Seismic structure of the European upper mantle based on adjoint tomography

H. Zhu, J. Tromp, E. Bozdağ

Geophysical Journal International

Geophys. J. Int. (2015) **201,** 18–52 doi: 10.1093/gji/ggu492
GJI Seismology

Seismic structure of the European upper mantle based on adjoint tomography

Hejun Zhu,[1,*] Ebru Bozdağ[1,†] and Jeroen Tromp[1,2]

[1]*Department of Geosciences, Princeton University, Princeton, NJ* 08544, *USA. E-mail: hejunzhu@Princeton.EDU*
[2]*Program in Applied & Computational Mathematics, Princeton University, Princeton, NJ* 08544, *USA*

Accepted 2014 December 24. Received 2014 December 22; in original form 2014 May 23

SUMMARY

We use adjoint tomography to iteratively determine seismic models of the crust and upper mantle beneath the European continent and the North Atlantic Ocean. Three-component seismograms from 190 earthquakes recorded by 745 seismographic stations are employed in the inversion. Crustal model EPcrust combined with mantle model S362ANI comprise the 3-D starting model, EU_{00}. Before the structural inversion, earthquake source parameters, for example, centroid moment tensors and locations, are reinverted based on global 3-D Green's functions and Fréchet derivatives. This study consists of three stages. In stage one, frequency-dependent phase differences between observed and simulated seismograms are used to constrain radially anisotropic wave speed variations. In stage two, frequency-dependent phase and amplitude measurements are combined to simultaneously constrain elastic wave speeds and anelastic attenuation. In these two stages, long-period surface waves and short-period body waves are combined to simultaneously constrain shallow and deep structures. In stage three, frequency-dependent phase and amplitude anomalies of three-component surface waves are used to simultaneously constrain radial and azimuthal anisotropy. After this three-stage inversion, we obtain a new seismic model of the European curst and upper mantle, named EU_{60}. Improvements in misfits and histograms in both phase and amplitude help us to validate this three-stage inversion strategy. Long-wavelength elastic wave speed variations in model EU_{60} compare favourably with previous body- and surface wave tomographic models. Some hitherto unidentified features, such as the Adria microplate, naturally emerge from the smooth starting model. Subducting slabs, slab detachments, ancient suture zones, continental rifts and backarc basins are well resolved in model EU_{60}. We find an anticorrelation between shear wave speed and anelastic attenuation at depths < 100 km. At greater depths, this anticorrelation becomes relatively weak, in agreement with previous global attenuation studies. Furthermore, enhanced attenuation is observed within the mantle transition zone beneath the North Atlantic Ocean. Consistent with typical radial anisotropy in 1-D reference models, the European continent is dominated by features with a radially anisotropic parameter $\xi > 1$, indicating predominantly horizontal flow within the upper mantle. In addition, subduction zones, such as the Apennines and Hellenic arcs, are characterized by vertical flow with $\xi < 1$ at depths greater than 150 km. We find that the direction of the fast anisotropic axis is closely tied to the tectonic evolution of the region. Averaged radial peak-to-peak anisotropic strength profiles identify distinct brittle-ductile deformation in lithospheric strength beneath oceans and continents. Finally, we use the 'point-spread function' to assess image quality and analyse trade-offs between different model parameters.

Key words: Seismic anisotropy; Seismic attenuation; Seismic tomography; Computational seismology; Europe.

1 INTRODUCTION

Owing to advances in high-performance computing and numerical methods (e.g. the spectral-element method, SEM; Komatitsch & Vilotte 1998; Komatitsch & Tromp 1999), seismologists are now

*Now at: BEG 2.112H, Jackson School of Geosciences, The University of Texas at Austin, Austin, TX 78757, USA.
†Now at: Laboratoire Géoazur, Université de Nice, Sophia-Antipolis, 250 rue Albert Einstein, 06560 Valbonne, France.

able to accurately simulate wave propagation in complex 3-D earth models, ranging from local (Komatitsch *et al.* 2004; Peter *et al.* 2011) to global scales (Komatitsch & Tromp 2002a,b). Synthetic seismograms based on the SEM and the latest 3-D earth models, for example, S362ANI (Kustowski *et al.* 2008b) or S40RTS (Ritsema *et al.* 2011), can match observed seismograms quite well at longer periods (> 60 s; Komatitsch *et al.* 2002; Tromp *et al.* 2010). However, at shorter periods there are remaining differences between observed and simulated seismograms due to unmodelled source complexity and 3-D heterogeneity. How to use these remaining differences to improve 3-D earth models is currently a very active research area in seismology (Chen *et al.* 2007; Fichtner *et al.* 2009, 2010; Tape *et al.* 2009, 2010; Lekic & Romanowicz 2011; Zhu *et al.* 2012, 2013; Fichtner *et al.* 2013; Zhu & Tromp 2013).

Adjoint methods, first introduced in seismology by Lailly (1983) and Tarantola (1984), enable us to numerically compute the gradient of a misfit function in complex 3-D earth models. This method is widely used to perform 'full waveform inversion' in exploration seismology (Gauthier *et al.* 1986; Mora 1987; Luo & Schuster 1991; Pratt *et al.* 1998; Brossier *et al.* 2009; Virieux & Operto 2009). Tromp *et al.* (2005) showed that the adjoint method is closely related to finite-frequency 'banana-doughnut' theory, which has started to replace classical ray-based tomography in recent years (Marquering *et al.* 1998, 1999; Dahlen *et al.* 2000; Hung *et al.* 2000; Montelli *et al.* 2004). Liu & Tromp (2006, 2008) applied the SEM and adjoint methods to numerically compute 3-D sensitivity kernels on both local and global scales. In combination with gradient-based optimization algorithms—for example, a preconditioned conjugate gradient approach (Fletcher & Reeves 1964) or the Broyden-Fletcher-Goldfarb-Shanno (BFGS; Broyden 1970; Fletcher 1970; Goldfarb 1970; Shanno 1970) quasi-Newton algorithm, in particular its limited-memory version (L-BFGS; Matthies & Strang 1979; Nocedal 1980)—adjoint methods can be employed to iteratively improve images of the Earth's interior by progressively minimizing discrepancies between observed and simulated seismograms (Akçelik *et al.* 2002, 2003; Tape *et al.* 2007). 'Adjoint tomography', a tomographic procedure based on the adjoint method, has been successively used to constrain crustal structure in southern California (Tape *et al.* 2009, 2010), as well as the upper-mantle structure of Australia (Fichtner *et al.* 2009, 2010), Europe (Zhu *et al.* 2012, 2013; Fichtner *et al.* 2013; Zhu & Tromp 2013) and the North Atlantic (Rickers *et al.* 2013).

Seismic heterogeneities investigated in this study involve 3-D variations in elasticity, anelasticity and anisotropy. Over the past several decades, most tomographic studies have focused on mapping lateral heterogeneities in (transversely) isotropic elastic wave speeds based on traveltimes of body waves, dispersion of surface waves or splitting of free oscillations. However, the propagation of seismic waves is influenced by other physical properties besides (transversely) isotropic elastic heterogeneity, such as attenuation and anisotropy. Anelastic attenuation leads to physical dispersion and dissipation, which affect both the phase and amplitude of seismic waveforms (Liu *et al.* 1976). In contrast to elastic wave speed tomography, progress in attenuation tomography has been relatively slow (Dalton *et al.* 2008), and there are significant discrepancies between anelastic models determined by different groups (Romanowicz 1995; Billien *et al.* 2000; Gung & Romanowicz 2004; Lawrence & Wysession 2006; Dalton *et al.* 2008; Wiens *et al.* 2008). Waveform amplitudes, which are usually employed to constrain attenuation, are relatively difficult to extract and are affected by a host of other factors besides intrinsic attenuation, such as earthquake magnitude, radiation pattern, elastic focusing and defocusing as well as scattering (Ruan & Zhou 2010, 2012). Therefore, in order to investigate attenuation it is preferable to simultaneously invert for elastic wave speeds and anelastic attenuation, using frequency-dependent phase and amplitude information.

Anisotropy is another important factor which affects the propagation of seismic waves. The constituent minerals of the crust and upper mantle, such as mica, amphibole and olivine, are highly anisotropic in terms of seismic wave speeds. The fast 'a' axes of these minerals are aligned with the directions of flow or principal extension depending on the state of deformation (Zhang & Karato 1995). Thus, seismic wave speeds vary with direction and polarization because of this lattice preferred orientation (LPO; Ekström & Dziewonski 1998; Simons *et al.* 2002; Gung *et al.* 2003; Marone & Romanowicz 2007; Yuan & Romanowicz 2010). Mapping these variations constrains the deformation state and history within the Earth's interior (Park & Levin 2002). Shear wave splitting measurements have been widely employed to extract azimuthal anisotropy in terms of splitting times and fast propagation directions (Silver 1996). However, owing to the relatively poor depth resolution, it is difficult to infer the vertical distribution of deformation and motion. This ambiguity has generated a long-lasting debate about the origin of seismic anisotropy within the uppermost mantle (Silver & Chan 1991; Vinnik *et al.* 1992; Silver 1996). Surface wave tomography provides an important complementary tool—with better depth resolution—for mapping azimuthal anisotropy within the crust and upper mantle (Simons *et al.* 2002; Debayle *et al.* 2005; Marone & Romanowicz 2007; Yuan & Romanowicz 2010; Lin *et al.* 2011; Endrun *et al.* 2011).

The purpose of this study is to construct a reference seismic model for the crust and upper mantle beneath Europe and the North Atlantic by fully exploiting three-component seismic waveforms and utilizing modern numerical simulations. This was one of the main goals of the Initial Training Network in computational seismology named QUEST (`quest-itn.org`), funded within the 7th Framework *People* Programme by the European Commission. In Section 2 and Section 3, we introduce the basic tectonic structure of Europe and review previous tomographic investigations in the study region. In Section 4, we discuss the distribution of earthquakes and seismographic stations used in the inversion. 3-D starting model EU_{00} is presented in Section 5. An initial source inversion, aimed at correcting any biases in routine global CMT solutions, is discussed in Section 6. A three-stage inversion strategy for constraining elastic, anelastic and anisotropic heterogeneities is discussed in Section 7, Section 9 and Section 10, respectively. In Section 11, we discuss the behaviour of various misfit functions and histograms of phase and amplitude anomalies. In Sections 12–14, various cross-sections are used to illustrate 3-D variations in elastic wave speeds, anelastic attenuation and radial and azimuthal anisotropy, and several previous surface and body-wave tomographic models are compared with model EU_{60}. Finally, the 'point-spread function' (Fichtner & Trampert 2012) is used in Section 15 to analyse resolution in the tomographic images and trade-offs between different model parameters.

2 TECTONIC SETTING

The northeastern part of the European continent is dominated by the East European Craton (EEC), which is composed of the Baltic Shield, the Ukrainian Shield and the East European Platform (Fig. 1). This region is of Archaean–Proterozoic age and has been stable over a long geological time. The Tornquist-Tesseyre Suture

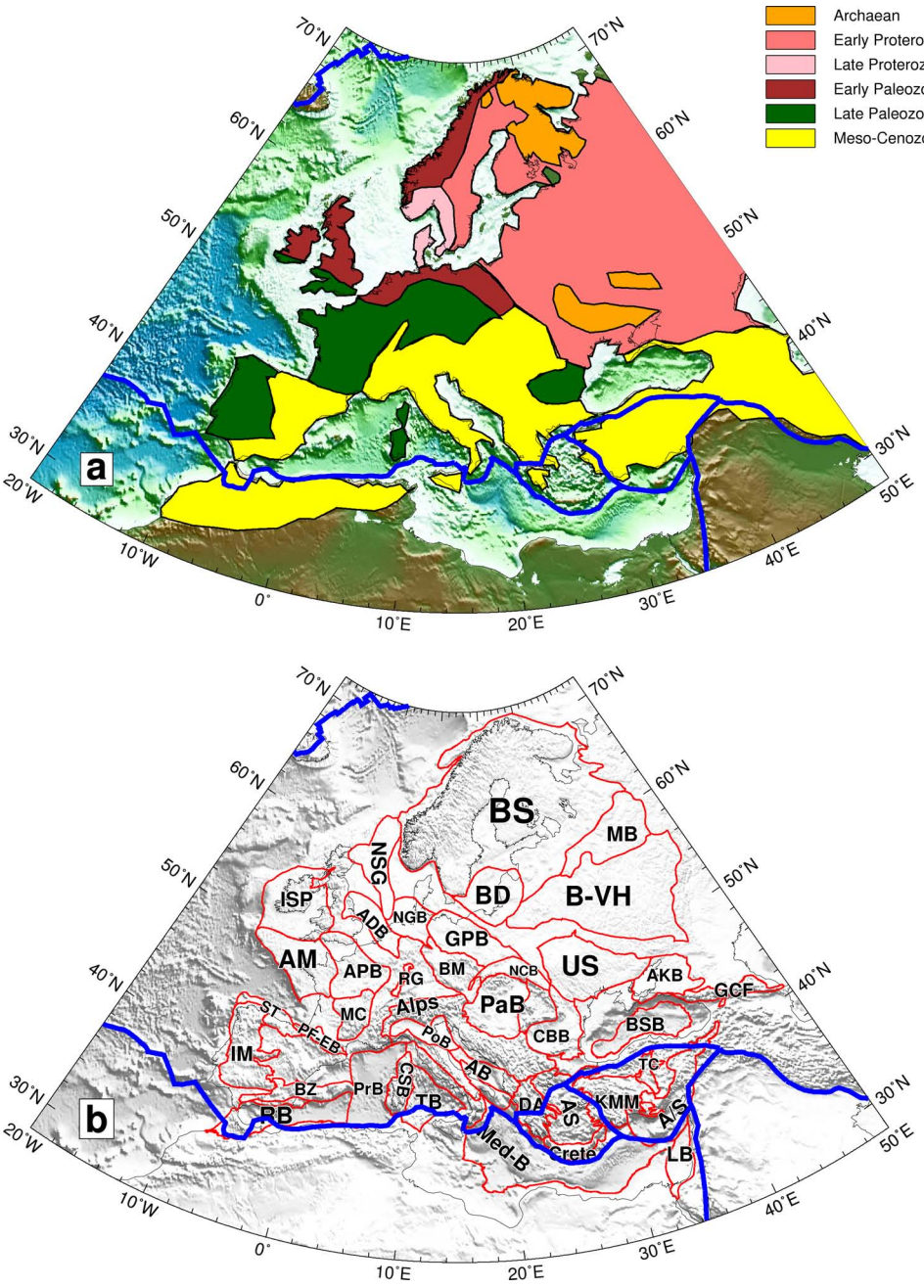

Figure 1. Tectonic structures of the European continent. (a) Simplified geological provinces of the European continent, modified from Artemieva *et al.* (2006). (b) Main tectonic structures of the European continent. Blue lines denote global plate boundaries (Bird 2003). AB, Adriatic Basin; ADB, Anglo-Dutch Basin; AKB, Azov-Kuban Basin; AM, Armorican Massif; APB, Anglo-Paris Basin; AS, Aegean Sea; A/S, Adana/Sivas; BD, Baltic Depression; BM, Bohemia Massif; BSB, Black Sea Basin; BS, Baltic Shield; B-VH, Belorussian-Voronezh High; BZ, Betic Zone; CBB, Carpathians-Balkanian Basin; CSB, Corsican-Sardinian Basin; DA, Dinaric Alps; GCF, Great Caucasus Foldbelt; GPB, German-Polish Basin; IM, Iberian Massif; ISP, Ireland-Scotland Platform; KMM, Kardiff/Menders Massif; LB, Levantine Basin; MC, Massif Central; Med-B, Mediterranean Basin; MB, Moscow Basin; NCB, North Carpathians Basin; NGB, Northwest German Basin; NSG, North Sea Graben; PaB, Pannonian Basin; PF-EB, Pyrenean Foothills-Ebro Basin; PoB, Po Basin; PrB, Provence Basin; RB, Rif Basin; RG, Rhine Graben; ST, Spanish Trough; TB, Tyrrhenian Basin; TC, Tuz/Corum; US, Ukrainian Shield.

Zone (TTSZ; Zielhuis & Nolet 1994) separates this Precambrian craton from the Phanerozoic parts of Europe. During the Palaeozoic, central and western Europe were mainly shaped by the Caledonides and Variscides Orogenies, which are currently distributed along the western coast of Scandinavia, the British Isles, Germany and France (see Fig. 1).

Mesozoic and Cenozoic tectonic activities were primarily driven by the convergency of the Eurasian and African-Arabian plates. This broad-scale compression resulted in subduction of the Tethys ocean. Numerous mountain belts with arcuate shapes, such as the Maghrebides, Apennines, Alps, Dinarides and Hellenides, were created under a tectonic background of convergence. Backarc extension accompanied by the subduction of oceanic lithosphere occurred in the Algero-Provençal Sea, Tyrrhenian Sea, Pannonian Basin and Aegean Sea. Slab roll-back and trench retreat are important mechanisms for the development of these arcuate mountain belts and

backarc basins (Wortel & Spakman 2000). For instance, the Calabrian arc migrated to its current position due to a roll-back of slabs that started approximately 30 Ma, resulting in the opening of the Algero–Provençal Basin and the Tyrrhenian Sea, as well as the rotation of the Corsica and Sardinia blocks.

The European Cenozoic Rift System (ECRS; Ziegler 1992) is responsible for extension and volcanism in western Europe during the Cenozoic. This rift system extended from northern Africa to the North Sea, and created several massifs, grabens and hotspots, for example, the Massif Central, Rhine Graben and Eifel Hotspot (Goes *et al.* 1999).

In eastern Europe, due to the closure of the Tethys, the northward moving Arabian Plate collided with the stable EEC and turned its trajectory westerly. This movement contributed to the counter-clockwise rotation of the Anatolian Plate, accommodated by strike-slip motion along the North Anatolian Fault. The main tectonic structures of the European continent are illustrated in Fig. 1. A detailed overview of the tectonic evolution of the Alpine–Mediterranean region may be found in Dercourt (1986) and Dewey *et al.* (1989).

3 PREVIOUS TOMOGRAPHIC STUDIES

Over the past several decades, body-wave tomography based on *P*- and *S*-wave arrival times has been used to construct 3-D compressional and shear wave speed models of the European continent (Spakman 1986, 1990, 1991; Spakman *et al.* 1993; Lippitsch *et al.* 2003; Piromallo & Morelli 2003; Amaru 2007; Koulakov *et al.* 2009; Mitterbauer *et al.* 2011). In spite of good constraints on deep mantle structure, teleseismic arrival times have relatively poor resolutions at shallow depths because of steeply incident rays and an uneven distribution of earthquakes and stations.

In contrast, surface wave tomography based on the dispersion of Rayleigh and Love waves provides relatively good lateral resolution on variations in shear wave speed at shallow depths. Global-scale surface wave tomographic studies constrain long-wavelength structures in Europe (Shapiro & Ritzwoller 2002; Kustowski *et al.* 2008a; Boschi *et al.* 2009), and regional-scale surface wave tomography has been employed to determine 3-D variations in smaller-scale shear wave speed variations (Boschi *et al.* 2004; Pasyanos 2005; Weidle & Maupin 2008; Schivardi & Morelli 2009, 2011). Due to the limited depth sensitivity of surface waves, deep structures in these models are poorly constrained compared to shallower features. In order to simultaneously constrain shallow and deep structures, it is preferable to jointly invert body and surface waves. Partitioned waveform inversion was developed to fit both shear body waves and surface waveforms (Nolet 1990). It has been successfully employed to construct 3-D shear wave speed models of Europe (Zielhuis & Nolet 1994; Marone *et al.* 2004; Schmid *et al.* 2008; Chang *et al.* 2010a,b). Finally, ambient noise tomography also provides good constraints on crust and uppermost mantle structure beneath the European continent (Yang *et al.* 2007).

4 DATABASE

In this study, 190 earthquakes are used to illuminate the region of interest. These earthquakes are evenly distributed along the North Atlantic Ridge (NAR) and the Mediterranean-Himalayan Belt (Fig. 2a). Most occurred between 1996 and 2011, with magnitudes ranging from 4.5 to 6.5 (Figs 2e and b). The majority of events are shallower than 30 km (Fig. 2c). Initial source parameters, for example, origin times, locations and moment tensor solutions, are collected from the global Centroid-Moment-Tensor (CMT) website (`globalcmt.org`). Before the structural inversion, we perform source inversions for all 190 earthquakes (Section 6). For this purpose, good azimuthal coverage is crucial. Three-component seismic waveforms recorded by 239 global seismographic stations (from networks operated by IRIS/IDA, IRIS/USGS, GEOFON and GEOSCOPE) are collected from the Incorporated Research Institutions for Seismology (`iris.edu`) .

In the structural inversion, seismic waveforms recorded by 338 seismographic stations of 40 European networks are collected from the Observatories and Research Facilities for European Seismology (`orfeus-eu.org`). In addition, several IRIS/PASSCAL arrays are included in the database. For example, the HOTSPOT array (Foulger *et al.* 2001; Allen *et al.* 2002a,b) is incorporated in order to illuminate suburface structures beneath Iceland and the NAR. We also assimilate some seismographic stations from the Kandilli Observatory (`koeri.boun.edu.tr`) in order to better constrain structures beneath the Anatolian Plate. In total, three-component seismic waveforms recorded by 745 seismographic stations are collected (Fig. 3a). The instrument response is removed from the raw seismic data to obtain bandpassed ground displacement, which is subsequently rotated into vertical, radial and transverse components.

5 3-D STARTING MODEL EU_{00}

Modern numerical techniques in combination with high-performance computing enable us to accurately and effectively calculate synthetic seismograms and misfit gradients in 3-D earth models. Therefore, instead of performing tomographic inversions in 1-D spherically symmetric earth models based on approximate, asymptotic methods, we are able to use a 3-D model as the starting model, and iteratively improve it using gradient-based optimization techniques. In this study, EPcrust and S362ANI are chosen as initial crust and mantle models, respectively, which are combined to form starting model EU_{00}.

5.1 3-D crustal model: EPcrust

The crust is a highly heterogeneous region of the Earth, which can strongly affect seismic wave propagation. In tomography, mantle images can be severely distorted if 3-D crustal structure is not taken into account properly (Waldhauser *et al.* 2002; Bozdag & Trampert 2008). Previous tomographic studies have frequently relied on 'crustal corrections' to remove crustal effects before imaging mantle structure (Lekic *et al.* 2010; Panning *et al.* 2010). Since SEM and adjoint methods enable us to incorporate 3-D crustal models in both forward and gradient calculations, no additional crustal corrections are required in our inversion. In addition, crust and upper-mantle structures are updated simultaneously according to the behaviour of 3-D sensitivity kernels, which helps us to reduce trade-offs between crust and mantle heterogeneity. There are a variety of crustal models for the European continent, ranging from global to local scales (Bassin *et al.* 2000; Tesauro *et al.* 2008; Molinari & Morelli 2011). Considering our simulation region, we choose EPcrust (Molinari & Morelli 2011) as our starting crustal model. EPcrust is a 3-D crustal model with $0.5^\circ \times 0.5^\circ$ resolution, covering the entire European continent as well as Greenland and the North Atlantic Ocean. EPcrust is described in terms of three layers: sediments, upper and lower crust. It provides the following seismic

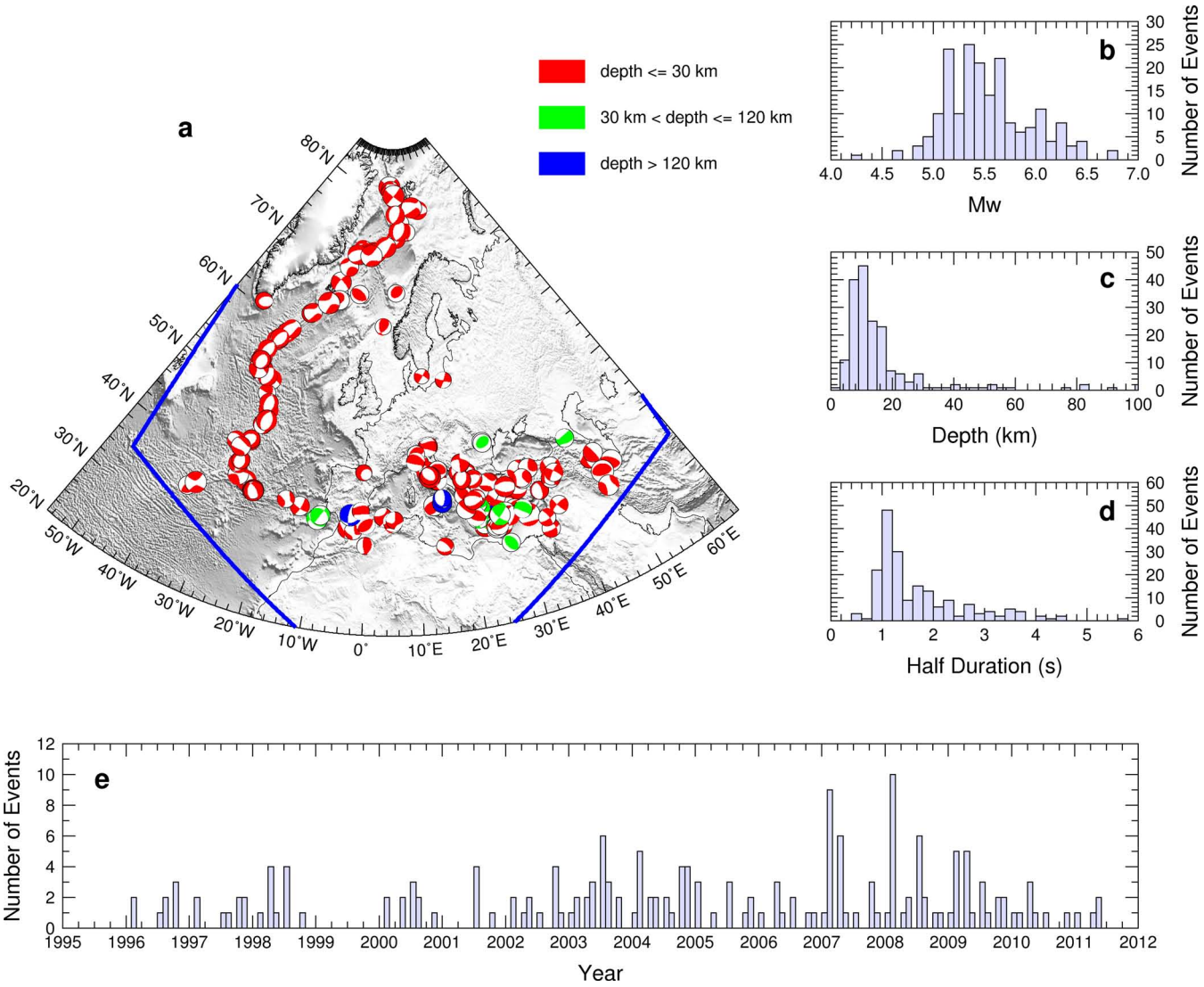

Figure 2. 190 earthquakes used in this study. (a) Distribution of earthquakes. The blue quadrilateral denotes the SEM simulation region. Colours of beach balls indicate earthquake depths. (b)–(e) Histograms of earthquake moment magnitudes, depths, half durations and origin times.

parameters: the Moho, sedimentary thickness, mass density, as well as isotropic compressional and shear wave speeds. Since EPcrust incorporates several local-scale European crustal studies, in some areas it involves much more refined structures than Crust2.0, a global crustal model with $2^\circ \times 2^\circ$ resolution (Bassin *et al.* 2000).

5.2 SEM mesh

The Moho is one of the most important boundaries within the Earth's interior. From a numerical perspective, for weak-form implementations of the equation of motion—employed in the SEM—it is very important to design a spectral-element mesh which honours such a first-order discontinuity (Komatitsch & Vilotte 1998; Komatitsch & Tromp 1999). In our study region, Moho depth varies from 7 to 10 km beneath the North Atlantic Ocean to greater than 50 km beneath the EEC (Fig. 4a). It is challenging to honour such a dramatically varying discontinuity based on a hexahedral mesh. Following Tromp *et al.* (2010), the spectral-element mesh is stretched to honour the Moho discontinuity when it is shallower than 15 km and between 25 and 45 km, dominating the distribution of Moho depth variations in the area of interest (Fig. 4b). The goal of this stretching is to employ one layer of spectral-elements to capture oceanic crust and two or three layers of spectral-elements to represent continental crust. The stretched mesh provides an adequate grid sampling for the entire crust, as illustrated in Fig. 4(d) in vertical cross-section A–A′. In Fig. 4(c), the areas in which the Moho is honoured are coloured blue (ocean) and red (continent), whereas white regions denote areas where the Moho runs through spectral elements.

The simulation domain has lateral dimensions of $65^\circ \times 65^\circ$, ranging from northern Africa to the North Pole, and from Greenland to the Urals (Fig. 2a). One 'cubed sphere' chunk of a global mesh (Komatitsch & Tromp 2002a) is used in this study. The total number of elements is 4 692 600, and the horizontal element size is approximately 42 km on the free surface, resulting in an average spacing of ∼10 km between Gauss-Lobatto-Legendre (GLL) interpolation points. The minimum period resolved by a forward calculation is ∼12 s. Using 100 cores on a Dell Intel Nehalem PC cluster, it takes approximately 1 hr for a forward simulation (a 30-min record) and 2 hr for a gradient calculation.

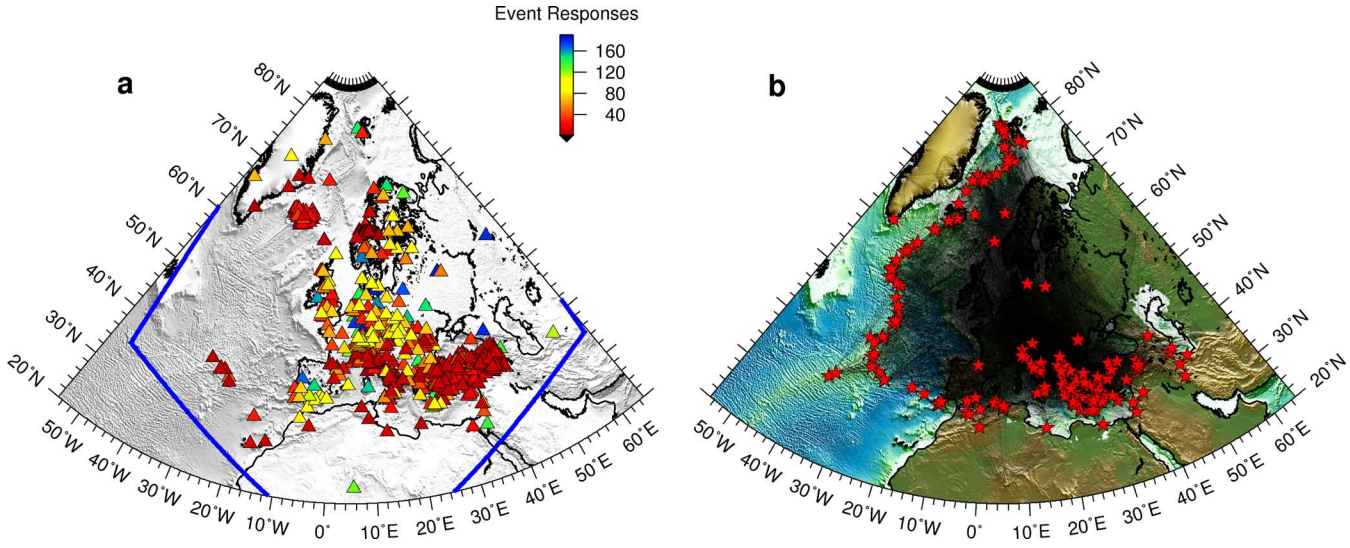

Figure 3. Distribution of seismographic stations and ray coverage map. (a) Distribution of 745 stations. Colours denote the number of events for which they contributed waveforms to the inversion. The blue quadrilateral denotes the SEM simulation region. (b) Ray coverage. Red stars denote the locations of the 190 earthquakes used in this study.

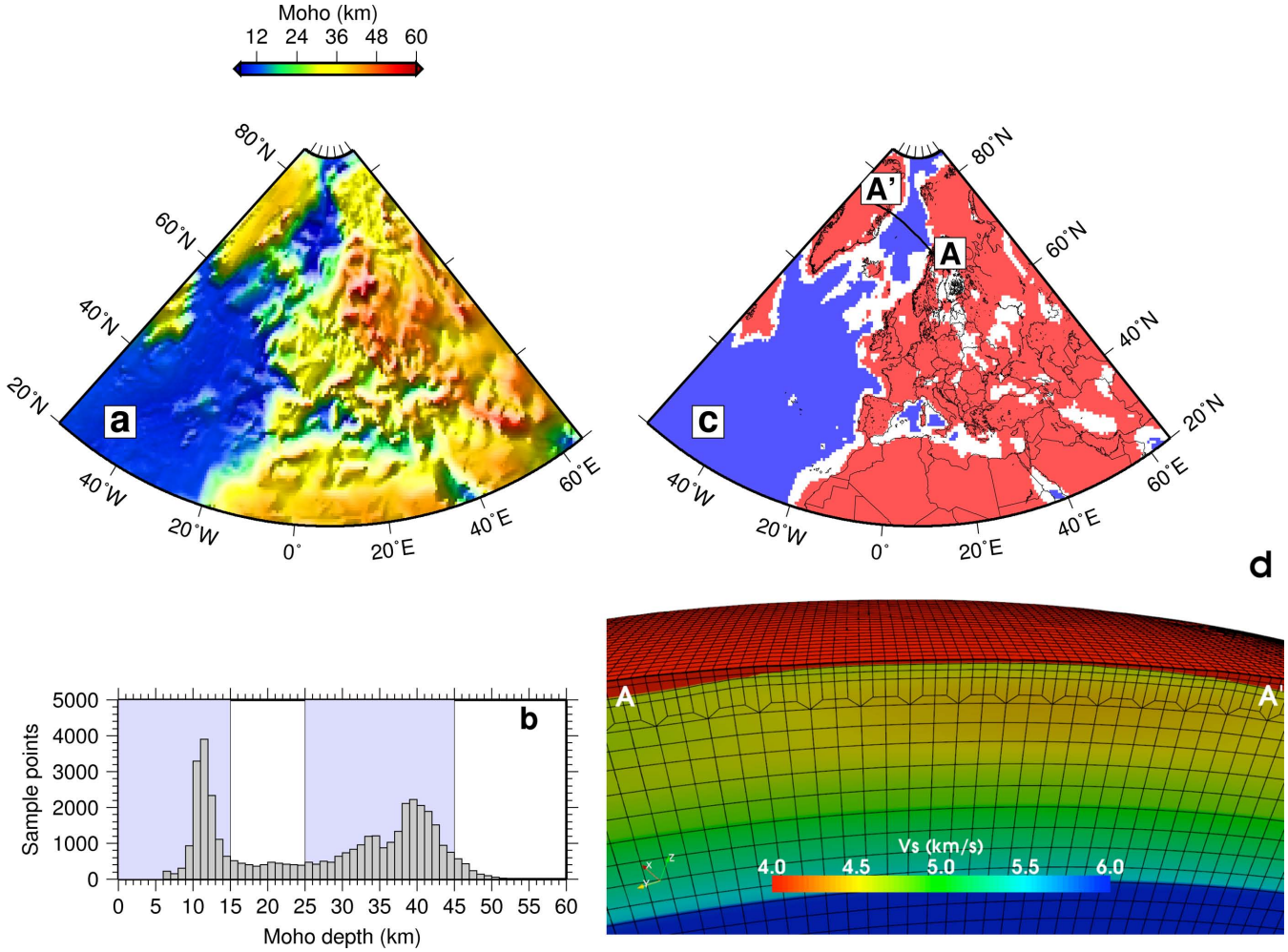

Figure 4. Spectral-element mesh for crustal model EPcrust (Molinari & Morelli 2011). (a) Moho depths of EPcrust. (b) Histogram of Moho depths. (c) Areas where the Moho is honoured by the spectral-element mesh. Red regions indicate the crust is captured by two or three layers of spectral elements. Blue regions indicate the crust is captured by one layer of spectral elements. White regions indicate the Moho is not honoured by the mesh. (d) Spectral-element mesh and isotropic shear wave speed perturbations in the starting model along cross-section A–A′ in (c).

5.3 3-D mantle model: S362ANI

S362ANI (Kustowski *et al.* 2008b), a global radially anisotropic shear wave speed model, is chosen as the 3-D starting mantle model. It provides six model parameters: mass density (ρ), two compressional wave speeds (α_v and α_h), two shear wave speeds (β_v and β_h) and a dimensionless parameter (η). Instead of describing 3-D perturbations with respect to PREM (Dziewonski & Anderson 1981) or IASP91 (Kennett & Engdahl 1991), S362ANI has its own 1-D reference model, namely STW105 (Fig. 5a). In contrast to PREM, radial anisotropy in STW105 extends to depths in excess of 300 km. Considering the complexity of the area of interest, in our inversion, radial anisotropy is allowed from the bottom of the crust to the bottom of the transition zone, that is, to depths of 660 km. The 1-D Q model is shown in Fig. 5(b), which involves a fairly strongly attenuating asthenosphere between 80 and 220 km. This 1-D attenuation model is fixed in the elastic and anisotropic inversions (Sections 7 and 10), and iteratively updated in the anelastic inversion (Section 9). Figs 5(c) and (d) show relative perturbations in β_v and β_h from S362ANI at a depth of 75 km. Since S362ANI is a global model, it only describes large-scale lateral variations, such as the old and cold EEC as well as the relatively young Mediterranean Sea and Anatolian Plate.

6 SOURCE INVERSION

Hjörleifsdóttir & Ekström (2010) estimated uncertainties associated with routine global CMT solutions. One of their conclusions is that the CMT procedure tends to locate sources deeper than their actual depths; usually this depth bias ranges from 5 to 8 km. They attributed this bias mainly to 3-D crustal heterogeneity, which is only

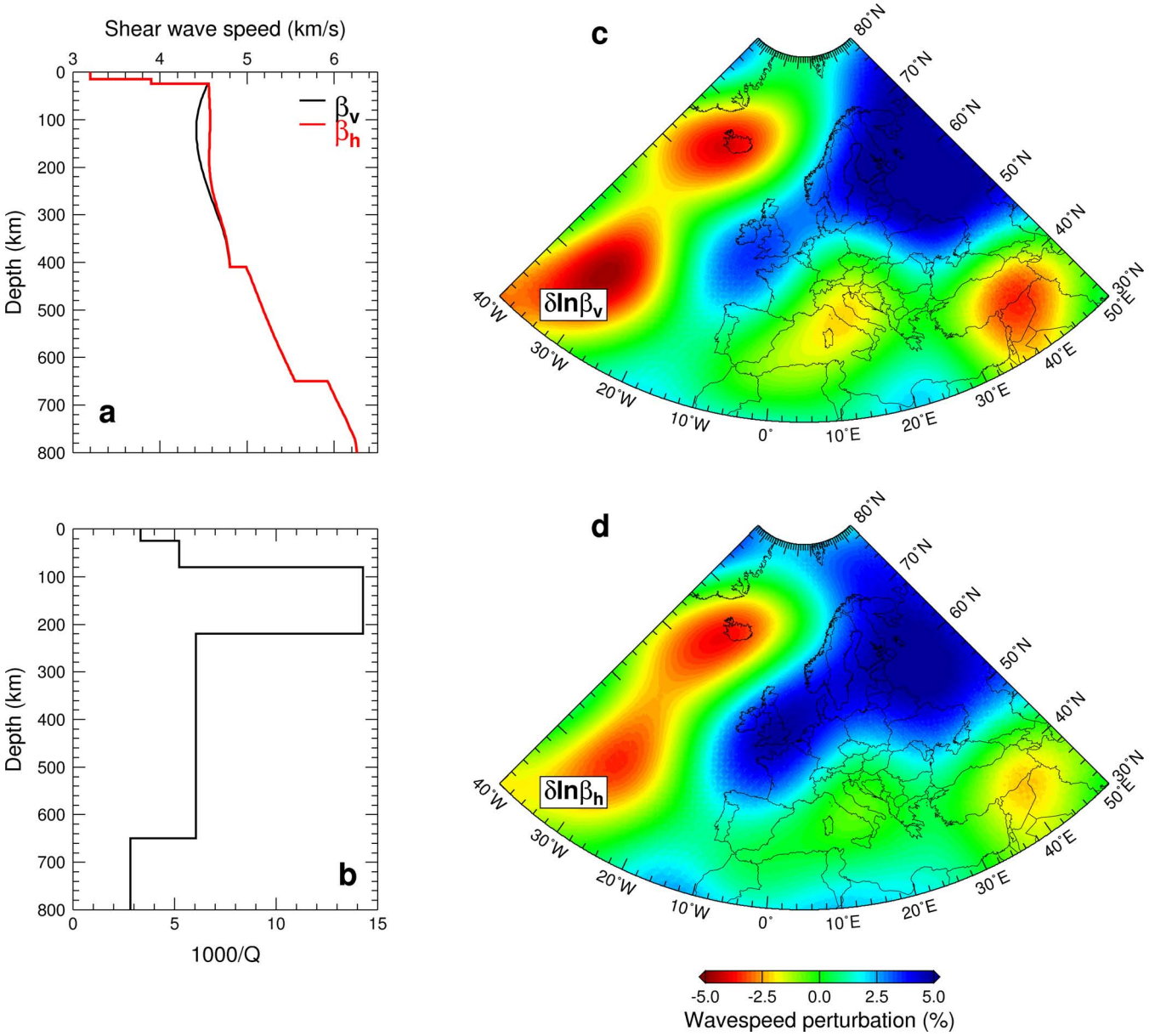

Figure 5. Shear wave speeds and attenuation for starting mantle model S362ANI (Kustowski *et al.* 2008b). (a) 1-D wave speed model STW105. Black and red lines denote speeds of horizontally travelling and vertically (β_v) and horizontally (β_h) polarized shear waves, respectively. (b) 1-D shear attenuation Q model. (c) & (d) Horizontal cross-sections of relative perturbations in β_v and β_h at 75 km depth in model S362ANI.

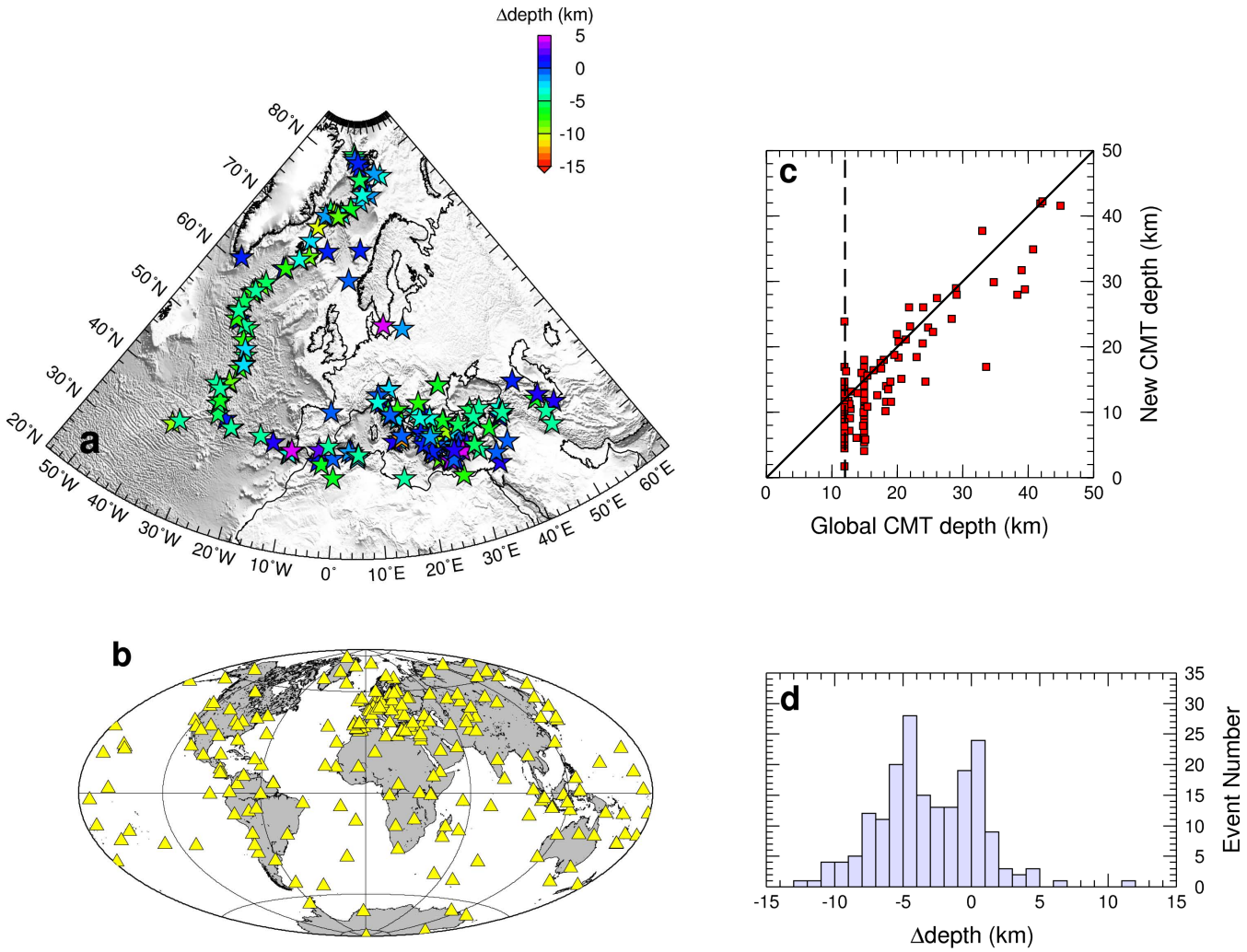

Figure 6. Depth changes for the 190 earthquakes used in this study after source inversion. (a) Map view of the depth changes for all events. (b) Distribution of global seismographic stations employed in the source inversion (see Section 4). (c) Depth comparison between global CMT solutions and reinverted sources; only events with depths shallower than 50 km are shown. (d) Histogram of depth changes, indicating a general shallowing compared to CMT depths.

approximately considered in the CMT algorithm. This systematic depth bias necessitates a source parameters inversion prior to the structural inversion.

We use the source inversion algorithm of Liu *et al.* (2004). Targeted least-squares waveform differences between observed, **d**, and simulated, **s**, seismograms are used to define the non-dimensional misfit function

$$\chi = \sum_{c=1}^{N_c} w_c \sum_{m=1}^{N_m} w_m \frac{\int [\mathbf{d}_m(t) - \mathbf{s}_m(t - \Delta t, \mathbf{m}_0)]^2 \, \mathrm{d}t}{\int [\mathbf{d}_m(t)]^2 \, \mathrm{d}t}, \qquad (1)$$

where N_c denotes the number of categories that define the misfit. In this study, three-component body waves with periods between 30 s and 80 s and three-component surface waves with periods between 80 s and 120 s are combined to constrain the source parameters, that is, $N_c = 6$. The quantity w_c represents a weighting term associated with each category, which is equal to the reciprocal of the number of measurements in each category, that is, N_m. The quantity w_m refers to a weighting factor for each measurement m, which is related to azimuth, distance and component (for details see Liu *et al.* 2004). Cross-correlation traveltime differences between observed and synthetic seismograms, Δt, are used to correct simulated seismograms. These traveltime anomalies are attributed to unmodelled lateral heterogeneities, which are the target of the structural inversion. The initial source model, $\mathbf{m}_0$, is the global CMT solution. Waveform differences are weighted by the integrated data power within each measurement window.

The SEM is used to calculate synthetic seismograms and Fréchet derivatives in the 3-D global model. Source parameters may include latitude, longitude, depth and the six components of the moment tensor. Fréchet derivatives with respect to these source parameters are calculated based on a finite-difference approximation. In order to obtain good azimuthal coverage, we use global data sets (Section 4 and Fig. 6b) and global simulations based on crustal model Crust2.0 and mantle model S362ANI. Thus, the source inversion for each event requires 10 global forward calculations (one forward calculation and nine calculations for Fréchet derivatives with respect to nine model parameters). Each simulation takes approximately 5 hr on 150 cores for a 100-min record.

The source inversion procedure routinely provides results for three combinations of model parameters: moment tensor only, moment tensor plus depth and moment tensor plus depth, longitude and latitude. In addition, each combination involves either zero-trace or double-couple constraints. Therefore, in total six scenarios are considered in the source inversion procedure. When depth, longitude

and latitude are included in the inversion, the problem becomes non-linear and should be solved iteratively (Liu *et al.* 2004). In this study, source parameters are updated only once because the changes are small compared to structural effects on phase and amplitude, and the computational cost is high. However, between the elastic and anelastic inversions of stages I & II, we perform a grid search for the scalar moment and centroid time of each earthquake to limit structural trade-offs with the source (see details in Section 8). Variance reduction is calculated for each scenario by comparing waveform differences between data and synthetics with original and updated source parameters. In principle, the solution with the largest variance reduction is chosen as the new source model. In some specific areas, for example, along the NAR, the double-couple constraint is always preferred.

After carefully selecting the best result from the six scenarios, we obtain a new source model for each event. Fig. 6 presents depth differences between updated source models and original global CMT solutions. Most new depths are shallower than the global CMT solutions, and the average depth difference is approximately 3–8 km (Figs 6c and d), in agreement with the conclusions of Hjörleifsdóttir & Ekström (2010). In Fig. 6(a), new depths of earthquakes beneath the NAR are 3–5 km shallower than global CMT solutions, which is attributed to a bias due to a uniform PREM crust (which is too thick underneath the oceans) used in the global CMT algorithm.

In the following three sections, we discuss inversion strategies for determining 3-D variations in elastic (Section 7), anelastic (Section 9) and anisotropic (Section 10) heterogeneities.

7 STAGE I: ELASTIC INVERSION

As discussed in the Introduction, our goal is to constrain elastic, anelastic and anisotropic heterogeneities in a three-stage inversion. In this section we discuss the Stage I elastic inversion.

7.1 Misfit function

In Stage I, we only use phase differences between observed and simulated seismograms to constrain elastic wave speeds. The total misfit at this stage is

$$\chi_{\rm I} = \chi^{\phi}, \tag{2}$$

where χ^{ϕ} refers to the phase misfit. Three-component short-period body waves and three-component long-period surface waves are combined to simultaneously constrain deep and shallow structures. Therefore, the total phase misfit consists of six categories: *P-SV* body waves on vertical and radial components, *SH* body waves on transverse components, Rayleigh waves on vertical and radial components and Love waves on transverse components. FLEXWIN (Maggi *et al.* 2009), an automated time window selection tool, is used to select suitable measurement windows. A multitaper technique (Laske & Masters 1996; Zhou *et al.* 2004) is used to measure frequency-dependent phase differences between data and synthetics. Thus, the phase misfit may be expressed as

$$\chi^{\phi} = \sum_{c=1}^{N_c} w_c \sum_{m=1}^{N_m} \int w_m \left[\frac{\Delta\tau_m(\omega)}{\sigma_m^{\phi}(\omega)} \right]^2 {\rm d}\omega, \tag{3}$$

where N_c denotes the number of categories, that is, $N_c = 6$, w_c is a weighting term associated with each category, and ω denotes angular frequency. Since the misfit values in each category are balanced in this study, we set w_c to the reciprocal of the number of measurements in each category, that is, N_m. The quantity w_m is a weighting term associated with each measurement. The quantities $\Delta\tau_m(\omega)$ and $\sigma_m^{\phi}(\omega)$ are angular frequency-dependent phase differences and associated uncertainties for multitaper measurement m.

7.2 Model parameters

General anisotropic materials are described by a fourth-order elastic tensor $c_{ijk\ell}$, which involves 21 independent elements. For an anisotropic material with a radially symmetric axis, the number of independent model parameters is reduced to five: A, C, L, N and F, the so-called Love parameters (Love 1927). In seismic tomography, we prefer to use wave speeds as model parameters rather than the Love parameters, because traveltime measurements are more sensitive to wave speeds. Therefore, radially anisotropic earth models, such as PREM, are usually described in terms of six parameters, as explained in Section 5.3: ρ, $\alpha_{\rm v}$, $\alpha_{\rm h}$, $\beta_{\rm v}$, $\beta_{\rm h}$ and η. The relationships between wave speeds and Love parameters are given in eq. (A1).

In our inversion, we assume that the bulk modulus, κ, remains isotropic, and that radial anisotropy is solely due to shear anisotropy, that is, the parameters L and N. In this case, we may use the isotropic bulk sound wave speed, $c = \sqrt{\kappa/\rho}$, together with the two shear wave speeds, $\beta_{\rm v} = \sqrt{L/\rho}$ and $\beta_{\rm h} = \sqrt{N/\rho}$. Thus, rather than considering four wave speeds ($\alpha_{\rm v}$, $\alpha_{\rm h}$, $\beta_{\rm v}$ and $\beta_{\rm h}$), we consider just three (c, $\beta_{\rm v}$ and $\beta_{\rm h}$), such that $\alpha_{\rm v}^2 = c^2 + \beta_{\rm v}^2$ and $\alpha_{\rm h}^2 = c^2 + \beta_{\rm h}^2$.

In general, traveltimes of seismic waves are much more sensitive to wave speeds than mass density, ρ, which is usually constrained by free oscillation and the Earth's moments of inertia. In this study, an empirical relationship between relative perturbation in mass density and isotropic shear wave speed is used to update mass density (Montagner & Anderson 1989), namely,

$$\delta\ln\rho = 0.33\,\delta\ln\beta, \tag{4}$$

where β refers to the Voigt average of the radially anisotropic shear wave speeds (Babuska & Cara 1991):

$$\beta = \sqrt{\frac{2\beta_{\rm v}^2 + \beta_{\rm h}^2}{3}}. \tag{5}$$

In summary, in the elastic inversion we consider four model parameters: the isotropic bulk sound wave speed (c), the wave speeds of horizontally travelling and vertically and horizontally polarized shear waves ($\beta_{\rm v}$ and $\beta_{\rm h}$) and the dimensionless radial anisotropic parameter (η). Perturbations in total misfit are expressed as a volume integral over relative perturbations in these four model parameters:

$$\delta\chi_{\rm I} = \int_V {\rm K}_c\,\delta\ln c + {\rm K}_{\beta_{\rm v}}\,\delta\ln\beta_{\rm v} + {\rm K}_{\beta_{\rm h}}\,\delta\ln\beta_{\rm h} + {\rm K}_{\eta}\,\delta\ln\eta\,{\rm d}V, \tag{6}$$

where V denotes the Earth's volume. The quantities ${\rm K}_{\rm c}$, ${\rm K}_{\beta{\rm v}}$, ${\rm K}_{\beta{\rm h}}$ and ${\rm K}_{\eta}$ are sensitivity kernels with respect to relative perturbations in the radially anisotropic model parameters. These four kernels can be derived from primary kernels based on relationships given in eq. (A3).

7.3 Frequency-band selection

In this study, we use a simple multiscale strategy to reduce the non-linearity of our problem. In the first several iterations, we start by fitting long-period signals, for example, body waves with periods between 15 s and 50 s and surface waves with periods between 50 s and 150 s. As the model improves and the misfit between data and synthetics diminishes, we gradually decrease the short-period corner of the bandpass filter. For instance, the short-period corner

of long-period surface wave measurements is gradually reduced from 50 s to 25 s. This strategy allows us to first resolve large-scale features based on long-period signals, and gradually map smaller-scale features based on shorter-period data. This multifrequency procedure is also used in the anelastic (Section 9) and anisotropic inversions (Section 10).

7.4 Kernel pre-conditioning

For an inverse problem, if the Hessian—that is, the second derivative of the misfit function—is available, model updates are equal to the dot product between the (generalized) inverse of the Hessian and the negative gradient (Tarantola 2005; Tape *et al.* 2007), as dictated by the Newton method. However, it is usually prohibitively expensive to compute and store the full Hessian (Chen *et al.* 2007). Therefore, a variety of pre-conditioners has been designed to approximate the Hessian or its diagonal terms (Pratt *et al.* 1998). In adjoint tomography, pre-conditioning is an important procedure since it is infeasible to access the Hessian, which would involve the very expensive calculation of 'banana-doughnut' kernels for every single measurement at every iteration. Here, we employ a pre-conditioner which involves the vector dot product and convolution of the forward and adjoint accelerations:

$$P(\mathbf{x}) = 1/\int \partial_t^2 \mathbf{s}(\mathbf{x}, t) \cdot \partial_t^2 \mathbf{s}^\dagger(\mathbf{x}, T-t)\,\mathrm{d}t, \tag{7}$$

where $\mathbf{s}$ and $\mathbf{s}^\dagger$ are the forward and adjoint displacement wavefields, respectively. Both are readily accessible during an adjoint calculation, and therefore there are no additional costs associated with the calculation of this pre-conditioner. It has a small magnitude in the vicinity of sources and receivers, allowing us to reduce relatively large magnitudes of sensitivity kernels near earthquakes and stations. In order to avoid division by a very small value, a water level is applied in the denominator of (7). Event kernels for individual earthquakes are first summed to construct misfit kernels, which are then multiplied by the pre-conditioner (7) to obtain the pre-conditioned misfit gradient.

7.5 Kernel smoothing

Regularization is generally used to stabilize an inverse problem, especially if it is ill-posed (Aster *et al.* 2005). In adjoint tomography, we do not explicitly incorporate a regularization term in the misfit function (Section 7.1). Instead, a 3-D Gaussian function is used to smooth pre-conditioned misfit gradients, which may be regarded as a regularization procedure (Tape *et al.* 2010). In classical global or regional tomography, seismic models are often parametrized by a spline function in the radial direction and spherical harmonics in the lateral directions. In this study, a 3-D Gaussian function with different half widths in the radial and azimuthal directions is employed to smooth the pre-conditioned gradients. For every GLL point $\mathbf{r}$, the kernels may be smoothed based on

$$\mathrm{K}(\mathbf{r}) = \frac{1}{W(\mathbf{r})}\int_V \mathrm{K}(\mathbf{r}')\,\exp[-(r'\Delta)^2/(2\sigma_\Delta^2)] \times \exp[-(r-r')^2/(2\sigma_r^2)]\,\mathrm{d}^3\mathbf{r}', \tag{8}$$

where Δ denotes the azimuthal distance between points $\mathbf{r}$ and $\mathbf{r}'$, $r = ||\mathbf{r}||$, $r' = ||\mathbf{r}'||$ and where $W(\mathbf{r})$ is the normalization factor

$$W(\mathbf{r}) = \int_V \exp[-(r'\Delta)^2/(2\sigma_\Delta^2)]\,\exp[-(r-r')^2/(2\sigma_r^2)]\,\mathrm{d}^3\mathbf{r}'. \tag{9}$$

The quantities σ_Δ and σ_r are the half widths of the Gaussian function in the azimuthal and radial directions, respectively. They are chosen based on the wavelengths of structure which can be resolved by the current frequency bands. These two widths are reduced as the model improves. For instance, σ_Δ is gradually decreased from 100 to 50 km, reflecting the incorporation of shorter-period surface wave measurements.

7.6 Conjugate gradient method

Gradient-based optimization approaches are used to iteratively update model parameters based on pre-conditioned and smoothed misfit gradients (Tromp *et al.* 2005; Tape *et al.* 2007). In the first iteration, the steepest descent method is used to update the model parameters, that is, search directions are equal to negative misfit gradients. In subsequent iterations, a conjugate gradient method is used to compute search directions (Fletcher & Reeves 1964), which are equal to combinations of current misfit gradients and previous search directions:

$$\mathbf{d}_i = -\mathbf{g}_i + \beta \mathbf{d}_{i-1}, \tag{10}$$

where

$$\beta = \frac{\mathbf{g}_i^T \cdot (\mathbf{g}_i - \mathbf{g}_{i-1})}{\mathbf{g}_{i-1}^T \cdot \mathbf{g}_{i-1}}. \tag{11}$$

The search direction is denoted by $\mathbf{d}$, and $\mathbf{g}$ refers to the misfit gradient. There are alternative formulae to compute β (Tarantola 2005). In this study, the Polak-Ribière formula is employed (Tromp *et al.* 2005; Tape *et al.* 2007, 2010). One advantage of this formula is that β can be reset to zero when it is negative. Thus, the algorithm is able to automatically forget previous search directions and restart as a new steepest descent search. We have also experimented with quasi-Newton optimization methods, specifically the limited-memory version of the BFGS algorithm (L-BFGS; Matthies & Strang 1979; Nocedal 1980). These methods give comparable results and convergence rates (Luo *et al.* 2013).

7.7 Line search and model update

A model update is obtained based on the expression

$$\ln \frac{\mathbf{m}_{i+1}}{\mathbf{m}_i} = \alpha\,\mathbf{d}_i, \tag{12}$$

where α denotes a step length in the ith search direction $\mathbf{d}_i$. Tape *et al.* (2007) used quadratic and cubic interpolations to determine the optimal step length along the search direction. In this study, using a representative subset of earthquakes, we generate several test models by choosing different values of α, compute synthetic seismograms and evaluate the misfit function for each test model. The model with the minimum misfit value is selected as the new model. We usually test five α values ranging from 0.01 to 0.05, corresponding to maximum 1–5 per cent model perturbations along the search direction. The behaviour of the total misfit, as well as its behaviour in the six subcategories (discussed in Section 7.1), are monitored to determine the best test model. Fig. 7 shows how we select the value of α in the first iteration based on this strategy.

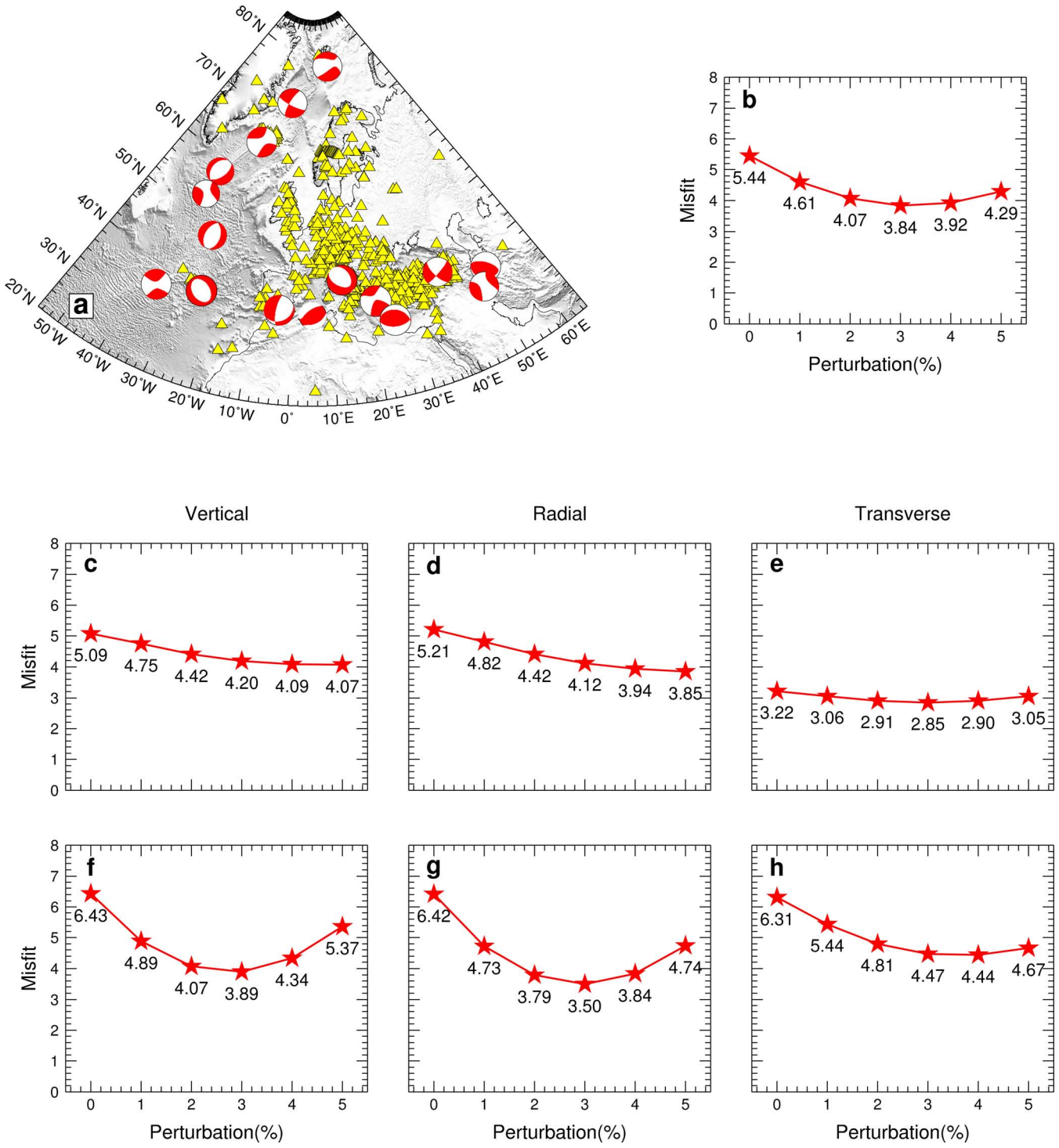

Figure 7. Line search to determine the step length in the first iteration. (a) Distribution of 16 earthquakes used in the line search. Yellow triangles denote seismographic stations. (b) Evolution of total misfit for the five test models with different α values ranging from 0.01 to 0.05. (c)–(e) Evolution of misfit for *P-SV* body waves on vertical (c) and radial (d) components, and *SH* body waves on transverse components (e). (f)–(h) Evolution of misfit for Rayleigh waves on vertical (f) and radial (g) components, and Love waves on transverse components (h). A 3 per cent update is chosen in the first iteration.

Since it is very expensive to perform forward calculations for all 190 earthquakes, we use a representative subset of 16 earthquakes (Fig. 7a). Therefore, for each line search, we perform another 16×5 forward calculations.

The inversion strategy and computational costs of adjoint tomography for elastic iterations are summarized in Fig. 8. Intensive computations are required for the source inversions, forward and adjoint calculations. In comparison, pre- and post-processing requires a limited amount of storage and computation. A new elastic wave speed model for the crust and upper mantle beneath Europe and the North Atlantic, named EU_{30}, is constructed based on 30 preconditioned conjugate gradient iterations, which required more than 17 100 wavefields simulations and 2.3 million central processing unit hours (Zhu *et al.* 2012).

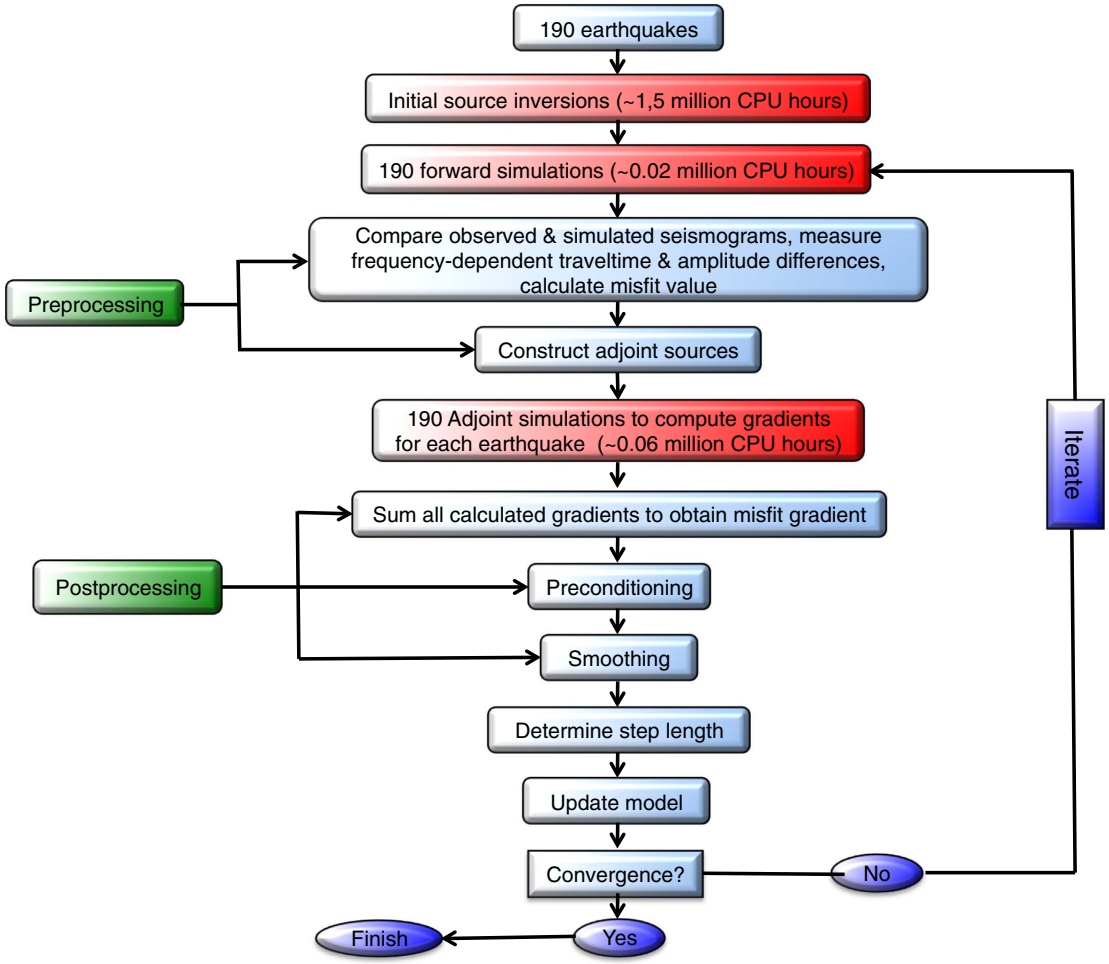

Figure 8. Adjoint tomography workflow for the Stage I elastic inversion. Computational requirements for the source inversion as well as for forward and adjoint calculations are indicated.

8 SOURCE CORRECTION

Before the Stage II anelastic inversion, we perform a grid search for the scalar moment, M_0, and centroid time, t_0, of each earthquake to minimize trade-offs between structure and source. For every event, we define a 2-D misfit function of corrections in origin time, Δt_0, and relative scalar moment, $\Delta \ln M_0$. Phase and amplitude anomalies between observed seismograms, $\mathbf{d}(t)$, and corrected synthetic seismograms, $(1 + \Delta \ln M_0)\,\mathbf{s}(t + \Delta t_0)$, are combined in the misfit function. A simple grid search is used to determine a pair of corrections which minimizes the misfit. The search dimensions are $-5\,s < \Delta t_0 < 5\,s$ and $-0.8 < \Delta \ln M_0 < 0.8$. An example of the source correction for event 200905241617A is shown in Fig. 9. The mean values of both phase and amplitude histograms are moderately improved after the source corrections. However, there are no significant improvements in their standard deviations, which must be reduced by considering 3-D anelastic heterogeneity.

9 STAGE II: ANELASTIC INVERSION

After 30 iterations, we have significantly reduced the phase anomalies between observed and simulated seismograms. However, there are no obvious improvements in the amplitude anomalies (see Section 11.2). Anelastic attenuation is an important factor affecting the amplitudes of waveforms. Therefore, at this stage we combine amplitude anomalies with remaining phase differences to simultaneously constrain elastic and anelastic heterogeneities. Anelasticity involves physical dispersion and attenuation, and therefore we fit phase and amplitude anomalies simultaneously to limit trade-off between elasticity and anelasticity. Model EU_{30} from the previous elastic inversion is chosen as the starting model for the current stage.

The same inversion strategy as described in Section 7 and Fig. 8 is used to simultaneously constrain elastic and anelastic heterogeneity. 20 additional pre-conditioned conjugate gradient iterations are performed to construct a new anelastic model of Europe and the North Atlantic, namely EU_{50}, which required more than 18 050 wavefields simulations and 2.5 million central processing unit hours.

9.1 Misfit function

The total misfit function of Stage II is defined as

$$\chi_{\mathrm{II}} = w_\phi\, \chi^\phi + w_A\, \chi^A, \tag{13}$$

where χ^ϕ and χ^A refer to phase and amplitude misfits, and w_ϕ and w_A are weighting factors associated with these two misfits, which are chosen to balance the relative contributions of phase and amplitude.

As in Section 7.1, time windows are selected with FLEXWIN (Maggi *et al.* 2009), and in these windows multitaper phase and amplitude anomaly measurements are made (Laske & Masters 1996;

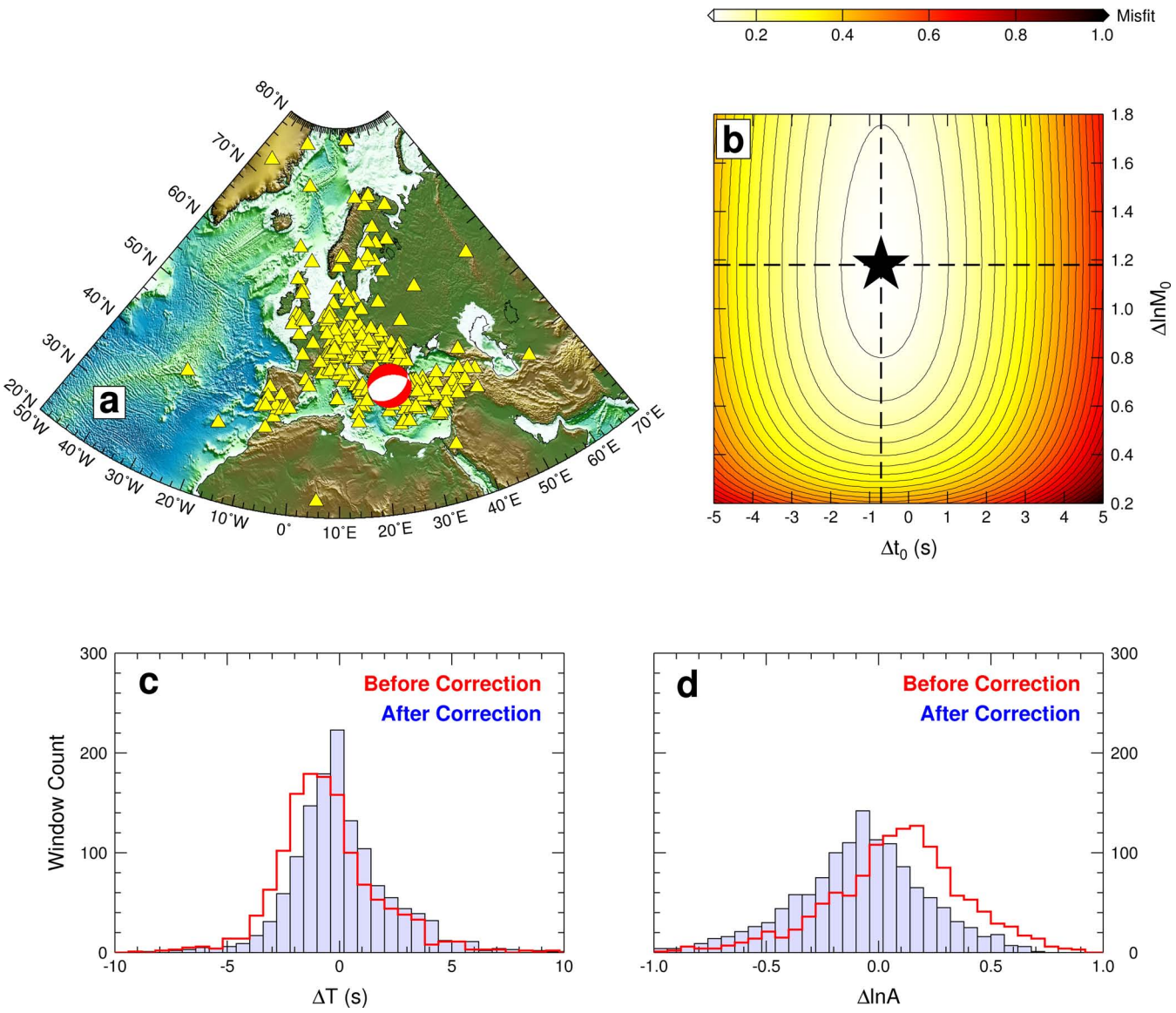

Figure 9. Example of a source correction for event 200905241617A. (a) Locations of event (indicated by the beach ball) and stations (indicated by yellow triangles) used in the source correction. (b) 2-D misfit function as a function of centroid time, Δt_0, and relative scalar moment, $\Delta \ln M_0$. The black star denotes the pair of parameters used to correct the source. (c) & (d) Histograms of phase anomalies (c) and amplitude anomalies (d) before (red) and after (blue) the source correction.

Zhou *et al.* 2004). Phase and amplitude misfits χ^{ϕ} and χ^{A} are calculated based on the expressions

$$\chi^{\phi} = \sum_{c=1}^{N_c} w_c \sum_{m=1}^{N_m} \int w_m \left[\frac{\Delta \tau_m(\omega)}{\sigma_m^{\phi}(\omega)} \right]^2 \mathrm{d}\omega, \tag{14}$$

$$\chi^{A} = \sum_{c=1}^{N_c} w_c \sum_{m=1}^{N_m} \int w_m \left[\frac{\Delta \ln A_m(\omega)}{\sigma_m^{A}(\omega)} \right]^2 \mathrm{d}\omega, \tag{15}$$

where $\Delta \ln A_m(\omega)$ and $\sigma_m^A(\omega)$ are frequency-dependent amplitude differences and associated uncertainties for multitaper measurement m. The other parameters are the same as defined in Section 7.1. Both short-period body waves and long-period surface waves in three-component seismograms are combined in the misfit functions to simultaneously constrain deep and shallow structures. Thus, the number of categories N_c in eqs (14) and (15) equals six, that is, *P-SV* body waves on vertical and radial components, *SH* body waves on transverse components, Rayleigh waves on vertical and radial components and Love waves on transverse components.

Based on the definition of the total misfit function in eq. (13), the corresponding adjoint sources are

$$f^{\dagger} = f_{\phi}^{\dagger} + f_{A}^{\dagger}, \tag{16}$$

where $f_{\phi}^{\dagger}$ represents the adjoint source for the phase misfit defined in eq. (14) and $f_{A}^{\dagger}$ denotes the contribution related to the amplitude misfit defined in eq. (15).

9.2 Model parameters

The model parameters at the current stage include elastic and anelastic contributions. For the elastic part, we use the same radially anisotropic model parameters as defined in Section 7.2, including the wave speeds of horizontally travelling and vertically and horizontally polarized shear waves β_{v} and β_{h}, the isotropic bulk sound

wave speed c, and the dimensionless parameter η. As before, mass density is scaled to the isotropic shear wave speed via eq. (4).

The inverse quality factor Q^{-1} is used to quantify anelasticity. Since the magnitude of bulk attenuation Q_κ^{-1} is usually negligible compared to shear attenuation Q_μ^{-1} (Dalton *et al.* 2008), only shear attenuation is considered in this study. For brevity, in the following expressions, we use Q^{-1} rather than Q_μ^{-1} to denote the inverse shear quality factor.

Thus, the five model parameters considered in this stage are c, β_v, β_h, η and Q^{-1}. The Stage II misfit variation $\delta\chi_\mathrm{II}$ may be expressed as

$$\delta\chi_\mathrm{II} = \int_V \mathrm{K}_c\,\delta\ln c + \mathrm{K}_{\beta_\mathrm{v}}\,\delta\ln\beta_\mathrm{v} + \mathrm{K}_{\beta_\mathrm{h}}\,\delta\ln\beta_\mathrm{h} + \mathrm{K}_\eta\,\delta\ln\eta + \mathrm{K}_{Q^{-1}}\,\delta Q^{-1}\,\mathrm{d}V, \tag{17}$$

where K_c, $\mathrm{K}_{\beta_\mathrm{v}}$, $\mathrm{K}_{\beta_\mathrm{h}}$, K_η and $\mathrm{K}_{Q^{-1}}$ are sensitivity kernels for the elastic and anelastic model parameters. Absolute perturbations in Q^{-1} are used in (17) in order to balance the relative contributions of elastic and anelastic gradients.

9.3 Anelastic kernels

Liu *et al.* (1976) demonstrated that for an absorption band solid the complex, frequency-dependent shear modulus may be expressed as

$$\mu(\omega) = \mu(\omega_0)[1 + (2/\pi)\,Q^{-1}\ln(|\omega|/\omega_0) + i\,\mathrm{sgn}(\omega)Q^{-1}], \tag{18}$$

where ω_0 denotes a reference angular frequency. The second and third contributions on the right-hand side of eq. (18) denote the effects of physical dispersion and dissipation due to anelasticity, respectively.

Upon perturbing both sides of eq. (18), we obtain a relationship between $\delta\mu$ and δQ^{-1}, namely

$$\delta\mu(\omega) = \mu(\omega_0)[(2/\pi)\ln(|\omega|/\omega_0) + i\,\mathrm{sgn}(\omega)]\,\delta Q^{-1}. \tag{19}$$

Tromp *et al.* (2005) showed that the same expression for calculating the shear modulus sensitivity kernel, that is, K_μ, may be employed to determine the shear attenuation sensitivity kernel

$$\mathrm{K}_{Q^{-1}} = -\int_0^T 2\mu\,\mathbf{D}^\dagger(T-t) : \mathbf{D}(t)\,\mathrm{d}t, \tag{20}$$

where $\mathbf{d}$ and $\mathbf{d}^\dagger$ denote the forward and adjoint traceless strain deviators, respectively. In this case, however, the adjoint wavefield used to calculate the shear attenuation kernel is determined by the anelastic adjoint source

$$\tilde{f}_i^\dagger(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}[(2/\pi\,\ln(|\omega|/\omega_0 - i\,\mathrm{sgn}(\omega)]^*\,f_i^\dagger(\omega)\exp(i\omega t)\,\mathrm{d}\omega, \tag{21}$$

where $f^\dagger$ denotes the elastic adjoint source. Therefore, two separate adjoint simulations are required to simultaneously determine the elastic and anelastic kernels.

10 STAGE III: ANISOTROPY INVERSION

As described in the Introduction, anisotropy is another important factor affecting the propagation of seismic waves. In this stage, we use long-period surface waves to map both radial and azimuthal anisotropy in the crust and upper mantle beneath Europe and the North Atlantic.

10.1 Misfit function

Similar to eq. (13), phase and amplitude anomalies are combined in the Stage III total misfit:

$$\chi_\mathrm{III} = w_\phi\chi^\phi + w_A\chi^A, \tag{22}$$

where χ^ϕ and χ^A are frequency-dependent phase and amplitude misfits as defined in eqs (14) and (15), respectively. However, only three categories are used in this stage, namely, Rayleigh waves on vertical and radial components and Love waves on transverse components.

10.2 Model parameters

Smith & Dahlen (1973) and Montagner & Nataf (1986) showed that in a weakly anisotropic medium the phase speed of surface waves, c, is a function of both angular frequency, ω, and azimuth, θ, and may be expressed as the Fourier series

$$c(\omega,\theta) = c_0(\omega) + c_1(\omega)\cos(2\theta) + c_2(\omega)\sin(2\theta) + c_3(\omega)\cos(4\theta) + c_4(\omega)\sin(4\theta), \tag{23}$$

where radial anisotropy is captured by c_0, which involves combinations of mass density ρ and the five Love parameters: A, C, L, N and F, as described in Section 7.2. The functions c_1 and c_2 are 2θ-dependent components, and are combinations of $G_{c,s}$, $H_{c,s}$ and $B_{c,s}$ in surface wave tomography (Smith & Dahlen 1973; Montagner & Nataf 1986), where subscripts c and s denote cosine and sine dependence, respectively. The functions c_3 and c_4 are 4θ-dependent components, and are combinations of $E_{c,s}$. Thus, it requires 13 model parameters to describe surface wave anisotropy. However, in practical applications one cannot resolve all 13 parameters. Most surface wave tomographic studies focus on mapping lateral variations in two radially anisotropic parameters, namely, L and N, and two azimuthally anisotropic parameters, namely, G_c and G_s. We adopt the same strategy at this stage to map radial and azimuthal anisotropy in the crust and upper mantle.

The perturbation in the total misfit may be expressed as

$$\begin{aligned}\delta\chi_\mathrm{III} &= \int_V \mathrm{K}_L\delta L + \mathrm{K}_N\delta N + \mathrm{K}_{G_c}\delta G_c + \mathrm{K}_{G_s}\delta G_s\,\mathrm{d}V,\\ &= \int_V \mathrm{K}_{\beta_\mathrm{v}}\,\delta\ln\beta_\mathrm{v} + \mathrm{K}_{\beta_\mathrm{h}}\,\delta\ln\beta_\mathrm{h} + \mathrm{K}_{G_c'}\,\delta G_c' + \mathrm{K}_{G_s'}\,\delta G_s'\,\mathrm{d}V,\end{aligned} \tag{24}$$

where K_L, K_N, K_{G_c} and K_{G_s} are sensitivity kernels for the four model parameters L, N, G_c and G_s, whereas $\mathrm{K}_{\beta_\mathrm{v}}$, $\mathrm{K}_{\beta_\mathrm{h}}$, $\mathrm{K}_{G_c'}$ and $\mathrm{K}_{G_s'}$ are sensitivity kernels for the dimensionless model parameters $\delta\ln\beta_\mathrm{v}$, $\delta\ln\beta_\mathrm{h}$, G_c' and G_s'. The dimensionless parameters G_c' and G_s' are defined as

$$G_c' = G_c/(\rho\beta_0^2), \tag{25}$$

$$G_s' = G_s/(\rho\beta_0^2), \tag{26}$$

where β_0 denotes the isotropic shear wave speed in the 1-D reference model. Relationships between these kernels and the primary kernels may be found in eqs (A5) and (A6).

We use the same pre-conditioned conjugate gradient approach to iteratively update these four anisotropic model parameters. Isotropic shear wave speed β and the radially anisotropic model parameter ξ

may be derived based on the updated L and N Love parameters as follows:

$$\beta = \sqrt{(2L + N)/3\rho}, \tag{27}$$

$$\xi = N/L. \tag{28}$$

The direction of the fast anisotropic axis, ζ, and the strength of the azimuthal anisotropy, G_0, may be calculated based on G_c and G_s via

$$G_0 = \sqrt{G_s^2 + G_c^2}, \tag{29}$$

$$\zeta = \frac{1}{2}\arctan(G_s/G_c). \tag{30}$$

10 additional pre-conditioned conjugate gradient iterations are performed to construct a new anisotropic model of Europe and the North Atlantic, named EU_{60}, which required more than 5700 wavefields simulations and 0.8 million central processing unit hours.

11 IMPROVEMENTS IN MISFITS AND HISTOGRAMS

11.1 Misfit function evolution

Fig. 10 shows the evolution of the phase misfit during the three-stage inversion. During the first 50 iterations, the total misfit involves six categories (see Sections 7.1 and 9.1). As described in Section 7.3, the short-period corner of the surface wave bandpass filter is gradually reduced from 50 s to 25 s to progressively resolve smaller-scale structures. The period range of body waves is changed from 15–50 s to 15–40 s at iteration 4 and fixed in subsequent iterations. The total misfit—as well as the misfit in each of the six subcategories—is gradually reduced over the 50 iterations, except for several slight increases when the short-period corner is reduced. For example, at iteration 4 the total misfit increases from 2.84 to 3.19 due to the incorporation of 40 s surface waves. Thus, it is important to recognize that the 'misfit function' is a moving target that changes every several iterations. The increase in misfit at iteration 18 is due to the assimilation of several data sets recorded by stations from temporary IRIS/PASSCAL experiments and the Kandilli Observatory. Over the last 10 iterations, only surface waves are employed to constrain radial and azimuthal anisotropy in the crust and upper mantle (Section 10.1). Therefore, we only monitor the behaviour of the misfit for three-component surface waves.

Amplitude measurements are used in the inversion after 30 elastic iterations. In Fig. 11, we monitor the behaviour of the amplitude misfits from iteration 30 to 60. Misfits in the first 20 anelastic iterations involve six categories, that is, three-component body- and surface wave seismograms. Over the last 10 anisotropic iterations, there are only three categories, that is, just three-component surface waves. The overall amplitude misfit and the contributions from each of its subcategories are gradually reduced over the 30 iterations.

11.2 Comparisons of histograms

In this section, we compare phase and amplitude histograms for starting model EU_{00} and final model EU_{60}. Phase histograms in all six categories are summarized in Fig. 12. Compared with starting model EU_{00}, both mean values and standard deviations of phase histograms for the final model are significantly improved, for example, for Rayleigh waves on the vertical component the mean value is reduced from −1.36 to 0.02 and the standard deviation is reduced from 3.93 to 2.40. These improvements demonstrate that synthetic seismograms based on the final model EU_{60} are able to simultaneously match observed short-period body waves and long-period surface waves.

In Fig. 13, we compare amplitude histograms for the starting and final models. Similar to the phase histograms shown in Fig. 12, amplitude histograms in all six categories are improved after the last 30 iterations. For instance, for Rayleigh waves on the vertical component the mean value and standard deviation are reduced from 0.07 to 0.05 and 0.36 to 0.29, respectively. In contrast, we compare amplitude histograms for starting model EU_{00} and model EU_{30} after the elastic inversion in Fig. 14. There are no significant improvements in the amplitudes after 30 elastic iterations, because only phase measurements are used to constrain elastic wave speeds in Stage I (Section 7). Amplitudes for some specific paths may be improved due to focusing and defocusing effects, however, in order to significantly reduce amplitude anomalies, shear attenuation Q^{-1} has to be considered in the structural inversion.

In the following three sections, we present lateral variations in elastic wave speeds (Section 12), anelastic attenuation (Section 13) and azimuthal anisotropy (Section 14) of model EU_{60}.

12 ELASTIC WAVE SPEEDS

12.1 Isotropic shear wave speed

In Fig. 15, we present lateral variations in relative isotropic shear wave speed in model EU_{60}. Model STW105 (Kustowski *et al.* 2008b), a transversely isotropic model shown in Fig. 5, is used as a 1-D reference model. Long-wavelength structures at shallow depths are consistent with previous surface wave models (Shapiro & Ritzwoller 2002; Boschi *et al.* 2004; Kustowski *et al.* 2008a; Chang *et al.* 2010a; Schivardi & Morelli 2011). The EEC is characterized by faster-than-average wave speeds (> 4 per cent) down to depths in excess of 250 km, representing a cold and old continental lithospheric lid. The TTSZ is imaged as a sharp boundary between the EEC and western and central Europe down to depths greater than 250 km (Zielhuis & Nolet 1994). The Alpine–Himalaya orogenic belt, starting from the western Mediterranean, continuing through the Pannonian Basin, and extending to Anatolia, is resolved as a slow wave speed anomaly at depths shallower than 150 km.

Within the mantle transition zone, central Europe is dominated by fast wave speed anomalies (> 3 per cent), which are related to slab roll-back associated with the Apennines-Calabrian-Maghrebides, Carpathians-Vrancea-Adria and Hellenic-Cyprus arcs (Zhu *et al.* 2012). These fast wave speed anomalies are in excellent agreement with results from body-wave traveltime tomography (Wortel & Spakman 2000; Piromallo & Morelli 2003). Since in our inversion, short-period body waves and long-period surface waves are combined to simultaneously constrain deep and shallow structures, there is generally good agreement between model EU_{60} and previous surface wave inversions at shallow depths and body-wave images at greater depths (for details see Section 12.5).

Numerous short-wavelength model features are resolved in EU_{60}. Most of these are missing in large-scale surface wave tomographic models because of their inherently limited lateral resolution. Fast wave speed anomalies associated with the Central Graben and the Armorican Massif are well imaged down to depths in excess of 200 km. The Alps, Adriatic and Hellenides are imaged as a continuous belt with fast wave speeds down to 200 km. At greater depths,

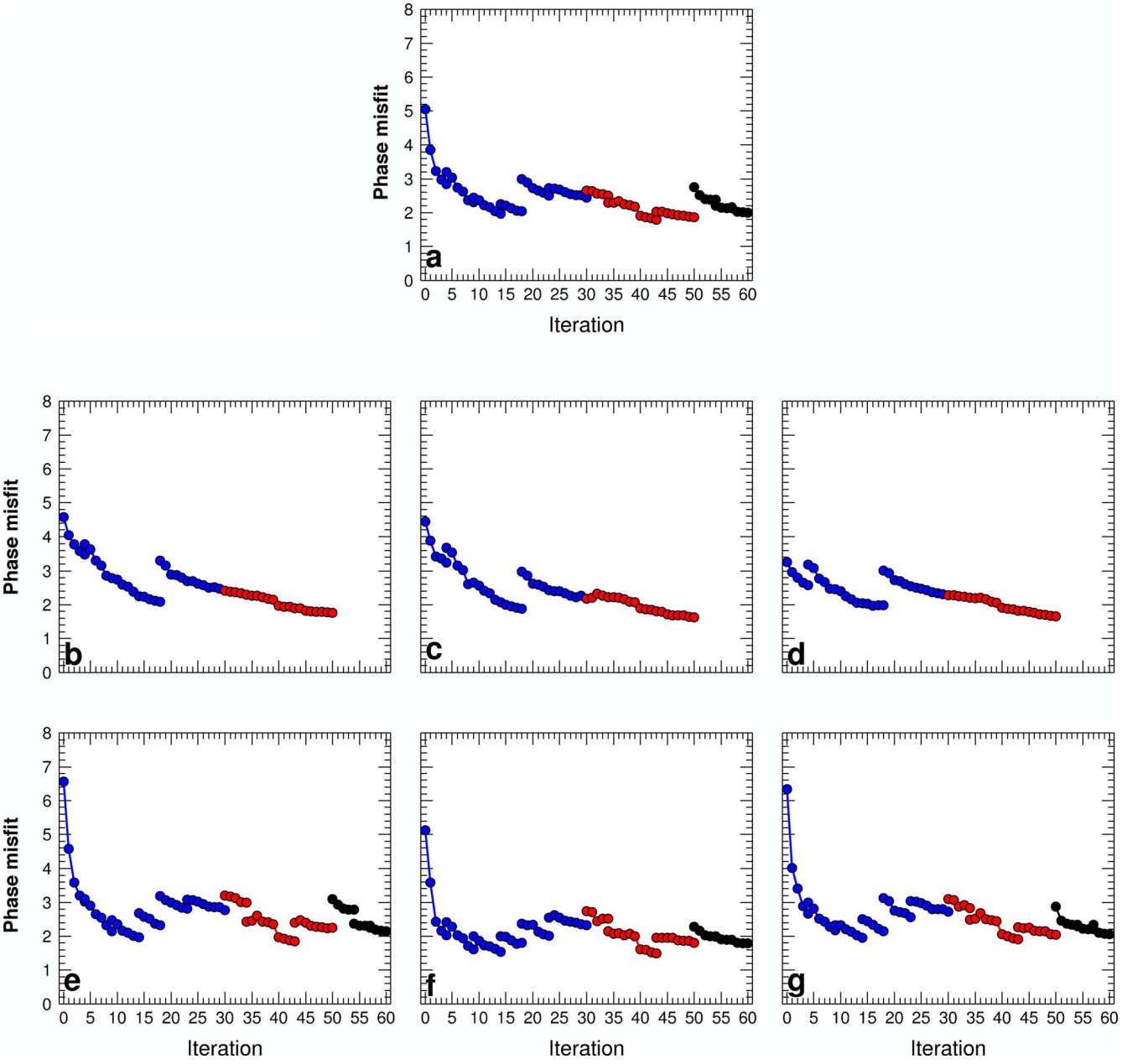

Figure 10. Evolution of phase misfits during the three-stage inversion. (a) Evolution of the total phase misfit, where blue, red and black dots label elastic, anelastic and anisotropic iterations, respectively. (b)–(d) Evolution of the phase misfit for *P-SV* body waves on vertical (b) and radial (c) components, and *SH* body waves on transverse components (d). (e)–(g) Evolution of the phase misfits for Rayleigh waves on vertical (e) and radial (f) components, and Love waves on transverse components (g). A simple multiscale strategy is used in the iterations, that is, the short-period corner of the bandpass filter is gradually reduced as the models improve, allowing us to steadily resolve smaller-scale structures (see Section 7.3). Note that the 'misfit function' is a moving target that changes every several iterations. The important observation is that after each change, the new misfit is gradually reduced in all six categories.

this continuous belt disappears in most places, except the Hellenic arc, where the subducting slab is clearly traceable down into the lower mantle. At depths greater than 250 km, a localized fast wave speed anomaly related to the Calabrian arc is resolved. Within the mantle transition zone, there are two fast wave speed anomalies. The one beneath the western Mediterranean is related to slab roll-back associated with the Apennines-Calabrian-Maghrebides arc. The second one beneath central Europe is related to subducting slabs associated with the Carpathians-Vrancea-Adria and Hellenic-Cyprus arcs (Zhu *et al.* 2012). Small-scale slow wave speed anomalies related to the ECRS (Ziegler 1992), such as the Massif Central, Eifel Hotspot, Bohemian Massif and Central Slovakia Volcanic Field, are mapped at depths shallower than 300 km.

12.2 Radial anisotropy

Both vertically (β_v) and horizontally (β_h) polarized shear wave speeds are considered in this study. Thus, radial anisotropy may be captured based on the parameter

$$\xi = (\beta_h/\beta_v)^2. \quad (31)$$

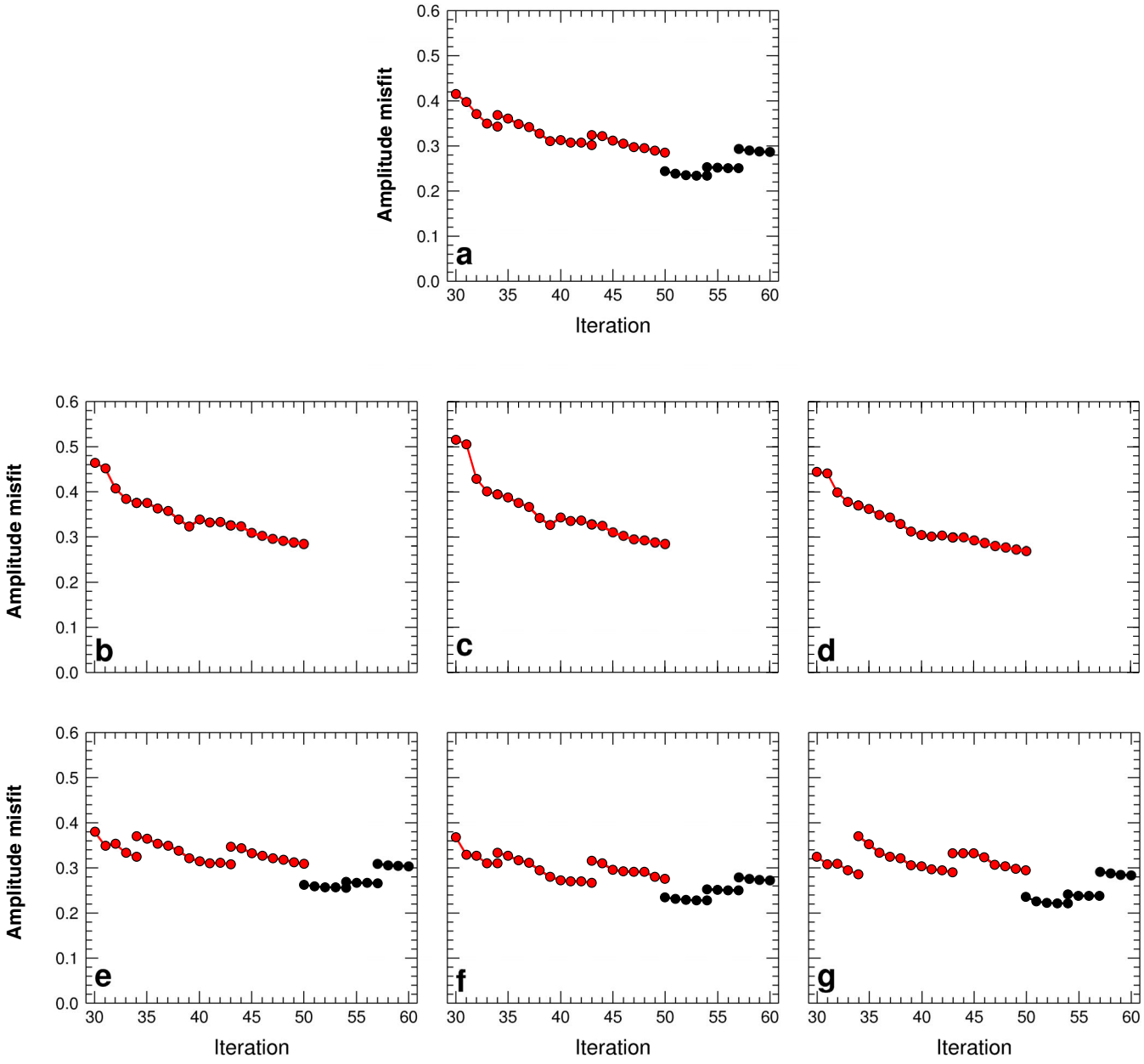

Figure 11. Evolution of amplitude misfits during the anelastic and anisotropic inversions, Stages II & III. (a) Evolution of the total amplitude misfit from iteration 30 to 60. Red and black dots label the anelastic and anisotropic inversions, respectively. (b)–(d) Evolution of amplitude misfits for *P-SV* body waves on vertical (b) and radial (c) components, and *SH* body waves on transverse components (d). (e)–(g) Evolution of amplitude misfits for Rayleigh waves on vertical (e) and radial (f) components, and Love waves on transverse components (g).

Variations in the ξ parameter may be used to infer vertical and horizontal flows in the crust and upper mantle (Ekström & Dziewonski 1998; Gung *et al.* 2003; Shapiro *et al.* 2004; Moschetti *et al.* 2010). Fig. 16 shows lateral variations in ξ at various depths in model EU_{60}. At depths shallower than 150 km, the horizontally polarized shear wave speed (β_h) is generally faster than the vertically polarized shear wave speed (β_v), that is, $\xi > 1$, in agreement with reference model STW105 (Kustowski *et al.* 2008b; see Fig. 5). Large values of ξ may indicate the presence of horizontal flow within the upper mantle. At greater depths (>150 km), some regions, such as the Calabrian and Hellenic arcs, are characterized by $\beta_h < \beta_v$, which may indicate vertical flow induced by subducting slabs. The small values of ξ at a depth of 220 km beneath the North Atlantic Ocean may be related to upwelling within the upper mantle driven by the divergency of the Eurasian and North American plates.

12.3 Other model parameters

We simultaneously determine 3-D variations in bulk sound wave speed (c), and vertically (β_v) and horizontally (β_h) polarized shear wave speeds. In Fig. 17, we compare lateral variations in different model parameters at a depth of 100 km. Long-wavelength features, such as the EEC, are imaged in both compressional (α) and shear (β) wave speeds. However, fast wave speed anomalies beneath the Adria microplate and the Hellenic arc are much broader in the compressional wave speed image. In contrast, slow anomalies related to the western Mediterranean and the Pannonian Basin are relatively

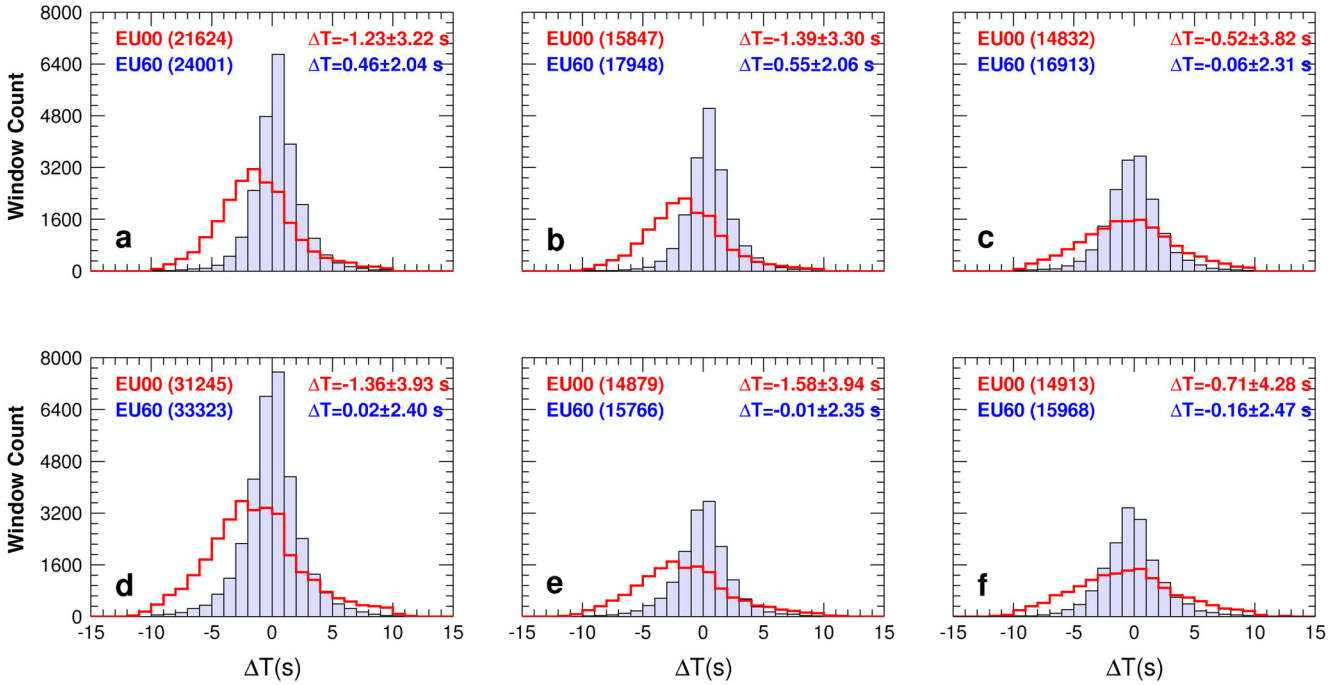

Figure 12. Comparisons of phase histograms between starting model EU_{00} (red) and final model EU_{60} (blue) for 15–40 s body waves and 25–100 s surface waves. (a)–(c) Comparisons of phase histograms for *P-SV* body waves on vertical (a) and radial (b) components, *SH* body waves on transverse components (c). (d)–(f) Comparisons of phase histograms for Rayleigh waves on vertical (d) and radial (e) components, Love waves on transverse components (f). The number of windows are compared between EU_{00} and EU_{60}. Both mean values and standard deviations of phase measurements are denoted in the figures.

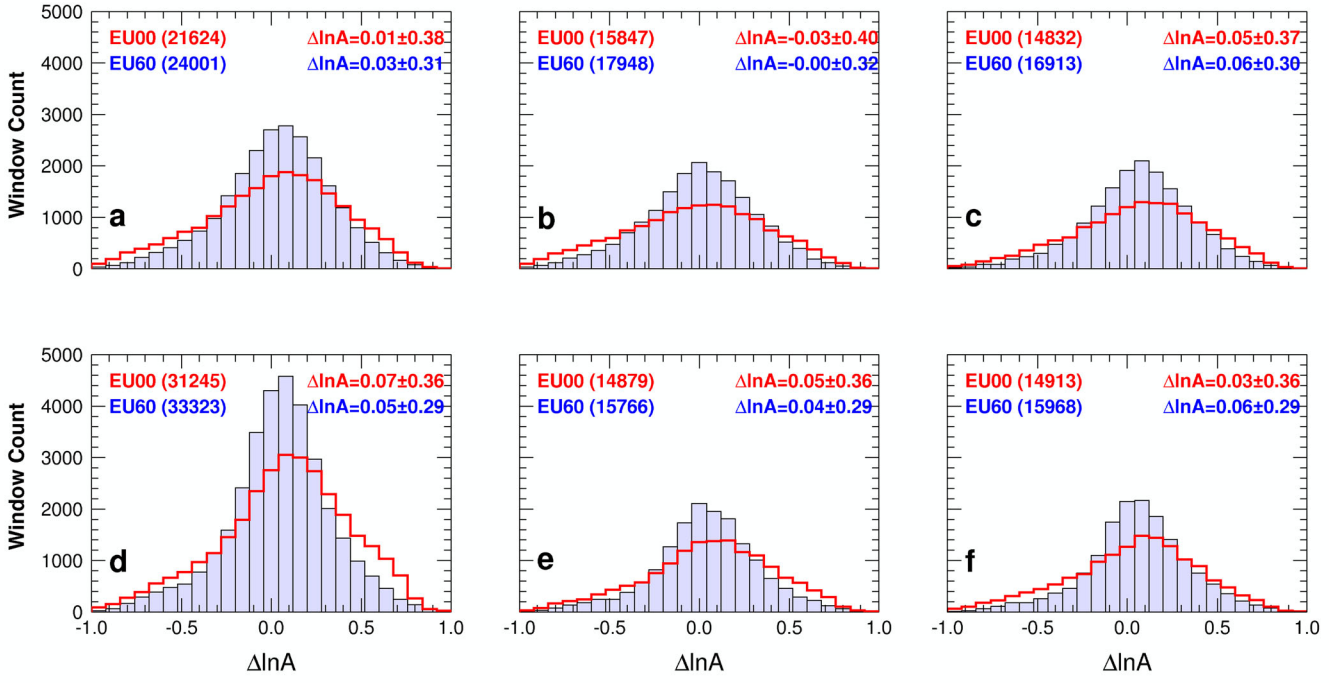

Figure 13. Same as Fig. 12 except for amplitudes.

weak in compressional wave speeds. Based on the compressional and shear wave speeds, we are able to calculate the V_p/V_s ratio. Fig. 17 illustrates the relative perturbation of V_p/V_s ratio at a depth of 100 km. There is a good correlation between regions with slow shear wave speed and a high V_p/V_s ratio, for instance, beneath the Alpine–Himalaya orogenic belt. High V_p/V_s ratios might be indicative of the presence of partial melt in this regions (Zhang *et al.* 2004). In contrast, the EEC is characterized by low V_p/V_s ratios at this depth.

12.4 Model evolution

A three-stage inversion strategy is used to constrain 3-D variations in elastic, anelastic and anisotropic model parameters. In Stage I,

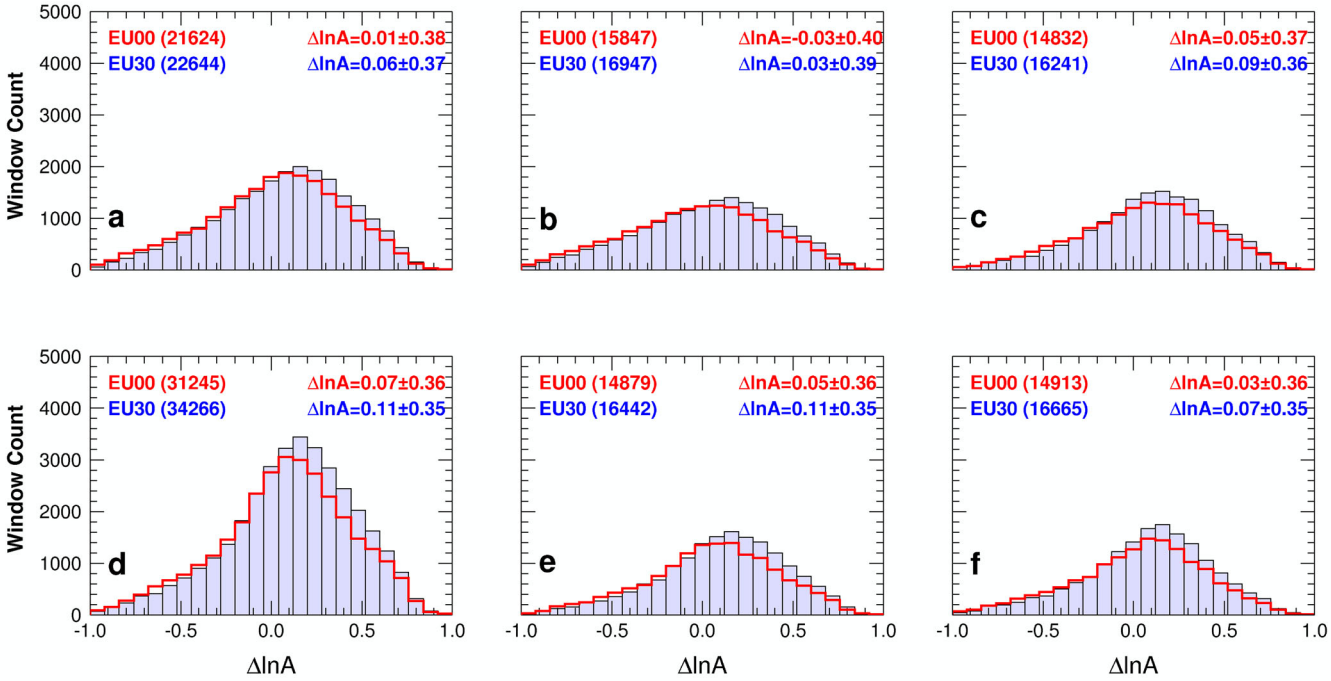

Figure 14. Comparisons of amplitude histograms between starting model EU_{00} and elastic model EU_{30}. Amplitude anomalies are not used in the Stage I elastic inversion, which does not reduce these significantly. Same as Fig. 12 except for amplitudes.

only frequency-dependent phase differences between observed and simulated seismograms are employed to map lateral variations in radially anisotropic elastic wave speeds. In Fig. 18, we observe significant improvements in isotropic shear wave speed from the starting model EU_{00} to the new elastic model EU_{30} at both shallow (100 km) and great (600 km) depths. Since the starting model is based on a global-scale study (Kustowski *et al.* 2008b), it only involves long-wavelength features. In EU_{30}, numerous small-scale features, as described in Section 12.1, naturally emerge from the smooth starting model.

In Stage II, that is, iterations 30–50, we combine frequency-dependent amplitude anomalies with remaining phase differences from Stage I to simultaneously estimate elastic wave speeds and anelastic attenuation. Most long-wavelength elastic wave speed features of EU_{30} are preserved in model EU_{50}, except for slight modifications of some small-scale heterogeneities. For instance, slow anomalies beneath the western Mediterranean and Eifel Hotspot become more prominent after the anelastic inversion. At a depth of 600 km, fast wave speed slab features beneath central Europe become stronger.

In Stage III, that is, iterations 50–60, only surface waves phase and amplitude anomalies are employed to constrain variations in radial and azimuthal anisotropy at shallow depths. The change in β_v and β_h from EU_{50} to EU_{60} is relatively modest because of small updates in wave speeds from iteration 50 to 60. No body waves are incorporated in the anisotropic inversion, therefore, there is no change in shear wave speed at great depths, for example, 600 km.

12.5 Comparisons with previous tomographic images

As discussed in Section 3, a variety of compressional and shear wave speed models of the European upper mantle has been developed based on different data sets and approaches. In order to assess similarities and discrepancies between model EU_{60} and these complementary studies, we compare elastic shear wave speed structures in EU_{60} with five previous surface wave tomographic models, namely EPmantle (Schivardi & Morelli 2011), Chang2010 (Chang *et al.* 2010a), S2.9EA (Kustowski *et al.* 2008a), LRSP30EU02 (Boschi *et al.* 2009) and CUSDT1.0 (Shapiro & Ritzwoller 2002), as well as body-wave traveltime model LLNL-G3Dv3 (Simmons *et al.* 2012).

EPmantle (Schivardi & Morelli 2011) was constructed based on fundamental mode Rayleigh and Love group wave speeds; only regional earthquakes were used in this study. Chang *et al.* (2010a) combined regional *S* and Rayleigh waveforms, teleseismic arrival times and Rayleigh group wave speeds to constrain upper-mantle shear wave speeds along the Tethyan margin. S2.9EA (Kustowski *et al.* 2008a) is a shear wave speed model constructed based on surface wave phase wave speeds, long-period waveforms and body-wave traveltimes. LRSP30EU02 (Boschi *et al.* 2009) is a shear wave speed model based on Rayleigh and Love fundamental-mode phase anomalies. Both S2.9EA and LRSP30EU02 are large-scale tomographic inversion with a finer parametrization beneath Eurasia (S2.9EA) and the Mediterranean (LRSP30EU02), respectively. CUSDT1.0 (Shapiro & Ritzwoller 2002) is a global shear wave speed model based on fundamental-mode phase and group wave speeds. All models are radially anisotropic.

In Fig. 19, we compare relative perturbations in isotropic shear wave speed at a depth of 100 km. At long wavelength, the level of agreement among these six models is very good. For instance, the EEC and the Alpine–Himalaya orogenic belt are imaged as fast and slow anomalies in all six models, respectively. Model EU_{60} involves more short-wavelength features which are not well resolved by the other five models. For instance, the continuous belt with fast wave speed connecting the Alps, Adriatic and Hellenides is not obvious in EPmantle and Chang2010, and is only resolved as a very smooth feature in S2.9EA, LRSP30EU02 and CUSDT1.0. A similar observation can be made for fast wave speed anomalies associated with the Central Graben and the Armorican Massif. In

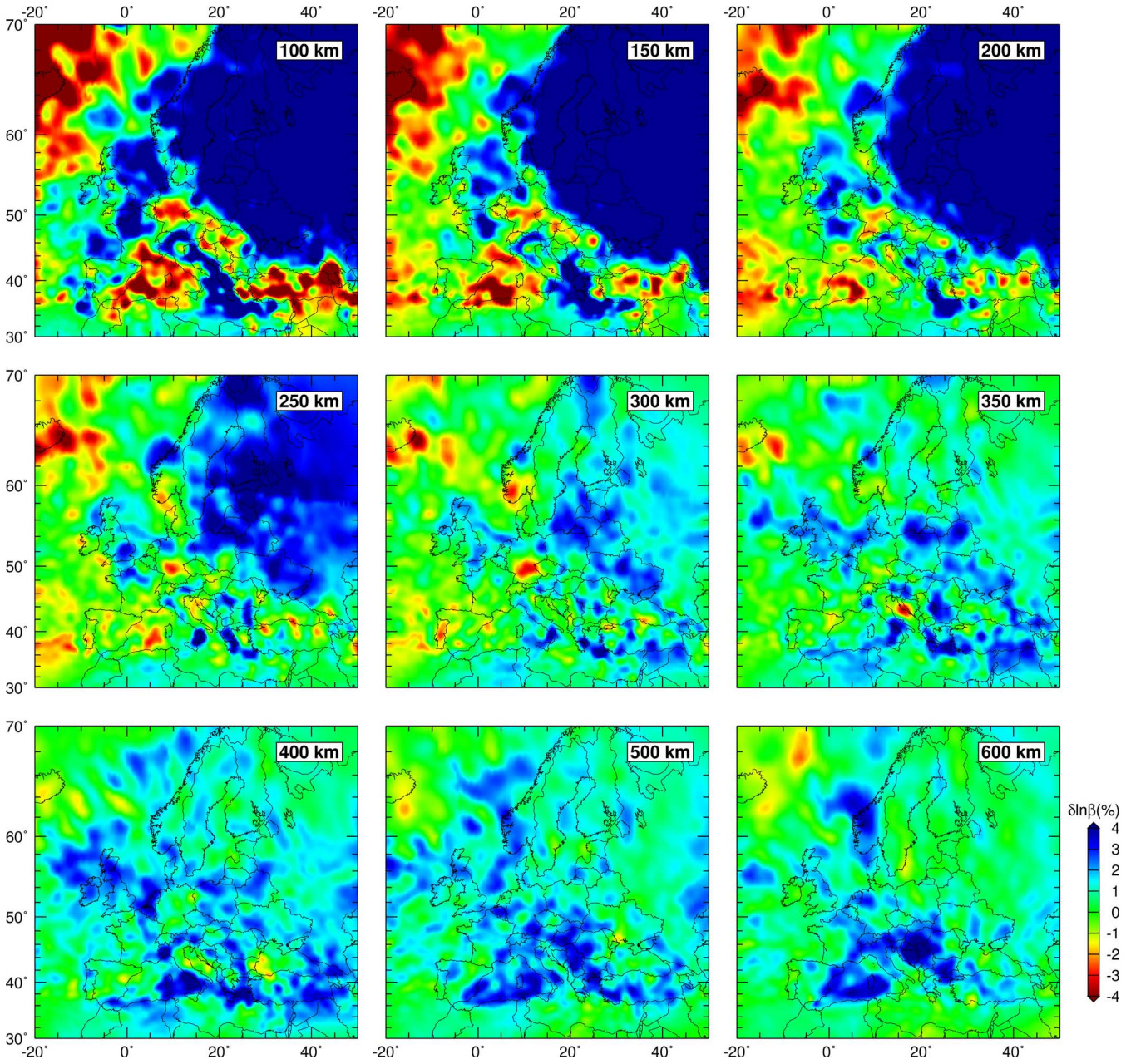

Figure 15. Relative perturbations in isotropic shear wave speed at various depths in model EU_{60}. Depths are denoted in the right top corner.

addition, the amplitudes of slow anomalies beneath the western Mediterranean, Pannonian Basin and Anatolian Plate in EPmantle and Chang2010 are relatively large compared to the other four models.

In Fig. 20, we compare radial anisotropy in these six models at a depth of 150 km. In contrast to the reasonable consensus in isotropic shear wave speeds among the various models, the level of agreement in radial anisotropy is relatively poor. This discrepancy is due to different weights assigned to Rayleigh and Love waves in the various inversions. At long wavelengths, all models involve large values of ξ beneath the European continent, indicating the presence of horizontal flow as discussed in Section 12.2. Model EU_{60} identifies small values of ξ beneath the Apennines, Adria an Hellenic arcs, which are only weakly imaged in model LRSP30EU02.

Finally, in Fig. 21 we compare model EU_{60} with body-wave traveltime model LLNL-G3Dv3 (Simmons *et al.* 2012). There is good agreement between shear wave speed signatures in EU_{60} and compressional wave speed features in LLNL-G3Dv3 at both shallow (80 km) and great (600 km) depths. The EEC is not well imaged in model LLNL-G3Dv3 due to vertically incident body wave and a lack of stations and earthquakes in this region. Both models reveal strong fast wave speed anomalies beneath central Europe within the mantle transition zone.

13 ATTENUATION

Horizontal cross-sections of relative perturbations in shear attenuation Q^{-1} are displayed in Fig. 22. The radial Q profile from STW105 (Kustowski *et al.* 2008b) is used as 1-D reference model to calculate relative perturbations (Fig. 5). At shallow depths, for example, 100 km, there is a clear anticorrelation between elastic wave speeds and anelastic attenuation (compare with Fig. 15). For instance, the

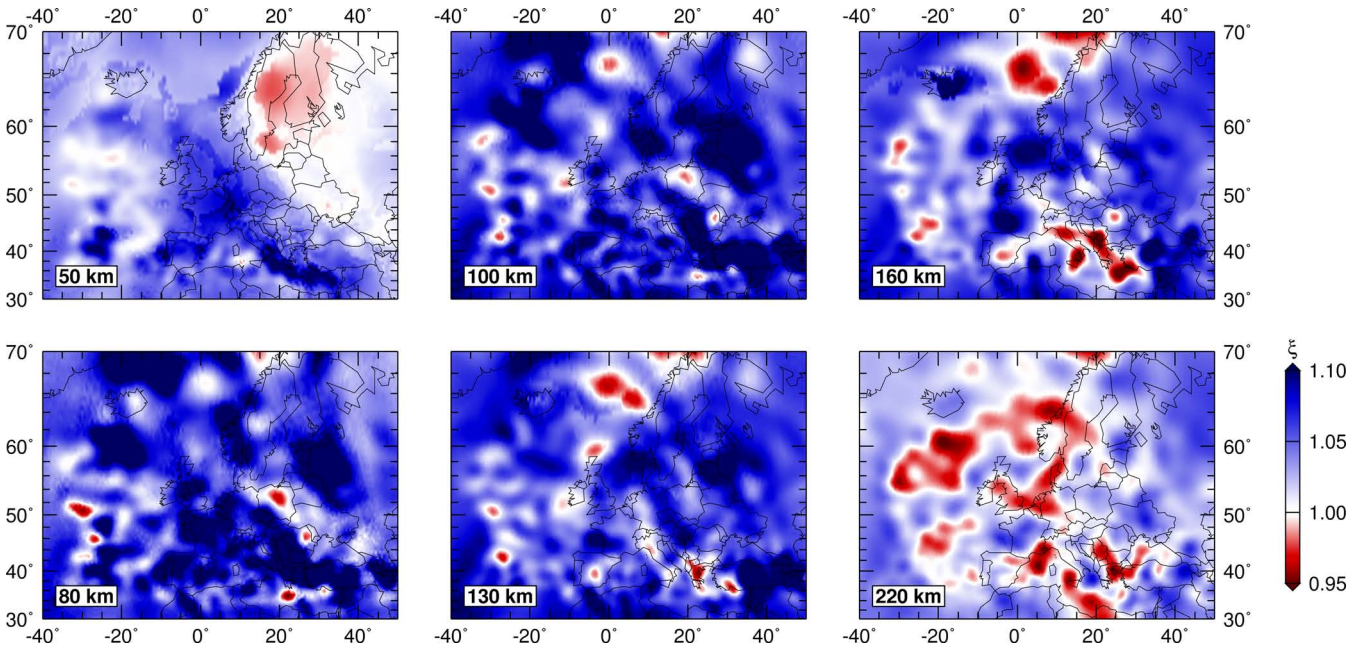

Figure 16. Radially anisotropic model parameter $\xi = (\beta_h/\beta_v)^2$ at various depths in model EU_{60}. A value of $\xi > 1$ indicates the presence of horizontal flow, while values $\xi < 1$ indicate vertical flow. Depths are denoted in the left bottom corner. Values of $\xi < 1$ at 50 km beneath the Baltic Shield might indicate radial anisotropy within the lower crust.

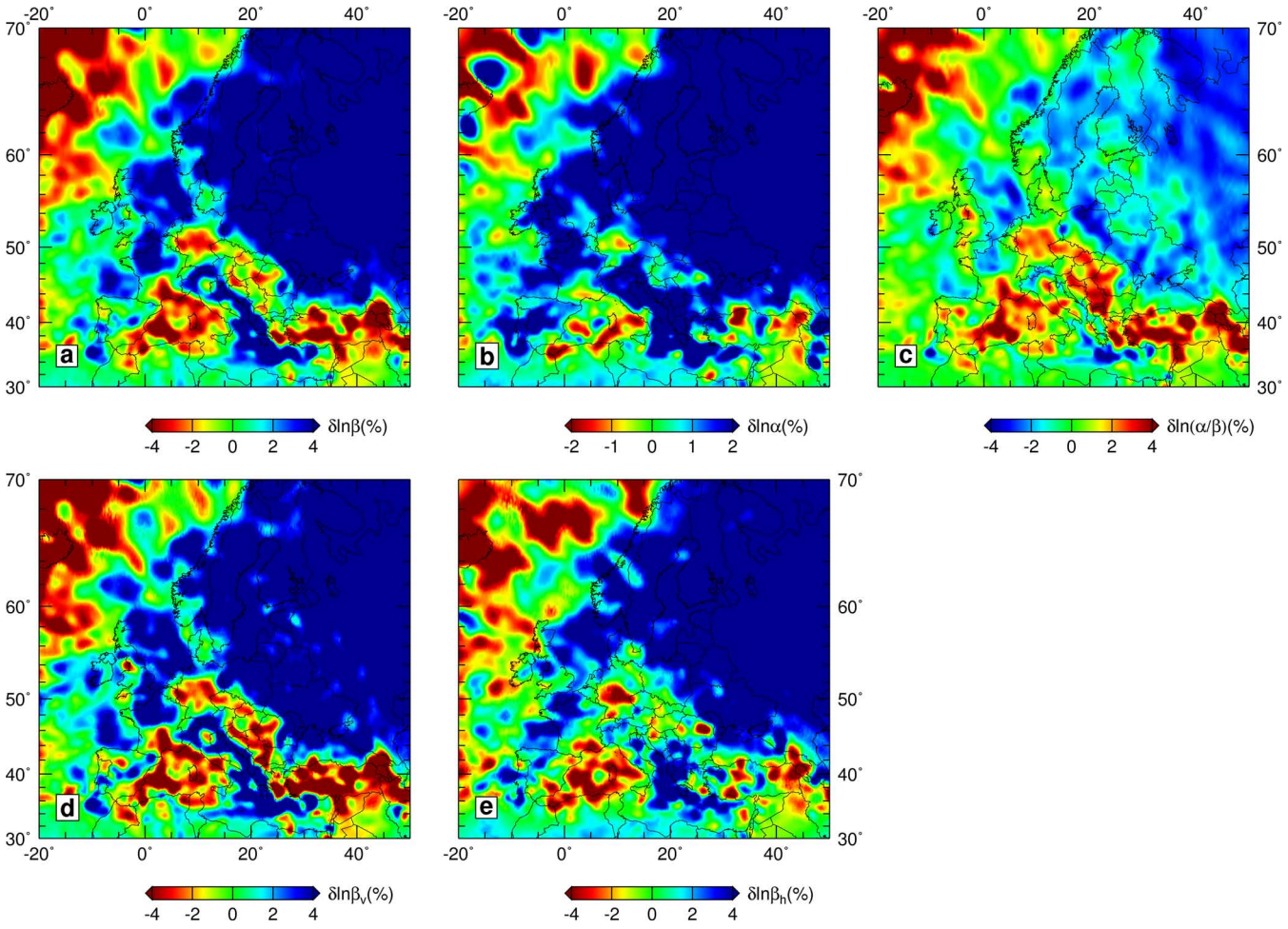

Figure 17. Various EU_{60} model parameters at a depth of 100 km. (a)–(e) Relative perturbations in isotropic shear wave speed (a), compressional wave speed (b), V_p/V_s ratio (c), vertically (d) and horizontally (e) polarized shear wave speeds.

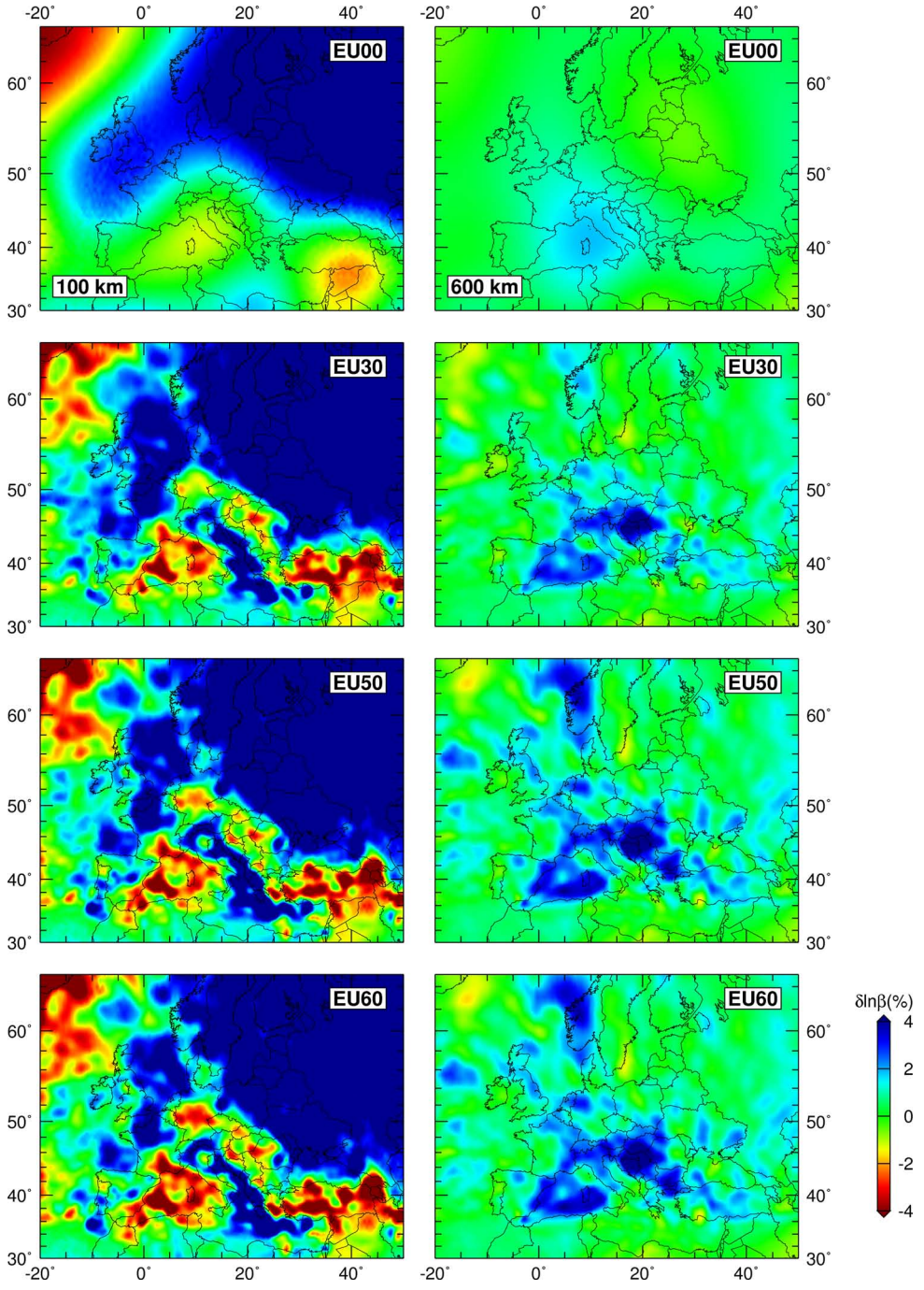

Figure 18. Evolution of isotropic shear wave speed from starting model EU_{00} to purely elastic model EU_{30} (Stage I) to anelastic model EU_{50} (Stage II) to final model EU_{60} (Stage III) at depths of 100 km (left column) and 600 km (right column).

EEC is revealed as a region with low attenuation and fast wave speed. Below 200 km, the anticorrelation between wave speeds and attenuation becomes relatively weak, in agreement with observations of global surface wave attenuation tomography (Romanowicz 1995; Billien *et al.* 2000; Gung & Romanowicz 2004; Lawrence & Wysession 2006; Dalton *et al.* 2008).

As discussed in Section 12.1, within the mantle transition zone central Europe is dominated by several fast wave speed anomalies related to subducting slabs in the Mediterranean–Alpine region. However, there is no obvious anticorrelation between anelastic attenuation and elastic wave speeds at these depths (Fig. 22).

In Fig. 23, three vertical cross-sections (N1– N3) are used to explore shear attenuation beneath the North Atlantic. The lithosphere generally exhibits weak attenuation whereas the asthenosphere is characterized by relatively high attenuation, and we recognize these characteristics in the cross-sections. Perhaps surprisingly, enhanced attenuation is revealed within the mantle transition zone, but no significant accompanying reduction in shear wave speed is observed. This feature might be related to the presence of water. Major minerals of the mantle transition zone, namely Wadsleyite and Ringwoodite, have lager water solubilities than minerals in the shallow upper mantle (Kohlstedt *et al.* 1996). These minerals might be reservoirs

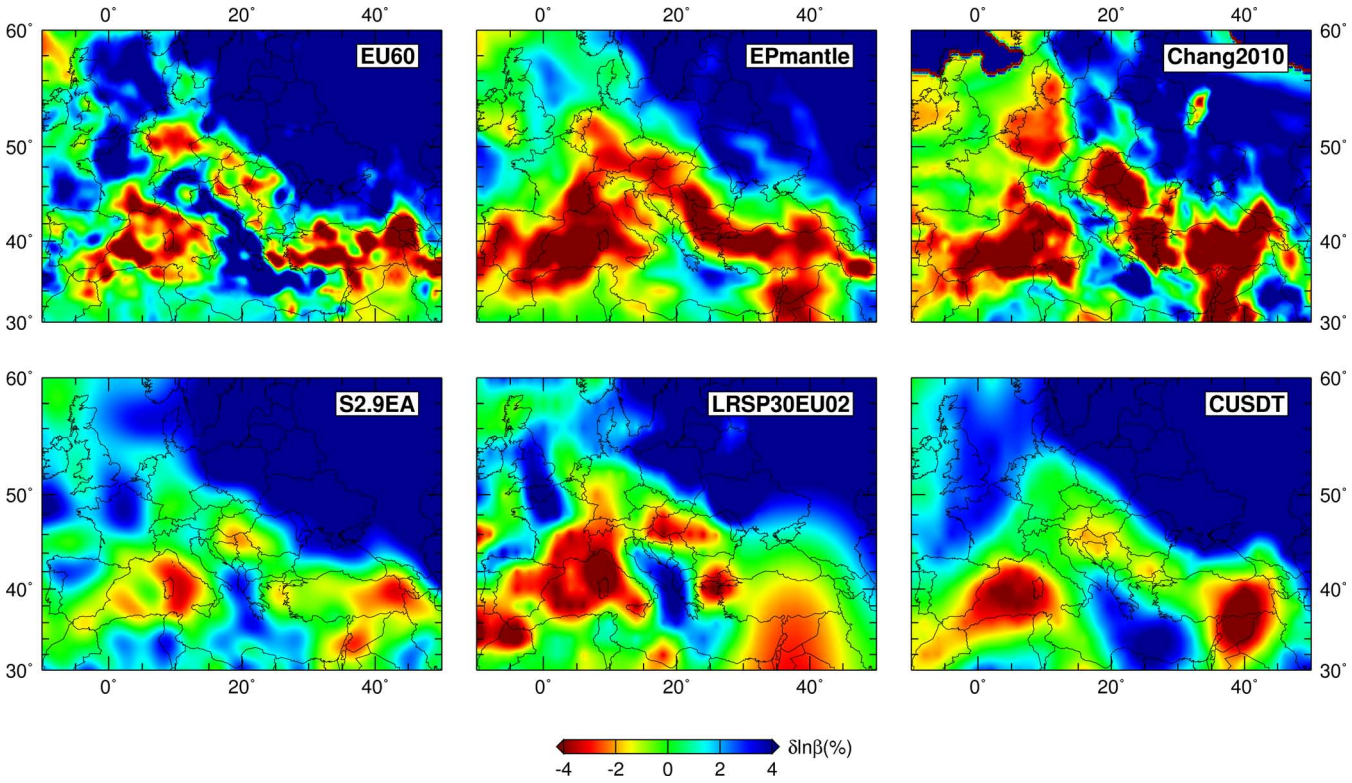

Figure 19. Comparison of relative perturbations in isotropic shear wave speed at a depth of 100 km for six different tomographic models, namely, EU_{60} (this study), EPmantle (Schivardi & Morelli 2011), Chang2010 (Chang *et al.* 2010a), S2.9EA (Kustowski *et al.* 2008a), LRSP30EU02 (Boschi *et al.* 2009) and CUSDT1.0 (Shapiro & Ritzwoller 2002). STW105 (Kustowski *et al.* 2008b) is used as a reference model to calculate relative perturbations.

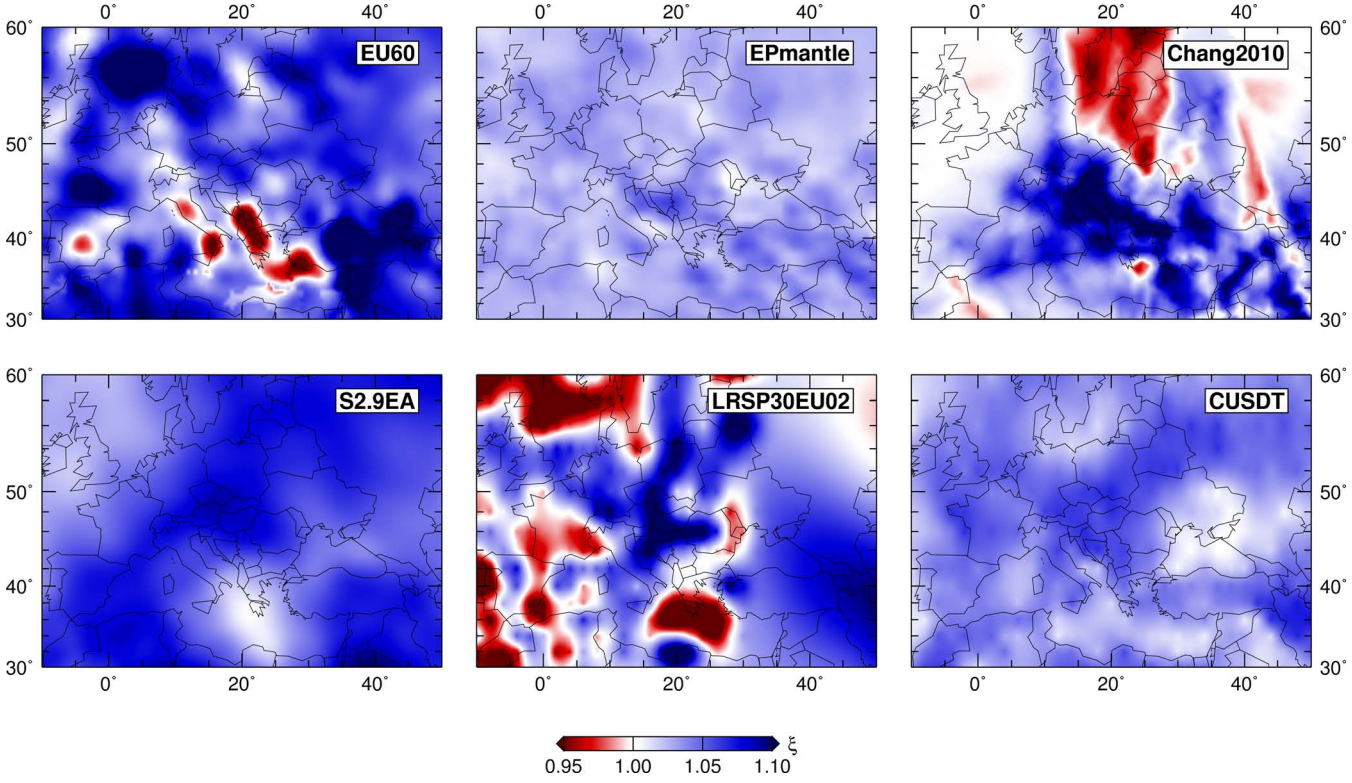

Figure 20. Same as Fig. 19, except for a comparison of the radially anisotropic model parameter ξ at a depth of 150 km.

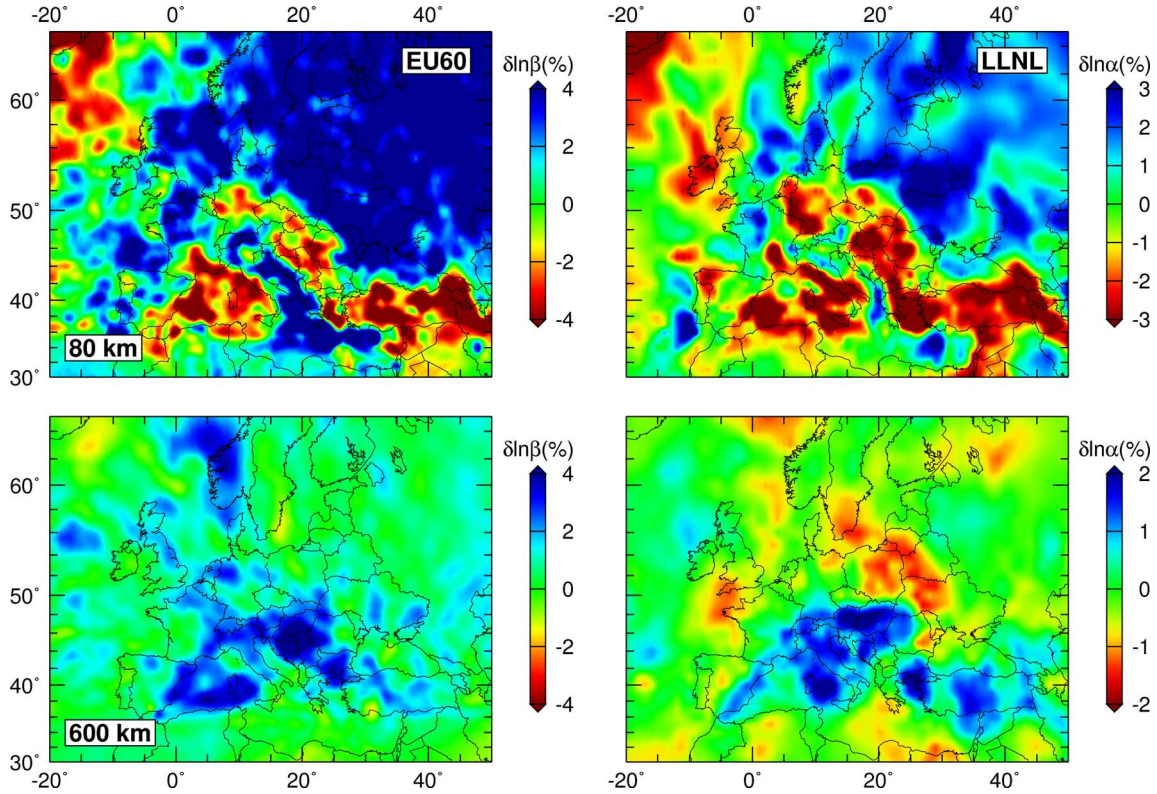

Figure 21. Comparison of EU_{60} with body-wave traveltime model LLNL-G3Dv3 (Simmons *et al.* 2012) at depths of 80 km (top panels) and 600 km (bottom panels). STW105 (Kustowski *et al.* 2008b) is used as a 1-D reference model to calculate relative perturbations in isotropic shear wave speed for EU_{60}, while PREM (Dziewonski & Anderson 1981) is used as a 1-D reference model to calculate relative perturbations in compressional wave speed for LLNL-G3Dv3 (Simmons *et al.* 2012).

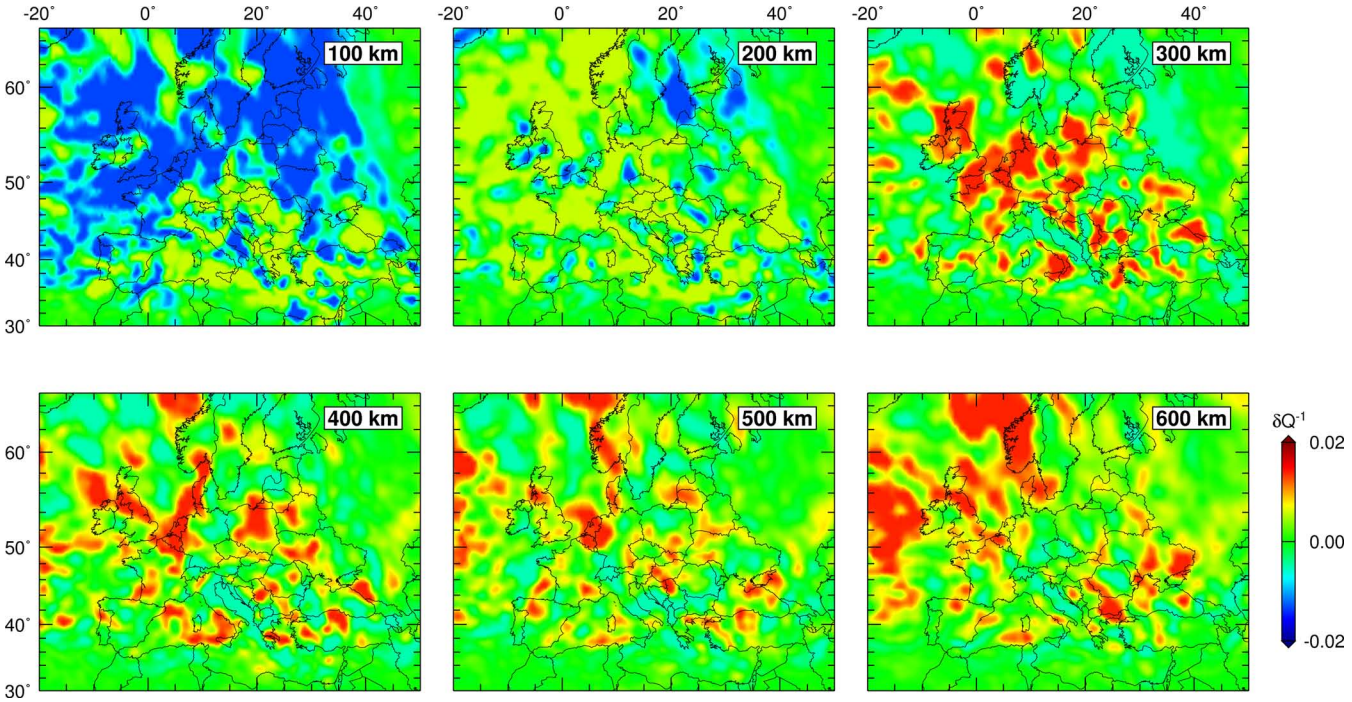

Figure 22. Relative perturbations in shear attenuation Q^{-1} at various depths in model EU_{50}. The radial Q model of STW105 (Kustowski *et al.* 2008b, see Fig. 5) is used as 1-D reference model to calculate relative perturbations.

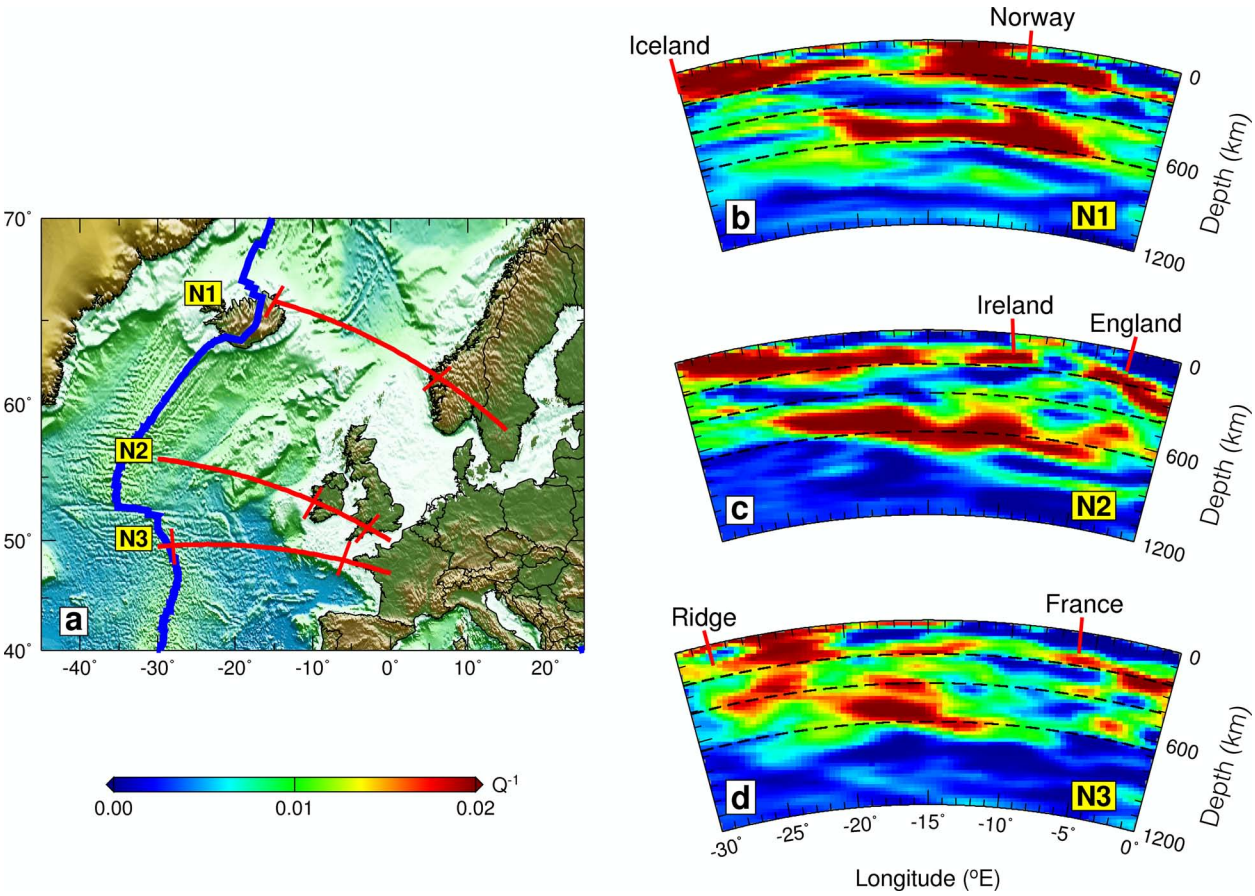

Figure 23. Vertical cross-section of shear attenuation Q^{-1} beneath the North Atlantic Ocean. (a) Locations of the three vertical cross-sections N1– N3. (b)–(d) Shear attenuation in the three vertical cross-sections N1– N3. The dashed black lines in N1–N3 denote the 220, 410 and 660 km discontinuities.

for significant amounts of water in the Earth's mantle (Bercovic & Karato 2003; Karato 2011). A detailed discussion of the implications of a water-enriched mantle transition zone may be found in Zhu *et al.* (2013).

14 ANISOTROPY

Azimuthal anisotropy in EU_{60} between 50 and 220 km is illustrated in Fig. 24. It involves complex lateral and depth variations in the directions and amplitudes of the fast anisotropic axis beneath continental Europe and the North Atlantic. Because of the limited depth sensitivity of surface waves, the strength of azimuthal anisotropy decreases below 200 km. Above 200 km, anisotropic fabrics are well correlated with regional tectonic evolution. Along the NAR, the fast axis runs parallel to the extensional direction of the ridge system within the upper mantle. Small-scale complexities, such as the convergency of the fast axis along the western coast of England, are intriguing and require further investigations involving other geophysical observables, such as gravity and electrical conductivity. The fast axis beneath the western Mediterranean follows the opening trajectories of the Algero–Provençal and Tyrrhenian Seas, suggesting trench retreat of the Apennines–Calabrian arc.

At shallow depths, for example, 50–150 km, the EEC involves complex azimuthally anisotropic patterns, which might be correlated with ancient continental rifts, as discussed in Zhu & Tromp (2013). At greater depths, relatively weak anisotropy is observed within the EEC. Around the EEC, the fast axis is well correlated with tectonic activities which are closely related to the accretion of the EEC since the Palaeozoic. The fast axis along the western coast of Scandinavia follows the trend of the Caladonian Orogeny during the Palaeozoic (see Fig. 1). The TTSZ is delineated by the fast anisotropic direction throughout the upper mantle, separating the EEC and western and eastern Europe. To the southern border of the EEC, the fast direction indicates northward motion and counterclockwise rotation of the Arabian Plate, reflecting collision between Eurasia and Arabia.

In Fig. 25, we compare averaged radial peak-to-peak anisotropic strength profiles for several regions. For oceanic regions, such as the North Atlantic Ocean (profile 1) and the Mediterranean (profile 2), strong azimuthal anisotropy is observed at a depth of 100 km. Within the upper mantle, the anisotropic strength increases monotonically, and then steadily decreases with depth below its maximum value. The same feature is observed beneath the Pannonian Basin, except that the depth of maximum anisotropic strength is approximately 150 km (profile 3).

In contrast, beneath the Aegean and Anatolian plates (profiles 4 and 5), two peaks in anisotropic strength are observed. The first peak indicates weak and ductile lower crust. The second peak reflects strong mantle flow within the lithosphere and asthenosphere. They are accommodated by a transition zone with relatively weak anisotropic strength. These profiles are consistent with mineral physics experiments involving distinct brittle-ductile deformation in lithospheric strength beneath oceans and continents (Kohlstedt *et al.* 1995).

Beneath the EEC and Ukrainian Shield (profiles 6 and 7), strong azimuthal anisotropy is observed within the lower crust. However,

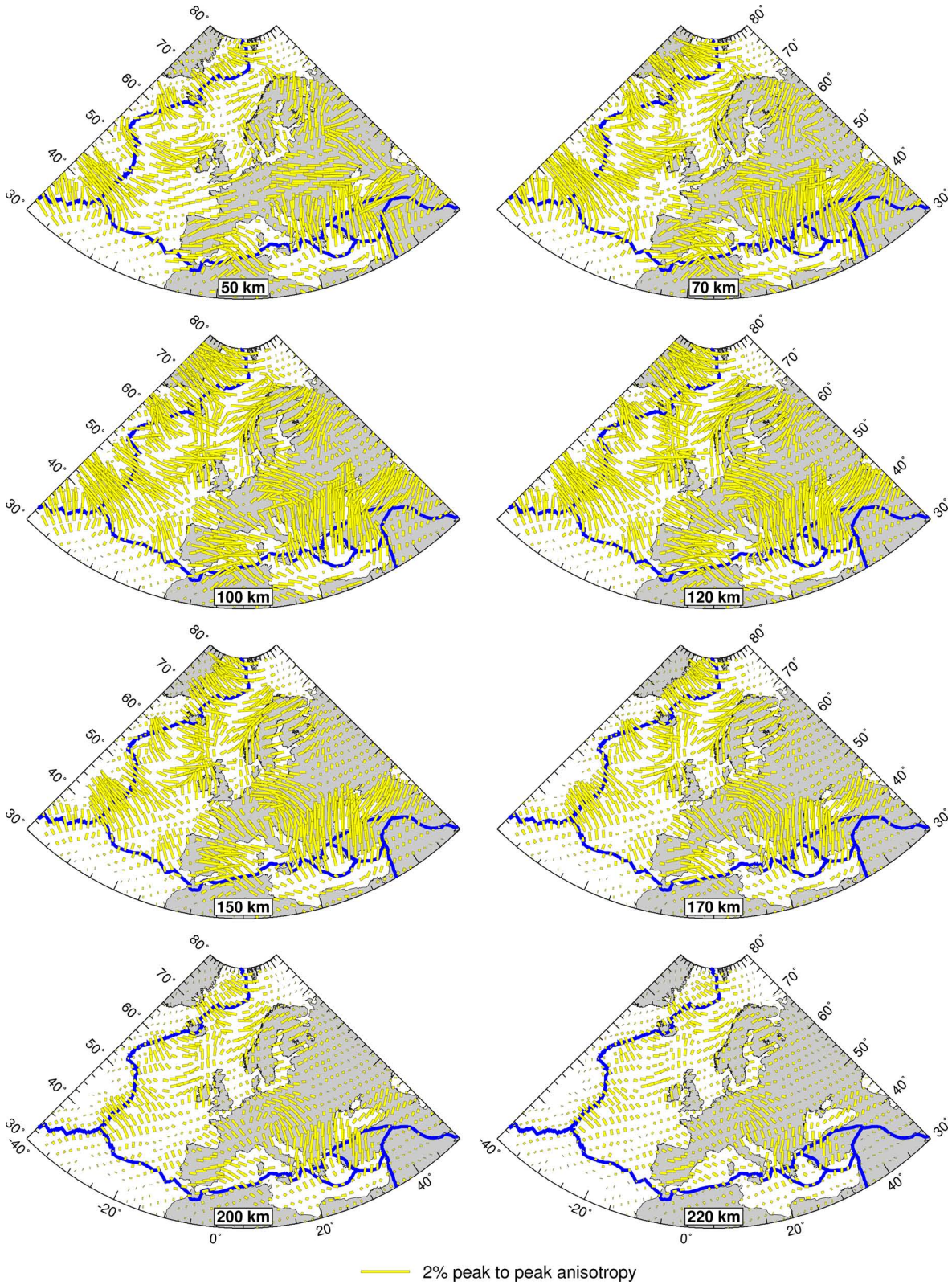

Figure 24. Azimuthal anisotropy at various depths in model EU_{60}. The direction and amplitude of the fast axis are given by the orientation and length of the yellow bar. Blue lines denote global plate boundaries (Bird 2003).

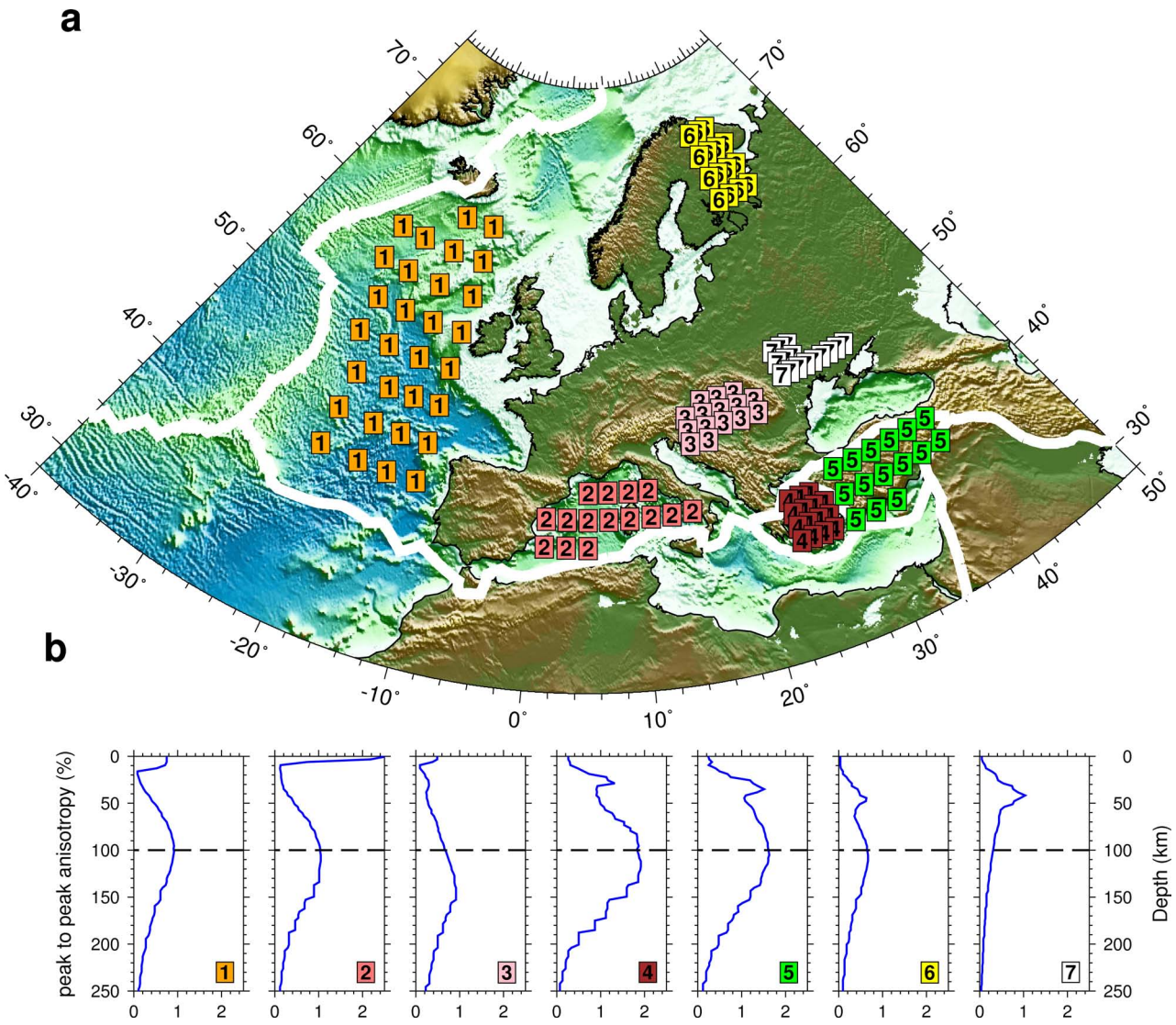

Figure 25. Comparisons of averaged radial peak-to-peak anisotropic strength profiles for different regions. (a) Locations of seven averaged profiles. Different colours refer to different regions. White lines denote global plate boundaries (Bird 2003). (b) Averaged radial peak-to-peak anisotropic strength profiles for seven regions: 1, the North Atlantic Ocean; 2, the Mediterranean; 3, the Pannonian Basin; 4, the Aegean Sea; 5, the Anatolian Plate; 6, the EEC; 7, the Ukrainian Shield. Black dashed lines denote a reference depth of 100 km.

the lower lithosphere (depth of ∼100 km) is observed to have relatively weak anisotropic strength, which might be indicative of a plunging axis of symmetry in the continental lithosphere (Debayle *et al.* 2005). The current model parametrization does not accommodate this type of dipping anisotropy. The digital model is available as the supplementary file.

15 RESOLUTION ANALYSIS

Because of the demanding computational requirements of adjoint tomography, it is very expensive to perform traditional 'checkboard' tests for assessing 3-D image quality, which require the same amount of computational resources as an actual structural inversion. Such experiments have been performed successfully in two dimensions—see for example, Tape *et al.* (2007) and Luo *et al.* (2013)—with the conclusion that success is basically controlled by ray coverage. For this reason, ray density maps are an excellent indicator of resolution, as one might expect (Luo *et al.* 2013).

In this study, the 'point-spread function' (Fichtner & Trampert 2011, 2012) is used to assess image quality in models EU_{30}, EU_{50} and EU_{60}, as well as for analysing trade-offs between elastic, anelastic and anisotropic model parameters. A finite-difference approximation is used to calculate the action of the Hessian on a localized model perturbation:

$$\mathbf{H} \cdot \delta\mathbf{m} \approx \mathbf{g}(\mathbf{m} + \delta\mathbf{m}) - \mathbf{g}(\mathbf{m}), \tag{32}$$

where $\mathbf{H}$ denotes the Hessian and $\delta\mathbf{m}$ refers to a localized model perturbation with respect to the current model $\mathbf{m}$. The misfit gradient $\mathbf{g}$ is evaluated for both models $\mathbf{m}$ and $\mathbf{m} + \delta\mathbf{m}$. Based on the action of the Hessian on the model perturbation, $\mathbf{H} \cdot \delta\mathbf{m}$, we are able to assess the curvature of the misfit function at a particular 'point' in the model space, reflecting the degree of 'blurring' of that point. Since we have to calculate the misfit gradient $\mathbf{g}(\mathbf{m} + \delta\mathbf{m})$ for the perturbed model and we already have the gradient $\mathbf{g}(\mathbf{m})$ for the current model, the computational requirements for a single spot analysis are the same as one full iteration. To perform this test numerically, we employ a −1 per cent 3-D Gaussian model

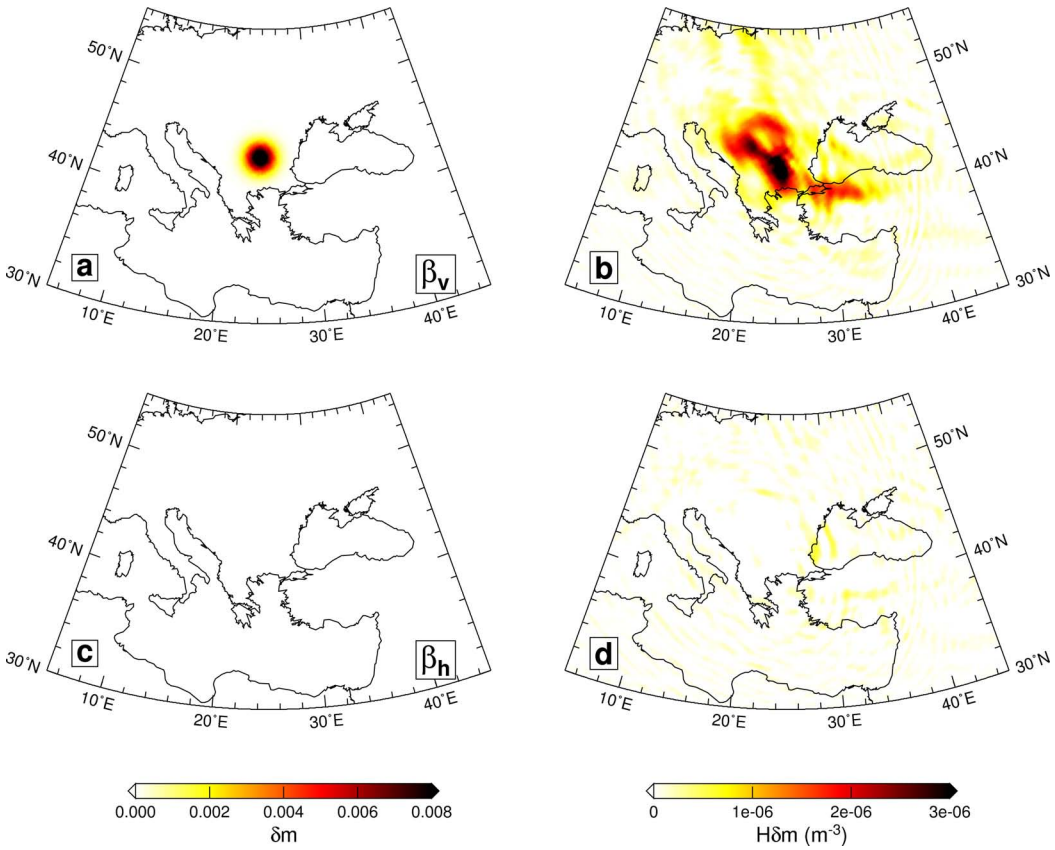

Figure 26. Resolution analysis beneath eastern Europe after the Stage I elastic inversion. (a) and (c) Model perturbations in β_v and β_h with respect to elastic model EU_{30} at a depth of 480 km. (b) and (d) Corresponding 'point-spread function' for β_v and β_h, respectively. The half width of the Gaussian is 120 km.

perturbation with an ∼120-km half width in Figs 26–28 for models EU_{30}, EU_{50} and EU_{60}. Experiments with centre differencing rather than forward differencing or with smaller (positive or negative) perturbations lead to comparable results and conclusions.

In Fig. 26, β_v is perturbed with respect to elastic model EU_{30} at a depth of 480 km beneath eastern Europe, while the remaining three model parameters, β_h, c and η, are held fixed. Although there is some smearing, the main features of the 3-D Gaussian perturbation are preserved in the 'Hessian kernel', $\mathbf{H} \cdot \delta\mathbf{m}$, thereby confirming image quality at this location. By comparing $\mathbf{H} \cdot \delta\mathbf{m}$ for β_v and β_h, we conclude that there is limited trade-off between the two shear wave speeds in our tomographic images.

In Fig. 27, we perturb Q^{-1} with respect to anelastic model EU_{50} at the same location as in Fig. 26, while keeping the other parameters fixed. There is modest smearing, while the trade-off between elastic and anelastic model parameters is relatively weak. However, this resolution test indicates that the Q^{-1} model shown in Fig. 22 contains smaller-scale features which are unresolved, and its interpretation should be limited to its long-wavelength characteristics.

The same behaviour is observed when we perturb the azimuthally anisotropic parameter G_c with respect to the final anisotropic model EU_{60} in Fig. 28 (the depth of the model perturbation is changed from 480 to 120 km, because only surface waves are used for mapping azimuthally anisotropic heterogeneity). Based on the results shown in Figs 26–28, we conclude that features imaged in EU_{30}, EU_{50} and EU_{60} are robust, and that trade-offs between different model parameters are relatively weak.

Finally, we use the 'approximate Hessian'—a scalar field—to assess ray coverage in this study (Luo *et al.* 2013). It may be expressed as a product of forward and adjoint acceleration fields (see eq. 7). Fig. 29 illustrates the depth dependence of the approximate Hessian for starting model EU_{00}. Within the upper mantle, we have very good ray coverage for the entire European continent and North Atlantic Ocean. At greater depths, for example, 600 km, our data set is still able to illuminate the European continent very well because of the incorporation of body waves.

16 CONCLUSIONS

In this paper, we use adjoint tomography to image the crust and upper mantle beneath Europe and the North Atlantic. A three-stage inversion strategy is designed to determine 3-D variations in elastic, anelastic and anisotropic model parameters. In Stage I, only phase information is used to image elastic wave speeds. After 30 pre-conditioned conjugate gradient iterations, we determine a new elastic model, named EU_{30}, which is simultaneously constrained by three-component short-period body waves and long-period surface waves. In Stage II, we combine phase and amplitude differences between observed and simulated seismograms to simultaneously constrain elastic wave speeds and anelastic attenuation. A new anelastic model, namely EU_{50}, is constructed based on 20 additional iterations. In Stage III, remaining phase and amplitude anomalies for three-component surface waves are used to constrain radial and azimuthal anisotropy, culminating in anisotropic model EU_{60}. Gradual reductions in misfit and significant improvements in phase and amplitude histograms help us to validate our inversion strategy.

Elastic wave speed variations in the final model, named EU_{60}, confirm most long-wavelength features determined in previous

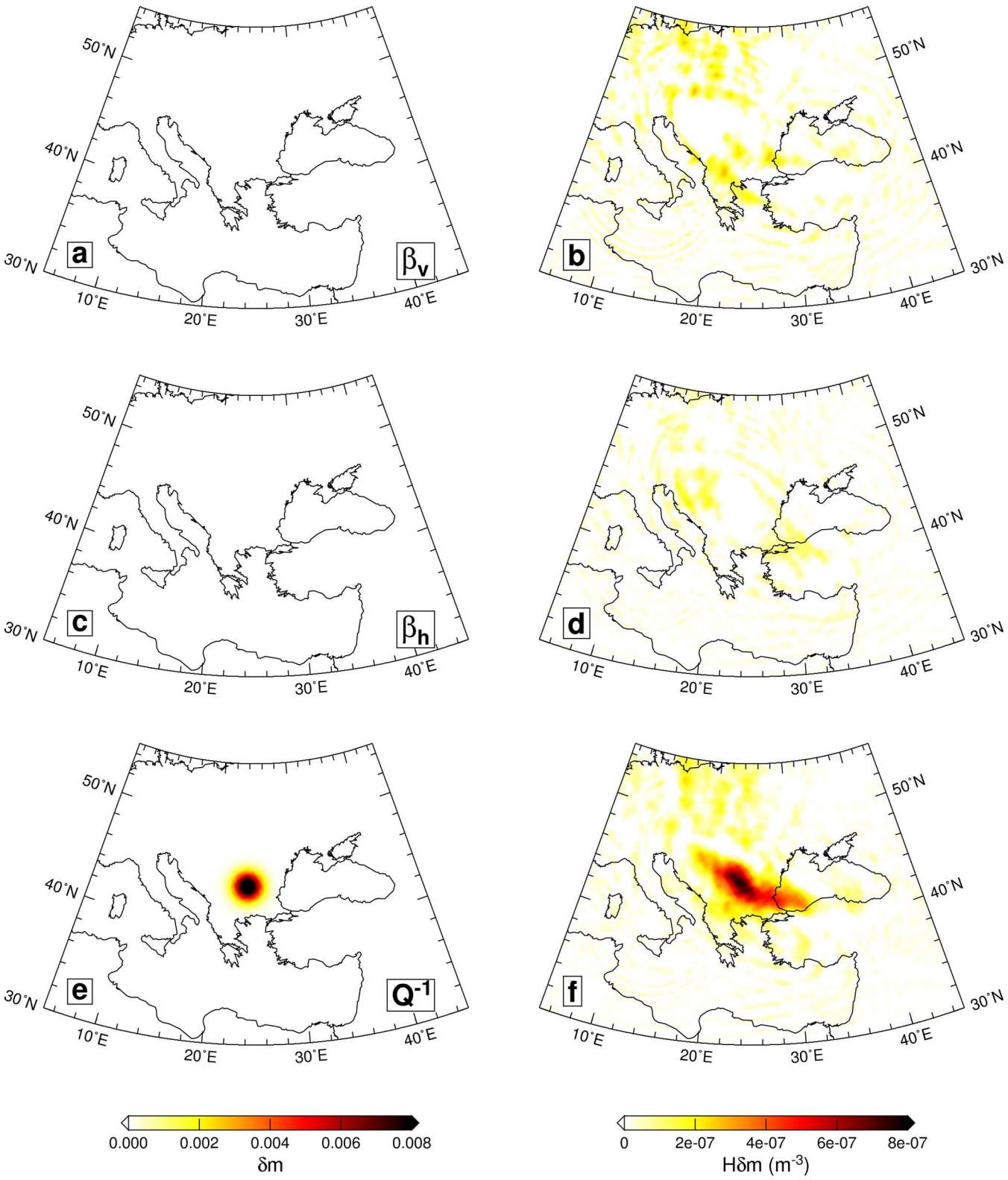

Figure 27. Resolution analysis beneath eastern Europe after the Stage II anelastic inversion. (a), (c) and (e) Model perturbations in β_v, β_h and Q^{-1} with respect to anelastic model EU_{50} at a depth of 480 km. (b), (d) and (f) Corresponding 'point-spread function' with respect to β_v, β_h and Q^{-1}.

body- and surface wave tomographic studies. However, numerous hitherto unidentified, small-scale structures are gradually revealed in our iterative inversion. These features are related to subducting slabs, slab detachments, roll-back and backarc basins.

Our images show an anticorrelation between shear wave speed and anelastic attenuation at shallow depths, where both parameters are correlated with surface tectonic provinces, such as the EEC and the western Mediterranean. At greater depths, this anticorrelation

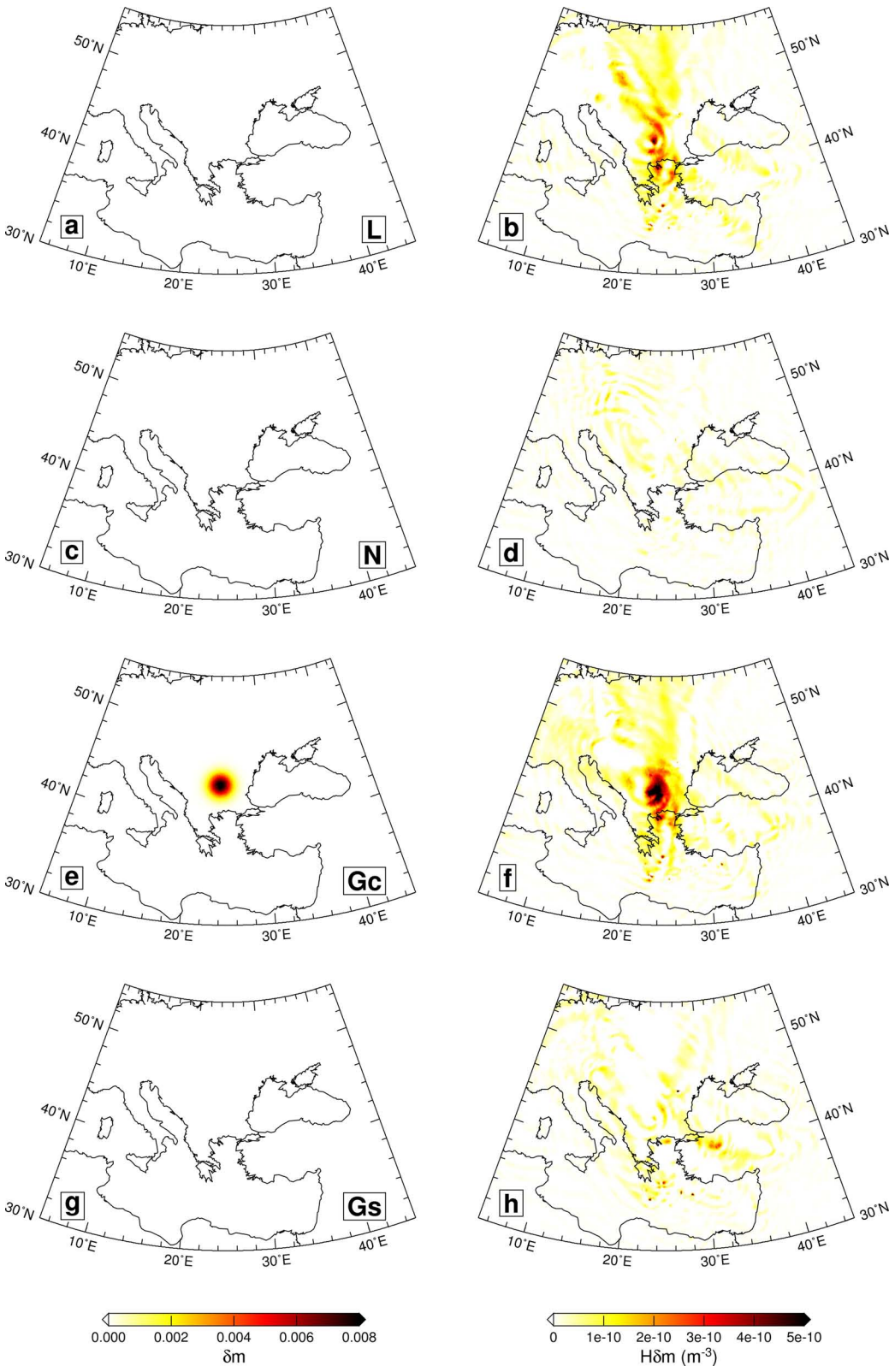

Figure 28. Resolution analysis beneath eastern Europe after the Stage III anisotropic inversion. (a), (c), (e) and (g) Model perturbations in L, N, G_c and G_s with respect to anisotropic model EU_{60} at a depth of 120 km. (b), (d), (f) and (h) Corresponding 'point-spread function' with respect to L, N, G_c and G_s.

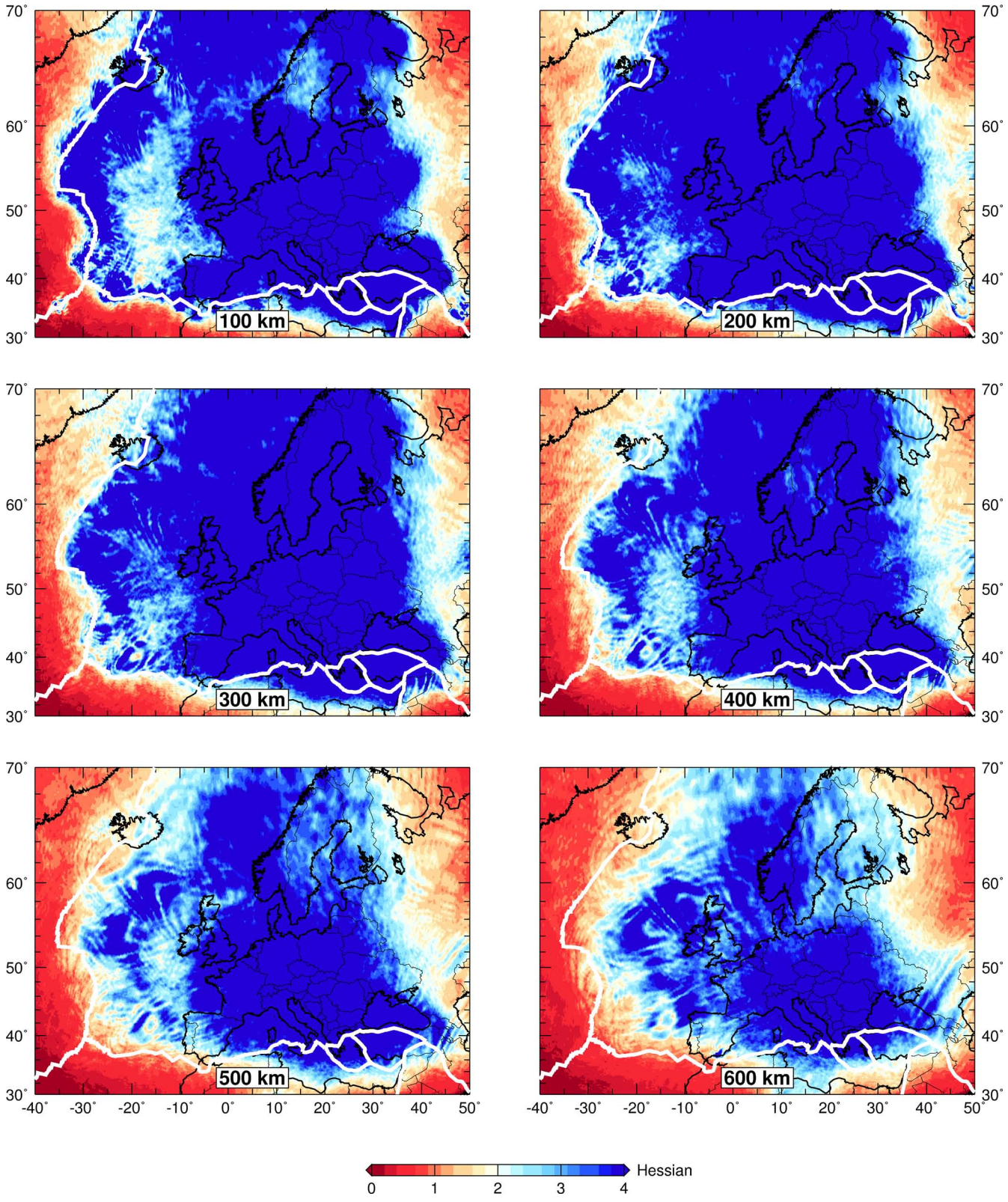

Figure 29. Approximate Hessian for starting model EU_{00} at various depths. This 3-D map may be used to assess ray coverage. White lines denote global plate boundaries (Bird 2003). Depths are denoted on the bottom.

becomes relatively weak, in agreement with conclusions from previous global attenuation tomography. Enhanced attenuation is observed within the mantle transition zone beneath the North Atlantic Ocean, which might be indicative of the presence of water in this region.

Most parts of the European continent are characterized by a radially anisotropic parameter $\xi > 1$, indicating the presence of horizontal flow within the upper mantle. Beneath subduction zones, such as the Apennines and the Hellenic arc, radially anisotropic features with $\xi < 1$ are resolved at depths greater than 150 km, indicating predominantly vertical flow.

Azimuthally anisotropic fabrics in EU_{60} are well correlated with the tectonic evolution of Europe over the past hundreds of millions of years, identifying extension along the NAR and counterclockwise rotation of the Anatolian Plate. Averaged radial peak-to-peak anisotropic strength profiles reveal different regimes of brittle-ductile deformation in lithospheric strength beneath oceans and continents, in agreement mineral physical experiments.

Simultaneously analysing these elastic, anelastic and anisotropic variations will improve our understanding of variations in temperature, water content, chemical composition as well as ancient and present deformation within the crust and upper mantle.

ACKNOWLEDGEMENTS

We thank two anonymous reviewers and editor Gabi Laske for comments and suggestion which helped to improve the paper. We acknowledge IRIS (`iris.edu`), ORFEUS (`orfeus-eu.org`) and the Kandilli Observatory (`koeri.boun.edu.tr`) for providing the data used in this study. We thank Daniel Stich for providing Spanish data, which help us to illuminate the Iberian Peninsula. Numerical simulations for this paper were performed on a Dell cluster built and maintained by the Princeton Institute for Computational Science & Engineering (PICSciE). Data and synthetics processing were accomplished based on the Seismic Analysis Code (SAC; Goldstein *et al.* (2002). All maps and cross-sections were made with the Generic Mapping Tool (GMT; Wessel & Smith 1991). The open source spectral-element software package SPECFEM3D_GLOBE and the seismic measurement software package FLEXWIN used for this paper are freely available for download via the Computational Infrastructure for Geodynamics (CIG; `geodynamics.org`). This research is supported by NSF grants 1063057 and 1112906.

REFERENCES

Akçelik, V., Biros, G. & Ghattas, O., 2002. Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propoagation, in *Proceedings of the ACM/IEEE Supercomputing SC'2002 Conference*. Available at: http://www.sc-conference.org/sc2002/. [Published on CD-ROM]

Akçelik, V. *et al.*, 2003. High resolution forward and inverse earthquake modeling on terascale computers, in *Proceedings of the ACM/IEEE Supercomputing SC'2003 Conference*. Available at: http://www.sc-conference.org/sc2003/paperpdfs/pap298.pdf. [Published on CD-ROM]

Allen, R. *et al.*, 2002a. Imaging the mantle beneah Iceland using integrated seismological techniques, *J. geophys. Res.,* **107,** 2325, doi:10.1029/2001JB00595.

Allen, R. *et al.*, 2002b. Plume-driven plumbing and crustal formation in Iceland, *J. geophys. Res.,* **107**(B8), 2163, doi:10.1029/2001JB000584.

Amaru, M., 2007. Global travel time tomography with 3-D reference models, *PhD thesis,* Utrecht University.

Artemieva, I., Thybo, H. & Kaban, M., 2006. Deep Europe today: geophysical synthesis of the upper mantle structure and lithospheric processes over 3.5 Ga, in *European Lithosphere Dynamics,* Vol. 32, pp. 11–41, eds Gee, D. & Stephenson, R., Geol. Soc. London, Special Publication.

Aster, R., Borchers, B. & Thurber, C., 2005. *Parameter Estimation and Inverse Problems,* Elsevier Academic Press.

Babuska, V. & Cara, M., 1991*Seismic Anisotrophy in the Earth,* Kluwer.

Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *EOS, Trans. Am. geophys. Un.,* **81,** 48.

Bercovic, D. & Karato, S., 2003. Whole-mantle convection and the transition-zone water filter, *Nature,* **425,** 39–44.

Billien, M., Leveque, J. & Trampert, J., 2000. Global maps of Rayleigh wave attenuation for periods between 40 and 150 seconds, *Geophys. Res. Lett.,* **27,** 3619–3622.

Bird, P., 2003. An update digital model of plate boundaries, *Geochem. Geophys. Geosyst.,* **4**(3), doi:10.1029/2001GC000252.

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

Boschi, L., Fry, B., Ekström, G. & Giardini, D., 2009. The European upper mantle as seen by surface waves, *Surv. Geophys.,* **30,** 463–501.

Bozdag, E. & Trampert, J., 2008. On crustal corrections in surface wave tomography, *Geophys. J. Int.,* **172,** 1066–1082.

Brossier, R., Operto, S. & Virieux, J., 2009. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion, *Geophysics,* **74,** WCC105–WCC118.

Broyden, C., 1970. The convergence of a class of double-rank minimization algorithms, *J. Inst. Math. Appl.,* **6,** 76–90.

Chang, S.J. *et al.*, 2010a. Joint inversion for three-dimensional S velocity mantle structure along the Tethyan margin, *J. geophys. Res.,* **115,** B08309, doi:10.1029/2009JB007204.

Chang, S.J., der Lee, S.V., Matzel, E. & Bedle, H., 2010b. Radial anisotropy along the Tethyan margin, *Geophys. J. Int.,* **182,** 1013–1024.

Chen, P., Zhao, L. & Jordan, T., 2007. Full 3D tomography for the Crustal structure of the Los Angeles region, *Bull. seism. Soc. Am.,* **97,** 1094–1120.

Dahlen, F. & Tromp, J., 1998. *Theoretical Global Seismology,* Princeton Univ. Press.

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

Dalton, C., Ekström, G. & Dziewoński, A., 2008. The global attenuation structure of the upper mantle, *J. geophys. Res.,* **113,** B09303, doi:10.1029/2007JB005429.

Debayle, E., Kennett, B. & Priestiey, K., 2005. Global azimuthal seismic anisotropy and the unique plate-motion deformation of Australia, *Nature,* **433,** 509–512.

Dercourt, J., 1986. Geological evolution of the Atlantic to the Pamirs sine the Lias, *Tectonophysics,* **123,** 241–315.

Dewey, J., Helman, M., Turco, E., Hutton, D. & Knott, S., 1989. Kinematics of the western Mediterranean, in *Alpine Tectonics,* Vol. 45, pp. 265–283, eds Coward, M.P., Dietrich, D. & Park, R.G., Geol. Soc. Special Publication.

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

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

Endrun, B., Lebedev, S., Meier, T., Tirel, C. & Friederich, W., 2011. Complex layered deformation within the Aegean crust and mantle revealed by seismic anisotropy, *Nature Geosci.,* **4,** 203–207.

Fichtner, A. & Trampert, J., 2011. Hessian kernels of seismic data functionals base upon adjoint techniques, *Geophys. J. Int.,* **185,** 775–798.

Fichtner, A. & Trampert, J., 2012. Resolution analysis in full waveform inversion, *Geophys. J. Int.,* **187,** 1604–1624.

Fichtner, A., Kennett, B., Igel, H. & Bunge, H., 2009. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.,* **179,** 1703–1725.

Fichtner, A., Kennett, B., Igel, H. & Bunge, H., 2010. Full waveform tomography for radially anisotropic structure: new insights into present and past states of the Australasian upper mantle, *Earth planet. Sci. Lett.,* **290,** 270–280.

Fichtner, A., Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y. & Villaseñor, A., 2013. Multiscale full waveform inversion, *Geophys. J. Int.,* **194**(1), 534–556.

Fletcher, R., 1970. A new approach to variable metric algorithms, *Comp. J.,* **13,** 317–322.

Fletcher, R. & Reeves, C., 1964. Function minimization by conjugate gradients, *Comp. J.,* **7,** 149–154.

Foulger, G. *et al.*, 2001. Seismic tomography shows that upwelling beneath Iceland is confined to the upper mantle, *Geophys. J. Int.,* **146,** 504–530.

Gauthier, O., Virieux, J. & Tarantola, A., 1986. Two-dimensional nonlinear inversion of seismic waveforms: numerical results, *Geophysics,* **51,** 1387–1403.

Goes, S., Spakman, W. & Bijwaard, H., 1999. A lower mantle source for central European volcanism, *Science,* **286,** 1928–1931.

Goldfarb, D., 1970. A family of variable metric updates derived by variational means, *Math. Comput.,* **24,** 23–26.

Goldstein, P., Dodge, D., Firpo, M. & Minner, L., 2002. Sac2000: signal processing and analysis tools for seismologists and engineers, *Int. Handbook earthq. Eng. Seismol.,* **81,** 1613–1614.

Gung, Y. & Romanowicz, B., 2004. Q tomography of the upper mantle using three-component long-period waveforms, *Geophys. J. Int.,* **157,** 813–830.

Gung, Y., Panning, M. & Romanowicz, B., 2003. Global anisotropy and the thickness of continents, *Nature,* **422,** 707–711.

Hjörleifsdóttir, V. & Ekström, G., 2010. Effects of three-dimensional Earth structure on CMT earthquake parameters, *Phys. Earth planet. Inter.,* **179,** 178–190.

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

Karato, S., 2011. Water distribution across the mantle transition zone and its implication for global material circulation, *Earth planet. Sci. Lett.,* **301,** 413–423.

Kennett, B.L.N. & Engdahl, E.R., 1991. Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.,* **105,** 429–465.

Kohlstedt, D., Evans, B. & Mackwell, S., 1995. Strength of the lithosphere: constraints imposed by laboratory experiments, *J. geophys. Res.,* **100,** 17 587–17 602.

Kohlstedt, T., Keppler, H. & Rubie, D., 1996. The solubility of water in α, β and γ phases of $(Mg,Fe)_2SiO_4$, *Contrib. Minearl. Petrol,* **123,** 345–357.

Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral-element method for 3-D seismic wave propagation, *Geophys. J. Int.,* **139,** 806–822.

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

Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.,* **150,** 308–318.

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

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

Komatitsch, D., Liu, Q., Tromp, J., Suss, J., Stidham, P. & Shaw, J., 2004. Simulations of ground motion in the Los Angeles basin based upon the spectral-element method, *Bull. seism. Soc. Am.,* **94,** 187–206.

Koulakov, I., Kaban, M., Tesauro, M. & Cloetingh, S., 2009. P- and S-velocity anomaly in the upper mantle beneath Europe from tomographic inversion of ISC data, *Geophys. J. Int.,* **179,** 345–366.

Kustowski, B., Ekström, G. & Dziewonski, A., 2008a. The shear-wave velocity structure in the upper mantle beneath Eurasia, *Geophys. J. Int.,* **174,** 978–992.

Kustowski, B., Ekström, G. & Dziewonski, A., 2008b. Anisotropic shear-wave velocity structure of the Earth's mantle: a global model, *J. geophys. Res.,* **113,** B06306, doi:10.1029/2007JB005169.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migration, in *Conference on Inverse Scattering: Theory and Application,* pp. 206–220, ed. Bednar, J., Society for Industrial and Applied Mathematics.

Laske, G. & Masters, G., 1996. Constraints on global phase velocity maps from long-period polarization data, *J. geophys. Res.,* **101,** 16 059–16 075.

Lawrence, J. & Wysession, M., 2006. Seismic evidence for subduction-transported water in the lower mantle, *Geophys. Monogr.: Earth's Deep Water Cycle,* **168,** 251–261.

Lekic, V. & Romanowicz, B., 2011. Inferring upper-mantle structure by full waveform tomography with the spectral element method, *Geophys. J. Int.,* **185,** doi:10.1111/j.1365–246X.2011.04969.x.

Lekic, V., Panning, M. & Roamnowicz, B., 2010. A simple method for improving crustal correction in waveform tomography, *Geophys. J. Int.,* **182,** 265–278.

Lin, F., Ritzwoller, M., Yang, Y., Moschetti, M. & Fouch, M., 2011. Complex and variable crustal and uppermost mantle seismic anisotropy in the western United States, *Nature Geosci.,* **4,** 55–61.

Lippitsch, R., Kissling, E. & Ansorge, J., 2003. Upper mantle structure beneath the Alpine orogen from high-resolution teleseismic tomography, *J. geophys. Res.,* **108,** doi:10.1029/2002JB002016.

Liu, Q. & Tromp, J., 2006. Finite-frequency sensitivity kernels based upon adjoint methods, *Bull. seism. Soc. Am.,* **96,** 2383–2397.

Liu, Q. & Tromp, J., 2008. Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods, *Geophys. J. Int.,* **174,** 265–286.

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

Liu, Q.Y., Polet, J., Komatitsch, D. & Tromp, J., 2004. Spectral-element moment tensor inversion for earthquakes in Southern California, *Bull. seism. Soc. Am.,* **94,** 1748–1761.

Love, A.E.H., 1927. *A Treatise on the Theory of Elasticity,* Cambridge Univ. Press.

Luo, Y. & Schuster, G., 1991. Wave-equation traveltime inversion, *Geophysics,* **56,** 645–653.

Luo, Y., Tromp, J., Denel, B. & Calandra, H., 2013. 3D coupled acoustic-elastic migration with topography and bathymetry based on spectral-element and adjoint methods, *Geophysics,* **78,** S193–S202.

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

Marone, F. & Romanowicz, B., 2007. The depth distribution of azimuthal anisotropy in the continental upper mantle, *Nature,* **447,** 198–201.

Marone, F., der Lee, S.V. & Giardini, D., 2004. Three-dimensional upper-mantle S-velocity model for the Eurasia-Africa plate boundary region, *Geophys. J. Int.,* **158,** 109–130.

Marquering, H., Nolet, G. & Dahlen, F., 1998. Three-dimensional waveform sensitivity kernels, *Geophys. J. Int.,* **132,** 521–534.

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

Matthies, H. & Strang, G., 1979. The solution of nonlinear finite element equation, *Int. J. Numer. Methods Eng.,* **14,** 1613–1626.

Mitterbauer, U. *et al.*, 2011. Shape and origin of the East-Alpine slab constrained by the ALPASS teleseismic model, *Tectonophysics,* **510,** 195–206.

Molinari, I. & Morelli, A., 2011. EPcrust: a reference crustal model for the European Plate, *Geophys. J. Int.,* **185,** 352–364.

Montagner, J.P. & Anderson, D., 1989. Petrological constrains on seismic anisotropy, *Phys. Earth planet. Inter.,* **54,** 82–105.

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

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

Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics,* **52,** 1211–1228.

Moschetti, M., Ritzwoller, M., Lin, F. & Yang, Y., 2010. Seismic evidence for widespread western-US deep-crustal deformation caused by extension, *Nature,* **464,** 885–889.

Nocedal, J., 1980. Updating quasi-Newton matrices with limited storage, *Math. Comput.,* **35,** 773–782.

Nolet, G., 1990. Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs, *J. geophys. Res.,* **95,** 8499–8512.

Panning, M., Lekic, V. & Romanowicz, B., 2010. The importance of crustal corrections in the development of a new global model of radial anisotropy, *J. geophys. Res.,* **115,** doi:10.1029/2010JB007520.

Park, J. & Levin, V., 2002. Seismic anisotropy: tracing plate dynamics in the mantle, *Science,* **296,** 485–489.

Pasyanos, M., 2005. A variable resolution surface wave dispersion study of Eurasia, North Africa, and surrounding regions, *J. geophys. Res.,* **110,** B12301, doi:10.1029/2005JB003749.

Peter, D. *et al.*, 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.,* **186,** 721–739.

Piromallo, C. & Morelli, A., 2003. P wave tomography of the mantle under the Alpine-Mediterranean area, *J. geophys. Res.,* **108,** doi:10.1029/2002JB001757.

Pratt, R., Shin, C. & Hicks, G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophys. J. Int.,* **133,** 341–362.

Rickers, F., Fichtner, A. & Trampert, J., 2013. The Iceland-Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: evidence from full-waveform inversion, *Earth planet. Sci. Lett.,* **367,** 39–51.

Ritsema, J., Deuss, A., Heijst, H. & Woodhouse, J., 2011. S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.,* **184,** 1223–1236.

Romanowicz, B., 1995. A global tomographic model of shear attenuation in the upper mantle, *J. geophys. Res.,* **100,** 12 375–12 394.

Ruan, Y. & Zhou, Y., 2010. The effects of 3-D anelasticity (Q) structure on surface wave phase delays, *Geophys. J. Int.,* **181,** 479–492.

Ruan, Y. & Zhou, Y., 2012. The effects of 3-D anelasticity (Q) structure on surface wave amplitudes, *Geophys. J. Int.,* **189,** 967–983.

Schivardi, R. & Morelli, A., 2009. Surface wave tomography in the European and Mediterranean region, *Geophys. J. Int.,* **177,** 1050–1066.

Schivardi, R. & Morelli, A., 2011. Epmantle: a 3-D transversely isotropic model of the upper mantle under the European Plate, *Geophys. J. Int.,* **185,** 469–484.

Schmid, C., der Lee, S.V., VanDecar, J.C., Engdahl, E.R. & Giardini, D., 2008. Three-dimensional S velocity of the mantle in the Africa-Eurasia plate boundary region from phase arrival times and regional waveforms, *J. geophys. Res.,* **113,** B03306, doi:10.1029/2005JB004193.

Shanno, D., 1970. Conditioning of quasi-newton methods for function minimization, *Math. Comput.,* **24,** 675–656.

Shapiro, N. & Ritzwoller, M., 2002. Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle, *Geophys. J. Int.,* **151,** 88–105.

Shapiro, N., Ritzwoller, M., Molnar, P. & Levin, V., 2004. Thinning and flow of Tibetan crust constrained by seismic anisotropy, *Science,* **305,** 233–236.

Sieminski, A., Liu, Q.Y., Trampert, J. & Tromp, J., 2007a. Finite-frequency sensitivity of body waves to anisotropy based upon adjoint methods, *Geophys. J. Int.,* **171,** 368–389.

Sieminski, A., Liu, Q.Y., Trampert, J. & Tromp, J., 2007b. Finite-frequency sensitivity of surface wave to anisotropy based upon adjoint methods, *Geophys. J. Int.,* **168,** 1153–1174.

Silver, P., 1996. Seismic anisotropy beneath the continents: probing the depths of geology, *Annu. Rev. Earth planet. Sci.,* **24,** 385–432.

Silver, P. & Chan, W., 1991. Shear wave splitting and subcontinental mantle deformation, *J. geophys. Res.,* **96,** 16 429–16 454.

Simmons, N., Myers, S., Johannesson, G. & Matzel, E., 2012. Llnl-g3dv3: global P wave tomography model for improved regional and teleseismic travel time prediction, *J. geophys. Res.,* **117,** doi:10.1029/2012JB009525.

Simons, F., Van der Hilst, R., Montagner, J. & Zielhuis, A., 2002. Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle, *Geophys. J. Int.,* **151,** 738–754.

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

Spakman, W., 1986. Subduction beneath Eurasia in connection with the Mesozoic Tethys, *Geologie en Mijnbouw,* **65,** 145–153.

Spakman, W., 1990. Images of the upper mantle of central Europe and the Mediterranean, *Terra Nova,* **2,** 542–553.

Spakman, W., 1991. Delay-time tomography of the upper mantle below Europe, the Mediterranean and Asia minor, *Geophys. J. Int.,* **107,** 309–332.

Spakman, W., van der Lee, S. & van der Hilst, R., 1993. Travel-time tomography of the European-Mediterranean mantle down to 1400 km, *Phys. Earth planet. Inter.,* **79,** 3–74.

Tape, C., Liu, Q. & Tromp, J., 2007. Finite-frequency tomography using adjoint methods: methodology and examples using membrane surface waves, *Geophys. J. Int.,* **168,** 1105–1129.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the southern California crust, *Science,* **325,** 988–992.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.,* **180,** 433–462.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics,* **49,** 1259–1266.

Tarantola, A., 2005. *Inverse Problem Theory and Method for Model Parameter Estimation,* SIAM.

Tesauro, M., Kaban, M. & Cloetingh, S., 2008. Eucrust-07: a new reference model for the European crust, *Geophys. Res. Lett.,* **35,** L05313, doi:10.1029/2007GL032244.

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

Tromp, J. *et al.*, 2010. Near real-time simulations of global CMT earthquakes, *Geophys. J. Int.,* **183,** 381–389.

Vinnik, L., Makeyeva, L., Milev, A. & Usenko, A., 1992. Global patterns of azimuthal anisotropy and deformations in the continental mantle, *Geophys. J. Int.,* **111,** 433–447.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics,* **74,** WCC1–WCC26.

Waldhauser, F., Lippitsch, R., Kissling, E. & Ansorge, J., 2002. High-resolution teleseismic tomography of upper-mantle structure using an a priori three-dimensional crustal model, *Geophys. J. Int.,* **150,** 403–414.

Weidle, C. & Maupin, V., 2008. An upper-mantle S-wave velocity model for Northern Europe from Love and Rayleigh group velocities, *Geophys. J. Int.,* **175,** 1154–1168.

Wessel, P. & Smith, W., 1991. Free software helps map and display data, *EOS, Trans. Am. geophys. Un.,* **72**(41), 441–446.

Wiens, D., Conder, J. & Faul, U., 2008. The seismic structure and dynamics of the mantle wedge, *Annu. Rev. Earth Planet. Sci.,* **36,** 421–455.

Wortel, M.J.R. & Spakman, W., 2000. Subduction and slab detachment in the Mediterranean-Carpathian region, *Science,* **290,** 1910–1917.

Yang, Y., Ritzwoller, M., Levshin, A. & Shapiro, N., 2007. Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.,* **168,** 259–274.

Yuan, H. & Romanowicz, B., 2010. Lithospheric layering in the North American craton, *Nature,* **466,** 1063–1068.

Zhang, S. & Karato, S., 1995. Lattice preferred orientation of olivine aggregates deformed in simple shear, *Nature,* **375,** 774–777.

Zhang, H., Thurber, C., Shelly, D., Ide, S., Beroza, G. & Hasegawa, A., 2004. High-resolution subducting-slab structure beneath northern Honshu, Japan, revealed by double-difference tomography, *Geology,* **32,** 361–364.

Zhou, Y., Dahlen, F. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.,* **158,** 142–168.

Zhu, H. & Tromp, J., 2013. Mapping tectonic deformation in the crust and upper mantle beneath Europe and the North Atlantic Ocean, *Science,* **341,** 871–875.

Zhu, H., Bozdag, E., Peter, D. & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography, *Nature Geosci.,* **5,** 493–498.

Zhu, H., Bozdag, E., Duffy, T. & Tromp, J., 2013. Seismic attenuation beneath Europe and the North Atlantic: implications for water in the mantle, *Earth planet. Sci. Lett.,* **381,** 1–11.

Ziegler, P., 1992. European Cenozoic rift system, *Tectonophysics,* **208,** 91–111.

Zielhuis, A. & Nolet, G., 1994. Deep seismic expression of an ancient plate boundary in Europe, *Science,* **265,** 79–81.

APPENDIX: ANISOTROPIC MODEL PARAMETERS AND SENSITIVITY KERNELS

A1. Radial anisotropy

According to section 8.9 of Dahlen & Tromp (1998), the relationships between radially anisotropic wave speeds and the Love parameters (Love 1927) are

$$\alpha_{\rm h} = \sqrt{\frac{A}{\rho}},$$

$$\alpha_{\rm v} = \sqrt{\frac{C}{\rho}},$$

$$\beta_{\rm h} = \sqrt{\frac{N}{\rho}},$$

$$\beta_{\rm v} = \sqrt{\frac{L}{\rho}},$$

$$\eta = \frac{F}{A-2L}. \tag{A1}$$

According to appendix B1 of Sieminski *et al.* (2007a), sensitivity kernels for radially anisotropic wave speeds are related to the Love parameter sensitivity kernels via

$$\mathrm{K}_{\alpha\rm h} = 2A\,\mathrm{K}_A + 2A\eta\,\mathrm{K}_F,$$

$$\mathrm{K}_{\alpha\rm v} = 2C\,\mathrm{K}_C,$$

$$\mathrm{K}_{\beta\rm h} = 2N\,\mathrm{K}_N,$$

$$\mathrm{K}_{\beta\rm v} = 2L\,\mathrm{K}_L - 4L\eta\,\mathrm{K}_F,$$

$$\mathrm{K}_\eta = F\,\mathrm{K}_F. \tag{A2}$$

If the isotropic bulk sound wave speed, c, is chosen as a model parameter instead of compressional wave speed, the relationships (A2) may be rewritten as

$$\mathrm{K}_c = \frac{c^2}{\alpha_{\rm h}^2}\mathrm{K}_{\alpha\rm h} + \frac{c^2}{\alpha_{\rm v}^2}\mathrm{K}_{\alpha\rm v},$$

$$\mathrm{K}_{\beta_{\rm h}'} = \mathrm{K}_{\beta\rm h} + \frac{4}{3}\frac{\beta_{\rm h}^2}{\alpha_{\rm h}^2}\mathrm{K}_{\alpha\rm h},$$

$$\mathrm{K}_{\beta_{\rm v}'} = \mathrm{K}_{\beta\rm v} + \frac{4}{3}\frac{\beta_{\rm v}^2}{\alpha_{\rm v}^2}\mathrm{K}_{\alpha\rm v},$$

$$\mathrm{K}_\eta = F\,\mathrm{K}_F. \tag{A3}$$

A2. Azimuthal anisotropy

For the azimuthally anisotropic model parameters discussed in Section 10.2, relationships between L, N, G_c and G_s and the elastic tensor C_{ij} may be found in Sieminski *et al.* (2007b), namely

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

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

$$G_c = \frac{1}{2}(C_{55} - C_{44}),$$

$$G_s = -C_{45}. \tag{A4}$$

Sensitivity kernels with respect to the above four anisotropic model parameters can be derived from the sensitivity kernels for the elastic tensor C_{ij} via

$$\mathrm{K}_L = \mathrm{K}_{c_{44}} + \mathrm{K}_{c_{55}},$$

$$\mathrm{K}_N = \mathrm{K}_{c_{66}} - 2\mathrm{K}_{c_{12}},$$

$$\mathrm{K}_{G_c} = \mathrm{K}_{c_{55}} - \mathrm{K}_{c_{44}},$$

$$\mathrm{K}_{G_s} = -\mathrm{K}_{c_{45}}. \tag{A5}$$

Sensitivity kernels with respect to dimensionless model parameters $\delta\ln\beta_{\rm v}$, $\delta\ln\beta_{\rm h}$, G_c' and G_s' may be expressed as

$$\mathrm{K}_{\beta_{\rm v}} = 2L\,\mathrm{K}_L - 4L\eta\,\mathrm{K}_F,$$

$$\mathrm{K}_{\beta_{\rm h}} = 2N\,\mathrm{K}_N,$$

$$\mathrm{K}_{G_c'} = \rho\beta_0^2\,\mathrm{K}_{G_c},$$

$$\mathrm{K}_{G_s'} = \rho\beta_0^2\,\mathrm{K}_{G_s}, \tag{A6}$$

where $\eta = F/(A-2L)$.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

(http://gji.oxfordjournals.org/lookup/suppl/doi:10.1093/gji/ggu492/-/DC1)

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.