submitted to *Geophys. J. Int.* (2001)

Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle

N.M. Shapiro and M.H. Ritzwoller

Center for Imaging the Earth's Interior

Department of Physics

University of Colorado

Campus Box 390, Boulder, CO 80309, USA

SUMMARY

We describe a method to invert surface wave dispersion data for a model of shear velocities with uncertainties in the crust and uppermost mantle. The inversion is a multi-step process, constrained by a-priori information, that culminates in a Markov-chain Monte-Carlo sampling of model space to yield an ensemble of acceptable models at each spatial node. The model is radially anisotropic in the uppermost mantle to an average depth of about 200 km and is isotropic elsewhere. The method is applied on a $2^\circ \times 2^\circ$ grid globally to a large data set of fundamental mode surface wave group and phase velocities (Rayleigh group velocity, 16 - 200 s; Love group velocity, 16 - 150 s; Rayleigh and Love phase velocity, 40 - 150 s). The middle of the ensemble (Median Model) defines the estimated model and the half-width of the corridor of models provides the uncertainty estimate. Uncertainty estimates allow the identification of the robust features of the model which, typically, persist only to depths of ∼250 km. We refer to the features that appear in every member of the ensemble of acceptable models as "persistent". Persistent features include sharper images of the variation of oceanic lithosphere and asthenosphere with age, continental roots, extensional tectonic features in the upper mantle, the shallow parts of subducted lithosphere, and improved resolution of radial anisotropy. In particular, we find no compelling evidence for "negative anisotropy" ($v_{sv} > v_{sh}$) anywhere in the world's lithosphere.

1. Introduction

Even as earth models proliferate, reports of model uncertainties are rare in contemporary seismic tomography at all scales. There are several reasons for this. First, the construction of seismic models is typically a nonlinear inverse problem solved using regularized, weighted linear or iterative regression analysis. Classical error analyses with normally distributed data errors tend to underestimate model variances. Covariance estimates are similarly suspicious and are commonly ignored. Problems arise when the forward model and the data are inconsistent, in which case errors are usually not random, let alone normally distributed. In addition, classical error estimates may not account for an array of choices made by the "tomographer" during inversion (e.g., arbitrariness of damping, parameterization, etc.) which can strongly affect both the qualitative and quantitative character of the resulting model. These choices are, in effect, a-priori constraints on the model imposed by the tomographer, but their affects on the resulting model are frequently only dimly perceived. Second, the error of a seismic model is scale dependent. The large scale or average characteristics of the model may be well known, but uncertainties grow as the scale of the model decreases. A full report of uncertainties, therefore, would include error estimates over a spectrum of length-scales at each spatial location in the model. This compounds the difficulty of the problem, and seismic tomographers commonly respond by reporting their model devoid of meaningful quantitative estimates of the model's reliability.

This state of affairs is far from ideal. Model uncertainties are needed to identify features that are worthy of interpretation and to guide the use of predictions made from the seismic model (e.g., gravity, heat flow, temperature, travel times, etc.). The predictive capabilities of seismic models are currently underexploited, in our opinion, not because the predictions are inaccurate but because the accuracy of the predictions is unknown. All too often, therefore, seismic models only provide images that evoke the Earth's interior rather than fundamental information that is used to understand the physics of the Earth's interior. Perhaps most importantly, without uncertainties, discord between models is difficult to resolve and scientific progress can be seriously impeded.

Model space sampling methods, such as Monte-Carlo methods (e.g., Mosegaard and Tarantola 1995), help to address some of these issues. The basic idea of a Monte-Carlo "inversion" is to select models randomly and to retain only the subset of models that satisfies acceptability criteria. These criteria usually include some combination of data fit and a-priori information that defines a physically plausible model (Figure 1). The outcome of this inversion is an ensemble of acceptable models whose variability provides some information about model uncertainty.

There are a number of advantages to Monte-Carlo inversions. First, Monte-Carlo sampling is simple to effect and change, involving only model selection and forward modeling. A-priori constraints typically define the region of model space to be searched and are imposed explicitly during model selection. Second, the model can be over-parameterized which may be desired to assess how trade-offs between different types of structures affect the range of acceptable models. Third, the resulting uncertainties may be what the modeler really wants. They summarize the range of models that will fit the data and encorporate the modeler's prejudices.

The major disadvantage of the method is computational expense, and Monte-Carlo methods work best when the volume of model space searched is small. A second disadvantage is that the uncertainty estimates depend strongly on the choice of the acceptance criterion which may be ad hoc. In particular, the a-priori constraints, which largely reflect the modeler's prejudices, strongly influence the uncertainties. This may be good from the modeler's perspective but undesirable for the user of the model. Finally, systematic errors in the data and the a-priori information, as well as inconsistencies between the data and the a-priori information, may bias the mean of the ensemble of acceptable models in a way that will not be reflected in the uncertainties. It is important then to attempt to identify and, to the maximum extent possible, eliminate systematic errors and inconsistencies prior to the Monte-Carlo inversion, as in all inversions. These caveats aside, we find that in most cases the ensemble of acceptable models accurately reflects our degree-of-belief in the model and usefully guide the interpretation of the model, predictions made from the model, as well as comparisons with other models.

In this paper we discuss the application of a Monte-Carlo method to estimate the shear velocity structure of the crust and upper mantle world-wide. The study has three key innovations: (1) simultaneous inversion of a large data set of fundamental mode group and phase velocity (Trampert and Woodhouse 1995; Ekström *et al.* 1997) dispersion information, (2) the way in which we introduce a-priori information in the inversion, and (3) the global scale of the Monte-Carlo inversion. Although Monte-Carlo and related inversions have a long history and are now common in surface wave seismology (e.g., Levshin *et al.* 1966; Keilis-Borok and Yanovskaya 1967; Press 1968; Lomax and Snieder 1994; Shapiro *et al.* 1997), we are unaware of previous studies that have applied this method on a global scale. The result is a shear velocity model on a $2^\circ \times 2^\circ$ grid around the globe with uncertainties everywhere. We present below a brief discussion of the data set and the method and results of surface wave tomography. Data processing is more fully discussed by Ritzwoller & Levshin (1998), the tomographic method is presented by Barmin et al. (2001), and an application of tomography is discussed by Ritzwoller et al. (2001). The main purpose of this paper is to describe the inversion procedure, although for completeness we will summarize certain aspects of the model both for isotropic and radially anisotropic structures.

2. Data

The data are surface wave group and phase velocities. Although phase ($C = \omega/k$) and group ($U = d\omega/dk$) velocities are simply related by

$$U(\omega) = \frac{C(\omega)}{1 - \frac{\omega}{C(\omega)}\frac{dC}{d\omega}} \approx C(\omega) + \omega\frac{dC(\omega)}{d\omega}, \tag{1}$$

the simultaneous inversion of U and C is substantially better than the use of either alone, as shown below. There are two reasons for this. First, group velocity measurements typically extend to much shorter periods than the phase velocities and, therefore, provide unique constraints on shallow structures that help to resolve the trade-off between crustal and mantle structures in the inversion. Second, phase and group velocities are measured differently. Group velocities are measured on the amplitude of the surface wave-packet and phase velocities on the phase, so the error processes in the measurements are largely independent.

We measured the group velocities using frequency-time analysis (Levshin et al. 1989) in which an analyst defines the frequency band of measurement for every waveform and guides the procedure to separate the signal from a variety of noise sources (e.g., overtones, fundamental modes of different type, other earthquakes, multipaths, scattered arrivals). We used broadband waveforms following earthquakes that occurred from 1977 - 1999 at stations from global networks (GDSN, GSN, GEOSCOPE) as well as temporary regional arrays (e.g., KNET, Saudi Arabian Network, SKIPPY). The phase velocities were measured at Harvard University and Utrecht University, separately, and donated to this study. These phase velocity data sets are described by Ekström et al. (1997) and Trampert & Woodhouse (1995). We merged all phase velocity measurements into a single data set following the procedure described by Ritzwoller et al. (2001). A cluster analysis (e.g., Ritzwoller & Levshin 1998) is applied to both the group and phase velocity measurements to reject out-liers and estimate the rms-variation in the measurements. The results of this analysis for the southern hemisphere, presented by Ritzwoller et al. (2001), show that typical measurement errors are 20 - 25 m/s for group velocities, except for very short periods, and 10 - 15 m/s for phase velocities, except for long period Love waves. This is a conservative estimate of measurement errors for the whole Earth because, on average, surface wave excitation is weaker and paths typically longer in the southern hemisphere.

The phase velocity data are decimated onto the same discrete grid of periods for both Rayleigh and Love waves (40, 50, 60, 70, 80, 90, 100, 125, 150 s) , and the group velocity data are discretized onto a grid from 20 s to 150 s for Love waves (20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 125, 150 s) and four additional periods for Rayleigh waves 16, 18, 175, 200 s). There are, therefore, 50 measurements for each sourc–receiver path.

Data coverage is highly heterogeneous, being imposed by the distribution of receiving stations and earthquakes. It is generally better for Rayleigh waves than for Love waves, is better at intermediate than at very short or very long periods, and is better in the northern than in the southern hemisphere. Data coverage is most dense in Eurasia and is currently sparsest across Africa, the central Pacific, parts of the Indian Ocean, and Antarctica. At present, the data set consists of more than 100,000 group velocity paths and 50,000 phase velocity paths. Figure 2 presents examples of Rayleigh wave path density for several periods.

3. Method of Inversion

The relation between surface-wave dispersion and the seismic velocity structure of the earth is nonlinear. There are two common approaches to resolve this nonlinearity. One is waveform fitting in which the relation between the model and the seismic waveforms is linearized and the model is iteratively estimated (e.g., Snieder 1988; Nolet 1990; Marquering et al. 1996). Our approach, in contrast, is based on direct measurements of surface wave dispersion rather than fitting waveforms so that the nonlinear inverse problem is divided into two steps: Step 1 is a nearly linear part called surface wave tomography to estimate 2-D dispersion maps and Step 2 is a nonlinear inversion of the dispersion curves at each geographical point for a shear velocity model of the crust and upper mantle. (The measurement procedure itself may also be nonlinear.) In regions of poor data coverage, the tomographic maps and the shear velocity model will revert to a common reference (the “Initial Model” defined below) while the uncertainties will increase to limits imposed by a-priori constraints.

3.1 Forward problem

The forward problem, i.e., the prediction of the frequency dependent surface-wave travel times from the three-dimensional shear-velocity model, is similarly divided into two steps, mirroring the steps in the inverse problem. The first step is the prediction of the Rayleigh and Love wave dispersion curves from a three dimensional (3-D) model at each geographical point. The second step is the prediction of surface-wave travel times for different source-receiver pairs. We follow tradition here by using geometrical ray theory in the latter step. The limitations and effects of this approximation are discussed further by Ritzwoller et al. (2002a).

At each geographical point (θ, ϕ), the Rayleigh and Love wave dispersion curves extracted from the 2-D tomographic maps compose the data vector:

$$\mathbf{d} = \left[U^R(\omega), C^R(\omega), U^L(\omega), C^L(\omega)\right]^T \tag{2}$$

where ω is frequency, C is the phase velocity, U is the group velocity, and T denotes transpose. The indices R and L refer to Rayleigh and Love waves, respectively.

The dispersion curves are assumed to result from the 3-D earth model at (θ, ϕ),

$$\mathbf{m} = [c_{ijkl}(z), \rho(z), Q(z)]^T \tag{3}$$

where z is depth, $c_{ijkl}(z)$ is the elastic tensor, $\rho(z)$ is density, and $Q(z)$ is the shear quality factor. The forward problem can then be written schematically as:

$$\mathbf{d} = \mathcal{F}(\mathbf{m}) \tag{4}$$

which can be solved with a number of algorithms. We use the method and computer code of Woodhouse (1988) which operates on a radially anisotropic earth model. A radially anisotropic (or transversely isotropic) medium consists of five mutually independent elastic moduli (Smith and Dahlen 1973), $A = \rho v_{ph}^2, C = \rho v_{pv}^2, F/(A-2L) = \eta, L = \rho v_{sv}^2$, and $N = \rho v_{sh}^2$, so that

$$\mathbf{m} = [v_{sh}(z), v_{sv}(z), v_{ph}(z), v_{pv}(z), \eta(z), \rho(z), Q(z)]^T \,. \tag{5}$$

For an isotropic solid, $A = C = \kappa + 4\mu/3, N = L = \mu, F = \kappa - 2\mu/3$, and $\eta = 1$ where κ and μ are bulk modulus and rigidity, respectively.

3.2 Step 1. Surface wave tomography

Surface wave tomography is the inference of maps of surface wave velocities for each wave type (Rayleigh, Love) and period from surface wave travel times. We use the method described in detail by Barmin et al. (2001) to construct maps on a $2^\circ \times 2^\circ$ grid world-wide at the discrete periods identified in section 2. We estimate a total of 50 global tomographic maps. Examples of tomographic maps using the same method have been published in previous papers for specific regions (e.g., Barmin et al. 2001; Levshin et al. 2001; Ritzwoller et al. 2001; Villaseñor et al. 2001). These maps typically fit the measured group velocities with rms misfits of 40 - 50 m/s and the measured phase velocities to 20 - 40 m/s, or about twice the measurement error. The few examples shown here in Figure 3 are similar to those presented by Ritzwoller et al. (2001), although the data set has evolved.

We assume that surface waves propagate along great-circle paths, which linearizes surface wave tomography. The lateral resolution of the shear velocity model will be determined by the damping and regularization of the tomographic maps, which is effected through a Gaussian smoothing condition applied to the model $\mathbf{m}(\mathbf{r})$ in the inner product matrix

$$\mathbf{m}(\mathbf{r}) \sim \int\limits_S S(\mathbf{r},\mathbf{r}')\mathbf{m}(\mathbf{r}')d\mathbf{r}' \ , \tag{6}$$

where

$$S(\mathbf{r},\mathbf{r}') = K \exp\left(-\frac{|\mathbf{r}-\mathbf{r}'|^2}{2\sigma^2}\right) \tag{7}$$

$$\int\limits_S S(\mathbf{r},\mathbf{r}')d\mathbf{r}' = 1, \tag{8}$$

where $\mathbf{m}$ is a dispersion map, $\mathbf{r}$ is the position vector in 2-D, and K is a normalization constant. This regularization procedure is similar to the use of "fat rays" in 2-D with a Gaussian cross-section of standard deviation $\sqrt{2}\sigma$, but is done more to reduce artifacts that would otherwise contaminate the maps than to model the spatial sensitivity of the data. Ritzwoller et al. (2002a) show that better sensitivity kernels are more likely to affect estimates of resolution than the tomographic maps themselves. Ritzwoller et al. (2002a) provide more meaningful globally averaged resolution estimates based on "diffraction tomography": for Rayleigh wave group velocities resolution increases with period from about 350 km for the shortest periods to 700 km at long periods, and from about 500 km to more than 1000 km for Love wave group velocities. We argue, therefore, that the model is resolved to regional or tectonic length scales even though the model is global in extent. This fact, together with the improvements in vertical resolution that derive from the simultaneous inversion of group and phase velocities, have driven the development of the model presented here.

3.3 Step 2. Inversion of dispersion curves

The primary purpose of this paper is to describe the second step of the inversion for a shear velocity model of the crust and upper mantle. Because $\mathcal{F}$ in equation (4) is a non-linear function, $\mathcal{F}^{-1}$ is not well defined. It is common to consider the Taylor series expansion of the forward solution around a reference model $\hat{\mathbf{m}}$:

$$\mathbf{d} = \mathcal{F}(\hat{\mathbf{m}}) + \sum_i \left(\partial \mathcal{F}/\partial m_i\right) \delta m_i + \frac{1}{2} \sum_{i,j} \left(\partial^2 \mathcal{F}/\partial m_i \partial m_j\right) \delta m_i \delta m_j + \mathcal{O}(\delta m^3), \tag{9}$$

where each component of the model vector is $m_i = \hat{m}_i + \delta m_i$. If one drops the nonlinear terms, the first partial derivatives form a matrix which can be inverted with regularization constraints to estimate the perturbations δm_i. Because surface wave dispersion is dominantly affected only by v_{sv} and v_{sh}, the quantities Q, ρ, v_{pv}, and v_{ph} are commonly fixed in surface wave inversions or are set to scale in some way with the estimates of v_{sv} and v_{sh}. Villaseñor et al. (2001) presents an example of this approach.

We generalize this approach here into three separate stages (Figure 1). The input and output of each stage are summarized in Table 1. The first stage is the linearized inversion, and has been applied in Central Asia by Villaseñor et al. (2001). The inversion begins with the Initial Model, denoted $\mathbf{m}^0$. In this stage we simply parameterize the estimated model and call it the Simple Reference Model, $\mathbf{m}^1$. The primary aim is to identify the region of model space for detailed Monte-Carlo sampling and to speed the forward solution. In Stage 2, we generalize the parameterization of the model and perform simulated annealing to construct the Best Fit Model ($\mathbf{m}^2$) on which the Monte-Carlo acceptance criterion is based. In Stage 3, we randomly sample model space in the volume surrounding $\mathbf{m}^2$ using a Markov-chain (random-walk) algorithm (Gilks *et al.* 1996) to construct the ensemble of models that are judged to be acceptable, $\mathbf{m}^3$. We summarize the ensemble with the “Median Model” $\mathbf{m}^3$. The half-width of the corridor of models defined by the ensemble specifies the model uncertainties.

Lateral smoothing constraints are applied in Step 1, the surface wave tomography. Lateral smoothing is not additionally applied in Step 2, but vertical smoothness constraints are applied. The final model retains much of the lateral smoothness of the tomographic maps that are input into Step 2 of the inversion, but some of the smoothness is lost.

3.3.1 Initial Model $\mathbf{m}^0$

The Initial Model is based on a variety of sources of global information, including the sediment model of Laske & Masters (1997), the crustal model CRUST5.1 of Mooney et al. (1998), and the shear-wave velocity model of the upper mantle S20A of Ekström & Dziewonski (1998). In Eurasia, we introduced regional information, including maps of sediment and crustal thicknesses constructed by the Russian Institute of Physics of the Earth which was converted to digital form by the Cornell Digital Earth project (Seber et al. 1997) and a recent model of crustal thickness over part of Eurasia based on seismic profiles, which was compiled by G. Laske (personal communication). The result is a model of the crust and upper mantle that includes a water layer where appropriate, topography on the solid surface and Moho, and 3-D variations in v_s and v_p in the sediments and crystalline crust. Shear velocities in the mantle are from the isotropic part of the model S20A modified with radial anisotropy from PREM (Dziewonski & Anderson 1981). The average of the shear velocity model has been replaced with the 1-D model ak135 (Kennett et al. 1995) in order to remove the discontinuity at 220 km in PREM. Density and isotropic compressional velocity in the mantle scale with variations in v_s using $d\ln v_p/d\ln v_s = 0.5, d\ln \rho/d\ln v_s = 0.25$. Radial anistropy is introduced into the P-wave velocities by analogy with PREM and η is set to the PREM value without reintroducing the 220 km discontinuity. The Q model is also from PREM.

3.3.2 Parameterization and a-priori constraints

Model parameterization strongly affects the Median Model and the uncertainties, which both depend on the size of the model subspaces considered in the inversion. If, for example, the inversion is too weakly constrained, there will be a broad subset of models that will fit the data and large uncertainties will result at each depth. Much tighter constraints on the model space reduce the uncertainty in the estimated parameters, but the model will be increasingly subject to systematic errors. Our approach is to attempt to over-parameterize but then to apply physically motivated constraints on each model parameter. The effect is to mitigate against considering grossly aphysical models, but to allow physically realistic components from the null-space into the space of models considered.

We use a uniform parameterization over the whole globe. In Stage 1, we follow Villaseñor et al. (2001) and use eight parameters in the crust and upper mantle of which we can estimate 4-5 linear combinations. We generalize the model in Stages 2 and 3 to 14 parameters, seven coefficients in the crust and seven in the mantle as shown in Figure 4. Isotropic P- and S-velocities in three crustal layers and crustal thickness are changed during the inversion. Isotropic mantle S-wave velocity structure is parameterized with four cubic B-splines. The remaining three coefficients parameterize the radially anisotropic part of the upper mantle with two different shear velocities, v_{sh} and v_{sv}. We use a simple parameterization for radial anisotropy similar to PREM in which two of the three coefficients are the values of v_{sh} and v_{sv} in the uppermost mantle directly beneath the crust and the third coefficient is the thickness of the anisotropic layer. The strength of anisotropy reduces monotonically with depth from the Moho to the base of the anisotropic layer. Because of the relatively large number of crustal parameters, this parameterization is tuned more for continental than oceanic areas.

Because Rayleigh waves are predominantly sensitive to v_{sv} and Love waves to v_{sh}, we have constraints on only two of the five elastic moduli that compose a radially anisotropic model. The model needs to be completed in order to solve the forward problem, however. For want of a better solution, we set η to the PREM value at each depth and compute v_{pv} and v_{ph} using a logarithmic scaling relation from v_{sv} and v_{sh}; $d \ln v_{ph} / d \ln v_{sh} = d \ln v_{pv} / d \ln v_{sv} = 0.5$. Because the surface wave velocities are only weakly dependent on compressional velocities and η in the mantle, the arbitrariness of this procedure has little affect on the results of the inversion for v_{sh} and v_{sv}. As in the Initial Model, density scales with v_s and Q remains fixed at the PREM value.

We impose two types of constraints. First, we limit the range of perturbations for some of the parameters; i.e., $\delta m_i^{min} < \delta m_i < \delta m_i^{max}$. This type of constraint is illustrated in Figure 1 which shows that we search only a subspace of the model space around a reference model. For example, we constrain the depth of the Moho to be varied within ±5 km of the Initial Model. This constraint on the Moho depth reduces the trade-off between the crustal and upper-mantle velocities. Second, we impose monotonicity constraints on the velocities in the crust; i.e., crustal velocities must satisfy $m_i < m_{i+1}$ where m_i is the velocity of a layer directly over a layer with velocity m_{i+1}. The explicit constraints are listed in Table 2. We constrain the crustal velocities and the depths of Moho and the bottom of the anisotropic mantle. There is no explicit constraint on perturbations to the mantle velocities, but these perturbations are implicitly constrained by the selected parameterization; i.e., the use of cubic B-splines imposes a degree of vertical smoothness.

We have applied these a-priori constraints uniformly over the whole globe. The inversion method, however, allows regional tuning. In particular, because the oceanic crust is younger and more homogeneous than the continental crust, it may be useful to apply stronger constraints on the crustal parameters in the oceans. In contrast, for poorly known continental regions, like Africa or Antarctica, larger allowed variations in crustal thickness and velocities may be beneficial.

3.3.3 Stage 1. Preliminary linearized inversion for model $\mathbf{m}^1$

This preliminary inversion is used to improve the Initial Model in order to accelerate the forward problem, as discussed further below. The linearized inversion also helps to define the region of model space for Monte-Carlo sampling. At this stage, we use the iterative linearized inversion described by Villaseñor et al. (2001). On each iteration, partial derivatives are calculated as the difference between the dispersion curves computed for the reference and perturbed models. The inversion typically converges in 5-8 iterations.

3.3.4 Stages 2 and 3. Simulated annealing ($\mathbf{m}^2$*) and Monte-Carlo (*$\mathbf{m}^3$*) inversion*

Model space sampling methods, such as simulated annealing and Monte-Carlo, require a fast solution to the forward problem to sample model space adequately. The forward problem is accelerated if we replace the exact solution by its truncated Taylor series expansion. James & Ritzwoller (1999) suggest retaining at least selected second and third-order terms in equation (9). They assumed, however, that the reference model $\hat{\mathbf{m}}$ would be a poor approximation to local structure. As Figure 5 shows, the use of the Simple Reference Model for $\hat{\mathbf{m}}$ produces sufficient accuracy with only second order terms in the Taylor series expansion. The Simple Reference Model is used to calculate all partial derivatives up to second order for all 14 model parameters used in equation (9).

In Stage 2, we apply a simulated annealing (SA) inversion which uses a random sampling of the model space based on an analogy with the annealing of solids (e.g., Metropolis et al. 1953; Kirkpatrick et al. 1983), in which allowed perturbations are subjected to the a-priori constraints (Table 2). In SA, each model realization is generated as a random perturbation to a previous model. The probability distribution of the perturbation is modified during the inversion by slowly reducing its amplitude, analogous to the annealing of solids where the temperature of a solid is slowly reduced to reach a state with minimum internal energy. In SA optimization, the 'statistical temperature' is slowly reduced to find a model realization with minimal cost function. Therefore, the cost function E is analogous to the physical energy and the 'statistical temperature' T is a parameter that controls the amplitude of the random model perturbation. More exactly, the SA method consists of three functional relations: (1) $g(m)$, the probability density of the model space which is used to generate a new model realization, (2) $h(E)$, the probability of the acceptance of the generated model based on the new and old values of the cost-function, and (3) $T(k)$, the schedule of 'annealing' the 'statistical temperature' on step k. We use Boltzmann annealing (e.g., Ingber 1989) in which:

$$g(m_k) = (2\pi T)^{-\frac{D}{2}} \exp(-|m_k - m_{k-1}|^2/(2T)) \tag{10}$$

$$h(E) = 1/(1 + \exp((E_k - E_{k-1})/T) \tag{11}$$

$$T(k) = T_0/\ln(k), \tag{12}$$

where m_k and E_k are the model realization under test and its cost function, respectively, and D is dimension of model space ($D = 14$ in our case). It has been proven (e.g., Ingber 1989) that the sampling of model space controlled by equations (10) - (12) converges to a global minimum when k goes to infinity. However, we stop the sampling after the cost function E drops below some threshold; an average misfit of ∼30 m/s.

In Stage 3, we perform a random, Monte-Carlo (MC) sampling of the model space using an algorithm described by Shapiro *et al.* (1997). The MC inversion begins with the Best Fit Model, $\mathbf{m}^2$, which is also used as the reference model for the second-order truncated solution of the forward problem. The MC steps define a Markov-chain in which the current model is randomly perturbed to find the next model which is tested for acceptability. In the following step, the random search is reinitiated in the vicinity of the new acceptable model. As a result, the sampling walks randomly through model space subjected to the a-priori constraints (Table 2) forming a Markov-chain similar to Brownian motion, as shown in Figure 1. Therefore, we refer to this algorithm as Markov-chain sampling. It combines speed with efficiency in sampling the model space.

In both the SA and MC inversions, we use a cost-function defined as follows:

$$E = \sum_i W_i^{UR} \frac{|U^R_{obs}(\omega_i) - U^R_{pred}(\omega_i)|}{\sigma_i^{R,U}} + \sum_j W_j^{UL} \frac{|U^L_{obs}(\omega_j) - U^L_{pred}(\omega_j)|}{\sigma_j^{L,U}} + \\ \sum_k W_k^{CR} \frac{|C^R_{obs}(\omega_k) - C^R_{pred}(\omega_k)|}{\sigma_k^{R,C}} + \sum_l W_l^{CL} \frac{|C^L_{obs}(\omega_l) - C^L_{pred}(\omega_l)|}{\sigma_l^{L,C}} \tag{13}$$

where σ denotes the estimated uncertainties in the dispersion maps and *obs* and *pred* refer to observed and predicted velocities, respectively. The L_1-norm is used for robustness to outliers. The uncertainties in the dispersion maps are equated with the rms-misfit between the predicted and observed velocities averaged world-wide and are shown in Figure 6a. W denotes additional location-dependent weights that summarize the local quality of the dispersion maps relative to the global average. At the beginning of Stage 2, this estimate is based exclusively on the local path density of each map. However, after 5,000 steps of the SA inversion, if the cost function remains higher than an established threshold value, we modify the weight by using information about the ability to fit each datum. So, $W/\sigma \sim 1$ for a particular datum if path density is good and if the SA is able to find a model that fits that datum. An ideal case is shown in Figure 6b. If either condition is not met, however, the weight is reduced as Figures 6b-d show. In regions of poor data coverage, the path density weights vanish and the

Median Model will revert to the Initial Model and the uncertainty will grow to limits imposed by the a-priori constraints.

In both Stages 2 and 3, the a-priori constraints are applied to ensure that the selected models are physically plausible. In Stage 3, a model is considered acceptable on the value of the cost-function for the Best Fitting Model, E_2. Typically, the acceptance criterion is $1.2E_2$; that is, we accept a model if its cost (or fit) is no worse than 20% higher than the Best Fit Model. At each geographical point we test about 20,000 realizations to find an ensemble of 2,000 acceptable models. This ensemble characterizes the average properties of the structure and the uncertainty of the model.

3.3.5 Summarizing the ensemble of acceptable models

An example of the Monte-Carlo inversion is shown in Figure 7 in which the ensemble of acceptable models forms a corridor of models. The middle of this corridor at each depth defines the Median Model. Features in the Median Model are characterized in terms of perturbations to a reference model. A perturbation is considered to be "persistent" if it appears in every member of the ensemble. In particular, the perturbation at a particular depth is persistent if its value is larger than the half-width of the corridor. The model ak135 (Kennett et al. 1995) is used here as the reference model here.

The statistical properties of the ensemble of acceptable models at a point in the East European Platform (54N 30E) are shown in Figure 8. The velocity distribution of the ensemble at each depth is approximately Gaussian (Figure 8b), but the characteristics of the Gaussian are not simply related to the uncertainty of the model. In particular, the standard deviation of the distribution at each depth underestimates the model uncertainty. We use a conservative estimate of uncertainty given by the half-width of the corridor of acceptable values which, as Figure 8c shows, turns out to be about three times larger than the standard deviation of the ensemble.

3.3.6 Importance of a-priori constraints

A-priori constraints on the models selected by Monte-Carlo sampling during Stages 2 and 3 are important not only to speed the inversion by limiting the volume of model space searched, but also define what we judge to be physically reasonable or plausible candidate models. There are a number of noteworthy examples of the importance of these physical constraints. One involves the way in which v_p in the crust trades-off with radial anisotropy in the mantle. Love waves are insensitive to v_p while the Rayleigh waves have some sensitivity to v_p down up to about one eighth of a wavelength (e.g., Dahlen & Tromp 1998), which is typically in the crust. Thus, v_p variations in the crust affect long period Rayleigh waves but not Love waves.

This is illustrated in Figure 9 which shows the results of two Monte-Carlo inversions at the same location. The normal inversion is shown in Figure 9a, where the results for v_p in the crust are included. The Rayleigh-Love discrepancy is resolved mostly by radial anisotropy in the upper mantle. In the second inversion (Figure 9b), we modified the a-priori constraint on v_p in the crust and allowed much lower P-velocities. As a result, we are able to reduce the Rayleigh wave phase and group velocities without affecting the Love wave velocities, and the estimated model is nearly isotropic. It is possible to resolve the Rayleigh-Love discrepancy in this way without introducing radial anisotropy in the upper mantle. To do so, however, requires reducing middle and lower crustal P velocities to 5.4 - 6.5 km/s. Because the Rayleigh/Love discrepancy is ubiquitous, this solution would require these velocities as averages in continental regions world-wide. Studies of local and regional body wave travel times show clearly, however, that these velocities are much too low on average (e.g., Ritzwoller et al. 2002b). Therefore, models such as that in Figure 9b are physically implausible, on average, and crustal P-wave velocities must be tightly constrained in the inversion. The allowed variations in crustal P-wave speeds trade-off with the strength of radial anisotropy in the upper mantle and increase the estimated uncertainties in upper mantle anisotropy.

3.3.7 Potential problems

The inversion method described above produces a global shear-velocity model of the crust and upper mantle with estimated point-wise uncertainties. The uncertainties, however, reflect only part of the possible errors in the model, not accounting, for example, for bias in the dispersion maps caused by unmodeled wave-propagation effects such as off-great-circle propagation or scattering (e.g., Wielandt 1987; Laske 1995; Nolet & Dahlen 2000; Spetzler et al. 2001), inconsistencies that may arise from differences in the resolution of different data, errors in a-priori information, or inadequacies in the parameterization. A particular concern is that the resolution of Rayleigh and Love waves may significantly differ in strongly heterogeneous regions, such as near continent-ocean transitions. Dispersion maps at different periods may also be inconsistent. At long periods, which provide sensitivity to deep structures, Fresnel zones are larger than at short periods and the spatial resolution is worse. Therefore, the resolution of the shear-velocity model tends to degrade with depth. Our method of inversion does not account for these differences in resolution with dept, and small-scale artifacts may appear in the deep structures, particularly where there are small-scale, large amplitude structures in the shallowest mantle (e.g., near mid-oceanic ridges).

3.3.8 Computation time

The entire inversion at one geographical point requires a few hundred exact solutions of the forward problem to compute the partial derivatives and about 50,000 solutions based on the truncated approximation (simulated annealing, Monte-Carlo sampling). The number of iterations during the linearized inversion and the number of model realizations tested during the Monte-Carlo sampling varies, but, on average, the inversion at a single location takes about 3 minutes of CPU time on a current generation scientific workstation. Running the inversion for the whole Earth on a $2^\circ \times 2^\circ$ grid, therefore, requires about a month on a single processor. The inversion can be run concurrently on several processors and completed in a few days. In contrast, if the full forward solution were used rather than the second-order truncated Taylor series approximation (James & Ritzwoller 1999), the inversion would require several years on a single processor.

4. Overview of Results of Inversion

The inversion produces an ensemble of acceptable models at each spatial node on a $2^\circ \times 2^\circ$ grid world-wide. We summarize this ensemble of models with the "Median Model", which is the center of the corridor defined by the ensemble, and the uncertainties, which are identified with the half-width of the corridor at each depth. Global averages of the Median Model beneath continents and oceans are shown in Figure 10. The features of the model that are worthy of interpretation are the "persistent" features that appear in every member of the ensemble of acceptable models. An exhaustive discussion of the persistent features of the model is beyond the scope of this paper, but we will highlight some of the isotropic and anisotropic characteristics of the model and the uncertainties while concentrating discussion on the mantle part of the model.

4.1 Isotropic structure

Although we concentrate discussion on the mantle part of the model, Table 3 presents the uncertainty of estimated crustal parameters averaged over the globe. This uncertainty is the average of the half-width of the corridor of acceptable models, which should be interpreted in terms of the allowed perturbations for each parameter, presented in Table 2. For example, the average uncertainty for Moho depth is about 2.8 km although 5 km is allowed in the inversion. This means that, on average, only a sub-range of allowed Moho depths fits the data. Crustal velocities, particularly P-wave speeds, are more poorly estimated than Moho depth in that their uncertainties are a larger fraction of the allowed range of values.

Figure 11 presents horizontal slices of the Median Model at several mantle depths. The large-scale anomalies are well known from previous global tomographic studies (e.g., Woodhouse & Dziewonski 1984; Nataf et al. 1986; Montagner & Tanimoto 1991; Zhang & Tanimoto 1992, 1993; Masters et al. 1996; Ekström & Dziewonski 1998; Mégnin & Romanowicz 2000). High-velocities appear beneath all shields. Continental low-velocities appear in tectonically deformed regions such as the Red Sea rift and in back-arc regions adjacent to subduction zones. Age dependent lithospheric thickening and asthenospheric thinning is also evident beneath oceans.

A more detailed inspection of the model (e.g., Figure 12a) reveals smaller scale features that are not apparent in previous global tomographic models. Some of these features at high southern latitudes are discussed by Ritzwoller et al. (2001) and in Central Asia by Villaseñor et al. (2001). Prior to interpretation, however, it must be determined if these anomalies are persistent, are mere accidents, or are artifacts of the inversion. We plot in Figure 12b the corresponding uncertainties in the estimated S-wave velocities. At a depth of 80 km the uncertainties are smaller than the amplitudes of most of the anomalies so that most of the anomalies are persistent and worthy of interpretation. Figure 13 presents the world-wide average of uncertainties together with the rms-amplitude of v_s as a function of depth. Typically, uncertainties grow with depth in the mantle whereas anomalies decrease, so most of the features worthy of interpretation are above a depth of about 250 km.

This can also be seen clearly in Figure 14 where we present several vertical slices of isotropic upper mantle v_s. Figure 14b shows the average velocity perturbations for profile $A-A'$ crossing India, Tibet, the Tarim Basin, the Tien-Shan, Kazakhstan, and Southern Siberia. At depths less than 200 km, there is a high-velocity zone corresponding to thickened lithosphere that is especially well developed beneath India and Tibet. A strong low-velocity zone can be seen beneath India at depths larger than 200 km. The shallow low-velocity zone underlying northern Tibet is also apparent on the 80 km depth slice (Fig 12a). Figure 14c presents the uncertainties of the velocities along profile $A-A'$. Consistent with the world-wide average, the amplitude of the uncertainty increases from ~1% at the top of the mantle to more than 3% at 400 km, but the amplitude of velocity anomalies decreases with depth. The solid black contours on the vertical slices (Figures 14b, d, e) encloses the persistent model features; i.e., those features with amplitudes larger than the uncertainty. Few model features below 250 km are persistent.

Figures 14d and 14e show two other vertical slices, one across northern Eurasia and the other across the northern Pacific. The northern Eurasian profile crosses two major shields, the European platform and the Siberian shield, and shows high-velocity lithosphere beneath both shields. The thickness of the lithosphere varies along the profile and at some points appears to reach ~250 km. The structure below the lithosphere is not resolved. The northern Pacific slice reveals a number of persistent features, including: (1) a low velocity zone beneath the western United States, (2) a high-velocity oceanic lithosphere with systematically increasing thickness with age underlain by a low-velocity asthenosphere, (3) a high-velocity anomaly corresponding to the lithosphere subducting beneath Japan, and (4) a back-arc low-velocity zone beneath the Sea of Japan. Figure 10b shows how average lithospheric and asthenospheric shear velocities vary across the Pacific as a function of age (Mueller et al. 1997). Although the model parameterization has been tuned for continents, oceanic lithosphere and asthenosphere appear prominently.

4.2 Radial anisotropy

Over most of the Earth, long period Rayleigh and Love wave dispersion curves are 'inconsistent' in the sense that they cannot be fit simultaneously using a simple isotropic model. Similar to many previous studies (e.g., McEvilly 1964; Dziewonski & Anderson 1981; Gaherty & Jordan 1995; Montagner & Jobert 1988; Montagner & Tanimoto 1991; Ekström & Dziewonski 1998; Villaseñor et al. 2001) we resolve this Rayligh-Love discrepancy by introducing radial anisotropy in the upper mantle. It is true that this discrepancy can be resolved either by introducing very low P-wave speeds in the crust (e.g., Figure 9) or by allowing fine-scale oscillations in S-wave speed in the uppermost mantle (e.g., Mitchell 1984). We do not consider either alternative to be physically plausible, first, because they would have to be ubiquitous features of the upper mantle and, second, because independent evidence for anisotropy in the upper mantle is now strong (e.g., from receiver function amplitudes versus azimuth and shear wave splitting). In addition, while an oscillatory upper mantle can be successful in some locations, it cannot produce a satisfactory model everywhere. In particular, in tectonically deformed regions the inversion with the isotropic parameterization produces a high-velocity subcrustal lid with an extremely high S_n velocity ($>$ 5 km/s) that is inconsistent with recent models of S_n velocities (Ritzwoller *et al.* 2002b).

In Figure 15a we show the distribution of the strength of radial anisotropy in the Median Model, described by parameter ζ defined as:

$$\zeta = \frac{v_{sh} - v_{sv}}{v_{sv}}, \tag{14}$$

where v_{sv} and v_{sh} are taken at the top of the radially anisotropic upper mantle (Smith &

Dahlen 1973). In most regions, $\zeta \sim 4\%$ near the top of the mantle, which is similar to the value in PREM. There is no significant anomaly with negative radial anisotropy ($v_{sv} > v_{sh}$). The local uncertainty in the strength of the anisotropy is shown in Figure 15b and averages about 2%. In most places radial anisotropy in the upper mantle is a persistent model feature, but in a few regions radial anisotropy cannot be resolved. While radial anisotropy is generally persistent, its uncertainty is relatively large, averaging about 50% of the observed value. This is clearly seen in Figure 16, which presents the world-wide average of uncertainties together with the rms-amplitude of ζ as a function of depth. Very strong anisotropy is found in some oceanic regions and in some tectonically deformed zones within continents (e.g. Tibet, Iran, eastern Africa). Consistent with Ekström & Dziewonski (1998), radial anisotropy is very strong in the central Pacific. Strong anisotropy may be caused in some places by differential lateral resolution of Rayleigh and Love waves if the Rayleigh waves resolve a low velocity features but the Love waves do not.

As shown in Figures 15 and 16, positive radial anisotropy ($v_{sh} > v_{sv}$) is a persistent feature of the Earth's upper mantle. Radial anisotropy can only be resolved unambiguously by incorporating short and intermediate period measurements in our whole data-set. This is illustrated by Figure 17, which shows uncertainties in the strength of radial anisotropy beneath Eastern Antarctica with different subsets of data. The estimated uncertainty is very large when we use only long-period data (i.e., phase velocities with periods >70 s) as in Figure 17a) and reduces by almost a factor of three when we use the entire data set of broad-band group and phase velocity measurements.

In summary, a (PREM-like) parameterization of radial anisotropy in which the strength of anisotropy decreased monotonically with depth below the Moho is sufficient to fit Rayleigh- and Love-wave dispersion measurements world-wide and resolve the Rayleigh-Love discrepancy. Even using this simple parameterization, we obtain very large ($\sim 50\%$) uncertainties in the strength of radial anisotropy. Increasing the number of parameters describing the radial anisotropy would increase uncertainties further. Therefore, more complicated vertical patterns of radial anisotropy cannot be resolved using surface-wave data alone. This appears to be true also for the unique anisotropy of the central Pacific suggested by Ekström & Dziewonski (1998). Support for the more complicated vertical pattern of the strength of radial anisotropy in this region may come, however, from the other types of data and longer period data employed by Ekström & Dziewonski (1998).

5. Conclusions

Two main characteristics distinguish the global model that we present here from previous global models of the crust and upper mantle. First, both vertical and lateral resolution are improved as a result of inverting a new broad-band data set of group and phase velocities and using a-priori information to restrict the range of physically plausible models. A novel characteristic of the data set is the use of a large number of group velocity measurements for both Rayleigh and Love waves. The group velocity data contain short and intermediate period information that improves vertical resolution for both isotropic (e.g., Figure 18) and radially anisotropic (Figure 17) structures. As Figure 18 illustrates, using phase velocities alone produces large uncertainties in the crust and, consequently, also in the upper mantle. Inverting group velocities alone produces smaller uncertainties in the crust and uppermost mantle due to the measurements at periods shorter than 40s, but uncertainties deeper in the upper mantle are larger. When phase and the group velocities are inverted simultaneously, however, uncertainties are significantly reduced at all depths. A-priori information is required because surface wave data alone are insufficient to resolve all of the model parameters unambiguously. Therefore, we constrain crustal structures in order to help resolve the trade-off between crustal and upper mantle velocities and have identified and attempt to resolve other important trade-offs, such as that between the strength of radial anisotropy in the upper mantle and crustal P-wave velocities.

The second important characteristic of the model is that it contains estimates of uncertainties. The model, therefore, is perhaps the first global model with meaningful “error bars”. The uncertainty estimates derive from a multi-step inversion procedure that culminates in a Monte-Carlo sampling of model space to produce an ensemble of acceptable models. The features that appear in every member of this ensemble are termed “persistent” and only these features are deemed to be worthy of interpretation. The uncertainty analysis indicates that our surface wave data resolve upper mantle structures to depths of about 250 km.

The procedure we describe here, and the data set to which it is applied, open new possibilities for the study of the crust and upper mantle structure. Because of its relatively high lateral resolution, the model reveals anomalies at scales that are relevant to regional tectonics. The breadth of the frequency band improves vertical resolution which also is important for interpretation and use of the model (e.g., Levshin & Ritzwoller 2002). Finally, the uncertainty analysis allows identification of those features of the model that are worthy of interpretation in the framework of regional tectonic and geodynamical processes.

ACKNOWLEDGMENTS

We gratefully acknowledge the staffs at the IRIS-DMC and the GEOSCOPE data center for providing most of the waveform data on which the dispersion measurements were obtained. We are also particularly grateful to Jeannot Trampert at Utrecht University and Michael Antolik, Adam Dziewonski, and Goran Ekstrom at Harvard University for providing phase velocity measurements and to Robert van der Hilst and Brian Kennett for donating waveform data from the SKIPPY and KIMBA arrays in Australia. The SKIPPY and KIMBA data sets were collected by the Research School of Earth Sciences, Australian National University. Keith McLaughlin and Jeannot Trampert provided insightful technical reviews and Steven Ward's editorial review also improved the manuscript. Gabi Laske provided sedimentary and crustal thickness models prior to publication. All maps were generated with the Generic Mapping Tools (GMT) data processing and display package (Wessel and Smith 1991, 1995). Aspects of this work were supported by grants from the Office of Polar Programs at the U.S. National Science Foundation, NSF-OPP-9615139 and NSF-OPP-9818498, and Defense Threat Reduction Agency contracts DTRA01-99-C-0019 and DTRA01-00-C-0013.

REFERENCES

Barmin, M.P., Levshin, A.L., & Ritzwoller, M.H., 2001. A fast and reliable method for surface wave tomography, *Pure appl. Geophys.*,158(8), 1351 - 1375, 2001.

Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton University Press, Princeton, New Jersey.

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

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

Ekström, G., Tromp, J., & Larson, E.W.F., 1997, Measurements and global models of surface waves propagation, *J. geophys. Res.,* **102**, 8137-8157.

Engdahl, E.R., R. van der Hilst, and R. Buland, 1998. Global teleseismic earthquake relocation with improved travel time and procedures for depth determination, *Bull. Seism. Soc. Am.*, **88**, 722 - 743.

Gaherty, J.B. & Jordan, T.H., 1995. Lehmann discontinuity as the base of an anisotropic layer beneath continents, *Science*, **268**, 1468-1471.

Gilks, W.R., Richardson, S., and D.J. Spiegelhalter, Markov chain Monte Carlo in practice, eds., 486pp., Chapman and Hall, London, 1996.

Ingber, L., 1989. Very fast simulated re-annealing, *J. Math. Comput. Model.*, **12**, 967-973.

James, M.B. & Ritzwoller, M.H., 1999. Feasibility of truncated perturbation expansions to approximate Rayleigh wave eigenfrequencies and eigenfunctions in heterogeneous media, *Bull. seism. Soc. Am.,* **89**, 433-442.

Keilis-Borok, V.I. and Yanovskaya, T.B., 1967. Inverse problems of seismology (structural review), *Geophys. J. R. astr. Soc.,* , **13**, 223-234.

Kennett, B.L.N., Engdahl, E.R., & Buland, R., 1995. Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.,* **122**, 403-416.

Kirkpatrick, S., Gelatt, C.D., & Vecchi, M.P., 1983. Optimization by simulated annealing, *Science*, **220**, 671-680.

Laske, G., 1995. Global observation of off-great-circle propagation of long-period surface waves, *Geophys. J. Int.,* **123**, 245-259.

Laske, G. & Masters, G., 1997. A global digital map of sediments thickness (abstract), *EOS Trans. AGU*, **78**, F483.

Levshin, A.L., Sabitova, T.M., & Valus, V.P., 1966. Joint interpretation of body and surface waves data for a district in Middle Asia, *Geophys. J. R. astr. Soc.,* **11**, 57-66.

Levshin, A.L. and M.H. Ritzwoller, Application of a global 3-D model to improve regional event locations, *Stud. Geophys. Geod.*, in press.

Levshin, A.L., Yanovskaya, T.B., Lander, A.V., Bukchin, B.G., Barmin, M.P., Ratnikova, L.I., & Its, E.N., 1989. Recording, identification, and measurement of surface wave parameters. In: *Seismic surface waves in a laterally inhomogeneous Earth* (Keilis-Borok, V.I., editor). Kluwer Academic

Publisher, Dordrecht, 131-182.

Levshin, A.L., Ritzwoller, M.H., Barmin, M.P., & Villaseñor, A., 2001. New constraints on the Arctic crust and uppermost mantle: Surface wave group velocities, P_n, and S_n, *Phys. Earth planet. Inter.,* **123**, 185-204.

Lomax, A. & Snieder, R., 1994. Finding sets of acceptable solutions with a genetic algorithm with application to surface wave group velocity dispersion in Europe, *Geophys. Res. Lett.,* **21**, 57-66.

Marquering, H., Snieder, R., & Nolet, G., 1996. Waveform inversions and the significance of surface mode coupling, *Geophys. J. Int.,* **124**, 258-270.

Masters, G., Johnson, S. Laske, G., & Bolton, H., 1996. A shear-velocity model of the mantle, *Phil. Trans. R. Soc. Lond. A.*, **354**, 1385-1411.

McEvilly, T.V., 1964. Central U.S. crust - upper mantle structure from Love and Rayleigh wave phase velocity inversion, *Bull. seism. Soc. Am.,* **54**, 1997-2015.

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

Metropolis, N, Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., & Teller, E., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.*, **21**, 1087-1092.

Mitchell, B.J., 1984. On the inversion of Love- and Rayleigh-wave dispersion and implications for the Earth structure and anisotropy, *Geophys. J. R. astr. Soc.,* **76**, 233-241.

Montagner, J.-P. & Jobert, N., 1988. Vectorial tomography. II: Application to the Indian Ocean, *Geophys. J. R. astr. Soc.,* **94**, 309-344.

Montagner, J.-P. & Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies, *J. geophys. Res.,* **96**, 20,337-30,351.

Mooney, W.D., Laske, G., & Masters, G., 1998. CRUST5.1: A global crustal model at 5° × 5°, *J. geophys. Res.,* **103**, 727-747.

Mosegaard, K. and Tarantola, A., 1995. Monte-Carlo sampling of solutions to inverse problems, *J. geophys. Res.,* **100**, (B7) 12431-12447.

Mueller, R.D., W.R. Roest, J.-Y. Royer, L.M. Gahagan, and J.G. Sclater, 1997. Digital isochrons of the world's ocean floor, *J. geophys. Res.,* **102**, 3211-3214.

Nataf, H.-C., Nakanishi, I., & Anderson, D.L., 1986. Measurements of mantle wave velocities and inversion for lateral heterogeneities an anisotropy, *J. geophys. Res.,* **91**, 7261-7307.

Nolet, G., 1990. Partitioned wave-form inversion and 2-dimensional structure under the network of autonomously recording seismographs, *J. geophys. Res.,* **95**, 8499-8512.

Nolet, G. & Dahlen, F.A., 2000. Wave front healing and the evolution of seismic delay times, *J.*

geophys. Res., **105**, 19,043-19,054.

Press, F., 1968. Earth models obtained by Monte-Carlo inversion, *J. geophys. Res.,* **67**, 1647-1658.

Ritzwoller, M.H. & Levshin, A.L., 1998. Eurasian surface wave tomography: group velocities, *J. geophys. Res.,* **103**, 4839-4878.

Ritzwoller, M.H., Shapiro, N.M., Levshin, A.L. and Leahy, G.M., 2001. The structure of the crust and upper mantle beneath Antarctica and the surrounding oceans, *J. geophys. Res.,* **106**(B12), 30645 - 30670.

Ritzwoller, M.H., Shapiro, N.M., Barmin, M.P., and Levshin, A.L., 2002a. Global surface wave diffraction tomography, submitted to *J. geophys. Res.*

Ritzwoller, M.H., M.P. Barmin, A. Villasenor, A.L. Levshin, and E.R. Engdahl, 2002b. P_n and S_n tomography across Eurasia, *Tectonophys.*, in press.

Seber, D., Vallve, M., Sandvol, E.A., Steer, D.N., & Barazangi, M., 1997. Middle East tectonics; applications of geographical information systems (GIS), *GSA Today*, **7**, 1-6.

Shapiro N.M., Campillo, M., Paul, A., Singh, S.K., Jongmans, D., & Sánchez-Sesma, F.J., 1997. Surface wave propagation across the Mexican Volcanic Belt and the origin of the long-period seismic wave amplification in the Valley of Mexico, *Geophys. J. Int.,* **128**, 151-166.

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

Snieder, R., 1988. Large-scale waveform inversions of surface waves for lateral heterogeneities, *J. geophys. Res.,* **93**, 12,055-12,065.

Spetzler, J., Trampert, J., & Snieder, R., 2001 Are we exceeding the limits of the great circle approximation in global surface wave tomography? *Geophys. Res. Lett.,* , **28**, 2341-2345.

Trampert, J. & Woodhouse, J.H., 1995. Global phase velocity maps of Love and Rayleigh waves between 40 and 150 s period, *Geophys. J. Int.,* **122**, 675-690.

Villaseñor, A., Ritzwoller, M.H., Levshin, A.L., Barmin, M.P., Engdahl, E.R., Spakman, W., & Trampert, J., 2001. Shear velocity structure of Central Eurasia from inversion of surface wave velocities, *Phys. Earth planet. Inter.,* **123**, 169-184, 2001.

Wessel, P., and W.H.F. Smith, 1991. Free software helps map and display data, *EOS, Trans. Am. geophys. Un.,* **72**, 441.

Wessel, P., and W.H.F. Smith, 1995. New version of the Generic Mapping Tools released, *EOS, Trans. Am. geophys. Un.,* **76**, 329, 1995.

Wielandt, E., 1987. On the validity of the ray approximation for interpreting delay times, in: *Seismic Tomography* (G. Nolet, ed.), 85-98, Norwell, Mass..

Woodhouse, J. H., 1988. The calculation of the eigenfrequencies and eigenfunctions of the free oscillations of the Earth and the Sun, in: *Seismological Algorithms* (D. J. Doornbos, ed.), 321-370.

Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle: three dimensional modeling of Earth's structure by inversion of seismic waveforms, *J. geophys. Res.,* **89**, 5953-5986.

Zhang, Y.S. & Tanimoto, T., 1992. Ridges, hot spots, and their interaction as observed in seismic velocity maps, *Nature*, **355**, 45-49.

Zhang, Y.S. & Tanimoto, T., 1993. High-resolution global upper mantle structure and plate tectonics, *J. geophys. Res.,* **98**, 9793-9823.

Table 1. Outline of the Shear Velocity Inversion

Substep	Input Model	Method	Output Model(s)
Substep 1	Initial Model ($\mathbf{m^0}$)	Linearized Inversion	Simple Reference Model ($\mathbf{m^1}$)
Substep 2	Simple Reference Model ($\mathbf{m^1}$)	Simulated Annealing	Best Fit Model ($\mathbf{m^2}$)
Substep 3	Best Fit Model ($\mathbf{m^2}$)	Monte-Carlo	Ensemble of Acceptable Models ($\mathbf{m^3}$)

Table 2. A Priori Constraints on Allowed Models

Feature	Allowed Range	Reference
crustal thickness	±5 km	CRUST5.1
crustal v_s	± 300 m/s	Simple Reference Model
upper & middle crustal v_p	± 200 m/s	CRUST5.1
lower crustal v_p	± 300 m/s	CRUST5.1
crustal v_s	monotonic increase	-
crustal v_p	monotonic increase	-
v_{sh} & v_{sv} in uppermost mantle	±250 m/s	Simple Reference Model
bottom of zone of radial anisotropy	± 30 km	220 km
mantle isotropic v_s velocity	unconstrained	-

Table 3. Average Uncertainty of Estimated Crustal Parameters

Feature	Uncertainty
crustal thickness	2.76 km
crustal v_s, Upper Layer	188 m/s
crustal v_s, Middle Layer	198 m/s
crustal v_s, Lower Layer	209 m/s
crustal v_p, Upper Layer	152 m/s
crustal v_p, Middle Layer	162 m/s
crustal v_p, Lower Layer	242 m/s

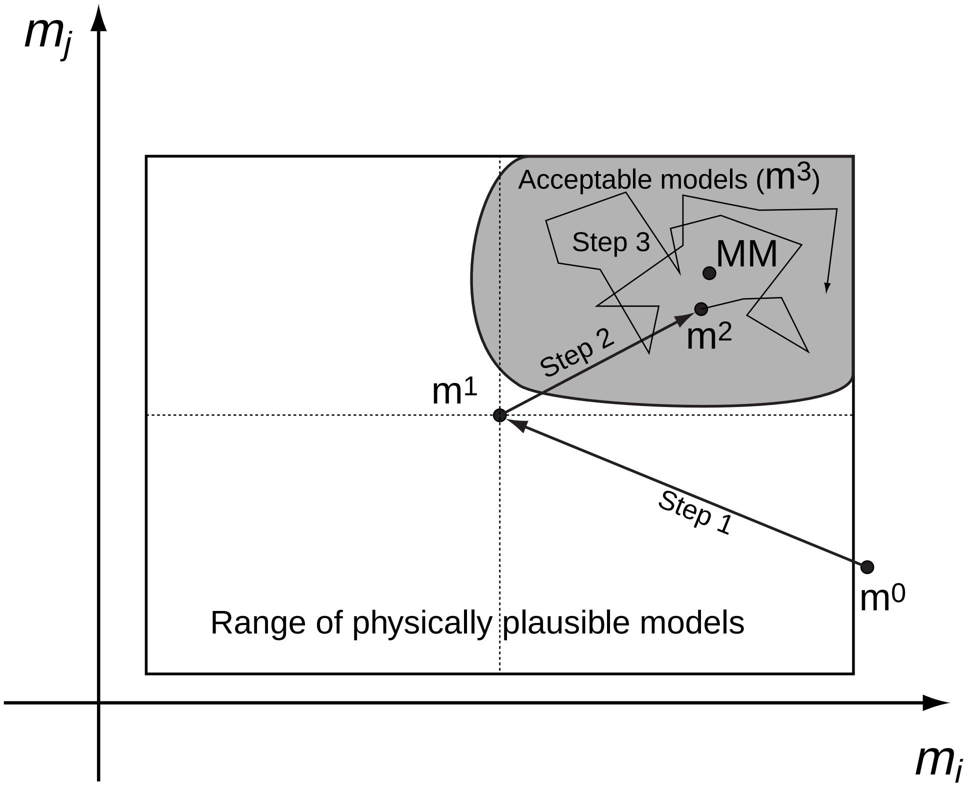


Figure 1. Schematic representation of the three-step inversion procedure projected onto a 2-D subspace of model space. Step 1 is the Linearized Inversion, which begins with the Initial Model ($\mathbf{m}^0$) and produces the Simple Reference Model ($\mathbf{m}^1$). In Step 2, the Best-Fit Model ($\mathbf{m}^2$) is found by simulating annealing. In Step 3, the inversion procedure culminates with Monte-Carlo resampling of model space producing an ensemble of acceptable models ($\mathbf{m}^3$). This ensemble is used to estimate the Median Model (MM) and the spread of acceptable models, which is interpreted as the model uncertainty.

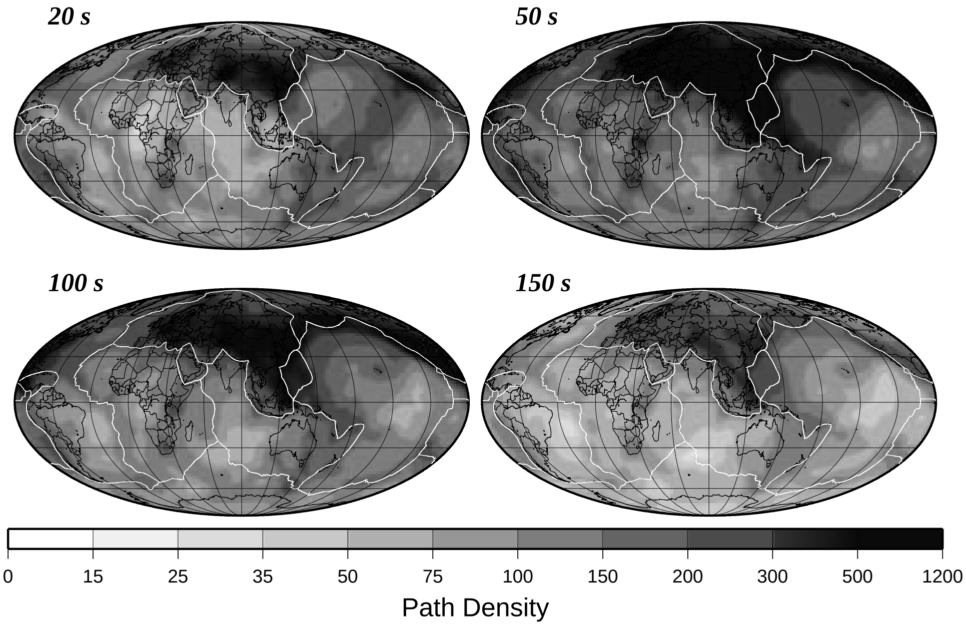


Figure 2. Examples of path density for Rayleigh wave group velocities, plotted as the number of paths intersecting each 2° × 2° cell ($\sim 50{,}000$ km^2).

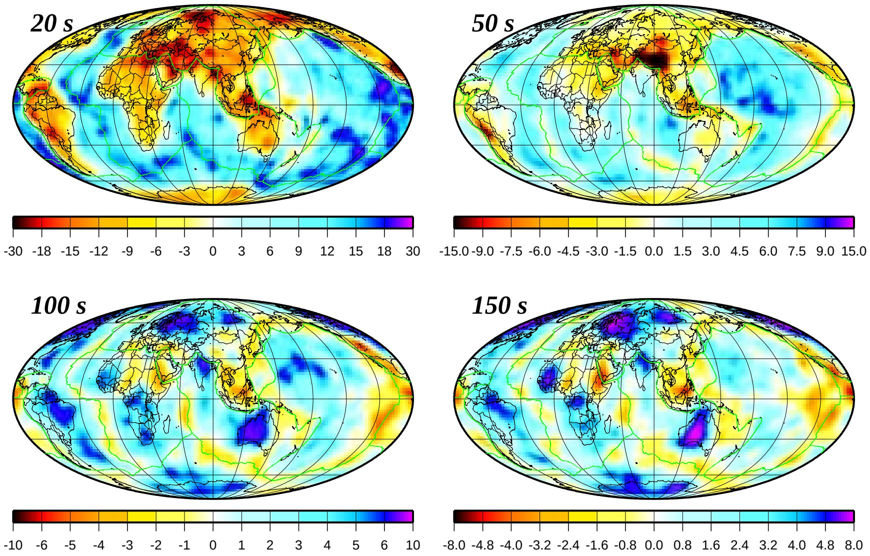


Figure 3. Examples of Rayleigh wave group velocity maps at the four indicated periods displayed as percent perturbation to the PREM velocity.

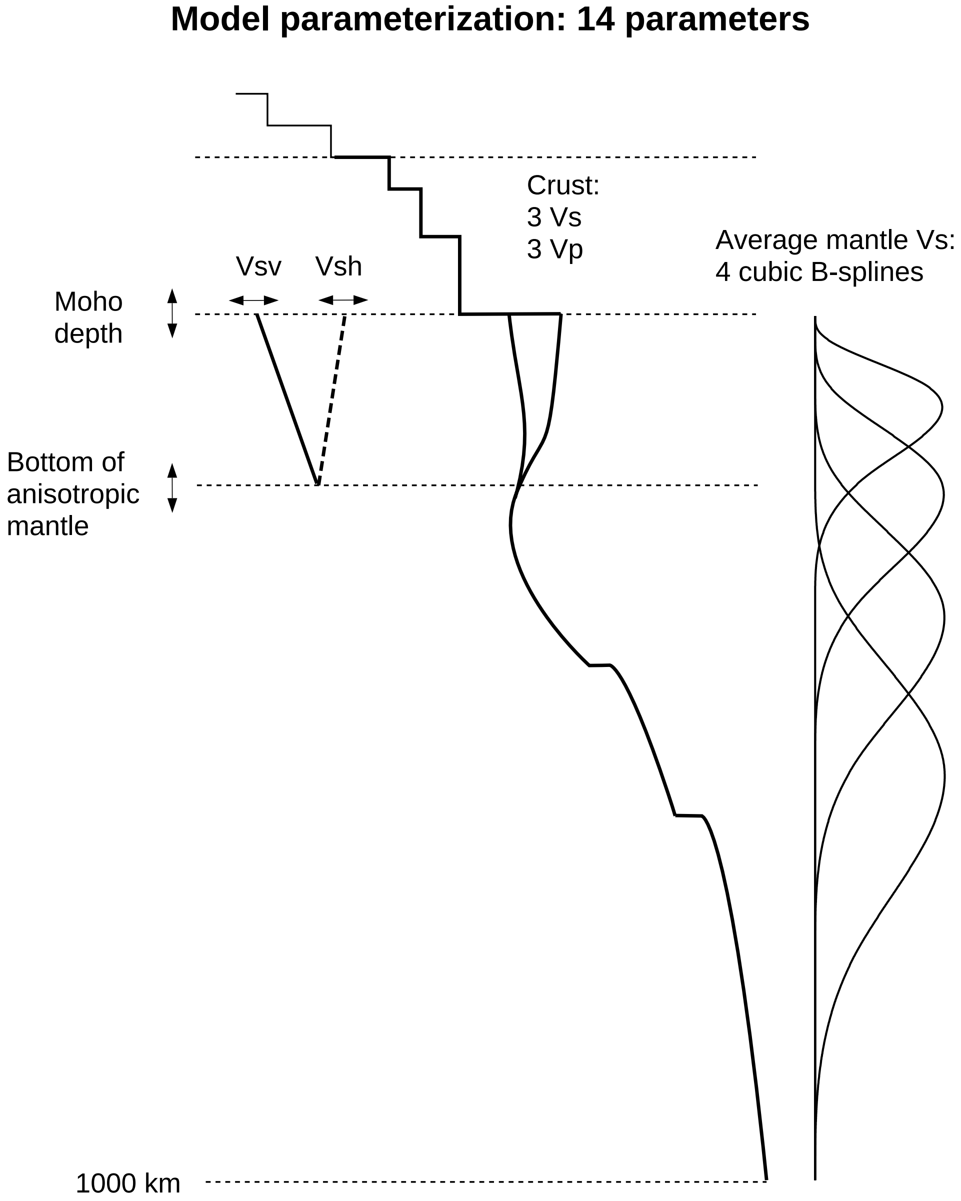


Figure 4. Model parameterization including 14 parameters: (1-3) Crustal S-wave velocities, (4-6) crustal P-wave velocities, (7) Moho depth, (8) v_{sv} beneath Moho, (9) v_{sh} beneath Moho, (10) depth of the bottom of the anisotropic mantle, (11-14) cubic B-spline perturbations to the average mantle S-wave velocity.

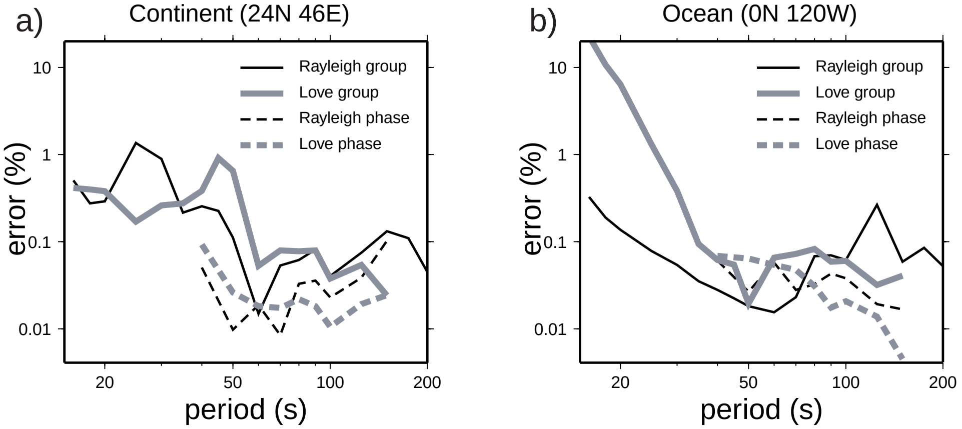


Figure 5. Relative errors in the dispersion curves computed by using the truncated second-order Taylor series approximation (eqn. 9). (a) Example of errors at a continental point (Arabian Peninsula). Shear velocities have been perturbed by ±5% at all depths, which is larger than perturbations considered during the simulated annealing and the Monte-Carlo inversions. At all periods the errors are less than $\sim$ 1%. (b) Example of errors at an oceanic point (Central Pacific). Shear velocities have been perturbed by ±5% in the mantle and by ±2.5% in the crust. The errors are small except for Love-wave group velocities at periods less than $\sim$25 s, but measurements of short period Love waves are rare in oceanic areas.

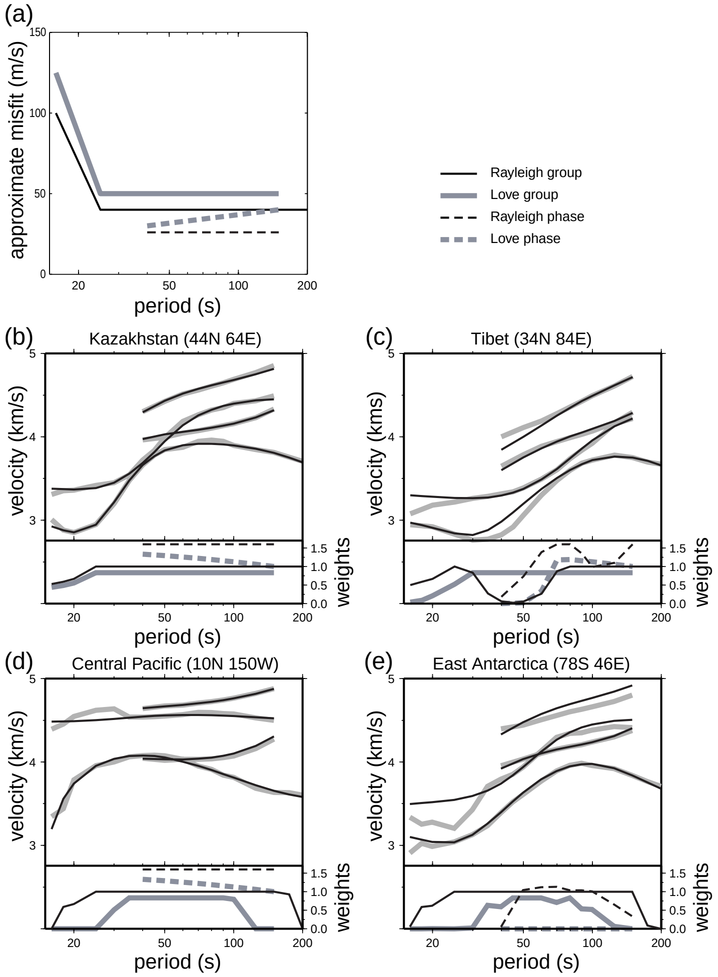


Figure 6. (a) Measurement uncertainties (σ, eqn (13)) defined as the average RMS misfit of the tomographic maps to dispersion measurements. (b)-(e) Each plot is divided into upper and lower frames: (upper) the observed (gray lines) and predicted (black lines) dispersion curves and (lower) corresponding location-specific weights, W/σ. (b) A point in Kazakhstan exemplifying an ideal situation with high path density and little inconsistency between different data types. The weights are then controlled only by σ shown in (a). (c) A point in Tibet where the inversion cannot fit all the data simultaneously and, therefore, weights down certain data types in particular period bands. (d) A point in the Central Pacific where the Love-wave group velocities are down-weighted at short and long periods because of low path density. (e) A point in Antarctica where Rayleigh-wave phase velocities and long-period Lowe-wave group velocities are down-weighted because of low path density. Love-wave phase velocity path density is so low that data weights are set to zero at all periods. The short-period Love waves are down-weighted because the inversion cannot fit the dispersion maps, probably because of inaccuracies in the a-priori crustal model.

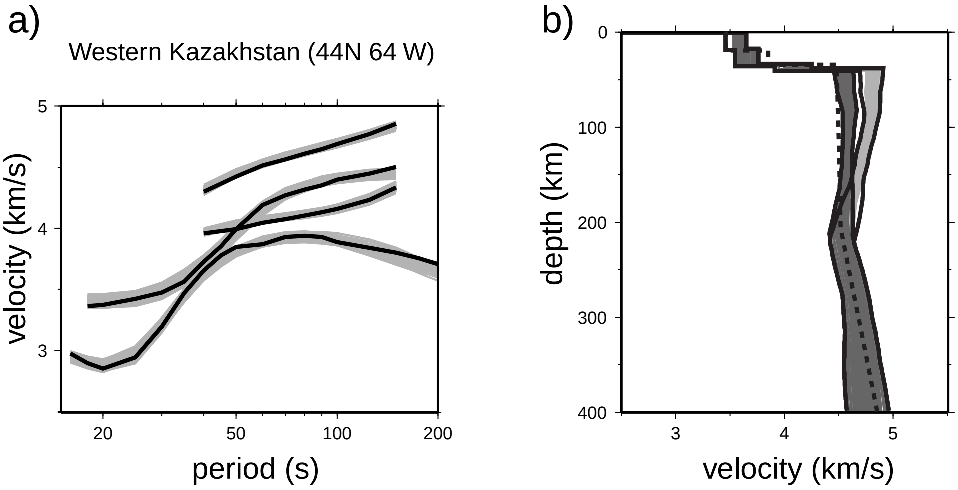


Figure 7. Example of the inversion for an ensemble of acceptable shear velocity models at a point in Western Kazakhstan (44 N, 64 E). (a) Four dispersion curves obtained from tomographic velocity maps (thick black lines) and the predictions from the ensemble of acceptable models (gray lines). (b) The ensemble of acceptable models where SV and SH velocities are presented with dark and light gray shades, respectively. The corridor of acceptable values is indicated with the solid black lines. The S-wave velocity from the global reference model ak135 is plotted as the dashed line.

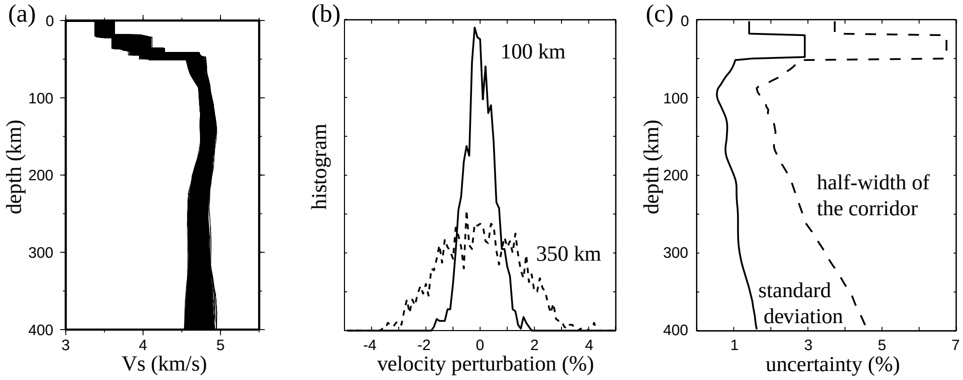


Figure 8. (a) Ensemble of acceptable models at a point in the East European Platform (54N 30E). Only the isotropic part of the model is plotted. (b) Histograms of the velocity perturbations in the ensemble at two depths: 100 km (solid line) and 350 km (dashed line). (c) Estimates of uncertainty obtained using the ensemble of acceptable models. The standard deviation of velocity at each depth is shown with the solid line and the half-width of the corridor of acceptable values is shown with the dashed line.

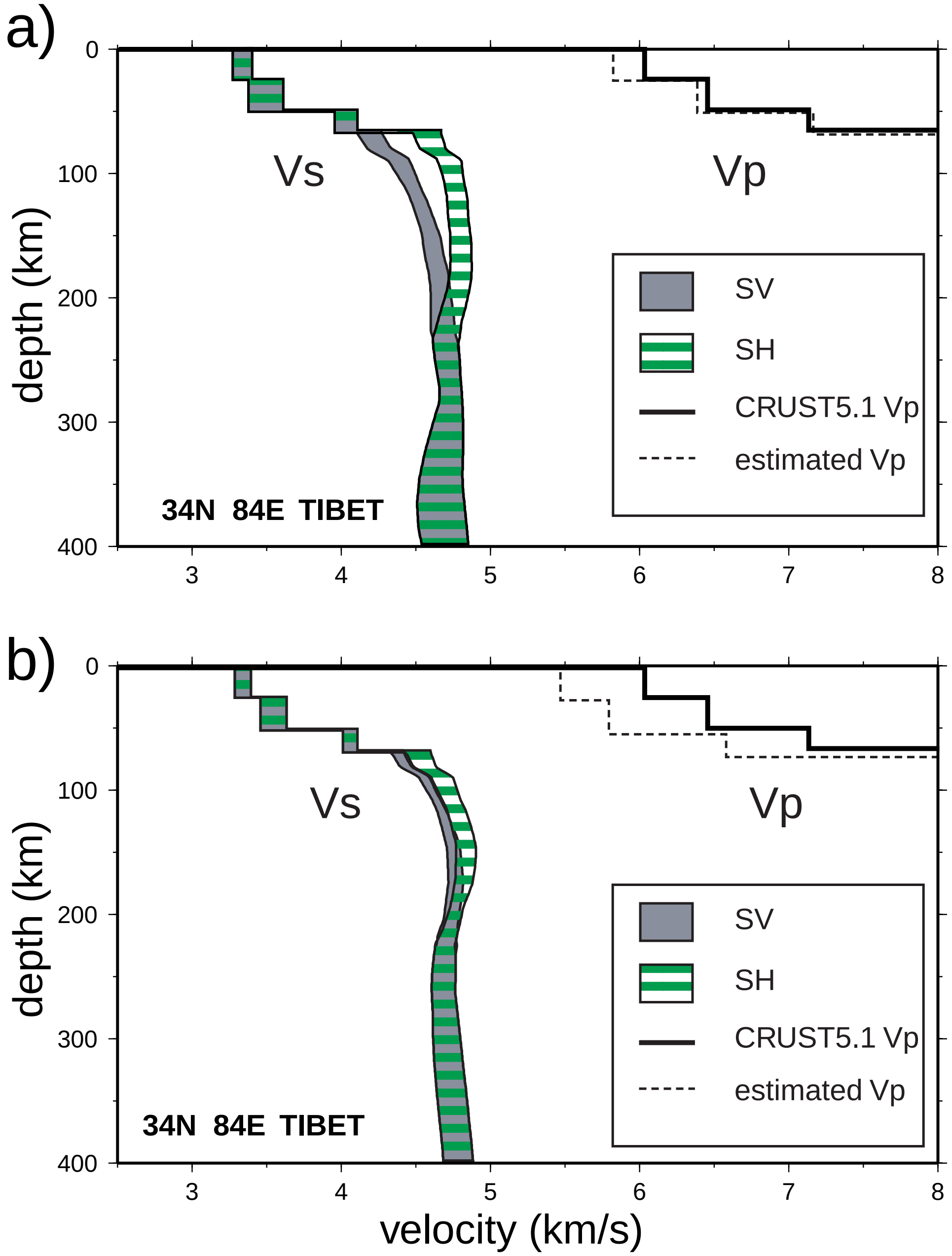


Figure 9. Examples of inversions at a point in Tibet illustrating the trade-off between the strength of radial anisotropy and the crustal P-wave speed. (a) The tolerances in Table 2 are applied to crustal P-wave velocities from CRUST5.1. (b) The same tolerances are applied to a crustal reference model in which the CRUST5.1 P-wave velocities are reduced by 10%. Reduced P-wave speeds in the crust can resolve the Rayleigh-Love discrepancy, but produce a nonphysical model.

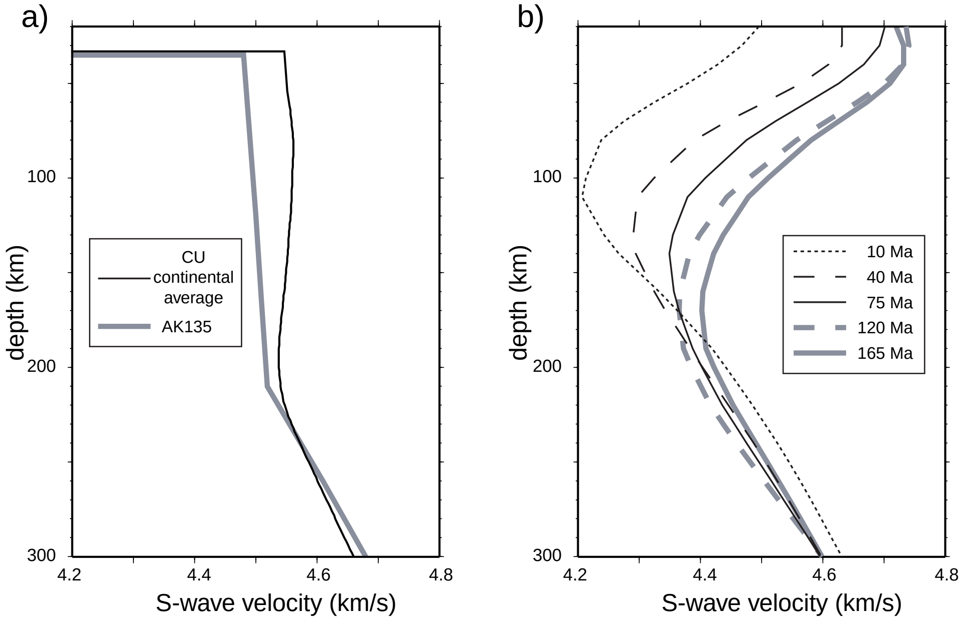


Figure 10. (a) Global average of the Median Model under continents compared with the S-wave velocity from the 1-D model ak135. (ak135 is more similar to tectonic continental structures than average structures across continents.) (b) Vertical profiles of isotropic shear velocity averaged as a function of lithospheric age in the Pacific Ocean. Five profiles are shown at ages of 10, 40, 75, 120, and 165 Ma, averaged within an age window of age $\pm\sqrt{\text{age}}$.

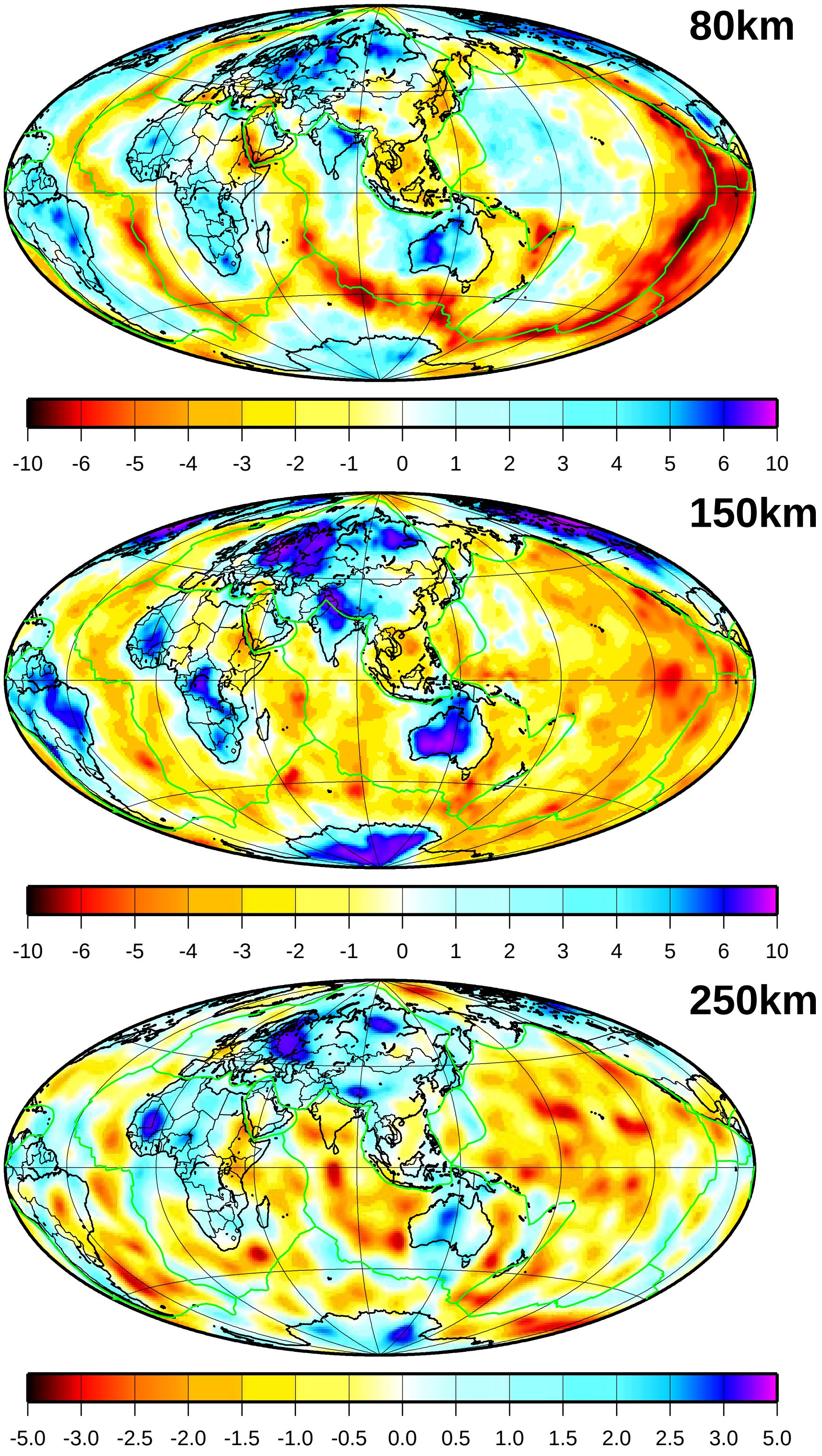


Figure 11. Isotropic S-wave velocities ($v_s = (v_{sv} + v_{sh})/2$) from the Median Model at three depths in the upper mantle, presented as percent deviation from model ak135.

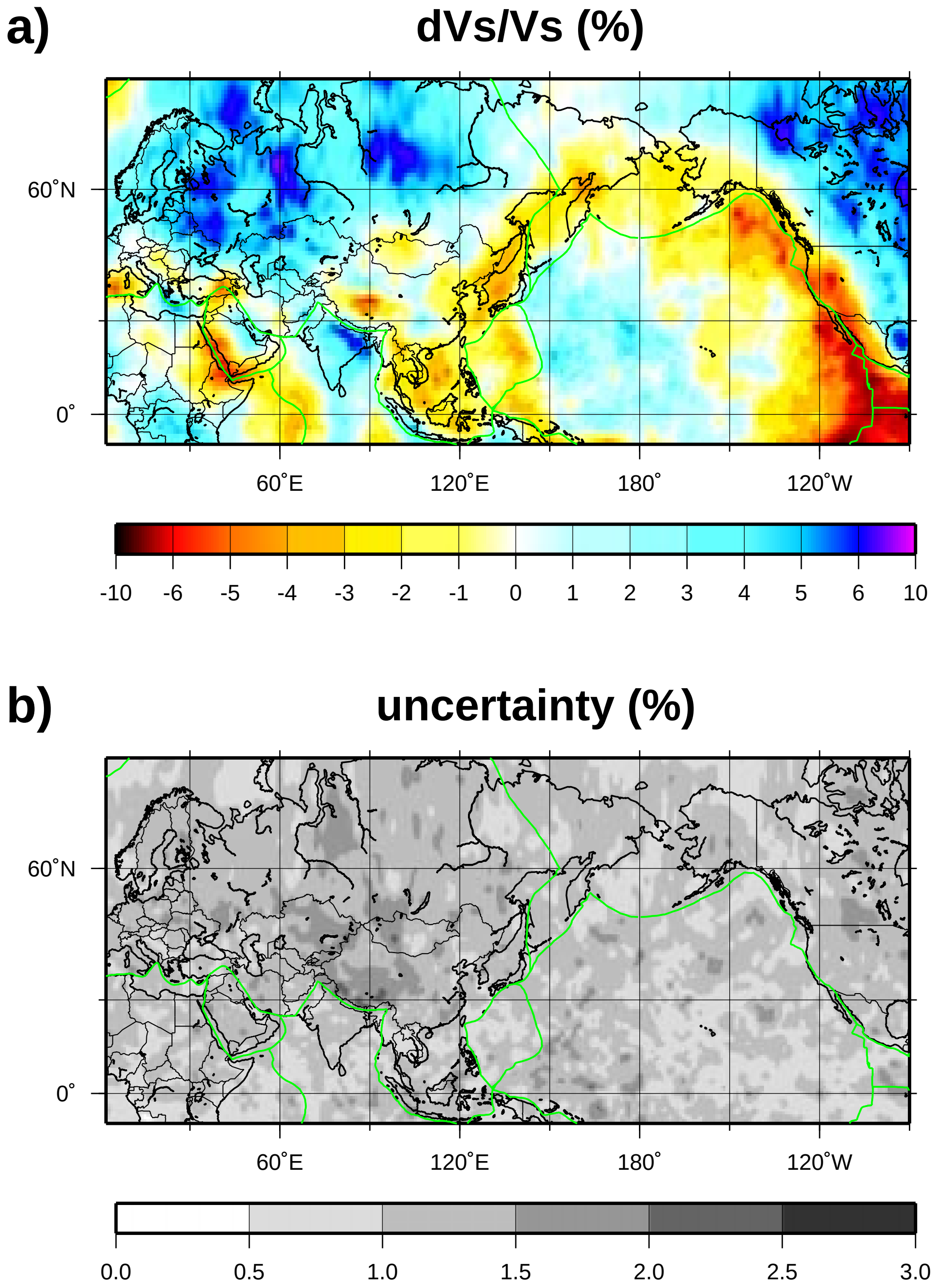


Figure 12. (a) The isotropic S-wave velocity from the Median Model at 80 km presented as percent deviation from model ak135. (b) Uncertainty in isotropic S-wave velocity at 80 km defined as the half-width of the corridor of acceptable values.

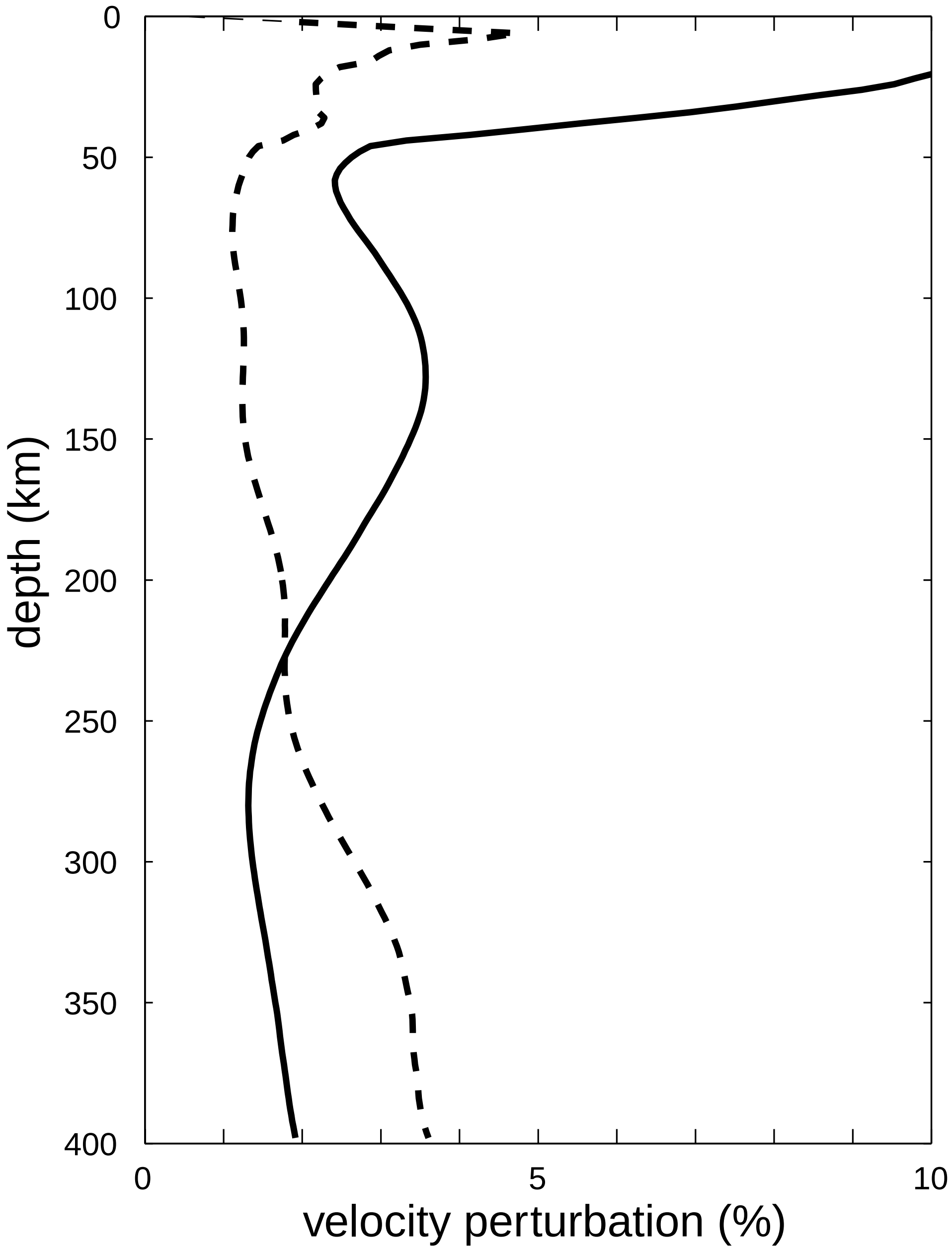


Figure 13. Global RMS shear-velocity perturbations in isotropic S-wave velocity (solid line) and global average shear-velocity uncertainty (dashed line) presented as percent deviation from the S-wave velocities in the 1-D model ak135.

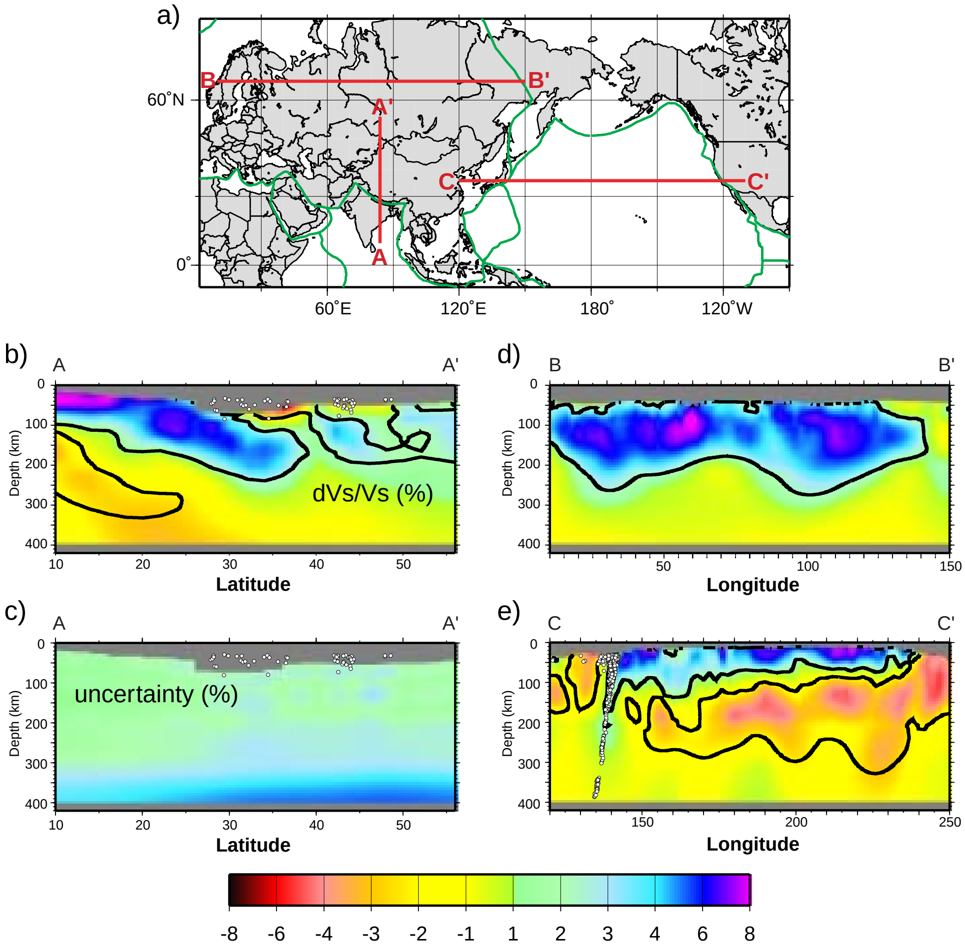
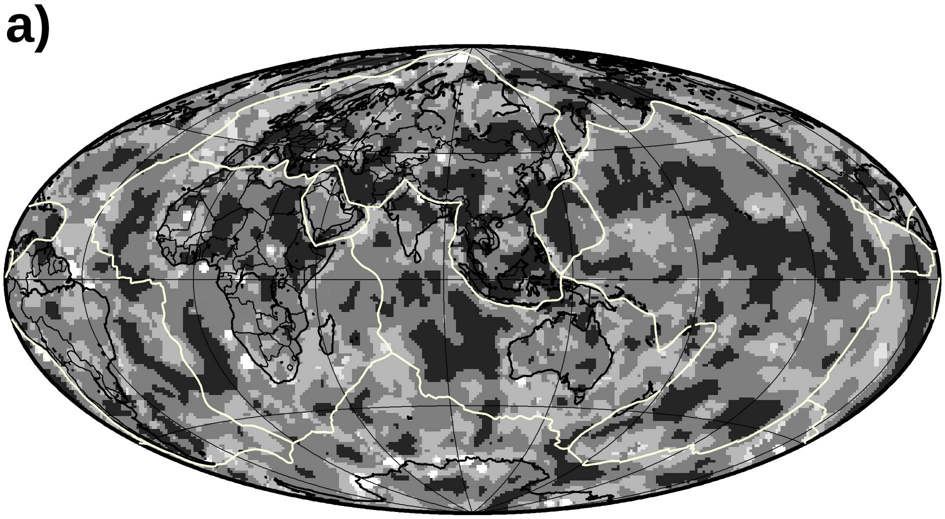


Figure 14. Vertical slices of S-wave velocity. (a) Map showing the locations of three profiles. (b), (d), and (e) Isotropic S-wave velocity ($v_s = (v_{sv} + v_{sh})/2$) beneath the three profiles shown in (a), presented as percent deviation from the S-wave speed in the 1-D model ak135. Black contours outline the persistent velocity anomalies. (c) Uncertainties in the isotropic S-wave velocity for the profile A-A'. The same color scale is used for estimated velocities and uncertainties. Earthquake locations, shown as small circles in (b) - (e), are taken from Engdahl et al. (1998).



strength of the radial anisotropy (%)

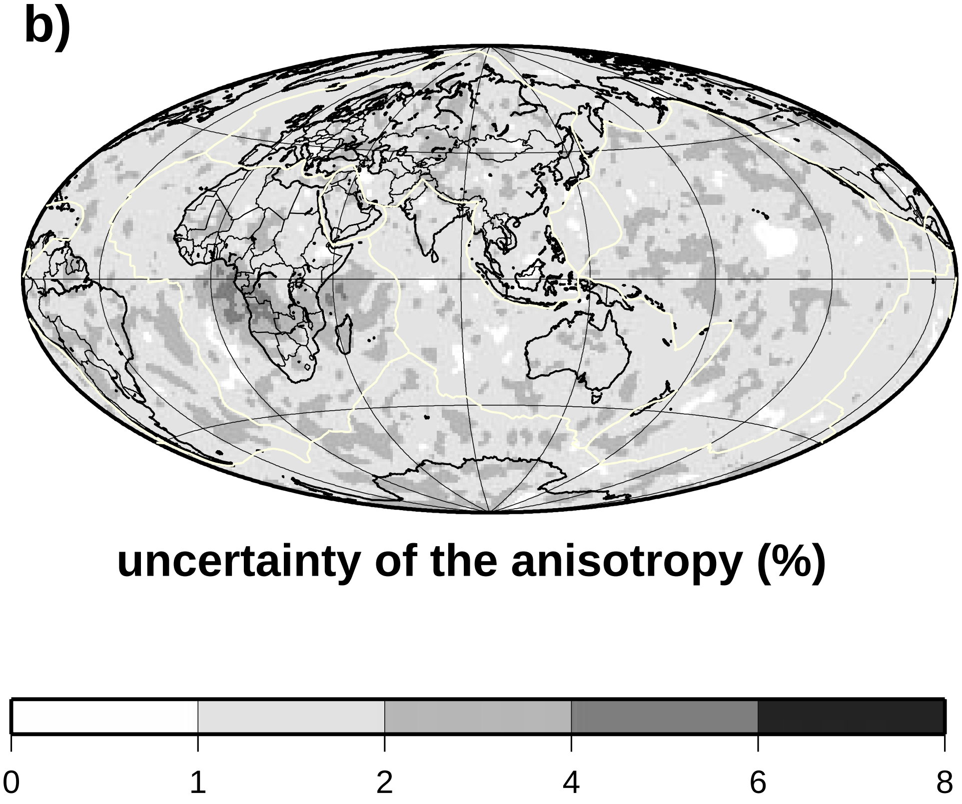


Figure 15. (a) Global distribution of the strength of radial anisotropy at the top of the mantle: $\zeta = (v_{sh} - v_{sv})/v_{sv}$. (b) Uncertainty in the strength of radial anisotropy. The same color scale is used for both (a) and (b).

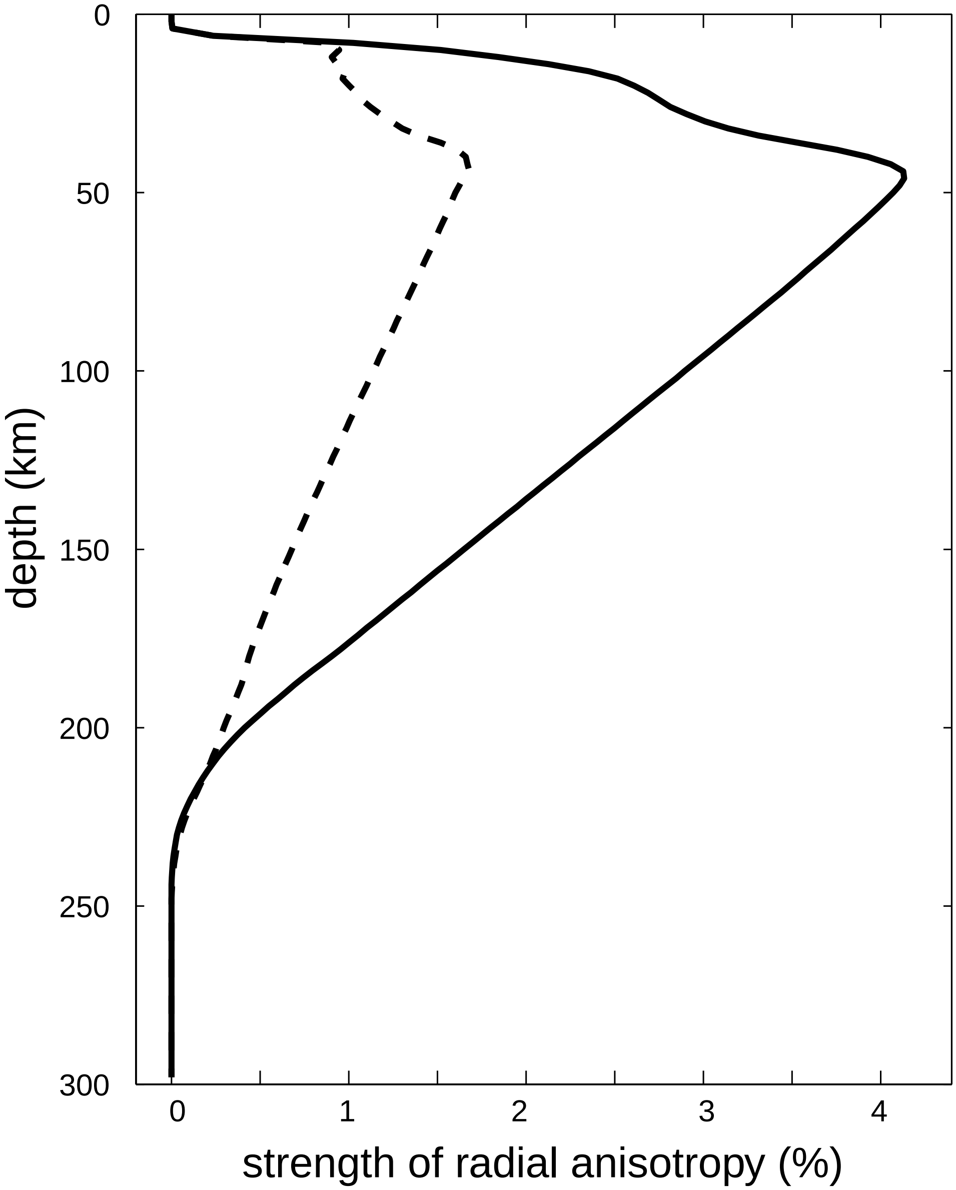


Figure 16. Global average strength of radial anisotropy (solid line) and its uncertainty (dashed line) as functions of depth.

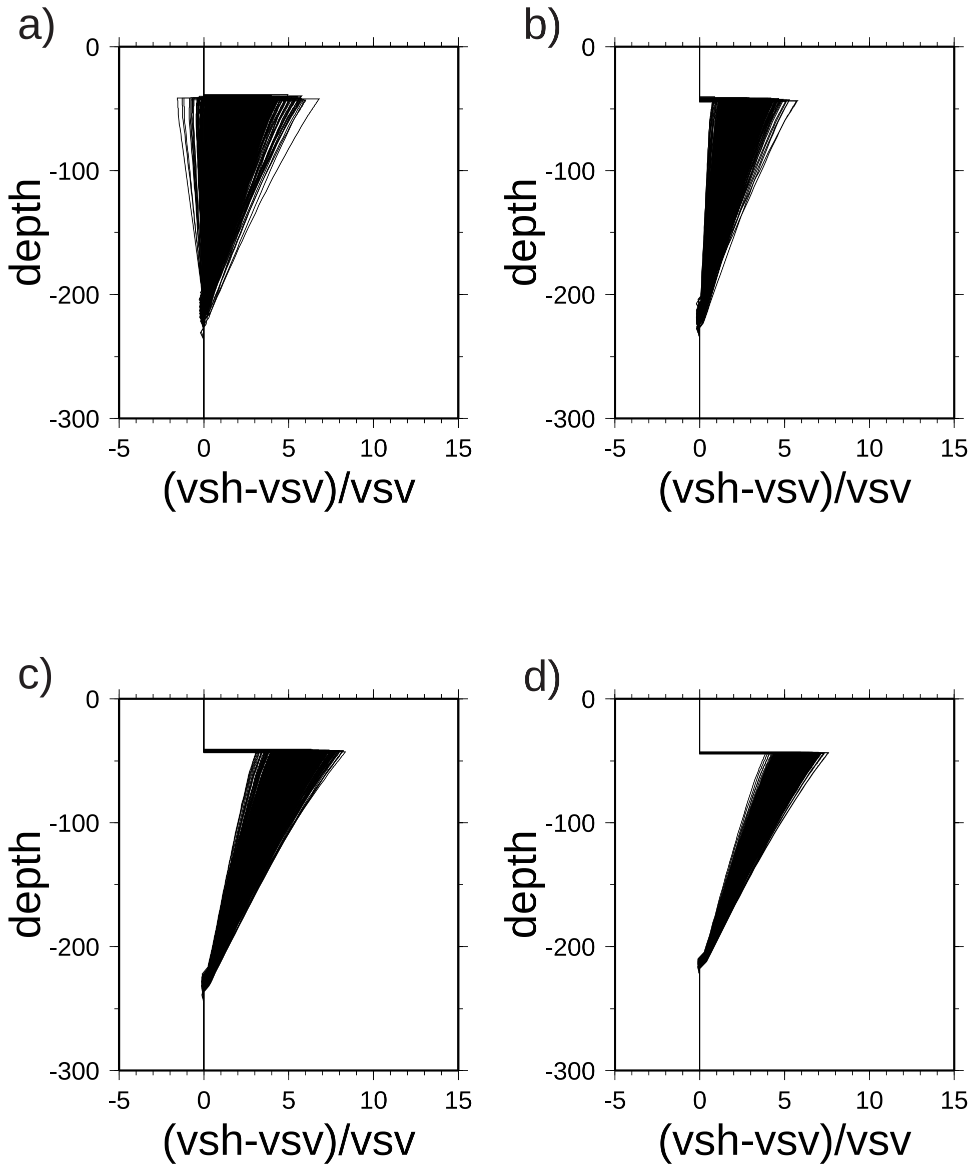


Figure 17. Strength of radial anisotropy beneath Eastern Antarctica (80 S, 90 E) obtained with different subsets of data: (a) phase velocities at periods greater than 70 s, (b) phase velocities between 40 s and 150 s, (c) group velocities between 16 s and 200 s, (d) the combined data-set.

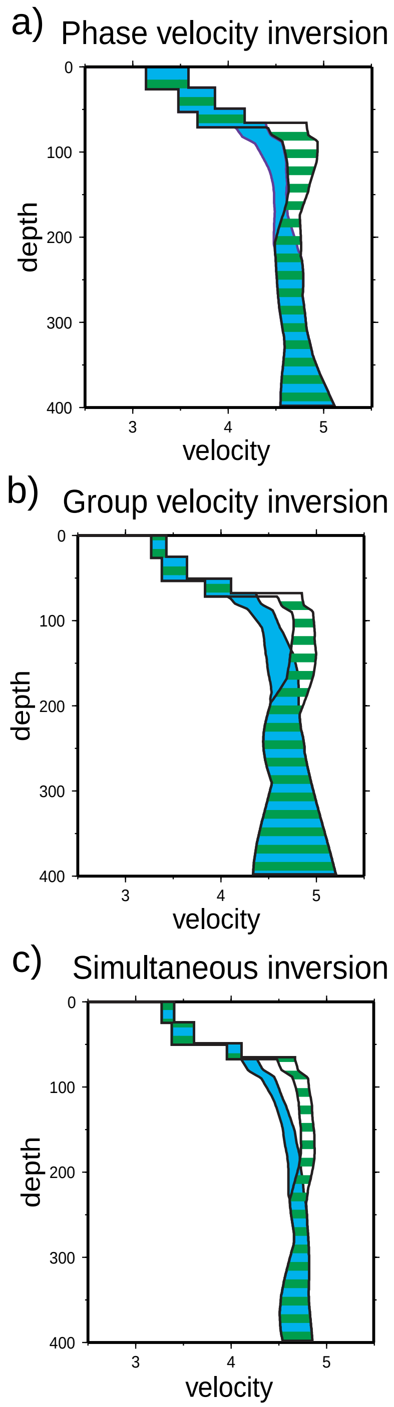


Figure 18. Examples of inversions at a point located in Tibet (34°N, 84°E) illustrating the relative importance of phase and group velocities: (a) inversion of the phase velocities only, (b) inversion of the group velocities only, (c) simultaneous inversion of the combined data-set. The corridor of acceptable v_{sv} velocities is plotted in gray and the v_{sh} velocities with horizontal hatching.