SASSY21: A 3-D seismic structural model of the lithosphere and underlying mantle beneath Southeast Asia from multi-scale adjoint waveform tomography

Deborah Wehner[1], Nienke Blom[2], Nicholas Rawlinson[2], Daryono Daryono[3], Christian Boehm[4], Meghan Samantha Miller[5], Pepen Supendi[6], and Sri Widiyantoro[7]

[1]Univeristy of Cambridge
[2]University of Cambridge
[3]Agency for Meteorology, Climatology, and Geophysics (BMKG), Indonesia
[4]Department of Earth Sciences, Institute of Geophysics, ETH Zürich
[5]Australian National University
[6]Agency for Meteorology, Climatology, and Geophysics (BMKG)
[7]Global Geophysics Research Group, Faculty of Mining and Petroleum Engineering, Institut Teknologi Bandung, Indonesia

November 24, 2022

Abstract

We present the first continental-scale seismic model of the lithosphere and underlying mantle beneath Southeast Asia obtained from adjoint waveform tomography (often referred to as full-waveform inversion or FWI), using seismic data filtered at periods from 20 - 150s. Based on >3,000h of analyzed waveform data gathered from ˜13,000 unique source-receiver pairs, we image isotropic P-wave velocity, radially anisotropic S-wave velocity and density via an iterative non-linear inversion that begins from a 1-D reference model. At each iteration, the full 3-D wavefield is determined through an anelastic Earth, accommodating effects of topography, bathymetry and ocean load. Our data selection aims to maximize sensitivity to deep structure by accounting for body-wave arrivals separately. SASSY21, our final model after 87 iterations, is able to explain true-amplitude data from events and receivers not included in the inversion. The trade-off between inversion parameters is estimated through an analysis of the Hessian-vector product. SASSY21 reveals detailed anomalies down to the mantle transition zone, including multiple subduction zones. The most prominent feature is the (Indo-)Australian plate descending beneath Indonesia, which is imaged as one continuous slab along the 180-degree curvature of the Banda Arc. The tomography confirms the existence of a hole in the slab beneath Mount Tambora and locates a high S-wave velocity zone beneath northern Borneo that may be associated with subduction termination in the mid-late Miocene. A previously undiscovered feature beneath the east coast of Borneo is also revealed, which may be a signature of post-subduction processes, delamination or underthrusting from the formation of Sulawesi.

***SASSY21*: A 3-D seismic structural model of the lithosphere and underlying mantle beneath Southeast Asia from multi-scale adjoint waveform tomography**

Deborah Wehner[1], Nienke Blom[1], Nicholas Rawlinson [1], Daryono[2], Christian Böhm[3,4], Meghan S. Miller[5], Pepen Supendi[2], and Sri Widiyantoro[6,7]

[1]Bullard Laboratories, Department of Earth Sciences, University of Cambridge, Cambridge, UK
[2]Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG)
[3]ETH Zürich, Zurich, Switzerland
[4]Mondaic AG, Zurich, Switzerland
[5]Australian National University, Canberra, Australia
[6]Global Geophysics Research Group, Faculty of Mining and Petroleum Engineering, Institut Teknologi Bandung, Bandung, Indonesia
[7]Faculty of Engineering, Maranatha Christian University, Bandung, Indonesia

Key Points:

- A new FWI tomographic model illuminates the upper mantle beneath Southeast Asia
- Use of an extensive dataset allows key features, including multiple subduction zones, to be identified
- A high-velocity structure beneath Borneo, likely associated with post-subduction processes, is clearly visible

Corresponding author: Deborah Wehner, dwehner@esc.cam.ac.uk

Abstract

We present the first continental-scale seismic model of the lithosphere and underlying mantle beneath Southeast Asia obtained from adjoint waveform tomography (often referred to as full-waveform inversion or FWI), using seismic data filtered at periods from 20 – 150 s. Based on > 3,000 h of analyzed waveform data gathered from ∼13,000 unique source-receiver pairs, we image isotropic P-wave velocity, radially anisotropic S-wave velocity and density via an iterative non-linear inversion that begins from a 1-D reference model. At each iteration, the full 3-D wavefield is determined through an anelastic Earth, accommodating effects of topography, bathymetry and ocean load. Our data selection aims to maximize sensitivity to deep structure by accounting for body-wave arrivals separately. *SASSY21*, our final model after 87 iterations, is able to explain true-amplitude data from events and receivers not included in the inversion. The trade-off between inversion parameters is estimated through an analysis of the Hessian-vector product. *SASSY21* reveals detailed anomalies down to the mantle transition zone, including multiple subduction zones. The most prominent feature is the (Indo-)Australian plate descending beneath Indonesia, which is imaged as one continuous slab along the 180° curvature of the Banda Arc. The tomography confirms the existence of a hole in the slab beneath Mount Tambora and locates a high S-wave velocity zone beneath northern Borneo that may be associated with subduction termination in the mid-late Miocene. A previously undiscovered feature beneath the east coast of Borneo is also revealed, which may be a signature of post-subduction processes, delamination or underthrusting from the formation of Sulawesi.

Plain Language Summary

Southeast Asia is one of the world's most tectonically active regions, as evidenced by frequent large earthquakes and volcanic eruptions. We present a large-scale 3-D seismic structural model of this region down to a depth of 800 km that reveals a variety of primary features, including beneath the poorly understood islands of Borneo and Sulawesi. This is possible thanks to the use of a sizable dataset of earthquakes recorded by a large number of permanent and temporary stations located in Southeast Asia, and advanced imaging methodology that is better able to capture the true physics of seismic wave propagation compared to more traditional methods. Our new model is capable of resolving variations in seismic properties associated with ongoing subduction (when one tectonic plate descends into the mantle below another plate), particularly along the northern margin of the Australian plate beneath the Sunda Arc. More subtle anomalies associated with remnant subduction, which correspond to plate fragments that remain once subduction stops, can also be imaged. These results are important for achieving a better understanding of the subduction cycle, which plays a central role in plate tectonics, and has important implications for, among other things, the evolution of the continents, the global carbon budget, and volcanic and earthquake hazard.

1 Introduction

Seismic tomography has played a crucial role in the illumination of deep Earth structure since the first pioneering studies of the mid 1970's (e.g. Aki et al., 1977; Dziewonski et al., 1977). A wide range of tomographic methods now exist, but these are mostly based on seismic ray theory and hence do not fully account for the true physics of wave propagation. In particular, seismic waves propagate at finite frequencies and sample extensive regions outside the geometric ray path. Adjoint waveform tomography, often referred to as full-waveform inversion (FWI), embraces the full complexity of seismic wave propagation, by accurately solving the 3-D seismic wave equation numerically. It can account for effects such as wavefront healing, interference and (de)focusing, which are not accurately modeled with ray theory (e.g. Rickers et al., 2012). As a result, FWI promises high-resolution images and a more reliable quantification of anomalies, which opens up new avenues for more robust interpretation of seismic models in terms of composition, temperature, melt and other material properties (Tromp, 2020).

The mathematical background of FWI has been known since the 1980s (Lailly & Bednar, 1983; Tarantola, 1984), but its comprehensive application has not been computationally feasible until recently. The method was first developed in seismic exploration (Gauthier et al., 1986; Pratt & Worthington, 1990) and has proven its ability in a wide range of applications in this field (e.g. Sirgue et al., 2010). It has also been successfully applied in other areas such as medicine using ultrasound measurements (e.g. Schreiman et al., 1984; Guasch et al., 2020) and engineering using ultraseismic waveforms (e.g. Jalinoos et al., 2017). The first applications of FWI in earthquake seismology include imaging the Californian crust and the Australasian upper mantle (Chen et al., 2007; Fichtner et al., 2009; Tape et al., 2010; Zhu et al., 2012). Since then, the method has demonstrated its ability to produce high-resolution, multi-parameter subsurface images across all scales (e.g. Lei et al., 2020), thus providing new opportunities for geophysical and geochemical interpretation. However, issues such as high computational requirements, significant non-linearity of the inverse problem, data selection and sensitivity to multiple parameter types typically makes the implementation of this iterative process much more challenging compared to ray-based methods. A more detailed technical review, including FWI in the context of seismic exploration, is provided in Virieux and Operto (2009) and Tromp (2020).

FWI is especially suitable for imaging tectonically active parts of the Earth, where large contrasts in elastic properties are likely to be present, and the assumptions of ray theory become less valid. Southeast Asia is one such region, where significant tectonic complexity is caused by its location at the junction of three converging tectonic plates. This has produced a network of subduction zones, which makes the region vulnerable to natural hazards such as large-magnitude shallow earthquakes that can lead to tsunamis (e.g. 2004 Sumatra earthquake, McCaffrey, 2009) and volcanic eruptions (e.g. 2018 Krakatoa eruption, Petley, 2019). Overall, Southeast Asia provides a unique setting to investigate a variety of primary tectonic processes, including subduction initiation, ongoing subduction, subduction termination, collision (both arc-continent and continent-continent), orogen collapse and tectonic escape (e.g. Hall, 2013).

So far, studies that investigate the seismic structure of Southeast Asia as a whole are either global or regional, using body or surface wave tomography methods only (e.g. Widiyantoro & van der Hilst, 1996; Bijwaard et al., 1998; Lebedev & Nolet, 2003; Fukao & Obayashi, 2013; Schaeffer & Lebedev, 2013; Miller et al., 2016; Zenonos et al., 2019; Harris et al., 2020). The resultant models all tend to agree on low velocities in the upper 200 km beneath the region encompassing the Thai-Malay Peninsula and Borneo. They also agree on high velocities along the Indonesian volcanic arc and around the northward continuation of the North Australian craton in the Timor Sea. Furthermore, several of these studies have identified a number of subducting slabs in Southeast Asia, mainly around the Sunda and Banda Arcs as well as Sulawesi and Borneo. However, discrepancies exist regarding the geometry and depth extent of the subducted slab segments and previous studies lack constraints in key regions, in particular around the poorly imaged islands of Borneo and Sulawesi. While several smaller-scale features have been imaged in this region, they have tended to be treated as artifacts due to poor data coverage (Hall & Spakman, 2015; Zenonos et al., 2019).

We present a new large-scale model of the entire Southeast Asian lithosphere and underlying mantle, defined by both P-wave and S-wave structure, and constrained by inversion of both body and surface waveforms. This is achieved through the application of adjoint waveform tomography to a large regional dataset that permits the imaging of structures down to the mantle transition zone. This paper elaborates on the inversion setup and assessment of the robustness of the final model. Furthermore, we highlight and discuss some of the key features of the tomographic model.

2 Tectonic setting of Southeast Asia

Southeast Asia is located at the triple junction of three key tectonic plates: the Eurasian, (Indo-)Australian, and Philippine Sea plates (see Figure 1). Seismicity occurs at the highly active boundaries between these plates, where extensive subduction zones feature slabs descending at rates between 5 – 10 centimeter per year (e.g, Simons et al., 2007), and generate frequent earthquakes to depths of up to 700 km, thus providing an excellent dataset for regional tomography. The study region is largely comprised of a shallow, continental shelf that includes Borneo, Peninsular Malaysia, Sumatra, Java and parts of the South China Sea (see Figure 1). This continental promontory of the Eurasian Plate, often referred to as the Sundaland block (see Figure 1), includes a large number of thick Cenozoic sedimentary basins (e.g. Hall & Morley, 2004). Overall, it experiences low levels of seismicity within its interior, but evidence for a complex pattern of subsidence and elevation indicates that the region has been far from stable during most of the Cenozoic (Hall & Morley, 2004; Yang et al., 2016). GPS measurements demonstrate that the Sundaland block moves independently from Eurasia towards the east while rotating clockwise, with an average velocity of several millimeters per year (Simons et al., 2007).

In the western and southern regions of the Sundaland block, the descent of the oceanic (Indo-)Australian plate forms an active subduction system beneath the Indonesian volcanic arc. Ongoing subduction along the Sunda Arc represents a significant natural hazard due to associated earthquakes, tsunamis and volcanoes, which is why it is the focus of ongoing research (e.g. Métrich et al., 2017; Wang & He, 2020). However, discrepancies exist among previous studies regarding the geometry and depth extent of the subducted slab segments (Li et al., 2021). Several previous studies suggest that the subducted slab only extends in depth to the mantle transition zone (e.g. Gudmundsson & Sambridge, 1998; Amaru, 2007), while others advocate for its penetration into the lower mantle (e.g. Huang et al., 2015; Fukao & Obayashi, 2013).

Borneo is the largest island within Southeast Asia and lies in the eastern region of the Sundaland block. In the Miocene, two sequential but apposed subduction systems were in operation in the northern part of Borneo, which featured southeast subduction of the proto-South China Sea, and northwest subduction of the Celebes Sea (Hall, 2013). Termination of the proto-South China Sea subduction at 23 Ma coincided with continent-continent collision and formation of the Crocker Range, and termination of the Celebes Sea subduction at $\sim$9 Ma was followed by southerly subduction beneath northern Sulawesi (Spakman & Hall, 2010). Sulawesi itself only formed in the Miocene, and its unique k-shape arises from being formed by an assemblage of Gondwana and Sundaland fragments, along with island arc remnants (e.g. Katili, 1978; Hall, 2011).

The region to the east of Sundaland is characterized by a system of microplates and features earthquakes that occur up to 700 km depth. This complexity is driven by the Southeast Asia-Australia collision zone, where the Sunda Arc subduction transitions to an arc-continent collision, resulting in the spectacular 180° curvature of the Banda Arc (e.g. Audley-Charles, 1968; Carter et al., 1976; Harris, 2011). Whether the oblique subduction that occurs here is caused by a single (e.g. Hamilton, 1979) or two opposing slabs from the north and south (e.g. Hall, 2002) has long been debated.

3 Methodological background: Adjoint waveform tomography

Ray tracing has traditionally been the standard data prediction approach in seismic tomography due to its mathematical simplicity and computational efficiency (Červený, 2001; Rawlinson et al., 2008). The main issues with this approximation include its inability to account for certain wave-like behavior (e.g. diffraction, scattering) and hence the requirement for smooth media (e.g. Nolet, 2008), i.e. seismic wavelength much smaller than the scale length of structure. Furthermore, ray tomographic methods only use a limited por-

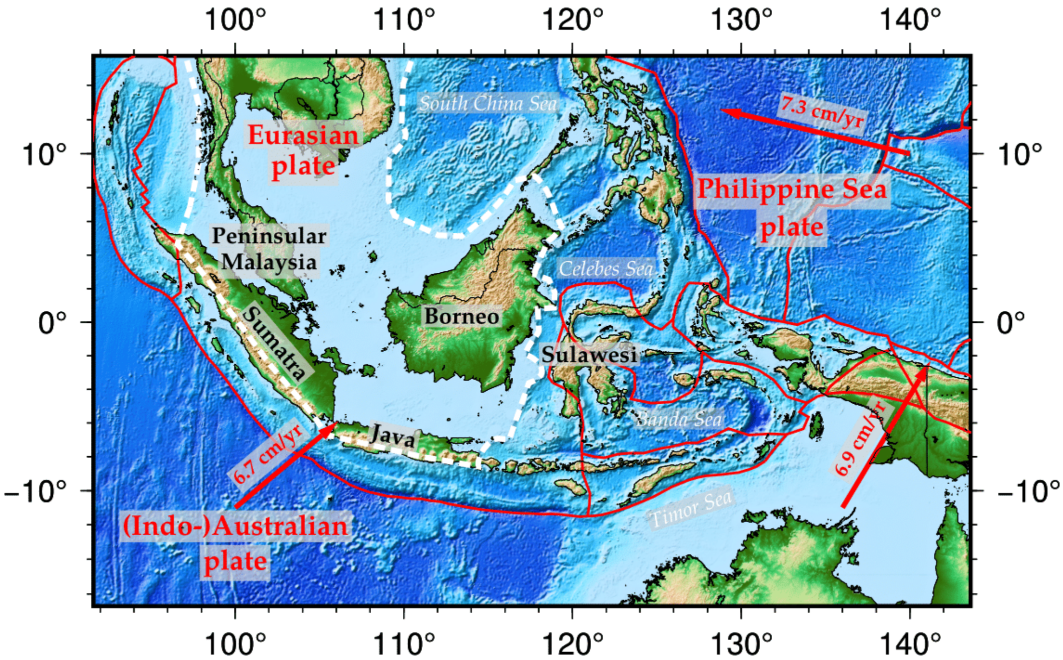

Figure 1. Map of the study area, showing the interaction of the three primary tectonic plates in Southeast Asia. The white dotted line indicates the outline of the Sundaland block. Plate tectonic boundaries are taken from Bird (2003). Plate motions are taken from *ITRF2014* (Altamimi et al., 2016). Topographic variations are taken from *ETOPO1* (Amante & Eakins, 2008).

tion of a seismogram such as phase arrival times. Adjoint waveform tomography overcomes the limitations of ray theory by solving the 3-D seismic wave equation numerically, thereby taking the often complex, volumetric sensitivity of seismic waves into account. In theory, it allows the exploitation of the full information content of seismograms and is thus frequently referred to as full-waveform inversion or FWI (Tromp, 2020).

Adjoint waveform tomography is one of the most challenging methods for obtaining information on Earth structure due to the complex, tangled workflow and non-linearity of the inverse problem. The first step is to obtain accurate synthetic seismograms from an initial Earth model for a set of specified sources by solving the 3-D seismic wave equation numerically. The synthetic waveforms are compared to the observed data using a suitable misfit measure. Then, the gradient of the misfit function is used to update the initial model in order to reduce the waveform misfit. This process is iterated until the waveform match is deemed sufficient according to some criteria.

3.1 Obtaining synthetics: 3-D seismic wave propagation

Synthetic seismograms – that is the time- and space-dependent solution of the wave equation at specified locations – are obtained by computing the 3-D wavefield through a region of interest. Seismic wave propagation through the solid Earth is governed by the elastic wave equation (e.g. Aki & Richards, 2002) and can be expressed as:

$$\rho(\mathbf{x})\frac{\partial^2 \mathbf{u}(\mathbf{x},t)}{\partial t^2} - \nabla \cdot \boldsymbol{\sigma}(\mathbf{x},t) = \mathbf{s}(\mathbf{x},t), \tag{1}$$

where ρ is density, $\boldsymbol{u}$ is displacement, $\boldsymbol{\sigma}$ is the stress field and $\boldsymbol{s}$ represents a source term. The parameters $\boldsymbol{x}$ and t indicate space and time dependencies, respectively.

A wealth of numerical techniques to calculate the 3-D wavefield have been developed over the past few decades. In full-waveform inversion, spectral-element methods (a form of finite-element methods) are currently considered to provide an optimal balance between simulation accuracy and efficiency in earthquake seismology (e.g. Komatitsch et al., 2003; Afanasiev et al., 2019), while finite-difference methods are popular in seismic exploration (e.g. Virieux, 1984; Operto et al., 2015). The spectral-element method is also preferred in earthquake seismology because of its ability to accommodate topography, bathymetry and fluid–solid boundaries, such as the ocean-crust boundary (e.g. Komatitsch & Vilotte, 1998). Throughout this study, we employ the spectral-element wave propagation solver *Salvus* (Afanasiev et al., 2019) to obtain accurate 3-D synthetic seismograms.

3.2 Quantification of waveform differences: Misfit function

The misfit function quantifies the differences between observed and predicted waveforms and is used to measure the consistency between a model and the observables used to constrain it. There are many different ways to define the difference between two seismograms and the choice can have a significant effect on the tomographic result. Consequently, quantifying waveform differences remains an active area of research in waveform tomography (e.g. Yuan et al., 2020).

The most common misfit functions used in waveform tomographic studies include a summation of the least-squares differences of the waveforms (L2, e.g. Bamberger et al., 1982) and time-shift measurements (cross-correlation and multi-taper misfit functions, e.g. Tape et al., 2010; Zhou et al., 2004). The main drawbacks are usually considerable sensitivity to outliers for the former and the assumption of similar waveforms for the latter. Consequently, time- and frequency-dependent phase misfits were proposed, where phase and amplitude information are separated (Kristeková et al., 2006; Fichtner et al., 2008). To date, most FWI studies in earthquake seismology exploit phase information from selected seismogram portions, disregarding amplitude information for reasons of source uncertainty, inadequate instrument response information and contamination caused by site effects (e.g. Tromp, 2020). However, there are ongoing developments towards true-amplitude FWI (e.g. Wang et al., 2020).

In this study, we use a time-frequency phase misfit function following Fichtner et al. (2008). It is based on the transformation of both observed and synthetic seismograms to the time-frequency domain, and makes use of both phase and relative amplitude information. The time-frequency phase misfit measure has the advantage that individual seismic phases do not need not be identified and isolated. Nevertheless, it requires the separation of small and large amplitudes, and a selection of suitable seismogram portions to avoid cycle skips and noisy portions of the data. The phase misfit χ_p can be formulated as a weighted L_2 norm of the phase difference $\phi^\mathrm{syn} - \phi^\mathrm{obs}$ for a single waveform component $\boldsymbol{u}$ as follows:

$$\chi_\mathrm{p}{}^2(\mathbf{u^{syn}}, \mathbf{u^{obs}}) = \int_{\mathbb{R}^2} W_\mathrm{p}{}^2(t,\omega)[\phi^\mathrm{syn}(t,\omega) - \phi^\mathrm{obs}(t,\omega)]^2 dt\, d\omega, \tag{2}$$

where ω denotes the angular frequency linking the phase difference $\Delta\phi$ to a time shift Δt via $\Delta\phi = \omega\Delta t$. Furthermore, W_p represents a positive weighting function that is necessary for the stability of the measurement and suppresses phase differences when no physically meaningful measurement is possible, e.g. when the signal is below the noise level (see Fichtner, 2010).

3.3 Model update: Gradient-based optimization

We aim to minimize the waveform deviation (see Section 3.2) using an iterative non-linear approach, and thus seek the first derivative of the misfit function with respect to the model parameters, which corresponds to the gradient. The misfit gradient combines all possible source-receiver combinations and is constructed from sensitivity kernels, which are obtained using adjoint techniques (Chavent, 1974). The adjoint method is a convenient and computationally feasible way of computing the gradient (e.g. Tromp et al., 2005; Fichtner et al., 2010); one of its main computational advantages is that for each source, only two numerical simulations are needed, which can utilize the same wave propagation solver. Thus the computational cost scales linearly with the number of events. While the source term for the forward wavefield is given by a seismic source, the adjoint source is fully determined by the misfit, giving rise to a fictitious wavefield. The interaction between both wavefields defines the sensitivity kernels.

The model update is computed using a gradient-based optimization scheme. In this study, we use the L-BFGS method (e.g. Nocedal & Wright, 2006), which is generally regarded as the most efficient method for waveform tomography problems (e.g. Modrak & Tromp, 2016). The L-BFGS method is a quasi-Newton method, because it employs an approximation of the inverse Hessian to obtain curvature information on the misfit landscape. Here, the Hessian approximation is based on the history of the past ten gradients since FWIs are relatively convex compared to other optimization problems and thus, the change in curvature between iterations is small. In order to determine the step size, we employ a trust-region method, which does not require any additional simulations compared to line search methods. The misfit function is quadratically approximated within a local region and this region is automatically adjusted based on the quality of the approximation that was observed in the previous iterations (e.g. Conn et al., 2000; van Herwaarden et al., 2020), that is, the region is expanded if an adequate model was found within the trust-region. Thus, no additional simulations are required to determine the step length.

In this study, the inversion parameters are restricted to those well-constrained by the intermediate-period waveform data, i.e. isotropic P-wave velocity (v_P), radially anisotropic S-wave velocity (v_{SH} and v_{SV}) and density (ρ).

4 Southeast Asian waveform tomography

4.1 Model domain

The chosen study region is centered around Borneo and encompasses Malaysia and Indonesia (see Figure 1). It comprises an area of approximately 6,000 km in the east-west, 3,500 km in the north-south and 800 km in the depth direction. For gradient-based optimization schemes, the starting model needs to be sufficiently close to the true model in order to avoid entrapment in local minima. For this study, we adopt the *Collaborative Seismic Earth Model (CSEM)* introduced by Fichtner et al. (2018), which is a modified version of the 1-D anisotropic *PREM* (Dziewonski & Anderson, 1981), since no region-specific model is currently available. The model is designed to be conservative in the sense that it only contains the least complex structure that seismic data are sensitive to, e.g. the Lehmann discontinuity was replaced by a linear gradient and the elastic properties of the lower crust have been extended to the surface. For Southeast Asia, *CSEM* is an acceptable starting model since it still matches our longest-period (100 – 150 s) data to within half a cycle. The starting model is presented in the Supplementary Material in Section 1.

To further mitigate the risk of converging towards a local minimum and to avoid cycle skips, a multi-scale approach (Bunks et al., 1995) is employed, where the longest periods are inverted for first (100 – 150 s), and shorter period content is successively added (down to 20 – 150 s). It follows that the simulation mesh needs to become denser as the iterations progress to accurately sample the wavefield at shorter periods. Here, we use the Python package *MultiMesh* (Thrastarson et al., 2021) for the mesh interpolation between different period bands.

4.2 Event and data selection

Most seismicity within the region occurs along the highly active plate boundaries, resulting in a heterogeneous event distribution (see Figure 1). In order to select suitable events for waveform tomography, we observe that only those earthquakes of $M_w \geqslant 5.5$ have sufficient energy to generate high signal-to-noise-ratio waveforms at distant receivers within the domain. Furthermore, we preferentially select earthquakes with a large number of useful recordings to enhance the efficiency of the adjoint-based inversion. We also removed events potentially affected by interference; that is if another event of $M_w \geqslant 7.0$ occurred elsewhere in the world, or a $M_w \geqslant 5.0$ event occurred in an extensive area encompassing the domain at a similar time. Moreover, using multiple earthquakes that occur in a similar location does not improve the inversion result, but will significantly increase compute time (see Section 3.3), so we are careful to include only those events that are likely to contribute meaningfully to the final model. Based on these criteria, the final event catalog contains 143 earthquakes with magnitudes between $5.5 \leqslant M_w \leqslant 7.5$, which occurred between 2008 and 2020. Earthquakes with $M_w > 7.5$ are disregarded since the point source assumption is strictly not valid anymore; many large-magnitude earthquakes are characterized by rupture durations of several seconds and rupture lengths of tens of kilometers (e.g. Wells & Coppersmith, 1994).

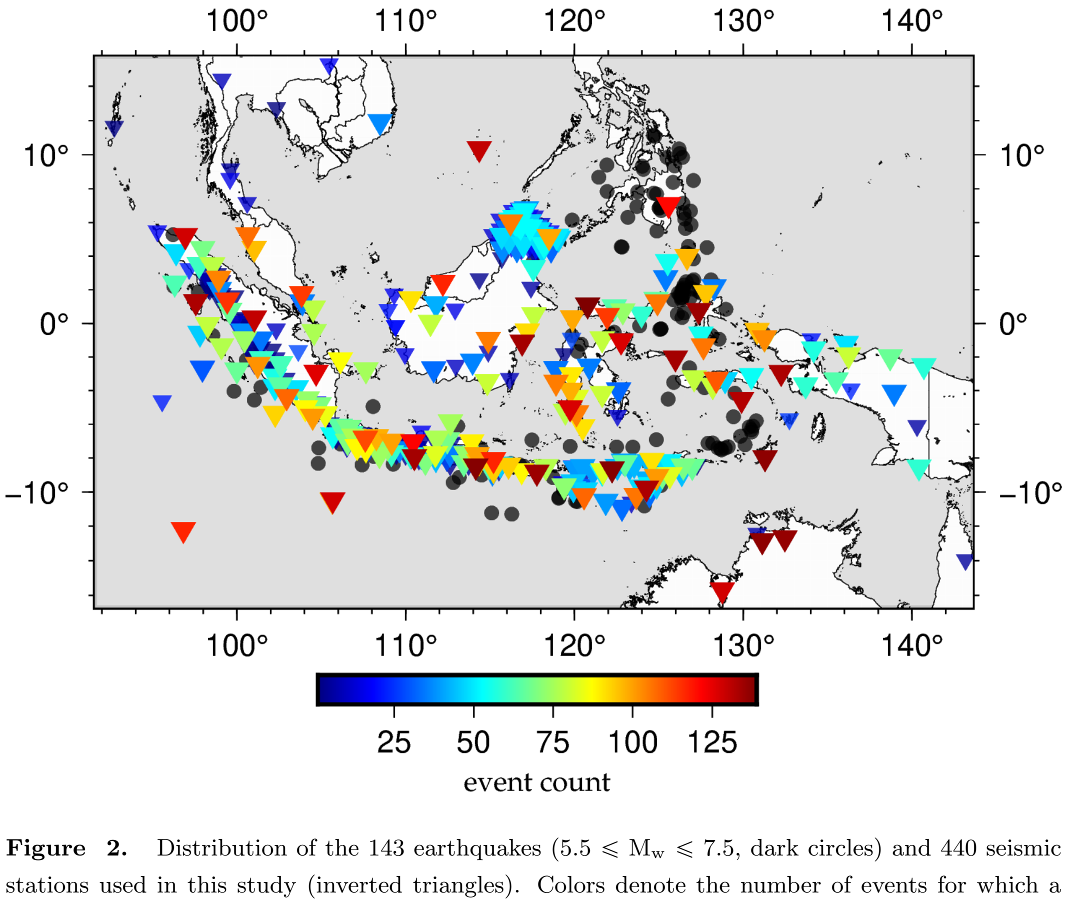

Figure 2. Distribution of the 143 earthquakes ($5.5 \leqslant M_w \leqslant 7.5$, dark circles) and 440 seismic stations used in this study (inverted triangles). Colors denote the number of events for which a given station contributes waveforms to the inversion. Stations with a number of events < 30 are plotted in a smaller size. These are usually temporary arrays deployed over a short period of time. The maximum source-receiver distance is ∼5,600 km.

Event locations and moment tensors are retrieved from the *GCMT* catalog (Ekström et al., 2012) and remain constant throughout the inversion. To mitigate finite-source effects contaminating the tomography, large-magnitude earthquakes are removed and lower-magnitude events are added as the period content is decreased. Furthermore, source time functions are reviewed using *SCARDEC* (Vallée, 2013) and are removed, if necessary. We find that long source time functions correlate with large event misfits. The number of events used for each period band can be found in Table 1.

To date, data from only a relatively small proportion of permanent network stations in Southeast Asia have been made publicly available. We have been able to include data from several networks with restricted access, resulting in an unprecedented dataset that comprises recordings from 440 on-shore stations within this region. Figure 2 shows the station and event distribution for this study. A detailed overview of the selected events, available stations and how waveform data was accessed is available in the Supplementary Material in Section 2.1 and Section 3.

The inclusion of temporary networks within the region results in a highly uneven geographical station distribution since they tend to target features of particular interest and are therefore closely spaced. Thus, we implement geographical station weighting as proposed by Ruan et al. (2019) in order to minimize the effect of dense regional networks. Under this scheme, a station is assigned a larger weight if it has few nearby stations, and vice versa.

4.3 Inversion setup

We use the *Salvus* software package (Afanasiev et al., 2019) for the mesh generation, forward and adjoint simulations and non-linear optimization, within its integrated workflow. Accurate synthetics are obtained using *Salvus*' built-in spectral-element wave propagation solver, which approximates the frequency-dependence of attenuation with five linear solids (e.g. van Driel & Nissen-Meyer, 2014; Afanasiev et al., 2019). Furthermore, topography and bathymetry are implemented across all period bands using *Earth2014* (Hirt & Rexer, 2015). The fluid ocean is approximated by the weight of its water column (Komatitsch & Tromp, 2002). We find this to be a valid assumption for this study compared to explicitly modeling the fluid ocean by replacing it by acoustic elements, which is computationally more expensive.

Synthetic and observed seismograms are compared using time-frequency dependent phase misfits as described in Section 3.2. This still requires us to define the parts of a seismogram suitable for waveform comparison (*windows*), avoiding noisy portions of the seismograms as well as phase jumps, which would contaminate the tomography results. Furthermore, many misfit functions favor large-amplitude signals, which in most cases come from surface waves, thus making the recovery of deep structure challenging. Therefore, we maximize sensitivity to deep structure by specifically accounting for body wave signals in a separate window (see Figure 4). Note that the challenge of resolving deep structure is also a consequence of the relatively long periods currently considered in FWI.

The data selection algorithm *FLEXWIN* (Maggi et al., 2009) is employed using its Python port *Pyflex* (Krischer & Casarotti, 2015) to suggest windows, but selecting meaningful windows on noisy traces and an automated separation of body and surface wave arrivals is challenging. Thus, we found it necessary to manually review the suggested windows for each period band in order to 1) exploit as many waveforms as possible, and 2) properly separate small and large-amplitude arrivals to enhance depth sensitivity. This is by far the most time-consuming part of the inversion setup, but it triples the analyzed window lengths compared to the tuned *FLEXWIN* algorithm.

We use ∼13,000 unique source-receiver pairs and a total analyzed time window length between 1,000 – 3,000 h per period band (see Table 1). We attribute the increasing number of measurements during the initial period bands (I – III) to the increasing number of windows

Table 1. Summary of the data selection including the number of events, seismogram traces, selected windows, average number of windows per event, percentage of traces with windows, total window length in hours, the average window length per event in hours and the number of unique source-receiver pairs per period band.

period band	# events	# traces	# windows	avg. # windows per event	% traces w/ windows	$\sum$ window length [h]	avg. window length per event [h]	# unique s-r pairs
100 s (I)	118	67,401	20,594	175	22,4	2,306	19.5	10,312
80 s (II)	118	67,317	25,614	217	27,5	2,995	25.4	11,604
65 s (III)	118	68,460	26,988	229	28,6	3,103	26.3	12,269
50 s (IV)	117	64,449	25,583	219	28,7	2,711	23.2	12,060
40 s (V)	106	58,464	32,081	302	38,1	2,586	24.4	12,960
30 s (VI)	83	44,787	26,679	321	40,9	1,519	18.3	10,279
20 s (VII)	71	38,727	22,683	319	40,6	1,064	15.0	8,656

that meet the selection criteria after the model has improved. From 50 s onwards, body wave signals become clearly identifiable, contributing approximately 30 % of the total windows in the final period band. However, windows around body wave arrivals are much shorter and the surface wave train becomes more compact as the minimum period is decreased. Thus, the overall analyzed window length per event decreases despite the increasing number of windows. From 30 s onwards, the 3-D wavefield becomes increasingly complex (e.g. due to crustal scattering), which in turn allows us to use a smaller number of events and select fewer windows. Nevertheless, the number of windows per event and the number of traces with windows almost double from the initial to the final period bands, indicating that we are successively including more data per event as the iterations progress.

For each event, the waveform misfits for all windows and traces are summed to produce the event gradient. The raw gradients are preconditioned before the descent direction for the model update is computed in order to help mitigate the ill-posedness of the inverse problem. Furthermore, preconditioning can provide significant overall computational savings by accelerating the convergence of the optimization algorithm (e.g. Modrak & Tromp, 2016; Liu et al., 2020). Here, we apply a two-stage preconditioning:

1. Source and receiver imprints are removed for each event gradient because they usually show strong localized sensitivity in these areas.
2. The event gradients are summed to produce the misfit gradient before applying an anisotropic, depth-dependent, diffusion-based smoothing operator as described by Boehm et al. (2019), preventing sub-wavelength structure from entering the model.

An example plot as well as a table presenting the smoothing lengths and removed imprint radii per period band are provided in Section 4 of the Supplementary Material.

The inversion parameters are restricted to those well-constrained by the intermediate-period waveform data, i.e. isotropic P-wave velocity (v_P) and radially anisotropic S-wave velocity (v_{SH} and v_{SV}). We also include density (ρ) as an inversion parameter in order to avoid artifacts (Blom et al., 2017), but do not interpret these results (e.g. Blom et al., 2020). More information on technical parameters of the inversion setup can be found in Section 5 in the Supplementary Material.

5 Results

5.1 Misfit development

A total of 87 inversion iterations divided over seven period bands between 20 and 100 s were carried out (see Table 1). The inversion process was performed on a supercomputer and required > 50,000 CPU hours, half of which were used during the final period band, which can be attributed to the denser wavefield sampling at shorter periods. Shorter period data is added once the misfit decrease stagnates or the number of events that decrease their misfit significantly drops. Each broadening of the period band is accompanied by a mesh interpolation and data review (events and windows).

The misfit development for all seven period bands used in this study is displayed in Figure 3. The overall misfit decrease is remarkable, which we partially attribute to the 1-D starting model leaving a lot of room for improvement. The first period band yields the greatest misfit decrease of > 40 %; the initial model updates focus on including a regional, low-velocity zone, the need for which was already apparent from strongly delayed observed waveforms (see Figure 4). 30 % of the misfit decrease within this period band is achieved during the first iteration, indicating that regional updates can be accounted for within one or two iterations as previously suggested by Fichtner et al. (2018).

For many period bands (e.g. 80 s, 65 s, 50 s, 40 s), we observe a strong misfit decrease for the second iteration, which we believe to be the result of the trust-region based L-BFGS optimization scheme used in this study (see Section 3.3). In this scheme, the initial search direction is equivalent to the steepest descent method since no additional information about the misfit landscape, other than the gradient, has yet been obtained. From the second iteration onwards, the approximation of the Hessian is taken into account and the trust-region is adjusted, which speeds up convergence.

In the final period band, no single event (out of 71) contributes more than ∼3 % to the misfit decrease between initial and final model, indicating that the inversion is not dominated by a few events. None of the events increase their misfit, and no geographical misfit pattern is identifiable, nor is any correlation with depth, magnitude or focal mechanism (see Section 2.2 in the Supplementary Material).

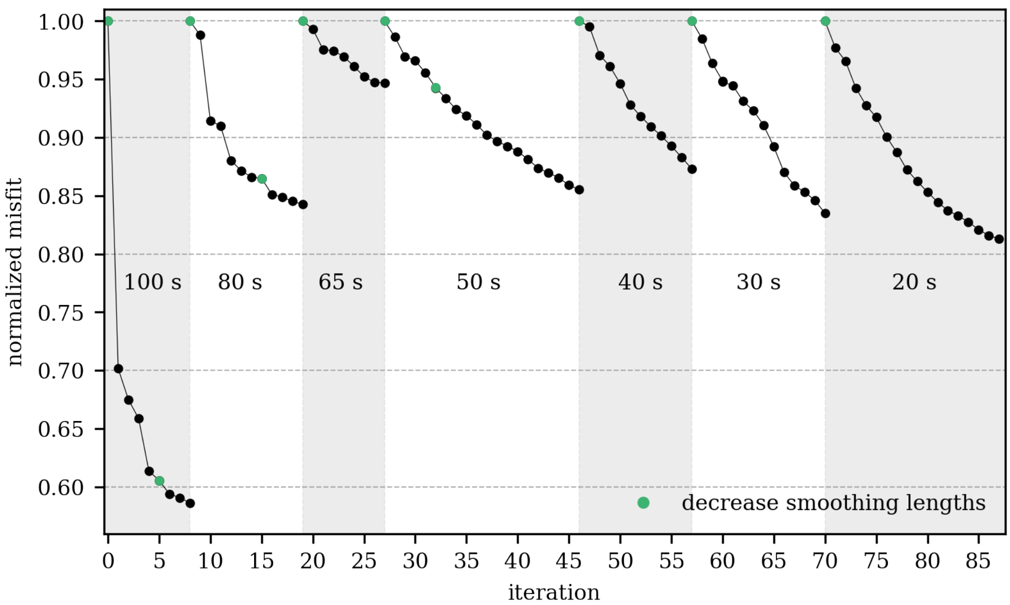

Figure 3. Misfit development across 87 iterations, normalized by the initial misfit within each period band. Green dots indicate a smoothing length decrease. Each broadening of the period content is accompanied by a mesh interpolation and data review.

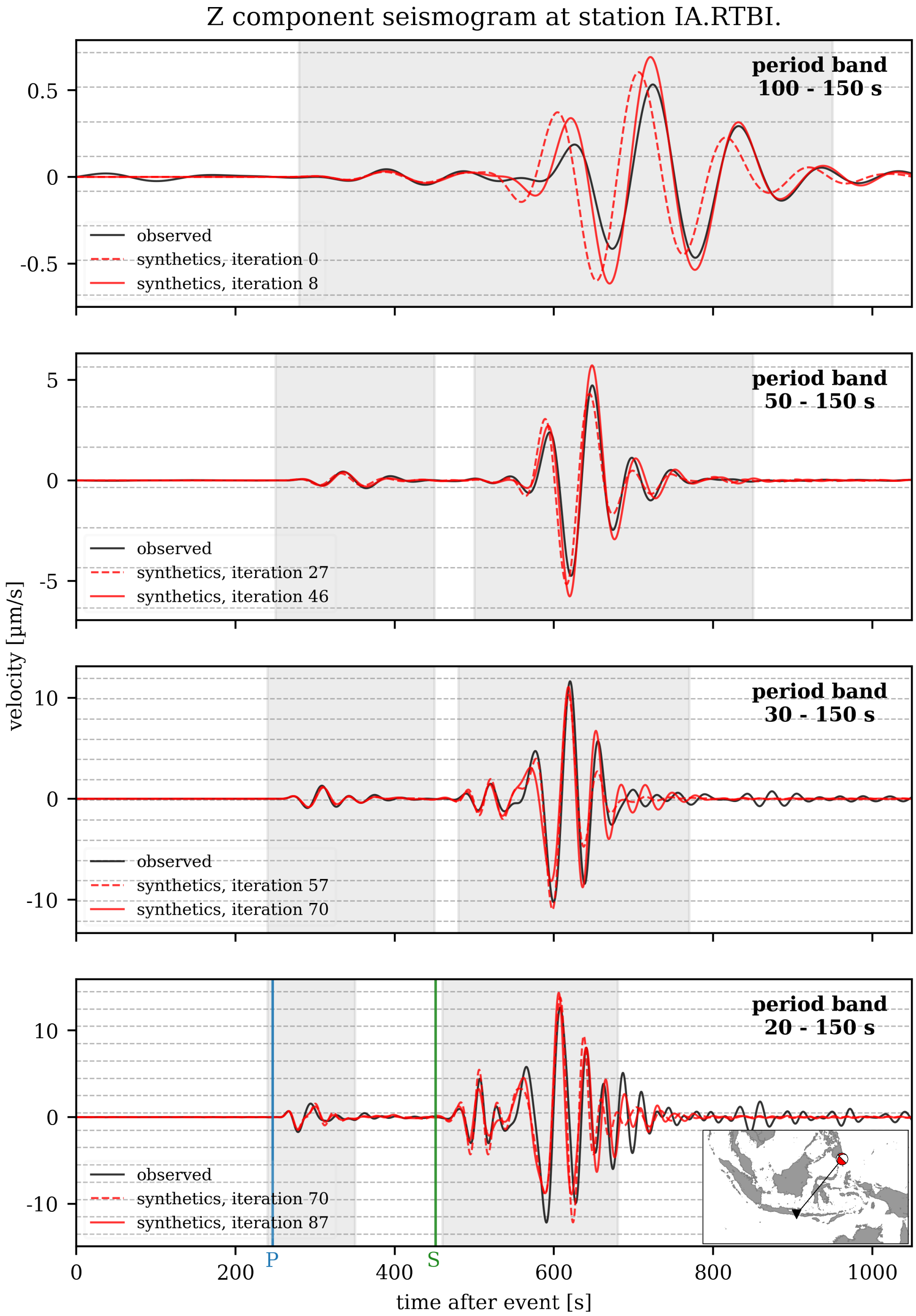

Figure 4. Waveform match improvement across four of the seven period bands for the vertical component of a station in Bali, Indonesia, which recorded a $M_w6.2$ event south of the Philippines. For each period band, the final synthetics (solid red) match the observed waveforms (black) better than the synthetics from the initial iteration (dashed red). From 50 s onwards, an additional window around a smaller amplitude arrival can be selected. Vertical lines indicate predicted P- (blue) and S-wave (green) first arrival times obtained from the *TauP* toolkit (Crotwell et al., 1999) for *PREM* (Dziewonski & Anderson, 1981).

5.2 Waveform match improvement

The misfit development described in the previous section is entirely driven by a waveform match improvement. Figure 4 presents the waveform comparison across four of the seven period bands for the vertical component of a single source-receiver pair. While the majority of windows are selected on the vertical component, 33 – 42 % of the windows per period band are selected on horizontal components. Strong initial delays of observed waveforms with a particularly large time shift at 100 s are observed, indicating that the starting model is too fast for the region. From 50 s onwards, the data fit is already excellent for the initial iteration and we are able to include an additional window around a smaller amplitude arrival. For the final period band at 20 s, we achieve an overall misfit decrease of > 50 % for the entire dataset compared to the initial model. Note that we are able to explain true-amplitudes despite only utilizing relative amplitude information throughout the inversion (see Section 3.2). More waveform fits are provided in Section 6 of the Supplementary Material.

5.3 Model assessment

In traditional ray theory tomography, the checkerboard test is popular (e.g. Rawlinson & Spakman, 2016), but it is computationally prohibitive in FWI. Consequently, obtaining reliable information on model uncertainty information remains an active area of research in adjoint waveform tomography (e.g. Liu et al., 2020). To date, many studies employ spike tests and random probing (Fichtner & Leeuwen, 2015) for resolution analysis. However, it is possible to pursue more data-driven approaches towards validating the model, as described below.

5.3.1 Misfit decrease and analyzed window lengths

The waveform match improvement across the ensemble of period bands (see Figure 4) and the associated misfit decrease of > 50 % indicate that the new model satisfies the data significantly better than the starting model. This is reinforced by computing *FLEXWIN* windows for the starting and final model at 20 s (in order to avoid time-consuming manual window picking for the starting model), which results in a doubling of window lengths in the latter case, thus indicating that our final model explains observed waveforms significantly better than the starting model.

5.3.2 Ability to satisfy unused data

We tested the validity of our model by selecting ten earthquakes (M_w5.5 – 6.5) that were not used in the tomography, including events in unique locations around Sulawesi and Western New Guinea. The 3-D synthetics through the final model result in an event misfit decrease that is only 3 % lower compared to data used in the actual inversion. Figure 5 shows that synthetics obtained from our final model are able to explain horizontal and vertical components as well as body and surface wave arrivals. For comparison, we also show the synthetics obtained from the starting model at this period.

5.3.3 Hessian-vector product analysis

Uncertainty quantification based on exploiting the inverse Hessian H^{-1} is currently prohibitively expensive to handle in FWI. Consequently, several studies have analyzed the Hessian-vector product $H\delta m$ for a test function δm (e.g. Fichtner & Leeuwen, 2015), e.g. by approximating $H\delta m$ with gradient differences (e.g. Krischer et al., 2018; Gao et al., 2021). However, this is built upon the assumption that the inversion has reached convergence and requires additional simulations. Since we have already constructed an approximation of the Hessian with L-BFGS during the inversion, we can directly apply this to a model perturbation in order to obtain a qualitative analysis of inter-parameter trade-offs.

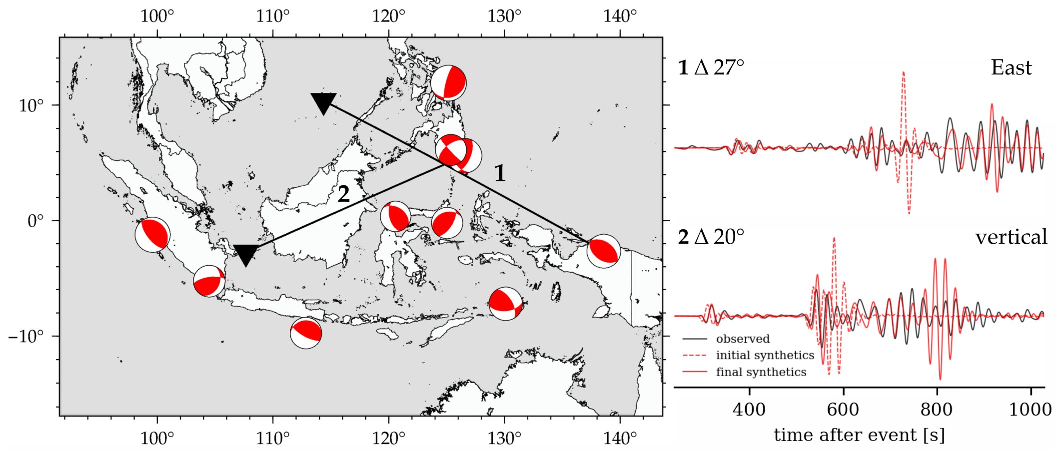

Figure 5. *Left:* Map of the validation dataset consisting of ten earthquakes of M_w5.5 – 6.5 with a relatively even spatial distribution. *Right:* Horizontal and vertical component seismograms at two different stations with epicentral distances of 20 and 27 degrees.

Figure 6 presents a visualization of the Hessian-vector product for a v_{SV} perturbation. This reveals that the model is most sensitive to changes beneath the Sundaland block and around the northward continuation of the North Australian craton, as expected from the data coverage (see Figure 2). The inversion appears to suffer from some cross-talk between parameters, which is more pronounced for v_{SH} and density than for v_P, and is weaker at greater depths.

5.4 *SASSY21*

After 87 iterations, the model is updated considerably for all inversion parameters down to the transition zone. Figure 7 shows the depth-averaged perturbations, which reveal mostly negative anomalies for seismic wave parameters. P-wave structure is updated the least – around -1 % in the upper 200 km –, while horizontal shear-wave velocity and density exhibit similar behavior in their updates. This lack of suspicious behavior is reassuring, because both parameters are difficult to constrain during the inversion, since they are less sensitive to the data than v_{SV}.

The model updates are strongest near the surface, and decrease in strength with depth. This can be attributed to most sources and all receivers being located near the surface and the sensitivity of surface waves decaying with depth. We attribute the somewhat linear variation in elastic parameters in the upper ∼70 km (see Figure 7) to the wavelength at 20 s, that is seismic waves at this period are sensitive to the bulk crustal structure (e.g. Capdeville et al., 2010). The kink at 70 km does not coincide with a mesh element boundary.

We observe strong perturbations in v_S, in particular for the v_{SV} parameter in the upper ∼100 km. This is because at the relatively long periods considered, the wavefield is dominated by surface waves, which are strongly sensitive to shear-wave structure. Thus, the subsequent discussion will be based on the S-wave model since it is better constrained. In the following, v_S is defined as the Voigt average: $v_\mathrm{S} = \sqrt{(2{v_{SV}}^2 + {v_{SH}}^2)/3}$ (e.g. Babuska & Cara, 1991; Panning & Romanowicz, 2006). The results for other inversion parameters are presented in Section 7 in the Supplementary Material.

Figure 8 shows v_s depth slices from 50 to 700 km through the final model, which is dominated by low v_S at shallow depths. At 50 km depth, the oceanic lithosphere beneath

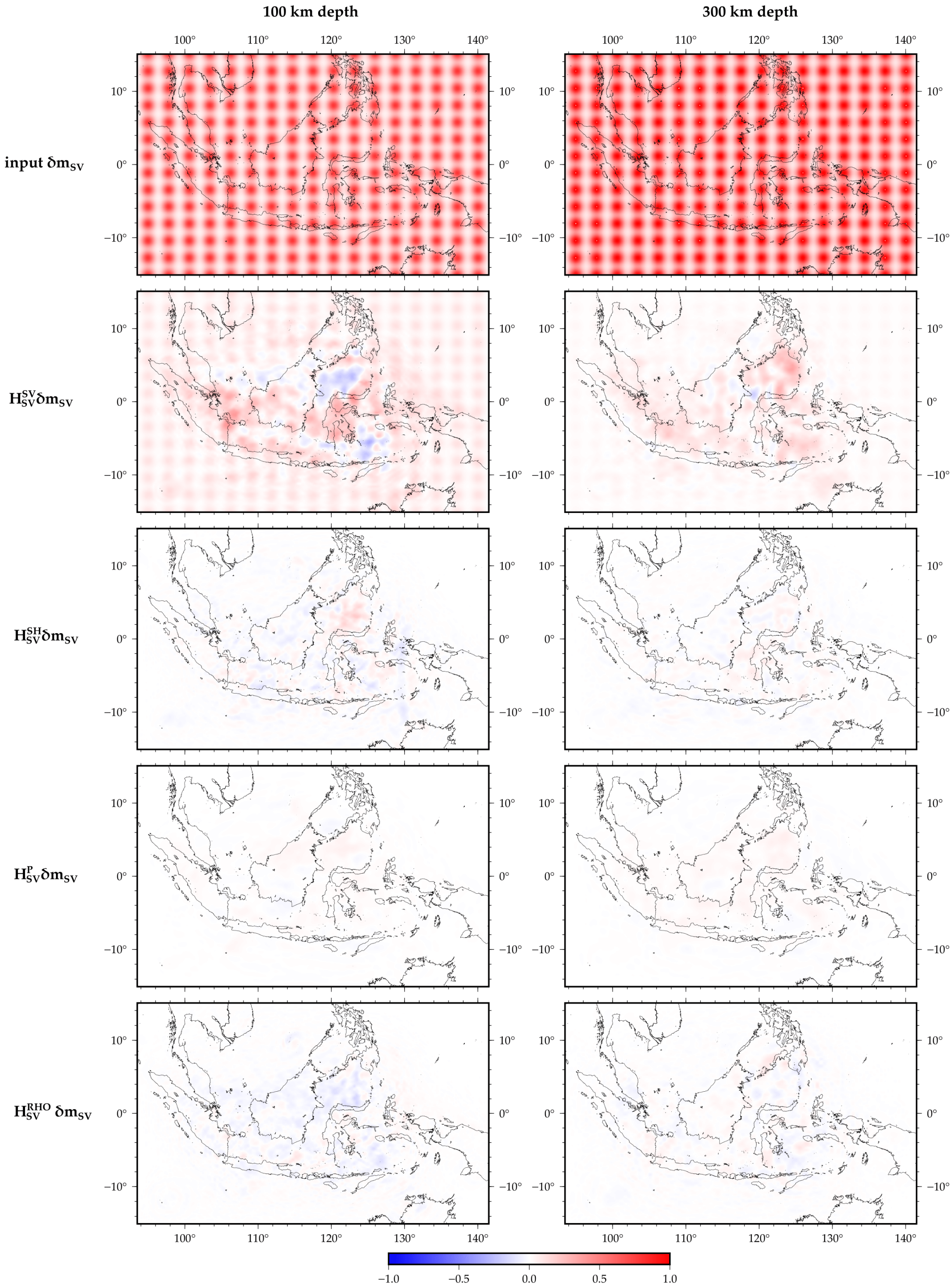

Figure 6. Visualization of $H\delta m_{SV}$ for all inversion parameters at 100 and 300 km depth. *Top panel:* Depth slices of the input perturbation: a 3-D checkerboard pattern of Gaussian v_{SV} spheres with a standard deviation of 70 km. *Panels below:* $H\delta m$ for all inversion parameters (v_{SV}, v_{SH}, v_P and ρ) relative to the input and normalized to $H_{SV}^{SV}\delta m$.

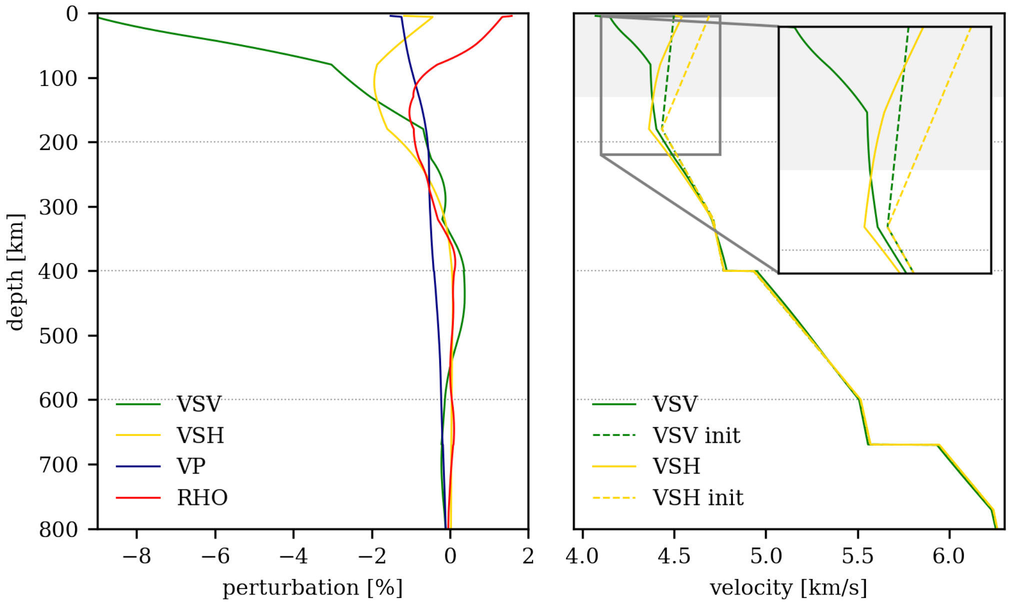

Figure 7. *Left:* Depth average of the magnitude of the relative difference between initial and final model for all inversion parameter classes. *Right:* The depth-averaged absolute v_{SV} and v_{SH} values for the initial and final model, including a zoom-in for the upper 220 km. The grey highlighted area denotes depth values with positive radial anisotropy ($v_{SH} > v_{SV}$) in the final model. The absolute values for v_P and ρ can be found in the Supplementary Material in Section 1.

the Banda Sea in the east, the (Indo-)Australian plate in the southwest and the Celebes Sea north of Sulawesi are faster than the Sundaland block, which mainly consists of continental crust. We assume the slab is not visible at shallower depths due to the limited data coverage. Further tests confirmed that this is not a result of the source and receiver imprint removal applied to the gradients described in Section 4.3. At greater depths, the most prominent feature is a high-velocity zone that follows the Indonesian volcanic arc, which is interpreted as the descending (Indo-)Australian plate. In the following, we will discuss some of the key features of the final model in more detail.

6 Discussion

6.1 Regional, anisotropic low-velocity zone

The initial model updates focus on including regional-scale, low velocities for P- and S-wave structure in the upper ∼200 km, with particularly strong perturbations in the upper ∼150 km (see Figure 7). The low lithospheric velocities are consistent with previous tomographic studies (e.g. Van der Hilst et al., 1997; Lebedev & Nolet, 2003; Zenonos et al., 2019) and other measurements such as high heat flow (e.g. Artemieva & Mooney, 2001). This suggests a thin, warm and weak lithosphere, which may be the result of long-term subduction beneath Sundaland (e.g. Hall & Morley, 2004).

The low-velocity zone is characterized by strong radially anisotropic values of up to 18 %. For the upper 130 km, we observe overall positive radial anisotropy ($v_{SH} > v_{SV}$), which transitions to negative radial anisotropy ($v_{SV} > v_{SH}$) at greater depths. The absolute, depth-averaged v_{SH} and v_{SV} values can be found in Figure 7, while Figure 9 presents lateral variations at 150 and 250 km depth. These results reveal negative radial anisotropy along the

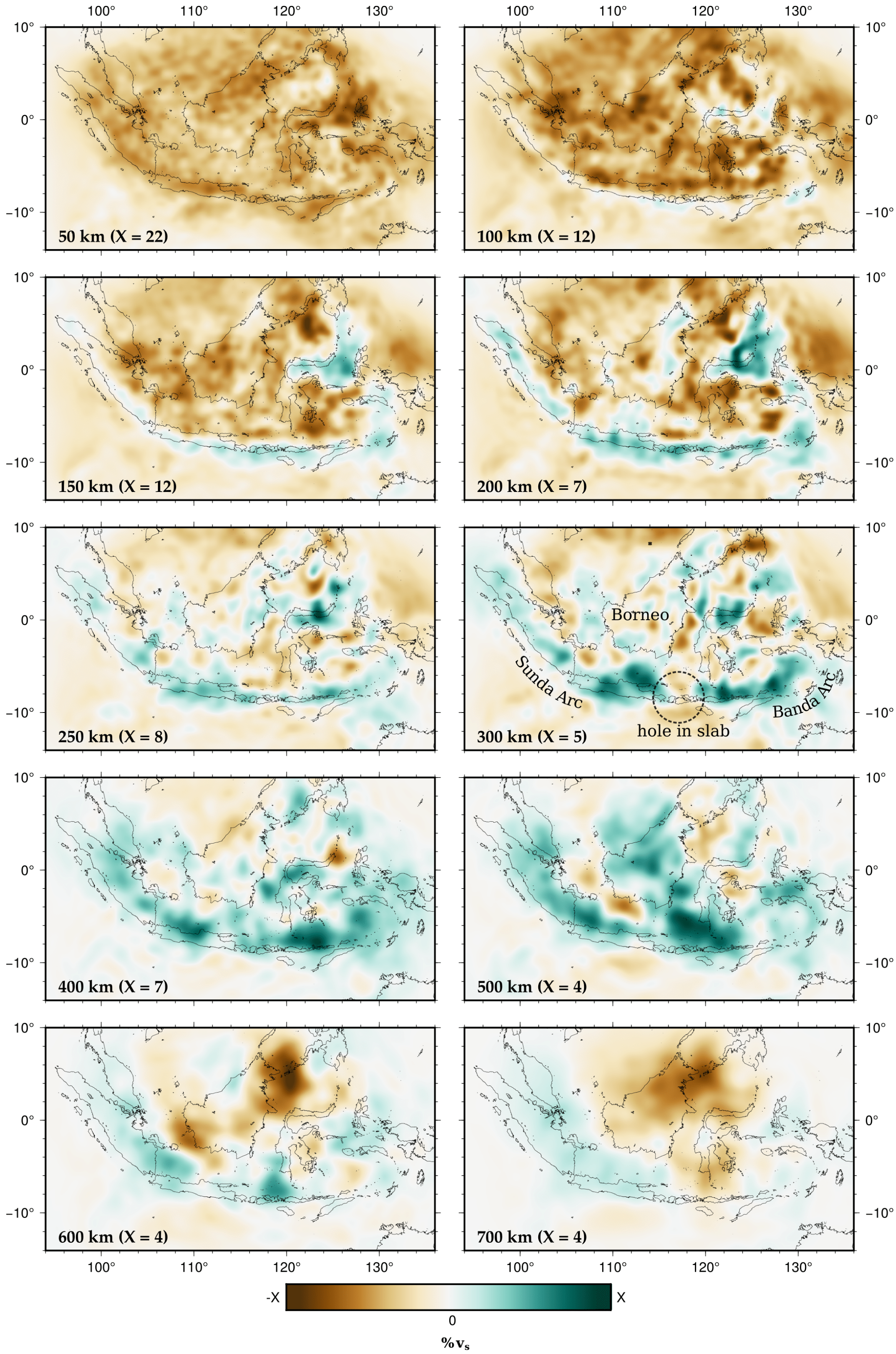

Figure 8. Shear-wave (v_S) depth slices between the range 50 and 700 km. Perturbations are in % relative to the initial model. The limits of the colourscale X are shown in the lower left corner of each plot.

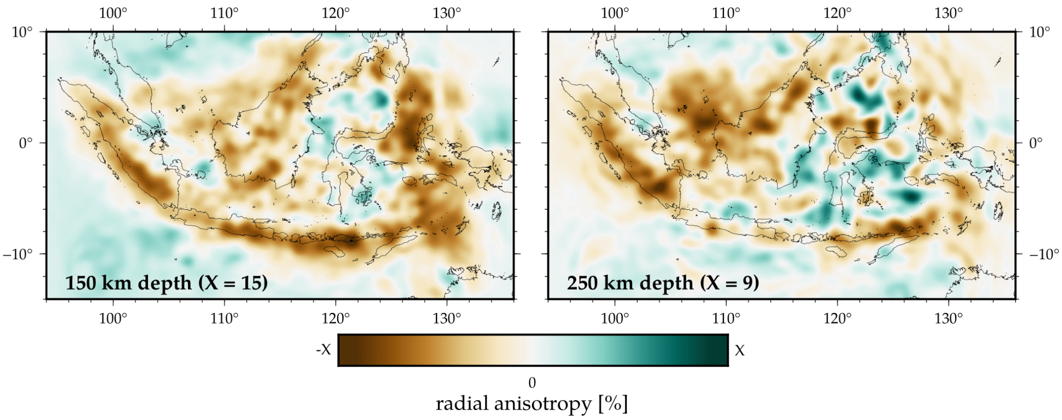

Figure 9. Radial anisotropy ($\frac{v_{\mathrm{SH}}-v_{\mathrm{SV}}}{v_{\mathrm{S}}}$) in % for the final model at (left) 150 km (right) and 250 km depth. The limits of the colourscale X are shown in the lower left corner of the plots.

slabs (which is in good agreement with Sturgeon et al., 2019) and beneath Sundaland, and positive radial anisotropy around the Celebes Sea, Sulawesi and the Banda Sea. We believe this to be the result of two different mechanisms: 1) the oceanic (Indo-)Australian plate consists of horizontally aligned minerals, which then rotate into subvertical orientations during subduction and/or entrain the surrounding mantle and induce vertical flow, thus explaining negative values along the slab (Song & Kawakatsu, 2012), and 2) negative frozen-in anisotropy of continental-lithosphere roots during formation (Priestley et al., 2021), thus explaining negative values beneath the Sundaland block.

However, it should be noted that a detailed interpretation of the anisotropy pattern is complicated by the differing sensitivities of Love and Rayleigh waves. Furthermore, it has been shown that the current resolving power of seismic tomography is insufficient to distinguish between “intrinsic“ (produced by the crystallographic preferred orientation of minerals) and “extrinsic“ (produced by other mechanisms such as fluid inclusions, fine layering or partial melting) seismic anisotropy (Fichtner et al., 2013). Thus, we refrain from a more detailed geological interpretation of the radial anisotropy.

6.2 Subduction along the Indonesian volcanic arc

The most prominent feature of the final model is a high-velocity structure following the Sunda Arc and the 180° curvature of the Banda Arc, which can be associated with the descent of the (Indo-)Australian plate (see Figure 8). The slab first becomes apparent at 50 s, which we largely attribute to body wave arrivals becoming clearly identifiable at this period (see Section 4.3). They become sharper and more intense as the dominant period is decreased (see Figure 10).

The depth slices in Figure 8 show the Sunda slab descending at depths $\geqslant$ 100 km down to the mantle transition zone. Further east, the bending of the Banda Arc is imaged as one continuous slab at 200 km depth. A geodynamic modeling study by Moresi et al. (2014) potentially supports the interpretation of a single bent and deformed slab by modeling how the curvature of this system could have developed from northward motion of the (Indo-)Australian plate. In the southeast, hints of this northward continuation of the North Australian craton can be observed, which is in good agreement with Fichtner et al. (2010).

The bottom panel of Figure 10 presents an east-west cross-section through Java and the bending point of the Banda Arc for the final model (20 s), which shows the continuation of the Sunda slab in the west down to the mantle transition zone. In the east, the Banda slab is associated with deep seismicity and appears to stagnate before penetrating through the 410 km discontinuity, although this does not align with the seismicity. For the upper ∼100 km, we can distinguish between high velocities arising from the oceanic lithosphere of the (Indo-)Australian plate and low velocities within the Sundaland block as expected from the large, thick Cenozoic sedimentary basins in this region. Figure 11a shows a south-north cross-section through Sumatra, revealing a steeply dipping Sunda slab and low velocities for the Sundaland block. Further east, Figure 11d shows opposed subducting slabs around the Banda Sea, and an oblique view of the descending slab along the Philippine Trench in the north, which is associated with elevated seismicity.

6.3 Hole in slab beneath Mount Tambora

The 300 km v_s depth slice in Figure 8 reveals a hole in the slab east of Java, roughly beneath Mount Tambora. The existence of this hole was previously suggested based on ray tomographic studies (e.g. Widiyantoro et al., 2011; Hall & Spakman, 2015; Zenonos et al., 2019) as a feature caused by slab necking and hence tearing as a result of the transition from oceanic to continental crust towards the Southeast Asia-Australia collision zone. However, based on a regional finite-frequency teleseismic P-wave tomographic model, Harris et al. (2020) concluded that there is no evidence for slab tearing in this transition region. Instead, the hole may be associated with the pertusion of continental lithosphere via entrainment of subducted plateau material (e.g. Keep & Haig, 2010). This would align with isotopic signatures indicating continental contamination in this region as previously observed by Turner et al. (2003) and Elburg et al. (2004). Figure 11b shows the hole in a cross-section, which has dimension of ∼ 300 x 100 km.

6.4 High-velocity zone(s) beneath Borneo and Sulawesi

We image a high-velocity zone beneath northern Borneo, which extends from 100 to 300 km depths (see Figure 8). A similar anomaly was imaged previously in ray tomographic studies (Hall & Spakman, 2015; Zenonos et al., 2019), but was regarded as suspicious owing to the poor data coverage. However, our study uses data from a dense, regional network in this region (see Figure 2) and we thus argue that this feature is likely not an artifact. Previous studies suggest that this anomaly may be associated with remnant subduction (e.g. Cottam et al., 2013; Hall, 2013).

Further south, the tomography reveals an S-shaped anomaly in Kalimantan (southern Borneo), which has not been imaged previously and extends from 150 to 300 km depth (see Figure 8). The anomaly appears connected with the one identified beneath northern Borneo. The absence of seismicity in the area suggests that both features may indicate remnant subduction, which is consistent with the known Neogene history of northern Borneo (e.g. Cottam et al., 2013; Hall, 2013). The S-shaped anomaly beneath Kalimantan may be associated with underthrusting from the accretion of Sulawesi in the east during the Miocene (e.g. Hall & Wilson, 2000).

Sulawesi itself is seismically highly active and located within the tectonically most complex part of the study region. Figure 11c shows the (Indo-)Australian plate descending beneath Timor in the South and the almost vertically dipping slab beneath the North Sulawesi Trench at the northern arm of Sulawesi. The slab extends down to 410 km depth, while *SLAB2.0* (Hayes et al., 2018) tracks it down to only 240 km.

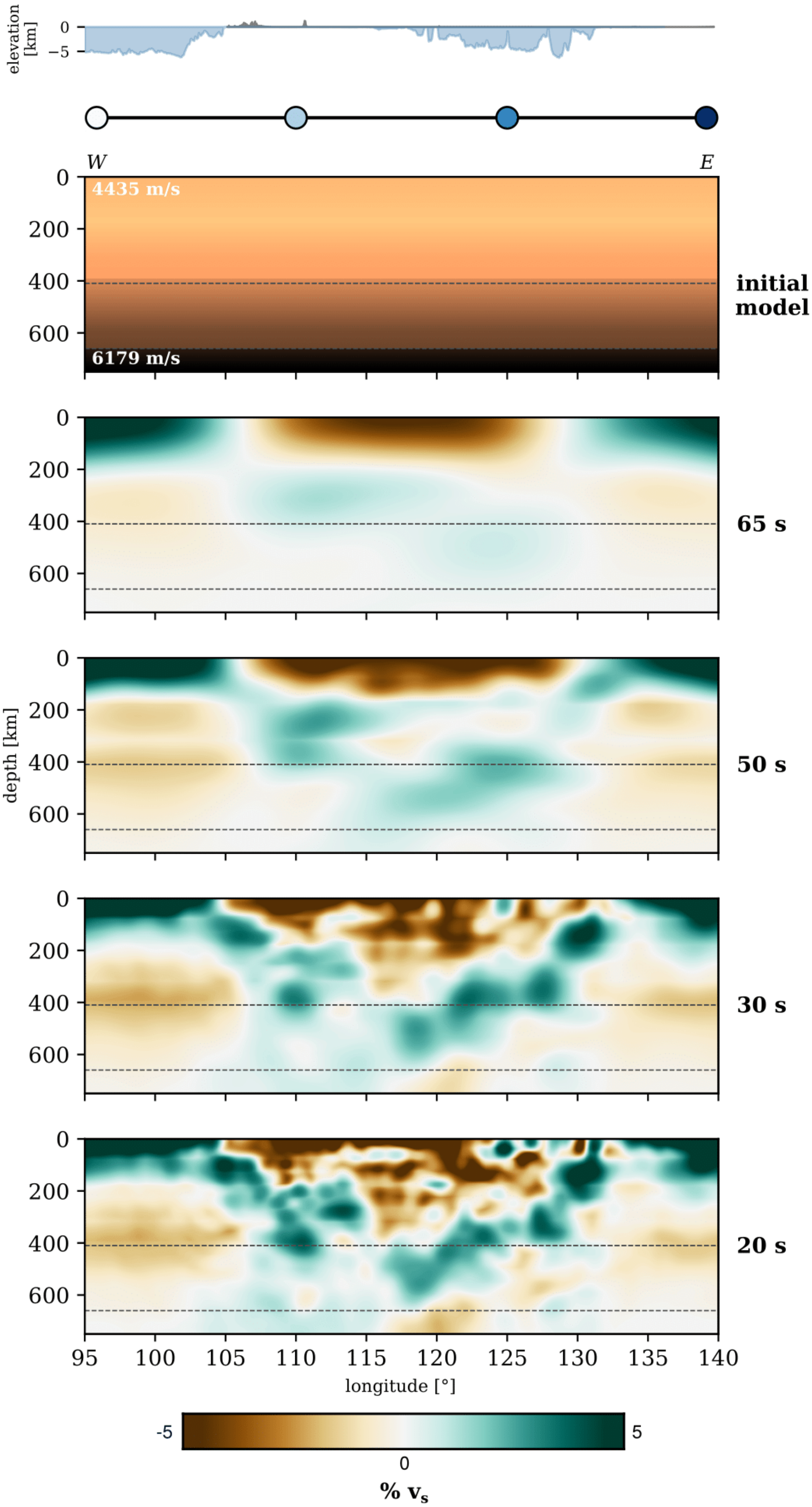

Figure 10. West-east v_S cross-section across different period bands. The section's location corresponds to the red dotted section in Figure 11. The top plot shows the absolute values of the initial model, while the other plots show perturbations from the depth-average in % for the final iteration within the respective period band. Earthquake locations (red dots in bottom plot) are taken from the *ISC* catalog (International Seismological Centre, 2016) and are within 25 km from the cross-section slice. The vertical-horizontal ratio is 1:2.

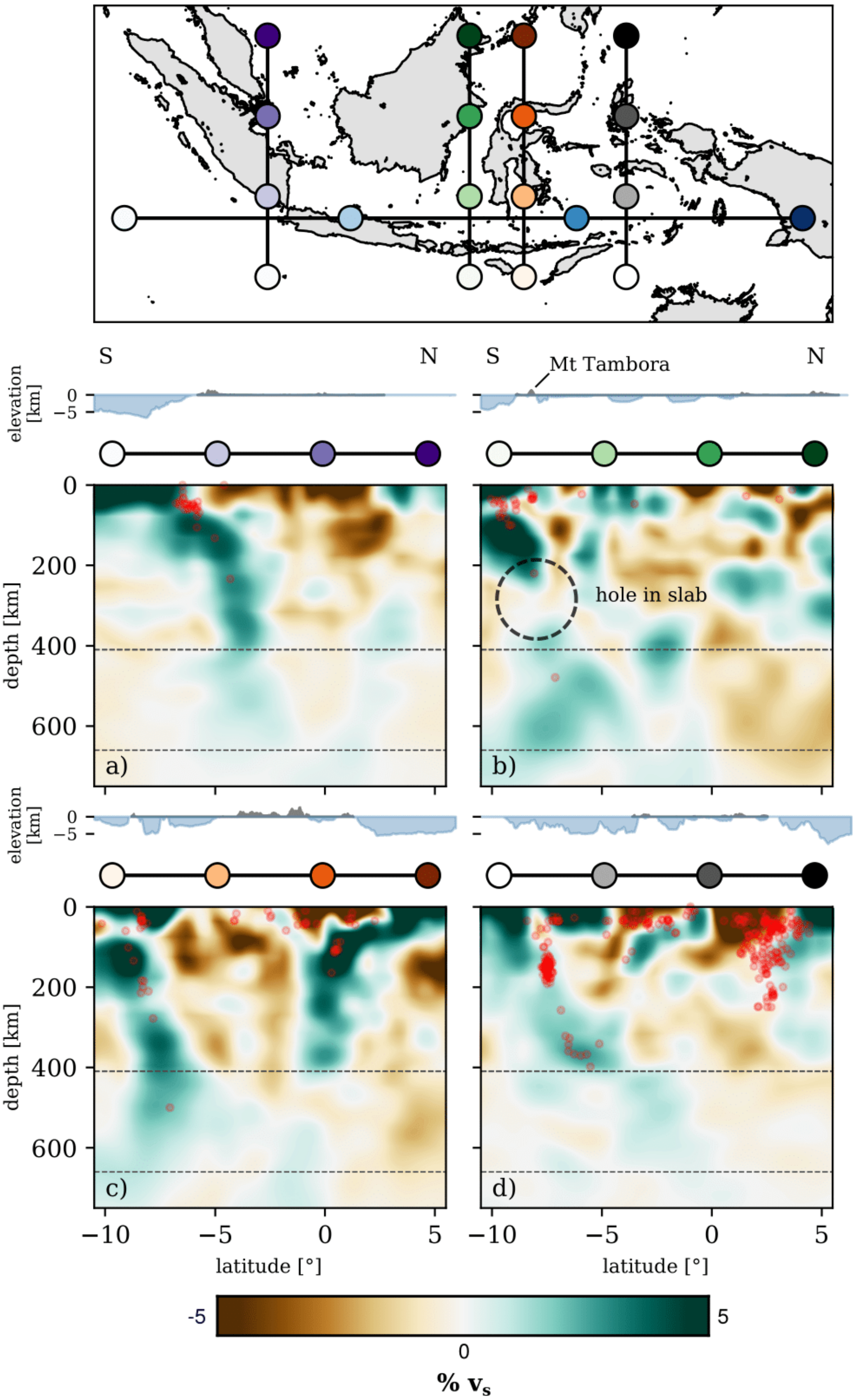

Figure 11. South-north v_S cross-sections and a map showing their locations. Perturbations are in % relative to the depth-average. Earthquake locations (red dots) are taken from the *ISC* catalog (International Seismological Centre, 2016) and are within 25 km from the cross-section slice. The vertical-horizontal ratio is 1:2.

6.5 Comparison to other models

Figure 12 shows a comparison between *SASSY21* and other S-wave tomographic models at 200 km depth. For all models, the high-velocity zone along the Indonesian arc is the most prominent feature, even though they differ in extent and anomaly amplitude. All models agree on a high-velocity zone in the southeast, which is associated with the Southeast Asia-Australia collision zone. Our model is able to resolve smaller-scale features, in particular around Borneo, Sulawesi and along the Banda Arc. One of the main factors contributing to this difference is the availability of regional earthquake data from the dense seismic networks used in our study.

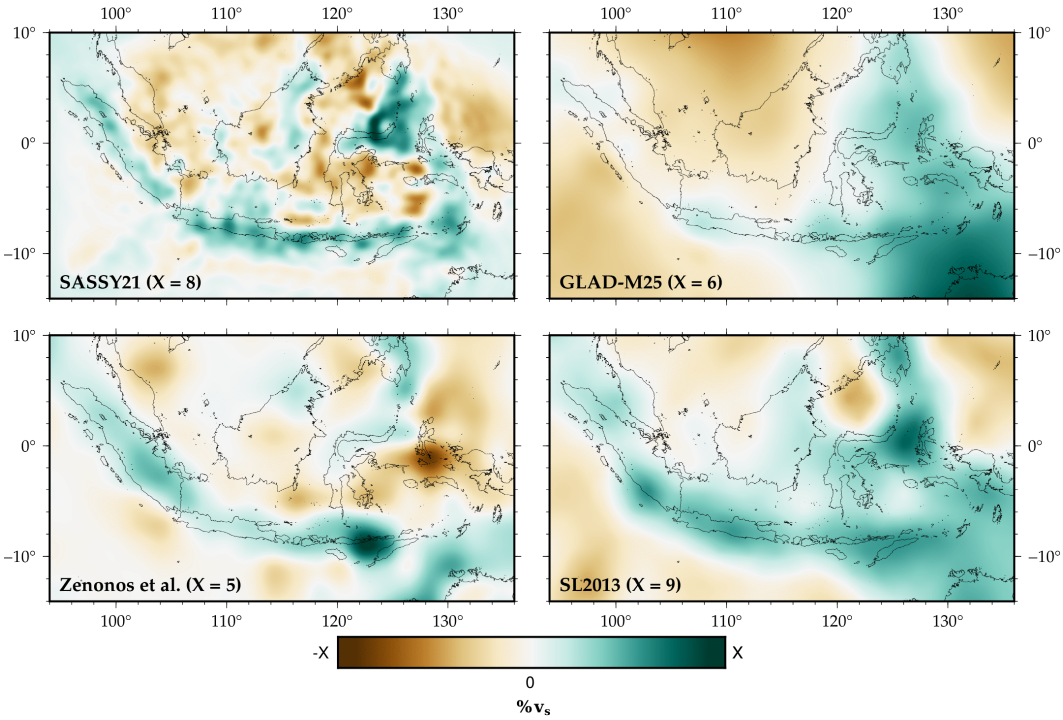

Figure 12. Shear-wave depth slices at 200 km depth for four different models. Perturbations are in % relative to the depth-average within the region. The limits of the colorscale X are shown in the lower left corner of each plot. *Top left:* This study. *Top right:* *GLAD-M25* – A global adjoint waveform tomography model by Lei et al. (2020). *Bottom left:* A continental-scale S-wave travel time tomography model by Zenonos et al. (2019). *Bottom right:* *SL2013* – A global shear-wave model of the upper mantle by Schaeffer and Lebedev (2013).

6.6 Limitations

An obvious limitation of FWI is the high computational cost of the forward problem, which translates to the use of a smaller dataset and relatively long periods compared to ray tomographic studies. Furthermore, we would ideally invert for other physical parameters such as attenuation and for more complex forms of anisotropy in order to mitigate parameter trade-off. However, almost all studies only attempt to constrain v_p and v_s and their anisotropic counterparts (e.g. Fichtner et al., 2010; Simutė et al., 2016). Few studies have investigated the benefits of reconstructing other properties (e.g. density, Blom et al., 2017), which is mainly a result of the difficulty to determine the optimal observables for constraining a specific parameter and the lack of data constraints. For the latter reason, we are also not inverting for source parameters, despite the potential for source errors to map as artifacts in the tomographic model (Blom et al., submitted). However, we believe we somewhat mitigate this by careful event selection and monitoring throughout the inversion (see Section 4.2) and the fact that the lowest period considered is 20 s, which corresponds to a wavelength much longer than anticipated rupture lengths and source durations of the events that are used.

We briefly mentioned strong updates in the upper ∼100 km (see Section 5.4), which may be somewhat mitigated by only allowing crustal updates at shorter periods (Morency et al., 2020). In this study, we are not implementing crustal structure explicitly (e.g. *CRUST2.0*) because 1) it increases the compute time as a result of small mesh elements along the surface, and 2) Southeast Asia is very complex and we are not confident that a global crustal model properly captures this. Thus, we decided not to add any prior information about the crustal structure. However, we believe we only start to separate out crust and mantle structure in the final period band at 20 s.

7 Conclusions

We have imaged the lithosphere and underlying mantle beneath Southeast Asia at periods between 20 − 150 s using multi-scale adjoint waveform tomography. The inversion parameters were restricted to isotropic P-wave velocity (v_P), vertically (v_{SV}) and horizontally (v_{SH}) polarized shear-wave velocity, and density (ρ). A sophisticated spectral-element solver was implemented to produce realistic synthetic seismograms by implementing topography, bathymetry, attenuation and approximating the fluid ocean by the weight of its water column. Furthermore, we enhanced depth sensitivity by separating small and large-amplitude arrivals. Our final model, *SASSY21*, was reached after 87 iterations and is most reliable for shear-wave velocity due to the natural dominance of surface wave signals in adjoint waveform tomography. We are able to resolve mantle structure, including multiple subduction zones, down to the transition zone, with v_{SV} exhibiting the strongest perturbations. The final model is able to explain true-amplitude data from events and receivers not included in the inversion. The trade-off between inversion parameters is estimated through an analysis of the Hessian-vector product. The most prominent feature is the (Indo-)Australian plate descending beneath Indonesia, with a steeply dipping Sunda slab in the west. Further east, we image the Southeast Asia-Australia collision zone, indicated by high velocities that reflect the presence of the northward moving North Australian continental lithosphere. The 180° curvature of the Banda Arc is imaged as one continuous slab. We observe overall positive radial anisotropy ($v_{SH} > v_{SV}$) for the upper 130 km, which transitions to negative radial anisotropy ($v_{SV} > v_{SH}$) at greater depths. Lateral variations in radial anisotropy reveal negative values along the slabs and beneath Sundaland, which we attribute to lattice-preferred orientation of mantle minerals and frozen-in anisotropy. *SASSY21* confirms the existence of a hole in the slab beneath Mount Tambora, which may be associated with the pertusion of continental lithosphere via entrainment of subducted plateau material. We further image a high-velocity zone around northern Borneo and reveal a previously undiscovered feature beneath the east coast of Borneo. While two subduction systems terminated in the Neogene around northern Borneo, which may have left upper mantle remnants, the origin of the high-velocity zone in eastern Borneo remains enigmatic, but may be associated with underthrusting from the formation of Sulawesi.

Model availability

The final model is available as *NetCDF* and *HDF5* files, with the former being readable by e.g. *xarray* (Hoyer & Hamman, 2017) and the latter suitable for viewing with *ParaView* (Ahrens et al., 2005) and interaction with *Salvus* (Afanasiev et al., 2019). We further provide *SASSY21* in *CSV* format. The final model and a 3-D model fly-through can be found on a *Zenodo* repository at `https://doi.org/10.5281/zenodo.5166488` (Wehner et al., 2021).

Competing interests

No competing interests are present.

Acknowledgments

We use the *Salvus* software package (release 0.11.23 – 0.11.33, `www.mondaic.com`) for the mesh generation, forward and adjoint simulations and non-linear optimization, within its integrated workflow. Simulations were run using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (`www.csd3.cam.ac.uk`), and facilitated by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council (`www.dirac.ac.uk`).

Data processing was done using *NumPy* (C. R. Harris et al., 2020) and *ObsPy* (Beyreuther et al., 2010). Waveform data are handled in the *Adaptable Seismic Data Format (ASDF)* (Krischer et al., 2016). Visualizations were created using *PyGMT* (Uieda et al., 2021) and *Matplotlib* (Hunter, 2007).

This research is funded by the Engineering and Physical Sciences Research Council (EPSRC) project reference 2073302, BP, BPI and Schlumberger, Global Challenges Research Fund (GCRF) G102642 and National Science Foundation (NSF) grant EAR-1250214.

We would like to thank Lion Krischer, Tim Greenfield, Michael Afanasiev, Ya-Jian Gao, Keith Priestley, Nepomuk Boitz, Sölvi Thrastarson, Dirk-Philip van Herwaarden and Minghao Zhang for fruitful discussions.

References

Afanasiev, M., Boehm, C., van Driel, M., Krischer, L., Rietmann, M., May, D. A., ... Fichtner, A. (2019). Modular and flexible spectral-element waveform modelling in two and three dimensions. *Geophysical Journal International*, *216*(3), 1675–1692.

Ahrens, J., Geveci, B., & Law, C. (2005). Paraview: An end-user tool for large data visualization. *The visualization handbook*, *717*(8).

Aki, K., Christofferson, X., & Husebye, Y. (1977). Three-dimensional seismic structure of the lithosphere. *J. geophys. Res.*, *82*, 277–296.

Aki, K., & Richards, P. G. (2002). *Quantitative seismology.*

Altamimi, Z., Rebischung, P., Métivier, L., & Collilieux, X. (2016). Itrf2014: A new release of the international terrestrial reference frame modeling nonlinear station motions. *Journal of Geophysical Research: Solid Earth*, *121*(8), 6109–6131.

Amante, C., & Eakins, B. (2008). Etopo1 1 arc-minute global relief model: Procedures, data sources and analysis, national geophysical data center, nesdis, noaa, us dept. *Commerce, Boulder, CO, USA*.

Amaru, M. (2007). *Global travel time tomography with 3-d reference models* (Vol. 274). Utrecht University.

Artemieva, I. M., & Mooney, W. D. (2001). Thermal thickness and evolution of precambrian lithosphere: A global study. *Journal of Geophysical Research: Solid Earth*, *106*(B8), 16387–16414.

Audley-Charles, M. G. (1968). The geology of the portuguese timor.

Babuska, V., & Cara, M. (1991). *Seismic anisotropy in the earth* (Vol. 10). Springer Science & Business Media.

Bamberger, A., Chavent, G., Hemon, C., & Lailly, P. (1982). Inversion of normal incidence seismograms. *Geophysics*, *47*(5), 757–770.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). Obspy: A python toolbox for seismology. *Seismological Research Letters*, *81*(3), 530–533.

Bianchi, M., Evans, P., Heinloo, A., & Quinteros, J. (2015). Webdc3 web interface.

Bijwaard, H., Spakman, W., & Engdahl, E. R. (1998). Closing the gap between regional and global travel time tomography. *Journal of Geophysical Research: Solid Earth*, *103*(B12), 30055–30078.

Bird, P. (2003). An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems*, *4*(3).

Blom, N., Boehm, C., & Fichtner, A. (2017). Synthetic inversions for density using seismic and gravity data. *Geophysical Journal International*, *209*(2), 1204–1220.

Blom, N., Gokhberg, A., & Fichtner, A. (2020). Seismic waveform tomography of the central and eastern mediterranean upper mantle. *Solid Earth*, *11*(2), 669–690.

Blom, N., Hardalupas, P.-S., & Rawlinson, N. (submitted). Mitigating the effect of errors in earthquake parameters on seismic (waveform) tomography. *Geophysical Journal International*.

Boehm, C., Afanasiev, M., Krischer, L., van Driel, M., & Fichtner, A. (2019). Anisotropic diffusion-based smoothing filters for full-waveform inversion. In *Geophysical research abstracts* (Vol. 21).

Bunks, C., Saleck, F. M., Zaleski, S., & Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, *60*(5), 1457–1473.

Capdeville, Y., Guillot, L., & Marigo, J.-J. (2010). 1-d non-periodic homogenization for the seismic wave equation. *Geophysical Journal International*, *181*(2), 897–910.

Carter, D. J., Audley-Charles, M. G., & Barber, A. (1976). Stratigraphical analysis of island arc—continental margin collision in eastern indonesia. *Journal of the Geological Society*, *132*(2), 179–198.

Červený, V. (2001). *Seismic ray theory, cambridge univ.* Press Cambridge).

Chavent, G. (1974). Identification of functional parameters in partial differential equations: Identification of parameter distributed systems: Re goodson, and polis. *New York, ASME*.

Chen, P., Zhao, L., & Jordan, T. H. (2007). Full 3d tomography for the crustal structure of the los angeles region. *Bulletin of the Seismological Society of America*, *97*(4), 1094–1120.

Clayton, R., & Engquist, B. (1977). Absorbing boundary conditions for acoustic and elastic wave equations. *Bulletin of the seismological society of America*, *67*(6), 1529–1540.

Conn, A. R., Gould, N. I., & Toint, P. L. (2000). *Trust region methods.* SIAM.

Cottam, M. A., Hall, R., Sperber, C., Kohn, B. P., Forster, M. A., & Batt, G. E. (2013). Neogene rock uplift and erosion in northern borneo: evidence from the kinabalu granite, mount kinabalu. *Journal of the Geological Society*, *170*(5), 805–816.

Crotwell, H. P., Owens, T. J., & Ritsema, J. (1999). The taup toolkit: Flexible seismic travel-time and ray-path utilities. *Seismological Research Letters*, *70*(2), 154–160.

Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference earth model. *Physics of the earth and planetary interiors*, *25*(4), 297–356.

Dziewonski, A. M., Hager, B. H., & O'Connell, R. J. (1977). Large-scale heterogeneities in the lower mantle. *Journal of Geophysical Research*, *82*(2), 239–255.

Ekström, G., Nettles, M., & Dziewoński, A. (2012). The global cmt project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, *200*, 1–9.

Elburg, M., Van Bergen, M., & Foden, J. (2004). Subducted upper and lower continental crust contributes to magmatism in the collision sector of the sunda-banda arc, indonesia. *Geology*, *32*(1), 41–44.

Fichtner, A. (2010). *Full seismic waveform modelling and inversion.* Springer Science & Business Media.

Fichtner, A., De Wit, M., & van Bergen, M. (2010). Subduction of continental lithosphere in the banda sea region: Combining evidence from full waveform tomography and isotope ratios. *Earth and Planetary Science Letters*, *297*(3-4), 405–412.

Fichtner, A., Kennett, B. L., Igel, H., & Bunge, H.-P. (2008). Theoretical background for continental-and global-scale full-waveform inversion in the time–frequency domain.

Geophysical Journal International, *175*(2), 665–685.

Fichtner, A., Kennett, B. L., Igel, H., & Bunge, H.-P. (2009). Full seismic waveform tomography for upper-mantle structure in the australasian region using adjoint methods. *Geophysical Journal International*, *179*(3), 1703–1725.

Fichtner, A., Kennett, B. L., & Trampert, J. (2013). Separating intrinsic and apparent anisotropy. *Physics of the Earth and Planetary Interiors*, *219*, 11–20.

Fichtner, A., & Leeuwen, T. v. (2015). Resolution analysis by random probing. *Journal of Geophysical Research: Solid Earth*, *120*(8), 5549–5573.

Fichtner, A., van Herwaarden, D.-P., Afanasiev, M., Simutė, S., Krischer, L., Çubuk-Sabuncu, Y., ... others (2018). The collaborative seismic earth model: generation 1. *Geophysical research letters*, *45*(9), 4007–4016.

Fukao, Y., & Obayashi, M. (2013). Subducted slabs stagnant above, penetrating through, and trapped below the 660 km discontinuity. *Journal of Geophysical Research: Solid Earth*, *118*(11), 5920–5938.

Gao, Y., Tilmann, F., van Herwaarden, D.-P., Thrastarson, S., Fichtner, A., Heit, B., ... Schurr, B. (2021). Full waveform inversion beneath the central andes: Insight into the dehydration of the nazca slab and delamination of the back-arc lithosphere. *Journal of Geophysical Research: Solid Earth*, *126*(7), e2021JB021984. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JB021984` (e2021JB021984 2021JB021984) doi: https://doi.org/10.1029/2021JB021984

Gauthier, O., Virieux, J., & Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms: Numerical results. *Geophysics*, *51*(7), 1387–1403.

Greenfield, T., Widiyantoro, S., & Rawlinson, N. (2018). *Kalimantan temporary network.* International Federation of Digital Seismograph Networks. Retrieved from `http://www.fdsn.org/networks/detail/9G_2018/` doi: 10.7914/SN/9G_2018

Guasch, L., Agudo, O. C., Tang, M.-X., Nachev, P., & Warner, M. (2020). Full-waveform inversion imaging of the human brain. *NPJ digital medicine*, *3*(1), 1–12.

Gudmundsson, Ó., & Sambridge, M. (1998). A regionalized upper mantle (rum) seismic model. *Journal of Geophysical Research: Solid Earth*, *103*(B4), 7121–7136.

Hall, R. (2002). Cenozoic geological and plate tectonic evolution of se asia and the sw pacific: computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences*, *20*(4), 353–431.

Hall, R. (2011). Australia–se asia collision: plate tectonics and crustal flow. *Geological Society, London, Special Publications*, *355*(1), 75–109.

Hall, R. (2013). Contraction and extension in northern borneo driven by subduction rollback. *Journal of Asian Earth Sciences*, *76*, 399–411.

Hall, R., & Morley, C. K. (2004). Sundaland basins. *Continent–Ocean Interactions within East Asian Marginal Seas. AGU Geophysical Monograph Series*, *149*, 55–86.

Hall, R., & Spakman, W. (2015). Mantle structure and tectonic history of se asia. *Tectonophysics*, *658*, 14–45.

Hall, R., & Wilson, M. (2000). Neogene sutures in eastern indonesia. *Journal of Asian Earth Sciences*, *18*(6), 781–808.

Hamilton, W. B. (1979). *Tectonics of the indonesian region* (Vol. 1078). US Government Printing Office.

Harris. (2011). The nature of the banda arc–continent collision in the timor region. In *Arc-continent collision* (pp. 163–211). Springer.

Harris, Miller, M. S., Supendi, P., & Widiyantoro, S. (2020). Subducted lithospheric boundary tomographically imaged beneath arc-continent collision in eastern indonesia. *Journal of Geophysical Research: Solid Earth*, *125*(8), e2019JB018854.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... others (2020). Array programming with numpy. *Nature*, *585*(7825), 357–362.

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.

Hirt, C., & Rexer, M. (2015). Earth2014: 1 arc-min shape, topography, bedrock and ice-sheet models–available as gridded data and degree-10,800 spherical harmonics. *International Journal of Applied Earth Observation and Geoinformation*, *39*, 103–112.

Hosseini, K., & Sigloch, K. (2017). obspydmt: a python toolbox for retrieving and processing of large seismological datasets. *Solid Earth*, *8*.

Hoyer, S., & Hamman, J. (2017). xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, *5*(1). Retrieved from `http://doi.org/10.5334/jors.148` doi: 10.5334/jors.148

Huang, Z., Zhao, D., & Wang. (2015). P wave tomography and anisotropy beneath southeast asia: Insight into mantle dynamics. *Journal of Geophysical Research: Solid Earth*, *120*(7), 5154–5174.

Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *IEEE Annals of the History of Computing*, *9*(03), 90–95.

Institute Of Earth Sciences, A. S. (1996). *Broadband array in taiwan for seismology.* Institute of Earth Sciences, Academia Sinica, Taiwan. Retrieved from `http://www.fdsn.org/doi/10.7914/SN/TW` doi: 10.7914/SN/TW

International Seismological Centre. (2016). On-line bulletin. *Internatl. Seis. Cent. Thatcham.*

Jalinoos, F., Tran, K. T., Nguyen, T. D., & Agrawal, A. K. (2017). Evaluation of bridge abutments and bounded wall type structures with ultraseismic waveform tomography. *Journal of Bridge Engineering*, *22*(12), 04017104.

Katili, J. A. (1978). Past and present geotectonic position of sulawesi, indonesia. *Tectonophysics*, *45*(4), 289–322.

Keep, M., & Haig, D. W. (2010). Deformation and exhumation in timor: Distinct stages of a young orogeny. *Tectonophysics*, *483*(1-2), 93–111.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., ... Willing, C. (2016). Jupyter notebooks – a publishing format for reproducible computational workflows. In F. Loizides & B. Schmidt (Eds.), *Positioning and power in academic publishing: Players, agents and agendas* (p. 87 - 90).

Komatitsch, D., & Tromp, J. (2002). Spectral-element simulations of global seismic wave propagation—ii. three-dimensional models, oceans, rotation and self-gravitation. *Geophysical Journal International*, *150*(1), 303–318.

Komatitsch, D., Tsuboi, S., Ji, C., & Tromp, J. (2003). A 14.6 billion degrees of freedom, 5 teraflops, 2.5 terabyte earthquake simulation on the earth simulator. In *Sc'03: Proceedings of the 2003 acm/ieee conference on supercomputing* (pp. 4–4).

Komatitsch, D., & Vilotte, J.-P. (1998). The spectral element method: an efficient tool to simulate the seismic response of 2d and 3d geological structures. *Bulletin of the seismological society of America*, *88*(2), 368–392.

Kosloff, R., & Kosloff, D. (1986). Absorbing boundaries for wave propagation problems. *Journal of Computational Physics*, *63*(2), 363–376.

Krischer, L., & Casarotti, E. (2015, September). *pyflex: 0.1.4.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.31607` doi: 10.5281/zenodo.31607

Krischer, L., Fichtner, A., Boehm, C., & Igel, H. (2018). Automated large-scale full seismic waveform inversion for north america and the north atlantic. *Journal of Geophysical Research: Solid Earth*, *123*(7), 5902–5928.

Krischer, L., Smith, J., Lei, W., Lefebvre, M., Ruan, Y., de Andrade, E. S., ... Tromp, J. (2016). An adaptable seismic data format. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, *207*(2), 1003–1011.

Kristeková, M., Kristek, J., Moczo, P., & Day, S. M. (2006). Misfit criteria for quantitative comparison of seismograms. *Bulletin of the seismological Society of America*, *96*(5), 1836–1850.

Lailly, P., & Bednar, J. (1983). The seismic inverse problem as a sequence of before stack migrations. In *Conference on inverse scattering: theory and application* (pp. 206–220).

Lebedev, S., & Nolet, G. (2003). Upper mantle beneath southeast asia from s velocity tomography. *Journal of Geophysical Research: Solid Earth*, *108*(B1).

Lei, W., Ruan, Y., Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., … Pugmire, D. (2020). Global adjoint tomography—model glad-m25. *Geophysical Journal International*, *223*(1), 1–21.

Li, J., Ding, W., Lin, J., Xu, Y., Kong, F., Li, S., … Zhou, Z. (2021). Dynamic processes of the curved subduction system in southeast asia: A review and future perspective. *Earth-Science Reviews*, 103647.

Liu, Q., Beller, S., Lei, W., Peter, D., & Tromp, J. (2020). Preconditioned bfgs-based uncertainty quantification in elastic full waveform inversion. *arXiv preprint arXiv:2009.12663*.

Maggi, A., Tape, C., Chen, M., Chao, D., & Tromp, J. (2009). An automated time-window selection algorithm for seismic tomography. *Geophysical Journal International*, *178*(1), 257–281.

McCaffrey, R. (2009, 05). The tectonic framework of the sumatran subduction zone. *Earth Planet. Sci. Annu. Rev. Earth Planet. Sci*, *3737*, 345-66. doi: 10.1146/annurev.earth.031208.100212

Métrich, N., Vidal, C. M., Komorowski, J.-C., Pratomo, I., Michel, A., Kartadinata, N., … Surono (2017). New insights into magma differentiation and storage in holocene crustal reservoirs of the lesser sunda arc: The rinjani–samalas volcanic complex (lombok, indonesia). *Journal of Petrology*, *58*(11), 2257–2284.

Miller, M. S. (2014). *Transitions in the banda arc-australia continental collision.* International Federation of Digital Seismograph Networks. Retrieved from `http://www.fdsn.org/doi/10.7914/SN/YS_2014` doi: 10.7914/SN/YS_2014

Miller, M. S., O'Driscoll, L. J., Roosmawati, N., Harris, C. W., Porritt, R. W., Widiyantoro, S., … Joshua West, A. (2016). Banda arc experiment—transitions in the banda arc-australian continental collision. *Seismological Research Letters*, *87*(6), 1417–1423.

Modrak, R., & Tromp, J. (2016). Seismic waveform inversion best practices: regional, global and exploration test cases. *Geophysical Journal International*, *206*(3), 1864–1889.

Morency, C., Matzel, E., Afanasiev, M., Krischer, L., Boehm, C., & Rodgers, A. J. (2020). Improved seismic tomography of the brady geothermal field, nevada, based upon full waveform inversion using salvus. In *Agu fall meeting abstracts* (Vol. 2020, pp. S063–0013).

Moresi, L., Betts, P. G., Miller, M. S., & Cayley, R. A. (2014). Dynamics of continental accretion. *Nature*, *508*(7495), 245–248.

Nocedal, J., & Wright, S. (2006). *Numerical optimization.* Springer Science & Business Media.

Nolet, G. (2008). A breviary of seismic tomography. *bst*.

Operto, S., Miniussi, A., Brossier, R., Combe, L., Métivier, L., Monteiller, V., … Virieux, J. (2015). Efficient 3-d frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to valhall in the visco-acoustic vertical transverse isotropic approximation. *Geophysical Journal International*, *202*(2), 1362–1391.

Panning, M., & Romanowicz, B. (2006). A three-dimensional radially anisotropic model of shear velocity in the whole mantle. *Geophysical Journal International*, *167*(1), 361–379.

Petley, D. (2019). *The anak krakatau landslide and tsunami.* Retrieved 2019-03-12, from `https://blogs.agu.org/landslideblog/2018/12/26/anak-krakatau-1/`

Pratt, R. G., & Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. part 1: Acoustic wave-equation method 1. *Geophysical prospecting*, *38*(3), 287–310.

Priestley, K., Ho, T., & McKenzie, D. (2021). The formation of continental roots. *Geology*, *49*(2), 190–194.

Rawlinson, N. (2018). *North borneo orogeny seismic survey.* International Federation of Digital Seismograph Networks. Retrieved from `http://www.fdsn.org/doi/10.7914/SN/YC_2018` doi: 10.7914/SN/YC_2018

Rawlinson, N., Hauser, J., & Sambridge, M. (2008). Seismic ray tracing and wavefront tracking in laterally heterogeneous media. *Advances in geophysics*, *49*, 203–273.

Rawlinson, N., & Spakman, W. (2016). On the use of sensitivity tests in seismic tomography. *Geophysical Journal International*, *205*(2), 1221–1243.

Rickers, F., Fichtner, A., & Trampert, J. (2012). Imaging mantle plumes with instantaneous phase measurements of diffracted waves. *Geophysical Journal International*, *190*(1), 650–664.

Ruan, Y., Lei, W., Modrak, R., Örsvuran, R., Bozdağ, E., & Tromp, J. (2019). Balancing unevenly distributed data in seismic tomography: a global adjoint tomography example. *Geophysical Journal International*, *219*(2), 1225–1236.

Schaeffer, A., & Lebedev, S. (2013). Global shear speed structure of the upper mantle and transition zone. *Geophysical Journal International*, *194*(1), 417–449.

Schreiman, J., Gisvold, J., Greenleaf, J. F., & Bahn, R. (1984). Ultrasound transmission computed tomography of the breast. *Radiology*, *150*(2), 523–530.

Simons, W., Socquet, A., Vigny, C., Ambrosius, B., Haji Abu, S., Promthong, C., ... others (2007). A decade of gps in southeast asia: Resolving sundaland motion and boundaries. *Journal of Geophysical Research: Solid Earth*, *112*(B6).

Simutė, S., Steptoe, H., Cobden, L., Gokhberg, A., & Fichtner, A. (2016). Full-waveform inversion of the japanese islands region. *Journal of Geophysical Research: Solid Earth*, *121*(5), 3722–3741.

Sirgue, L., Barkved, O., Dellinger, J., Etgen, J., Albertin, U., & Kommedal, J. (2010). Thematic set: Full waveform inversion: The next leap forward in imaging at valhall. *First Break*, *28*(4).

Song, T.-R. A., & Kawakatsu, H. (2012). Subduction of oceanic asthenosphere: Evidence from sub-slab seismic anisotropy. *Geophysical Research Letters*, *39*(17).

Spakman, W., & Hall, R. (2010). Surface deformation and slab–mantle interaction during banda arc subduction rollback. *Nature Geoscience*, *3*(8), 562–566.

Sturgeon, W., Ferreira, A. M., Faccenda, M., Chang, S.-J., & Schardong, L. (2019). On the origin of radial anisotropy near subducted slabs in the midmantle. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5105–5125.

Tape, C., Liu, Q., Maggi, A., & Tromp, J. (2010). Seismic tomography of the southern california crust based on spectral-element and adjoint methods. *Geophysical Journal International*, *180*(1), 433–462.

Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, *49*(8), 1259–1266.

Thrastarson, S., van Herwaarden, D.-P., & Fichtner, A. (2021, feb). *solvithrastar/MultiMesh: MultiMesh - Python-based interpolations between discretizations.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.4564523` doi: 10.5281/zenodo.4564523

Tromp, J. (2020). Seismic wavefield imaging of earth's interior across scales. *Nature Reviews Earth & Environment*, *1*(1), 40–53.

Tromp, J., Tape, C., & Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, *160*(1), 195–216.

Turner, S., Foden, J., George, R., Evans, P., Varne, R., Elburg, M., & Jenner, G. (2003). Rates and processes of potassic magma evolution beneath sangeang api volcano, east sunda arc, indonesia. *Journal of Petrology*, *44*(3), 491–515.

Uieda, L., Tian, D., Leong, W. J., Toney, L., Schlitzer, W., Yao, J., ... Wessel, P. (2021, March). *PyGMT: A Python interface for the Generic Mapping Tools.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.4592991` doi: 10.5281/zenodo.4592991

Vallée, M. (2013). Source time function properties indicate a strain drop independent of earthquake depth and magnitude. *Nature communications*, *4*(1), 1–6.

Van der Hilst, R. D., Widiyantoro, S., & Engdahl, E. (1997). Evidence for deep mantle circulation from global tomography. *Nature*, *386*(6625), 578–584.

van Driel, M., & Nissen-Meyer, T. (2014). Optimized viscoelastic wave propagation for weakly dissipative media. *Geophysical Journal International*, *199*(2), 1078–1093.

van Herwaarden, D. P., Boehm, C., Afanasiev, M., Thrastarson, S., Krischer, L., Trampert, J., & Fichtner, A. (2020). Accelerated full-waveform inversion using dynamic mini-batches. *Geophysical Journal International*, *221*(2), 1427–1438.

Virieux, J. (1984). Sh-wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, *49*(11), 1933–1942.

Virieux, J., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, *74*(6), WCC1–WCC26.

Wang, & He, X. (2020). Seismic anisotropy in the java-banda and philippine subduction zones and its implications for the mantle flow system beneath the sunda plate. *Geochemistry, Geophysics, Geosystems*, *21*(4), e2019GC008658.

Wang, Singh, S., & Noble, M. (2020). True-amplitude versus trace-normalized full waveform inversion. *Geophysical Journal International*, *220*(2), 1421–1435.

Wehner, D., Blom, N., Rawlinson, N., Daryono, Böhm, C., Miller, M. S., ... Widiyantoro, S. (2021, August). SASSY21: A 3-D seismic structural model of the lithosphere and underlying mantle beneath Southeast Asia from multi-scale adjoint waveform tomography. *Journal of Geophysical Research: Solid Earth*. Retrieved from `https://doi.org/10.5281/zenodo.5166488` doi: 10.5281/zenodo.5166488

Wells, D. L., & Coppersmith, K. J. (1994, 08). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America*, *84*(4), 974-1002.

Widiyantoro, S., Pesicek, J., & Thurber, C. (2011). Subducting slab structure below the eastern sunda arc inferred from non-linear seismic tomographic imaging. *Geological Society, London, Special Publications*, *355*(1), 139–155.

Widiyantoro, S., & van der Hilst, R. (1996). Structure and evolution of lithospheric slab beneath the sunda arc, indonesia. *Science*, *271*(5255), 1566–1570.

Yang, T., Gurnis, M., & Zahirovic, S. (2016). Mantle-induced subsidence and compression in se asia since the early miocene. *Geophysical Research Letters*, *43*(5), 1901–1909.

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. *Geophysical Journal International*, *221*(2), 1145–1164.

Zenonos, A., De Siena, L., Widiyantoro, S., & Rawlinson, N. (2019). P and s wave travel time tomography of the se asia-australia collision zone. *Physics of the Earth and Planetary Interiors*, *293*, 106267.

Zhou, Y., Dahlen, F., & Nolet, G. (2004). Three-dimensional sensitivity kernels for surface wave observables. *Geophysical Journal International*, *158*(1), 142–168.

Zhu, H., Bozdağ, E., Peter, D., & Tromp, J. (2012). Structure of the european upper mantle revealed by adjoint tomography. *Nature Geoscience*, *5*(7), 493–498.

Supporting Information for "*SASSY21*: A 3-D seismic structural model of the lithosphere and underlying mantle beneath Southeast Asia from multi-scale adjoint waveform tomography"

Deborah Wehner[1], Nienke Blom[1], Nicholas Rawlinson [1], Daryono[2], Christian Böhm[3,4], Meghan S. Miller[5], Pepen Supendi[2], and Sri Widiyantoro[6,7]

[1]Bullard Laboratories, Department of Earth Sciences, University of Cambridge, Cambridge, UK

[2]Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG)

[3]ETH Zürich, Zurich, Switzerland

[4]Mondaic AG, Zurich, Switzerland

[5]Australian National University, Canberra, Australia

[6]Global Geophysics Research Group, Faculty of Mining and Petroleum Engineering, Institut Teknologi Bandung, Bandung, Indonesia

[7]Faculty of Engineering, Maranatha Christian University, Bandung, Indonesia

Contents of this file

1. S1 presents the starting model.

2. Table S2.1 contains an overview of the events used in this study. S2.2 shows the event misfit reduction for the final model relative to the starting model.

3. S3 provides a description and map of the station availability.

4. S4 gives more detail on the preconditioning steps applied to the gradients.

---

5. S5 elaborates further on technical details not described in the main text.

6. S6 presents further waveform fits.

7. S7 presents depth slices from 50 to 700 km for all inversion parameters (v_{SH}, v_{SV}, v_P and density).

Additional supporting information (files uploaded separately)

The final model is provided as *NetCDF* and *HDF5* files, with the former being readable by e.g. *xarray* (Hoyer & Hamman, 2017) and the latter suitable for viewing with *ParaView* (Ahrens et al., 2005) and interaction with *Salvus*. We further provide *SASSY21* in *CSV* format, and a Jupyter Notebook (Kluyver et al., 2016) demonstrating how to interact with the different file formats.

Introduction

In this Supplementary Material, we present the starting model (Section S1), provide additional detail about the earthquake data used throughout the inversion (Section S2.1), event misfits compared to hypocentral depth, magnitude and focal mechanism (Section S2.2), the stations used throughout this study (Section 3), the processing steps applied to the raw gradients (Section S4) and an overview of further technical parameters (Section 5). We also show further waveform fits in Section 6 and extra depth slices for v_{SV}, v_{SH}, v_P and density ρ (Section 7). The final model is provided separate to this document.

S1 Starting model

Figure S1 presents the absolute values of the 1-D starting model, which was taken from *CSEM* (Fichtner et al., 2018).

S2.1 Event overview

Table S2 contains an overview of the events used in this study. Event locations and moment tensors are retrieved from the *GCMT* catalog and remain constant throughout the inversion. The source time function is approximated by a Butterworth bandpass filtered Heaviside step function, representing an instantaneous rupture process. 15 events with depths > 300 km were selected to ensure a diversity of data coverage.

Table S2: List of events used throughout this study. The period bands used for each event are indicated by roman numerals in the last column, following the notation from Table S5. The first 50 events are used across all period bands.

#	Focaltime	M_w	Longitude	Latitude	Depth [km]	Period bands
1	2014-09-10T02:46:11.7	6.27	125.06	-0.36	29.2	I - VII
2	2014-12-02T05:11:37.2	6.58	123.17	6.31	631.7	I - VII
3	2014-12-06T22:05:14.8	6.04	130.57	-6.12	137.8	I - VII
4	2015-02-27T13:45:08.9	6.97	122.50	-7.35	551.5	I - VII
5	2015-03-03T10:37:35.7	6.18	98.58	-0.72	23.6	I - VII
6	2015-03-17T22:12:32.1	6.28	126.48	1.78	41.9	I - VII
7	2015-03-28T22:28:52.4	5.92	122.00	0.43	130.6	I - VII
8	2015-05-15T20:26:58.3	6.04	102.14	-2.61	158.4	I - VII
9	2015-07-03T06:43:24.4	6.11	126.25	10.08	43.8	I - VII
10	2015-07-26T07:05:09.9	5.90	112.82	-9.45	43.7	I - VII
11	2015-08-20T11:00:11.3	5.81	126.50	0.63	71.7	I - VII
12	2015-09-16T07:41:02.6	6.32	126.47	2.01	33.0	I - VII
13	2015-11-11T23:36:22.0	5.84	128.93	-7.41	137.0	I - VII
14	2015-11-21T09:06:16.2	6.04	130.11	-7.22	100.4	I - VII
15	2015-12-24T23:10:59.7	5.81	129.11	-7.34	132.1	I - VII
16	2016-01-11T16:38:11.6	6.49	127.05	3.84	12.0	I - VII
17	2016-02-12T10:02:29.4	6.24	119.35	-9.87	38.0	I - VII
18	2016-04-05T08:29:39.2	5.92	126.63	4.21	29.7	I - VII
19	2016-04-06T14:45:35.3	6.05	107.42	-8.41	41.9	I - VII
20	2016-06-05T16:25:36.5	6.30	125.56	-4.51	449.0	I - VII
21	2016-09-04T02:38:13.9	5.77	125.85	8.38	19.0	I - VII
22	2016-09-23T22:53:11.3	6.30	126.49	6.55	63.2	I - VII

Continued on next page

Table S2 – *Continued from previous page*

#	Focaltime	M_w	Longitude	Latitude	Depth [km]	Period bands
23	2016-10-19T00:26:04.8	6.61	108.07	-4.95	622.8	I - VII
24	2016-10-27T08:17:52.2	5.78	125.88	1.40	67.8	I - VII
25	2016-12-05T01:13:07.2	6.28	123.46	-7.36	531.9	I - VII
26	2016-12-29T22:30:21.9	6.26	118.74	-9.16	98.4	I - VII
27	2017-04-28T20:23:23.6	6.85	124.89	5.49	31.4	I - VII
28	2017-07-15T12:12:22.5	5.94	121.95	0.44	125.8	I - VII
29	2017-08-13T03:08:17.8	6.47	101.43	-3.81	43.3	I - VII
30	2017-12-15T16:48:00.7	6.55	108.11	-7.91	109.4	I - VII
31	2018-01-23T06:34:57.0	6.02	106.16	-7.18	53.1	I - VII
32	2018-02-26T13:34:58.8	6.00	126.82	-2.65	12.9	I - VII
33	2018-03-02T02:20:14.5	5.93	130.35	-6.17	151.9	I - VII
34	2018-03-25T20:14:50.2	6.43	129.84	-6.72	181.7	I - VII
35	2018-04-05T03:53:42.0	6.06	126.88	6.69	45.5	I - VII
36	2018-05-10T18:02:29.8	5.88	123.70	6.95	543.1	I - VII
37	2018-08-17T15:35:04.1	6.51	119.75	-7.31	538.9	I - VII
38	2018-11-04T07:55:29.9	5.99	123.75	7.82	599.3	I - VII
39	2018-12-01T13:27:25.2	6.47	128.67	-7.47	146.0	I - VII
40	2018-12-29T03:39:14.8	6.98	126.91	5.87	54.4	I - VII
41	2019-02-08T11:55:12.6	5.90	126.41	9.85	20.8	I - VII
42	2019-03-06T00:13:04.8	5.86	127.05	8.49	18.1	I - VII
43	2019-03-08T15:06:16.4	6.06	126.20	10.35	43.3	I - VII
44	2019-04-06T21:55:04.1	6.28	124.86	-6.92	546.5	I - VII
45	2019-05-31T10:12:33.1	6.15	126.54	6.22	87.9	I - VII
46	2019-08-02T12:03:34.8	6.89	104.85	-7.40	51.9	I - VII
47	2019-09-21T19:53:15.4	5.88	130.50	-6.46	87.8	I - VII
48	2019-09-29T02:02:53.4	6.25	126.58	5.65	77.3	I - VII
49	2019-10-29T01:04:49.0	6.61	125.05	6.87	18.0	I - VII
50	2020-04-05T18:37:14.9	6.02	126.33	1.53	41.3	I - VII
51	2014-09-10T05:16:56.8	5.89	125.12	-0.33	26.5	I - III
52	2014-11-26T14:33:50.0	6.77	126.44	2.11	35.2	I - III
53	2014-11-29T19:40:15.3	5.77	126.99	2.51	27.4	I - III
54	2014-12-29T09:29:40.9	6.14	121.45	8.68	15.0	I - III
55	2015-03-15T23:17:28.2	6.06	122.35	-0.53	25.1	I - III
56	2015-11-04T03:44:21.2	6.54	124.95	-8.20	12.0	I - III
57	2015-12-09T10:21:54.6	6.79	129.51	-4.16	12.2	I - III
58	2016-06-07T19:15:19.5	6.36	126.35	1.41	31.4	I - III
59	2017-02-10T14:03:47.5	6.47	125.49	9.85	12.0	I - III
60	2018-08-05T11:46:44.7	6.94	116.24	-8.33	17.8	I - III
61	2018-09-08T07:16:52.7	6.13	126.43	7.14	15.1	I - III
62	2018-09-28T10:02:59.4	7.57	119.86	-0.72	12.0	I - III
63	2018-10-10T18:44:59.0	5.97	114.48	-7.45	13.5	I - III
64	2018-12-28T03:03:35.5	5.81	134.01	-1.41	48.8	I - III
65	2019-07-14T09:11:04.6	7.19	128.13	-0.72	12.0	I - III

Continued on next page

Table S2 – *Continued from previous page*

#	Focaltime	M_w	Longitude	Latitude	Depth [km]	Period bands
66	2020-01-19T16:58:22.9	6.19	123.87	-0.15	129.4	I - III
67	2014-01-25T05:14:22.8	6.15	109.27	-8.36	76.1	I - IV
68	2014-11-21T10:10:25.4	6.54	127.08	2.60	30.1	I - IV
69	2014-12-17T06:10:10.0	5.79	99.84	-4.04	14.9	I - IV
70	2015-06-04T23:15:46.6	5.99	116.65	6.17	12.3	I - IV
71	2015-06-15T17:41:00.1	5.81	125.12	-9.62	18.6	I - IV
72	2017-11-18T16:07:05.0	5.81	128.10	2.59	14.1	I - IV
73	2018-08-09T05:25:34.9	5.91	116.20	-8.38	21.9	I - IV
74	2018-10-01T23:59:47.5	5.97	120.16	-10.57	22.0	I - IV
75	2019-07-08T18:52:38.1	5.88	126.38	0.35	19.6	I - IV
76	2019-10-14T22:23:59.9	6.08	101.04	-4.57	12.0	I - IV
77	2014-05-15T10:16:47.5	6.25	121.92	9.40	24.0	I - V
78	2014-08-06T11:45:28.7	6.19	127.92	-7.13	19.5	I - V
79	2015-09-24T15:53:33.7	6.58	131.23	-0.62	18.9	I - V
80	2015-12-20T18:47:38.1	6.05	117.56	3.66	12.0	I - V
81	2016-02-17T17:26:05.0	6.09	128.98	0.84	15.5	I - V
82	2016-06-09T04:13:11.2	6.06	116.29	-11.30	31.5	I - V
83	2016-10-09T14:46:28.1	5.82	127.48	1.82	141.1	I - V
84	2016-11-07T21:31:30.5	5.78	104.83	-8.32	41.8	I - V
85	2016-12-06T22:03:39.5	6.56	96.22	5.28	17.5	I - V
86	2017-04-11T21:21:01.5	5.83	124.70	7.74	12.0	I - V
87	2017-07-06T08:04:00.6	6.48	124.68	11.15	12.0	I - V
88	2017-07-10T01:41:52.6	5.80	124.76	11.08	13.6	I - V
89	2017-07-27T12:08:41.9	5.78	125.89	-3.52	20.7	I - V
90	2018-04-15T19:30:47.4	6.02	126.85	1.51	40.2	I - V
91	2018-08-19T14:56:35.6	6.93	116.75	-8.40	23.5	I - V
92	2018-08-28T07:08:17.9	6.18	124.14	-10.82	12.0	I - V
93	2018-10-02T00:16:48.8	5.92	120.07	-10.53	24.4	I - V
94	2019-01-21T23:59:28.3	6.09	119.09	-10.32	20.4	I - V
95	2019-01-22T05:10:09.4	6.44	119.07	-10.37	19.4	I - V
96	2019-07-07T15:08:47.3	6.91	126.10	0.55	30.5	I - V
97	2019-07-12T20:42:58.5	5.79	125.94	9.35	12.0	I - V
98	2019-09-14T16:21:32.2	5.86	128.57	-0.94	12.0	I - V
99	2019-09-25T23:46:48.4	6.47	128.39	-3.54	12.7	I - V
100	2019-10-16T11:37:10.3	6.42	125.01	6.86	17.1	I - V
101	2019-10-31T01:11:21.4	6.47	125.10	6.98	12.0	I - V
102	2019-11-15T01:17:43.0	5.98	126.25	1.69	28.8	I - V
103	2019-11-16T10:19:19.5	5.86	126.16	1.80	27.3	I - V
104	2019-11-18T13:22:12.8	5.90	124.87	7.69	1	I - V
105	2019-12-15T06:11:57.1	6.74	125.14	6.72	12.0	I - V
106	2020-01-07T06:05:24.9	6.34	96.27	2.21	12.0	I - V
107	2020-03-18T17:45:43.8	6.25	115.10	-11.23	12.0	I - V
108	2020-03-28T15:43:20.2	5.84	120.18	-1.68	18.9	I - V

Continued on next page

Table S2 – *Continued from previous page*

#	Focaltime	M_w	Longitude	Latitude	Depth [km]	Period bands
109	2014-02-03T22:36:42.4	5.87	128.20	-7.12	12.0	I - VI
110	2014-05-01T14:35:42.3	5.85	97.72	1.88	43.5	I - VI
111	2014-11-15T02:31:49.8	7.05	126.37	1.98	38.1	I - VI
112	2014-12-21T11:34:18.3	6.39	126.51	2.29	33.4	I - VI
113	2016-06-01T22:56:05.0	6.67	100.57	-2.18	28.9	I - VI
114	2017-01-10T06:13:55.9	7.27	122.78	4.57	621.5	I - VI
115	2017-05-29T14:35:28.3	6.58	120.40	-1.24	12.0	I - VI
116	2017-10-31T11:50:52.4	6.10	127.71	-3.83	12.0	I - VI
117	2019-01-06T17:27:24.2	6.63	126.63	2.48	34.9	I - VI
118	2019-07-01T16:59:26.1	5.93	124.09	9.15	545.8	I - VI
119	2014-04-17T04:38:20.0	5.76	122.82	4.55	575.0	IV
120	2018-03-08T13:06:14.5	5.53	116.65	6.15	12.0	IV, V
121	2018-08-25T18:33:18.7	5.54	116.99	-8.48	12.0	IV - VI
122	2014-10-30T12:11:36.8	5.76	117.48	-6.94	547.4	IV - VII
123	2016-03-19T08:51:26.5	5.70	129.43	-5.56	282.0	IV - VII
124	2016-04-15T04:50:12.9	5.59	126.98	2.06	108.7	IV - VII
125	2016-11-16T15:10:13.1	5.71	113.18	-9.14	105.7	IV - VII
126	2016-11-17T16:56:46.3	5.57	130.48	-6.33	127.5	IV - VII
127	2016-12-04T05:24:08.2	5.73	127.86	4.52	161.6	IV - VII
128	2017-03-21T23:10:28.1	5.69	115.27	-8.75	130.2	IV - VII
129	2018-03-25T08:58:12.6	5.73	128.50	-7.40	160.1	IV - VII
130	2018-12-03T14:00:09.3	5.54	128.72	-7.52	142.9	IV - VII
131	2018-12-30T08:39:14.2	5.80	102.25	-2.68	175.5	IV - VII
132	2019-07-16T00:18:38.3	5.78	114.50	-9.01	102.7	IV - VII
133	2020-02-05T18:12:36.8	6.23	113.09	-6.11	597.0	IV - VII
134	2017-12-28T17:20:23.4	5.75	126.83	4.10	32.5	VI
135	2008-09-11T00:00:06.8	6.58	127.34	1.91	119.6	VI, VII
136	2015-02-25T01:31:44.7	5.67	119.87	6.15	18.4	VI, VII
137	2016-04-13T18:21:55.9	5.97	121.94	7.84	24.2	VI, VII
138	2017-05-20T01:06:16.4	5.98	124.02	9.33	544.6	VI, VII
139	2018-02-02T00:20:43.6	5.60	125.13	-0.32	30.9	VI, VII
140	2018-06-02T16:29:03.2	5.80	126.76	4.59	28.2	VI, VII
141	2019-02-07T04:15:33.3	5.72	126.39	1.53	40.5	VI, VII
142	2019-03-24T04:37:39.1	6.15	126.36	1.77	41.9	VI, VII
143	2019-06-14T20:10:55.2	5.71	130.77	-5.80	129.2	VI, VII

S2.2 Event misfits

In Figure S2.1, we show the event misfit reduction for the final model relative to the starting model, normalized to the total misfit decrease. No single event contributes more than 3 % to the total misfit decrease, indicating that the inversion is not driven by data from only a small subset of events. Furthermore, no patterns associated with a dependency on focal mechanisms or hypocentral depths are identifiable.

S3 Station availability

Figure S3 presents a map of all 440 stations used in this study. Publicly available waveforms including instrument responses were downloaded automatically using *obspyDMT* (Hosseini & Sigloch, 2017), which accesses over 20 data centers via the *International Federation of Digital Seismograph Networks (FDSN)* and *ArcLink* interfaces. To date, only a small proportion of permanent network stations have their data made publicly accessible within the region. Thus, the majority of our dataset consists of stations from several networks with restricted access:

- *IA*, accessed via the *Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) WebDC3* web interface (Bianchi et al., 2015)
- Most of the *MY* network
- *YC* (Rawlinson, 2018)
- *YS* (Miller, 2014) accessed via the *Australian Passive Seismic Server (AusPass) WebDC3* web interface (Bianchi et al., 2015)
- *9G* (Greenfield, 2018)

Data from a Taiwanese station in the South China Sea (*TW.VNAS*) is recorded as part of the *Broadband Array in Taiwan for Seismology (BATS)*, and was requested from the *Academia Sinica, Institute of Earth Sciences* in Taipei since it is publicly available only before 2014. More information about individual networks can be found here:

`https://www.fdsn.org/networks/`.

S4 Gradient preconditioning

Event kernels usually show large sensitivities around the source region, with values typically around five times higher than the surrounding region in this study. Thus, these imprints have to be removed to avoid a strong localization of model updates (see Figure S4a). We favor not applying the source imprint removal to the summed gradient, but to the event kernels individually, otherwise the gradient will turn into a "Swiss cheese" and constraint around *all* event hypocentres is lost. The removal region is defined by a sphere with the radii for the source imprint removal shown in Table 1, and a radius of 50 km for each receiver. However, receiver imprints are smaller and usually wiped out by the smoothing operator described in the next section.

Initial model updates (100 – 65 s) use an anisotropic smoothing operator (horizontal and vertical smoothing lengths are fixed across the model). From 50 s onwards, depth-dependent smoothing is applied in order to account for the local wavelengths of the model. The respective wavelengths are based on the shear wave velocity of the prior model. The effect of the smoothing operator is presented in Figure S4b.

S5 Technical details

In Table S5, an overview of the technical parameters is given. Throughout the entire inversion, 1.5 elements per minimum wavelength are used and velocity seismograms are considered. Note that each period decrease is accompanied by a decrease in the smoothing lengths since the smoothing operator is based on the minimum wavelength considered.

The simulation time is decreased since the surface wave train becomes more compact, which spares computational resources.

To account for the remaining non-physical boundaries of the computational domain, a first order Clayton-Enquist boundary condition (Clayton & Engquist, 1977) is applied and the 3-D wavefield is attenuated within absorbing boundary layers following Kosloff and Kosloff (1986). The absorbing layer width is based on 3.5 minimum wavelengths at a reference velocity of 6 km/s.

S6 Waveform fits

Figure S6 shows additional three-component waveform fits not shown in the main text.

S7 Depth slices

Figures S7a – S7d present depth slices from 50 to 700 km for all inversion parameters (v_{SV}, v_{SH}, v_P and ρ).

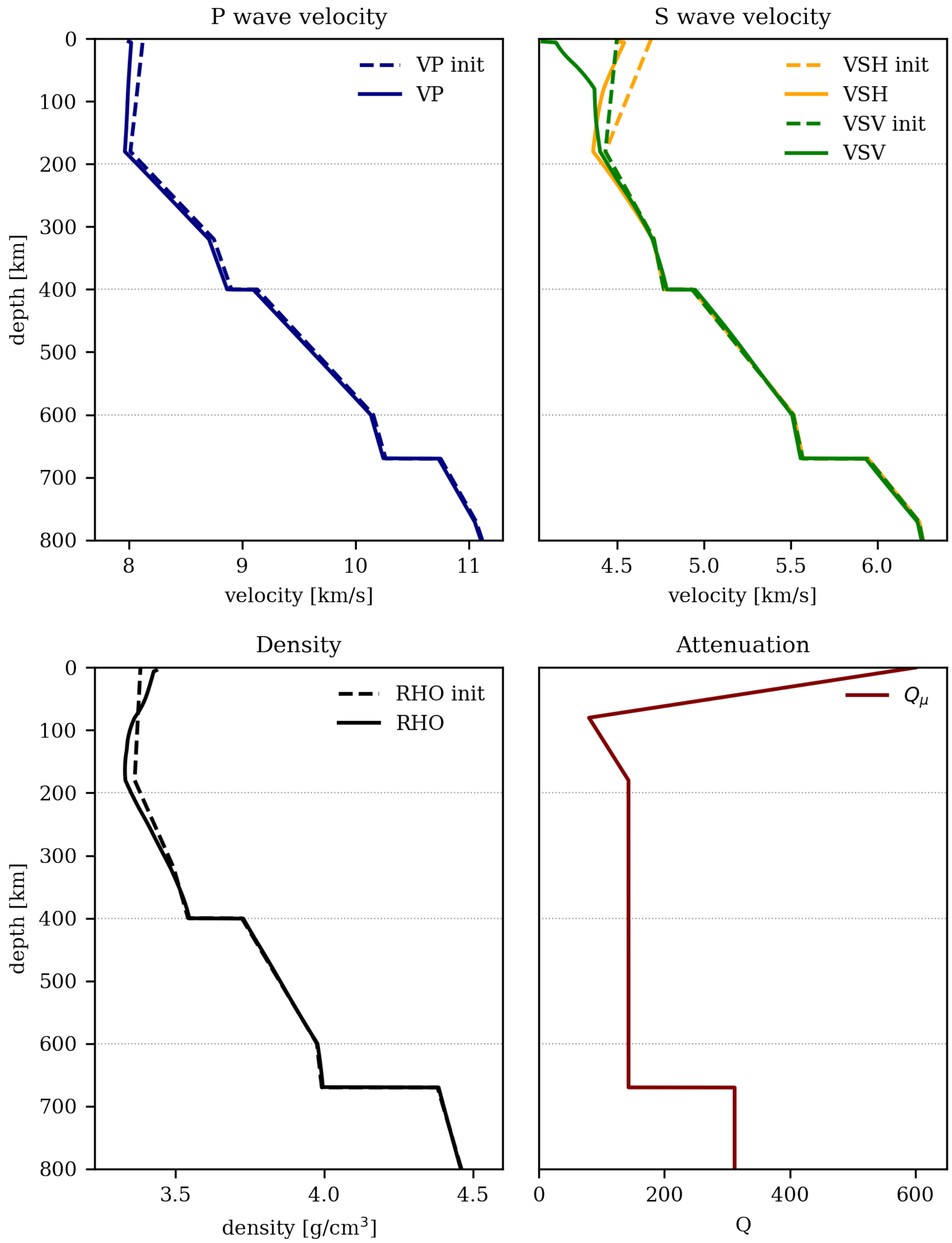

Figure S1. Absolute values for the starting model (dashed lines) and final model (solid lines) for the upper 800 km, which is equivalent to the mesh depth in this study (without absorbing layers). Q_μ and Q_κ remain constant throughout the inversion. Q_κ is not shown, but has a constant value of 57,823.

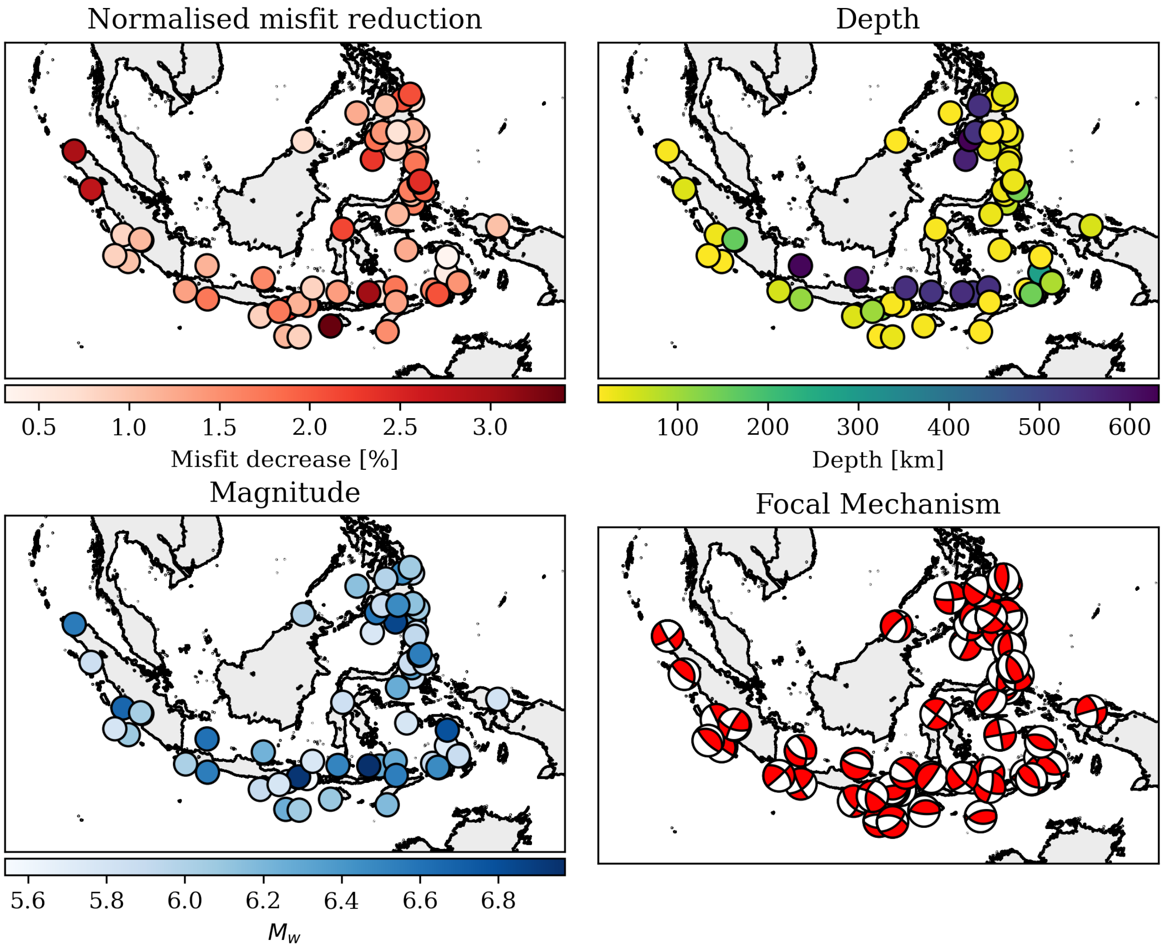

Figure S2.1. *Top left:* The normalized event misfit reduction for each event. *Top right:* Events colored by depth. *Bottom left:* Events colored by magnitude. *Bottom right:* Focal mechanisms.

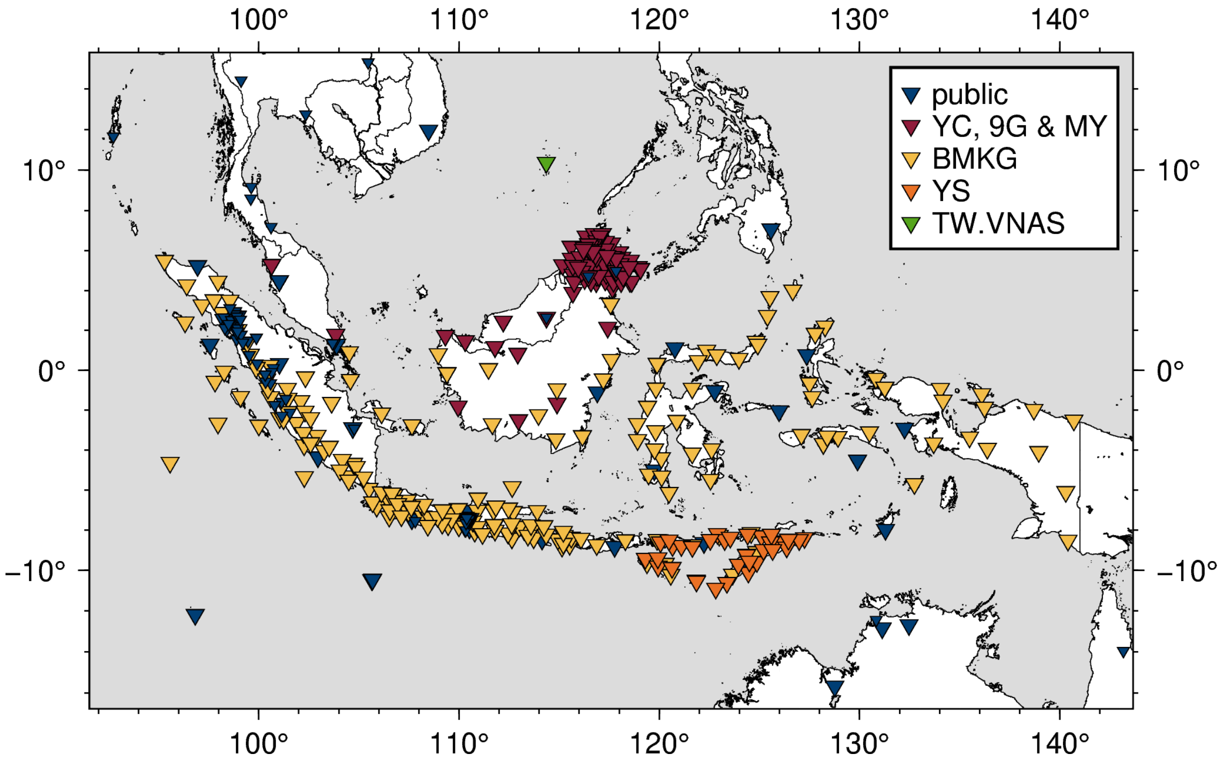

Figure S3. Map showing the 440 on-shore stations used in this study. Publicly available stations are shown in blue. Public stations contributing < 30 waveforms to the inversion are plotted in smaller size, e.g. temporary networks on Java and Sumatra.

Table S4. Overview of the smoothing lengths chosen throughout this study. During the initial period bands (100 – 65 s), a purely anisotropic, diffusion-based smoothing (PA) is applied. From 50 s onwards, a depth-dependent, anisotropic, diffusion-based smoothing (DD) is used.

period band	**smoothing type**	**smoothing lengths – horizontal, vertical**	**source imprint removal [km]**
100 s (Ia)	PA	450, 100 km	500
100 s (Ib)	PA	375, 100 km	500
80 s (IIa)	PA	375, 80 km	450
80 s (IIb)	PA	300, 80 km	450
65 s (III)	PA	300, 65 km	400
50 s (IVa)	DD	1.0, 0.2 $\lambda_{\min}$	350
50 s (IVb)	DD	0.75, 0.2 $\lambda_{\min}$	350
40 s (V)	DD	0.5, 0.2 $\lambda_{\min}$	300
30 s (VI)	DD	0.5, 0.2 $\lambda_{\min}$	300
20 s (VII)	DD	0.5, 0.2 $\lambda_{\min}$	300

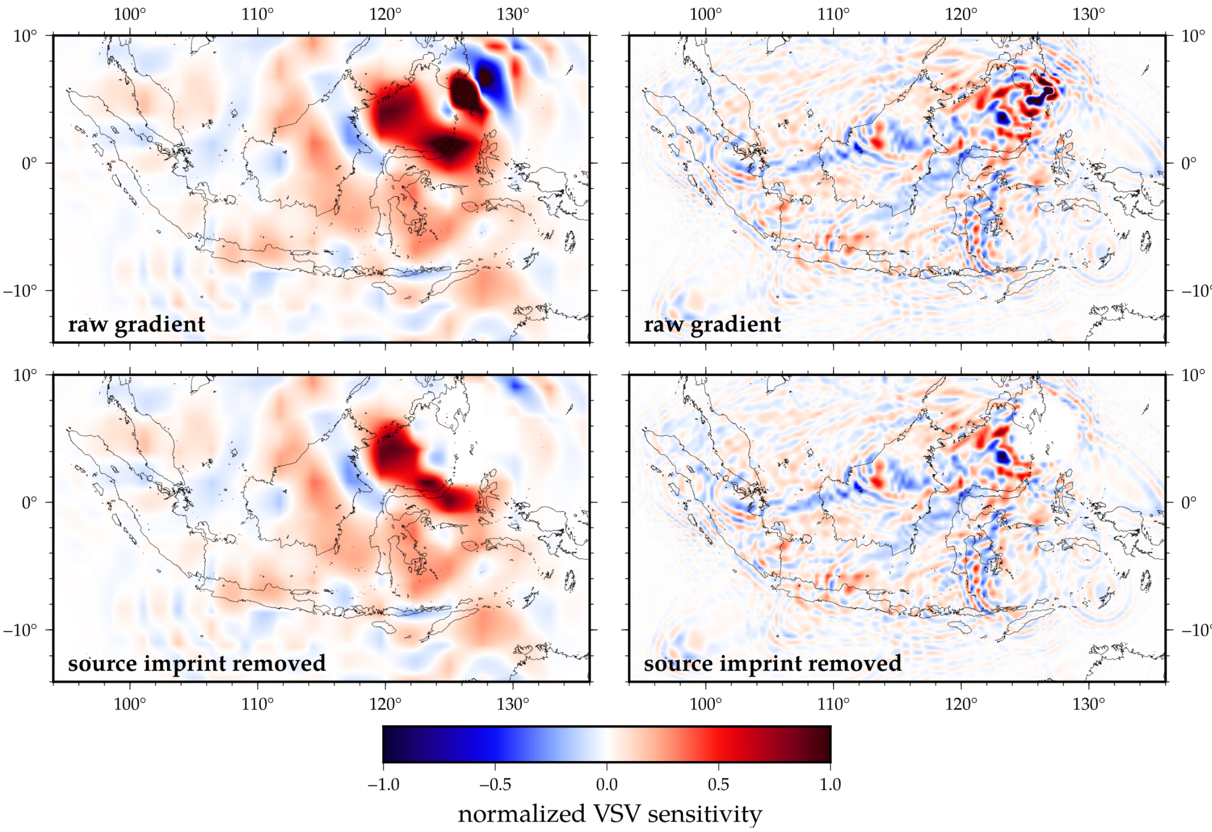

Figure S4a. Source imprint removal for a M_w6.3 event southeast of the Philippines at 100 s (left column, iteration 5) and 20 s (right column, iteration 87) for a depth slice at 75 km, which is the event's hypocentral depth taken from *GCMT*. The upper row shows the raw v_{SV} event kernel, and the bottom row shows the event kernel after the source imprints have been removed. Note the radius decrease of the source imprint removal and the overall smaller scale structure as we consider shorter periods. The sensitivities are normalized per period band since the gradients of 100 and 20 s vary by two orders of magnitude. The receiver imprints have not yet been removed.

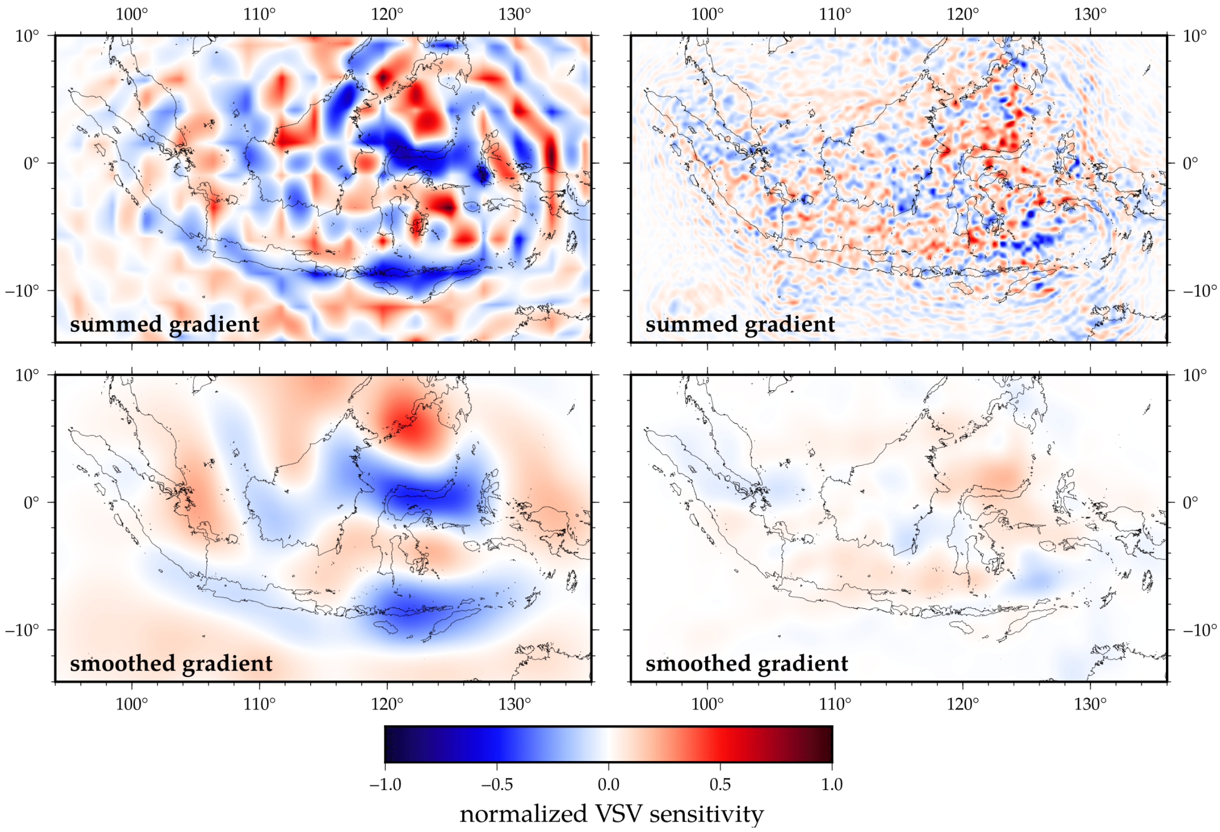

Figure S4b. Smoothing of the misfit gradient, at 100 s (left column, iteration 5) and 20 s (right column, iteration 87) for a depth slice at 75 km. The upper row shows the summed v_{SV} gradient after the source imprint has been removed, and the bottom row shows the smoothed gradient. The sensitivities are normalized per period band since the gradients of 100 and 20 s vary by two orders of magnitude. Note the sensitivity to smaller scale structure as the period is decreased.

Table S5. Overview of technical parameters.

period band	**# iterations**	**# mesh elements**	**absorbing layer width [km]**	**simulation time [s]**	**time step [s]**
100 s (Ia)	0 - 5	14,250	2,100	1,600	0.55
100 s (Ib)	5 - 8	14,250	2,100	1,600	0.55
80 s (IIa)	8 - 15	17,600	1,680	1,600	0.55
80 s (IIb)	15 - 19	17,600	1,680	1,600	0.55
65 s (III)	19 - 27	23,400	1,365	1,600	0.55
50 s (IVa)	27 - 32	33,866	1,050	1,600	0.5
50 s (IVb)	32 - 46	33,866	1,050	1,600	0.5
40 s (V)	46 - 57	49,680	840	1,500	0.45
30 s (VI)	57 - 70	84,796	630	1,250	0.375
20 s (VII)	70 - 87	207,636	420	1,100	0.28

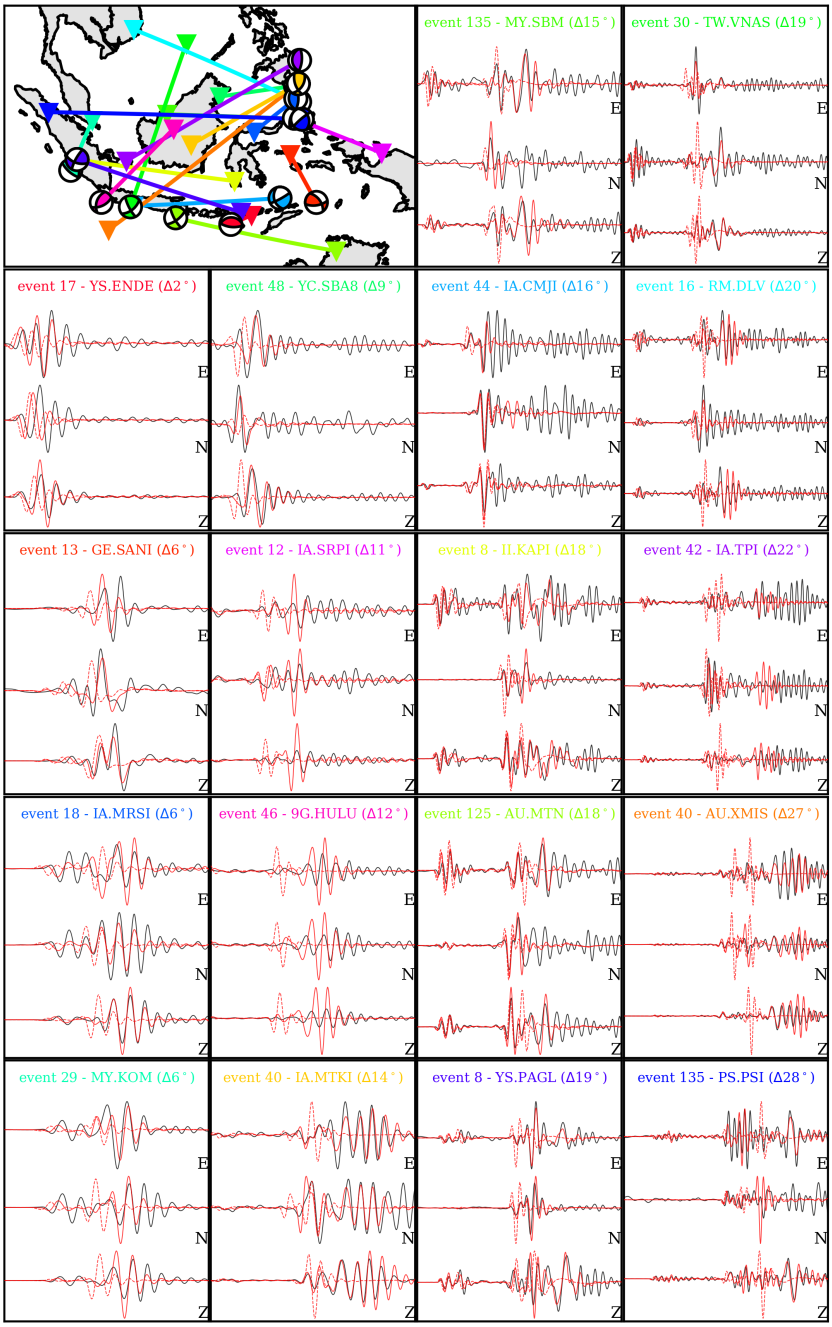

Figure S6. Three-component waveform match for the initial model (iteration 0, dashed red), the final synthetics (iteration 87, solid red) and observed waveforms (black) for 18 source-receiver pairs. The event numbers are taken from Table S2.

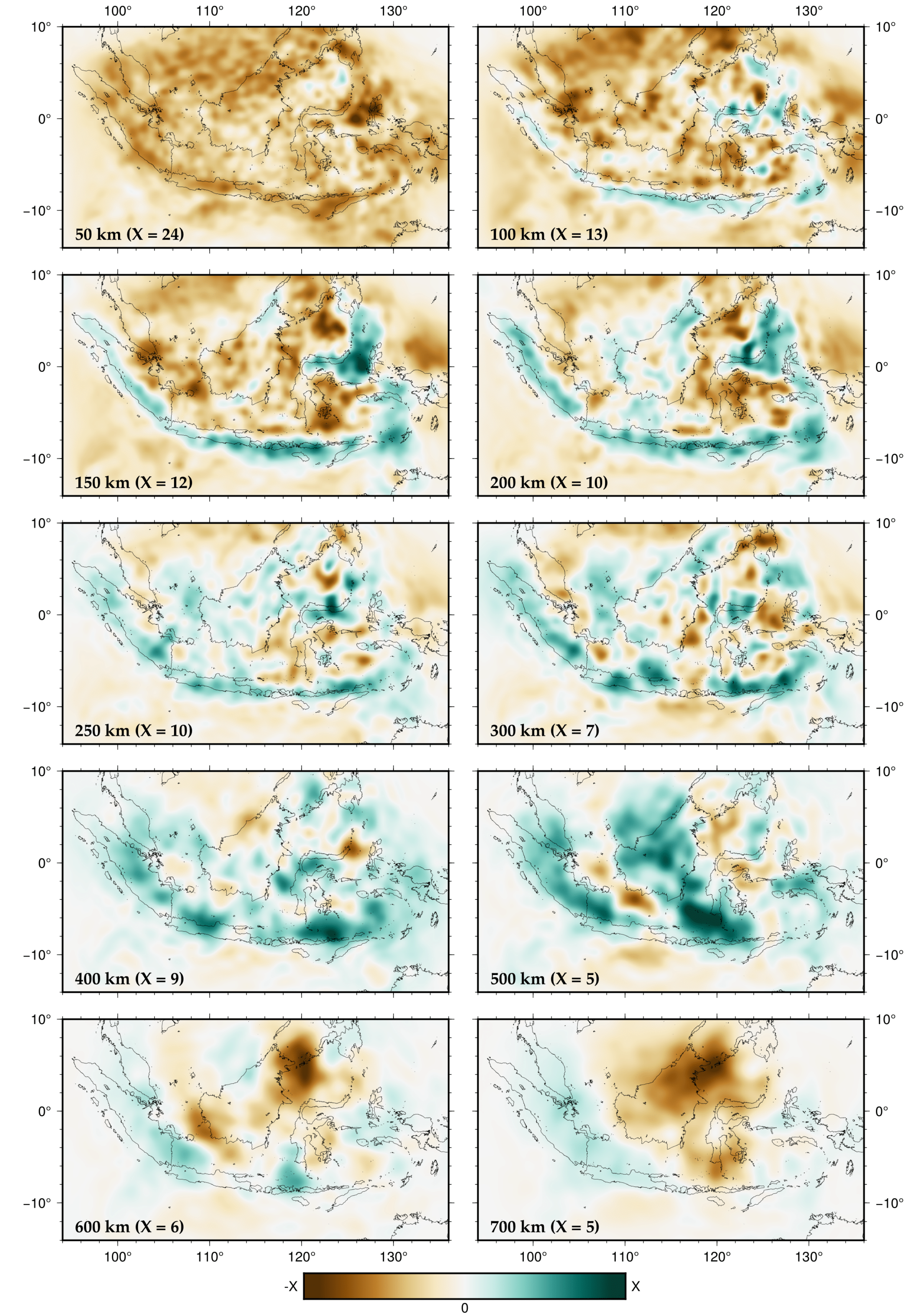

Figure S7a. v_{SV} depth slices from 50 to 700 km. Perturbations are in % relative to the initial model. The limits of the colorscale X are shown in the lower left corner of each plot.

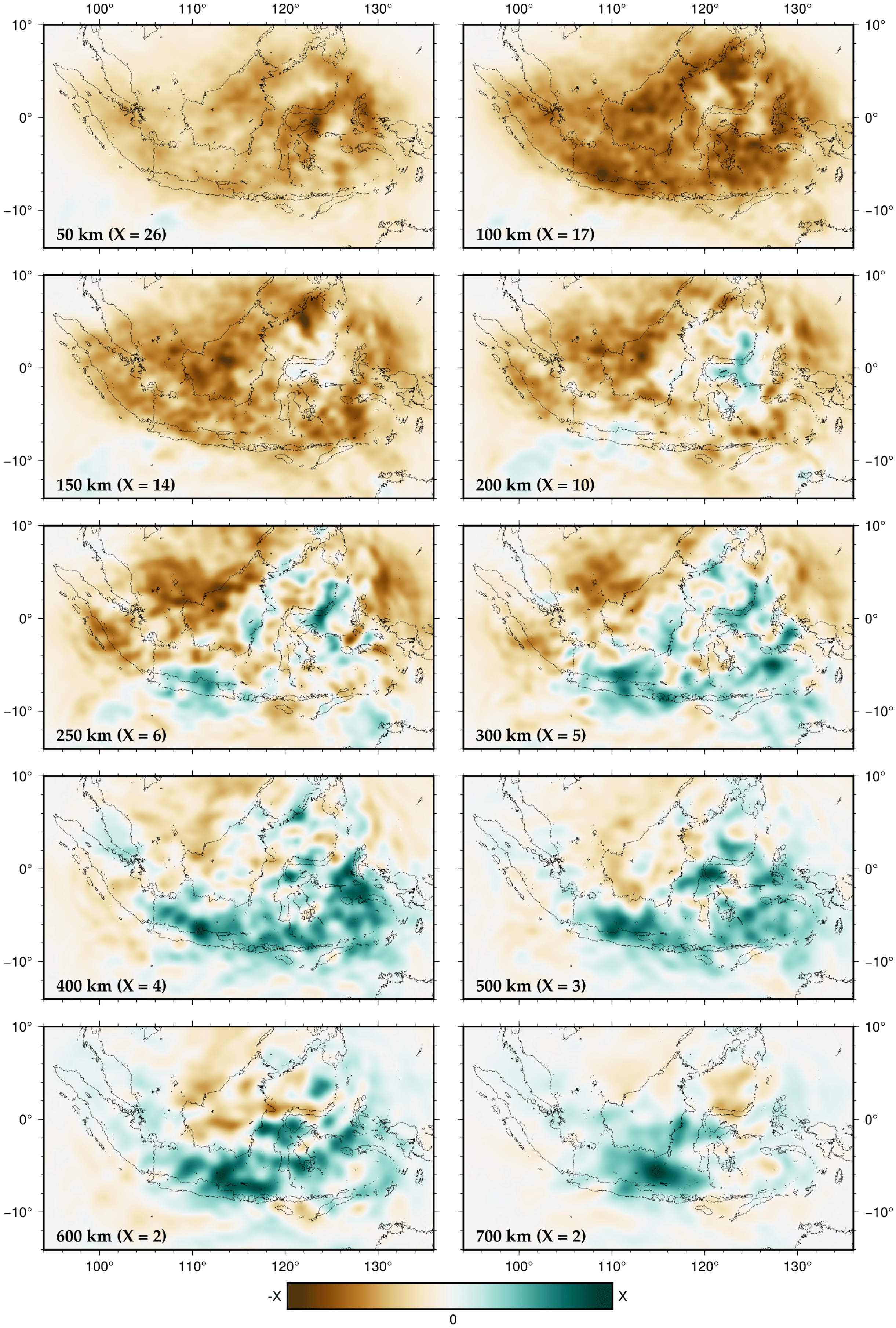

Figure S7b. v_{SH} depth slices from 50 to 700 km. Perturbations are in % relative to the initial model. The limits of the colorscale X are shown in the lower left corner of each plot.

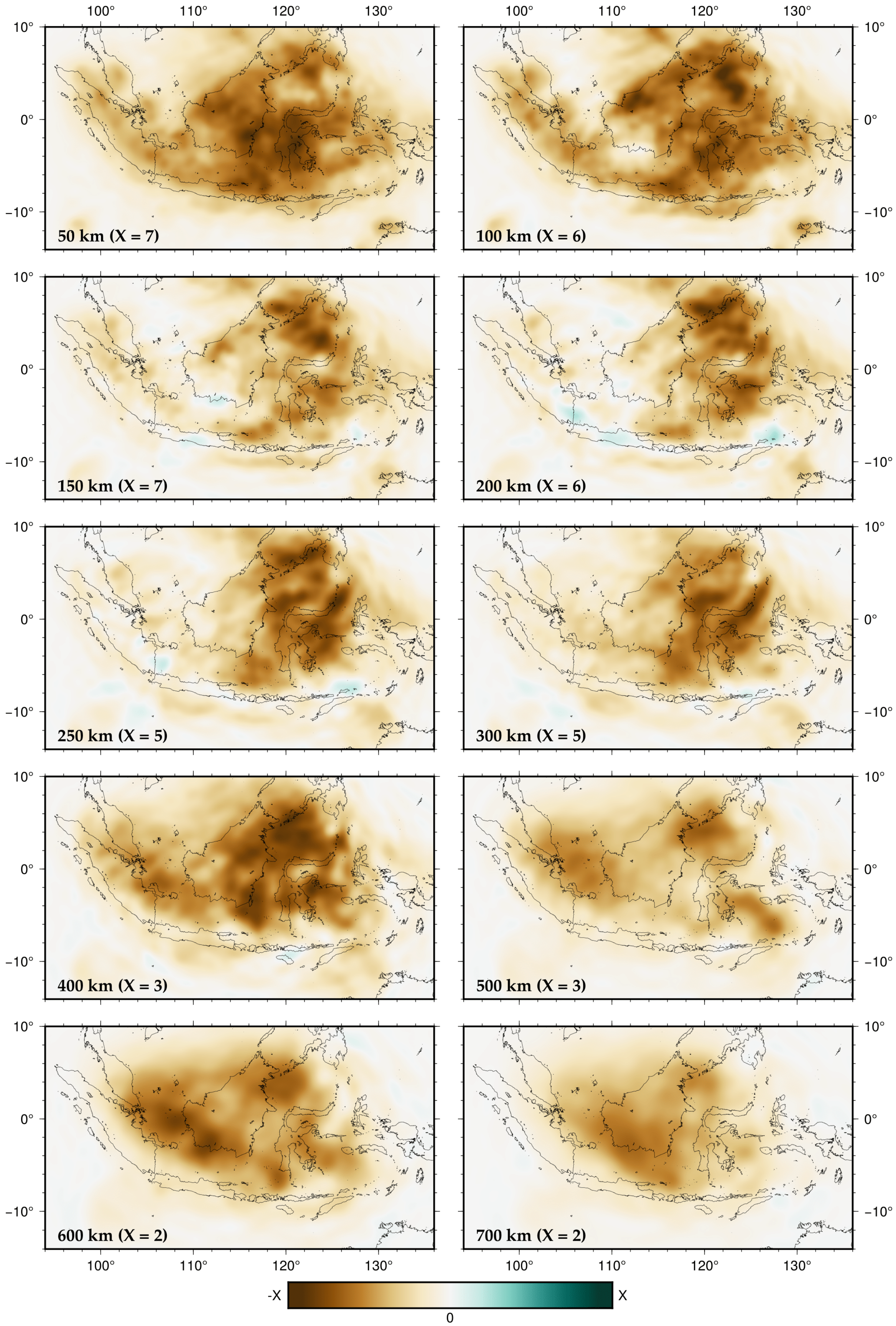

Figure S7c. v_P depth slices from 50 to 700 km. Perturbations are in % relative to the initial model. The limits of the colorscale X are shown in the lower left corner of each plot.

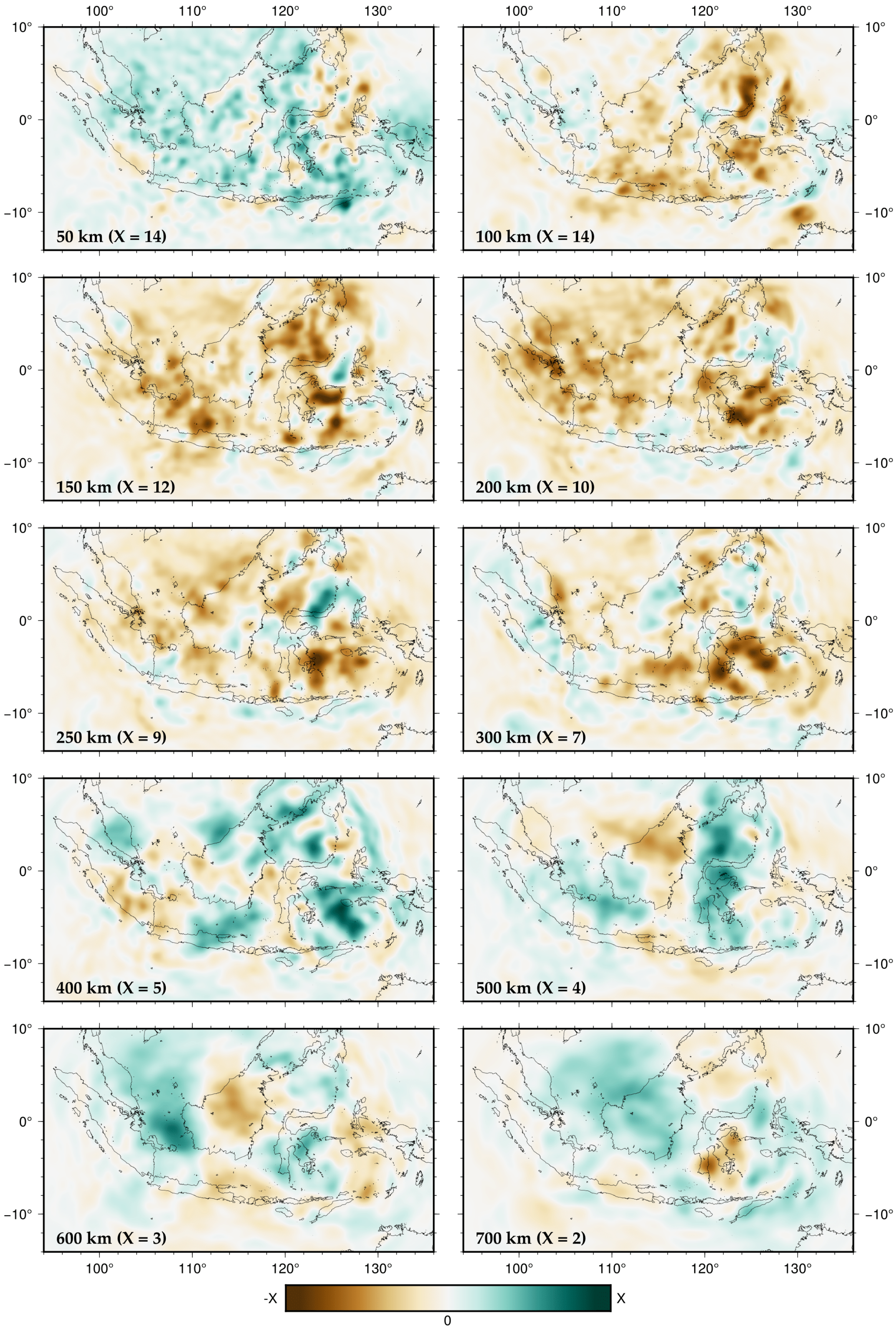

Figure S7d. Density (ρ) depth slices from 50 to 700 km. Perturbations are in % relative to the initial model. The limits of the colorscale X are shown in the lower left corner of each plot.