

EARA2024: a new radially anisotropic seismic velocity model for the crust and upper mantle beneath East Asia and Northwestern pacific subduction zones

Ziyi Xi¹, Min Chen^{1,2,*}, Songqiao Shawn Wei^{1,2}, Jiaqi Li^{1,2}, Tong Zhou,^{1,†}, Baoshan Wang^{3,4} and YoungHee Kim⁵

¹Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI 48824, USA. E-mail: xiziyi@msu.edu

²Department of Earth and Environmental Sciences, Michigan State University, East Lansing, MI 48824, USA. E-mail: swei@msu.edu

³School of Earth and Space Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China

⁴Mengcheng National Geophysical Observatory, University of Science and Technology of China, Hefei, Anhui 230026, China

⁵School of Earth and Environmental Sciences, Seoul National University, Seoul 08826, Republic of Korea

Accepted 2024 August 16. Received 2024 June 17; in original form 2023 June 21

SUMMARY

We present a new 3-D radially anisotropic seismic velocity model EARA2024 of the crust and mantle beneath East Asia and the northwestern Pacific using adjoint full-waveform inversion tomography. We construct the EARA2024 model by iteratively minimizing the waveform similarity misfit between the synthetic and observed waveforms from 142 earthquakes recorded by about 2000 broad-band stations in East Asia. Compared to previous studies, this new model renders significantly improved images of the subducted oceanic plate in the upper mantle, mantle transition zone, and uppermost lower mantle along the Kuril, Japan, Izu-Bonin and Ryukyu Trenches. Complex slab deformation and break-offs are observed at different depths. Moreover, our model provides new insights into the origins of intraplate volcanoes in East Asia, including the Changbaishan, Datong-Fengzhen, Tengchong and Hainan volcanic fields.

Key words: Asia; Computational seismology; Waveform inversion; Subduction zone processes.

1 INTRODUCTION

East Asia and northwestern Pacific subduction zones consist of various tectonic structures and host diverse and active geological processes in the crust and upper mantle (Fig. 1). In the east, the Pacific Plate subducts westwards along the Kuril, Japan, Izu-Bonin, Mariana and Yap Trenches, and the Philippine Sea Plate subducts along the Nankai, Ryukyu, Malina and Philippine Trenches. These subduction zones, with numerous earthquakes and arc volcanoes, are one of the most geologically active regions in the world. In the middle, the Xing'an-East Mongolia Block, North China Block and South China Block form the central-eastern end of the Eurasian continent (Zhang 1997). In the west, the collision between the Indian and Eurasian Plates forms the Tibetan Plateau and Himalayan Mountains (Allégre *et al.* 1984; Yin & Harrison 2000). The complex geological structure and active tectonic activities in this region provide a natural laboratory for studying inter-plate and

intra-plate tectonics. For instance, the collision of the Indian Plate and Eurasian Plate results in the Karakorum, Kunlun, Altyn-Tagh Faults and the Tarim, Junggar and Qaidam Basins (Bruce Watson & Brenan 1987). The resistance of the Yangtze Block in southern China to the subducted Indian Plate causes numerous intraplate earthquakes in southwestern China (Tappornier *et al.* 1986). Furthermore, the subducted Pacific Plate in the mantle transition zone (MTZ) also has significant influences on the evolution of the East Asian lithosphere (Zhao *et al.* 2007, 2009; Lei *et al.* 2013; He *et al.* 2014; Tang *et al.* 2014; Guo *et al.* 2016, 2018).

High-resolution seismic images provide valuable information about the crustal and upper mantle structures and geological processes in this region. Since decades ago, body-wave traveltimes tomography, based on the high-frequency approximation or 3-D finite-frequency kernels of seismic wave propagation, has been widely used to image East Asia (Aki & Lee 1976; Dziewonski *et al.* 1977; Friederich 2003; Lei & Zhao 2005, 2016; Huang & Zhao 2006; Zhao *et al.* 2007, 2009; Kustowski *et al.* 2008b; Lei *et al.* 2009; Tian *et al.* 2009; Li & Van Der Hilst 2010; Obrebski *et al.* 2012; Li *et al.* 2021). In addition, ambient noise, surface wave dispersion and Rayleigh wave ellipticity also have been used to image the subsurface structure beneath East Asia (Yao *et al.* 2006, 2008; Zheng *et al.*

*Deceased

†Now at: Aramco Research Center–Beijing, Aramco Asia, Beijing 100102, China.

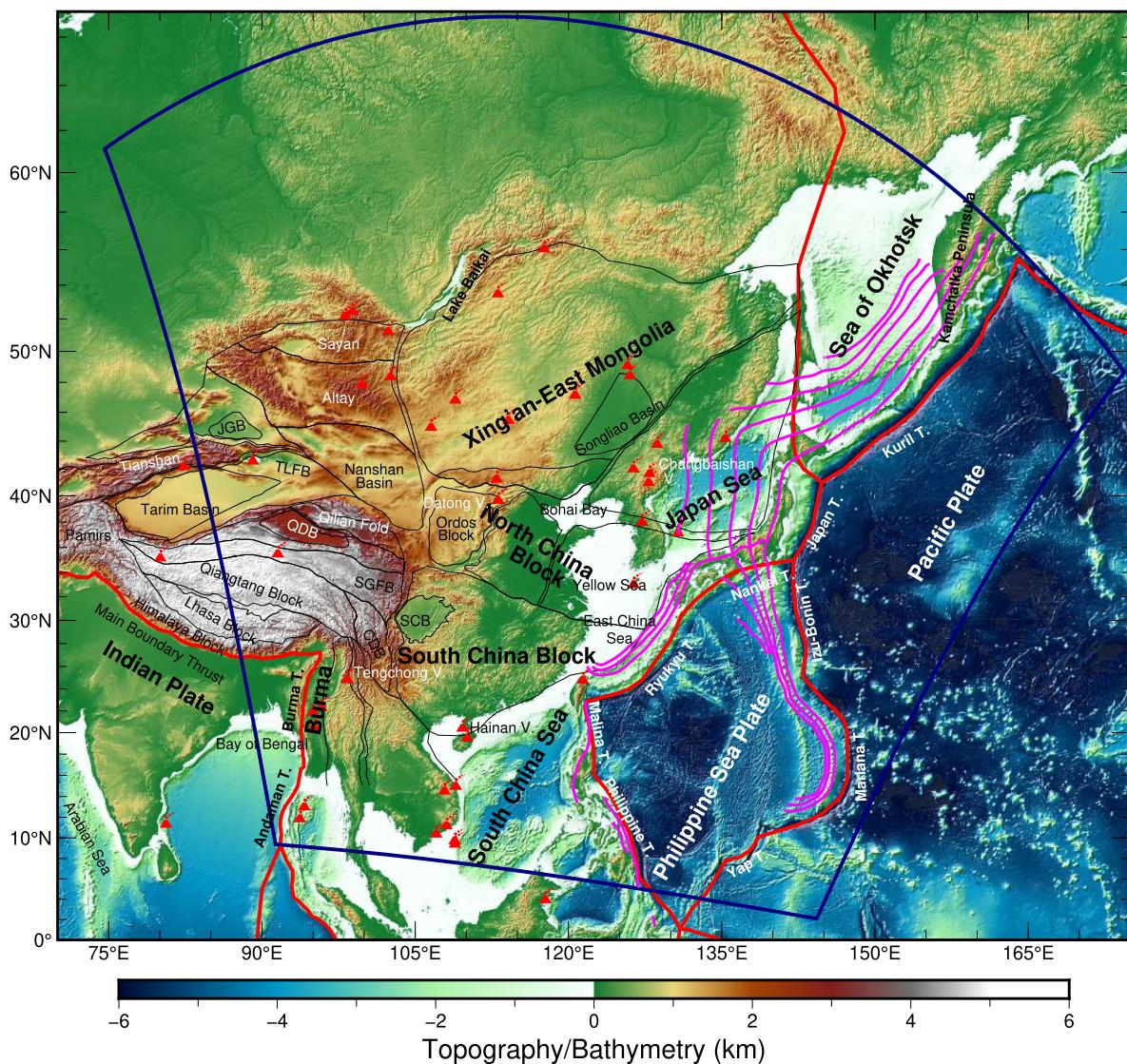


Figure 1. Topography/bathymetry map of East Asia and the northwestern Pacific. The simulation region is outlined by navy blue lines. Magenta curves represent slab depth contours with a 100-km interval based on the Slab2 model (Hayes *et al.* 2018). The tectonic units are outlined by black curves, and intraplate volcanoes are marked with red triangles. Red thick lines indicate plate boundaries from the NUVEL-1 model (DeMets *et al.* 1990). Acronyms of certain tectonic units: CDB, Chuandian Block; JGB, Junggar Basin; QDB, Qaidam Basin; SCB, Sichuan Basin; TLFB, Tulin Basin; SGFB, Songpan Ganzi Fold Belt.

2006; Witek *et al.* 2014; Bao *et al.* 2015; Li *et al.* 2016; Yang *et al.* 2019). Finite-difference methods of simulating wave propagation provide another tool for studying the East Asian crust and upper mantle (Wang *et al.* 2013).

The development of a spectral-element method (SEM) provides another way to simulate wave propagation and naturally considers the finite-frequency effects (Komatsitsch & Tromp 2002a, b). Its combination with the adjoint state method provides an effective way to compute the gradients of 3-D model parameters and thus can be used for tomographic imaging. The full-waveform inversion (FWI) method based on SEM and adjoint state methods images the Earth's interior by iteratively applying the SEM to simulate seismic waveforms in 3-D models and then updating the model with finite-frequency kernels. By considering all usable segments of seismic waveforms, this method uses much more information than other seismic tomography methods, thus potentially enhancing the model resolution. Furthermore, there is no need for a crustal correction,

a common practice in conventional tomography methods, because the crust is explicitly included in the simulations. Previous studies have suggested that imperfect crustal corrections could have large impacts on radial anisotropy (Bozdağ & Trampert 2008; Ferreira *et al.* 2010), even in the lower mantle (Panning *et al.* 2010). Therefore, we choose to invert simultaneously for mantle and crust to correctly map anisotropy. This approach also makes it easy to consider the effects of the Earth's ellipticity, bathymetry, topography and internal discontinuities in the full waveform simulation. Several global or regional models have been developed using this method (Tape *et al.* 2007, 2009, 2010; Fichtner *et al.* 2009, 2010; Zhu *et al.* 2012a, b; Lei *et al.* 2013, 2020; French & Romanowicz 2014; Chen *et al.* 2015; Bozdağ *et al.* 2016; Tao *et al.* 2018; Zhou *et al.* 2022; Ma *et al.* 2022).

In this work, we use SPECFEM3D_GLOBE (Komatsitsch & Tromp 2002a, b; Peter *et al.* 2011) to invert 10 yr of high-quality seismic data collected from East Asia and the northwestern Pacific

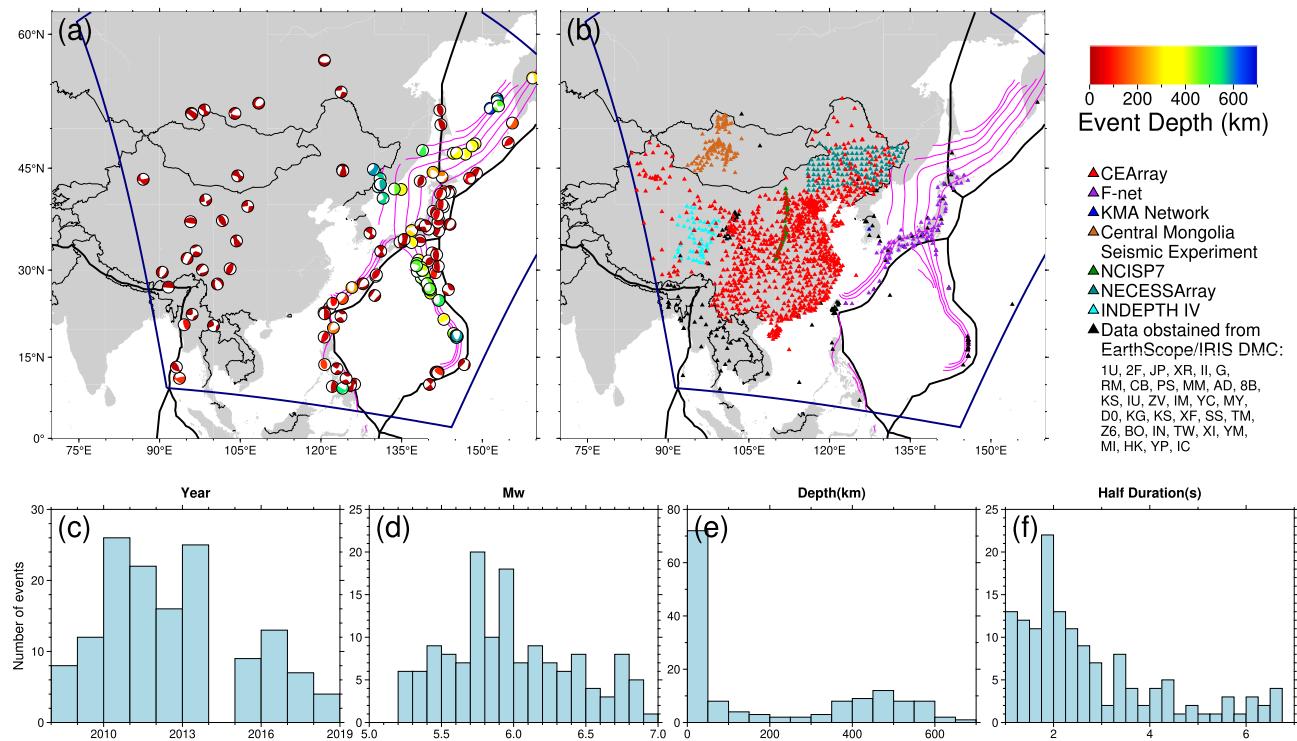


Figure 2. Distribution of earthquakes (a) and seismic stations (b) used in this study. Earthquakes in (a) are colour-coded by depth. The focal mechanisms are from the Global Centroid Moment Tensor (GCMT) project (Dziewonski & Anderson 1981; Ekström *et al.* 2012). Stations in (b) are colour-coded by networks. Other features in (a) and (b) are the same as Fig. 1. (c–f) Histograms of earthquake origin time (c), moment magnitude (d), depth (e) and half duration time of the source time function (f).

to image the crustal and mantle structure. This paper presents our latest FWI tomography model of this region: EARA2024. First, we will introduce our dataset and methods. Secondly, we will assess our model with resolution tests and waveform comparisons. Thirdly, we will describe the model and compare our model with previous models. Lastly, we will discuss slab morphology in northwestern Pacific subduction zones and the origin of four intraplate volcanoes based on our new model.

2 DATA AND METHOD

2.1 Data selection and processing

We defined our study region in East Asia and the northwestern Pacific as a nearly square region centred at 36°N , 123°E , with a length of 60° on each side, and anticlockwise rotated by 10° (Fig. 1). We selected 142 earthquakes with M_w 5.2–7.0 from the global centroid moment tensor (CMT) catalogue (Ekström *et al.* 2012) between 2008 and 2018 in our study region (Fig. 2). These earthquakes were carefully chosen to provide balanced data coverage. We excluded the events less than 2° away from the study region boundary to further reduce the numerical noise from the boundary effect, even though Clayton–Engquist–Stacey absorbing conditions (Clayton & Engquist 1977; Stacey 1988) had been used in our simulations. 58 of these 142 earthquakes were deeper than 150 km for better constraining deep mantle structures, such as the subducted slabs and MTZ. We also excluded events with M_w larger than 7.0 or source half duration time longer than 7 s to avoid complex rupturing processes that might complicate the structural inversion. Consequently,

our earthquake catalogue consisted of 142 events in the Kuril, Japan, Izu-Bonin, Mariana, Nankai, Ryukyu, Manila and Philippine subduction zones down to about 650 km depth, as well as continental events shallower than 30 km in the Eurasian lithosphere.

We used three-component seismic waveforms recorded by about 2000 stations from the China Earthquake Administration Array (CEArray) (Zheng *et al.* 2009), F-net (Okada *et al.* 2004), NECESSArray (Tang *et al.* 2014), INDEPTH IV project (Wei *et al.* 2010; Yue *et al.* 2012), the central Mongolia seismic experiment (Meltzer *et al.* 2019), the Korean seismic network and other regional or global seismic networks. More than 90 per cent of the stations had broad-band (BH and HH) seismometers, whereas other stations had short-period high-gain (SH) seismometers. After detrending and tapering, we first removed the instrumental response to get displacement waveforms in three components: vertical (Z), east (E) and north (N). We then rotated the two horizontal components (E and N) to the radial (R) and tangential (T) components. In the next step, the R and T components of stations with misorientation were corrected according to the methods described in previous studies for CEArray (Niu & Li 2011) and the Korean seismic network (Lim *et al.* 2018). In the first inversion stage, we applied a second-order zero-phase Butterworth bandpass filter between 10 and 40 s for body waves and 40–120 s for surface waves. In the second stage, we tried to achieve higher resolution with higher-frequency signals and filtered body waves at 8–40 s and surface waves at 30–120 s. In the third stage, we kept the same bandpass filter range for body waves as the second stage, and filtered surface waves at 20–120 s. Finally, we down-sampled the filtered data to 10 samples per second.

2.2 Configuration of the spectral-element-method simulation

We use a spectral-element method (Komatitsch & Tromp 2002a, b) to simulate wave propagation and compute sensitivity kernels based on the adjoint state method. In addition, we aim to use a conjugate gradient method (Hestenes & Stiefel 1952) to maximize the waveform similarity and phase matching between data and synthetics. The 3-D model space extends from the Earth's surface, as delineated by the Crust1.0 model (Laske *et al.* 2013), down to the centre of the Earth. We construct the initial model by combining two regional models, FWEA18 (Tao *et al.* 2018) and EARA2014 (Chen *et al.* 2015), embedded in two global models, S362ANI (Kustowski *et al.* 2008b) for the mantle and CRUST1.0 for the crust. FWEA18, a radially anisotropic model, contributes to the majority of this initial model. In regions not covered by FWEA18, the initial model relies on EARA2014 instead. This approach leverages the consistency and compatibility of the methodologies behind these models, and is driven by the progression in model development, where each new model builds upon the previous one, supplemented by additional data. It is likely that the most recent model provides a better estimation of the real Earth, although this may not always be the case due to various potential numerical issues. By starting with a model that is as close to reality as possible, we aim to enhance the efficiency and accuracy of our inversion process. Our model construction process involved merging the velocity perturbations of FWEA18 and EARA2014 relative to S362ANI in the mantle and to Crust1.0 in the crust. To facilitate a smooth integration of these models and mitigate numerical instability, we apply Gaussian smoothing to the combined perturbations. The extends of these models are shown in Fig. S12. We fix the topography of the 410- and 660-km discontinuities as that from S362ANI and the attenuation structure as that from QL6 (Durek & Ekström 1996) during the inversions.

In the crust and upper mantle, we use five parameters to parametrize our model: density ρ , isotropic bulk sound speed V_C , horizontally propagating S wave speeds V_{SV} (polarized in the vertical direction) and V_{SH} (polarized in the horizontal direction) and a non-dimensional parameter η measuring the anellipticity of P -wave phase slowness. Radial anisotropy is introduced to the crustal part of our model, because previous studies indicate the presence of radial anisotropy in this region (Xie *et al.* 2013; Tao *et al.* 2018; Witek *et al.* 2021). Since previous studies (e.g. Panning & Romanowicz 2006) suggest that the upper part of the lower mantle is nearly isotropic, we parametrize it with isotropic P wave speed V_P , isotropic S wave speed V_S , and density ρ . Because density ρ cannot be well constrained by our data, we scale ρ with isotropic shear wave speed perturbation $\delta \ln V_S$ as $\delta \rho = 0.338 \ln V_S$ (Panning & Romanowicz 2006). Here the perturbation reference model STW105 (Kustowski *et al.* 2008a) is crucial in the inversion process, as we update model perturbations rather than the absolute values of model parameters.

The horizontal dimension of the spectral element is set as ~ 20 km in the crust and doubled in the mantle, and the vertical dimension is ~ 10 km in the crust and ~ 35 km in the mantle. The shortest period for waveform forward modelling is 8 s. Each spectral element contains 125 Gauss–Lobatto–Legendre interpolation points. In total, the simulation involves about 2.3 million elements. To simulate a 30-min wave propagation for one event, the SPECFEM.GLOBE takes ~ 30 min wall time for the forward simulation and ~ 70 min for sensitivity kernel calculation using 441 cores on the SKX nodes of the Stampede2 supercomputer at the Texas Advanced Computing Center (TACC). The total volume of seismic data and synthetics in

each iteration is ~ 800 GB. To improve I/O efficiency, we develop a Python-based data analysis and inversion workflow based on the Adaptable Seismic Data Format (ASDF, Krischer *et al.* 2016) and ObsPy (Beyreuther *et al.* 2010) to store and process the seismic data.

2.3 Waveform misfit measurement

We group our measurements into six categories: P – SV body waves on the vertical and radial components, SH body waves on the transverse component, Rayleigh waves on the vertical and radial components, and Love waves on the transverse component. We use a conjugate gradient method (Hestenes & Stiefel 1952) to iteratively minimize the misfit between recorded waveforms and synthetics. More specifically, we define the misfit function as a normalized zero-lag cross-correlation coefficient (NZCC). Compared to other traveltimes-only-based misfit functions, NZCC captures both waveform similarity and traveltimes information. Although NZCC introduces non-linearity to the inversion, this problem can be mitigated with a relatively small step length and a relatively fine starting model in the conjugate gradient update. The NZCC for each measurement window is defined as:

$$\text{NZCC}^{c,e,s,w} = \frac{\int s^{c,e,s,w}(t) \cdot u^{c,e,s,w}(t) dt}{\sqrt{\int |s^{c,e,s,w}(t)|^2 dt \int |u^{c,e,s,w}(t)|^2 dt}}, \quad (1)$$

where $s(t)$ is the synthetic and $u(t)$ is the observed waveform in category c and window w for event e recorded by station s . We measure all major body-wave and surface-wave phases, including P , pP , sP , PP , S , sS , SS , ScS , Rayleigh wave and Love wave.

Compared with other misfit measurements that utilize dispersion or frequency-dependent phase arrival time, the NZCC is a fixed frequency measurement method. Previous studies introduce several waveform inversion strategies to address dispersion, including generalized seismological data functionals (Gee & Jordan 1992), multitaper delay time measurements (Tape *et al.* 2010), instantaneous phase and envelope misfits (Bozdağ *et al.* 2011; Rickers *et al.* 2012), exponentiated phase (Yuan *et al.* 2020) and time–frequency phase misfit for waveform misfit objectives (Fichtner *et al.* 2009, 2010, 2013; Krischer *et al.* 2018). These methods focus on dispersed signals and measure frequency-dependent traveltimes at various frequencies, utilizing phase and waveform information. Multi-scale measurements, which are sensitive to a broadband of dispersed signals (long, intermediate and short periods), are considered robust in improving model resolution on different scales (Rodgers *et al.* 2022). In contrast, NZCC focuses on the maximum amplitude parts of the waveforms, which are more sensitive to rapid velocity contrasts, often resulting in sharper boundaries in the inversion results (Tao *et al.* 2017, 2018; Zhou *et al.* 2021, 2022). Although NZCC has band-limited sensitivity and does not naturally incorporate longer period measurements from earlier inversion stages, its focus on waveform shape, especially the maximum amplitude parts in the measurement window, offers a significant advantage in detecting rapid changes in subsurface properties. Therefore, we use NZCC in this study focusing on the Western Pacific subduction zones, where sharp velocity contrasts are expected. To mitigate the limitation of the band-limited measurements, we expand the surface wave frequency band from 40–120 s in the first stage to 30–120 s in second stage, so that the long-period information can be honoured through both stages.

Amplitude information helps increase the model resolution but can suffer from non-linearity and attenuation effects. Consequently,

Table 1. Summary of weighting terms in the misfit function.

Weighting term	Formula	Applied on	Comment
Categorical weighting W_c	$\frac{1}{\sum_{e=1}^{N_e} \sum_{s=1}^{N_{e,s}} N_{c,e,s,w}}$	Each category	e : event; s : station; c : category; w : window; N_e : number of events; $N_{e,s}$: number of stations for event e ; $N_{c,e,s,w}$: number of measurement windows in category c for event e and station s
Geographical weighting $W_{e,s}$	$\frac{1}{\sum_{j=1}^{N_{e,s}} e^{-\left(\frac{\Delta_{sj}}{\Delta_0}\right)^2}}$	Each station	s : current station; j : other stations but include s itself; e : event; $N_{e,s}$: number of stations for event e ; Δ_{sj} : Epicenter distance between station s and station j ; Δ_0 : reference distance, with detail in note ^a
Quality weighting W_w	$f_1(\text{CC})f_2(\text{SNR})f_3(\delta t_{\text{CC}})$	Each measurement window	CC: cross-correlation coefficient; SNR: signal to noise ratio; δt_{CC} : time-shift obtained from cross-correlation; f_1 , f_2 , and f_3 are defined in the same manner as eq. (3) to eq. (6) in Tao et al. (2018)

Note:^a For the geographical weighting, we select Δ_0 as $\frac{1}{\sum_{j=1}^{N_{e,s}} e^{-\left(\frac{\Delta_{sj}}{\Delta_0}\right)^2}} = \frac{1}{3} \max_x \frac{1}{\sum_{j=1}^{N_{e,s}} e^{-\left(\frac{\Delta_{sj}}{x}\right)^2}}$. This definition is consistent with Ruan et al. (2019).

many studies exclude amplitude information (e.g. Bozdağ *et al.* 2011). In this study, we use relative, rather than absolute, amplitude to construct NZCC, because this approach helps eliminate the influence of attenuation on waveform amplitude. Although attenuation also affects waveform shape, we believe this effect is not severe enough to outweigh the benefits of NZCC in our study.

Following Tao *et al.* (2018), we select a measurement window based on a predicted travelttime because this strategy is more computationally efficient compared to the conventional FLEXWIN method (Maggi *et al.* 2009). We define the body-wave measurement windows from 20 s before to 50 s after the predicated phase arrival times based on the AK135 model (Kennett *et al.* 1995). If multiple body-wave windows overlap with each other, we merge them into a single one. For surface waves, the measurement window starts 50 s before a predicted phase arrival time, assuming a wave speed of 4.6 km s⁻¹, and ends 50 s after a predicted phase arrival time, assuming a wave speed of 3.3 km s⁻¹. We only measure surface waves from events shallower than 150 km. In total, we measure ~650 000 body-wave windows and ~150 000 surface-wave windows that satisfy our criteria of signal-to-noise ratio (SNR), cross-correlation coefficient (CC) and time-shift. The three criteria are described in Appendix A in detail. It is worth noticing that the absolute amplitude information is not taken into account because the NZCC is normalized. Therefore, our inversion is insensitive to seismic attenuation even though it may slightly distort waveforms. Therefore, we only focus on the Earth's elastic properties in this study.

We apply categorical and geographical weighting (Table 1) to balance the misfits from different categories and regions. The weighted misfit function is

$$\chi = \sum_{c=1}^{N_c} W_c \sum_{e=1}^{N_e} \sum_{s=1}^{N_{e,s}} W_{e,s} \sum_{w=1}^{N_{c,e,s,w}} W_w (1 - \text{NZCC}^{c,e,s,w}), \quad (2)$$

where N_c represents the number of categories (i.e. 6) used in our inversion, N_e indicates the total number of events, $N_{e,s}$ is the number of stations for event e and $N_{c,e,s,w}$ is the number of measurement windows in the corresponding category c for event e and station s . W_c is a categorical weighting term, balancing the misfit contribution from different categories. $W_{e,s}$ is a geographical weighting term, balancing the misfit contribution from different regions and mitigating the bias due to non-uniform station distribution (Ruan *et al.* 2019). W_w is a weighting term to control the waveform quality in each measurement window.

The corresponding adjoint source of NZCC is defined as

$$\frac{\delta \chi}{\delta s} = \sum_{c=1}^{N_c} W_c \sum_{e=1}^{N_e} \sum_{s=1}^{N_{e,s}} W_{e,s} \sum_{w=1}^{N_{c,e,s,w}} W_w \frac{\delta \chi^{c,e,s,w}}{\delta s^{c,e,s,w}}, \quad (3)$$

where

$$\frac{\delta \chi^{c,e,s,w}}{\delta s^{c,e,s,w}} = \frac{s^{c,e,s,w}(t) - A^{c,e,s,w} u^{c,e,s,w}(t)}{W^{c,e,s,w}} \quad (4)$$

$A^{c,e,s,w}$ is defined as

$$A^{c,e,s,w} = \frac{\int s^{c,e,s,w}(t) dt \int u^{c,e,s,w}(t) dt}{\int |u^{c,e,s,w}(t)|^2 dt}. \quad (5)$$

And the normalization factor $W^{c,e,s,w}$ is defined as

$$W^{c,e,s,w} = \sqrt{\int |s^{c,e,s,w}(t)|^2 dt \int |u^{c,e,s,w}(t)|^2 dt}. \quad (6)$$

The structural inversion details will be described below in Section 2.5.

2.4 Source inversion

Earthquake source parameters also influence seismic waveforms, and thus we have to invert for source and structure iteratively. Here we use the steepest descent method to optimize the source model and the adjoint state method to compute the gradients following Kim *et al.* (2011). The misfit function and the adjoint source of event e for the source inversion are defined in similar ways:

$$\chi_e = \sum_{c=1}^{N_c} W_c \sum_{s=1}^{N_{e,s}} W_{e,s} \sum_{w=1}^{N_{c,e,s,w}} W_w (1 - \text{NZCC}^{c,e,s,w}) \quad (7)$$

$$\frac{\partial \chi_e}{\partial t} = \sum_{c=1}^{N_c} W_c \sum_{s=1}^{N_{e,s}} W_{e,s} \sum_{w=1}^{N_{c,e,s,w}} W_w \frac{\partial \chi^{c,e,s,w}}{\partial t}. \quad (8)$$

The subscripts and superscripts here have the same definitions as those in Section 2.3. Categorical weighting is defined only for a single event:

$$W_c = \frac{1}{\sum_{s=1}^{N_{e,s}} N_{c,e,s,w}}. \quad (9)$$

The invertible source parameters include centroid moment tensor M , 3-D centroid location X , and source time function $S(t)$, which includes the event origin time. Readers are referred to Appendix B for technical details.

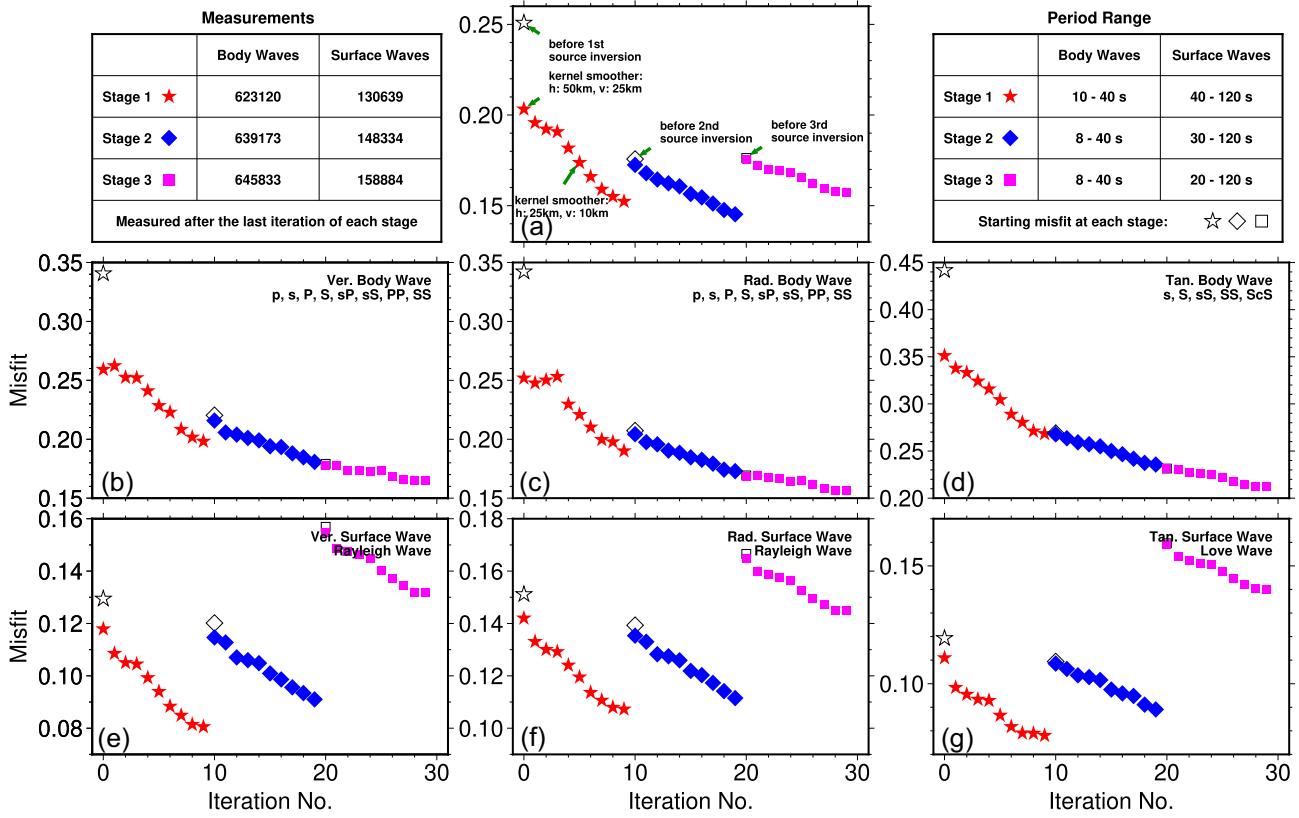


Figure 3. Misfit reduction curves in three stages (stage 1 in red stars, stage 2 in blue diamonds, and stage 3 in magenta squares). The open symbols represent the misfit before the source inversion. (a) Overall misfits. (b) Misfits for body waves on the vertical component. (c) Misfits for body waves on the radial component. (d) Misfits for body waves on the transverse component. (e) Misfits for Rayleigh waves on the vertical component. (f) Misfits for Rayleigh waves on the radial component. (g) Misfits for Love waves on the transverse component.

2.5 Inversion process

The inversion process is divided into three stages, and each stage includes source inversions followed by structure inversions using a preconditioned conjugate gradient method (Polak & Ribiere 1969). In the first stage, after 3 iterations of source inversion based on the initial model m_{00} , the source parameters remain unchanged, and we perform 5 iterations of structural inversion, using body waves filtered at 10–40 s and surface waves filtered at 40–120 s. Then we change the dimension of Gaussian smoother to sensitivity kernels from 50 km horizontally and 25 km vertically to 25 km horizontally and 10 km vertically for another 5 iterations. In the second stage, we start from the output model from the first stage and use higher-frequency body waves filtered at 8–40 s and surface waves filtered at 30–120 s. In this stage, after 3 iterations of source inversion, we perform another 10 iterations of structure inversion. In the third stage, we have also performed the same number of inversions as the second stage, but the surface waves are further filtered at 20–120 s. In sum, we run 9 iterations of source inversion and 30 iterations of structure inversion. Readers are referred to Appendix C for technical details about the structure inversion. We conclude each stage after 10 iterations for efficiency. Although the inversion does not necessarily converge at the end of each stage, the misfit reduction is small with additional iterations (Fig. S8). More importantly, the subsequent stage can further reduce the data misfit with the frequency band used in the previous stage. Fig. 3 shows that the inversions converge after 30 iterations where misfit reduction becomes minimal. The whole inversion process consumes about 4.5 million CPU hours on Stampede2.

3 RESULTS AND MODEL ASSESSMENT

3.1 Average V_P and V_S models

The final structure model is named EARA2024 (East Asia Radial Anisotropic Model 2024). Average V_P and V_S are calculated following Panning & Romanowicz (2006):

$$V_S^2 = \frac{2V_{SV}^2 + V_{SH}^2}{3} \quad (10)$$

$$V_P^2 = \frac{V_{PV}^2 + 4V_{PH}^2}{5}. \quad (11)$$

The model shows $\delta \ln V$ perturbations as large as 6 per cent. The $\delta \ln V_S$ perturbations are typically larger than the $\delta \ln V_P$ perturbations, but both show similar large-scale features (Figs 4 and S1). S-wave radial anisotropy $(V_{SH} - V_{SV})/V_S$ is about 3 per cent in the lithosphere and varies from -3 per cent to 3 per cent in the asthenosphere (Fig. 5). For the rest of this paper, we focus on the S-wave model because it is better resolved than the P-wave model due to the constraints from surface waves. As discussed by Tao *et al.* (2018), thin anomalies ~60 km above and below the 410- and 660-km discontinuities are likely artefacts due to the discontinuity topography model from S36ANI (Kustowski *et al.* 2008a) that is fixed during inversions. For the same reason, some high-velocity anomalies, interpreted as subducted slabs, seem to be split into two parts separated by a thin low-velocity layer in between, with the upper part in the MTZ and the lower part in the uppermost lower mantle. These are common artefacts shown in previous models using the same approach of handling discontinuities (Tao *et al.* 2018;

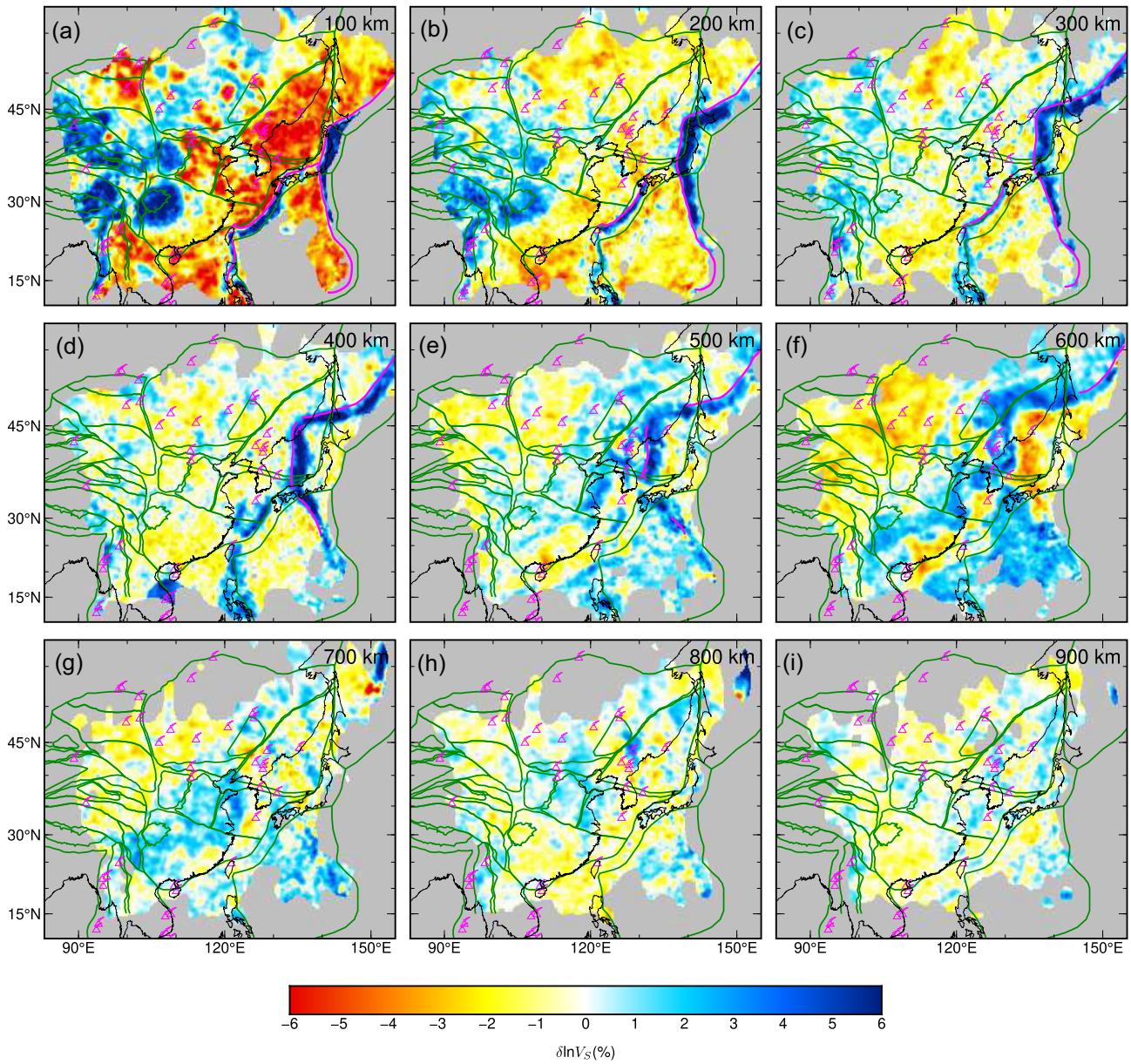


Figure 4. Mapviews of $\delta \ln V_S$ perturbation at various depths. The reference model is the 1-D average of the 3-D model. Regions with low resolution based on the point spread function resolution test are masked out. Magenta curves represent the slab depth contour at each depth based on the Slab2 model (Hayes *et al.* 2018). Other features are the same as Fig. 1.

Lei *et al.* 2020), and should not be overinterpreted. To avoid overinterpretation, we place semi-transparent bands ± 60 km around the 410- and 660-km discontinuities to mask potential artefacts due to the discontinuity topography.

The EARA2024 model reveals several important tectonic structures similar to previous models in this region (Chen *et al.* 2015; Tao *et al.* 2018). Fig. 4 shows high-velocity anomalies beneath the Kuril, Japan, Izu-Bonin and Ryukyu subduction zones, as well as the Sichuan Basin, eastern Tibet Plateau and Ordos Block in China at lithospheric depths. Compared to previous models, one obvious improvement is the high-resolution structure of the Japan slab beneath the Changbaishan volcanoes in the MTZ. Fig. 6 compares cross-sections of multiple tomography models EARA2024, FWEA18 and EARA2014 (Tao *et al.* 2018), EARA2014 (Chen *et al.* 2015), GLAD_M25 (Lei

et al. 2020), SPiRaL (Simmons *et al.* 2021) and GAP_P4 (Obayashi *et al.* 2013). In global models (GLAD_M25, SPiRaL and GAP_P4), the flat slab appears to be continuous in the MTZ. However, Tang *et al.* (2014) used regional traveltome tomography to first image a gap in the flat slab beneath the Changbaishan volcanoes. Over the course of developing FWI models EARA2014, FWEA18 and EARA2024, this gap structure gradually emerges and becomes more complex than previously thought. The EARA2024 model shows that there are at least 2 holes in the flat slab at 125°E and 129°E. Their implications will be discussed in Section 4.2.

Here we only highlight a few complex crustal and upper mantle structures beneath the Tibetan Plateau, Sichuan Basin and North China Craton (Fig. 7) while keeping thorough interpretations for future studies. The Tibetan Plateau is the world's highest and largest

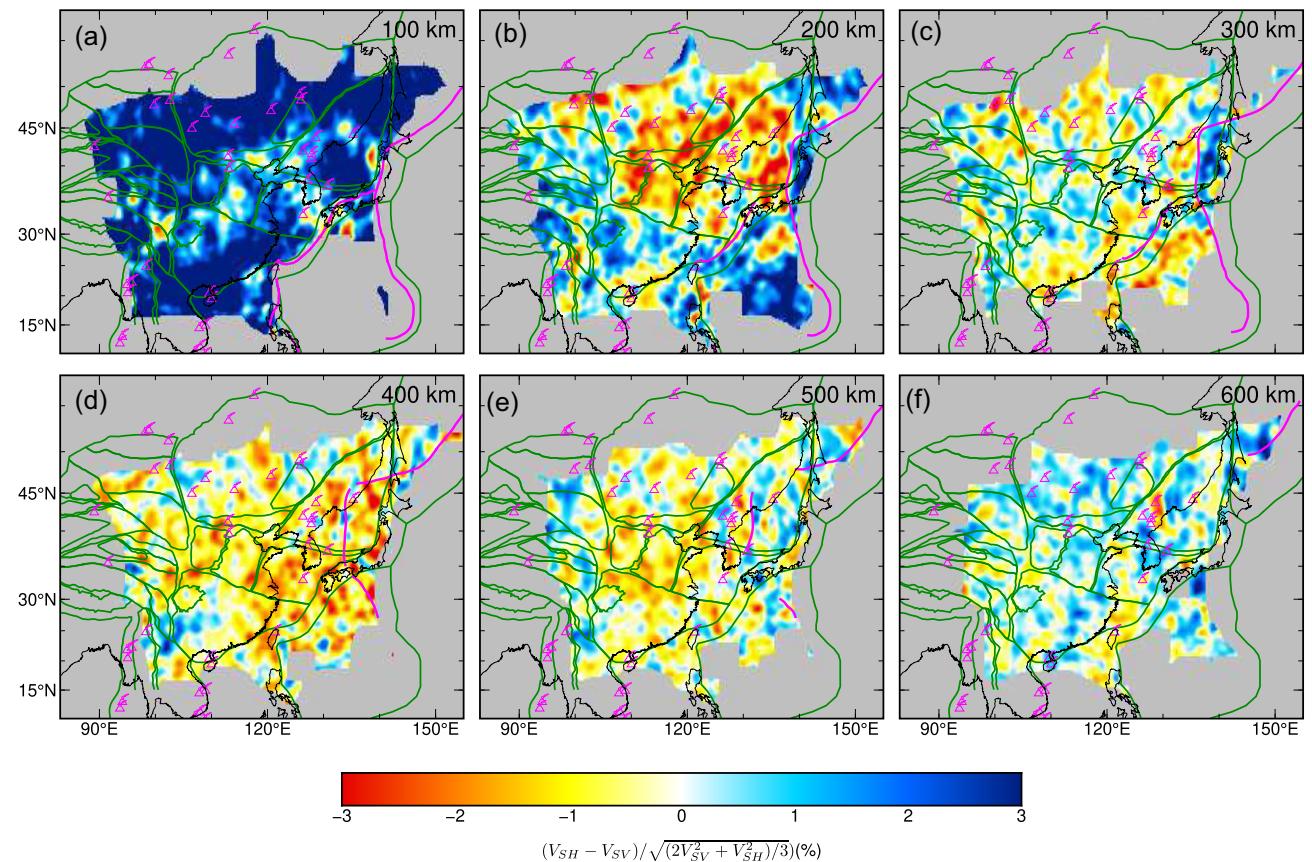


Figure 5. Mapviews of radial anisotropy at various depths. Other features are the same as Fig. 4.

plateau, resulting from the ongoing collision of the Indian and Eurasian Plates. The Sichuan Basin, located to the east of the plateau, is a sedimentary and tectonic basin. We observe high-velocity anomalies in the upper mantle beneath the Tibetan Plateau and Sichuan Basin (Figs 7a, b and c). The upper crust of the Sichuan Basin is characterized by high-velocity anomalies (Fig. S7). West of 100°E, we image a large high-velocity anomaly zone with a maximum $\delta \ln V_S$ perturbation of 4 per cent, potentially indicating a subducted plate reaching the MTZ and uppermost mantle (see discussions in Section 4.1). Moving eastwards, another high-velocity anomaly is observed below the Sichuan Basin at about 100 km depth, with a maximum $\delta \ln V_S$ perturbation of 6 per cent, representing the western edge of the South China Block and extending to a depth of around 300 km. We find significant low-velocity anomalies in the crust and uppermost mantle beneath the North China Craton (NCC, Figs 7d and e), a stable continental region that has experienced significant tectonic deformation and lithosphere thinning (Zhu *et al.* 2011). Beneath Huabei Plain, the eastern part of the NCC, the strong low-velocity anomaly at 80–150 km depths reveals a thinned lithosphere above it. In contrast, the lithosphere beneath the Ordos Block, indicated by a high-velocity anomaly, extends to at least 200 km depth.

In summary, our study provides a detailed model of the crustal and upper mantle structures beneath different tectonic features in East Asia and the northwestern Pacific. These observations offer valuable insights into the underlying geological processes in this region.

3.2 Model assessment

To assess the model, we first compute the phase time residual (ΔT) between synthetic (Iteration #30) and observed waveforms. Fig. 8 shows that the EARA2024 model has significantly reduced phase time residuals, particularly for surface waves, compared to the initial model. Additionally, the final model shows improved waveform similarity, as measured by normalized zero-lag cross-correlation coefficient (NZCC) and cross-correlation coefficient (CC). NZCC measures both phase time-shift and similarity, while CC only measures phase similarity. These results demonstrate that the final model provides a more accurate representation of the seismic structure in this region.

To further assess the model resolution, we conduct a point spread function (PSF) test (Fichtner & Trampert 2011b), a widely used method for resolution tests in FWI studies (Chen *et al.* 2015; Tao *et al.* 2018; Zhou *et al.* 2022). Compared to traditional checkerboard methods, which are computationally expensive for FWI, the PSF test provides a more efficient way to evaluate smoothing, distortion and parameter trade-offs in the inversion process. The PSF test results are shown in Fig. 9. For a checkerboard-shaped model perturbation to V_{SV} (Fig. 9a), we expect the PSF response for V_{SV} to resemble the input in high-resolution regions. The V_{SV} response (Fig. 9b) shows that the EARA2024 model has good resolutions in central-eastern China, Korea, the Japan Sea, the Yellow Sea and the East China Sea at 100 km depth. However, due to the event distribution, the V_{SV} PSF response is distorted in Mongolia, the South China Sea and the Izu–Bonin–Mariana Arc. At greater depths, the V_{SV}

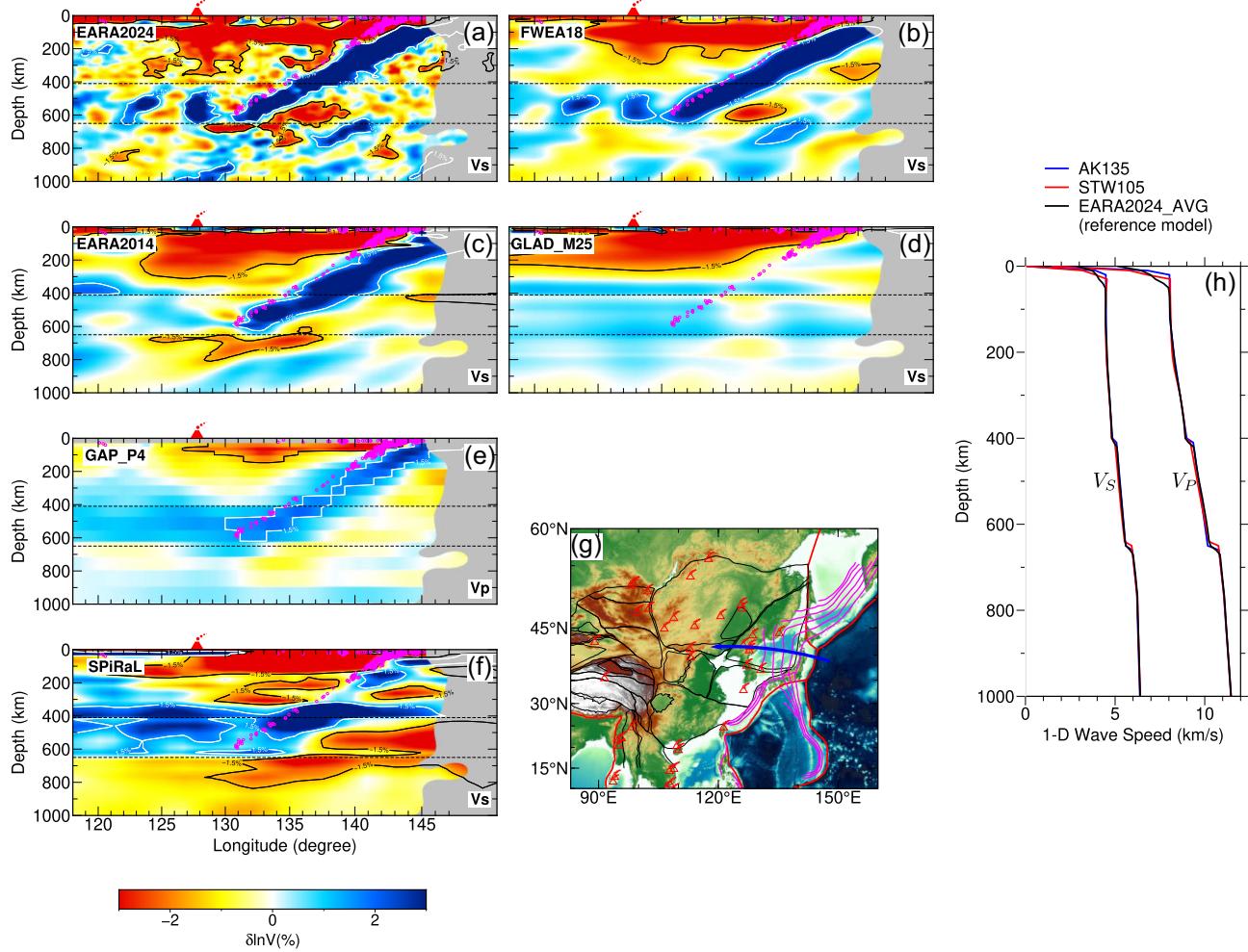


Figure 6. (a–f) Comparison of different models along the same cross-section across the Changbaishan volcanoes. The EARA2024, FWEA18 (Tao *et al.* 2018), EARA2014 (Chen *et al.* 2015), GLAD_M25 (Lei *et al.* 2020) and SPIRaL (Simmons *et al.* 2021) models present the $\delta \ln V_S$ perturbation with respect to EARA2024_AVG (averaged EARA2024), whereas the GAP_P4 model (Obayashi *et al.* 2013) displays the V_P perturbation with respect to its own reference model. The 410- and 660-km depths are shown as dashed lines. Earthquakes from 1964 to 2019 from the ISC-EHB catalogue (Weston *et al.* 2018) within 100 km of the cross-section are shown as magenta circles. Regions with the low resolution based on the point spread function resolution test are masked out. Semi-transparent bands ± 60 km around the 410- and 660-km discontinuities mask potential artefacts due to the discontinuity topography in (a). (g) The location of the cross-section is shown as the blue arrow. Other features are the same as Fig. 1. (h) Comparison of 1-D velocity models AK135 (Kennett *et al.* 1995), STW105 (Kustowski *et al.* 2008a) and EARA2024_AVG.

PSF response remains similar to that at 100 km but with smaller amplitudes. At depths greater than 900 km, the resolution is low, and the model should not be over-interpreted. Ideally, perturbations to V_{SV} should not affect V_{SH} , and thus the PSF response for V_{SH} should be 0. However, the results show non-zeros V_{SH} response (Fig. 9c), suggesting weak trade-offs between V_{SV} and V_{SH} , as they both rely on the same elastic moduli. In other words, a small portion of V_{SV} heterogeneity is artificially mapped into V_{SH} , slightly biasing radial anisotropy. Similarly, Fig. 9(d) shows the bulk velocity V_C response being nearly 0 everywhere (Fig. 9d), suggesting that the inversion can perfectly distinguish V_S and V_C . More technical details of the PSF tests are presented in Appendix D.

Waveform comparisons between observed waveforms, synthetic waveforms based on the initial structure and source models, and synthetic waveforms based on the final EARA2024 and source models also demonstrate the improvements. Fig. 10 shows an example of such comparisons for a 21-km-deep event on 7 May 2008 (GCMT ID: 200805071602A) in the Japan subduction zone recorded by a

CEArray station (JX.JIJ) in southeastern China. The results show that the final models provide better waveform fitting than the initial models, particularly for surface waves. Fig. S6 offers another example of fitting triplicated waveforms with the final models. Additionally, Figs S9 and S10 provide independent validation of our models using recent earthquakes (GCMT ID: 202307011438A and 202306220124A) which are not used for inversions. The waveforms of these earthquakes at several GSN stations are better fitted with our final models.

Since the final models integrate both source and structural inversions, it is crucial to distinguish the improvements attributable to source inversions and structural adjustments. Fig. 8 compares histograms of ΔT , NZCC, and CC for three different cases: (i) using the initial structural and source models; (ii) using the initial structural model with inverted source models after three iterations of the source inversion updates in the first stage of inversions and (iii) using the final structural and source models. This approach highlights that while the source inversions mainly adjust phase time-shifts ΔT ,

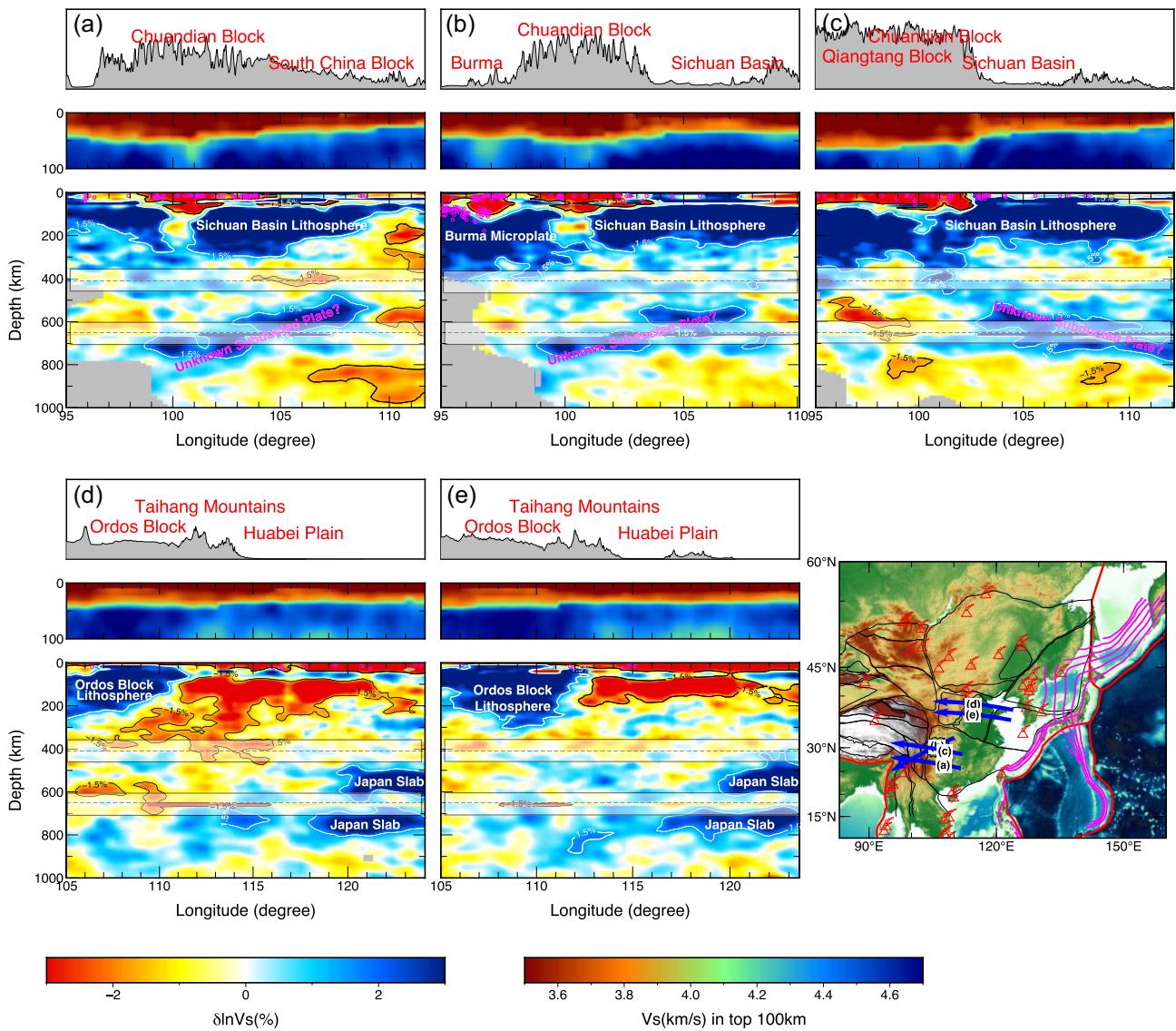


Figure 7. Cross-Sections of $\delta \ln V_S$ perturbation beneath Eastern Tibet (a–c) and North China Craton (d–e). The locations of the cross-sections are marked as thick blue arrows on the map. The annotations on the map are similar to the previous figures. Each cross-section includes a topography panel on the top with a 60-times vertical exaggeration, an absolute velocity panel from the surface to 100 km depth, and a panel of $\delta \ln V_S$ perturbation with respect to EARA2024_AVG (the 1-D average of the 3-D model) down to 1000 km depth. Earthquakes from the ISC-EHB catalogue (Weston *et al.* 2018) within 100 km of the cross-section are shown as magenta circles. White and black curves mark the 1.5 per cent contours. Regions with low resolution based on the point spread function resolution test are masked out. Semi-transparent bands ± 60 km around the 410- and 660-km discontinuities mask potential artefacts due to the discontinuity topography.

structural inversion plays a significant role in improving NZCC and CC. Additionally, Fig. S11 offers a comparative analysis between observed waveforms and synthetics produced for the three different cases mentioned above. This comparison shows the collaborative improvements for the waveform fitting through both the source and structure inversions.

Moreover, we use phase time residual ΔT as a key indicator of source parameter variations, as ΔT is mostly influenced by the change of event origin time in the source inversions. For body waves, the source inversions greatly reduce the mean and standard deviation of ΔT . For surface waves, which are more sensitive to the crustal and mantle structures, our analysis leads to insignificant improvement of the mean and standard deviation of ΔT in the source inversions. Therefore, the structure inversions should have contributed most of the improvements for the crustal and mantle

structures. Therefore, we are confident with the EARA2024 model in capturing seismic structures beneath East Asia and northwestern Pacific subduction zones.

4 DISCUSSION

4.1 Slab morphology of northwestern pacific subduction zones

The northwestern Pacific region hosts a number of subduction zones. The Pacific Plate subducts westwards beneath the Eurasian Plate along the Kuril, Japan Trenches and beneath the Philippine Sea Plate along the Izu-Bonin, Mariana and Yap Trenches, whereas the

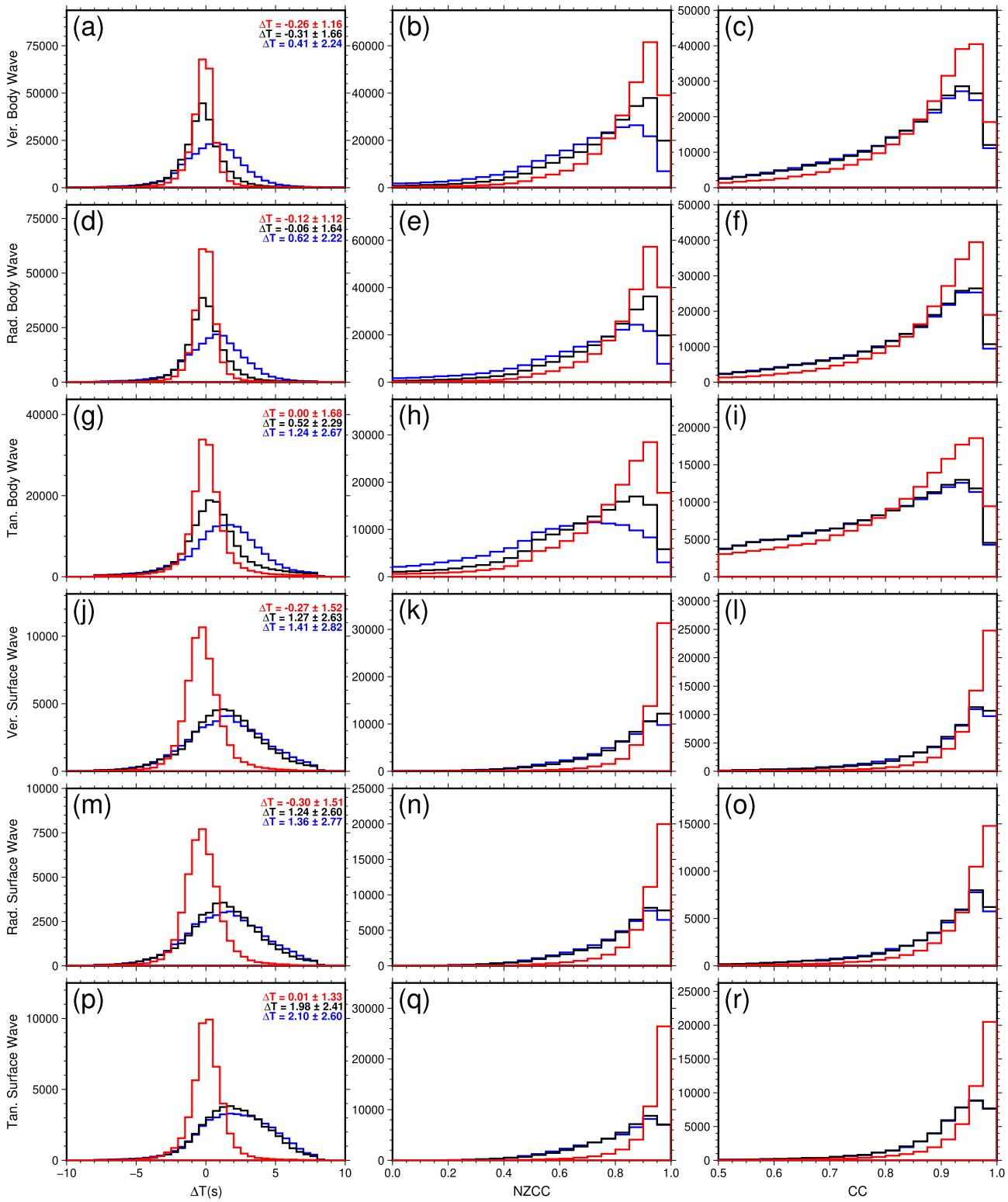


Figure 8. Histograms of time-shifts, normalized zero-lag cross-correlation coefficients (NZCC) and cross-correlation coefficients (CC) for all six measurement categories. Blue histograms represent the initial structural and source models, black histograms represent the initial structural model with inverted source models after three iterations in the first stage of source inversions, whereas red ones represent the final structural and source models. For time-shifts, also shown the mean and standard deviation.

Philippine Sea Plate subducts beneath the Eurasian Plate northwestwards along the Nankai, Ryukyu, Malina and Philippine Trenches. Numerous earthquakes and arc volcanic activities occur in these subduction zones. Fig. 11 presents vertical cross-sections of the

$\delta \ln V_S$ perturbation for the Kuril, Japan, Izu-Bonin, Nankai and Ryukyu subduction zones. We also create a 3-D slab model in this region by identifying all fast- V_S anomalies of > 1.5 per cent (Fig. 12 and [Supplementary Video](#)).

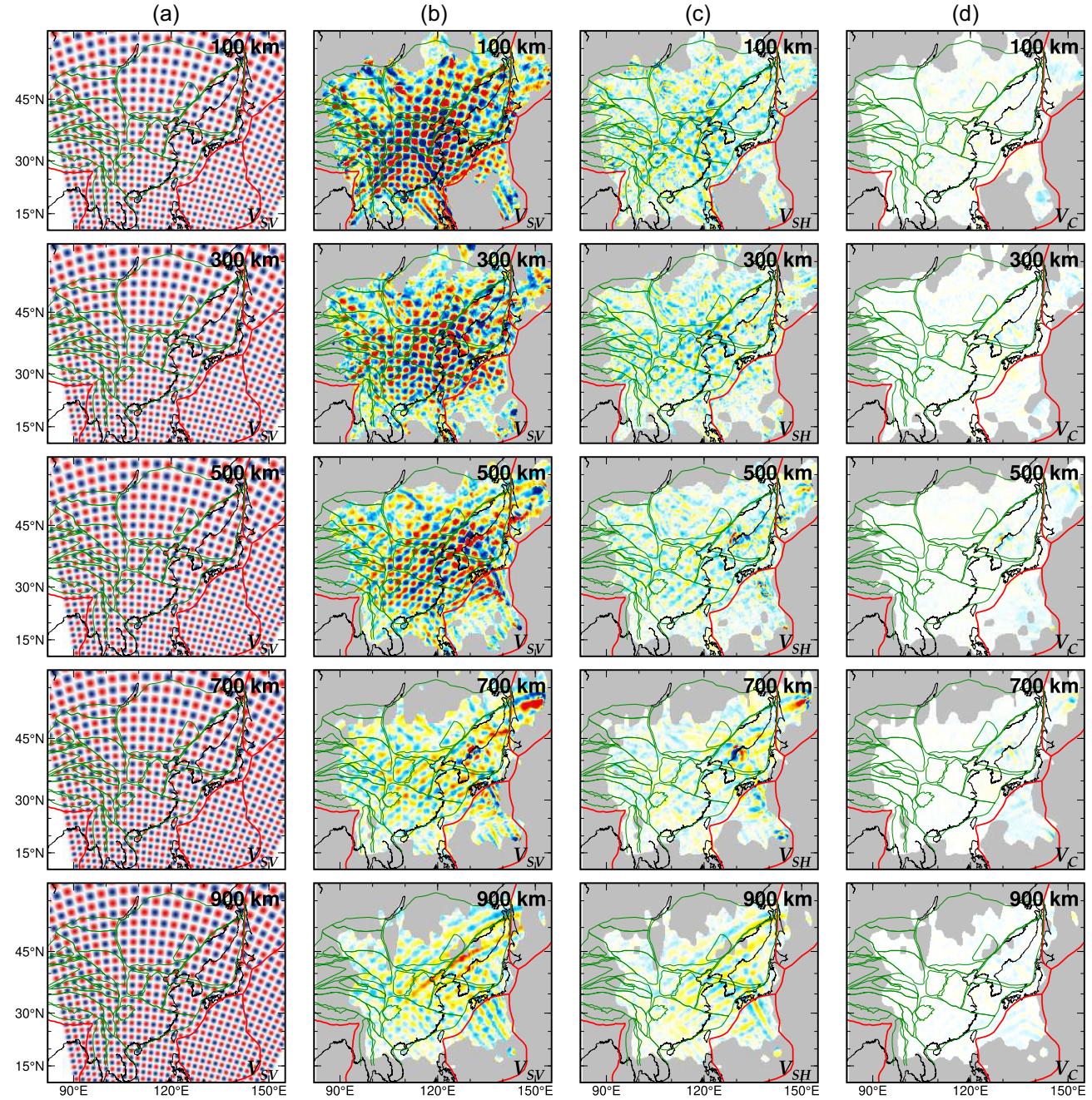


Figure 9. Point spread function (PSF) test for V_{SV} perturbation at various depths. (a) The input model perturbation applied to V_{SV} for the final model. (b) The PSF response for V_{SV} , which should mimic the input model in an ideal case. (c) The PSF response for V_{SH} , which should be null in an ideal case. (d) The PSF response for bulk velocity V_C , which should be null in an ideal case.

The Kuril, Japan and Izu-Bonin slabs are the subducted Pacific Plate along the Kuril, Japan and Izu-Bonin Trenches, respectively. The Kuril slab is characterized by a nearly constant dipping angle of $\sim 36^\circ$ and a significant maximum 8 per cent $\delta \ln V_S$ perturbation reaching the 660-km discontinuity (Fig. 11a). A substantial low-velocity anomaly zone above 200 km depth represents the vast mantle wedge, with a maximum $\delta \ln V_S$ perturbation of 4 per cent. The Japan slab, characterized by high-velocity anomalies, subducts with a shallower dipping angle of $\sim 25^\circ$ in the upper mantle and is deflected and extends horizontally in the mantle transition zone (MTZ) for at least 800 km. There are two holes within the slab at $\sim 125^\circ E$ and $129^\circ E$ (Figs 11c). Their implications and relationship

to Changbaishan volcanoes will be discussed in Section 4.2. The Japan slab may extend further to $\sim 115^\circ E$, with a big slab hole between $\sim 120^\circ E$ and $\sim 123^\circ E$. Fig. 11(b) shows the transition between the Kuril slab to the Japan slab, where the subducted Pacific Plate starts to flatten in the MTZ. It is worth noting that the Japan slab beneath the East China Sea (Fig. 11g) is much thicker than that beneath Northeast China (Fig. 11c), extending into the uppermost lower mantle.

The mantle wedge above the Japan slab shows a low $\delta \ln V_S$ anomaly of as much as ~ 6 per cent, reaching depths between 200 and 300 km. Another low-velocity anomaly is detected beneath the Japan slab in the MTZ, with a peak $\delta \ln V_S$ anomaly of ~ 2 per cent.

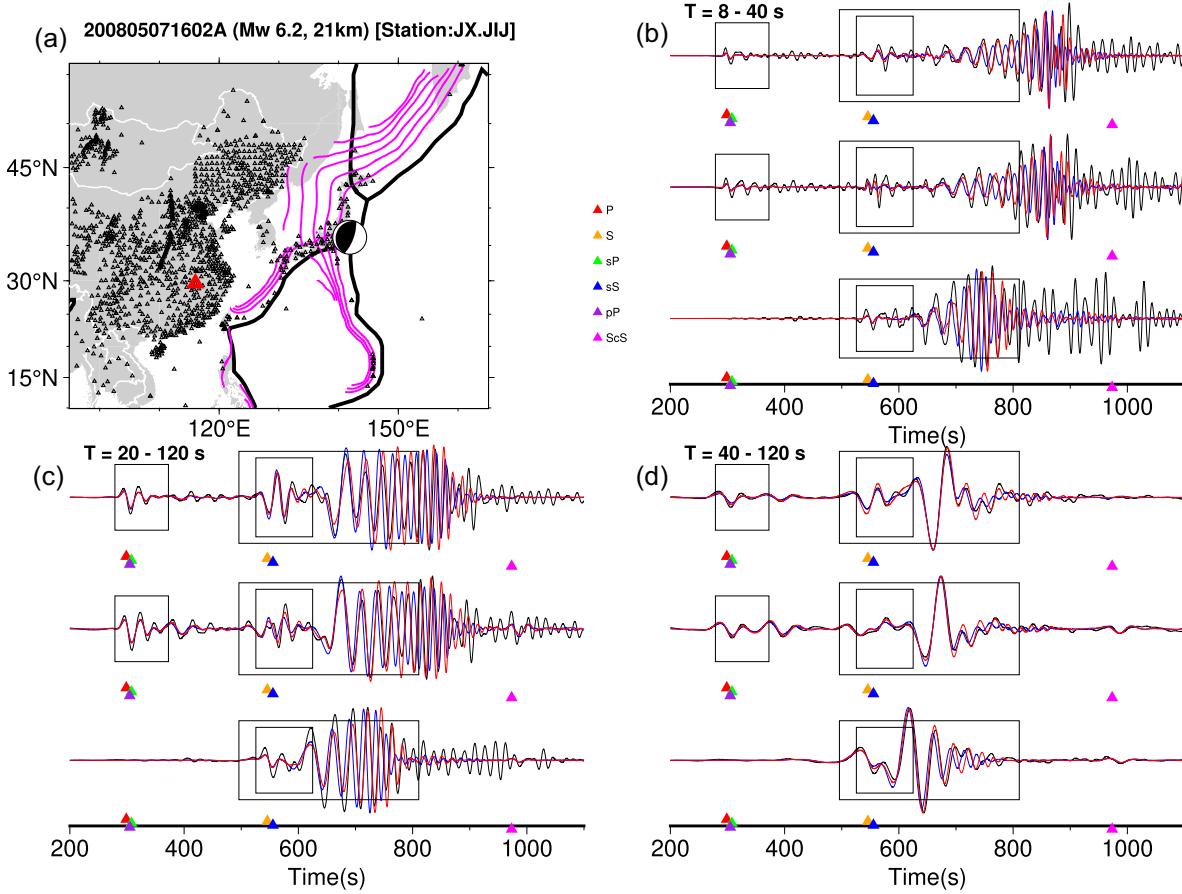


Figure 10. Waveform comparisons for a selected earthquake (GCMT ID: 200805071602A) recorded by a station indicated by the red triangle in (a). All other stations used in the inversion are shown as black triangles for reference. Other features are the same as Fig. 2(b). (b–d) Waveform comparisons between observations (black), synthetics based on the starting model (blue) and synthetics based on the final model (red) at different frequency bands for various seismic phases. Black boxes indicate the measurement windows based on the theoretical phase arrival times predicted by the AK135 model (Kennett *et al.* 1995).

Because the major part of this anomaly is well above the 660-km discontinuity, we believe it is a real feature and may indicate sub-slab partial melting that is entrained by subduction from the shallow depths (Kawakatsu *et al.* 2009; Wang *et al.* 2020).

The Izu-Bonin slab, the Pacific Plate subducted beneath the Philippine Sea Plate along the Izu-Bonin Trench, is characterized by a maximum $\delta \ln V_S$ anomaly greater than 6 per cent, and has a steeper dipping angle of $\sim 40^\circ$ in the upper mantle. The slab appears flat but heavily deformed in the mid-mantle, with pieces of high-velocity anomalies at varying depths in the MTZ and uppermost lower mantle (Figs 11d and e). The slab deformation is particularly clear in the southern part of the Izu-Bonin slab, although we lose the resolution of the Mariana slab to the further south. Our observations of high-velocity anomalies are consistent with a previous study (Zhang *et al.* 2019) that interpreted the anomalies as evidence of a single contiguous slab undergoing tearing.

The Philippine Sea Plate is much younger and thus warmer than the Pacific Plate in this region. As the Philippine Sea Plate subducts beneath the Eurasian Plate along the Nankai Trough and Ryukyu Trench, the Nankai and Ryukyu slabs show distinct characteristics. The Nankai slab, one of the warmest slabs globally, subducts with a dipping angle of $\sim 20^\circ$ to 200 km depth, with a maximum $\delta \ln V_S$ perturbation of only 1 per cent (Fig. 11f). Consequently, the Nankai slab does not appear in our 3-D slab morphology model (Fig. 13), which shows the 1.5 per cent isosurface of high-velocity anomalies.

The Ryukyu slab subducts with a nearly constant dipping angle of $\sim 50^\circ$ and extends to ~ 500 km depth (Fig. 11g), consistent with previous studies (Li *et al.* 2008; Tao *et al.* 2018). Wu *et al.* (2016) interpret the thick high-velocity anomaly in the MTZ and uppermost lower mantle beneath the Yellow Sea as the subducted Philippine Sea Plate overlying the Pacific Plate. However, our model shows a ~ 300 -km-wide gap between the Ryukyu slab beneath the East China Sea and the Pacific Plate beneath the Yellow Sea. Therefore, the thick high-velocity anomaly beneath the Yellow Sea is more likely the Japan slab ponding around the 660-km discontinuity, as suggested by Tao *et al.* (2018).

Similar to Tao *et al.* (2018), we also image a NE–SW striking high-velocity anomaly in the MTZ beneath the South China Block. It appears to connect to the Japan slab beneath the Yellow Sea. However, this connection may be an artefact due to low resolution, as there is no evidence suggesting that the subducted Pacific Plate reaches this far to the southwest. Tao *et al.* (2018) speculate that this anomaly is related to the subduction of a proto-South China Sea, although a plate reconstruction model places this structure further to the east (Wu *et al.* 2016).

4.2 Intraplate volcanoes

Intraplate volcanoes are intriguing because they cannot be easily explained by the classic theory of plate tectonics. Some of them,

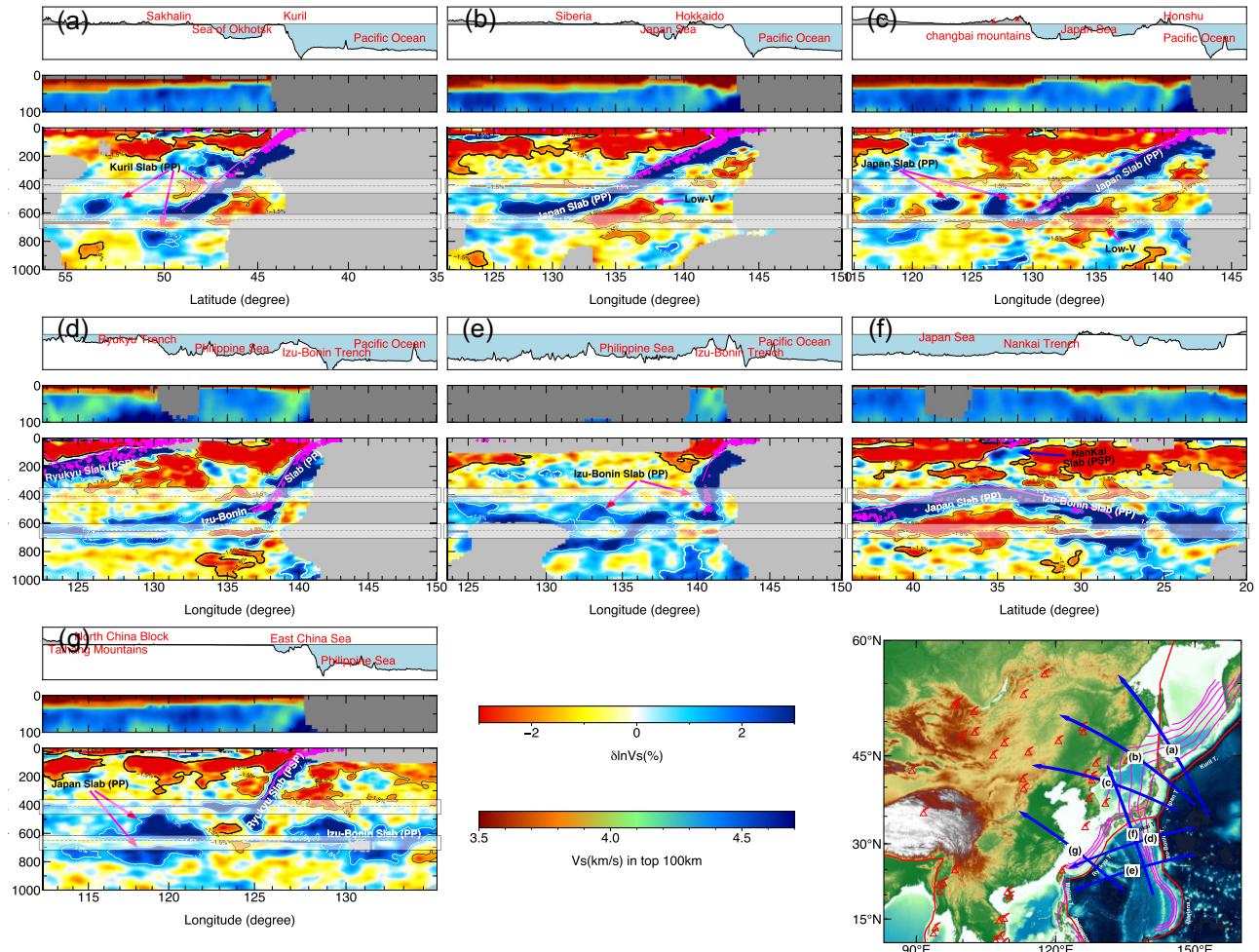


Figure 11. Cross-sections of $\delta \ln V_S$ perturbation for northwestern Pacific subduction zones, plotted similar to Fig. 7. The reference model is EARA2024_AVG. Acronyms: PP: Pacific Plate, PSP: Philippine Sea Plate.

for example the Hawaiian volcanoes, are related to mantle plumes originating from the lower mantle (e.g. Wolfe *et al.* 2009; French & Romanowicz 2015). Some of them, for example Bermuda Island, are rooted in the MTZ (e.g. Mazza *et al.* 2019). However, the origins of other intraplate volcanoes, especially those on continents, for example the Yellowstone volcano, are under debate (Nelson & Grand 2018; Zhou 2018; Zhou *et al.* 2018). Here we focus on four intraplate volcanic fields, Datong–Fengzhen, Tengchong, Changbaisan and Hainan in East Asia, and try to explore their origins. These volcanoes are far away from any active subduction zones and thus are unlikely to be related to arc or backarc volcanism. They are also distant from any known divergent plate boundaries.

The Datong–Fengzhen volcanoes in northern China predominantly consist of Cenozoic basaltic lava flows and pyroclastic deposits (Xu *et al.* 2005; Smithsonian-Institution 2023). Lei (2012) image the upper mantle beneath the North China Craton with teleseismic wave traveltimes tomography and suggest that Datong–Fengzhen volcanoes are generated by hot materials upwelling from the lower mantle. In contrast, Kim *et al.* (2021) image the S-wave velocity structure of East Asia with teleseismic traveltimes tomography and suggest that these volcanoes are generated by slab-induced dehydration of the hydrous mantle transition zone. Moreover, by imaging mantle discontinuities beneath eastern North China with

receiver functions, Zuo *et al.* (2020) find a normal MTZ thickness in this region, ruling out the possibility of an abnormally hotter mantle plume. Instead, they suggest that the mantle return flow induced by lithospheric delamination beneath the southern Great Xing'an Range generates decompression melting to feed these volcanoes. Our results show a low-velocity zone immediately below the Datong–Fengzhen volcanoes, extending to ~ 400 km depth (Figs 13a and S5a). The origin of the Datong–Fengzhen volcanoes thus appears to be not a mantle plume from greater depths. In addition, the subducted Pacific Plate, characterized by high-velocity anomalies in the MTZ and uppermost lower mantle beneath the Bohai Sea, is hundreds of kilometres away from the low-velocity anomalies beneath Datong–Fengzhen volcanoes. Instead, we observe high-velocity anomalies in the upper mantle and MTZ beneath the southern Great Xing'an Range, possibly indicating lithospheric delamination. Therefore, we propose that the mantle return flow induced by the delaminated lithosphere is the most plausible cause for the Datong–Fengzhen volcanoes, as suggested by Zuo *et al.* (2020).

The Tengchong volcanic field in southwestern China is characterized by Cenozoic basalts and andesites (Wang *et al.* 2007; Smithsonian-Institution 2023). Previous traveltimes tomography images (Wei *et al.* 2012; Huang *et al.* 2015; Lei & Zhao 2016; Lei *et al.* 2019) show a high-velocity anomaly, interpreted as the subducted Burma microplate beneath Tengchong, suggesting that the

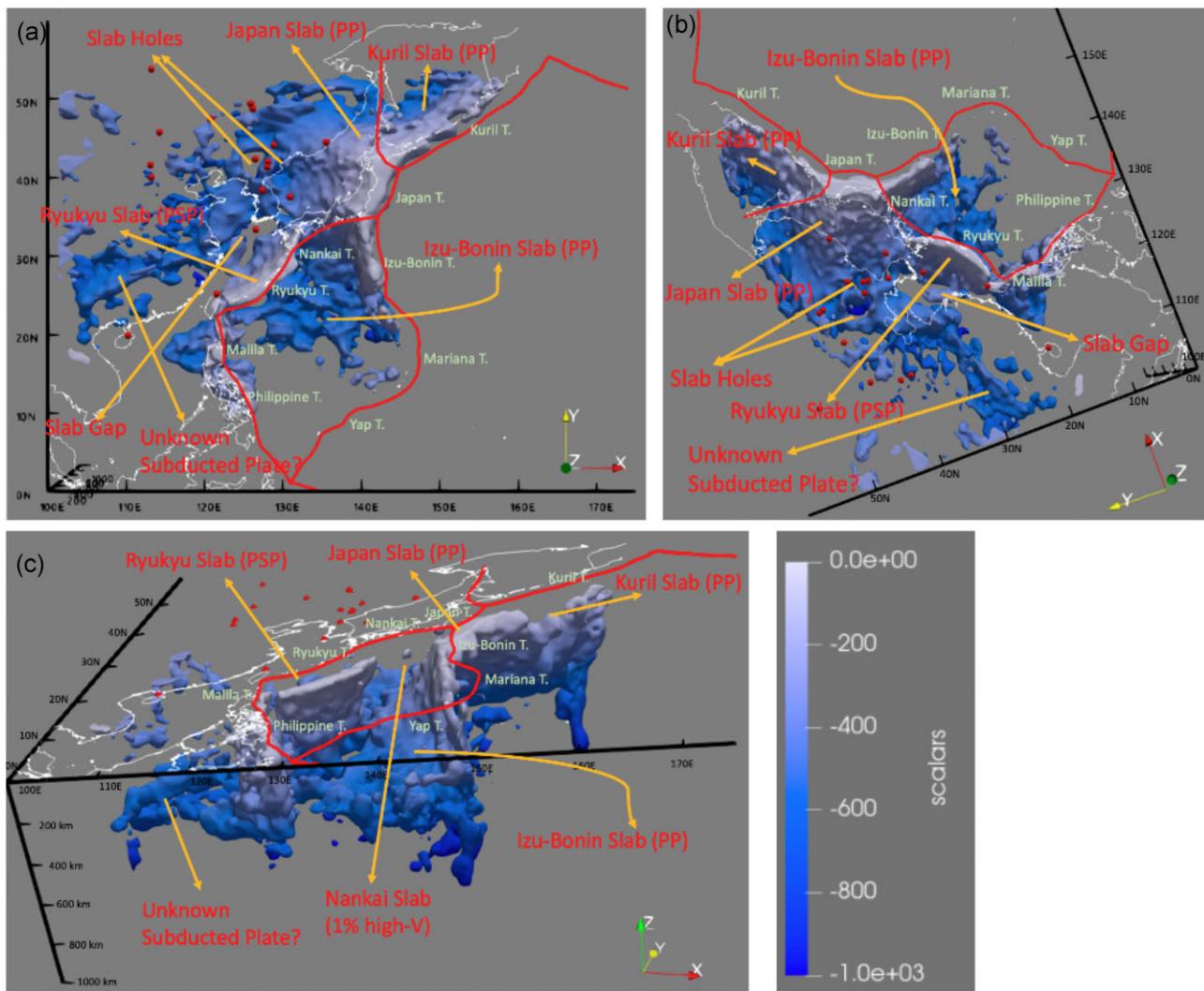


Figure 12. 3-D images of slab morphology in the northwestern Pacific subduction zones. The slabs are represented by the 1.5 per cent isosurface of high-velocity anomalies and colour-coded by depth. Red thick curves indicate the plate boundaries from the NUVEL-1 model (DeMets *et al.* 1990), white curves show the coastline, and red dots represent intraplate volcanoes. (a) Mapview. (b) Viewing from the northwest. (c) Viewing from the southeast.

volcanism is related to slab dehydration and partial melting in the so-called big mantle wedge (BMW). However, our results show no strong high-velocity anomalies in the MTZ beneath Tengchong, and the Burma slab subducts with a high dip angle to only the 410-km discontinuity northwest of Tengchong (Fig. 13b). Our observation is consistent with the global tomography model UU-P07 (Amaru 2007). Therefore, we conclude that the Tengchong volcanic field is not caused by slab dehydration of the Burma slab. Instead, we suspect the volcanism is related to small-scale mantle convection and upwelling between the subducting Burma slab and the 200–300 km thick lithosphere of the Sichuan Basin.

The Changbaishan volcanoes, located in northeastern China along the border with North Korea, are primarily composed of trachytic and basaltic lava flows, as well as pyroclastic deposits (Zhang *et al.* 2018; Smithsonian-Institution 2023). These stratovolcanoes have been active since the late Pleistocene epoch and are renowned for their large-scale eruptions. Early seismic traveltimes tomography studies of Northeast Asia, such as Huang & Zhao (2006) and Lei & Zhao (2005), image a flattened slab in the MTZ, suggesting the dehydration of the stagnant Pacific slab releases water to the so-called big mantle wedge (BMW) and triggers partial melting

in the upper mantle beneath these volcanoes. However, this BMW hypothesis cannot explain the specific location of these volcanoes. In contrast, recent tomography studies empowered by regional permanent and temporal seismic stations (e.g. Tang *et al.* 2014; Tao *et al.* 2018) have uncovered a slab gap in the MTZ, suggesting that the subslab asthenosphere entrained by subduction escape through the slab gap and upwells to trigger decompression melting to feed the Changbaishan volcanoes. A new receiver-function study (Wang *et al.* 2020) observes hydrous melting atop the MTZ and hydrous slab crust segments in the MTZ, supporting the BMW hypothesis. However, Li *et al.* (2023) compares seismic waveform modelling with an electrical resistivity model in this region, arguing that seismic anomalies atop the MTZ are likely caused by moduli reduction during phase transformation rather than hydrous melting. Our results show low-velocity anomalies immediately beneath the Changbaishan volcanoes, extending into two slab holes in the MTZ (Fig. 13c). This observation is consistent with that from Tang *et al.* (2014) and Tao *et al.* (2018), suggesting that the Changbaishan volcanoes are fed by decompression melting from upwelling materials through the slab holes. The origin of the upwelling materials remains uncertain. Since there is no surface geological evidence of a

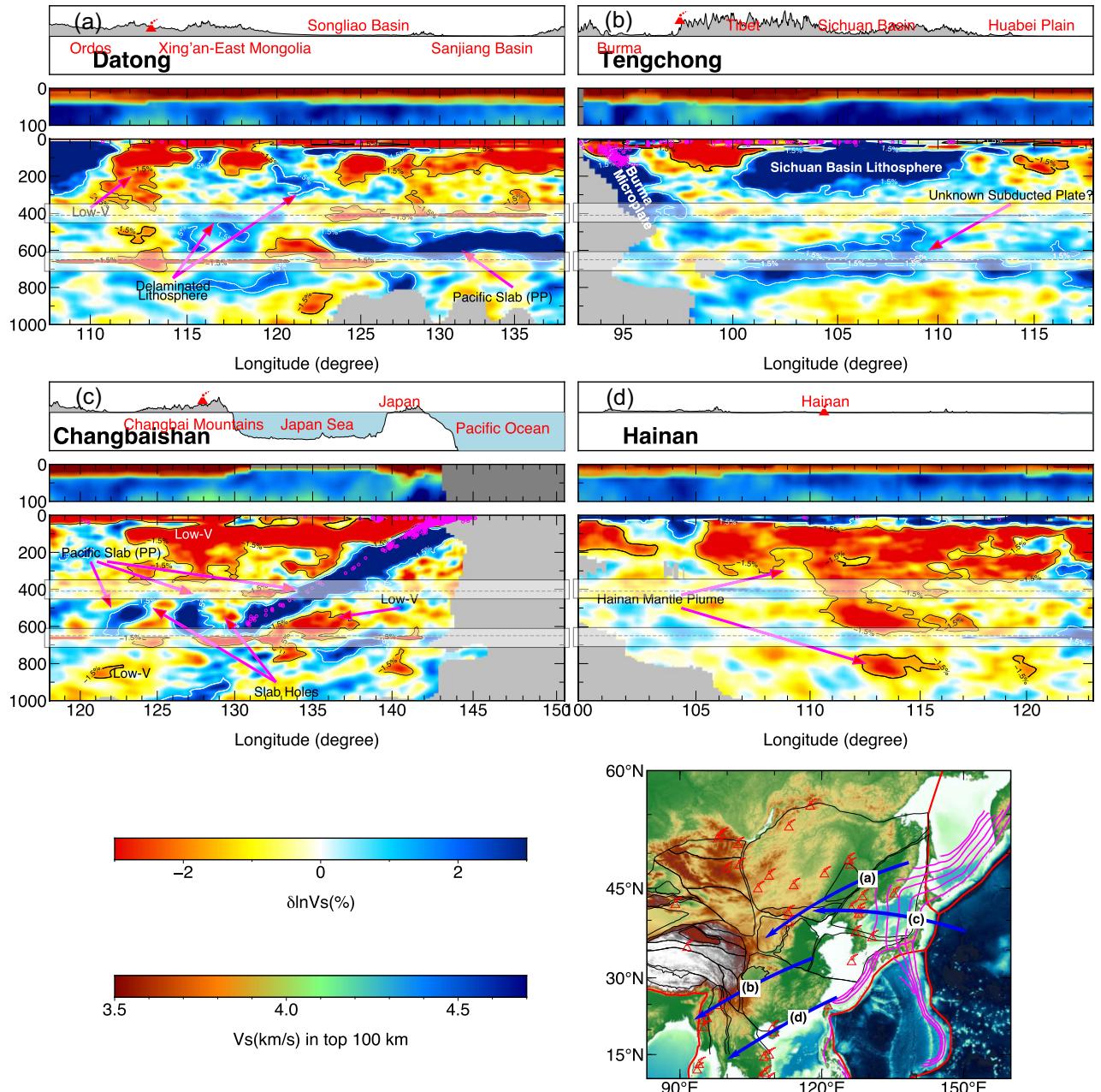


Figure 13. Cross-sections of $\delta \ln V_s$ perturbation for four intraplate volcanic fields in East Asia, plotted similar to Fig. 7. The reference model is EARA2024_AVG. (a) Datong-Fengzhen volcanoes. (b) Tengchong volcanic field. (c) Changbaishan volcanoes. (d) Hainan volcanic field.

deep-rooted mantle plume, Tang *et al.* (2014) hypothesize that the subslab asthenosphere is entrained by the subduction of the Pacific Plate and escapes through the slab holes. Our results show a subslab low-velocity anomaly in the MTZ at $\sim 135^\circ\text{E}$, possibly indicating the entrained asthenosphere. On the other hand, we cannot rule out another possibility that the upwelling materials come from the lower mantle, as evidenced by the low-velocity anomalies immediately below the slab holes extending down to at least 1000 km depth. To verify the source of the upwelling materials, more tomographic images of the lower mantle in this region are needed.

The Hainan volcanic field in the Leizhou Peninsula and Hainan Island in southern China is covered by late Cenozoic basalts (Zhou

& Armstrong 1982; Smithsonian-Institution 2023). An early global finite-frequency tomography model by Montelli *et al.* (2004) proposed a mantle plume in this region. Later regional travelttime tomography studies with higher resolution (Huang & Zhao 2006; Lei *et al.* 2009) showed more details of this Hainan mantle plume. Furthermore, a receiver function study imaged a thinner-than-normal MTZ beneath this region, suggesting that the Hainan mantle plume is about 170–200 °C hotter than the surrounding mantle (Wei & Chen 2016). In our study, Fig. 13(d) shows strong low-velocity anomalies (more than a 2 per cent reduction for V_s) extending from the surface to the lower mantle, suggesting a deep-rooted Hainan mantle plume.

5 CONCLUSIONS

We present a new 3-D radially anisotropic seismic velocity model EARA2024 of the crust and mantle down to 1000 km depth beneath East Asia and northwestern Pacific subduction zones using full waveform inversion (FWI) tomography. Adjoint tomography based on a spectral element method is applied to a large dataset of about 640 000 body-wave and 148 000 surface-wave measurement windows. The new model provides high-resolution images of the subducted plates, continental lithosphere, asthenosphere and mantle transition zone (MTZ) in this region. Compared to the previous FWI models of East Asia (FWEA18 and EARA2014), EARA2024 shows more small-scale velocity anomalies, better delineating the continental lithosphere thickness variations, slab morphology at different depths and mantle upwelling. In particular, the shear wave velocity structures of EARA2024 provide the following new insights into slab morphology and origins of intraplate volcanoes in this region.

(i) The Japan slab is flattened in the MTZ west of the Japan Sea. At least two slab holes, ~200 km wide each, are imaged in the subducted Pacific Plate beneath the Changbaishan volcanoes. These holes provide the pathway for mantle upwelling that feeds the intraplate volcanoes (Figs 6, 11c and 13c).

(ii) The Izu-Bonin slab is heavily deformed and fragmented, with pieces of subducted Pacific Plate residing in the MTZ or uppermost lower mantle (Figs 11 d and e).

(iii) The Philippine Sea Plate subducts along the Ryukyu Trench and reaches 500 km depth. Although both are in the MTZ, the subducted Philippine Sea Plate beneath the East China Sea is not attached to the subducted Pacific Plate beneath the Yellow Sea (Fig. 11f).

(iv) Beneath the Datong volcanoes, low-velocity anomalies are imaged from the lithosphere down to ~400 km depth. The mantle upwelling is more likely induced by lithospheric delamination beneath the southern Great Xing'an Range (Fig. 13a).

(v) The low-velocity anomalies beneath the Tengchong volcanic field are confined within the top 200 km. There is no obvious connection between the Tengchong volcano and the subducted Burma microplate. The intraplate volcanism is more likely to be triggered by small-scale asthenospheric convection and upwelling between the subducting Burma slab and the 200–300 km thick lithosphere of the Sichuan Basin (Fig. 13b).

(vi) We clearly image the Hainan mantle plume, characterized by low-velocity anomalies extending from the surface to the lower mantle (Fig. 13d).

ACKNOWLEDGMENTS

We appreciate the insightful discussions and support from Allen McNamara, Jeffrey Freymueller, Jeroen Tromp and Rongrong Wang. We also thank the editor Ebru Bozdağ, the reviewers Kai Tao and Arthur Rodgers for their constructive comments. This work was supported by the National Science Foundation grants EAR-1802247 and EAR-1942431 to M.C. and S.S.W. and their start-up funds at Michigan State University. Y. Kim acknowledges support from the NRF grant funded by the Korean government (MSIT, 2022R1A5A1085103). We appreciate the Extreme Science and Engineering Discovery Environment (XSEDE, supported by NSF Grant ACI-1053575) for providing high-performance computing

resources through XSEDE Grant TG-EAR140030 on the Stampede2 supercomputer hosted by the Texas Advanced Computing Center (TACC). We acknowledge the transition from XSEDE to ACCESS (Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support) and their continued commitment to connecting researchers with valuable cyberinfrastructure resources. We also thank the Institute for Cyber-Enabled Research (ICER) at Michigan State University for providing additional computational resources for data processing.

Lastly, we would like to dedicate this paper to the memory of our esteemed co-author Dr Min Chen who passed away in July 2021. Min was an extraordinary seismologist who advanced our understanding of plate tectonics and Earth's interior structure with high-resolution seismic imaging. Min's enthusiasm for life and research, unwavering commitment to diversity, equity and inclusion, and generosity in supporting colleagues and friends around the world will always be remembered and cherished.

SUPPORTING INFORMATION

Supplementary data are available at [GJIRAS](#) online.

[supplementary.pdf](#)

[slab3d.mp4](#)

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

DATA AVAILABILITY

All codes associated with this paper are openly accessible and based on open-source resources. The spectral-element seismic wave simulation code package SPECFEM3D_GLOBE utilized in this study is available through the Computational Infrastructure for Geodynamics (CIG) (<https://geodynamics.org/>). We used open-source Python packages for data processing and visualization, including ObsPy (<https://github.com/obspy/obspy/>), PyGMT (<https://github.com/GenericMappingTools/pygmt>) and Pyasdf (<https://github.com/SeismicData/pyasdf>). Parts of the broad-band seismic data were acquired from the EarthScope/IRIS Data Management Center (<https://ds.iris.edu/ds/nodes/dmc/>) under network codes 1U, 2F, JP, XR, II, G, RM, CB, PS, MM, AD, 8B, K5, IU, ZV, IM, YC, MY, D0, KG, KS, XF, SS, TM, Z6, BO, IN, TW, XI, YM, MI, HK, YP, IC, YL, XL, X4 and Z1. Other seismic data were obtained from the China Earthquake Administration Array (CEArray, Zheng *et al.* 2009), F-net (Okada *et al.* 2004) at <https://www.fnet.bosai.go.jp/> and the Korean Seismic Network at <https://necis.kma.go.kr>. We acknowledge that a large portion of the data used in this study, including CEArray, are not openly available due to regulatory constraints. To acquire CEArray data, readers should contact the China Earthquake Administration directly, understanding that an approval process is needed. We uploaded the CEArray waveform data used for plotting the figures to <https://osf.io/zgv37/>. The Global Centroid Moment Tensor (GCMT) project data (Dziewonski & Anderson 1981; Ekström *et al.* 2012) can be accessed at <https://www.globalcmt.org/>. ISC-EHB catalogue (Weston *et al.* 2018) can be accessed at <http://isc-mirror.iris.washington.edu/isc-ehb/>. The EARA2024 model is available at the EarthScope/IRIS Earth Model Collaboration (EMC) repository (<https://doi.org/10.17611/dp/emc.2024.eara2024.1>).

REFERENCES

- Aki, K.** & Lee, W.H.K., 1976. Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model, *J. geophys. Res.*, **81**(23), 4381–4399.
- Allégre, C.J.** et al., 1984. Structure and evolution of the Himalaya–Tibet orogenic belt, *Nature*, **307**(5946), 17–22.
- Amaru, M.L.**, 2007. *Global Travel Time Tomography with 3-D Reference Models*, Vol. **274**, Utrecht Univ.
- Bao, X.**, Song, X. & Li, J., 2015. High-resolution lithospheric structure beneath Mainland China from ambient noise and earthquake surface-wave tomography, *Earth planet. Sci. Lett.*, **417**, 132–141.
- Beyreuther, M.**, Barsch, R., Krischer, L., Megies, T., Behr, Y. & Wassermann, J., 2010. ObsPy: a Python Toolbox for Seismology, *Seismol. Res. Lett.*, **81**(3), 530–533.
- Bozdağ, E.** & Trampert, J., 2008. On crustal corrections in surface wave tomography, *Geophys. J. Int.*, **172**(3), 1066–1082.
- Bozdağ, E.**, Trampert, J. & Tromp, J., 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.*, **185**(2), 845–870.
- Bozdağ, E.**, Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N. & Pugmire, D., 2016. Global adjoint tomography: first-generation model, *Geophys. J. Int.*, **207**(3), 1739–1766.
- Bruce Watson, E.** & Brenan, J.M., 1987. Fluids in the lithosphere, 1. Experimentally-determined wetting characteristics of CO₂H₂O fluids and their implications for fluid transport, host-rock physical properties, and fluid inclusion formation, *Earth planet. Sci. Lett.*, **85**(4), 497–515.
- Chen, M.**, Niu, F., Liu, Q., Tromp, J. & Zheng, X., 2015. Multiparameter adjoint tomography of the crust and upper mantle beneath East Asia: 1. Model construction and comparisons, *J. geophys. Res.*, **120**(3), 1762–1786.
- Clayton, R.** & Engquist, B., 1977. Absorbing boundary conditions for acoustic and elastic wave equations, *Bull. seism. Soc. Am.*, **67**(6), 1529–1540.
- DeMets, C.**, Gordon, R.G., Argus, D.F. & Stein, S., 1990. Current plate motions, *Geophys. J. Int.*, **101**(2), 425–478.
- Durek, J.J.** & Ekström, G., 1996. A radial model of anelasticity consistent with long-period surface-wave attenuation, *Bull. seism. Soc. Am.*, **86**(1A), 144–158.
- Dziewonski, A.M.** & Anderson, D.L., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.*, **25**(4), 297–356.
- Dziewonski, A.M.**, Hager, B.H. & O’Connell, R.J., 1977. Large-scale heterogeneities in the lower mantle, *J. geophys. Res.*, **82**(2), 239–255.
- Ekström, G.**, Nettles, M. & Dziewoński, A., 2012. The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.*, **200–201**, 1–9.
- Ferreira, A. M.G.**, Woodhouse, J.H., Visser, K. & Trampert, J., 2010. On the robustness of global radially anisotropic surface wave tomography, *J. geophys. Res.*, **115**(B4), 2009JB006716, doi:10.1029/2009JB006716.
- Fichtner, A.** & Trampert, J., 2011a. Hessian kernels of seismic data functionals based upon adjoint techniques, *Geophys. J. Int.*, **185**(2), 775–798.
- Fichtner, A.** & Trampert, J., 2011b. Resolution analysis in full waveform inversion, *Geophys. J. Int.*, **187**(3), 1604–1624.
- Fichtner, A.**, Kennett, B. L.N., Igel, H. & Bunge, H.-P., 2009. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.*, **179**(3), 1703–1725.
- Fichtner, A.**, Kennett, B. L.N., Igel, H. & Bunge, H.-P., 2010. Full waveform tomography for radially anisotropic structure: New insights into present and past states of the Australasian upper mantle, *Earth planet. Sci. Lett.*, **290**(3–4), 270–280.
- Fichtner, A.**, Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y. & Villaseñor, A., 2013. Multiscale full waveform inversion, *Geophys. J. Int.*, **194**(1), 534–556.
- French, S.W.** & Romanowicz, B., 2015. Broad plumes rooted at the base of the Earth’s mantle beneath major hotspots, *Nature*, **525**(7567), 95–99.
- French, S.W.** & Romanowicz, B.A., 2014. Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography, *Geophys. J. Int.*, **199**(3), 1303–1327.
- Friederich, W.**, 2003. The S -velocity structure of the East Asian mantle from inversion of shear and surface waveforms, *Geophys. J. Int.*, **153**(1), 88–102.
- Gee, L.S.** & Jordan, T.H., 1992. Generalized seismological data functionals, *Geophys. J. Int.*, **111**(2), 363–390.
- Guo, Z.**, Chen, Y.J., Ning, J., Yang, Y., Afonso, J.C. & Tang, Y., 2016. Seismic evidence of on-going sublithosphere upper mantle convection for intra-plate volcanism in Northeast China, *Earth planet. Sci. Lett.*, **433**, 31–43.
- Guo, Z.**, Wang, K., Yang, Y., Tang, Y., John Chen, Y. & Hung, S.-H., 2018. The origin and mantle dynamics of quaternary intraplate volcanism in northeast China from joint inversion of surface wave and body wave, *J. geophys. Res.*, **123**(3), 2410–2425.
- Hayes, G.P.**, Moore, G.L., Portner, D.E., Hearne, M., Flamme, H., Furtney, M. & Smoczyk, G.M., 2018. Slab2, a comprehensive subduction zone geometry model, *Science*, **362**(6410), 58–61.
- He, C.**, Dong, S., Chen, X., Santosh, M. & Niu, S., 2014. Seismic evidence for plume-induced rifting in the Songliao Basin of Northeast China, *Tectonophysics*, **627**, 171–181.
- Hestenes, M.R.** & Stiefel, E., 1952. *Methods of Conjugate Gradients for Solving Linear Systems*, Vol. **49**, NBS Washington, DC.
- Huang, J.** & Zhao, D., 2006. High-resolution mantle tomography of China and surrounding regions, *J. geophys. Res.*, **111**(B9), doi:10.1029/2005JB004066.
- Huang, Z.** et al., 2015. Mantle structure and dynamics beneath SE Tibet revealed by new seismic images, *Earth planet. Sci. Lett.*, **411**, 100–111.
- Kakawatsu, H.**, Kumar, P., Takei, Y., Shinohara, M., Kanazawa, T., Araki, E. & Suyehiro, K., 2009. Seismic evidence for sharp lithosphere–asthenosphere boundaries of oceanic plates, *Science*, **324**(5926), 499–502.
- Kennett, B. L.N.**, Engdahl, E.R. & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, **122**(1), 108–124.
- Kim, Y.**, Liu, Q. & Tromp, J., 2011. Adjoint centroid-moment tensor inversions, *Geophys. J. Int.*, **186**(1), 264–278.
- Kim, Y.**, Chang, S., Witek, M., Ning, J. & Wen, J., 2021. S-velocity mantle structure of east Asia from teleseismic traveltimes tomography: inferred mechanisms for the Cenozoic intraplate volcanoes, *J. geophys. Res.*, **126**(3), doi:10.1029/2020JB020345.
- Komatitsch, D.** & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation-I. Validation, *Geophys. J. Int.*, **149**(2), 390–412.
- Komatitsch, D.** & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation-II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.*, **150**(1), 303–318.
- Krischer, L.**, Smith, J., Lei, W., Lefebvre, M., Ruan, Y., De Andrade, E.S., Podhorszki, N., Bozdağ, E. & Tromp, J., 2016. An adaptable seismic data format, *Geophys. J. Int.*, **207**(2), 1003–1011.
- Krischer, L.**, Fichtner, A., Boehm, C. & Igel, H., 2018. Automated large-scale full seismic waveform inversion for North America and the North Atlantic, *J. geophys. Res.*, **123**(7), 5902–5928.
- Kustowski, B.**, Ekström, G. & Dziewoński, A.M., 2008a. Anisotropic shear-wave velocity structure of the Earth’s mantle: a global model, *J. geophys. Res.*, **113**(B6), doi:10.1029/2007JB005169.
- Kustowski, B.**, Ekström, G. & Dziewoński, A.M., 2008b. The shear-wave velocity structure in the upper mantle beneath Eurasia, *Geophys. J. Int.*, **174**(3), 978–992.
- Laske, G.**, Masters, G., Ma, Z. & Pasyanos, M., 2013. Update on CRUST1.0—a 1-degree global model of Earth’s crust, in *Proceedings of the EGU General Assembly 2013*, held 7–12 April 2013, Vienna, Austria, id. EGU2013-2658.
- Lei, J.**, 2012. Upper-mantle tomography and dynamics beneath the North China Craton, *J. geophys. Res.*, **117**(B6), doi:10.1029/2012JB009212.
- Lei, J.** & Zhao, D., 2005. P-wave tomography and origin of the Changbai intraplate volcano in Northeast Asia, *Tectonophysics*, **397**(3–4), 281–295.
- Lei, J.** & Zhao, D., 2016. Teleseismic P-wave tomography and mantle dynamics beneath Eastern Tibet, *Geochim. Geophys. Geosyst.*, **17**(5), 1861–1884.

- Lei, J.**, Zhao, D., Steinberger, B., Wu, B., Shen, F. & Li, Z., 2009. New seismic constraints on the upper mantle structure of the Hainan plume, *Phys. Earth planet. Inter.*, **173**(1–2), 33–50.
- Lei, J.**, Xie, F., Fan, Q. & Santosh, M., 2013. Seismic imaging of the deep structure under the Chinese volcanoes: an overview, *Phys. Earth planet. Inter.*, **224**, 104–123.
- Lei, J.**, Zhao, D., Xu, X., Xu, Y.-G. & Du, M., 2019. Is there a big mantle wedge under eastern Tibet?, *Phys. Earth planet. Inter.*, **292**, 100–113.
- Lei, W.** et al., 2020. Global adjoint tomography—model GLAD-M25, *Geophys. J. Int.*, **223**(1), 1–21.
- Li, C.** & Van Der Hilst, R.D., 2010. Structure of the upper mantle and transition zone beneath Southeast Asia from travelttime tomography, *J. geophys. Res.*, **115**(B7), doi:10.1029/2009JB006882.
- Li, C.**, Van Der Hilst, R.D., Engdahl, E.R. & Burdick, S., 2008. A new global model for *P* wave speed variations in Earth's mantle, *Geochem. Geophys. Geosyst.*, **9**(5), doi:10.1029/2007GC001806.
- Li, G.**, Chen, H., Niu, F., Guo, Z., Yang, Y. & Xie, J., 2016. Measurement of Rayleigh wave ellipticity and its application to the joint inversion of high-resolution *S* wave velocity structure beneath northeast China, *J. geophys. Res.*, **121**(2), 864–880.
- Li, J.**, Chen, M., Koper, K.D., Zhou, T., Xi, Z., Li, S. & Li, G., 2021. FastTrip: a fast MPI-accelerated 1D triplication waveform inversion package for constraining mantle transition zone discontinuities, *Seismol. Res. Lett.*, **92**(4), 2647–2656.
- Li, S.** et al., 2023. Deep geophysical anomalies beneath the Changbaishan Volcano, *J. geophys. Res.*, **128**(4), e2022JB025671, doi:10.1029/2022JB025671.
- Lim, H.**, Kim, Y., Song, T.-R.A. & Shen, X., 2018. Measurement of seismometer orientation using the tangential *P*-wave receiver function based on harmonic decomposition, *Geophys. J. Int.*, **212**(3), 1747–1765.
- Ma, J.**, Bunge, H.-P., Thrastarson, S., Fichtner, A., Herwaarden, D.-P.V., Tian, Y., Chang, S.-J. & Liu, T., 2022. Seismic full-waveform inversion of the crust-mantle structure beneath China and adjacent regions, *J. geophys. Res.*, **127**(9), doi:10.1029/2022JB024957.
- Maggi, A.**, Tape, C., Chen, M., Chao, D. & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.*, **178**(1), 257–281.
- Mazza, S.E.**, Gazel, E., Bizimis, M., Moucha, R., Béguelin, P., Johnson, E.A., McAleer, R.J. & Sobolev, A.V., 2019. Sampling the volatile-rich transition zone beneath Bermuda, *Nature*, **569**(7756), 398–403.
- Meltzer, A.**, Stachnik, J.C., Sodnomsambuu, D., Munkhuu, U., Tsagaan, B., Dashdondog, M. & Russo, R., 2019. The central Mongolia seismic experiment: multiple applications of temporary broadband seismic arrays, *Seismol. Res. Lett.*, **90**(3), 1364–1376.
- Montelli, R.**, Nolet, G., Dahlen, F.A., Masters, G., Engdahl, E.R. & Hung, S.-H., 2004. Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, **303**(5656), 338–343.
- Nelson, P.L.** & Grand, S.P., 2018. Lower-mantle plume beneath the Yellowstone hotspot revealed by core waves, *Nat. Geosci.*, **11**(4), 280–284.
- Niu, F.** & Li, J., 2011. Component azimuths of the CEArray stations estimated from *P*-wave particle motion, *Earthq. Sci.*, **24**(1), 3–13.
- Obayashi, M.**, Yoshimitsu, J., Nolet, G., Fukao, Y., Shiobara, H., Sugioka, H., Miyamachi, H. & Gao, Y., 2013. Finite frequency whole mantle *P* wave tomography: Improvement of subducted slab images, *Geophys. Res. Lett.*, **40**(21), 5652–5657.
- Obrebski, M.**, Allen, R.M., Zhang, F., Pan, J., Wu, Q. & Hung, S.-H., 2012. Shear wave tomography of China using joint inversion of body and surface wave constraints, *J. geophys. Res.*, **117**(B1), doi:10.1029/2011JB008349.
- Okada, Y.**, Kasahara, K., Hori, S., Obara, K., Sekiguchi, S., Fujiwara, H. & Yamamoto, A., 2004. Recent progress of seismic observation networks in Japan —Hi-net, F-net, K-NET and KiK-net—, *Earth, Planets Space*, **56**(8), xv–xxviii.
- Panning, M.** & Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.*, **167**(1), 361–379.
- Panning, M.P.**, Lekić, V. & Romanowicz, B.A., 2010. Importance of crustal corrections in the development of a new global model of radial anisotropy, *J. geophys. Res.*, **115**(B12), doi:10.1029/2010JB007520.
- Peter, D.** et al., 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.*, **186**(2), 721–739.
- Polak, E.** & Ribiere, G., 1969. Note sur la convergence de méthodes de directions conjuguées, *Revue française d'informatique et de recherche opérationnelle. Série rouge*, **3**(16), 35–43.
- Powell, M.J.D.** & Powell, P.O.A.N.A.M.J.D., 1981. *Approximation Theory and Methods*, Cambridge Univ. Press.
- Rickers, F.**, Fichtner, A. & Trampert, J., 2012. Imaging mantle plumes with instantaneous phase measurements of diffracted waves, *Geophys. J. Int.*, **190**(1), 650–664.
- Rodgers, A.**, Krischer, L., Afanasiev, M., Boehm, C., Doody, C., Chiang, A. & Simmons, N., 2022. WUS256: an adjoint waveform tomography model of the crust and upper mantle of the western United States for improved waveform simulations, *J. geophys. Res.*, **127**(7), e2022JB024549, doi:10.1029/2022JB024549.
- Ruan, Y.**, Lei, W., Modrak, R., Örsuvuran, R., Bozdağ, E. & Tromp, J., 2019. Balancing unevenly distributed data in seismic tomography: a global adjoint tomography example, *Geophys. J. Int.*, **219**(2), 1225–1236.
- Simmons, N.A.**, Myers, S.C., Morency, C., Chiang, A. & Knapp, D.R., 2021. SPiRaL: a multiresolution global tomography model of seismic wave speeds and radial anisotropy variations in the crust and mantle, *Geophys. J. Int.*, **227**(2), 1366–1391.
- Smithsonian-Institution**, 2023. *Global Volcanism Program*, 2023, [Database]Volcanoes of the World (v. 5.0.4; 17 Apr 2023), Distributed by Smithsonian Institution, compiled by Venzke, E.
- Stacey, R.**, 1988. Improved transparent boundary formulations for the elastic-wave equation, *Bull. seism. Soc. Am.*, **78**(6), 2089–2097.
- Tang, Y.** et al., 2014. Changbaishan volcanism in northeast China linked to subduction-induced mantle upwelling, *Nat. Geosci.*, **7**(6), 470–475.
- Tao, K.**, Grand, S.P. & Niu, F., 2017. Full-waveform inversion of tripli-cated data using a normalized-correlation-coefficient-based misfit function, *Geophys. J. Int.*, **210**(3), 1517–1524.
- Tao, K.**, Grand, S.P. & Niu, F., 2018. Seismic structure of the upper mantle beneath Eastern Asia from full waveform seismic tomography, *Geochem. Geophys. Geosyst.*, **19**(8), 2732–2763.
- Tape, C.**, Liu, Q. & Tromp, J., 2007. Finite-frequency tomography using adjoint methods—methodology and examples using membrane surface waves, *Geophys. J. Int.*, **168**(3), 1105–1129.
- Tape, C.**, Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the southern California crust, *Science*, **325**(5943), 988–992.
- Tape, C.**, Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.*, **180**(1), 433–462.
- Tapponnier, P.**, Peltzer, G. & Armijo, R., 1986. On the mechanics of the collision between India and Asia, *Geol. Soc. Lond. Spec. Publ.*, **19**(1), 113–157.
- Tian, Y.**, Zhao, D., Sun, R. & Teng, J., 2009. Seismic imaging of the crust and upper mantle beneath the North China Craton, *Phys. Earth planet. Inter.*, **172**(3–4), 169–182.
- Wang, S.**, Niu, F. & Zhang, G., 2013. Velocity structure of the uppermost mantle beneath East Asia from *Pn* tomography and its dynamic implications, *J. geophys. Res.*, **118**(1), 290–301.
- Wang, X.** et al., 2020. Distinct slab interfaces imaged within the mantle transition zone, *Nat. Geosci.*, **13**(12), 822–827.
- Wang, Y.**, Zhang, X., Jiang, C., Wei, H. & Wan, J., 2007. Tectonic controls on the late Miocene–Holocene volcanic eruptions of the Tengchong volcanic field along the southeastern margin of the Tibetan plateau, *J. Asian Earth Sci.*, **30**(2), 375–389.
- Warner, M.** et al., 2013. Anisotropic 3D full-waveform inversion, *Geophysics*, **78**(2), R59–R80.
- Wei, S.** et al., 2010. Regional earthquakes in northern Tibetan Plateau: implications for lithospheric strength in Tibet, *Geophys. Res. Lett.*, **37**(19), doi:10.1029/2010GL044800.
- Wei, S.S.** & Chen, Y.J., 2016. Seismic evidence of the Hainan mantle plume by receiver function analysis in southern China, *Geophys. Res. Lett.*, **43**(17), 8978–8985.

- Wei, W., Xu, J., Zhao, D. & Shi, Y., 2012. East Asia mantle tomography: new insight into plate subduction and intraplate volcanism, *J. Asian Earth Sci.*, **60**, 88–103.
- Weston, J., Engdahl, E.R., Harris, J., Di Giacomo, D. & Storchak, D.A., 2018. ISC-EHB: reconstruction of a robust earthquake data set, *Geophys. J. Int.*, **214**(1), 474–484.
- Witek, M., van der Lee, S. & Kang, T.-S., 2014. Rayleigh wave group velocity distributions for East Asia using ambient seismic noise, *Geophys. Res. Lett.*, **41**(22), 8045–8052.
- Witek, M., Chang, S.-J., Lim, D.Y., Ning, S. & Ning, J., 2021. Radial anisotropy in East Asia from multimode surface wave tomography, *J. geophys. Res.*, **126**(7), e2020JB021201, doi:10.1029/2020JB021201.
- Wolfe, C.J., Solomon, S.C., Laske, G., Collins, J.A., Detrick, R.S., Orcutt, J.A., Bercovici, D. & Hauri, E.H., 2009. Mantle shear-wave velocity structure beneath the Hawaiian hot spot, *Science*, **326**(5958), 1388–1390.
- Wu, J., Suppe, J., Lu, R. & Kanda, R., 2016. Philippine Sea and East Asian plate tectonics since 52 Ma constrained by new subducted slab reconstruction methods: Philippine sea plate tectonics-slabs, *J. geophys. Res.*, **121**(6), 4670–4741.
- Xie, J., Ritzwoller, M.H., Shen, W., Yang, Y., Zheng, Y. & Zhou, L., 2013. Crustal radial anisotropy across Eastern Tibet and the Western Yangtze Craton, *J. geophys. Res.*, **118**(8), 4226–4252.
- Xu, Y.-G., Ma, J.-L., Frey, F.A., Feigenson, M.D. & Liu, J.-F., 2005. Role of lithosphere–asthenosphere interaction in the genesis of Quaternary alkali and tholeiitic basalts from Datong, western North China Craton, *Chem. Geol.*, **224**(4), 247–271.
- Yang, Y. et al., 2019. Crustal structure beneath Northeast China from ambient noise tomography, *Phys. Earth planet. Inter.*, **293**, doi:10.1016/j.pepi.2019.04.008.
- Yao, H., Van Der Hilst, R.D. & De Hoop, M.V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - I. Phase velocity maps, *Geophys. J. Int.*, **166**(2), 732–744.
- Yao, H., Beghein, C. & Van Der Hilst, R.D., 2008. Surface wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - II. Crustal and upper-mantle structure, *Geophys. J. Int.*, **173**(1), 205–219.
- Yin, A. & Harrison, T.M., 2000. Geologic evolution of the Himalayan–Tibetan Orogen, *Annu. Rev. Earth planet. Sci.*, **28**(1), 211–280.
- Yuan, Y.O., Bozdağ, E., Ciardelli, C., Gao, F. & Simons, F.J., 2020. The exponentiated phase measurement, and objective-function hybridization for adjoint waveform tomography, *Geophys. J. Int.*, **221**(2), 1145–1164.
- Yue, H. et al., 2012. Lithospheric and upper mantle structure of the northeastern Tibetan Plateau, *J. geophys. Res.*, **117**(B5), doi:10.1029/2011JB008545.
- Zhang, H., Wang, F., Myhill, R. & Guo, H., 2019. Slab morphology and deformation beneath Izu-Bonin, *Nat. Commun.*, **10**(1), doi:10.1038/s41467-019-09279-7.
- Zhang, K.-J., 1997. North and South China collision along the eastern and southern North China margins, *Tectonophysics*, **270**(1–2), 145–156.
- Zhang, M. et al., 2018. The intraplate Changbaishan volcanic field (China/North Korea): a review on eruptive history, magma genesis, geo-dynamic significance, recent dynamics and potential hazards, *Earth-Sci. Rev.*, **187**, 19–52.
- Zhao, D., Maruyama, S. & Omori, S., 2007. Mantle dynamics of Western Pacific and East Asia: insight from seismic tomography and mineral physics, *Gondwana Res.*, **11**(1–2), 120–131.
- Zhao, D., Tian, Y., Lei, J., Liu, L. & Zheng, S., 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**(3–4), 197–206.
- Zheng, T., Chen, L., Zhao, L., Xu, W. & Zhu, R., 2006. Crust–mantle structure difference across the gravity gradient zone in North China Craton: seismic image of the thinned continental crust, *Phys. Earth planet. Inter.*, **159**(1–2), 43–58.
- Zheng, X.-F., Ouyang, B., Zhang, D.-N., Yao, Z.-X., Liang, J.-H. & Zheng, J., 2009. Construction of the technical system of data backup centre for China Seismograph Network and its support to Wenchuan earthquake researches, *Chinese J. Geophys.*, **52**(3), 657–662.
- Zhou, Q., Liu, L. & Hu, J., 2018. Western US volcanism due to intruding oceanic mantle driven by ancient Farallon slabs, *Nat. Geosci.*, **11**(1), 70–76.
- Zhou, T., Xi, Z., Chen, M. & Li, J., 2021. Assessment of seismic tomographic models of the contiguous United States using intermediate-period 3-D wavefield simulation, *Geophys. J. Int.*, **228**(2), 1392–1409.
- Zhou, T., Li, J., Xi, Z., Li, G. & Chen, M., 2022. CUSRA2021: a radially anisotropic model of the contiguous US and surrounding regions by full-waveform inversion, *J. geophys. Res.*, **127**(8), doi:10.1029/2021JB023893.
- Zhou, X. & Armstrong, R., 1982. Cenozoic volcanic rocks of eastern China—secular and geographic trends in chemistry and strontium isotopic composition, *Earth planet. Sci. Lett.*, **58**(3), 301–329.
- Zhou, Y., 2018. Anomalous mantle transition zone beneath the Yellowstone hotspot track, *Nat. Geosci.*, **11**(6), 449–453.
- Zhu, H., Bozdağ, E., Peter, D. & Tromp, J., 2012a. Seismic wavespeed images across the Iapetus and Tornquist suture zones, *Geophys. Res. Lett.*, **39**(18), doi:10.1029/2012GL053053.
- Zhu, H., Bozdağ, E., Peter, D. & Tromp, J., 2012b. Structure of the European upper mantle revealed by adjoint tomography, *Nat. Geosci.*, **5**(7), 493–498.
- Zhu, H., Bozdağ, E. & Tromp, J., 2015. Seismic structure of the European upper mantle based on adjoint tomography, *Geophys. J. Int.*, **201**(1), 18–52.
- Zhu, R., Chen, L., Wu, F. & Liu, J., 2011. Timing, scale and mechanism of the destruction of the North China Craton, *Sci. China Earth Sci.*, **54**(6), 789–797.
- Zuo, J., Wang, L. & Niu, F., 2020. Multiple source downwellings beneath eastern North China revealed by 3-D CCP migration of receiver function data, *J. Asian Earth Sci.*, **192**, doi:10.1016/j.jseaes.2020.104266.

APPENDIX A: DATA QUALITY CONTROL WEIGHTING

As shown in eq.(2) of the main texts, W_w is the weighting term to control the data quality in each measurement window. Following Tao et al. (2018), W_w is defined as

$$W_w = f_1(\text{CC})f_2(\text{SNR})f_3(\delta t_{\text{CC}}) \quad (\text{A1})$$

where

$$f_1(\text{CC}) = \begin{cases} 0, & \text{CC} < 0.5 \\ \frac{1}{2} + \frac{1}{2} \cos\left(\pi \frac{0.7-\text{CC}}{0.7-0.5}\right), & 0.5 \leq \text{CC} < 0.7 \\ 1, & \text{CC} \geq 0.7 \end{cases} \quad (\text{A2})$$

$$f_2(\text{SNR}) = \begin{cases} 0, & \text{SNR} < 10 \\ \frac{1}{2} + \frac{1}{2} \cos\left(\pi \frac{15-\text{SNR}}{15-10}\right), & 10 \leq \text{SNR} < 15 \\ 1, & \text{SNR} \geq 15 \end{cases} \quad (\text{A3})$$

$$f_3(\delta t_{\text{CC}}) = \begin{cases} 0, & |\delta t_{\text{CC}}| \geq 10 \\ \frac{1}{2} - \frac{1}{2} \cos\left(\pi \frac{10-|\delta t_{\text{CC}}|}{10-8}\right), & 8 \leq |\delta t_{\text{CC}}| < 10 \\ 1, & |\delta t_{\text{CC}}| < 8 \end{cases} \quad (\text{A4})$$

SNR is the signal-to-noise ratio for each measurement window:

$$\text{SNR} = 10 \log \frac{\frac{1}{t_{\text{signal}}} \int u_{\text{signal}}^2(t) dt}{\frac{1}{t_{\text{noise}}} \int u_{\text{noise}}^2(t) dt} \quad (\text{A5})$$

where $u_{\text{signal}}(t)$ represents displacement within the measurement window, t_{signal} is the corresponding window length, $u_{\text{noise}}(t)$ is displacement in the noise window defined as from the event origin time to 10 s before the predicted first P-wave arrival time based on the AK135 model (Kennett et al. 1995), and t_{noise} is the corresponding window length. CC and δt_{CC} are the cross-correlation coefficient and time-shift between the data and synthetics in the measurement window.

APPENDIX B: SOURCE INVERSION DETAILS

The invertible source parameters include centroid moment tensor M , 3-D centroid location X , and source time function $S(t)$. Assuming a Gaussian source time function, we have

$$S(t) = \frac{1}{\sqrt{\pi}\tau} e^{-\left(\frac{t-t_0}{\tau}\right)^2} \quad (\text{B1})$$

where τ means the half duration time and t_0 represents the event's origin time. Because the frequency band of our inversion cannot well constrain τ , we only invert for t_0 and use τ determined by the magnitude in the Global CMT solution (Ekström *et al.* 2012).

Following Kim *et al.* (2011), we have ten source parameters to invert, including six elements of the moment tensor $M_{11}, M_{22}, M_{33}, M_{12}, M_{13}, M_{23}$, 3 elements of the 3-D centroid location x_1, x_2, x_3 , and the centroid origin time t_0 . The gradient to the data misfit χ_e is

$$g = \left\{ \frac{\partial \chi_e}{\partial M_{11}}, \frac{\partial \chi_e}{\partial M_{22}}, \frac{\partial \chi_e}{\partial M_{33}}, \frac{\partial \chi_e}{\partial M_{12}}, \frac{\partial \chi_e}{\partial M_{13}}, \frac{\partial \chi_e}{\partial M_{23}}, \frac{\partial \chi_e}{\partial x_1}, \frac{\partial \chi_e}{\partial x_2}, \frac{\partial \chi_e}{\partial x_3}, \frac{\partial \chi_e}{\partial t_0} \right\} \quad (\text{B2})$$

By using $\sigma_M = \sqrt{2}((\frac{\partial \chi_e}{\partial M_{11}})^2 + (\frac{\partial \chi_e}{\partial M_{22}})^2 + (\frac{\partial \chi_e}{\partial M_{33}})^2 + (\frac{\partial \chi_e}{\partial M_{12}})^2 + (\frac{\partial \chi_e}{\partial M_{13}})^2 + (\frac{\partial \chi_e}{\partial M_{23}})^2)^{-\frac{1}{2}}$ and $\sigma_X = ((\frac{\partial \chi_e}{\partial x_1})^2 + (\frac{\partial \chi_e}{\partial x_2})^2 + (\frac{\partial \chi_e}{\partial x_3})^2)^{-\frac{1}{2}}$ to non-dimensionalize M and X , we have non-dimensional parameters $\hat{M} = \frac{M}{\sigma_M}$, $\hat{X} = \frac{X}{\sigma_X}$ and gradients $\frac{\partial \chi_e}{\partial \hat{M}} = \sigma_M \frac{\partial \chi_e}{\partial M}$, $\frac{\partial \chi_e}{\partial \hat{X}} = \sigma_X \frac{\partial \chi_e}{\partial X}$. The non-dimensionalization will guarantee different types of parameters contribute equally to the gradient's magnitude.

Each source inversion step contains three iterations of the steepest descent update, after which the source parameters will almost not change. To do line search of the steepest descent update in each iteration, we first calculate the waveform synthetics of the perturbed source model $\{M - \epsilon \frac{\partial \chi_e}{\partial \hat{M}}, X - \epsilon \frac{\partial \chi_e}{\partial \hat{X}}, t_0\}$, where $\epsilon = 0.01$. By assuming a linear relationship between the source model and the waveform perturbations, we approximate the waveform synthetics $s_\alpha(x, t)$ for the source model $\{M - \alpha \frac{\partial \chi_e}{\partial \hat{M}}, X - \alpha \frac{\partial \chi_e}{\partial \hat{X}}, t'\}$ with arbitrary step length α and origin time t' as

$$s_\alpha(x, t) = s_0(x, t - (t' - t_0)) + \frac{\alpha}{\epsilon}(s_\epsilon(x, t - (t' - t_0)) - s_0(x, t - (t' - t_0))) \quad (\text{B3})$$

where $s_0(x, t)$ is the synthetics for the starting source model $\{M, X, t_0\}$ of this iteration, $s_\epsilon(x, t)$ is the perturbed synthetics for the source model $\{M - \epsilon \frac{\partial \chi_e}{\partial \hat{M}}, X - \epsilon \frac{\partial \chi_e}{\partial \hat{X}}, t_0\}$. $s_0(x, t - (t' - t_0))$ and $s_\epsilon(x, t - (t' - t_0))$ shift the synthetics by $t' - t_0$ to align the waveforms. We use a Bayesian optimization method to efficiently find an optimized step length α_{opt} and the origin time t'_{opt} to minimize the data misfit.

APPENDIX C: OPTIMIZATION OF THE STRUCTURAL MODEL PARAMETERS

We use the conjugate gradient descent to update our structure model. In each iteration, the search direction is determined by the gradient of the current iteration g_i and the search direction of the previous iteration d_{i-1} :

$$d_i = -Pg_i + \beta d_{i-1} \quad (\text{C1})$$

where P is the pre-conditioner and

$$\beta = \frac{(d_i - d_{i-1}) \cdot Pg_i}{d_{i-1} \cdot Pg_{i-1}} \quad (\text{C2})$$

In the first iteration, β is set to 0. In the later iterations, we calculate the Powell ratio (Powell & Powell 1981) defined as:

$$R = \frac{Pg_i \cdot Pg_{i-1}}{Pg_{i-1} \cdot Pg_{i-1}} \quad (\text{C3})$$

If $R > 0.2$ for all the invertible structure parameters, we reset β to 0 to restart the optimization when a critical point is found, thus speed up our inversion.

The pre-conditioner P is chosen to be close to the inverse Hessian matrix (Fichtner & Trampert 2011a). In practice, it is difficult to directly calculate the Hessian matrix. Instead, we use the so-called pseudo-Hessian matrix (Bozdağ *et al.* 2016):

$$H(x) = \sum_{e=1}^{N_e} \int \frac{\partial^2 s(x, t)}{\partial t^2} \cdot \frac{\partial^2 s^\dagger(x, T-t)}{\partial t^2} dt \quad (\text{C4})$$

where $s(x, t)$ is the wave field of the forward simulation and $s^\dagger(x, T-t)$ is the backward wave field of the adjoint source. N_e denotes the total number of events.

By setting an appropriate threshold δ of the inverse hessian, we have

$$P(x) = \begin{cases} \frac{1}{\delta} & H(x) < \delta \\ \frac{1}{H(x)} & H(x) \geq \delta \end{cases} \quad (\text{C5})$$

δ is necessary as P acts as weighting to the gradient. We need δ to suppress a large P . It should be noticed that the choice of δ has a significant impact on the inversion. As P can be regarded as the weighting of the gradient in different regions, a smaller δ will tolerate a larger value of $P(x)$. A region with a larger $P(x)$ usually means a smaller Hessian matrix, i.e., less data coverage. To balance model updating in different regions, we should use appropriate weighting to enhance the gradient for the area with less data coverage (Fichtner & Trampert 2011a).

For the calculated kernels from the adjoint simulation of each event, we first use Gaussian spheres centred at the sources and receivers to damp out regions with potential singularity (Zhu *et al.* 2015). The diameters of the Gaussian spheres are 50 km for the source and 10 km for the receivers. Then we apply the preconditioning after summing up the kernels for all events. In the next step, we apply Gaussian smoothing on the kernel to help stabilize the inversion. In the first 5 iterations of the inversion, the Gaussian smoothing spheres are configured as 50 km horizontally and 25 km vertically. In the subsequent 25 iterations, we choose Gaussian smoothing spheres with 25 km horizontally and 10 km vertically to capture small-scale structures.

We use line search in each iteration to find the optimal step length. Based on the line search direction d_i , which is the negation of the normalized gradient, we forwardly simulate synthetics for the perturbed model $m_i + \epsilon d_i$, where $\epsilon = 0.01$. Using the linear dependency approximation of waveform to structure (Warner *et al.* 2013), we estimate the perturbed synthetics for arbitrary model $m_i + \alpha d_i$ as

$$s(x, t; m_i + \alpha d_i) = s(x, t; m_i) + \frac{\alpha}{\epsilon}(s(x, t; m_i + \epsilon d_i) - s(x, t; m_i)) \quad (\text{C6})$$

where $s(x, t; m_i)$ represents the synthetics for model m_i , and $s(x, t; m_i + \epsilon d_i)$ represents the synthetics for model $m_i + \epsilon d_i$. It is computationally efficient to calculate $s(x, t; m_i + \alpha d_i)$ without performing the SEM simulation. We can simply test different values of α and find the optimized one.

APPENDIX D: POINT SPREAD FUNCTION TESTS

Following Fichtner & Trampert (2011b), if we apply a single point perturbation to one of the invertible model parameters, the gradient from the adjoint state method can be regarded as the negation of the Hessian matrix, also called the point spread function (PSF), of the final model. This function can quantify the model's blurring and distortion by showing how it tends to update when only one point is

perturbed. For example, when we apply a single point perturbation to V_{SV} , the PSF response for V_{SV} and V_{SH} are $H_{V_{SV}, V_{SV}}$, $H_{V_{SV}, V_{SH}}$, respectively, where H represents the Hessian matrix.

In practice, we apply Gaussian sphere perturbations with a radius of 50 km and a separation of 2° horizontally and 200 km vertically, located at depths of 100, 300, 500, 700 and 900 km, to the final model. By using the adjoint state method, the corresponding PSFs are calculated and shown in the main text.