AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE

10.1002/2014JB011161

Key Points:
- We developed a new method for delay time tomography of triplicated *P* waves
- The method improved resolution in the vicinity of mantle discontinuities
- The results provide support for hot material beneath the Changbai Volcano

Correspondence to:
N. Takeuchi,
takeuchi@eri.u-tokyo.ac.jp

Citation:
Takeuchi, N., et al. (2014), Upper mantle tomography in the Northwestern Pacific region using triplicated *P* waves, *J. Geophys. Res. Solid Earth*, *119*, 7667–7685, doi:10.1002/2014JB011161.

Received 1 APR 2014
Accepted 11 SEP 2014
Accepted article online 16 SEP 2014
Published online 6 OCT 2014

Upper mantle tomography in the northwestern Pacific region using triplicated *P* waves

Nozomu Takeuchi[1], Hitoshi Kawakatsu[1], Satoru Tanaka[2], Masayuki Obayashi[2], Y. John Chen[3], Jieyuan Ning[3], Steve P. Grand[4], Fenglin Niu[5], James Ni[6], Rryohei Iritani[1,7], Koki Idehara[1,8], and Takashi Tonegawa[1,9]

[1]Earthquake Research Institute, Tokyo University, Bunkyo, Japan, [2]Department of Deep Earth Structure and Dynamics Research, Japan Agency for Marine-Earth Science and Technology, Yokosuka, Japan, [3]School of Earth and Space Sciences, Peking University, Beijing, China, [4]Department of Geological Sciences, University of Texas at Austin, Austin, Texas, USA, [5]Department of Earth Science, William Marsh Rice University, Houston, Texas, USA, [6]Department of Physics, New Mexico State University, Las Cruces, New Mexico, USA, [7]Now at Schlumberger K.K., Sagamihara, Japan, [8]Now at Planning, Review and Research Institute for Social Insurance and Medical Program, Chiyoda, Japan, [9]Now at Research and Development Center for Earthquake and Tsunami, Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

Abstract We conducted delay time tomography of the upper mantle beneath the northwestern Pacific using *P* data from NorthEast China Extended SeiSmic Array, F-net, and nearby available stations. To improve resolution and accuracy in the vicinity of mantle discontinuities, we extracted traveltimes of both initial and secondary phases observed at triplication distances by using a waveform fitting technique. Compared with the model obtained by using only the initial phase, the resolution just above the 410 km discontinuity is especially improved, and low-velocity anomalies beneath the Changbai Volcano are clearly observed down to the 410 km discontinuity. Compared with previous models, low-velocity anomalies atop the 410 are more pronounced. The results of this study together with the previous receiver function analysis provide further support that we have hot material beneath the Changbai Volcano.

1. Introduction

Structures associated with subduction zones have been extensively studied by seismic tomography. Many high-resolution models have been obtained using body waves [e.g., *Fukao et al.*, 2001; *Gorbatov and Kennett*, 2003; *Li and van der Hilst*, 2010; *Zhao*, 2012; *Wei et al.*, 2012] and surface waves [e.g., *Yoshizawa et al.*, 2010; *Zheng et al.*, 2011] for the northwestern Pacific region as well as other subduction zones. The deeper structures (such as those in the transition zone) have been primarily constrained by delay time tomography of body wave traveltimes, and the resultant models have contributed considerably to our understanding of subduction dynamics [e.g., *Fukao et al.*, 2001, 2009]. However, the existence of triplicated phases (caused by mantle discontinuities) appears to be one of the most significant obstacles to further progress [e.g., *Tajima and Grand*, 1995]. Even if multiple phases are expected to simultaneously arrive at a single station, in conventional tomography, usually only one traveltime is used without deliberate attempts at phase associations. This approach degrades the resolution and the accuracy of the obtained models, especially in the vicinity of mantle discontinuities.

To overcome these problems, several efforts have been made. For shorter-period analysis, numerous studies [e.g., *Grand and Helmberger*, 1984a, 1984b; *Tajima and Grand*, 1995, 1998; *Tajima et al.*, 2009] have obtained structure models by trial and error through comparison of detailed features of observed and synthetic seismograms for triplicated phases. *Gao et al.* [2006] conducted waveform inversion of triplicated *P* and *S* waveforms after careful data processing. These approaches are useful for obtaining insight into the structural features in the studied region, but the systematic and quantitative retrieval of structural information from these approaches appears to be limited owing to the relatively large cost for analysis and the assumption of lateral homogeneity for each subregion. For longer-period analysis, *Fuji et al.* [2010] conducted systematic waveform inversion using a large amount of triplicated *SH* waveform data. *Stähler et al.* [2012] computed finite frequency kernels for frequency-dependent traveltimes of composite waveforms of triplicated *P* that are applicable to the inversion of traveltimes measured by using a cross-correlation

©2014. American Geophysical Union. All Rights Reserved.

technique. However, these approaches suffer from limited spatial resolution. Efforts to extend methods for longer-period analysis to shorter-period analysis are in progress [e.g., *Fuji et al.*, 2012].

In this study, we achieve systematic, quantitative, and higher-resolution structural inversion by extending conventional delay time tomography. We measure onset times to both initial and secondary phases of triplicated *P*. In most conventional methods, we pick onset times of various phases. To resolve difficulties for handpicking emergent onsets, we sometimes measure arrival times using the cross-correlation technique [e.g., *Woodward and Masters*, 1991]. In such conventional methods, problems emerge when multiple phases overlap each other. *Garcia et al.* [2004] and *Iritani et al.* [2010] have developed a method for measuring differential traveltimes of overlapping core phases observed by seismograph arrays. Their method is based on waveform fitting and is applicable to overlapping phases. Because our primary interest is to study upper mantle structure, we extend their method to the triplication phases caused by upper mantle discontinuities. We slightly modify their method to enable the measurements of absolute traveltimes rather than differential traveltimes.

The applicability of these methods depends on the availability of array data. From 2009 to 2011, we deployed the broadband seismic array NECESSArray (NorthEast China Extended SeiSmic Array) in Northeast China [e.g., *Grand et al.*, 2006]. In the Japanese Islands, we use the dense broadband array F-net [*Okada et al.*, 2004]. We extracted traveltimes of triplicated *P* in these array data together with other available data (data from nearby Incorporated Research Institutions for Seismology (IRIS) GSN stations). We then conducted delay time tomography to determine the three-dimensional *P* structures of upper mantle beneath the northwestern Pacific. We discuss the correlations between the distributions of the off-arc volcanoes in Northeast China and heterogeneous structures beneath them.

2. Method for Measuring Traveltime of Triplicated Phases

If we have array data and focus on one particular phase, we expect that all stations should record essentially identical waveforms after corrections for attenuation. We also expect that different phases (e.g., PKPab, PKPbc, and PKPdf) should have more or less identical waveforms after corrections for attenuation and $\pi/2$ phase shifts for retrograde branches [*Choy and Richards*, 1975]. Using these assumptions, *Garcia et al.* [2004] and *Iritani et al.* [2010] synthesized seismograms by introducing several parameters (such as arrival times, amplitudes, and a reference waveform). They then determined optimal parameters to minimize the misfit between the observed and synthetic seismograms. To avoid problems associated with local minima, they used simulated annealing (SA) in the optimization. They applied the above method to measure differential arrival times and differential attenuation parameters of core phases. In this study, we slightly modify their formulation to enable measurements of absolute traveltimes of triplicated *P*.

2.1. Parameterization

We focus on array waveform data for one event. We synthesize the seismogram for the *i*th station, $u^{(i)}(t)$, as follows:

$$u^{(i)}(t) = \sum_{j=1}^{3} w_j^{(i)} A_j^{(i)}(t) f(t - t_j^{(i)}), \tag{1}$$

where $f(t)$ is the reference waveform, A_j and t_j are the amplitude and the arrival time of the *j*th phase, respectively, and $w_j^{(i)}$ is the weighting factor, which is either 1 or 0. Because of existence of the 410 and 660 km discontinuities, multiple phases simultaneously arrive at a single station (Figure 1). In this study, we take the following three phases into account: the AB ($j = 1$), CD ($j = 2$), and EF ($j = 3$) branches in the traveltime curves shown in Figure 1a. They correspond to direct *P* waves bottoming in the upper mantle, the transition zone, and the lower mantle (Figure 1b). B and D are the points where *P* bottoms just above the 410 km and the 660 km depth, respectively, and C and E are the points where *P* bottoms just below the 410 km and the 660 km depth, respectively.

In the formulation of equation (1), we introduced several approximations. Following the method of *Garcia et al.* [2004] and *Iritani et al.* [2010], we introduce the following approximations: (1) we ignore the branches turning at the discontinuities (in this study, BC and DE in Figure 1); (2) we ignore the finite frequency effects (such as the phase shifts and the existence of diffracted waves) in the vicinity of the cusps (in this study, B, C, D, and E in Figure 1); and (3) we ignore the differential apparent velocity between the direct and the depth

©2014. American Geophysical Union. All Rights Reserved.

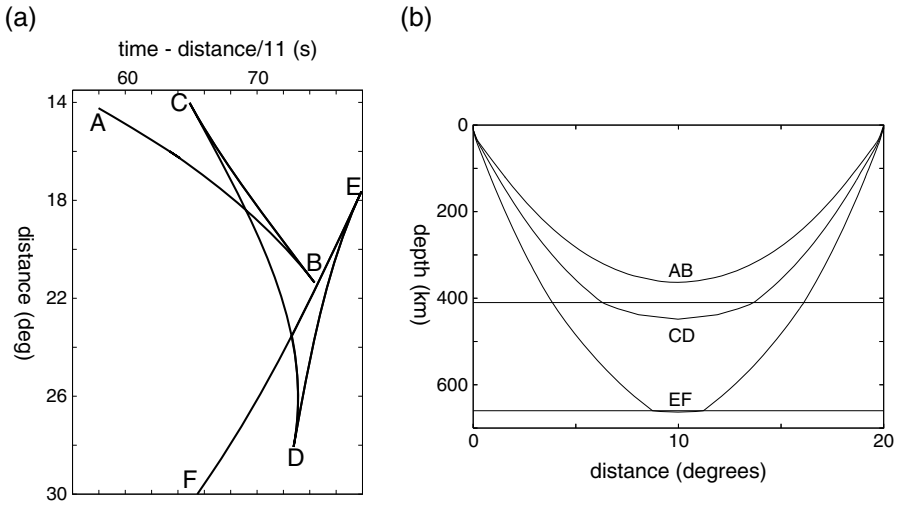

Figure 1. (a) Typical *P* traveltime curve for shallow events. Shown are traveltime curves for an event at 10 km depth and for the AK135 structural model [*Kennett et al.*, 1995]. (b) Raypaths of the AB, CD, and EF branches at 20° for an event at 10 km depth.

phases for shallower events. Approximation (3) means that the contributions of depth phases to the AB, CD, and EF branches are not explicitly included by j but implicitly included in $f(t)$ in equation (1). The reference waveform $f(t)$ is therefore the overlapping waveforms of direct and depth phases. The reliability of these approximations has been discussed in *Iritani et al.* [2010].

In this study, we further introduced the following approximations: (4) we ignore the effects of differential t^* (attenuation) and (5) we ignore the contribution of phases whose waveforms are not clearly confirmed by visual inspection. The former approximation is justified by considering that the synthetic differential t^* computed for the preliminary reference Earth model (PREM) [*Dziewonski and Anderson*, 1981] is 0.02 at most. The latter approximation means that we set the weighting factor $w_j^{(i)}$ in equation (1) to 0 when the corresponding phase cannot be confirmed in the record section. In our experience, the traveltime measurements are not stable without this approximation.

2.2. Optimization

Following the method of *Garcia et al.* [2004] and *Iritani et al.* [2010], we determine the optimal parameters $A_j^{(i)}$, $t_j^{(i)}$, and $f(t)$ in equation (1) to minimize the L1 norm of the residuals:

$$\sum_i \int \left| d^{(i)}(t) - u^{(i)}(t) \right| \mathrm{d}t, \tag{2}$$

where $d^{(i)}(t)$ is the observed waveform for the ith station. The optimization scheme is the same as that used in *Iritani et al.* [2010] except for the points described below.

To stabilize the problem, we introduced the following two additional constraints. We first checked the record section, and, for each station (i.e., for each waveform trace), we a priori specify the sequence of the phases appearing in the trace. The identification is based on visual inspection. The imposed constraint on $t_j^{(i)}$ is

$$t_{j_1}^{(i)} < t_{j_2}^{(i)}, \tag{3}$$

if we initially identify the j_1th phase to be earlier than the j_2th phase.

The other constraint is for minimizing the trade-off between the arrival times $t_j^{(i)}$ and the reference waveform $f(t)$. Equation (1) shows that we have a trade-off between origin time of $f(t)$ (or shift in t) and a systematic shift in $t_j^{(i)}$. The optimal $t_j^{(i)}$ can therefore have common offset (for all i and j) compared with actual arrival times, and this is why previous studies [*Garcia et al.*, 2004; *Iritani et al.*, 2010] could discuss only differential traveltimes. For removing this offset to enable measurements of absolute traveltimes, we use traveltime data of 1 Hz because we want to obtain velocity models at a reference period of 1 Hz. We first handpick

©2014. American Geophysical Union. All Rights Reserved.

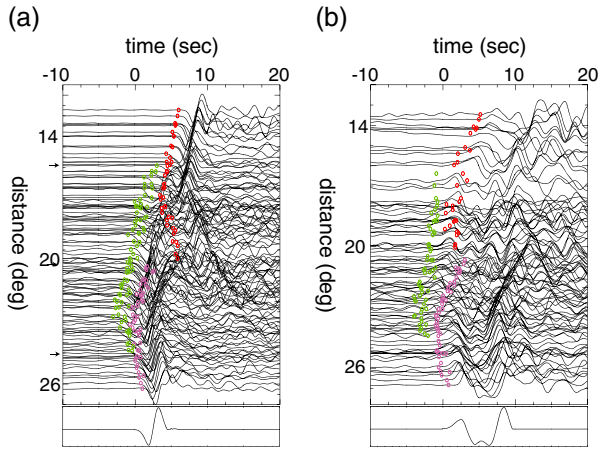

Figure 2. (a) Example of the observed waveforms and measured traveltimes (top) together with the obtained reference waveform $f(t)$ (bottom). The waveforms are for the event at 21:24:38 on 10 October 2009 in the Kuril Islands region ($M_w = 5.9$; 112 km depth) to which a low-pass filter with corner frequency of 0.5 Hz has been applied. The time is relative to the first arrival time predicted by the AK135 model [*Kennett et al.*, 1995]. Red, green, and purple circles denote the measured arrival times for AB, CD, and EF branches, respectively. The measured arrival times are plotted only when the estimated errors are within the threshold. (b) The same as Figure 2a except that the waveforms are for the event at 01:02:16 on 18 January 2010 in the Kuril Islands region ($M_w = 5.6$; 7 km depth).

the onset times at 1 Hz as much as possible and then determine the optimal traveltimes by referring to the handpicked times. The explicit form of the constraint is written as follows:

$$\sum_i \sum_j r_j^{(i)} \left(t_j^{(i)} - s_j^{(i)} \right) = 0, \qquad (4)$$

where $s_j^{(i)}$ is the handpicked onset time for the ith station and the jth phase, and $r_j^{(i)}$ is a parameter set to 1 when we could pick the onset visually and set to 0 when we could not. Equation (4) means that the average of the offset $t_j^{(i)} - s_j^{(i)}$ should be zero, where average is computed for all available handpicked arrival times. Note that $r_j^{(i)}$ in equation (4) is the parameter for visibility of onsets, while $w_j^{(i)}$ in equation (1) is the parameter for visibility of wave packets. For a pair of i and j with $w_j^{(i)} = 0$, $r_j^{(i)}$ is always 0, but for a pair with $w_j^{(i)} = 1$, $r_j^{(i)}$ can be either of 0 and 1.

In this study, the onset time $s_j^{(i)}$ is picked from the seismograms to which a low-pass filter with a corner frequency of 1 Hz has been applied. By using $s_j^{(i)}$ as a reference, the obtained $t_j^{(i)}$ can be assumed to be the absolute arrival times for the reference frequency of 1 Hz. In other words, our optimization consists of two steps: (i) conventional handpick and (ii) correction by using SA, and the second step is conducted with guaranteeing no systematic shifts. By using such an $s_j^{(i)}$, the obtained $f(t)$ in equation (1) can be assumed to be the incoming wave to the array whose onset time is $t = 0$. Such a function can be used, for example, for an initial model of the input waveform for future waveform inversion for structural parameters.

Another option to remove the offset is evaluating the offset from the onset time of the $f(t)$ obtained by using the method in the previous studies [*Garcia et al.*, 2004; *Iritani et al.*, 2010]. Using this approach, we should analyze waveforms of 1 Hz to obtain travel times of 1 Hz. In some problems (including analyses in this study), waveforms of 1 Hz are too complex to conduct waveform analyses. In such cases, both of this approach and our approach require use of longer-period waveforms, which causes systematic bias. However, our approach is expected to suffer from smaller amount of bias for most of applications because longer-period data are used only for the correction stage. Quantitative comparison of bias is shown in the next subsection.

In the actual optimization, we assume $f(t) = 0$ outside the time window specified a priori. We set $f(t) = 0$ for $t < 0$ and $t > t_w$, where t_w is the appropriate length of the time window. Because we impose the condition of equation (4), we can assume that the onset of $f(t)$ is at $t = 0$, and the assumption of $f(t) = 0$ for $t < 0$ will not have significant impact on obtained solutions. In contrast, the choice of t_w may significantly affect solutions. We therefore choose appropriate t_w by trial and error. Note that t_w is different for each event. The explicit method for making the choice is presented in the next subsection.

The initial values of each parameter and the range of the parameter space are also identical to those in *Iritani et al.* [2010] except for the following points. In this study, we set the initial value of $t_j^{(i)}$ as $s_j^{(i)}$ where we can pick the onset (i.e., when $r_j^{(i)} = 1$). When we cannot (i.e., when $r_j^{(i)} = 0$), we estimate unknown onset times from other handpicked onset times and use the estimated $s_j^{(i)}$ as the initial value. The procedure for the estimation is as follows: For each event and phase type ($j = 1, j = 2$ or $j = 3$) we determine the optimal polynomial functions, which best explains the handpicked onset times, as a function of distance; we then estimate $s_j^{(i)}$ at the distance of this station using the optimal polynomial functions for the phase j. We search

©2014. American Geophysical Union. All Rights Reserved.

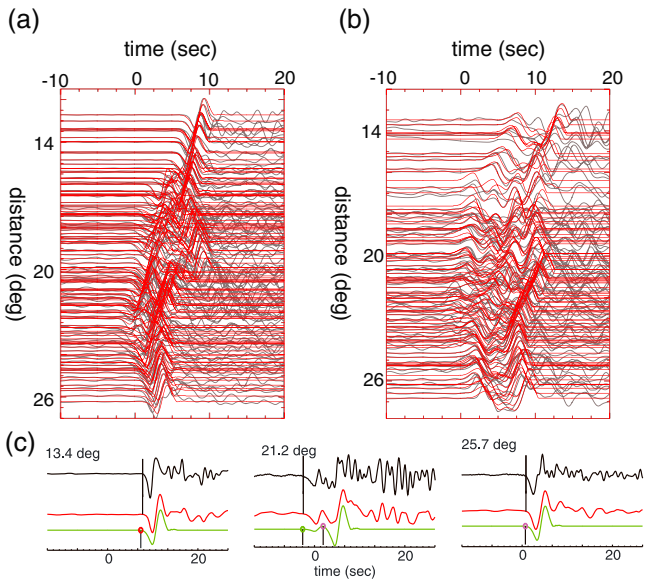

Figure 3. (a) Comparison of observed and synthetic waveforms (shown in light gray and red, respectively) for the event of Figure 2a. (b) The same as Figure 2a except that the waveforms are for the event of Figure 2b. (c) Comparison for representative three stations in the record section of Figure 3a. In each panel, the number on the top denote the distance, the black and red traces are observed seismograms to which a low-pass filter with the corner frequency of 1 and 0.5 Hz, respectively, is applied, and the green trace is a synthetic seismogram for the optimal parameters to explain the red trace. The vertical line on the top two traces denote the onset time handpicked from the black trace. The vertical lines and colored circles on the bottom trace denote the optimal arrival times, where the same colors as those in Figure 2 are used to identify phase type.

the optimal values of $t_j^{(i)}$ between $\left|t_j^{(i)} - s_j^{(i)}\right| < 0.75$ and $\left|t_j^{(i)} - s_j^{(i)}\right| < 3.0$ for the cases of $r_j^{(i)} = 1$ and $r_j^{(i)} = 0$, respectively. The wider range of the parameter space in the latter case is based on the fact that we can have large errors in $s_j^{(i)}$ because of uncertainties in the extrapolations.

2.3. Example

Examples of the observed waveforms, the measured traveltimes, and the obtained reference waveform are shown in Figure 2. Figure 2a is for a deep event (112 km depth) for which depth phases are not in the record section, while Figure 2b is for a shallow event (7 km depth) for which both direct and depth phases are in the section. In general, we invert only direct *P* phases when direct and depth phases are separated, while we invert both direct and depth phases when they overlap with each other. We also show comparison between observed and synthetic seismograms in Figure 3.

For the actual waveform fitting to measure the arrival times, a causal filter with a corner frequency of 0.5 Hz is applied to the observed waveforms. Note that $d^{(i)}(t)$ in equation (2) is a seismogram to which this filter is applied. As we discussed in the previous subsection, the obtained optimal onset time, $t_j^{(i)}$ in equation (1), is not that for 1 Hz in a strict sense; however, it is assumed to be essentially identical to it because of the following reason. First of all, differential arrival times between 1 and 0.5 Hz are not very large (the maximum is about 0.17 s for the PREM attenuation model of *Dziewonski and Anderson* [1981], while observed traveltime residuals shown later in Figure 7 range between −9.2 and 8.1 s with an root-mean-square amplitude of 2.1 s). If we do not use our two-step optimization, these differential times are the expected amount of bias. However, because our approach uses the constraints of equation (4), most of the bias is expected to be corrected. Note that differential times are completely corrected by equation (4) if they are identical among all stations and phases (i.e., if t^* are all identical). The errors are only caused by differential t^* within the array. The largest bias will never exceed 0.1 s for the array size used in this study.

Table 1. Parameters Used in Traveltime Measurements for the Waveform Data in Figure 2[a]

Distance	w_1	w_2	w_3	Sequence	r_1	r_2	r_3
Event at 21:24:38 10/10/2009							
$\Delta \le 15.4$	1	0	0	-	1	0	0
$15.4 \le \Delta \le 20.1$	1	1	0	$t_2 < t_1$	0	1	0
$20.1 \le \Delta \le 24.6$	0	1	1	$t_2 < t_3$	0	1	0
$24.6 \le \Delta$	0	0	1	-	0	0	1
Event at 01:02:16 01/18/2010							
$\Delta \le 16.0$	1	0	0	-	1	0	0
$16.0 \le \Delta \le 20.5$	1	1	0	$t_2 < t_1$	0	1	0
$20.5 \le \Delta \le 24.5$	0	1	1	$t_2 < t_3$	0	1	0
$24.5 \le \Delta$	0	0	1	-	0	0	1

[a]See the text for the definitions for each variable.

The parameters used in traveltime measurements are summarized in Table 1. Explicit procedures to determine these parameters are shown below using the event in Figure 2a as an example. By visual inspection, we assume that the AB, CD, and EF branches are observed only before 20.1°, between 15.4° and 24.6°, and beyond 24.6°, respectively.

©2014. American Geophysical Union. All Rights Reserved.

The parameters $w_j^{(i)}$ in equation (1) are uniquely determined from these observations. Secondary CD branches should be before 20.1° and beyond 24.6°, but they are not clearly confirmed in the record section in Figure 2a. BC and DE branches are not clearly confirmed either. We ignore contributions of these later phases. Because the largest phase shift by phase overlap is $\pi/2$ (0.5 s in this study) even for cases when two phases have comparable amplitudes, neglection of minor later phases will not alter the main results of this paper. By visual inspection, we also assume that the initial phase types are AB, CD, and EF for stations before 15.4°, between 15.4° and 24.6°, and beyond 24.6°, respectively. We therefore impose the corresponding constraints of the sequence of equation (3). For the constraint of equation (4), we could handpick the onset only for the initial phase. The parameters $r_j^{(i)}$ are thus set to be 1 only for the initial phase and 0 for the others.

Using the above parameters, we obtained $t_j^{(i)}$, $A_j^{(i)}$, and $f(t)$ to minimize the norm of equation (2). To conduct the optimization for minimizing equation (2), we have to a priori specify the length of the time series $f(t)$, t_w, defined in the previous subsection. We first inspect the record section and roughly estimate the length of common waveforms for every phase and station. For example, for the case of Figure 2a, we can see that the length is between 5 and 8 s from the duration time of wave packets with larger amplitudes in the smallest and largest distance ranges. (Although the choice of the common part is based on coherence and amplitudes, it is inevitably ad hoc as it was so in the previous studies by *Garcia et al.* [2004] and *Iritani et al.* [2010].) After several trial and error iterations, we found that the optimal length is $t_w = 6.0$ s, because $f(t)$ becomes almost zero at the end of the time window (i.e., at $t = t_w$; see the bottom traces in Figures 2a and 2b). Such an $f(t)$ can be assumed to be a plausible input waveform to the array. The obtained $f(t)$ is a waveform of the direct phase for the event of Figure 2a, but it is a composite waveform of the direct and depth phases for the event of Figure 2b.

In the optimization, we search for optimal values by using 16 random seeds: For each event, we conduct 16 optimizations using different sets of random numbers, and, for each event, station and phase, we obtain 16 different optimal onset times $t_j^{(i)}$. We compute their average and standard deviation to evaluate the measured onset time and its error. When the standard deviation is greater than 0.15 s, we discard it because the measurement is considered to be unstable. Because the observed data $d^{(i)}(t)$ in equation (2) are velocity seismograms of 0.5 Hz, we expect time shifts of more than 2 sec if we have a cycle skip problem. The threshold of 0.15 s roughly corresponds to the criterion that we have such a cycle skip problem for 1, at most, of 16 trials. Even if the standard deviation is small, there are several optimal values that obviously have cycle skip problems, but such data are also discarded. The cycle skip problems can be easily identified by visual inspection because we a priori identify where all of the focused phases are in the record section. In other words, our traveltime measurements are based on visual inspection with the aid of the waveform fitting procedures to obtain explicit numbers of the onset time.

The obtained $t_j^{(i)}$ are shown by red, green, and purple circles in Figure 2. We can see that the obtained arrival times trend continuously without any discontinuities at the phase cross-over distances (e.g., at $\Delta = 15.4°$ and $\Delta = 24.6°$ for the case of Figure 2a), which provides confirmation of the accuracy of our phase associations specified by the parameters in Table 1. If we have implausible discontinuities in the trends, we revise the parameters until continuous trends are obtained.

The comparison between observed and synthetic seismograms (Figures 3a and 3b) shows that overall features of waveforms are well simulated both for initial and later phases, which shows overall accuracy of our traveltime measurements. However, close inspection of waveforms (light gray and red lines in Figures 3a and 3b) shows some detailed features (e.g., sharp onsets at larger distances for the event of Figure 3a and complex waveforms at closer distances for the event of Figure 3b) are not well modeled, which is probably due to distortion of waveforms by complex structures in subduction zones. Such distortion can be possible source of traveltime measurement errors which are drawbacks of systematic and semi-automatic measurements.

Comparison for individual seismograms (Figure 3c) shows that, when we observe an isolated wave packet (for 13.4 and 25.7°) our inversion is comparable to traveltime measurements using conventional cross-correlation techniques. However, for cases when multiple phases overlap with each other (for 21.2°), our method should improve time picking. Because, in these seismograms, the secondary phases have much larger amplitudes than the initial phases, a traveltime for the maximum cross-correlation can be biased by or may be essentially equivalent to a traveltime of the secondary phase. Our inversion therefore has merits not

©2014. American Geophysical Union. All Rights Reserved.

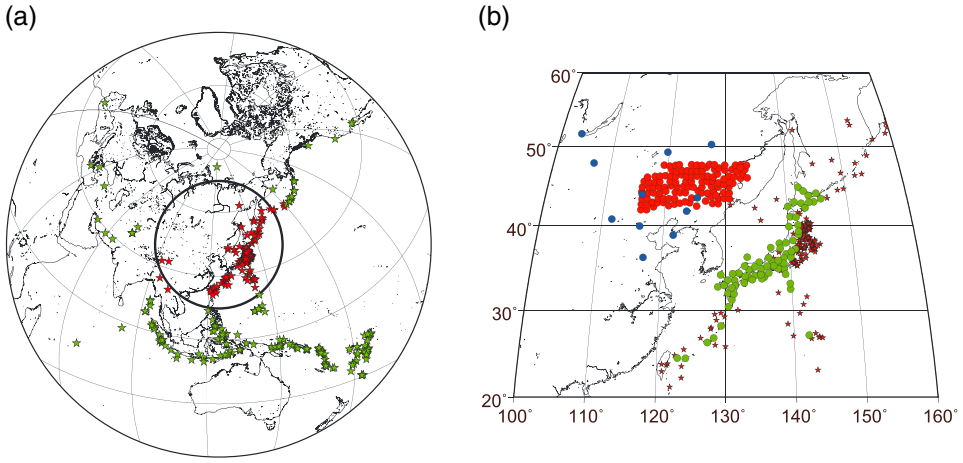

Figure 4. (a) Distribution of events used in this study. Red stars denote events for which traveltimes are measured by the waveform fitting described in section 2. Green stars denote events for which traveltimes are measured by handpicking the onset time. Thick black lines denote the 30° from the centroid of the NECESSArray stations. (b) Distribution of stations used in this study. Red, green, and blue circles denote NECESSArray, F-net and other stations, respectively. The nearby event distribution is also shown by small red stars for reference.

only in its ability to measure the secondary phases but also in its accuracy for both traveltime measurements and phase association.

3. Observed Traveltime Data

3.1. Data Set Categories and Traveltime Measurement Methods

The distribution of the events used in this study is shown in Figure 4a. To categorize the data set, we classify the events as (i) events at triplication distances (Figure 4a, red stars) and (ii) events at teleseismic distances (Figure 4a, green stars).

The distribution of the stations used in this study is shown in Figure 4b. To categorize the data set, we classify the stations as (i) stations of NECESSArray (Figure 4b, red circles), (ii) stations of F-net (Figure 4b, green circles), and (iii) other stations from IRIS GSN or CDSN (Figure 4b, blue circles).

According to the event and station category, we measured traveltimes by using different methods, as summarized in Table 2. For the events at triplication distances, we measured traveltimes by waveform fitting described in the previous section. Optimization is separately conducted for each station category. We therefore obtain different reference functions, $f(t)$, for F-net and NECESSArray for the same event. This is because directivity effects are considered to be azimuthal dependent. The F-net stations sometimes have wide azimuthal variations because some events are inside or very close to the network. For such cases, we divide F-net into several subnetworks and invert for different reference functions, $f(t)$, for each subnetwork. For the IRIS GSN or CDSN stations, we do not invert for reference functions, $f(t)$, but fix it to that obtained for the stations of NECESSArray. This is because the station density for this category of the dataset is insufficient to independently determine $f(t)$. Because the azimuthal coverage of this data set is similar to that of NECESSArray (see the event distribution in Figure 4b), the reference function can be considered to be more or less similar. For teleseismic data, we measured arrival times by handpicking the onset because the waveforms are not contaminated by phase overlap.

Table 2. Methods for Traveltime Measurements for Each Data Set Category[a]

	NECESSArray	F-net	Others
Triplication	waveform fitting	waveform fitting	waveform fitting*
Teleseismic	handpicking	handpicking	-

[a]For the data set denoted by *, optimization is conducted in a simplified way. See the text for the details.

©2014. American Geophysical Union. All Rights Reserved.

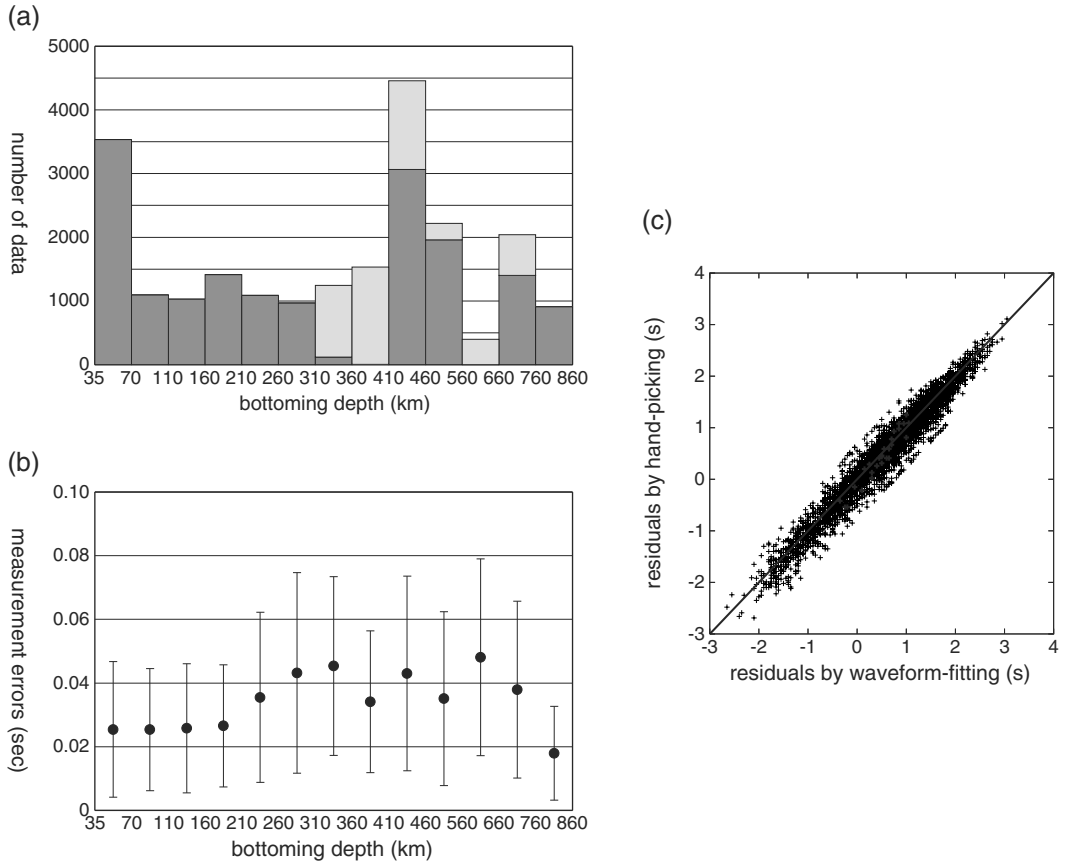

Figure 5. (a) The number of measured traveltimes in our triplication data set as a function of bottoming depth. Dark and light gray denote data for initial and secondary phases, respectively. (b) Average and standard deviation of measurement errors of traveltimes in our triplication data set as a function of bottoming depth. (c) Scatterplot between residuals measured by using waveform fitting (horizontal axis) and hand picking (vertical axis).

3.2. Observed Traveltimes

We present distributions of observed traveltime residuals together with the path coverage of our data set. Our data set consists of data for events at triplication distances (about 22,000 measurements) and data for events at teleseismic distances (about 19,000 measurements). In this section, we primarily focus on the former data set to see the improvements made by the use of the secondary phases.

Figure 5a shows the distribution of bottoming depths of our triplication data set. The secondary phase data comprise only 24% of the data set; however, they significantly fill the gaps where initial phases do not sample. In the data set of the initial phases, we have significant gaps of bottoming depths just above the 410 and the 660. By using the secondary phases, the gap above the 410 is well filled. The gap above the 660 is also filled, but the number of observations is not large because secondary CD branches (see Figure 1 for the phase name convention) were, in general, hard to confirm in record sections. We therefore expect that the main improvement in the resolution should be observed just above the 410 km discontinuity.

The top figure in Figure 6 shows path coverages of the initial and the secondary phases bottoming between depths of 310 and 410 km. We see that the secondary phases densely sample the regions beneath the Japan Sea, the Korean Peninsula, and the Yellow Sea, whereas the initial phases sparsely sample there. Primary improvements of the resolution are therefore expected in these regions.

Figure 5b shows average and standard deviation of measurement errors of traveltimes in our triplication data set as a function of bottoming depth. The measurement errors are estimated from standard deviations of optimal traveltimes as presented in section 2.2. Traveltimes for *P* waves bottoming between 210 and 760 km depths have larger errors compared with *P* waves bottoming in other depth ranges. The initial and the secondary phases have comparable errors in this depth range (0.036 $\pm$ 0.028 s for the initial phase

©2014. American Geophysical Union. All Rights Reserved.

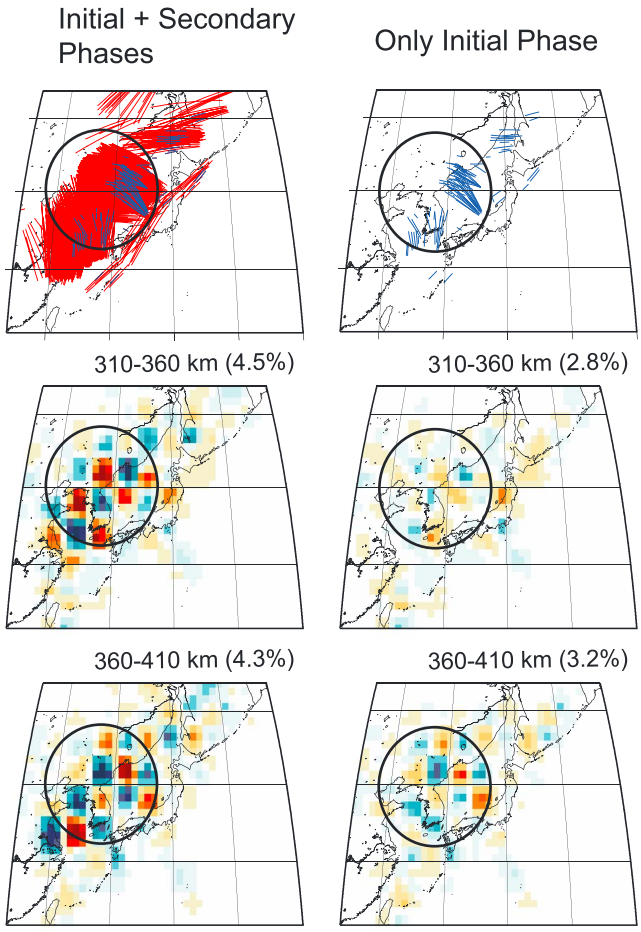

Figure 6. Comparison of path coverage of *P* waves bottoming (top) between 310 and 410 km depth and comparison of the recovered checkerboard patterns at the depth ranges of (middle) 310–360 km and (bottom) 360–410 km for the case when we use both the initial and secondary phases (left) and when we use only the initial phase (right). In Figures 6 (top left) and Figures 6 (top right), portions where rays pierce the region between 310 and 410 km depths are shown by blue and red lines for the initial and secondary *P*, respectively. In all figures, the regions with notable differences are shown by thick black circles.

and 0.042 ± 0.028 s for the secondary phase). The results suggest that complexities due to phase overlap cause degradation of traveltime measurements both for the initial and secondary phases. However, because we applied a low threshold (0.15 s was the maximum of acceptable errors), such degradation will not seriously affect the accuracy of the tomography models.

Figure 7 shows the distribution of the observed residuals projected at the bottoming point. Cap averages over 2° are shown. Only the points with more than three observations are plotted. The residuals are relative to the prediction of the AK135 model [*Kennett et al.*, 1995]. In the upper mantle (above 410 km depth), the region above the subducting slabs, in general, shows large positive anomalies, which suggests strong low-velocity anomalies in the wedge regions. In the transition zone, we can see abrupt decreases of the residuals from 410–510 to 510–660 km depth, which suggests strong high-velocity anomalies in the lower half of the transition zone. In the top of the lower mantle (660–760 km depths), residuals are, in general, small, which suggests weak heterogeneities there. These features are retrieved in the velocity model presented below (in Figure 9).

To measure traveltime residuals of the teleseismic data, we used handpicking rather than the method presented in this paper. This is because the primary purpose of this study is to improve the resolution in the vicinity of mantle discontinuities. To test measurement accuracies, we compare the residuals which are measured by using these two methods (Figure 5c). We compared 4528 residuals for all 40 teleseismic events that occurred in 2009. Standard deviation of differences is 0.21 s which is comparable to the acceptable errors for triplication data (0.15 s), and, more importantly, we do not confirm any systematic bias between two methods.

4. Velocity Models

4.1. Tomography Method

The initial velocity model is the AK135 model [*Kennett et al.*, 1995] with slight modification. The modifications are made to avoid complexity of the traveltime curves caused by the second-order discontinuity at 120 km depth in AK135. We applied smoothing to the structures between the Moho and 210 km depth. The discrepancy between the traveltimes predicted by the modified model and the original AK135 is 0.2 s at most, which will not alter the conclusions of this study. The initial source parameters are those in the PDE.

©2014. American Geophysical Union. All Rights Reserved.

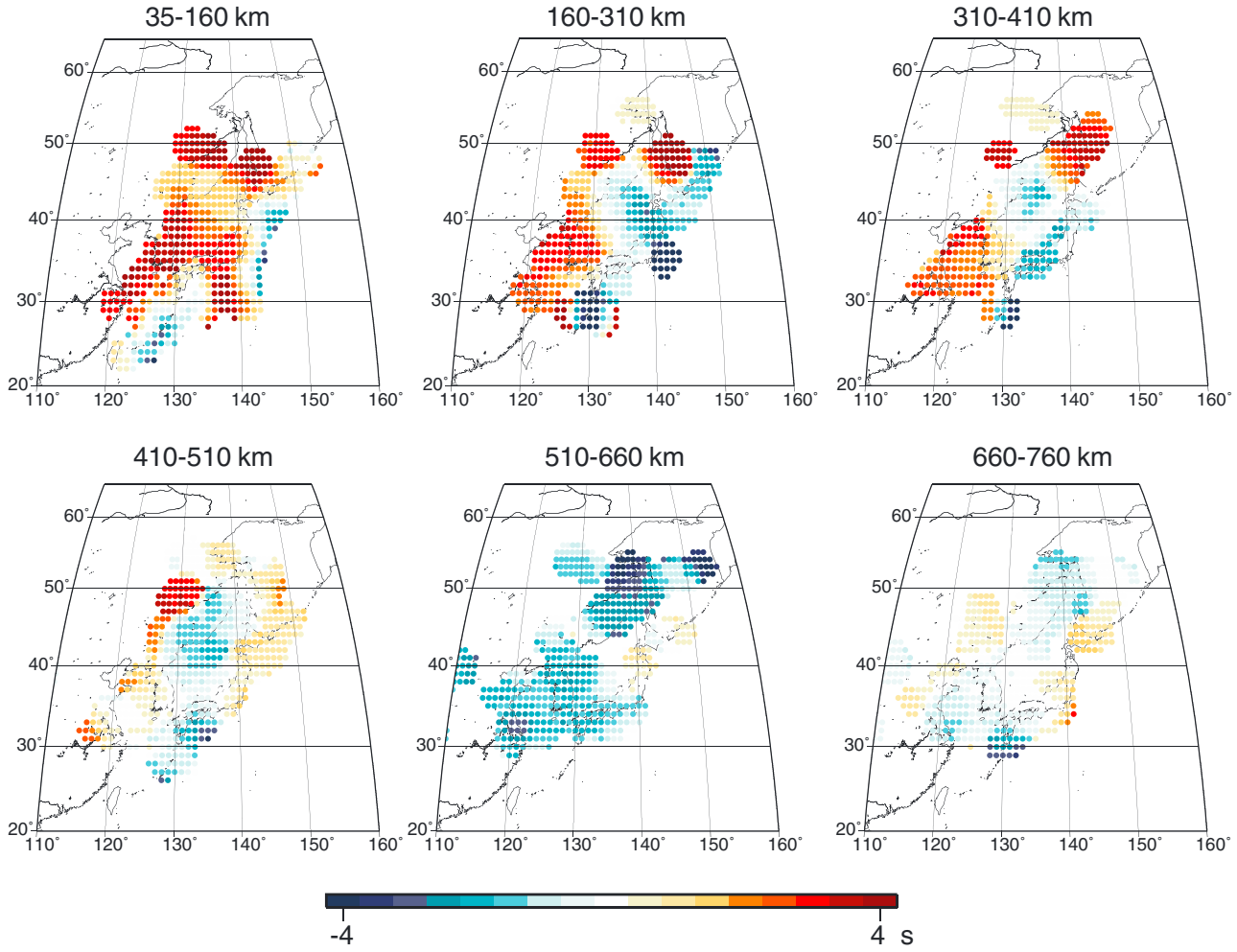

Figure 7. Observed traveltime residuals projected at the bottoming points. Cap-averaged residuals over 2° in radius are shown for each depth range. Residuals are relative to the traveltimes predicted by the AK135 model [*Kennett et al.*, 1995].

The slowness perturbations with respect to the initial velocity model are expanded in terms of blocks. For regions where we expect higher resolution (the region 96°E–162°E, 18°N–60°N and 0–860 km depth), we use finer blocks with a horizontal dimension of 2° × 2° and a vertical thickness of 35–100 km. Outside these regions, we use coarse blocks with a horizontal dimension of 6° × 6° and a vertical thickness of 400–429 km. In other words, we conduct global tomography but use finer grids only in regions where we expect good resolution. Perturbations with respect to the initial source parameters are allowed only for the origin time, because of difficulties in tightly constraining all the source parameters since the events are mostly outside of the station networks used in this study.

The inversion is based on conventional damped least squares. We iteratively update the velocity and the source parameters until we obtain satisfactory convergence of the model. We fixed the ray to that for the initial model and did not take perturbations due to model alternation into account. The effects of this approximation are discussed in section 5. Variances of traveltime residual for the initial and final models are 1.55^2 s^2 and 0.49^2 s^2, respectively, and variance reduction is 90%. If we modify only source parameters and fix structural parameters to the initial model, variance reduction was only 23%. Variance reduction is comparable for each data type: 89% and 93% are achieved for triplication data and teleseismic data, respectively, and 91%, 87%, and 91% are achieved for AB, CD, and EF branches in Figure 1a, respectively.

©2014. American Geophysical Union. All Rights Reserved.

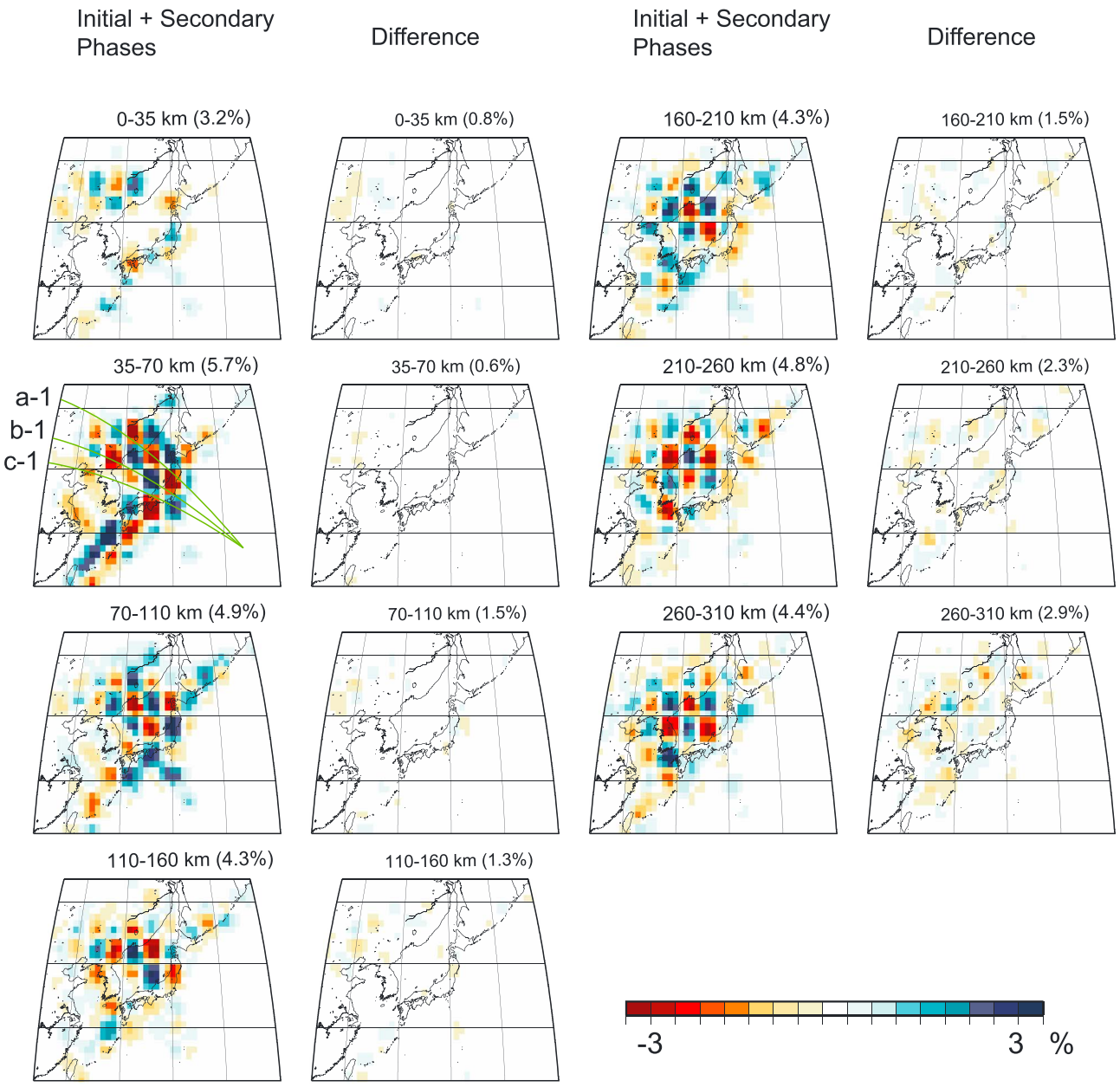

Figure 8. (left) Recovered checkerboard patterns when we use both initial and secondary phases and (right) differential patterns to the case when we use only the initial phase. The numbers in the upper right corner of each map denote the maximum absolute amplitudes of the perturbations for that depth range. Green lines denote the locations of the vertical cross sections shown in Figure 10. The regions where we see notable differences are shown by thick black circles.

4.2. Resolution

We first compare the resolution of our model with that of the model obtained by using only initial phases. The recovered checkerboard patterns of slowness perturbations of ±5% are compared in Figures 6 and 8. Note that the checkerboard patterns exist in both the horizontal and vertical directions: i.e., the sign of the perturbations of one block is always opposite to that of the next blocks. The observed patterns therefore show the resolutions for the minimum spatial scales to be expressed by our model parameters.

As shown in Figure 8, for the shallower region above 260 km depth, we do not see any significant differences in the recovered patterns. However, for the depth range between 310 and 410 km, we see clear improvements in the recovered pattern (in the regions shown by thick black circles in Figure 6). This is due to the

©2014. American Geophysical Union. All Rights Reserved.

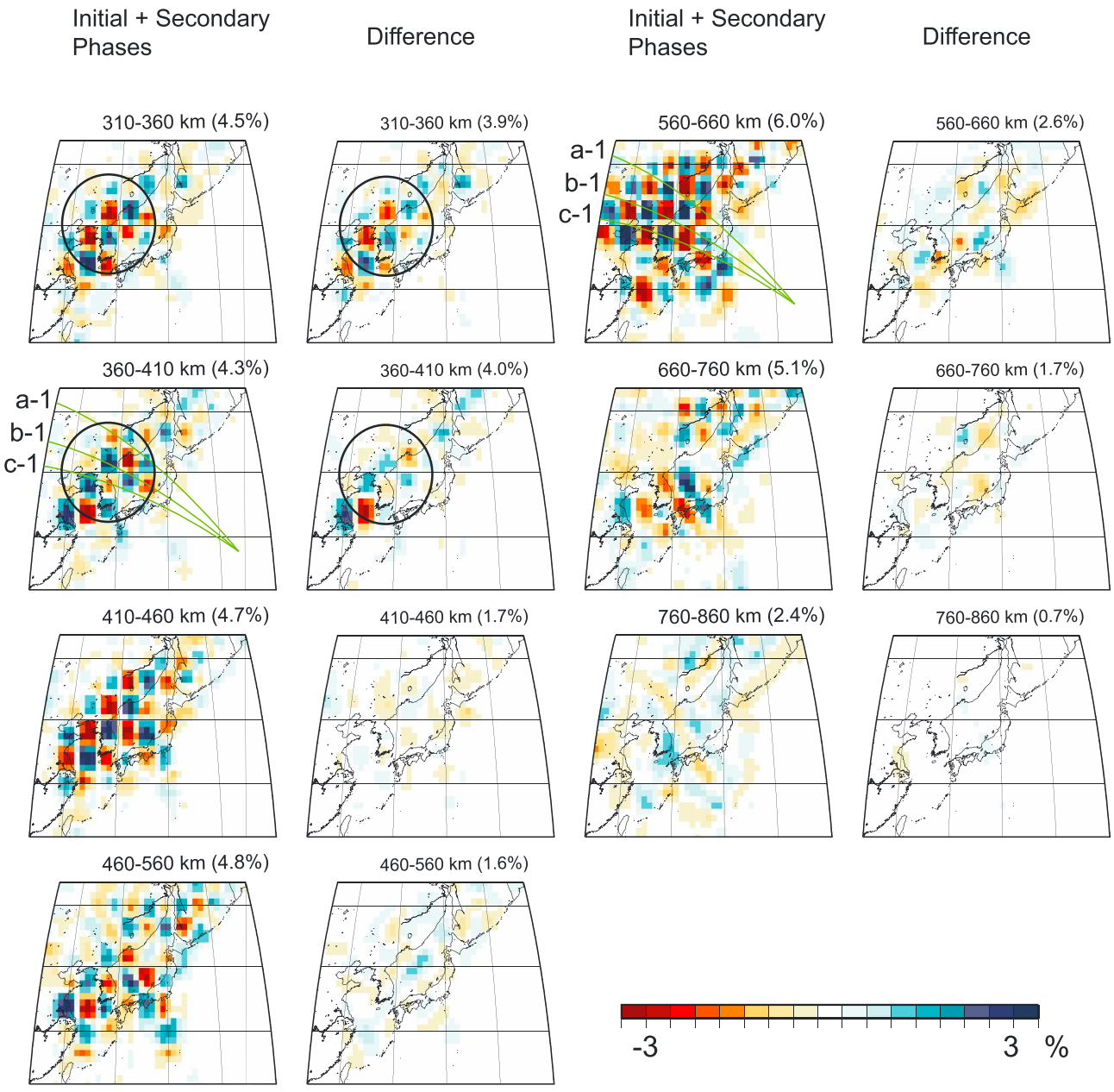

Figure 8. Continued

improvements in the path coverage just above the 410 km discontinuity (shown in Figure 6 (top left) and Figure 6 (top right)). Similar improvements are observed for the depth range between 560 and 660 km (Figure 8), but the improvements are marginal. This is probably due to the relatively smaller number of secondary phases bottoming just above the 660 compared with that for the 410 (see Figure 5a). Note that smaller improvements do not necessarily mean poor resolution. We actually have good resolution in the depth range of 560–660 km, as can be seen in Figure 8.

4.3. Comparison of Models

We next compare our model with the model obtained by using only the initial phase data (Figure 9). As expected from the resolution test in Figure 8, we do not see significant differences in the depth ranges

©2014. American Geophysical Union. All Rights Reserved.

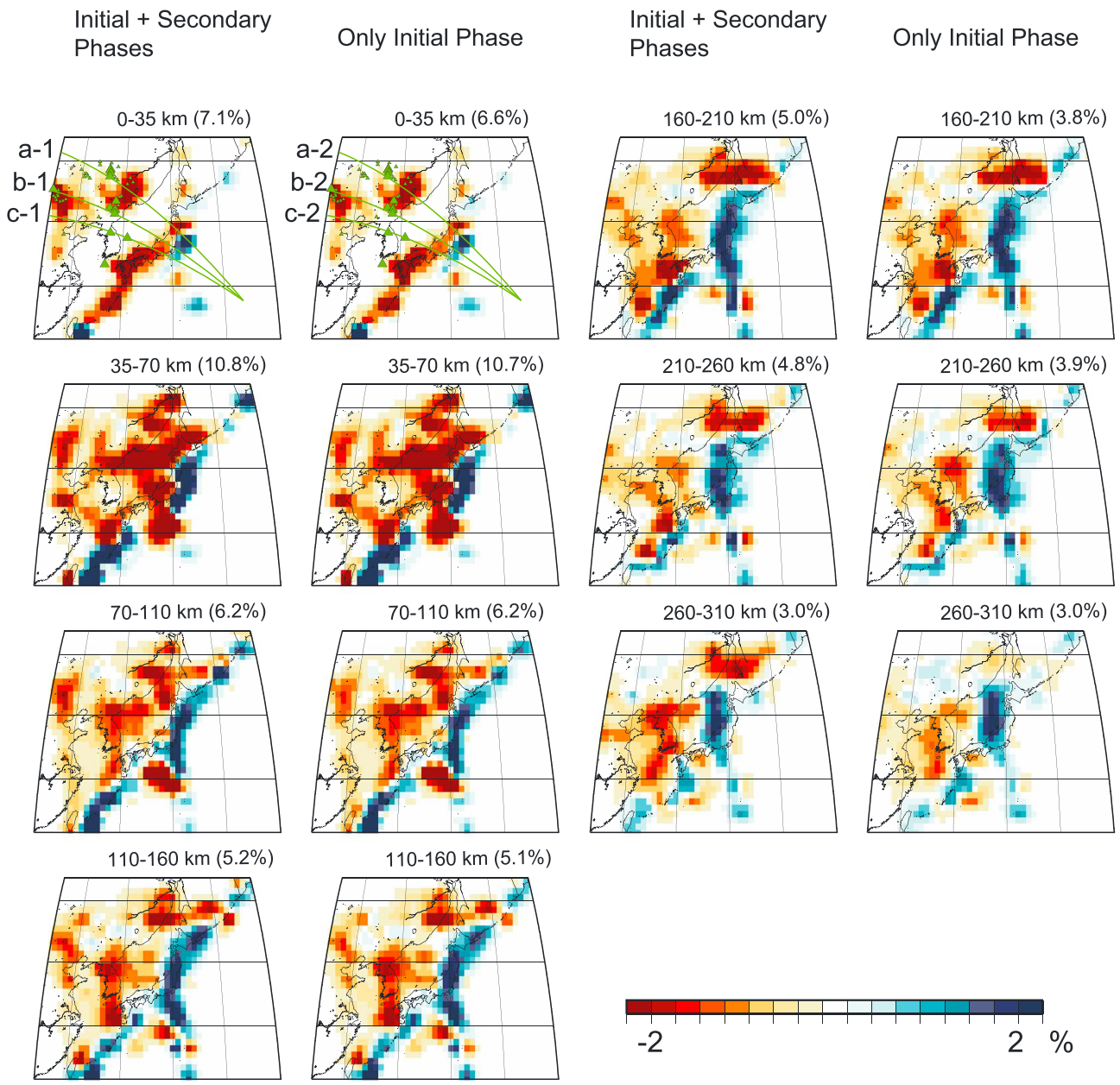

Figure 9. Comparison of the obtained *P* velocity models when we use (left) both the initial and secondary phases and (right) the initial phase only. Velocity perturbations with respect to the modified AK135 model (see the text for details) are shown for each depth range. The percentages shown in the upper right are the maximum absolute values of the velocity perturbation for that depth range. Green lines denote the locations of the vertical cross sections shown in Figure 10. Green triangles denote the locations of off-arc Holocene volcanoes (denoted by large triangles) and Pleistocene volcanoes (denoted by small triangles). The regions where we see notable differences are denoted by thick black circles.

above 260 km. However, for the structures between depths of 310 and 410 km, we see clear differences. Lower velocity anomalies beneath the Japan Sea and the Korean Peninsula are more pronounced in our model.

Such differences are also clearly observed in the vertical cross sections shown in Figure 10. Strong low-velocity anomalies are clearly observed down to the 410 in sections b-1 and c-1 (in the regions shown by thin black circles). However, the low-velocity anomalies just above the 410 are less clear in sections b-2 and c-2 obtained using only initial phase data. Such differences can affect interpretations for the origin of the off-arc volcanoes in this region, which will be discussed in the next section.

©2014. American Geophysical Union. All Rights Reserved.

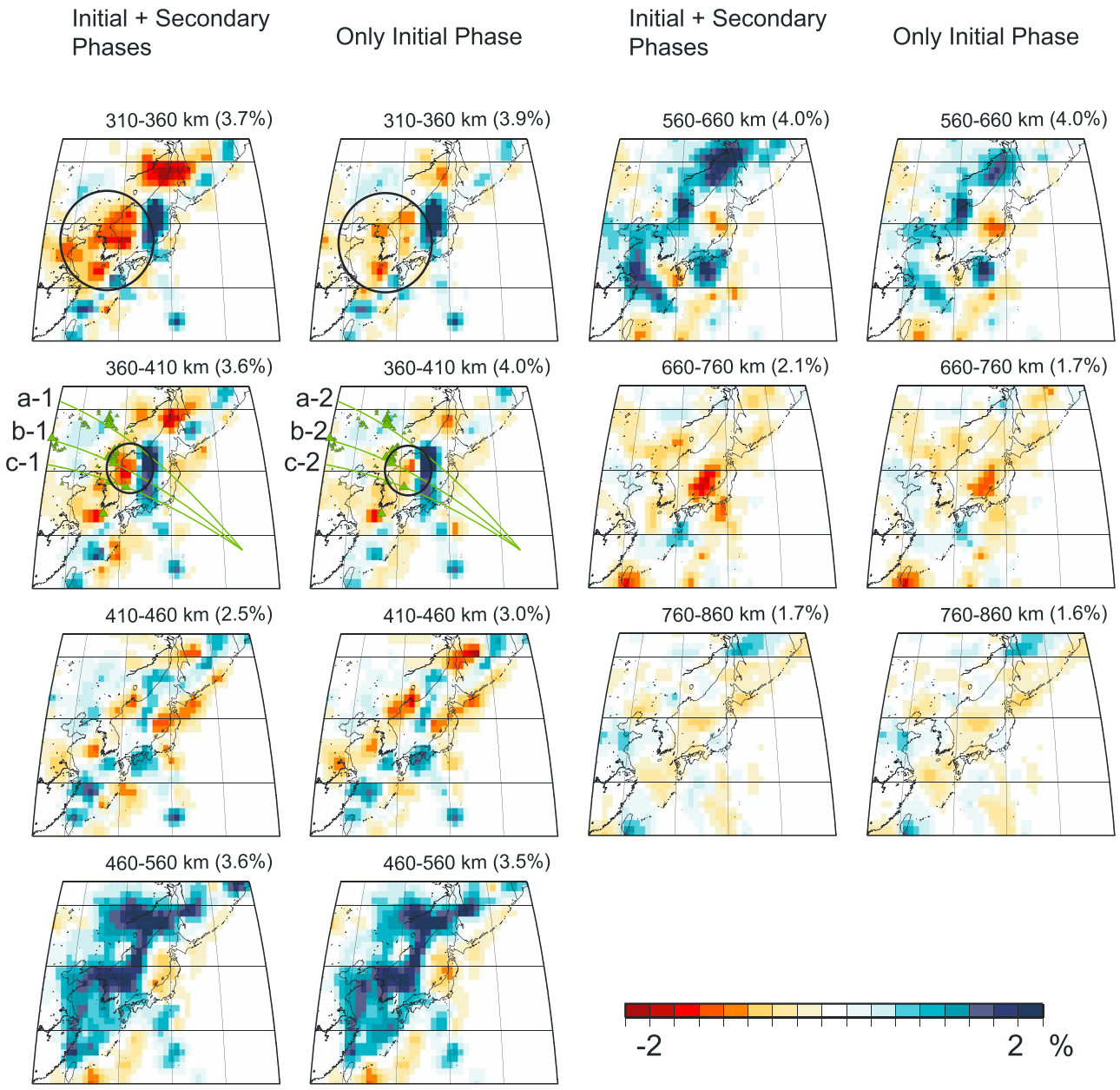

Figure 9. Continued

5. Discussion

In Figures 9 and 10, we show the distribution of off-arc Holocene and Pleistocene Volcanoes presented in *Chen et al.* [2007]. Because many of these volcanoes are located in the vicinity of the three sections in Figure 10, these three sections should give insight for the origin of off-arc volcanoes in this region.

We first compare the sections of our model (sections a-1, b-1, and c-1 in Figure 10) with the corresponding sections of the model obtained by using initial phase data only (sections a-2, b-2, and c-2 in Figure 10). In the second and third sections, we observe significant discrepancies (shown by thin black circles in Figure 10). We observe strong low-velocity anomalies just above the 410 km discontinuity in sections b-1 and c-1, whereas such anomalies are less prominent in sections b-2 and c-2. We therefore confirm the importance of the use of the secondary phases for discussing possible link between deep anomalies and the off-arc volcanism in this region.

©2014. American Geophysical Union. All Rights Reserved.

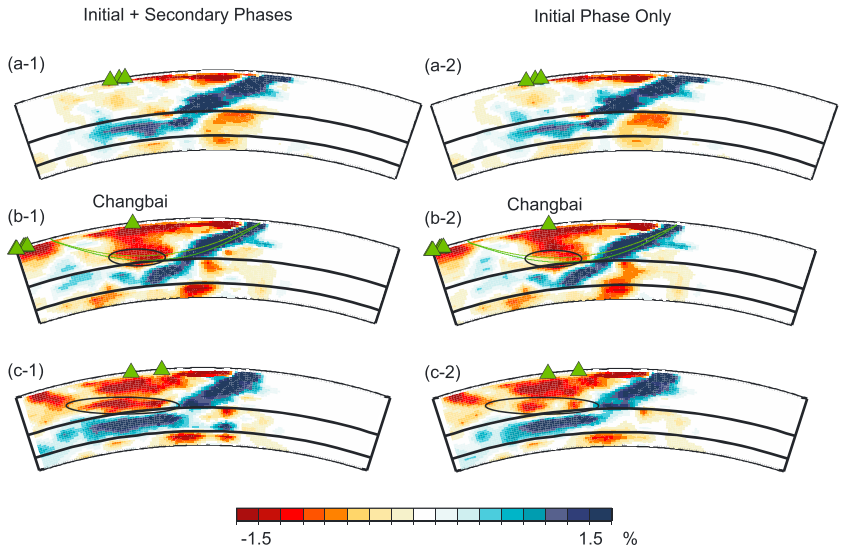

Figure 10. Comparison of the vertical cross sections for the locations specified in Figures 8 and 9. The model obtained by using (a-1, b-1, and c-1) both the initial and secondary phases and that obtained by using (a-2, b-2, and c-2) the initial phase only are shown. Green triangles denote the locations of off-arc volcanoes in Figure 9. The regions where we see notable differences are denoted by thin black circles. The green lines denote a typical raypath sampling the low-velocity anomalies beneath the Changbai Volcano for the initial model (upper lines) and the final model (lower lines). See the text for details.

Comparing the vertical sections in a-1, b-1, and c-1, a difference in the heterogeneity patterns is obvious. First, strong low-velocity anomalies extend down to the 410 km discontinuity in sections b-1 and c-1, whereas no anomalies are seen in section a-1. Next, the stagnant slab is clearly visible in the transition zone in sections a-1 and c-1, but the flat slab is unclear in section b-1, as has been pointed out by *Obayashi et al.* [2011] and *Tang et al.* [2014]. These regional variations suggest that the off-arc volcanoes in this region do not have a common origin.

Sections b-1 and b-2 cut into the vicinity of the Changbai Volcano. The structure in this region have been extensively discussed in previous studies, but the detailed features remain controversial. For example, some models [e.g., *Lei and Zhao*, 2005; *Zhao et al.*, 2009; *Li and van der Hilst*, 2010; *Tang et al.*, 2014] show low-velocity anomalies down to 410 km depth, while other models [e.g., *Huang and Zhao*, 2006; *Obayashi et al.*, 2011; *Wei et al.*, 2012] suggest low-velocity anomalies are confined to the uppermost 200–300 km mantle. Our model (section b-1) clearly shows low-velocity anomalies down to 410 km, which supports a deep origin for the magmatism.

A weakness in our approach is that we have assumed a 1-D velocity model for ray tracing. We do not believe this approximation significantly affects our main conclusions. To examine this we first compare raypaths, which sample the focused low-velocity anomalies, for the initial and the final models (Figure 10, green lines). We conducted 3-D ray tracing for the event and station shown by the blue star and circle in Figure 11, respectively. At a distance of 19.5° the bottoming depth for the 1-D initial model and for the 3-D final model are 351 km depth and 401 km depth, respectively, which will cause bias in the obtained model. Nonetheless, we believe that the observed low-velocity anomalies are robust because of the following evidence. We plot the observed traveltime residuals for events and stations (shown by the red stars and yellow circles in Figure 11a) in the vicinity of those in the above example. Note that the distance is corrected to the case when the event depth is 0 km. The observed scatter in the residuals for the different events probably results from using the source parameters in PDE rather than relocated sources. However, because we cannot precisely tell to which extent the errors in the source parameters contribute to the observed residuals, we here discuss only the overall features of the residuals for the initial parameters (i.e., the source parameters in PDE). The increase of the observed residuals almost disappear beyond the distance of 17° (Figure 11b) although the raypath travel for longer distance in the higher-velocity slab (see the raypath in Figure 10). Such features are predicted by our final model (Figure 11b, green line), while they are not predicted if we excluded

©2014. American Geophysical Union. All Rights Reserved.

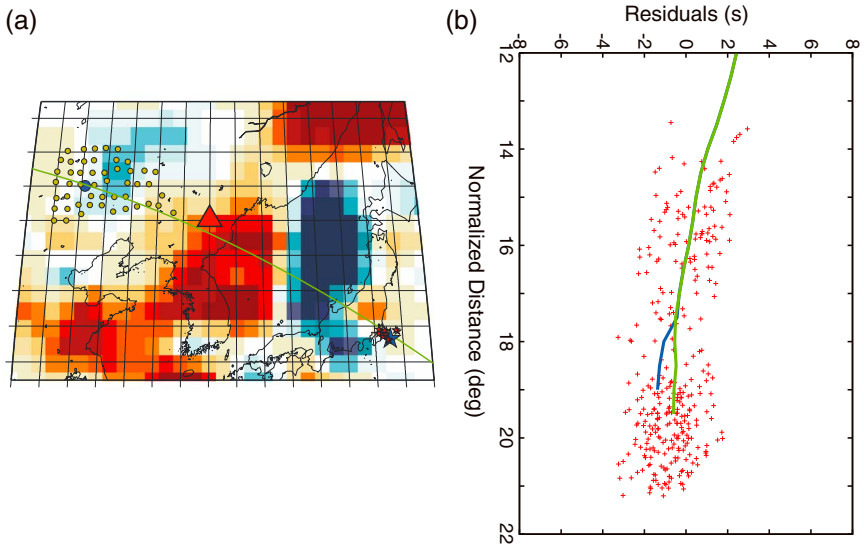

Figure 11. (a) Events (red stars) and stations (yellow circles), for which residuals are measured in Figure 11b. The red triangle and the green line denote the location of the Changbai Volcano and the section of Figure 10b-1, respectively. The *P* velocity model in Figure 9 at 310–360 km depth is overplotted. The event and station for which the raypath is shown in Figure 10 are also plotted by the blue star and the blue circle, respectively. (b) The observed *P* residuals as a function of distance. The distance is corrected to that for the event with 0 km depth. Colored lines denote predicted residuals computed for our final model (green) and our final model without heterogeneities between 310 and 410 km depth (blue). In predicting residuals, we used 3-D ray tracing.

heterogeneities between 310 and 410 km depth from our final model (Figure 11b, blue line). The results suggest the need of low-velocity anomalies at the bottom of the upper mantle to explain the observations.

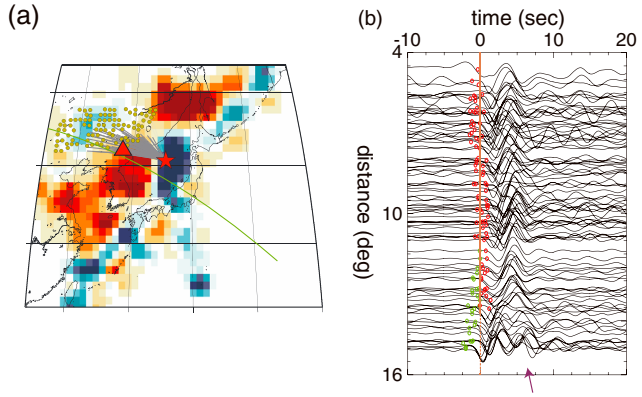

Figure 12. (a) Sampling region of *P* waves recorded with NECESSArray for the event at 06:48:11 on 5 February 2010 (M_w = 5.1; 382 km depth). The yellow circles and the red star denote the location of stations and the event, respectively. The heavy gray lines denote raypaths of *P* or diffracted *P* piercing the region between 310 and 410 km depths. The red triangle and the green line denote the location of the Changbai Volcano and the section of Figure 10b-1, respectively. The *P* velocity model in Figure 9 at 310–360 km depth is overplotted. (b) The observed *P* waveforms for the event of Figure 12a. The horizontal axis denotes the time relative to the synthetic arrival times of *P* or diffracted *P* bottoming above the 410 km discontinuity. The synthetic times are computed for the AK135 model [*Kennett et al.*, 1995]. The red and green circles are observed onset times, and the purple arrow denotes rough estimation of apparent velocity of the observed diffracted *P*.

A deep origin for the Changbai Volcano is directly suggested by observed waveforms (Figure 12). *P* waves for the event shown by the red star in Figure 12a (M_w = 5.1; 382 km depth) sample the region beneath the Changbai Volcano. In Figure 12b, we plot the observed *P* waveforms aligned by synthetic onset times of *P* or diffracted *P* (diffracted waves beyond the cusp B in Figure 1) bottoming above or on the 410 km discontinuity. Note that the origin time of the plot in Figure 12b is different from that in Figure 2: The origin time in Figure 2 was the synthetic arrival times of the initial *P* rather than AB or its diffracted waves. The observed onset times (shown by red circles in Figure 12b) increase as distance increases, which shows that the apparent velocity of *P* is slower than the prediction. The diffracted *P* bottoming on the 410 is also visible in Figure 12b, and its apparent velocity is also slower than predicted. These observations points to a low-velocity anomaly above the 410.

©2014. American Geophysical Union. All Rights Reserved.

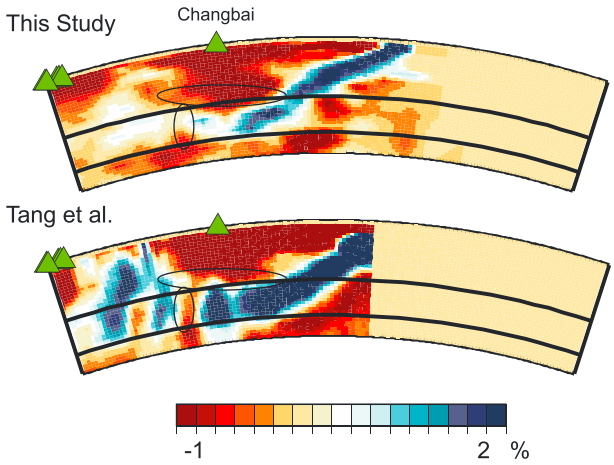

Figure 13. Comparison of the vertical cross sections between the *P* model by this study (upper) and the *S* model by *Tang et al.* [2014] for the location of b-1 specified in Figures 8 and 9. Similar color scales are used as those used in *Tang et al.* [2014]. The regions where we see notable differences are denoted by thin black circles.

Revenaugh and Sipkin [1994] have suggested that a low-velocity zone (LVZ) atop the 410 about 80 km thick exists above the stagnant slab in easternmost Asia. Our model suggests strong low-velocity anomalies at 310–410 km depths that are located above the regions of high-velocity anomalies at 460–660 km depths (Figure 9). Note that, because they analyzed longer-period waves (longer than 25 s), the boundary at the top of the LVZ is not necessarily sharp. Most of their corridors (except for their Corridor 1) sample our low-velocity regions, and their LVZ may correspond to our low-velocity anomalies. *Bagley et al.* [2009] suggested a similar LVZ exists in regions oceanward of the subducting Honshu slab, but we cannot observe pronounced corresponding slow anomalies in our model, which is probably due to insufficient resolution there (Figure 8). Receiver function analyses beneath the Japanese islands [e.g., *Tonegawa et al.*, 2006, 2008; *Kawakatsu and Yoshioka*, 2011] could not detect such an LVZ. Our model does not have strong low-velocity anomalies in those regions (Figure 9), but, because the resolution of our model is not so high there (Figure 8), we cannot conclude the absence of an LVZ.

The unclear (or a gap of) stagnant slab beneath the Changbai Volcano has only been explicitly pointed out in a few studies [e.g., *Obayashi et al.*, 2011; *Tang et al.*, 2014]. Regional array analyses by *Wang and Niu* [2010] detected weaker higher-velocity anomalies in the northern part of their study region (their subregion C) than those in the southern part (their subregion A), which implicitly suggests the unclear stagnant slab. Robustness of this feature has been discussed by *Tang et al.* [2014] who showed a larger amplitude signal of teleseismic *S* waves passing through this region. Our model and the model by *Obayashi et al.* [2011] show normal or slightly higher velocities, while the *S* wave model by *Tang et al.* [2014] shows lower velocity anomalies there. The discrepancies should be caused by a difference in analysis methods: *Obayashi et al.* [2011] and this study analyzed absolute *P* traveltimes for events at both triplication and teleseismic distances, while the *S* wave model of *Tang et al.* [2014] was obtained by analyzing relative *S* traveltimes within NECESSArray for events at teleseismic distances. Because the transition zone beneath NECESSArray, in general, has higher velocities, the stagnant slab could be imaged as a relatively low velocity anomaly.

Based on their *S* model, *Tang et al.* [2014] suggested a low-velocity conduit to the Changbai Volcano. In our model, such features are clear in the upper mantle but are not in the transition zone (the lower black circle in Figure 13). The discrepancies in the transition zone are probably due to the difference in analysis method as it was described above. The low-velocity anomalies atop the 410 are more pronounced in our model (the upper black circle in Figure 13). The initial model of Tang et al.'s *S* model was scaled from a global *P* model (GAP-P2 [*Obayashi et al.*, 2009]). This means that their *S* model was essentially obtained by using both of the *P* data set by *Obayashi et al.* [2009] and the *S* data set by *Tang et al.* [2014]. The observed difference suggests that both of the previous data sets have poor resolution in this region.

In our model, the strongest low-velocity anomalies atop 410 beneath the Changbai is −2.2% located at 39°N and 131°E, and pronounced anomalies are observed just beneath or slightly southeastward of the Changbai. The low-velocity anomalies there roughly range between −1 and −2% (see Figure 10b-1), which corresponds to about 100–200°C temperature anomalies if we use the scaling relation by *Karato* [1993]. Receiver function analysis by *Liu et al.* (Receiver function images of the mantle transition zone beneath NE China: New constraints on intraplate volcanism, deep subduction, and their potential link, submitted to *Earth Planet. Sci. Lett.*, 2014) (see Figure S4 of *Tang et al.* [2014] for the results) shows about 10–20 km depression of the 410 km discontinuity in the southeastward of the Changbai, although their studied region covers only a part of the low-velocity region of our model. The results can be consistently interpreted as thermal

©2014. American Geophysical Union. All Rights Reserved.

anomalies, which provides further support that we have hot material beneath the off-arc volcano which cannot be interpreted by the conventional plate tectonics.

Acknowledgments
We greatly thank Koji Miyakawa as well as other participants for their aids in our installing and maintaining seismographs in Northeast China. We also thank the Associate Editor and two anonymous reviewers for their constructive comments. We used an SGI ALTIX4700 installed at the Earthquake Research Institute, University of Tokyo, and an HA8000 installed at the Information Technology Center, University of Tokyo. This research is partially supported by JSPS KAKENHI grants 19104011, 21740323, 22000003, and 25400439. The NECESSArray project was supported by NSF, JSPS, and Peking University. All of the data from F-net were provided by the National Research Institute for Earth Science and Disaster Prevention, Japan (http://www.fnet.bosai.go.jp/fnet/). All of the other waveform data used in this study are available via IRIS DMC (http://www.iris.edu/dms/nodes/dmc/) and OHPDMC (http://ohpdmc.eri.u-tokyo.ac.jp).

References

Bagley, B., A. Courtier, and J. Revenaugh (2009), Melting in the deep upper mantle oceanward of the Honshu slab, *Phys. Earth Planet. Inter.*, *175*, 137–144.

Chen, Y., Y. X. Zhang, D. Graham, S. G. Su, and J. F. Deng (2007), Geochemistry of Cenozoic basalts and mantle xenoliths in Northeast China, *Lithos*, *96*, 108–126.

Choy, G. L., and P. G. Richards (1975), Pulse distortion and Hilbert transformation in multiply reflected and refracted body waves, *Bull. Seismol. Soc. Am.*, *65*, 55–70.

Dziewonski, A. M., and D. L. Anderson (1981), Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, *25*, 297–356.

Fuji, N., S. Chevrot, L. Zhao, R. J. Geller, and K. Kawai (2012), Finite-frequency structural sensitivities of short-period compressional body waves, *Geophys. J. Int.*, *190*, 522–540.

Fuji, N., K. Kawai, and R. J. Geller (2010), A methodology for inversion of broadband seismic waveforms for elastic and anelastic structure and its application to the mantle transition zone beneath the northwestern Pacific, *Phys. Earth Planet. Inter.*, *180*, 118–137.

Fukao, Y., M. Obayashi, T. Nakakuki, and The Deep Slab Project Group (2009), Stagnant slab: A review, *Ann. Rev. Earth Planet. Sci.*, *37*, 19–46.

Fukao, Y., S. Widiyantro, and M. Obayashi (2001), Stagnant slabs in the upper and lower mantle transition region, *Rev. Geophys.*, *39*, 291–323.

Gao, W., E. Matzel, and S. Grand (2006), Upper mantle structure beneath eastern Mexico determined from P and S waveform inversion and its implications, *J. Geophys. Res.*, *111*, B08307, doi:10.1029/2006JB004304.

Garcia, R., S. Chevrot, and M. Weber (2004), Nonlinear waveform and delay time analysis of triplicated core phases, *J. Geophys. Res.*, *109*, B01306, doi:10.1029/2003JB002429.

Gorbatov, A., and B. L. N. Kennett (2003), Joint bulk-sound and shear tomography for Western Pacific subduction zones, *Earth Planet. Sci. Lett.*, *210*, 527–543.

Grand, S. P., and D. V. Helmberger (1984a), Upper mantle shear structure of North America, *Geophys. J. R. Astron. Soc.*, *76*, 399–438.

Grand, S. P., and D. V. Helmberger (1984b), Upper mantle shear structure beneath the northwest Atlantic Ocean, *J. Geophys. Res.*, *89*, 1465–1475.

Grand, S. P., Y. Chen, H. Kawakatsu, Q. Chen, J. Ni, F. Niu, M. Obayashi, and S. Tanaka (2006), NorthEast China Extended SeiSmic Array (NECESSArray): Deep subduction, mantle dynamics and continental evolution beneath northeast China, *Eos Trans. AGU*, *87*(36), Western Pac. Geophys. Meet. Suppl., Abstract S41B-03.

Huang, J. L., and D. P. Zhao (2006), High-resolution mantle tomography of China and surrounding regions, *J. Geophys. Res.*, *111*, B09305, doi:10.1029/2005JB004066.

Iritani, R., N. Takeuchi, and H. Kawakatsu (2010), Seismic attenuation structure of the top half of the inner core beneath the northeastern Pacific, *Geophys. Res. Lett.*, *37*, L19303, doi:10.1029/2010GL044053.

Karato, S. (1993), Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.*, *20*, 1623–1626.

Kawakatsu, H., and S. Yoshioka (2011), Metastable olivine wedge and deep dry cold slab beneath southwest Japan, *Earth Planet. Sci. Lett.*, *303*, 1–10.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, *122*, 108–124.

Lei, J. S., and D. P. Zhao (2005), P-wave tomography and origin of the Changbai intraplate volcano in Northeast Asia, *Tectonophysics*, *397*, 281–295.

Li, C., and R. D. van der Hilst (2010), Structure of the upper mantle and transition zone beneath Southeast Asia from traveltime tomography, *J. Geophys. Res.*, *115*, B07308, doi:10.1029/2009JB006882.

Obayashi, M., J. Yoshimitsu, and Y. Fukao (2009), Tearing of stagnant slab, *Science*, *324*, 1173–1175.

Obayashi, M., et al. (2011), P-wave tomography of Northeastern China observed with NECESSArray, *Abstract S24B-06* presented at 2011 Fall Meeting, AGU, San Francisco, Calif., 5–9 Dec.

Okada, Y., K. Kasahara, S. Hori, K. Obara, S. Sekiguchi, H. Fujiwara, and A. Yamamoto (2004), Recent progress of seismic observation networks in Japan—Hi-net, F-net, K-NET and KiK-net, *Earth Planet Space*, *56*, 15–28.

Revenaugh, J., and S. Sipkin (1994), Seismic evidence for silicate melt atop the 410-km discontinuity, *Nature*, *369*, 474–476.

Stähler, S. C., K. Sigloch, and T. Nissen-Meyer (2012), Triplicated P-wave measurements for waveform tomography of the mantle transition zone, *Solid Earth*, *3*, 339–354.

Tajima, F., and S. P. Grand (1995), Evidence of high-velocity anomalies in the transition zone associated with southern Kurile subduction zone, *Geophys. Res. Lett.*, *22*, 3139–3142.

Tajima, F., and S. P. Grand (1998), Variation of transition zone high-velocity anomalies and depression of 660 km discontinuity associated with subduction zones from the southern Kuriles to Izu-Bonin and Ryukyu, *J. Geophys. Res.*, *103*, 15,015–15,036.

Tajima, F., I. Katayama, and T. Nakagawa (2009), Variable seismic structure near the 660 km discontinuity associated with stagnant slabs and geochemical implications, *Phys. Earth Planet. Sci.*, *172*, 183–198.

Tang, Y., M. Obayashi, F. Niu, S. P. Grand, Y. J. Chen, H. Kawakatsu, S. Tanaka, J. Ning, and J. F. Ni (2014), Changbaishan volcanism in northeast China linked to subduction-induced mantle upwelling, *Nat. Geosci.*, *7*, 470–475.

Tonegawa, T., K. Hirahara, T. Shibutani, and S. Katsuhiko (2006), Upper mantle imaging beneath the Japan Islands by Hi-net tiltmeter recordings, *Earth Planet Space*, *58*, 1007–1012.

Tonegawa, T., K. Hirahara, T. Shibutani, H. Iwamori, H. Kanamori, and S. Katsuhiko (2008), Water flow to the mantle transition zone inferred from a receiver function image of the Pacific slab, *Earth Planet. Sci. Lett.*, *274*, 346–354.

Wang, B., and F. Niu (2010), A broad 660 km discontinuity beneath northeast China revealed by dense regional seismic networks in China, *J. Geophys. Res.*, *115*, B06308, doi:10.1029/2009JB006608.

Wei, W., J. D. Xu, D. P. Zhao, and Y. L. Shi (2012), East Asia mantle tomography: New insight into plate subduction and intraplate volcanism, *J. Asian Earth Sci.*, *60*, 88–103.

Woodward, R. L., and G. Masters (1991), Global upper mantle structure from long-period differential travel times, *J. Geophys. Res.*, *96*, 6351–6377.

©2014. American Geophysical Union. All Rights Reserved.

Yoshizawa, K., K. Miyake, and K. Yomogida (2010), 3D upper mantle structure beneath Japan and its surrounding region from inter-station dispersion measurements of surface waves, *Phys. Earth Planet. Inter.*, *183*, 4–19.
Zhao, D. (2012), Tomography and dynamics of western-Pacific subduction zones, *Monogr. Environ. Earth Planets*, *1*, 1–70.
Zhao, D. P., Y. Tian, J. S. Lei, L. Liu, and S. Zheng (2009), Seismic image and origin of the Changbai intraplate volcano in East Asia: Role of big mantle wedge above the stagnant Pacific slab, *Phys. Earth Planet. Inter.*, *173*, 197–206.
Zheng, Y., W. Shen, L. Zhou, Y. Yang, Z. Xie, and M. H. Ritzwoller (2011), Crust and uppermost mantle beneath the North China Craton, northeastern China, and the Sea of Japan from ambient noise tomography, *J. Geophys. Res.*, *116*, B12312, doi:10.1029/2011JB008637.

©2014. American Geophysical Union. All Rights Reserved.