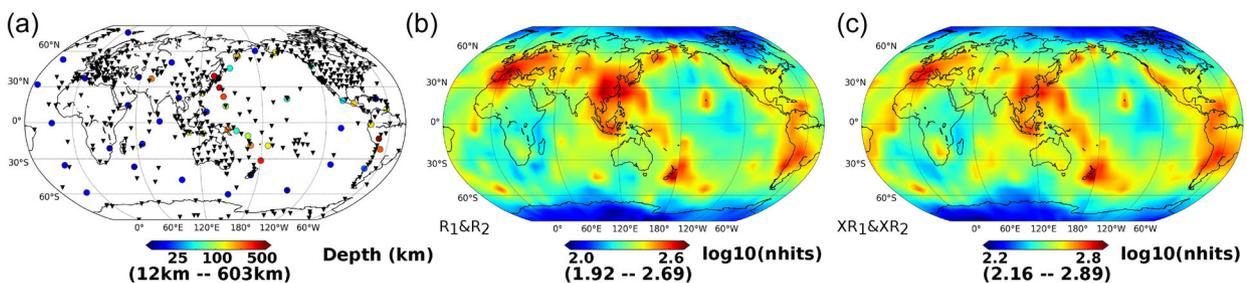
$$\frac{\delta m}{m} = \begin{cases} A_{\text{peak}} \begin{bmatrix} 1.0 + \cos\left(\frac{2\pi(r-r_0)}{h}\right) \\ 1.0 + \cos\left(\frac{2\pi\Delta}{w}\right) \end{bmatrix} \\ 0, \quad \text{if } |r-r_0| > h/2 \\ \Delta > w/2 \end{cases} \quad (16)$$

where  $A_{\text{peak}}$  is the maximum perturbation at the centre of the anomaly located at  $(r_0, \theta_0, \phi_0)$ ,  $h$ ,  $w$  represent its height and width, respectively, and  $\Delta$  is the distance from the center of the anomaly.

We established our synthetic waveform data set using 50 selected events out of 273 events in the event catalogue that was used in developing the SEMUCB-WM1 model. The moment magnitudes ( $M_w$ ) of these 50 events vary from 6 to 7.3 and their depths range from 12 to 603 km. We used CSEM to compute the seismic wavefield in the period range of 60–400 s, which was the period range targeted in building the SEM-based upper-mantle elastic global models SEMum (Lekić & Romanowicz 2011) and SEMum2 (French *et al.* 2013). CSEM includes the frequency independent  $Q_\mu$  approximation based on the standard linear solid model (Liu *et al.* 1976)



**Figure 6.** Comparison of the inversion results for the three approaches (WF, E-WF, SA) along the source–receiver path (SR Path) of the  $R_1$  wave packet. The cross-section views of the reference models are shown in the top-right panel. In each subplot, the first and second rows correspond to the inversions with and without anelastic dispersion respectively (see Appendix A). In each subplot, the tests with negative (positive) perturbations from the reference models are shown on the left(right). The min, max perturbation values are shown at the top of each figure. The perturbation profiles at the midpoint of the source–receiver path are compared in (d) with the reference model profile plotted in black.

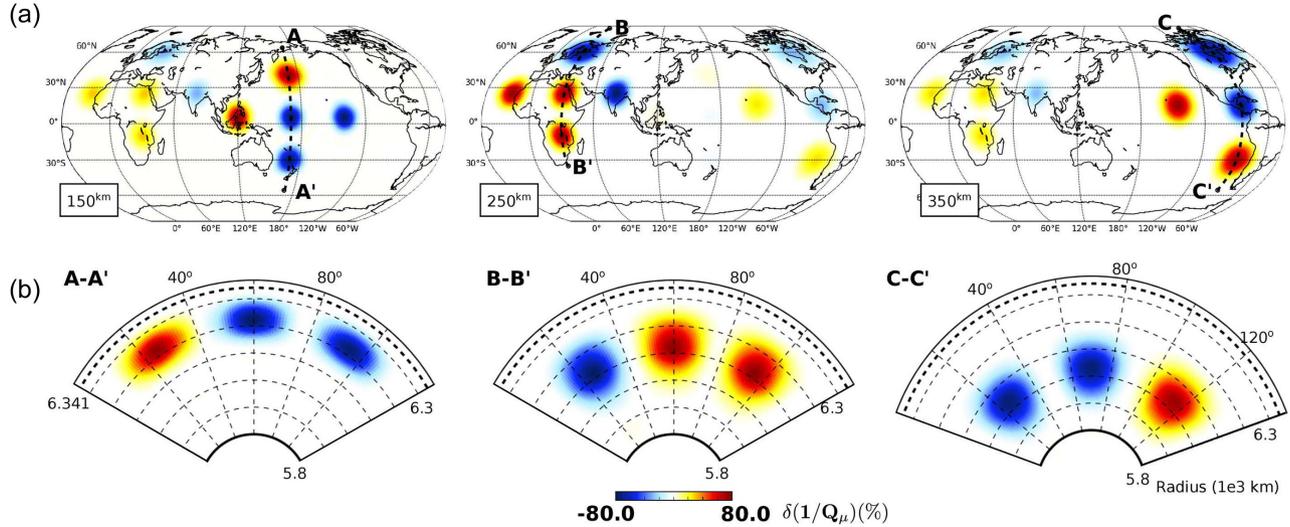


**Figure 7.** (a) Distribution of the selected 50 earthquake epicentres (circles), coloured as a function of their hypocentre depth, and 495 receivers (black inverted triangles). Ray density in  $10 \times 10$  degree cells for Z-component (b) fundamental mode ( $R_1$ ,  $R_2$ ) and (c) overtone ( $XR_1$ ,  $XR_2$ ) Rayleigh wave packets.

coherent with our assumptions (eqs 5 and 6). Here, we restricted the full-waveform inversion to only the first two arriving fundamental and overtone Rayleigh wave phases ( $R_1$ ,  $R_2$ ,  $XR_1$ ,  $XR_2$ ) in Z-component records, and selected the same wave packets as were picked in the construction of SEMUCB-WM1 model. This limits us in terms of the ray coverage, yet provides a realistic basis in assessing the performance of the misfit functionals tested in our synthetic experiments. Fig. 7 shows the distribution of events and receivers together with the ray coverage of the 18 087 (7697 fundamental,

10 390 overtone) wave packets selected. We used the same wave packets for all the inversion results presented below.

We decided on the number of iterations based on our preliminary tests which showed that using our hybrid full-waveform inversion technique, in one iteration, we can recover perturbations with at most 40 per cent amplitude anomalies keeping the inverse problem well-posed. As we will explain further below, the introduced shear attenuation perturbation amplitudes in the 3-D synthetic tests require at least two iterations.



**Figure 8.** Hypothetical model for the 3-D  $Q_\mu$  synthetic test, with attenuation perturbations added to the radially symmetric model with zero-degree  $V_S$  and  $\xi$  profiles of SEMUCB-WM1 and the unperturbed  $Q_\mu$  profile presented in Section 3.1. (a) Lateral distribution of perturbations at three depths (150, 250, 350 km) corresponding to the depths where the anomaly amplitudes peak (80 per cent) and (b) depth cross-sections along the dashed lines marked on the maps.

While the synthetic seismograms for the starting spherically symmetric model (1-D model) are computed by normal mode summation, for the second iteration, they are computed by using CSEM in the heterogeneous 3-D model recovered in the first iteration.

### 3.2.1 3-D $Q_\mu$ test

For the first 3-D test, a hypothetical model was built with no elastic 3-D anomalies and with multiple positive and negative  $Q_\mu$  perturbations with widths ranging from 4000 to 4800 km, heights from 150 to 350 km and peak amplitude 80 per cent of the reference starting model at the corresponding depths ( $\delta \ln(1/Q_\mu)$ ) (see eq. 16). The anomaly amplitudes are chosen in such a way, that they are not too far from the values present in global mantle attenuation models, such as QRLW8 of Gung & Romanowicz (2004) with peak values around 55 per cent, and QRFSI12 of Dalton *et al.* (2008) with peak values of 133 per cent, and also so that we can expect to recover them in a few iterations using our hybrid full-waveform inversion technique.

The dimensions of the heterogeneity are chosen in the specified ranges to lower the computational cost in the inversion step by using a coarse model parametrization. The smaller the dimension of the heterogeneity, the more the computational cost, in addition to a possible requirement for more iterations. We introduced heterogeneities with peaks at three depths (150, 250, 350 km). Fig. 8 shows the lateral distribution and depth cross-sections of  $Q_\mu$  heterogeneities of the target model. Given the ray coverage, with our model parametrization (Fig. 1), these heterogeneities should be recoverable, although we might expect better resolution in the northern hemisphere. This test provides insight on the limitations of the misfit functional choice in a more realistic way than the 1-D test presented above, and allows us to test the NACT implementation when perturbations in the attenuation model are considered. As in the case of 1-D tests, in the inversion, we use a sampling rate of 10 s in the time domain, ensuring a dense enough sampling for the targeted period range of 60–400 s. In the frequency domain, the sampling rate is 0.1 mHz.

In these inversions for the attenuation parameter ( $Q_\mu$ ), we computed partial derivatives using NACT, and we used an a priori

model covariance matrix ( $C_M^{-1}$ ) with correlation lengths specified as 1600 km laterally and varying values from 50 to 200 km radially, in accordance with the distance between neighbouring radial nodes (see Section 3.1 for further detail on radial correlation lengths). The correlation lengths were determined based on preliminary tests and the average distance between the spherical nodes ( $\sim 9^\circ$ ). The inversions were conducted using a range of norm damping parameter values ( $\alpha$  in eq. 1), which were reduced until the inversion became unstable. This was decided qualitatively, observing radial and lateral fluctuations of the recovered anomalies. Below, we present the best models recovered after two iterations using the lowest norm damping parameter that preserves the stability of the inversions, for each of the 3 misfit functionals considered.

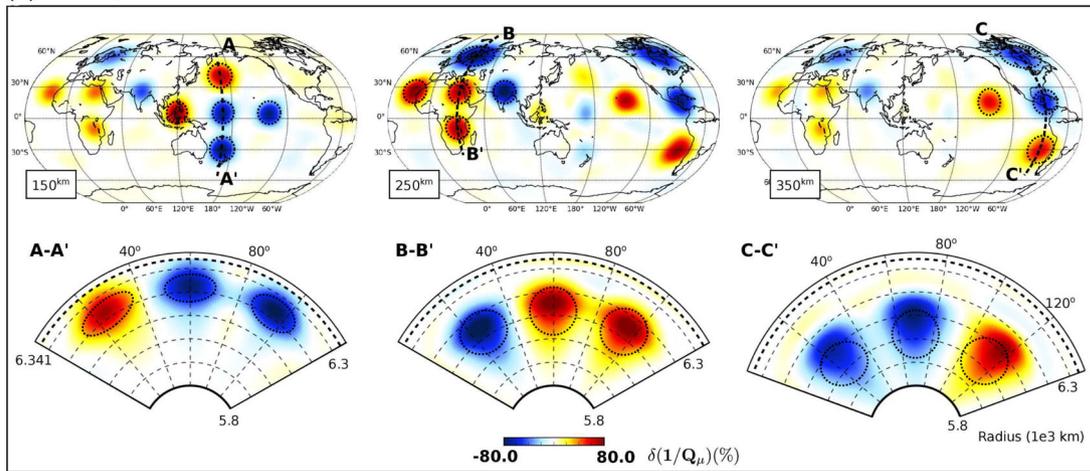
E-WF and SA both performed best as they resulted in almost full recovery of the target model, with a little more lateral and in depth smearing in the case of E-WF, as shown in Fig. 9. Here, anelastic dispersion was ignored in the computation of the partial derivatives (**G** matrix—see Appendix A).

For the WF misfit functional, ignoring the anelastic dispersion in the computation of the partial derivatives leads to a significant loss in depth resolution as shown in Fig. 10. Additionally, a higher regularization parameter value was required for the stability of the inversion, compared to the case with anelastic dispersion. Relatively higher level of background fluctuations in the recovered model and the fact that some heterogeneities (both positive and negative) leak to very shallow depths, even when dispersion corrections are considered in the calculation of the partial derivatives, indicates, again, that this approach is the least favourable of all. The importance of anelastic dispersion when using WF further reflects the fact that this misfit functional emphasizes misfit in the phase part of the waveform. This makes it a suitable misfit functional to image elastic heterogeneity, yet one should be cautious when employing it for attenuation.

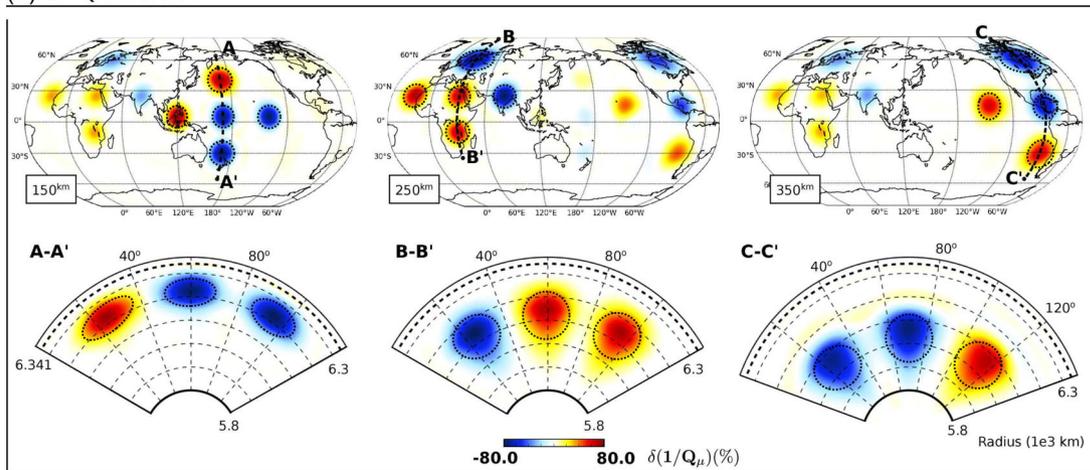
### 3.2.2 The significance of overtones

As a complement to the previous section, we present the models recovered using only the fundamental mode Rayleigh waves. This serves to clarify the importance of including Rayleigh wave

(a) E-WF Q Inversion



(b) SA Q Inversion



**Figure 9.** The recovered models after two iterations using the envelope waveform (a) and acceleration spectral amplitude (b) misfit functionals. The inversions are carried out using the Z-component fundamental mode and overtone Rayleigh wave time windows ( $R_1$ ,  $R_2$ ,  $XR_1$ ,  $XR_2$ ). Top panels: map views of distribution of heterogeneities at 150, 250 and 350 km depth; bottom panels: corresponding depth cross-sections along the A-A', B-B' and C-C' paths indicated on the maps. The locations of the  $Q_\mu$  anomalies in the target model are circled by dashed lines. Lateral and in depth smearing of the recovered perturbations is expected due to limited ray coverage and the small number of iterations.

overtone for attenuation mapping. Fig. 11 presents the recovered models using only the Z-component fundamental mode Rayleigh waves, with E-WF and SA misfit functionals.

Excluding overtone Rayleigh wave data, we significantly lose the  $Q_\mu$  anomalies below  $\sim 300$  km depth, with both SA and E-WF. In the recovered models, the anomalies in the reference hypothetical model that peak at 350 km depth appear with significantly reduced amplitudes and the anomalies that should peak at 250 km are mapped closer to the surface. Including overtone Rayleigh wave data improves the depth resolution of the recovered models and also provides additional constraints on near-surface anomalies, reducing their smearing as a function of depth. Fig. 12 clarifies the contribution of the overtones by comparing the recovered 1-D  $Q_\mu$  profiles at the centre of the anomalies located in the middle of A-A', B-B' and C-C' paths indicated on Figs 9 and 11.

### 3.2.3 3-D $Q_\mu$ test with noise included

Next, we test the E-WF and SA in the presence of seismic noise. In this test, real noise records collected from the stations considered are added to the synthetic waveforms. As presented in Fig. 13, for

the recovered model with E-WF, the effect of added real noise is negligible. This is reassuring, as the wave packets we used in the  $Q_\mu$  inversions are the same as the ones used in building SEMUCB-WM1 model, except that they are calculated for the hypothetical models. These wave packets were picked through a careful procedure that eliminates those with low signal-to-noise ratio (see Section 2.2).

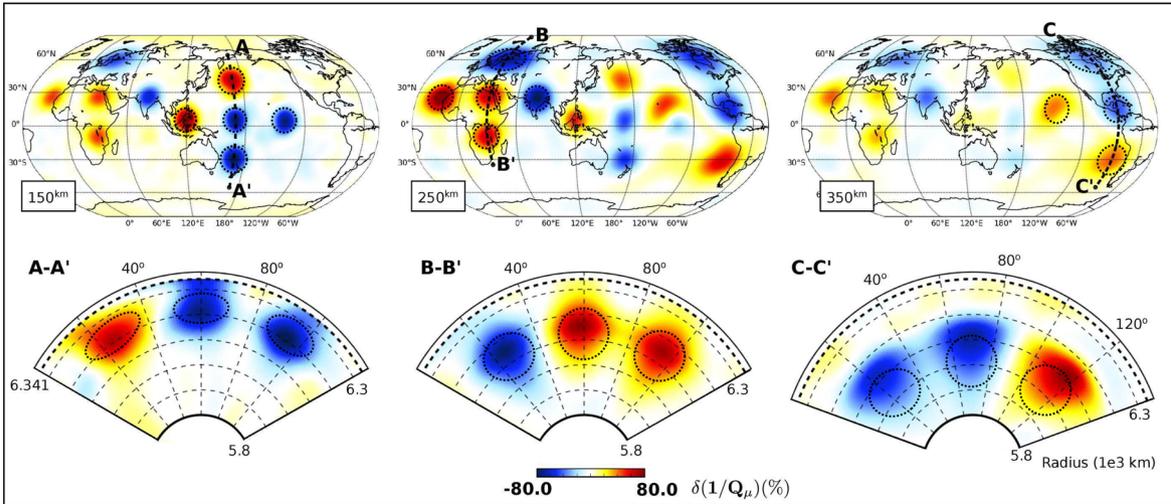
As the recovered model in the presence of noise is very similar when using the SA misfit functional as it is for E-WF (Fig. 9), we do not show it here for brevity.

The test with real noise illustrates the validity of our wave packet picking for the  $Q_\mu$  inversions and the robustness of E-WF and SA misfit functionals, working with carefully selected data in the presence of real noise. To test the E-WF and SA at their limits, we present recovered models with high levels of noise defined through a scaling parameter in appendix B.

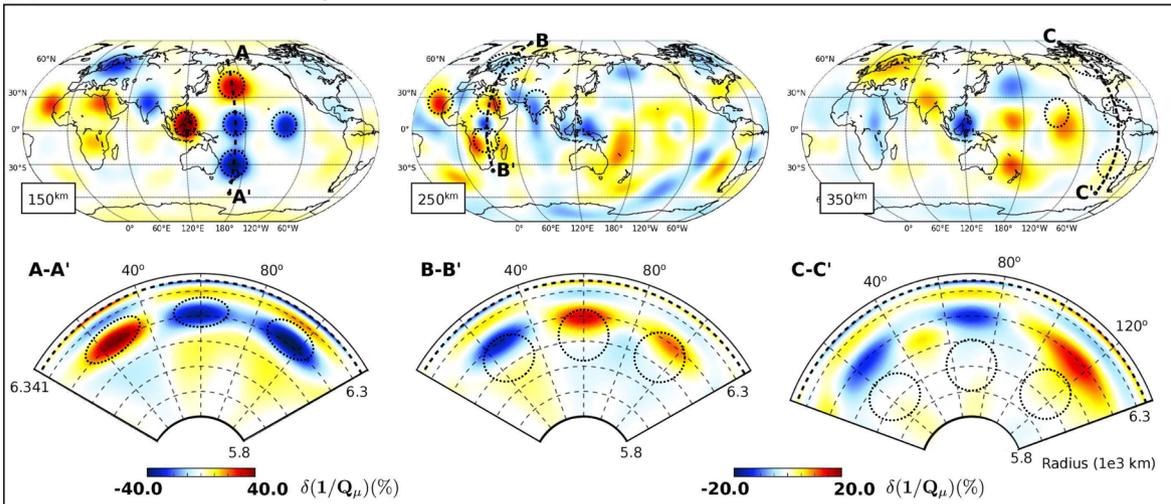
### 3.2.4 Test including 3-D $V_S$ and $Q_\mu$ anomalies

Finally, we evaluate the attenuation imaging performance of the three misfit functionals tested in the presence of unresolved  $V_S$  heterogeneities without seismic noise. By unresolved  $V_S$  anomalies, we

(a) WF Q Inversion - With Anelastic Dispersion



(b) WF Q Inversion - Without Anelastic Dispersion



**Figure 10.** The recovered models after two iterations using the waveform misfit functional with (a) and without (b) the anelastic dispersion correction as mentioned in the methodology section. The locations of the  $Q_\mu$  anomalies in the hypothetical model are circled by dashed lines. Ignoring the anelastic dispersion leads to a significant loss in depth resolution. For this case, we had to use a higher norm damping value for stability. Additionally, the recovered model perturbations are much weaker and they become less pronounced with depth, as indicated in the change in colour scale in (b).

mean that although both  $V_S$  and  $Q_\mu$  3-D heterogeneities are present in the target model, we invert only for  $Q_\mu$ , ignoring the presence of unresolved lateral variations in  $V_S$ . This is an important scenario to test potential leaks from unresolved elastic heterogeneities into the  $Q_\mu$  images.

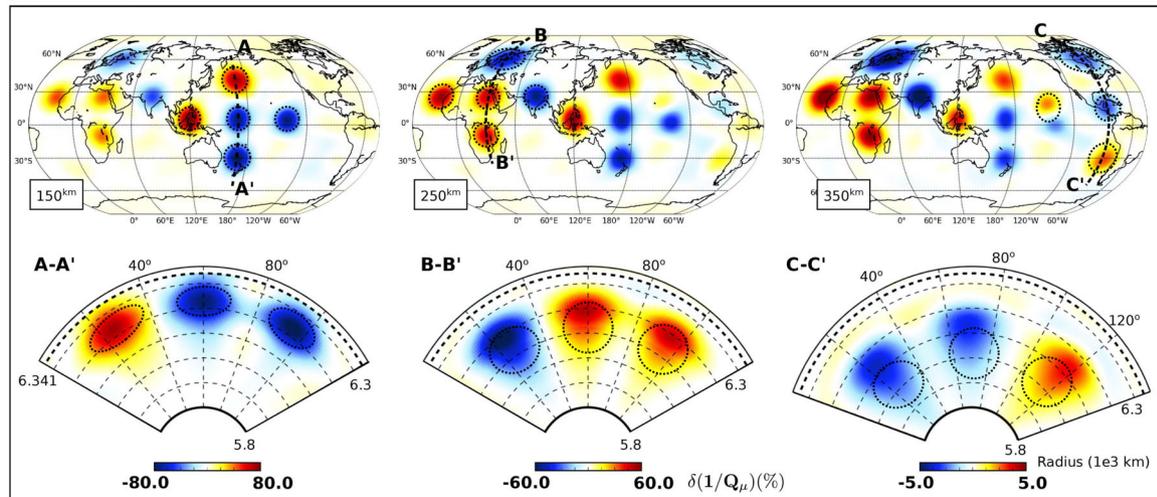
The hypothetical model built for this test with both  $V_S$  and  $Q_\mu$  heterogeneities is presented in Fig. 14. We utilized the same  $Q_\mu$  heterogeneities as in the 3-D  $Q_\mu$  test presented in the previous section with maximum anomaly amplitudes of 80 per cent, but without the deepest anomalies peaking at 350 km depth.  $V_S$  heterogeneities are centred at 200 km and spatially distributed very close to  $Q_\mu$  perturbations and even overlap with them in the western Pacific and south-east Asia, in order to evaluate leakage from one to the other in an extreme scenario. The  $V_S$  heterogeneities are 5000 km in width and 400 km in height (see eq. 16) and 3 per cent ( $\delta \ln(V_S)$ ) in amplitude with respect to the 1-D average model. This is meant to represent underestimated amplitudes of  $V_S$  anomalies recovered from seismic tomography. In the target model, we kept  $\xi$  as in our starting model, which is a radially symmetric model, and we

did not invert for  $\xi$  as we work only with Z-component records of Rayleigh waves. This simplifies the test by reducing the possible interactions between  $V_S$ ,  $\xi$  and  $Q_\mu$ , and to the first order presents the key strengths and weaknesses of the misfit functionals due to the interaction between the elastic and attenuation structure.

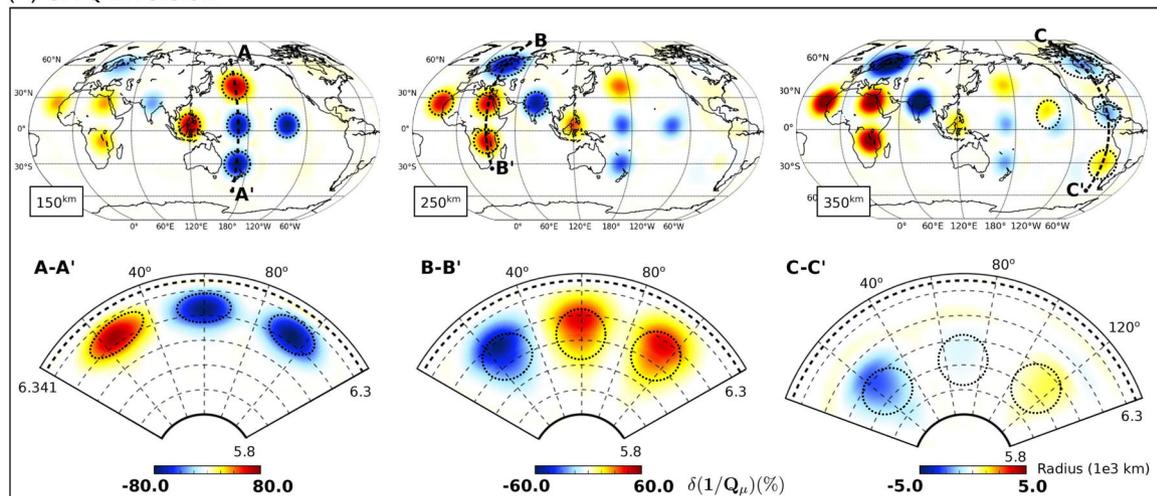
We present a summary of the best recovered models in Fig. 15. Both E-WF and SA are able to recover  $Q_\mu$  heterogeneities to a certain degree with a slight leak from  $V_S$  heterogeneities, as visible both in the lateral and depth cross-sectional views. Where both types of heterogeneities overlap, the recovery of the  $Q_\mu$  anomalies seems to deteriorate with increased lateral and in-depth smearing (see the anomalies located in the middle of B-B' and C-C' paths in Fig. 15).

With WF,  $Q_\mu$  heterogeneities are poorly recovered even with the anelastic dispersion accounted for in the partial derivatives. Additionally a stable inversion with WF requires a relatively high norm-damping parameter value ( $\alpha$  in eq. 1)—leading to the recovery of lower perturbation amplitudes. The recovered model is far worse in the absence of the anelastic dispersion term, since  $V_S$  heterogeneities are mapped into  $Q_\mu$ .

## (a) E-WF Q Inversion



## (b) SA Q Inversion



**Figure 11.** The recovered models after two iterations using the envelope waveform (a) and acceleration spectral amplitude (b) misfit functionals. The inversions are carried out using only the Z-component fundamental mode Rayleigh wave time-windows ( $R_1$ ,  $R_2$ ). Top panels: map views of distribution of heterogeneities at 150, 250 and 350 km depth; bottom panels: corresponding depth cross-sections along the A-A', B-B' and C-C' paths indicated on the maps. The locations of the  $Q_\mu$  anomalies in the target model are circled by dashed lines.

The leakage from  $V_S$  to  $Q_\mu$  is manifested by mapping of high attenuation in fast zones of limited lateral extent and vice versa. It is in line with expectations from focusing effects, as seismic waves passing through slow and fast zones, increase and decrease in amplitude, respectively (e.g. Romanowicz 1995; Dalton & Ekström 2006).

The leakage from  $V_S$  to  $Q_\mu$  heterogeneities is most pronounced in the WF case, especially in the absence of the anelastic dispersion term in the partial derivatives matrix. To further illustrate this last point, we present the A-A' depth cross-sections in Fig. 15 (bottom panel), where  $V_S$  perturbations peak in the target model, and clearly dominate the imaged  $Q_\mu$  structure for WF (with reversed 'colour'). However, these effects are seen even in the case of E-WF and SA, implying that the elastic heterogeneity must be accounted for as best as possible in the forward modelling step (e.g. Romanowicz 1995; Dalton & Ekström 2006; Bao *et al.* 2016).

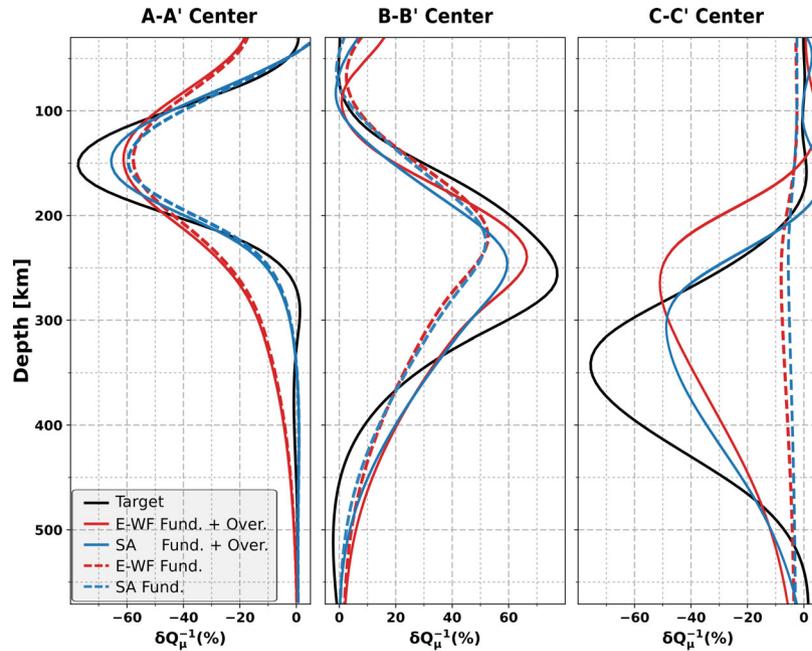
#### 4 DISCUSSION AND CONCLUSIONS

We have carried out a comparative assessment of three misfit functionals based on waveform (WF), waveform envelope (E-WF) and

spectral amplitude (SA) differences, within the framework of a hybrid full-waveform inversion scheme for attenuation imaging. With these misfit functionals, we also evaluated the relative significance of the two effects of intrinsic attenuation on seismic records, namely amplitude decay and anelastic dispersion. The goal was to identify the best approach in isolating attenuation anomalies from other effects such as seismic noise and elastic heterogeneities. To that end, we carried out inversions targeting only  $Q_\mu$  without inverting for any unresolved elastic heterogeneity.

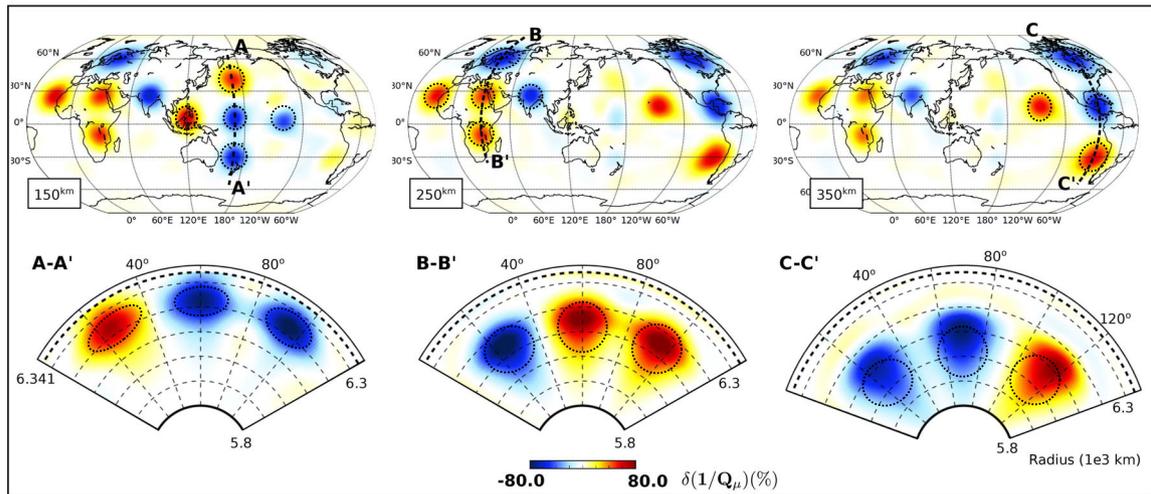
In these tests, the earth's structure is modelled as radially anisotropic and anelastic with frequency independent  $Q_\mu$ . All the wavefield simulations are performed using CSEM, and anelastic behaviour is introduced by means of standard linear solid models. SEM simulations provide high accuracy in addressing (de)focusing effects due to elastic heterogeneities, thus providing a realistic basis for the comparison of the relative performance of the three misfit functionals considered.

We have presented 1-D and 3-D synthetic tests in a period range of 60–400 s, as was considered in building previous upper-mantle



**Figure 12.** The 1-D  $Q_\mu$  profiles along the centre of the anomalies located at the mid-points of the A-A', B-B' and C-C' paths shown in Figs 9 and 11. The contribution of the overtones in reducing the smearing in the radial direction and recovering the deep anomalies below  $\sim 300$  km is clearly visible.

#### E-WF Q Inversion



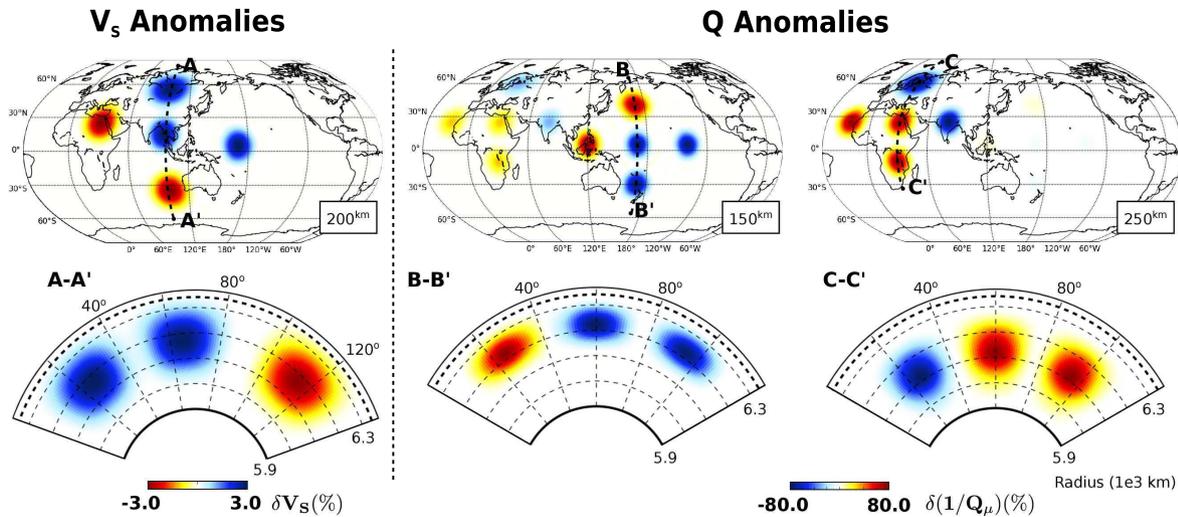
**Figure 13.** The recovered models in two iterations using E-WF in the presence of seismic noise. Real noise records collected from the stations are added to the synthetic seismograms computed for the target model before the inversion. The effect of the noise is negligible for E-WF and SA. For the latter, we recovered a model very similar to the one shown in Fig. 9. The locations of the  $Q_\mu$  anomalies in the target model are circled by dashed lines.

seismic velocity models in our group. In the 3-D tests, inversions were conducted using a realistic and limited ray coverage replicated from a subset of Z-component Rayleigh wave data ( $R_1$ ,  $R_2$ ,  $XR_1$ ,  $XR_2$ ) used in building the SEMUCB-WM1 global elastic model. Using the synthetic waveform data set computed for target models with  $Q_\mu$  and, in some tests,  $V_S$  anomalies, we tested the three misfit functionals for recovering the anomalies starting with a spherically symmetric Earth model. The size of anomalies in the hypothetical models vary from 4000 km to 5000 km laterally and 100 to 200 km radially. The sizes of anomalies are chosen in these ranges to allow us to employ a coarse model parametrization grid, and to achieve convergence in a few iterations reducing the computational cost. The peak amplitudes of anomalies, on the other hand, are chosen such as to match those in existing global attenuation and seismic

velocity models. The conclusions reached can be generalized to the cases with smaller size heterogeneities, which would require more refined model parametrizations, computations to shorter periods, and, likely, more iterations.

In our first 1-D synthetic test, we verified the accuracy of first order normal mode perturbation theory in the PAVA approximation, and carried out a comparative assessment of the three misfit functionals for attenuation imaging along a source-to-receiver path, using a single  $R_1$  wave packet. This exercise gave initial insights into the performance of the misfit functionals with encouraging results for E-WF and SA misfit functionals.

Moving to a more realistic setting, in the first 3-D synthetic test, we considered the recovery of a hypothetical model with both positive and negative 3-D  $Q_\mu$  heterogeneities, but no heterogeneity in



**Figure 14.** Hypothetical target model including both 3-D  $V_S$  and  $Q_\mu$  perturbations.  $V_S$  perturbations are shown in the left panel (peaking at 200 km depth with 3 per cent amplitude), while  $Q_\mu$  perturbations are shown in the middle and right panels (peaking at 150 and 250 km depth with 80 per cent amplitude). These perturbations are added to the 1-D unperturbed model presented in Section 3.1. The velocity perturbations are relatively larger in size and depth range as visible in the map (top) and depth-section (bottom) views. The two types of perturbations are located very close to each other in some areas, and some of them even coincide, in order to evaluate the leakage from one to another as an extreme scenario.

shear velocity. In this test, all the misfit functionals performed very well except WF in the absence of anelastic dispersion. In this case, although we were able to recover the sign of the attenuation perturbations, the depth resolution deteriorated with significant lateral and in-depth smearing. Including anelastic dispersion in the computation of the partial derivatives for the inversion led to somewhat improved images.

The significance of including overtone Rayleigh wave data in attenuation imaging is illustrated by comparing inversions with and without overtone wave packets. Excluding overtones led to significantly poorer recovery of the deep  $Q_\mu$  anomalies (below  $\sim 250$  km depth), for both E-WF and SA misfit functionals. We conclude from this test that overtones are needed to map 3-D  $Q_\mu$  structure at transition zone depths and to provide additional constraints on near-surface anomalies.

The second 3-D test was designed to evaluate the performance of E-WF and SA in the presence of seismic noise. Both methods performed well when realistic noise levels were added. Increasing the noise level to test the limits of both methods, we observed that larger norm-damping parameter values were necessary to obtain a stable model, thus reducing the recovered amplitude anomalies. Also, as the noise level increased, SA performed better than E-WF especially in recovering the deep structures. This illustrates the drawback of working in the time domain, as the introduced noise distorts not only the amplitude but also the phase of the signal. Although E-WF is less sensitive to phase anomalies than WF, it is still affected by anelastic dispersion. On the other hand, the SA approach relies only on amplitude anomalies in the spectral domain, isolating the attenuation effect more accurately, provided the appropriate time-window chosen.

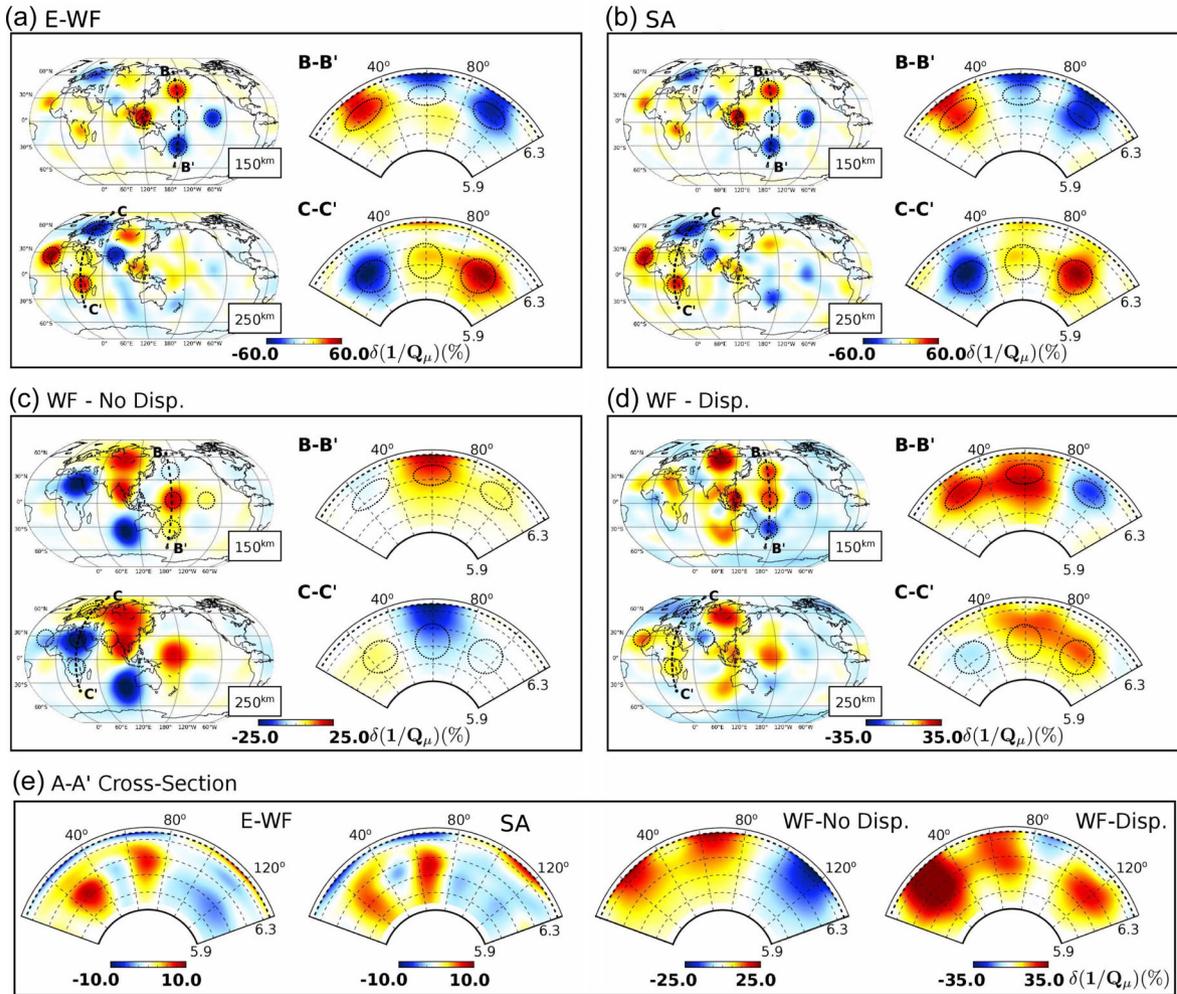
Notably, the resolution of deep structure is lost with increased noise level, as presented in appendix B, for both E-WF and SA misfit functionals. This is due to the definition of our noise level which leads to the suppression of low amplitude overtones at a higher rate than fundamental mode Rayleigh waves.

For the final 3-D test, we built a more complex target model including both  $V_S$  and  $Q_\mu$  heterogeneities. Leaving  $V_S$  heterogeneities

as unknown, we tried to recover the  $Q_\mu$  model. As it turns out, the leakage from unresolved  $V_S$  anomalies to the  $Q_\mu$  model is inevitable regardless of the misfit functionals chosen. The leakage appears as positive attenuation heterogeneities for slow regions and vice versa, as expected for the effects of focusing. However, the level of leakage is very low in the E-WF and SA cases, whereas it is quite dominant for WF. In fact, in the absence of anelastic dispersion terms in the **G** matrix, WF leads to a  $Q_\mu$  model visibly dominated by  $V_S$  leaks.

In general, we conclude that for attenuation imaging SA and E-WF are the more appropriate misfit functionals, with the former performing relatively better with increasing noise level and frequency content. WF prioritizes phase information over amplitude information, and is more sensitive to elastic heterogeneities. For all the three misfit functionals, it remains important to try and first determine the best possible elastic 3-D model, in order to accurately account for focusing effects. In accordance with these conclusions, we suggest updating seismic velocity and attenuation models sequentially using WF for the former and E-WF or SA for the latter, as well as including overtones in the inversions to map attenuation heterogeneities below  $\sim 300$  km and to better constrain the upper-mantle structure.

We note that in our previous efforts at 3-D upper mantle attenuation mapping using waveforms in the time domain, we used the WF misfit functional and corrected for anelastic dispersion effects due to 1-D  $Q_\mu$  structure but not due to 3-D  $Q_\mu$  structure (Gung & Romanowicz 2004). However, in that study, a very rigorous data selection was performed, keeping only those waveforms which were in phase with the synthetic seismograms computed in a 3-D elastic model, and considering only very long wavelength 3-D  $Q_\mu$  structure. This resulted in realistic 3-D  $Q_\mu$  images down to transition zone depths, albeit with likely unreliable amplitudes of lateral variations. In more recent work, Dalton & Ekström (2006); Dalton *et al.* (2008) and Bao *et al.* (2016) adopted a SA misfit functional - albeit only for fundamental mode Rayleigh waves—applied to a large data set (much larger than the one considered in our tests, in which we restricted the number of events due to the heavy SEM computations involved). This plus corrections for



**Figure 15.** Summary of results of synthetic test with  $V_S$  and  $Q_\mu$  perturbations in the target model, after two iterations. The presence of  $V_S$  perturbations is ignored in the inversions. E-WF (a) and SA (b) succeed very well in recovering the  $Q_\mu$  perturbations with relatively small leakage from  $V_S$  perturbations compared to WF. Without considering the anelastic dispersion, WF maps  $V_S$  perturbations into  $Q_\mu$  perturbations (c). Although considering anelastic dispersion improves the performance of WF (d), the leakage is still quite visible and dominant. In both cases, we have to apply higher damping to keep the inversion stable. The level of leakage from  $V_S$  to  $Q_\mu$  is shown in the depth cross-sections of the recovered models along A-A' (e). The leakage is worst with WF, as can be seen from the amplitude of the imaged perturbations in  $Q_\mu$ , where there should be none. In all the panels, the locations of the  $Q_\mu$  anomalies in the target model are circled by dashed lines.

focusing likely allowed them to reach higher lateral resolution for  $Q_\mu$  in the upper 200 km of the mantle. Still, as shown here, fundamental modes are not sufficient to resolve 3-D  $Q_\mu$  structure in the transition zone.

It is well known that significant amplitude uncertainty in seismic waveforms stem from inaccurately known source parameters and instrument responses. However, these uncertainties cannot be avoided through the definition of the misfit functional. To that end, within the scope of waveform comparison, approaches that rely on differential comparison of similar waveforms that travel through overlapping paths have been suggested by several studies (e.g. Romanowicz 1990; Bhattacharyya *et al.* 1993; Ford *et al.* 2012). Using our hybrid full-waveform inversion through the comparison of individual waveforms in time or frequency domain, we suggest a joint or iterative inversion for source parameters and receiver terms (scalar or frequency dependent as in the work of Dalton & Ekström 2006), for 3-D  $Q_\mu$  perturbations, and for perturbations in elastic structure.

This study is part of an effort towards building a new global upper-mantle attenuation model based on long-period full-waveform in-

version. By evaluating our hybrid full-waveform inversion scheme through synthetic tests, we have developed physical and operational insights into the recovery of anelastic heterogeneities in the case of real data, the result of which are presented in a companion paper (Karaoğlu & Romanowicz 2017).

## ACKNOWLEDGEMENTS

The authors thank Gabi Laske and three anonymous reviewers for their constructive comments on the manuscript. This work was supported by the European Research Council under the EC's 7th Framework Programme (FP7-IDEAS-ERC)/ERC Advanced Grant WAVETOMO. The computations were performed on OCCIGEN of the HPC national facility CINES in France, on EDISON and CORI at the National Energy Research Scientific Computing Center (NERSC), USA, and the local cluster (MALBEC) of Institute de Physique du Globe de Paris. BR acknowledges support from NSF grant EAR-1497229.

## REFERENCES

- Adenis, A., Debayle, E. & Ricard, Y., 2017. Seismic evidence for broad attenuation anomalies in the asthenosphere beneath the Pacific ocean, *Geophys. J. Int.*, **209**(3), 1677–1698.
- Anderson, J.G. & Hough, S.E., 1984. A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bull. seism. Soc. Am.*, **74**(5), 1969–1993.
- Auer, L., Boschi, L., Becker, T., Nissen-Meyer, T. & Giardini, D., 2014. Savani: a variable resolution whole-mantle model of anisotropic shear velocity variations based on multiple data sets, *J. geophys. Res.*, **119**(4), 3006–3034.
- Bao, X., Dalton, C.A. & Ritsema, J., 2016. Effects of elastic focusing on global models of Rayleigh wave attenuation, *Geophys. J. Int.*, **207**(2), 1062–1079.
- Bhattacharyya, J., Shearer, P. & Masters, G., 1993. Inner core attenuation from short-period PKP (BC) versus PKP (DF) waveforms, *Geophys. J. Int.*, **114**(1), 1–11.
- Bhattacharyya, J., Masters, G. & Shearer, P., 1996. Global lateral variations of shear wave attenuation in the upper mantle, *J. geophys. Res.*, **101**(B10), 22 273–22 289.
- Bozdag, E., Trampert, J. & Tromp, J., 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.*, **185**(2), 845–870.
- Cammarano, F. & Romanowicz, B., 2008. Radial profiles of seismic attenuation in the upper mantle based on physical models, *Geophys. J. Int.*, **175**(1), 116–134.
- Capdeville, Y., To, A. & Romanowicz, B., 2003. Coupling spectral elements and modes in a spherical earth: an extension to the sandwich case, *Geophys. J. Int.*, **154**(1), 44–57.
- Chang, S.-J., Ferreira, A.M., Ritsema, J., Heijst, H.J. & Woodhouse, J.H., 2015. Joint inversion for global isotropic and radially anisotropic mantle structure including crustal thickness perturbations, *J. geophys. Res.*, **120**(6), 4278–4300.
- Dalton, C.A. & Ekström, G., 2006. Global models of surface wave attenuation, *J. geophys. Res.*, **111**(B5), doi:10.1029/2005JB003997.
- Dalton, C.A. & Faul, U.H., 2010. The oceanic and cratonic upper mantle: clues from joint interpretation of global velocity and attenuation models, *Lithos*, **120**(1), 160–172.
- Dalton, C.A., Ekström, G. & Dziewoński, A.M., 2008. The global attenuation structure of the upper mantle, *J. geophys. Res.*, **113**(B9), doi:10.1029/2007JB005429.
- Dalton, C.A., Hjörleifsdóttir, V. & Ekström, G., 2014. A comparison of approaches to the prediction of surface wave amplitude, *Geophys. J. Int.*, **196**(1), 386–404.
- Dalton, C.A., Bao, X. & Ma, Z., 2017. The thermal structure of cratonic lithosphere from global Rayleigh wave attenuation, *Earth planet. Sci. Lett.*, **457**, 250–262.
- Durek, J.J. & Ekström, G., 1996. A radial model of anelasticity consistent with long-period surface-wave attenuation, *Bull. seism. Soc. Am.*, **86**(1A), 144–158.
- Durek, J.J., Ritzwoller, M.H. & Woodhouse, J.H., 1993. Constraining upper mantle anelasticity using surface wave amplitude anomalies, *Geophys. J. Int.*, **114**(2), 249–272.
- Dziewoński, A.M. & Anderson, D.L., 1981. Preliminary reference earth model, *Phys. Earth planet. Inter.*, **25**(4), 297–356.
- Fichtner, A., Kennett, B.L., Igel, H. & Bunge, H.-P., 2008. Theoretical background for continental- and global-scale full-waveform inversion in the time–frequency domain, *Geophys. J. Int.*, **175**(2), 665–685.
- Ford, S.R., Garnero, E.J. & Thorne, M.S., 2012. Differential  $t^*$  measurements via instantaneous frequency matching: observations of lower mantle shear attenuation heterogeneity beneath western central America, *Geophys. J. Int.*, **189**(1), 513–523.
- French, S. & Romanowicz, B., 2014. Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography, *Geophys. J. Int.*, **199**(3), 1303–1327.
- French, S., Lekić, V. & Romanowicz, B., 2013. Waveform tomography reveals channelled flow at the base of the oceanic asthenosphere, *Science*, **342**(6155), 227–230.
- Gung, Y. & Romanowicz, B., 2004. Q tomography of the upper mantle using three-component long-period waveforms, *Geophys. J. Int.*, **157**(2), 813–830.
- Gung, Y., Panning, M. & Romanowicz, B., 2003. Global anisotropy and the thickness of continents, *Nature*, **422**(6933), 707–711.
- Hwang, Y., Ritsema, J. & Goes, S., 2011. Global variation of body-wave attenuation in the upper mantle from teleseismic p wave and s wave spectra, *Geophys. Res. Lett.*, **38**(8), doi:10.1029/2011GL046812.
- Jackson, I., Faul, U.H., Gerald, F., John, D. & Tan, B.H., 2004. Shear wave attenuation and dispersion in melt-bearing olivine polycrystals: 1. specimen fabrication and mechanical testing, *J. geophys. Res.*, **109**(B6), doi:10.1029/2003JB002406
- Jordan, T.H., 1978. A procedure for estimating lateral variations from low-frequency eigenspectra data, *Geophys. J. Int.*, **52**(3), 441–455.
- Kanamori, H., 1967. Spectrum of p and pcp in relation to the mantle-core boundary and attenuation in the mantle, *J. geophys. Res.*, **72**(2), 559–571.
- Kanamori, H. & Anderson, D.L., 1977. Importance of physical dispersion in surface wave and free oscillation problems: Review, *Rev. Geophys.*, **15**(1), 105–112.
- Karaoğlu, H. & Romanowicz, B., 2017. Inferring global upper-mantle shear attenuation structure by waveform tomography using the spectral element method, *Geophys. J. Int.*, in revision.
- Karato, S.-i., 1993. Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.*, **20**(15), 1623–1626.
- Karato, S.-i., 2003. Mapping water content in the upper mantle, in *Inside the Subduction Factory*, pp. 135–152, ed. Eiler, J., American Geophysical Union.
- Kennett, B., Engdahl, E. & Buland, R., 1995. Constraints on seismic velocities in the earth from traveltimes, *Geophys. J. Int.*, **122**(1), 108–124.
- Komatitsch, D. & Vilotte, J.-P., 1998. The spectral element method: An efficient tool to simulate the seismic response of 2d and 3d geological structures, *Bull. seism. Soc. Am.*, **88**(2), 368–392.
- Lawrence, J.F. & Wyssession, M.E., 2006. Seismic evidence for subduction-transported water in the lower mantle, *Earth's Deep Water Cycle*, pp. 251–261.
- Lekić, V. & Romanowicz, B., 2011. Inferring upper-mantle structure by full waveform tomography with the spectral element method, *Geophys. J. Int.*, **185**(2), 799–831.
- Lekić, V., Matas, J., Panning, M. & Romanowicz, B., 2009. Measurement and implications of frequency dependence of attenuation, *Earth planet. Sci. Lett.*, **282**(1), 285–293.
- Li, X.-D. & Romanowicz, B., 1995. Comparison of global waveform inversions with and without considering cross-branch modal coupling, *Geophys. J. Int.*, **121**(3), 695–709.
- Li, X.-D. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. geophys. Res.*, **101**(B10), 22 245–22 272.
- Liu, H.-P., Anderson, D.L. & Kanamori, H., 1976. Velocity dispersion due to anelasticity; implications for seismology and mantle composition, *Geophys. J. Int.*, **47**(1), 41–58.
- Ma, Z., Masters, G. & Mancinelli, N., 2016. Two-dimensional global Rayleigh wave attenuation model by accounting for finite-frequency focusing and defocusing effect, *Geophys. J. Int.*, **204**(1), 631–649.
- Masters, G. & Gilbert, F., 1983. Attenuation in the earth at low frequencies, *Phil. Trans. R. Soc. A*, **308**(1504), 479–522.
- Matheny, M.P. & Nowack, R.L., 1995. Seismic attenuation values obtained from instantaneous-frequency matching and spectral ratios, *Geophys. J. Int.*, **123**(1), 1–15.
- Mégnin, C. & Romanowicz, B., 1999. The effects of the theoretical formalism and data selection on mantle models derived from waveform tomography, *Geophys. J. Int.*, **138**(2), 366–380.
- Mégnin, C. & Romanowicz, B., 2000. The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms, *Geophys. J. Int.*, **143**(3), 709–728.
- Moulik, P. & Ekström, G., 2014. An anisotropic shear velocity model of the earth's mantle using normal modes, body waves, surface waves and long-period waveforms, *Geophys. J. Int.*, **199**(3), 1713–1738.
- Oki, S., Fukao, Y. & Obayashi, M., 2004. Reference frequency of teleseismic body waves, *J. geophys. Res.*, **109**(B4), doi:10.1029/2003JB002821.

- Panning, M. & Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.*, **167**(1), 361–379.
- Panning, M.P., Capdeville, Y. & Romanowicz, B.A., 2009. Seismic waveform modelling in a 3-D earth using the born approximation: potential shortcomings and a remedy, *Geophys. J. Int.*, **177**(1), 161–178.
- Park, J., 1987. Asymptotic coupled-mode expressions for multiplet amplitude anomalies and frequency shifts on an aspherical earth, *Geophys. J. Int.*, **90**(1), 129–169.
- Romanowicz, B., 1987. Multiplet-multiplet coupling due to lateral heterogeneity: asymptotic effects on the amplitude and frequency of the earth's normal modes, *Geophys. J. Int.*, **90**(1), 75–100.
- Romanowicz, B., 1990. The upper mantle degree 2: constraints and inferences from global mantle wave attenuation measurements, *J. geophys. Res.*, **95**(B7), 11 051–11 071.
- Romanowicz, B., 1994. On the measurement of anelastic attenuation using amplitudes of low-frequency surface waves, *Phys. Earth planet. Inter.*, **84**(1-4), 179–191.
- Romanowicz, B., 1995. A global tomographic model of shear attenuation in the upper mantle, *J. geophys. Res.*, **100**(B7), 12 375–12 394.
- Romanowicz, B. & Gung, Y., 2002. Superplumes from the core-mantle boundary to the lithosphere: implications for heat flux, *Science*, **296**(5567), 513–516.
- Romanowicz, B., Roul, G. & Kohl, T., 1987. The upper mantle degree two pattern: constraints from geoscope fundamental spheroidal mode eigenfrequency and attenuation measurements, *Geophys. Res. Lett.*, **14**(12), 1219–1222.
- Romanowicz, B.A., Panning, M.P., Gung, Y. & Capdeville, Y., 2008. On the computation of long period seismograms in a 3-D earth using normal mode based approximations, *Geophys. J. Int.*, **175**(2), 520–536.
- Roul, G., 1975. Attenuation of seismic waves of very low frequency, *Phys. Earth planet. Inter.*, **10**(2), 159–166.
- Roul, G. & Clévéde, E., 2000. New refinements in attenuation measurements from free-oscillation and surface-wave observations, *Phys. Earth planet. Inter.*, **121**(1), 1–37.
- Roul, G., Romanowicz, B. & Montagner, J., 1990. 3-D upper mantle shear velocity and attenuation from fundamental mode free oscillation data, *Geophys. J. Int.*, **101**(1), 61–80.
- Sailor, R.V. & Dziewonski, A.M., 1978. Measurements and interpretation of normal mode attenuation, *Geophys. J. Int.*, **53**(3), 559–581.
- Selby, N.D. & Woodhouse, J.H., 2000. Controls on rayleigh wave amplitudes: attenuation and focusing, *Geophys. J. Int.*, **142**(3), 933–940.
- Tarantola, A., 1984. Linearized inversion of seismic reflection data, *Geophys. Prospect.*, **32**(6), 998–1015.
- Teng, T.-L., 1968. Attenuation of body waves and the q structure of the mantle, *J. geophys. Res.*, **73**(6), 2195–2208.
- Tonn, R., 1991. The determination of the seismic quality factor q from vsp data: a comparison of different computational methods, *Geophys. Prospect.*, **39**(1), 1–27.
- Wang, Z. & Dahlen, F., 1995. Spherical-spline parameterization of three-dimensional earth models, *Geophys. Res. Lett.*, **22**(22), 3099–3102.
- Warren, L.M. & Shearer, P.M., 2002. Mapping lateral variations in upper mantle attenuation by stacking *P* and *PP* spectra, *J. geophys. Res.*, **107**(B12), doi:10.1029/2001JB001195.
- Widmer, R., Masters, G. & Gilbert, F., 1991. Spherically symmetric attenuation within the earth from normal mode data, *Geophys. J. Int.*, **104**(3), 541–553.
- Woodhouse, J., 1988. The calculation of eigenfrequencies and eigenfunctions of the free oscillations of the earth and the sun, in *Seismological Algorithms*, pp. 321–370, ed. Laske, G., Oxford Journals. Geophysical Journal International.
- Woodhouse, J. & Dahlen, F., 1978. The effect of a general aspherical perturbation on the free oscillations of the earth, *Geophys. J. Int.*, **53**(2), 335–354.
- Woodhouse, J. & Gernius, T., 1982. Surface waves and free oscillations in a regionalized earth model, *Geophys. J. Int.*, **68**(3), 653–673.
- Woodhouse, J. & Wong, Y., 1986. Amplitude, phase and path anomalies of mantle waves, *Geophys. J. Int.*, **87**(3), 753–773.

- Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle: Three-dimensional modeling of earth structure by inversion of seismic waveforms, *J. geophys. Res.*, **89**(B7), 5953–5986.
- Zhou, Y., Dahlen, F. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.*, **158**(1), 142–168.
- Zhu, H., Bozdağ, E., Duffy, T.S. & Tromp, J., 2013. Seismic attenuation beneath Europe and the North Atlantic: Implications for water in the mantle, *Earth planet. Sci. Lett.*, **381**, 1–11.

## APPENDIX A

The Fréchet derivative computations are based on the time domain acceleration seismogram  $u(t)$  computed using the NACT approximation. NACT accounts for the coupling of multiplets  $k$  along the same mode branch, as does the PAVA (Woodhouse & Dziewonski 1984; Romanowicz 1987), and for coupling of multiplets  $k$  and  $k'$  belonging to different mode branches. For a 3-D anelastic model,  $u(t)$  is computed as follows in the PAVA approximation (e.g. Li & Romanowicz 1995):

$$u(t) = \Re \left\{ \sum_k e^{i\tilde{\omega}_k t} e^{-\tilde{\alpha}_k t} \left[ \sum_m R_k^m S_k^m \right] \right\} \quad (\text{A1})$$

where  $m$  is the azimuthal order of a singlet within a multiplet  $k$  of radial order  $n$  and angular order  $l$ .  $R_k^m$  and  $S_k^m$  are the receiver and source terms computed for singlet  $m$  within the multiplet  $k$  (Woodhouse & Gernius 1982). Multiplet frequency ( $\tilde{\omega}_k$ ) and decay rate/attenuation factor ( $\tilde{\alpha}_k$ ) computed in the 3-D model are related to those ( $\omega_k$  and  $\alpha_k$ ) computed in the 1-D reference model as follows:

$$\tilde{\omega}_k = \omega_k + \frac{1}{\Delta} \int_S^R \delta\omega_k d\Delta \quad (\text{A2})$$

$$\tilde{\alpha}_k = \alpha_k + \frac{1}{\Delta} \int_S^R \delta\alpha_k d\Delta \quad (\text{A3})$$

where the source to receiver path-averaged integrals contain the effects of local frequency and attenuation perturbations (Jordan 1978), and  $\Delta$  is epicentral distance.

In the case of NACT, coupling across multiplets  $k \neq k'$  is considered to first order, and the expression for  $u(t)$  becomes

$$u(t) = \Re \left\{ \sum_k \left[ (1 - it\tilde{\omega}_k) e^{i\tilde{\omega}_k t} e^{-\tilde{\alpha}_k t} \sum_m R_k^m S_k^m + \sum_{k' \geq k} \frac{(e^{-\tilde{\alpha}_k t} e^{i\tilde{\omega}_k t} - e^{-\tilde{\alpha}_{k'} t} e^{i\tilde{\omega}_{k'} t}) A_{kk'}}{(\omega_k + \omega_{k'}) + ((\tilde{\omega}_k + \tilde{\omega}_{k'}) + i(\tilde{\alpha}_k + \tilde{\alpha}_{k'}))} \right] \right\} \quad (\text{A4})$$

$$A_{kk'} = \frac{1}{2\pi} \left[ Q_{kk'}^1 \int_0^{2\pi} (\omega_k + \omega_{k'}) (\delta\omega_{kk'} + i\delta\alpha_{kk'}) \cos((l-l')\varphi) d\varphi + Q_{kk'}^2 \int_0^{2\pi} (\omega_k + \omega_{k'}) (\delta\omega_{kk'} + i\delta\alpha_{kk'}) \sin((l-l')\varphi) d\varphi \right] \quad (\text{A5})$$

where  $A_{kk'}$  is the amplitude term of the across branch multiplet coupling contribution, and the integral is calculated along the great circle path containing source ( $S$ ) and receiver ( $R$ ). The expressions for  $Q_{kk'}^1$  and  $Q_{kk'}^2$  are given in the appendix A of Li & Romanowicz (1995). Also,

$$\delta\alpha_{kk'}(\theta, \phi) = \frac{1}{\omega_k + \omega_{k'}} \left[ \int_0^a \delta \left( \frac{1}{Q_\mu(r, \theta, \phi)} \right) K_{kk'}^{Q_\mu}(r) + \delta \left( \frac{1}{Q_\nu(r, \theta, \phi)} \right) K_{kk'}^{Q_\nu}(r) r^2 dr \right] \quad (\text{A6})$$

$$\delta\omega_{kk'}(\theta, \phi) = \frac{1}{\omega_k + \omega_{k'}} \left[ \int_0^a \delta m(r, \theta, \phi) K_{kk'}^m(r) r^2 dr \right] + \frac{2}{\pi} \ln \left[ \frac{(\omega_k + \omega_{k'})}{2\omega_0} \right] \delta\alpha_{kk'}. \quad (\text{A7})$$

where  $a$  is the earth's radius. Excluding shear and bulk attenuation perturbations ( $\delta Q_\mu^{-1}$ ,  $\delta Q_\kappa^{-1}$ ) from these expressions leaves us with perturbations in the parameters of the elastic model  $m$ . These parameters are reduced to two of the 5 radially anisotropic elastic parameters ( $V_{Siso}$  and  $\xi$ ) in the present case. The corresponding kernel expressions ( $K_{kk'}^m$ ) are given by Woodhouse & Dahlen (1978) and Romanowicz (1987) for the cases with  $k = k'$  and  $k \neq k'$  respectively. The kernel expressions for  $Q_\mu^{-1}$  and  $Q_\kappa^{-1}$  are directly related to those for shear ( $\mu$ ) and bulk ( $\kappa$ ) moduli respectively, and can be written in terms of wave speed kernels computed for Voigt average isotropic  $S$ - and  $P$ -wave velocities ( $V_S$ ,  $V_P$ ) as follows:

$$K_{Q_\mu} = \mu_0 K_\mu = \frac{1}{2} \left( V_S K_S + \frac{4}{3} \frac{V_S^2}{V_P^2} V_P K_P \right) \quad (\text{A8})$$

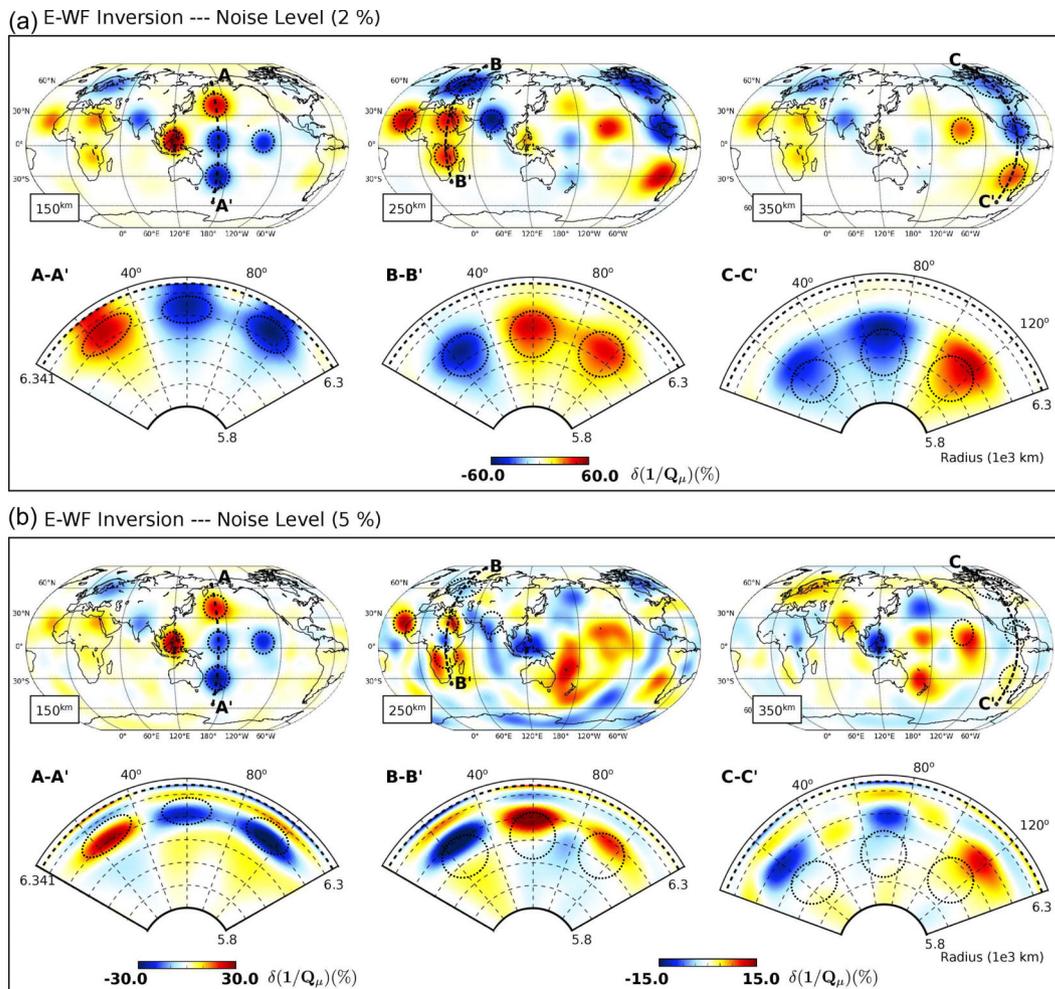
$$K_{Q_\kappa} = \kappa_0 K_\kappa = \frac{1}{2} \left( \frac{(V_P^2 - \frac{4}{3} V_S^2)}{V_P^2} V_P K_P \right), \quad (\text{A9})$$

where we have dropped subscripts ( $k, k'$ ) and  $\mu_0$ ,  $\kappa_0$  denote the shear and bulk moduli for the 1-D average model, respectively. Expressions for  $K_S$  and  $K_P$  are provided by Panning & Romanowicz (2006) in terms of the velocity kernels of horizontally and vertically polarized  $P$ - and  $S$ -waves.

The attenuation perturbation ( $\delta\alpha_{kk'}$ ) that appears in the frequency perturbation ( $\delta\omega_{kk'}$ ) in eq. (A7) leads to physical dispersion. In the presented synthetic tests, we test the significance of this term when using the three misfit functionals considered, which turns out to be crucial for WF misfit functional.

## APPENDIX B

We tested the E-WF and SA misfit functionals in the presence of real levels of seismic noise in Section 3.2.3. Here, we increase the noise level to test the misfit functionals, E-WF and SA, in extreme



**Figure B1.** The recovered models using E-WF with two different noise levels. Real noise records collected from the stations are added to the synthetic data (computed in the target model) after multiplication by a factor of 2 per cent (a) and 5 per cent (b) of the peak amplitude value in the full record (3 hr). The peak values in these tests correspond to the first arriving Rayleigh wave. Thus, the noise effect is more pronounced for the overtones which are more sensitive to deeper structure. The recovered model for the 2 per cent noise level shows more smearing laterally and in depth and weaker perturbation amplitudes compared to the case with real noise (Fig. 13). Increasing the noise level to 5 per cent, we lose the depth resolution below 250 km.

scenarios until they break down  $Q_\mu$  inversions. In this test, real noise records collected from the stations considered are added to the synthetic data after scaling their root-mean-square amplitudes by a factor defined as the percentage of the peak amplitude value in the full 3 hr synthetic record used as data. By taking the peak amplitude in the full synthetic waveform data as reference for the added noise, we ensure that it is at the same level for all the synthetic data, independent of the event magnitude or the noise level at the station. Next, we increase the noise level until we lose the recovered model accuracy significantly. In what follows, we present the best recovered models for two levels of noise (2 per cent and 5 per cent). The first noise level at 2 per cent is relatively low, and both misfit functionals perform well, whereas the second one (5 per cent) is a high noise level that suppresses the overtones and shows the difference in the performance of the two misfit functionals.

Fig. B1 shows the recovered models using E-WF. For a 2 per cent noise level, the recovery of the position and shape of heterogeneities is satisfactory, although we lose the perturbation amplitude accuracy achieved in the absence of noise, with some more smearing. Once the noise level is increased to 5 per cent though, there is significant loss of resolution, especially in depth.

The peak amplitudes in the synthetic data correspond to the first arriving Rayleigh waves, therefore the noise effect is more pronounced for the overtones. This leads to a loss of resolution at larger depths with increased noise level as the overtones are more sensitive to them than the fundamental modes.

Lowering the signal-to-noise ratio for the overtones leads to a leakage of near-surface heterogeneities to greater depths. This is clear in the case with 5 per cent noise level, where heterogeneities peaking at 150 km reappear at 350 km with the reversed sign. We also see the mapping of deep structure heterogeneities to shallower depths as the ones at 350 km weakly appear at 250 km. Although it might be possible to lower the amplitudes of such leaks and fluctuations by imposing a higher damping level ( $\alpha$  in eq. 1) in the  $Q_\mu$  inversion at the expense of losing recovered perturbation amplitudes, we prefer to present the figures for this case at the chosen damping level to show: (1) A pattern that can be misinterpreted when working with real data. (2) The significance of including reliable overtone waveforms when imaging the attenuation below  $\sim 250$  km depth.

SA performs better with seismic noise compared to E-WF. The recovered models are presented Fig. B2. The relatively better

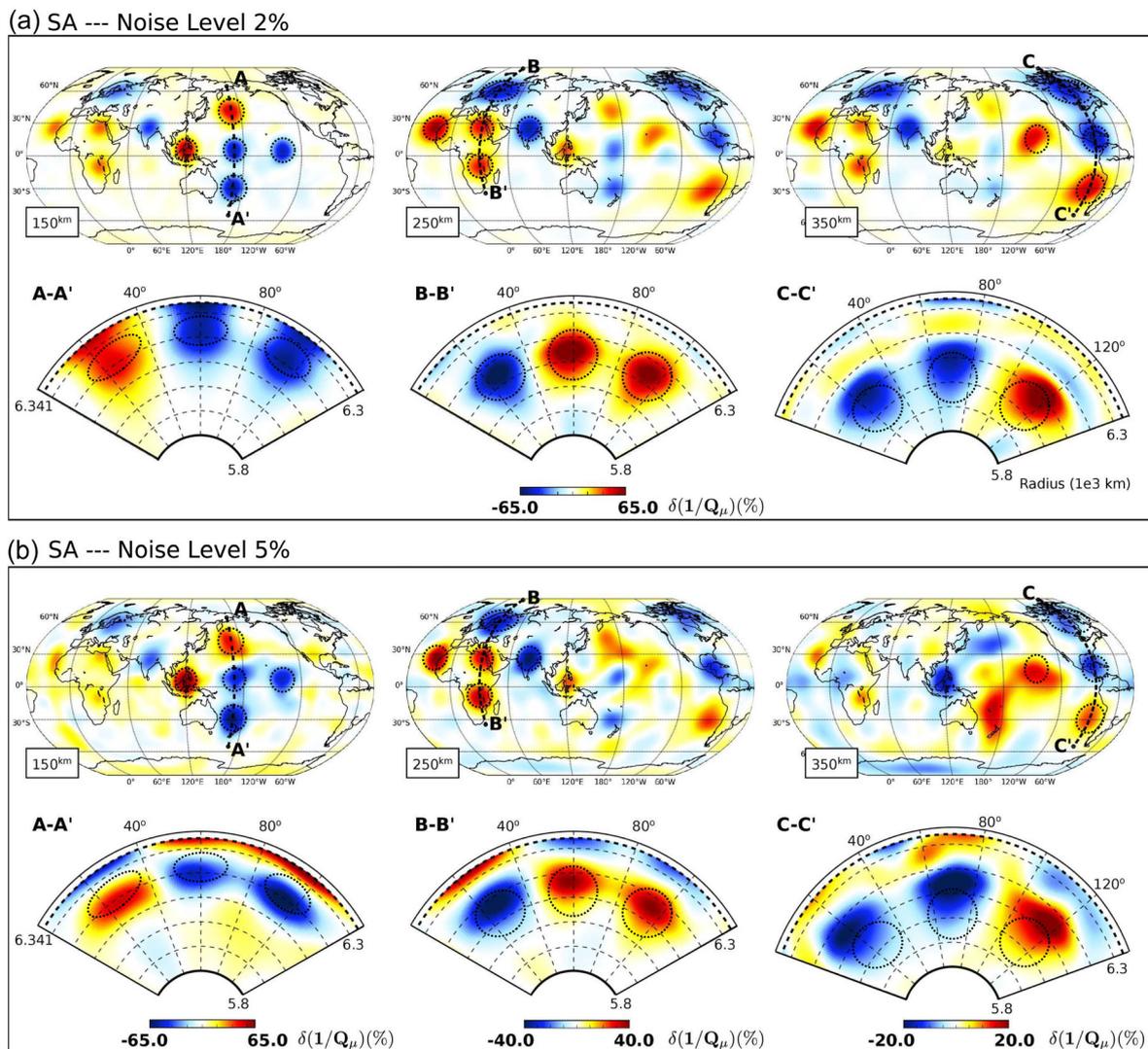


Figure B2. The recovered models using SA with the same noise levels as in Fig. B1. In general, we observe a better performance compared to E-WF for the same level of noise. This is so both for the depth resolution and the amplitude of the recovered perturbations.

performance of SA is clear for the case with 5 per cent noise level. As opposed to E-WF, SA succeeds in recovering the deep  $Q_\mu$  heterogeneities below  $\sim 250$  km, although we lose the accuracy in the recovery of perturbation amplitudes.

The relatively higher performance of SA compared to E-WF shows an inherent difficulty of working in the time domain when imaging attenuation. Although E-WF suppresses the phase anomaly

dominance that we see in WF, one needs to be careful as it cannot fully isolate the amplitude anomalies in the presence of high levels of noise.

Common to both the E-WF and SA cases is the necessity of increasing the value of  $\alpha$  (eq. 1) in the inversion step in the presence of noise, as expected. The primary effect of that is to reduce the amplitudes of the recovered anomalies.