JOINT INVERSION OF CRUSTAL AND UPPERMOST MANTLE STRUCTURE IN WESTERN CHINA

Xiaodong Song¹, Zhen Xu¹, Lupei Zhu², and Yuming Zhou²

University of Illinois at Urbana-Champaign¹ and Saint Louis University²

Sponsored by the Air Force Research Laboratory

Award No. FA9453-10-C-0216 Proposal No. BAA10-55

ABSTRACT

The objective of this research is to develop joint-inversion methods involving P travel times (including secondary P), receiver functions, and surface wave dispersion measurements from ambient noise correlation and to apply them to the western China region to obtain 3D models of P and S structures of the crust and upper mantle.

This is the second year of the three-year project. Our joint inversion methods (Song et al., MRR, 2010) involve obtaining crustal P model and Pn station delays, which are then used with surface wave dispersions and receiver functions at each station to invert jointly for S structure, the Moho depth, and Vp/Vs ratios in the crust. Our main accomplishments include the following three areas. (1) We successfully implemented a search-based algorithm (neighborhood algorithm) for joint inversion of surface wave dispersion data, receiver function, and Pn delay time. The implementation uses parallel programming with MPI calls, making it possible for massive data processing. (2) We tested joint inversion methods to HiClimb data, which improves resolution of crust and the Moho and shows features of mid-crustal low velocity and Moho variations. (3) We made a number of improvements on existing methods or lines of works. We constructed new dispersion maps using a combination of both ambient noise data. We increased the total number of dispersion measurements by more than four folds. We implemented parallel algorithm for speeding up computation of correlation Green functions by using multi-processors at the same time. We are also systematically processing receiver functions in western China using a semi-automatic procedure.

OBJECTIVES

The objective of this research is to develop joint-inversion methods involving P travel times, receiver functions, and surface-wave dispersion measurements from ambient noise correlation and to apply them to the western China region to obtain 3D models of P and S structures of the crust and upper mantle.

RESEARCH ACCOMPLISHED

1) Development of joint inversion method of receiver functions, surface wave dispersions, and Pn station delay for S velocities and Moho

A major effort of this project is to include P data in the joint inversion of surface wave dispersion and receiver functions as described previously (Song et al., 2010). The combination of receiver functions and surface-wave dispersions bridges some resolution gaps associated with either dataset (e.g., Julia et al., 2000; Ozalaybey et al., 1997). However, important ambiguity remains. In the P receiver function, the Moho P-to-S converted phase (Ps) is generally the strongest. The Moho depth determined from the Ps delay (relative to P) trades off strongly with crustal Vp/Vs ratio. Adding well-constrained P information will help the joint inversion of receiver function and surface-wave dispersion to find the right S model when the Vp/Vs ratio is not known (Song et al., 2010). One type of P information is the Pn station delay time, which provides a constraint on Moho depth and average crustal velocity.

We have focused on an implementation of the joint inversion method that includes surface wave dispersion data, receiver functions, and Pn delay time that are associated with the same station of interest. The implementation includes the following key elements. (1) We use fast global search method, neighborhood algorithm (NA) (Sambridge, 1999a, 1999b). Using a global parameter search approach, compared with more traditional linearized iterative inversion approach, has two key advantages: (a) We have more control on parameterization, for example, to include Pn data, we need to define a Moho; and (b) we are able to explore parameter space more fully and to quantity the resolution and uncertainty of the solution more easily. (2) Our parameterization includes the Moho, the depth of which is allowed to vary. It can include variable crustal layers as well. Regions between layers are parameterized as smooth functions (linear or splines) to reduce the number of search parameters. (3) The implementation uses parallel programming with MPI calls, which greatly speeds up the model search with multi-processor or parallel machines. Figure 1 shows a synthetic example using the global search method for joint inversion of surface wave and receiver functions and Pn delay time. The method finds a solution with perfect fit to the input model and data.

2) Data Preparation

We use the HiClimb data to test our joint inversion implementation. This is a large volume of data (some 180 stations), we need a effective procedure to obtain each data set for the joint inversion. **First**, we obtained high resolution dispersion curves along the array using a linear array inversion procedure. Because, it's a near-linear dense array, we can reduce the 3D structure inversion problem to 2D problem along the array, thus allowing us to refine the grid size (to 0.1 degree). We implemented a version of ambient noise surface wave tomography that is suitable for a dense linear array. **Second**, we constructed receiver functions for all the stations using the semi-automatic procedure described. The interactive procedure allowed us to control data quality but also to increase efficiency that is essential to process a large volume of data. For each receiver function, we also extract the delay time of the Moho-converted phase Ps relative to the reference phase P from the H-K stacking procedure (Zhu and Kanamori, 2000). Although the crustal thickness from the receiver function trades off strongly with Vp/Vs ratio, the Ps delay time is robust as long as the converted phase is strong enough. The delay time provides an efficient constraint on the model search by ruling out models with significant difference in the delay time. **Third**, our joint inversion method includes Pn station delay times. This effort is underway separately. In the current runs, we have not included Pn delays.

3) An Implementation for Real Data

Below we describe practical considerations in the implementation for real data (with HiClimb data as an example). Particular effort was made in formulating effective procedures and in model parameterization. Although a searchbased approach provides a great flexibility in fitting different types of data, it is generally very expensive computationally. The selection of parameterization requires a particular care because of the need to limit the number of parameters.

Derivation of starting model. Our joint inversion is done in two steps. In the first step, we invert only the dispersion data to obtain a starting S velocity profile at each station, which is then used as the starting model for the model searches. The main purpose of the starting model is to reduce the model space to help the search algorithm converge faster. We use cubic splines for model parameterization. The advantage of using splines is that it yields naturally smooth model and yet reduce the number of parameters for searching.

For dispersion inversion in the first step, we adopt 12 uniformly distributed spline nodes from the surface down to 150 km. We seek solution that satisfies the following minimizer:

$$\min\left(||\mathbf{D}^{obs} - \mathbf{D}^{pred}|| + \lambda ||\mathbf{m}|| + \varphi ||\mathbf{Lm}||\right)$$
(1)

where the first term is data misfit, D^{obs} is the data (dispersion curve) we are trying to fit, and D^{pred} is the predictions from the random model generated by the NA algorithm. The second term is the model constraint. Since the dispersion periods of the HiClimb data from the ambient noise extends only up to 45 s, which has a limited constraint on the deeper part of the structure, we introduce model constraint for the depth from 95 km down to the 150 km using the S model of China and surrounding region, which includes group velocity period up to 120 s. The third term is the smoothing constraint. L is the Laplacian smoothing operator (Lees and Crosson, 1989), which requires second derivative of the model to be zero. λ and φ are two real positive numbers that have been chosen so that different terms are well balanced. We seek the solution to minimize the L1 norm of above formula instead of traditional L2 norm. The choice of L1 norm reduces the risk that the model is biased by extreme values due to error in the data. We convert spline model into layered velocity model to calculate predicted Rayleigh group and phase velocity. We divide the whole 150 km into 30 layers with 5 km thickness for each layer. Velocity within each layer is constant. We fix the Vp/Vs ratio for all the layers to 1.75. Density is calculated by Nafe-Drake (1963) relationship. The forward calculation subroutine is adopted from Robert Herrmann's program (1991).

Joint inversion. The parameterization for the second step in the joint inversion is different from the first step. We use two sets of cubic splines to represent velocity model: one for the crust (first set) and one for the Moho and the mantle (second set). The last node of the first spline has the same value as the first node of the second spline so that the model is continuous throughout the depth. The node spacing for each set of spline is not uniform; instead, it is stretched over depth, with closer spacing between nodes at the top of the spline and larger spacing at the bottom of the spline. We use the following formula to map the normalized spline space (0 to 1) to the depth:

$$\mathbf{h} = \mathbf{D}\mathbf{x}^2 + \mathbf{h}_0 \tag{2},$$

where h is the depth from the top of the first node in the spline; D is the total thickness; h_0 is the depth of the first node; x is the coordinate in the normalized spline space between 0 and 1. This set up is effective in modeling the structure at shallow depth and near Moho discontinuity region by giving more freedom to the model where it is better constraint by short period dispersion data and receiver function, respectively. It also yields smoother model at the lower crust and deeper mantle, which are less constraint by the data. We use 12 nodes for the first spline in the crust and 10 nodes for the second spline in the Moho and the mantle. We also allow the depth of the connecting node between the first and the second splines to vary so that the Moho depth can be adjusted. The use of the Moho and mantle spline provides more flexibility compared with a Moho with imposed discontinuity or a linear gradient. For a Moho depth at 70 km, the node spacing between the first and second nodes in the mantle is ~0.99 km and the spacing between the second and third nodes is ~2.96 km, enough to model sharp velocity jump, thus mimic the Moho discontinuity. On the other hand, this parameterization also works well in the situation where Moho is not sharp but presents as a transition zone. We convert spline model into layered model as in the first step. We use 2 km layer thickness throughout the depth range of the first spline in the crust. In the mantle, we use fifteen 2 km layers

2011 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies

for the top 30 km as transition zone and followed by 5 km layer to the bottom at 150 km. We seek solution that satisfies the L1 norm of the following minimizer:

$$\min(w_1 \left\| \mathbf{D}_{disp}^{obs} - \mathbf{D}_{disp}^{pred} \right\| + w_2 \left\| \mathbf{D}_{rfs}^{obs} - \mathbf{D}_{rfs}^{pred} \right\| + \lambda \|\mathbf{m} - \mathbf{m}_0\| + \phi \|\mathbf{Lm}\|$$
(3),

where the first term is the dispersion misfit, the second term is the RFs misfit and the third and last terms are the same as formula (1). The w_1 and w_2 are the relative weighting for dispersion and RFs misfits, respectively.

We impose model-screening algorithms to speed up model convergence. Every time when NA generates a random model, we compare the model with starting model from dispersion inversion for each layer except those 15 layers in the transition zone. If the velocity difference in any given layer is larger than 0.5 km/s, this model is discarded by setting a large misfit value without going through the forward calculation. We also require the velocity to increase monotonically from the last two nodes in the crust to first two nodes in the mantle. This model screening process can reduce computational time by confining the model search in the model space where the likelihood of the model is the largest.

Another screening parameter is the measured Ps delay time. Because the converted Ps phase may be mis-identified in the H-K stacking process, we need to sort out the "outliers" of the Ps delay times before applying the screening. We fit the measured Ps delay times of all the stations with a smooth spline curve. If the measurement for a station is differed from the prediction of the spline fit by more than 1.5 s, then we use the spline prediction as the Ps delay time for that station, otherwise we use the original Ps delay measurement. For each model generated in the search, we calculate the velocity gradient at each layer to the next layer and pick the one with largest gradient as the "Moho". In most cases, it is the last layer in the first spline to the first layer in the mantle spline. We calculate the predicted Ps delay time using the "Moho" defined above and compare it with the measured Ps delay. If the difference is larger than 1 s, then the model is discarded (assigned a large misfit value). Our experience suggests that by introducing the Ps delay time, the model converges faster than without the Ps delay screening.

As we haven't included Pn delay time data, we fix the Vp/Vs ratio to 1.75 throughout the depth range. Density of each layer is calculated using the same Nafe-Drake relationship as described above. Because the lack of coherent crust multiple signals, only first 15 s of receiver functions are included in inversion. For each station, we select 4 receiver functions with different ray parameters for inversion. The move out of Ps signal at different ray parameters is helpful in determining the Moho depth. Subroutines of dispersion and receiver function forward calculation are adopted from Robert Herrmann's computer programs. We iterate 800 times to obtain the final model and 100 new models are generated each iteration in 5 Voronoi cells.

4) Application to HiCLIMB Data: Preliminary Results

We show the joint inversion results and fits to the data for a few stations in the HiClimb array (Figure 2). The stations are from south to north, showing different velocity structure, Moho depth and sharpness as reflected in different characteristics of the dispersion and receiver function data. There is no low velocity zone under station H0330 (near the Main Boundary Thrust) in the south. Moho is very clear with a velocity jump at ~50 km. Station H1220 is located at southern Lhasa block. The shear velocity profile shows a pronounced low velocity structure in mid-crust (near 25 km). Moho depth reaches ~75 km with a clear velocity jump. Velocity structures between 31.5° N and 33° N are complicated and seem to show no consistency between different stations. H1320 exhibits a weaker but still distinguishable mid crust low velocity zone and a weak Moho discontinuity, where the Ps converted phase is weak and inconsistent.

In summary, our joint inversion method appears to be able to capture the main features of a variety of different structures and data behaviors.

5) Computational Improvement

Because of tremendous of amount of continuous data involved, we spent considerable effort in speeding up the calculations of empirical Green functions. Because the task for each pair is quite independent, it can be parallelized efficiently. The main obstacle is lots of I/O operations, which require expensive reading and writing hard drives. We

2011 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies

have come up a solution that would speed up the computation by distributing tasks to multi-processors within the same computer (node) and execute them simultaneously. The method uses the LINUX POSIX message queue (mq) capabilities. Below is a test result for the main program calculating cross-correlations of station pairs in a desktop LINUX machine with dual-core and total of 8 processors. The speed up is almost linear as the number of processor increases and achieves more than 5 times reduction in computing time when all 8 processors are used.

number_	_of_processes	user_run_time(s)	speed_up
1	3045	1	
2	1533	1.97	
4	821	3.71	
8	535	5.69	

We have developed a semi-automatic procedure to generate receiver functions. Because of large number of stations and large amount of data, such a procedure is essential for productivity. The procedure includes interactive review and selection of good quality raw data using a graphic interface, picking and alignment of P waveforms using waveform cross-correlation and a graphic interface, automatic calculations of receiver functions and stackings according to ray parameters and azimuths, and interactive selections of good quality receiver functions using a graphic interface. Figure 3 shows the results of the receiver functions for a total of 182 stations of the HiClimb array.

CONCLUSIONS AND RECOMMENDATIONS

Our effort for this period has focused on two areas. One is the development of joint inversion methods. We successfully implemented a search-based algorithm (neighborhood algorithm) for joint inversion of surface wave dispersion data, receiver function, and Pn delay time. The implementation uses parallel programming with MPI calls, making it possible for massive data processing. We tested joint inversion methods on HiClimb data, which appeared effective in capturing the main features of a variety of different structures and data behaviors. In the second area, we made a number of improvements on existing methods or lines of works, including speeding up computations of Green functions and processing of receiver functions. Future effort will focus on improving the joint inversion methods and continuing systematic processing of massive data of the project.

REFERENCES

- Herrmann, R. B. (1991). *Computer Programs in Seismology, Version* 3.30. St. Louis University: Department of Earth and Atmosphere Sciences.
- Julia, J., C. J. Ammon, R. B. Herrmann, A. M. Correig (2000). Joint inversion of receiver functions and surfacewave dispersion observations, *Geophys. J. Int.* 143: 99–112.
- Lees, J. M., and R. S. Crosson (1989). Tomographic inversion for threedimensional velocity structure at Mount St. Helens using earthquake data, *J. Geophys. Res.* 94: 5716–5728.
- Nafe, J. E., and Drake, C. L. (1963). Physical properties of marine sediments. In Hill, M. N. (Ed.) *The Sea* (Vol. 3): New York (Interscience): 794–815.
- Ozalaybey, S., M. K. Savage, A. F. Sheehan, J. N. Louie, 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. Seism. Soc. Am.*, 87: 183–189.
- 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. Appraising the ensemble. *Geophys. J. Int.* 138: 727–746.
- Song, X. D., Z. J. Xu, L. P. Zhu, Y. M. Zhou (2010). Joint inversion of crustal and uppermost mantle structure in western China, in *Proceedings of the 2010 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, LA-UR-10-05578, Vol. 1, pp. 224–231.
- Zhu, L. P. and H. Kanamori (2000). Moho depth variation in southern California from teleseismic receiver functions, J. Geophys. Res. 105: 2969–2980.



Figure 1. Synthetic test of joint inversion including receiver function, surface wave dispersion and Pn delay time data using Neighborhood Algorithm. Input velocity model is shown in blue curve in (A). Input model consists of three layers. Moho discontinuity locates at 30 km depth. The Vp/Vs ratios in the crust layers are 1.75 and 1.85, respectively. Two sets of splines are used in inversion. Each set consists of 5 evenly spaced nodes. Output velocity model is plotted in red in (A). Note the velocity jump in the crust is smoother than input model due to the natural smoothness of spline function. Average Vp/Vs ratio of 1.80 in the crust is obtained during searches when Pn delay time data is included. The input (solid line) and output (dashed line) of receiver function waveforms are shown in (B). Only the first 10 second of receiver function is used in inversion. Rayleigh wave group and phase dispersion data (blue) from 5 s to 45 s are used in inversion (C). The corresponding dispersion predictions from output model are shown in red.

(**A**)

(B) Station H0330



Figure 2. Results of joint inversion of surface wave dispersion and receiver function data for Hi-CLIMB stations (A). Examples show stations H0330 (B), H1220 (C), and H1320 (D) from south to north of the profile, respectively. Observed receiver functions are in black solid curves and predicted waveforms are in dashed lines. Observed and predicted dispersion curves are in blue and red colors, respectively.

(A)



Figure 3. Results of receiver functions of HiClimb stations. (A) Stacked P receiver functions. From left to right and top to bottom are stations from the south to the north. Some prominent characters are obvious, e.g., Ps conversions changes from about 6 s in the south (top, left) to nearly 10 s (center, bottom) and to around 9s (right, bottom). However, Ps phases are very complicated at other locations (left, middle-bottom; center, top) or sometimes hard to observe (right, top). (B) Common-conversion point (CCP) stacking of receiver functions. The image shows clear Moho depth variation and complications along the along.