Pure Appl. Geophys. 174 (2017), 937–953

DOI 10.1007/s00024-016-1434-5

Pure and Applied Geophysics

Surface Wave Tomography with Spatially Varying Smoothing Based on Continuous Model Regionalization

Chuanming Liu[1,2] and Huajian Yao[1,2]

Abstract—Surface wave tomography based on continuous regionalization of model parameters is widely used to invert for 2-D phase or group velocity maps. An inevitable problem is that the distribution of ray paths is far from homogeneous due to the spatially uneven distribution of stations and seismic events, which often affects the spatial resolution of the tomographic model. We present an improved tomographic method with a spatially varying smoothing scheme that is based on the continuous regionalization approach. The smoothness of the inverted model is constrained by the Gaussian a priori model covariance function with spatially varying correlation lengths based on ray path density. In addition, a two-step inversion procedure is used to suppress the effects of data outliers on tomographic models. Both synthetic and real data are used to evaluate this newly developed tomographic algorithm. In the synthetic tests, when the contrived model has different scales of anomalies but with uneven ray path distribution, we compare the performance of our spatially varying smoothing method with the traditional inversion method, and show that the new method is capable of improving the recovery in regions of dense ray sampling. For real data applications, the resulting phase velocity maps of Rayleigh waves in SE Tibet produced using the spatially varying smoothing method show similar features to the results with the traditional method. However, the new results contain more detailed structures and appears to better resolve the amplitude of anomalies. From both synthetic and real data tests we demonstrate that our new approach is useful to achieve spatially varying resolution when used in regions with heterogeneous ray path distribution.

Key words: Surface wave tomography, continuous regionalization, spatially varying smoothing, correlation length.

1. Introduction

Surface wave tomography based on dispersion measurement from earthquake waveforms or ambient noise cross-correlations is an effective tool to study the structure of crust and upper mantle on both regional and global scales (e.g., Montagner and Tanimoto 1991; Ritzwoller et al. 2001; Shapiro et al. 2004; Debayle et al. 2005; Yang et al. 2007; Lin et al. 2008; Yao et al. 2010).

The classical surface wave tomography from dispersion data is usually performed in two stages. The first stage involves 2-D regionalization in which 2-D period-dependent phase or group maps are constructed based on the ray theory (e.g., Montagner 1986; Ekström et al. 1997; Barmin et al. 2001), 2-D finite frequency sensitivity kernels (e.g., Ritzwoller et al. 2002; Yoshizawa and Kennett 2004), or Eikonal or Helmholtz equations in regions with dense station distribution (Lin et al. 2009; Pollitz and Snoke 2010; Lin and Ritzwoller 2011). In the second stage, at each geographical location, the pure path dispersion curve is inverted to obtain a local 1-D shear velocity model, which together forms the fully 3-D shear wave speed model (e.g., Shapiro and Ritzwoller 2002; Yao et al. 2008). The dispersion data from available paths can be also directly inverted for 3-D shear velocity variations without construction of phase or group velocity maps (e.g., Boschi and Ekström 2002; An et al. 2009; Fang et al. 2015).

A number of traveltime tomographic methods based on the ray theory have been developed to invert surface-wave dispersion measurements on regional or global scales for 2-D isotropic and azimuthally anisotropic velocity maps, which differ in aspects such as geometry, model parameterization, and

[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, Anhui, China. E-mail: hjyao@ustc.edu.cn

[2] National Geophysical Observatory at Mengcheng, Mengcheng, Anhui, China.

regularization schemes. For example, based on parameterization of integral kernels, the 2-D Backus–Gilbert approach using the first-spatial gradient smoothness constraints (Ditmar and Yanovskaya 1987; Yanovskaya and Ditmar 1990) has been extensively used in regional 2D phase or group velocity tomography (e.g., Levshin et al. 1989; Ritzwoller and Levshin 1998). The tomographic technique presented by Barmin et al. (2001) based on minimizing a penalty function composed of data misfit, model smoothness, and the path coverage is widely used in regional or global scale applications (e.g., Yang et al. 2007; Lin et al. 2008). Surface wave tomography with continuous regionalization based on the Bayesian Theorem (Montagner 1986), which is derived from the continuous form proposed by Tarantola and Nercessian (1984), is also widely used to constrain phase velocity variations and azimuthal anisotropy (e.g., Silveira and Stutzmann 2002; Debayle and Sambridge 2004; Yao et al. 2005, 2010). The approach proposed by Debayle and Sambridge (2004) has dramatically increased the computational efficiency of the original method by Montagner (1986) with incorporation of some sophisticated geometrical algorithms.

For surface-wave tomographic problems, an inevitable issue is that the uneven distribution of ray paths often leads to the relatively ill-conditioned inverse problem. Consequently, damping and model regularization are introduced to stabilize the ill-posed inverse system, which usually sacrifices some model resolution. To overcome this problem, a useful way is to change the model parameterization scheme. One explicit strategy is irregular parameterization based on rectangles (or squares) with adaptive grid spacing (e.g., Abers and Roecker 1991; Spakman and Bijwaard 2001; Simons et al. 2002) or Delaunay triangles (or Voronoi diagram) with flexible shape and size (e.g. Sambridge and Gumundsson 1998; Böhm et al. 2000; Debayle and Sambridge 2004; Zhang and Thurber 2005). Using this kind of irregular parameterization, the irregular grids/blocks can match with the non-uniform ray path distribution, reduce the number of free parameters, and improve stability of traditional tomographic methods. Another popular strategy is the wavelet-based multi-resolution parameterization, which has been applied to compensate for the mismatch correlation between uneven ray path distribution and regular grids to resolve the model at different scales using the inherent multiscale nature of the wavelet transform in the spatial domain (e.g., Chiao and Kuo 2001; Chiao and Liang 2003; Loris et al. 2007; Delost et al. 2008; Hung et al. 2011; Simons et al. 2011; Fang and Zhang 2014). In addition, an adaptive parameterization method, the Bayesian trans-dimensional tomography, in which the number of unknowns is an unknown itself, has been shown its feasibility in most 2-D and some 3-D problems (e.g., Bodin and Sambridge 2009; Hawkins and Sambridge 2015; Saygin et al. 2016). And this method can give a quantitative access to the reliability of the solution model.

Besides changing the model parameterization, another way to fulfill tomography with spatial varying resolution is to allow for spatially variable priors or smoothing constraints in the model regularization. In this paper, following the continuous regionalization inversion method with a least squares criterion proposed by Tarantola and Valette (1982), we modify the fixed correlation length in the Gaussian a priori covariance function to be a spatial function of ray path density. In this way, smoothness of the inverted model in regions with different ray path coverage is constrained by the a priori covariance function with spatially varying correlation lengths. We use both synthetic and real data to evaluate the performance of our algorithm.

2. *Methodology*

The continuous regionalization inversion approach takes the function of a continuous variable itself as the unknown where the theoretical relationship between data and unknowns is assumed to be linear (Tarantola and Valette 1982). This method has been applied to 3-D seismic velocity tomography using arrival times of body waves (Tarantola and Nercessian 1984) and surface waves (Montagner 1986; Yao et al. 2005). In this section, we describe the inversion method for regional surface wave tomography used in Montagner (1986), and then present the spatially varying smoothing scheme on this basis.

2.1. Spatially Varying Smoothing Based on Continuous Regionalization

The surface wave tomography method (Montagner 1986) based on continuous regionalization of model parameters (Tarantola and Nercessian 1984) is widely used to invert for 2-D surface wave phase and group velocity maps for its clear physical meaning when adjusting inversion parameters. With the least squares criterion, the inversion results not only have the best data fitting but also can keep as close as possible to the a priori model. When the data errors exhibit a Gaussian distribution, the objective function for the generalized inverse problem (Tarantola and Valette 1982) is expressed as:

$$\begin{aligned}\phi(\mathbf{m}) = &(\mathbf{d}_0 - \mathbf{d})^{\mathrm{T}}\mathbf{C}_{d_0}^{-1}(\mathbf{d}_0 - \mathbf{d}) \\ &+ (\mathbf{m}_0 - \mathbf{m})^{\mathrm{T}}\mathbf{C}_{m_0}^{-1}(\mathbf{m}_0 - \mathbf{m}),\end{aligned} \tag{1}$$

where $\mathbf{d}_0$ is a vector of the observed arrival time data at a fixed period, $\mathbf{d}$ is a vector of the arrival time data predicted from the actual slowness model $\mathbf{m}$ through a forward equation of the form $\mathbf{d} = g(\mathbf{m})$, $\mathbf{m}_0$ is the a priori slowness model, $\mathbf{C}_{d_0}$ represents the a priori data covariance matrix, and $\mathbf{C}_{m_0}$ is the a priori model covariance matrix for the a priori model $\mathbf{m}_0$.

A classical least squares solution of Eq. (1) could be obtained (Tarantola and Valette 1982) as:

$$\mathbf{m} = \mathbf{m}_0 + \mathbf{C}_{m_0}\mathbf{G}^{\mathrm{T}}(\mathbf{G}\mathbf{C}_{m_0}\mathbf{G}^{\mathrm{T}} + \mathbf{C}_{d_0})^{-1}(\mathbf{d}_0 - \mathbf{G}\mathbf{m}_0). \tag{2}$$

In the continuous form, $\mathbf{G}$ stands for the sensitivity along the ray paths, which can be represented by integrals along the ray paths, the unknowns are the functions of a continuous variable r (spatial location on the spherical surface here) and the solution (Tarantola and Nercessian 1984) can be written as

$$m(r) = m_0(r) + \sum_i W_i \int_{R_i(m)} \mathrm{d}s_i C_{m_0}(r, r_i), \tag{3a}$$

$$W_i = \sum_j (S_{ij}^{-1}) V_j, \tag{3b}$$

$$S_{ij} = (C_{d_0})_{ij} + \int_{R_i(m)} \mathrm{d}s_i \int_{R_j(m)} \mathrm{d}s_j C_{m_0}(r_j, r_i), \tag{3c}$$

$$V_j = d_{0j} - \int_{R_j(m)} \mathrm{d}s_j m_0(r_j). \tag{3d}$$

Here, i and j are the path indexes and the integral path $R_i(m)$ is assumed to be along the ith great circle path for surface wave propagation. It is usually assumed that phase or group velocity measurements are independent for each path. In this case, $\boldsymbol{C}_{d_0}$ is a diagonal matrix and the diagonal term represents the square of the estimated data error (σ_d^i) (i.e., variance) in the ith measurement.

In generally, the model a priori covariance function describes the confidence in the a priori model $\mathbf{m}_0$ and controls the correlation between neighboring grids. The Gaussian a priori covariance function $C_{m_0}(r, r')$ used by Tarantola and Nercessian (1984) and Montagner (1986) represents the covariance between points r and r' with the analytical form of:

$$C_{m_0}(r, r') = \sigma_m(r)\sigma_m(r') \exp\left(\frac{-\Delta_{r,r'}^2}{2L_{\mathrm{corr}}^2}\right), \tag{4}$$

where $\Delta_{r,r'}$ is the distance between grid points r and r', the $\sigma_m(r)$ represents the a priori model uncertainty at each grid point that controls the amplitude perturbation of the model parameters, and L_{corr} represents the correlation length of the model parameters that acts as the spatial smoothing filter in the model space. For many tomographic problems, we do not know the exact model prior $\sigma_m(r)$. However, we could roughly estimate $\sigma_m(r)$ based on the data, here phase (or group) velocities measured at each ray paths, which is commonly referred as the "Empirical Bayesian" approach (e.g., Gelman et al. 2004). In general, we set $\sigma_m(r)$ as about twice that of the standard deviation of all observed phase or group velocities at each period.

The correlation length L_{corr} is a fixed value for each period in the previous continuous regionalization approach (Montagner 1986; Yao et al. 2005), and its value mainly depends on the spatial and azimuthal coverage of ray paths and the wavelength of surface wave. Because of the fixed L_{corr}, the size of heterogeneity that the inversion can retrieve is greater

or equal to $2L_{\text{corr}}$, approximately. Hence, to avoid small-scale inversion artifacts and get reliable results, L_{corr} cannot be too small, thus tends to inevitably smear out sharp boundaries and fine scale features in regions with dense ray path coverage.

To solve the problem and improve the local resolution in regions with enough ray path coverage, we change the fixed correlation length L_{corr} into spatially varying L_{corr} as a function of local ray path density, namely, assigning smaller values to L_{corr} in regions with higher path density so as to achieve spatially varying smoothing for the 2-D velocity model. And the new a priori model covariance function can be written as:

$$C_{m_0}(r, r') = \sigma_m(r)\sigma_m(r') \exp\left(\frac{-\Delta^2_{r,r'}}{2L_{\text{corr}}(p(r))L_{\text{corr}}(p(r'))}\right), \tag{5}$$

where the correlation length $L_{\text{corr}}(p(r))$ is a spatially varying function proportional to the ray path density $p(r)$ at the grid point r. This idea is similar to inversion of the 1-D shear velocity model from dispersion data, where smaller correlation lengths are chosen for model parameters at shallower depths while larger correlation lengths are given at greater depths, considering the decreased sensitivity of dispersion data to shear velocities at greater depths (e.g., Yao et al. 2010).

For the construction of the $L_{\text{corr}}(p(r))$, we first perform a series of checkerboard tests with the descending anomaly size to assess local resolution in different regions and find the corresponding appropriate correlation length. In this way, we determine the minimum anomaly size that we can recover and the corresponding appropriate correlation length that should be used. Then, in this study, we set the $L_{\text{corr}}(p(r))$ varying linearly from the minimum correlation length $L_{\min}$ determined in the study region associated with the largest ray path density to some appropriate maximum value $L_{\max}$ determined in regions with small ray path density, although other functional forms between the ray path density and the correlation length can be chosen. Besides, we will illustrate this process in detail in the synthetic test part.

2.2. Two-Step Inversion for Suppressing Data Outliers

The standard error in surface wave dispersion measurements is typically $\sim 1\%$. However, in practice, there may exist bad measurements or even data outliers, which to some extent result in artifacts in the tomographic models. To suppress the effects of bad measurements and data outliers, we propose a two-step inversion scheme in this study. First, we perform the inversion with the diagonal terms of the a priori data covariance matrix $\boldsymbol{C}_{d_0}$ set as the square of estimated measurement errors or 1% of the observed data for paths without confident error estimates. In the second step inversion, we calculate the data misfits (differences between the observed and predicted velocities) from the first inversion result, and if the relative data misfits of the corresponding paths are more than twice of the standard error of all data misfits, we increase the original data uncertainty σ_d^i exponentially as the new data uncertainty $\hat{\sigma}_d^i$ in $\boldsymbol{C}_{d_0}$ following the expression:

$$\{\hat{\sigma}_d^i\}^2 = \{\sigma_d^i\}^2 \exp\left\{\left(\frac{\epsilon^i}{2\bar{\sigma}} - 1\right)\right\}, \tag{6}$$

where ϵ^i is the absolute value of the corresponding misfit of the ith path, and $\bar{\sigma}$ is the standard error of all data misfits. Here in practice, we save the matrix of the double-integration (the most time-consuming computational part) in (3c) in the first step inversion and update the S_{ij} in (3c) in the second step inversion using (6), which is very simple to implement in the program. In the application and discussion, we use an example to demonstrate the effectiveness of this two-step approach.

3. Synthetic Data Tests

Actual spatial resolution of surface wave tomography is always a combined effect of geometrical constraint and physical limitations (Debayle and Sambridge 2004). The geometric resolution is mainly constrained by the ray path coverage and azimuthal distribution, and physical resolution is dominated by

the wavelength. In general, the upper limit of spatial resolution is determined by the physical resolution, that is, the Fresnel zone width of the particular wave type, under the assumption of ray theory with single scatting approximation. This is also the main reason restricting the lateral resolution of global surface wave tomography (Spetzler and Snieder 2001), in which the period range is usually between 40 and 300 s and the wavelength range is about 160–1500 km. In the following synthetic tests, we use the Rayleigh wave phase velocity dispersion data at the period of 20 s, at which the wavelength is about 60–80 km, and we only take the geometric resolution into account.

The model a priori covariance function of the traditional method is generally in a Gaussian form (Eq. (4)). When the distance between the two points at r and r' is set as the size of checkerboard anomalies, the spatial correlation $R(D, L_{\mathrm{corr}})$ of these two points can be written as

$$R(D, L_{\mathrm{corr}}) = \exp\left(-\frac{D^2}{2L_{\mathrm{corr}}^2}\right). \tag{7}$$

Here, D is the anomaly size, which means the target horizontal resolution scale or checkerboard pattern size. When the correlation length L_{corr} is set equal to D, the correlation $R(D, L_{\mathrm{corr}})$ is about 0.6, which implies that these two points are strongly correlated and it is difficult to resolve the anomaly pattern in this case. And when the correlation length L_{corr} is set as $\frac{D}{\sqrt{2}}$, the correlation $R(D, L_{\mathrm{corr}})$ is about 0.37 which is an intermediate correlation value. If the correlation length L_{corr} is set to $\frac{D}{2}$, the correlation $R(D, L_{\mathrm{corr}})$ is about 0.1, which implies very weak correlation between these two points, and the anomaly pattern can be recovered well in regions with enough ray path coverage, or artifacts will likely be produced in regions with sparse path coverage since the inversion may get ill-posed in this situation.

To compare the performance of our spatially varying smoothing algorithm with traditional algorithm and show the process of how to perform spatially varying smoothing algorithm, we use the traditional checkerboard tests although this method has some drawbacks (Lévêque et al. 1993; Rawlinson and Spakman 2016). The distribution of ray paths used in the inversion is the same as that at the period of 20 s in SE Tibet as in Yao et al. (2010) for interstation Rayleigh wave phase velocity tomography. Figure 1e presents the path coverage, and we can construct the corresponding ray path density ($p(r)$) map (Fig. 1f). It is obvious that the ray path density is larger in the Lehigh array area and smaller in the MIT array area due to denser station distribution in the Lehigh array area (Fig. 1). This kind of heterogeneous distribution of path density is ideal to evaluate the performance of our algorithm. For convenience, the Lehigh array area is denoted as block A and the MIT array area is denoted as block B hereinafter.

We first create two different checkerboard patterns with anomaly size of 0.4° × 0.4° and 1° × 1° (Figs. 2a, 3a) and assess the local horizontal resolution by evaluating how well the pattern can be retrieved in different regions with different ray path density distributions (Fig. 1f) of the data set. Synthetic phase velocity dispersion data are generated with these input patterns and we also add 1% Gaussian random errors to the synthetic data. When the anomaly size is set as 1° × 1° (Fig. 2a), the most appropriate fixed L_{corr} is 40 km with the least model difference between the input and recovered models (Fig. 4a). With this correlation length, the anomaly pattern is almost correctly reconstructed for most of the study area (Fig. 2d), except for the central region between the two arrays with non-uniform azimuthal ray path coverage (Fig. 1f). If we choose a larger correlation length of 70 km (i.e., $L_{\mathrm{corr}} \sim \frac{D}{\sqrt{2}}$), the recovered pattern in block B is pretty smooth and anomaly amplitudes in block A with larger ray path density are less recovered than results with the correlation length of 40 km. And when we choose the correlation length of 100 km (i.e. $L_{\mathrm{corr}} \sim D$), the results of the whole study region get smeared, because the correlation length is similar to the anomaly size that causes strong spatial correlation between model parameters.

Then as for the 0.4° × 0.4° anomaly size, the corresponding fixed correlation length with the least model difference is about 20 km (i.e., $L_{\mathrm{corr}} \sim \frac{D}{2}$) (Fig. 4a), and the anomaly pattern is retrieved in block A with good resolution, while artificial anomalies appear in block B (Fig. 3b) due to relatively sparse coverage of ray paths and the choice of small L_{corr} (Fig. 1f). And as we choose larger

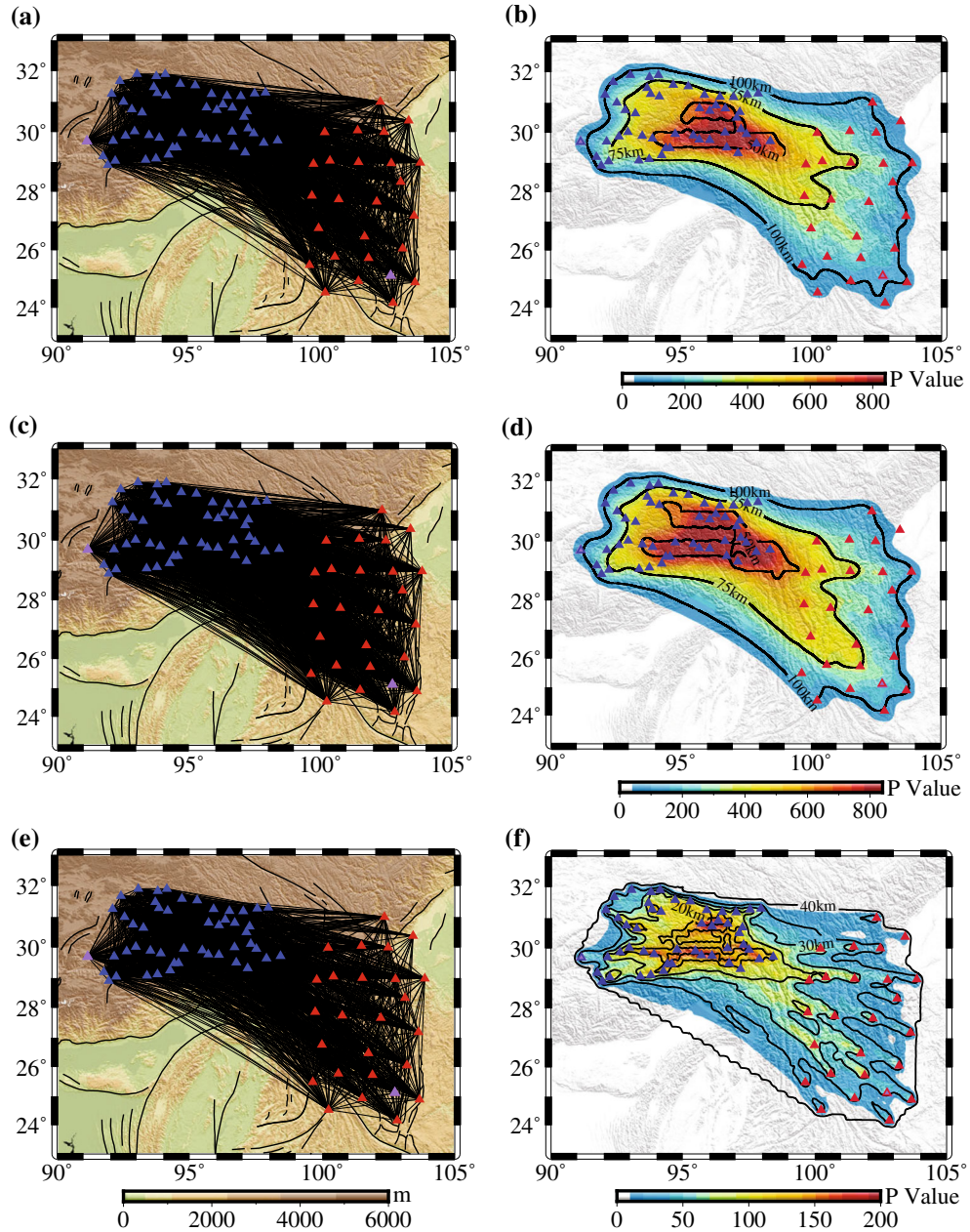

Figure 1

Ray path coverage of Rayleigh-wave phase velocity measurements at different periods (10, 30, 20 s) (**a**, **c**, **e**) (from Yao et al. 2010) and corresponding ray path density (*P* value) maps (**b**, **d**, **f**; the background image with the *colorbar* for *P* values) with spatially varying correlation lengths (*black contours*). The path density value and spatially varying correlation length maps are calculated based on 0.5° × 0.5° grids for (**b**, **d**) and 0.2° × 0.2° grid for (**f**). In Fig. 1, the *red* and *blue triangles* are the temporary stations deployed by MIT and Lehigh University, respectively; the two permanent stations (*dark purple triangles*) are located at Kunming (KMI) and Lhasa (LSA), China

correlation lengths (e.g., $L_{\text{corr}} = 25\,\text{km}, L_{\text{corr}} = 40\,\text{km}$), the amplitude of anomaly in block A is less recovered and the artificial anomalies in block B get smoother (Fig. 3c, d).

So, for ray path coverage of the data set we use, we can well retrieve the 1° × 1° anomalies pattern for almost the entire study region but can only retrieve the 0.4° × 0.4° anomalies in block A with

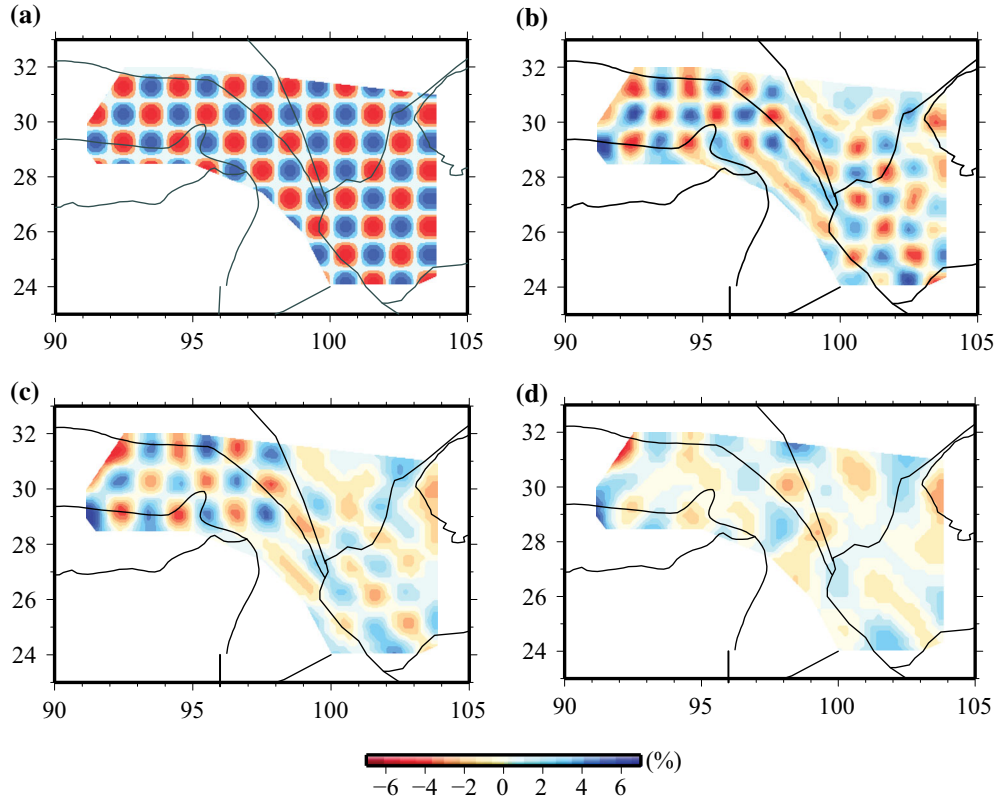

Figure 2
Checkerboard tests for the ray path coverage at 20 s with different anomaly sizes: **a** the input 1° × 1° checkerboard model; **b–d** recovery of the 1° × 1° checkerboard model using the traditional method with the fixed correlation length of 40, 70, and 100 km, respectively. The corresponding path coverage and ray path density maps are shown in Fig. 1e, f. *The black lines* show the block boundaries in this region. *The colorbar* shows the amplitude of the phase velocity anomaly in percent

enough path density, when we choose the appropriate correlation length (i.e., $L_{corr} = 40$ km for 1° × 1° pattern; $L_{corr} = 20$ km for 0.4° × 0.4° pattern). This implies that the horizontal resolution in block A can reach about 0.4° and only about 1° in block B. Hence, to avoid spurious tomographic results, we will typically choose L_{corr} to be 40 km or larger for the traditional approach if only considering the geometric resolution. This will certainly sacrifice the resolution in the area with dense ray path coverage, such as block A here.

To check the resolvability of the spatially varying smoothing method, we compare the inversion results of the new and traditional algorithms using a new 2-D synthetic model. We set this new synthetic model with 0.4° × 0.4° checkerboard anomalies in block A and 1° × 1° checkerboard anomalies in bock B (Fig. 5a), and then generate the synthetic phase velocity data with 1 percent Gaussian random noise. For the traditional approach, the correlation length L_{corr} is set to be a series of values from 5 to 70 km. By contrast, for our spatially varying smoothing algorithm, since we have known the appropriate correlation lengths to achieve expected resolution in block A and block B are 20 and 40 km correspondingly, we let $L_{corr}(p(r))$ vary linearly with $p(r)$ from L_{min} (20 km) to L_{max} (40 km) (see the corresponding contour map shown in Fig. 1f) to take full advantage of the information of path coverage. Then, we conduct the traditional inversion and the spatially varying smoothing inversion with the same synthetic data and a priori model and the results are shown in Fig. 5.

Then, we calculate the standard deviations of model differences for the results of the traditional and

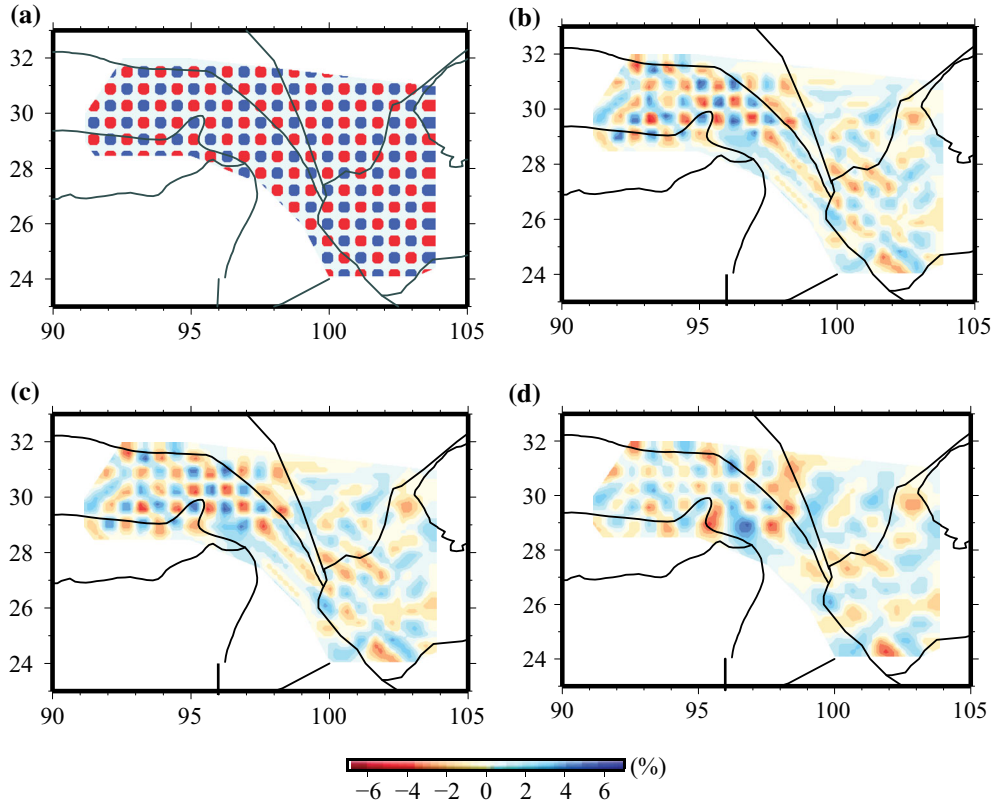

Figure 3
Similar as Fig. 2 but for **a** the input 0.4° × 0.4° checkerboard model and **b–d** recovery of the 0.4° × 0.4° checkerboard model using the traditional method with the fixed correlation length of 20, 25, and 40 km, respectively

new methods in block A, block B, and the whole study region, respectively. For the traditional method, the inversion results can achieve the least model differences in block A, block B, and the whole study region accordingly, when the corresponding correlation lengths are set to be 20, 40, and 30 km, respectively (Fig. 4b).

For the traditional inversion with a fixed correlation length of 20 km, it can recover most of the 0.4° × 0.4° anomalies (Fig. 5b) with the least model difference in block A (Fig. 4b), while artificial anomalies appear in block B since the very small correlation length will result in relatively weak regularization of model parameters and thus unstable inversion results in regions with sparse path coverage. When we set the correlation length as 40 km for the traditional method, it can recover the 1° × 1° anomalies almost perfectly (Fig. 5c) with the least model difference in block B (Fig. 4b), but the 0.4° × 0.4° anomalies get smeared (Fig. 5c) because the correlation length is similar as the anomaly size. We can get the recovered model with the least model difference in the whole study region when the correlation length is set to be 30 km (Fig. 5d), which somewhat compromises the resolution in both block A and B. However, the retrieved amplitude of anomalies in block A (Fig. 5d) is worse than the result with the correlation length of 20 km (Fig. 5b) and the recovery of the 1° × 1° anomalies (Fig. 5d) is less constrained than the result with the correlation length of 40 km in block B (Fig. 5c).

By contrast, the retrieved model of the spatially varying smoothing method can resolve the 1° × 1° anomalies properly in block B and have preferable resolution for the 0.4° × 0.4° anomalies with better recovery of amplitude of anomalies (Fig. 5e) than

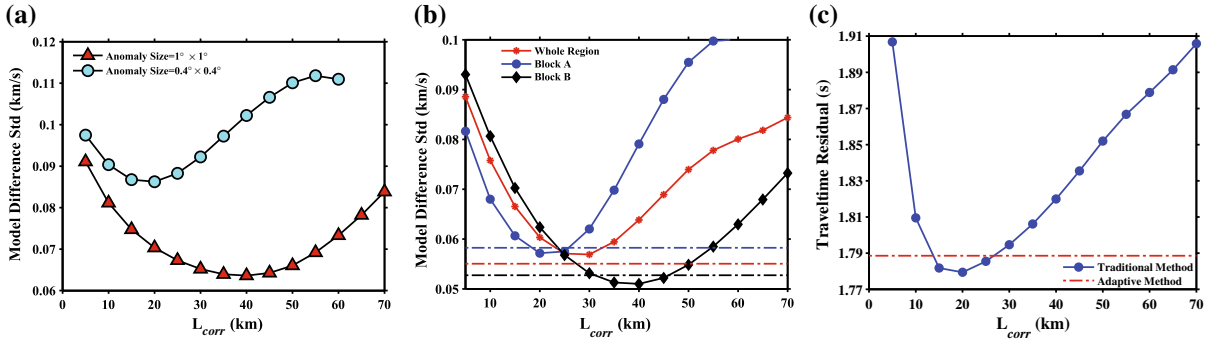

Figure 4

The standard deviation of the phase velocity model differences and data residuals: **a** the standard deviation of the model differences between the input and the recovered checkerboard models with the anomaly size 0.4° × 0.4° (*circle*) and 1° × 1° (*triangle*) using different correlation lengths in the traditional tomography. **b** The standard deviation of the model differences between the input and recovered checkerboard models with the checkerboard model of mixed anomaly sizes in block A (*blue*), block B (*black*), and the whole study region (*red*). *The solid lines* with symbols of *dot* (*blue*), *diamond* (*black*) and *asterisk* (*red*) represent the model differences of the traditional method results using different correlation lengths in block A, block B, and the whole study region, respectively, and *the dashed lines* indicates the model differences of the new method results with the line color for the region same as the traditional method. **c** The standard deviation of the traveltime residuals between the input (observed) and synthetic traveltime data with the checkerboard model of mixed anomaly sizes using the new method (*red dashed line*) and the traditional method (*blue dotted line*) with different correlation lengths

that of the traditional method with the correlation length of 30 km (Fig. 5d). From the perspective of model difference and data fitting, the result of the spatially varying smoothing method achieves a small model difference in block A close to the result with the fixed correlation of 20 km and a relative close model difference in block B with the result using the fixed correlation of 40 km (Fig. 4b). At the same time, it gets the least model difference in the whole study area than all the results of the traditional method with good data fitting (Fig. 4b, c).

Tarantola and Nercessian (1984) can obtain the posterior model covariance for continuous regionalization, which gives meaningful uncertainty estimates of model parameters. The posterior model errors in the results of the traditional method and spatially varying smoothing method (Fig. 6) are calculated from the posterior covariance matrix. The use of large correlation lengths reduces ill-conditioning problems and results in smooth but reliable results with smaller posterior errors (Fig. 6b). Compared with the results of the traditional method with correlation lengths of 20 and 30 km, the spatially varying smoothing algorithm can recover the 1° × 1° pattern better with smaller posterior errors. And the spatially varying smoothing algorithm significantly improves the amplitude recovery of the 0.4° × 0.4° anomalies in block A with relatively small posterior errors by contrast with the traditional method with the correlation length of 40 km. Through assessing the model posterior errors, we can ensure that the increased resolution of the spatially varying smoothing method is not at the cost of larger model uncertainties.

4. Application and Discussion

4.1. Application to Data in SE Tibet

As a proof of concept, we apply our method using the same dispersion data set as in Yao et al. (2010) to obtain Rayleigh-wave phase velocity variations at periods 10 and 30 s. The ray path distribution (see Fig. 1a, c) for these two different periods is excellent, and the corresponding path density $p(r)$ maps are shown in Fig. 1b, d using 0.5° grid spacing. The study region is also meshed with a grid separation of 0.5° for the inversion. For the traditional inversion approach the correlation length L_{corr} is set to be 110 km at 10 and 30 s similar as in Yao et al. (2010), and for the spatially varying smoothing inversion method, $L_{\mathrm{corr}}(p(r))$ is varying from 50 to 110 km linearly based on path density (see Fig. 1b, d).

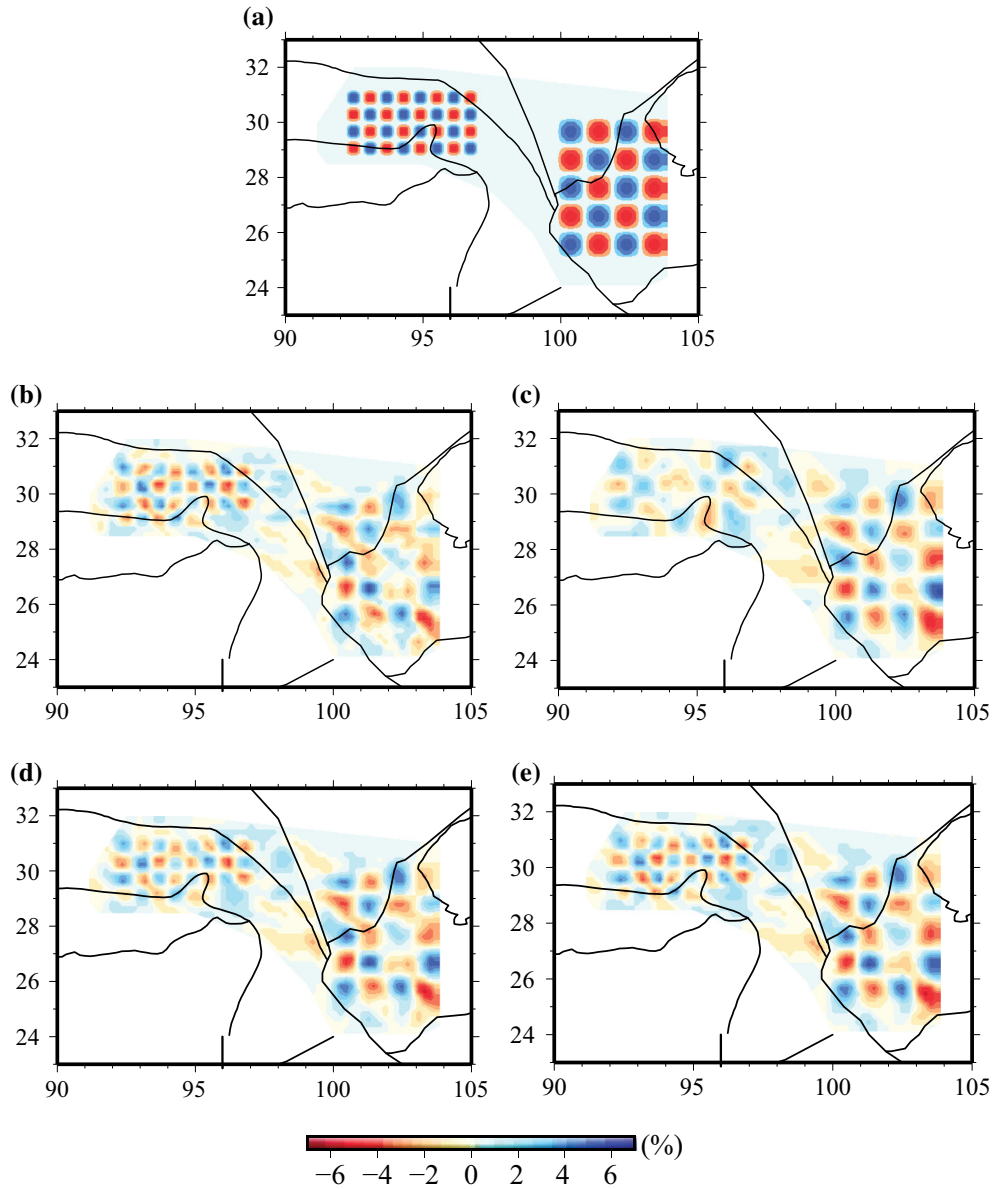

Figure 5
Checkerboard tests for the ray path coverage at 20 s with the checkerboard model of mixed anomaly sizes: **a** the checkerboard model with the mixed 0.4° × 0.4° and 1° × 1° anomalies; **b–d** the recovery using the traditional method with the fixed correlation length of 20, 30, and 40 km, respectively; **e** the recovery using the new method. The corresponding path coverage and ray path density map are shown in Fig. 1e, f

Compared with the phase velocity maps derived from the traditional inversion and spatially varying smoothing inversion method, the results show very similar features with a good correspondence to the geological features, but the spatially varying smoothing inversion results show more details and relatively larger velocity variations (Fig. 7). Generally speaking, the fundamental mode Rayleigh wave phase velocity is mostly sensitive to shear wave speed at depths around 1/3 of its corresponding wavelength. At the period of 10 s, Rayleigh wave phase velocity is mostly sensitive to shear wave speed structure between ∼5 and 15 km depth, thus the whole structure is very similar to the map of shear

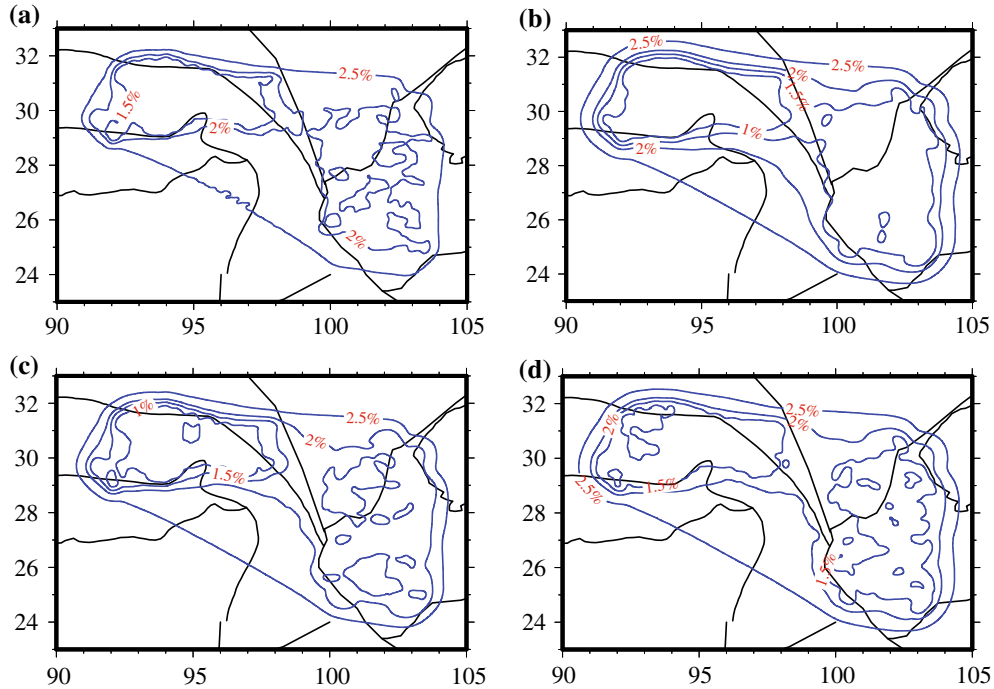

Figure 6
Posterior errors (in percent) of phase velocities at 20 s with the checkerboard model of mixed anomaly sizes: **a–c** posterior errors for the recovered phase velocity map using the traditional method with the fixed correlation length of 20, 30 and 40 km, respectively; **d** posterior errors for the recovered results using the new method. The corresponding recovery results are shown in Fig. 5b–e

wave speeds at 10 km (Yao et al. 2010). The low phase velocities appear near major faults zones, such as Xianshuihe Fault, Batang Fault, Lijiang Fault, Red River Fault, and Xianjiang Fault zone; and the boundaries of the low velocity regions seem to coincide roughly with some major faults in the study area, e.g., Red River Fault and Lijiang Fault.

The phase velocity at the period of 30 s samples primarily the shear velocity structure in the lower crust in the high plateau area and lower crust and uppermost mantle in Yunnan, according to its sensitivity kernel. Both 30 s phase velocity maps (Fig. 7c, d) show similar features: low phase velocities appear beneath the plateau area with crustal thickness of about 50–75 km (Yao et al. 2010) and high phase velocities beneath the Sichuan Basin and Yunnan in SW China, where the crustal thickness is typically around 40–50 km (Yao et al. 2010). Besides the thick crustal thickness that contributes to low phase velocities at 30 s in the high plateau area, previous tomographic results have revealed apparent mid-lower crust low velocities in SE Tibet (e.g., Yao et al. 2010; Yang et al. 2012; Liu et al. 2014; Bao et al. 2015), which may also significantly decrease the phase velocity at 30 s in these regions.

In contrast with the result of the original inversion method, the spatially varying inversion reveals finer structures in the Lehigh array area and western part of the MIT array as the result of high path density and hence shorter correlation length rather than the artifacts produced in the inversion; for example, in the traditional results at 10 s, the low velocity anomaly seems continuous near Batang Fault, but for the new method, there are two separated low velocity regions. And the amplitude of some anomalies also gets larger in the new model such as the high velocity anomaly near Luzhijiang Fault (Fig. 7d). We also calculate the posterior errors of the results to access model uncertainties (Fig. 8), and in most regions the posterior error of the spatially varying smoothing method lies within the 1% error contour line (Fig. 9b, d) except for the central region between

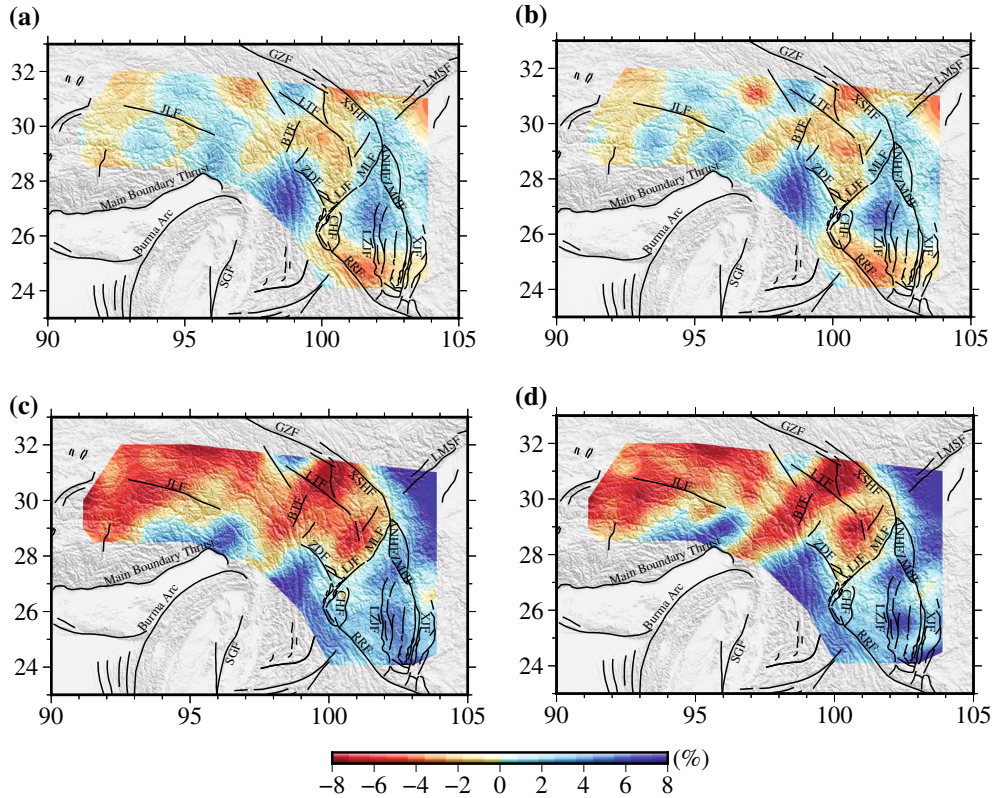

Figure 7

Rayleigh-wave phase velocity maps at periods 10 and 30 s using the traditional inversion and spatially varying smoothing inversion methods. **a** and **c** give the phase velocity maps at 10 and 30 s using the traditional method with a fixed correlation length of 110 km, respectively. **b** and **d** are the phase velocity maps at 10 and 30 s using the spatially varying smoothing method with the corresponding spatial correlation length shown in Fig. 1b, d, respectively. The *colorbar* shows the amplitude of the velocity anomaly in percent. The major faults are depicted as *thin black lines* (after Wang et al. 1998; Wang and Burchfiel 2000; Shen et al. 2005), and the corresponding abbreviations are as follows: *JLF* Jiali Fault, *GZF* Ganzi Fault, *LMSF* Longmenshan Fault, *XSHF* Xianshuihe Fault, *LTF* Litang Fault, *BTF* Batang Fault, *ANHF* Anninghe Fault, *ZMHF* Zemuhe Fault, *ZDF* Zhongdian Fault, *LJF* Lijiang Fault, *MLF* Muli Fault, *CHF* Chenghai Fault, *LZJF* Luzhijiang Fault, *XJF* Xiaojiang Fault, *RRF* Red River Fault

the two arrays with non-uniform azimuthal path coverage, which suggests that the anomaly patterns of the results (Fig. 7b, d) using the spatially varying smoothing method are not the artifacts of the inversion. As for the data misfit, compared with the original inversion, the standard deviation of the traveltime residuals is reduced from 1.642 to 1.612 s at 10 s and from 1.505 to 1.468 s at 30 s using the spatially varying smoothing inversion, meanwhile the average of the residual is closer to zero. This is mainly due to the fact that the spatially varying smoothing inversion has smaller correlation lengths in regions with denser ray path coverage, thus resulting in less lateral smearing effects in the model space.

4.2. *Two-Step Inversion to Suppress Effect of Data Outliers*

To inspect the performance of our proposed two-step inversion scheme to suppress the effect of data outliers, we conduct another synthetic test at 10 s using the output model of the spatially varying smoothing method (Fig. 7b) as the "true" model. We first generate the synthetic phase velocity data from the "true" model. Among all the synthetic data, 97% (randomly picked) are added 1% random Gaussian noise, and the rest 3% (see Fig. 9 for the paths) are added 10% random noise, which stand for the potential data outliers. Then we conduct inversion adopting the spatially varying smoothing method with

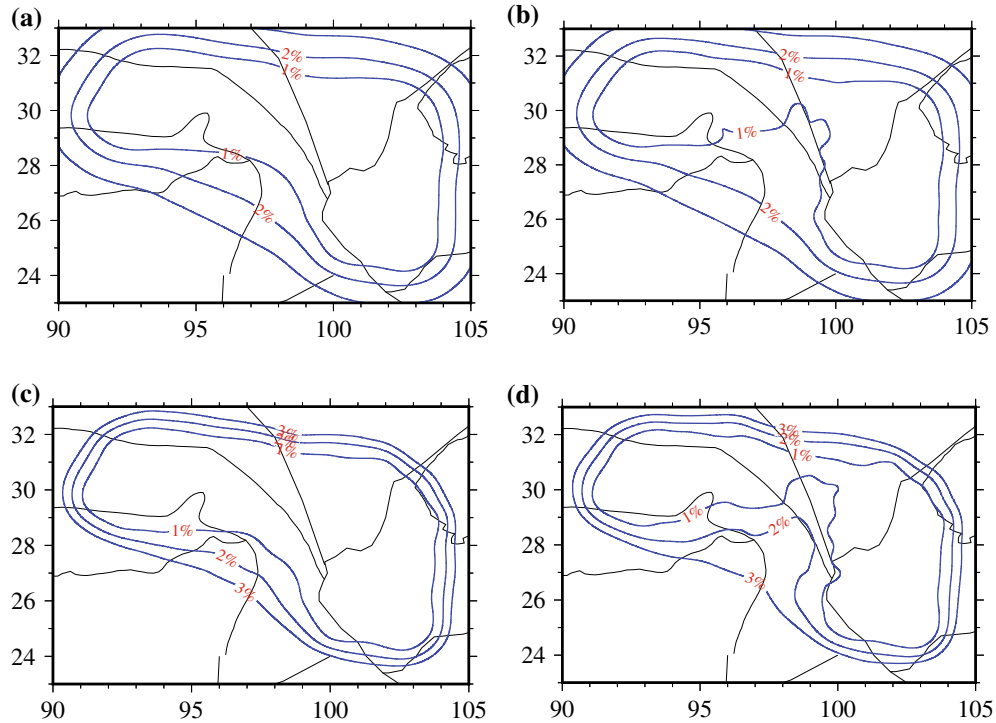

Figure 8

Posterior errors (in percent) of Rayleigh-wave velocities at periods 10 and 30 s using the traditional and new methods. **a** and **c** give the posterior error maps at 10 and 30 s using the traditional method with a fixed correlation length of 110 km, respectively. **b** and **d** are the posterior error maps at 10 and 30 s using the new method with the corresponding spatial correlation lengths shown in Fig. 1b, d, respectively. *The black lines* show the block boundaries in this region

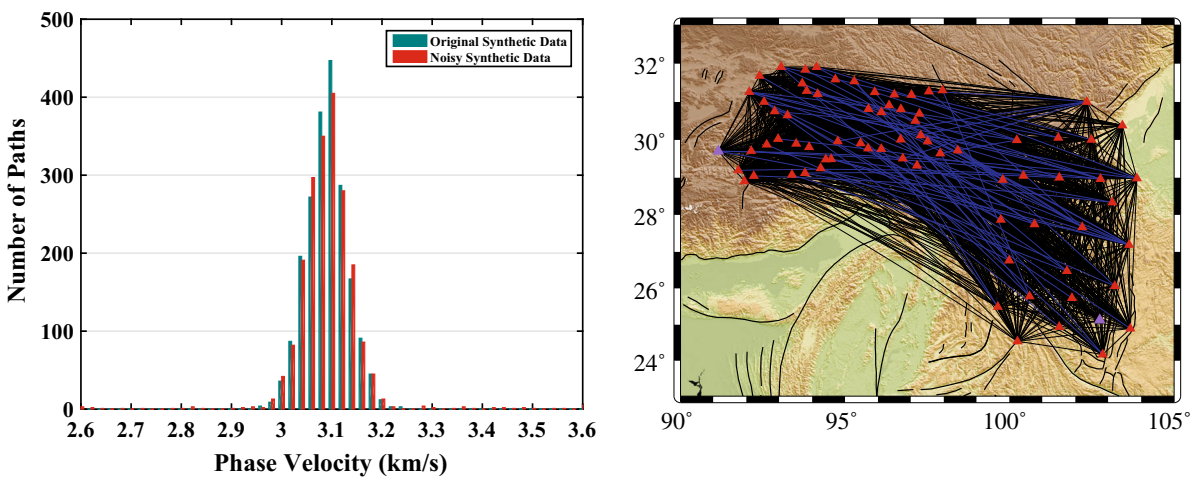

Figure 9

Statistics and ray path coverage of the synthetic data used in the test with data outliers. **a** Gives histograms of the synthetic data without data outliers (*green*, 1% random Gaussian noise added for each measurement) and with data outliers (*red*), in which 3% ray paths have 10% random noise (outliers) and only 1% random noise added for the other ray paths. **b** Shows the ray path coverage of the synthetic data in which the *black lines* show the 97% ray paths with 1% random noise while the *blue lines* show the rest 3% ray paths with 10% random noise

the spatial correlation length same as those used for Fig. 7b, and the first and second step inversion results are shown in Fig. 10a, b, respectively. The corresponding differences between the obtained model of each step and the "true" model (Fig. 7b) are shown in Fig. 10c, d, respectively. The first step inversion results show obvious artifacts and large relative model differences (above 6%) in the central region (Fig. 10c, d). In the second step inversion, we increase the a priori uncertainty for paths with very large misfit (see Eq. (6)), and the results show that the pattern as well as amplitude of anomalies can be well retrieved except the marginal regions with sparse data coverage (Fig. 10b, d). Moreover, in most of the study region, the relative model differences are below 2%. This test well demonstrates that our two-step inversion scheme can suppress the effect of data outliers effectively.

4.3. Comparisons and Limitations

Our spatially varying smoothing approach and the Voronoi diagram approach by Debayle and Sambridge (2004) both adopt the continuous regionalization scheme of Tarantola and Nercessian (1984) and Montagner (1986), and both methods can evaluate the variation of lateral resolution and achieve tomographic models with spatially varying resolution based on ray path coverage. The Voronoi diagram approach by Debayle and Sambridge (2004) uses the local ray path density and azimuthal coverage to construct spatially adaptive Voronoi cells (i.e., the model space) but with a fixed correlation length. However, in our approach, we use the fixed regular grids but with a spatially varying correlation length for each grid based on its ray path density. Hence, in the Voronoi diagram approach, the

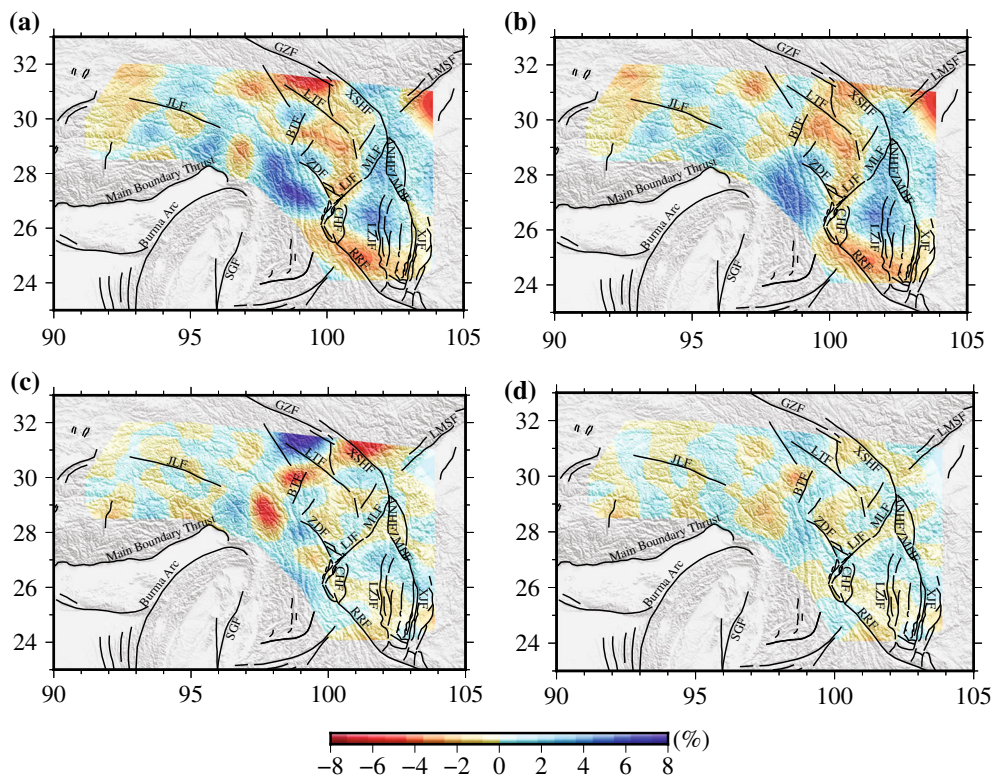

Figure 10

Recovery tests from the synthetic model (Fig. 7b) using the two-step method at period 10 s with data outliers shown in Fig. 9. **a** and **b** give the recovered models of first and second step inversion results with the same spatial correlation length shown as in Fig. 1b, respectively. **c** and **d** give the model differences between the recovered models of first and second step inversion and the "true" model (Fig. 7b), respectively. The *colorbar* shows the amplitude of the velocity anomaly in percent

highest resolution is controlled by the fixed correlation length, while in our approach the highest resolution is controlled by the lowest correlation length, which is provided by a series of resolution tests. In regions with sparse path coverage (i.e., lower resolution), the Voronoi diagram approach increases the size of the Voronoi cells, while in our approach we increase the correlation lengths.

There still exit some drawbacks in our spatially varying smoothing method. In our approach, the $L_{\mathrm{corr}}(p(r))$ is set to vary linearly with $p(r)$, the ray path density, with the largest and lowest correlation lengths determined by a series of checkerboard resolution tests. However, the traditional checkerboard tests have some potential drawbacks, for instance, it can neither provide quantitative measures of resolution nor reveal true structural distortion or smearing that can be caused by the data coverage (Rawlinson and Spakman 2016), and the shape of checkerboard pattern may also affect the recovery results (Lévêque et al. 1993). A more appropriate work in the future is to first determine the spatially varying resolution length via analysis of the resolution matrix (Yanovskaya and Ditmar 1990; Barmin et al. 2001) and then relates the spatial correlation length function linearly with the spatial resolution length. In addition, we have not considered the azimuthal path coverage in determining the ray path density, which also needs to be improved in the future, in particular for the inversion of azimuthal anisotropy. And the use of spatial resolution length estimated from the resolution matrix will be better than the ray path density for considering the azimuthal path coverage, which helps to give more reliable spatially varying correlation lengths.

5. Conclusion

We propose a new approach to achieve spatially varying smoothing tomography based on the traditional continuous regionalization method for surface wave tomography. This new method uses spatially varying correlation lengths based on the ray path density instead of a fixed-value correlation length. To suppress effects of data outliers on inversion results, we propose a two-step inversion procedure, in which the data standard errors for paths with very large misfits in the first step inversion will be significantly increased in the second step inversion. Our method inherits the merits of the continuous regionalization approach including its robustness and clear physical meanings. The checkerboard tests with different sizes of anomalies are used to design the range of the spatially varying correlation lengths. The synthetic data tests show that the spatially varying smoothing inversion provides better spatial resolution in regions with higher path density owing to the choice of smaller correlation lengths in the inversion. We applied the proposed method to real surface wave dispersion data set in SE Tibet. The new algorithm obtains similar phase velocity structures as the traditional inversion approach, but the results have higher spatial resolution in regions with better path coverage and appear to better resolve the amplitude of anomalies.

Acknowledgements

We appreciate three anonymous reviewers for their constructive comments, which help to significantly improve the original manuscript. Most figures are made from the Generic Mapping Tools (GMT) (http://gmt.soest.hawaii.edu/). This study is supported by National Natural Science Foundation of China (41574034, 41222028), and the Fundamental Research Funds for the Central Universities in China (WK2080000053).

References

Abers, G. A., & Roecker, S. W. (1991). Deep structure of an arc-continent collision: Earthquake relocation and inversion for upper mantle P and S wave velocities beneath Papua New Guinea. *Journal Geophysical Research, 96*(B4), 6379–6401.

An, M., Feng, M., & Zhao, Y. (2009). Destruction of lithosphere within the north China craton inferred from surface wave tomography. *Geochemistry, Geophysics, Geosystems, 10*(8), Q08016.

Bao, X., Sun, X., Xu, M., Eaton, D. W., Song, X., Wang, L., et al. (2015). Two crustal low-velocity channels beneath SE Tibet revealed by joint inversion of Rayleigh wave dispersion and receiver functions, Earth planet. *Science Letters, 415*, 16–24.

Barmin, M. P., Ritzwoller, M. H., & Levshin, A. L. (2001). A fast and reliable method for surface wave tomography. *Pure and Applied Geophysics, 158*, 1351–1375.

Bodin, T., & Sambridge, M. (2009). Seismic tomography with the reversible jump algorithm. *Geophysical Journal International, 178*(3), 1411–1436.

Böhm, G., Galuppo, P., & Vesnaver, A. (2000). 3D adaptive tomography using Delaunay triangles and Voronoi polygons. *Geophysical Prospecting, 48*(4), 723–744.

Boschi, L., & Ekström, G. (2002). New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies. *Journal of Geophysical Research: Solid Earth (1978–2012), 107*(B4), 2059.

Chiao, L. Y., & Kuo, B. Y. (2001). Multiscale seismic tomography. *Geophysical Journal International, 145*(2), 517–527.

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

Debayle, E., Kennett, B., & Priestley, K. (2005). Global azimuthal seismic anisotropy and the unique plate-motion deformation of Australia. *Nature, 433*(7025), 509–512.

Debayle, E., & Sambridge, M. (2004). Inversion of massive surface wave data sets: Model construction and resolution assessment. *Journal Geophysical Research, 109*, B02316.

Delost, M., Virieux, J., & Operto, S. (2008). First arrival traveltime tomography using second generation wavelets. *Geophysical Prospecting, 56*(4), 505–526.

Ditmar, P. G., & Yanovskaya, T. B. (1987). Generalization of the Backus–Gilbert method for estimation of the horizontal variations of surface wave velocity. *Izvestiya of the Academy of Sciences of the USSR Physics of the Solid Earth, English Translation, 23*(6), 470–477.

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

Fang, H., Yao, H., Zhang, H., Huang, Y. C., & van der Hilst, R. D. (2015). Direct inversion of surface wave dispersion for three-dimensional shallow crustal structure based on ray tracing: methodology and application. *Geophysical Journal International, 201*(3), 1251–1263.

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

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Boca Raton: Chapman & Hall/CRC.

Hawkins, R., & Sambridge, M. (2015). Geophysical imaging using trans-dimensional trees. *Geophysical Journal International, 203*(2), 972–1000.

Hung, S. H., Chen, W. P., & Chiao, L. Y. (2011). A data-adaptive, multiscale approach of finite-frequency, traveltime tomography with special reference to P and S wave data from central Tibet. *Journal of Geophysical Research: Solid Earth (1978–2012), 116*(B6), B06307. doi:10.1029/2010JB008190.

Lévêque, J. J., Rivera, L., & Wittlinger, G. (1993). On the use of the checker-board test to assess the resolution of tomographic inversions. *Geophysical Journal International, 115*, 313–318.

Levshin, A. L., Yanovskaya, T. B., Lander, A. V., Bukchin, B. G., Barmin, M. P., Ratnikova, L. I., et al. (1989). *Seismic surface waves in a laterally inhomogeneous Earth*. Norwell: Kluwer Academic.

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. *Geophysical Journal International, 173*(1), 281–298.

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

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

Liu, Q. Y., van der Hilst, R. D., Li, Y., Yao, H. J., Chen, J. H., Guo, B., et al. (2014). Eastward expansion of the Tibetan Plateau by crustal flow and strain partitioning across faults. *Nature Geoscience, 7*(5), 361–365.

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

Montagner, J.-P. (1986). Regional three-dimensional structures using long period surface waves. *Annales Geophysicae, 4*, 283–294.

Montagner, J.-P., & Tanimoto, T. (1991). Global upper mantle tomography of seismic velocities and anisotropies. *Journal of Geophysical Research: Solid Earth (1978–2012), 96*(B12), 20337–20351.

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. *Geophysical Journal International, 180*(3), 1153–1169.

Rawlinson, N., & Spakman, W. (2016). On the use of sensitivity tests in seismic tomography. *Geophysical Journal International, 205*(2), 1221–1243.

Ritzwoller, M. H., & Levshin, A. L. (1998). Eurasian surface wave tomography: Group velocities. *Journal of Geophysical Research: Solid Earth (1978–2012), 103*(B3), 4839–4878.

Ritzwoller, M. H., Shapiro, N. M., Barmin, M. P., & Levshin, A. L. (2002). Global surface wave diffraction tomography. *Journal of Geophysical Research: Solid Earth (1978–2012), 107*(B12), ESE-4.

Ritzwoller, M. H., Shapiro, N. M., Levshin, A. L., & Leahy, G. M. (2001). Crustal and upper mantle structure beneath Antarctica and surrounding oceans. *Journal of Geophysical Research: Solid Earth (1978–2012), 106*(B12), 30645–30670.

Sambridge, M., & Gumundsson, Ó. (1998). Tomographic systems of equations with irregular cells. *Journal Geophysical Research, 103*(B1), 773–781.

Saygin, E., Cummins, P. R., Cipta, A., Hawkins, R., Pandhu, R., Murjaya, J., et al. (2016). Imaging architecture of the Jakarta Basin, Indonesia with transdimensional inversion of seismic noise. *Geophysical Journal International, 204*(2), 918–931.

Shapiro, N. M., & Ritzwoller, M. H. (2002). Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle. *Geophysical Journal International, 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.

Shen, Z.-K., Lü, J., Wang, M., & Bürgmann, R. (2005). Contemporary crustal deformation around the southeast borderland of the Tibetan plateau. *Journal Geophysical Research, 106*(B4), 6793–6816.

Silveira, G., & Stutzmann, E. (2002). Anisotropic tomography of the Atlantic Ocean. *Physics of the Earth and Planetary Interiors, 132*(4), 237–248.

Simons, F. J., Loris, I., Nolet, G., Daubechies, I. C., Voronin, S., Judd, J. S., et al. (2011). Solving or resolving global tomographic models with spherical wavelets, and the scale and sparsity of seismic heterogeneity. *Geophysical Journal International, 187*(2), 969–988.

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. *Geophysical Journal International, 151*, 738–754.

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

Spetzler, J., & Snieder, R. (2001). The effect of small scale heterogeneity on the arrival time of waves. *Geophysical Journal International, 145*(3), 786–796.

Tarantola, A., & Nercessian, A. (1984). Three-dimensional inversion without blocks. *Geophysical Journal of the Royal Astronomical Society, 76*, 299–306.

Tarantola, A., & Valette, B. (1982). Generalized nonlinear inverse problem solved using the least squares criterion. *Reviews of Geophysics and Space Physics, 20*(2), 219–232.

Wang, E., & Burchfiel, B. C. (2000). Late Cenozoic to Holocene deformation in southwestern Sichuan and adjacent Yunnan, China, and its role in formation of the southeastern part of the Tibetan plateau. *Geological Society of America Bulletin, 112*(3), 413–423.

Wang, E., Burchfiel, B. C., Royden, L. H., Chen, L., Chen, J., Li, W., et al. (1998). Late Cenozoic Xianshuihe-Xiaojiang, Red River and Dali fault systems of southwestern Sichuan and central Yunnan, China. *Special Papers - Geological Society of America, 327*, 1–108.

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

Yang, Y., Ritzwoller, M. H., Zheng, Y., Shen, W., Levshin, A. L., & Xie, Z. (2012) A synoptic view of the distribution and connectivity of the mid-crustal low velocity zone beneath Tibet. *Journal of Geophysical Research: Solid Earth (1978–2012), 117*(B4), B04303. doi:10.1029/2011JB008810.

Yanovskaya, T. B., & Ditmar, P. G. (1990). Smoothness criteria in surface wave tomography. *Geophysical Journal International, 102*(1), 63–72.

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. *Geophysical Journal International, 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. *Journal of Geophysical Research: Solid Earth (1978–2012), 115*(B12), B12307. doi:10.1029/2009JB007142.

Yao, H., Xu, G., Zhu, L., & Xiao, X. (2005). Mantle structure from interstation Rayleigh wave dispersion and its tectonic implication in western China and neighboring regions. *Physics of the Earth and Planetary Interiors, 148*, 39–54.

Yoshizawa, K., & Kennett, B. (2004). Multimode surface wave tomography for the Australian region using a three-stage approach incorporating finite frequency effects. *Journal of Geophysical Research: Solid Earth (1978–2012), 109*(B2), B02310. doi:10.1029/2002JB002254.

Zhang, H., & Thurber, C. H. (2005). Adaptive mesh seismic tomography based on tetrahedral and Voronoi diagrams: Application to Parkfield, California. *Journal Geophysical Research, 110*(B4), B04303. doi:10.1029/2004JB003186.

(Received February 1, 2016, revised November 8, 2016, accepted November 15, 2016, Published online November 24, 2016)