Geophysical Journal International

Geophys. J. Int. (2013) doi: 10.1093/gji/ggt179

Geophysical Journal International Advance Access published May 30, 2013

Ray-parameter based stacking and enhanced pre-conditioning for stable inversion of receiver function data

Youlin Chen[1] and Fenglin Niu[2,3]

[1] *Advanced Technology Division, Array Information Technology, Inc., Greenbelt, MD 20770, USA. E-mail: youlin.chen@arrayinfotech.com*
[2] *State Key Laboratory of Petroleum Resource and Prospecting, and Unconventional Natural Gas Institute, China University of Petroleum, Beijing 102249, China*
[3] *Department of Earth Science, Rice University, Houston, TX 77251, USA*

Accepted 2013 May 2. Received 2013 April 19; in original form 2012 July 12

SUMMARY

While inversion of seismic velocity from receiver function data could be instable due to its intrinsic non-linearity and non-uniqueness, improper stacking of receiver function could also introduce significant biases to the resulting velocity structure. In a distance section of receiver functions, the Moho Ps conversion and the two reverberations possess a positive and negative moveout, respectively. Stacking receiver functions without moveout correction could significantly reduce and distort the amplitude and waveform of these phases. Inversion with these incorrectly stacked receiver functions will thus inevitably introduce artefacts to the resulting velocity structure. In this study, we have improved the inversion procedure in two ways. First, we introduce a ray-parameter based (RPB) stacking method to correctly construct receiver function data for inversion. Specifically we develop a 'four-pin' method that accounts for the moveout effect of the converted and reverberated phases in stacking individual receiver functions recorded at various distances. Secondly, we divide the receiver function trace into conversion and reverberation windows and assign different weights between the two windows in the inversion. More weight is given to the Ps conversion window in resolving the shallow structure, which can be nearly fixed in the successive inversion of deeper structure. We also employ other pre-conditioning proposed by previous studies, such as balancing the receiver function data being filtered with different Gaussian filters, smoothing the velocity model and further regulating the model based on existing information. We compute synthetic receiver functions at distances between 30° and 90° from a target model and then use the RPB stacking method to generate the input data for various inversions (iterative linear) with different initial models. Our inversions with enhanced pre-conditioning and RPB stacked data demonstrate a good capability in recovering the target model from generally more stable iterations. Applying these techniques to two broad-band stations in China indicates that the improvements on data stacking and inversion can eliminate potential stacking-induced artefacts, and yield models more consistent with surface geology.

Key words: Time-series analysis; Numerical solutions; Inverse theory; Body waves; Crustal structure; Asia.

1 INTRODUCTION

With the rapid increase of three-component broad-band seismographs around the globe, the receiver function analysis has become an effective tool extensively used in modelling lithospheric structure beneath the receiver sites. The advantage of receiver function analysis attributes to its capability of isolating the response of Earth's structure from source signatures by deconvolving the incident wave recorded on the vertical component from the mode-converted waves on the horizontal components. The deconvolution results in a series of *P* to *S* converted and reflected *S* waves associated with Earth's internal boundaries beneath the seismograph. Since the concept of receiver function was first introduced (e.g. Langston 1979), many methods and techniques have been developed to apply receiver function data to estimate crustal properties, image lithospheric structure and model depth profile of shear wave velocity.

Receiver function analysis provides a robust way to characterize the first-order properties of crustal structure such as crustal thickness (H), V_P/V_S ratio (κ), hence the Poisson's ratio, as well as velocity and density contrasts at the Moho discontinuity. The H–κ analysis is now a widely used method in simultaneously estimating crustal thickness and V_P/V_S ratio (e.g. Clarke & Silver 1993; Zandt & Ammon 1995; Zandt *et al.* 1995; Chevrot & van der Hilst 2000; Zhu & Kanamori 2000; Nair *et al.* 2006; Niu *et al.* 2007; Chen *et al.* 2010). The inherent problem within this method is the substantial trade-off between H and κ (Nair *et al.* 2006). Although it can be

partially reduced by including the major reverberation phases in the analysis (Chevrot & van der Hilst 2000; Zhu & Kanamori 2000; Nair *et al.* 2006), a 'band' area of amplitude peaks along the optimal H and κ pairs is always shown in the amplitude contour map, in which multiple acceptable H and κ values can be marked. Niu *et al.* (2007) improved the H–κ method by introducing a cross correlation between the Ps phase and the crustal multiple phases as additional weight factors. This modified H–κ method has proven successful in reducing the 'band' to a small area centred at the optimal H–κ pair when it was applied in various tectonic regions with complicated crustal structure (Niu *et al.* 2007; Chen *et al.* 2010; Pan & Niu 2011; Sun *et al.* 2012). An accurate estimate of H and κ plays an important role in this study. First, they are crucial parameters in computing the moveout corrections of the converted and reverberated phases for the ray-parameter based (RPB) stacking method developed here; secondly, they are used in constructing the starting model and the P-wave velocity in our inversion.

Inversion of receiver function data for 1-D velocity structure is generally regarded as a strongly non-linear and non-unique inverse problem (e.g. Langston 1979; Ammon *et al.* 1990; Ammon 1991; Jacobsen & Svenningsen 2008). The non-linearity arises mostly from the complicated relationship between the conversion/reflection coefficient and velocity contrast at the Earth's internal boundaries, such that for any reasonable perturbations to an initial model the higher-order terms in the Taylor expansion of receiver functions are not negligible (e.g. Jacobsen & Svenningsen 2008). The linearized inversion thus strongly depends on the choice of the starting model. Many authors have employed an iterative method to gradually approach the optimum model (e.g. Ammon *et al.* 1990; Ammon & Zandt 1993; Searcy *et al.* 1996; Julià *et al.* 1998; Mangino *et al.* 1999). In order to circumvent problems associated with linearization, various non-linear approaches, which combine forward modelling and Monte Carlo method, have been introduced to search for the global optimum in the parameter space. These non-linear methods include the interactive forward modelling of receiver functions (e.g. Tkalčić & Banerjed 2009), grid search (e.g. Sandvol *et al.* 1998; Tkalčić *et al.* 2006, 2011), genetic or niching genetic algorithm (e.g. Shibutani *et al.* 1996; Koper *et al.* 1999; Chang *et al.* 2004; Lawrence & Wiens 2004; Chang & Baag 2005; Lawrence *et al.* 2006; Ma *et al.* 2007), simulated annealing (e.g. Zhao *et al.* 1996; Vergne *et al.* 2002; Vinnik *et al.* 2004) and neighbourhood algorithm (e.g. Sambridge 1999a,b; Agostinetti *et al.* 2002). The non-uniqueness of receiver function inversion arises from the trade-off between depth and velocity, which makes the response of a shallow low-velocity model in the receiver function data to be very similar to that of a deep high-velocity model (Ammon *et al.* 1990; Jacobsen & Svenningsen 2008). The trade-off is caused by the fact that receiver functions are sensitive to gradients in seismic velocities, but weakly sensitive to absolute velocities. It has been shown that a joint inversion of receiver functions and surface wave dispersion data is able to provide better estimates of the overall velocity structure (e.g. Özalaybey *et al.* 1997; Du & Foulger 1999; Julià *et al.* 2000, 2003; Chang *et al.* 2004; Lawrence & Wiens 2004; Chang & Baag 2005; Lawrence *et al.* 2006; Tkalčić *et al.* 2006; Yoo *et al.* 2007). The alternative approach is to include *a priori* on the S- and P-wave velocity information from other independent geophysical studies to constrain a better starting velocity model before the start of the inversion. A well-constrained starting model can prevent the iterative inversion from falling into local minima. Tkalčić *et al.* (2011) made use of crustal S-wave velocity derived from a coarse grid search method, crustal thickness and V_P/V_S ratio from H–κ analysis, and the uppermost mantle velocity structure from tomography study to build up their starting models, and successfully inverted for the S-wave velocity profile from receiver function data recorded at 18 stations in southeast China.

The advantages and limitations of the receiver function technique have been discussed in numerous previous studies. On the other hand, properly and efficiently processing observed data is another challenging issue in utilizing receiver function technique, especially when large numbers of receiver functions are available. In performing receiver function inversion, it is usually preferred to have the stacked data as the input as noise level and 2-D/3-D effects are significantly suppressed in the stacked receiver function data. In this study, we first develop a RPB stacking method that is able to stack receiver functions generated from earthquakes occurring at different epicentral distances without distorting the waveform of the converted and reverberated phases provided that the Earth structure beneath the station may be reasonably well represented by a 1-D velocity model. We then optimize the iterative linearized inversion method of Ammon *et al.* (1990) by pre-conditioning both the data and the model. To show the effectiveness of our method, we conduct extensive inversion tests with synthetic receiver function data. We also examine the improvement on receiver function inversion using real seismic data recorded at two permanent broad-band stations in China, XAN and MDJ. These two stations are selected for two reasons: (1) the sites have unique tectonic settings and (2) the lithosphere structure beneath the two stations has been investigated by previous studies, allowing us to verify the inversion results.

2 METHODS

2.1 'Four-pin' moveout correction

The relative arrival times of the primary P to S converted phase, $0p1s$, and the two major crustal multiple phases, $2p1s$ and $1p2s$, with respect to the direct P wave are usually approximated by the following equations (e.g. Chevrot & van der Hilst 2000; Zhu & Kanamori 2000; Nair *et al.* 2006):

$$t_{0p1s} = H\left[\sqrt{\left(\frac{\kappa}{V_P}\right)^2 - p^2} - \sqrt{\left(\frac{1}{V_P}\right)^2 - p^2}\right]$$
$$= H\left[\left(\frac{\kappa}{V_P}\right)^2 - \left(\frac{1}{V_P}\right)^2\right] \Bigg/ \left[\sqrt{\left(\frac{\kappa}{V_P}\right)^2 - p^2} + \sqrt{\left(\frac{1}{V_P}\right)^2 - p^2}\right], \quad (1)$$

$$t_{2p1s} = H\left[\sqrt{\left(\frac{\kappa}{V_P}\right)^2 - p^2} + \sqrt{\left(\frac{1}{V_P}\right)^2 - p^2}\right], \quad (2)$$

$$t_{1p2s} = 2H\sqrt{\left(\frac{\kappa}{V_P}\right)^2 - p^2}. \quad (3)$$

We have followed Niu & James (2002) on the notation of the later arrivals, *npms*, where n and m are the numbers of P- and S-wave legs within the crust. Here H, V_P, κ and p are the crustal thickness, average P-wave velocity, V_P/V_S ratio and the P-wave ray-parameter, respectively. In eqs (1)–(3), we have made a plane wave assumption, which assumes that the Ps conversion and reverberation phases possess a ray-parameter similar to the direct P wave. This is

actually not true for a point source, as ray paths for different arrivals can never be exactly parallel. An increasing of ray-parameter, p (equivalent a decreasing in epicentral distance) in eqs (1)–(3) results in an increase of t_{0p1s} and a decrease of t_{2p1s} and t_{1p2s}. Eqs (1)–(3) thus suggest that the arrival times of the $0p1s$, and $2p1s/1p2s$ have a negative and positive relative ray-parameter with respect to the direct P wave, respectively, that is a negative and positive moveout.

As shown above, the moveout of the converted and reverberated S-wave phases with respect to the direct P wave is not zero; stacking receiver functions aligning to the direct P wave can thus significantly weaken the amplitude and distort the waveform of these later arrivals. Since the arrival moveout is larger for the multiples, they suffer more severe stacking induced distortions. This is particularly true for high-frequency receiver function data. The destructive stacking can become even worse when a thick crust is present and the distance coverage of the data is broad. For example, for a crust with a thickness of 50 km and an average V_P/V_S ratio of 1.78, the arrival time of the $0p1s$, $2p1s$ and $1p2s$ phases 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 s deg^{-1}. When receiver functions are constructed using a Gaussian filter width $a = 2.5$ or above, such large phase shifts can result in completely out-of-phase stacking if no moveout correction is made. In principle, this problem can be avoided by narrowing 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 heterogeneous, 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.

We therefore develop a 'four-pin' method to correct the moveout of the Ps converted phase, $0p1s$, as well as the following two crustal multiples, $2p1s$ and $1p2s$. Stacking the moveout corrected receiver functions then becomes a straightforward task and is obviously applicable to all the events occurring at the entire teleseismic distance range. To implement the 'four-pin' correction, first, we need to choose one ray-parameter, p_0, as a reference and compute the arrival times of the three phases, $t^0{}_{0p1s}$, $t^0{}_{2p1s}$ and $t^0{}_{1p2s}$ following eqs (1)–(3) using the accurate estimates of crustal thickness and V_P/V_S ratio derived from H-κ analysis (i.e. Niu *et al.* 2007) in advance. These three time points will be used as marks on a standard receiver function to rescale all individual receiver functions to be corrected. Next, for each receiver function, for example the ith receiver function with a ray-parameter of p_i, we compute the three later arrival times, $t^i{}_{0p1s}$, $t^i{}_{2p1s}$ and $t^i{}_{1p2s}$ also using eqs (1)–(3). We divide this receiver function into three segments bounded by time zero (P-wave arrival time) to t_{0p1s}, t_{0p1s} to t_{2p1s} and t_{2p1s} to t_{1p2s}, respectively. For each segment, we try to pin its two ends to the marks on the standard receiver function by stretching or contracting this segment to make its length equal to the same segment on the standard receiver function. In an evenly sampled digital seismogram, such stretching/contracting procedure changes the sample interval of the raw seismogram. We simply resample the stretched/contracted seismogram using the original sample rate to implement the stretching/contracting process. The stretching/contracting is performed in a consecutive manner, that is we first pin the direct P-wave phase on the ith receiver function onto the direct P wave on the standard receiver function, and stretch/contract the segment from the direct P wave to $0p1s$ to align the time point $t^i{}_{0p1s}$ to $t^0{}_{0p1s}$. Once this alignment is done, we next pin the corrected $t^i{}_{0p1s}$ to $t^0{}_{0p1s}$ and rescale the second segment along the time axis in order to align its right end $t^i{}_{2p1s}$ to the second mark $t^0{}_{2p1s}$. Finally, we repeat the process to the last segment of the data to match $t^i{}_{1p2s}$ to the third marker $t^0{}_{1p2s}$. An analogy of this process is to pin down four specific points of a spring to four marks of a ruler on a wall, thereby we refer this method as 'four-pin' correction method, which simultaneously aligns the three later arrivals on all the receiver functions to the reference arrival times as if they all occur at the same location as the reference source. Fig. 1(a) demonstrates the steps of the 'four-pin' method. In this example, the corrections are made with respect to a reference ray-parameter $p_0 = 6.5$ s deg^{-1}. The arrival times of the converted and reverberated phases before the corrections are indicated by red dots on each receiver function trace, and the traces are stretched or contracted so that they are aligned along three vertical lines.

2.2 RPB stacking

In order to demonstrate the effect of the RPB stacking, we generate a serial of synthetic receiver functions using the target velocity profile shown in Fig. 2. The model has a Moho discontinuity at 53 km deep and an average V_P/V_S ratio of 1.736 across the crust. The purpose of choosing a model with a deep Moho is to clearly illustrate the significance of the moveout on different receiver functions. We first compute 150 synthetic receiver functions randomly distributed in the ray-parameter range between 4.5 and 8.9 s deg^{-1} using the target model. We then apply the 'four-pin' correction to simultaneously align the converted and reverberated phases to a reference trace with $p_0 = 6.5$ s deg^{-1}. Since the ray-parameter is proportional to the incidence angle of the direct P wave, an increase of ray-parameter results in a nearly linear increase of the P-wave amplitude on receiver function data (Fig. 1b), as well as the following converted and reverberated S-wave amplitudes. We thus normalize the amplitude of each receiver function by dividing by its own ray-parameter and then multiplying the reference ray-parameter. Since the pre-stacking corrections of time and amplitude are made solely based on ray-parameters, we call the stacking after the corrections as RPB stacking to distinguish it from the direct stacking without any correction.

We show the stacked receiver functions using the RPB stacking and direct stacking methods in Fig. 1(c). To make comparison, we also show the receiver function computed using the reference ray-parameter of $p_0 = 6.5$ s deg^{-1}, which is also the average of ray-parameters of all the 150 receiver functions. The differences between the stacked receiver functions and the reference trace are shown in the lower panel of Fig. 1(c). The three later arrivals on the direct stacked trace (blue lines in Fig. 1c) all show a large deviation from the reference receiver function, mainly due to the out-of-phase stacking. In contrast the RPB stacked trace is very similar to the reference receiver function with slight difference because the later phases waveform of each receiver function is not exact the same. In practice, the reference ray-parameter does not necessarily correspond to a specific individual receiver function. We can choose any value for the reference ray-parameter as long as the value falls in the teleseismic distance range. This means that we can construct multiple RPB stacks using different reference ray-parameters.

2.3 Enhanced pre-conditioning for the iterative linearized inversion

As mentioned above, various techniques have been developed to extract lithospheric structure from receiver function data. In this study, we choose to use the iterative linearized inversion method based on

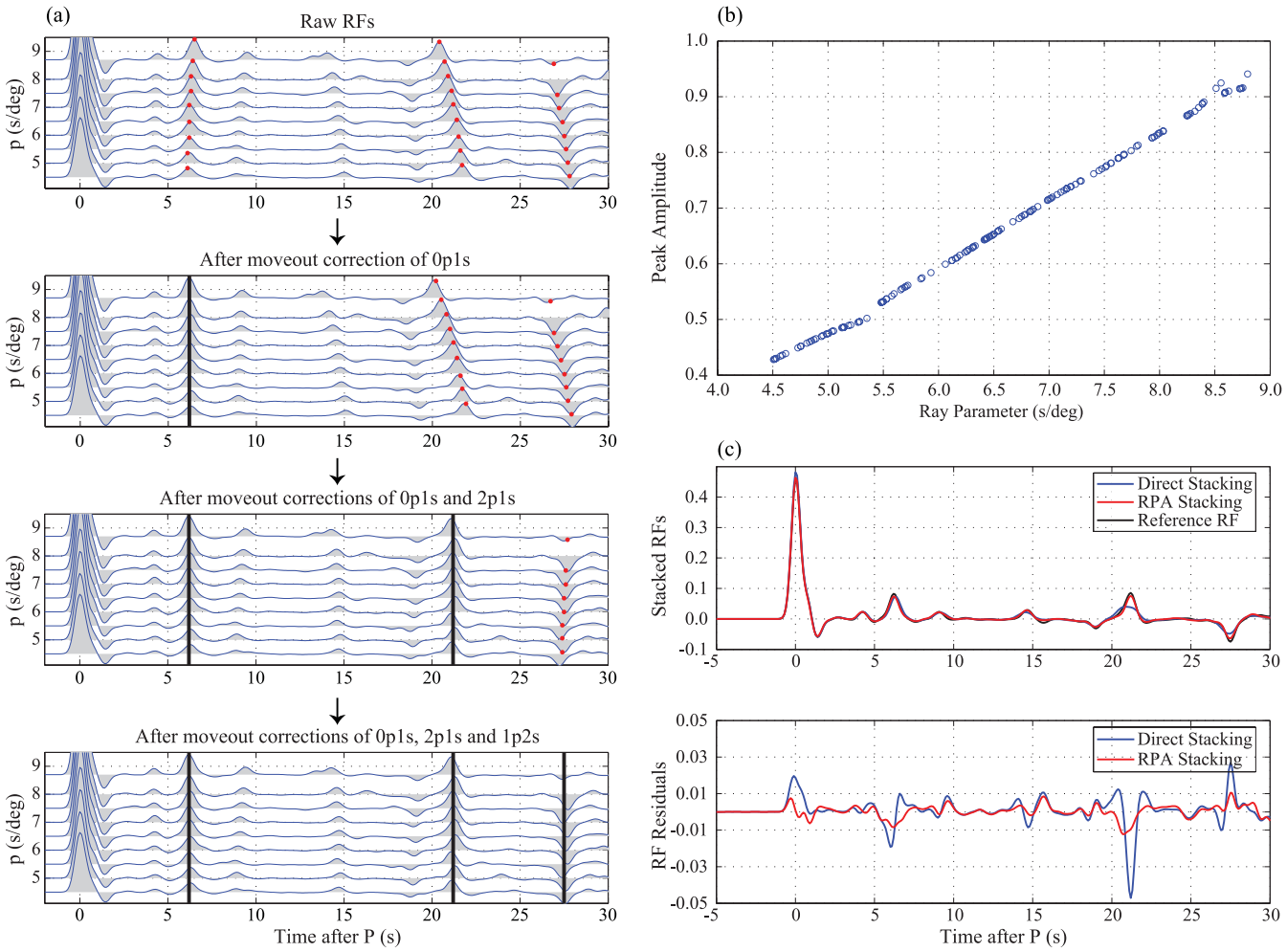

Figure 1. (a) Steps illustrating the 'four-pin' method used to correct the distance moveout of the $0p1s$, $2p1s$ and $1p2s$ phases. In this example, the correction is made with respect to $p_0 = 6.5\ \mathrm{s\ deg^{-1}}$. The arrival time of the $0p1s$, $2p1s$ and $1p2s$ phases are marked with red dots on each time trace. The vertical lines crossing these phases indicate their corresponding moveout correction in each step. (b) Peak amplitude of direct P wave is shown as a function of ray-parameters. (c) Top panel shows the direct stacking (blue line) and RPB stacking (red line) with respect to the reference receiver function (black line), and bottom panel shows the difference of two types of stacks from the reference receiver function.

the least-square technique to extract S-wave velocity structure beneath a seismic station. The method is known for its high-resolution power to determine crustal structure and also for its relatively low cost in computing (e.g. Snieder & Trampert 1999, 2000). Most of the non-linear inversions are based more or less on Monte Carlo search, and thus demand a lot of computation. For instance, the niching genetic algorithm adopted by Ma *et al.* (2007) requires approximately a 1-hr run on 11 CPUs in parallel implementation, or more than 10 hr on a single CPU to search 500 000 candidates for an optimal model. In contrast, an iterative linearized inversion method demands only several minutes of computing with a single CPU to perform the same inversion (Chen *et al.* 2011).

Following Ammon *et al.* (1990), we formulate the linearized inversion as

$$\mathbf{D}\cdot\mathbf{m} = \mathbf{r} + \mathbf{D}\cdot\mathbf{m}_0, \tag{4}$$

where $\mathbf{m}$ is the model, and $\mathbf{r}$ is the residual of receiver function. The matrix $\mathbf{D}$ is the sensitivity kernel defined as the partial derivatives of receiver function with respect to the model parameters, which are computed numerically using the efficient forward perturbation algorithm developed by Randall (1989). We use petrophysical relationships such as the empirical Birch's law (Birch 1961) and Poisson's ratio to tie density and P velocity to S velocity, leaving S velocity the only independent model variable in our inversion. Eq. (4) is performed iteratively to upgrade the model gradually.

The direct application of eq. (4) usually leads to rapid, large amplitude variation in the resultant model due to the lack of damping in the inversion. The smoothness control is thus always required to add on eq. (4). We adopt the second-order difference smoothing scheme (Ammon *et al.* 1990) as the first pre-conditioning in our study

$$\mathbf{S}\cdot\mathbf{\Delta}\cdot\mathbf{m} = \mathbf{0}, \tag{5}$$

where the matrix $\mathbf{\Delta}$ is a double difference operator applied on model $\mathbf{m}$ and the diagonal matrix $\mathbf{S}$ controls the trade-off between waveform fitness and model smoothness. In the inversion in this study, we simply use uniform values of $\mathbf{S}$, in which $\mathbf{S} = s\mathbf{I}$ ($\mathbf{I}$ is the identity matrix), and change s gradually to investigate the effect of smoothness controls on an inversion.

We choose velocity models down to 500 km deep for forward modelling, but only invert velocity structure of the top 200 km. Our receiver function usually has a time-window length of 30–50 s,

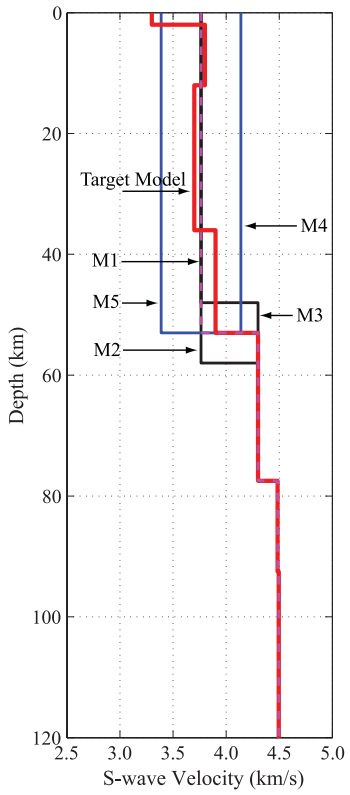

Figure 2. *S*-wave velocity models. The target model (red line) is used to generate synthetic receiver functions. Models M1 to M5 are initial velocity models used in inversion tests. Note the five models have very different Moho depth and average crustal velocity.

which is not long enough to model structure deeper than 200 km. The second pre-conditioning in our inversion is thus to fix the deep layers in the model to predetermined values, such as those of the AK135 or the PREM model. This can be achieved by adding the following equations into eq. (4) (e.g. Herrmann *et al.* 2000):

$$\mathbf{W} \cdot \mathbf{m} = \mathbf{W} \cdot \mathbf{m_0}, \tag{6}$$

where $\mathbf{W}$ is a diagonal matrix of weights and the model vector $\mathbf{m}_0$ contains predetermined velocity values. For layers shallower than 200 km, the corresponding weighting factors w_{ii} are zeros, and the corresponding equations are in fact excluded from our linear system. For the layers that require to be fixed to some extent, w_{ii} are assigned to adjustable values larger than zero, generally in the range of 1.0–10.0. The weight-adjustable scheme can also be applied to crustal layers when *a priori* determined information, such as the well-constrained sedimentary structure below Earth's surface, is available.

The construction of receiver function, no matter in frequency domain or in time domain, usually involves the convolution of a low-pass Gaussian filter with the expression, $\exp\left(\frac{-\omega^2}{4a^2}\right)$, where ω is radial frequency, and a is width factor controlling the cut-off frequency and also affecting the overall amplitude of the receiver function. For example, $a = 1.0$ roughly corresponds to a cut-off frequency at 0.5 Hz, whereas $a = 2.5$ corresponds to about 1.2 Hz. A higher a value thus yields a receiver function containing higher-frequency information. If an inversion involves multiple receiver functions generated with different values of a, a division by a is required to properly weight these receiver function data. Otherwise, a noisy receiver function with a higher a value can dominate the inversion. This division is the third pre-conditioning we used in our inversion.

In an individual receiver function, the Moho Ps converted phase is usually located within the first ~10 s after the *P* arrival, while the crustal multiples appear at the later part of the trace. We divide the receiver function trace into a Ps conversion segment and a reverberation segment. Since the two segments possess different signal-to-noise ratio (SNR) and sensitivity to the crustal structure, and hence different importance in an inversion, we assign different weights to the two segments as the last pre-conditioning in our inversion.

All of the pre-conditions discussed above are integrated in eq. (4), and our final linear inversion system turns to

$$\begin{bmatrix} \vdots & \vdots \\ (p/a)_i & \mathbf{D}^i_{\text{seg1}} \\ (q/a)_i & \mathbf{D}^i_{\text{seg2}} \\ \vdots & \vdots \\ s & \mathbf{\Delta} \\ & \mathbf{W} \end{bmatrix} \cdot \mathbf{m} = \begin{bmatrix} \vdots & \vdots \\ (p/a)_i & \mathbf{r}^i_{\text{seg1}} \\ (q/a)_i & \mathbf{r}^i_{\text{seg2}} \\ \vdots & \vdots \\ & 0 \\ & 0 \end{bmatrix} + \begin{bmatrix} \vdots & \vdots \\ (p/a)_i & \mathbf{D}^i_{\text{seg1}} \\ (q/a)_i & \mathbf{D}^i_{\text{seg2}} \\ \vdots & \vdots \\ & 0 \\ & \mathbf{W} \end{bmatrix} \cdot \mathbf{m_0}, \tag{7}$$

where index i denotes any one receiver function generated with Gaussian width a, and p and q are the weighing factors assigned to the conversion and the reverberation segments in this receiver function, respectively. In order to balance the two segments that usually have different length, we define the two weighting factors as

$$p = \frac{c}{N_1}, \qquad q = \frac{1-c}{N_2} \tag{8}$$

where N_1 is the number of points in the conversion segment, and N_2 is the numbers of points in the reverberation segment. The adjustable parameter c controls the relative weight between the two segments.

In summary, the regularizations we have employed for the linearized inversion can be grouped into two categories: (1) the model regularization, which include the model smoothness and the fixed-layers constraint; (2) data weighting scheme, which consist of internal waveforms balancing achieved by the division of a and segment weighting for individual receiver function data. In our inversion scheme, the model regularization limits the number of free parameters, while the data weighting balances the effects of receiver functions derived from different a values and also helps to emphasize specific phases. These pre-conditions may not necessarily increase, or even decrease waveform fit as discussed in Ammon *et al.* (1990). However, the regularization is necessary to stabilize the inversion and to generate models consistent with geological observations.

3 SYNTHETIC TESTS

The purpose of our synthetic tests is to examine the capability and stability of the inverse problem through linearized method with enhanced pre-conditioning on RPB stacked data. In all the synthetic tests, we intend to recover the target model shown in Fig. 2 by inverting either the RPB or direct stacks of receiver function data. We adopt various one-layer crustal models with different Moho depth and velocity value as the starting models. Since our receiver function data are filtered by low-pass Gaussian filters with the highest equivalent corner frequency of ∼1.0 Hz, which corresponds to *S*-wave wavelength of ∼4 km in crust, we thus divided the crust and another 20 km of the upper mantle into layers of 2 or 3 km thick. The rest upper mantle structure down to 500 km is taken from the AK135 model (Kennett *et al.* 1995), and is divided into layers with a thickness varying from 12 to 50 km with increasing depth. In general, the misfit decreases with increasing number of layers before it reaches to a stable value. Using ∼0.5 Hz low-pass filtered data, Tkalčić *et al.* (2011) showed that the waveform fit improved substantially when increasing the number of crustal layers from 4 to 8, but saturated at 12 with an average layer thickness of 2.5 km.

The synthetic receiver functions used for RPB and direct stacking are generated randomly in the ray-parameter range between 4.7 and 8.9 s deg^{-1} and have a mean ray-parameter of 6.5 s deg^{-1}. The inversions are conducted on both the direct (DRT case) and RPB (RPB case) stacks. The mean ray-parameter is used in the forward modelling part of the inversion with the directly stacked data. The RPB stacking can be generated with respect to different ray-parameters including this mean, such that the inversions could be conducted on any single RPB stack or on multiple RPB stacks simultaneously. We do not add noise on the synthetic receiver function data as we expect that a RPB stack possesses significantly less noise than an individual receiver function. The noise-free data can also help us to track down potential sources of misfit in the inversion. We use two Gaussian filters with width $a = 1.0$ and 2.5 in forming the synthetic receiver functions. Hereafter we refer them as the low- and high-frequency data, and both are used in the inversions. Including high-frequency data ensures proper parametrization of the crust and helps to reveal small-scale structure sensitive to high-frequency seismic wave.

In all the following tests, we balance the low- and high-frequency receiver function data by dividing by their Gaussian filter width factors, and include the regularization on the smoothness of the model, with the control parameter s varying from 0 to 1.0 in a step of 0.05. The detailed information and inversion results of all tests are summarized in Table 1. We compute the variance reduction (VR) to measure waveform misfit:

$$\mathrm{VR} = \left(1 - \sqrt{\frac{\sum \left(d_{\mathrm{obs}} - d_{\mathrm{syn}}\right)^2}{\sum d_{\mathrm{obs}}^2}}\right) \times 100\%, \tag{9}$$

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 Ps conversion and reverberation segments, of an individual receiver function. VR is not used as a threshold to stop iteration of an inversion; rather it is used as a criterion to determine the optimum model. Since the target model is known in advance in our synthetic tests, we introduce rms difference to evaluate how similar the inverted model is to the target model,

$$\mathrm{rms} = \sqrt{\frac{\sum \left(m_{\mathrm{inv}} - m_{\mathrm{tgt}}\right)^2}{M}}, \tag{10}$$

where m_{inv} and m_{tgt} denote the value at one layer in the inverted model and the target model, respectively. The rms is averaged over all M model layers. As we use a set of smoothness parameters to conduct inversion, we obtain a set of models for each data set derived from either DRT or RPB stacking. We further define a harmonic mean ⟨rms⟩:

$$\langle \mathrm{rms} \rangle = 1 \Big/ \sqrt{\frac{1}{K} \sum_{i}^{K} \frac{1}{\mathrm{rms}_i^2}}, \tag{11}$$

which is the average over a total of K models derived from inversions with different smoothness parameters. The harmonic mean reduces the effect of the outliers of the rms difference.

The first test case starts from the model M1 shown in Fig. 2, which has a constant crustal velocity that equals to the weighted average of the crustal velocity of the target model, and the depth of the Moho is kept as the same as the target model. In the DRT case 1.0, we employ neither the fixed-layer regularization on model nor the segment weighting on data. As shown in Fig. 3(a), the resulting model (red line) possesses a high-velocity zone in the upper crust and a low-velocity zone in the middle to lower crust down to ∼35 km. The velocity perturbation along the depth direction roughly follows the target model; however, the inversion apparently overestimates those velocity values. From ∼35 to ∼60 km, the *S*-wave velocity gradually increases with depth, without showing a clear velocity jump as the Moho discontinuity in the target model does. The model also shows a high-velocity zone in the uppermost mantle (Fig. 3a), which is apparently an artefact induced by inversion procedure. In the DRT case 1.1, we fix the upper mantle layers from 120 down to 500 km by assigning a gradually increasing weight factors from 0.5 to 1.0, and in DRT case 1.2, we adopt a segment weighting scheme to the data in the inversion. The inverted models of the two cases are shown in Figs 3(b) and (c), respectively. These two models do not show significant difference from the model of the DRT case 1.0. The fixed-layer regularization diminishes waveform fit and affects the estimate of the velocities of the deep layers, and forces the data to blend smoothly into all free layers. Adding more weight to the Ps conversion segment in the inversion only slightly increases the whole waveform fit, but has almost no influence on the inversion result. When the fixed-layer regularization on model and the segment weighting on data are combined, which is referred as the full constraints hereafter, the resulting model shows a high similarity to the target model (DRT case 1.3, Fig. 3d). Here, limiting changes of the velocity values on the lowermost upper mantle layers and the segment weighting scheme apparently boost the mean ⟨rms⟩ by approximately 30 per cent (Table 1). However, a notable feature presented in all the iterations is that the Moho discontinuity is not as sharp as in the target model. It is very likely that this feature of the Moho discontinuity results from the distortion of the converted and reverberated phases associated with the direct stacking. The stacking induced broadening of the lateral arrivals is better matched by a gradual crust–mantle boundary. This is particularly true with the high-frequency receiver function data, that is those filtered by a Gaussian width of $a = 2.5$. Fig. 3(e) shows the fitting of waveforms generated from starting and final models to the direct stack (observed data) of both low- and high-frequency bands. The shading areas in Fig. 3(e) mark the crustal multiples, clearly showing the difference between the observed data and the forward mapping of the resultant model.

We next repeat the inversions but using the RPB stacked synthetic data and taking the same initial model M1 (Fig. 2). In the RPB case 1.0, we apply neither velocity fixing nor segment

Table 1. Summary of synthetic tests.

Test case	Starting model	Inversion details	VR range	Rms range (mean)
DRT case 1.0	M1	Without fixed-layer and segment weight	81.98–90.97	0.1451–0.1983 (0.1740)
DRT case 1.1	M1	With fixed-layer	81.94–86.33	0.1201–0.2256 (0.1459)
DRT case 1.2	M1	With segment weight	90.38–92.47	0.1115–0.1575 (0.1365)
DRT case 1.3	M1	Full constraints	89.95–91.85	0.0795–0.1860 (0.1092)
RPB case 1.0	M1	Without fixed-layer and segment weight; $p = 6.5$	80.82–93.24	0.1037–0.1623 (0.1280)
RPB case 1.1	M1	Full constraints; $p = 5.5$	87.99–92.24	0.0696–0.1782 (0.0999)
RPB case 1.2	M1	Full constraints; $p = 6.5$	89.01–92.79	0.0375–0.0681 (0.0504)
RPB case 1.3	M1	Full constraints; $p = 8.0$	89.79–90.91	0.0739–0.0784 (0.0762)
RPB case 1.4	M1	Full constraints; $p = 5.5, 6.5, 8.0$	90.78–91.25	0.0483–0.0577 (0.0505)
DRT case 2	M2	Full constraints	90.66–92.30	0.0651–0.2432 (0.0898)
RPB case 2.1	M2	Full constraints; $p = 6.5$	89.49–92.77	0.0463–0.1428 (0.0636)
RPB case 2.2	M2	Full constraints; $p = 5.5, 6.5, 8.0$	90.50–91.09	0.0540–0.0660 (0.0569)
DRT case 3	M3	Full constraints	90.09–91.97	0.1402–0.1678 (0.1533)
RPB case 3.1	M3	Full constraints; $p = 6.5$	88.60–91.65	0.0710–0.1866 (0.0876)
RPB case 3.2	M3	Full constraints; $p = 5.5, 6.5, 8.0$	90.29–91.12	0.0547–0.0607 (0.0547)
DRT case 4	M4	Full constraints	89.75–91.08	0.1618–0.1966 (0.1755)
RPB case 4.1	M4	Full constraints; $p = 6.5$	88.77–90.77	0.0812–0.1936 (0.0946)
RPB case 4.2	M4	Full constraints; $p = 5.5, 6.5, 8.0$	90.45–91.04	0.0501–0.0666 (0.0501)
DRT case 5	M5	Full constraints	86.32–90.67	0.0672–0.3085 (0.1307)
RPB case 5.1	M5	Full constraints; $p = 6.5$	83.26–89.15	0.0675–0.3205 (0.1537)
RPB case 5.2	M5	Full constraints; $p = 5.5, 6.5, 8.0$	83.80–84.37	0.2599–0.2947 (0.2774)

weighting, just like the DRT case 1.0. The receiver function data with $a = 1.0$ and 2.5 are RPB stacks at $p_0 = 6.5\,\mathrm{s\,deg^{-1}}$. The resulting model (Fig. 4a) is, however, not greatly improved compared to the cases using the direct stacks (Fig. 3). Since the RPB stack is equivalent to a single receiver function with a ray-parameter $p_0 = 6.5\,\mathrm{s\,deg^{-1}}$, the inversion using the data at a single ray-parameter falls into the inherent limit of receiver function modelling, that is the depth–velocity trade-off due to its strong non-linearity and non-uniqueness.

The constant crustal velocity in the starting model is quite different from the target model, and it gives full freedom to allow the inversion to fall into any local minima with a decent waveform fit (Fig. 4b). The VR in RPB case 1.0 indicates waveform fit as good as 93.24 per cent, but the model rms of all iterations are all pretty poor (Table 1). In the RPB cases 1.1, 1.2 and 1.3, we employ the enhanced pre-conditioning inversion method which includes a regularization on model and segment weighting on data using RPB stacks at a ray-parameter of $p_0 = 5.5$, 6.5 and 8.0 $\mathrm{s\,deg^{-1}}$, respectively. The inversion method and weights are similar to those used in the DRT case 1.3. The results are plotted in Figs 4(c)–(e), showing that all of the three final *S*-wave velocity profiles are very close to the target model. In the RPB case 1.4, we use three RPB stacks at $p_0 = 5.5$, 6.5 and 8.0 $\mathrm{s\,deg^{-1}}$ as the input data to conduct the enhanced pre-conditioning inversion. The synthetic seismograms in the forward modelling are accordingly computed using these three ray-parameters. As shown in Fig. 4(f), the resulting model is also very close to the target model, but no substantial improvement is found from both VR and rms (Table 1) comparing to the RPB cases (1.1–1.3) using single ray-parameter. The result from the RPB case 1.4 in fact confirms the conclusion drawn by Ammon *et al.* (1990) that the range of ray-parameter typically used in receiver function inversion appears to be not broad enough to eliminate the depth–velocity ambiguity. However, the inversion using multiple rays greatly increases the convergence of the models for every smoothness control (Fig. 4f), and hence increases the stability of the solutions around the global optimum. This makes the RPB case 1.4 the best solution among all the test cases in terms of convergence and stability of the iterations although it may not generate the highest VR and lowest rms.

In order to further illustrate the effectiveness of our new method in stabilizing the inversion, we conduct inversions with other two starting models. The Moho depth in the two initial models is, respectively, 5 km deeper (M2) and shallower (M3) than the target model (Fig. 2). For all of our inversions hereafter, we start to use full enhanced preconditioning scheme. The inversions on the direct stacks are categorized as DRT case 2 and DRT case 3, and the results are shown in Fig. 5. Here in both cases, the Moho is rapidly rectified to the correct depth after several iterations. The major structure presented in the velocity profiles is the two high-velocity zones in the upper crust at ∼10 km and in the lower crust at ∼45 km. These artificial structures, on the other hand, disappear on the crustal models inverted from the RPB stacked data (RPB case 2.1 and RPB case 3.1, Figs 6a and c). In RPB case 2.1 and RPB case 3.1 we use the same initial models, M2 and M3, as DRT case 2 and DRT case 3, respectively. Since we employ the same inversion method and the only difference here is the data used in the inversion, we thus believe that the artificial structures in the DRT cases are associated with the direct stacking, which causes significant distortion to the data. In RPB case 2.2 and RPB case 3.2, we simultaneously invert three receiver functions stacked at $p_0 = 5.5$, 6.5 and 8.0 $\mathrm{s\,deg^{-1}}$ starting from the initial model M2 and M3, respectively (Figs 6b and d). As it has been shown in the RPB case 1.4, using multiple rays enhances model convergence and inversion stability.

To further test our method, we also use another type of initial models. Initial models M4 and M5 have a crustal velocity 10 per cent higher and 10 per cent lower, respectively, than the average crustal velocity of the target model (Fig. 2). Consequently, the velocity jump at the Moho of these two models is very different from the target model, leading to a very poor fit between the model synthetics and the data. Receiver function inversion is well known for its non-linearity, so we expect that it is likely to be difficult to converge to the target model from these two initial models. In the DRT case 4, we utilize the enhanced preconditioning method to invert

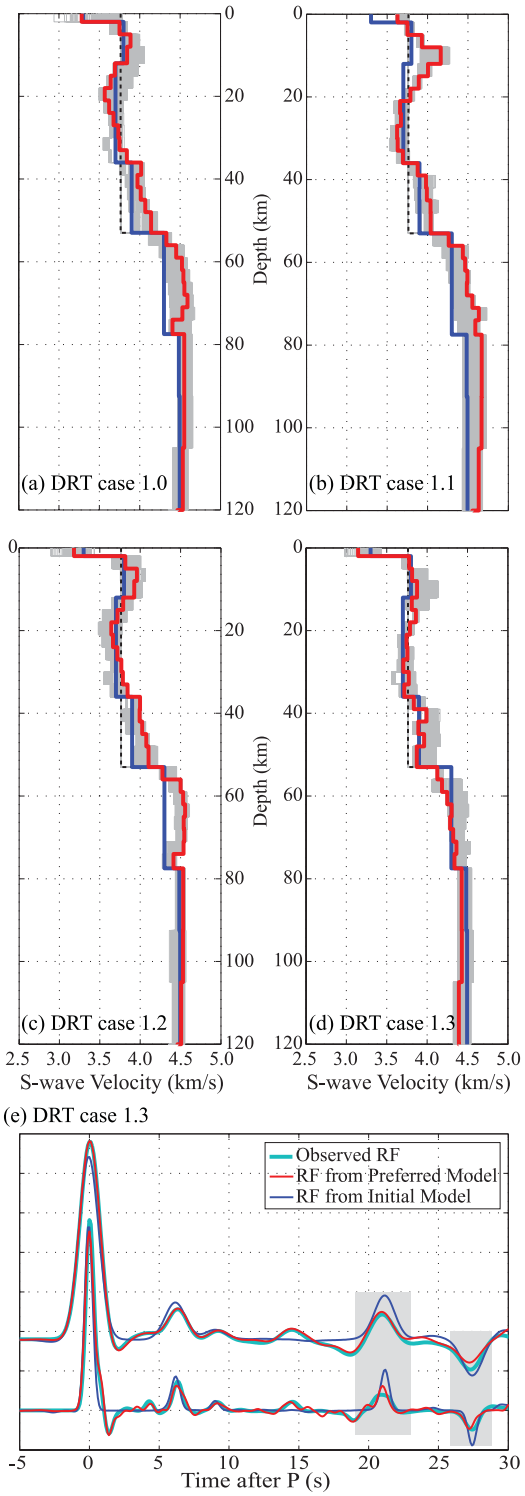

Figure 3. (a)–(d) Results of four inversions using different methods on directly stacked synthetic receiver function data and the initial model M1 in DRT cases 1.0–1.3. In each panel, initial model M1 is plotted in thin dashed black line and the target model is shown in blue line. Grey thin lines indicate all the inverted velocity profiles using different smoothness parameters s, whereas the red line marks the preferred model. (e) Receiver function waveforms in DRT case 1.3. Green thick lines are the 'observation' data, whereas thin lines in blue and red correspond to receiver functions generated from the initial and final velocity model, respectively. The low- and high-frequency receiver functions shown in the top and bottom, respectively, are processed by a Gaussian low-pass filter with $a = 1.0$ and 2.5.

the directly stacked data from the initial model M4. The resulting *S*-wave velocity model shown in Fig. 7(a) is higher than the target model in most of the 120-km depth range. It also has a deeper Moho than the one in the target model. The higher velocity and deeper Moho can be explained by the velocity–depth trade-off. The inversions with the RPB stacks, on the other hand, show a much better convergence to the target model during the iterations (Figs 8a and b). In both the RPB case 4.1 (inversion using a single stack at $p_0 = 6.5\,\mathrm{s\,deg^{-1}}$) and RPB case 4.2 (inversion using three stacks at $p_0 = 5.5$, 6.5 and $8.0\,\mathrm{s\,deg^{-1}}$), it appears that the depth–velocity trade-off affects only the first a few iterations. In the later iterations, the inversions are able to rectify the models to the target model for any given smoothness parameter (Figs 8a and b). Comparing the results between RPB case 4.1 and 4.2, the relatively high-velocity zones at ∼5–15, ∼40–55 and ∼55–75 km shown in RPB case 4.1 are corrected in RPB case 4.2, suggesting that using multiple receiver function stacks can result in a more stable inversion.

The initial model M5 is used in the inversions of DRT case 5, RPB case 5.1 and RPB case 5.2. In contrast to the DRT case 4, the inversion in the DRT case 5 results in a velocity profile lower than that of the target model, and a shallower Moho (Fig. 7b). Only when a very smooth parameter ($s \geq 0.85$) is used in the inversion, the crustal velocity can be rectified back to the expected values. Because of the large velocity jump at the Moho in M5, large smooth parameters are required to reduce velocity contrast between model layers. With this starting model, the inversion using one RPB stack (RPB case 5.1) performs just like DRT case 5 and fails in recovering the target model(Fig. 8c). Although inversions with high-smoothness parameters do generate acceptable models, the resultant Moho depth vibrates between the initial model and the target model when variable smoothness is given. The inversion of RPB case 5.2 uses multiple RPB stacks, but it is also unable to generate an acceptable model either. As shown in Fig. 8(d), all models tend to have lower velocity with a shallower Moho. Comparing the results of these two RPB cases, we have reason to believe the inversions in RPB case 5.1 are unstable in spite of the acceptable models occasionally derived at high-smoothness controls. For real seismic data, because the crustal and Moho structures are unknown in advance, we have no idea whether increasing smoothness makes the inverted model towards or away from the real Earth structure. In general, a change in Moho depth only results in a shift on the arrival times of the converted and reverberated phases in the synthetic receiver function trace, whereas changes in crustal velocity alter both the arrival time and amplitude of these phases. It thus makes the target model more difficult to be recovered from those starting models with large deviations in velocity than those with large variations in Moho depth.

Overall, the above synthetic tests have demonstrated the capability of the enhanced preconditioning inversion method with data from the RPB stacking. Before we further validate this new technique with real data, here we summarize what we have learned from the synthetic tests: (1) the regularization on both model and data reduces the number of free parameters and force an inversion towards the correct direction. In a rather 1-D structure beneath the station, the RPB stacked data possess more information about the converted and reverberated *S*-wave phases leading to a more realistic velocity structure through the inversion procedure. A combination of regulating the model on the inversion side and using more accurate observation on the data side is capable of stabilizing the inversion process and also significantly reducing

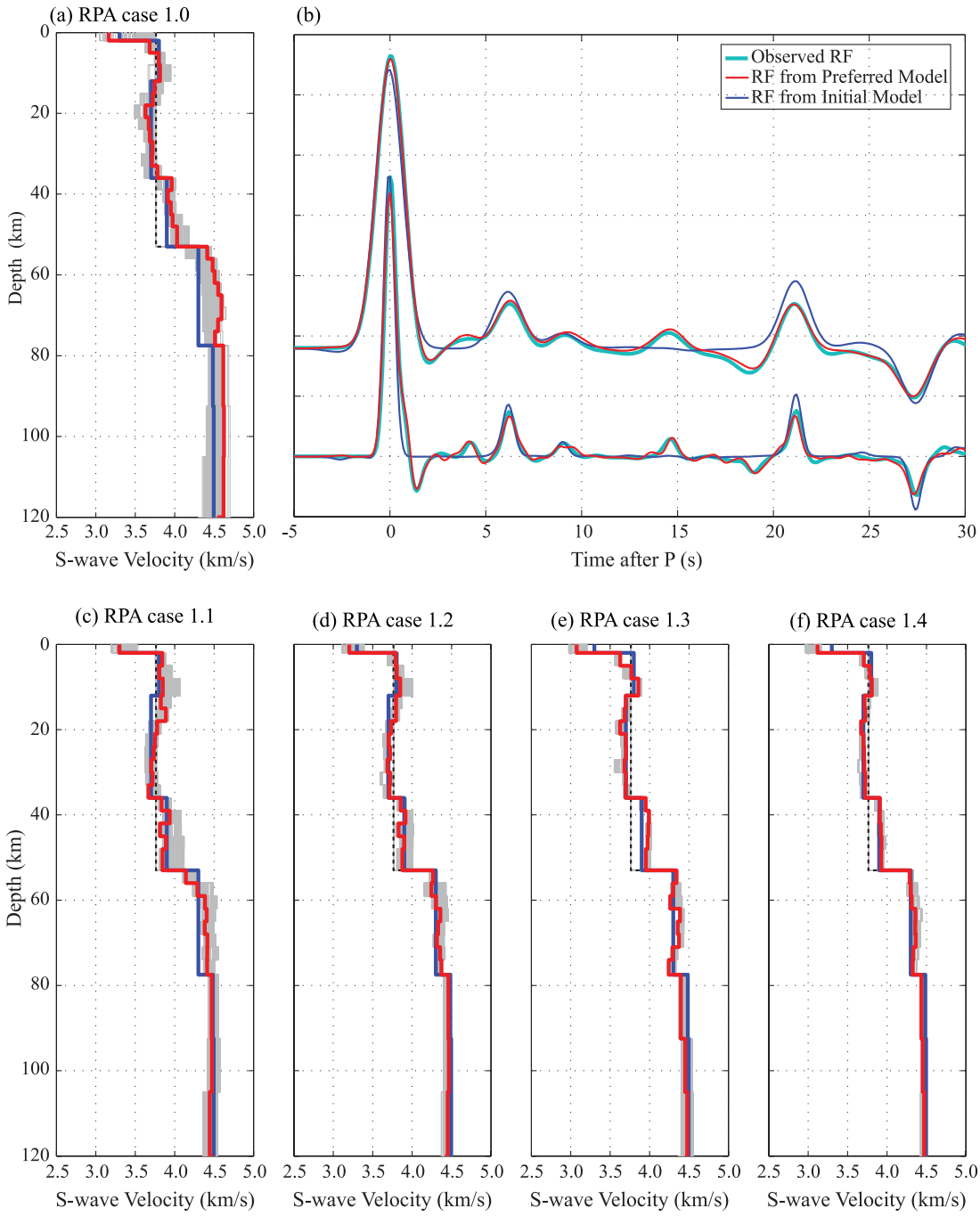

Figure 4. (a), (c)–(f) Results of five different inversions using RPB stacked synthetic receiver function data and the initial model M1 in RPB cases 1.0–1.4. They are plotted in the same way as Figs 3(a)–(d). (b) Receiver function waveforms in RBP case 1.0. They are plotted in the same manner as Fig. 3(e).

the non-uniqueness of the inversion. Resolving only one side of the problem may not be able to produce sufficient effect as anticipated; (2) simultaneously inverting multiple RPB stacks can substantially improve the model convergence, but helps little in resolving the inherent depth–velocity trade-off problem in receiver function modelling and (3) a well constructed starting model is always preferred in receiver function modelling. Even non-linear method cannot guarantee an inversion moving towards the correct direction if it starts from a poorly configured model. As shown in RPB case 5.2, all of our improvements cannot prevent the inversion from falling into a local minimum due to the strong depth–velocity trade-off.

4 APPLICATION TO DATA

In order to extend the success with synthetic data, we have applied our new methods on two broad-band stations situated in China with distinct geological environments. Station XAN is located at central China, and MDJ is located at northeast China (Fig. 9). Both stations belong to the IRIS GSN network installed in the last century. They have been recording a large amount of waveform data over their 20+ yr operation history. We collect teleseismic events with $M_w \geq 5.9$ and an epicentral distance between 30° and 90° occurring in the last 10 yr and generate individual receiver function traces using the time-domain iterative deconvolution method

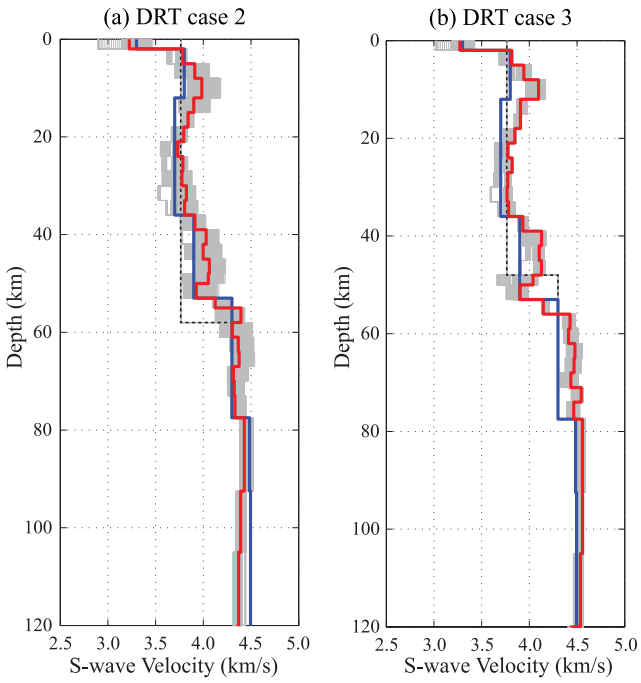

Figure 5. Results of the inversions using direct stacks in DRT case 2 (a) and in DRT case 3 (b). Initial model M2 is used in (a) and M3 is used in (b). Lines in each panel are the same as in Figs 3(a)–(d).

(i.e. Ligorría & Ammon 1999). We select over 300 highly coherent individual receiver functions for analysis at each station using the cross-correlation matrix method (Chen *et al.* 2010; Tkalčić *et al.* 2011). The cross-correlation coefficients of all the selected receiver functions are ≥85 per cent. The selected data set has a good back-azimuthal coverage (Fig. 9, insets), although most of the selected events are located in the Tonga-Fiji and the Indonesia subduction zones. We further filter the receiver functions with two types of Gaussian low-pass filters, one with a width parameter $a = 1.0$ and another one with a width of $a = 2.5$ to generate low- and high-frequency data.

4.1 Results from station XAN

Station XAN is situated at the Qinling Mountain Range, an orogenic belt located between the Sino-Korean Craton and the Yangtze Craton. The northward underthrusting of the Yangtze Craton to the Sino-Korean Craton occurring in the Permian-Triassic has resulted in the development of complicated crustal structure along the Qinling-Dabie suture. We first use the H–κ analysis of Niu *et al.* (2007) to determine the crustal thickness and V_P/V_S ratio. Two maximum amplitudes appear on the H–κ contour map (Fig. 10a). One is located at a depth of 39 km with a V_P/V_S value of 1.66, whereas the other is located at a depth of 34 km and a higher V_P/V_S value of 1.74. By comparing the results with previous studies (i.e. Sun & Toksöz 2006; Chen *et al.* 2010), we prefer to interpret the deeper one as the Moho, as it also shows a better prediction of the $0p1s$ and $2p1s$ arrival times (Fig. 10b). Fig. 10(a) shows the results using low-frequency ($a = 1.0$) receiver function data, but the same two peaks also present on the H–κ contour map using high-frequency ($a = 2.5$) receiver functions. The H–κ measurements derived from both frequency bands are consistent with each other (Table 2). The slight difference between them can result from either processing

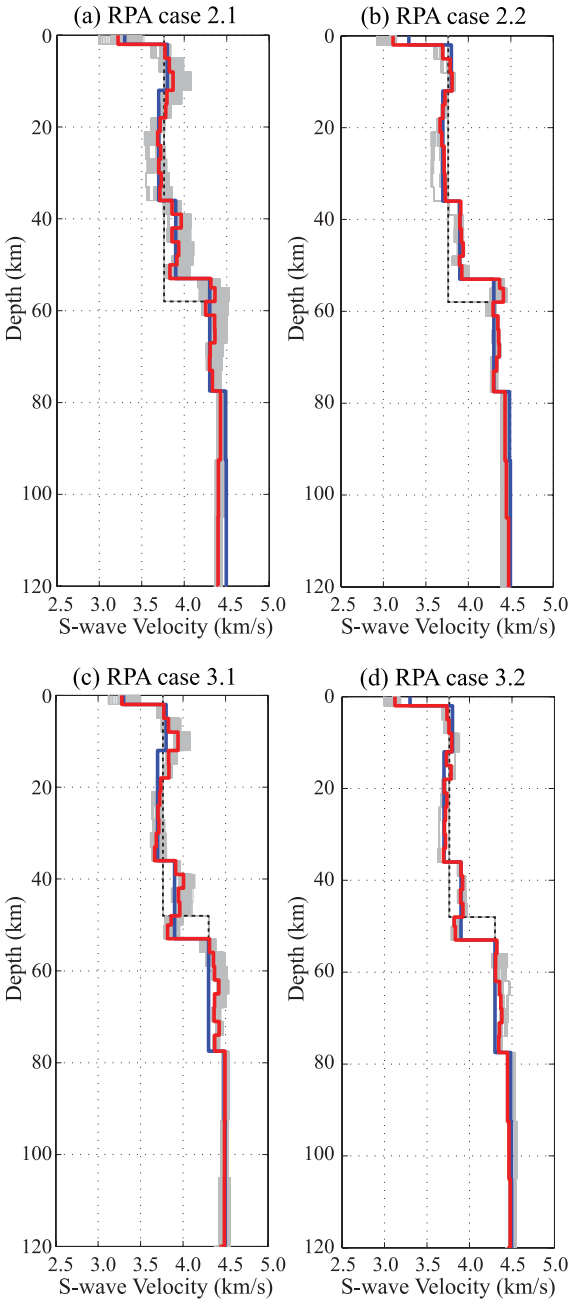

Figure 6. Results of the inversions using RPB stacks in RPB case 2.1 (a), RPB case 2.2 (b), RPB case 3.1 (c) and RPB case 3.2 (d). Initial model M2 is used in (a) and (b), and M3 is used in (c) and (d). Lines in each panel are the same as in Figs 3(a)–(d).

induced noise or the nature of the crust–mantle boundary, which may cause different frequency response.

Tkalčić *et al.* (2011) inverted the receiver function data of this station and found no sharp velocity jump corresponding to either depth observed in H–κ measurements. The velocity changes associate with a crustal–mantle transition starts from as shallow as 15 km and extends to almost 50 km deep, which makes it almost impossible to identify the Moho. We select the receiver functions in the backazimuth range of 90–240° to repeat the inversion analysis. The backazimuth range is determined for two reasons: (1) most of the teleseismic events fall in this range and (2) we would like to compare our results with Tkalčić *et al.*

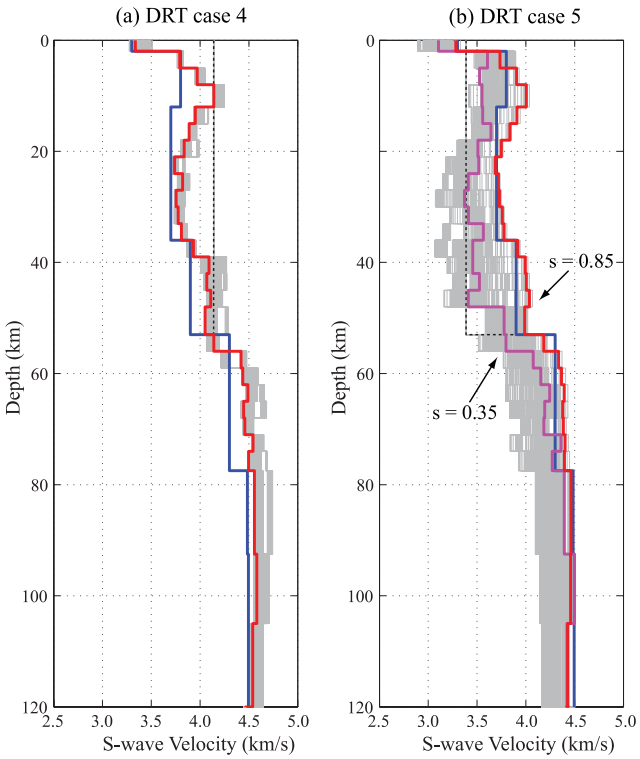

Figure 7. Results of the inversions using direct stacks in DRT case 4 (a) and DRT case 5 (b), which use the initial model M4 and M5, respectively. Initial model is plotted in thin dashed black line and target model is shown by blue line. Grey thin lines indicate the inverted velocity profiles using different smoothness parameters s. Red lines indicate the optimal models recovering target model. In (b), the magenta line indicates the model using $s = 0.35$, and the red line indicates the model with $s = 0.85$.

(2011), which also used earthquakes roughly in this backazimuth range.

Using the estimated H and κ values, we make RPB stacks at three reference ray-parameters, $p_0 = 5.5$, 6.5 and 8.0 s deg^{-1} (Fig. 10b). As noted before an increase of ray-parameter results in an increase and decrease of Ps and reverberation arrival time, respectively, together with an increase in amplitude. We also create the direct stack of receiver function data by simply averaging the receiver functions at each time point. In order to illustrate the effect of RPB stacking, we plot the standard deviation curbs (grey area) of all receiver functions involved in each type of stack around the stacked trace in Fig. 11(a). The average standard deviation of direct stack is 3.4 per cent, while the average standard deviation of RPB stack reduces to only 2.3 per cent. The decrease in variance, particularly around the Ps conversion and the reverberation phases is presumably related to the corrections on moveout and amplitude of the RPB stacking.

Following Tkalčić *et al.* (2011), we also make a great effort to construct the initial velocity. To do so, we refer multiple model sources and employ the following steps: (1) we conduct a coarse grid search to determine a preliminary S-wave velocity model with four layers in the crust (Sandvol *et al.* 1998), we then calculate the corresponding P-wave velocity and density using the estimated V_P/V_S ratio from the H–κ analysis and Birch's law; (2) we take the Pn and Sn models from tomography studies (e.g. Pei *et al.* 2004, 2007; Sun *et al.* 2004; Sun & Toksöz 2006) for the uppermost mantle velocity; (3) the velocities of the rest upper mantle are taken from the global AK135 model (Kennett *et al.* 1995); (4) we fix

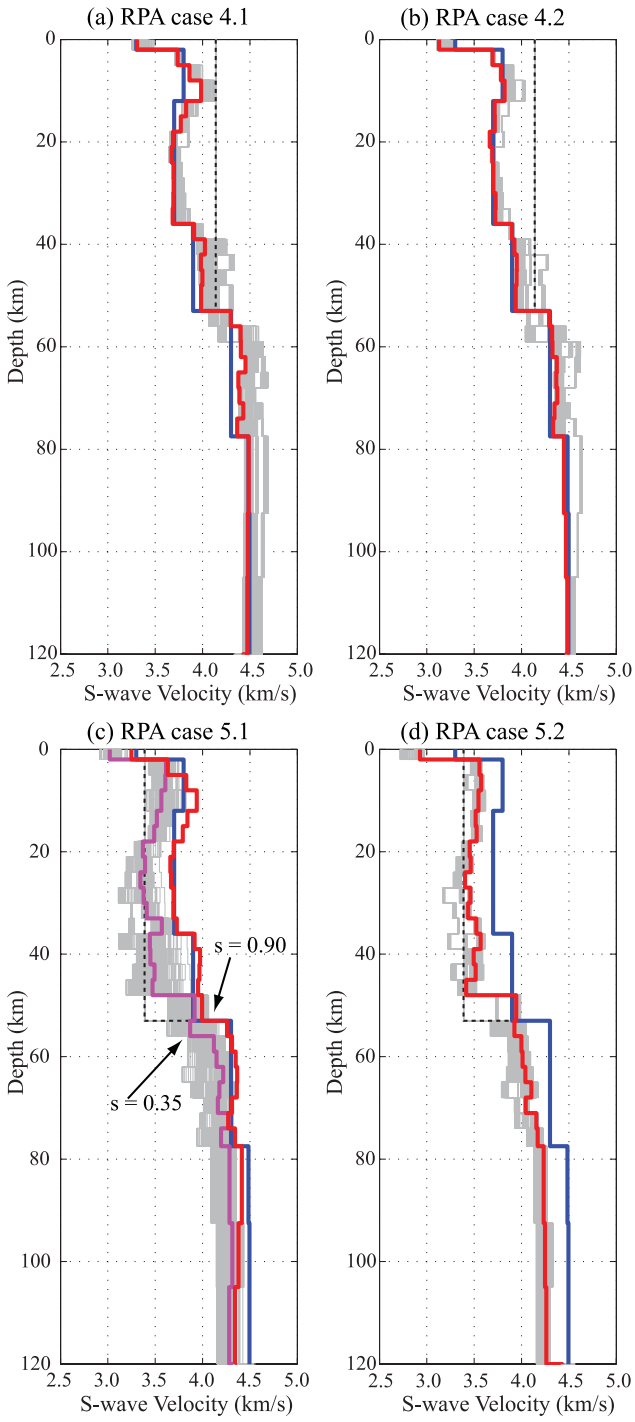

Figure 8. Results of the inversions using RPB stacks in RPB case 4.1 (a), RPB case 4.2 (b), RPB case 5.1 (c) and RPB case 5.2 (d). The initial model M4 is used in (a) and (b), and M5 is used in (c) and (d). Lines in each panel are the same as in Fig. 7.

the Moho at the depth determined by the H–κ analysis. Finally, we parametrize the crust and the uppermost mantle with thin constant velocity layers of 2–3 km thick. The initial velocity model has two velocity jumps at 30 and 40 km, respectively (black dashed line in Fig. 12c).

In addition to the initial model, our inversion with the direct stack also has some notable difference from Tkalčić *et al.* (2011). We use more teleseismic events, and generate both low- and high-frequency

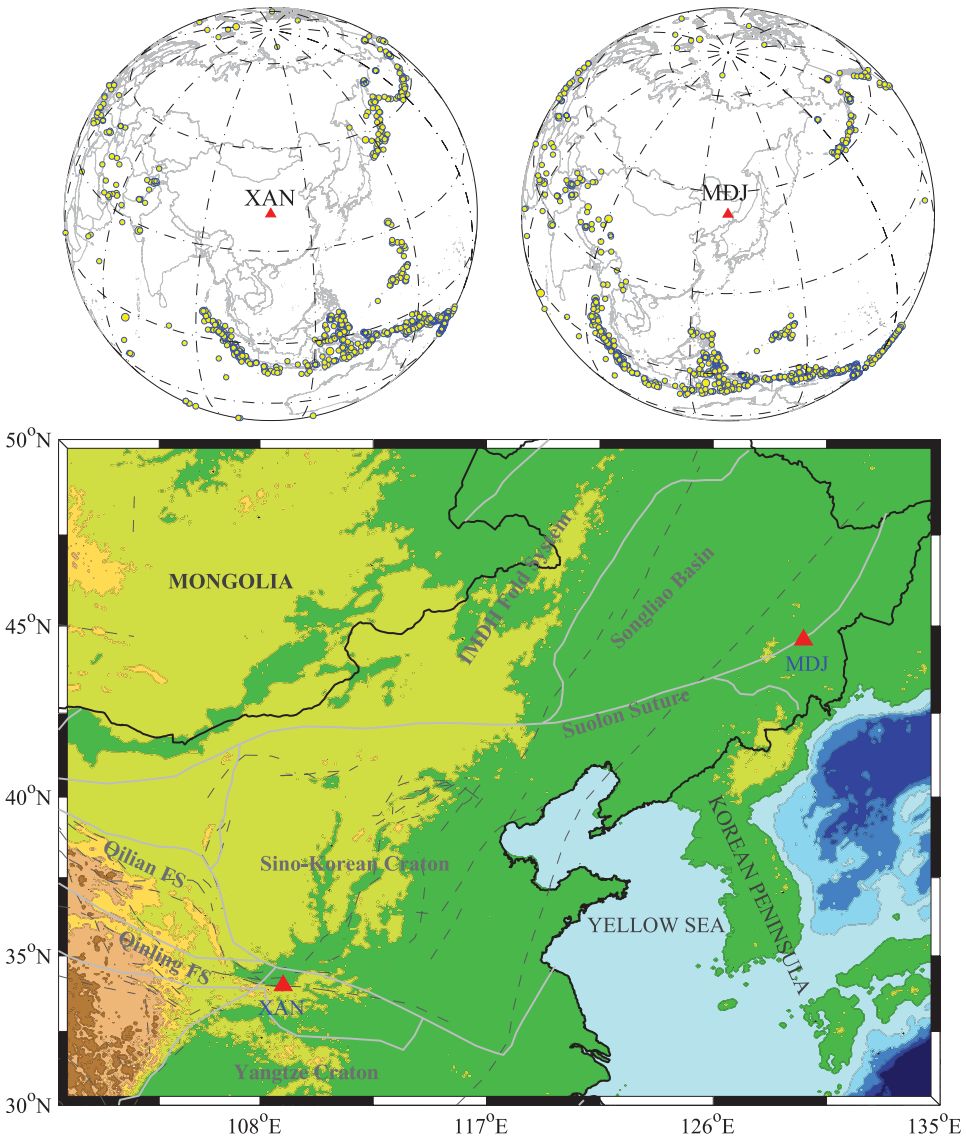

Figure 9. Map showing the geographic locations of two seismic stations XAN and MDJ used in this study. Two maps on the top show the distribution of teleseismic events recorded by the two stations.

receiver functions for the inversion as shown in Fig. 12(a). Most of all, we employ the enhanced preconditioning inversion scheme that regulates the model and weights different segments of the data. For each inversion, we use scalar smoothness control s, and vary s from 0.05 to 1.0 for different inversions, and the VR of each inversion is shown in Fig. 12(b) as a function of s. Although the VR reaches to the maximum at $s = 0.1$–0.2, the resultant model is rougher than the model of Tkalčić *et al.* (2011). We thus choose the one corresponding to $s = 0.3$ as the final model (red solid lines in Fig. 12c). For comparison, we also show the models of individual iterations (grey lines) together with the initial model (black dashed line) in Fig. 12(c). Our final model is actually quite similar to the model of Tkalčić *et al.* (2011). It possesses a broad transition zone featured by a gradual increase of *S*-wave velocity from 3.5 to 4.25 km s^{-1} at depths from approximately 30–40 km (Fig. 12c). Although the transition zone is much narrower than that of Tkalčić *et al.* (2011) (∼15–50 km), it is also not easy to define a Moho from this model.

In the RPB data inversion, we make use three RBP stacks of receiver functions aligned at reference ray-parameters of $p_0 = 5.5$, 6.5 and 8.0 s deg^{-1} and filtered with $a = 1.0$ and 2.5, respectively, and conduct the inversion with the enhanced preconditioning method as shown in Fig. 13. The final model, which is indicated by the red line in Fig. 13(c), is obtained from the inversion with $s = 0.45$. The most prominent feature revealed by the *S*-wave velocity profile is the increased sharpness of the Moho at ∼38 km. There is another mild velocity jump at ∼30 km, which may be related to the shallower peak in the H–κ contour map.

4.2 Results from station MDJ

The other station MDJ is located near the Suolon suture in northeast China (Fig. 9). Reports on the crustal thickness in the vicinity of MDJ from previous studies are really diverse, covering a depth range of 27–42 km. Regional *P*- and *PP*-waveform modelling by Beckers *et al.* (1994) suggested that the crust is rather thin, ∼27 km thick. Study of active source refraction data by Yuan *et al.* (1986) indicated a 34–36-km-thick crust near the station. Surface wave and receiver function studies, on the other hand, found that the Moho depth in the area is ∼42–43 (Feng *et al.* 1981) and 36–42 km (Mangino *et al.* 1999), respectively. Our H–κ analysis gives an estimate of

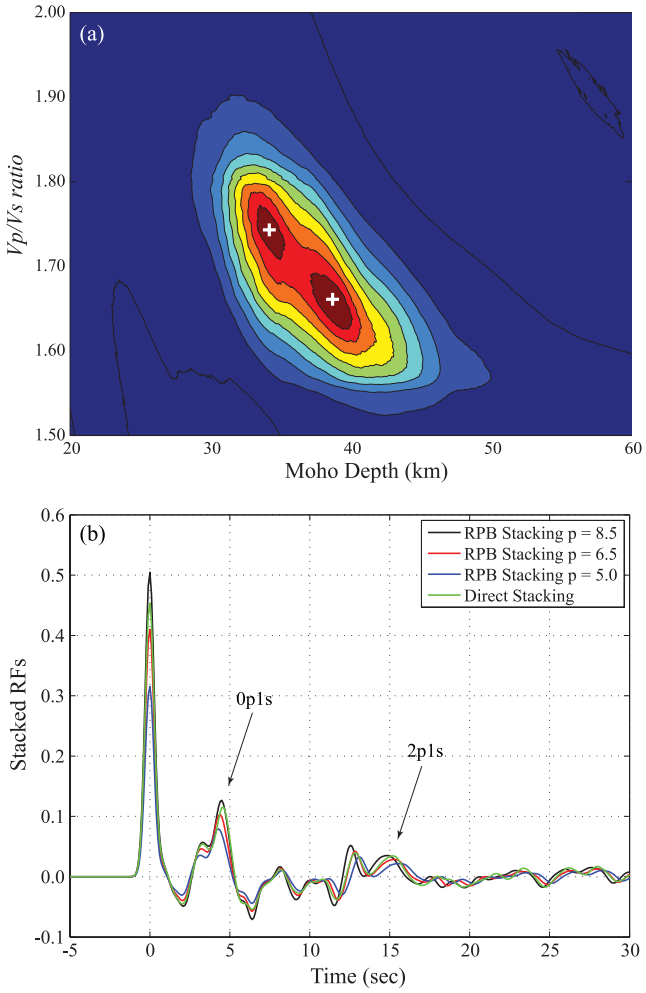

Figure 10. (a) Result of the H–κ analysis obtained from station XAN. Colour contours show the summed amplitude as a function of crustal thickness and V_P/V_S ratio. The peaks marked by the white plus signs indicate the optimal crustal thickness (H) and average V_P/V_S ratio (κ). (b) Direct stack (green) of receiver functions is plotted to compare with RPB stacks generated at different ray-parameters (black, red and blue). Events in the backazimuth range of 90° to 240° are used in the stacking.

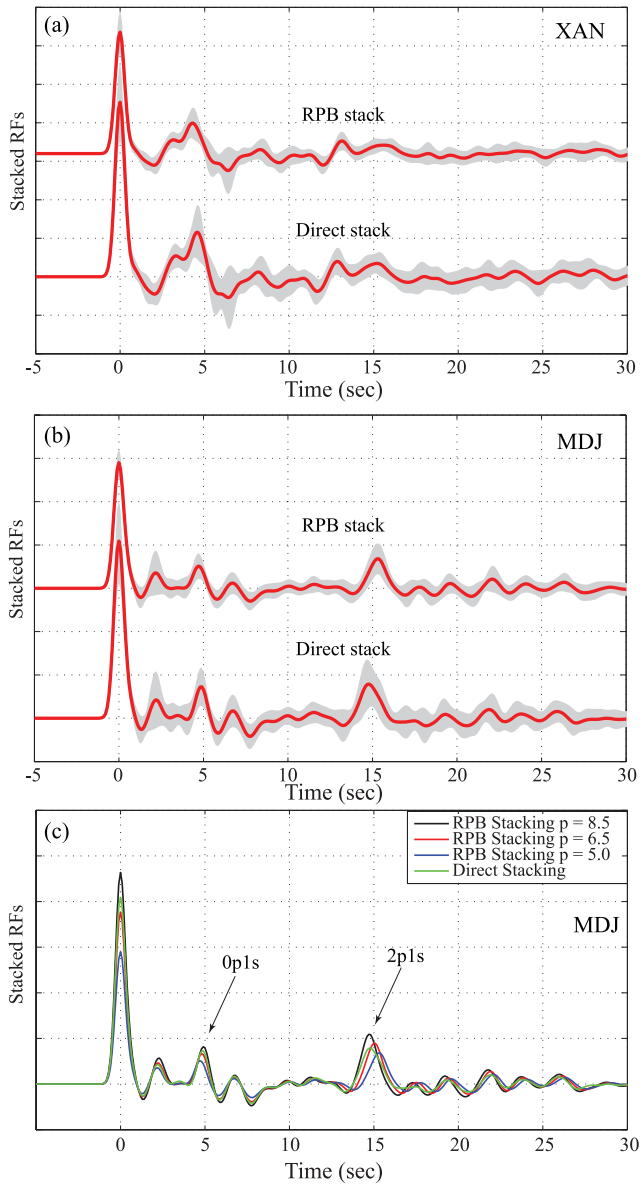

Figure 11. (a) Standard deviation curbs (grey areas) of the receiver functions in the backazimuth range between 90° and 240° recorded at XAN are shown with the stacked data. The RPB and direct stacks are shown in the top and bottom of the panel, respectively. (b) Same as (a) but for station MDJ, where events in the backazimuth range of 120–240° are used in the stacking. (c) Direct stack (green) of receiver functions at MDJ is plotted to compare with RPB stacks generated at different ray-parameters (black, red and blue).

the crustal thickness of 36–37 km and V_P/V_S ratio of 1.74–1.80 (Table 2). The diversity among these measurements may reflect the complexity of the Moho structure beneath the station, as well as the sensitivity of different types of the data. It may also result from the biases in data processing and in the analysis methods. Ma *et al.* (2007) performed both linearized and non-linear inversions using receiver function data to obtain an *S*-wave velocity profile beneath the station down to 200 km. They found that the two methods yield nearly similar velocity structure in the crust and the uppermost mantle. There are, however, significant differences in the upper mantle below. Models from the linearized inversion always show a low-velocity zone (LVZ) at approximately 140–160 km depths, which is not presented in the models from non-linear inversion. It was not revealed in Ma *et al.* (2007) whether the LVZ reflects the true structure of the Earth, or is just an artefact associated with the linearized inversion.

Chen *et al.* (2010, 2011) found evidence for dipping and anisotropic structure beneath the station and attributed their origins to the Jurassic Suolon suture, which was reactivated during the

Table 2. H–κ measurements on XAN and MDJ.

Station	Lat	Lon	$a = 1.0$		$a = 2.5$	
			H (km)	V_P/V_S (κ)	H (km)	V_P/V_S (κ)
XAN	34.03	108.92	38.88 ± 0.20	1.659 ± 0.005	40.12 ± 0.07	1.662 ± 0.001
MDJ	44.62	129.59	37.36 ± 0.08	1.739 ± 0.003	36.39 ± 0.05	1.801 ± 0.001

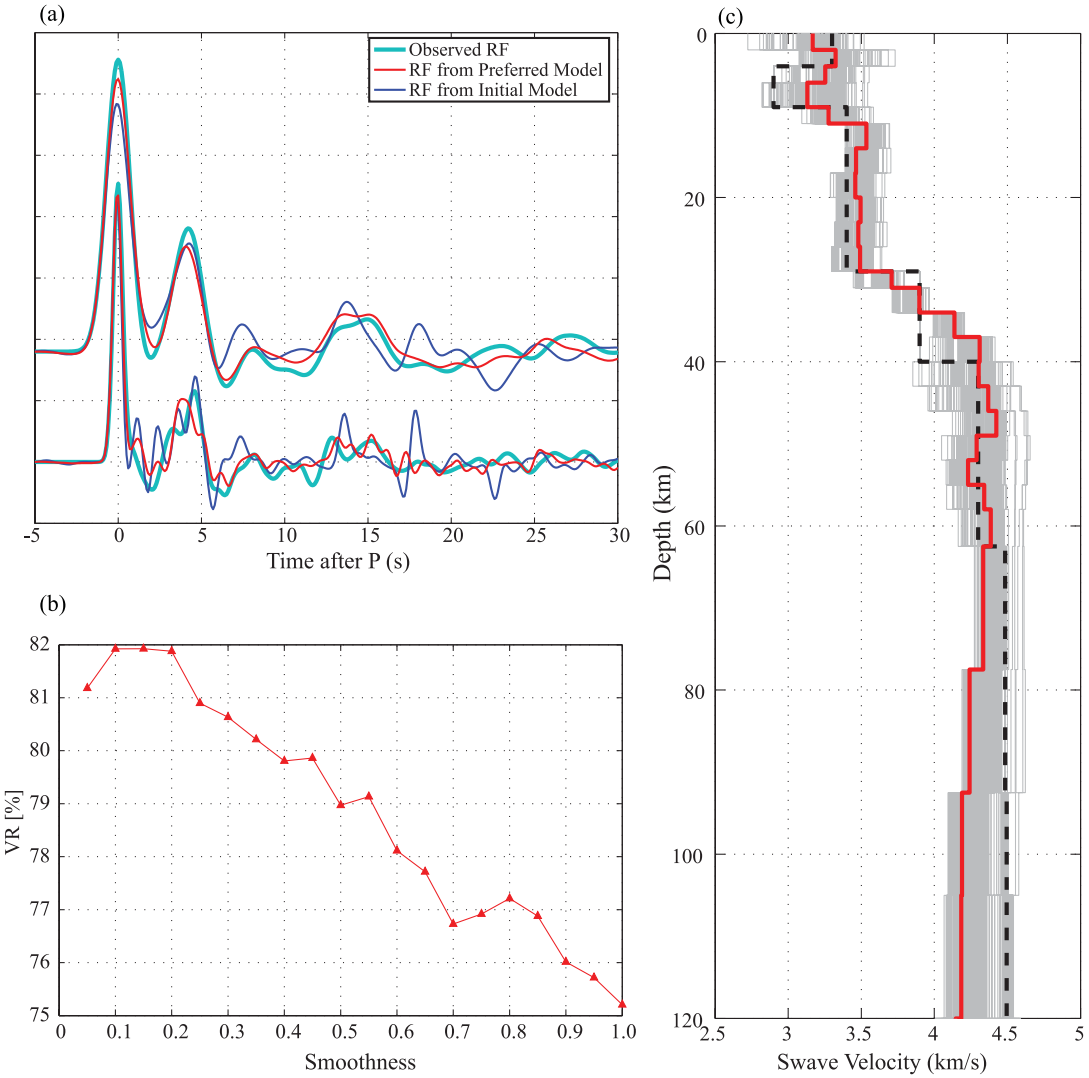

Figure 12. Results from the inversion using directly stacked receiver function data recorded at the XAN station. (a) Green thick lines indicate observation data, whereas thin lines in blue and red correspond to the synthetic receiver functions generated from the initial and final velocity model, respectively. Both low- and high-frequency receiver functions generated by a Gaussian low-pass filter with width $a = 1.0$ (the top ones) and $a = 2.5$ (the bottom ones) are employed in the inversions. (b) VR is shown as a function of smoothness parameter s. (c) S-wave velocity profiles. The initial model is plotted in thin dashed black line and the preferred model is shown in red line. The grey thin lines indicate all the inverted velocity profiles using different smoothness parameters s.

Cenozoic. The major heterogeneity here is thus the structural difference between the two sides of the NE–SW trending Suolon suture. To avoid this complexity, we choose earthquakes between backazimuth 120° and 240° where the receiver functions have a high pair cross-correlation coefficient for data stacking. The standard deviation of direct stack is ∼3 per cent, but reducing to 2.1 per cent after the RPB corrections (Fig. 11b). The comparison between three RPB stacks (with reference ray-parameters of $p_0 = 5.5$, 6.5 and 8.0 s deg^{-1}, respectively) and the direct stack is shown in Fig. 11(c).

We process the data in a similar manner as we do with station XAN. First, we filter the receiver functions with a Gaussian width $a = 1.0$ and 2.5. For each frequency band, we then conduct the RPB stacking at three different reference ray-parameters, $p_0 = 5.5$, 6.5 and 8.0 s deg^{-1}. We also used results from grid search, H–κ analysis and Pn/Sn tomography to build the initial model. We divide the crust and the uppermost mantle down to 60 km deep into layers with a thickness of 2 or 3 km, and divide the upper mantle between 60 and 200 km into layers of 5–9 km thick. Finally, we use the enhanced pre conditioning method to invert both the DRT and RPB stacked data.

The inverted S-wave velocity structure with the DRT and RPB data is shown in Figs 14 and 15, respectively. In general, the two models agree with each other quite well in the crustal and uppermost mantle depths. Both show a sharp velocity jump at ∼37 km, consistent with the H–κ analysis. Ma *et al.* (2007) also found a relatively sharp velocity jump at nearly the same depth, while the crust–mantle boundary in the model of Mangino *et al.* (1999) covers a much broader depth range. Both the DRT and RPB models also show a high-velocity zone at 20 km depth, followed by a low-velocity zone in the lowermost crust. However, the most significant difference is that the DRT model reveals a distinct LVZ in upper mantle at ∼140 km depth, but which is not seen in the RPB model. The preconditioning used in the DRT and RPB inversions is the same, and the only difference in these inversions is the input data. Since the RPB stack does a better job in preserving the true amplitude and the shape of the three later arrivals, we tend to conclude that that the mantle LVZ shown in the DRT model is an artificial structure associated with the improper stacking of receiver function data. The amplitude of the crustal multiples predicted by the sharp

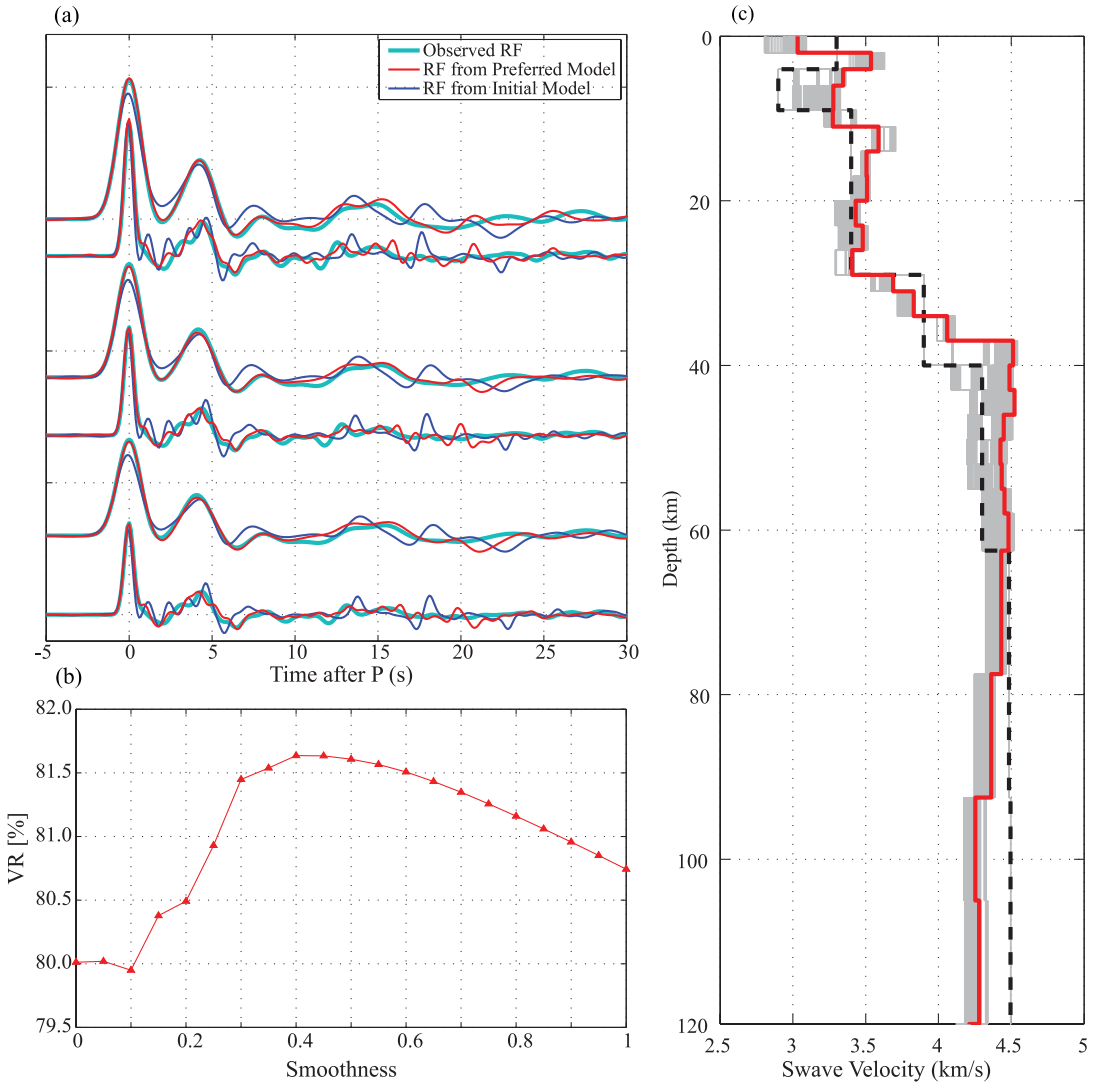

Figure 13. Results from the inversion using RPB stacked receiver function data of XAN. (a) Green thick lines are observation data, whereas thin lines in blue and red correspond to the synthetic receiver functions generated from the initial and final velocity model, respectively. Both low- and high-frequency receiver functions generated by a Gaussian low-pass filter with width $a = 1.0$ and 2.5 are employed in the inversions. Three reference ray-parameters, $p_0 = 5.5$, 6.5 and 8.0 s deg^{-1}, are used to form RPB stacks for per couple of low- and high-frequency data. The receiver functions corresponding to the three p_0 are plotted in the order form the bottom to the top. For per couple of waveforms, the top one is for $a = 1.0$ and the bottom one is for $a = 2.5$. A total of six receiver function stacks are used in the inversion. (b) VR is shown as a function of smoothness parameter s. (c) S-wave velocity profiles. The initial model is plotted in thin dashed black line and the preferred model is shown in red line. The grey thin lines indicate all the inverted velocity profiles using different smoothness parameters s.

Moho is larger than the one shown in the DRT stack, which is caused by out-of-phase stacking (Fig. 11b). For example, the relative amplitude of $2p1s$ with respect to direct P in the three RPB stacks is ∼24 per cent, whereas the relative amplitude in the DRT stack is only ∼19 per cent (Fig. 11c). To compensate the weak amplitude of the crustal multiples in the DRT data, an LVZ at ∼140 km is required to generate Ps conversion energy that partially cancels the $2p1s$ multiple. This explains why the mantle LVZ in the DRT model is unlikely a true structure.

5 DISCUSSION

Iterative linearized inversions of receiver function data are widely used nowadays to obtain high-resolution velocity structure of the crust and upper mantle beneath seismic stations. The input data of the inversions are usually the stacked receiver functions, as stacking can reduce the noise level and digital errors, and suppress 2-D/3-D structural effects, which otherwise may mislead the direction of an iterative inversion. As mentioned above the Ps converted phase and the two crustal reverberations possesses a slightly different slowness from the direct P wave, making it difficult to preserve the true waveform of the three phases simultaneously with a simple slant stacking method. The stacking induced error in data can be suppressed by carefully selecting and stacking receiver functions in a relatively narrow epicentral distance range. This practice is, however, not always possible when the heterogeneous global seismicity is taken into account. Using the entire individual receiver functions as the input data, and inverting them simultaneously is an alternative approach. However, as the SNR of the converted and reverberated waves in the individual receiver function data is generally not high, the large noise level can misguide the iteration direction.

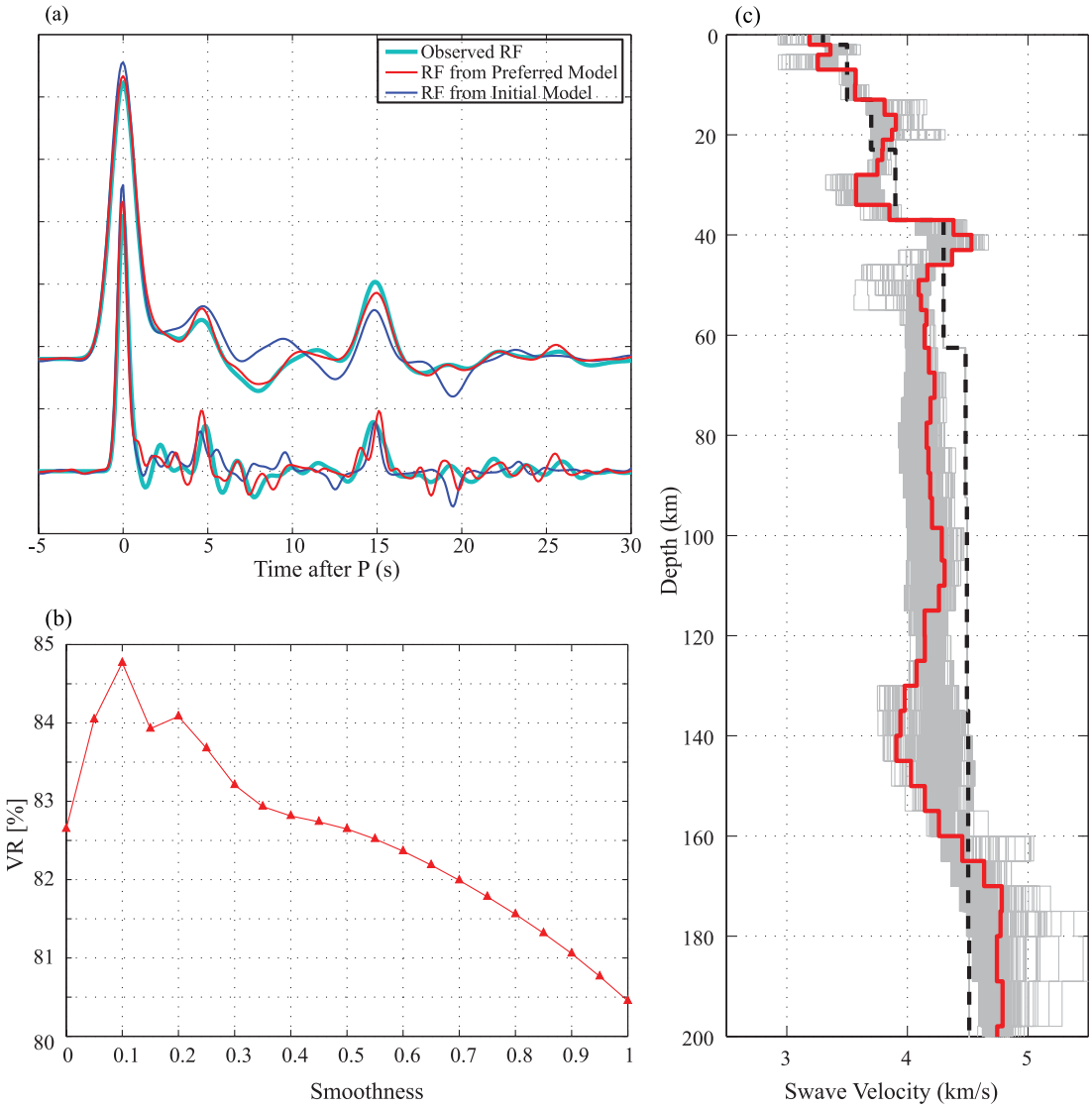

Figure 14. Same as Fig. 12, but for MDJ.

On the other hand, the out-of-phase stacking is less severe for low-frequency data compared to high-frequency data. Thus inversion with direct stacks of low-frequency receiver functions, for examples, those filtered with a Gaussian filter width $a = 1.0$, can work reasonably well as long as the distance coverage is not so large.

The 'four-pin' method developed in this study is able to correct the time-shift of the Moho conversion and two later reverberations with respect to a specific ray path, that is a reference ray-parameter as long as it is within the teleseismic distance range. The correction only requires the knowledge on the average crustal *P*-wave velocity V_P, crustal thickness, and average V_P/V_S ratio, which can be determined by the H–κ analysis, making it an automatic method to process all receiver function data. Stacking moveout corrected data is then straightforward avoiding the tedious tasks of data selection and analysis. The RPB stacking is thus easy to be implemented, particularly to large datasets such as the USArray. Although such a correction for distance moveout has not been used in receiver function inversion, it has been used in many other studies to separate variations of receiver functions related to distance and backazimuth. For example, analysing crustal anisotropy using multiple receiver functions requires the removal of the distance related time-shifts from the Ps arrival time (e.g. Liu & Niu 2011). Actually, the essence of the H–κ analysis is to find an optimal distance moveout for the Ps conversion and reverberation phases to obtain the maximum stacking amplitude of these phases.

We want to emphasize that the RPB correction and stacking proposed here are based on 1-D velocity models. When there are strong lateral heterogeneities within the crust, the arrival times of the Ps conversion phase and the major reverberation phases may significantly differ from those predicted by the 1-D reference model. If this is the case, then the 'four-pin' moveout corrections are not expected to be able to enhance signal coherence of receiver function data from different backazimuthal. However, if data from a single direction are sufficient, we expect that lateral heterogeneities have minimum effect on the data and therefore the 'four-pin' moveout corrections can still lead to in-phase stacking, as demonstrated by the MDJ station. Also it should be noted that the RPB stacking is a technique proposed for data processing, independent from the preconditioning inversion part, we expect it to be used in any inversion methods.

In order to stabilize the inversions, we have implemented four types of pre-conditions: (1) add smoothness to the models, (2) fix

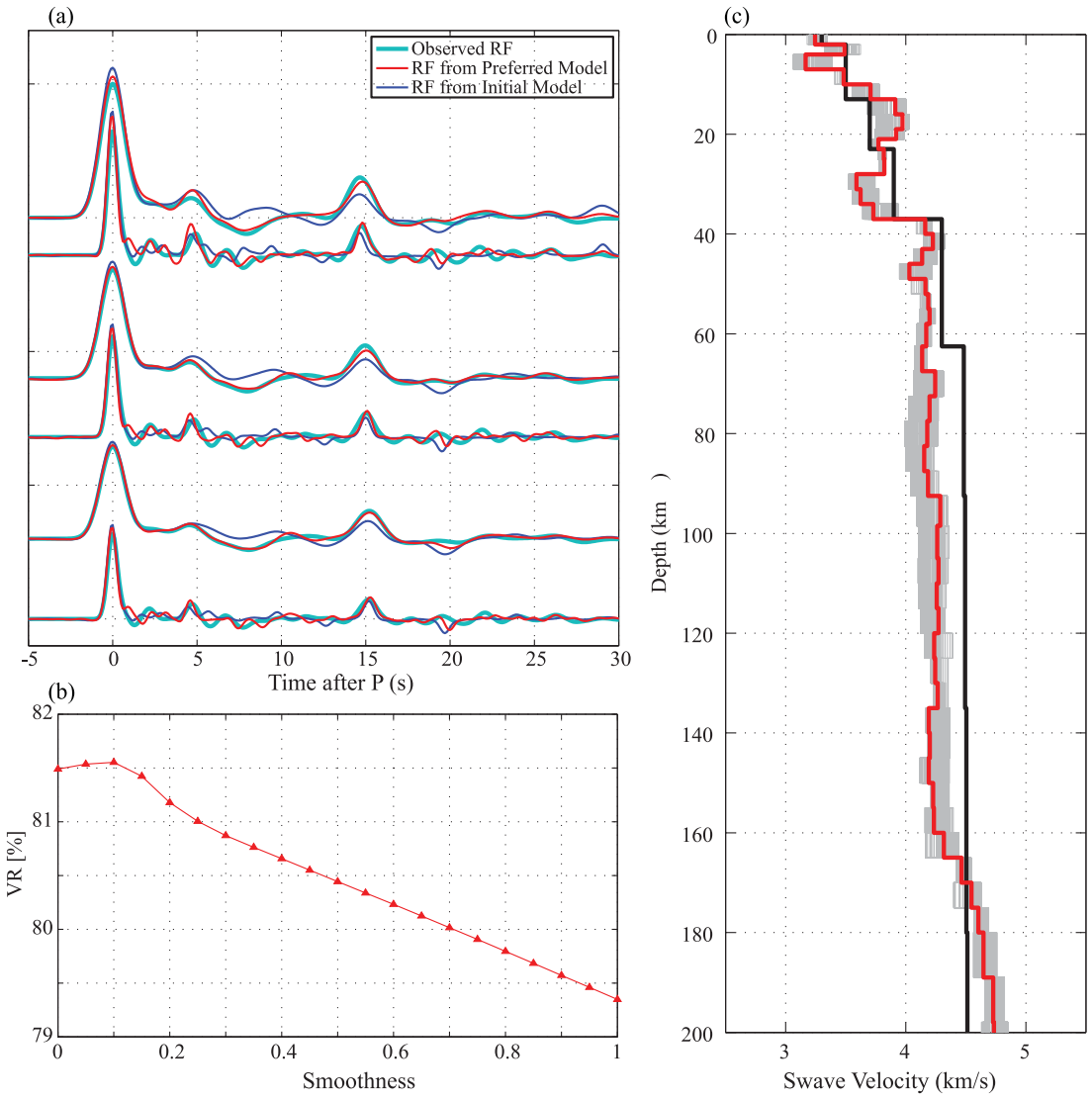

Figure 15. Same as Fig. 13, but for MDJ.

the deep layers, (3) balance the data filtered with different Gaussian functions and (4) weight differently between the conversion and reverberation segments. Among the four types of preconditions set up on model and data in this study, the regularization on the model smoothness and the balance between the high- and low-frequency receiver function data are always required in order to obtain physically meaningful velocity models. If these two preconditions are not applied, the resulting *S*-wave velocity may go to negative during the first few iterations. Selecting the smoothness parameter s is somewhat subjective. A useful criterion for choosing this parameter is to find the value that produces the inversion residual in the same order of the pre-signal noise level in the receiver function stacks. However, a poorly configured inversion may not be able to reduce the residual to the pre-signal noise level by only adjusting the smoothness parameter. In general, waveform fit (characterized by VR) usually decreases with increasing model smoothness, although this relation could become more complicated when multiple regularizations are implemented (Fig. 13b, for example). On the other hand, a set of depth-dependent smoothness parameters (eq. 5) can also be easily employed in the linear equations. In practical implementation, we can first search a wide range of smoothness parameters to determine the optimal value that gives a good balance between model roughness and waveform fit (VR). We have shown this type of plot (VR versus model smoothness) in the results of all the inversions conducted at the two stations in Figs 12(b) to 15(b). We then tune the depth-dependent smoothness parameters ($\boldsymbol{S}$ in eq. 5) to reduce the smoothness at depths with large velocity jumps and raise smoothness at other depths.

The choice of the weight parameter w_{ii} to fix velocity of a certain layer is also subjective if only it is large enough to keep the velocity almost intact during an inversion. In this study, we have set $w_{ii} = 1.0$ to fix velocity structure between 200 and 500 km in our inversions using both synthetic and field observed receiver function data. We must say this scheme is not limited to the deepest layers in the upper mantle, but also can be applied to crustal layers when their *S*-wave velocity is well constrained by previous studies.

Balancing the weight of different segments of a receiver function trace for the inversion has something to do with the length of each segment. Assuming that a receiver function trace is 50 s long, we divide it into a 10 s Ps segment and a 40 s reverberation segment. The second segment is thus three times longer than the first one. Given the different length of the two segments, we adopt

a similar method used in balancing receiver function and surface wave dispersion data (refer to eqs 10and 11 in Julià *et al.* 2000) to properly weight the two segments (eq. 8). Our numerical tests suggest that assigning a weighting factor without considering data length can lead to unstable inversion, and yield velocity structure that is even worse than those inverted from data without segment weighting. These preconditions provide us flexibility to regulate the linearized receiver function inversion. For an instance, the segments weight may introduce an Earth-stripping approach to determine the crustal and upper mantle structure from the top to the bottom. We first assign more weight ($c \geq 0.5$) to the first segment containing mainly the Ps converted waves to better constrain the shallow structure in an inversion. Once the shallow structure is determined such as by reviewing VR, we then shift our weight to the later segment comprised of reverberations of the shallow structure and the Ps conversions of the deep boundaries, fix the crustal layers, and continue the inversion to characterize the deep structure.

Besides the VR, the rms of the difference between the observed and predicted receiver function data or the residual of the least-square solution can also serve as the indicator of data misfit. However, a model yielding the maximum VR or minimum misfit may not necessarily be the best physical model. In practice, we may set up a critical value that defines whether a model is acceptable or not, and used the average of all the acceptable models as the preferred model. We also intend to select models with minimum number of free parameters as our final model. The simplicity of these models makes them relatively easy to be interpreted, and thus can be widely used in other studies.

The extensive tests with the synthetic and field data have demonstrated the effectiveness of our integrated approach to stabilize model convergence in receiver function inversion. Although our regularizations for the linearized inversion may not fundamentally solve the intrinsic problem of receiver function data regarding the trade-off between velocity and depth under certain circumstance as shown in the RPB case 5.1 and 5.2, it can be further integrated into other techniques (e.g. Jacobsen & Svenningsen 2008) to reduce the non-uniqueness and non-linearity of receiver function inversion. The non-linear methods may suffer less dependence on initial models and be more suitable for high level of automation (e.g. Bodin *et al.* 2012). However, the regularizations on models and data proposed here should also be important issues for those inversions.

6 CONCLUSION

In this study, we intend to improve the iterative linearized inversion of receiver function data in two aspects, from both the data processing and inversion procedure itself. We first develop a 'fourpin' method to simultaneously align the Moho Ps converted phase and two reverberated phases by stretching/contracting and rescaling each receiver function. The stretching/contracting and rescaling are based on the ray-parameter of the incoming *P* wave, and can be pinned to any standard receiver function travelling with a certain ray-parameter. The stack of the receiver functions after being corrected on both arrival time and amplitude with respect to the standard receiver function, referred as the RPB stack, is better than the direct stack (DRT) at least in two aspects: (1) it preserves the waveform of the three later arrivals, which is the accurate response of the underneath structure of the receiver and (2) it allows us to construct multiple stacks of receiver functions at any teleseismic distances. Our synthetic tests indicate that inversion with multiple RPB stacks can enhance model convergence, leading to speedy recovery of the target model.

In the inversion side, the integration of all preconditions discussed in this study yields an algorithm that is able to provide better stability in the iterative linearized inversion procedure. Applying this inversion algorithm to RPB data has demonstrated the capability in recovering the target model from our extensive synthetic tests with various types of starting models. We apply the integrated method to two broad-band seismic stations in China, XAN and MDJ. Our results show that the crust–mantle transition beneath XAN located at the Qinling orogenic belt occurs at a depth range not as broad as previously reported. On the other hand, the mantle LVZ beneath MDJ in northeast China suggested by previous studies is likely an artificial structure associated with the improperly constructed receiver function data.

ACKNOWLEDGEMENTS

This research was supported by the AFRL under contract number FA9453-10-C-0258. The project was also supported by the NSF grant EAR-063566. We are grateful to H. Liu, K. Tao and S. Grand for discussion and to the two anonymous reviewers for their constructive comments and suggestions. We also thank the IRIS DMC for providing the seismic data.

REFERENCES

Agostinetti, N.P., Lucente, F.P., Selvaggi, G. & Di Bona, M., 2002. Crustal structure and Moho geometry beneath the northern Appennines (Italy), *Geophys. Res. Lett.,* **29**(20), doi:10.1029/2002GL015109.

Ammon, C.J., 1991. The isolation of receiver effects from teleseismic P waveforms, *Bull. seism. Soc. Am.,* **81,** 2504–2510.

Ammon, C.J., Randall, G.E. & Zandt, G., 1990. On the non-uniqueness of receiver function inversions, *J. geophys. Res.,* **95,** 15 303–15 318.

Ammon, C.J. & Zandt, G., 1993. Receiver structure beneath the southern Mojave block, California, *Bull. seism. Soc. Am.,* **83,** 737–755.

Beckers, J., Schwartz, S.Y. & Lay, T., 1994. The velocity structure of the crust and upper mantle under China from broadband P and PP waveform analysis, *Geophys. J. Int.,* **119,** 574–594.

Birch, F., 1961. The velocity of compressional waves in rocks to 10 kilobars, part 2, *J. geophys. Res.,* **66,** 2199–2224.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K. & Rawlinson, N., 2012. Transdimensional inversion of receiver function and surface wave dispersion, *J. geophys. Res.,* **117,** B02301, doi:10.1029/2011JB008560.

Chang, S.-J., Baag, C.-E. & Langston, C.A., 2004. Joint analysis of teleseismic receiver functions and surface wave dispersion using the genetic algorithm, *Bull. seism. Soc. Am.,* **94,** 691–704.

Chang, S.-J. & Baag, C.-E., 2005. Crustal structure in southern Korea from joint analysis of teleseismic receiver functions and surface-wave dispersion, *Bull. seism. Soc. Am.,* **95,** 1516–1534.

Chen, Y., Niu, F., Liu, R., Huang, Z., Tkalčić, H., Sun, L. & Chan, W., 2010. Crustal structure beneath China from receiver function analysis, *J. geophys. Res.,* **115,** B03307, doi:10.1029/2009JB006386.

Chen, Y., Liu, R., Huang, Z. & Sun, L., 2011. Preliminary study of lateral variation in crustal structure of Northeast China from teleseismic receiver functions, *Earthq. Sci.,* **24,** 15–25.

Chevrot, S. & van der Hilst, R.D., 2000. The Poisson ratio of the Australian crust: geological and geophysical implications, *Earth planet. Sci. Lett.,* **183,** 121–132.

Clarke, T.J. & Silver, P.G., 1993. Estimation of crustal Poisson's ratio from band teleseismic data, *Geophys. Res. Lett.,* **20,** 241–244.

Du, Z.J. & Foulger, G.R., 1999. The crustal structure beneath the northwest fjords, Iceland, from receiver functions and surface waves, *Geophys. J. Int.,* **139,** 419–432.

Feng, R., Zhu, J. & Ding, Y., 1981. A study on the crustal structure of China with surface waves (in Chinese), *Acta Seismol. Sin.,* **3**(4), 335–350.

Herrmann, R.B., Ammon, C.J. & Julià, J., 2000. Joint inversion of receiver functions and surface-wave dispersion for crustal structure, in *Proceedings of the 22nd Annual DoD/DoE Seismic Research Symposium,* September 13–15, 2000, New Orleans, Louisiana.

Jacobsen, B.H. & Svenningsen, L., 2008. Enhanced uniqueness and linearity of receiver function inversion, *Bull. seism. Soc. Am.,* **98,** 1756–1767.

Julià, J., Vila, J. & Macià, R., 1998. The receiver structure beneath the Ebro basin, Iberian peninsula, *Bull. seism. Soc. Am.,* **88,** 1538–1547.

Julià, J., Ammon, C.J., Herrmann, R.B. & Correig, A.M., 2000. Joint inversion of receiver function and surface wave dispersion observations, *Geophys. J. Int.,* **143,** 1–19.

Julià, J., Ammon, C.J. & Herrmann, R.B., 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., Engdahl, E.R. & Buland, R., 1995. Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.,* **122,** 108–124.

Koper, K.D., Wysession, M.E. & Wiens, D.A., 1999. Multimodal function optimization with a niching genetic algorithm, *Bull. seism. Soc. Am.,* **89,** 978–988.

Langston, C.A., 1979. Structure under Mount Rainier, Washington, inferred from teleseismic body waves, *J. geophys. Res.,* **84,** 4749–4762.

Lawrence, J.F. & Wiens, D.A., 2004. Combined receiver function and surface wave phase-velocity inversion using a niching genetic algorithm: application to Patagonia, *Bull. seism. Soc. Am.,* **94,** 977–987.

Lawrence, J.F., Wiens, D.A., Nyblade, A.A., Anandakrishnan, S., Shore, P.J. & Voigt, D., 2006. Crust and upper mantle structure of the Transantarctic Mountains and surrounding region from receiver functions, surface waves, and gravity: implications for uplift models, *Geochem. Geophys. Geosyst.,* **7,** Q10011, doi:10.1029/2006GC001282.

Ligorría, J.P. & Ammon, C.J., 1999. Iterative deconvolution and receiver function estimation, *Bull. seism. Soc. Am.,* **89,** 1395–1400.

Liu, H. & Niu, F., 2011. Estimating crustal seismic anisotropy with a joint analysis of radial and transverse receiver function data, *Geophys. J. Int.,* **188,** 144–164.

Ma, Y., Niu, F., Zhou, H. & Chen, Y., 2007. Joint inversion of receiver functions and Rayleigh wave dispersion for crustal and upper mantle structures: a comparison of a nonlinear genetic approach and a linear inversion, *EOS, Trans. Am. geophys. Un.,* **88,** 52, Fall Meet. Suppl., Abstract S42A-02.

Mangino, S., Priestley, K. & Ebel, J., 1999. The receiver structure beneath the China digital seismograph network stations, *Bull. seism. Soc. Am.,* **89,** 1053–1076.

Nair, S.K., Gao, S.S., Liu, K.H. & Silver, P.G., 2006. Southern African crustal evolution and composition: constraints from receiver function studies, *J. geophys. Res.,* **111,** B02304, doi:10.1029/2005JB003802.

Niu, F. & James, D.E., 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., Bravo, T., Pavlis, G., Vernon, F., Rendon, H., Bezada, M. & Levander, A., 2007. Receiver function study of the crustal structure of the southeastern Caribbean plate boundary and Venezuela, *J. geophys. Res.,* **112,** doi:10.1029/2006JB004802.

Özalaybey, S., Savage, M.K., Sheehan, A.F., Louie, J.N. & Brune, J.N., 1997. Shear-wave velocity structure in the northern basin and range province from the combined analysis of receiver functions and surface waves, *Bull. seism. Soc. Am.,* **87,** 183–189.

Pan, S. & Niu, F., 2011. Large contrasts in crustal structure and composition between the Ordos plateau and the NE Tibetan plateau from receiver function analysis, *Earth planet. Sci. Lett.,* **303,** 291–298.

Pei, S., Xu, Z. & Wang, S., 2004. Sn wave tomography in the uppermost mantle beneath the China continent and adjacent regions, *Chin. J. Geophys.,* **47,** 250–256.

Pei, S. *et al.*, 2007. Upper mantle seismic velocities and anisotropy in China determined through Pn and Sn 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.

Sambridge, M., 1999a. Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space, *Geophys. J. Int.,* **138,** 479–494.

Sambridge, M., 1999b. Geophysical inversion with a neighbourhood algorithm—II. Apprasing the ensemble, *Geophys. J. Int.,* **138,** 727–746.

Sandvol, E., Seber, D., Calvert, A. & Barazangi, M., 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.

Searcy, C.K., Christensen, D.H. & Zandt, G., 1996. Velocity structure beneath college station Alaska from receiver functions, *Bull. seism. Soc. Am.,* **86,** 232–241.

Shibutani, T., Sambridge, M. & Kennett, B., 1996. Genetic algorithm inversion for receiver functions with application to crust and uppermost mantle structure beneath eastern Australia, *Geophys. Res. Lett.,* **23,** 1829–1832.

Snieder, R. & Trampert, J., 1999. Inverse problems in geophysics, in *Wavefield Inversion,* pp. 119–190, ed. Wirgin, A., Springer Verlag, New York.

Snieder, R. & Trampert, J., 2000. Linear and nonlinear inverse problems, in *Geomatic Methods for the Analysis of Data in the Earth Sciences,* pp. 93–164, eds Dermanis, A., Grun, A. & Sanso, F., Springer, Berlin.

Sun, Y., Li, X., Kuleli, S., Morgan, F.D. & Toksöz, M.N., 2004. Adaptive moving window method for 3-D P-velocity tomography and its application in China, *Bull. seism. Soc. Am.,* **94,** 740–746.

Sun, Y. & Toksöz, M.N., 2006. Crustal structure of China and surrounding regions from P wave traveltime tomography, *J. geophys. Res.,* **111,** B03310, doi:10.1029/2005JB003962.

Sun, Y., Niu, F., Liu, H., Chen, Y. & Liu, L., 2012. Crustal structure and deformation of the SE Tibetan plateau revealed by receiver function data, *Earth planet. Sci. Lett.,* **349,** 186–197.

Tkalčić, H., Pasyanos, M., Rodgers, A., Gök, R., Walter, W. & Al-Amri, A., 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. & Banerjed, D., 2009. Interactive receiver function forward modeler (IRFFM), software and manual available at: http://www.rses.anu.edu.au/∼hrvoje/IRFFMv1.1.html (last accessed on May, 15, 2013).

Tkalčić, H., Chen, Y., Liu, R., Huang, Z., Sun, L. & Chan, W., 2011. Multi-step modeling of teleseismic receiver functions combined with constraints from seismic tomography: crustal structure beneath southeast China, *Geophys. J. Int.,* **187,** 303–326.

Vergne, J., Wittlinger, G., Hui, Q., Tapponnier, P., Poupinet, G., Mei, J., Herquel, G. & Paul, A., 2002. Seismic evidence for stepwise thickening of the crust across the NE Tibetan plateau, *Earth planet. Sci. Lett.,* **203,** 25–33.

Vinnik, L.P., Reigber, C., Aleshin, I.M., Kosarev, G. L, Kaban, M.K., Oreshin, S.I. & Roecher, S.W., 2004. Receiver function tomography of the central Tien Shan, *Earth planet. Sci. Lett.,* **225,** 131–146.

Yoo, H.J., Herrmann, R.B., Cho, K.H. & Lee, K., 2007. Imaging the three-dimensional crust of the Korean peninsula by joint inversion of surface-wave dispersion, *Bull. seism. Soc. Am.,* **97,** 1002–1011.

Yuan, X., Wang, S., Li, L. & Zhu, J., 1986. A geophysical investigation of the deep structure in China, in *Reflection Seismology: A Global Perspective, Geodynamic Series,* Vol. 13, pp. 151–160, AGU, Washington, DC.

Zandt, G. & Ammon, C.J., 1995. Continental crust composition constrained by measurements of crustal Poisson's ratio, *Nature,* **374,** 152–154.

Zandt, G., Myers, S.C. & Wallace, T.C., 1995. Crust and mantle structure across the Basin and Range—Colorado plateau boundary at 37N latitude and implications for Cenozoic extensional mechanism, *J. geophys. Res.,* **100,** 10 529–10 548.

Zhao, L., Sen, M.K., Stoffa, P. & Frolich, C., 1996. Application of very fast simulated annealing to the determination of the crustal structure beneath Tibet, *Geophys. J. Int.,* **125,** 355–370.

Zhu, L. & Kanamori, H., 2000. Moho depth variation in southern California from teleseismic receiver functions, *J. geophys. Res.,* **105,** 2969–2980.