

11 Global Imaging of the Earth's Deep Interior: Seismic Constraints on (An)isotropy, Density and Attenuation

JEANNOT TRAMPERT AND ANDREAS FICHTNER

Department of Earth Sciences, Utrecht University, Utrecht, The Netherlands

Summary

Seismic tomography is our principal tool to probe the deep interior of the Earth. Over the past decades, it has drawn the picture of a vigorously convecting planet, with large-scale up- and down-wellings convincingly imaged by isotropic velocity variations. Models of seismic anisotropy induced by crystal alignment provide insight into the underlying convective motion, and variations of density allow us to discriminate between thermal and compositional heterogeneities.

However, despite substantial progress, enormous challenges remain: The strength of imaged anisotropy trades off nearly perfectly with the roughness of isotropic heterogeneity, which complicates any quantitative interpretation. Density variations in the Earth, except for spherical harmonic degrees 2, 4 and 6, are still largely unknown, and error estimates are dominated by subjective regularization. Concerning attenuation, any pair of global 3D models is uncorrelated, with mutual differences larger than the attributed uncertainties.

Most of these difficulties can be traced to the complicated nature of multi-observable/multi-parameter inverse problems that suffer from weak constraints and complex trade-offs. Promising directions to strengthen constraints and reduce trade-offs, include (1) the addition of new data at both the short-period (scattered waves) and

long-period (mode splitting) ends of the spectrum, (2) the implementation of fully probabilistic approaches, (3) the joint inversion of phase and amplitude information for elastic and anelastic structure, (4) the design of misfit functionals targeted at specific aspects of Earth structure, and (5) the continuing development of full waveform inversion techniques. Furthermore, much progress can be made by objectively quantifying the uncertainties of our inferences. This can lead to both more consistent models of the Earth and more reliable interpretations in terms of its thermo-chemical structure and evolution.

11.1 Introduction

Seismic tomography is central to the multi-disciplinary effort aimed at understanding the structure and thermo-chemical evolution of the Earth. Indeed, as seismic waves produced by earthquakes travel through the deep Earth, they absorb information from the (an)elastic and density structure along their way which can be extracted by seismic imaging. Compared to geological time scales, seismic waves record an instantaneous snapshot of the current Earth, although anisotropy provides some constraints on geological events in the Earth's recent history. The process of absorbing the information is a well posed forward problem and currently, with

the help of large computers, state-of-the-art techniques allow us to model quasi-exact wave propagation in quite complex Earth models. Seismic tomography is the corresponding ill-posed inverse problem and still quite a challenge. The main limiting factors are the uneven illumination of the Earth's interior by seismic waves and the computational resources. The former is addressed by using more of the available data, although the nonuniform distribution of earthquakes and seismic stations cannot entirely be undone. The computational burden is reduced by implementing various theoretical approximations during imaging.

The most common approach to seismic tomography is based on travel times of identified body and surface waves, which are inverted for velocities using approximations of elastic isotropy and high frequency wave propagation (ray theory). Recent reviews provide extensive reference lists going back to the beginning of this research field in the 1970s (e.g., Romanowicz, 2003; Trampert & van der Hilst, 2005; Nolet, 2008; Rawlinson *et al.*, 2010). Sometimes, a collection of normal-mode frequencies is added to provide a global long-wavelength constraint on the structure (Ritsema *et al.*, 2011). While these studies have shaped a clear picture of the Earth's internal structure, they have not provided unambiguous information on their origin and evolution. Trampert & van der Hilst 2005 argued that estimates of uncertainty and information on the density structure are essential to distinguish between temperature and chemical origins. A convincing case can further be made that information on intrinsic attenuation and anisotropic parameters significantly helps our quest to understand the thermo-chemical evolution of the Earth (e.g., Karato, 2008). Mantle discontinuities also provide interesting information and is discussed in chapter 10 by Deuss *et al.*

The purpose of this chapter is to review on which class of models there is agreement and where further investigation is required. The main emphasis will be on the imaging itself and not so much its thermo-chemical interpretation, although we will provide some indications. In

turn we will treat the case of isotropic velocity tomography, anisotropy, density and attenuation imaging in decreasing order of consensus. Finally, we will make some suggestions for future work.

11.2 An Introduction to Linearised Inverse Theory

Assume that we have a collection of measurements gathered in a data vector \mathbf{d} . First, we have to decide on a mathematical description of the parameter field we want to image. Often the model parameters are expanded onto a finite number of global or local basis functions, e.g. spherical harmonics, blocks and many others. The coefficients for these basis functions are called the model parameters and are collected in a vector \mathbf{m} . Mathematically, \mathbf{d} and \mathbf{m} belong to vector spaces, the data space D and the model space M , respectively. The mapping from M to D is called the forward operator. It is, in the case of a linear problem, represented by the matrix \mathbf{G} . For instance, if we use ray theory, the data could be travel time residuals, the model parameters constant slownesses in blocks, and a row of \mathbf{G} contains the lengths of a particular ray in each block. Formally, we may write

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{e}. \quad (11.1)$$

The data are not perfect and therefore we added an error vector \mathbf{e} which is implicitly recorded together with the data \mathbf{d} . The inverse problem consists of finding a linear mapping \mathbf{S} from D to M . Since \mathbf{G} is usually not invertible, \mathbf{S} is not the exact inverse of \mathbf{G} , and thus we only find an estimator $\tilde{\mathbf{m}}$ of \mathbf{m} . This is formally written as

$$\tilde{\mathbf{m}} = \mathbf{m}_0 + \mathbf{S}(\mathbf{d} - \mathbf{G}\mathbf{m}_0), \quad (11.2)$$

where we have introduced an optional starting model \mathbf{m}_0 . The interesting question is how our estimated model $\tilde{\mathbf{m}}$ is related to the true model \mathbf{m} that we intended to find. We therefore insert Equation (11.1) into Equation (11.2) to get

$$\tilde{\mathbf{m}} - \mathbf{m}_0 = \mathbf{R}(\mathbf{m} - \mathbf{m}_0) + \mathbf{S}\mathbf{e}, \quad (11.3)$$

where we have defined the resolution operator as $\mathbf{R} = \mathbf{S}\mathbf{G}$. It is more instructive to introduce the target model \mathbf{m} explicitly, and we find that

$$\tilde{\mathbf{m}} - \mathbf{m} = (\mathbf{I} - \mathbf{R})(\mathbf{m}_0 - \mathbf{m}) + \mathbf{S}\mathbf{e}, \quad (11.4)$$

with the identity matrix \mathbf{I} . The last equation clearly shows that our estimated model $\tilde{\mathbf{m}}$ deviates from the target model \mathbf{m} by two terms. The first one is due to imperfect resolution, meaning that the resolution matrix is not equal to the identity matrix ($\mathbf{R} \neq \mathbf{I}$). The second term is the result of data errors propagating into the estimated solution.

There is considerable freedom and choice involved in the construction of the approximate inverse operator \mathbf{S} (e.g., Parker, 1994; Tarantola, 2005). The most general approach starts by assigning a probability density $\sigma_M(\mathbf{m})$ to each model, i.e.

$$\sigma_M(\mathbf{m}) = k\rho_M(\mathbf{m})L(\mathbf{m}), \quad (11.5)$$

where ρ_M is the prior distribution in the model space, L the likelihood function which measures how well the model explains the data within their uncertainty, and k a normalizing constant (e.g., Tarantola, 2005). Assuming that both, data uncertainty and our prior knowledge, can adequately be described by Gaussian distributions, Equation (11.5) takes the form

$$\sigma_M(\mathbf{m}) = k e^{-\frac{1}{2} \chi(\mathbf{m})}, \quad (11.6)$$

with the misfit functional

$$\chi(\mathbf{m}) = (\mathbf{d} - \mathbf{G}\mathbf{m})^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{G}\mathbf{m}) + (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0). \quad (11.7)$$

The superscript T denotes vector transposition, and \mathbf{C}_d and \mathbf{C}_m are the covariance matrices in the data and model space, respectively. On the basis of Equations (11.6) and (11.7) we define our estimator $\tilde{\mathbf{m}}$ as the maximum-likelihood model, i.e. the model that maximizes (11.6) and minimizes (11.7). Requiring that the derivative of χ with respect to \mathbf{m} vanishes at the position of the

maximum-likelihood model $\tilde{\mathbf{m}}$, we find that \mathbf{S} is determined by

$$\mathbf{S} = \left(\mathbf{G}^T \mathbf{G} + \frac{\sigma_d^2}{\sigma_m^2} \mathbf{I} \right)^{-1} \mathbf{G}^T, \quad (11.8)$$

where we assumed, for simplicity, that the covariance matrices are diagonal, i.e. $\mathbf{C}_d = \sigma_d^2 \mathbf{I}$ and $\mathbf{C}_m = \sigma_m^2 \mathbf{I}$. The symbol σ_d denotes the standard deviation of the data uncertainty, and σ_m is the standard deviation of the prior model range. The posterior model covariance is then simply

$$\mathbf{C}_{\tilde{\mathbf{m}}} = \sigma_d^2 \left(\mathbf{G}^T \mathbf{G} + \frac{\sigma_d^2}{\sigma_m^2} \mathbf{I} \right)^{-1}, \quad (11.9)$$

where we note that such explicit expressions can only be obtained on the basis of Gaussian statistics. Equation (11.8) reveals a dilemma in the solution of inverse problems: For most realistic applications, the matrix $(\mathbf{G}^T \mathbf{G} + \frac{\sigma_d^2}{\sigma_m^2} \mathbf{I})$ is badly conditioned or not invertible at all, unless the ratio $\frac{\sigma_d^2}{\sigma_m^2}$ is artificially increased. In this case, the initial variances are used to regularize the inversion – and not to objectively quantify data errors and prior knowledge, as it was originally intended. Decreasing σ_m for the purpose of regularization also reduces the posterior covariance, therefore providing an unrealistically optimistic estimate of the errors in our inferred model $\tilde{\mathbf{m}}$. Expression (11.8) nicely reveals the effect that regularization has on the estimated model $\tilde{\mathbf{m}}$. A large data uncertainty (σ_d large) and a narrow search around the prior model (σ_m small), result in a small approximate inverse \mathbf{S} , and hence little update of \mathbf{m}_0 . This explains why most tomographic inversions recover only a fraction of the amplitudes of the actual heterogeneities. Furthermore, the regularization employed in the construction of \mathbf{S} reduces the resolution, because $\mathbf{R} = \mathbf{S}\mathbf{G}$. Moreover, as seen from Equation (11.4), regularization acts as a trade-off between the error propagation and the imperfect resolution. For a strong regularization, $\mathbf{S}\mathbf{e}$ is small and $(\mathbf{I} - \mathbf{R})(\mathbf{m}_0 - \mathbf{m})$ is large and vice versa. The knowledge of both terms

is needed to assess the quality of $\tilde{\mathbf{m}}$, unless one of them is dominant.

The only way to avoid possibly unrealistic Gaussian distributions and the effects that regularization has on our inferences, is to directly sample the probability distribution (11.5), using, for instance, Monte Carlo techniques (e.g. Sambridge & Mosegaard, 2002) or neural networks (e.g. Meier *et al.*, 2007). These are, however, computationally much more involved than solving the regularized analytical expressions based on Gaussian statistics.

Finally, we note that seismic tomography is in reality not a linear inverse problem because our data generally depend nonlinearly on the properties of the Earth. Within a probabilistic framework, non-linearity can naturally be accommodated in the model space sampling. In deterministic problems, the nonlinearity is most often addressed using a perturbation approach. This involves updating the model iteratively:

$$\mathbf{m}_n = \mathbf{m}_{n-1} + \mathbf{P}_{n-1}(\mathbf{d} - \mathbf{G}_{n-1}\mathbf{m}_{n-1}) + \mathbf{Q}(\mathbf{m}_{n-1} - \mathbf{m}_0), \quad (11.10)$$

where n indicates the iteration step and \mathbf{G}_{n-1} the Fréchet derivatives of the nonlinear forward functional with respect to \mathbf{m}_{n-1} . Commonly used algorithms are, for instance, conjugate gradient or Newton schemes which each define separate operators \mathbf{P}_{n-1} and \mathbf{Q} . It is worth noting that solution (11.5) holds for linear as well as nonlinear problems. In the following, we will refer to these different ways of solving the inverse problem in the context of (an)isotropic and (an)elastic seismic tomography.

11.3 Isotropic Velocity Tomography

Over the last three decades seismic tomography has produced a large number of models with a high degree of overlap as documented in many review articles (e.g., Woodhouse & Dziewonski, 1989; Ritzwoller & Lavelle, 1995; Masters *et al.*, 2000; Romanowicz, 2003; Trampert & van der Hilst, 2005; Rawlinson *et al.*, 2010).

P-velocity models are almost exclusively constructed from large collections of travel time residuals. The uneven distribution of earthquakes and seismic stations implies that mainly tectonically active regions are sampled by these data, and the main features to be imaged are descending lithospheric plates (e.g., van der Hilst *et al.*, 1997). The depth extent of the imaged slabs varies, some reaching the lower-most mantle, others stagnant in the transition zone (Fukao *et al.*, 2009). If S-wave travel time residuals are used, remarkably similar images are retrieved (Grand *et al.*, 1997). Travel time models are often referred to as high-resolution models. This is misleading since the fine block or pixel parametrization, usually employed in the construction of these models, only indicates a potential maximum resolution. The actual achieved resolution is frequently unknown, despite the many proposed synthetic tests.

S-velocity models are more often constructed from long-period waveforms and/or a combination of long-period body wave travel time residuals, surface wave dispersion measurements and normal-mode splitting functions (e.g. Mégnin & Romanowicz, 2000; Kustowski *et al.*, 2008; Ritsema *et al.*, 2011). The main robust features appearing are low-velocity mid-oceanic ridges recognizable down to a depth of 100–150 km (Figure 11.1, left). There is a clear ocean-continent difference disappearing at around 250 km depth. The transition zone and the lower-most mantle are dominated by large low-velocity zones beneath Africa and the Pacific Ocean and higher velocities in a circum-Pacific belt (Figure 11.1, middle and right). These features have been known since the pioneering studies of (Masters *et al.* 1982), (Woodhouse & Dziewonski 1984) and (Dziewonski 1984). They have since been confirmed by virtually all successive studies. These S-velocity models employ data with a more even coverage than high-frequency P-wave travel time residuals, and the parametrization emphasizes the long-wavelength structure by employing low-order spherical harmonic expansions. Overall, the resolution is more even over the globe, but that is not to say that it is known precisely.

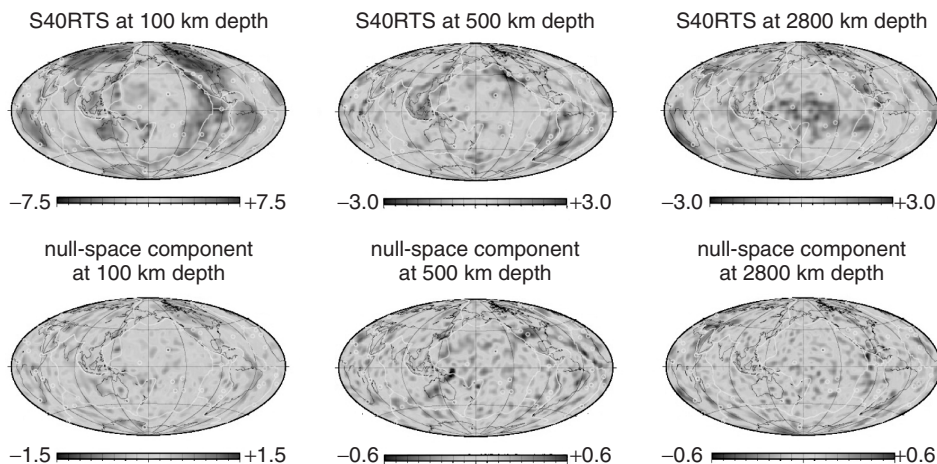


Fig. 11.1 Top Relative S velocity variations, $d \ln v_s$, in the global model S40RTS (Ritsema *et al.*, 2011) at 100, 500 and 2800 km depth. Bottom: The corresponding null-space component $\hat{\mathbf{m}}^{null}$. The null-space component contains short-wavelength structure that can be scaled and added to the model without changing the misfit. (See Color Plate 11).

With a few notable exceptions (e.g., Bijwaard & Spakman, 2000; Panning & Romanowicz, 2006), all classical P- and S-velocity models are obtained by a linear inversion using expression (11.2). The most important tool to assess the amplitude and shape of the models obtained by linearized inversions is the resolution operator, provided that the influence of the data errors is small (Equation 11.3). The latter is the case for most strongly regularized models. The resolution is an operator which tells us how the obtained model parameters are linearly related to each other. Ideally we would like to construct a model such that the resolution is the identity matrix, meaning that the data can constrain all the chosen parameters separately with the correct amplitude. A typical global seismic tomography model consists of few thousand to a few hundred thousand model parameters, and this number squared is the number of entries in the resolution matrix. It is therefore easily understandable that the latter can computationally be a challenge, although not impossible (e.g., Soldati *et al.*, 2006). The importance of calculating the resolution matrix was put forward

for comparing seismic tomography to geodynamic models (e.g., Ritsema *et al.*, 2007), but more often than not simple synthetic tests of the checker board type are used as a proxy for the resolution matrix. (Lévêque *et al.* 1993) convincingly argued that such a simple test has to be interpreted with caution as it can be quite misleading and hardly representative of the true resolution.

As a result of regularization, most resolution tests indicate that only a quarter to a third of the amplitude of heterogeneities is recovered (e.g., Li *et al.*, 2008; Ritsema *et al.*, 2007). This observation is crucial, but often forgotten, when combining seismic tomography and mineral physics data to estimate the thermo-chemical structure of the mantle. Equation-of-state modeling (e.g., Karato & Karki, 2001; Trampert *et al.*, 2001; Stacey & Davis, 2004; Stixrude & Lithgow-Bertelloni, 2005; 2011) of mineral physics data allows us to infer sensitivities (partial derivatives) of velocity variations to temperature and chemical variations. To make the conversion, amplitude and position of the seismic anomalies has to be known with great precision. For instance, if only temperature is changing at a depth of 2800 km,

a relative shear velocity variation $d \ln v_s = 0.01$ would correspond to a temperature variation of the order of 400 K, using the temperature sensitivities of (Trampert *et al.*, 2004), a plausible value from current geodynamic models. The underestimation of amplitude in tomography has profound consequences on this conversion. If we have to multiply the amplitudes by a factor of 3–4, the resulting temperature variation would be 1200–1600 K. This, however, is difficult to reconcile with mantle flow models.

The shape of the anomalies is mainly influenced by the off-diagonal elements of the resolution matrix. They can be very significant and usually oscillate strongly. This is most easily visualized by evaluating the null-space component of a given model. Every model vector can be decomposed into a part that lies in the range of the forward operator \mathbf{G} and in its null-space. The null-space component of a model is defined by finding the part for which $\mathbf{G}\tilde{\mathbf{m}} = 0$. It can be expressed as (Deal & Nolet, 1996)

$$\tilde{\mathbf{m}}^{null} = (\mathbf{I} - \mathbf{R}) \tilde{\mathbf{m}}. \quad (11.11)$$

By definition, this null-space component does not change the fit to the data. In practice, Equation (11.11) is only exact if the estimated model explains the data exactly. For most regularizations, such as in Equation (11.7) where there is a compromise between data fit and model norm, this is not the case. The data can then not distinguish between $\tilde{\mathbf{m}}$ and $\tilde{\mathbf{m}} + \gamma \tilde{\mathbf{m}}^{null}$ for a certain range of scalars γ determined by the data uncertainty, or equivalently the tolerance on the misfit (de Wit *et al.*, 2012). Nevertheless expression (11.11) is useful for illustrating certain properties of the resolution. If it is a scaled version of the original model, \mathbf{R} has little off-diagonal elements and it can be used to directly visualize the amplitude uncertainty mentioned above. If the resolution matrix oscillates, so does the null-space operator $(\mathbf{I} - \mathbf{R})$, and hence $(\mathbf{I} - \mathbf{R})\tilde{\mathbf{m}}$ also has a strong short wave-length component. Figure (11.1) shows, as an example, the null-space component of the shear velocity model S40RTS (Ritsema *et al.*, 2011). It has a significant short wave-length component which is mostly

determined by regularization. A similar result is seen for P velocity models (de Wit *et al.*, 2012).

It appears thus that the amplitude of isotropic wave speed models is significantly underestimated and the smallest wave-lengths in each model strongly distorted. In any thermo-chemical interpretation this should, but has not yet, been taken into account.

11.4 Anisotropic Velocity Tomography

When physical properties depend on the direction in which they are measured, the medium is said to be anisotropic. A medium with properties depending on position only is called isotropic. In seismology we treat the Earth as having isotropic mass density, but velocities are often found to be anisotropic which can easily be modeled with an appropriate elastic tensor, with two independent parameters for an isotropic medium and between 3 and 21 for an anisotropic medium. The effect of anisotropy on elastic wave propagation is three-fold: obviously wave propagation depends on the direction of propagation, it alters the wave polarization away from parallel or perpendicular to the propagation direction and it provokes shear wave splitting, a phenomenon similar to birefringence in optical media. All three properties are readily observed for seismic waves. We will only cite selected examples to show the diversity of observations without the intention of being exhaustive. An excellent overview on theory and observation of seismic wave propagation in anisotropic media is by Maupin and Park (2007).

There is large agreement amongst seismologists that seismic wave propagation is anisotropic. Although in global studies the crust is usually assumed to be isotropic, there are many reports of local wave propagation requiring an anisotropic description of the crust. Seismic lamination observed in many parts of the lower continental crust (e.g., Meissner *et al.*, 2006), polarization anomalies of Lg waves (e.g., Maupin, 1990) and S-wave splitting from local earthquakes (e.g., Paulssen, 2004) are all best explained with crustal anisotropy. Most observations concern waves propagating in the upper mantle or

interpretations concerning the upper mantle only. Significant azimuthal anisotropy has been observed for Pn waves beneath oceans (Hess, 1964) and continents (Smith & Ekström, 1999) as well as for travel time residuals of teleseismic P waves (e.g., Babuška *et al.*, 1998). There are clear observations of polarization anomalies of long period P-waves which have been used to infer upper mantle anisotropy (Schulte-Pelkum *et al.*, 2001), but the biggest wealth of observations comes from SKS-splitting measurements (for a review see e.g., Long & Silver, 2009). Surface waves also exhibit a clear anisotropic behaviour. Both Love and Rayleigh waves show azimuthal dependencies for fundamental modes (e.g., Trampert & Woodhouse, 2003; Ekström, 2011) as well as overtones (Visser *et al.*, 2008). The azimuthally averaged phase velocities of Love and Rayleigh waves see a different vertically averaged elastic structure. This is known as the Love-Rayleigh discrepancy and was first interpreted by (Anderson 1961) in terms of anisotropy. A transverse isotropic medium with vertical symmetry is sufficient to reconcile Love and Rayleigh wave propagation. Most joint inversions of Love and Rayleigh wave phase and/or group velocities for three-dimensional structure therefore employ a transverse isotropic description of the uppermost mantle (e.g., Montagner & Tanimoto, 1991; Gung *et al.*, 2003; Kustowski *et al.*, 2008). There are few clear inferences of anisotropy in the mantle transition zone, but surface wave overtones suggest azimuthal (Trampert & van Heijst, 2002) and radial (Visser *et al.*, 2008) anisotropy. There is a large consensus that the lower mantle is devoid of anisotropy with the exception of D'' (e.g., Maupin *et al.*, 2005; Panning & Romanowicz, 2006). Wave propagation through the inner core is best explained using anisotropy although the details are not fully mapped yet (e.g., Morelli *et al.*, 1986; Woodhouse *et al.*, 1986; Ishii & Dziewonski, 2002; Beghein & Trampert, 2003; Deuss *et al.*, 2010).

In mathematical terms, seismic anisotropy is easily understood. From continuum mechanics we know that if the local elastic tensor has more than two independent parameters,

wave propagation depends on the direction of propagation. With seismic observations we cannot estimate the elastic tensor at a specific point in space, but only spatial averages from tens to thousands of kilometres depending on the waves. This makes the seismic inferences difficult to interpret. From geodynamics and mineral physics modeling we know that anisotropic minerals or isotropic inclusions can align themselves in preferred strain orientations. The former is known as lattice preferred orientation (LPO) and the latter as shape preferred orientation (SPO) (e.g., Karato, 2008). Because the orientation of these crystals and inclusions is not instantaneous, the observation of seismic anisotropy potentially gives us valuable constraints on the geological history of the strain field (e.g., Wenk *et al.*, 2011). As mentioned above, seismic observations cannot image individual crystal or inclusion, but only spatial averages of them. The (an)isotropic description therefore is scale-dependent. Backus (1962) showed that a stack of thin (sub-wavelength) isotropic layers is seen by seismic waves as a large-scale anisotropic medium. This concept can be generalized, and in fact any large scale description of small scale isotropic heterogeneity has to include anisotropy (e.g., Capdeville *et al.*, 2010). While this is a fundamental property of the mathematical description at a limited wavelength, another ambiguity is often forgotten. To infer seismic anisotropy from travel time or polarization anomalies, an inverse problem is usually solved. Starting from Equation (11.1), let us explicitly partition the model parameters into isotropic (ⁱ) and anisotropic (^a) parts. Neglecting data errors and the starting model for notational simplicity, we find

$$\mathbf{d} = (\mathbf{G}^i | \mathbf{G}^a) (\mathbf{m}^i | \mathbf{m}^a)^T. \quad (11.12)$$

From Equation (11.8), we see that the inverse operator will also partition and hence

$$\begin{pmatrix} \tilde{\mathbf{m}}^i \\ \tilde{\mathbf{m}}^a \end{pmatrix} = \begin{pmatrix} \mathbf{S}^i \mathbf{G}^i & \mathbf{S}^i \mathbf{G}^a \\ \mathbf{S}^a \mathbf{G}^i & \mathbf{S}^a \mathbf{G}^a \end{pmatrix} \begin{pmatrix} \mathbf{m}^i \\ \mathbf{m}^a \end{pmatrix} \quad (11.13)$$

The ill-posedness of the inverse problem requires regularization which implies that the resolution

operator is not the identity matrix and hence trade-offs between isotropic and anisotropic parameters are artificially introduced. Partial derivatives for isotropic and anisotropic parameters are different and so are the partial inverse operators, and hence the resolution operator is not symmetric. This means that the trade-off from isotropic to anisotropic parameters, $\mathbf{S}^a \mathbf{G}^i$, is different from anisotropic to isotropic parameters, $\mathbf{S}^i \mathbf{G}^a$. The isotropic-anisotropic trade-off can have large consequences for our interpretation of the inferred anisotropy. If we construct global phase velocity maps using a purely isotropic description, i.e. \mathbf{S}^a forced to $\mathbf{0}$, the models appear blobby. Once anisotropy is introduced, both the isotropic and anisotropic parts of the models appear smooth, with a similar fit to the data (Trampert & Woodhouse, 2003). Hence, the strength of anisotropy trades off with the roughness of isotropic velocity variations. Another example is that significant apparent transverse isotropy is generated if the crustal model is inappropriate (Bozdag & Trampert, 2008), or when apparent splitting is seen on S_{diff} caused by isotropic velocity gradients in D'' (Komatitsch *et al.*, 2010).

We thus have to keep in mind, when seismologists report anisotropy, they explain the data using a description involving the least possible parameters. They implicitly or explicitly employ Occam's razor. In splitting measurements, the razor is implicit as the interpretation assumes a one- or multi-layered medium a priori (e.g., Long & Silver, 2009). A similar situation holds when normal-mode splitting functions are used as the medium is parametrized with a small number of unknowns (e.g., Beghein & Trampert, 2003; Deuss *et al.*, 2010). With surface wave measurements, the razor is more explicit, as we search for a parametrization which explains the data significantly better (e.g., Trampert & Woodhouse, 2003). Due to the wavelength and resolution operator averaging, this is the best the seismologists can do. However, the interpretation of the models evoking LPO or SPO is only meaningful if this averaging is taken into account. There are many studies which show that anisotropy can develop

during convection in the upper (e.g., Kaminski *et al.*, 2004; Becker *et al.*, 2008) and lower mantle (e.g., McNamara *et al.*, 2002). The comparisons to seismology, so far, are qualitative, but can be formulated in a quantitative way using the resolution operator. This is important, because the resolution strongly influences the recovered magnitude of anisotropy and properly accounts for the trade-offs.

This leaves the burning question: is the Earth anisotropic? As argued above, seismologists find that given the observations and the used model parametrization, the data are usually explained more efficiently, i.e. with less parameters, using anisotropy. They present mathematical equivalents of the physical Earth, filtered by a limited frequency band and a nonsymmetric resolution operator. There is consensus that the upper few hundred kilometres of the Earth are transversely isotropic, with decaying amplitude and a possible sign change around 200 km (Figure 11.2). There is also agreement that there is azimuthal anisotropy, although there is little similarity between the inferred models (e.g., Becker *et al.*, 2007). There are many reports of anisotropy in the D'' layer, but the different models tend to have few features in common. There is a large consensus that the inner core is anisotropic, but the details have yet to emerge. We concur with Becker *et al.* (2007) that the reason for this disparity are different strategies in the inverse modelling and we are currently comparing models with widely differing averaging properties. Unless the seismologists come to a similar degree of consensus as for isotropic structures, it seems difficult to infer any geodynamic constraints from the anisotropic models. Anisotropy needs far more parameters for a proper description than isotropy. A possible strategy is therefore to use implicit scalings or fix the symmetry a priori based on mineral physics arguments (e.g., Panning & Romanowicz, 2006; Chevrot, 2006; Long *et al.* 2008; Panning & Nolet, 2008). Another strategy is to be guided by the data and invert for parameters they are most sensitive to (e.g. Sieminski *et al.* 2009). The latter is easier to interpret because it is difficult

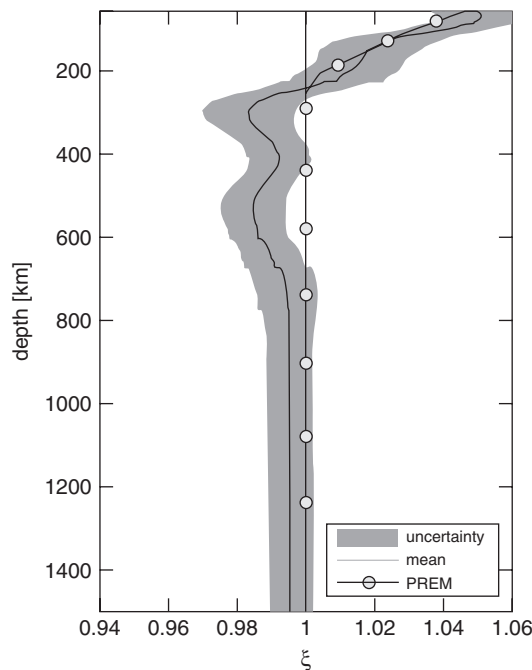


Fig. 11.2 Probability density function of the spherically averaged anisotropic parameter ξ , modified after (Visser *et al.* 2008). Plotted is the mean model and 2 standard deviations around it.

to quantify isotropic-anisotropic trade-offs if the parametrization is fixed a priori.

11.5 Density Tomography

Lateral variations in density are the source of mass transport in the Earth on all scales ranging from the upwelling of salt bodies to the subduction of lithospheric plates and the ascent of super-plumes. They are key to the resolution of long-lasting debates including the distribution of compositional heterogeneity, the nature of continental lithosphere, and the contribution of mantle convection to surface tectonics. However, despite its relevance, the detailed 3D density structure of the Earth remains largely unknown.

In contrast to seismic velocities that can be inferred from travel times, most surface observables

bear little information on density. While being most directly related to density, gravity provides only weak constraints, because solutions to the gravitational inverse problem are inherently non-unique. The use of free air gravity or the geoid is further complicated by the effects of boundary perturbations. Volumetric density heterogeneities by themselves affect gravity, but they also generate topography on the Earth's surface and internal interfaces, such as the core-mantle boundary. The strength of the boundary perturbations depends on viscosity, which is, however, poorly constrained. Furthermore, the volumetric and boundary perturbation effects nearly cancel at mid-mantle depths, which precludes any robust inference on 3D density structure.

The travel times of P and S waves - the major source of information on the Earth's velocity structure - are practically insensitive to density heterogeneities. This surprisingly complete absence of information is closely related to the scattering characteristics of velocity and density perturbations (e.g. Wu & Aki, 1985; Tarantola, 1986). When a body wave impinges upon a velocity perturbation, it generates a scattered wave that travels along with the incident wave. The interference with the scattered wave modifies the shape of the original wave. This modification is then perceived as a travel time shift that can be used in travel time tomography to track the causative velocity perturbation (e.g. Dahlen *et al.*, 2000). Density heterogeneities, in contrast, generate scattered waves that travel *opposite* to the direction of the incident wave. Therefore, the incident and the scattered waves can never interfere, meaning that the original pulse shape remains unaffected. It follows that travel times of the classical body wave phases are not influenced by density perturbations, and that information on density must be sought in the later parts of seismograms. This phenomenon is illustrated in Figure 11.3. We note that the backward scattering off density heterogeneities depends on the parametrization of the Earth model. For instance, choosing ρ together with the elastic parameters λ and μ , instead of ρ , v_p and v_s , leads to forward-scattering off density heterogeneities,

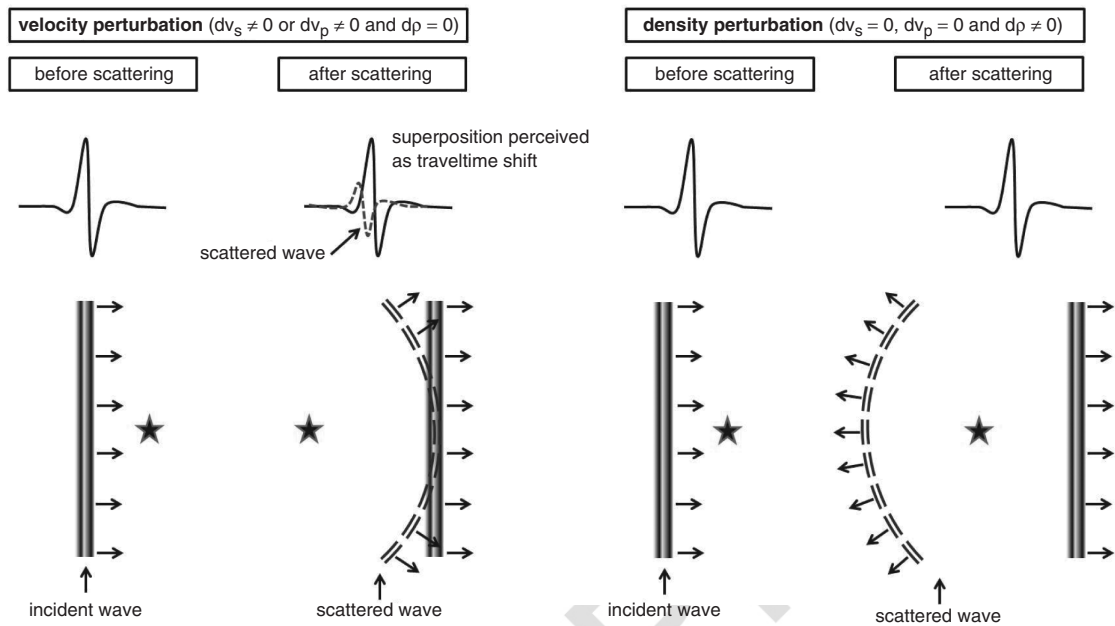


Fig. 11.3 Scattering characteristics of velocity and density perturbations (e.g. Wu & Aki, 1985; Tarantola, 1986).

Left: An incoming body wave impinges upon a velocity perturbation (black star) with either $dv_s \neq 0$ or $dv_p \neq 0$ and $d\rho = 0$. The scattered wave, indicated by the double dashed curve, propagates in the same direction as the incoming wave. The superposition of both waves is perceived as a travel time delay or advance, depending on the sign of the perturbation (e.g. Dahlen *et al.*, 2000). Right: The same as to the left, but for a density perturbation with $dv_s = 0$, $dv_p = 0$ and $d\rho \neq 0$. The scattered wave propagates opposite to the direction of the incoming wave. The pulse shape of the incoming wave remains unperturbed.

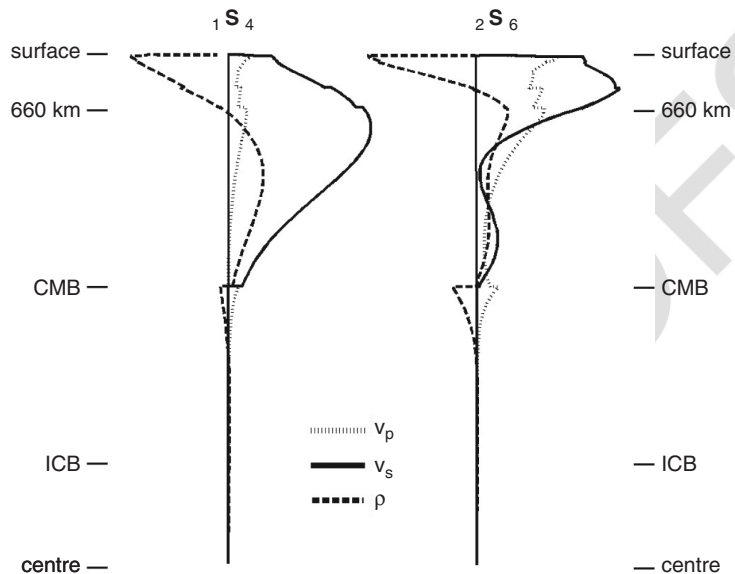
i.e., scattering in the propagation direction of the incoming wave. While this parametrization introduces body wave sensitivity to density, it also leads to undesirable trade-offs between the model parameters, similar to Equation (11.13).

The frequency-dependent travel times of surface waves (dispersion) are mildly sensitive to density variations. Unfortunately, however, this sensitivity tends to be oscillatory. A density perturbation may therefore interact with both positive and negative parts of the sensitivity distribution. The net effect is then nearly zero. Additional complications arise from the strong trade-offs between velocity structure and density in the multi-parameter inverse problem. Perturbations in seismic velocities and density influence surface wave dispersion simultaneously, and the two

effects cannot easily be distinguished. This led many authors to either ignore density variations or to scale v_s to ρ using various multiplicative factors (e.g. Panning & Romanowicz, 2006; Ritsema *et al.*, 2011).

At the long-period end of the seismic spectrum, around 0.5–3.0 mHz, the spheroidal free oscillations of the Earth are sensitive to the long-wavelength density distribution in the whole mantle because of the gravitational restoring force (Figure 11.4). Lateral heterogeneities lead to the splitting of normal-mode eigenfrequencies that can be used for tomography. However, since normal modes result from the constructive interference of waves traveling in opposite directions around the globe, they are primarily sensitive to even-degree structure, because odd-degree

Fig. 11.4 Sensitivity of the frequencies of the spheroidal modes ${}_1S_4$ and ${}_2S_6$ to perturbations in the radial distributions of v_p (dotted), v_s (solid) and ρ (dashed) away from the 1D reference Earth model ak135 (Kennett *et al.*, 1995). The peak amplitude of the density sensitivity is generally smaller than the sensitivity to v_s . Similar to surface waves, the normal-mode sensitivity to density is oscillatory, which further reduces the net effect of large-scale density variations.



contributions tend to cancel. While coupled modes provide some information on odd-degree structure, they have so far not been used in tomographic inversions for density.

The weak sensitivities and strong trade-offs map directly into large uncertainties, even in the radially symmetric density structure that is additionally constrained by the Earth's mass and moment of inertia. Exploring the space of one-dimensional Earth models with Monte Carlo sampling, Kennett (1998) found that density variations of around 1% over 200 km depth intervals are compatible with the frequencies of the gravest spheroidal modes and their associated errors. These uncertainties must be kept in mind when interpreting lateral variations relative to a one-dimensional average.

Compared to the abundance of tomographic inversions for velocity structure, few attempts have been made to exploit the scarce information on density from surface observations.

By far the most common approach to infer 3D density, is to transform tomographic P and S velocity models using a depth-dependent scaling. This approach relies on mineral physics modeling (e.g. Karato, 1993) and works under the

assumption that density heterogeneities are of purely thermal origin. While being a convenient choice in the absence of compositional information, the resulting density models are often inconsistent with fundamental geodynamic observations such as free air gravity anomalies (Forte, 2007). The neglect of compositional contributions to density heterogeneity can furthermore lead to incorrect predictions of mantle flow and the associated surface deformation.

To overcome such inconsistencies, Tondi *et al.* (2000, 2009) jointly inverted seismic travel times and Bouguer anomalies for lithospheric structure by imposing linear relations between velocities and density. The scaling links the two otherwise independent data sets by assuming that a given seismic velocity uniquely specifies density and vice versa. Using inverse problem terminology, the scaling reduces the number of free model parameters, thereby alleviating the inherent non-uniqueness of pure gravity inversions. This approach produces models of the lithosphere that explain both gravity and seismic data sets, but it does not allow for a decorrelation of velocities and density that is likely to result from compositional heterogeneities.

Along similar lines, Simmons *et al.*, 2010 assimilated seismic travel times, the global free air gravity field, the divergence of tectonic plates, dynamic surface topography and topography of the core-mantle boundary into a joint velocity and density model. Their method relies on the radial viscosity profile of the Earth (Mitrovica & Forte, 2004), and on the assumption that lateral heterogeneities are most likely of thermal origin. While being a significant step forward, this approach is limited by the bias towards thermally induced density variations, the large uncertainties in the viscosity profile, and the absence of a formal error analysis that accounts for plausible variations of the forward modeling assumptions.

Recognizing that enforced velocity-density correlations may prevent important inferences on thermo-chemical structure from the outset, Ishii & Tromp (1999, 2001, 2004) inverted normal-mode splitting and free air gravity for a global degree-6 model of velocities and density that were not related a priori. Their results suggest a decorrelation of velocity and density variations, and in particular the presence of high-density piles in the deep mantle beneath the African and Pacific super-plumes.

The robustness of Ishii and Tromp's density models was at the centre of a long debate that reflects the difficulty of solving a deterministic multi-parameter inverse problem that is characterized by weak constraints, strong trade-offs and the nearly complete absence of true physical prior information. Resovsky & Ritzwoller (1999) pointed out that the resolution analysis in a deterministic inverse problem depends on the prior knowledge, i.e. the choice of the prior model covariance σ_m^2 in Equations (11.8) and (11.9). Since we have little prior knowledge on density variations in the Earth, σ_m^2 corresponding to density should be very large. The weakness of our constraints furthermore implies that the entries of the forward modeling or sensitivity matrix \mathbf{G} are small. It follows that the matrix $\mathbf{G}^T \mathbf{G} + \sigma_d^2 \sigma_m^{-2} \mathbf{I}$ in Equation (11.8) may practically not be invertible, unless σ_m^2 is chosen smaller than it actually is on the basis of our prior information. The prior model covariance is then used to regularize

the inversion in a subjective way, and not to represent an objective state of knowledge. When the data constraints are weak, this subjectivity dominates the perceived resolution, expressed, for instance, in terms of the posterior covariance $\mathbf{C}_{\hat{m}}$ (Equation 11.9). Based on a series of test inversions with different choices of the prior information, Resovsky & Ritzwoller 1999 concluded that a decorrelation of S velocity and density could not be detected reliably. A similar conclusion was reached by Romanowicz (2001), who found that density in the lower mantle trades off strongly with the topography of the core-mantle boundary, and that gravity data hardly discriminate between different density models. However, Romanowicz (2001) was able to infer rough bounds for the depth-dependent relationship between the degree-2 structure of density and S velocity. The robustness of 3D density variations in deterministic inversions was furthermore questioned by Kuo & Romanowicz (2002) on the basis of synthetic experiments.

To circumvent the subjectivity of deterministic tomographies, Resovsky & Trampert (2003) worked with a probabilistic formulation of the inverse problem, that does not require explicit regularization to stabilize and reduce trade-offs. Using a combined set of normal-mode and surface-wave data they explored the model space with the help of a neighbourhood algorithm (Sambridge, 1999a; b) to produce probability distributions for velocity and density heterogeneities. These probability densities provide a complete description of our state of knowledge, including uncertainties. The results of Resovsky & Trampert (2003) suggest that long-wavelength variations in v_s and ρ are unlikely to be correlated as much as variations in v_p and v_s anywhere in the mantle. Especially within the transition zone the data favour a $d \ln v_s - d \ln \rho$ anti-correlation. Using mineral physics relations between seismic properties, temperature and composition, Trampert *et al.* 2004 were able to discriminate between thermal and compositional contributions to observed density variations, suggesting that high-density anomalies in the deep mantle (2000–2891 km depth) are most likely of compositional origin.

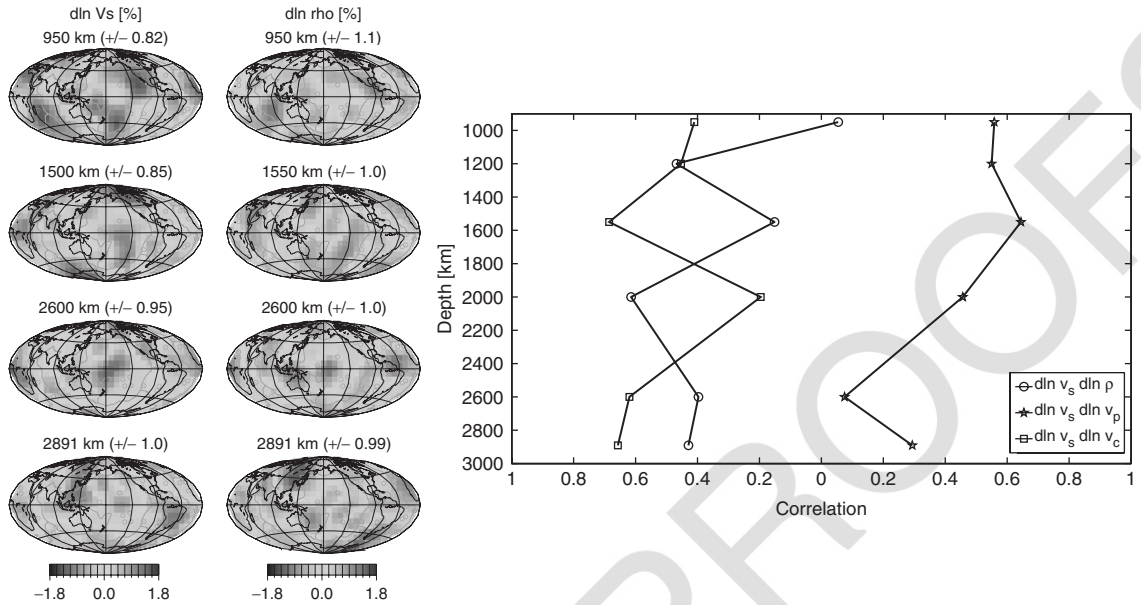


Fig. 11.5 Left: Lateral variations in V_s and ρ at various depths in the maximum-likelihood model of Mosca *et al.* 2012. The laterally averaged standard deviations are indicated in brackets. Note the pronounced anti-correlation of $d\ln V_s$ and $d\ln \rho$ around 2600 km depth beneath the central Pacific and Africa. (See Color Plate 12). Right: Correlation coefficients as a function of depth between the most likely models of $d\ln V_s$, $d\ln \rho$, $d\ln V_p$ and $d\ln V_c$, where V_c denotes the bulk sound velocity. Figure modified after Mosca *et al.* 2012.

Recently, the studies on probabilistic tomography were extended by Mosca *et al.* (2012) who incorporated improved splitting function measurements, as well as a large collection of body wave travel times (figure 11.5).

In the light of the weak constraints, it is not surprising that widely accepted seismological inferences on the Earth's density structure are relatively few in number. That subducting slabs and ascending plumes in the uppermost mantle are comparatively dense or light, respectively, follows from geodynamic arguments, but unequivocal seismic evidence on length scales above degree 6 is missing.

The most reliable results concern the correlation between velocity and density structure at degrees 2, 4 and 6. In the mid-mantle, between 660 and around 1800 km depth, $d\ln V_s$ and $d\ln \rho$ are most likely to be uncorrelated. Within the lower mantle, seismic data prefer a

mild anti-correlation of $d\ln V_s$ and $d\ln \rho$, with a correlation coefficient around -0.4 (see Figure 11.5). The only concrete structural inference relates to the presence of high-density material in the lowermost mantle beneath the African and Pacific superplumes that has been found consistently by several independent studies (e.g. Ishii & Tromp, 2001, 2004), Trampert *et al.*, 2004; Simmons *et al.*, 2010; Mosca *et al.*, 2012). All of these results must be interpreted with caution, given the relatively large standard deviations that range around 50% of the maximum relative density perturbations, even when large data sets of surface waves, normal modes and body waves are combined.

Finally, we note that robust seismic information on the density in the upper mantle is so far missing. This is because structure above 660 km depth cannot be adequately described in terms of the lowest-degree spherical harmonics that we are

able to recover reliably from long-period normal-mode data.

It is becoming increasingly evident that new inferences on the details of Earth's 3D density structure will force us to progress beyond both traditional data analysis and modeling. The accurate measurement of splitting functions on data from recent megathrust earthquakes, combined with proper mode coupling, has the potential to yield more robust information, also about odd-degree density structure (Deuss *et al.*, 2011). Several nonseismological developments may also help to improve our knowledge concerning density. Lateral variations of the Earth's body tides have recently been shown to be above the observational error, and may thus yield additional constraints on global density structure (Latychev *et al.*, 2009). Also on a global scale, neutrino tomography *et al.*, originally conceived in the early 1980s (e.g. De Rujula *et al.*, 1983) – may provide very direct insight into density structure, that is free from the trade-offs that plague seismological approaches (e.g. Gonzalez-Garcia *et al.*, 2008). On small scales (up to a few kilometres) muon transmission tomography has been used successfully to image low-density magma conduits inside active volcanoes (e.g. Tanaka *et al.*, 2003, 2010). Both tidal and elementary particle tomographies are, however, in their infancy.

11.6 Attenuation Tomography

Seismic waves propagating through the Earth are attenuated due to various relaxation mechanisms that lead to the transformation of elastic energy into heat. The large interest in the absorption properties of the Earth, parametrized in terms of the quality factor Q , is mostly related to their temperature dependence. While seismic velocities are quasi-linearly related to temperature, T , the temperature-dependence of Q is well approximated by an exponential Arrhenius-type law (e.g. Jackson, 2000)

$$Q^{-1} \propto e^{-\frac{E}{RT}}, \quad (11.14)$$

where E and R denote an activation energy and the gas constant, respectively. Before delving into the details of the Earth's Q structure, we give a brief summary of the description of seismic wave attenuation, complemented by the most fundamental observations.

11.6.1 Description of seismic wave attenuation, basic observables and observations

Assuming that all seismologically relevant relaxation mechanisms can be modeled by linear rheologies, shear attenuation is commonly described by the frequency-domain visco-elastic stress-strain relation

$$\sigma(\omega) = \tilde{\mu}(\omega)\varepsilon(\omega), \quad (11.15)$$

where $\tilde{\mu}$ denotes the complex shear modulus in the frequency domain. The ratio between the real and imaginary parts of $\tilde{\mu}$ defines the shear quality factor Q_μ :

$$Q_\mu(\omega) = \frac{\Re \tilde{\mu}(\omega)}{\Im \tilde{\mu}(\omega)}. \quad (11.16)$$

The bulk quality factor Q_κ is defined similarly on the basis of the complex bulk modulus $\tilde{\kappa}$. The observability of Q in the Earth is closely related to the amplitude decay of seismic waves. For instance, when $Q_\mu \gg 1$, the amplitude A of a plane S wave propagating through a homogeneous and isotropic medium behaves as

$$A \propto \exp\left(\frac{-\omega x}{2v_s Q_\mu}\right), \quad (11.17)$$

where x denotes the distance traveled. Relation (11.17) is frequently adapted in ray-theoretical studies that replace $x/Q_\mu v_s$ by

$$t^* = \int_C \frac{ds}{Q_\mu(\mathbf{x}) v_s(\mathbf{x})}, \quad (11.18)$$

with C and ds denoting the ray path and a path segment, respectively. Analogous relations can be derived for plane P waves. Most body wave

attenuation tomographies are based on variants of Equations (11.17) and (11.18) because they allow us to eliminate the contribution of the potentially uncertain seismic moment by measuring amplitude ratios of different seismic phases (e.g. Reid *et al.*, 2001; Cheng & Kennett, 2002; Kennett & Abdullah, 2011).

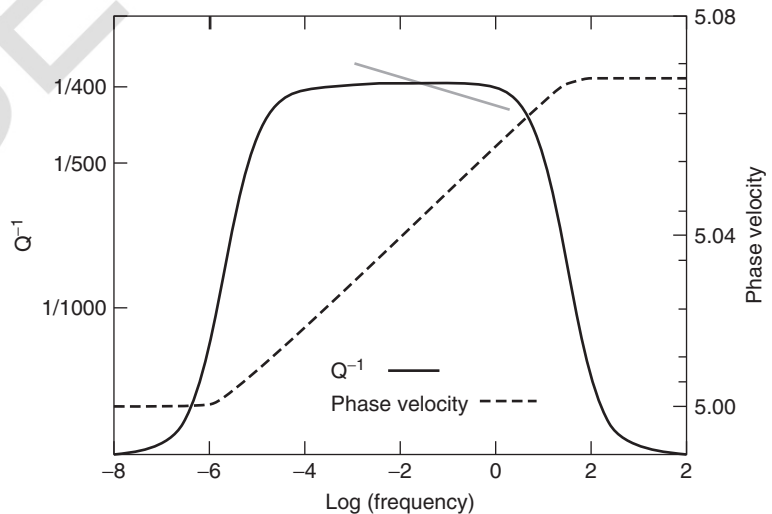
Early studies of Q in the Earth were primarily based on the decay of free oscillation peaks and long-period surface wave amplitudes (e.g. Anderson *et al.*, 1965; Canas & Mitchell, 1978; Anderson & Hart, 1978; Sailor & Dziewonski, 1978). They established that attenuation in shear largely dominates over attenuation associated to bulk deformation, meaning that $Q_\mu \ll Q_\kappa$ almost anywhere in the Earth. Attenuation measurements of the Chandler wobble at sub-seismic frequencies on the one hand, and surface waves on the other hand furthermore indicated that Q_μ is mildly frequency-dependent, with $Q_\mu \propto \omega^\alpha$ and α in the range of 0.1–0.5 (e.g. Anderson & Minster, 1979). These results were confirmed by numerous body wave analyses (e.g. Sipkin & Jordan, 1979; Flanagan & Wiens, 1998; Cheng & Kennett, 2002) and laboratory experiments (e.g. Jackson, 2000, 2007; Karato, 2008). In a recent study,

Lekic *et al.* (2009) developed a method that separates the depth- and frequency-dependencies of normal-mode and surface-wave attenuation measurements. Their results suggest that the effective α in the mantle is negative (≈ -0.4) at periods longer than 1000 s, transitioning to positive values around 0.3 for periods shorter than 500 s.

Despite the unequivocal evidence for a power-law frequency dependence, tomographic inversions mostly assume Q to be constant across the seismic frequency band, i.e. from ≈ 1 to $\approx 10^{-3}$ Hz. This simplification is motivated by the difficulty to robustly constrain Q variations in the Earth even within a narrow frequency band. The constant- Q model is closely related to the notion of a continuous absorption band, i.e. a continuous distribution of relaxation mechanisms that lead to nearly frequency-independent absorption (Liu *et al.*, 1976), as shown in Figure 11.6. This is in contrast to the appearance of isolated absorption peaks that are, for instance, commonly found in metals (e.g. Zener, 1948). Outside the absorption band, Q is predicted to be proportional to ω^{-1} at the low-frequency end, and to ω at the high-frequency end.

A direct consequence of attenuation in general is dispersion, i.e. the frequency-dependence

Fig. 11.6 Phase velocity (dashed) and Q^{-1} (solid) as a function of frequency in an absorption band model with a nearly constant Q of 400 in the frequency band from $\approx 10^{-5}$ Hz to 1 Hz. A frequency-dependent Q within the absorption band is shown schematically in the form of a gray line. A power-law frequency dependence, $Q \propto \omega^\alpha$ with a slightly positive α around 0.3 is consistently found in seismic and laboratory studies, but commonly ignored in tomographic inversions.



of phase velocity (Figure 11.6). In the constant- Q model (Liu *et al.*, 1976), the phase velocity dispersion relation is given by

$$\frac{v_{\text{phase}}(\omega)}{v_{\text{phase}}(\omega_{\text{ref}})} = 1 + \frac{1}{\pi Q} \ln \left(\frac{\omega}{\omega_{\text{ref}}} \right), \quad (11.19)$$

where ω_{ref} is a reference frequency. Equation (11.19) holds, provided that $Q \gg 1$, and that ω lies within the absorption band. Equation (11.19) suggests that phase velocity increases monotonically with increasing frequency; and this effect is most pronounced in highly attenuative media where Q is small.

It is important to make a clear distinction between velocity as a property of a traveling wave, and velocity as a material property. The phase velocity in Equation (11.19) is strictly speaking the propagation speed of a monochromatic plane wave in an isotropic and homogeneous medium. This is not generally identical with a velocity as material property, given by the square root of an elastic modulus divided by density.

11.6.2 The nature of the inverse problem

The dependence of the phase velocity on both Q and ω has profound implications on our ability to infer the absorption properties of the Earth. On the one hand, there is no seismological observable that is sensitive to variations in Q only, while being practically insensitive to variations in any other parameter. On the other hand, there is also no seismological observable that is unaffected by Q . The amplitudes of both long-period (50–200 s) surface waves and intermediate-period (5–30 s) body waves are found to be dominated by focusing induced by 3D variations of purely elastic structure (Sigloch, 2008; Zhou, 2009; Tian *et al.*, 2009). This does not mean that Equation (11.17) is incorrect, but it is a reminder that a proportionality derived from plane wave analysis in homogeneous media should not be mistaken for an equality that holds in the real Earth. The impact of attenuation on travel times should also not be underestimated. Based on the analysis of finite-frequency sensitivity kernels, Zhou (2009)

infers that 15–20% of the observed phase delays in long-period surface waves may in fact be the result of 3D heterogeneity in Q_{μ} . It follows, in conclusion, that any inference on the spatial distribution of Q requires the solution of an intrinsically coupled multi-observable/multi-parameter problem.

The inverse problem is complicated by the difficult nature of seismic wave amplitudes that are frequently used as the only measurement to constrain Q . In addition to attenuation, amplitudes are strongly affected by focusing. Waves traveling through a low-velocity region are focused and amplified, whereas defocusing and amplitude reduction occur in high-velocity zones. Furthermore, amplitude measurements are affected by miscalibrated instruments, scattering and seismic source characteristics. Small-scale heterogeneity in the immediate vicinity of the receiver, such as layers of nearly saturated sediments, may give rise to nonlinear amplitude effects. These factors add frequency-dependent source and receiver corrections to the list of unknowns.

The daunting complexity of the inverse problem for Q explains why substantial simplifications are common. Instead of jointly inverting for all the necessary free parameters, both elastic focusing effects and source/receiver corrections are frequently ignored. While convenient, this approach was shown to produce incorrect estimates of attenuation structure in body and surface wave studies (Dalton *et al.*, 2008; Tian *et al.*, 2009). To isolate the signature of attenuation, several studies proposed to remove the focusing effect from observed amplitudes with the help of 3D elastic velocity models derived from travel time tomography. While being a significant step forward, this strategy suffers from the subjective choice of an elastic model, and its smoothness properties in particular. Overly rough elastic models may lead unrealistically strong focusing, while overly smooth models will compensate the underestimated elastic effect by excessive Q variations. Interestingly, this scale-length problem is similar to the one encountered in anisotropic tomography where the roughness of isotropic velocity variations trades off with the strength of anisotropy. In any case, the resulting combination

of elastic velocity and attenuation models will be incompatible, because the deduced 3D variations in Q were not accounted for in the interpretation of travel times used to derive the velocity model that was initially needed to remove the elastic effect from amplitudes. This discrepancy underlines the intrinsic nonseparability of elastic and anelastic effects that precludes any consistent removal of focusing effects from seismic wave amplitudes.

Both the multi-observable/multi-parameter nature of the inverse problem for Q structure and the inconsistency of many simplifications explain why our progress in attenuation tomography has been rather slow, compared to tomography for elastic structure.

11.6.3 Constraints on the radial attenuation structure

Constraints on the spherically averaged attenuation in the Earth are mostly derived from the amplitude decay of body waves, long-period surface waves and free oscillation peaks.

Estimates of the radial shear attenuation by various authors, summarized in Figure 11.7, are relatively consistent. The only notable exception is the lithosphere where values range between $Q_\mu \approx 200$ (Durek & Ekström, 1996) and $Q_\mu \approx 600$ (Dziewonski & Anderson, 1981; Okal & Jo, 1990). This variability partly reflects the presence of strong heterogeneity in the lithosphere that prevents the determination of a physically meaningful average. Within a narrow asthenospheric layer, Q_μ is generally found to be low (60–90). This drop is followed by a rapid increase through the transition zone to $Q_\mu \approx 300$ near 660 km depth. Within the lower mantle some authors favour a slight decrease (Okal & Jo, 1990; Widmer *et al.*, 1991), while others prefer a constant $Q_\mu \approx 300$ down to the core-mantle boundary (Dziewonski & Anderson, 1981; Durek & Ekström, 1996). Reliable uncertainty estimates for the radial Q_μ profile were provided by Resovsky *et al.* (2005), who applied a probabilistic inversion to normal-mode and surface-wave data. Typically, the standard

deviations were found to range between 10% and 20%, depending on depth.

The distribution of bulk attenuation, required by the decay of radial free oscillations, is the subject of a long-lasting debate, in the course of which a large variety of models have been presented. Some of these are displayed in Figure 11.7. Dziewonski & Anderson (1981) confined a comparatively large bulk attenuation in PREM to the inner core, where $Q_\kappa = 1327.7$ compared to $Q_\kappa = 57823$ anywhere else in the Earth. They noted, however, that “this should be only understood as a way to lower Q of radial modes in order to make them more compatible with observations. The problem is highly non-unique and its early resolution is not likely.” Similar to PREM, Anderson & Hart (1978) and Okal & Jo 1990 locate high bulk attenuation (small Q_κ) in the inner core. In contrast, more recent models based on higher-quality data clearly prefer a distribution of low Q_κ in the upper mantle and the transition zone (Durek & Ekström, 1996), or in the whole mantle and the outer core (Widmer *et al.*, 1991; Resovsky *et al.*, 2005). The variety of Q_κ models is proportional to the ill-posedness of the inverse problem. As in the case of density tomography, the observational constraints on Q_κ are weak, so that the solution of a deterministic inverse problem is dominated by the unavoidable and largely subjective regularization. Any particular value of Q_κ should be interpreted with caution, since plausible values of bulk attenuation may span nearly one order of magnitude (Resovsky *et al.*, 2005).

All of the radial Q_μ models shown in Figure 11.7 are based on the assumption of frequency-independent attenuation, despite the well-known power law relation $Q_\mu \propto \omega^\alpha$ with α mostly between 0.1 and 0.5. Additional free parameters to describe frequency dependence are unlikely to be resolvable. As pointed out by Lekic *et al.* (2009), this simplification may lead to incorrect Q_μ profiles, because higher-frequency waves with higher Q_μ constrain shallower structure, whereas lower-frequency waves with lower Q_μ constrain deeper structure. Thus, attenuation estimates may be biased towards low Q_μ values in the deep mantle and towards high Q_μ values in the upper mantle.

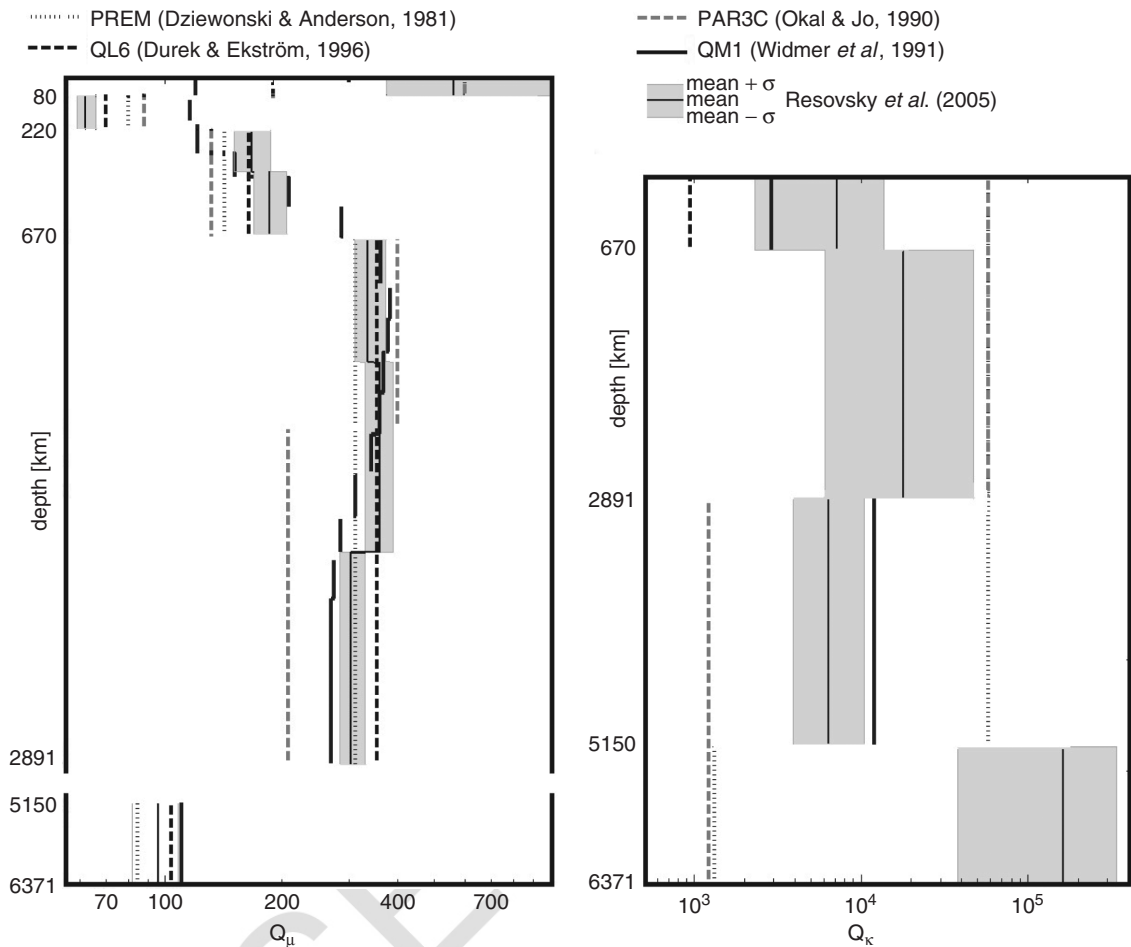


Fig. 11.7 Comparison of the radial Q models PREM (Dziewonski & Anderson, 1981), PAR3C (Okal & Jo, 1990), QM1 (Widmer *et al.*, 1991), QL6 (Durek & Ekström, 1996), and the model by Resovsky *et al.* (2005) with $\pm\sigma$ uncertainties indicated by the light-gray shaded areas.

11.6.4 3D attenuation models

Over the past four decades, numerous regional-scale studies made use of diverse data sets including surface wave amplitudes and differential t^* measurements on pairs of seismic phases. While individual results differ substantially due to the complexity of the inverse problem and the various ways of dealing with it, some robust links between near-surface tectonic features and the

attenuation properties of the uppermost mantle are emerging from the ensemble of available models: Low attenuation, i.e. large Q , is often found within stable continents (e.g. Chan & Der, 1988; Sipkin & Revenaugh, 1994; Tian *et al.*, 2009; Kennett & Abdullah, 2011). This is in contrast to the typically high attenuation in the vicinity of hotspots (e.g. Tian *et al.*, 2009; Kennett & Abdullah, 2011). Along convergent margins, several studies inferred high attenuation in the

mantle wedge above subducting slabs and lower attenuation within the downgoing plate (e.g. Flanagan & Wiens, 1990; Stachnik *et al.*, 2004). A decrease of uppermost mantle attenuation with increasing age of ocean floor was found, for instance, by Canas & Mitchell (1978) and Sheehan & Solomon (1992). These observations suggest that attenuation is primarily temperature-related, at least within the shallow mantle.

Global models of attenuation in the whole mantle have also been derived mostly from long-period surface wave amplitudes and amplitude ratios of various body wave phases. These models differ strongly in the treatment of elastic effects and source/receiver corrections. Romanowicz (1990) and Durek *et al.* (1993), for instance, considered the amplitudes of four consecutive arrivals of multiply-orbiting long-period Rayleigh waves (e.g. R1, R2, R3, R4), in order to reduce the effects of elastic focusing and source uncertainty. While this approach helps to pronounce the contribution of attenuation, it only yields information on even-degree Q_μ structure. An improved data analysis scheme was used by Romanowicz (1994), who selected both R1 and R2 amplitudes that did not appear to be contaminated by source errors and elastic effects. The selection criterion enforced consistency between the attenuation coefficients inferred from four consecutive wave trains, and those measured on minor and major arc amplitudes only. This strategy led to maps of both even- and odd-degree structure, that were used to construct the first global model of shear attenuation (Romanowicz, 1995).

Billien *et al.* (2000) and Selby & Woodhouse (2000; 2002) explicitly treated elastic focusing using a linear approximation derived from ray theory (Woodhouse & Wong, 1986). The contribution from focusing to the amplitudes of Rayleigh waves was found to be considerable, even for long periods between 70 s and 170 s. This motivated Billien *et al.* (2000) to jointly invert phase and amplitude measurements of Rayleigh waves for degree-20 maps of phase velocities and attenuation. However, being concerned that focusing may not be predicted with sufficient accuracy by current velocity models,

Selby & Woodhouse (2002) decided not to account for elastic effects in their inversion for 3D shear attenuation. A similar approach was taken by Gung & Romanowicz (2004) who neglected focusing as well as source/receiver corrections, because these factors seemed to have little effect on their degree-8 model in synthetic tests.

As an alternative to surface waves, Reid *et al.* (2001) estimated t^* from amplitude ratios of globally recorded S, SS and SSS waves. Also neglecting the effect of focusing, the t^* measurements were inverted for a degree-8 model of Q_μ in the upper mantle.

The currently most sophisticated study on whole-mantle shear attenuation was initiated by Dalton & Ekström (2006) who simultaneously inverted Rayleigh wave amplitudes and phase delays for maps of shear attenuation and phase velocity, as well as for amplitude correction factors for each source and receiver. They demonstrated that attenuation estimates depend significantly on each of the amplitude corrections, and that the neglect of elastic focusing translates into inaccurate Q structure. The phase delay maps were then used to remove the focusing effect from the amplitude data, which were then inverted for a global degree-12 Q_μ model (Dalton *et al.*, 2008). More than previous models, the work of Dalton *et al.* (2008) reveals a clear correlation between shallow-mantle attenuation and surface tectonics. In particular, high attenuation is found around 100 km depth beneath the circum-Pacific volcanic arc, the Lau basin, and along most mid-ocean ridges, including the East Pacific Rise, the Indian-Antarctic ridge and the Mid-Atlantic ridge. Precambrian lithosphere is characterized by lower than average attenuation.

In contrast to elastic P or S tomographies, there is little agreement between attenuation models. Dalton *et al.* (2008) compared their global Q_μ model to those of Romanowicz (1995), Reid *et al.* (2001), Warren & Shearer (2002), Selby & Woodhouse (2002) and Gung & Romanowicz (2004). When truncated at degree 8, the correlation of these models was found to be mostly below 0.4 throughout the upper mantle. Furthermore, Dalton *et al.* (2008) noted that global Q_μ models differ

most in the vicinity of mid-ocean ridges where characteristic low-velocity regions can lead to a strong focusing of wave energy that must be accounted for correctly. The strength of the lateral variations in global attenuation models varies by a factor of up to 4, which is significantly above the error estimates of around 10%, given, for instance, by Romanowicz (1995). The various models are therefore, technically speaking, inconsistent. The large discrepancies between Q_μ models and the overly optimistic error estimates are again related to the properties of an ill-posed inverse problem where weak constraints and strong trade-offs lead to a dominance of subjective and often implicit regularization.

Despite the difficulties involved in attenuation tomography, some large-scale robust features can be deduced from the combined analysis of all available global attenuation models. Above 250 km depth, most studies find a correlation between Q_μ and surface tectonics. The most notable consistency is the appearance of high Q_μ values in stable continental interiors, including Australia's Pilbara and Yilgarn cratons, the East European Platform and the Canadian Shield. Below 250 km depth the attenuation pattern changes, and low Q_μ can roughly be associated with the position of hotspots as imaged in isotropic velocity tomographies.

The variability of Q_μ models is due to a strong methodology-dependence that reflects the difficulty of extracting a robust signal from seismic wave amplitudes that can be attributed to attenuation with sufficient confidence. From the current perspective, much progress remains to be made in order to arrive at a stage where attenuation models can be interpreted quantitatively in terms of the Earth's thermo-chemical structure.

11.7 Promising Future Directions

From the above discussions it is clear that global tomography is, in general, a multi-observable/multi-parameter problem. The complicated nature of this problem has to be addressed very carefully, especially when inverting for weakly-constrained and scale-dependent properties such

as anisotropy, density and attenuation. The consensus amongst seismologists decreases from isotropic to anisotropic, density and attenuation models. The reason for this is the nonsymmetric nature of the resolution operator (Equation 11.13). The Fréchet derivatives are such that trade-offs to isotropic parameters are smallest. This explains the considerable overlap between different studies using different approximations and data. Good data exist to image the elastic part of the problem, but density is weakly constrained. The Fréchet derivatives to anelastic parameters are much smaller with common parametrizations and misfit functionals. We see mainly three general future directions: (1) progressing towards full waveform inversion with appropriate resolution analysis, (2) principal component analysis and (3) finding new observables which are sensitive to a limited range of parameters.

Full waveform inversion uses gradient-type implementations of Equation (11.10) with exact Fréchet derivatives computed with purely numerical solutions of the forward problem (Bamberger *et al.*, 1982; Tarantola, 1988; Igel *et al.*, 1996). This approach accounts for the nonlinear relation between structure, phase and amplitudes in the construction of tomographic models, thus progressing beyond linear approximations commonly used. Since the update of the model is just a scaled version of the misfit kernel, there is no explicit regularization required, beyond smoothing the kernels and terminating the inversion after a finite number of iterations. The first regional-to continental-scale seismological applications of full waveform inversion indicate that resolution indeed improves (e.g. Chen *et al.*, 2007; Fichtner *et al.*, 2009; 2010; Tape *et al.*, 2010). In particular, the amplitudes of the heterogeneities increase, and more small-scale features are robustly imaged. We anticipate that full waveform inversion techniques will make particularly valuable contributions to attenuation tomography. The replacement of source/receiver correction factors by more elaborate forward and inverse modeling techniques that correctly account for

local structure would constitute a very significant progress. This is technically possible thanks to advanced numerical wave propagation tools for 3D heterogeneous media (e.g. Moczo *et al.*, 2007; Dumbser *et al.*, 2007; Peter *et al.*, 2011). Potentially, we can calculate the misfit kernels for all parameters simultaneously and update the complete model. Not all parameters will, of course, be equally well determined, and trade-offs exist. This approach will require a detailed resolution analysis. Tools in the framework of full waveform inversion are just becoming available (Fichtner & Trampert, 2011a, b).

While a simultaneous inversion for all possible parameters is tempting, we must realize that not all Fréchet derivatives are equally strong. Principal component analysis (Sieminski *et al.*, 2009) opens the opportunity to determine a priori those linear combinations of elastic parameters that are best constrained by the data set, i.e., that generate the largest sensitivities. For the elastic problem, not more than 6 linear combinations mostly account for about 90% of the total sensitivity, meaning that all remaining linear combinations can hardly be constrained. Thus, rather than updating all elastic parameters, only the first 6 linear combinations of parameters could be updated instead. Principal component analysis is a valuable tool to optimize the design of an inverse problem by finding the maximum number of well-constrained parameters. While the optimal linear combinations of elastic parameters may not always have a specific meaning in terms of wave propagation, they can still be interpreted in a mineral physics context, thereby providing robust insight into the thermochemical and deformation state of the Earth.

A last promising direction is the incorporation of new observables, either in the form of completely new measurements or as specifically designed misfit functionals. Examples of the former include measurements of rotational ground motions (e.g. Ferreira & Igel, 2009), Bernauer *et al.*, 2009), Earth tides (Latychev *et al.*, 2009) and highly accurate long-period mode splitting (Deuss *et al.*, 2011). The design of targeted misfit functionals is based on the realization that not all

parametrizations are equivalent. It is well understood, for instance, that travel times are sensitive to velocities, but hardly to density. If the problem is reformulated with elastic parameters and density instead, the sensitivity to density increases, but now the problem has to be solved for elastic parameters and density simultaneously, with possible trade-offs. In the proposed approach it is further important to realize that the sensitivities are determined by the definition of the misfit functional or the measurement. The idea is then to find a misfit functional which maximizes, or even better, which diagonalizes the sensitivity to the parameter of interest. This can be set up as a design problem in combination with principal component analysis (Sieminski *et al.*, 2009) or in a fully nonlinear fashion (van den Berg *et al.*, 2003).

Most of these directions are certain to significantly increase our imaging capabilities. In the more distant future, sampling directly the posterior (i.e. solving Equation 11.5) including the full nonlinearity of the forward problem, should become feasible with exa-computing. Guided Monte Carlo or Neural Networks (e.g., Meier *et al.*, 2007) seem promising.

References

- Anderson, D. L., 1961. Elastic wave propagation in layered anisotropic media, *J. Geophys. Res.*, **66**, 2953–2963.
- Anderson, D. L., Ben-Menahem, A., & Archambeau, C. B., 1965. Attenuation of seismic energy in the upper mantle, *J. Geophys. Res.*, **70**, 1441–1448.
- Anderson, D. L. & Hart, R. S., 1978. Q of the Earth, *J. Geophys. Res.*, **83**, 5869–5882.
- Anderson, D. L. & Minster, J. B., 1979. The frequency dependence of Q in the Earth and implications for mantle rheology and the Chandler wobble, *Geophys. J. Roy. Astr. Soc.*, **58**, 431–440.
- Babuška, V., Montagner, J., Plomerov, J., & Girardin, N., 1998. Age-dependent large-scale fabric of the mantle lithosphere as derived from surface-wave velocity anisotropy, *Pure and Applied Geophysics*, **151**(2-4), 257–280.

- Backus, G. E., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. Geophys. Res.*, **67**, 4427–4440.
- Bamberger, A., Chavent, G., Hemons, C., & Lailly, P., 1982. Inversion of normal incidence seismograms, *Geophysics*, **47**, 757–770.
- Becker, T. W., Ekström, G., Boschi, L., & Woodhouse, J. H., 2007. Length scales, patterns and origin of azimuthal seismic anisotropy in the upper mantle as mapped by Rayleigh waves, *Geophys. J. Int.*, **171**(1), 451–462.
- Becker, T. W., Kustowski, B., & Ekström, G., 2008. Radial seismic anisotropy as a constraint for upper mantle rheology, *Earth Planet. Sci. Lett.*, **267**(1-2), 213–227.
- Beghein, C. & Trampert, J., 2003. Robust normal mode constraints on inner-core anisotropy from model space search, *Science*, **299**(5606), 552–555.
- Bernauer, M., Fichtner, A., & Igel, H., 2009. Inferring Earth structure from combined measurements of rotational and translational ground motions, *Geophysics*, **74**, WCD41–WCD47.
- Bijwaard, H. & Spakman, W., 2000. Non-linear global P-wave tomography by iterated linearized inversion, *Geophys. J. Int.*, **141**(1), 71–82.
- Billien, M., Leveque, J.-J., & Trampert, J., 2000. Global maps of Rayleigh wave attenuation for periods between 40 and 150 seconds, *Geophys. Res. Lett.*, **27**, 3619–3622.
- Bozdag, E. & Trampert, J., 2008. On crustal corrections in surface wave tomography, *Geophys. J. Int.*, **172**(3), 1066–1082.
- Canas, J. A. & Mitchell, B. J., 1978. Lateral variations of surface wave attenuation across the Pacific, *Bull. Seis. Soc. Am.*, **68**, 1637–1650.
- Capdeville, Y., Guillot, L., & Marigo, J. J., 2010. 2D nonperiodic homogenization to upscale elastic media for P-SV waves, *Geophys. J. Int.*, **182**, 903–922.
- Chan, W. W. & Der, Z. A., 1988. Attenuation of multiple ScS in various parts of the world, *Geophys. J.*, **92**, 303–314.
- Chen, P., Zhao, L., & Jordan, T. H., 2007. Full 3D tomography for the crustal structure of the Los Angeles region, *Bull. Seismol. Soc. Am.*, **97**, 1094–1120.
- Cheng, H.-X. & Kennett, B. L. N., 2002. Frequency dependence of seismic wave attenuation in the upper mantle beneath the Australian region, *Geophys. J. Int.*, pp. 45–57.
- Chevrot, S., 2006. Finite-frequency vectorial tomography: A new method for high-resolution imaging of upper mantle anisotropy, *Geophys. J. Int.*, **165**(2), 641–657.
- Dahlen, F., Hung, S.-H., & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes – I. Theory, *Geophys. J. Int.*, **141**, 157–174.
- Dalton, C. A. & Ekström, G., 2006. Global models of surface wave attenuation, *J. Geophys. Res.*, **111**, doi:10.1029/2005JB003997.
- Dalton, C. A., Ekström, G., & Dziewonski, A. M., 2008. The global attenuation structure of the upper mantle, *J. Geophys. Res.*, **113**, doi:10.1029/2007JB005429.
- De Rujula, A., Glashow, S. L., Wilson, R. R., & Chorpak, G., 1983. Neutrino exploration of the Earth, *Phys. Rep.*, **99**, 341–396.
- de Wit, R., Trampert, J., & van der Hilst, R., 2012. Towards quantifying uncertainty in travel time tomography using the null space shuttle, *J. Geophys. Res.*, **117**, B03301, doi:10.1029/2011JB008754.
- Deal, M. M. & Nolet, G., 1996. Nullspace shuttles, *Geophys. J. Int.*, **124**(2), 372–380.
- Deuss, A., Irving, J., & Woodhouse, J., 2010. Regional variation of inner core anisotropy from seismic normal mode observations, *Science*, **328**, 1018–1020.
- Deuss, A., Ritsema, J., & van Heijst, H. J., 2011. Splitting function measurements for Earth's longest-period normal-modes using recent earthquakes, *Geophys. Res. Lett.*, **38**, L04303, doi:10.1029/2010JB046115.
- Dumbser, M., Käser, M., & de la Puente, J., 2007. Arbitrary high-order finite volume schemes for seismic wave propagation on unstructured meshes in 2D and 3D, *Geophys. J. Int.*, **171**, 665–694.
- Durek, J. J. & Ekström, G., 1996. A radial model of anelasticity consistent with long-period surface wave attenuation, *Bull. Seis. Soc. Am.*, **86**, 144–158.
- Durek, J. J., Ritzwoller, M. H., & Woodhouse, J. H., 1993. Constraining upper mantle anelasticity using surface wave amplitude anomalies, *Geophys. J. Int.*, **114**, 249–272.
- Dziewonski, A. M., 1984. Mapping the lower mantle: Determination of lateral heterogeneity in P velocity up to degree and order 6, *J. Geophys. Res.*, **89**, 5929–5952.
- Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, **25**, 297–356.
- Ekström, G., 2011. A global model of Love and Rayleigh surface-wave dispersion and anisotropy, 25–250 seconds, *Geophys. J. Int.*, **187**, 1668–1686.

- Ferreira, A. M. G. & Igel, H., 2009. Rotational motions of seismic surface waves in a laterally heterogeneous Earth, *Bull. Seis. Soc. Am.*, **99**, 1429–1436.
- Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H.-P., 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. L. N., Igel, H., & Bunge, H.-P., 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., 2011a. Hessian kernels of seismic data functionals based upon adjoint techniques, *Geophys. J. Int.*, **185**, 775–798.
- Fichtner, A. & Trampert, J., 2011b. Resolution analysis in full waveform inversion, *Geophys. J. Int.*, **187**, 1604–1664.
- Flanagan, M. P. & Wiens, D. A., 1990. Attenuation structure beneath the Lau backarc spreading center from teleseismic S phases, *Geophys. Res. Lett.*, **17**, 2117–2120.
- Flanagan, M. P. & Wiens, D. A., 1998. Attenuation of broadband P and S waves in Tonga, *Pure. Appl. Geophys.*, **153**, 345–375.
- Forte, A., 2007. Constraints on seismic models from other disciplines – implications for mantle dynamics and composition, in: *Treatise of Geophysics*, **1**, 805–858.
- Fukao, Y., Obayashi, M., & Nakakuki, T., 2009. Stagnant slab: A review, *Annu. Rev. Earth Planet. Sci.*, **37**, 19–46.
- Gonzalez-Garcia, M. C., Halzen, F., Maltoni, M., & Tanaka, H. K. M., 2008. Radiography of Earth's core and mantle with atmospheric neutrinos, *Phys. Rev. Lett.*, **100**, doi:10.1103/PhysRevLett.100.061802.
- Grand, S., van der Hilst, R., & Widiyantoro, S., 1997. High resolution global tomography: a snapshot of convection in the Earth, *GSA Today*, **7**, 1–7.
- Gung, Y., Panning, M., & Romanowicz, B., 2003. Global anisotropy and the thickness of continents, *Nature*, **422**(6933), 707–711.
- Gung, Y. C. & Romanowicz, B., 2004. Q tomography of the upper mantle using three component long period waveforms, *Geophys. J. Int.*, **157**, 813–830.
- Hess, H., 1964. Seismic anisotropy of the uppermost mantle under the oceans, *Nature*, **203**, 629–631.
- Igel, H., Dijkpessse, H., & Tarantola, A., 1996. Waveform inversion of marine reflection seismograms for P impedance and Poisson's ratio, *Geophys. J. Int.*, **124**, 363–371.
- Ishii, M. & Dziewonski, A. M., 2002. The innermost inner core of the Earth: Evidence for a change in anisotropic behavior at the radius of about 300 km, *Proceedings of the National Academy of Sciences of the United States of America*, **99**(22), 14026–14030.
- Ishii, M. & Tromp, J., 1999. Normal-mode and free-air gravity constraints on lateral variations in velocity and density of Earth's mantle, *Science*, **285**, 1231–1236.
- Ishii, M. & Tromp, J., 2001. Even-degree lateral variations in the Earth's mantle constrained by free oscillations and the free-air gravity anomaly, *Geophys. J. Int.*, **145**, 77–96.
- Ishii, M. & Tromp, J., 2004. Constraining large-scale mantle heterogeneity using mantle and inner-core sensitive normal modes, *Phys. Earth Planet. Int.*, **146**, 113–124.
- Jackson, I., 2000. Laboratory measurements of seismic wave dispersion and attenuation: Recent progress, In: *Earth's Deep Interior: mineral physics and tomography from the atomic to the global scale*, *Geophysical Monograph*, **117**, 265–289.
- Jackson, I., 2007. Physical origins of anelasticity and attenuation in rock, In: *Treatise on Geophysics*, edited by G. D. Price, **2**, 496–522.
- Kaminski, E., Ribe, N. M., & Browaeys, J. T., 2004. Drex, a program for calculation of seismic anisotropy due to crystal lattice preferred orientation in the convective upper mantle, *Geophys. J. Int.*, **158**(2), 744–752.
- Karato, S.-I., 1993. Importance of attenuation in the interpretation of seismic tomography, *Geophys. Res. Lett.*, **20**, 1623–1626.
- Karato, S.-I., 2008. *Deformation of Earth materials: An introduction to the rheology of solid Earth*, Cambridge University Press, Cambridge.
- Karato, S.-I., & Karki, B. B., 2001. Origin of lateral variation of seismic wave velocities and density in the deep mantle, *J. Geophys. Res.*, **106**(B10), 21771–21783.
- Kennett, B. L. N., 1998. On the density distribution within the Earth, *Geophys. J. Int.*, **132**, 374–382.
- Kennett, B. L. N. & Abdullah, A., 2011. Seismic wave attenuation beneath the Australasian region, *Austr. J. Earth Sci.*, **58**, 285–295.
- Kennett, B. L. N., Engdahl, E. R., & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes., *Geophys. J. Int.*, **122**, 108–124.
- Komatitsch, D., Vinnik, L. P., & Chevrot, S., 2010. SHdiff-SVdiff splitting in an isotropic Earth, *J. Geophys. Res.*, **115**, B07312.

- Kuo, C. & Romanowicz, B., 2002. On the resolution of density anomalies in the Earth's mantle using spectral fitting of normal-mode data, *Geophys. J. Int.*, **150**, 162–179.
- Kustowski, B., Ekström, G., & Dziewonski, A. M., 2008. Anisotropic shear-wave velocity structure of the Earth's mantle: a global model, *J. Geophys. Res.*, **113**, B06306, doi:10.1029/2007JB005169.
- Latychev, K., Mitrovica, J. X., Ishii, M., Chan, N.-H., & Davis, J. L., 2009. Body tides on a 3D elastic Earth: Toward a tidal tomography, *Earth Planet. Sci. Lett.*, **227**, 86–90.
- Lekic, V., Matas, J., Panning, M., & Romanowicz, B., 2009. Measurement and implications of frequency dependence of attenuation, *Earth Planet. Sci. Lett.*, **282**, 285–293.
- Lévêque, J. J., Rivera, L., & Wittlinger, G., 1993. On the use of the checkerboard test to assess the resolution of tomographic inversions, *Geophys. J. Int.*, **115**, 313–318.
- Li, C., van der Hilst, R. D., Engdahl, E. R., & Burdick, S., 2008. A new global model for P wave speed variations in Earth's mantle, *Geochem Geophys Geosys*, **9**(5), doi:10.1029/2007GC001806.
- Liu, H.-P., Anderson, D. L., & Kanamori, H., 1976. Velocity dispersion due to anelasticity: implications for seismology and mantle composition, *Geophys. J. Roy. Astr. Soc.*, **47**, 41–58.
- Long, M. D., De Hoop, M. V., & van der Hilst, R. D., 2008. Wave-equation shear wave splitting tomography, *Geophys. J. Int.*, **172**(1), 311–330.
- Long, M. D. & Silver, P. G., 2009. Shear wave splitting and mantle anisotropy: Measurements, interpretations, and new directions, *Surv Geophys*, **30**(4-5), 407–461.
- Masters, G., Jordan, T., Silver, P. G., & Gilbert, F., 1982. Aspherical Earth structure from fundamental spheroidal mode data, *Nature*, **298**, 609–613.
- Masters, G., Laske, G., Bolton, H., & Dziewonski, A., 2000. The relative behavior of shear wave velocity, bulk sound speed, and compressional velocity in the mantle: implications for chemical and thermal structure, in *Earth's deep interior: mineral physics and tomography from the atomic to the global scale*, pp. 63–87, ed. *et al.*, S. K., American Geophysical Union, Washington.
- Maupin, V., 1990. Modeling of three-component Lg waves in anisotropic crustal models, *Bull. Seismol. Soc. Am.*, **80**(5), 1311–1325.
- Maupin, V., Garnero, E. J., Lay, T., & Fouch, M. J., 2005. Azimuthal anisotropy in the D'' layer beneath the Caribbean, *J. Geophys. Res.*, **110**(8), 1–20.
- Maupin, V. & Park, J., 2007. Theory and observations - wave propagation in anisotropic media, in *Treatise on geophysics, Volume 1*, pp. 289–321, ed. Schubert, G., Elsevier, Amsterdam.
- McNamara, A., van Keken, P., & Karato, S.-I., 2002. Development of anisotropic structure by solid state convection in the Earth's lower mantle, *Nature*, **416**, 310–314.
- Mégnin, C. & Romanowicz, B., 2000. The 3D shear velocity structure of the mantle from the inversion of body, surface and higher modes wave forms, *Geophys. J. Int.*, **143**, 709–728.
- Meier, U., Curtis, A., & Trampert, J., 2007. Fully non-linear inversion of fundamental mode surface waves for a global crustal model, *Geophys. Res. Lett.*, **34**, doi:10.1029/2007GL030989.
- Meissner, R., Rabbel, W., & Kern, H., 2006. Seismic lamination and anisotropy of the lower continental crust, *Tectonophysics*, **416**(1-2), 81–99.
- Mitrovica, J. X. & Forte, A. M., 2004. A new inference of mantle viscosity based upon joint inversion of convection and glacial isostatic adjustment data, *Earth Planet. Sci. Lett.*, **225**, 177–189.
- Moczo, P., Kristek, J., Galis, M., Pazak, P., & Balazovjech, M., 2007. The finite-difference and finite-element modelling of seismic wave propagation and earthquake motion, *Acta Physica Slovaca*, **57**, 177–406.
- Montagner, J.-P. & Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies, *J. Geophys. Res.*, **96**(B12), 20,337–20,351.
- Morelli, A., Dziewonski, A.M., & Woodhouse, J.H., 1986. Anisotropy of the inner core inferred from PKIKP travel times, *Geophys. Res. Lett.*, **13**, 1545–1548.
- Mosca, I., Cobden, L., Deuss, A., Ritsema, J., & Trampert, J., 2012. Seismic and mineralogical structures of the lower mantle from probabilistic tomography, *J. Geophys. Res.*, **117**, B06304, doi:10.1029/2011JB008851.
- Nolet, G., 2008. *A breviary of seismic tomography: Imaging the interior of the Earth and the Sun*, Cambridge University Press, Cambridge.
- Okal, E. A. & Jo, B.-G., 1990. Q measurements for phase X overtones, *Pure Appl. Geophys.*, **132**, 331–362.

- Panning, M. P. & Nolet, G., 2008. Surface wave tomography for azimuthal anisotropy in a strongly reduced parameter space, *Geophys. J. Int.*, **174**(2), 629–648.
- Panning, M. P. & Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.*, **167**, 361–379.
- Parker, R. L., 1994. *Geophysical inverse theory*, Princeton University Press.
- Paulssen, H., 2004. Crustal anisotropy in southern California from local earthquake data, *Geophys. Res. Lett.*, **31**(1), L01601 1–4.
- Peter, D., Komatitsch, D., Luo, Y., Martin, R., Le Goff, N., Casarotti, E., Le Loher, P., Magnoni, F., Liu, Q., Plitz, C., Nissen-Meyer, T., Basini, P., & Tromp, J., 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.*, **186**, 721–739.
- Rawlinson, N., Pogat, S., & Fishwick, S., 2010. Seismic tomography: A window to the deep Earth, *Phys. Earth Planet. Int.*, **178**, 101–135.
- Reid, F. J. L., Woodhouse, J. H., & van Heijst, H. J., 2001. Upper mantle attenuation and velocity structure from measurements of differential S phases, *Geophys. J. Int.*, **145**, 615–630.
- Resovsky, J. S. & Ritzwoller, M. H., 1999. Regularisation uncertainty in density models estimated from normal mode data, *Geophys. Res. Lett.*, **26**, 2319–2322.
- Resovsky, J. & Trampert, J., 2003. Using probabilistic seismic tomography to test mantle velocity-density relationships, *Earth Planet. Sci. Lett.*, **215**, 121–134.
- Resovsky, J. S., Trampert, J., & van der Hilst, R. D., 2005. Error bars for the global seismic Q profile, *Earth Planet. Sci. Lett.*, **230**, 413–423.
- Ritsema, J., Deuss, A., van Heijst, H. J., & Woodhouse, J. H., 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.
- Ritsema, J., McNamara, A. K., & Bull, A. L., 2007. Tomographic filtering of geodynamic models: Implications for models interpretation and large-scale mantle structure, *J. Geophys. Res.*, **112**(1), doi:10.1029/2006JB004566.
- Ritzwoller, M. H. & Lavelle, E. M., 1995. Three-dimensional seismic models of the Earth's mantle, *Rev. Geophys.*, **33**, 1–66.
- Romanowicz, B., 1990. The upper mantle degree 2: Constraints and inferences from global mantle wave attenuation measurements, *J. Geophys. Res.*, **95**, 11051–11071.
- Romanowicz, B., 1994. On the measurement of anelastic attenuation using amplitudes of low-frequency surface waves, *Phys. Earth Planet. Int.*, **84**, 179–191.
- Romanowicz, B., 1995. A global tomographic model of shear attenuation in the upper mantle, *J. Geophys. Res.*, **100**, 12375–12394.
- Romanowicz, B., 2001. Can we resolve 3D density heterogeneity in the lower mantle?, *Geophys. Res. Lett.*, **28**, 1107–1110.
- Romanowicz, B., 2003. Global mantle tomography: progress status in the past 10 years, *Ann. Rev. Earth Planet. Sci.*, **31**, 303–328.
- Sailor, R. V. & Dziewonski, A. M., 1978. Measurements and interpretation of normal mode attenuation, *Geophys. J. Roy. Astr. Soc.*, **53**, 559–581.
- Sambridge, M., 1999a. Geophysical inversion with a neighbourhood algorithm - I. Searching a parameter space, *Geophys. J. Int.*, **138**, 479–494.
- Sambridge, M., 1999b. Geophysical inversion with a neighbourhood algorithm - I. Appraising the ensemble, *Geophys. J. Int.*, **138**, 727–746.
- Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.*, **40**, doi: 10.1029/2000RG000089.
- Schulte-Pelkum, V., Masters, G., & Shearer, P. M., 2001. Upper mantle anisotropy from long-period P polarisation, *J. Geophys. Res.*, **106**(B10), 21917–21934.
- Selby, N. D. & Woodhouse, J. H., 2000. Controls on Rayleigh wave amplitudes: Attenuation and focusing, *Geophys. J. Int.*, **142**, 933–940.
- Selby, N. D. & Woodhouse, J. H., 2002. The Q structure of the upper mantle: Constraints from Rayleigh wave amplitudes, *J. Geophys. Res.*, **107**, doi:10.1029/2001JB000257.
- Sheehan, A. F. & Solomon, S. C., 1992. Differential shear-wave attenuation and its lateral variation in the North Atlantic region, *J. Geophys. Res.*, **97**, 15339–15350.
- Sieminski, A., Trampert, J., & Tromp, J., 2009. Principal component analysis of anisotropic finite-frequency sensitivity kernels, *Geophys. J. Int.*, **179**(2), 1186–1198.
- Sigloch, K., 2008. *Multiple-frequency body wave tomography*, Phd thesis, Princeton University.
- Simmons, N. A., Forte, A. M., Boschi, L., & Grand, S. P., 2010. GyPSuM: A joint tomographic model of mantle density and seismic wave speeds, *J. Geophys. Res.*, **115**, doi:10.1029/2010JB007631.

- Sipkin, S. A. & Jordan, T. H., 1979. Frequency-dependence of QScS, *Bull. Seis. Soc. Am.*, **69**, 1055–1079.
- Sipkin, S. A. & Revenaugh, J., 1994. Regional variation of attenuation and traveltime in China from analysis of multiple ScS phases, *J. Geophys. Res.*, **99**, 2687–2699.
- Smith, G. P. & Ekström, G., 1999. A global study of Pn anisotropy beneath continents, *J. Geophys. Res.*, **104**(B1), 963–980.
- Soldati, G., Boschi, L., & Piersanti, A., 2006. Global seismic tomography and modern parallel computers, *Annals of Geophysics*, **49**(4-5), 977–986.
- Stacey, F. D. & Davis, P. M., 2004. High pressure equations of state with applications to the lower mantle and core, *Phys. Earth Planet. Inter.*, **142**(3-4), 137–184.
- Stachnik, J. C., Abers, G. A., & Christensen, D. H., 2004. Seismic attenuation and mantle wedge temperatures in the Alaska subduction zone, *J. Geophys. Res.*, **109**, doi:10.1029/2004JB003018.
- Stixrude, L. & Lithgow-Bertelloni, C., 2005. Thermodynamics of mantle minerals—I. Physical properties, *GJI*, **162**(2), 610–632.
- Stixrude, L. & Lithgow-Bertelloni, C., 2011. Thermodynamics of mantle minerals—II. Phase equilibria, *Geophys. J. Int.*, **184**(3), 1180–1213.
- Tanaka, H., Nagamine, K., Kawamura, N., Nakamura, S. N., Ishida, K., & Shimomura, K., 2003. Development of a two-fold segmented detection system for near horizontally cosmic-ray muons to probe the internal structure of a volcano, *Nucl. Instr. Meth. A*, **507**, 657–669.
- Tanaka, H., Taira, H., Uchida, T., Tanaka, M., Takeo, M., Ohminato, T., Aoki, Y., Nishimata, R., Shoji, D., & Tsuiji, H., 2010. Three-dimensional computational axial tomography scan of a volcano with cosmic ray muon radiography, *J. Geophys. Res.*, **115**, B12332, doi:10.1019/2010JB007677.
- Tape, C., Liu, Q., Maggi, A., & Tromp, J., 2010. Seismic tomography of the southern California crust based upon spectral-element and adjoint methods., *Geophys. J. Int.*, **180**, 433–462.
- Tarantola, A., 1986. A strategy for nonlinear elastic inversion of seismic reflection data, *Geophysics*, **51**, 1893–1903.
- Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation, *Pure Appl. Geophys.*, **128**, 365–399.
- Tarantola, A., 2005. *Inverse problem theory and methods for model parameter estimation*, Society for Industrial and Applied Mathematics, Philadelphia.
- Tian, Y., Sigloch, K., & Nolet, G., 2009. Multiple-frequency SH-wave tomography for the western US upper mantle, *Geophys. J. Int.*, **178**, 1384–1402.
- Tondi, R., Achauer, U., Landes, M., davi, R., & Besutiu, L., 2009. Unveiling seismic and density structure beneath the Vrancea seismogenic zone, Romania, *J. Geophys. Res.*, **114**, doi:10.1029/2008JB005992.
- Tondi, R., de Franco, R., & Barzaghi, R., 2000. Sequential inversion of refraction and wide-angle reflection traveltimes and gravity data for two-dimensional velocity structures, *Geophys. J. Int.*, **141**, 679–698.
- Trampert, J., Deschamps, F., Resovsky, J., & Yuen, D., 2004. Probabilistic tomography maps chemical heterogeneities throughout the lower mantle., *Science*, **306**, 853–856.
- Trampert, J., Vacher, P., & Vlaar, N., 2001. Sensitivities of seismic velocities to temperature, pressure and composition in the lower mantle, *Phys. Earth Planet. Inter.*, **124**(3-4), 255–267.
- Trampert, J. & van der Hilst, R., 2005. Towards a quantitative interpretation of global seismic tomography, in *Earth's deep mantle: Structure, composition and evolution*, pp. 47–62, eds van der Hilst, R., Bass, J., Matas, J., & Trampert, J., American Geophysical Union, Washington.
- Trampert, J. & van Heijst, H. J., 2002. Global azimuthal anisotropy in the transition zone, *Science*, **296**, 1297–1299.
- Trampert, J. & Woodhouse, J. H., 2003. Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s, *Geophys. J. Int.*, **154**(1), 154–165.
- van den Berg, J., Curtis, A., & Trampert, J., 2003. Optimal nonlinear bayesian experimental design: An application to amplitude versus offset experiments, *Geophys. J. Int.*, **155**(2), 411–421.
- van der Hilst, R. D., Widiyantoro, S., & Engdahl, E. R., 1997. Evidence for deep mantle circulation from global tomography, *Nature*, **386**, 578–584.
- Visser, K., Trampert, J., Lebedev, S., & Kennett, B. L. N., 2008. Probability of radial anisotropy in the deep mantle, *Earth Planet. Sci. Lett.*, **270**(3-4) 241–250.
- Warren, L. M. & Shearer, P. M., 2002. Mapping lateral variations in upper mantle attenuation by stacking P and PP spectra, *J. Geophys. Res.*, **107**, doi:10.1029/2001JB001195.

- Wenk, H., Cottaar, S., Tomé, C. N., McNamara, A., & Romanowicz, B., 2011. Deformation in the lowermost mantle: From polycrystal plasticity to seismic anisotropy, *Earth Planet. Sci. Lett.*, **306**(1-2), 33–45.
- Widmer, R., Masters, G., & Gilbert, F., 1991. Spherically symmetric attenuation within the Earth from normal mode data, *Geophys. J. Int.*, **104**, 541–553.
- Woodhouse, J. H. & Dziewonski, A. M., 1984. Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. Geophys. Res.*, **89**, 5953–5986.
- Woodhouse, J. H. & Dziewonski, A. M., 1989. Seismic modelling of the Earth's large-scale three dimensional structure, *Philos. Trans. R. Soc. Lond. A*, **328**, 291–308.
- Woodhouse, J. H., Giardini, D., & Li, X. D., 1986. Evidence for inner core anisotropy from free oscillations, *Geophys. Res. Lett.*, **13**, 1549–1552.
- Woodhouse, J. H. & Wong, Y. K., 1986. Amplitude, phase and path anomalies of mantle waves, *Geophys. J. Roy. Astr. Soc.*, **87**, 753–773.
- Wu, R. & Aki, K., 1985. Scattering characteristics of elastic waves by an elastic heterogeneity, *Geophysics*, **50**, 582–595.
- Zener, C., 1948. *Elasticity and anelasticity of metals*, University of Chicago Press, Chicago.
- Zhou, Y., 2009. Surface-wave sensitivity to 3-D anelasticity, *Geophys. J. Int.*, **178**, 1403–1410.