

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE

10.1002/2015JB012450

Key Points:

- Model S velocity beneath China by jointly interpreting receiver function and surface wave data
- A multistep approach was developed and employed to obtain S velocity models
- The combined 3-D model reveals complicated velocity structure beneath China

Correspondence to:F. Niu,
niu@cup.edu.cn**Citation:**

Chen, Y., and F. Niu (2016), Joint inversion of receiver functions and surface waves with enhanced preconditioning on densely distributed CNDNS stations: Crustal and upper mantle structure beneath China, *J. Geophys. Res. Solid Earth*, 121, doi:10.1002/2015JB012450.

Received 13 AUG 2015

Accepted 30 DEC 2015

Accepted article online 5 JAN 2016

Joint inversion of receiver functions and surface waves with enhanced preconditioning on densely distributed CNDNS stations: Crustal and upper mantle structure beneath China

Youlin Chen¹ and Fenglin Niu^{2,3}

¹Advanced Technology Division, Array Information Technology, Inc., Greenbelt, Maryland, USA, ²State Key Laboratory of Petroleum Resources and Prospecting, and Unconventional Natural Gas Institute, China University of Petroleum, Beijing, China, ³Department of Earth Science, Rice University, Houston, Texas, USA

Abstract We present shear wave velocity structure beneath China by joint modeling of teleseismic receiver function and Rayleigh wave group velocity dispersion data observed at +1000 permanent broadband seismic stations in the Chinese National Digital Seismic Network (CNDNS). A ray-parameter-based stacking method is employed to minimize artifacts in stacking receiver functions from different sources. The Rayleigh wave dispersion curve is extracted from group velocity tomographic models at all applicable periods. Enhanced preconditions are applied on the linearized iterative inversion to regularize and balance multiple types of data. The velocity profile inversion at each station starts from an initial model derived from sediments, crustal thickness, V_p/V_s ratio and P_n/S_n models. This multistep approach not only reduces uncertainty and nonuniqueness of the velocity inversion but also efficiently fills information gap in each data set. We then generate a 3-D S velocity model by combining and smoothing all the 1-D models. The obtained 3-D model reveals crustal and upper mantle velocity structures that are well correlated with tectonic features of China, for example, our model shows a clear east-west bimodal distribution at 35 km deep, low velocity in the crust beneath central and eastern Tibetan plateau, and sedimentary structure in major cratons and basins. Our model is consistent with existing tomographic models in large scale but provides more structural details in regional and local scales.

1. Introduction

Receiver function imaging technique has become a powerful tool to probe Earth's interior structure, partially due to the rapid expansion of three-component broadband seismographs in recent years. The receiver function technique can effectively isolate the shear wave response of Earth's structure beneath the receiver site by deconvolving the vertical component with source signatures from the mode-converted waves on the horizontal components. In general, the shear wave response is more sensitive to radial changes rather than the absolute amplitude of seismic structures. Since *Langston* [1979] first introduced the concept in his pioneer work more than three decades ago, various methods have been developed to investigate crustal and upper mantle structure using receiver function data. They have been used to estimate crustal thickness, crustal V_p/V_s ratio, density contrast across the Moho [e.g., *Zhu and Kanamori*, 2000; *Chevrot and van der Hilst*, 2000; *Nair et al.*, 2006; *Niu et al.*, 2007; *Chen et al.*, 2010], to create images of discontinuities in the mantle and crust [e.g., *Owens et al.*, 2000; *Gilbert et al.*, 2003; *Chen et al.*, 2005] and to constrain dipping and anisotropic structures in the crust [e.g., *Cassidy*, 1992; *Frederiksen and Bostock*, 2000; *Liu and Niu*, 2012]. On the other hand, receiver functions have weak response to absolute velocities, leading to substantial uncertainties in modeling velocity profile. The inversion of receiver function for velocity is thus generally regarded as a strongly nonunique and nonlinear inverse problem [e.g., *Langston*, 1979; *Ammon et al.*, 1990; *Ammon*, 1991; *Jacobsen and Svensnigen*, 2008]. Although the nonunique problem can be partially reduced by incorporating tomographic models [e.g., *Tkalčić et al.*, 2011], using receiver function data alone cannot fully solve the issue no matter what methods are used [e.g., *Ammon et al.*, 1990; *Shibutani et al.*, 1996; *Zhao et al.*, 1996; *Sambridge*, 1999; *Bodin et al.*, 2012].

On the contrary, surface wave dispersion data have better sensitivity on the average of shear wave velocities but less sensitive to the style of velocity changes with depth. It has been shown that the joint inversion of receiver functions with surface wave dispersion data can improve the sensitivity to the true velocity structure due to the complementary nature of the two data sets. The joint inversion has been implemented either

linearly or nonlinearly and applied to data sets collected at various parts of the world [e.g., Özalaybey *et al.*, 1997; Du and Foulger, 1999; Julià *et al.*, 2000, 2003; Chang *et al.*, 2004; Lawrence and Wiens, 2004; Pasanos *et al.*, 2007; Bodin *et al.*, 2012].

As the major part of East Eurasia, the mainland of China is highly diverse in geology, consisting of various tectonic blocks from ancient Archean cratons to young and active orogens. The modern tectonics of China is the result of a long and complicated evolution of collision, subduction, and amalgamation of microcontinents and fold belts with Precambrian nuclei. During the later Permian to early Triassic, the North China Block (NCB) was first amalgamated with the central Asian orogen in the north and was collided by the South China Block (SCB) from the south. The collision between NCB and SCB occurred as a northward underthrusting of the SCB under the NCB, which likely led to the uplift of the Sulu and Qinling-Dabie orogens in the north side of the suture zone and the flexural subsidence in the south side [e.g., Vermeesch, 2003; Xiao *et al.*, 2003]. Almost immediately following the NCB-SCB collision, the Qiangtang block was underthrust by the amalgamated South China and Qaidam-Tarim blocks along the Jinsha suture in the northeast and was further underthrust by the Gondwana-derived Lhasa block from the south in Cretaceous [e.g., Yin and Nie, 1996; Yin and Harrison, 2000]. The collision of the India plate to the Asia plate occurred at approximately 45 Ma ago, leading to crustal shortening and thickening in the Himalaya-Tibet region, as well as the reactivation of a large area to the north such as Tianshan mountain range and the Mongolia plateau [e.g., Yin and Harrison, 2000; Tapponnier *et al.*, 2001]. In contrast, the subduction of the Pacific and Philippine plates led to the extension of the eastern part of China during the late Mesozoic and the entire Cenozoic [e.g., Ren *et al.*, 2002; Li *et al.*, 2012]. It is undoubting that these tectonic processes had produced strong seismic heterogeneities in the crust and mantle beneath China. Mapping out these seismic heterogeneities can thus provide constraints on the geodynamic evolution of this region.

The previous seismological studies regarding the lithospheric structure of China mainly relies on *P* and *S* travel time tomography [e.g., Liang *et al.*, 2004; Pei *et al.*, 2007; Sun and Toksöz, 2006; Li and Van der Hilst, 2010; Wang *et al.*, 2013]. While travel time tomography is an effective way to image lateral heterogeneities in the crust and mantle, its depth resolution is relatively low, usually on the order of tens of kilometers. Waveform-based studies are able to provide complementary information that is difficultly obtained from pure travel time tomography on the lithospheric structure, such as the deep seismic sounding (DSS) experiments [e.g., Li and Mooney, 1998; Li *et al.*, 2006; Teng *et al.*, 2013]. Furthermore, the rapid increase of three-component broadband seismic stations deployed in China during the recent decade has provided us opportunities to perform alternative waveform-based imaging studies that are cheaper and more stable than DSS can offer, such as using body wave receiver function [e.g., Mangino *et al.*, 1999; Kind *et al.*, 2002; Vergne *et al.*, 2002; Sherrington *et al.*, 2004; Chen *et al.*, 2010; Tkalcic *et al.*, 2011; He *et al.*, 2014; Li *et al.*, 2014] and surface wave data [e.g., Huang *et al.*, 2003; Zheng *et al.*, 2008; Yao *et al.*, 2008, 2010; Li *et al.*, 2013].

In order to build up a whole picture of lithospheric structure beneath the entire mainland China, the main goal of this study is to jointly interpret receiver function and Rayleigh wave group velocity dispersion data recorded on more than 1000 permanent broadband stations across China through a multistep approach. Specifically, the receiver function data are stacked using our recently developed ray-parameter-based (RPB) summing technique [Chen and Niu, 2013] to minimize artifacts introduced into observables from data processing. The Rayleigh wave group velocity dispersion of each station is extracted from group velocity tomographic models at all applicable periods. We choose an iterative linearized algorithm to implement the joint inversion because of its lower computational demands, as well as its higher depth resolution than those nonlinear methods. Since the linearized inversion depends highly on the initial model, we carefully construct the initial model for each station following the method proposed in Tkalcic *et al.* [2011]. We finally employ the enhanced preconditioning technique in the joint inversion to ensure reliable velocity models [Chen and Niu, 2013]. Processing seismic data through these multisteps provides us intuitive knowledge on how each part of the Earth model responds to a particular data set.

2. Seismic Waveform Data

During the last decade, the number of permanent broadband seismic stations in mainland China has rapidly increased under the great effort of the China Earthquake Administration. Being completed in early 2007, the Chinese National Digital Seismic Network (CNSN) is now the largest permanent seismic network in the

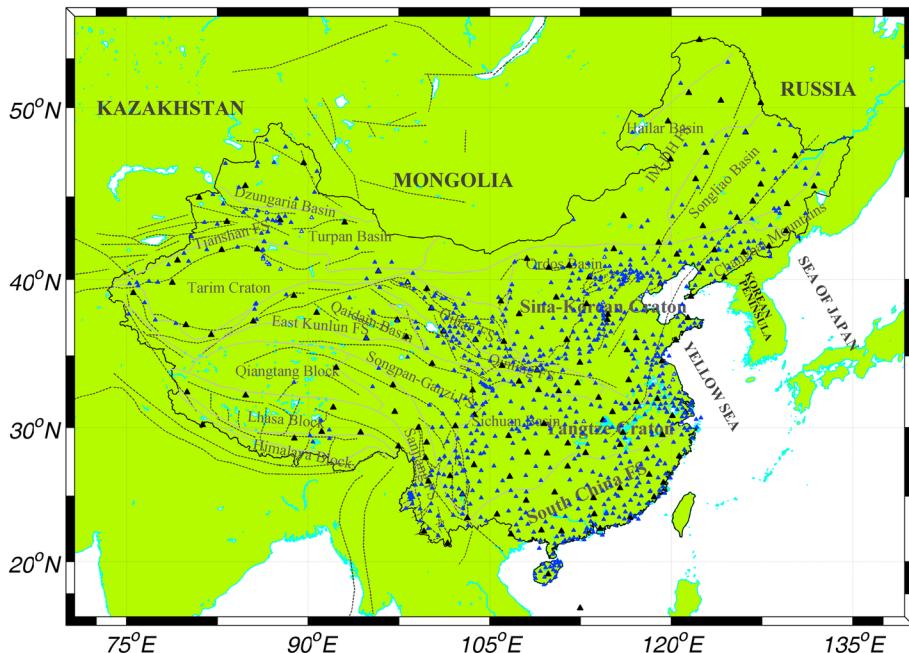


Figure 1. Geographic distribution of the CNDN stations used in this study. The black triangles indicate the 144 national stations, and the blue triangles indicate the regional stations of the CNDN. The open triangles stand for the stations without velocity models.

world with more than 1000 broadband stations, consisting of a backbone national network of 144 stations, 31 regional networks in every province that add up to more than 900 stations, and several small aperture arrays (Figure 1). These stations formed a large 2-D areal virtual array with an aperture of ~6700 km from east to west and ~3500 km from north to south, providing the data not only for the essential monitoring of seismicity, but also for the studies of seismic structures beneath China and adjacent regions. The waveform data are managed and provided by the China Earthquake Network Center (CENC).

Figure 1 illustrates the station distribution of national and regional networks used in this study. The spacing between stations varies drastically with locations, which reaches to only ~20 km in the eastern part of China but is ~350 km in western China. For most stations, we collected at least 3 years waveform data from August 2007 to October 2010. For those stations that were upgraded to digital broadband seismographs in earlier years, including the 11 IRIS-GSN stations, our data collection goes up to more than 5 years [Chen *et al.*, 2010]. However, we noticed that a number of stations had problems of misorientation and mislabeling on recording components in the early stage. In general, these problems existed for more than 1 year at some stations but disappeared after a point in time, which we assumed were fixed by maintenance. Using the method described in Niu and Li [2011], we identified and corrected misorientation and mislabeling errors on all of the problematic stations in preprocessing.

2.1. Receiver Function Data

2.1.1. Waveform Processing, Quality Control, and Data Selection

We searched the global earthquake catalog and selected ~ 440 teleseismic events with $M \geq 5.8$ that occurred between 30° and 90° to the CNDN stations within the 3 year period (Figure 2) to generate receiver functions. The pre-process of raw seismograms included removing baseline shifts, detrending, and high-pass filtering to remove long period ($\geq 15s$) fluctuations. And then the two horizontal components were rotated from observation coordinates (N-S and E-W) to radial (R) and transverse (T) components. Receiver functions are impulse responses of structures beneath seismic stations, which can be achieved through a deconvolution of vertical components from radial components in either frequency domain or time domain. Here we computed all the receiver functions using an iterative deconvolution technique in time domain [e.g., Niu and Kawakatsu, 1996; Ligorria and Ammon, 1999]. Comparing to the widely used “water level” deconvolution technique in

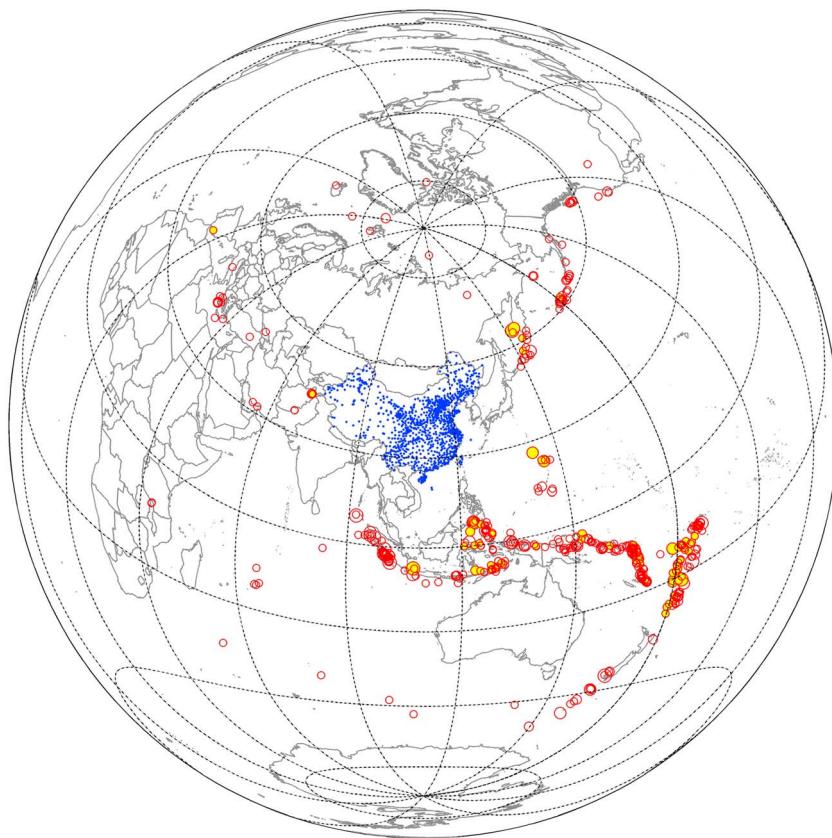


Figure 2. Distribution of teleseismic events used for receiver function analysis. The open and solid cycles represent earthquakes with a focal depth shallower and deeper than 100 km, respectively. The size of the cycles is proportional to the size of events.

frequency domain, this iterative method can produce more stable result for data with low signal-to-noise ratio but at the price of longer computation time. It also avoids complex relationships between water level values and the resultant receiver functions. We generated receiver functions in two frequency bands using a low-pass Gaussian filter, $\exp\left(-\frac{w^2}{4a^2}\right)$, with width factors of $a=1.0$ and 2.5 , both of which were used in our inversion. These two width factors correspond roughly to corner frequencies of 0.5 and 1.2 Hz filters, respectively. We did not perform strict quality control on the raw waveform data in order to allow as many waveforms as possible to be used in computing receiver functions. However, we adopted the statistical approach described in *Chen et al. [2010]* and *Tkalčić et al. [2011]* to select coherent receiver functions automatically from the large volume of data collected at each station. The selection is based on cross-correlation coefficients of radial receiver functions such that the noisy or erroneous receiver functions are separated from the coherent data set and are subsequently removed.

Approximately 70% receiver functions were finally selected at most stations, and the selected data generally exhibited a fair coverage in azimuth with clusters in the northeast to southeast directions corresponding to earthquakes occurring in western Pacific subduction zones. There are also groups of earthquakes that occurred in the Mediterranean region, and the Indian and South Atlantic Oceans from the west and southwest, respectively (Figure 2). For each station, we only chose the earthquakes coming from the southern azimuths, which appear to produce the majority of receiver functions among all the directions. By choosing only one azimuthal range of data, we have in fact set aside the issues in conjunction with a possible dipping Moho and/or crustal azimuthal anisotropy, which can cause directional dependence of receiver functions. Narrowing azimuthal range can minimize their influences even if crustal anisotropy and Moho dipping are not negligible. Determining azimuthal anisotropy and Moho dipping by using receiver function data is beyond the scope of this study and will be the subject of future work. Due to the large size of the study area,

the back azimuth from a single earthquake to the stations can vary significantly across the country. We did not fix the azimuthal range in the quadrant of 90°–180° as they did in *Tkalčić et al.* [2011]; instead, we chose the azimuths range of 90°–230° for stations in the east and 70°–200° for stations in the west.

After data quality control and back azimuth selection, we finally kept more than 200 receiver functions at most stations for the joint velocity inversion. However, due to different operational periods of individual stations, as well as inevitable technical, transferring, and processing failures that occasionally occurred at some stations, the number of usable events varies from several tens to several hundreds at different stations. This makes a small number of stations unused in case that they are unable to provide sufficient waveform data to obtain reliable results.

2.1.2. Ray-Parameter-Based Stacking as Observable Data

In performing receiver function inversion, it is usually preferred to use the stacked data as the input to reduce noise level and suppress 2-D/3-D effects. However, since the arrival times of the converted and reverberated S wave phases with respect to the direct P waves vary with epicentral distance, directly stacking the receiver function data aligned to the direct P waves can significantly reduce the amplitude and distort the waveform of these phases. The destructive stacking becomes more severe for the multiples and in the case of a thick crust and large distance coverage, particularly true for high-frequency receiver function data. For example, for a 50 km thick crust with an average V_p/V_s ratio of 1.78, the arrival time of the 0p1s, 2p1s, and 1p2s phases [Niu and James, 2002] can differ as much as ~0.4, 1.2, and 0.8 s, respectively, in the distance range between 30° and 90° (equivalent to a ray parameter range between 4.7 and 8.9 s⁰). When receiver functions are computed using a Gaussian filter width $a=2.5$ or above, such large phase shifts can result in completely out-of-phase summation if no moveout correction is made. In principle, this problem can be avoided by narrowing the distance range of the receiver functions used for stacking and constructing a set of receiver functions at different epicentral distances for inversion. However, the global seismicity is highly uneven, and it may not be possible to gather enough receiver functions at certain distance range, making this approach less feasible than what one might have expected. It is also inefficient in processing a large volume of data such as those used in this study.

We have developed a ray-parameter-based (RPB) method to stack receiver functions at all epicentral distances with proper corrections for moveouts and amplitudes. The technical details, merits, and its importance in receiver function inversions can be found in *Chen and Niu* [2013]. Here we only briefly summarized the two major steps employed in the RPB stacking method: (1) We use the so-called “four-pin” algorithm to correct the moveouts of the P_s converted and the two crustal multiples to a reference ray parameter by stretching/compressing different time sections of the receiver functions. (2) We normalize the amplitude of each receiver function by dividing its own ray parameter and then multiplying the reference ray parameter. The phase moveout and amplitude-corrected receiver functions are then stacked as observables for inversion. The RPB method is obviously applicable to all the events occurring at the entire teleseismic distance range in an automatic way.

The crustal thickness (H) and crustal V_p/V_s ratio (κ) are two crucial parameters in the calculation of the amount of stretching/compressing in the “four-pin” algorithm. To accurately estimate these two parameters, we employed an advanced $H\text{-}\kappa$ analysis method in this study, where the cross-correlation coefficients between the primary P_s conversion 0p1s mode and the two reverberation 2p1s and 1p2s modes are inserted into the standard $H\text{-}\kappa$ stacking as additional weighting factors [Niu et al., 2007; Chen et al., 2010]. The advanced $H\text{-}\kappa$ stacking efficiently reduces the trade-off between H and κ and is capable of correctly identifying the Moho phases beneath areas with complicated crustal structures, where the Moho phases could be mixed up with other arrivals and are difficult to be picked by the regular $H\text{-}\kappa$ method. We first used the advanced $H\text{-}\kappa$ method to measure the crustal thickness and V_p/V_s ratio beneath each station and then used them to perform the RPB stacking. They are also used in constructing the initial model, which are very crucial to the linearized inversion.

Using stations GX.GUL and LN.SNY as examples, we compared the results of the RPB and direct stacks (Figure 3). Three RPB stacks were created using three reference ray parameters, $p_0=5.5$, 6.5, and 8.0 s⁰, respectively (Figures 3a and 3c). Note the variations in arrival time of the P_s conversion and reverberations among stacks with different ray parameters, as well as the significant differences between the RPB and direct stacks (Figures 3b and 3d). In general, the P_s conversion and the multiples are less broadened on the RPB

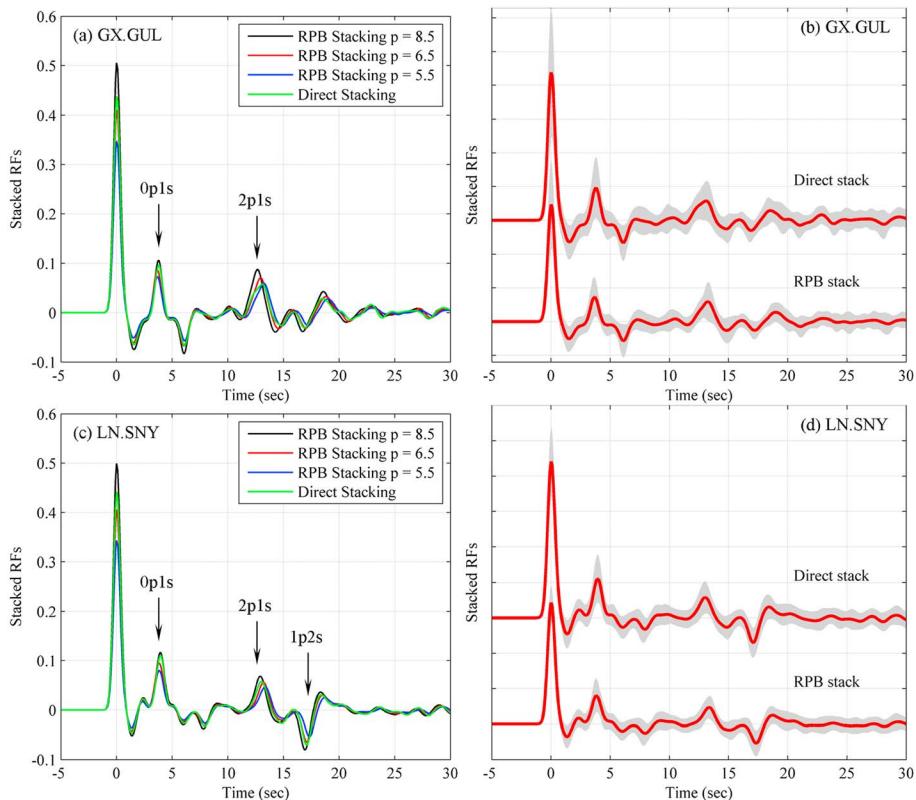


Figure 3. Comparison of RPB and direct stacking of receiver functions recorded at GX.GUL and LN.SNY. (a) RPB stacks created at three different ray parameters (black, red, and blue) of GX.GUL are plotted to compare with the direct stack (green). (b) Standard deviation curbs (gray areas) of the receiver functions used to create stacks are plotted around the direct and RPB stacks. (c) and (d) Same as Figures 3a and 3b, respectively, but for LN.SNY.

stacks as compared to the direct stacks. In Figures 3b and 3c, we also showed the standard deviation (gray area) of the two methods. For GX.GUL, the average standard deviation of the direct stack is 3.17%, while the average standard deviation of the RPB stack reduces to 2.50%. LN.SNY shows prominent 0p1s, 2p1s, and 1p2s phases, and the average standard deviation decreases from 2.60% on the direct stack to 2.08% on the RPB stack. The most substantial difference occurs at the direct P phase, where its standard deviation reduces from 9.81% to 5.38%.

2.2. Rayleigh Wave Group Velocity Dispersion Data

Receiver function is a point measurement in the geographic space, while surface wave dispersion is associated with the path between a source and a receiver. An important issue in simultaneously inverting these two types of data is to ensure that they sample the same area with a similar length scale. In order to reconcile the two data sets, i.e., to construct a dispersion curve associated with the location of a receiver, we need to take the average of all the dispersion curve measurements in an area around the station of interest and use it as the dispersion curve measurement of that station. Another more reliable method, as we adopted in this study, is first to carry out a 2-D tomographic inversion of surface wave velocities at each period and then construct a dispersion curve for each station based on the group velocities at the closest grid point.

2.2.1. Measurements of Rayleigh Wave Group Velocity Dispersion Curves

We chose about 100 regional earthquakes that occurred within and adjacent to China to conduct Rayleigh group velocity dispersion measurements. Since the group velocities of low-frequency surface wave are not expected to change rapidly among nearby stations, we only used the 144 backbone stations in CNDSN in measuring the dispersion curves. We processed over 10,000 vertical seismograms and obtained more than 3000 fundamental mode Rayleigh waves. Figure 4 shows the path coverage of all the Rayleigh

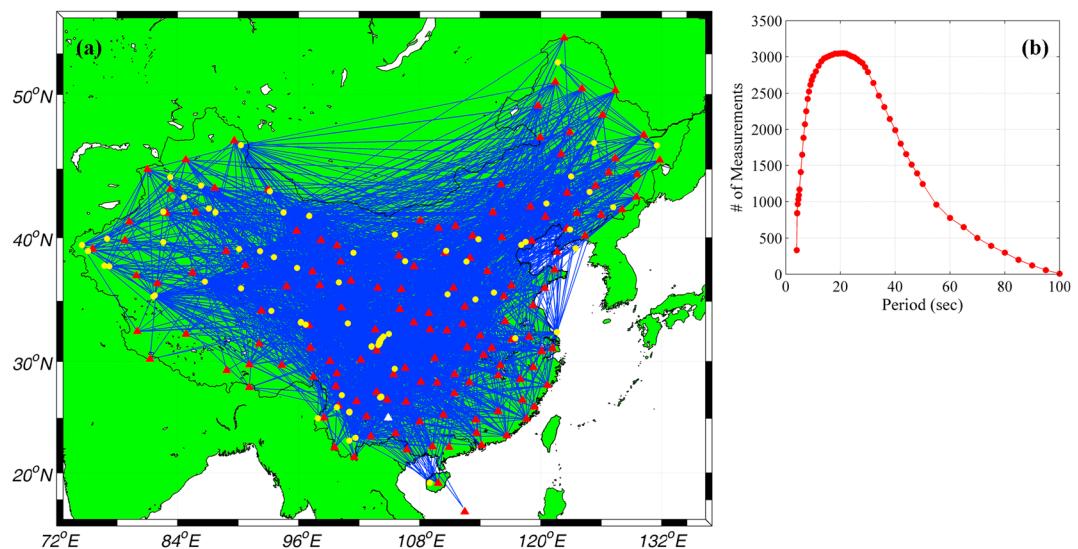


Figure 4. (a) Raypath coverage of Rayleigh wave at period of 20 s. The red triangles are the 144 national stations used to perform dispersion measurements with recordings of regional earthquakes (yellow circles). (b) The number of Rayleigh wave group velocity measurements is plotted as function of period.

wave group velocity measurements at a period of 20.0 s, together with the numbers of measurements at periods from 4.0 to 100.0 s. Overall, we have enough ray coverage in intermediate periods for the 2-D tomographic inversion.

2.2.2. Two-Dimensional Tomography of the Rayleigh Wave Group Velocity

We used the fast marching surface tomography (FMST) technique developed by Rawlinson and Sambridge [2005] to conduct the 2-D tomographic inversion of Rayleigh wave group velocity at each period. The FMST employs a fast-marching method (FMM) for the forward modeling part and a subspace inversion scheme for the inversion part. Repeated application of FMM and subspace inversion allows full account of the nonlinear relationship between velocity and travel time. The travel times from point sources to receivers are computed in 2-D spherical shell coordinates. The inversion step allows both smoothing and damping regularization to be imposed in order to address the nonunique problem of solution.

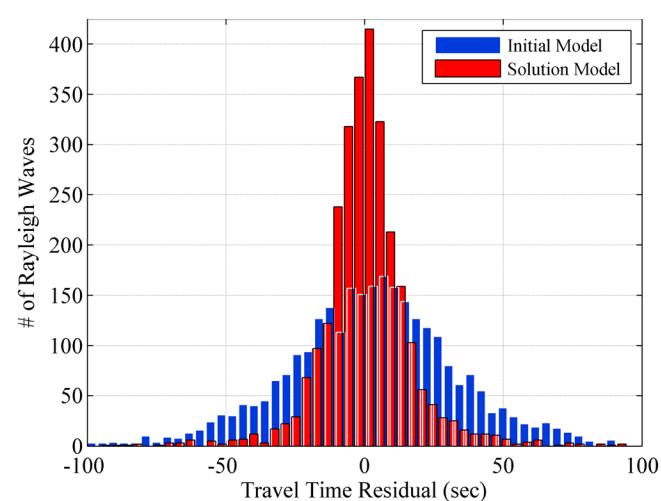


Figure 5. Histograms of travel time residuals of Rayleigh waves with respect to the initial model and the solution model are shown in blue and red, respectively, for comparison.

We chose the average of all measured Rayleigh wave group velocities as the starting model and ran the FMST several times for an inversion. After an initial running, we located and removed the event-station pairs with very large travel time residual. Large residuals can be caused by numerous reasons, and one of them is the location error of seismic events. The cleaned observation data were input into FMST for the next running. Figure 5 compares the histograms of travel time residuals between the initial and solution models. Because we use the average velocity of all observation as the initial model, the travel time residual distribution is of Gaussian type centered at zero. The residuals of the initial model show a fairly flat distribution between

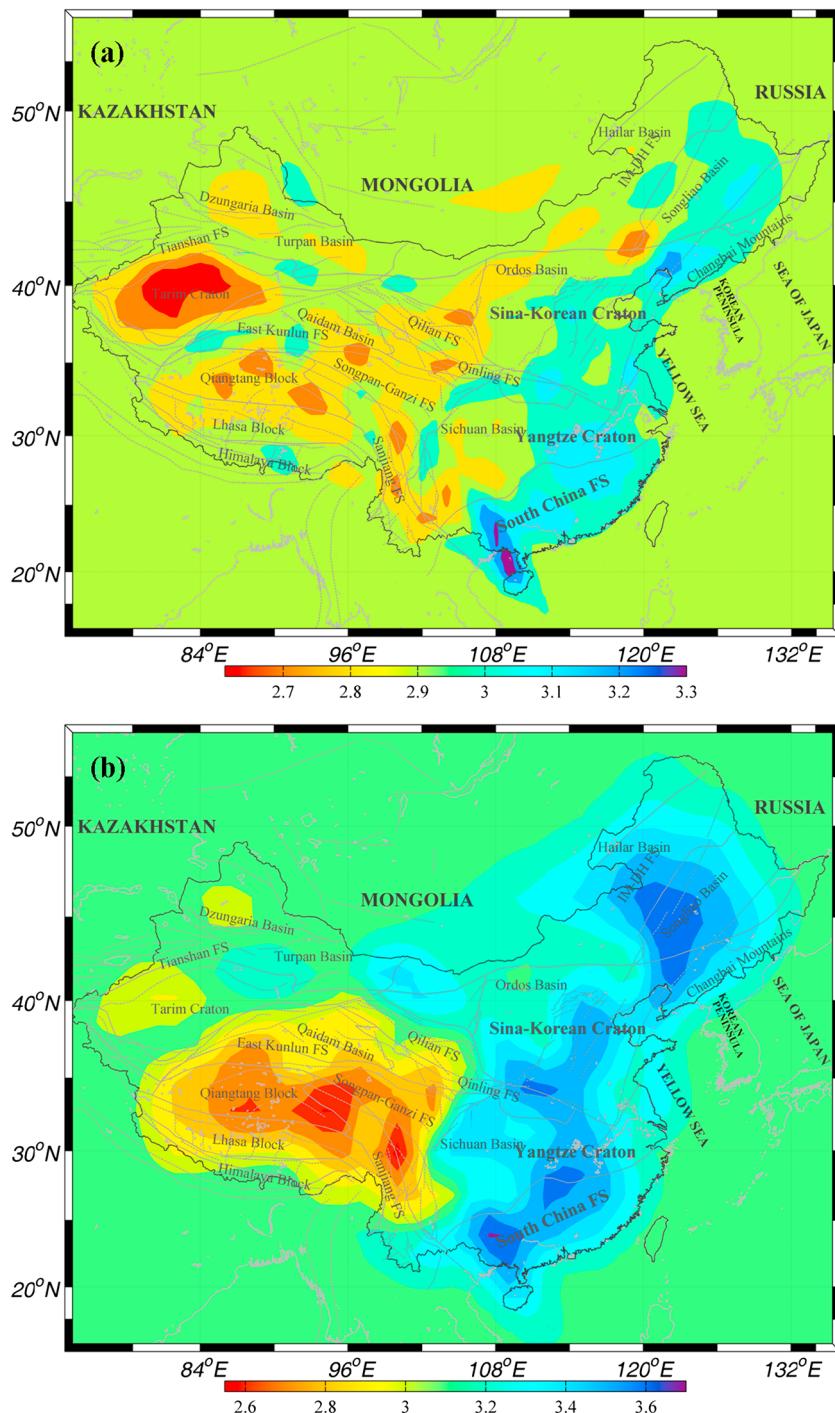


Figure 6. Lateral variations of Rayleigh wave group velocity at periods of (a) 24 s and (b) 36 s inverted with the FMST package.

–100 and 100 s with a standard deviation of 30.6 s, which decays significantly after a few iterations and reaches to 16.2 s in the final model.

Using FMST, we generated Rayleigh wave group velocity tomographic models at periods from 4.0 to 100.0 s. Based on the raypath counts shown in Figure 4b, we used a cell size of $1^\circ \times 1^\circ$ for the periods of 5.0–55.0 s and $1.5^\circ \times 1.5^\circ$ for the other periods, as suggested by checkerboard tests. Figures 6a and 6b show the Rayleigh

Table 1. Grid Search Scheme With Variable Crustal Thickness Defined Between 20 and 92 km

Layer	Thickness (km)	V_s (km/s)	ΔV_s (km/s)
1	$H1 = 2\text{--}21$	$V1 = 2.4\text{--}4.2$	0.3
2	$H2 = 2\text{--}36$	$V2 = 2.6\text{--}4.1$	0.3
3	$H3 = 5\text{--}41$	$V3 = 3.4\text{--}4.3$	0.3
4	$H4 = 5\text{--}41$	$V4 = 3.6\text{--}4.2$	0.3
5	-	$V5 = 4.2\text{--}4.6$	0.1

basins such as the Tarim, Qaidam, and Dzungaria basins in western China and the Ordos and Sichuan basins in central part of China (Figure 6a). In contrast, the South China fold system and the stable Yangtze craton are featured by high-velocity structures. At longer periods (≥ 36 s), the models reflect structures across the Moho down to the upper mantle. As shown in Figure 6b, the group velocity tomography displays a striking bimodal distribution with high velocity in the east and low velocity in the west. This bimodal pattern corresponds very well with our crustal thickness model that exhibits thin crust in the east and thick crust in the west.

We employed a bootstrap approach to estimate uncertainties in the solution models [Efron and Tibshirani, 1986]. In general, the uncertainties in the intermediate periods (15–55 s) are about 0.02–0.05 km/s, similar to those found by other studies [e.g., Julià *et al.*, 2000]. The model uncertainties in other periods, particularly in shorter periods, are relatively larger than those in the intermediate periods. However, they are expected to have little influence on our joint inversions since the weights we assign to short-period (≤ 15 s) dispersion data are much lower than those of intermediate and long periods (> 15 s). Therefore, we expect the upper crustal structure to be mainly constrained by receiver function data, which will be further discussed in section 3.3.

We notice that some boundary regions around China are not well covered by Rayleigh wave data, which may affect the inverted dispersion data for the very few stations located inside these regions. The obtained group wave velocities at most stations are expected to be robust for the following two arguable reasons: (1) the obtained group velocities are consistent with the results of previous studies [e.g., Huang *et al.*, 2003; Li *et al.*, 2013], which used much denser data coverage and (2) lateral variations of the obtained group velocities exhibit a good agreement with major tectonic terranes inside China.

3. Joint Inversion and Initial Model

3.1. Initial Model With Multiconstraints

We carefully built an initial model for each station to account for the strong dependence of linearized inversion on the initial model. We first adopted a grid search method [e.g., Sandvol *et al.*, 1998; Tkalcic *et al.*, 2006, 2011] to create a preliminary crustal S wave model for each station. Specifically for our study region, we designed a grid search scheme with a model space composed of four crustal layers plus a half-space mantle that covers the upper and lower bounds of the crustal parameters for Eastern Eurasian. The details of the grid search scheme are summarized in Table 1. The crustal thickness varies between 20 and 92 km, with a step of 1–2 km. The variation of S wave velocity in each crustal layer contains the possibility of high- or low-velocity zone in the crust. The values of S wave velocities in upper mantle are selected based on the P_n and S_n tomographic models [Sun and Toksoz, 2006; Pei *et al.*, 2007].

We generated totally 2,237,760 models and computed their corresponding synthetic receiver functions using the grid search scheme shown in Table 1. Because the RPB stacking has explicitly defined ray parameter, we only computed synthetics with one ray parameter of 5.6 s^{−1}. Since the uppermost mantle is constrained by P_n and S_n tomographic models, we only use receiver function data to constrain crustal model at this step. The waveform misfit is measured by variance reduction (VR) defined as

$$VR = 1 - \sqrt{\frac{\sum(d_{\text{obs}} - d_{\text{syn}})^2}{\sum d_{\text{obs}}^2}}, \quad (1)$$

where d_{obs} and d_{syn} are the observed and predicted values of receiver functions, respectively. The summation is taken over the entire time window, including both the P_s conversion and reverberation segments, of an individual receiver function.

group velocity tomographic models at periods of 24 and 36 s, respectively. Major anomalies in the group velocity maps are well correlated with large-scale tectonic structures of China. At the period of 24 s, Rayleigh waves are primarily sensitive to structure between 10 and 40 km in depth. The low velocities are associated with major

The procedure of constructing the initial models with multiple sources is summarized as follows:

1. The S wave velocity model for the crust is selected from grid search database and, the corresponding P wave velocity and density are calculated using the V_p/V_s ratio model derived from our advanced $H\text{-}\kappa$ analysis and Birch's law [Birch, 1961], respectively.
2. The P_n and S_n models from tomographic studies [Sun and Toksöz, 2006; Pei et al., 2007] are used as the uppermost mantle velocities.
3. The velocities of the rest of the upper mantles down to 500 km are copied from the global AK135 model [Kennett et al., 1995].
4. The Moho is fixed at the depth determined by the advanced $H\text{-}\kappa$ analysis. Steps 2 and 4 serve as extra criteria to rule out more than 95% of the models in the grid search database. The other 5% of the models are then through waveform fitting selection by calculating their VRs to the observed receiver function. To further narrow down the plausible models, we visually reviewed the top 20 to 50 synthetics with the highest VRs by comparing their first (direct P wave) and second (0p1s) peaks, as well as the later cycles (2p1s and 1p2s) to observed receiver function.
5. Finally, we reparameterize the crust and the uppermost mantle with thin constant velocity layers. Although the waveform misfit decreases with increasing number of layers in a linearized inversion, we do not want to overparameterize our model space with unnecessary number of layers. Using ~ 0.5 Hz low-pass filtered receiver function data, Tkalcic et al. [2011] found that the waveform fitting increases with reducing layer thickness but saturates at around 2.5 km. We thus divided the crust and another 20 km of the uppermost mantle into layers of 2 or 3 km thick. The rest of the upper mantle structures down to 500 km are divided into layers with a thickness varying from 12 to 50 km with increasing depth.

3.2. Sedimentary Layer

It has been known that many areas in China are covered by thick sediment due to its complicated tectonic evolution history [e.g., Xu, 2007; Zheng et al., 2012; Tao et al., 2014]. If unconsolidated sediments are present below a station, the strong reverberations generated at its base could mask the P_s conversions at deep structure, such as the Moho. The velocity structure of the sediment must be included in the initial model because the thickness of the first layer ($2 \sim 3$ km) is usually thicker than the unconsolidated sedimentary layer, limiting the ability to confine the sedimentary structure through the inversion.

Our raw receiver function data also show clear evidence of the presence of unconsolidated sediments beneath a number of stations. As the response of layered structure beneath a receiver, a receiver function is expected to be a series of Delta functions. If only the P_s converted phase of the first velocity jump below the surface is considered, the receiver function can be approximately expressed as [Ammon, 1991]

$$r(t) = \frac{r_0}{z_0} [\delta(t) + A_s \delta(t - t_s)], \quad (2)$$

where r_0 and z_0 are the amplitudes of P wave recorded at the radial and vertical components, respectively, A_s is the transmission coefficient of the P_s conversion at the bottom of the first layer, and t_s is the relative arrival time of P_s phase with respect to the direct P wave. This formula shows that the maximum amplitude of an R/Z receiver function is equal to $\frac{r_0}{z_0}$, synchronous with the arrival of P wave at $t=0$. However, when there is a thin sedimentary layer with very low velocity beneath the surface, the interval between the P_s phase that originated from the bottom of the thin sediment layer and the direct P wave, t_s , becomes so small that the P_s phase is mixed with the P phase. The overall effect looks as if the arrival of the P wave is delayed by a short time (<1.0 s) from time zero. Consequently, the short time delay of direct P wave on a receiver function can be regarded as an effective indicator of the existence of soft sediments beneath a station.

The stations with identified "abnormal" receiver functions are plotted as red cross signs in Figure 7. The size of the sign is proportional to the length of delay time. The delay times of most stations are between 0.1 and 0.4 s, implying sediment thickness of less than 2 km. Most of the "abnormal" stations are located within basins and alluvial plains, such as the Songliao, Tarim, Odors, and Sichuan basins, and the North China plain. We noticed that a few stations inside the Tibetan plateau also showed a short P wave delay (~ 0.1 s), indicating the presence of thin and soft sediments in certain parts of the plateau. Since the P wave delay time is affected by both sediment thickness and velocity, it is impossible to constrain sediment structure from the P wave

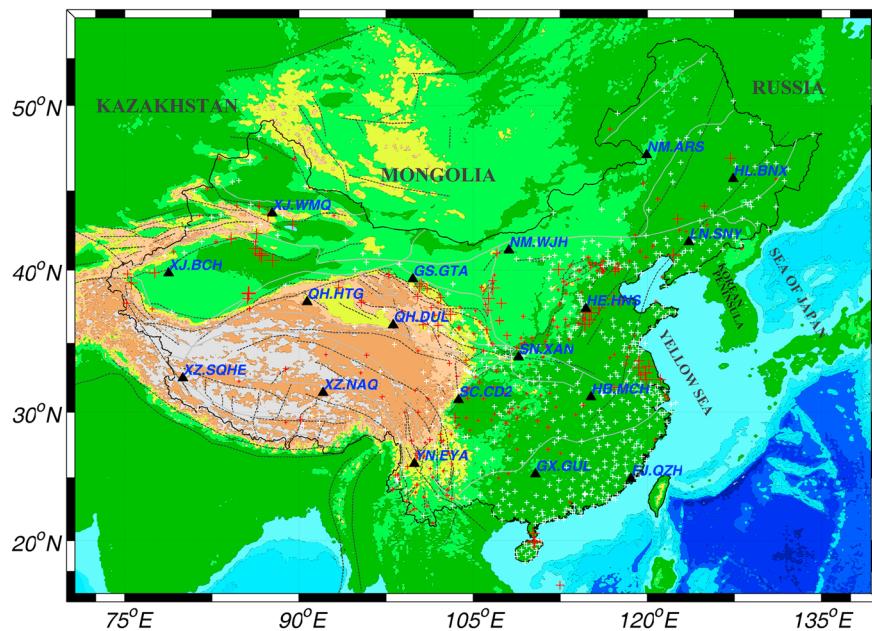


Figure 7. Map shows the location of the stations with normal (white crosses) and abnormal (red crosses) receiver functions. The abnormal receiver function data show a significant delay in P arrival time, which is indicated by the size of the red crosses. The map also shows the geographic locations of the 18 example stations across China (black triangles).

delay time alone. We thus first calibrated the P wave delay time with sediment thickness at some stations estimated by previous studies [e.g., *Shen and Zhou*, 2012; *Tao et al.*, 2014] and then build sediment models from the P wave delay time as shown in Table 2. For those stations with identified P wave delays, the top 1–2 km in their initial models was replaced with the sediment structure.

3.3. Linearized Iterative Inversion With Enhanced Preconditioning

We chose the linearized iterative inversion method based on the least squares technique to extract S wave velocity structure beneath a seismic station. Using the “jumping” scheme to implement the linearized inversion [*Ammon et al.*, 1990], we formulate the linear system as

$$\mathbf{D} \cdot \mathbf{m} = \mathbf{r} + \mathbf{D} \cdot \mathbf{m}_0, \quad (3)$$

where \mathbf{m} is the model, and \mathbf{r} is the residual between observables and synthetics. The matrix \mathbf{D} is the sensitivity kernel defined as the partial derivatives of synthetics with respect to the model parameters. For receiver function, they are computed numerically using the efficient forward perturbation algorithm developed by *Randall* [1989], whereas for Rayleigh wave, they are computed using Haskell propagator matrix [*Herrmann*, 2006]. Similar to the initial model, we used the petrophysical relationships of the empirical Birch’s law and Poisson’s ratio to tie density and P velocity to S velocity, leaving S velocity the only independent model variable in our inversion. The model is updated gradually by solving equation (3) iteratively.

In *Chen and Niu* [2013], we optimized the linearized iterative inversion of receiver functions with enhanced preconditioning on both model and data. The extensive synthetic tests as well as the application on field data demonstrated its effectiveness in improving the linearized inversion process and in reducing the nonuniqueness of the result. Generally, the enhanced preconditioning on receiver function data included the following: (1) Add smoothness to

Table 2. Parameter of Sedimentary Layer Used in Constructing Initial Model

Delay Time (s)	Layer Thickness (km)	V_s (km/s)	V_p/V_s	Density (g/cm^3)
0.1–0.2	0.5	0.9	2.5	1.30
	0.5	1.8	2.5	2.00
0.3–0.6	0.5	0.7	2.5	1.30
	0.5	1.4	2.5	1.65
≥ 0.7	0.5	1.8	2.0	2.00
	1.0	1.3	2.5	1.30
	1.0	2.0	2.0	2.00

the models, (2) fix the deep structure in the upper mantle, (3) balance the receiver functions filtered by Gaussian functions with different width factors, and (4) assign different weights to P_s conversion segment and later cycles. Details on the inversion technique can be found in Chen and Niu [2013]. We employed similar regularizations on the surface wave data in the joint inversions. Specifically, we divided the Rayleigh wave group velocity dispersion curve into two-period segments and assign different weights to account for their variance and impact on the inversion. The complete linear system can be expressed as

$$\begin{bmatrix} \vdots & \vdots \\ (p_{\text{rf}}/a) & \mathbf{D}_{\text{rf_seg1}} \\ (q_{\text{rf}}/a) & \mathbf{D}_{\text{rf_seg1}} \\ \vdots & \vdots \\ p_{\text{sw}} & \mathbf{D}_{\text{sw_seg1}} \\ q_{\text{sw}} & \mathbf{D}_{\text{sw_seg2}} \\ \vdots & \vdots \\ \mathbf{S} & \Delta \\ \mathbf{W} & \end{bmatrix} \cdot \mathbf{m} = \begin{bmatrix} \vdots & \vdots \\ (p_{\text{rf}}/a) & \mathbf{r}_{\text{rf_seg1}} \\ (q_{\text{rf}}/a) & \mathbf{r}_{\text{rf_seg1}} \\ \vdots & \vdots \\ p_{\text{sw}} & \mathbf{r}_{\text{sw_seg1}} \\ q_{\text{sw}} & \mathbf{r}_{\text{sw_seg2}} \\ \vdots & \vdots \\ 0 & 0 \\ \end{bmatrix} + \begin{bmatrix} \vdots & \vdots \\ (p_{\text{rf}}/a) & \mathbf{D}_{\text{rf_seg1}} \\ (q_{\text{rf}}/a) & \mathbf{D}_{\text{rf_seg1}} \\ \vdots & \vdots \\ p_{\text{sw}} & \mathbf{D}_{\text{sw_seg1}} \\ q_{\text{sw}} & \mathbf{D}_{\text{sw_seg2}} \\ \vdots & \vdots \\ 0 & 0 \\ \end{bmatrix} \cdot \mathbf{m}_0, \quad (4)$$

where the subscript rf refers to receiver function data, whereas sw refers to surface wave dispersion data, and the subscripts, seg1 and seg2, indicate the two segments of receiver function and surface wave dispersion data, respectively. Apparently, equation (4) is able to contain multiple receiver functions and surface wave dispersion curves in one simultaneous inversion. It also provides us large flexibilities to adjust free parameters for different purpose. The matrix Δ is a double-difference operator applied on model \mathbf{m} that implements the second-order-difference smoothing scheme on the model, and the diagonal matrix \mathbf{S} controls the trade-off between the waveform fitting and the model smoothness. If a depth-independent smoothness-controlling factor is used, \mathbf{S} degrades to a scalar s , in which $\mathbf{S}=s\mathbf{I}$ (\mathbf{I} is the identity matrix). $\mathbf{W}=w_{ii}\mathbf{I}$ is a diagonal matrix of weighting factors to fix the model layers to the predetermined velocity values contained in \mathbf{m}_0 . For free layers, the corresponding weighting factors w_{ii} are zeros, and the corresponding equations are in fact excluded from the linear system. For a layer that requires to be fixed to some extent, w_{ii} is assigned to a value larger than zero. The choice for the weighting factor is rather subjective, generally in the range of 1.0 to 20.0, as long as it is large enough to limit changes of a fixed layer.

The weighing factors, p_{rf} and q_{rf} are assigned to the P_s and the multiples of a receiver function trace, whereas p_{sw} and q_{sw} are assigned to short-period and long-period segments of a Rayleigh wave group velocity dispersion curve, respectively. Because each data set and segment is characterized by its own physical unit, variance, and number of samples, we define the four weighting factors as [Julia et al., 2000]

$$\begin{aligned} p_{\text{rf}} &= \sqrt{\frac{c_1}{N_{\text{rf_seg1}}\sigma_{\text{rf_seg1}}^2}}, \quad q_{\text{rf}} = \sqrt{\frac{c_2}{N_{\text{rf_seg2}}\sigma_{\text{rf_seg2}}^2}}, \\ p_{\text{sw}} &= \sqrt{\frac{c_3}{N_{\text{sw_seg1}}\sigma_{\text{sw_seg1}}^2}}, \quad q_{\text{sw}} = \sqrt{\frac{c_4}{N_{\text{sw_seg2}}\sigma_{\text{sw_seg2}}^2}}, \end{aligned} \quad (5)$$

where c_1 , c_2 , c_3 , and c_4 are influence factors used to control the relative importance of different segments of receiver function and surface wave dispersion curve, respectively. The summation of these four free parameters is subject to unit constraint. All σ^2 's and N 's, as indicated by their subscripts, are data variances and the number of samples in different segments, respectively. If no noise is present, the complementary nature in receiver functions and surface wave dispersions provides strong constraints on the S wave velocity profile. When noise is present in the data, the constraints on model parameters of the two types of data can become incompatible; thus, proper weighting that takes account of data variance and number of data samples is preferred, as it can effectively balance the impact of the two data sets in the joint inversion.

The uncertainty of receiver function data can be roughly estimated from the transverse component, and the typical value for most receiver functions (both segments) in our data set is about 0.01. The Rayleigh wave group velocity tomographic models suggest that the uncertainties of the intermediate to long periods (>15 s) group velocities are smaller than those of the shorter periods (≤ 15 s). We thus divided the Rayleigh wave group velocity dispersion curve at 15 s and set $\sigma_{sw_seg1} = 0.05$ km/s and $\sigma_{sw_seg2} = 0.03$ km/s for the short and long periods, respectively. These two values are consistent with the uncertainties in our Rayleigh wave group velocity tomographic models and are tested by trial-and-error experiments at several stations.

As mentioned before, we expect to solve the S velocity gradients mainly by receiver functions, whereas the absolute S velocity by surface wave dispersion. For this reason, the influence factors c_1 to c_4 must be carefully chosen to avoid yielding solutions that are dominated by any single data set. In order to solve shallow structure from receiver function data, we assigned more weight to the P_s segment than the later cycles. For the Rayleigh wave group velocity dispersion data, we gave more weight to the long-period segment than the short-period segment to better constrain the deep structure. The influence factors are thus chosen to ensure $p_{rf} > q_{rf}$ and $q_{sw} > p_{sw}$. The value of p_{rf} is set slightly larger than q_{sw} , and so does q_{rf} to p_{sw} , but, generally, they are approximately at the same level.

4. Results and Discussion

The techniques discussed in the previous sections were applied on all +1000 CNDSN stations to invert for 1-D S wave velocity profile of the crust and upper mantle. For receiver functions of each station, we produced six RPB stacks as observables using reference ray parameters of $p_0 = 5.5, 6.0, 6.5, 7.5$, and 8.0 s/ $^{\circ}$ and the harmonic mean of the ray parameters of all stacked data. The stacking was carried out on two frequency bands with Gaussian filter width parameter $a = 1.0$ and 2.5 , respectively. Including high-frequency data helps to reveal small-scale structures that are sensitive to high-frequency seismic waves. Therefore, the total number of receiver functions is 12 (or 10 if the harmonic mean of the ray parameters of all stacked data is equal to any value of $5.5, 6.0, 6.5, 7.5$, or 8.0 s/ $^{\circ}$). The purpose of including multiple receiver functions from different distances is to enhance convergence of the inversion, although it cannot efficiently resolve the trade-off between depth and velocity [Chen and Niu, 2013].

The misfit between observable and synthetic receiver functions is computed using the VR (equation (1)). While for surface wave dispersion, the root-mean-square (RMS) is a more appropriate measure of misfit as discrete dispersion measurements have much less samples than the continuous receiver function data.

$$\text{RMS} = \sqrt{\frac{\sum(d_{\text{obs}} - d_{\text{syn}})^2}{N}}, \quad (6)$$

where N is the number of selected periods for a Rayleigh wave dispersion curve, and the summation is taken over the entire data window. The maximum of VR and the minimum of RMS indicate the best fit of observables and synthetics for receiver function and surface wave dispersion, respectively. The VR and RMS are not used as thresholds to terminate iterations of an inversion but serve as effective criteria used to decide the optimal model from an inversion with determined parameters.

Most parameters in equation (4) are not necessarily changed from station to station once they are determined empirically. The only variable parameter is the smoothness-controlling factor, which can vary significantly in yielding the model with the best waveform fitting from station to station. Ideally, we want to apply a depth-dependent smoothness-controlling matrix \mathbf{S} in an inversion and reduce the smoothness at the depth range with large velocity gradients while raising smoothness at other depths. However, it is difficult to decide the values of \mathbf{S} empirically as the choice of \mathbf{S} depends on data quality and subsurface structures. In practice, we simply used a depth-independent smoothness-controlling scalar s and conducted individual inversions using a wide range of s values. We found that the smoothness factors between 0.05 and 2.0 cover the range that gave the best waveform fitting for both data sets at most stations.

The solution model error is the direct mapping of data errors. However, it is difficult to quantitatively estimate model error because the accurate estimate of data errors is usually unavailable. The data errors not only can come from the computation of individual receiver functions and measurement of Rayleigh wave group velocity dispersion curves but also can be introduced by receiver function stacking and group velocity

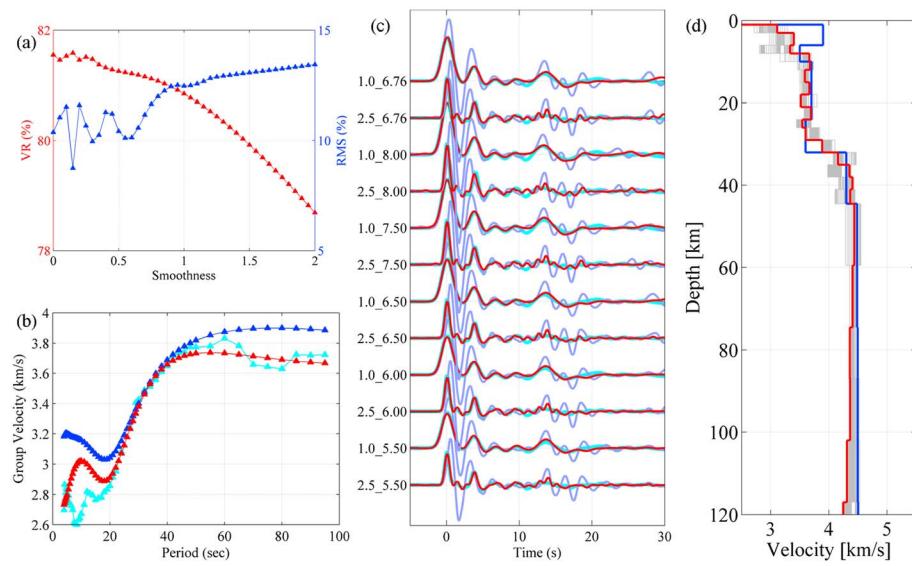


Figure 8. Results of the joint inversion at HE.HNS. (a) VR of receiver function misfit (red triangles) and RMS of Rayleigh group velocity dispersion misfit (blue triangle) are shown as functions of the smoothness parameter. See the main text for the explanation of the smoothness parameter. (b) The observed Rayleigh wave group velocity dispersion (cyan) is plotted with the synthetic dispersions computed with the initial (blue) and the solution model (red). (c) The observed receiver functions stacked with different ray parameters and filtered with different Gaussian filters (cyan) are shown together with synthetic receiver functions computed from the initial (blue) and the solution model (red). The α value of the Gaussian filter and the ray parameter (in s^{-1}) are labeled on the left side of the receiver functions. (d) S velocity profile versus depth. The solution model is plotted in red and the initial model is plotted in blue. The grey lines indicate the intermediate models during all iterations using all smoothness parameters.

tomographic models. Those typical values used in configuring weighting factors cannot include all data variances precisely. On the other hand, it seems that the intermediate solutions generated under acceptable smoothness parameters during all the iterations provide a reasonably well sample of the model space, thus delimiting the lower and upper bound of the final model. They approximately reflect the model uncertainty, and we take the weighted average of the models inverted with all acceptable smoothness parameters as the optimal model and their standard deviation as the model uncertainty.

4.1. One-Dimensional S Wave Velocity Profiles

We obtained 1-D S wave velocity profile at all the +1000 stations. Below, we show the results of several stations located at different tectonic units (Figure 7). All the selected stations belong to the national network as they usually have better data quality than regional stations.

4.1.1. Station HE.HNS

The results of the joint inversion at station HE.HNS are shown in Figure 8. The variation of VR of receiver functions and RMS of Rayleigh velocity dispersion are plotted in Figure 8a as functions of smoothness factor, s . The data fit of Rayleigh velocity dispersion and receiver functions are plotted in Figures 8b and 8c, respectively. The final velocity profile (red line) is plotted in Figure 8d, together with the initial model (blue line) and the intermediate results inverted with all acceptable smoothness factors in all the iterations (grey lines). With a smoothness parameter $s < 0.55$, the inversion appears to be unstable, which is characterized by rapid and complicated variations in VR and RMS (Figure 8a). The inversion becomes stable when s increases to 0.6 and the resultant models have ~5%–8% variations from the optimal model.

HE.HNS is located at the central NCB (Figure 7). Comparing to the model obtained from receiver function data alone [Tkalić et al., 2011], we found that both models revealed a clear thin (~1–2 km) sedimentary layer at the surface, a sharp discontinuity at about 8 km, and identified gradient Moho structure at 32–33 km. The improvements through the joint inversions can be found in the middle crust and upper mantle. In our model, the middle crustal structure is relatively simple, which remains nearly constant at about 3.7 km/s from 8 to 30 km. There are small alternative jumps in this depth range (~8–30 km). Since they are smaller than the model uncertainties (Figure 8d), they probably do not reflect the true structures but are more likely the

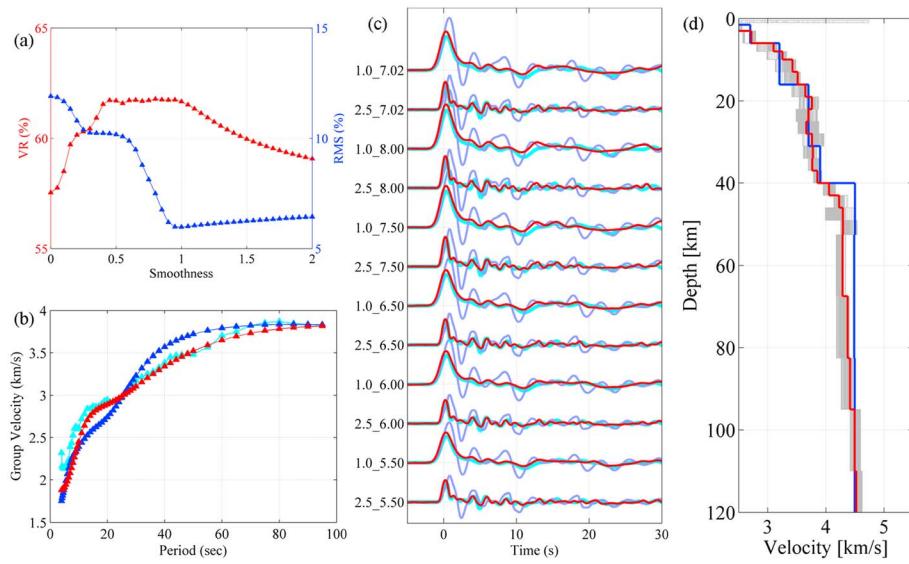


Figure 9. The same as Figure 8 but for the joint inversion results at the station SC.CD2.

artifacts resulting from model parameterization and smoothness control. The receiver-function-based model, however, shows a gradual increase of S wave velocity in the same depth range. It also has a lid-like structure in the upper mantle between 50 and 70 km, which disappears in our model. We speculated that the lid-like feature is an artifact due to inappropriate data stacking and lack of sufficient regulation on inversions in the work of Tkalcic et al. [2011].

4.1.2. Station SC.CD2

Station SC.CD2 is located near Chengdu, the provincial capital of Sichuan province and is only 25 km away from the epicenter of the 2008 Wenchuan earthquake. It is roughly located at the boundary of the Tibetan plateau to the west and the Sichuan basin to the east (Figure 7). The crustal structure beneath this station is believed to vary significantly with back azimuths. Since all the earthquakes used in generating receiver functions come from the southeast, the resultant model is thus expected to mainly reflect the structure of the Sichuan basin. We notice a ~0.4 s P wave delay on the receiver function data (Figure 9c), suggesting that it is underlain by a thick sedimentary layer of ~2–3 km (Figure 9d). The VR increases and RMS decreases with increasing smoothness in a rather complicated way and reach a maximum and a minimum at $s=0.95$

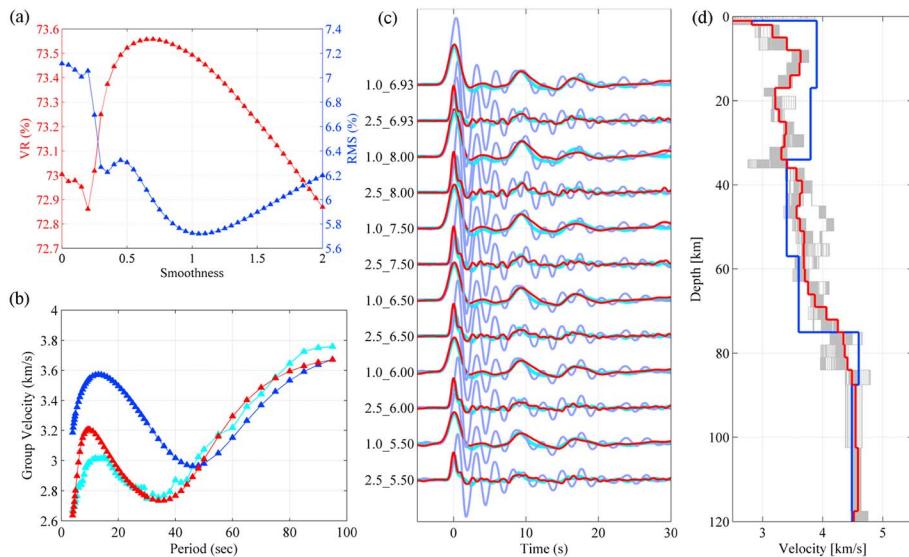


Figure 10. The same as Figure 8 but for the joint inversion results at the station XZ.NAQ.

(Figure 9a), respectively. The complication is likely related to the nature of the inversion, which tries to adjust the sediment structure to fit the data. Consequently, the S velocities in the top 1 km of the crust swing rapidly when the smoothness is not yet strong enough to stabilize the inversion (Figure 9d). The S velocity monotonically increases from the sediment to about 20 km in the middle crust. From middle crust down to the Moho, the velocity is kept nearly constant. The Moho is a gradient structure centered at ~42 km, in good agreement with our $H\text{-}\kappa$ estimate.

4.1.3. Station XZ.NAQ

XZ.NAQ is located on the Qiangtang block of the Tibetan plateau (Figure 7). This part of the Tibetan plateau is likely to have the thickest crust on the Earth. Our $H\text{-}\kappa$ analysis indicates that the crust beneath this station is as thick as ~76 km. The VR reaches to the maximum at $s = 0.75$, while the RMS reaches to the minimum at $s = 1.05$ (Figure 10a). Within a large range of s between 0.5 and 2.0, all inversions are stable and able to yield acceptable models. Although the intermediate results in some layers present large deviations from the final model, they quickly converge to the final model after a few more iterations (Figure 10d). From the surface to the upper crust at about 20 km, the S velocity model suggests a thin sedimentary layer on the top and a high-velocity zone at about 10 km. The rest of crust structure can be roughly divided into two parts, one part from 20 to 35 km and the other part from 35 to 60 km. Within each part, the S velocities are kept nearly constant at about 3.3 and 3.7 km/s, respectively. The velocity below 60 km continues to increase until it approaches to 4.5 km/s in the upper mantle. As a result, the structure of Moho is presented as a transition zone at the center of about 75 km.

4.1.4. S Velocity Profiles of Other Stations

The S velocity profiles of the other 15 stations labeled in Figure 7 are shown in Figure 11. Stations NM.ARS and NM.WJH are located at Inner Mongolia. Their crustal structure is relatively simple without significant jumps. The solution model of NM.ARS indicates a nearly constant crustal and upper mantle structure similar to the initial model, with a sharp Moho at about 38 km. The crustal velocity beneath NM.WJH increases gradually from the surface to the Moho but present a high-velocity zone at ~20 km deep. Stations HL.BNX, LN.SNY, HE.HNS (Figure 8), HB.MCH, GX.GUL, and FJ.QZH are all located in the easternmost China, spreading from northeast China to the South China fold system. The velocity models beneath these stations share a more or less similar pattern throughout the crust. They all exhibit a gradient Moho from about 25 to 35 km, consistent with the observation of a weak P_s on these stations. The Moho depth determined by our advanced $H\text{-}\kappa$ analysis corresponds well with the largest gradient within the transition zone. The crust beneath these stations presents a high-velocity zone in the upper crust followed by a low-velocity zone in the middle crust. The upper mantle also shows a broad low-velocity zone from about 40 to 120 km. These features were not clearly shown on the receiver function data only based models of GX.GUL and FJ.QZH obtained by *Tkalčić et al. [2011]*. The other common difference is that the upper mantle velocities in our models are slightly lower than velocities in their models. These discrepancies are likely caused by the different techniques and data used in the two studies.

Station SN.XAN is located on the Qinling mountain range between the Sino-Korean Craton and the Yangtze Craton (Figure 7). *Tkalčić et al. [2011]* found a very broad crustal mantle transition region from 15 to 50 km. The Moho in our model also has a gradual boundary, but it is much clearer and shallower than their model. This considerable discrepancy very likely resulted from the inappropriate receiver functions stacking, minimum control on inversion, and the lack of surface wave dispersion data in the previous work.

The two northwestern stations XJ.WMQ and XJ.BCH are located at the eastern Tianshan mountain range and the Tarim basin, respectively, and exhibit distinct Earth structure. The velocity model of XJ.BCH is relatively simple. It has a clear and sharp middle/lower crust boundary at ~32 km, separating the middle crust with a velocity of ~3.4 km/s from a 3.7 km/s at lower crust. XJ.WMQ is surrounded by the desert to the south and southeast, from which directions our receiver function data are assembled. We observed a small amount delay in the direct P wave and attributed it to the soft layers beneath the southeast side of the station. The middle to lower crustal structure of XJ.WMQ is quite similar to the previous models [e.g., *Mangino et al., 1999*] but exhibit some differences in the upper crust.

We have six stations located at different parts of the Tibetan plateau, which showed very different velocity structures. GS.GTA is situated on the Qilian fold system at the north boundary of Tibetan plateau (Figure 7). The crust consists a 3.6 km/s middle crust from 10 to 28 km and a 3.8 km/s lower crust from 35 to 50 km. The Moho is a transitional boundary at the depths between 50 and 53 km. QH.HTG is another station located at the Qilian fold system near the western end of the Qaidam basin (Figure 7). The S velocity on

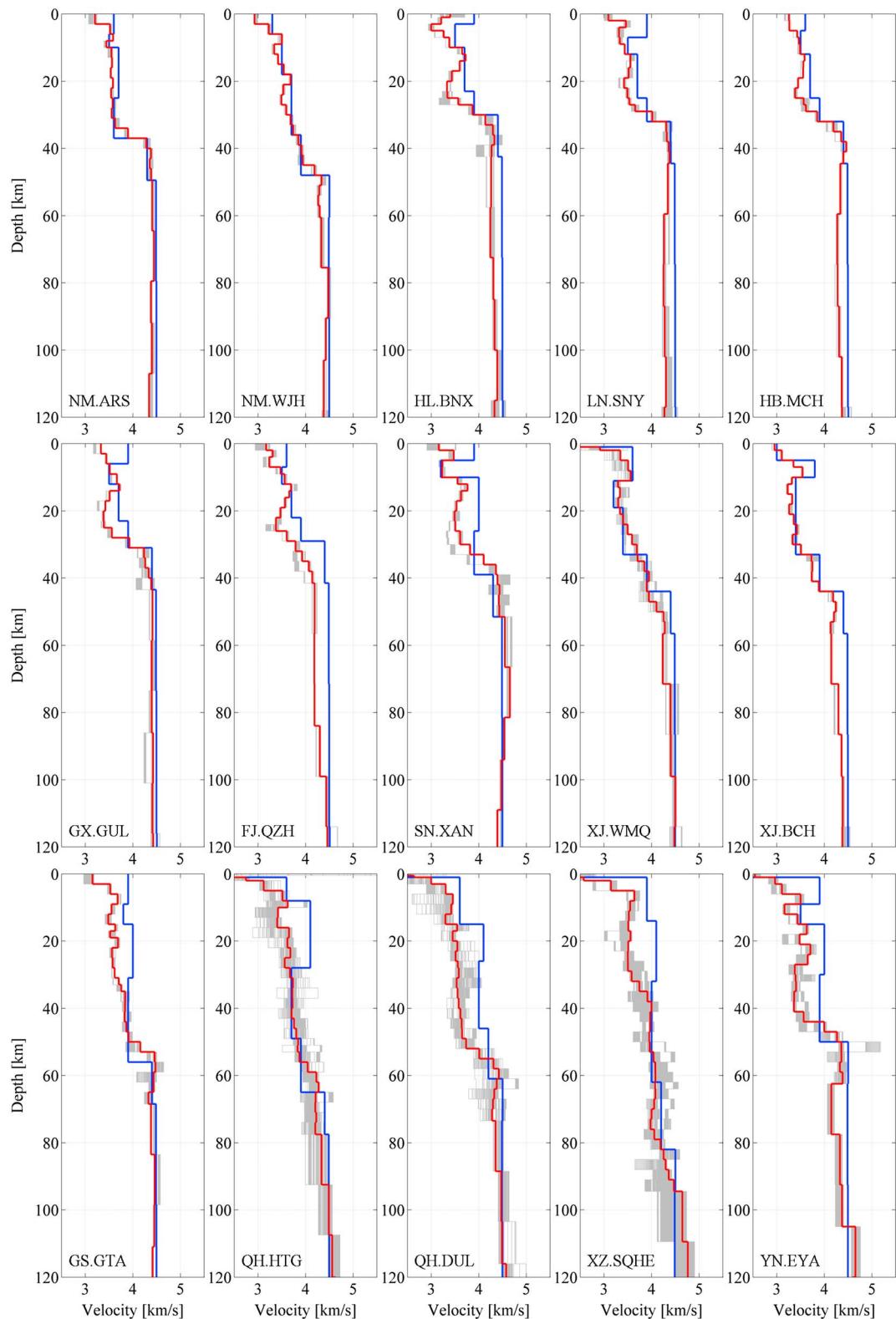


Figure 11. S velocity profiles inverted at the other 15 example stations shown in Figure 7. For each profile, the initial and solution models are plotted in red and blue lines, respectively. The grey lines indicate the intermediate models obtained at different iterations with various smoothness parameters.

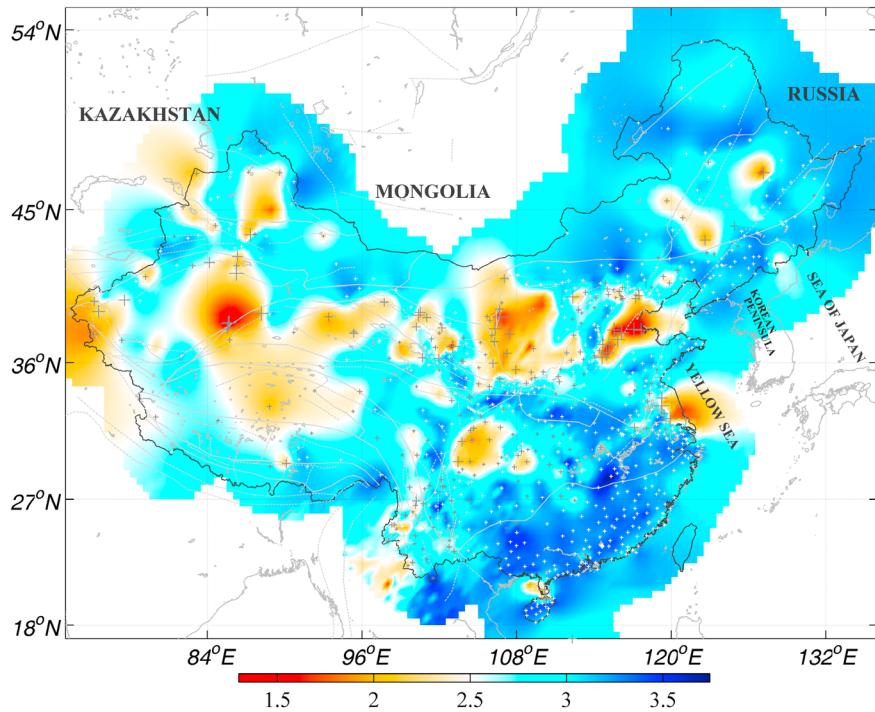


Figure 12. Lateral variations of the S velocities of the topmost 2 km of the crust are shown with CNDN stations. Stations with normal and abnormal receiver functions are plotted in white and red crosses, respectively.

the top of the crust presents apparent sedimentary structure and then rapidly increases to more than 3.5 km/s at about 10 km. From 10 down to 55 km, the velocity remains nearly constant with a high value of ~3.7 km/s. The velocity jump at the Moho is thus not so large and occurs at a depth range of ~6 km centered at 60 km deep. QH.DUL is separated from QH.HTG by the Qaidam basin (Figure 7). The velocity structure of QH.DUL seems to be similar to that of QH.HTG but has a relatively strong velocity jump at the Moho. The velocity model of XZ.SQHE has a similar pattern to XZ.NAQ (Figure 10). Both have a soft sediment layer at the top and a transition zone in the middle of the crust that separates the crust into two parts. The velocities in the middle and lower crust beneath XZ.SQHE are slightly higher than beneath XZ.NAQ, and the Moho beneath XZ.SQHE is ~10 km deeper than that of XZ.NAQ. YN.EYA is located at the center of Sanjiang fold system at the southeast corner of Tibetan plateau (Figure 7). The velocity model beneath this station also shows a sedimentary layer at the top of crust and then gradually increases to ~3.5 km/s in the upper and middle crustal depths. Below ~27 km, the S velocity remains nearly constant at a low value of 3.4 km/s. The Moho has a complicated gradient structure and is few kilometers shallower than the $H\text{-}\kappa$ estimate.

4.2. Inverted 2-D/3-D S Velocity Models

Taking the advantage of the very dense distribution of stations across China, we generated 3-D S velocity models by combining all 1-D models on the +1000 stations. Assuming a constant but unknown mean and enough observations to estimate the variogram, we adopted the ordinary kriging method [e.g., Isaaks and Srivastava, 1989] to interpolate the results onto 0.5° by 0.5° grids covering our study region from 17° to 55° in latitude and from 73° to 136° in longitude. The basic idea behind kriging is to predict the value at a given point by computing a weighted average of the known values in the neighborhood of the point. Because the mean and the covariance function of the random field are usually unknown, the ordinary kriging computes the best linear unbiased estimator of the given point based on a stochastic model of the spatial dependence quantified by the variogram of the random field.

For the final 3-D model, we hope the depth profile maintains velocity discontinuities revealed by all 1-D models, while the horizontal direction is smoothed by interpolation. For this reason, we did not generate 3-D model directly, instead we generated a series of 2-D models at different depths as illustrated in Figures 12–14. In all 2-D S velocity maps, we did not interpolate the points that are far more than 400 km from the nearest sample point.

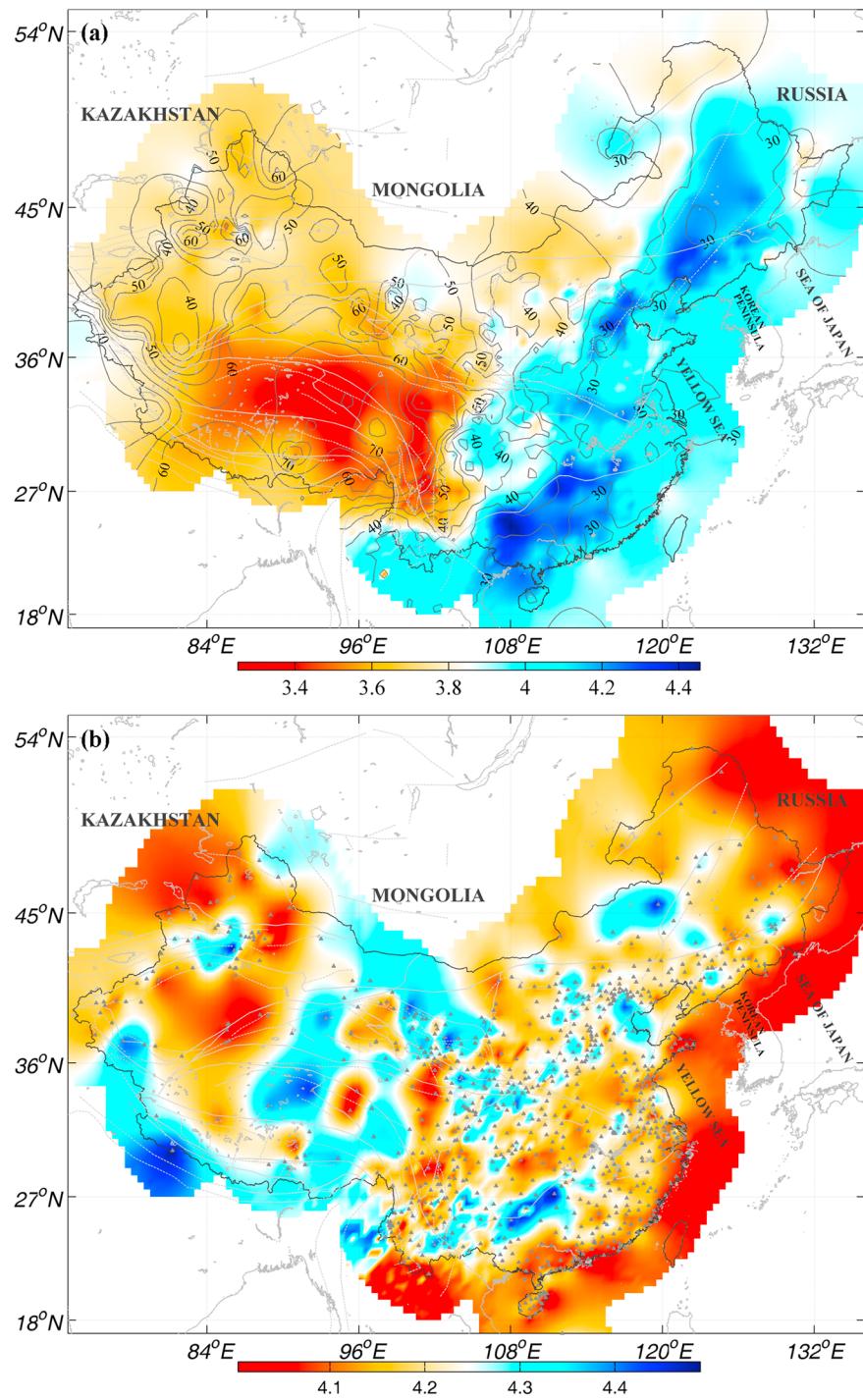


Figure 13. (a) S velocity variations at depth of 35 km are plotted with the contour distribution of Moho depth obtained from receiver function data. (b) Lateral variations of S_n velocities defined as the S velocities at the layer immediately below Moho.

4.2.1. S Velocity Model of the Top Sedimentary Layer

The average velocity structure of the uppermost 2 km crust is shown in Figure 12. The regions with sedimentary structure are generally well correlated with the stable cratons in China, such as the Tarim craton, the Dzungaria basin and the Qaidam basin in west of China, the Odors basin in north China, and the Sichuan basin (part of Yangtze craton) in southwest China. The Songliao basin in northeast China, the north China

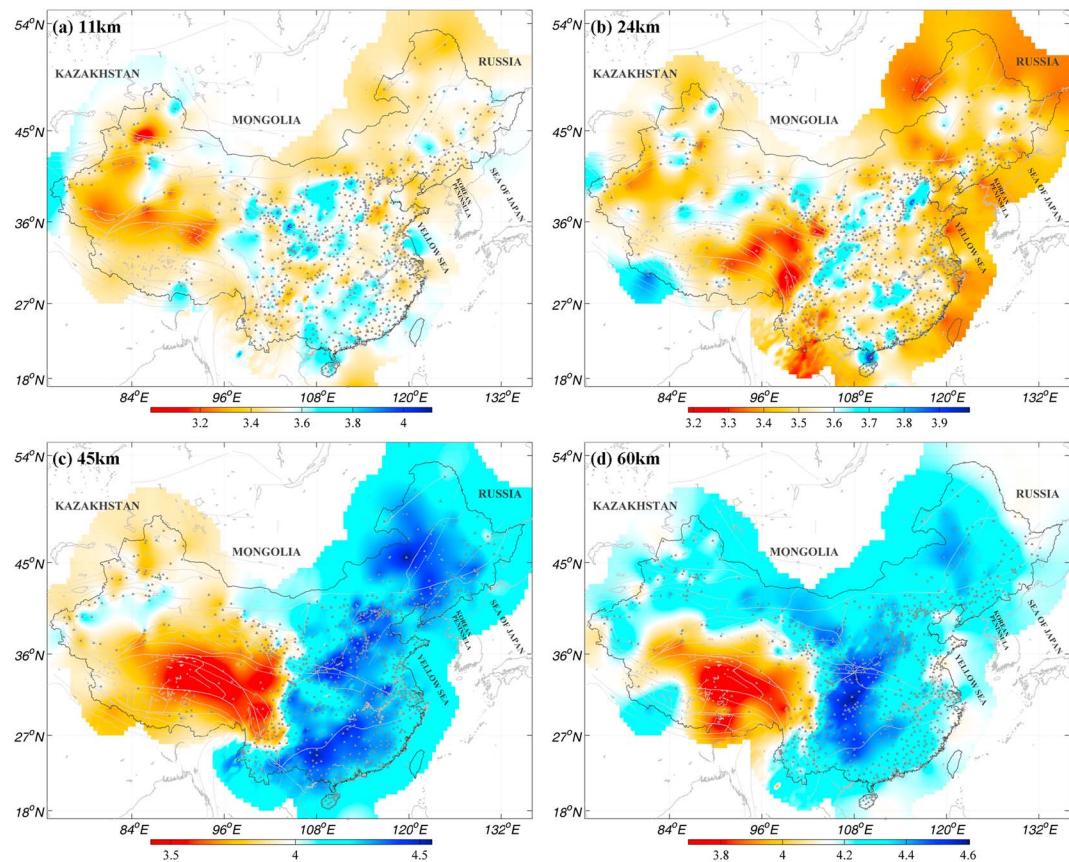


Figure 14. Variations of S velocities at two depths (a) 11 and (b) 24 km within the crust and two depths (c) 45 and (d) 60 km in the uppermost mantle.

plain, and east China plain are also covered by prominent sedimentary layers with very low S velocities. For the rest of the Yangtze craton and the entire south China fold system, the average S velocity is about 3.0–3.5 km/s without prominent signs of sedimentary structure on the top of crust. We also note that a large area in the northern Tibetan plateau, including the Qiangtang block, East Kunlun, and Qilian fold belts, is also covered by a thin sedimentary layer with low velocities between 2.0 and 2.5 km/s. Compared to the other parts of China, the Tibetan plateau is less instrumented; thus, the velocity structure is undersampled and could be biased by the very few stations. On the other hand, we found a good agreement in velocity structure among neighboring stations, suggesting that measurements at these stations can reasonably well represent shallow structure of the plateau.

4.2.2. S Velocity Model at the Moho and S_n Velocity Model

The S velocity structure at 35 km is shown in Figure 13a. It reflects either crustal structure or upper mantle structure depending on the Moho depth of the sampled area. The contour map of the Moho depth, which is calculated from our $H\text{-}\kappa$ measurements at all the stations using kriging interpolation, is also plotted in Figure 13a. The most prominent feature of the seismic structure at this depth is the large velocity contrast between the east and the west bounded by the so-called “N-S trending belt” at $\sim 105^\circ\text{E}$ with complicated shape delimited by our ample samples. Since the crust beneath the eastern China is generally thinner than 35 km, the S velocities of most areas, including the eastern part of the northeast China and the NCB, the Sichuan basin, the Yangtze craton, and the South China fold system reflect the structure of the uppermost mantle with S velocities higher than 4.0 km/s. The Songliao basin in northeast China has the highest S velocities and the shallowest crust comparing to the surrounding areas, in agreement with the rifting, lithospheric thinning, and mantle upwelling in this region. The highest S velocities at 35 km in the south are not located at the eastern coastline of the SCB but exist in somewhere inland. This area corresponds to the thinnest crust beneath China. In contrast, the region to the west of the “N-S trending belt” shows much more variable

crustal properties. The crust in most areas of the western China is thicker than 40 km and drastically increases to more than 75 km beneath the central Tibetan plateau. Therefore, the S velocities in the western China reflect crustal structure with an average of ~3.6 km/s. Moreover, a prominent low-velocity zone in the upper to middle crust beneath central Tibet is also observed. The S velocity is only about 3.4 km/s, ~5% lower than the surrounding areas.

The good agreement of S velocities at this depth with crustal thickness also verifies our crustal thickness model obtained from receiver function data. Moreover, our crustal thickness model is consistent with most recent models that were also derived from dense data sets in continental scale [e.g., Teng *et al.*, 2013; He *et al.*, 2014; Li *et al.*, 2014]. Meanwhile, it is also not surprising to find differences among these models in local scale, largely due to different data types and data frequency ranges and techniques used in these studies. For example, the active source reflection data [e.g., Teng *et al.*, 2013] sample the Moho from the upper side using high-frequency seismic waves, while the receiver functions usually focus on relatively lower-frequency teleseismic waves and illuminate the Moho from the lower side. These differences are expected to yield different estimates on Moho depth in the case of a gradient crust-to-mantle transition. The apparent differences among the receiver function models [e.g., Li *et al.*, 2014] are mainly located at the Tianshan range, western Tibet region, and the Tarim craton, which could be caused by insufficient station coverage, as well as the complexity in crustal and Moho structure.

Using the Moho depths shown in Figure 13a, we selected the S velocity of the layer immediately below Moho as S_n velocity of each station and generated the 2-D S_n velocity map in Figure 13b. In general, the S_n velocities extracted from our 3-D model are about 5% lower than the previous tomographic models [e.g., Pei *et al.*, 2007; Sun *et al.*, 2008]. The lateral variations of S_n velocity in our model are rather complicated; they do not form a very clear bimodal distribution, with a high velocity in western China versus a low velocity in eastern China. Although the eastern China has relatively low S_n velocities, it is embedded with small-scale areas with high S_n velocities. Except for the part of the Dzungaria and Qaidam basins, we did not observe high S_n velocities beneath the major basins in western China [e.g., Pei *et al.*, 2007; Sun *et al.*, 2008]. This is very likely because we are lacking of samples within these basins, such as the Tarim basin. Most western stations are deployed on the Tianshan and Kunlun mountains surrounding the Tarim basin, and the S_n velocity of Tarim basin is actually the result of surrounding stations through interpolation. Probably for the same reason, there also exist large discrepancies in the Tibetan plateau between our and tomographic models. In our model, the Tibetan region does not show a uniformly high S_n velocity as observed by Pei *et al.* [2007] and Sun *et al.* [2008]. Instead, we observed high- S_n -velocity zone beneath east Kunlun and Qilian areas and the southeastern Tibetan. If we assume that the lateral variations of P_n velocities follow the pattern of the S_n velocity variations, our results are somewhat in agreement with the P_n tomographic models of Hearn *et al.* [2004] and Liang *et al.* [2004] beneath the Tibetan plateau. However, we want to emphasize that it might be not suitable to compare these models since they may represent P or S wave velocity at different depths around the Moho. For example, we defined the Moho depth independently in our model, whereas the tomographic studies [e.g., Pei *et al.*, 2007; Sun *et al.*, 2008] inverted for Moho depth simultaneously with S_n and P_n velocities. Given the complexity in Moho structure and propagation of S_n and P_n waves across the continental China, both types of studies may only reflect partial structural response of the Moho and the uppermost mantle.

4.2.3. The Crustal and Upper Mantle S Velocity Model

The strong S velocity variations observed in this study (Figure 14) indicate the existence of significant structural heterogeneities in the crust and uppermost mantle beneath the mainland of China. In the upper crust depth (11 km, Figure 14a), the low-velocity zones focus on the cratons and basins, such as the Songliao, Ordos, Sichuan, Tarim, and Dzungaria basins, and the eastern part of Yangtze craton. Most of them remain low velocities through the crust. The high-velocity anomalies can be seen beneath orogenic zones, such as Qilian and Qinling fold systems in central China, the eastern Tibetan plateau, and the South China fold system in the southeast (Figure 14a). These high-velocity anomalies appear to extend to middle crustal depth (24 km, Figure 14b) except for the South China fold systems. At the depth of 11 km, the Tibetan plateau exhibits distinct patterns with low velocity in the west and high velocity in the east. This pattern seems to be reversed in the middle crustal depth (24 km), with a prominent low-velocity anomaly beneath the eastern Tibetan plateau (Figure 14b).

Figures 14c and 14d show the S velocity structure at depths of 45 and 60 km, mainly in the upper mantle. At these depths and below, the Rayleigh group velocity dispersion data become dominant in the joint inversion

due to the weak response of receiver function data to deep structure. Comparing to the lateral variations in the crust, the large-scale velocity structure is relatively simple in the upper mantle, generally exhibiting a bimodal distribution of a fast east versus a slow west. If we assume the S velocity of the upper mantle is larger than 4.0 km/s, the entire eastern China falls into the upper mantle at depth of 45 km. The average S velocities of most areas of northeast China, NCB, and SCB are about 4.4–4.5 km/s, whereas the eastern coastline and the boundaries of these tectonic blocks have relatively low S velocities of about 4.2–4.3 km/s. The Tarim craton is the only area entering the upper mantle at 45 km deep as it has the thinnest crust in the western China. On the other hand, the entire Tibetan plateau are still situating within the crust at 45 km, thus is featured by a relatively low S wave velocity. This situation is also largely true at 60 km (Figure 14d).

Overall, the velocity images of the shallow layers correlate well with the surface geology, topography, and lithology, while the velocity patterns of the deep layers are in good agreement with the variations of the Moho depth. In general, our models are consistent with previous tomographic velocity models [e.g., Liang et al., 2004; Sun and Toksöz, 2006; Pei et al., 2007; Sun et al., 2008] at large scale, but show considerable differences at local scale. These discrepancies are likely caused by differences in imaging methods, data sets, and resolution between our joint inversion and previous tomographic studies. Such discrepancies also suggest that the crust and upper mantle beneath China are complicated in both regional and local scales. Given the heterogeneity of the stations' location in this study, we may not be able to reveal the structure of the areas without sufficient data coverage. The investigation of the local structure of these areas requires more densely covered temporal deployment data [e.g., Kind et al., 2002; Vergne et al., 2002; Yao et al., 2008; Li et al., 2012; Guo et al., 2015].

5. Summary and Conclusions

In this study, we applied a multistep approach to simultaneously invert the teleseismic receiver function and Rayleigh wave group velocity dispersion data observed at +1000 permanent broadband seismic stations in China. We used the RPB method to generate more reliable receiver function observables for inversion. Rayleigh wave group velocity dispersion curves (4.0–100.0 s) at all stations were extracted from high-resolution Rayleigh wave group velocity tomographic models. The initial model of each station was constructed by integrating variety types of data. For the joint inversion of velocity profile beneath each station, we applied enhanced preconditioning on the linearized iterative inversion to regularize and balance multiple data with different types and frequency bands. Finally, we generated a 3-D S wave velocity model by combining all the 1-D velocity models.

In summary, our S velocity models exhibit the following features:

1. An unconsolidated sediment layer between a few hundreds of meters to 2 km with extremely low S velocities (~1.5–2.0 km/s) are present under the major basins, as well as a large area in the northern Tibetan plateau.
2. Low-velocity zones in the upper crust (11 km deep) focus on the major basins. High-velocity anomalies at this depth are found mainly under major orogenic zones. The correlation disappears in the middle crustal depths (24 km).
3. Due to the large difference in Moho depth between the eastern and western China, the S wave velocity maps at the depth range of 35 to 60 km show a strong velocity contrast between the fast east and the slow west bounded by the so-called "N-S trending belt" located at ~105°E.
4. The S_n velocity of the uppermost mantle, on the other hand, exhibits a fast west and slow east due to the difference in sampling depth at large scale, mixed with some small-scale anomalies.

Acknowledgments

We greatly thank Z. Huang, R. Liu, L. Sun, and other staffs at CENC for providing data service and technical support. We also thank the Associate Editor and one anonymous reviewer for their constructive comments and suggestions, which significantly improved the quality of this paper. This work is supported by the contracts FA9453-10-C-0258 in responding to Broad Agency Announcement and NSFC 41274099. The model data to reproduce the figures and results of this article in text format are available from <http://www.globalseismology.rice.edu/Downloads/MODELS/ChenNiu2016.JGR/>.

References

- Ammon, C. J. (1991), The isolation of receiver effects from teleseismic P waveforms, *Bull. Seismol. Soc. Am.*, 81, 2504–2510.
Ammon, C. J., G. E. Randall, and G. Zandt (1990), On the non-uniqueness of receiver function inversions, *J. Geophys. Res.*, 95, 15,303–15,318.
Birch, F. (1961), The velocity of compressional waves in rocks to 10 kilobars, part 2, *J. Geophys. Res.*, 66, 2199–2224.
Bodin, T., M. Sambridge, H. Tkalcic, P. Arroucau, K. Gallagher, and N. Rawlinson (2012), Transdimensional inversion of receiver function and surface wave dispersion, *J. Geophys. Res.*, 117, B02301, doi:10.1029/2011JB008560.
Cassidy, J. F. (1992), Numerical experiments in broadband receiver function analysis, *Bull. Seismol. Soc. Am.*, 82, 1453–1474.
Chang, S.-J., C.-E. Baag, and C. A. Langston (2004), Joint analysis of teleseismic receiver functions and surface wave dispersion using the genetic algorithm, *Bull. Seismol. Soc. Am.*, 94, 691–704.
Chen, L., L. Wen, and T. Zheng (2005), A wave equation migration method for receiver function imaging: 1. Theory, *J. Geophys. Res.*, 110, B11309, doi:10.1029/2005JB003665.

- Chen, Y., and F. Niu (2013), Ray-parameter based stacking and enhanced preconditioning for stable inversion of receiver function data, *Geophys. J. Int.*, 194(3), 1682–1700, doi:10.1093/gji/ggt179.
- Chen, Y., F. Niu, R. Liu, Z. Huang, H. Tkalcic, L. Sun, and W. Chan (2010), Crustal structure beneath China from receiver function analysis, *J. Geophys. Res.*, 115, B03307, doi:10.1029/2009JB006386.
- Chevrot, S., and R. D. van der Hilst (2000), The Poisson ratio of the Australian crust: Geological and geophysical implications, *Earth Planet. Sci. Lett.*, 183, 121–132.
- Du, Z. J., and G. R. Foulger (1999), The crustal structure beneath the northwest fjords, Iceland, from receiver functions and surface waves, *Geophys. J. Int.*, 139, 419–432.
- Efron, B., and R. Tibshirani (1986), Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Stat. Sci.*, 1, 54–77.
- Frederiksen, A. W., and M. G. Bostock (2000), Modeling teleseismic wave in dipping anisotropic structure, *Geophys. J. Int.*, 141, 401–412.
- Gilbert, H. J., A. F. Sheehan, K. G. Dueker, and P. Molnar (2003), Receiver functions in the western United States, with implications for upper mantle structure and dynamics, *J. Geophys. Res.*, 108(B5), 2229, doi:10.1029/2001JB001194.
- Guo, Z., Y. J. Chen, J. Ning, Y. Feng, S. P. Grand, F. Niu, H. Kawakatsu, S. Tanaka, M. Obayashi, and J. Ni (2015), High resolution 3-D crustal structure beneath NE China from joint inversion of ambient noise and receiver functions using NECESSArray data, *Earth Planet. Sci. Lett.*, 416, 1–11, doi:10.1016/j.epsl.2015.01.044.
- He, R. Z., X. F. Shang, C. Q. Yu, H. J. Zhang, and R. D. Van der Hilst (2014), A unified map of Moho depth and V_p/V_s ratio of continental China by receiver function analysis, *Geophys. J. Int.*, 199(3), 1910–1918, doi:10.1093/gji/ggu365.
- Hearn, T. M., S. Wang, J. F. Ni, Z. Xu, Y. Yu, and X. Zhang (2004), Uppermost mantle velocities beneath China and surrounding regions, *J. Geophys. Res.*, 109, B11301, doi:10.1029/2003JB002874.
- Herrmann, R. B. (2006), Surface waves, receiver functions and crustal structure, in computer programs in seismology Version 3.30, Saint Louis Univ., St. Louis, Mo. [Available at <http://www.eas.slu.edu/People/RBHerrmann/CPS330.html>.]
- Huang, Z., W. Su, Y. Peng, Y. Zheng, and H. Li (2003), Rayleigh wave tomography of China and adjacent regions, *J. Geophys. Res.*, 108(B2), 2073, doi:10.1029/2001JB001696.
- Isaaks, E. H., and R. M. Srivastava (1989), *An Introduction to Applied Geostatistics*, pp. 278–322, Oxford Univ. Press, New York.
- Jacobsen, B. H., and L. Svennningsen (2008), Enhanced uniqueness and linearity of receiver function inversion, *Bull. Seismol. Soc. Am.*, 98, 1756–1767, doi:10.1785/0220070180.
- Julià, J., C. J. Ammon, R. B. Herrmann, and A. M. Correig (2000), Joint inversion of receiver function and surface wave dispersion observations, *Geophys. J. Int.*, 143, 1–19.
- Julià, J., C. J. Ammon, and R. B. Herrmann (2003), Lithospheric structure of the Arabian shield from the joint inversion of receiver functions and surface wave dispersion, *Tectonophysics*, 371, 1–21.
- Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.*, 122, 108–124.
- Kind, R., et al. (2002), Seismic images of crust and upper mantle beneath Tibet: Evidence for Eurasian plate subduction, *Science*, 298, 1219–1221.
- Langston, C. A. (1979), Structure under Mount Rainier, Washington, inferred from teleseismic body waves, *J. Geophys. Res.*, 84, 4749–4762.
- Lawrence, J. F., and D. A. Wiens (2004), Combined receiver function and surface wave phase-velocity inversion using a niching genetic algorithm: Application to Patagonia, *Bull. Seismol. Soc. Am.*, 94, 977–987.
- Li, C., and R. D. van der Hilst (2010), Structure of the upper mantle and transition zone beneath South East Asia from travel time tomography, *J. Geophys. Res.*, 115, B07308, doi:10.1029/2009JB006882.
- Li, S. L., and W. D. Mooney (1998), Crustal structure of China from deep seismic sounding profiles, *Tectonophysics*, 288, 105–113.
- Li, S. L., W. D. Mooney, and J. Fan (2006), Crustal structures of mainland China from deep seismic sounding data, *Tectonophysics*, 420, 239–252.
- Li, S. Z., G. C. Zhao, L. M. Dai, X. Liu, L. H. Zhou, M. Santosh, and Y. H. Suo (2012), Mesozoic basins in eastern China and their bearing on the deconstruction of the North China craton, *J. Asian Earth Sci.*, 47, 64–79.
- Li, Y., Q. Wu, J. Pan, F. Zhang, and D. Yu (2013), An upper-mantle S-wave velocity model for East Asia from Rayleigh wave tomography, *Earth Planet. Sci. Lett.*, 377–378, 367–377.
- Li, Y., M. Gao, and Q. Wu (2014), Crustal thickness map of the Chinese mainland from teleseismic receiver functions, *Tectonophysics*, 610, 51–60, doi:10.1016/j.tecto.2013.11.019.
- Liang, C., X. Song, and J. Huang (2004), Tomographic inversion of P_n travel times in China, *J. Geophys. Res.*, 109, B11304, doi:10.1029/2003JB002789.
- Ligorria, J. P., and C. J. Ammon (1999), Iterative deconvolution and receiver-function estimation, *Bull. Seismol. Soc. Am.*, 89, 1395–1400.
- Liu, H., and F. Niu (2012), Estimating crustal seismic anisotropy with a joint analysis of radial and transverse receiver function data, *Geophys. J. Int.*, 188, 144–164, doi:10.1111/j.1365-246X.2011.05249.x.
- Mangino, S., K. Priestley, and J. Ebel (1999), The receiver structure beneath the China digital seismograph network stations, *Bull. Seismol. Soc. Am.*, 89, 1053–1076.
- Nair, S. K., S. S. Gao, K. H. Liu, and P. G. Silver (2006), Southern African crustal evolution and composition: Constraints from receiver function studies, *J. Geophys. Res.*, 111, B02304, doi:10.1029/2005JB003802.
- Niu, F., and D. E. James (2002), Fine structure of the lowermost crust beneath the Kaapvaal craton and its implications for crustal formation and evolution, *Earth Planet. Sci. Lett.*, 200, 121–130.
- Niu, F., and H. Kawakatsu (1996), Complex structure of the mantle discontinuities at the tip of the subducting slab beneath the northeast China: A preliminary investigation of broadband receiver functions, *J. Phys. Earth*, 44, 701–711.
- Niu, F., and J. Li (2011), Component azimuths of the CEArray stations estimated from P -wave particle motion, *Earthquake Sci.*, 24, 3–33, doi:10.1007/s11589-011-0764-8.
- Niu, F., T. Bravo, G. Pavlis, F. Vernon, H. Rendon, M. Bezada, and A. Levander (2007), Receiver function study of the crustal structure of the southeastern Caribbean plate boundary and Venezuela, *J. Geophys. Res.*, 112, B11308, doi:10.1029/2006JB004802.
- Owens, T. J., A. A. Nyblade, H. Gurrola, and C. A. Langston (2000), Mantle transition zone structure beneath Tanzania, East Africa, *Geophys. Res. Lett.*, 27, 827–830.
- Özalaybey, S., M. K. Savage, A. F. Sheehan, J. N. Louie, and J. N. Brune (1997), Shear-wave velocity structure in the northern basin and range province from the combined analysis of receiver functions and surface waves, *Bull. Seismol. Soc. Am.*, 87, 183–189.
- Pasyanos, M., H. Tkalcic, R. Gök, and A. Rodgers (2007), Seismic structure of Kuwait from joint inversion of surface wave group velocities and RFs, *Geophys. J. Int.*, 170, 299–312.

- Pei, S., J. Zhao, Y. Sun, Z. Xu, S. Wang, H. Liu, C. A. Rowe, M. N. Toksöz, and X. Gao (2007), Upper mantle seismic velocities and anisotropy in China determined through P_n and S_n tomography, *J. Geophys. Res.*, 112, B05312, doi:10.1029/2006JB004409.
- Randall, G. E. (1989), Efficient calculation of differential seismograms for lithospheric receiver functions, *Geophys. J. Int.*, 99, 469–481.
- Rawlinson, N., and M. Sambridge (2005), The fast marching method: An effective tool for tomographic imaging and tracking multiple phases in complex layered media, *Explor. Geophys.*, 36, 341–350.
- Ren, J. Y., K. Tamaki, S. T. Li, and J. X. Zhang (2002), Late Mesozoic and Cenozoic rifting and its dynamic setting in eastern China and adjacent areas, *Tectonophysics*, 344, 175–205.
- Sambridge, M. (1999), Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space, *Geophys. J. Int.*, 138, 479–494.
- Sandvol, E., D. Seber, A. Calvert, and M. Barazangi (1998), Grid search modeling of receiver functions: Implications for crustal structure in the Middle East and North Africa, *J. Geophys. Res.*, 103, 26,899–26,917.
- Shen, X., and H. Zhou (2012), Receiver functions of CCDSN and crustal structure of Chinese mainland, *Earthquake Sci.*, 25, 3–16, doi:10.1007/s11589-012-0826-6.
- Sherrington, H. F., G. Zandt, and A. Frederiksen (2004), Crustal fabric in the Tibetan Plateau based on waveform inversions for seismic anisotropy parameters, *J. Geophys. Res.*, 109, B02312, doi:10.1029/2002JB002345.
- Shibutani, T., M. Sambridge, and B. Bennett (1996), Genetic algorithm inversion for receiver functions with application to crust and uppermost mantle structure beneath eastern Australia, *Geophys. Res. Lett.*, 23, 1829–1832.
- Sun, Y., and M. N. Toksöz (2006), Crustal structure of China and surrounding regions from P wave traveltimes tomography, *J. Geophys. Res.*, 111, B03310, doi:10.1029/2005JB003962.
- Sun, Y., M. N. Toksöz, S. Pei, and F. D. Morgan (2008), The layered shear-wave velocity structure of the crust and uppermost mantle in China, *Bull. Seismol. Soc. Am.*, 98, 746–755, doi:10.1785/0120050222.
- Tao, K., T. Liu, J. Ning, and F. Niu (2014), Estimating sedimentary and crustal structure using wavefield continuation: Theory, techniques and applications, *Geophys. J. Int.*, 197, 443–457, doi:10.1093/gji/ggt515.
- Tapponnier, P., Z. Xu, F. Ronger, B. Meyer, N. Arnaud, G. Wittlinger, and J. Yang (2001), Oblique stepwise rise and growth of the Tibet Plateau, *Science*, 294, 1671–1677.
- Teng, J., Z. Zhang, X. Zhang, C. Wang, R. Gao, B. Yang, Y. Qiao, and Y. Deng (2013), Investigation of the Moho discontinuity beneath the Chinese mainland using deep seismic sounding profiles, *Tectonophysics*, 609, 202–216, doi:10.1016/j.tecto.2012.11.024.
- Tkalčić, H., M. Pasanos, A. Rodgers, R. Gök, W. Walter, and A. Al-Amri (2006), A multi-step approach in joint modelling of surface wave dispersion and teleseismic receiver functions: Implications for lithospheric structure of the Arabian peninsula, *J. Geophys. Res.*, 111, B11311, doi:10.1029/2005JB004130.
- Tkalčić, H., Y. Chen, R. Liu, Z. Huang, L. Sun, and W. Chan (2011), Multistep modeling of teleseismic receiver functions combined with constraints from seismic tomography: Crustal structure beneath southeast China, *Geophys. J. Int.*, 187, 303–326, doi:10.1111/j.1365-246X.2011.05132.x.
- Vergne, J., G. Wittlinger, Q. Hui, P. Tapponnier, G. Poupinet, J. Mei, G. Herquel, and A. Paul (2002), Seismic evidence for stepwise thickening of the crust across the NE Tibetan plateau, *Earth Planet. Sci. Lett.*, 203, 25–33.
- Vermeesch, P. (2003), A second look at the geologic map of China: The “Sloss approach”, *Int. Geol. Rev.*, 45, 119–132.
- Wang, S., F. Niu, and G. Zhang (2013), Velocity structure of the uppermost mantle beneath East Asia from Pn tomography and its dynamic implications, *J. Geophys. Res. Solid Earth*, 118, 290–301, doi:10.1002/jgrb.50085.
- Xiao, W., B. F. Windley, J. Hao, and M. Zhai (2003), Accretion leading to collision and the Permian Solonker suture, Inner Mongolia, China: Termination of the central Asian orogenic belt, *Tectonics*, 22(6), 1069, doi:10.1029/2002TC001484.
- Xu, Y. G. (2007), Diachronous lithospheric thinning of the North China Craton and formation of the Daxin'anling-Taihangshan gravity lineament, *Lithos*, 96, 281–298.
- Yao, H., C. Beghein, and R. D. van der Hilst (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.*, 163, 205–219, doi:10.1111/j.1365-246X.2007.03696.x.
- Yao, H., R. D. van der Hilst, and J. P. Montagner (2010), Heterogeneity and anisotropy of the lithosphere of SE Tibet from surface wave array tomography, *J. Geophys. Res.*, 115, B12307, doi:10.1029/2009JB007142.
- Yin, A., and T. M. Harrison (2000), Geologic evolution of the Himalayan-Tibetan orogen, *Annu. Rev. Earth Planet. Sci.*, 28, 211–280.
- Yin, A., and S. Y. Nie (1996), A Phanerozoic palinspastic reconstruction of China and its neighboring regions, in *The Tectonic Evolution of Asia*, edited by A. Yin and T. M. Harrison, pp. 442–485, Cambridge Univ. Press, New York.
- Zhao, L., M. K. Sen, P. Stoffa, and C. Frolich (1996), Application of very fast simulated annealing to the determination of the crustal structure beneath Tibet, *Geophys. J. Int.*, 125, 355–370.
- Zheng, S., X. Sun, X. Song, Y. Yang, and M. H. Ritzwoller (2008), Surface wave tomography of China from ambient seismic noise correlation, *Geochim. Geophys. Geosyst.*, 9, Q05020, doi:10.1029/2008GC001981.
- Zheng, T., R. Zhu, L. Zhao, and Y. Ai (2012), Intralithospheric mantle structures recorded continental subduction, *J. Geophys. Res.*, 117, B03308, doi:10.1029/2011JB008873.
- Zhu, L., and H. Kanamori (2000), Moho depth variation in southern California from teleseismic receiver functions, *J. Geophys. Res.*, 105, 2969–2980.