Geophysical Journal International

Geophys. J. Int. (2015) **201,** 1251–1263
GJI Seismology

doi: 10.1093/gji/ggv080

Direct inversion of surface wave dispersion for three-dimensional shallow crustal structure based on ray tracing: methodology and application

Hongjian Fang,[1,2] Huajian Yao,[1,2] Haijiang Zhang,[1,2] Yu-Chih Huang[3,4] and Robert D. van der Hilst[5]

[1] *Laboratory of Seismology and Physics of Earth's Interior, School of Earth and Space Sciences, University of Science and Technology of China, Hefei* 230026, *China. E-mail: hjyao@ustc.edu.cn*
[2] *National Geophysical Observatory at Mengcheng, Anhui, China*
[3] *Institute of Earth Sciences, Academia Sinica, Taipei, Taiwan*
[4] *Taiwan Volcano Observatory at Tatun, Ministry of Science and Technology, Taipei, Taiwan*
[5] *Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA* 02139, *USA*

Accepted 2015 February 16. Received 2015 February 8; in original form 2014 August 6

SUMMARY

We propose a method to invert surface wave dispersion data directly for 3-D variations of shear wave speed, that is, without the intermediate step of phase or group velocity maps, using frequency-dependent ray tracing and a wavelet-based sparsity-constrained tomographic inversion. A fast marching method is used to compute, at each period, surface wave traveltimes and ray paths between sources and receivers. This avoids the assumption of great-circle propagation that is used in most surface wave tomographic studies, but which is not appropriate in complex media. To simplify the problem we consider quasi-stratified media with smoothly varying seismic properties. We represent the 3-D shear wave speed model by means of 1-D profiles beneath grid points, which are determined from all dispersion data simultaneously using a wavelet-based sparsity-constrained tomographic method. The wavelet coefficients of the wave speed model are estimated with an iteratively reweighted least squares algorithm, and upon iteration the surface wave ray paths and the data sensitivity matrix are updated using the newly obtained wave speed model. To demonstrate its feasibility, we apply the method to determine the 3-D shallow crustal shear wave speed variations in the Taipei basin of Taiwan using short period interstation Rayleigh wave phase velocity dispersion measurements extracted from the ambient noise cross-correlation method. The results are consistent with previous studies and reveal strong shallow crustal heterogeneity that correlates with surface geology.

Key words: Wavelet transform; Inverse theory; Surface waves and free oscillations; Seismic tomography.

1 INTRODUCTION

Surface wave tomography based on dispersion or waveforms from earthquake data plays an important role in studies of the structure of the Earth's crust and upper mantle on both regional and global scales (e.g. Woodhouse & Dziewonski 1984; Nolet 1990; Montagner & Tanimoto 1991; Simons *et al.* 1999; Ritzwoller *et al.* 2001; Boschi & Ekström 2002; An *et al.* 2009; Yao *et al.* 2010). In the past decade, surface wave tomography based on ambient noise cross-correlations has greatly improved our understanding of regional crustal structures (e.g. Shapiro *et al.* 2004; Sabra *et al.* 2005; Yao *et al.* 2006; Yang *et al.* 2007; Lin *et al.* 2008). Crustal studies based on ambient noise tomography are typically conducted in the period band of 5–40 s, but shorter period surface waves (∼1 s, using station spacing of ∼20 km or less) have been used to investigate shallow crustal or even near surface shear wave speed variations (e.g. Picozzi *et al.* 2009; Huang *et al.* 2010; Young *et al.* 2011; Pilz *et al.* 2012; Lin *et al.* 2013; Shirzad & Shomali 2014). This line of research is important for the prediction of ground strong motion in seismically active regions (Zhang *et al.* 2008) and for the characterization of overburden structure in oil and gas reservoir fields (e.g. Gouédard *et al.* 2012).

Surface wave or ambient noise tomography using dispersion measurements is usually performed in two steps: (i) construction of 2-D phase (or group) velocity maps using the traveltime tomographic method either based on the ray theory (e.g. Cara & Lévêque 1987; Trampert & Woodhouse 1995; Ekström *et al.* 1997; Barmin *et al.* 2001; Yao *et al.* 2006) or 2-D finite frequency sensitivity kernels

(e.g. Ritzwoller *et al.* 2002; Yoshizawa & Kennett 2004), followed by (ii) point-wise inversion of dispersion data for 1-D profiles of shear wave speed as a function of depth at each grid/cell point, which together form the 3-D shear wave speed model (e.g. Shapiro & Ritzwoller 2002; Yao *et al.* 2008). In regions with dense station distribution (such as USArray), localized estimation of 2-D phase velocity variations with azimuthal anisotropy can be efficiently obtained from tomographic methods based on the eikonal or Helmholtz equations from earthquake surface waves or ambient noise cross-correlations (Pollitz 2008; Liang & Langston 2009; Lin *et al.* 2009; Pollitz & Snoke 2010; Lin & Ritzwoller 2011).

Direct linearized inversion of path-specific surface wave dispersion data without constructing intermediate phase or group velocity maps for 3-D shear wave speed structure can be achieved, for instance, using ray-based approaches (Boschi & Ekström 2002; An *et al.* 2009; Feng & An 2010; Pilz *et al.* 2012) or analytical 3-D finite frequency sensitivity kernels based on a 1-D reference model (e.g. Zhou *et al.* 2006). However, these approaches do not usually update the ray paths and sensitivity kernels for the newly obtained 3-D models, and one-step linearization may lead to biased wave speed estimation in complex media, for instance, in the shallow crust where wave speed variations can exceed 20 per cent (e.g. Picozzi *et al.* 2009; Young *et al.* 2011; Lin *et al.* 2013). Adjoint tomography and full waveform inversion using earthquake waveforms (Tape *et al.* 2010; Zhu *et al.* 2012) or ambient noise cross-correlations (Chen *et al.* 2014; Gao & Shen 2014) directly solve for 3-D wave speed models with 3-D sensitivity kernels updated iteratively. This type of approach is accurate in theory, but it comes at a high computational cost, in particular at high frequencies, unless sufficiently accurate 3-D reference models are used to facilitate convergence.

To balance accuracy and computational efficiency, we propose a ray tracing-based direct inversion method that inverts (in a single inversion) all path-specific surface wave dispersion data for the 3-D shear wave speed model. The linearization scheme is generally similar to that proposed by Boschi & Ekström (2002) or Feng & An (2010), but we perform surface wave ray tracing at each period using a fast marching method (Rawlinson & Sambridge 2004) and iteratively update the sensitivity kernels of period-dependent dispersion measurements to updates of the 3-D wave speed model. Accounting for the effect of heterogeneity on wave propagation is particularly useful for short period dispersion measurements, which are sensitive to the strong and complicated shallow crustal structure. There is no formal obstruction for the incorporation of finite frequency effects, but in the implementation presented here we use simple ray geometry.

In order to facilitate multiscale tomography and overcome ill-posedness of the inverse problem due to uneven path coverage, we use a wavelet-based sparsity-constrained seismic tomography method (e.g. Simons *et al.* 2011; Fang & Zhang 2014) for the direct inversion. To show its validity, we apply the direct inversion method to the short period dispersion data set obtained from ambient noise cross-correlation in the Taipei Basin of Taiwan (Huang *et al.* 2010).

2 METHODOLOGY

In this section, we describe the strategy for the direct inversion of dispersion measurements for 3-D shear wave speed variations using a fast marching ray tracing method (Rawlinson & Sambridge 2004) and a wavelet-based sparsity-constrained inversion technique (Fang & Zhang 2014). We first provide the mathematical representations of the forward and inverse problems and then illustrate the inversion procedure with real and synthetic data.

2.1 The forward problem

Traditionally in global seismology, we measure the average surface wave phase velocity $c_{\rm AB}(\omega)$ at an angular frequency ω between source A and receiver B with an assumption of great circle propagation of surface waves. Therefore, the actual traveltime measurement $t_{\rm AB}(\omega)$ from A to B can be obtained from

$$t_{\rm AB}(\omega) = \frac{L_{\rm AB}}{c_{\rm AB}(\omega)}, \tag{1}$$

with $L_{\rm AB}$ the great circle distance between A and B.

Due to structural heterogeneity, however, the real surface wave ray path may deviate from the great circle path between source and receiver. The traveltime can thus be expressed as:

$$t_{\rm AB}(\omega) = \int_{l_{\rm AB}} S(l,\omega)\,{\rm d}l, \tag{2}$$

where $S(l, \omega)$ is the (unknown) local slowness along (the unknown) actual ray path $l_{\rm AB}$ between A and B. Eq. (2), which assumes the high-frequency approximation, can be discretized as:

$$t_{\rm AB}(\omega) = \sum_{p=1}^{P} S_p(\omega)\Delta l_{\rm AB}, \tag{3}$$

where $S_p(\omega)$ represents the phase slowness for path segment $\Delta l_{\rm AB}$ along AB, and P is the number of path segments. In view of Fermat's Principle, the problem can be linearized around a (known) reference path that can be computed, for instance, along the great circle in a laterally homogeneous medium or (as will be done here) by ray tracing in a heterogeneous medium.

For the forward problem we parametrize the area under study using a regular grid of K points (Fig. 1a), and underneath each grid point k we consider a 1-D model Θ_k. This 1-D model, which itself is parametrized by means of depth nodes (Fig. 1b), is used to calculate local phase (or group) velocity as a function of frequency for that grid point. We construct the 2-D distribution of the phase slowness estimates $\hat{S}_k(\omega)(k = 1, 2, \ldots, K)$ at frequency ω for the study region and then use bilinear interpolation to estimate the phase slowness at any point along the path $l_{\rm AB}$:

$$S_p(\omega) = \sum_{k=1}^{K} \upsilon_{pk}\hat{S}_k(\omega), \tag{4}$$

where υ_{pk} are bilinear interpolation coefficients. The slowness at the kth grid point, that is $\hat{S}_k(\omega)$, is obtained from a 1-D model Θ_k as:

$$\hat{S}_k(\omega) = g(\Theta_k, \omega), \tag{5}$$

where $g(\Theta_k, \omega)$ is a forward function that maps Θ_k to frequency dependent phase or group velocities based on the layered model according to Haskell (1953) and Thomson (1950), with modifications due to Dunkin (1965) and Herrmann (2001). Therefore, we need to transform our model represented by vertical grid nodes to a layered model (a stack of homogeneous elastic isotropic layers) using a simple linear interpolation method (Fig. 1b).

The traveltime $t_{\rm AB}(\omega)$ at frequency ω along the path AB can then be expressed as

$$t_{\rm AB}(w) = \sum_{p=1}^{P} S_p(w)\Delta l_{\rm AB} = \sum_{p=1}^{P}\sum_{k=1}^{K} \upsilon_{pk}\hat{S}_k(w)\Delta l_{\rm AB}. \tag{6}$$

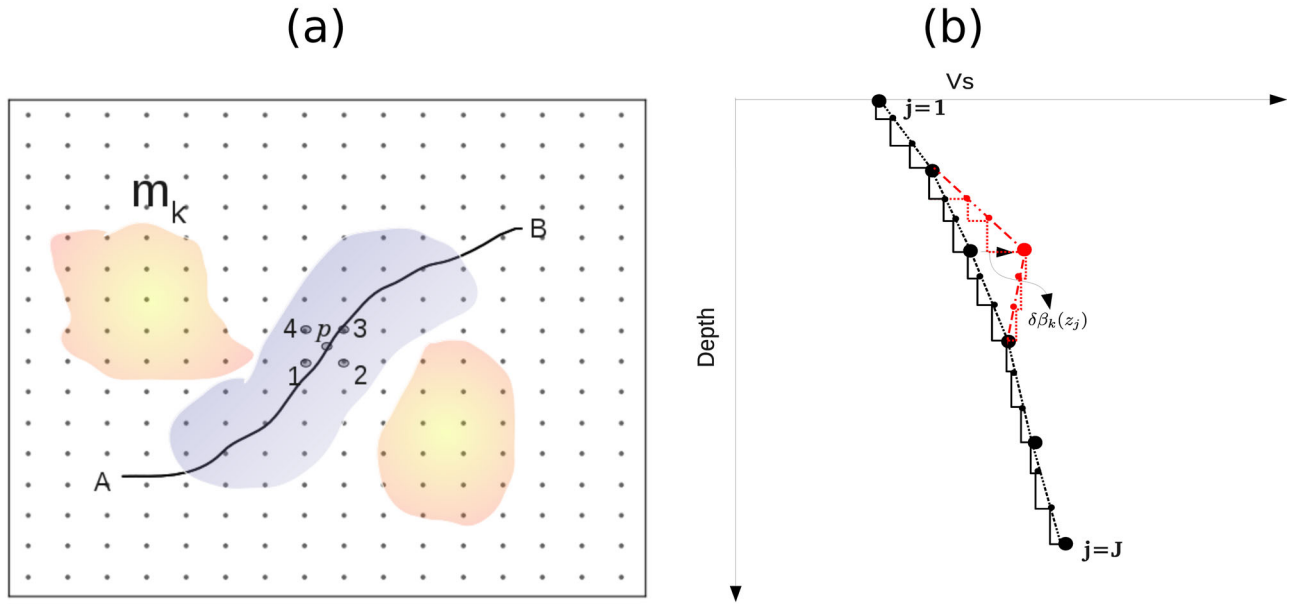

Figure 1. Illustration of the model discretization in the horizontal plane (a) and in the vertical direction (b). In (a) the black solid line represents the propagation path between two stations A and B for the surface wave at some period. The phase slowness at any point p along the path is determined from the values at four surrounding horizontal grid points using a bilinear interpolation method. In (b) the vertical grid model (bigger black dots with finer layers shown as the black line) is perturbed (as shown by the red dot and red dashed lines) to compute the depth sensitivity of dispersion data to model parameters (bigger black dots) via a difference method (see texts for more details).

For the ith surface wave traveltime measurement, we rewrite eq. (6) as:

$$t_i(w) = \sum_{k=1}^{K} \nu_{ik} \hat{S}_k(w), \tag{7}$$

where ν_{ik} is $\sum_{p=1}^{P} \upsilon_{pk}^{(i)} \Delta l_i$ and $\upsilon_{pk}^{(i)}$ are the bilinear interpolation coefficients along the ray path associated with the ith traveltime data.

2.2 The inverse problem

The objective of the tomographic inversion is to find a model $\mathbf{m}$ that minimizes the differences $\delta t_i(\omega)$ between the observed times $t_i^{\rm obs}(w)$ and the model predictions $t_i(w)$ for all frequencies w. The traveltime difference for path i is given by

$$\delta t_i(w) = t_i^{\rm obs}(w) - t_i(w) = \sum_{k=1}^{K} \nu_{ik} \delta \hat{S}_k(w) \approx -\sum_{k=1}^{K} \nu_{ik} \frac{\delta C_k(w)}{C_k^2(w)}, \tag{8}$$

where $t_i(w)$ is the calculated traveltime using a reference model that can be updated in the inversion, and C_k and $\delta C_k(\omega)$ are, respectively, the phase velocity and phase-velocity perturbation at the kth grid point. The phase velocity $C_k = 1/\hat{S}_k$ and its perturbation can be expressed as:

$$\delta C_k(w) = \int \left[\left.\frac{\partial C_k(w)}{\partial \alpha_k(z)}\right|_{\Theta_k} \delta\alpha_k(z) + \left.\frac{\partial C_k(w)}{\partial \beta_k(z)}\right|_{\Theta_k} \delta\beta_k(z) + \left.\frac{\partial C_k(w)}{\partial \rho_k(z)}\right|_{\Theta_k} \delta\rho_k(z) \right] {\rm d}z, \tag{9}$$

where Θ_k denotes the 1-D reference model and $\alpha_k(z)$, $\beta_k(z)$ and $\rho_k(z)$ are the compressional wave speed, shear wave speed and mass density, respectively. The sensitivities of phase velocity to each model parameter at depth can be computed using a simple difference method by performing two forward dispersion calculations with a small perturbation to that parameter (Fig. 1b). Typically, surface wave dispersion is mostly sensitive to shear wave speed. In the shallow crust, however, short period Rayleigh wave dispersion also has a significant sensitivity to compressional wave speed. We relate compressional wave speed and density to shear wave speed using the empirical relationships $\alpha(z) = \sum_n \chi_n^{[\alpha]} \beta^n(z)$ and $\rho(z) = \sum_n \chi_n^{[\rho]} \alpha^n(z)$ given by Brocher (2005), where $\chi_n^{[\alpha,\rho]}$ represent the fitting polynomial coefficients. Then, $\delta\alpha(z) = \sum_n n\chi_n^{[\alpha]} \beta^{n-1}(z)\delta\beta(z) = R_\alpha(z)\delta\beta(z)$ and $\delta\rho(z) = \sum_n n\chi_n^{[\rho]} \alpha^{n-1}(z)\delta\alpha(z) = R_\rho(z)\delta\beta(z)$, where $R_\alpha(z) = \sum_n n\chi_n^{[\alpha]} \beta^{n-1}(z)$ and $R_\rho(z) = R_\alpha(z) \sum_n n\chi_n^{[\rho]} \alpha^{n-1}(z)$, which are the corresponding partial derivatives with respect to β. Finally, after discretizing the integration along depth in eq. (9) and inserting it into eq. (8), we obtain:

$$\delta t_i(w) = \sum_{k=1}^{K} \left(-\frac{\nu_{ik}}{C_k^2(\omega)} \right) \sum_{j=1}^{J} \left[R_\alpha(z_j) \frac{\partial C_k(\omega)}{\partial \alpha_k(z_j)} + R_\rho(z_j) \frac{\partial C_k(\omega)}{\partial \rho_k(z_j)} + \left.\frac{\partial C_k(\omega)}{\partial \beta_k(z_j)}\right] \right|_{\Theta_k} \delta\beta_k(z_j) = \sum_{l=1}^{M} G_{il} m_l, \tag{10}$$

where J is the number of grid nodes in the depth direction, and $M = KJ$, which is the total number of grid points. Through eq. (10) we explicitly include the perturbation of compressional wave speed and mass density on the calculation of surface wave dispersion.

Eq. (10) can be written in the matrix form:

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \tag{11}$$

where $\mathbf{d}$ is the surface wave traveltime residual vector for all paths and frequencies, $\mathbf{G}$ is the data sensitivity matrix and $\mathbf{m}$ is the model

parameter vector expressed as:

$$\mathbf{m} = [\delta\beta_1(z_1)...\delta\beta_1(z_J)\,\delta\beta_2(z_1)...\delta\beta_2(z_J)...\delta\beta_K\,(z_1)...\delta\beta_K(z_J)]^T\,. \tag{12}$$

A solution to eq. (11) can be found by minimizing the following cost function:

$$\Phi(\mathbf{m}) = ||\mathbf{d} - \mathbf{G}\mathbf{m}||_2^2 + \lambda||\mathbf{L}\mathbf{m}||_2^2, \tag{13}$$

where the first term on the right-hand side gives the ℓ_2-norm data misfit and the second term denotes the ℓ_2-norm model regularization term, $\mathbf{L}$ is a model smoothing operator and λ is the weighting parameter balancing data fitting and model regularization. $\mathbf{L}$ is usually chosen as the first- or second-order spatial derivative operator (Simons *et al.* 2002; Aster *et al.* 2013), although this tends to smear out sharp boundaries and fine scale features even in regions with high ray path density. An estimate of the solution $\mathbf{m}$ can then be obtained from

$$\tilde{\mathbf{m}} = (\mathbf{G}^\mathbf{T}\mathbf{G} + \lambda\mathbf{L}^\mathbf{T}\mathbf{L})^{-1}\mathbf{G}^\mathbf{T}\mathbf{d}, \tag{14}$$

using, for instance, the LSQR algorithm by Paige & Saunders (1982).

2.3 Wavelet-based sparsity-constrained inversion

Ray coverage is often highly uneven due to non-uniform station and/or source geometry, missing data and propagation effects due to lateral variations in elastic properties. Ideally, one would adapt model parametrization to data distribution in a way that it will have good resolution in regions with good data constraint, while keeping the long wavelength features of the model in regions that are poorly constrained by the data. A model parametrization based on data constraints is difficult to achieve using regular grids or cells and conventional regularization methods. For this reason, irregularly distributed nodes or cells are often used in seismic tomography to achieve data adaptive spatial resolution (e.g. Karason & Van Der Hilst 2000; Spakman & Bijwaard 2001; Debayle & Sambridge 2004; Zhang & Thurber 2005).

Here we use a wavelet-based sparsity-constrained method (Fang & Zhang 2014) to invert surface wave frequency-dependent traveltime measurements for the model parameters, which is data adaptive so that model parametrization is inherently adaptive to data distribution. This method takes advantage of the multiresolution representation of the wavelet transform to solve the model parameters (that is, wave speed anomalies) in the wavelet domain, and it resolves features of different scales depending on the strength of the local data constraints (Chiao & Liang 2003). The main idea of the multiscale tomography in the wavelet domain is to resolve both the approximation coefficients and the detail coefficients in regions where data constraint is good, thus producing higher resolution. For regions with poor data constraint only the approximation coefficients can be resolved, and detail coefficients that cannot be constrained by the data will be zero due to the constraint of ℓ_1-norm regularization, that is, sparsity, by assuming that wave speed models are sparse in the wavelet domain (Simons *et al.* 2011; Charléty *et al.* 2013).

Different from (Fang & Zhang 2014) which minimize the data residuals in the ℓ_2-norm, we use the ℓ_1-norm-based minimization for the data misfit function since it will be more robust to outliers than the ℓ_2-norm or least squares solution (Shearer 1997).

For a wavelet transform on a basis represented by a matrix $\mathbf{W}$ (Daubechies 1992), a wave speed model $\mathbf{m}$ in the wavelet domain can be expressed as $\hat{\mathbf{m}} = \mathbf{W}\mathbf{m}$, where $\hat{\mathbf{m}}$ contains the wavelet coefficients. We perform the wavelet transform using a lifting scheme, which decomposes the original model into approximation coefficients and detail coefficients at different scales (Sweldens 1996). The approximation coefficients represent the smooth components of the model and are obtained by representing the model in the space formed by the scaling functions at different scales. Detail coefficients describe the fine scale features of the model and are obtained by expanding the model using wavelet functions.

For an orthogonal wavelet basis the inverse wavelet transform is $\mathbf{W}^{-1} = \mathbf{W}^T$ so that $\mathbf{W}^T\mathbf{W} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix and superscript T denotes the transpose. We choose a D4 wavelet (Daubechies 1992) because it is orthogonal and its support size is small compared to the dimension of our model (thus minimizing edge effects). Moreover, it has a vanishing moment of 2 which leads to sparse wavelet coefficients. A 3-D wavelet transform is approximated by applying 1-D wavelet transform sequentially to each dimension (Prochazka *et al.* 2011). Therefore, the original inverse problem without model constraints or regularizations becomes

$$\text{minimize} \quad ||\mathbf{G}\mathbf{W}^{-1}\mathbf{W}\mathbf{m} - \mathbf{d}||_1, \tag{15}$$

which can be simplified to

$$\text{minimize} \quad ||\hat{\mathbf{G}}\hat{\mathbf{m}} - \mathbf{d}||_1, \tag{16}$$

where $\hat{\mathbf{G}} = \mathbf{G}\mathbf{W}^\mathbf{T}$ and $\hat{\mathbf{m}} = \mathbf{W}\mathbf{m}$. Now the inverse problem is recast as first seeking the wavelet coefficients of the model and then using the inverse wavelet transform matrix $\mathbf{W}^{-1}$ to infer the model parameters.

As shown by Fang & Zhang (2014), the sensitivity matrix is less ill-posed in the wavelet domain than in the spatial domain. However, some regularization is still needed to stabilize the inversion. We follow Loris *et al.* (2007) and Fang & Zhang (2014) and apply the ℓ_1-norm regularization by assuming that the model parameters in the wavelet domain are generally sparse (Simons *et al.* 2011). Now the inverse problem becomes

$$\text{minimize} \quad ||\hat{\mathbf{G}}\hat{\mathbf{m}} - \mathbf{d}||_1 + \lambda||\hat{\mathbf{m}}||_1, \tag{17}$$

where Langrange multiplier λ is used to balance data misfit and sparsity. If one knows the noise level in the data λ can be determined from the discrepancy principle (Scherzer 1993). If one does not know the data noise, like in seismic tomography, one can use generalized cross validation or the L-curve method to find a solution (Aster *et al.* 2013).

The minimization problem shown in eq. (17) can be solved using an iteratively reweighted least squares (IRLS) method (Scales *et al.* 1988). For this purpose we first rewrite eq. (17) as

$$\text{minimize} \quad \left\| \begin{bmatrix} \hat{\mathbf{G}} \\ \lambda\mathbf{I} \end{bmatrix} \hat{\mathbf{m}} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_1, \tag{18}$$

and then use a residual vector

$$\mathbf{r} = \tilde{\mathbf{G}}\hat{\mathbf{m}} - \tilde{\mathbf{d}}, \tag{19}$$

where $\tilde{\mathbf{G}}$ is $\left[\begin{smallmatrix} \hat{\mathbf{G}} \\ \lambda\mathbf{I} \end{smallmatrix}\right]$ and $\tilde{\mathbf{d}}$ is $\left[\begin{smallmatrix} \mathbf{d} \\ \mathbf{0} \end{smallmatrix}\right]$ to simplify the expression. We thus want to minimize

$$f(\hat{\mathbf{m}}) = \|\mathbf{r}\|_1 = \sum_{l=1}^{N} |r_l|. \tag{20}$$

The function f is non-differentiable at points where one of the elements of $\mathbf{r}$ is zero, but we can compute the derivatives at other points:

$$\frac{\partial f(\hat{\mathbf{m}})}{\partial \hat{m}_n} = \sum_{l=1}^{N} \frac{\partial |r_l|}{\partial \hat{m}_n} = \sum_{l=1}^{N} \tilde{G}_{l,n} \frac{r_l}{|r_l|}. \tag{21}$$

Approximately, the gradient of f is

$$\nabla f(\hat{\mathbf{m}}) \approx \tilde{\mathbf{G}}^{\mathrm{T}} \mathbf{R} \mathbf{r} = \tilde{\mathbf{G}}^{\mathrm{T}} \mathbf{R} (\tilde{\mathbf{G}} \hat{\mathbf{m}} - \tilde{\mathbf{d}}), \tag{22}$$

where $\mathbf{R}$ is a diagonal weighting matrix with diagonal elements that are the absolute values of the reciprocals of the residuals, so that

$$R_{i,i} = \begin{cases} \frac{1}{\epsilon} & |r_i| < \epsilon \\ \frac{1}{|r_i|} & |r_i| \geq \epsilon \end{cases}, \tag{23}$$

where ϵ is a tolerance below which we consider the residuals to be effectively zero. According to eq. (22), optimization (that is solving $\nabla f(\hat{\mathbf{m}}) = 0$) gives

$$\tilde{\mathbf{G}}^{\mathrm{T}} \mathbf{R} \tilde{\mathbf{G}} \hat{\mathbf{m}} = \tilde{\mathbf{G}}^{\mathrm{T}} \mathbf{R} \tilde{\mathbf{d}}. \tag{24}$$

Effectively, these are the normal equations for the least squares problem

$$\text{minimize} \quad \left\| \sqrt{\mathbf{R}} \begin{bmatrix} \hat{\mathbf{G}} \\ \lambda \mathbf{I} \end{bmatrix} \hat{\mathbf{m}} - \sqrt{\mathbf{R}} \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_2. \tag{25}$$

When $\tilde{\mathbf{G}}$ is large and sparse it can be advantageous to apply LSQR (Paige & Saunders 1982) to solve the least squares problem (25) instead of solving the system of equations in (24) directly. In the IRLS procedure, we first solve eq. (25) by setting $\mathbf{R}$ to the identity matrix and then update it iteratively to get a sparse solution.

3 APPLICATION TO TAIPEI BASIN

As a proof of concept, we apply our method using the same dispersion data set as in Huang *et al.* (2010) to estimate the crustal structure of Taipei basin, Taiwan. Using the method due to Yao *et al.* (2006), Huang *et al.* (2010) measured Rayleigh wave phase velocity dispersion for periods ranging from 0.5 to 3.0 s from the empirical Green's functions estimated from ambient noise cross-correlation. Based on eq. (1), we estimated the traveltimes at periods from 0.5 to 3.0 s between two stations from dispersion data. Fig. 2 shows the station distribution as well as some major geological features in the

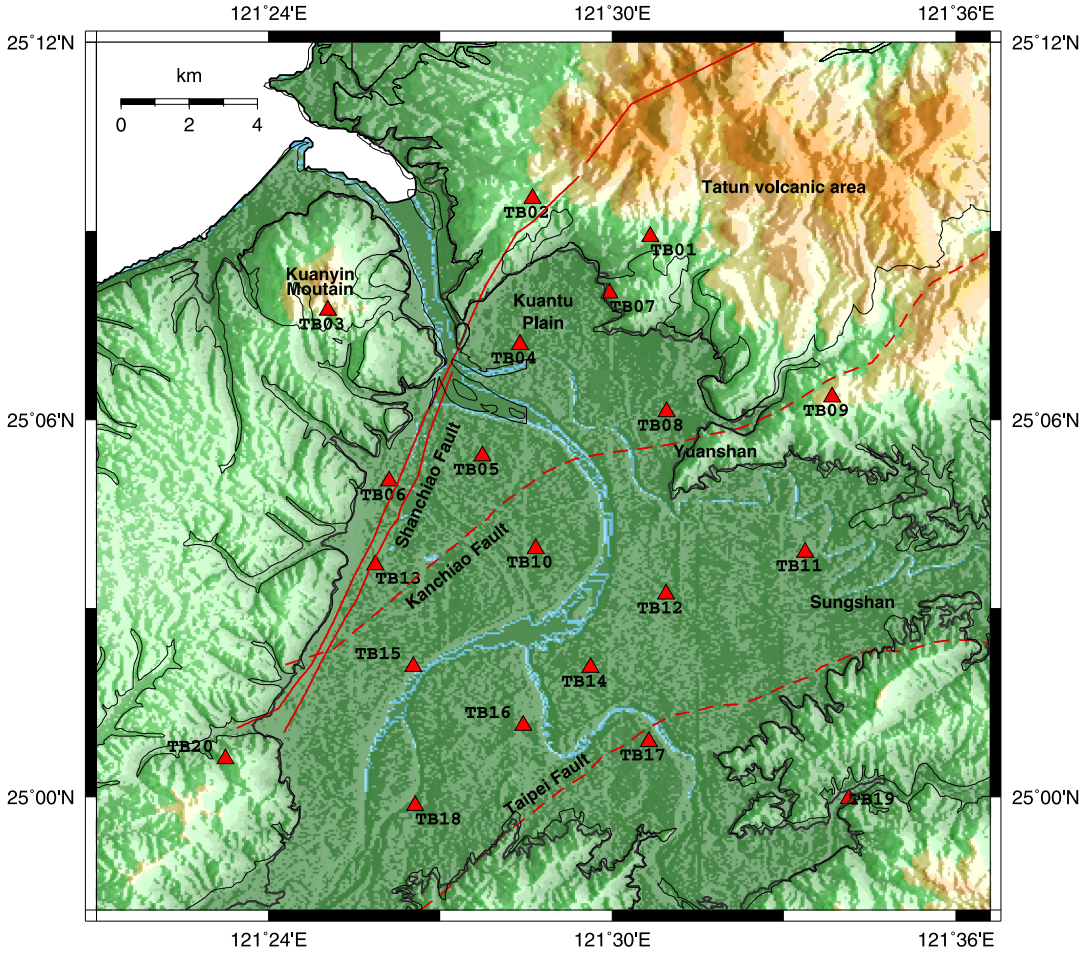

Figure 2. Topography, tertiary basement, faults, river systems and station locations in the Taipei basin. There are one active fault shown as the red solid lines and two suspected active faults shown as the red dash lines. Station locations are shown as red triangles with station names marked.

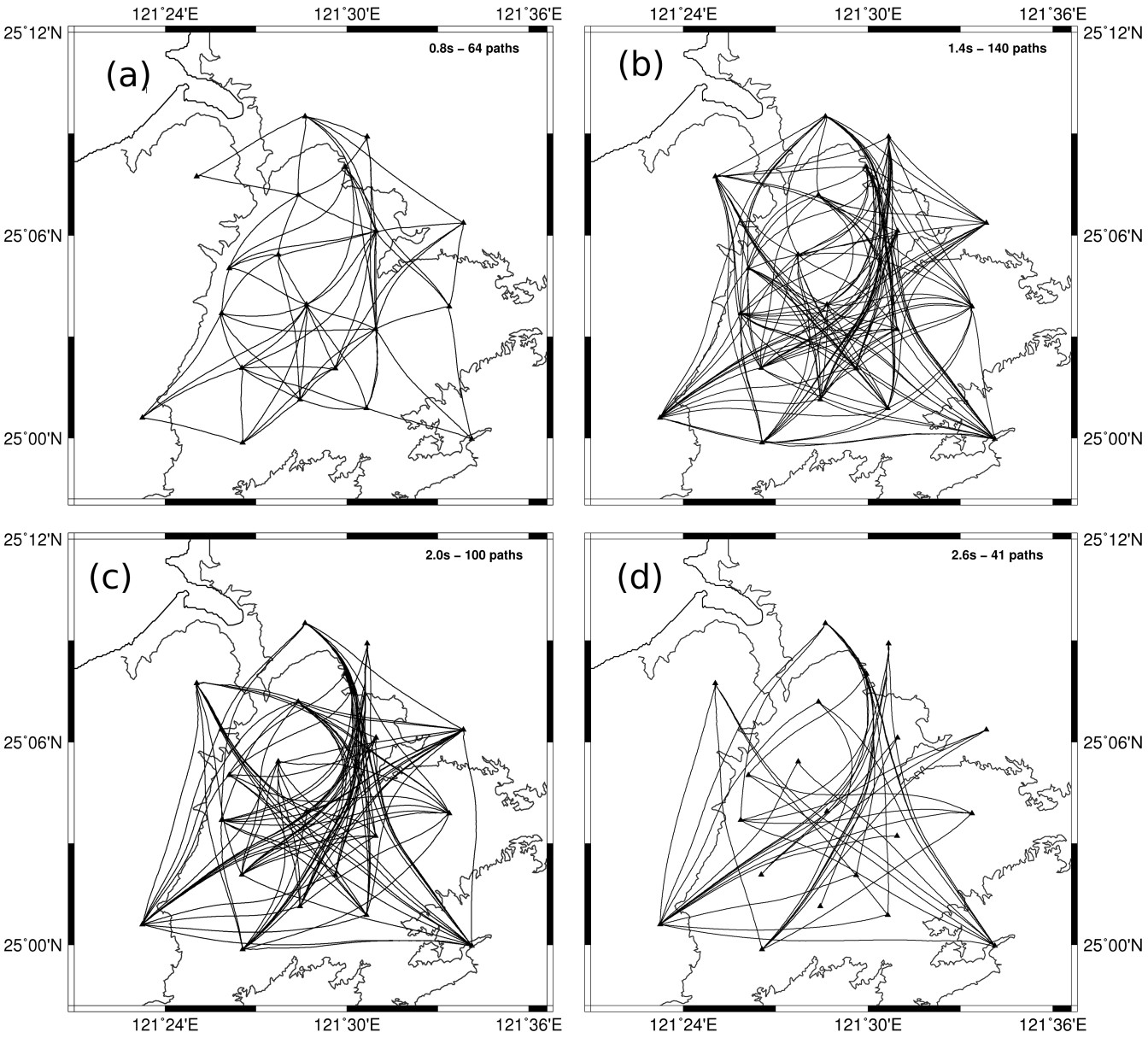

Figure 3. The ray paths obtained from the final 3-D model (see Figs 6 and 7) at four periods by the fast marching method: (a) 0.8 s, (b) 1.4 s, (c) 2.0 s and (d) 2.6 s.

study region. The ray paths for four different periods are shown in Fig. 3. These paths are obtained (with the fast marching method) from the final 3-D model (shown in Figs 6 and 7). The study region is meshed with 16 by 16 grid points with an interval of 0.015° in latitude and 0.017° in longitude and eight grid nodes at 0.3 km spacing in depth. In our ℓ_1-norm inversion based on eq. (25), we choose ϵ (in eq. 23) to be 10^{-6} considering the machine accuracy for single precision floating-point number. As a rule of thumb, the fundamental mode Rayleigh wave phase velocity is mostly sensitive to shear wave speed at depths around 1/3 of its corresponding wavelength. For a uniform half-space Poisson solid, the phase velocity $c = 0.92Vs$ (shear wave speed, Shearer 2009). Therefore, we choose an initial shear wave speed model that is close to the average of measured phase velocities times 1.1 ($Vs \approx 1.1c$) at the depth of 1/3 of the wavelength (Fig. 4). The empirical relationships according to Brocher (2005) are used to obtain the reference compressional wave speed and density models for each surface grid point.

To investigate the performance of our algorithm for the (spatial and spectral) sampling produced by our data we perform checkerboard resolution tests. We construct a 3-D checkerboard model in order to assess the lateral and vertical resolution (Figs 5a and b), compute inter-station Rayleigh wave traveltimes for the same source–receiver distribution as in the real data distribution (using the fast marching method), add 1 per cent random noise to the synthetic traveltime data, and invert the data using the method described above. The results (Figs 5a and b) suggest that structures can be resolved well in regions where the data constraints are good. Fig. 6 shows the final inversion results at different depths in the shallow crust. The results are generally consistent with the phase velocity maps at the corresponding periods in Huang *et al.* (2010) and show

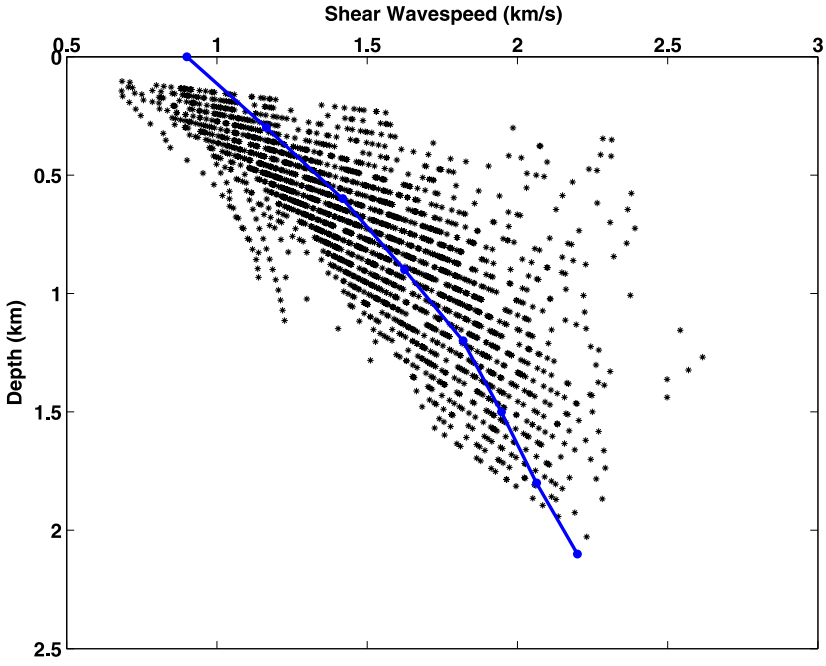

Figure 4. Transformation of all interstation Rayleigh wave phase velocity dispersion curves measured from the ambient noise cross-correlation method (Huang *et al.* 2010) to a depth-shear wave speed approximation profile. For each Rayleigh wave phase velocity measurement, we replace period with 1/3 of the corresponding wavelength (phase velocity times period) at the vertical axis and replace phase velocity with the approximated shear wave speed (1.1 times phase velocity) at the horizontal axis. The average wave speed model from this profile (blue line) is used as the initial reference shear wave speed model in the inversion.

a good correspondence to the geological features, with relatively low shear wave speed in the basin and high wave speed in mountain ranges (Fig. 2), which is also evident from vertical crustal sections (Fig. 7). Upon our iterative, non-linear inversion the standard deviation of the traveltime residuals is reduced from 1.44 to 0.91 s, and the average of the residuals for the final model is close to zero (Fig. 8).

4 DISCUSSION

We have developed an iterative, wavelet-based non-linear method to invert surface wave dispersion data directly for 3-D shear wave speed structure. Compared to conventional methods our new method has the following advantages: (i) it does not require the construction of 2-D phase (or group) velocity maps; (ii) a 3-D reference model can be included; (iii) constraints on the lateral variations in structure can be incorporated, (iv) it is computationally efficient compared to full waveform inversion or adjoint tomography; (v) the regularization used to stabilize the inversion is less subjective than common applications of damping and smoothing, and (vi) the wavelet-based model parametrization is inherently adaptive to data coverage.

We note that the formulation and implementation used here is based on the high-frequency approximation (ray theory) and assumes weak heterogeneity and smooth lateral variation of the elastic properties. There is no obstruction, however, for the incorporation of quasi-3D finite frequency kernels by combining 2-D lateral finite frequency kernels of surface wave traveltimes (e.g. Ritzwoller *et al.* 2002; Yoshizawa & Kennett 2004; Zhou *et al.* 2004) with 1-D depth sensitivity kernels as used in this study. In the computation of 2-D lateral finite frequency kernels, we can use the 3-D wave speed model at each inversion step to obtain the 2-D phase velocity maps and construct the 2-D lateral kernel with respect to each ray path. Therefore, the 2-D lateral kernels can also be updated during the inversion. More accurate kernels can be computed from the adjoint method (Tape *et al.* 2007; Chen *et al.* 2014), but the computational cost at high frequencies (e.g. $\sim$1 Hz) is still large. Therefore, our current approach will be especially useful to investigate shallow crustal or near surface wave speed structure using higher frequency surface wave dispersion data, for instance, from ambient noise cross-correlations or active source data in exploration seismology or deep seismic sounding. If so desired, results from inversions using the method described here can serve as a starting point for full waveform inversion.

An important further advantage and a motivation for its development is its potential to exploit data redundancy though joint surface and body wave inversions (Zhang *et al.* 2014). Indeed, due to its formulation as a traveltime inversion it is easy to combine surface and body wave traveltimes, and as long as the appropriate sensitivity kernels are used it can jointly handle Rayleigh and Love wave (phase and or group) velocities. Since shorter period Rayleigh wave dispersion data still have quite large sensitivity to compressional wave speed structure in the shallow crust, the joint inversion of surface wave and body wave traveltime data may provide better constraints on both compressional wave speed and shear wave speed structures in the shallow crust.

Problems may arise when model uncertainty is required for ℓ_1 norm direct inversions for millions of parameters, which is common in 3-D global or regional tomography. Uncertainty estimation is relatively straightforward in the traditional two-step inversion, which breaks the inverse problem into a linear step to construct the phase/group velocity maps followed by a non-linear step to produce

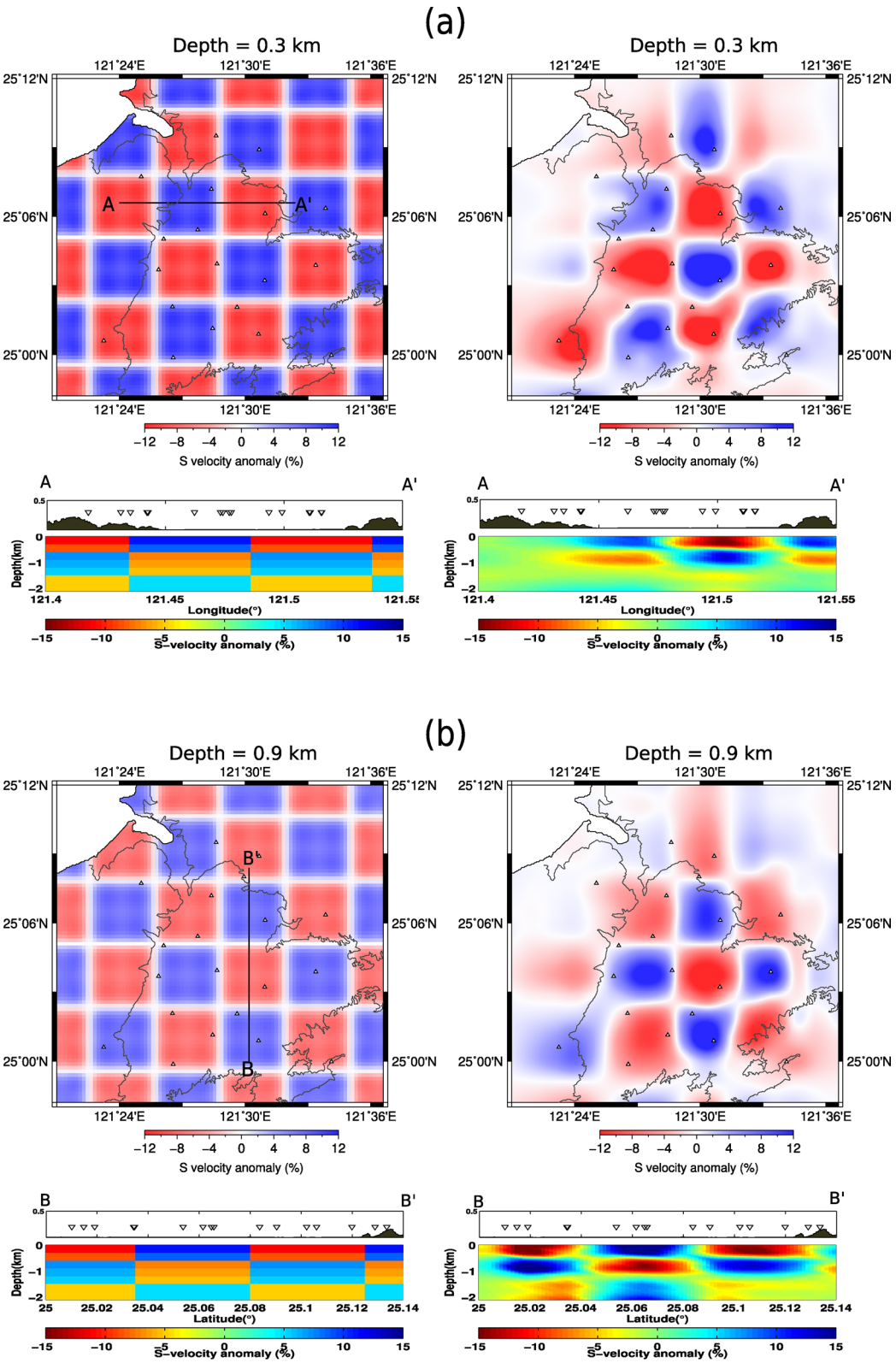

Figure 5. Lateral and vertical shear wave speed slices of the checkerboard test results. The checkerboard models at 0.3 and 0.9 km depths are shown in (a) and (b). The top panels on vertical cross-section slices show the topography along the cross-section and the bottom panel gives the shear wave speed anomaly in per cent with the colour bar given at the bottom. The triangles on each vertical cross-section slice mark the location of the stations close to the corresponding profile.

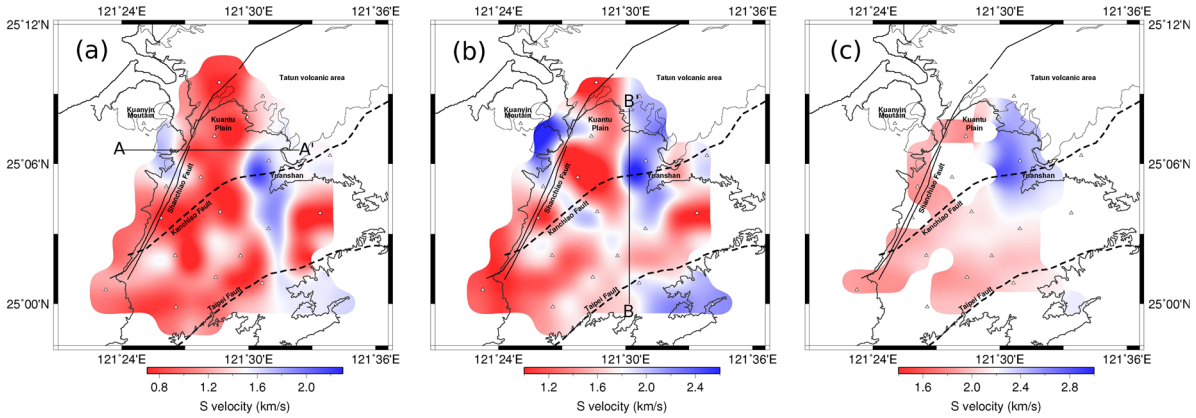

Figure 6. Horizontal slices of shear wave speed model of the Taipei Basin at depths of: (a) 0.3 km, (b) 0.9 km and (c) 1.5 km. Only zones of good resolution based on the model uncertainty results (see Fig. 9) are shown.

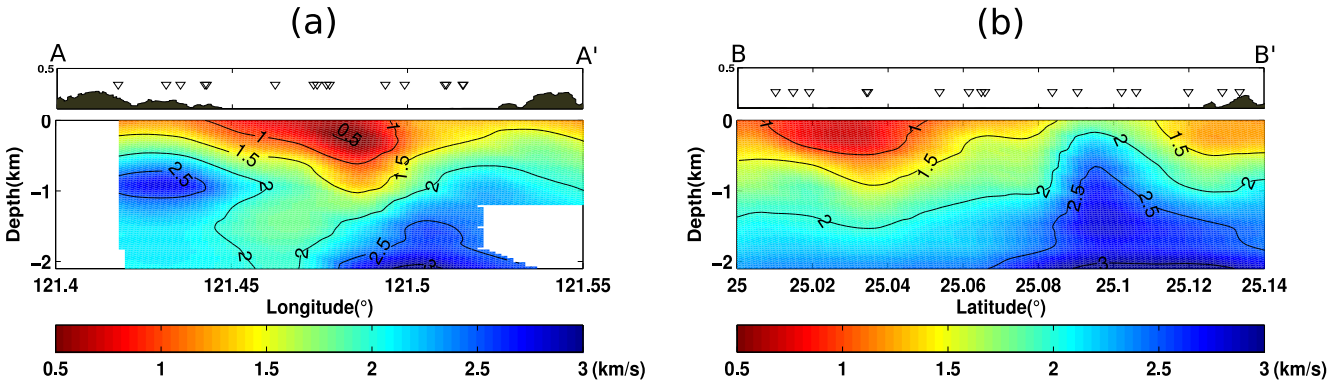

Figure 7. Shear wave speed model along the two vertical cross-sections of (a) A–A′; (b) B–B′. Locations of the cross-sections are shown in Figs 6(a) and (b). In each plot the top panel shows the topography along the cross-section and the bottom panel gives the shear wave speed (km s^{-1}) with the colour bar given at the bottom. Only zones of good resolution based on the model uncertainty results (see Fig. 10) are shown.

a 3-D model (e.g. Shapiro & Ritzwoller 2002; Yao *et al.* 2008). It is still computationally prohibitive for 3-D tomography problems but will become feasible with the increasing power of computers. Since our data set is small, we estimate uncertainty using a Monte Carlo error propagation technique (Aster *et al.* 2013). We first calculated the synthetic traveltimes using the inverted wave speed model, and then inverted the data many times each time adding 2 per cent Gaussian random noise. Finally, we calculated the standard deviation using the obtained wave speed models. We also used different initial models to take into account the under estimation of uncertainty due to the use of regularization term in regions with poor or even no data constraint. Figs 9 and 10 show the uncertainty (standard deviation) distribution for the final 3-D model. It shows very large uncertainties where the data coverage is poor, but in regions with good data constraints the uncertainty is, as expected, small.

We also compare the phase velocity maps between the two-step inversion (Huang *et al.* 2010) and direct inversion methods in Fig. 11. The phase velocity map for the direct inversion is constructed by computing the phase velocities for each grid point using the final 3-D model. The results show very similar features with low wave speed in basin regions and high wave speed in mountain regions, except in regions with poor data coverage where artefacts seem to appear in the two-step inversion.

5 CONCLUSION

We propose an approach for ray tracing based direct (non-linear) inversion of surface wave dispersion data (that is, without the intermediate step of phase or group velocity maps) for 3-D shear wave speeds using a wavelet-based sparsity-constrained tomographic technique. During the iterative inversion, ray paths at each period (and, thus, the sensitivity matrix in the inversion system) is updated using the newly obtained wave speed model at each iteration. This is particularly relevant for tomographic reconstructions of complex media, for instance, in the shallow crust, where great-circle propagation of surface waves is typically no longer valid. The wavelet-based sparsity-constrained inversion naturally helps to achieve multiresolution of the tomographic models with higher resolution in regions better constrained by data. As a proof of concept, we successfully applied our algorithm to obtain the shallow crustal shear wave speed structure in the Taipei Basin of Taiwan using the short period Rayleigh wave dispersion data from ambient noise cross-correlations. The large wave speed variations in the study region strongly affect wave propagation of short period surface waves. In the future, incorporation of quasi-3D sensitivity kernels of surface wave dispersion data and joint inversion with body wave traveltimes will be implemented to resolve P and S wave speed variations in complex media better.

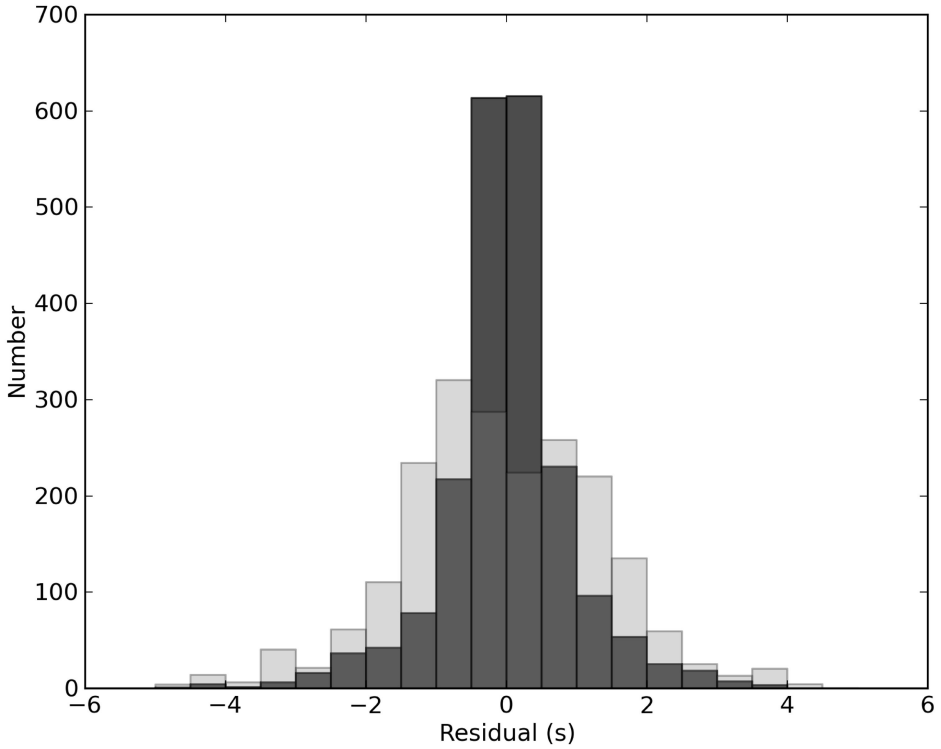

Figure 8. Distribution of surface wave traveltime residuals before (light grey, with the average $\mu = -0.07$ s and the standard deviation $\sigma = 1.44$ s) and after inversion (dark grey, with $\mu = 0.01$ s and $\sigma = 0.91$ s).

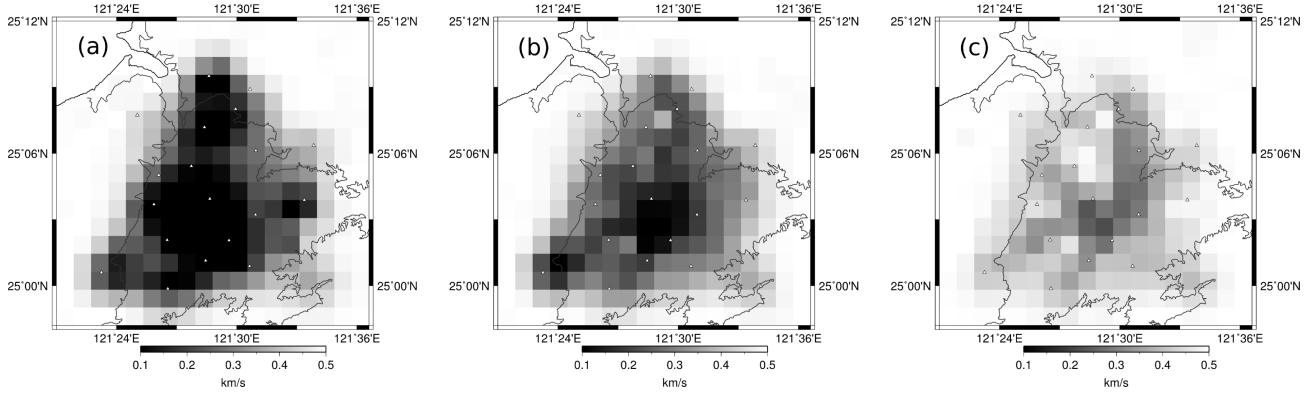

Figure 9. Estimated uncertainties of the wave speed model at depths of: (a) 0.3 km, (b) 0.9 km and (c) 1.5 km based on the Monte Carlo error propagation method (see texts for details).

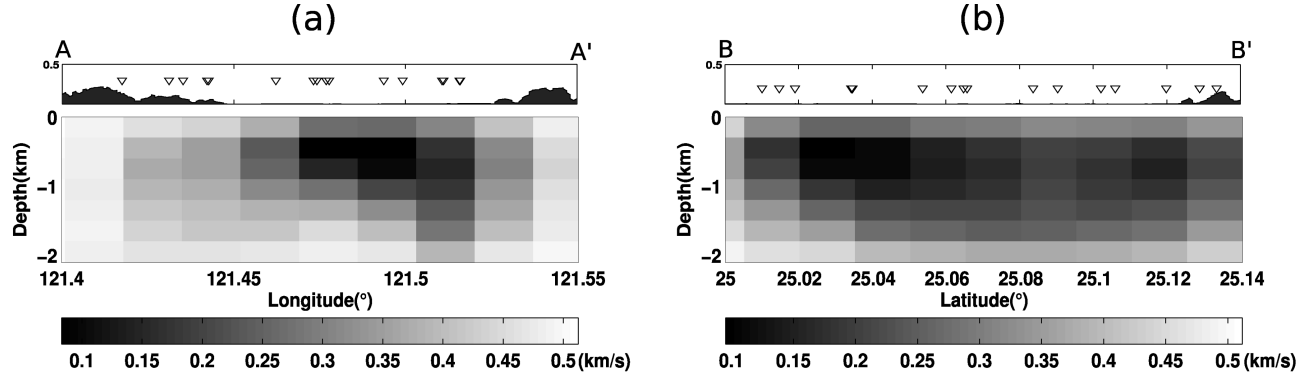

Figure 10. Estimated uncertainties of the wave speed model along two cross sections of (a) A–A′ and (b) B–B′ corresponding to Fig. 7.

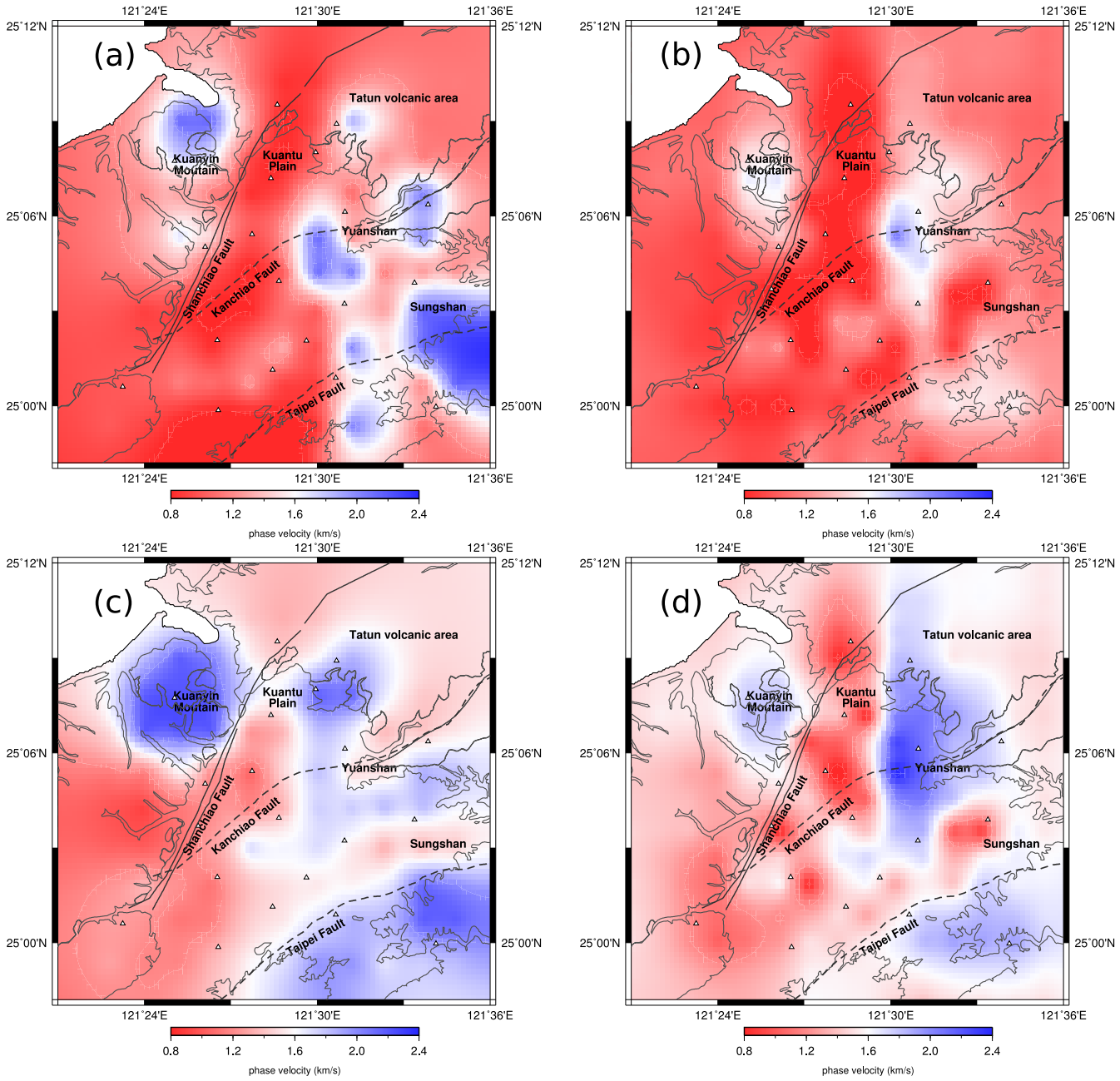

Figure 11. Comparison of phase velocity maps at 0.8 s and 2.0 s between the two step inversion (Huang *et al.* 2010) and direct inversion (this study) methods. (a) 0.8 s of two-step inversion, (b) 0.8 s of direct inversion, (c) 2.0 s of two-step inversion and (d) 2.0 s of direct inversion.

ACKNOWLEDGEMENTS

We would like to thank the editor and two anonymous reviewers who helped to improve the paper. We also thank Dr. Nick Rawlinson who provided us the fast marching method to compute the surface wave ray paths. This research is supported by the National Natural Science Foundation of China (41222028, 41474039), China National Special Fund for Earthquake Scientific Research in Public Interest (201508008) and the Fundamental Research Funds for the Central Universities (WK2080000053) and also by DE-NA0001523 from the U.S. Department of Energy.

REFERENCES

An, M., Feng, M. & Zhao, Y., 2009. Destruction of lithosphere within the north China craton inferred from surface wave tomography, *Geochem. Geophys. Geosyst.,* **10**(8), doi:10.1029/2009GC002562.

Aster, R.C., Borchers, B. & Thurber, C.H., 2013. *Parameter Estimation and Inverse Problems,* Academic Press.

Barmin, M., Ritzwoller, M. & Levshin, A., 2001. A fast and reliable method for surface wave tomography, in *Monitoring the Comprehensive Nuclear-Test-Ban Treaty: Surface Waves,* eds Levshin, A.L. & Ritzwoller, M., pp. 1351–1375, Springer-Verlag.

Boschi, L. & Ekström, G., 2002. New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies,

J. geophys. Res.: Solid Earth (1978–2012), **107**(B4), 2059, doi:10.1029/2000JB000059.ESE–1.

Brocher, T.M., 2005. Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. seism. Soc. Am.,* **95**(6), 2081–2092.

Cara, M. & Lévêque, J., 1987. Waveform inversion using secondary observables, *Geophys. Res. Lett.,* **14**(10), 1046–1049.

Charléty, J., Voronin, S., Nolet, G., Loris, I., Simons, F.J., Sigloch, K. & Daubechies, I.C., 2013. Global seismic tomography with sparsity constraints: comparison with smoothing and damping regularization, *J. geophys. Res.: Solid Earth,* **118**(9), 4887–4899.

Chen, M., Huang, H., Yao, H., Hilst, R. & Niu, F., 2014. Low wave speed zones in the crust beneath SE Tibet revealed by ambient noise adjoint tomography, *Geophys. Res. Lett.,* **41**(2), 334–340.

Chiao, L.-Y. & Liang, W.-T., 2003. Multiresolution parameterization for geophysical inverse problems, *Geophysics,* **68**(1), 199–209.

Daubechies, I., 1992. *Ten Lectures on Wavelets,* Vol. 61, SIAM.

Debayle, E. & Sambridge, M., 2004. Inversion of massive surface wave data sets: model construction and resolution assessment, *J. geophys. Res.: Solid Earth (1978–2012),* **109**(B2), doi:10.1029/2003JB002652.

Dunkin, J.W., 1965. Computation of modal solutions in layered, elastic media at high frequencies, *Bull. seism. Soc. Am.,* **55**(2), 335–358.

Ekström, G., Tromp, J. & Larson, E.W., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.: Solid Earth (1978–2012),* **102**(B4), 8137–8157.

Fang, H. & Zhang, H., 2014. Wavelet-based double-difference seismic tomography with sparsity regularization, *Geophys. J. Int.,* **199**(2), 944–955.

Feng, M. & An, M., 2010. Lithospheric structure of the Chinese mainland determined from joint inversion of regional and teleseismic Rayleigh-wave group velocities, *J. geophys. Res.: Solid Earth (1978–2012),* **115**(B6), doi:10.1029/2008JB005787.

Gao, H. & Shen, Y., 2014. Upper mantle structure of the Cascades from full-wave ambient noise tomography: evidence for 3D mantle upwelling in the back-arc, *Earth planet. Sci. Lett.,* **390,** 222–233.

Gouédard, P., Yao, H., Ernst, F. & van der Hilst, R.D., 2012. Surface wave eikonal tomography in heterogeneous media using exploration data, *Geophys. J. Int.,* **191**(2), 781–788.

Haskell, N.A., 1953. The dispersion of surface waves on multilayered media, *Bull. seism. Soc. Am.,* **43**(1), 17–34.

Herrmann, R.B., 2001. Computer programs in seismology-an overview of synthetic seismogram computation Version 3.1, Department of Earth and Planetary Sciences, St Louis Univ.

Huang, Y.-C., Yao, H., Huang, B.-S., van der Hilst, R.D., Wen, K.-L., Huang, W.-G. & Chen, C.-H., 2010. Phase velocity variation at periods of 0.5–3 seconds in the Taipei Basin of Taiwan from correlation of ambient seismic noise, *Bull. seism. Soc. Am.,* **100**(5A), 2250–2263.

Karason, H. & Van Der Hilst, R.D., 2000. Constraints on mantle convection from seismic tomography, in *The History and Dynamics of Global Plate Motions,* pp. 277–288, eds Richards, M.A., Gordon, R.G. & van der Hilst, R.D., John Wiley and Sons.

Liang, C. & Langston, C.A., 2009. Wave gradiometry for USArray: Rayleigh waves, *J. geophys. Res.: Solid Earth (1978–2012),* **114**(B2), doi:10.1029/2008JB005918.

Lin, F.-C. & Ritzwoller, M.H., 2011. Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure, *Geophys. J. Int.,* **186**(3), 1104–1120.

Lin, F.-C., Moschetti, M.P. & Ritzwoller, M.H., 2008. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.,* **173**(1), 281–298.

Lin, F.-C., Ritzwoller, M.H. & Snieder, R., 2009. Eikonal tomography: surface wave tomography by phase front tracking across a regional broadband seismic array, *Geophys. J. Int.,* **177**(3), 1091–1110.

Lin, F.-C., Li, D., Clayton, R.W. & Hollis, D., 2013. High-resolution 3D shallow crustal structure in Long Beach, California: application of ambient noise tomography on a dense seismic array, *Geophysics,* **78**(4), Q45–Q56.

Loris, I., Nolet, G., Daubechies, I. & Dahlen, F., 2007. Tomographic inversion using 1-norm regularization of wavelet coefficients, *Geophys. J. Int.,* **170**(1), 359–370.

Montagner, J.-P. & Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies, *J. geophys. Res.: Solid Earth (1978–2012),* **96**(B12), 20 337–20 351.

Nolet, G., 1990. Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs, *J. geophys. Res.: Solid Earth (1978–2012),* **95**(B6), 8499–8512.

Paige, C.C. & Saunders, M.A., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software (TOMS),* **8**(1), 43–71.

Picozzi, M., Parolai, S., Bindi, D. & Strollo, A., 2009. Characterization of shallow geology by high-frequency seismic noise tomography, *Geophys. J. Int.,* **176**(1), 164–174.

Pilz, M., Parolai, S., Picozzi, M. & Bindi, D., 2012. Three-dimensional shear wave velocity imaging by ambient seismic noise tomography, *Geophys. J. Int.,* **189**(1), 501–512.

Pollitz, F.F., 2008. Observations and interpretation of fundamental mode Rayleigh wavefields recorded by the Transportable Array (USArray), *J. geophys. Res.: Solid Earth (1978–2012),* **113**(B10), doi:10.1029/2007JB005556.

Pollitz, F. & Snoke, J.A., 2010. Rayleigh-wave phase-velocity maps and three-dimensional shear velocity structure of the western US from local non-plane surface wave tomography, *Geophys. J. Int.,* **180**(3), 1153–1169.

Prochazka, A., GráfováL. Vyšata, O. & Caregroup, N., 2011. Three-dimensional wavelet transform in multi-dimensional biomedical volume processing, in *Proceedings of the IASTED International Conference on Graphics and Virtual Reality,* Cambridge, UK.

Rawlinson, N. & Sambridge, M., 2004. Wave front evolution in strongly heterogeneous layered media using the fast marching method, *Geophys. J. Int.,* **156**(3), 631–647.

Ritzwoller, M.H., Shapiro, N.M., Levshin, A.L. & Leahy, G.M., 2001. Crustal and upper mantle structure beneath Antarctica and surrounding oceans, *J. geophys. Res.: Solid Earth (1978–2012),* **106**(B12), 30 645–30 670.

Ritzwoller, M.H., Shapiro, N.M., Barmin, M.P. & Levshin, A.L., 2002. Global surface wave diffraction tomography, *J. geophys. Res.: Solid Earth (1978–2012),* **107**(B12), 2335, doi:10.1029/2002JB001777.

Sabra, K.G., Gerstoft, P., Roux, P., Kuperman, W. & Fehler, M.C., 2005. Surface wave tomography from microseisms in Southern California, *Geophys. Res. Lett.,* **32**(14), doi:10.1029/2005GL023155.

Scales, J.A., Gersztenkorn, A. & Treitel, S., 1988. Fast Lp solution of large, sparse, linear systems: application to seismic travel time tomography, *J. Comput. Phys.,* **75**(2), 314–333.

Scherzer, O., 1993. The use of Morozov's discrepancy principle for Tikhonov regularization for solving nonlinear ill-posed problems, *Computing*, **51**(1), 45–60.

Shapiro, N. & Ritzwoller, M., 2002. Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle, *Geophys. J. Int.,* **151**(1), 88–105.

Shapiro, N.M., Ritzwoller, M.H., Molnar, P. & Levin, V., 2004. Thinning and flow of Tibetan crust constrained by seismic anisotropy, *Science,* **305**(5681), 233–236.

Shearer, P.M., 1997. Improving local earthquake locations using the L1 norm and waveform cross correlation: application to the Whittier Narrows, California, aftershock sequence, *J. geophys. Res.: Solid Earth (1978–2012),* **102**(B4), 8269–8283.

Shearer, P.M., 2009. *Introduction to Seismology,* Cambridge Univ. Press.

Shirzad, T. & Shomali, Z.H., 2014. Shallow crustal structures of the Tehran basin in Iran resolved by ambient noise tomography, *Geophys. J. Int.,* **196**(2), 1162–1176.

Simons, F.J., Zielhuis, A. & van der Hilst, R.D., 1999. The deep structure of the Australian continent from surface wave tomography, *Develop. Geotect.,* **24,** 17–43.

Simons, F.J., Van Der Hilst, R.D., Montagner, J.-P. & Zielhuis, A., 2002. Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle, *Geophys. J. Int.,* **151**(3), 738–754.

Simons, F.J. *et al.*, 2011. Solving or resolving global tomographic models with spherical wavelets, and the scale and sparsity of seismic heterogeneity, *Geophys. J. Int.,* **187**(2), 969–988.

Spakman, W. & Bijwaard, H., 2001. Optimization of cell parameterizations for tomographic inverse problems, *Pure appl. Geophys.,* **158**(8), 1401–1423.

Sweldens, W., 1996. The lifting scheme: a custom-design construction of biorthogonal wavelets, *Appl. Comput. Harmon. Anal.,* **3**(2), 186–200.

Tape, C., Liu, Q. & Tromp, J., 2007. Finite-frequency tomography using adjoint methods: methodology and examples using membrane surface waves, *Geophys. J. Int.,* **168**(3), 1105–1129.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.,* **180**(1), 433–462.

Thomson, W.T., 1950. Transmission of elastic waves through a stratified solid medium, *J. Appl. Phys.,* **21**(2), 89–93.

Trampert, J. & Woodhouse, J.H., 1995. Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds, *Geophys. J. Int.,* **122**(2), 675–690.

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.: Solid Earth (1978–2012),* **89**(B7), 5953–5986.

Yang, Y., Ritzwoller, M.H., Levshin, A.L. & Shapiro, N.M., 2007. Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.,* **168**(1), 259–274.

Yao, H., van Der Hilst, R.D. & Maarten, V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis—I. Phase velocity maps, *Geophys. J. Int.,* **166**(2), 732–744.

Yao, H., Beghein, C. & Van Der Hilst, R.D., 2008. Surface wave array tomography in SE Tibet from ambient seismic noise and two-station analysis—II. Crustal and upper-mantle structure, *Geophys. J. Int.,* **173**(1), 205–219.

Yao, H., van der Hilst, R.D. & Montagner, J.-P., 2010. Heterogeneity and anisotropy of the lithosphere of SE Tibet from surface wave array tomography, *J. geophys. Res.: Solid Earth (1978–2012),* **115**(B12), doi:10.1029/2009JB007142.

Yoshizawa, K. & Kennett, B., 2004. Multimode surface wave tomography for the Australian region using a three-stage approach incorporating finite frequency effects, *J. geophys. Res.: Solid Earth (1978–2012),* **109**(B2), doi:10.1029/2002JB002254.

Young, M., Rawlinson, N., Arroucau, P., Reading, A. & Tkalčić, H., 2011. High-frequency ambient noise tomography of southeast Australia: new constraints on Tasmania's tectonic past, *Geophys. Res. Lett.,* **38**(13), doi:10.1029/2011GL047971.

Zhang, H. & Thurber, C., 2005. Adaptive mesh seismic tomography based on tetrahedral and Voronoi diagrams: application to Parkfield, California, *J. geophys. Res.: Solid Earth (1978–2012),* **110**(B4), doi:10.1029/2004JB003186.

Zhang, W., Shen, Y. & Chen, X., 2008. Numerical simulation of strong ground motion for the *M*s 8. 0 Wenchuan earthquake of 12 May 2008, *Sci. China Series D: Earth Sci.,* **51**(12), 1673–1682.

Zhang, H., Maceira, M., Roux, P. & Thurber, C., 2014. Joint inversion of body-wave arrival times and surface-wave dispersion for three-dimensional seismic structure around SAFOD, *Pure appl. Geophys.,* **171**(11), 3013–3022.

Zhou, Y., Dahlen, F. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.,* **158**(1), 142–168.

Zhou, Y., Nolet, G., Dahlen, F. & Laske, G., 2006. Global upper-mantle structure from finite-frequency surface-wave tomography, *J. geophys. Res.: Solid Earth (1978–2012),* **111**(B4), doi:10.1029/2005JB003677.

Zhu, H., Bozdağ, E., Peter, D. & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography, *Nat. Geosci.,* **5**(7), 493–498.