Geophys. J. Int. (2003) **153,** 88–102

The *S*-velocity structure of the East Asian mantle from inversion of shear and surface waveforms

Wolfgang Friederich

Institut für Geophysik der Universität Stuttgart Richard-Wagner-Str. 44 70184 *Stuttgart, Germany. E-mail: wolle@geophys.uni-stuttgart.de*

Accepted 2002 September 23. Received 2002 September 20; in original form 2002 April 19

SUMMARY

The structure of the mantle beneath East Asia down to 800 km depth is investigated using full waveforms of seismic shear and surface waves. Epicentral distances are limited to less than 40°. In contrast with previous waveform inversions, we avoid ray-theoretical or path-integral approaches. Instead, we use (1) exact 3-D waveform sensitivity kernels that correctly reflect off-path sensitivity and the existence of Fresnel zones. We apply (2) an accurate 3-D forward modelling technique based on a coupled-mode, multiple-forward-scattering approach, allowing us (3) to iterate the inversion procedure through several 3-D models and (4) to evaluate the true misfit between the data and the synthetics for the 3-D model. Average lateral resolution of the model in regions with good path coverage is 400 km throughout the upper mantle. In the depth range from 100 to 250 km the lateral resolution even approaches 200 km. Since wave front healing is taken into account, the amplitudes of velocity perturbations are larger than in other tomographic models. Moreover, the waveform sensitivity kernels provide an intrinsic physical smoothing of the model, stabilizing the inversion. Finally, owing to the use of exact sensitivities, better resolution can be achieved with a given set of seismograms.

Major features of the model are high-velocity subducting slabs along the West Pacific subduction zones stagnating at the 660 km discontinuity, a strong low-velocity zone between 80 and 250 km depth in the West Pacific backarc regions, a plume-like low-velocity feature beneath the southern tip of the Baikal rift zone extending into the transition zone, a low-velocity region under the Tien Shan with connection into the transition zone and thick crust under Tibet reaching its maximum thickness of approximately 80 km close to the 35th parallel. The lithospheric mantle underneath southern Tibet is very fast indicating underthrusting of the Indian lithosphere. In contrast, the upper mantle beneath northern Tibet exhibits average-to-slow values from the Moho down to the 660 km discontinuity. Very high *S* velocities are again observed beneath the Tarim Basin. The high-velocity mantle lithospheres under southern Tibet and the Tarim join beneath the Karakorum range at the western tip of the Tibet Plateau.

Key words: East Asia, mantle, mode coupling, scattering, seismic waves, tomography.

INTRODUCTION

Records of the Earth's seismic motion form the raw material used by seismologists to infer the internal structure of the Earth. With modern broad-band seismometers (Wielandt & Streckeisen 1982), all seismic signals relevant for structural interpretations can be measured. Since the broad-band seismic signal, however, is much too complex for a direct structural inference, it is subjected to further processing with the aim of extracting the information relevant for the intended study. This processing may involve filtering, picking of onset times of selected phases, extraction of time windows, determination of dispersion curves, stacking, beam forming and various other operations. Seismologists usually refer to the result of the processing as their 'data', and it is these 'data' from which the actual inferences on the Earth's structure are made. A basic prerequisite for doing this is a relation between the data and the earth model that in most cases allows one to uniquely predict the data from a given model but not vice versa. For this reason, it is usually called the forward modelling method. Depending on the data this relation may be exact or may contain severe approximations. Thus, inference of the Earth's structure from its seismic motion is unsatisfactory in three aspects: first, the data are not the full seismograms; secondly, the relation between data and Earth structure is only approximate; and thirdly, this relation is not invertible in a strict sense. While the third aspect is a principal property of the geophysical inverse problem, there is much room for improvement with regard to the first two points.

As early as 1926, Wiechert stated that it should be the aim of seismology to explain every wiggle in the seismogram. Following this utopia, many studies have been undertaken that go some way towards explaining the full seismic waveform. Pioneering work was done by Woodhouse & Dziewonski (1984) who inverted long-period waveforms for a global *S*-wave velocity model of the Earth's mantle. In 1990, Nolet presented the partitioned waveform inversion method (PWI) intended for regional-scale studies of mantle structure from long-period shear and surface waveforms. It has found many applications in the meantime (Zielhuis & Nolet 1994; Lebedev & Nolet 1997; Van der Lee & Nolet 1997a,b; Das & Nolet 1998). The peculiarity of the PWI is the division of the inversion process into two successive stages: first, a laterally homogeneous model for each source–receiver path is searched. Secondly, linear constraints are constructed using a diagonalization of the Hessian of the waveform misfit function that are transformed into a single 3-D model. The theoretical basis of the mentioned studies is a representation of the seismic displacement field by a superposition of normal modes that provides an exact solution for a spherically symmetric Earth model. The normal modes are characterized by a specific dependence of displacement on depth and by a specific frequency–wavenumber relation. The effect of lateral heterogeneity is treated in a path-integral approximation the basic assumptions of which are: (1) that the sensitivity of the seismic waveform is confined to those regions of the medium underlying the great-circle path connecting source and receiver; (2) that the effect of aspherical structure on the seismic waveform can be modelled by taking averages over the great-circle path; and (3) that the constituting modes remain independent even in the presence of asphericity.

While the first two assumptions of the path-integral approach can be shown to be valid if synthetic seismograms are calculated for Earth models with heterogeneities where the scalelength is much greater than the considered wavelengths, it was first demonstrated by Marquering & Snieder (1995) that the third assumption leads to incorrect sensitivity kernels if body waves are modelled. Sensitivity kernels are crucial in the inversion process because they determine where heterogeneities should be put in order to explain discrepancies between observed and synthetic waveforms. Incorrect sensitivity kernels may eventually lead to biased Earth models (Marquering *et al.* 1996). Moreover, with the growing number of worldwide deployed broad-band seismographs it has become possible to resolve heterogeneities with scalelengths that are comparable to or even smaller than the considered wavelengths. Path-integral approximations must be considered unsatisfactory to predict synthetic waveforms for such models.

Snieder (1986) was the first to go beyond the path-integral approximation by introducing a Born scattering theory for seismic surface waves. Born scattering takes into account waves scattered from off-path heterogeneities but neglects the repeated scattering of these waves. The theory also accounts for the coupling of normal modes by lateral heterogeneities and thus provides a means for computing exact sensitivity kernels. The theory of Snieder (1986) originally formulated for a half-space was later generalized to Born scattering on a laterally heterogeneous sphere (Snieder & Nolet 1987; Snieder & Romanowicz 1988). Unfortunately, Born scattering yields inaccurate results if the waves travel several wavelengths through the heterogeneous medium (Friederich *et al.* 1993). For this reason, structural inversions using Born scattering (Alsina *et al.* 1996) have been performed in two steps: a smooth 3-D model was obtained with a path-integral formulation and used as a reference model for a second inversion step using Born scattering theory. However, since the reference model is laterally heterogeneous, excitation and propagation of the scattered waves can only be computed within the path-integral formulation. A remarkable alternative to this approach is an extension of the standard PWI method proposed by Meier *et al.* (1997) in which Born scattering is used in the second stage of the PWI.

Combining the path-integral approach with Born scattering can be completely avoided if a multiple-scattering approach is used. The theory was first formulated by Friederich *et al.* (1993) for surface waves on a half-space and later extended (Friederich 1999) to shear and surface waves propagating in an aspherical Earth model. Since the theory takes into account mode coupling it is a highly accurate forward modelling method and at the same time allows the computation of waveform sensitivity kernels even with respect to a laterally heterogeneous initial model. Owing to this property, an inversion can be iterated through several laterally heterogeneous models, and the true misfit between synthetic and observed waveform can be evaluated.

In this paper, we apply the multiple-scattering theory of Friederich (1999) to infer the *S*-velocity structure of the East Asian mantle from vertical-component waveforms of seismic shear and surface waves. We start out with a brief review of multiple-scattering theory, continue with a presentation of the data and a detailed description of the inversion procedure, move on to resolution tests and end by displaying and discussing the East Asian mantle structure.

1 ESSENTIALS OF MULTIPLE-SCATTERING THEORY

The computation of synthetic seismograms and sensitivity kernels for an aspherical Earth model uses the multiple-scattering theory as formulated by Friederich (1999). The seismic displacement field in the frequency domain is split into a field for an appropriate, spherically symmetric reference model and a scattered field generated by deviations of the aspherical Earth model from the spherically symmetric one. The basic results of the theory are as follows. (1) The total displacement field can be constructed once the normal-mode problem of the reference model is solved for a fixed frequency. Each non-trivial solution is termed a mode that is characterized by its depth-dependent vertical eigenfunction, its phase velocity and quality factor. (2) The total displacement field can be expanded into a series of these depth-dependent, vertical eigenfunctions with expansion coefficients that are functions on the unit sphere. (3) The expansion coefficients can be determined by solving a system of coupled integral equations.

To be more specific, we write for the vertical component using an epicentral spherical coordinate system

$$u_z(r,\Delta,\phi) = \sum_n U_n(r)\Phi_n(\Delta,\phi), \tag{1}$$

where r, Δ and ϕ are, respectively, the radius, the epicentral distance and the longitude, and n indicates the mode number. $U_n(r)$ is a vertical eigenfunction of the nth mode computed for the spherically symmetric reference model and $\Phi_n(\Delta,\phi)$ is a potential defined on the unit sphere, here playing the role of an expansion coefficient. In a spherically symmetric Earth model the potentials take the form

$$\Phi_n^{\mathrm{ref}}(\Delta,\phi) = \frac{-i}{4c_n c_{gn}}\left(\frac{\Delta}{\sin\Delta}\right)^{1/2} \times \sum_{s=0}^{2}(-1)^s \nu_n^s H_s^{(2)}(\zeta_n\Delta)\,\mathrm{Re}\left[(\sqrt{2}\mathrm{e}^{i\phi})^s q_{ns}\right], \tag{2}$$

where c_n is the phase velocity, c_{gn} is the group velocity, ν_n is the angular wavenumber, $H_s^{(2)}$ is a Hankel function of the second kind and q_{ns} are excitation coefficients defined by the seismic source. Attenuation is taken into account via the argument of the Hankel function, where $\zeta_n = \nu_n - i\omega/(2Q_n c_{gn})$, with Q_n being the quality factor of mode n and ω the angular frequency. Eqs (1) and (2) represent the travelling-wave decomposition of the seismic displacement field, which is equivalent to the standing-wave representation commonly used in free oscillation theory. To make the expansion complete, an infinite number of terms with complex wavenumbers are required, which represent evanescent waves that are relevant in the near field (Pollitz 2001). They will be neglected in the present study because of the large scattering distances involved. Note also that in contrast with other work on surface wave scattering (e.g. Snieder 1986), the far-field approximation of the Hankel functions is not used in this work.

In an aspherical Earth model, the potentials obey a system of coupled integral equations that in the case of spheroidal–spheroidal coupling can be written as

$$\Phi_{n'}(\Delta_R, \phi_R) = \Phi_{n'}^{\mathrm{ref}}(\Delta_R, \phi_R) + \sum_n \int_\Omega d\Omega \sum_{N=0}^{2} \big[G_{n'}^{(N)}(\gamma)\, \delta\omega_{n'n}^{(N)}(\Delta, \phi) \times \mathrm{Re}\big\{[\sqrt{2}\mathcal{H}_+(\xi)]^N\big\} \Phi_n(\Delta, \phi)\big]. \quad (3)$$

This equation can be interpreted as follows: the potential of mode n' at the receiver (Δ_R, ϕ_R) is constructed from the value of the reference potential of mode n' at the receiver plus signal scattered from heterogeneities in Ω. The strength of the scattered signal is determined by the values of the modal potentials $\Phi_n(\Delta, \phi)$, the scattering angle ξ and the coupling functions $\delta\omega_{n'n}^{(N)}(\Delta, \phi)$ at the scatterers. The scattered signal is propagated over the distance γ between scatterer and receiver using the Green function

$$G_n^{(N)}(\gamma) = \frac{i}{4c_n c_{gn}} \left(\frac{\gamma}{\sin\gamma}\right)^{1/2} (-\nu_n)^N H_N^{(2)}(\zeta_n \gamma). \quad (4)$$

The coupling functions are determined from radial integrals of products of vertical eigenfunctions and perturbations of density and elastic moduli under the scatterer (see Appendix A of Friederich 1999). The operator $\mathcal{H}_+(\xi)$ contains first derivatives with respect to the angular coordinates (see Appendix D of Friederich 1999).

The system of coupled integral equations is solved numerically as described in detail in Friederich (1999). The number of modes used in the computation must be adjusted to the kind of signal to be modelled. To reduce the computational effort contributions from multiple backscattering are neglected. This is justified for media without vertical discontinuities or very strong lateral gradients. Since during the inversion roughness constraints are applied to the 3-D S-velocity model, synthetic seismograms predicted for the 3-D model are certainly not affected by this assumption.

A formula for the computation of sensitivity kernels can be derived from eq. (3): let us assume that we have obtained an aspherical Earth model from previous inversion steps with potentials Φ_n and coupling functions $\delta\omega_{n'n}^{(N)}$. The value of the S-velocity sensitivity kernel at some point in the medium is defined as the change of vertical displacement as a result of an infinitesimal change of S velocity at this point. Let $\epsilon_{n'n}(\Delta, \phi)$ be the coupling functions associated with this change in S velocity. Then the modal potentials change by

$$\begin{aligned}\delta\Phi_{n'}(\Delta_R, \phi_R) = \sum_n \int_\Omega d\Omega \sum_{N=0}^{2} \big[& G_{n'}^{(N)}(\gamma)\, \epsilon_{n'n}^{(N)}(\Delta, \phi) \\ & \times \mathrm{Re}\big\{[\sqrt{2}\mathcal{H}_+(\xi)]^N\big\} \Phi_n(\Delta, \phi) \\ & + G_{n'}^{(N)}(\gamma)\, \delta\omega_{n'n}^{(N)}(\Delta, \phi) \\ & \times \mathrm{Re}\big\{[\sqrt{2}\mathcal{H}_+(\xi)]^N\big\} \delta\Phi_n(\Delta, \phi)\big]. \end{aligned} \quad (5)$$

The first term describes scattered waves emanating from the perturbation of the aspherical model. However, they are propagated to the receiver with the Green function of the spherically symmetric reference medium. The second term accounts for the modification of the scattered waves owing to their propagation through the aspherical model. The first term can be evaluated directly because all quantities under the integral are known. Evaluation of the second term could be accomplished with the help of the reciprocity principle by additional forward computations with a source at the receiver. This is computationally too expensive and we therefore omit the second term when calculating sensitivities. The second term becomes relevant if the spatial average of $\delta\omega_{n'n}$ in the vicinity of the great circle is not zero. In this case, the signals in the sensitivity seismogram might be time-shifted and thus might lead the inversion in the wrong direction. By an appropriate choice of the reference model this problem can be avoided. The change of displacement is finally obtained by inserting the $\delta\Phi_n$ into eq. (1):

$$\begin{aligned}\delta u_z(r_R, \Delta_R, \Phi_R) &= \sum_{n'} U_{n'}(r_R)\, \delta\Phi_{n'}(\Delta_R, \Phi_R) \\ &\approx \sum_{n'} U_{n'}(r_R) \sum_n \int_\Omega d\Omega \sum_{N=0}^{2} \big[G_{n'}^{(N)}(\gamma)\, \epsilon_{n'n}^{(N)}(\Delta, \phi) \\ &\quad \times \mathrm{Re}\big\{[\sqrt{2}\mathcal{H}_+(\xi)]^N\big\} \Phi_n(\Delta, \phi)\big], \end{aligned} \quad (6)$$

where r_R denotes the radius of the receiver. Back transformation into the time domain yields the sensitivity seismograms.

2 DATA

The multiple-scattering approach is computationally much more expensive than Born scattering or path-integral methods. The numerical effort is mainly controlled by the number of modes used in the expansion in eq. (1), the epicentral distance and the Nyquist frequency. In the study presented here we use vertical-component seismograms filtered to a Nyquist frequency of 50 mHz. The maximum epicentral distance is 40° and the number of modes used to compute the synthetics is 15. This choice permits modelling of the complete waveform from the onset of the S wave to the end of the Rayleigh wave train. However, it confines the depth down to which structural information can be obtained to approximately 1000 km.

Waveforms have been collected from broad-band stations of the IRIS/IDA and GEOSCOPE networks. In addition, data from broadband stations deployed during the Tibetan Plateau temporary experiment in 1991/1992 have been included in the data set. All waveforms except the GEOSCOPE data have been obtained from the IRIS DMC. Fig. 1 shows the locations of the stations from which data have been included in the data set as black triangles. Owing to the large number of earthquakes occurring in East Asia it was not feasible to use all seismograms available even if the magnitude is restricted to 5 and greater. It is also certainly not wise to do so as earthquakes tend to form clusters with focuses in the West Pacific subduction zones and the Hindu Kush. To achieve a certain amount of declustering we attempted to select no more than one earthquake per $4 \times 4\ \mathrm{deg}^2$ cell at the surface and per 50 km depth interval. As a result, 138 earthquakes were chosen occurring during 1987–2001, for which approximately 1600 seismograms could be collected. In Fig. 1, the beachballs indicate the epicentres of the earthquakes and the white lines display the great-circle paths associated with the 1600 available waveforms. Path coverage is particularly good in a broad swath reaching from the West Pacific subduction zones to the

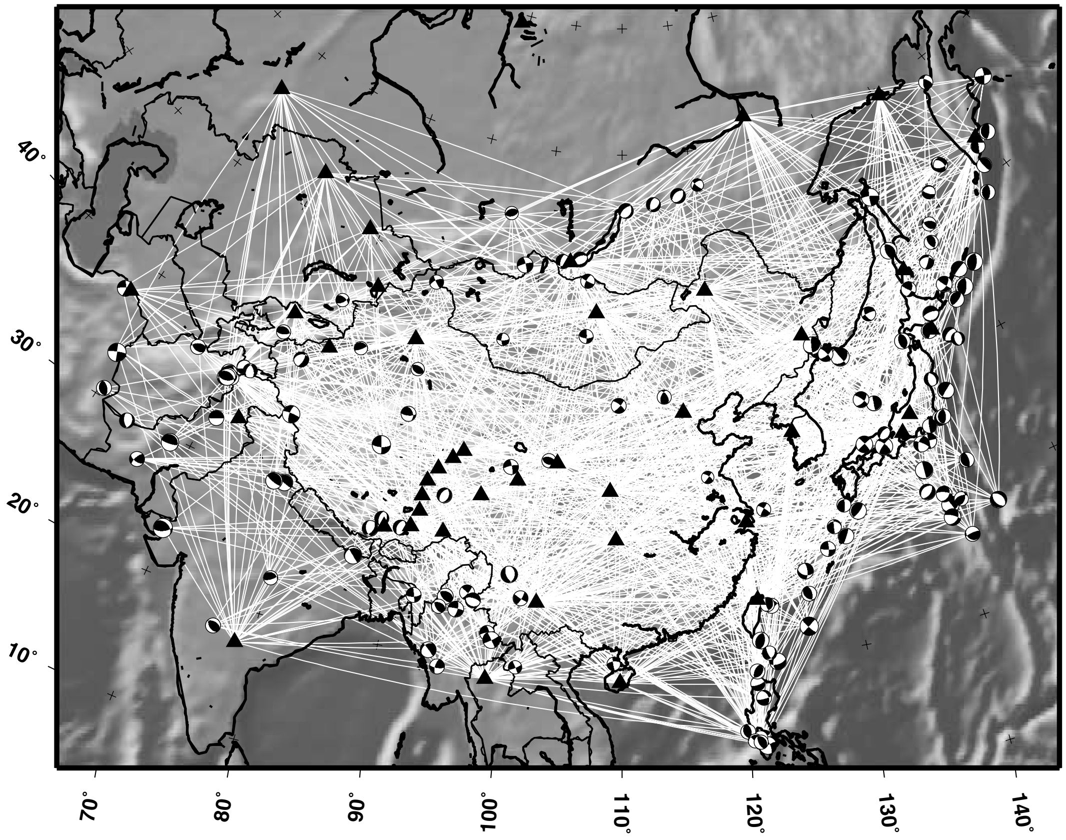

Figure 1. Station locations (black triangles), earthquake epicentres (beachballs) and all great-circle paths (white lines) for which a seismogram has been included into the data set.

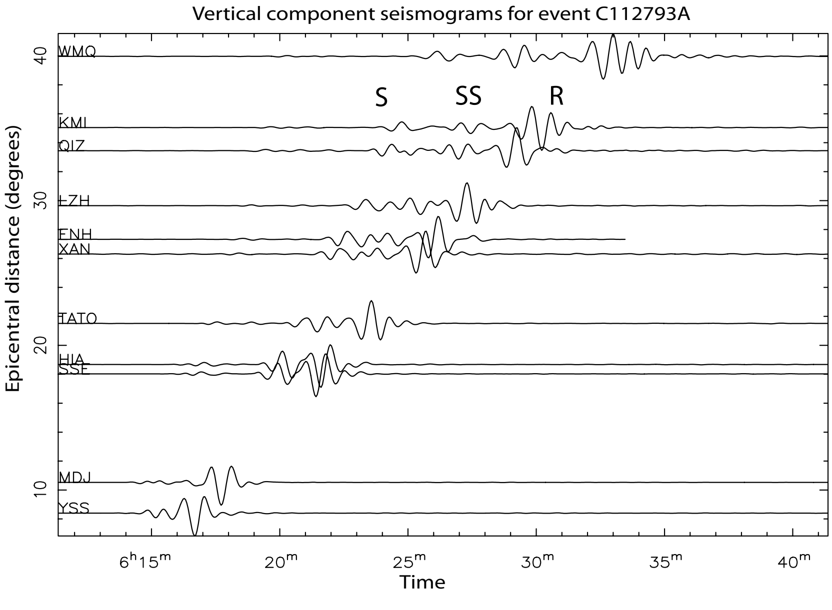

Figure 2. Vertical-component seismogram section for event C112793A near east coast of Honshu, depth 108 km, $M_b = 6.0$. Seismograms are low-pass filtered with corner frequency of 25 mHz. Major phases are the direct *S* wave (*S*), the surface-reflected *S* wave (*SS*) and the Rayleigh wave (*R*).

Hindu Kush with obvious deficiencies at the margins. In Fig. 2, we show a record section of typical low-pass filtered waveforms for an event near the east coast of Honshu at a depth of 110 km. Major phases in the seismograms are direct *S*, *SS* and the Rayleigh wave train.

3 INVERSION METHOD

In principle, implementation of a waveform inversion based on multiple scattering is quite straightforward: assume that the *S* velocity of the aspherical model is represented by parameters β_m and that we have already found an aspherical initial model M_i with S velocity $\beta_m(M_i)$. For this initial model, compute synthetic seismograms $s_p(t, M_i)$ for all paths p and, according to eq. (6) sensitivity seismograms $S_{pm}(t, M_i)$ for all paths p and parameters m. Then, with $d_p(t)$ denoting the observed seismograms, set up a system of equations of the form

$$d_p(t) - s_p(t, M_i) = \sum_m S_{pm}(t, M_i)[\beta_m - \beta_m(M_i)] \tag{7}$$

and solve this for β_m using a least-squares approach with some additional constraints on the parameters. Finally, iterate this procedure using the new 3-D model as the initial model.

There are several disadvantages with such a direct approach as we will outline below. The major difficulty is how to find the initial 3-D model. In most cases, such a model is not available *a priori*. We might be able to find an appropriate spherically symmetric (1-D) initial model, but for a large and tectonically active region such as East Asia, such a model would be very unsatisfactory. Large discrepancies between synthetic and observed waveforms would prohibit a linearized waveform inversion. The way out of this dilemma has been shown by Nolet (1990) with the partitioned waveform inversion and by Meier *et al.* (1997) who formulated an extension of the original PWI that uses correct sensitivity seismograms obtained from Born scattering theory: to start the inversion we do not need a single 3-D or 1-D reference model for the whole study region; it is sufficient to find 1-D reference models for each path separately! Ideally, these models should in some way represent the average structure along the path, and synthetic seismograms for these models will already fit the data fairly well. The very important additional benefit is that we can use these path-specific 1-D models as reference models for the scattering calculations. Instead of eq. (7) we set up a system of equations of the form

$$d_p(t) - s_p(t, M_p) = \sum_m S_{pm}(t, M_p)[\beta_m - \beta_m(M_p)], \tag{8}$$

where M_p is now the path-specific 1-D model and $s_p(t, M_p)$ and $S_{pm}(t, M_p)$ are synthetic seismograms and sensitivity seismograms, respectively, calculated for this model. Since the unknowns $\beta_m - \beta_m(M_p)$ depend on the path number, the above system cannot be solved as it stands. This deficiency can easily be mended by introducing a single 1-D reference model M_{ref} representing some average of the path-specific 1-D models. The above equation may then be rewritten as

$$d_p(t) - s_p(t, M_p) + \sum_m S_{pm}(t, M_p)[\beta_m(M_p) - \beta_m(M_{\text{ref}})]$$
$$= \sum_m S_{pm}(t, M_p)[\beta_m - \beta_m(M_{\text{ref}})]. \tag{9}$$

This equation is equivalent to eq. (41) of Meier *et al.* (1997). The vector of unknowns is now $\beta_m - \beta_m(M_{\text{ref}})$ and no longer path-dependent. All quantities on the left-hand side are known. Solving eq. (9) yields the first 3-D model from which further iterations can be started. Since the misfit between data and synthetics has already been greatly reduced when finding the path-specific 1-D models, the effect of solving eq. (9) is mainly to find a 3-D model that approaches the path-specific 1-D models in the vicinity of the paths where the dominant contributions to the sensitivity seismograms come from.

To perform further iterations, we use the resulting aspherical model as a new initial model. To make sure that the path-specific 1-D reference models used for performing the scattering calculation are appropriate averages of the new 3-D model over the vicinity of the path, we derive new path-specific 1-D models from the initial 3-D model. Then, a 3-D forward computation is conducted and synthetics and sensitivity seismograms for the new 3-D initial model are computed. The inversion is then continued according to eq. (9) after replacing there the models M_p by the new initial model M_i:

$$d_p(t) - s_p(t, M_i) + \sum_m S_{pm}(t, M_i)[\beta_m(M_i) - \beta_m(M_{\text{ref}})]$$
$$= \textstyle\sum_m S_{pm}(t, M_i)[\beta_m - \beta_m(M_{\text{ref}})]. \tag{10}$$

It should be noted that the synthetics $s_p(t, M_i)$ for the first 3-D model may fit the data considerably less well than the synthetics for the path-specific 1-D models do. For, while the path-specific 1-D models are optimized for each path separately, the 3-D model is a compromise of these many, partially contradictory models.

Both eqs (9) and (10) are solved subject to roughness constraints that are added to the system of equations to stabilize the inversion. The constraints are formulated as linear combinations of the model parameters:

$$0 = \sum_m C_{km}[\beta_m - \beta_m(M_{\text{ref}})], \tag{11}$$

where the index k runs through the number of constraints and C_{km} is a matrix by which the desired linear combinations are realized.

Another difficulty of any waveform inversion is the fact that the amplitude of a seismic signal does not necessarily reflect its information content. For example, for a shallow source, the surface waves dominate the seismogram but only contain information concerning relatively shallow mantle structure. Deeper mantle structure can be resolved with the body waves, which are much weaker signals. To avoid a dominance of the surface waves in the waveform inversion, we multiply observed, synthetic and sensitivity seismograms by the same time–dependent weight function that magnifies the signal at a given time instant to maximum amplitude if the envelope of the observed seismogram exceeds a predefined threshold. In addition, to avoid a dominance of large events in the waveform inversion, observed and synthetic seismograms are normalized to the maximum amplitude of the observed seismogram.

Finally, a numerical realization of eq. (10) with the constraints eq. (11) requires a time sampling of the observed, synthetic and sensitivity seismograms. A straightforward implementation would lead to one equation per time sample. However, what is the appropriate sampling rate? The sampling rate used to plot and store the seismograms would certainly lead to a very redundant set of equations. A better way to set up the system of equations for each path is to find independent linear combinations of the time samples that carry the major share of information concerning 3-D structure, a procedure proposed by Meier *et al.* (1997). This approach is very similar to the construction of linear constraints from the 1-D models obtained for each path in the standard PWI method with the difference that linear combinations of time samples are formed. Let N_{T} and N_{P} denote the number of time samples and number of parameters, respectively. Let us define a sensitivity matrix $\mathbf{S}_p^{im}$ obtained from the values of the sensitivity seismograms $S_{pm}(t)$ at the time samples t_i. By a singular-value decomposition of the sensitivity matrix,

$$\mathbf{S}_p = \mathbf{V}_p \Lambda_p \mathbf{U}_p^{\text{T}}, \tag{12}$$

we obtain an $N_{\text{T}} \times N_{\text{T}}$ matrix $\mathbf{V}_p$, the columns of which contain the coefficients of the desired independent linear combinations. The relative amount of information on 3-D structure of any linear combination is measured by the singular values associated with each column of $\mathbf{V}_p$. They appear as elements of the $N_{\text{T}} \times N_{\text{T}}$ diagonal matrix Λ_p. The rows of the $N_{\text{T}} \times N_{\text{P}}$ matrix $\mathbf{U}_p^{\text{T}}$ define the relation between the linear combinations of the time samples and the model parameters. If we define the function

$$z_p(t, M_i) = \sum_m S_{pm}(t, M_i)[\beta_m(M_i) - \beta_m(M_{\text{ref}})], \tag{13}$$

and construct vectors $\mathbf{d}_p$, $\mathbf{s}_p$ and $\mathbf{z}_p$ from the time samples of $d_p(t)$, $s_p(t)$ and $z_p(t)$, respectively, and also define a parameter vector $\boldsymbol{\beta}$, eq. (10) can be written as

$$\mathbf{d}_p - \mathbf{s}_p + \mathbf{z}_p = \mathbf{V}_p \Lambda_p \mathbf{U}_p^{\mathrm{T}} \boldsymbol{\beta}. \tag{14}$$

The desired independent linear combinations of the time samples are obtained by multiplication with $\mathbf{V}_p^{\mathrm{T}}$ on both sides. Using the orthogonality of the columns of $\mathbf{V}_p$, this leads to

$$\mathbf{V}_p^{\mathrm{T}}(\mathbf{d}_p - \mathbf{s}_p + \mathbf{z}_p) = \Lambda_p \mathbf{U}_p^{\mathrm{T}} \boldsymbol{\beta}. \tag{15}$$

This system still contains N_{T} equations but can easily be reduced to a much smaller size if all equations associated with singular values below some threshold are eliminated. The final system of equation that is solved for 3-D structure is built from the equations obtained for each path.

4 INVERSION DETAILS

4.1 Preliminary inversion for path-specific 1-D models

As outlined in the previous section, a prerequisite of a 3-D inversion is to find a 1-D model for each path. The inversion for such a model basically follows eq. (11) with the difference that now the initial and final model are 1-D, and that the sensitivities are calculated for 1-D perturbations of the initial model. For this reason, it is hampered by the same problems as the 3-D inversion: how to find an initial model that permits an iterative waveform inversion? Clearly, we need different initial models for paths running through different tectonic regimes. Since tectonics expresses itself in heterogeneities down to approximately 200–400 km it is mainly the surface waves that show large variations of phase and also amplitude. The two most important parameters that control the phase and the amplitude of surface waves are the crustal thickness and the average *S* velocity in the uppermost 200 km of the mantle. We exploit this fact and perform a grid search for these two parameters. Values of crustal thickness from 5 to 80 km in steps of 5 km and values of *S* velocity from 4.1 to 4.9 km s^{-1} in steps of 0.05 km s^{-1} have been searched for the best fit to the observed waveform. With the exception of events deeper than approximately 150 km, reasonable values of crustal thickness and uppermost mantle *S* velocity could be obtained for nearly all waveforms of the data set. In Fig. 3, we show a typical result of the grid search.

Another major problem of any waveform inversion is the accuracy of hypocentre locations and moment tensors. We use the improved hypocentre locations published by Engdahl *et al.* (1998) and moment tensors from the Harvard catalogue (Dziewonski *et al.* 1981). For some events, the Harvard moment tensors predict amplitudes that are consistently too large or too small. In many cases, however, a slight modification of the moment tensor leads to a substantial reduction of misfit between synthetic and observed waveforms. For this reason, we adjust the moment tensors after the grid search and before inverting for the path-specific 1-D models. This is done by inverting all waveforms of an event for slight changes of the six moment tensor elements under the constraint that they stay very close to the Harvard moment tensor solution.

With the improved moment tensors and the initial models obtained from the grid search, an iterative waveform inversion then leads to the path-specific 1-D models. In Fig. 4, we show the *S* velocity of all the resulting models in a grey-shade coded histogram. The variations of *S* velocity are greatest in the upper 250 km of the mantle with a steady decrease as the depth increases. The average 1-D model is very close to model ak135 (Kennett *et al.* 1995). To quantify the discrepancy between observed and synthetic seismograms, for each path *p* we compute a normalized misfit defined as the ratio of the waveform misfit over the energy of the data,

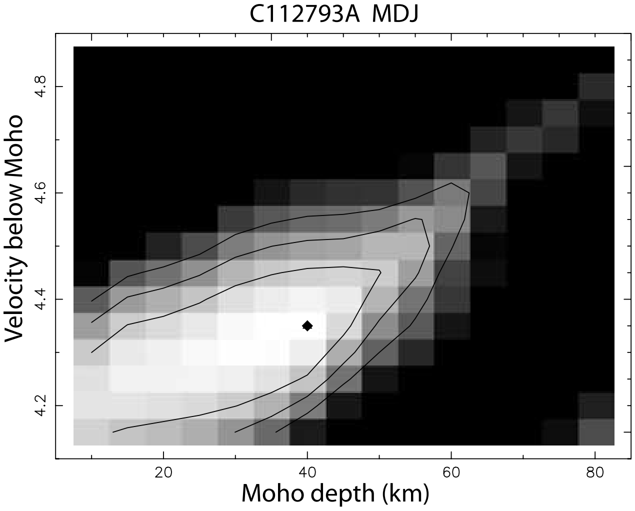

Figure 3. Typical result of the gridsearch for Moho depth and subMoho *S* velocity, here for event C112793A and station MDJ. The black diamond marks the pair of values for which the misfit between data and synthetics is minimum.

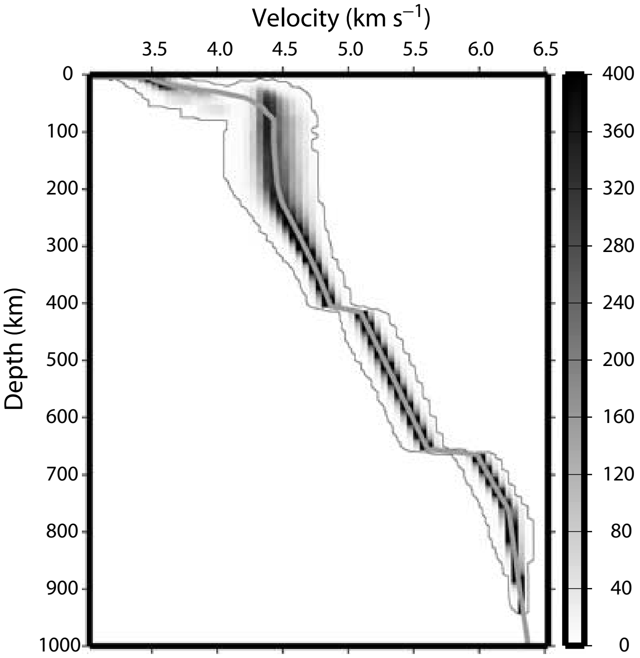

Figure 4. Grey-shaded histogram of *S*-velocity distribution as a function of depth for the 1600 1-D models. The shading indicates how often a certain velocity value occurs at a certain depth. For the region inside the single isoline at least one 1-D model exists with corresponding values for depth and velocity. The thick grey line represents the average model that is later used as a whole-area 1-D reference model.

$$\chi^2_N(M_p) = \frac{\left\| \mathbf{V}_p^{\mathrm{T}}(\mathbf{d}_p - \mathbf{s}_p) \right\|^2}{\left\| \mathbf{V}_p^{\mathrm{T}} \mathbf{d}_p \right\|^2}. \tag{16}$$

The misfit over energy for all paths is 1.65 for the whole-area reference model M_{ref} and 0.49 for the path-specific 1-D models M_p.

4.2 3-D inversion

The only quantity we invert for is the shear wave velocity β. The P-wave velocity α and density ρ are bound to the S velocity according to the correlation factors (Masters *et al.* 2000)

$$\frac{d\ln\alpha}{d\ln\beta} = 0.6 \quad \text{and} \quad \frac{d\ln\rho}{d\ln\beta} = 0.4. \tag{17}$$

The shear wave velocity in the model is specified at nodes with vertical distances ranging from 5 km in the upper 80 km of the model to 30 km down to 720 km depth to 60 km beyond this depth. Horizontally, the nodes are defined as centres of cells with edges parallel to circles of constant longitude and constant latitude. The northeast extension of a cell is fixed to 200 km, while the east–west extension is set to 200 km at 40° north and varies approximately with this value according to latitude. For this study, we use 1062 cells and 44 depth nodes amounting to a total of 46 728 parameters. The depth of the Moho is not a separate parameter in the inversion. The inversion is free to place a gradient zone of at least 5 km width anywhere in the upper 80 km of the model. For this reason, the 3-D model will not have a clearly defined Moho in accord with the fact that a sharp boundary cannot be resolved with the waveforms entering the inversion.

After singular-value decomposition of the sensitivity matrix and reduction of the data vector for each path to linear combinations of time samples associated with singular values greater than 5 per cent of the largest singular value, we obtain approximately 36 500 constraints from the 1600 waveforms in the data set. Theoretically, the waveforms are sensitive to structural perturbations over the whole study area. In practice, however, we may exploit the fact that the sensitivity is confined to a broad region surrounding the great circle connecting the source and the receiver. To simplify the construction of this sensitivity region, we define it as the interior of an ellipse with foci at the source and the receiver. The eccentricity of the ellipse is defined by the condition that the focal rays to the point on the ellipse halfway between source and receiver should intersect at a right angle. Restriction of the sensitivity to this region reduces the number of parameters constrained by a given waveform to approximately a quarter. Nevertheless, the number of non-vanishing matrix elements of the resulting system of equations still amounts to approximately 400 × 10^6. In Fig. 5, we show a plot of the sensitivity matrix $\mathbf{S}_p^{im}$ for parameter nodes m lying at 571 km depth, a receiver at 34° epicentral distance and a time instant t_i within the direct S wave that demonstrates the typical form of the sensitivity ellipse. The bright-spots close to the great circle mark the intersection points of the S-ray path with the cross-section plane. Note also the asymmetric form of the sensitivity kernel caused by the radiation pattern of the source (for more on sensitivity kernels see Friederich 1999; Zhao & Jordan 1998; Marquering *et al.* 1996; Li & Tanimoto 1993).

Since the inverse problem is underdetermined in both a formal and a mathematical sense, we add roughness constraints to the system of equations as formulated in eq. (11). We have implemented a Laplacian smoothing that penalizes the difference between the S velocity at a node and the average of the S velocity taken over the neighbouring nodes. Different weights for horizontal and vertical smoothing are allowed. They are expressed as reciprocal standard deviations σ_H and σ_V. Moreover, since the distance between nodes varies in the vertical direction, σ_V is normalized to a depth interval of 30 km.

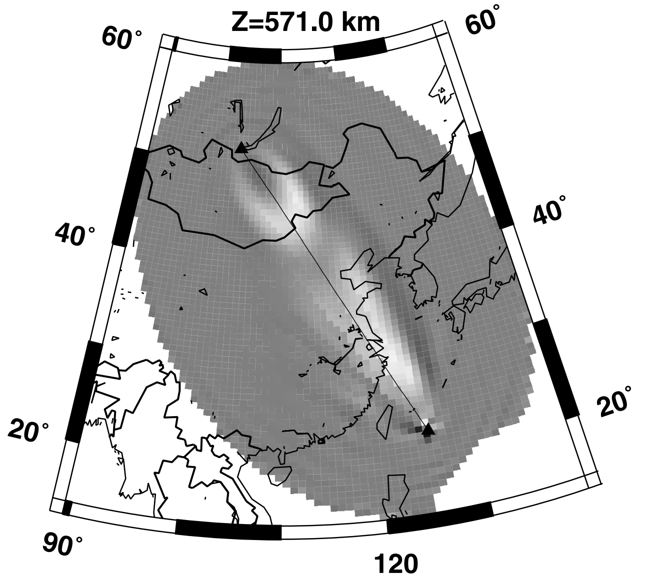

Figure 5. Horizontal cross-section through the sensitivity matrix at 571 km depth for an instant in the time window of the direct S wave. Scattered signal from bright areas contribute with a negative sign to the reference seismogram, the signal from dark areas contributes with a positive sign at the selected time instant.

Because of the large number of equations and parameters, the linear system cannot be solved with classical methods. We use a parallelized version of an iterative algorithm based on the conjugate gradient method termed CGLS1 by Bjoerck *et al.* (1998) who show that this algorithm is of equal efficiency and stability as the better known LSQR method (Paige & Saunders 1982). In general, approximately 100 iterations are sufficient to achieve a convergence in misfit reduction.

The final model of East Asia we will show and discuss below is the fourth of a sequence of 3-D models. The first one is obtained by solving eq. (9) with a whole-area 1-D reference model M_{ref} constructed by averaging over all path-specific 1-D models. It is displayed in Fig. 4 as the thick grey line. In this first step from the 1-D path models to a 3-D model, we choose to apply a rather rigid smoothing: for horizontal smoothing we set $\sigma_H = 0.05$ km s^{-1} at all depth levels and for vertical smoothing we used the same value with the exception of the upper 80 km where we set $\sigma_V = 0.5$ km s^{-1} to allow for an appropriate migration of the Moho from the 1-D path models into the 3-D model. We apply this strong smoothing because we expect the true 3-D model to exhibit large deviations from the 1-D path models. Since, in this first step, the sensitivity seismograms are computed for the 1-D path models, they might not give a very accurate estimate of the change in waveform involved with these deviations. Our experience shows that an unconstrained inversion in this first step leads to models with unreasonably large perturbations that are not required to fit the data.

The smoothing and, as mentioned before, the compromise between partially conflicting 1-D path models take their toll: a multiple-scattering forward computation with the new 3-D model shows that the normalized misfit for all paths has grown from 0.49

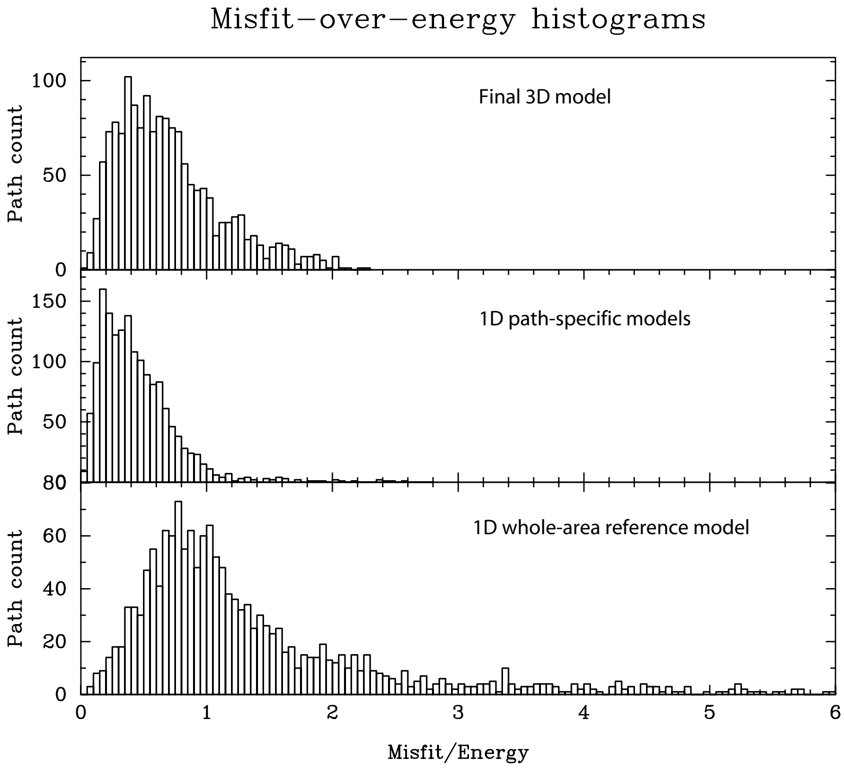

Figure 6. Histograms of misfit versus energy for all seismograms. Top, final 3-D model; centre, 1-D path-specific models; bottom, 1-D whole-area reference model.

to 0.84. However, the new 3-D model gives us the opportunity to update the sensitivity seismograms. A subsequent, second inversion with the same smoothing parameters results in a model with a much more modest level of heterogeneity in the upper 150 km than the first one, but, as a forward computation shows, an even smaller misfit of 0.80. This result confirms our strategy of not jumping from the 1-D path models to a high-resolution 3-D model.

This is done in the third inversion and is justified in the subsection concerning resolution. We increase σ_H to values of 0.10 km s^{-1} for $0 < z < 80$, to 0.14 km s^{-1} for $80 < z < 195$, to 0.21 km s^{-1} for $195 < z < 445$ and to 0.28 km s^{-1} for greater depths. For σ_V, we use values of 0.5, 0.2, 0.05 and 0.05 km s^{-1} in the same depth intervals. The normalized misfit is reduced to 0.726. A further inversion with the same smoothing parameters produces the 3-D model shown in this paper. The normalized misfit of 0.722 is negligibly smaller than for the third model and indicates a convergence of the inversion process if the smoothing parameters are kept fixed. Fig. 6 shows histograms of the misfit distribution for different models. Since these distributions have fairly long tails, it might be more appropriate to use the median instead of the mean to quantify the overall misfit. The numbers look more favourable then: for the whole-area 1-D reference model the median is 1.1, for the path-specific 1-D models it is 0.4 and for the final 3-D model it is 0.62.

5 RESOLUTION

Resolution is usually quantified in terms of the resolution length that is defined as the minimum spatial distance by which two anomalies of the same sign may be separated to be distinguished in the inverted model. Resolution depends on the amount of independent information concerning the model that is provided by the data and on the relative importance of the roughness constraints. Resolution is diminished by any kind of errors if we are searching for the least detailed model that explains the data within the error bounds. Some errors may be easily quantified, such as the amount of noise in a seismogram, while others, such as location errors and errors in the moment tensors, are difficult to assess. There are also errors in the synthetic seismograms caused by approximations when solving the forward problem. The model may also be biased owing to approximate sensitivity seismograms. By resolution tests, we can only estimate the information content of the data and the effect of the roughness constraints. This is usually done by generating a synthetic data set for a given structure and inverting it with exactly the same roughness constraints as were used when inverting the real data set.

Various test structures have been proposed in the literature, most notably chequerboard tests and spike tests. Spike tests serve for demonstrating the significance of a well-localized anomaly and for estimating the resolution length that is equal to the half-width of the image of the spike. Since the resolution length may vary for different parts of the model space, spike tests should be performed at many locations in the model, a procedure that might be too expensive in waveform inversions. Here we follow Bijwaard *et al.* (1998) and compute synthetic data for test structures composed of box-like anomalies of alternating sign horizontally separated from each other by zero structure areas. The boxes have a depth extent of 90 km in the mantle and 15 km in the crust. They are placed at various depths to study the variation of resolution with depth. In Fig. 7, we show the original and retrieved structures for box sizes of 8 × 8, 4 × 4 and 2 × 2 deg^2 spaced from each other by 8°, 4° and 2°, respectively. The 8 × 8 and 4 × 4 deg^2 boxes are reasonably well retrieved at all depths down to 720 km in the core region from the West Pacific to the Hindu Kush where the path density is highest. There are obvious

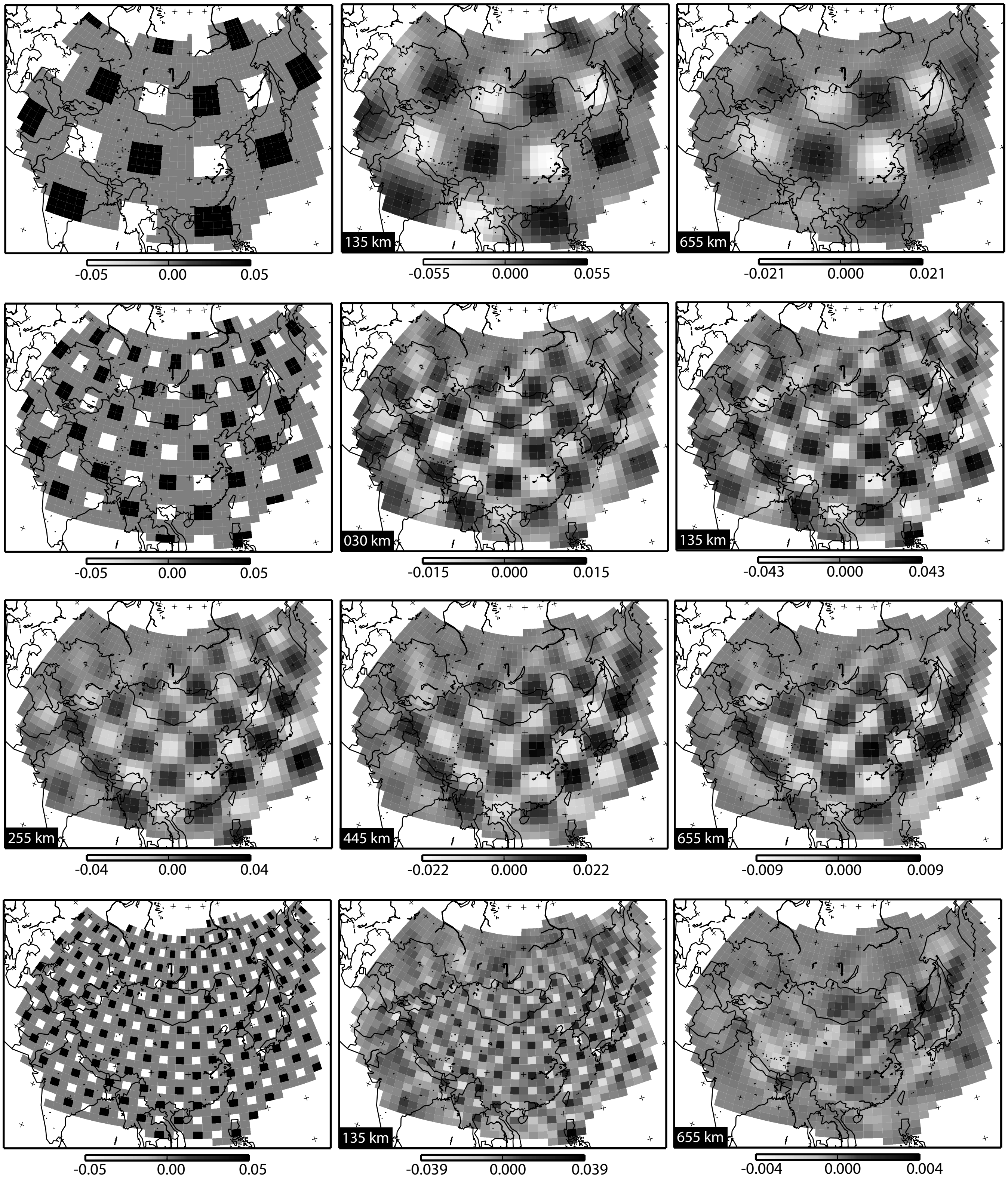

Figure 7. Resolution test with three different test structures. Test structures are emplaced at selected depths and the reconstructions are shown at the depths of emplacement. Top row, leftmost panel shows test structure consisting of 8 × 8 deg^2 by 90 km boxes, the right two panels show reconstructions of test structure at 135 and 655 km depth. Second and third rows, leftmost panel in the first row shows test structure consisting of 4 × 4 deg^2 by 90 km boxes, and the following five panels show the reconstructions at 30, 135, 255, 445 and 655 km depth (at 30 km depth we use 4 × 4 deg^2 by 15 km boxes). Bottom row, leftmost panel shows test structure with 2 × 2 deg^2 by 90 km boxes and the panels to the right show the reconstructions at 135 and 655 km depth.

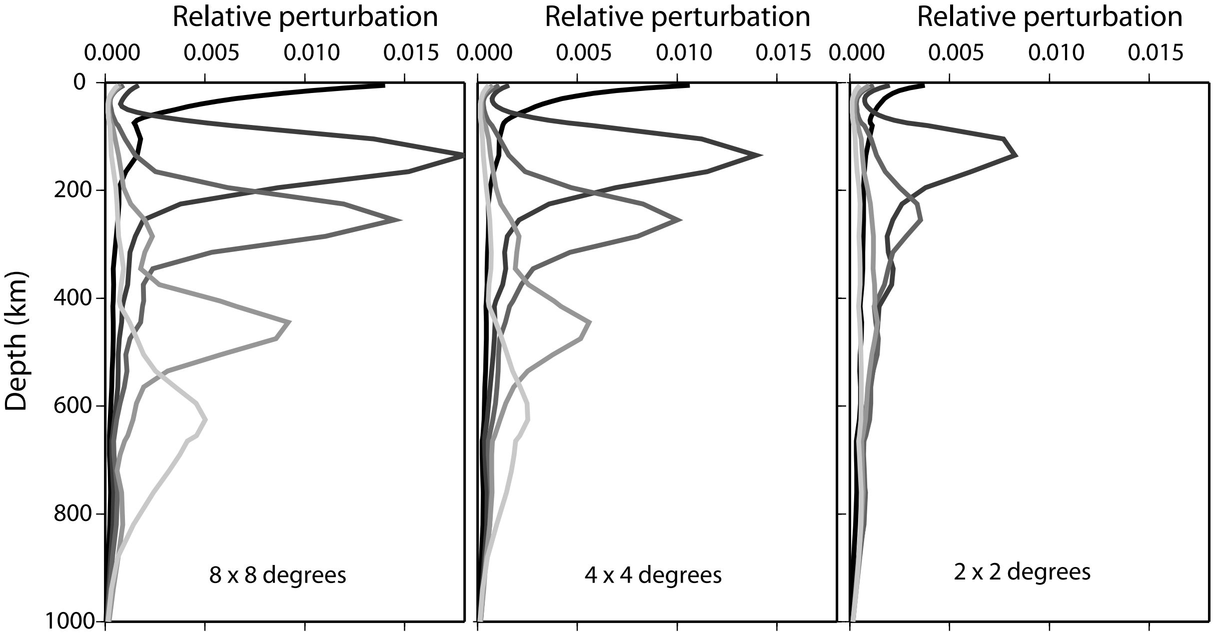

Figure 8. Vertical resolution: test structures are emplaced at selected depths (30, 135, 255, 445 and 655 km) and horizontal rms perturbations are computed for the reconstructed structures as a function of depth. The corresponding rms-perturbation curves are plotted with increasing brightness beginning with black for 30 km depth. Left-hand panel, rms perturbations for test structures with 8 × 8 deg^2 boxes; middle panel, same for test structures with 4 × 4 deg^2 boxes; right-hand panell same for test structures with 2 × 2 deg^2 boxes.

resolution deficits at the margins caused by either a lack of stations or a lack of earthquakes. At the shallow depths above 250 km the input anomaly of 5 per cent is nearly completely recovered, while at greater depth smearing reduces the recovered anomaly. For the 2 × 2 deg^2 pattern, we only obtain a satisfactory reconstruction down to 255 km depth. Below, the original pattern can no longer be recognized and the recovered anomaly falls to 10 per cent of the original one. Fig. 8 shows rms perturbations as a function of depth for the reconstructed models. The vertical resolution decreases with the size of the anomaly and vanishes below 255 km for the 2 × 2 deg^2 boxes. For the 4 × 4 deg^2 boxes, the input anomaly of 90 km depth extent is very well recovered with only small vertical smearing above 255 km depth. At 655 km depth, the input anomaly is spread to approximately twice its original width.

We conclude that the horizontal resolution is approximately 200 km at the best resolved depth of 135 km and increases to approximately 400 km at 655 km depth. The vertical resolution is of the order of 50–100 km.

6 A 3-D MANTLE MODEL OF EAST ASIA

In the following, we show both horizontal and vertical cross-sections through the 3-D mantle model of East Asia. In all sections, we display relative perturbations of the *S*-wave velocity. As a reference model, we use the average of all 1-D path models shown in Fig. 4. In addition, earthquakes occurring close to the cross-section plane are depicted as circles. The geographical locations of the vertical sections are displayed in Fig. 11, below.

The major features of the model are as follows.

(1) A difference in crustal thickness of oceanic and continental regions as becomes evident from Fig. 9 (30 km), where a horizontal section at 30 km depth is shown. At that depth we observe mid-crustal velocity values beneath the Tibet Plateau, the Tien Shan and the Altai ranges, while we generally have lower-crustal values in the surrounding continental regions. The oceanic regions in the western Pacific already exhibit mantle velocities. In spite of horizontal smearing, island arcs such as Japan are recognized as narrow zones of lower velocities compared with the surrounding oceanic areas (Pacific Ocean and Japan Sea).

(2) Very thick crust beneath the Tibet Plateau with Moho depths of up to 80 km in the central part of the plateau extending from 80° to 100° longitude (see Figs 9 , 80 km and 10A, B). Maximum thickness is attained along the 33rd parallel, with decreasing thickness both to the north and south.

(3) Very fast, gently northward-dipping Indian lithosphere under southern Tibet, and also fast, weakly southward-dipping Asian lithosphere under the northern margin of the plateau down to 250 km. In between is a zone of approximately 500 km width and approximately 800 km east–west extension where velocity perturbations are zero or slightly negative (see Figs 9, 135 km and 195 km and 10A, B). The velocity anomalies become more negative at greater depths (see Figs 9, 255–655 km and 10A, B). There is also fast Indian lithosphere extending way beyond the eastern rim of the Tibet plateau (Fig. 9, 135 and 195 km), which might indicate lithospheric extrusion to the east. The Asian lithosphere is particularly fast and thick beneath the Tarim basin (Figs 9, 135 and 195 km and 10A). A more detailed discussion and interpretation of Tibetan mantle structure is given elsewhere.

(4) Low *S* velocities in the backarc regions of the western Pacific subduction zones that reach from the South China Sea in the south to the Okhotsk Sea in the north and from the island arcs of Japan, Kuriles and Ryukyu in the east to the Asian continental margin in the west. These zones are relatively shallow and attain their lowest velocities at 135 km depth (see Figs 9, 80 and 135 km). At the latitude of central Japan they display some internal structure with

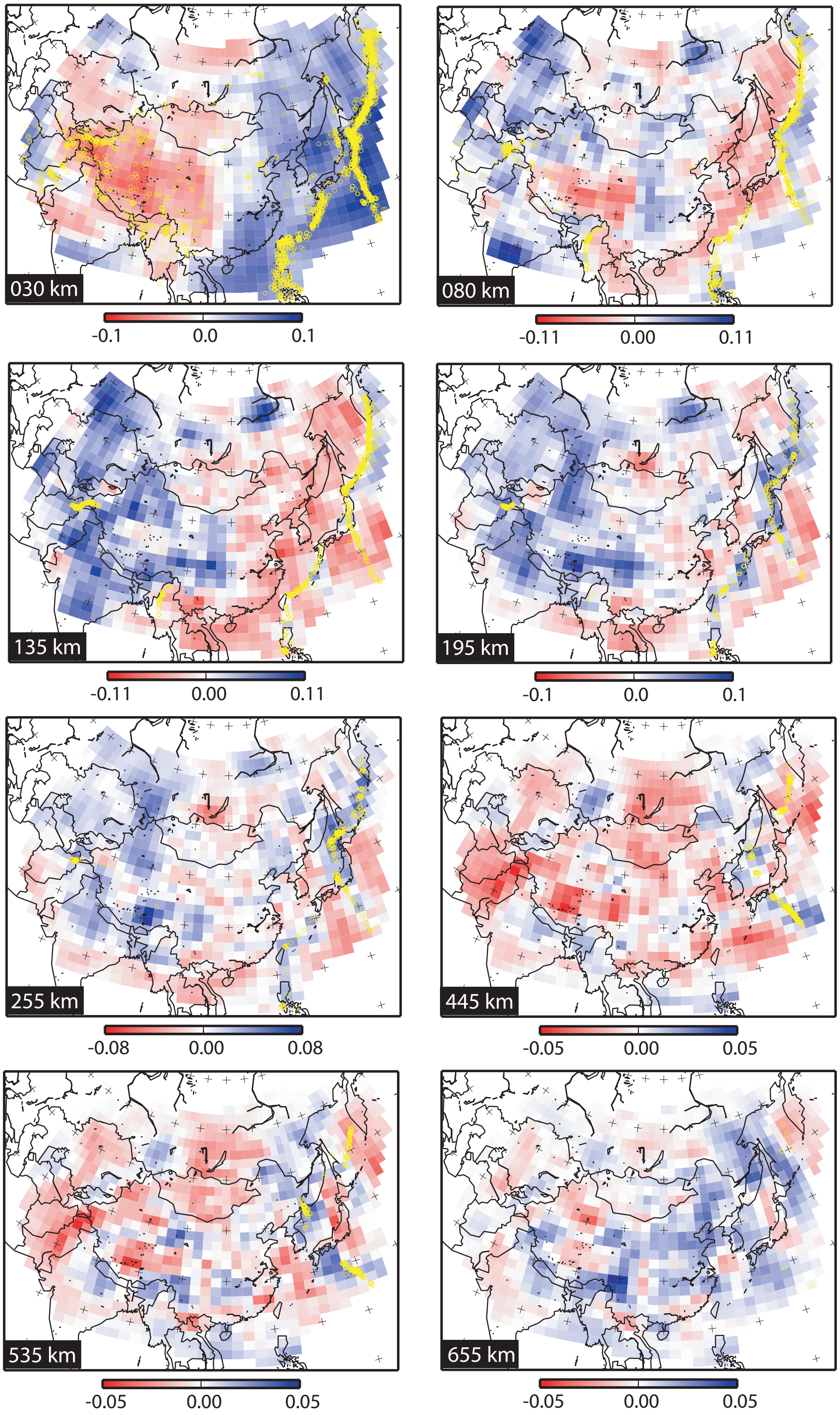

Figure 9. Horizontal cross-sections through 3-D model at depths indicated at the bottom left-hand corner of each panel. Yellow circles represent earthquake hypocentres. Shades represent relative perturbation of *S*-wave velocity with respect to the whole-area 1-D reference model.

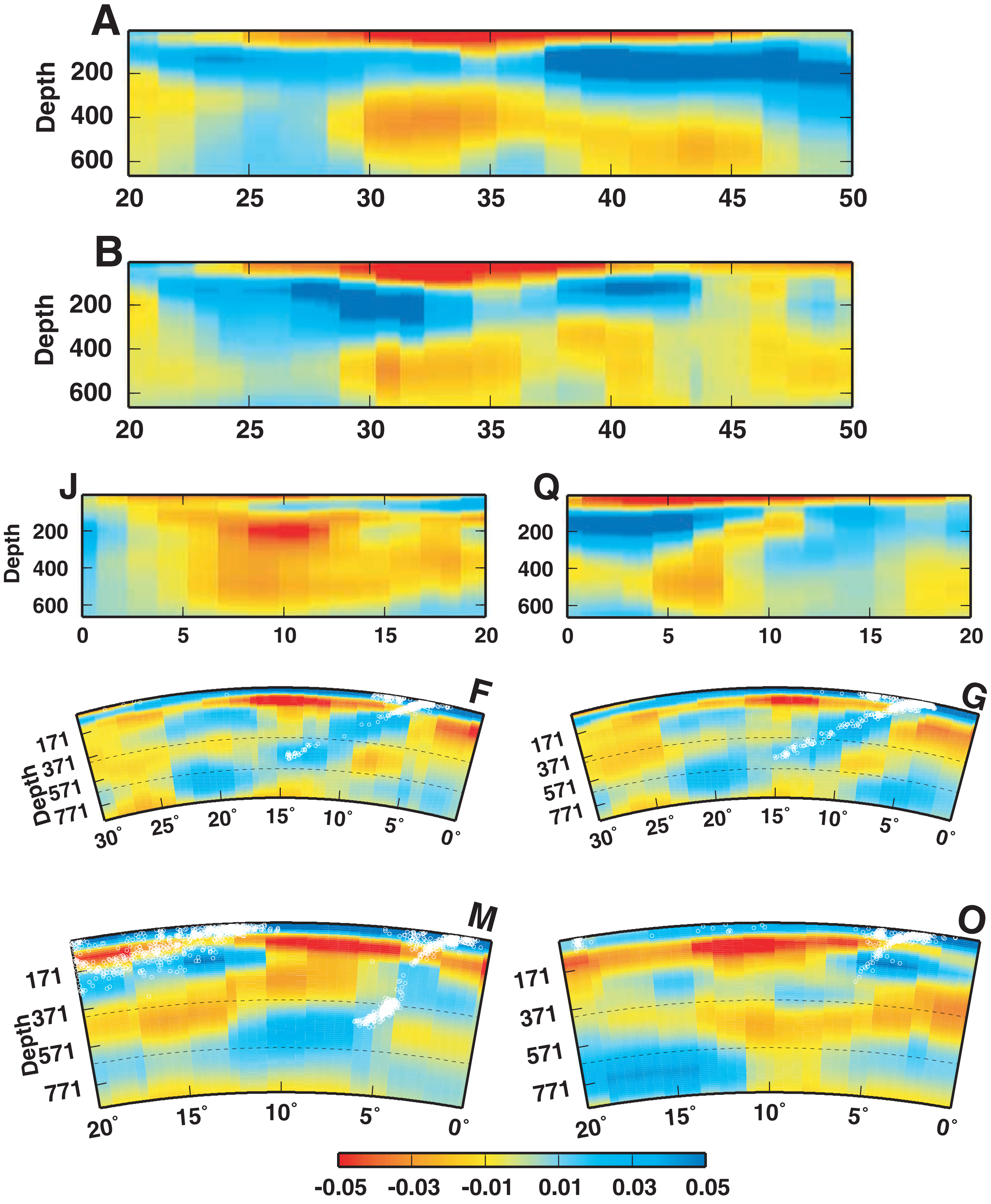

Figure 10. Vertical cross-sections through 3-D model along great-circle segments shown in Fig. 11. White circles represent earthquake hypocentres. Shades represent relative perturbation of *S*-wave velocity. In the sections A and B the annotation of the abscissae gives the geographical latitude. For the other sections, the annotation gives the distance along the great circle in degrees. Note that in most sections the colour scale is saturated in the crust and the top parts of the mantle.

the lowest velocities just west of the Japan island arc and under North Korea (Fig. 9, 80 km).

(5) A sharp transition from the backarc low-velocity zones to narrow, elongated high-velocity anomalies that represent the subducting plate (Figs 9, 80 and 135 km). Notably, there is no fast anomaly east of the southern part of Japan (Kyushu) at 80 and 135 km depth. This is consistent with an apparent lack of earthquake activity in this region at these depths and appears to be related to the low-velocity zone above the Izu–Bonin slab. In a certain sense, southern Japan is a backarc region with respect to subduction along the Izu–Bonin arc.

(6) The Japan slab (Figs 10F, G) and the Izu–Bonin slab (Fig. 10M) stagnate around the 660 km discontinuity and form bulky, horizontally elongated high-velocity anomalies that extend further away from the trench. For the Japan slab, these anomalies are well beneath the Asian continent (Fig. 9, 255–655 km). The location of the high-velocity anomalies representing the slabs clearly correlates with earthquake hypocentres. At shallow depths (80–195 km) the hypocentres lie on top of the slab while they move into the high-velocity anomalies for greater depths. The Ryukyu slab (Fig. 10O) does not penetrate the 410 km discontinuity. The Kuril slab is imaged down to 400 km but is lost at greater depths owing to lack of resolution.

(7) A well-confined, nearly cylindrical low-velocity anomaly beneath the southern tip of the Baikal rift (Figs 9, 135–655 km) reaching down into the transition zone. Remarkably, the low-velocity anomaly does not extend to the surface right below Lake Baikal but there appears to be a connection to the surface south-west of the Baikal rift in Northern Mongolia (Fig. 10J). This is consistent with observations of high heat flow and Cenozoic magmatism in this region (Barry & Kent 1998).

(8) Low velocities beneath the Tien Shan down to 200 km (Figs 9, 30–195 km). The anomaly extends into the transition zone in a southerly direction. At 500 km depth it is located beneath the Pamirs (Fig. 10Q).

7 COMPARISON WITH PREVIOUS WORK

The finding of fast velocities beneath southern Tibet and average-to-slow velocities under northern Tibet is consistent with results of other authors: Ni & Barazangi (1983) studied high-frequency S_n waves propagating across the Indian Shield, the Himalayas and Tibet and found efficient propagation of S_n beneath southern Tibet but inefficient propagation under northern Tibet. Wittlinger *et al.* (1996) conducted a teleseismic *P*-wave tomography across the Jinsha Suture and the Kunlun Fault (see Fig. 11) and detected a zone of low velocities compared with the surrounding mantle beneath an area reaching from the Jinsha Suture to the Kunlun Fault near the 35th parallel and 93° longitude. A large-scale group velocity study by Ritzwoller & Levshin (1998) also indicates lower velocities beneath the Qiangtang block in northern Tibet. Corroborating evidence for an anomalous mantle under northern Tibet also comes from studies of shear wave splitting (Huang *et al.* 2000), which show significant anisotropy with a roughly east–west polarization of the fast shear wave beneath the Qiangtang block north of the Bangong–Nujiang Suture. The anisotropy is assumed to be caused by mantle material squeezed out to the east by advancing Indian and Asian lithosphere. Investigations of the crust beneath Tibet (Owens & Zandt 1997) show low velocities in the northern part that are interpreted as a consequence of heating from an anomalously hot mantle below. Another line of evidence are magnetotelluric studies (Wei *et al.* 2001), which find a deep-reaching (>100 km) region of high electrical conductivity beneath the Quiangtang Terrane. There is one recent surface wave study of Tibet by Griot *et al.* (1998), in which it is argued that the low surface wave phase velocities beneath northern Tibet found in previous surface wave studies (Brandon & Romanowicz 1986; Bourjot & Romanowicz 1992; Wu & Levshin 1994) up to periods of 70 s are caused by the very thick and slow crust under Tibet and not by a slow mantle. After correcting their phase velocity maps for crustal structure, Griot *et al.* find high-velocity zones beneath the thick crust that appear to gently dip inwards under the Tibet plateau. While these results are in agreement with the results of this study for most parts of the Tibet plateau, a lack of resolution owing to insufficient path coverage might have prevented the detection of low velocities under northern Tibet.

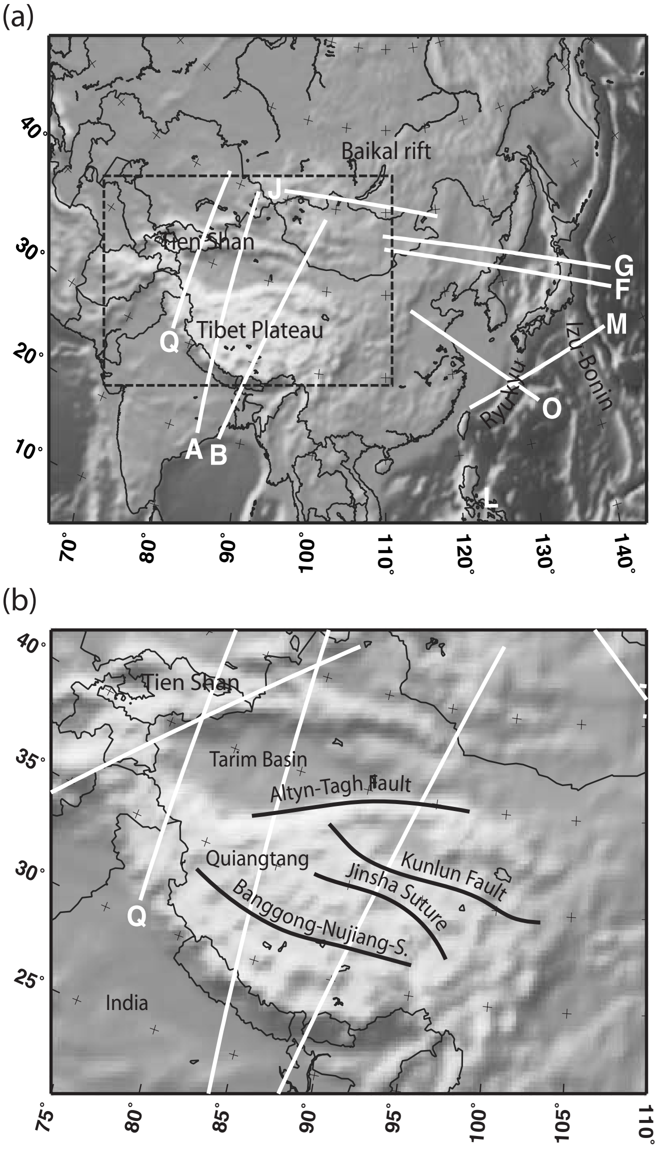

Figure 11. (a) Location of the cross-sections shown in Fig. 10. (b) enlarged view of the Tibet Plateau and surroundings with some tectonic elements mentioned in the text.

Our results on crustal thickness under Tibet are in fair agreement with wide-angle reflection data interpreted by Zhao *et al.* (2001). They observe a gradual thinning of the crust as one moves northward from the Banggong–Nujiang Suture at 32° northern latitude.

Maximum crustal thickness is attained in their model at approximately 30°N. Our model also shows the general northward thinning of the crust, but the maximum crustal thickness is reached under the Banggong–Nujiang Suture.

According to our 3-D model, the Baikal rift zone is underlain by a nearly vertical, plume-shaped low-velocity zone that, however, does not reach the surface. It extends from approximately 100 km depth down to the 660 km discontinuity but not further. Whether it is an active plume or rather a passive upwelling of asthenosphere caused by extension across the rift zone cannot be decided from the tomographic model. The fact that it does not reach the surface is consistent with the absence of volcanism in the Baikal rift zone. Studies on xenoliths brought to the surface 1 Ma in the Tariat region south-west of the Lake Baikal (Ionov *et al.* 1998) indeed indicate that the source of the magmatism there is deeper than 70 km. In contrast with these findings, Gao *et al.* (1994) interpret teleseismic traveltime residuals observed along a profile across the Baikal rift zone as an upwarp of the lithosphere–asthenosphere boundary to a minimum depth of 40 km under the eastern rim of Lake Baikal. Such a two-layer model, however, appears to be fairly simplistic and is certainly not the only way to interpret these data. GPS measurements by Calais *et al.* (1998) show a much greater extension across the Baikal rift zone than expected from the Indo-Asian collision alone. Calais *et al.* conclude that the deformation is also influenced by the western Pacific subduction zones. The surplus extension might, however, just as well be caused by an active plume as suggested by our model.

Since the slabs in the West Pacific subduction zones have been mapped earlier by *P*-wave delay-time tomographic studies (Fukao & Obayashi 1992; Van der Hilst *et al.* 1993; Bijwaard *et al.* 1998) with partially higher resolution than provided by the long-period *S* and surface waves, these structures may serve as a test for our waveform inversion method that provides an independent model of *S*-wave velocity. The agreement with the models from these studies concerning the location of the subducted slabs is excellent with the exception of the Kuril slab, which is not mapped in our model below 400 km depth. This can, however, be explained with loss of resolution at greater depths owing to lack of paths crossing this region. Other well-known features such as stagnation of the Japan and Izu–Bonin slab at the 660 km discontinuity and the varying dip angles of the subduction zones are well retrieved in our model and thus confirm the findings of the delay-time tomographies.

In a classical PWI study of Southeast Asia, Lebedev & Nolet (2002) found a low-velocity anomaly beneath the island of Hainan that they termed the Hainan Plume. This anomaly also appears in our model but it does not reach deeper than 450 km. Since Hainan is close to the southern boundary of the study area, lack of resolution may have prevented a mapping at greater depths.

8 CONCLUSIONS

Our study demonstrates the feasibility of a large-scale waveform inversion based on an accurate wave-theoretical forward algorithm and exact sensitivity kernels. With the increasing performance of computers, we will be able to extend the frequency range of the modelled seismograms and gain in resolution, if the station density also increases. The numerical effort is much greater than with ray theoretical or pure-path approaches. As a reward, we obtain excellent resolution with a relatively ‘small’ number of seismograms. Each seismogram contributes approximately 20–30 constraints on the *S*-velocity structure. Given sufficient path coverage, our waveform inversion method produces high-resolution *S*-velocity models of the upper mantle that are useful for tectonic and geodynamic interpretations.

A 3-D *S*-velocity model of the East Asian mantle was derived with average resolution of approximately 400 km throughout the upper mantle and a peak resolution of approximately 200 km at depths shallower than 250 km beneath Tibet and the subduction zones of the Western Pacific. Subducted slabs are imaged as narrow, linear high-velocity zones. Their locations correlate very well with earthquake hypocentres. Behind the subduction zones is an extended backarc region associated with very low shear wave velocities at shallow mantle depths. There is a plume-like low-velocity zone beneath the southern tip of the Baikal rift, indicating an additional active component controlling the extension of the rift. Tibet is underlain by fast mantle lithosphere in its southern part, indicating underthrusting of the Indian lithosphere under Tibet. Beneath most of northern Tibet, however, velocities exhibit average-to-slow values. Fast Asian lithosphere is imaged to the north of Tibet under the Tarim basin. The relatively low velocities beneath northern Tibet reach down into the transition zone. Indian and Asian high-velocity lithospheres join beneath the Karakorum at the western tip of the Tibet plateau. To the north, under the Tien Shan, the lithosphere again appears to be slow with a possible connection into the transition zone.

ACKNOWLEDGMENTS

We thank the IRIS-DMC for making the seismological data available. Data from the Tibetan Plateau were collected in a joint project between the University of South Carolina (T. Owens and G. Randall), SUNY-Binghampton (F. Wu) and the research group of R. Zeng, State Seismological Bureau, China. I thank F. Pollitz and S. van der Lee for their constructive reviews.

REFERENCES

Alsina, D., Woodward, R.L. & Snieder, R.K., 1996. Shear-wave velocity structure in North America from large-scale waveform inversions of surface waves, *J. geophys. Res.,* **101,** 15 969–15 986.

Barry, T.L. & Kent, R.W., 1998. Cenozoic magmatism in Mongolia and the origin of Central and East Asian basalts, *Geodynamic series,* **27,** 347–364.

Bijwaard, H., Spakman, W. & Engdahl, E.R., 1998. Closing the gap between regional and global traveltime tomography, *J. geophys. Res.,* **103,** 30 055–30 078.

Bjoerck, A., Elfving, T. & Strakos, Z., 1998. Stability of conjugate gradient and Lanczos methods for linear least squares problems, *SIAM J. Matrix Anal. Appl.,* **19,** 720–736.

Bourjot, L. & Romanowicz, B., 1992. Crust and upper-mantle tomography in Tibet using surface waves, *Geophys. Res. Lett.,* **19,** 881–884.

Brandon, C. & Romanowicz, B., 1986. A no-lid zone in the central Chang-Tang platform of Tibet: evidence from pure path phase velocitiy measurements of long-period Rayleigh waves, *J. geophys. Res.,* **91,** 6547–6564.

Calais, E. *et al.*, 1998. Crustal deformation in the Baikal rift from GPS measurements, *Geophys. Res. Lett.,* **25,** 4003–4006.

Das, T. & Nolet, G., 1998. Crustal thickness map of the western United States by partitioned waveform inversion, *J. geophys. Res.,* **103,** 30 021–30 038.

Dziewonski, A., Chou. T.-A. & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.,* **86,** 2825–2852.

Engdahl, E.R., van der Hilst, R. & Buland, R., 1998. Global teleseismic earthquake relocation with improved traveltimes and procedures for depth determination, *Bull. Seismol. Soc. Am.,* **88,** 722–743.

Friederich, W., 1999. Propagation of seismic shear and surface waves in a laterally heterogeneous mantle by multiple forward scattering, *Geophys. J. Int.,* **136,** 180–204.

Friederich, W., Wielandt, E. & Stange, S., 1993. Multiple forward scattering of surface waves: comparison with an exact solution and Born single-scattering methods, *Geophys. J. Int.,* **112,** 264–275.

Fukao, Y. & Obayashi, M., 1992. Subducting slabs stagnant in the mantle transition zone, *J. geophys. Res.,* **7,** 4809–4822.

Gao, S. *et al.*, 1994. Asymmetric upwarp of the asthenosphere beneath the Baikal rift zone, Siberia, *J. geophys. Res.,* **99,** 15 319–15 330.

Griot, D.-A., Montagner, J.-P. & Tapponnier, P., 1998. Phase velocity structure from Rayleigh and Love waves in Tibet and its neighbouring regions, *J. geophys. Res.,* **103,** 21 215–21 232.

Huang, W.-C. *et al.*, 2000. Seismic polarization anisotropy beneath the central Tibetan Plateau, *J. geophys. Res.,* **105,** 27 979–27 989.

Ionov, D.A., O'Reilly S.Y. & Griffin, W.L., 1998. A geotherm and lithospheric section for central Mongolia (Tariat region), *Geodynam. Ser.,* **27,** 127–151.

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.

Lebedev, S. & Nolet, G., 1997. The upper mantle beneath the Philippine Sea region from waveform inversions, *Geophys. Res. Lett.,* **24,** 1851–1854.

Lebedev, S. & Nolet, G., 2002. The upper mantle beneath Southeast Asia, *J. geophys. Res.,* in press.

Li, X.-D. & Tanimoto, T., 1993. Waveforms of long-period body waves in a slightly aspherical Earth model, *Geophys. J. Int.,* **112,** 92–102.

Marquering, H. & Snieder, R., 1995. Surface-wave mode coupling for efficient forward modelling and inversion of body-wave phases, *Geophys. J. Int.,* **120,** 186–208.

Marquering, H., Snieder, R. & Nolet. G., 1996. Waveform inversion and the significance of surface-wave mode coupling, *Geophys. J. Int.,* **124,** 258–278.

Masters, G., Laske, G., Bolton, H. & Dziewonski, A., 2000. The relative behaviour of shear velocity, bulk sound speed and compressional velocity in the mantle: implications for chemical and thermal structure, *AGU Monograph Series*, **117,** 63–87.

Meier, T., Lebedev, S., Noelt, G. & Dahlen, F.A., 1997. Diffraction tomography using multimode surface waves, *J. geophys. Res.,* **102,** 8255–8267.

Ni, J. & Barazangi, M., 1983. High-frequency seismic wave propagation beneath the Indian Shield, Himalayan Arc, Tibetan Plateau and surrounding regions: high uppermost mantle velocities and efficient S_n propagation beneath Tibet, *Geophys. J. R. astr. Soc.,* **72,** 665–689.

Nolet, G., 1990. Partitioned waveform inversion and twodimensional structure under the Network of Autonomously Recording Seismographs, *J. geophys. Res.,* **95,** 8499–8512.

Owens, T.J. & Zandt, G., 1997. Implications of crustal property variations for models of Tibetan plateau evolution, *Nature,* **387,** 37–43.

Paige, C.C. & Saunders, M.A., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Soft.,* **8,** 43–71.

Pollitz, F.F., 2001. Remarks on the travelling wave decomposition, *Geophys. J. Int.,* **144,** 233–246.

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

Snieder, R., 1986. 3D-linearized scattering of surface waves and a formalism for surface wave holography, *Geophys. J. R. astr. Soc.,* **84,** 226–241.

Snieder, R. & Nolet, G., 1987. Linearized scattering of surface waves on a spherical Earth, *J. Geophys.,* **61,** 55–63.

Snieder, R. & Romanowicz, B., 1988. A new formalism for the effect of lateral heterogeneity on normal modes and surface waves—I. isotropic perturbations, perturbations of interfaces and gravitational perturbations, *Geophys. J.,* **92,** 207–222.

Van der Hilst, R., Engdahl, E.R. & Spakman, W., 1993. Tomographic inversion of *P* and *Pp* data for aspherical mantle structure below the northwest Pacific region, *Geophys. J. Int.,* **115,** 264–302.

Van der Lee, S. & Nolet, G., 1997a. Upper mantle *S*-velocity structure of North America, *J. geophys. Res.,* **102,** 22 815–22 838.

Van der Lee, S. & Nolet, G., 1997b. Seismic image of the subducted trailing fragments of the Farallon plate, *Nature,* **386,** 266–269.

Wei, W. *et al.*, 2001. Detection of widespread fluids in the Tibetan crust by magnetotelluric studies, *Science,* **292,** 716–718.

Wiechert, E., 1926. Untersuchung der Erdrinde mit Hilfe von Sprengungen, *Geolog. Rund.,* **17,** 339–346.

Wielandt, E. & Streckeisen, G., 1982. The leaf-spring seismometer: design and performance, BSSA, **72,** 2349–2367.

Wittlinger, G. *et al.*, 1996. Seismic tomography of northern Tibet and Kunlun: evidence for crustal blocks and mantle velocity contrasts, *Earth planet. Sci. Lett.,* **139,** 263–279.

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.

Wu, F.T. & Levshin, A., 1994. Surface-wave group velocity tomography of East Asia, Phys., *Earth Planet. Int.,* **84,** 59–77.

Zhao, L. & Jordan, T.H., 1998. Sensitivity of frequency-dependent traveltimes to laterally heterogeneous, anisotropic Earth structure, *Geophys. J. Int.,* **133,** 683–704.

Zhao, W. *et al.*, 2001. Crustal structure of central Tibet as derived from project INDEPTH wide-angle seismic data, *Geophys. J. Int.,* **145,** 486–498.

Zielhuis, A. & Nolet, G., 1994. Shear-wave velocity variations in the upper mantle beneath central Europe, *Geophys. J. Int.,* **117,** 605–715.