St John's College
University of Cambridge

PhD Thesis

Seismic structure beneath Southeast Asia from adjoint waveform tomography

---

Author:
Deborah Wehner

Supervisor:
Nick Rawlinson

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
in the

Department of Earth Sciences

1 June 2022

Für Mama

Declaration

I, Deborah Wehner, hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 275 numbered pages, of which not more than 225 pages are text, appendices, illustrations, and bibliography.

Deborah Wehner
1 June 2022

UNIVERSITY OF CAMBRIDGE

Abstract

Department of Earth Sciences

Doctor of Philosophy

Seismic structure beneath Southeast Asia from adjoint waveform tomography

by Deborah WEHNER

Seismic tomography has played a crucial role in the illumination of deep Earth structure. Most existing tomographic methods are based on seismic ray theory and hence do not fully account for the true physics of wave propagation. Recent computational advances allow us to embrace the full complexity of seismic wave propagation by accurately solving the 3-D seismic wave equation numerically. This can account for effects such as wavefront healing, interference, scattering and (de)focusing, which are often ignored or not properly captured by other methods such as ray tracing. Thus, such methodologies are particularly suitable for strongly heterogeneous regions such as Southeast Asia, where large variations in elastic parameters are likely to be present. Here, an unprecedented dataset and access to sizeable computational resources allow their application to Southeast Asia for the first time.

In the first part of this thesis, a 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) is presented. FWI is a non-linear imaging method, where an initial model is updated in order to minimise the difference between observed and predicted waveforms. Based on $>$ 3,000 h of analysed waveform data gathered from $\sim$13,000 unique source-receiver pairs and filtered at periods between 20 – 150 s, isotropic P-wave velocity, radially anisotropic S-wave velocity and density are imaged 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.

SASSY21, the final model after 87 iterations, appears to be robust since it is able to explain true-amplitude data from events and receivers not included in the inversion. The new model 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.

In the second part of this thesis, *SASSY21* is used as a starting model to obtain a more refined image of the eastern Indonesian region, using seismic data filtered at periods from 15 – 150 s. In this study, the fluid ocean is accounted for explicitly by solving a coupled system of the acoustic and elastic wave equation. This is computationally more expensive but allows seismic waves within the water layer to be simulated, which becomes important at shorter periods. The effects arising from surface topography, bathymetry and the fluid ocean on synthetic waveforms become pronounced at periods $\leq$ 20 s. In particular, surface elevation can result in a considerable phase advance and change in amplitude of the surface wave train, and has an effect on both horizontal and the vertical seismogram components for this simulation setup. The fluid ocean results in a phase delay as well as a change in amplitudes and duration of the surface wave train, and affects both the radial and vertical components. At periods $\leq$ 20 s, accounting for the fluid ocean explicitly can lead to more realistic lithospheric velocities and a more refined image compared to the commonly used ocean load approximation, even at greater depths. Furthermore, it allows for an improved waveform match for source-receiver paths passing partially or entirely through oceanic regions.

The final model, *SASSIER22*, after 34 iterations reveals a convergent double-subduction along the southern segment of the Philippine Trench, which was not evident in the starting model and transitions to a divergent system in the Molucca Sea further south. A more detailed illumination of the slab beneath the North Sulawesi Trench subduction zone reveals a pronounced positive wavespeed anomaly down to $\sim$200 km depth, consistent with the maximum depth of seismicity, and a more diffuse but aseismic positive wavespeed anomaly that continues to the 410 km discontinuity.

Publications

Journal publications related to this dissertation

Wehner, D., Blom, N., Rawlinson, N., Daryono, Böhm, C., Miller, M. S., Supendi, P & Widiyantoro, S. (2022). *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.

Wehner, D., Rawlinson, N., Greenfield, T., Daryono, Miller, M. S., Supendi, P., Lü, C. & Widiyantoro, S. (2022). *SASSIER22*: Full-waveform tomography of the eastern Indonesian region that includes topography, bathymetry and the fluid ocean. Geochemistry, Geophysics, Geosystems. *Under review.*

Bacon, C. A., Rawlinson, N., Gilligan, A., Pilia, S., **Wehner, D.**, Cornwell, D. G. & Tongkul, F. (2022). Lithospheric anisotropy characterises the post-subduction setting of northern Borneo: new results from XKS splitting analysis. Geochemistry, Geophysics, Geosystems. *Under review.*

Noe, S., van Herwaarden, D.-P., Thrastarson, S., Masouminia, N., van Driel, M., Böhm, C., Ma, J., Bunge, H.-P., Gao, Y., Tilmann, F., **Wehner, D.**, Rawlinson, N. & Fichtner., A. Collaborative Seismic Earth Model: Generation 2. *In preparation.*

Cornwell, D. G., Gilligan, A., Rawlinson, N., Tongkul, F., Pilia, S., Hall, R., Bacon, C. A., Greenfield, T., Baker, S., **Wehner, D.** & Linang, H. T. Intricate post-subduction crust and upper mantle structure of Sabah, northern Borneo, from teleseismic receiver functions. *In preparation.*

Contents

Preamble

In 2013, the Society of Exploration Geophysicists (SEG) hosted a workshop, which was advertised as follows:

> *"Full-waveform inversion (FWI) has emerged as the final and ultimate solution to the Earth resolution and imaging objective."*[1]

Almost a decade later, FWI remains a buzzword in the petroleum industry and has become an established part of earthquake seismology and imaging. However, it still seems poorly understood outside a small group of people and traditional tomography methods remain the standard approach to image the Earth's interior. This is mainly a result of the high computational requirements of FWI and the fact that its application is still subject to ongoing development and improvement. Thus, only a limited (but increasing) volume of literature and (automated) workflows are publicly available, and they often lack implementation details. I hope this thesis can provide guidance for the interested reader who is getting started with FWI in earthquake seismology. Many aspects of the methodology and workflow can be applied to other regions (or even planets) at different scales, while others will depend on the given dataset. Let's get started!

A note on the nomenclature: I use the terms *full-waveform inversion (FWI)* and *(adjoint) waveform tomography* interchangeably throughout this thesis. There are good reasons for avoiding the term *full-waveform inversion* entirely since we hardly exploit the entire information content of a seismogram. This is mainly a result of the long periods currently considered, data selection procedures and the fact that we do not invert for all parameters governing the seismic wave equation (more on this in Chapter 3). However, the term is frequently encountered in publications and at conferences and thus, I believe it makes sense to add it to the vocabulary. Note that I will only use these terms for studies that solve the 3-D seismic wave equation numerically and compute accurate sensitivity kernels relative to a 3-D model at each iteration using adjoint techniques. However, it should be noted that other finite frequency studies are also often referred to as *(full-)waveform tomography*, including studies that approximate sensitivity volumes (e.g. French and Romanowicz, 2014; Celli et al., 2020). Furthermore, some FWI studies obtain sensitivity kernels using the scattering-integral method (e.g. Zhao et al., 2005), which is closely related to the adjoint-state method (e.g. Tromp et al., 2005; Chen et al., 2007).

[1] `https://seg.org/Events/Past-Events/Full-Waveform-Inversion-From-Near-Surface-to-Deep`

1 Introduction

1.1 Motivation

Seismic tomography has played a crucial role in the illumination of deep Earth structure since the first pioneering studies of the mid 1970s (e.g. Aki et al., 1976; Dziewoński 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. Recent computational advances have enabled us to overcome the limitations of ray theory, thus promising high-resolution images and a more reliable quantification of anomalies, which opens up new avenues for a more robust interpretation of seismic models in terms of composition, temperature, melt and other material properties (Tromp, 2020).

Adjoint waveform tomography – often referred to as *full-waveform inversion* or *FWI* – embraces the full complexity of seismic wave propagation, by accurately solving the 3-D seismic wave equation numerically. It is a non-linear imaging method that most commonly updates an Earth model iteratively in order to minimise the difference between observed and synthetic waveforms (see Figure 1.1). The method permits direct comparison with waveforms recorded by seismometers in the field, making use of the information contained in

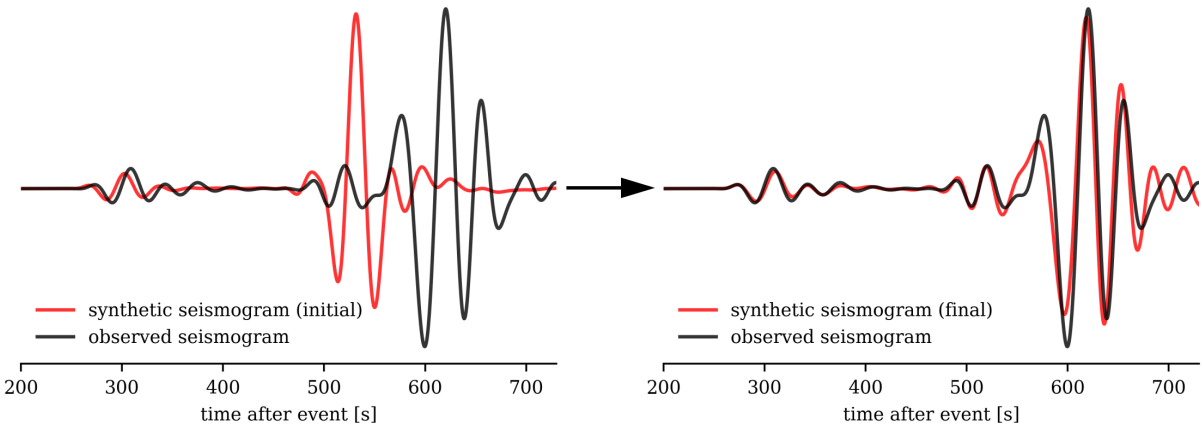


Figure 1.1: Adjoint waveform tomography objective. The aim is to determine a physically meaningful model that minimises the misfit between synthetic and observed waveforms.

the amplitudes and phases of, in theory[1], entire seismograms to resolve the state of the Earth. FWI coherently incorporates body and surface waveforms and allows the joint imaging of crust and mantle. There is no need to identify seismic phases (such as P- or S-wave arrivals) since FWI effectively deals with interfering phases by using information contained in the full wavefield. Furthermore, it can account for effects such as wavefront healing, scattering, interference and (de)focusing, which are not accurately modelled with ray theory (e.g. Rickers et al., 2012). This makes it a particularly suitable imaging method for tectonically complex settings such as Southeast Asia, where large variations in elastic parameters are likely to be present, and the assumptions of ray theory become less valid.

The pronounced tectonic activity in Southeast Asia is underscored by the fact that approximately 20 % of earthquakes of magnitude $M_w \geq 5$ worldwide are generated along its highly active subduction zones (see Figure 1.2). The region has received increased attention after the 2004 Sumatra-Andaman earthquake, which triggered a devastating tsunami causing $>$ 200,000 fatalities (Marano et al., 2010). Today, the area is of growing economic and geopolitical importance[2], yet it faces the ongoing risk of severe natural hazards such as floods, forest fires, large-magnitude earthquakes (such as the 2012 Indian Ocean earthquakes, Muttarak and Pothisiri, 2013), tsunamis and volcanic eruptions (such as the 2018 Anak Krakatoa eruption, Petley, 2019).

[1] In reality, only selected waveform portions are compared for reasons elaborated later in this thesis. However, the ultimate aim is to explain broadband seismograms "wiggle by wiggle" with hardly any human intervention, which is known as the *Tarantolian black box*.

[2] `https://geopoliticalfutures.com/aseans-growing-economic-importance/`

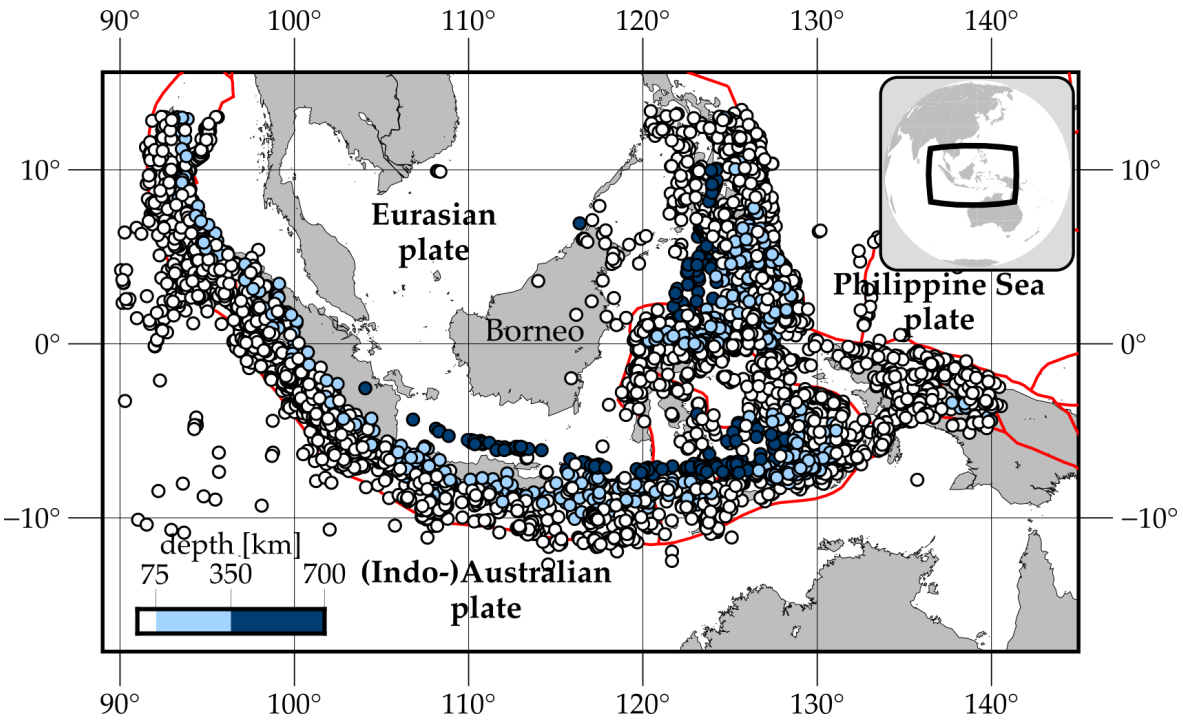


Figure 1.2: Southeast Asia's location and seismicity ($M_w > 5$) taken from the *ISC* catalogue (International Seismological Centre, 2016), coloured by depth. Red lines indicate plate tectonic boundaries according to Bird (2003).

Southeast Asia's tectonic complexity is mainly the result of its location at the confluence of three key tectonic plates: the Eurasian, (Indo-)Australian, and Philippine Sea plates. In the west, large-magnitude earthquakes are generated by the oceanic crust of the Australian plate subducting beneath Indonesia. This transitions to an arc-continental collision along the curved Banda Arc. 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. The east is characterised by events with depths of up to 700 km and a complex configuration of minor tectonic plates, resulting in several subduction zones with a diverse range of ages, geometries and lengths (e.g. Hall, 2019). The westward subduction of the Philippine Sea plate is associated with a slab of unknown depth extent. It is believed to be associated with convergent double subduction (e.g. Rangin, 1991; Lagmay et al., 2009) but this is neither evident in the subduction zone geometry model *Slab2* (Hayes et al., 2018) nor the plate tectonic boundary model by Bird (2003). Furthermore, an unusual divergent double subduction zone within the region has been imaged in P-wave tomographic studies (Amaru, 2007; Obayashi et al., 2013) but its extension to the north remains unclear.

Overall, the region provides a unique setting to investigate a variety of primary tectonic processes such as ongoing subduction, collision (both arc-arc and arc-continent) and a post-subduction setting around northern Borneo, where two opposing subduction systems ceased in the Miocene ($\sim$23 – 5 Myr), which may be key to understanding the termination phase of the tectonic subduction cycle (e.g. Pilia et al., 2019). So far, studies that investigate the seismic structure of Southeast Asia are either global or regional, using body or (multimode) surface wave tomography methods only (e.g. Widiyantoro and Hilst, 1996; Bijwaard et al., 1998; Lebedev and Nolet, 2003; Fukao and Obayashi, 2013; Schaeffer and Lebedev, 2013; Miller et al., 2016; Zenonos et al., 2019; Harris et al., 2020). Several of these studies have identified a number of subducting slabs in Southeast Asia, mainly along the Indonesian volcanic arc. 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 artefacts due to poor data coverage. Hence, a new large-scale image of the entire region including the upper mantle is required to illuminate all the region's complexity, which will be the objective of the first part of this thesis.

The success of adjoint waveform tomography is strongly dependent on being able to produce realistic synthetic waveforms from potentially complex models of the Earth's interior. In particular, surface topography, bathymetry and the ocean can have a significant effect on ground motion; for example, surface topography can lead to seismic wave scattering (Lee et al., 2008) and the fluid ocean elongates the surface wave train (Todoriki et al., 2016; Fernando et al., 2020). However, these effects are not routinely accounted for in surface wave and waveform tomography, largely due to the need for sophisticated meshing techniques, plus the added computational burden associated with simulating waves in the low-velocity fluid ocean. Instead, one common approach for dealing with this issue is to approximate the fluid ocean by the weight of its water column ("ocean load", Lei et al., 2020; Rodgers

et al., 2022) but this is only a valid assumption when periods are long compared to the water layer thickness (Komatitsch and Tromp, 2002; Zhou et al., 2016). In the second part of this thesis, the continental-scale waveform tomography model is used as a starting model for the eastern Indonesian region, which exhibits a particularly good data coverage and thus allows for a more refined image. In this study, the fluid ocean is accounted for explicitly, which is computationally more expensive but becomes important when considering shorter-period data. The final model is compared to the tomographic result obtained with the frequently used ocean loading approximation, and results from investigating path-dependent effects of surface elevation (topography and bathymetry) and the fluid ocean on synthetic waveforms are presented.

1.2 Objectives of this dissertation

This project aims to understand Southeast Asia as a whole, including the highly active subduction boundaries and the poorly understood regions of Borneo and Sulawesi. Access to restricted data from several seismic networks results in an unprecedented dataset that allows for the application of adjoint waveform tomography to Southeast Asia for the first time. The research objectives of this thesis can be summarised as follows:

- What are the best practices for performing full-waveform inversion (FWI) in Southeast Asia? How do they translate across different scales?
- Can FWI reveal previously undiscovered features and help to improve the quantification of anomalies?
- Can we illuminate the strongly curved Banda Arc slab geometry?
- Can we unravel the mysteries related to subduction termination in the region?
- Can we provide new insights into the complex subduction system around the Philippines?
- What is the effect of the fluid ocean on synthetic waveforms and how does it affect the tomographic result?

This thesis is organised as follows:

Chapter 2 provides an overview of the geological setting and tectonics of Southeast Asia.

Chapter 3 is dedicated to the methodological background of adjoint waveform tomography, i.e. I will elaborate on how an initial model is updated based on the waveform differences between observed and synthetic data in order to reduce the waveform misfit.

Chapter 4 will elaborate on the inversion setup of the Southeast Asian adjoint waveform tomography, placing a particular focus on producing realistic synthetic seismograms by accounting for topography, bathymetry and ocean load. The inversion setup also includes the choice of a starting model and the event and data evaluation.

Chapter 5 presents the new 3-D seismic model of Southeast Asia obtained from adjoint waveform tomography – *SASSY21* – including its model validation, discussion of key highlights and limitations. Furthermore, *SASSY21* is compared to other tomography models that span Southeast Asia.

In **Chapter 6**, *SASSY21* is used as a starting model for the eastern Indonesian region, which is characterised by particularly good data coverage and thus facilitates the recovery of a more refined image. The fluid ocean is meshed explicitly and the minimum period is decreased to 15 s (compared to 20 s for *SASSY21*). The final model, *SASSIER22*, is validated and discussed, and it is compared to the tomographic result obtained with the more commonly used ocean loading approximation.

In **Chapter 7**, the transferability of the inversion setup described in Chapter 4 to a contrasting and significantly smaller study region is examined. This chapter contains preliminary results for adjoint waveform tomography of Iceland at a regional scale, which includes the inversion setup and a seismic moment inversion.

Chapter 8 summarises the results of this dissertation and concludes with a discussion of future research directions.

2 Southeast Asia: Geological setting and tectonics

Southeast Asia lies amidst one of the world's most complex tectonic settings, primarily due to its location at the confluence of three large tectonic plates: the (Indo-)Australian, Eurasian and Philippine Sea plates (see Figure 2.1). Most seismicity occurs at the highly active boundaries between these plates (see Figure 1.2), which have generated thousands of kilometres of subducted material[1] and feature slabs descending at rates between 5 – 10 cm/yr (e.g. Simons et al., 2007). Consequently, the area is known to be vulnerable to natural hazards such as large-magnitude earthquakes that can lead to tsunamis (e.g. Sumatra earthquake in 2004,

[1] Li et al. (2021) estimate the annual amount of lithosphere subducted along the Sumatra-Java and Philippine subduction zones – nearly 8,000 kilometers of subducting plate boundaries – to be ~30 km^3.

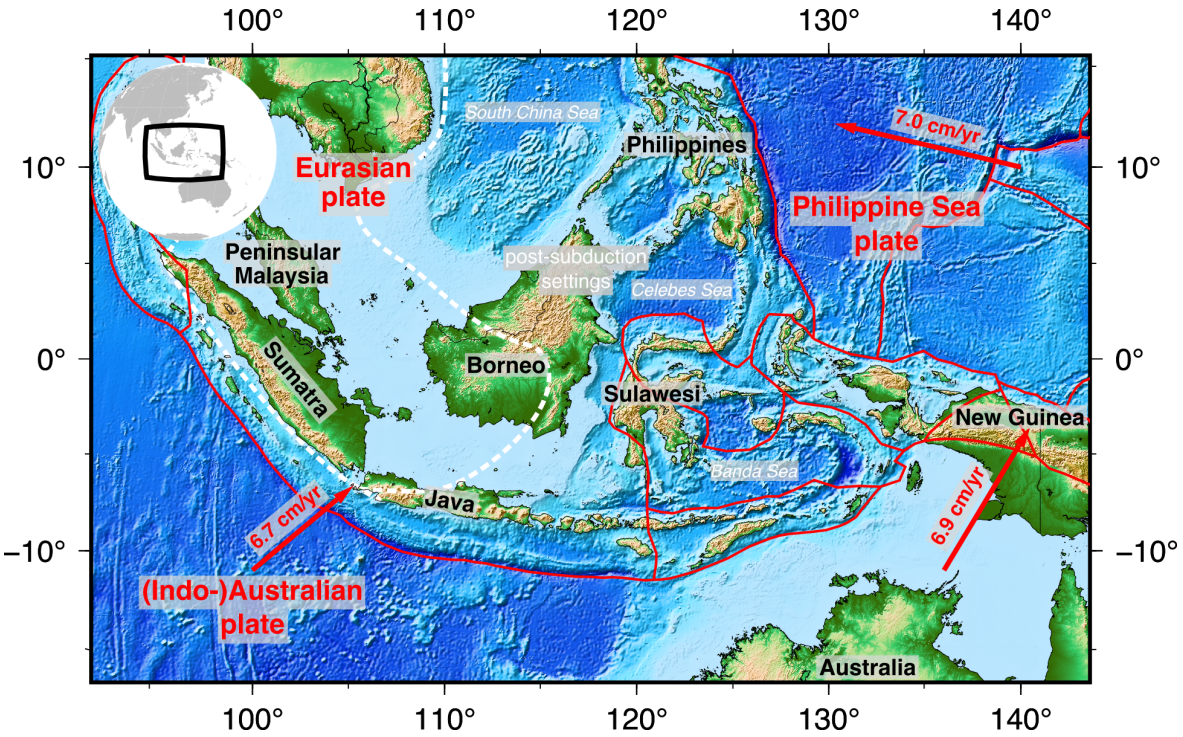


Figure 2.1: Map of the study area, showing the interaction of the three primary tectonic plates in Southeast Asia. The white dotted line indicates the outline of the Sundaland block (Hall, 2014). Plate tectonic boundaries are retrieved from Bird (2003). Plate motions are taken from *ITRF2014* (Altamimi et al., 2016) and are relative to the Eurasian plate. Topographic variations are taken from *ETOPO1* (Amante and Eakins, 2009).

Marano et al., 2010) and volcanic eruptions (e.g. Anak Krakatoa eruption in 2018, Petley, 2019).

The study region in the first part of this dissertation stretches from Sumatra (Indonesia) in the west to New Guinea in the east, and from the Philippines in the north to northern Australia in the south, thus encompassing an area of approximately 6,000 km in the east-west and 3,700 km in the north-south directions. In this chapter, I will elaborate on some of the key features of this region that are of particular relevance to the subsequent tomography study, including a brief summary of the tectonic history. However, it should be noted that Southeast Asia's tectonics are not well understood and many features remain debated.

2.1 Late Jurassic-Cenozoic tectonic reconstruction

The western part of Southeast Asia was largely formed during the closure of the Tethyan oceans in the Jurassic (∼163 – 145 Myr) and Cretaceous (∼145 – 66 Myr), with material sourced from east Asia and continental blocks that rifted from Gondwana (Hall, 2002), and by interaction with the Pacific plate. Figure 2.2 and Figure 2.3 present plate tectonic reconstructions at 90, 45, 25 and 5 Myr as inferred by Hall (2012)[2]. As India moved northward, the Ceno-Tethys ocean descended beneath Indonesia until ∼90 Myr ago, with subduction in this region not resuming until ∼45 Myr. At this point, Australia begins to move northward and the eastern part of Southeast Asia starts to form, giving rise to multiple new subduction zones during the Neogene (∼23 – 2 Myr). Thus, subduction zones in Southeast Asia were activated in different geological times, and features in the west are generally older than in the east.

2.2 Sundaland shelf

Southeast Asia is largely comprised of a shallow continental shelf, which is surrounded by several active tectonic plate boundaries. This continental promontory of the Eurasian Plate includes Borneo, Peninsular Malaysia, Sumatra, Java and parts of the South China Sea (see Figure 2.1) and is often referred to as the Sundaland block. It was assembled during the Triassic Indosinian orogeny approximately 250 Myr ago, when fragments rifted from Gondwana. The block was essentially in its present form, and in a similar position with respect to Asia by the Cretaceous (∼145 – 66 Myr, e.g. Metcalfe, 1990), but evidence for a complex pattern of subsidence, extension and elevation indicate that the region has been far from stable during most of the Cenozoic (∼66 Myr – today, e.g. Hall and Morley, 2004; Yang et al., 2016).

Today, the Sundaland shelf includes a large number of thick Cenozoic sedimentary basins and shallow seas (Hall and Morley, 2004), and experiences low levels of seismicity within its interior. GPS measurements indicate that the Sundaland block moves independently from Eurasia towards the east while rotating clockwise with an average velocity of several mm/yr (Simons et al., 2007).

[2]Hall (2012) presents 5 Myr intervals from 160 Myr – today.

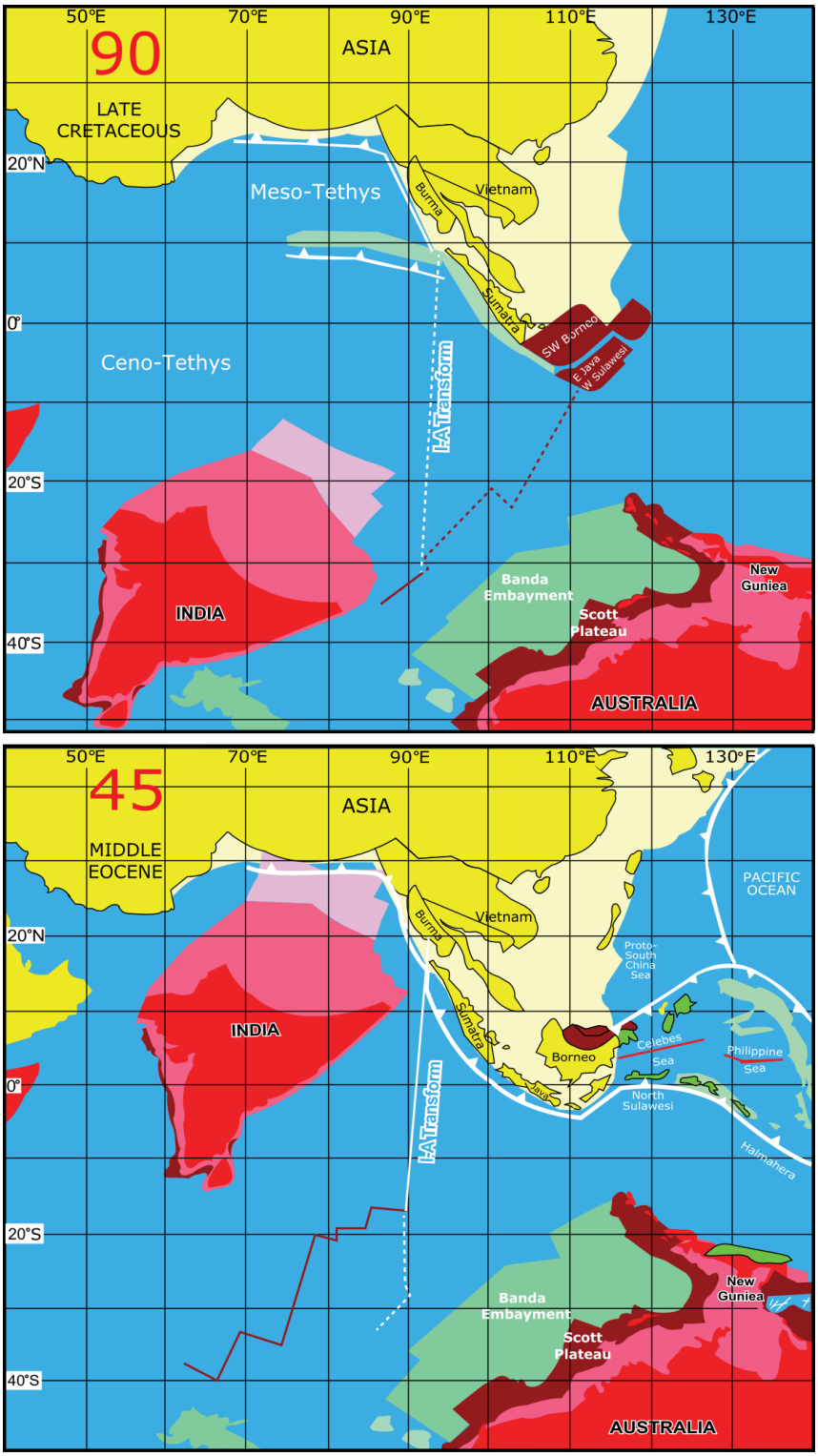


Figure 2.2: Simplified tectonic reconstruction at (*top*) 90 Myr and (*bottom*) 45 Myr, modified from Hall (2012). Areas filled with green are mainly arc, ophiolitic, and accreted material formed at plate margins. Eurasian crust is coloured in shades of yellow. Areas that were part of Gondwana in the Jurassic are coloured in shades of red. **90 Myr:** The India-Australia spreading centre becomes extinct (red line) and a new transform is initiated to form the new India–Australia plate boundary (*I-A Transform*) as India continues to move north. Subduction beneath Sundaland terminates and does not resume until 45 Myr. **45 Myr:** Australia and the Philippines move rapidly northward. The current subduction zone beneath Java begins and eastern Indonesia starts to form. The Celebes Sea and Philippine Sea open in a backarc setting.

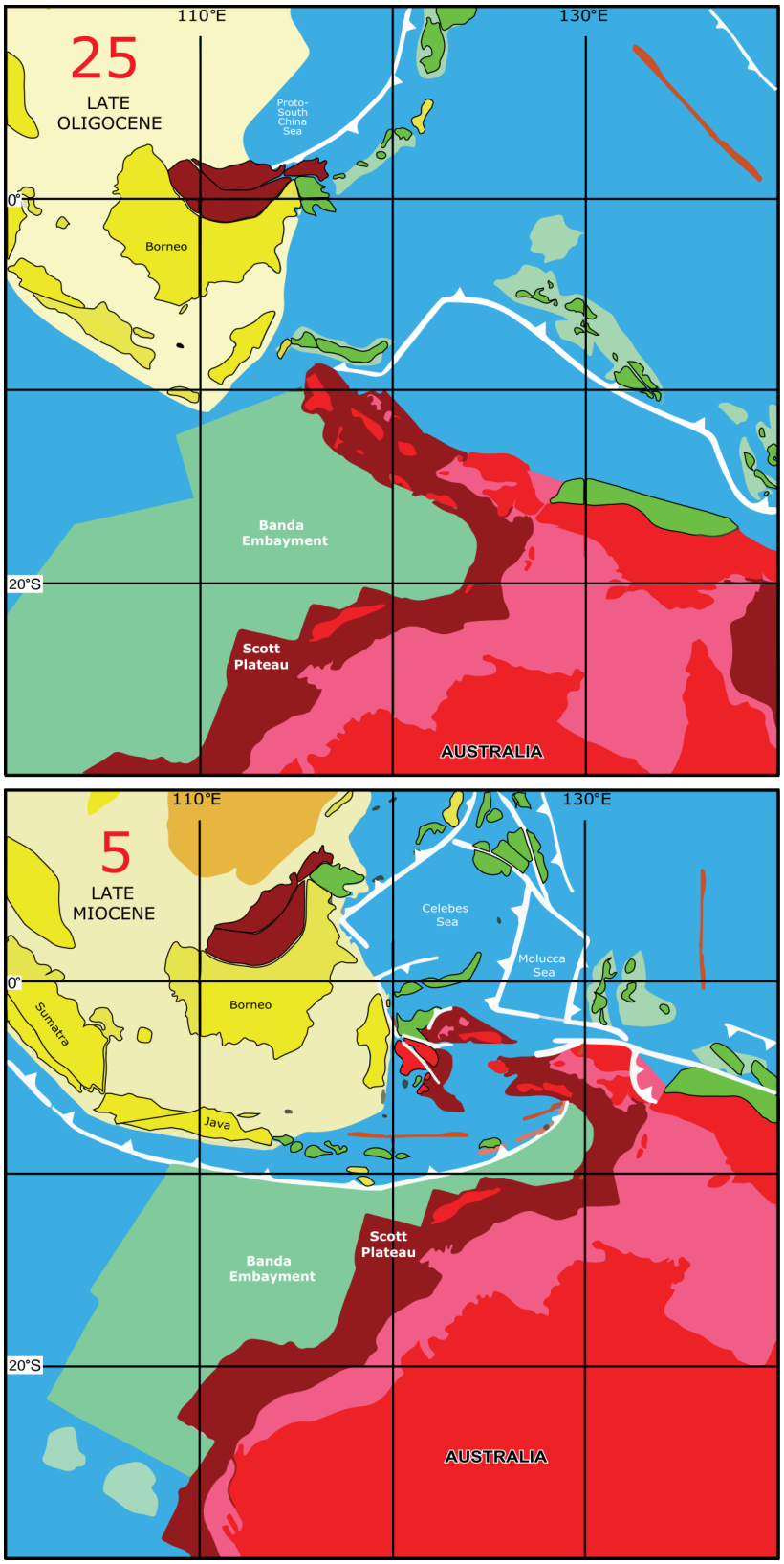


Figure 2.3: Simplified tectonic reconstruction at (*top*) 25 Myr and (*bottom*) 5 Myr, modified from Hall (2012). Areas filled with green are mainly arc, ophiolitic, and accreted material formed at plate margins. Eurasian crust is coloured in shades of yellow. Areas that were part of Gondwana in the Jurassic are coloured in shades of red. **25 Myr:** Sulawesi begins to form from several fragments. Southeast subduction of the proto-South China Sea beneath northern Borneo occurs. The Australia-Banda Sea collision is initiated. Continuing northward motion of the Philippines. **5 Myr:** Southeast subduction around northern Borneo has ceased but northwest subduction of the Celebes Sea occurs and is about to terminate. Subduction around northern Sulawesi has been initiated. The divergent double subduction around the Molucca Sea is apparent and the associated ocean basin becomes narrower over time.

It has long been known that the Sundaland block is a region characterised by low seismic velocities at relatively shallow depths (e.g. van der Hilst et al., 1997; Lebedev and Nolet, 2003; Bozdağ et al., 2016; Roberts et al., 2018). This suggests a thin, warm and weak lithosphere, an inference that is also supported by high heat flow values (e.g. Artemieva and Mooney, 2001) and gravity observations (e.g. Higgs et al., 1999), which may be the result of long-term subduction beneath the Sundaland block (e.g. Hall and Morley, 2004), producing a warmer and more hydrated mantle.

2.2.1 Borneo (Brunei, Indonesia and Malaysia)

Borneo lies in the eastern part of the Sundaland block and is the largest island within Southeast Asia. The island is shared by Brunei, Indonesia and Malaysia and is not particularly seismically active (see Figure 1.2). While the Sundaland shelf is not significantly elevated with depths considerably less than 200 m, the Bornean topography – with the highest peak being Mt Kinabalu at 4,095 m (see Figure 2.4) – is believed to have grown rapidly during Neogene times (∼23 – 2 Myr, e.g. Roberts et al., 2018). This uplift may be associated with Borneo's unique post-subduction setting (e.g. Hall, 2012; Gilligan et al., 2018; Pilia et al., 2019; Greenfield et al., 2022), which motivated our working group's deployment of a dense, temporary seismic network across northern Borneo between March 2018 and January 2020 (*northern Borneo Orogeny Seismic Survey*, Rawlinson, 2018). In the Miocene (∼23 – 5 Myr), two sequential but opposed subduction systems were in operation in the northern part of Borneo. They featured southeast subduction of the proto-South China Sea, and northwest subduction of the Celebes Sea, which terminated at 23 and 9 Myr, respectively (see Figure 2.3, e.g. Hall, 2013; Pilia et al., 2019).

Figure 2.4: View from the top of Mt Kinabalu (Borneo) at ∼4,100 m – one of the tallest mountains in Southeast Asia. Photo taken by author in January 2020.

So far, Borneo has received little attention from seismic tomography. A high-velocity anomaly beneath northern Borneo at depths of ∼100 – 300 km was previously imaged in ray tomographic studies (Amaru, 2007; Zenonos et al., 2019), but was regarded as suspicious owing to the poor data coverage (e.g. Hall and Spakman, 2015). While Greenfield et al. (2022) image the upper ∼150 km beneath northern Borneo using two-plane wave tomography and conclude that there is no evidence for any slabs, Pilia et al. (submitted) image an upper mantle high-velocity perturbation using P- and S-wave traveltime tomography and suggest it is associated with remnants of the proto-South China Sea slab.

2.3 Sulawesi (Indonesia)

Sulawesi is a k-shaped island to the east of the Sundaland block; its unique geometry arises from its assemblage from several crustal fragments (Katili, 1978; Hall and Wilson, 2000). It is thought to have formed during the Miocene (∼23 – 5 Myr, see Figure 2.3) from crust sourced from both Sundaland and Australia, juxtaposed by convergence (e.g. Hall, 2011).

This part of the study region is known to be particularly complex since it is located at the confluence of several minor tectonic plates and as a result, comprises multiple fault zones with a diverse range of lengths, ages and geometries. For example, the Palu-Koro transform

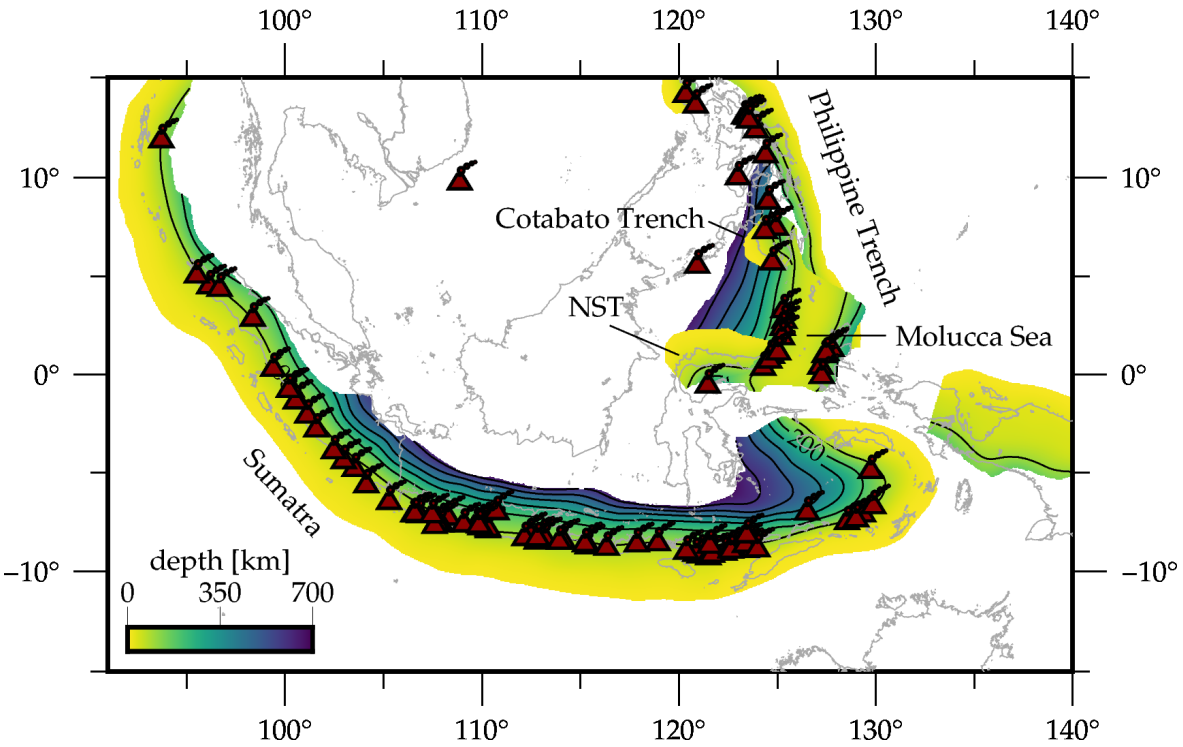


Figure 2.5: Depths to the upper surface of subducting slabs beneath Southeast Asia taken from the subduction zone geometry model *Slab2* (Hayes et al., 2018), where NST indicates subduction along the North Sulawesi Trench. The black line contours are drawn every 100 km. For Southeast Asia, *Slab2* is derived from the distribution of earthquake locations. In addition, relocated events and active source seismic data are used around Sumatra and Java. Dark red symbols indicate volcanoes that erupted during the Holocene (the past ∼12,000 years) as taken from the *Global Volcanism Program* (Venzke, 2013).

fault cuts through the centre of Sulawesi and caused the 2018 strike-slip Palu earthquake, which unexpectedly triggered a destructive tsunami (e.g. Socquet et al., 2019). Linked to the northern end to this strike-slip fault, the North Sulawesi Trench stretches along the northern end of the island, where subduction of the Celebes Sea was initiated at ∼9 Myr (see Figure 2.3, Hall, 2011). The slab exhibits earthquakes to a depth of 250 km (Hayes et al., 2018), which, given their dip, suggests at least 300 km of subduction. Song et al. (2022) present a numerical modelling study, which reveals an east-west oriented shallow–deep–shallow subduction style as a result of rotation along the northern Sulawesi arm. However, it should be noted that their study is built-upon *Slab2* (Hayes et al., 2018), which only uses the earthquake distribution to derive slab depths in this region. However, a shorter slab is consistent with an absence of volcanoes along the trench (see Figure 2.5).

2.4 Subduction along the Indonesian volcanic arc

The descent of the oceanic (Indo-)Australian plate beneath the Eurasian plate forms an active subduction system beneath the Indonesian volcanic arc, which transitions to an arc-continental collision along the curved Banda Arc (see Figure 2.1). The associated subducting slabs are usually the most prominent feature in tomographic studies of this region (e.g. Amaru, 2007; Obayashi et al., 2013; Zenonos et al., 2019).

2.4.1 Sumatra and Java region

In the west, subduction of the oceanic lithosphere of the Australian plate beneath the elongated Indonesian islands of Sumatra and Java was initiated during the Cretaceous (∼145 – 66 Myr) and – however discontinuous – controlled the Cenozoic (∼66 Myr – today) tectonic history of the region (see Figure 2.2 and Figure 2.3, Hall, 2012). Today, high plate velocities of 6 – 7 cm/yr and large-magnitude earthquakes are observed (e.g. Kopp et al., 2006; Spakman and Hall, 2010), and the region is characterised by a chain of active volcanoes on these islands (see Figure 2.5, McCaffrey, 2009). As a result, this region represents a significant natural hazard, which is one of the main reasons it is the focus of ongoing research (e.g. Métrich et al., 2017; Walter et al., 2019; Wang and He, 2020).

Several studies have imaged descending slabs along the Sunda Arc and Java Arc, but discrepancies exist among these studies regarding the geometry and depth extent of the subducted slab segments; several previous studies suggest that the subducted slab only extends in depth to the mantle transition zone (e.g. Gudmundsson and Sambridge, 1998; Amaru, 2007; Li et al., 2018), while others advocate for its penetration into the lower mantle (e.g. Huang et al., 2015). Cross-sections across the Java Arc by Fukao and Obayashi (2013) reveal that the slab of the (Indo-)Australian plate is trapped in the neighbourhood of the 660 km discontinuity.

2.4.2 Banda Sea region

The eastern region of the study area comprises fragments rifted from the Australian and South China margins during the Cenozoic (∼66 Myr – today), which – together with subduction complexes, island arcs and marginal seas – form a complex heterogeneous basement now largely covered by Cenozoic sediments (see Figure 2.2 and Figure 2.3, Metcalfe, 1990). This part of the study area is characterised by a complex configuration of minor tectonic plates and features earthquakes down to 700 km depth (see Figure 1.2), while fewer volcanoes are present (see Figure 2.5). This complexity is driven by the Southeast Asia-Australia collision zone that was formed as a consequence of the northward movement of the Australian plate (e.g. Hall, 2011; Fichtner et al., 2010), where Sunda Arc subduction beneath Sumatra transitions to an arc-continent collision due to a change from oceanic to continental lithosphere. Based on a regional finite frequency teleseismic P-wave tomographic model, Harris et al. (2020) conclude that there is no evidence for slab tearing in this transition region.

The northward motion of the North Australian continental lithosphere results in the spectacular 180° curvature of the > 1,000 km long Banda Arc (e.g. Audley-Charles, 1968; Carter et al., 1976; Harris, 2011), where subduction began ∼15 Myr ago (Spakman and Hall, 2010). 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. Based on combining mantle tomography with a kinematic plate model and active surface deformation, Spakman and Hall (2010) favour the hypothesis of a single deformed slab caused by subduction rollback. Furthermore, a geodynamic modelling study by Moresi et al. (2014) demonstrates how such a single bent and deformed slab could have developed from northward movement of the Australian plate. They find evidence for this process in eastern Australia but suggest that their model can explain retreat with further extension around the arcs of Java and Banda.

Tomographic images of the subducted Banda slab suggest that it is entirely confined to the upper mantle (Widiyantoro and Hilst, 1996; Richards et al., 2007; Amaru, 2007; Spakman and Hall, 2010). This may be a result of the much younger subduction in the Banda Sea region compared to the regions to the west, which only began in the Neogene (∼23 – 2 Myr, e.g. Spakman and Hall, 2010).

2.5 Subduction south of the Philippines

The tectonic history along the southwestern margin of the Philippine Sea plate is even more difficult to constrain as there is little evidence since critical areas are off-shore (Hall, 2012). It is generally assumed that the Philippine Sea plate travelled northward from a near-equatorial position since ∼40 Myr ago (see Figure 2.2). The majority of this northward motion occurred before 25 Myr, with a smaller northward motion after 15 Myr (see Figure 2.3, Yamazaki et al., 2010; Wu et al., 2016). Today, the plate is believed to move rapidly to the west (> 7 cm/yr, e.g. Altamimi et al., 2016) but it should be noted that plate motions in this region are difficult to constrain because the plate has been isolated from seafloor spreading and hot spot-based global reconstruction circuits (Seton et al., 2012).

2.5.1 Philippine Trench

The westward subduction of the Philippine Sea plate generates events down to ∼200 km depth along the north-south trending Philippine Trench (see Figure 2.5). Based on geochemical measurements, the age of the subducting slab was inferred as > 45 Myr (Ishizuka et al., 2013; Wu et al., 2016), while Fan and Zhao (2018) suggest subduction was initiated ∼20 – 25 Myr ago based on the depth at which slabs are resolved in their P-wave tomography study.

Wu et al. (2016) suggest that the slab extends to depths of ∼230 – 400 km beneath the Philippine Trench as inferred from seismicity and previous seismic tomography models. Hall and Spakman (2015), however, state that there is little evidence of a deep slab in tomography and propose a shorter slab length of slightly over ∼100 km, which is consistent with an absence of volcanoes along the trench (see Figure 2.5). Furthermore, convergent double subduction zone(s) formed by eastward dipping slab(s) from the Celebes Sea and the westward subduction along the Philippine Trench have previously been suggested (e.g. Rangin, 1991; Lagmay et al., 2009), but this complexity is not evident in the slab model *Slab2* (Hayes et al., 2018) nor the plate tectonic boundary model by Bird (2003).

2.5.2 Molucca Sea

Further south, a divergent double subduction zone with an inverted U-shape and slabs subducting to the east (Halmahera slab) and west (Sangihe slab) in the Molucca Sea (see Figure 2.5) has long been a point of discussion. Previous models suggest that the western limb penetrates into the lower mantle, while the eastern limb is imaged down to ∼100 – 150 km depth in P-wave tomography studies (Amaru, 2007; Obayashi et al., 2013; Wu et al., 2016; Fan and Zhao, 2018). This is consistent with the seismicity distribution but its extension to the north remains unclear. Hall (2012) suggests that the Molucca Sea was formed ∼20 Myr ago and became narrower until the present-day arc-arc collision (e.g. Soesoo et al., 1997) formed around ∼5 Myr ago (see Figure 2.3).

In summary, Southeast Asia provides a unique setting to investigate a variety of primary tectonic processes, including but not limited to, subduction initiation, ongoing subduction, subduction termination and both arc-continent and continent-continent collision (e.g. Hall, 2013). While several tomographic studies identify a number of subduction slabs within the region, discrepancies exist regarding their geometry and depth extent. In particular, previous studies lack constraints in key regions such as Borneo, Sulawesi, the transition from divergent to convergent double subduction south of the Philippines and the Banda Arc curvature. At crustal depths, Southeast Asia is also not very well understood; for example, Laske et al. (2013) show that model predictions from their global crustal model *CRUST1.0* (see Figure 5.16) compared against 25 s Rayleigh wave group velocity maps exhibit some of the largest misfits on the globe. The lack of constraints can largely be attributed to the complexity of the region and the limited data availability.

A note on plate tectonic boundaries: For Southeast Asia, there is no definite answer as to the number of tectonic plates and where their boundaries lie (Hall and Spakman, 2015). In particular, the existence of a single (Indo-)Australian plate and the number and location of the minor tectonic plate boundaries in the east have been a point of discussion (e.g. Hall, 1998). Throughout this study, I plot the boundaries as interpreted by Bird (2003). Bird (2003) has recognised four large plates (Australia, Eurasia, Pacific and Philippine Sea) and several small plates for this region.

3 Method: Adjoint waveform tomography

The following chapter provides an overview of adjoint waveform tomography (often referred to as *full-waveform inversion* or *FWI*) as applied to earthquake data. First, I will briefly review important milestones in seismic imaging, with a particular focus on the transition from ray to waveform tomographic methods in the subsequent section. I will then present the theoretical background of adjoint waveform tomography in more detail.

3.1 Introduction

3.1.1 A brief history of seismic imaging using earthquake sources

Seismic imaging aims to illuminate the Earth's structure by linking ground motion observations at the Earth's surface to variations in material properties of the subsurface. Although its origins are somewhat nebulous, one could argue that the science of earthquake seismology was born in the 18th century when John Michell suggested – following the 1755 Lisbon earthquake – that earthquakes generate waves, which travel through the Earth (Michell, 1761). A century later, the first modern instrument for registering ground motion was proposed by Robert Mallet who is now known as the founder of observational seismology (Mallet, 1862). In 1889, Ernst von Rebeur-Paschwitz (accidentally) recorded an earthquake in Potsdam, Germany, that occurred near Japan (Rebeur-Paschwitz, 1889), which heralded the birth of teleseismic methods. Ten years later, Richard Dixon Oldham identified separate seismic wave arrivals, namely the fast compressional P-wave (*primary*), the slower shear S-wave (*secondary*) and surface waves (Oldham, 1899).

In the early 20th century, seismological discoveries began to accelerate with the first pioneering studies exploring the radial structure of the Earth. In 1906, Oldham published the first evidence that the Earth must have a large core, which he inferred from seismic wave shadow zones (Oldham, 1906). In 1909, Andrija Mohorovičić demonstrated the existence of the crust-mantle discontinuity (*Moho*) based on a distinct set of P- and S-wave arrivals (known as *head waves*) that travel along this boundary (Mohorovicic, 1909). Five years later, Beno Gutenberg published an accurate determination of the depth of the core-mantle-boundary at 2,900 km below the surface, as inferred from an abrupt decrease in seismic velocities (Gutenberg, 1914). However, it was Harold Jeffreys who provided the first unequivocal evidence that the outer core is liquid (Jeffreys, 1926) and Inge Lehmann in 1936 who proved

that the Earth has a solid inner core surrounded by a liquid outer core (Lehmann, 1936). In 1940, the first 1-D Earth model built upon traveltime tables was released by Harold Jeffreys and Keith Bullen (Jeffreys and Bullen, 1940). While multiple reference models have been produced[1], three have been most commonly used to date (see Figure 3.1):

Preliminary Reference Earth Model (PREM): The first 1-D Earth model that included a realistic representation of attenuation and radial anisotropy was published by Adam Dziewonski and Don Anderson (Dziewoński and Anderson, 1981). It was obtained from body waves, normal modes, surface waves, surface wave dispersion and astronomical observations, and is recommended for long-period studies. Furthermore, it incorporates radial anisotropy down to the 220 km Lehmann discontinuity, which is mainly constrained by joint analysis of Rayleigh and Love wave data.

IASP91: An isotropic model by Kennett and Engdahl (1991) inferred from body wave data, where P-wave structure is better resolved than S-wave structure. It is particularly suitable for body wave studies and earthquake relocation but it is not valid at all epicentral distances.

ak135: The isotropic *ak135* velocity model (Kennett et al., 1995) has been augmented with a density and Q model by combining the study of traveltimes with those of free oscillations (*ak135f*, Montagner and Kennett, 1996). It is particularly suitable for body wave studies, including those that involve the core since it is well resolved by *PKP* differential traveltimes.

The earliest applications of 3-D seismic tomography were conducted in separate studies by Keiiti Aki based on relative traveltime residuals and Adam Dziewonski using P-wave delay times in the late 1970s (Aki et al., 1976; Dziewoński et al., 1977). Since then, a broad spectrum of local, regional and global tomographic studies in 2-D and 3-D have been conducted that provide new insight into various phenomena, including mantle convection (e.g. van der Hilst et al., 1997), the tectonic evolution of continents (e.g. Nolet and Zielhuis, 1994), hotspots (e.g. Montelli et al., 2004) and subduction zones (e.g. Dziewoński, 1984; Fukao and Obayashi, 2013).

To date, a wide range of tomographic methods have been developed including teleseismic tomography, local earthquake tomography, ambient noise tomography as well as regional and global body and surface wave tomography. Traditionally, they have exploited ray theory for computing the observables, but in recent times, finite frequency and waveform tomography methods have become more established. So far, nearly all studies have focused on understanding Earth structure and processes. However, there are also studies that have investigated the Moon using moonquakes (e.g. Zhao et al., 2008; Steinberger et al., 2015), and the 2018 *InSight* mission deployed a seismometer on Mars (e.g. Mimoun et al., 2017), which paves the way to eventually probing the interior of other planets using seismic tomography. For a more detailed overview of seismic tomography, and in particular different tomographic methods, I refer the interested reader to Rawlinson et al. (2010).

[1]For a recent and more detailed overview, I refer the interested reader to Kennett (2020). This study also presents a new model, *ek137*, which includes constraints on core structure.

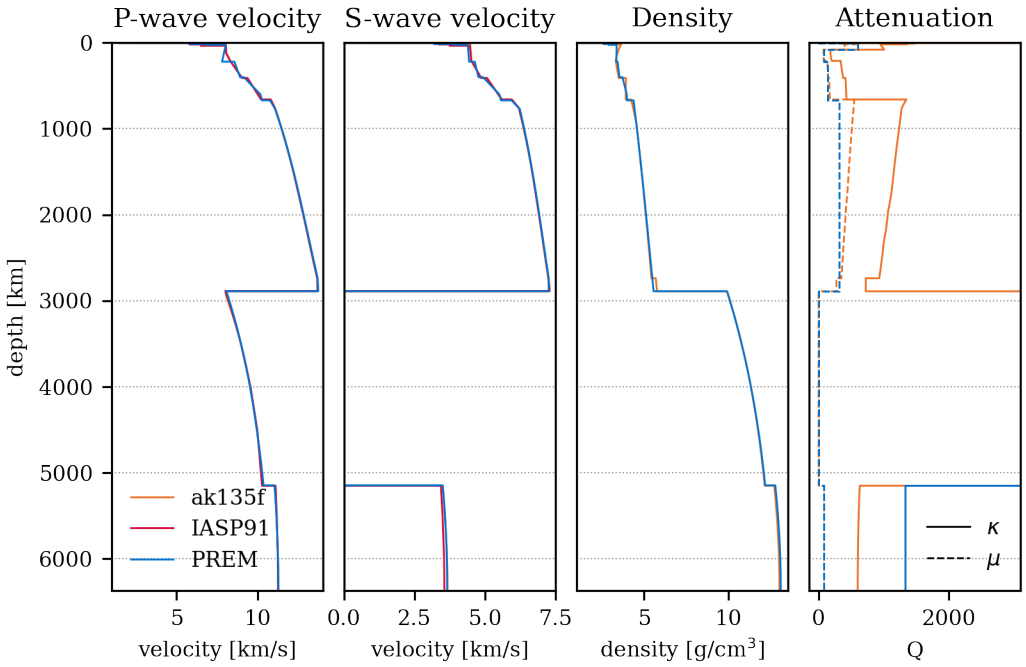


Figure 3.1: P-wave and S-wave velocity, density, and bulk and shear attenuation (Q_κ and Q_μ) for the three most commonly used reference models (*ak135*, *IASP91* and *PREM*). The attenuation scale is truncated at 3,000 since bulk attenuation is extremely low in the outer core through which S-waves do not propagate.

3.1.2 From ray tomography to adjoint waveform tomography

Most ground-breaking discoveries mentioned in Section 3.1.1 are based on simplified theories of seismic wave propagation such as geometric ray theory. Ray tomography has traditionally been the standard approach since it is mathematically simple and computationally efficient. The method has played a crucial role in determining earthquake hypocentre locations and Earth structure across a range of scales since the mid 1970s (e.g. Aki et al., 1976; Dziewoński et al., 1977).

Ray tomographic studies assume that seismic waves travel along infinitely narrow paths that connect source and receiver. Strictly speaking, this scenario is only valid if seismic waves propagate at infinite frequency, but in reality, they propagate at finite frequencies[2] and sample extensive regions outside the geometric ray path (e.g. Marquering et al., 1999; Dahlen et al., 2000). The main issue with this simplification is its inability to account for certain wave-like behaviour (e.g. diffraction, scattering) and the requirement of smooth media (e.g. Nolet, 2008). As such, it begins to break down when the seismic wavelength approaches or exceeds the scale length of heterogeneity (e.g. Červený, 2001; Panning et al., 2009). Furthermore, ray tomographic methods only use a limited portion of the information contained in a seismogram such as phase arrival times, which are not a directly recorded parameter. In addition, many ray tomographic studies employ crustal corrections to account for the behaviour of seismic waves in the upper few kilometres (e.g. Bozdağ and Trampert, 2008).

[2]For example, the frequency content of a recorded teleseismic P-wave generally does not exceed 1 Hz since higher frequencies get suppressed by attenuation.

Overcoming the limitations of ray tomography has been the objective of many studies over the past few decades. The mathematical background has been known since the 1980s (e.g. Tarantola, 1984) and generally involves taking the often complex, volumetric sensitivity of seismic waves into account. This typically requires numerically solving the partial differential equation that governs 3-D seismic wave propagation but this has not been computationally feasible until recently. A common approach to account for the volumetric sensitivity of seismic waves uses scattering theory based on the single-scattering first Born approximation (e.g. Gilbert and Dziewonski, 1975). Early waveform inversion studies matched long-period waveforms (Woodhouse and Dziewonski, 1984; Li and Tanimoto, 1993) and subsequent studies developed the finite-frequency "banana-doughnut" theory (e.g. Yomogida, 1992; Marquering et al., 1999; Dahlen et al., 2000). These techniques provide an extension of geometrical ray theory and are computationally much more efficient than FWI. While they take effects such as wavefront healing and single scattering into account (e.g. Hung et al., 2001; Montelli et al., 2004), they are not valid for strongly heterogeneous media and the sensitivity volumes of seismic waves are approximated. The potential benefits of finite frequency methods are discussed in more detail in Sieminski et al. (2004).

Rapid increases in computing power and storage, such as cheap random access memory (RAM) and parallel computing, have enabled us to accurately solve the 3-D seismic wave equation and compute exact sensitivity kernels using adjoint techniques in practical applications over the last decade or so. This allows the full complexity of seismic wave propagation to be simulated and thus can account for effects such as wavefront healing, interference, scattering and (de)focusing, which are not accounted for by ray theory (e.g. Rickers et al., 2012). The method has its origins in seismic exploration (e.g. Gauthier et al., 1986; Pratt and 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 fields such as medicine (e.g. Schreiman et al., 1984; Guasch et al., 2020; Bachmann and Tromp, 2020) and engineering (e.g. Jalinoos et al., 2017). First applications of FWI in earthquake seismology include imaging of 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 gradually become more accepted as a viable approach for producing high-resolution images of the Earth's interior using passive source seismic data. One example is an image of the African superplume as part of a global model that reveals strong shear-wave perturbations (Bozdağ et al., 2016; Lei et al., 2020), with implications for the associated thermochemical interpretation. For a more detailed technical review, including full-waveform inversion in the context of seismic exploration, I refer the interested reader to Virieux and Operto (2009) and Tromp (2020).

It should be noted that ray tomography is still a popular imaging method. In many applications, the vast amounts of data that can be included outperform the advantages of full-waveform inversion. However, adjoint waveform tomography promises higher resolution of multi-parameter (v_P, v_S, ...) models and can account for full wave effects that can hamper other seismological methods. Hence, it is particularly suitable for strongly heterogeneous regions, thereby providing new opportunities for geophysical and geochemical interpretation.

3.1.3 The workflow: Adjoint waveform tomography in a nutshell

Adjoint waveform tomography is one of the most challenging ways of obtaining information on Earth structure due to the complex, tangled workflow and non-linearity of the inverse problem, which is computationally expensive to solve. The first step is to obtain accurate synthetic seismograms from an initial Earth model for a set of specified sources and receivers by solving the 3-D seismic wave equation numerically. In the next step, 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.

In the following two sections, I will explain the theoretical background of adjoint waveform tomography by stepping through the workflow schematic shown below (see Figure 3.2). I will address the forward and inverse problem in the context of adjoint waveform tomography, but it should be noted that any problem, where a system or model is reconstructed from observations, can be formulated as an inverse problem.

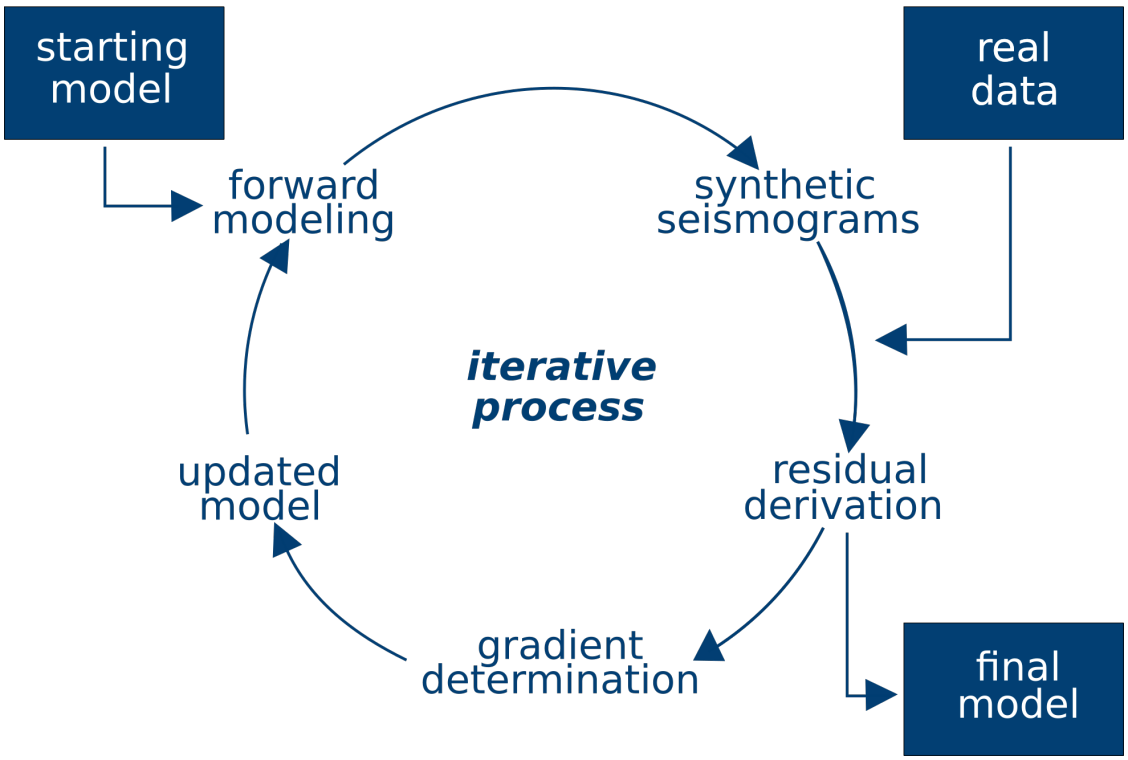


Figure 3.2: Full-waveform inversion workflow. This process is iterated until the waveform match is deemed sufficient according to some criteria. Most often, a multi-scale approach (Bunks et al., 1995) is applied in which longest periods are inverted for first and higher frequency content is successively added.

3.2 The forward problem

The ultimate aim in geophysical inversion is to determine an Earth model m that fits our observations d. This can be described by the following equation:

$$d = G(m), \tag{3.1}$$

where G represents the physical theory linking model space and data space (e.g. Maxwell's equations, wave equation, ...). This so-called *forward problem* can be used to obtain synthetic data from an initial model ($d_{syn} = G(m_0)$). In waveform tomography, solving the right-hand side of this equation requires the computation of the 3-D wavefield – that is the time- and space-dependent solution of the wave equation – through a given model.

3.2.1 Model domain discretisation in space and time

To use Equation 3.1, we first need to define a region of interest through which the seismic waves propagate (domain) that contains material properties (model) as defined by a discrete set of parameters. The initial values given to these parameters are often taken from one of a variety of publicly available seismic reference models (e.g. IRIS, 2011; Hutko et al., 2017). I will discuss the choice used in this study in more detail in the next chapter (see Section 4.1.1).

One typically divides the model region into a finite number of mesh elements, with the frequency content of the waveform primarily governing the mesh spacing in waveform tomography[3]. Thus, the spatial extent of a mesh element is largely based on the shortest wavelength $\lambda_{\min}$ of the signal, which is inversely related to the highest frequency f_{max} (or shortest period) and can be estimated via

$$\lambda_{\min} = \frac{v}{f_{\max}} \tag{3.2}$$

for a given velocity v. As such, the size of an element is proportional to the velocity considered. Thus, elements are usually smaller where velocities are lower (e.g. the Earth's crust) and larger at greater depths. As f_{max} increases, the elements become smaller, resulting in a denser sampling of the wavefield, which is accompanied by an increase in computational cost. In the three-dimensional case, this increase is proportional to the fourth power of the highest frequency of the signal (three orders in space plus one in time). Insufficient sampling will lead to instabilities in the simulation, often manifest as unphysical oscillations. This numerical dispersion introduces errors in our synthetic seismograms. To enhance accuracy, it is common practice to also define the number of elements used per $\lambda_{\min}$. In earthquake seismology, most studies employ values ranging between 1.0 and 2.0, depending on data quality and the scale of the application.

[3] Although there have been recent developments towards wavefield adapted meshes, which are refined radially based on the azimuthal symmetry of wavefields, i.e. elements are larger where the wavefield varies less (van Driel et al., 2020; Thrastarson et al., 2020).

The ideal temporal discretisation maximises the possible temporal step dt for the wavefield sampling, which improves the computational efficiency. The Courant– Friedrichs–Lewy condition can be used to compute the limit, where simulations become unstable[4] (Courant et al., 1928). This stability criterion is based on the element size h and can be written as

$$dt \leq C_{\mathrm{CFL}} \frac{h}{v}, \tag{3.3}$$

where C_{CFL} is the method-dependent Courant number. Once a domain has been defined, along with optimum spatial and temporal discretisation, the associated seismic wavefield can be computed. This will be discussed in the following sections.

3.2.2 Elastic waves in the solid Earth

Seismic wave propagation is governed by the elastic wave equation (e.g. Aki and Richards, 2002), which is solved numerically to compute synthetic seismograms. Derived from Newton's laws of motion, the elastic wave equation (neglecting self-gravitational and rotational effects) can be expressed as

$$\rho(\mathbf{x}) \frac{\partial^2 u(\mathbf{x}, t)}{\partial t^2} - \nabla \cdot \sigma(\mathbf{x}, t) = f(\mathbf{x}, t), \tag{3.4}$$

where ρ is density, u stands for displacement (related to the ground motion, which is measured by a seismometer). σ is the stress tensor and f represents a source term. The x and t indicate space and time dependencies, respectively. Typically, the stress tensor σ is obtained from Hooke's Law, which is an empirical expression of the elastic stress-strain relation via the fourth order elasticity tensor c_{ijkl} under the assumption of a linear elastic rheology:

$$\sigma_{\mathrm{ij}} = c_{\mathrm{ijkl}} e_{\mathrm{kl}} \tag{3.5}$$

noting the implied summation of k and l. For the three-dimensional case, c_{ijkl} can have between two (isotropic medium) and 21 (fully anisotropic medium) independent components based on inherent symmetries of c, σ and e. In many cases, seismic waves have measurable sensitivity to only a few of them (Sieminski et al., 2009).

Regarding the source term on the right-hand side of Equation 3.4, seismic waves can be generated by a wide variety of source types (e.g. explosions, landslides), but we will focus on earthquakes in this thesis. Earthquakes emit elastic waves from a potentially complex rupture surface at a local scale, and modelling this rupture process requires considerable effort and resources (e.g. Simons et al., 2011). At larger scales, earthquakes are usually approximated by an instantaneous point source, which neglects the temporal and spatial extent of the rupture. This point source is described by a hypocentre, an origin time and a description of the spatial radiation of energy (moment tensor), which are usually taken from existing catalogues (e.g. *GCMT* catalogue by Ekström et al., 2012; *SCARDEC* by Vallée et al., 2011). Most waveform tomographic studies assume source information to be fixed prior

[4] If the time step is too large, the simulation will blow up.

information, but errors in this description can result in significant artefacts in the tomographic image (e.g. Valentine and Woodhouse, 2010). Blom et al. (2022) show that source parameter errors can result in significant wavefield perturbations and corresponding time shifts of seismic waveforms, and present a number of mitigation strategies to minimise the formation of source error artefacts in the tomographic image.

3.2.3 Acoustic waves in the fluid ocean

In the previous section, I discussed wave propagation in the solid Earth. However, depending on the domain, modelling acoustic waves in the ocean may be desirable. In this case, the ocean is replaced by acoustic elements, so that it can be represented as a fluid medium, thus requiring the solution of a coupled system of acoustic and elastic wave equations (e.g. Afanasiev et al., 2019). The acoustic wave equation can be expressed as

$$\frac{1}{\rho(\mathbf{x})c^2(\mathbf{x})}\frac{\partial^2 u(\mathbf{x},t)}{\partial t^2} - \nabla\left(\frac{1}{\rho(\mathbf{x})}\nabla u(\mathbf{x},t)\right) = f(\mathbf{x},t), \tag{3.6}$$

where c is the speed of sound in the ocean.

It is worth noting that it is common practice to use an acoustic approximation for elastic waves in exploration, which is computationally efficient (e.g. Pratt et al., 1998; Berkhout, 1984; Plessix et al., 2010). This means only P-waves are considered and since explosive sources are typically used, the dominant wave type will be compressional, with S-waves largely coming from mode conversions at interfaces. However, accounting for S- and surface waves is an elementary part of any seismogram analysis in earthquake seismology.

3.2.4 Modelling the seismic wavefield

A wealth of numerical techniques have been developed to solve the forward problem of seismic wave propagation (see Equation 3.4 and Equation 3.6) over the past few decades[5]. The aim is to apply an accurate, efficient, flexible and scalable seismic wave propagation solver to calculate the forward wavefield. In full-waveform inversion, spectral-element methods (a form of finite-element method) are currently considered to provide an optimal balance between simulation accuracy and efficiency in earthquake seismology, while finite-difference methods are popular in seismic exploration.

The finite-difference method (e.g. Virieux, 1984) used in seismic exploration approximates the spatial derivative by difference quotients and can result in an increase in efficiency compared to the finite-element method. However, a major drawback for seismological applications is its inability to adapt the numerical grid to complex structures such as topography. Furthermore, the implementation of the free-surface condition is more difficult.

The spectral-element method is preferred in earthquake seismology because of its ability to accommodate topography, bathymetry and fluid–solid boundaries, such as the ocean-crust boundary. The method originates in fluid mechanics (Patera, 1984) and solves the weak form

[5]The Coursera course *Computers, Waves, Simulations: A Practical Introduction to Numerical Methods using Python* by Igel (n.d.) offers a comprehensive introduction to this topic.

of Equation 3.4 and Equation 3.6. Thus, the free-surface boundary condition is implicitly taken into account, thereby providing stable and accurate simulations even in regions with complex topography (e.g. Komatitsch and Vilotte, 1998; Komatitsch and Tromp, 1999). The wavefield is represented by Lagrange polynomials, which are expressed using the Gauss-Lobatto-Legendre (GLL) collocation points defined within each element. This results in a diagonal mass matrix, which means no complex matrix inversion needs to be carried out.

The first applications of the spectral-element method to the elastic wave equation can be found in Priolo et al. (1994) and Komatitsch and Vilotte (1998). In recent years, several 3-D spectral-element wave propagation solvers have been introduced, e.g. *SES3D* (Fichtner and Igel, 2008; Gokhberg et al., 2016), *SPECFEM* (Peter et al., 2011) and *Salvus* (Afanasiev et al., 2019).

Finally, synthetic seismograms – the solution of the forward wavefield at specified locations – can be obtained from complete wavefield simulations and their deviation from observed waveforms can be used to update the initial model. In the next section, I will elaborate on the inversion component of adjoint waveform tomography, which seeks a better model given a set of synthetic and observed waveforms. This includes the definition of a waveform misfit and how to update the current model in order to minimise this quantity.

3.3 The inverse problem

The inverse problem, which falls under the more general category of data inference, involves the recovery of a model or physical state from a set of observations (see Figure 3.3), and is formulated and solved across a multitude of disciplines, e.g. diagnosing tumours from X-ray images in medicine (e.g. Hounsfield, 1980) or determining the point of origin of an impact from bloodstain patterns in forensics (e.g. Cecchetto and Heidrich, 2011). In geophysics, the inverse problem provides an avenue for reconstructing the state of the Earth's[6] interior from data measured at the surface. The problem geophysicists face is that volumetric subsurface properties are reconstructed from point-localised, and often uneven, surface measurements. Furthermore, the data only indirectly constrain quantities of interest.

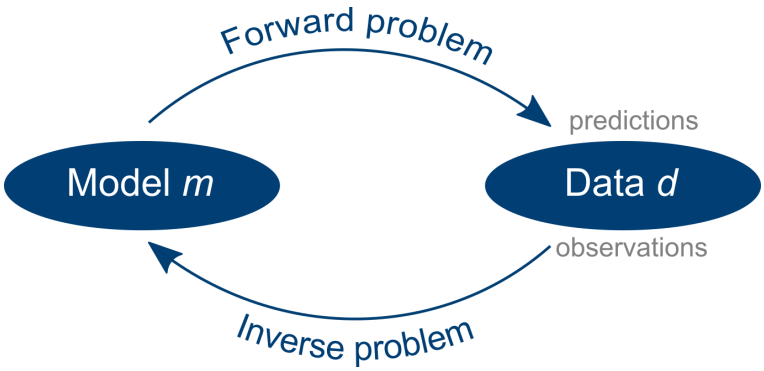


Figure 3.3: The conventional definition of the forward and inverse problems.

[6]Or any other system we have a set of observations for, e.g. Mars or Moon.

For purely linear problems, the forward relationship (see Equation 3.1) simplifies to $d = Gm$ and ostensibly, one might be tempted to formulate the inverse problem as

$$m = G^{-1}(d) \tag{3.7}$$

to estimate a model m based on observed data d. However, in practice G is likely to be large, very sparse and not square, which means that it will not be invertible. Even if it was reformulated to give the least-squares solution, which can deal with the non-squareness of G, the inverse problem will still likely be underdetermined. As a result, a wealth of techniques to solve linear problems have been developed along with several regularisation approaches (e.g. the Moore-Penrose inverse – Moore, 1920; Penrose, 1955; Tikhonov regularisation – Tikhonov, 1943), which is beyond the scope of this thesis to describe. For a more detailed introduction to this topic, I refer the interested reader to Chapter 14 in Nolet (2008). Unfortunately, most of the time, the forward operator G is not linear and hence the elegant machinery of linear inverse theory is no longer applicable. Thus, a more indirect approach is required for non-linear problems. A common approach is to iteratively update a candidate model based on the differences between its data predictions and the observations in order to improve the mismatch. Such an iterative non-linear approach is widely used in both ray and waveform tomography applications. In ray-based traveltime tomography, the problem is frequently locally linearised around a starting model using a Taylor series expansion and a regularised linear least-squares inversion is carried out, usually formulated as a perturbation to the starting model (e.g. Woodhouse and Dziewonski, 1984; Rawlinson and Sambridge, 2003), which can be iterated to address the non-linearity of the problem. In waveform tomography, the inverse problem is also non-linear and typically solved using this kind of iterative approach under the assumption of weak non-linearity[7]. Depending on the computational cost of the forward problem and number of model parameters (both are particularly large in waveform tomography), it follows that most often only one model is evaluated – a so-called deterministic approach – but ideally, we would like to be able to assess the performance of many models to evaluate their similarities, differences and presence of robust features.

The underdetermined nature of geophysical inverse problems arises from their inherent ill-posedness. An inverse problem is *ill-posed* (Hadamard, 1902) if at least one of the following three properties is not satisfied:

Existence

A solution exists – there exists a model m that fits the observed data d.

Uniqueness

The solution is unique – only one model m fits the observed data d.

Stability

The solution depends continuously on the initial conditions – small changes in the observed data d result in small changes in the estimated model m.

[7]Note that there are also fully non-linear problems, which cannot be solved using locally linearised techniques and thus, stochastic methods are often employed. One example is the inversion of seismic receiver functions for which the misfit landscape is very rough and a linearised approach will only explore a limited area near the starting model, which is unlikely to produce a satisfactory solution (e.g. Sambridge, 1999; Dreiling et al., 2020).

This ill-posedness makes geophysical inverse problems hard to solve. This can be mainly attributed to the non-uniqueness (Backus and Gilbert, 1970) of the problem; in other words, multiple models will explain the observations equally well. This has different origins, such as an inherent non-uniqueness caused by the physics behind the observation, where different models produce indistinguishable data. A common example is a gravity anomaly, which can be explained by an infinite number of density models (e.g. Snieder and Trampert, 1999) due in part to a depth/size/magnitude trade-off (see Figure 3.4). In seismic imaging, the main culprits are the sensitivity to initial parameters, insufficient data coverage and errors e in the data. One usually distinguishes between randomised errors, which are often easy to detect, but difficult to fix (e.g. seismic noise caused by road traffic), and systematic errors, which may be fixable, but are difficult to detect (e.g. instrument errors such as clock drifts, earthquake hypocentre mislocations). Hence a more realistic formulation of Equation 3.7 would be $d = G(m) + e$.

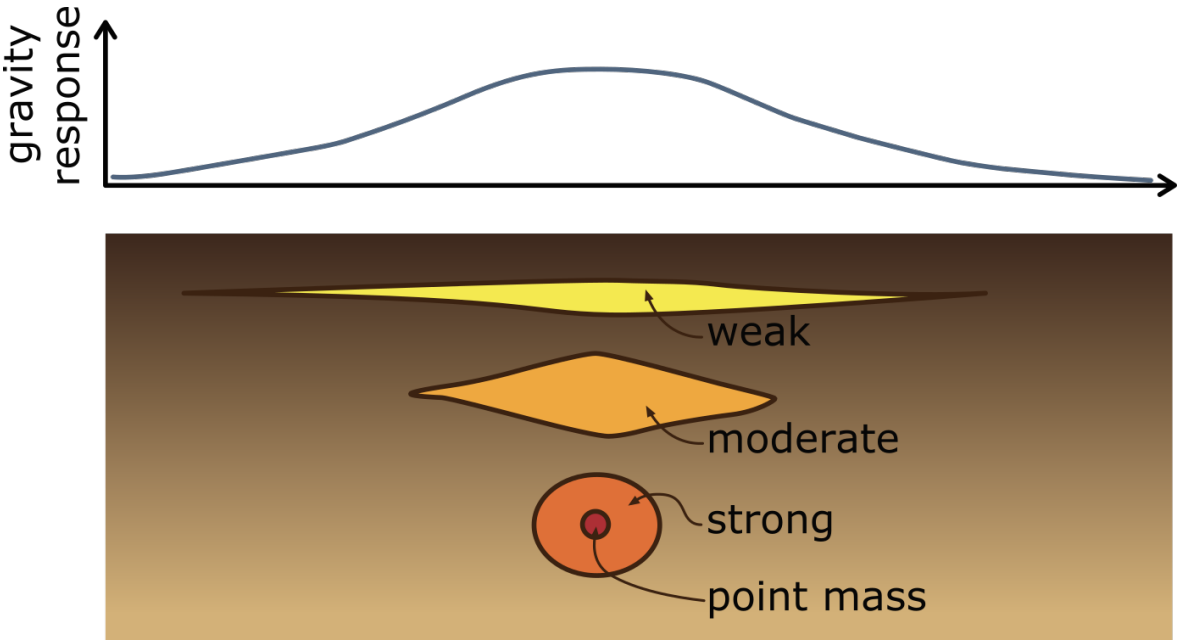


Figure 3.4: A hypothetical gravity survey illustrating the non-uniqueness of the gravity inverse problem. All four models are equally capable of explaining the data. Modified from Saltus and Blakely (2011) and Blom (2018).

In an ideal world, we would seek the true model m_{true} such that $d(m_{true}) = d_{obs}$. For reasons discussed above and a finite number of data points that contain observational errors, we can only ever hope to get an estimate of the true model. In fact, a misfit value below observational and forward modelling errors indicates that the model is responding to noise, and hence may result in artefacts (e.g. Rawlinson and Spakman, 2016). Nevertheless, the overall aim is to determine a physically meaningful model that minimises the misfit between synthetic and observed data, subject to any prior information that may be available. In the next section, I will discuss the quantification of the difference between synthetic and observed data.

3.3.1 Defining a waveform misfit: The objective function

The objective in waveform tomography is to update a model m in order to improve the match between synthetics d and observed waveforms d_{obs}. The misfit function or objective function χ summarises this difference for all sources and receivers in a single number (> 0), which we aim to minimise

$$\min \quad J(m) := \chi(d(m), d_{\text{obs}}) \tag{3.8}$$

with respect to changes in the model m.

The misfit function quantifies the differences between observed and predicted waveforms. It defines the type of measurement(s) we are making on a seismogram and can thus have a significant effect on the tomographic result. We usually seek a complete quantification of seismogram differences, that is a misfit of zero implies the two compared waveforms are identical. Finding a suitable misfit function in waveform tomography remains an active area of research (e.g. Yuan et al., 2020) and the choice often depends on the given dataset and objective of the study. Below, I will give a brief overview of the most common misfit functions used in waveform tomographic studies:

L_2-norm misfit — "L_2"

The L_2-norm, or least-squares norm, is the squared difference between observed and synthetic waveforms, which inherently makes use of all amplitude and phase information contained in a seismogram. The main drawbacks are that it is naturally dominated by large amplitudes and is strongly sensitive to outliers (i.e. it is non-robust). Furthermore, a simple time-shift can lead to a disproportionate misfit increase, which makes it less robust for real-world applications. However, it has been successfully used in exploration studies (e.g. Bamberger et al., 1982; Igel et al., 1996).

Cross-correlation travel-time misfit — "CC"

Luo and Schuster (1991) proposed to explicitly account for phase information using a cross-correlation misfit, which subsequently became popular in both exploration (e.g. Zhou et al., 1995) and earthquake seismology (e.g. Tape et al., 2010). The cross-correlation misfit computes a time shift between two waveforms corresponding to the maximum cross-correlation coefficient. This time shift is only a secondary observable and hence similar waveforms are required to ensure that the time shift is a proper replacement. To extract further information, many studies employ frequency-dependent cross-correlation misfit measures in order to capture wave dispersion (e.g. Sigloch et al., 2008). Multi-taper misfits, which calculate frequency-dependent transfer functions that transform the synthetic into the observed waveform (e.g. Zhou et al., 2004), represent a further extension of this approach.

Time-frequency (phase) misfit — "TF"

Both misfit functions described above do not yield physically meaningful results when the waveforms are not similar and phases are not clearly identifiable, e.g. when several phases interfere. Thus, Kristeková et al. (2006) – further extended in Kristeková et al. (2009) – and Fichtner et al. (2008) proposed the use of time- and frequency-dependent

misfits, where phase and amplitude information are separated. A particular advantage of this approach is that individual seismic phases do not need to be identified and isolated. However, careful data selection is still essential to avoid cycle skips – the misalignment of the two waveforms – which would result in phase jumps.

The misfit is based on the transformation of both observed and synthetic seismograms to the time-frequency domain, which can be done in different ways: Kristeková et al. (2006) use a wavelet transform while Fichtner et al. (2008) employ a windowed Fourier transform to analyse how the frequency content of the data evolves with time. In this study, the approach proposed by Fichtner et al. (2008) is adapted to characterise a signal in the time-frequency domain, where data can be represented by one term for phase and one term for the envelope information. Usually, only phase misfits are considered, but relative amplitude information is still incorporated as a result of Fourier transforming the data. The phase misfit χ_p is used throughout this thesis and can be formulated as a weighted L_2 norm of the phase difference $\phi^{\mathrm{syn}} - \phi^{\mathrm{obs}}$ for a single waveform component u as follows:

$$\chi_\mathrm{p}{}^2(u^{\mathrm{syn}}, u^{\mathrm{obs}}) = \int_{\mathbb{R}^2} W_\mathrm{p}{}^2(t,\omega)[\phi^{\mathrm{syn}}(t,\omega) - \phi^{\mathrm{obs}}(t,\omega)]^2 dt\, d\omega, \tag{3.9}$$

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). W_p attains values close to one when a meaningful measurement is possible and zero when it is not. It should be noted that it is assumed that suitable waveforms have been chosen, which were filtered within the frequency range of interest before being compared to each other. For a more detailed derivation, including information on the time-frequency envelope misfit, the reader is referred to Fichtner et al. (2008).

Instantaneous phase and exponentiated phase misfits — "IP" and "EP"

A similar misfit function to *TF* is the instantaneous phase, which extracts pure phase measurements and is thus independent of amplitudes (Bozdağ et al., 2011). As opposed to the previously mentioned time-frequency misfit, it is derived in the time domain only, which circumvents the data processing related to the transformation of the data to the frequency domain. Just like with any other phase measurements, great care should be taken to avoid cycle skips (e.g. Rickers et al., 2012).

Closely tied to the instantaneous phase, Yuan et al. (2020) proposed the exponentiated phase misfit to measure phases more uniquely and consistently. This is achieved by extracting phase information implicitly using a complex-valued phase representation.

Regarding amplitudes, it should be noted that it is a common procedure in earthquake seismology to disregard absolute amplitude information for reasons of source uncertainty, inadequate instrument response information and contamination owing to site effects (see Tromp, 2020). Still, relative amplitude differences can have a significant effect on the misfit distribution; many objective functions favour large-amplitude signals, which in most cases come from surface waves. Thus, sophisticated weighting and data selection algorithms have been introduced to balance amplitude differences, which I will discuss in more detail in Section 4.2.5.

As noted previously, the final choice of the misfit function often depends on the given dataset and should be made carefully because it strongly dictates the information that is extracted, e.g. a cross-correlation misfit measure can be suitable for long-period seismograms, where a simple time-shift is able to account for the waveform difference. Combining various objective functions can be sensible in order to optimise the results and such hybridisation approaches have occasionally been suggested, most recently by Yuan et al. (2020).

3.3.2 Obtaining the gradient: The adjoint method

Once a misfit function that quantifies the difference between synthetic and observed waveforms has been defined, the aim is to find a model that minimises the misfit function under the assumption of local linearity. Thus, we seek the first derivative of the objective function χ (see Equation 3.8) with respect to the model parameters m. This can be written as

$$\frac{\delta\chi}{\delta m_\mathrm{i}} = \frac{\delta\chi}{\delta u_\mathrm{j}} \frac{\delta u_\mathrm{j}}{\delta m_\mathrm{i}} = \nabla\chi^\mathrm{T} \frac{\delta u_\mathrm{j}}{\delta m_\mathrm{i}} \tag{3.10}$$

and corresponds to the gradient connecting changes in the model m to changes in the wavefield u. The obstacle in waveform tomography is to circumvent the direct computation of $\frac{\delta u_\mathrm{j}}{\delta m_\mathrm{i}}$ because it is computationally prohibitive to do so. This is due to the large number of model parameters, each of which would require the separate calculation of the forward problem. Avoiding this can be achieved by using the so-called *adjoint state method*, or simply *adjoint method*, which I will present in this section.

In many cases, the adjoint method is a convenient and computationally feasible way of computing the gradient, which points in the direction of misfit increase[8]. In a nutshell, the method is based on the time reversal of the wave equation: Source and receiver switch positions, the forward problem is solved again and the interaction between both wavefields is computed, thus "painting" the gradient (Tromp et al., 2005). The method was first introduced in the theory of inverse problems by Chavent (1974) and has been applied to a wide range of fluid mechanical problems in hydrology, meteorology and oceanography in the 1970s and 1980s (e.g. Yeh, 1986; Navon, 1998). Lailly and Bednar (1983), Bamberger et al. (1977), and Tarantola (1984) formulated the adjoint method for seismic problems, originally for exploration purposes. Later, its applicability for regional 3-D Earth models was demonstrated by Liu and Tromp (2006).

[8]Thus, we consider the negative gradient since we are interested in decreasing the misfit between waveforms.

In order to understand how powerful the method is, we will first show that the second wavefield mentioned above can indeed be described by a wave equation, and then elaborate on the benefits this yields. In the interest of subsequent legibility, the forward problem defined in Equation 3.4 can be rewritten more generally as

$$L(m)u(m) = f, \tag{3.11}$$

where f denotes a source term. L represents the laws of physics (here, the seismic wave equation) and u is the observed field (here, the wavefield), both depending on the Earth model m, which is implied in the following. In the next step, we differentiate the forward Equation 3.11 with respect to m, keeping in mind that we ultimately seek to eliminate the term $\frac{\delta u}{\delta m}$. Taking the chain rule into account, this yields:

$$\frac{\delta L}{\delta m}u + L\frac{\delta u}{\delta m} = 0 \tag{3.12}$$

Now, we introduce an arbitrary function h:

$$h^{\mathrm{T}}\frac{\delta L}{\delta m}u + h^{\mathrm{T}}L\frac{\delta u}{\delta m} = 0 \tag{3.13}$$

and add this to Equation 3.10:

$$\frac{\delta \chi}{\delta m} = \nabla\chi^{\mathrm{T}}\frac{\delta u}{\delta m} + h^{\mathrm{T}}\frac{\delta L}{\delta m}u + h^{\mathrm{T}}L\frac{\delta u}{\delta m} = h^{\mathrm{T}}\frac{\delta L}{\delta m}u + (h^{\mathrm{T}}L + \nabla\chi^{\mathrm{T}})\frac{\delta u}{\delta m} \tag{3.14}$$

In order to satisfy Equation 3.13, we can specify the following requirement for the last term to be zero:

$$L^{\mathrm{T}}h = -\nabla\chi, \tag{3.15}$$

which is the so-called adjoint wave equation – i.e., a wave equation (compare to Equation 3.11), with the fictitious adjoint wavefield h. While the source term for the forward wavefield is given by a seismic source, the adjoint source is fully determined by the misfit (see right-hand side of Equation 3.15). Each misfit measurement has its own adjoint source and its formulation depends on the chosen misfit function (see Fichtner, 2010). The adjoint sources are injected simultaneously at the receiver position backwards in time in order to satisfy terminal conditions (while the forward wavefield satisfies initial conditions). Note that we only inject misfit measurements in reverse time; the adjoint wavefield still propagates forward in time, and is – like the forward wavefield – governed by the seismic wave equation (see Equation 3.4).

Finally, by inserting Equation 3.15 into Equation 3.10, the gradient is given by

$$\frac{\delta \chi}{\delta m} = h^{\mathrm{T}}\frac{\delta L}{\delta m}u \tag{3.16}$$

which can be solved using Equation 3.15. This yields the gradient for a single event, the so-called event kernel, and the sum over all events produces the final gradient of the objective function. The 3-D gradient – or *sensitivity kernel* in loose terms – can be visualised as the

volumetric sensitivity of seismic waves to Earth structure. The sensitivity is built up by the interaction between the forward and adjoint wavefield, which highlights where the synthetic model differs from the true model. Examples for sensitivity kernels can be found in Figure 4.15 and Figure 4.16.

In summary, the gradient our model update will be based on can be obtained as follows:

1. Solve the forward problem to obtain the forward wavefield.
2. Compute the adjoint sources determined by the misfit between synthetic and observed waveforms.
3. Solve the adjoint wave equation in order to obtain the adjoint wavefield.
4. Obtain the misfit gradient by computing the interaction between the forward and adjoint wavefield.

Computational cost of the adjoint method

One of the main computational advantages of the adjoint method is that for each source only two numerical simulations are needed to compute the gradient of a misfit function. Thus, the computational cost scales linearly with the number of events. Furthermore, the same wave propagation solver can be used for both the forward and adjoint simulations. As we compute the wave propagation through the entire domain, we can extract its solution at all desired locations and thus, the computational cost is independent of the number of receivers.

One drawback of the method is that simultaneous access to the forward and adjoint wavefield is required in order to compute their convolution. The problem is that these wavefields have considerable data volumes (often several tens of terabytes) and storing them on disk is currently not feasible, even on powerful supercomputers. One way of circumventing this is the so-called *checkpointing* method (e.g. Anderson et al., 2012), where snapshots of the forward wavefield are stored. The forward wavefield is then restarted at these checkpoints when computing the adjoint wavefield. This obviously requires extra computations, but in particular on GPU clusters, a lot of time is spent on communicating data or waiting for memory access rather than the actual computations. Thus, checkpointing is an efficient method to avoid storing the wavefield on disk by trading memory requirements for additional computations.

3.3.3 The model update: Gradient-based inversion methods

The model update is based on the gradient. In a deterministic approach, the model update proceeds iteratively until some predefined criterion is fulfilled such as a stagnation of the model updates. The general approach is to go down the misfit slope for a current model m_i and re-evaluate the misfit for the new model m_{i+1}

$$m_{\mathrm{i+1}} = m_{\mathrm{i}} + \gamma_1 s_{\mathrm{i}} \quad \text{with} \quad \chi(m_{\mathrm{i+1}}) < \chi(m_{\mathrm{i}}) \tag{3.17}$$

with the step length γ and descent direction s.

Strictly speaking, the new model m_{i+1} is a trial model and will usually only be accepted if it indeed leads to an overall misfit decrease. Assessing this requires us to calculate the objective function for the trial model, which means we have to solve the forward problem again.

Descent direction

The descent direction s is (partially or entirely) determined by the negative gradient $-\nabla\chi(m)$ since the gradient is mathematically defined to point in the direction of the steepest ascent. It can be expressed as

$$s_i = -A_i \nabla\chi(m_i) \tag{3.18}$$

for any positive definite matrix A. The choice of A determines the type of gradient-based descent method used. Figure 3.5 gives an overview of current gradient-based optimisation schemes commonly used in FWI, ranging from steepest descent to Newton's method, which typically trade-off between being simple to implement and exhibiting rapid convergence[9]. In the following, I will briefly introduce the most widely-used inversion strategies that determine the path through model space. Note that they do not change the position of the global minimum we seek.

Figure 3.5: Schematic showing the spectrum of gradient-based descent methods commonly used in FWI. Adapted from "Lecture 2: Non-linear optimisation and adjoint methods" by Andreas Fichtner, 2020, retrieved from *10th Munich Earth Skience School* material.

Steepest descent method

In the simplest case, the descent direction is determined by the negative gradient only, which means A equals the identity matrix in Equation 3.18. Thus, the method is easy to implement and useful for testing purposes. However, simply searching in the direction of steepest descent is not necessary the fastest way of reaching the minimum. One common example is the Rosenbrock function (Rosenbrock, 1960), where the minimum lies inside a long, narrow, parabolic flat valley. Since the gradient is perpendicular to the contour lines, the algorithm zig-zags along the misfit valley in a very inefficient manner (see Figure 3.6). This results in slow convergence, i.e. the algorithm requires many iterations to converge to a minimum. Due to the high computational cost of the forward problem in waveform tomography, this is not favourable and has only been applied in early studies (e.g. Bamberger et al., 1982).

[9]Note that the *No Free Lunch Theorem* suggests that all optimisation algorithms perform equally well when averaged over all optimisation problems (Wolpert and Macready, 1997). Thus, this review is in the context of FWI only.

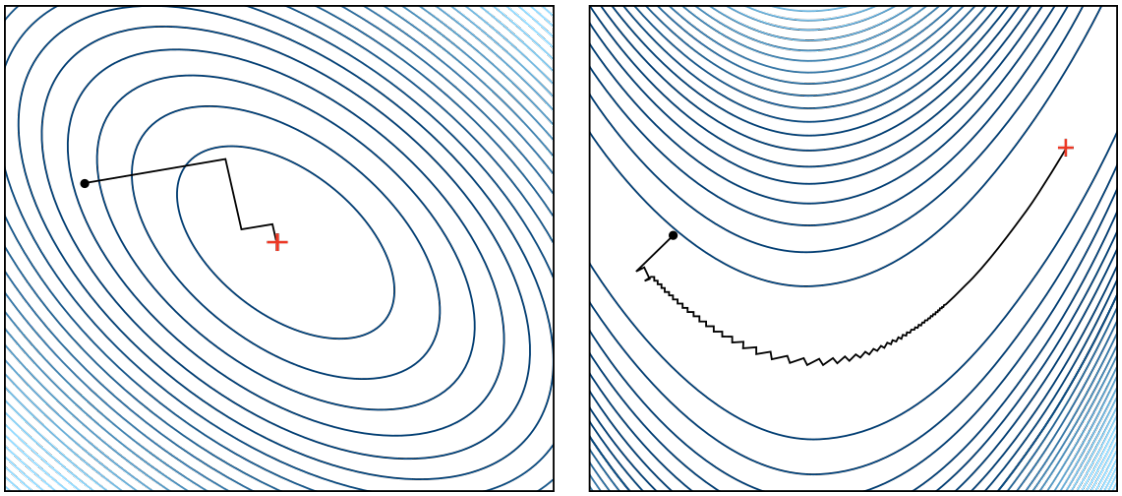


Figure 3.6: Illustration of the successive steps taken by the method of steepest descent for optimising (*left*) a quadratic function and (*right*) the Rosenbrock function. Note that the successive search directions are orthogonal.

Newton's method

Newton's method is the other end member of gradient-based descent methods since it is algorithmically complex, but results in fast convergence, e.g. a quadratic convergence rate for any smooth quadratic function (see Figure 3.7). This is achieved by using information about the curvature of the misfit valley, which can be obtained from the second derivative of the misfit function with respect to the model parameters. In more general terms, this is a Hessian matrix H and thus $A{=}H^{-1}$ in Equation 3.18. Computing the exact Hessian is currently prohibitively expensive in waveform tomography and to our knowledge, has so far only been applied to 1-D synthetic cases (e.g. Santosa and Symes, 1988).

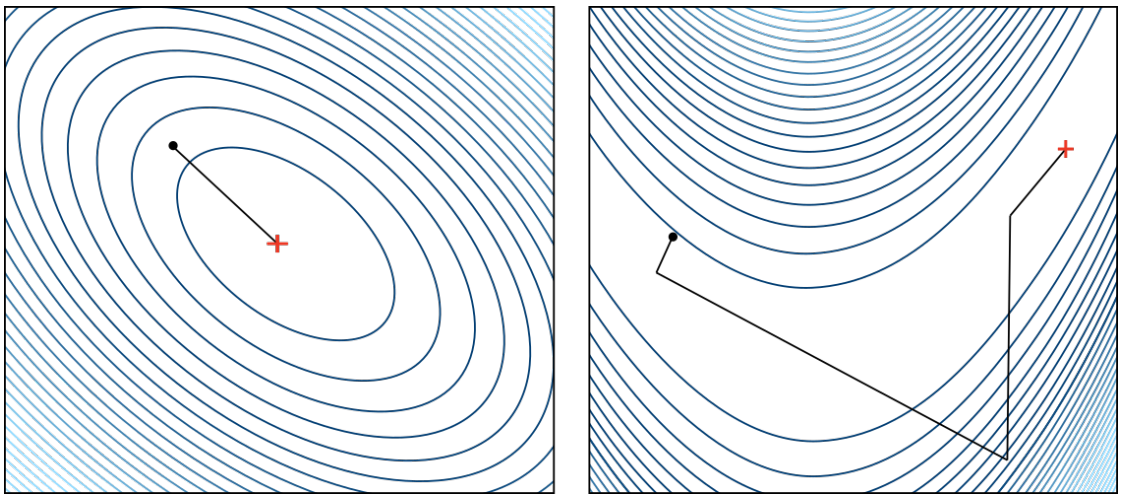


Figure 3.7: Illustration of the successive steps taken by Newton's method for optimising (*left*) a quadratic function and (*right*) the Rosenbrock function. Note the rapid convergence for the Rosenbrock function compared to the steepest descent approach described above.

In order to speed up convergence while bypassing the computational cost of computing the full Hessian, several intermediate approaches[10] have been proposed (see Figure 3.5). I will briefly discuss the two most popular ones below:

Conjugate gradient methods

Conjugate gradient methods use previous gradients in order to search in a direction that is conjugate to all previous search directions, without considering the Hessian. The initial search direction is equivalent to the steepest descent method and several methods have been developed to compute the subsequent descent directions resulting in different definitions of A (e.g. Hestenes and Stiefel, 1952; Fletcher and Reeves, 1964).

For quadratic misfit functions, the method converges in a number of iterations that is at most the number of model parameters. For general misfit functions, the method is an approximation, but tends to converge more quickly than the steepest descent method for many applications (e.g. Modrak and Tromp, 2016).

BFGS and L-BFGS methods

For the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970), the descent direction is determined by the negative gradient, preconditioned by the approximation of the inverse Hessian ($A=\tilde{H}^{-1}$). This approximation is constructed from changes in the gradient as we march through the model space (Nocedal and Wright, 2006). Hence the initial search direction is equivalent to the steepest descent method since no additional information has yet been obtained. As the iterations progress, more information is accumulated and consequently, the BFGS method requires a considerable amount of storage to retain the "full history". This motivated the development of the *limited-memory BFGS* (*L-BFGS*) method, which only stores the most recent gradient evaluations (Nocedal, 1980; Liu and Nocedal, 1989).

One should bear in mind that the method is a quasi-Newton method because it constructs an approximation of the Hessian. Fichtner and Trampert (2011) show that there are significant differences between the full and pseudo-Hessian, which are most relevant in resolution and parameter trade-off analysis. Nonetheless, the L-BFGS method is nowadays the most popular and generally regarded as the most efficient method for waveform tomographic problems (e.g. Modrak and Tromp, 2016) but it is more difficult to implement than other methods (see Figure 3.5).

Step length

The step length γ determines how far to move in the descent direction. Most often it is selected using a line search method involving two trial step lengths, calculating the corresponding misfit, fitting a quadratic to them and selecting the minimum. The quadratic fit usually works for updates of at most a few per cent.

Another option to determine a step length is to solve the trust-region subproblem (e.g. Conn et al., 2000), which updates the model in one step and does not require any

[10] They are inherently linked, e.g. one can prove that the L-BFGS method is a conjugate gradient method.

additional simulations compared to line search methods. The misfit function is quadratically approximated within a local region and from the second iteration onward, this region is automatically adjusted based on the quality of the approximation that was observed in the previous iterations (e.g. van Herwaarden et al., 2020; Thrastarson et al., 2022), that is, the region is expanded if an adequate model was found within the trust-region. Thus, I will use a trust-region based L-BFGS method throughout the remainder of this thesis.

The local nature of gradient-based optimisation schemes

So far, we have assumed that we are converging towards the global minimum, i.e. the model resulting in the overall lowest misfit that we also hope corresponds to the model most closely aligned with the real Earth (there is no guarantee of this, of course). However, all gradient-based optimisation methods hold the risk of getting trapped in a local minimum since they can only go downhill within a misfit valley. Thus, these algorithms rely on an initial model that is sufficiently close to the true model to validate the assumption of local linearity (see Figure 3.8).

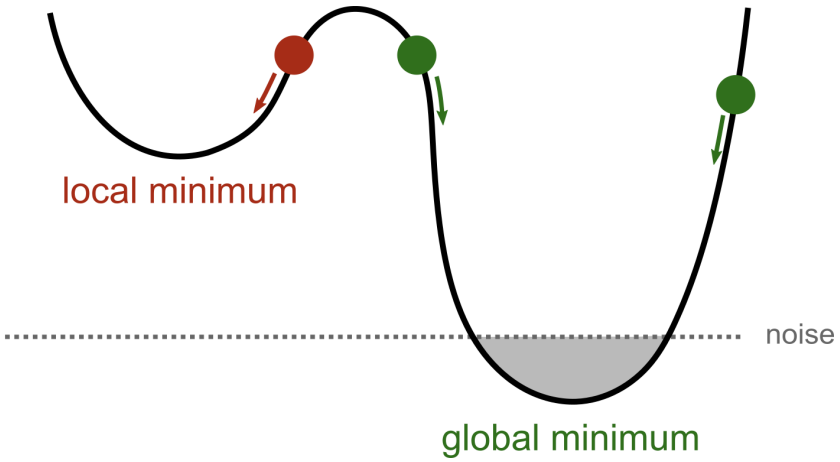


Figure 3.8: Schematic demonstrating the difference between local and global minima, and the importance of the starting model being located in the same misfit valley as the global minimum. A model in the left misfit valley (red dot) will converge to a local minimum while both models in the right valley (green dots) converge to the global minimum from different sides. Note that any model in the grey zone is considered valid due to the non-uniqueness of the problem.

To mitigate the risk of converging to a local minimum, many waveform tomography studies apply a multi-scale – originally referred to as "multi-grid" – approach (Bunks et al., 1995), where the longest periods are inverted for first, and higher frequency content is successively added. This is because higher frequency content dominates the signal in seismic tomography, which leads to a rougher misfit landscape (e.g. Gauthier et al., 1986) and hence increases the likelihood of entrapment in local minima. It should be noted that the multi-scale approach is built upon the assumption of weak non-linearity and does not guarantee convergence to the global minimum, but has been successfully applied over the past two decades (e.g. Bijwaard et al., 1998; Tape et al., 2007; Virieux and Operto, 2009).

In this context, I would also like to revisit the principle of non-uniqueness mentioned previously. The presence of noise and lack of data coverage means that there will always be many models that explain the observed data equally well, and the misfit landscape may contain many minima that are below the noise level (see grey zone in Figure 3.8), and hence without prior information one could not be favoured over any other. Furthermore, the direction from which the valley is entered is also important, and depends on the starting model and the iterative method chosen. The models within the right misfit valley in Figure 3.8 will likely not converge to the exact same model. Thus, the same dataset can yield significantly varying results depending on the starting model and inversion strategy chosen, and in fact many other choices such as the type of misfit function (see Section 3.3.1) and regularisation techniques applied (see Section 4.5). Ultimately, we do not know whether we are stuck in a local minimum since we only target one model using a deterministic approach. However, I will investigate strategies for assessing solution robustness in the following section.

3.4 Robustness tests and uncertainty quantification

Once a model has been determined, we would like to assess the robustness of the solution. Uncertainty arises from the ill-posedness of the inverse problem (see Section 3.3), the physics of the forward problem, the inversion setup, multi-parameter sensitivity[11], which includes the choice of parameterisation, noise etc. Ideally, we would like to assess many models that satisfy the data, and their associated parameter trade-offs and resolution lengths, which vary throughout the model as a result of heterogeneous data coverage (e.g. Yanovskaya, 1997). However, uncertainty estimates are difficult to quantify and as a result, this has remained an underdeveloped field to date. A comprehensive overview of this topic is provided in Rawlinson et al. (2014).

Uncertainty estimates for linear inverse problems are associated with the covariance and resolution matrices (e.g. Wiggins, 1972; Nolet et al., 1999). The earliest resolution tests in seismic tomography date back to the first studies in seismology; Aki and Lee (1976) and Dziewoński et al. (1977) estimate posterior covariance and resolution for a local and teleseismic study, respectively. Furthermore, Aki and Lee (1976) used synthetic reconstruction tests to estimate the resolution limits for a given dataset. Nowadays, spike tests (e.g. Walck and Clayton, 1987; Rawlinson and Spakman, 2016) and checkerboard tests (e.g. Hearn and Clayton, 1986; Rawlinson and Kennett, 2008) are popular synthetic tests, where synthetic data obtained from a short-wavelength anomaly and a regular alternating pattern of positive and negative anomalies, respectively, are inverted, which provides insight into smearing. Furthermore, statistical measures such as bootstrapping and jackknifing have received some attention (e.g. Gung and Romanowicz, 2004), where inversions using different data subsets are repeated. However, all of these methods have drawbacks regarding their insight into solution non-uniqueness (e.g. Rawlinson et al., 2014) and most importantly, they are currently computationally unfeasible in FWI.

[11] This can manifest as cross-talk in a tomography – more on this in Section 4.7.3.

For deterministic approaches applied in this thesis (see Section 3.3.3), the posterior covariance is related to the inverse Hessian, relying on a quadratic approximation of the misfit function. In principle, statistical methods based on random sampling such as Markov chain Monte Carlo or Hamiltonian Monte Carlo algorithms (e.g. Mosegaard and Tarantola, 1995; Mosegaard and Sambridge, 2002) could elegantly handle FWI problems and would be preferred since they provide information on the complete *posteriori* model distribution. However, both approaches are currently prohibitively expensive in FWI due to the size of the model space and the cost of the forward problem; consequently, obtaining uncertainty information remains an active area of research (e.g. Fichtner and Simutė, 2018; Gebraad et al., 2020; Liu et al., 2022).

Early resolution analysis in full-waveform tomography have mostly been limited to synthetic recovery tests (e.g. Chen et al., 2007) and estimates of composite volumetric sensitivity (e.g. Tape et al., 2010). In recent years, variants of spike tests have been used, e.g. Blom et al. (2020) added several point-localised anomalies δm to the final model and ran a few additional iterations to investigate to what extent they are removed. Furthermore, we recall that the full Hessian H is computationally prohibitively expensive to handle, but computing the Hessian-vector product $H\delta m$ can be feasible (e.g. Fichtner and Trampert, 2011; Krischer et al., 2018). The output is referred to as point-spread functions and gives an insight into spatial resolution and amplitude recovery but only allows specific regions of interest to be targeted. Random probing (Fichtner and van Leeuwen, 2015) permits quantitative resolution lengths to be obtained by sampling the Hessian using a small number of random test models. The Hessian-model applications are auto-correlated in different directions, yielding the corresponding direction-dependent resolution lengths. However, this approach assumes Gaussian errors and also requires additional simulations, which can be of the order of 25 % more than what is actually required in the inversion (e.g. Simutė et al., 2016).

3.5 Status quo and outlook

Recent computational advances have facilitated full-waveform inversion across a wide range of applications (see Section 3.1.2). However, we are still far from using all of the information contained in a seismogram, which is mainly a result of the method's computational constraints. Thus, FWI continues to be developed and improved, including but not limited to producing more realistic synthetics, extracting more information from seismograms and reducing the computational burden of FWI. In the following, I will highlight some recent advances within the field. For a comprehensive overview, the interested reader is referred to Tromp (2020).

As a rule of thumb, with today's computational resources, it is possible to simulate wavefields on the order of $\sim$100 wavelengths per domain in all areas where FWI is carried out. In earthquake seismology, FWI applications go down to periods of around $\sim$5 – 30 seconds on regional (e.g. Fichtner et al., 2009; Rickers et al., 2013; Krischer et al., 2015; Simutė et al., 2016; Tao et al., 2018; Blom et al., 2020; Gao et al., 2021; Chow et al., 2022) or even global scales (e.g. Bozdağ et al., 2016; Thrastarson et al., 2022; Noe et al., in prep.), using between

tens to a few hundred events. Reducing the computational burden of FWI would permit us to use significantly larger datasets, go to shorter periods, target areas with low seismicity rates and resolve deeper structures. Furthermore, statistical approaches would become feasible for larger problems (e.g. Gebraad et al., 2020; Zhang and Curtis, 2020), enabling us to sample the entire model space in order to explore the misfit landscape, which allows for a more sophisticated uncertainty quantification (see Section 3.4).

In order to reduce computational requirements, numerical advances including improvements and alternatives to the checkpointing method described in Section 3.3.2 (Böhm et al., 2016; Kukreja et al., 2020) as well as wavefield adapted meshes (van Driel et al., 2020; Thrastarson et al., 2020) have been made. Furthermore, there are developments towards teleseismic FWI – sometimes referred to as "box tomography" – (e.g. Roecker et al., 2010; Monteiller et al., 2015; Wang et al., 2016; Masson and Romanowicz, 2017; Möller and Friederich, 2021) and the simultaneous wavefield propagation simulation of several events have been proposed (e.g. Tromp and Bachmann, 2019; Romanowicz et al., 2020). To further decrease the number of required simulations, van Herwaarden et al. (2020) propose a gradient approximation using a dynamic subset of data, naturally exploiting redundancies in observed data.

Most waveform tomographic studies assume that differences between synthetic and observed waveforms arise from subsurface structure only. Hence realistic simulations are required to validate this assumption – for example by accounting for topography and the fluid ocean (e.g. Komatitsch and Tromp, 2002) –, otherwise modelling artefacts can be introduced by waveform residuals that are associated with these features. For example, van Driel et al. (2021) propose initial steps to include the full gravitational response, which is important for long-period seismology, but neglected in Equation 3.4. Furthermore, several other factors can influence waveforms such as data error and source effects. In particular, errors in source parameters can map as artefacts into the tomographic model (e.g. Blom et al., 2022). Thus, the ideal tomographic approach is a joint inversion of subsurface structure and source parameters (Valentine and Woodhouse, 2010), something that is the subject of ongoing developments by the waveform tomography community (e.g. Zhao et al., 2006; Hejrani et al., 2017; Fichtner and Simutė, 2018).

Another possible avenue for improvement is the multi-parameter sensitivity of seismic waves, which can readily be explored with FWI. To date, most studies only invert for v_P and (radially anisotropic) v_S, and few studies have investigated the benefits of imaging other properties (e.g. density, Blom et al., 2017; attenuation, Xing and Zhu, 2021), which requires the determination of the optimal observables to constrain a specific parameter (e.g. Bernauer et al., 2014). Furthermore, using an increasing amount of information from seismograms has been the aim of several studies, e.g. there are developments towards true-amplitude FWI (e.g. Wang et al., 2020) and the extraction of "hidden" phases on seismograms, which allow exploration of deeper parts of the Earth (e.g. Koroni et al., 2021).

Many waveform tomography studies reveal smaller and stronger perturbations compared to ray-based tomography, which indicate that the Earth may be more heterogeneous than we thought. FWI promises a more reliable quantification of anomalies, thus opening up new avenues for increasingly robust interpretation of seismic models in terms of composition,

temperature, melt and other material properties. This could be further improved by the inversion of multiple datasets, something that is already done with other seismic datasets, e.g. traveltimes jointly inverted with surface wave dispersion, gravity, heat flow, elevation etc (e.g. Afonso et al., 2016).

4 Southeast Asian waveform tomography: Inversion setup and dataset

This chapter provides a detailed overview of the inversion setup for the Southeast Asian waveform tomography, which requires a sophisticated workflow that involves numerous choices that need to be justified. The high computational cost of the forward problem means the setup needs to be treated with great care in order to avoid rerunning the inversion. The inversion setup includes but is not limited to the model domain, event and data evaluation, producing realistic synthetic waveforms and gradient preconditioning. It is explained in the context of the multi-scale approach (Bunks et al., 1995) applied throughout this study, which represents a gradual increase in the frequency content as the iterations progress (see Section 3.3.3). The new model will be presented and discussed in Chapter 5.

During the course of the PhD project, the *Salvus* software package (release 0.11.23 – 0.11.44) for full-waveform modelling and inversion is used for the mesh generation, forward and adjoint simulations as well as the non-linear optimisation, within its integrated workflow. *Salvus* was developed by *Mondaic AG/Ltd* (an ETH Zürich spin-off), and employs the spectral-element method to obtain a numerical solution of the wave equation (see Section 3.2.4). For more information about *Salvus*, I refer the interested reader to Afanasiev et al. (2019) and `https://www.mondaic.com`.

4.1 Model domain

In this section, I will elaborate on the mesh generation, which produces the 3-D model (a domain with medium properties attached) through which the propagation of seismic waves is computed (see Section 3.2). The extent of the geographical domain depends on the data availability and will be discussed in Section 4.2.

4.1.1 Choice of starting model

For gradient-based optimisation schemes, the starting model needs to be sufficiently close to the true seismic Earth in order to avoid entrapment in local minima (see Section 3.3.3). In case no region-specific model is available, a suitable 1-D background model can be chosen as the starting model. Several 1-D models of the whole Earth have been introduced over the last few decades (see Section 3.1.1), which mainly differ in the way the crust and discontinuities are represented and can thus vary by several per cent in these regions. However, Figure 4.1 shows that these 1-D velocity models produce similar synthetic waveforms in the period

range 100 – 150 s. This is due to the extensive sensitivity kernels at long periods and as a result, the wavefield only "sees" the overall 1-D structure of the upper few hundreds of kilometres. The time shift between the waveforms can be accounted for by regional updates, which usually only require one or two iterations (e.g. Fichtner et al., 2018), thus keeping the computational cost at a reasonable level.

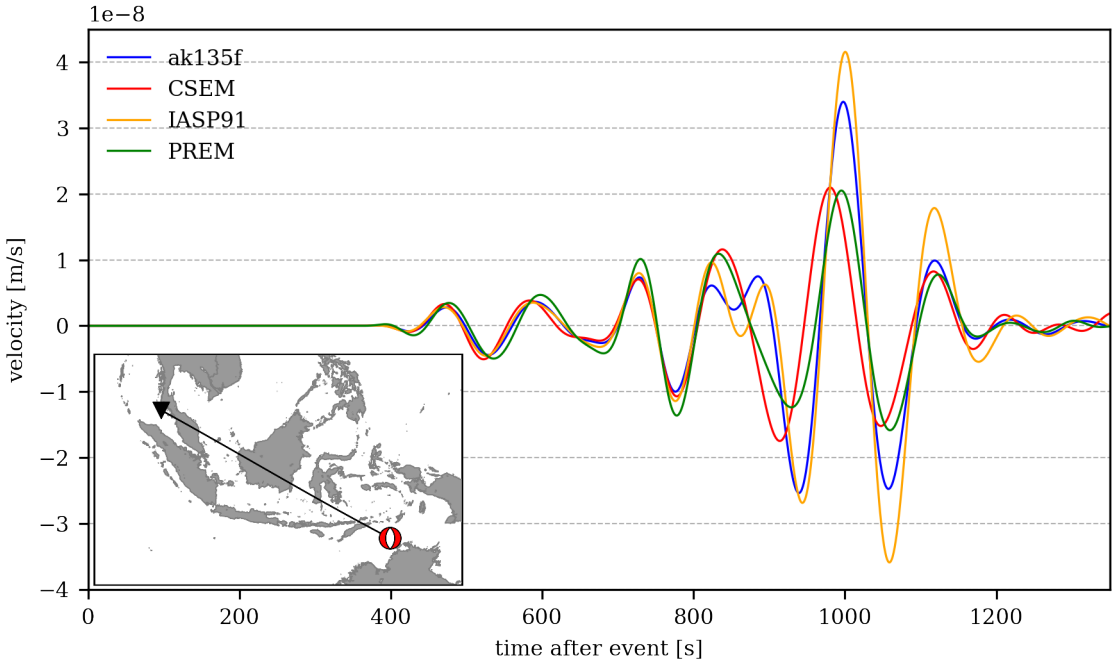


Figure 4.1: Vertical component synthetic waveforms in the period band 100 – 150 s recorded by a station at an epicentral distance of 36° from a synthetic M_w6.0 event for several 1-D velocity models (see Section 3.1.1): *ak135* (Kennett et al., 1995) in blue, *CSEM* (Fichtner et al., 2018) in red, *IASP91* (Kennett and Engdahl, 1991) in yellow and *PREM* (Dziewoński and Anderson, 1981) in green.

It should be noted that these 1-D models are usually obtained, inter alia, from significantly higher frequencies than currently considered in adjoint waveform tomography (see Section 3.1.1). This means that inhomogeneities at scales smaller than the minimum wavelength are included, which may lead to a needlessly fine mesh discretisation for our purposes. This results in a smaller time step for the wavefield simulation (see Equation 3.3) and thus, a higher computational cost. The homogenisation approach introduced by Capdeville et al. (2010a) and Capdeville et al. (2010b) addresses this problem by up-scaling the physical properties of the medium to what is really "seen" by the wavefield. For small perturbations across finite length scales, this is equivalent to a smoothed version of the elastic parameters. Sharp discontinuities, however, are known to lead to imprints in the tomographic result. Some discontinuities remain the subject of debate and if their presence in the starting model is erroneous, they will only be removed very slowly during an inversion, if at all. For example, the Lehmann discontinuity at ∼220 km depth has been suggested to be only a regional feature (e.g. Gu et al., 2001).

For this study, I adapt the *Collaborative Seismic Earth Model (CSEM): Generation 1* by Fichtner et al. (2018), which is built upon a smoothed version of the anisotropic *PREM* (hence the similarities between the respective waveforms in Figure 4.1). This has the advantage that the elastic parameters are obtained within a coherent framework. The *CSEM* model (see Figure 4.2) is designed to be conservative in the sense that it does not contain structure that seismic data cannot easily modify, e.g. the Lehmann discontinuity was replaced by a linear gradient. Furthermore, the elastic properties of the crust of *PREM* are replaced by the crustal model of Meier et al. (2007), which is derived from surface wave inversion. These simplifications mean that a lower number of mesh elements and hence a larger time step can be used for the wavefield simulation compared to the other aforementioned models. *CSEM* is assumed to be sufficiently close to the true Earth model since it still matches the longest-period data at 100 – 150 s to within half a cycle. I have chosen not to explicitly include crustal structure (e.g. *CRUST1.0*, Laske et al., 2013) in the starting model largely due to the added computational burden, but this is discussed in more detail in Section 5.4.3.

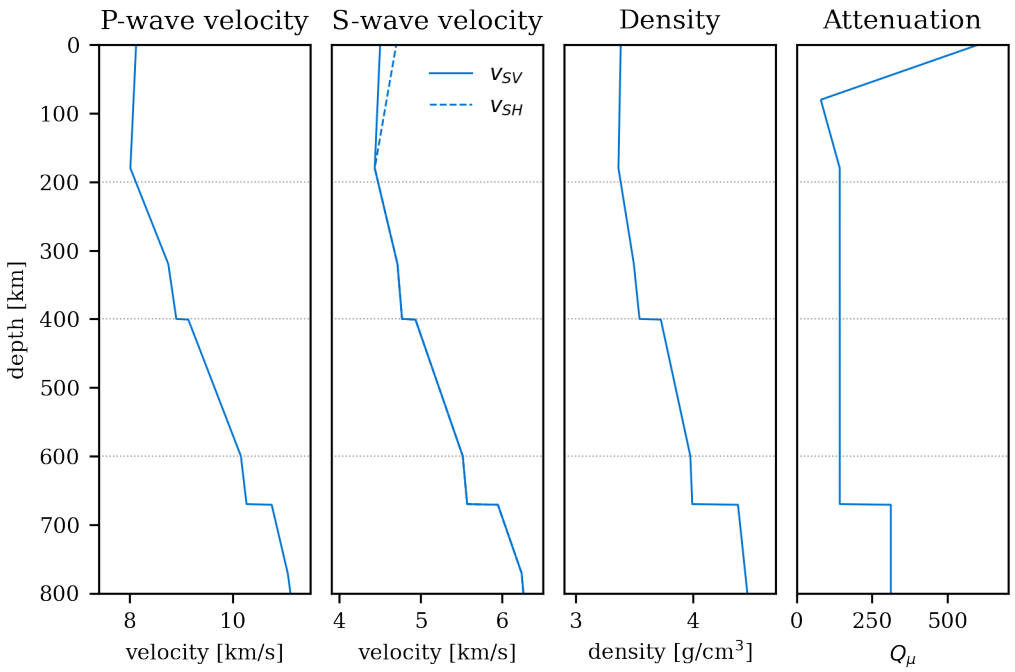


Figure 4.2: *CSEM: Generation 1* (Fichtner et al., 2018) absolute values as a function of depth for the upper 800 km for this region, which represents the starting model that is used in this study. Q_μ and Q_κ remain constant throughout the inversion. Q_κ is not shown, but has a constant value of 57,823.

4.1.2 Absorbing boundaries

For studies not undertaken at the global scale, the computational domain needs to be bounded on all sides since we only want to compute the wave propagation through the target region. The free-surface boundary condition is readily implemented in spectral-element methods (see Section 3.2.4). To account for the remaining non-physical boundaries of the computational domain, a two-stage approach is employed in *Salvus* to avoid artificial reflections: First, a first order Clayton-Enquist boundary condition (Clayton and Engquist, 1977) is applied, which relates the normal traction on the boundary to the normal and tangential components of the velocity field in elastic media. This boundary condition is computationally inexpensive and reduces reflections over a wide range of incident angles. However, it does not perform well for waves impinging on the boundaries at small angles and will still produce small reflections. These reflections can have amplitudes similar to reflections from heterogeneities in the medium. Thus, absorbing boundary layers – sometimes referred to as "sponge layers" – are implemented in a second step following the work of Kosloff and Kosloff (1986). Within these layers, the wavefield is attenuated based on a modification of the wave equation. The drawback of this method is the requirement to enlarge the computational domain significantly, where the width of the cushion region is proportional to the minimum wavelength the mesh can resolve. The effect of absorbing boundaries on synthetic waveforms is illustrated in Figure 4.3.

For this study, absorbing boundary widths based on 3.5 minimum wavelengths at a reference velocity of 6 km/s[1] are sufficient to cancel out artificial boundary reflections. Based on Equation 3.2, this results in absorbing boundary widths of 2,100 km at 100 s (first period band) and 420 km at 20 s (last period band) as summarised in Table 4.1 and illustrated in Figure 4.3. Here, implementing absorbing boundary layers represents an increase by a factor of two to five in the number of elements relative to the actual model domain. Further care was taken to ensure that sources and receivers are not located close to the edges of the domain since this can lead to low-frequency reflections from the edges.

It should be noted that the Gaussian damping functions proposed by Cerjan et al. (1985), which are frequently implemented in wavefield modelling methods (e.g. *SES3D* – Fichtner and Igel, 2008), are efficient for finite-difference methods, but lead to unacceptably large absorbing boundary layers for spectral-element methods. Another approach to remove reflections emanating from the boundaries are Perfectly Matched Layers (PMLs), which require a complex mathematical calculation including coordinate stretching. This is currently incorporated e.g. in *ExaHyPE* (Reinarz et al., 2020), and is on the wish list for *Salvus*.

4.1.3 Mesh generation, simulation accuracy and computational cost

The computational cost of the wavefield simulation is largely driven by the number of mesh elements and the number of time steps. The number of mesh elements depends on several factors such as the background model (see Section 4.1.1) and the number of elements per minimum wavelength (λ_{min}) for the shortest period considered (see Equation 3.2). For

[1] This translates to 21 elements in the cushion region in all directions in which absorbing boundaries are used.

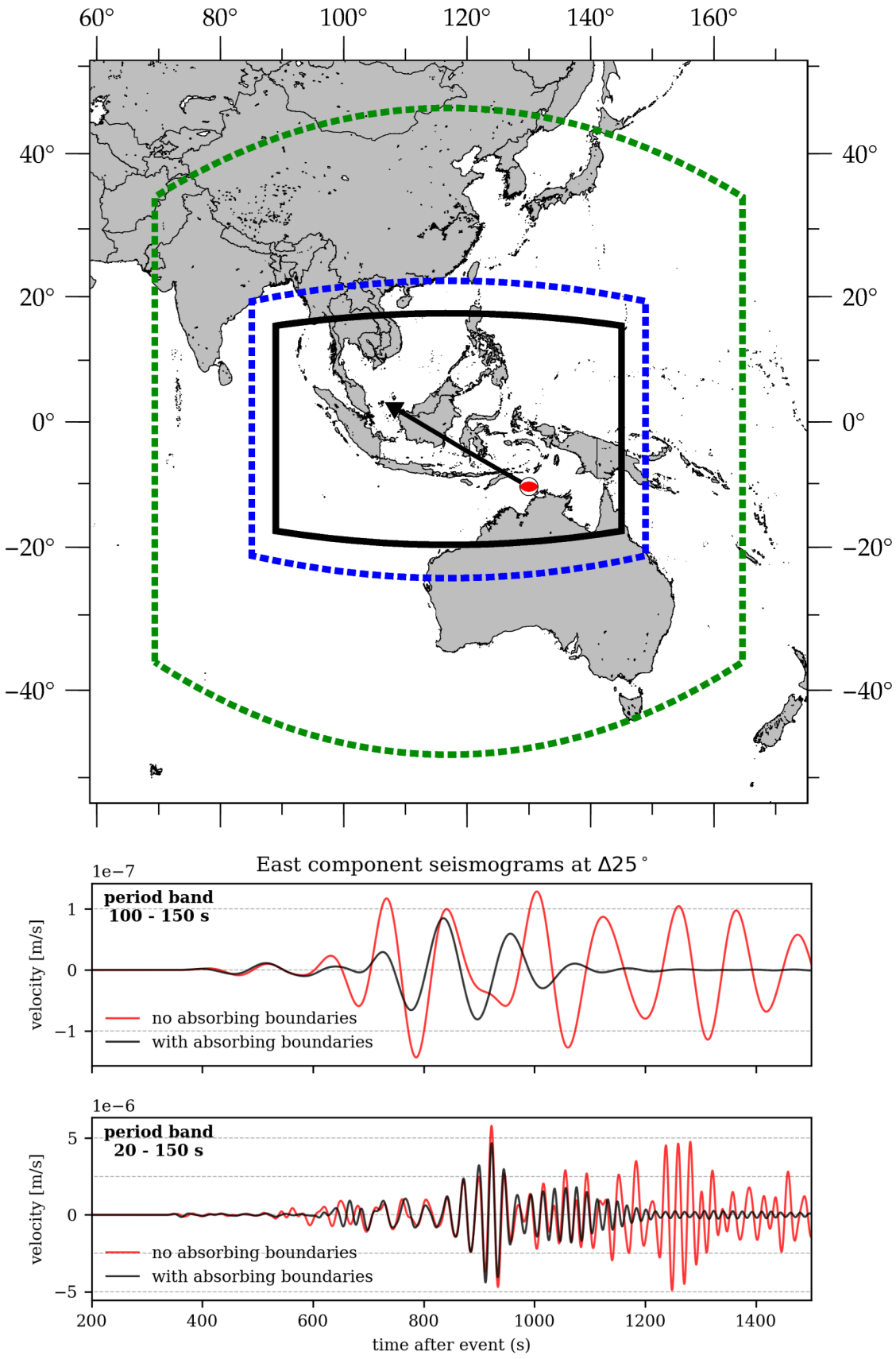


Figure 4.3: *Top panel:* Absorbing boundary widths at 100 – 150 s (green) and 20 – 150 s (blue) period ranges for this study region (black), based on 3.5 λ_{min} at a reference velocity of 6 km/s. *Bottom panels:* East component synthetic waveforms illustrating the effect of the Clayton-Enquist boundary condition and absorbing boundary layer at 100 – 150 s and 20 – 150 s. The location of the source and receiver are shown in the map in the top panel.

seismological applications, usually 1 – 2 elements per $\lambda_{\min}$ are chosen. Throughout this study, the mesh is generated based on 1.5 elements per $\lambda_{\min}$, but a synthetic test has shown that 1.2 would likely provide sufficient accuracy. This results in an element size of approximately 270 km and 60 km in the horizontal direction in the first period band (100 – 150 s) and last period band (20 – 150 s), respectively. In the vertical direction, elements are smaller where velocities are lower (e.g. the Earth's crust) and larger at greater depths (see Section 3.2.1). Furthermore, thin layers of elements are implemented around discontinuities. Table 4.1 presents the total number of mesh elements at each period band.

The other important parameter governing the computational cost is the number of time steps. Table 4.1 presents the time step per period band, which represents the temporal sampling of the wavefield (see Equation 3.3). At long periods, the wavefield simulation length is set to 1,600 s, which corresponds to the duration of seismic wave propagation through the entire domain. As the iterations progress, the surface wave train becomes more compact and thus, the simulation length is decreased to spare computational resources. However, once the data selection has taken place (see Section 4.2.5), one could base the wavefield simulation length on the maximum window end time in order to save computational effort.

Table 4.1: Overview of mesh and wavefield parameters: The total number of mesh elements (incl. absorbing boundaries), the absorbing boundary layer width, the duration of each earthquake simulation in seconds ("simulation length"), the time step of the wavefield sampling and the total number of time steps per period band. The computation time in the last column was obtained for the values in this table, accounting for attenuation but not storing checkpoints for the adjoint simulation (see Section 3.3.2).

period band	mesh elements	abs. layer width [km]	simulation length [s]	time step [s]	# time steps	comp. time on 1 core [min]
100 – 150 s (I)	14,250	2,100	1,600	0.55	2,909	∼11
80 – 150 s (II)	17,600	1,680	1,600	0.55	2,909	∼13
65 – 150 s (III)	23,400	1,365	1,600	0.55	2,909	∼19
50 – 150 s (IV)	33,866	1,050	1,600	0.5	3,200	∼29
40 – 150 s (V)	49,680	840	1,500	0.45	3,333	∼44
30 – 150 s (VI)	84,796	630	1,250	0.38	3,333	∼74
20 – 150 s (VII)	207,636	420	1,100	0.28	3,928	∼204

Table 4.1 also presents the computation time on one core for the mesh at each period band, which emphasises the increase in computational cost as the iterations progress due to the denser wavefield sampling. It should be noted that these compute times – and thus, any information regarding CPU hours used throughout an inversion – are indicative since a simulation with exactly the same settings can have very different run times depending on the specifications of the computer it was executed on. All tomographic inversions presented in this thesis were performed on the University of Cambridge's high-performance computing CPU-based cluster (`https://www.hpc.cam.ac.uk`), where the computation of the wave propagation was paralleled on 10 – 40 cores per event simulation with decreasing period. For

the last few period bands, I find that a good rule of thumb to enhance efficiency is to use one core per ∼5,000 mesh elements.

Note that accounting for attenuation does not change the number of time steps or mesh elements but it increases the computational cost by ∼30 % and significantly increases data volumes produced by the wavefield simulations. Attenuation is needed to account for the Earth's anelasticity and encompasses effects such as absorption and scattering, which have a significant effect on seismic waveforms (time delay and amplitude damping). *Salvus* approximates the frequency-dependence of attenuation with five linear solids (e.g. van Driel and Nissen-Meyer, 2014; Afanasiev et al., 2019). In this study, the attenuation quality factors are taken from the starting model *CSEM* (see Section 4.1.1) and remain constant throughout the inversion since inverting for attenuation would require a high density of data. However, one of the attractions of attenuation tomography is its strong sensitivity to temperature variations, and therefore its potential to image hot spots, mantle plumes and subduction zones (e.g. Eberhart-Phillips and Chadwick, 2002; Liu and Zhao, 2015).

4.1.4 Mesh interpolation

The multi-scale approach (Bunks et al., 1995) employed for this study involves a gradual increase in frequency content as the inversion progresses. It follows that the simulation mesh needs to be adapted to accurately sample the wavefield at these higher frequencies. Thus, the mesh becomes denser and consequently, the simulations more expensive as the iterations progress (see Sections 3.2.1 and 4.1.3).

The crux of the above process is that interpolation is not uniquely defined and should therefore be treated with caution. The easiest approach – without any loss of information – is to halve the size of each element in each direction. In 3-D, splitting each element evenly into eight (2^3) elements is equivalent to doubling the frequency content (e.g. decreasing the period from 100 to 50 s). However, this would be a large jump data-wise, likely resulting in phase jumps and problems obtaining longer wavelength structure. One option is to use the refined mesh for longer period simulations and thus highly oversample the synthetics, but this is likely not feasible due to the computational expense.

Here, I use the Python package *MultiMesh* (Thrastarson et al., 2021) for the mesh interpolation between different period bands. The package is built upon a nearest neighbour interpolation using *KDTree*, which is part of the *SciPy* library (Virtanen et al., 2020). I create a new mesh at the desired period, including bathymetry, topography and the ocean (or rather the approximation of it, see Section 4.4.2) at this period. In the next step, the difference between final and initial model of the last period band is interpolated onto the new mesh and the starting model (iteration 0) is added (as recommended by Dirk-Philip van Herwaarden, pers. comm., September 2020). I review the interpolated mesh at each period band and find that it deviates by less than one per cent for the absolute parameter values.

4.2 Quality control: Event and data evaluation

The model domain covers most of Southeast Asia, including some regions of particular geological interest such as the Indonesian volcanic arc, the Banda Arc curvature and the relatively poorly understood regions of Borneo and Sulawesi (see Chapter 2). The domain extension for this study is mainly driven by the data availability since all events and receivers have to be located within the model region (see Section 3.5). While a prodigious number of earthquakes occur in Southeast Asia, data from relatively few seismic stations have been made publicly available within the region to date. However, this study benefits from access to several restricted networks that allows an unprecedented dataset comprising recordings from 440 on-shore stations within Southeast Asia to be exploited. This represents a significant improvement in data coverage compared to what has previously been possible, which allows for the application of adjoint waveform tomography to this region for the first time. In the following, I will elaborate on the data availability and the compiled event catalogue, and provide information on the data processing and source parameterisation.

4.2.1 Data availability

Although almost twenty per cent of $M_w \geq 5$ earthquakes in the world are generated in Southeast Asia, only a handful of seismic stations were deployed within the region before 2005. The devastating 2004 Sumatra-Andaman earthquake prompted initiation of the *German-Indonesian Cooperation for a Tsunami Early Warning System* (*GITEWS*[2], Rudloff et al., 2009) program, with the first sensors installed in Indonesia only months after the severe tsunami catastrophe. Additional seismic deployments emerged within Southeast Asia over the following 15 years, mainly targeting hazardous regions such as the Sunda Arc.

Publicly available data[3] including instrument responses were downloaded automatically using *obspyDMT* (Hosseini and Sigloch, 2017), which accesses over 20 data centres via the *International Federation of Digital Seismograph Networks (FDSN)* and *ArcLink* interfaces. To our knowledge, this is currently the only toolbox offering an "update" mode for waveforms, response files and metadata information, where only new, modified, or previously failed data will be retrieved from the data centre(s) when a subsequent request is made. Furthermore, I access data from > 350 restricted stations allowing significant gaps in the publicly available data within the region to be filled, which leads to a substantial increase in data coverage. Restricted data were already available (Cambridge and Malaysian networks) or accessed via *Badan Meteorologi, Klimatologi, dan Geofisika (BMKG)* and the *Australian Passive Seismic Server (AusPass)*. I further request data from the *Academia Sinica, Institute of Earth Sciences* in Taipei for a station in the South China Sea (*TW.VNAS*), which is part of the *Broadband Array in Taiwan for Seismology (BATS)*.

[2]This project has been completed (Münch et al., 2011) and an extended system (*InaTEWS*) is currently operational under the Indonesian government (e.g. Häner and Kriegel, 2008; Harig et al., 2020).

[3]This includes data from the following networks: *AU* (Glanville and Geoscience Australia, 2021), *GE* (GEOFON Data Centre, 1993), *II* (Scripps Institution of Oceanography, 1986), *IU* (Albuquerque Seismological Laboratory/USGS, 1988) and *YR* (Metaxian et al., 2018). Furthermore, waveforms from networks *7A*, *MY*, *PS* and *ZB* are used, with further information available on `https://www.fdsn.org/networks/`.

The station map in Figure 4.4 shows the different seismic networks used in this study. Permanent, publicly available stations are mainly located on Peninsular Malaysia, Sumatra and in northern Australia. The Cambridge networks comprise two temporary deployments between 2018 and 2021: the *north Borneo Orogeny Seismic Survey* (*nBOSS*, Rawlinson, 2018), which was a deployment of 46 seismic instruments across northern Borneo, and a separate deployment of 10 stations across Kalimantan (in southern Borneo, Greenfield et al., 2018). Furthermore, our working group was granted access to several Malaysian stations and I take advantage of data from the *Banda Arc Project*, which targeted the Banda Arc-Australia continental collision using 33 stations recording between 2014 and 2016/2018 (Miller, 2014). In May 2020, I was granted access to the largest data set used in this study: the permanent Indonesian network of the *Meteorology, Climatology, and Geophysical Agency (BMKG)*, which comprises $>$ 200 stations in recent years.

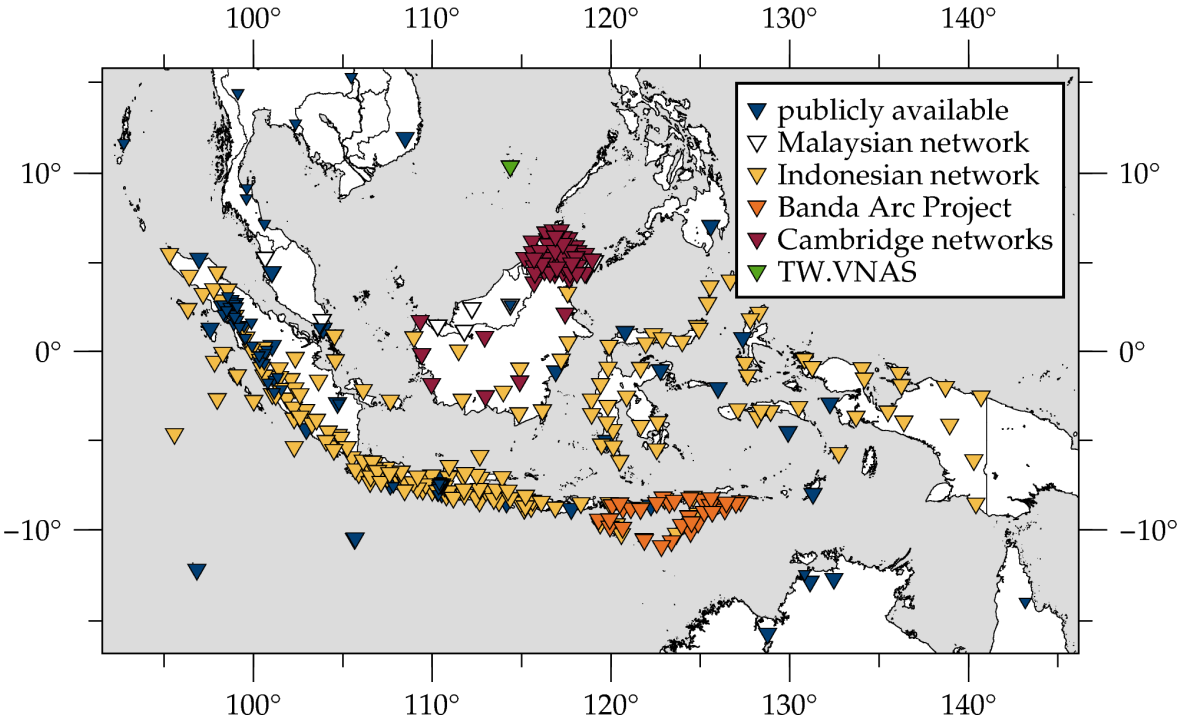


Figure 4.4: Map showing the 440 on-shore stations used throughout this study: publicly available stations (blue), Malaysian network (white), Cambridge networks (dark red), Indonesian network (yellow), Banda Arc project (orange) and *TW.VNAS* (green). Public stations contributing $<$ 30 waveforms to the inversion are plotted in smaller size, e.g. temporary networks on Java and Sumatra.

Several other temporary deployments within the region between 2004 and 2020 have been evaluated. This includes three deployments on Java, Indonesia: *Meramex* (Z6) in 2004, *TaskForce Java* (XN) in 2006 and *DOMERAPI* (YR) between 2013 and 2015 (Metaxian et al., 2018), as well as three deployments on and off the coast of Sumatra: *Simeulue* (ZB) in 2005/2006, *Sumatra Segmentation Aftershocks Deployment* (ZB) between 2007 and 2009, and *Lake Toba* (7A) in 2008 (Ryberg and Haberland, 2008). Overall, they do not yield sufficient signal-to-noise-ratio waveforms and thus, very few waveforms were actually used. This is mainly due to their lack of instrument sensitivity in the period ranges currently considered in

adjoint waveform tomography. For example, the Singapore Seismological Network (network code: MS), Australian Seismometers in Schools (network code: S1, Australian National University, 2011) and half of the *nBOSS* stations have sensors with a corner period of 30 s (e.g. Güralp CMG-6T) or lower. Thus, they only yield suitable waveforms at very late stages of the inversion, if at all. Note that there is no need to remove unsuitable waveforms yet since the data selection procedure will at a later stage define the portions of waveforms used during the inversion (see Section 4.2.5) and the computational cost of the method does not notably scale with the number of receivers used (see Section 3.3.2).

4.2.2 Data processing

Only basic processing steps are applied to the raw observed data using *ObsPy* (Beyreuther et al., 2010). Most importantly, the observed waveforms need to be filtered in the same way as the source. Note that it is crucial to use a prefiltered source rather than filtered synthetics[4] since the forward and adjoint wavefield will be combined during the gradient computation (see Section 3.3.2). After rotating to a ZNE coordinate system and fixing incorrect station rotations if flagged in the metadata, I detrend, demean and taper the data. Subsequently, the instrument response is removed to convert units from digital counts to actual ground motion in velocity (this is non-trivial, see Havskov and Alguacil, 2004). In this study, velocity traces are used, noting that the time derivative needs to be taken into account during the computation of the adjoint source as demonstrated by e.g. *Mondaic: Misfit functionals involving velocity*. In the final step, I apply a Butterworth bandpass filter for the desired period range (see first column in Table 4.1) because high frequencies that are not simulated need to be removed. The maximum period is set to 150 s since rotational and gravitational effects are neglected in Equation 3.4. This means that long-period surface waves ($>$ 120 s) – sometimes referred to as "mantle waves" – are still considered to some extent. Once the data is processed, it is written to the *Adaptable Seismic Data Format* (*ASDF*, Krischer et al., 2016), which is readable by *Salvus*.

An example that demonstrates the processing steps for the observed waveforms of an event used in the inversion can be found on a *Zenodo* repository at `https://doi.org/10.5281/zenodo.5573139` (Wehner et al., 2021). Furthermore, the repository contains the filtered waveforms used throughout this study. Note that only windowed waveforms are available since access to the majority of the dataset is currently restricted (see Section 4.2.1).

4.2.3 Source parameterisation

In most tomographic studies, earthquakes are parameterised by a point source in time and space. This is a valid assumption for small earthquakes and/or when considering long periods, where rupture effects (namely rupture length and duration) can be neglected. In this study, data from stations in the close proximity of the hypocenter are not considered since the source region is removed from the event gradient before a model update is computed in

[4]However, it can be shown that this yields equivalent synthetic waveforms, see *Mondaic: Equivalence of Filtering Source Time Functions and Filtering Synthetics*.

order to mitigate singularities in the gradient (see Section 4.5.1). Thus, one can assume to be in the "far field" at distances several wavelengths or more from the source (Nolet, 2008).

The source parameterisation has a significant effect on synthetic waveforms. However, seismic source inversion is a challenging task due to the non-linearity of the problem. Most catalogues determine only a location and a moment tensor solution for events, but it is generally also possible to invert for the temporal evolution of the rupture, which has been addressed in recent studies (e.g. Stähler and Sigloch, 2014). Yet, only a few catalogues estimate for the so-called source time function (e.g. Vallée, 2013; Garcia et al., 2013).

For this study, locations and moment tensors are retrieved from the *Global Centroid*[5] *Moment Tensor* (*GCMT*) catalogue (Ekström et al., 2012), which is based on long-period body and surface wave data ($\geq$ 40 s) and makes use of waveform matching between global seismic records and synthetic waveforms. Note that source parameters remain constant throughout the inversion for reasons outlined in Section 5.4.1.

The source time function is approximated by a Butterworth bandpass filtered Heaviside step function (see Figure 4.5), which represents an instantaneous rupture process. This is a valid assumption at long periods since the rupture duration of M5 – 6 earthquakes is on the order of seconds and hence small compared to the periods considered here. Note that the source time function should always start smoothly from zero to avoid artefacts in the wavefield (Lion Krischer, pers. comm., February 2020).

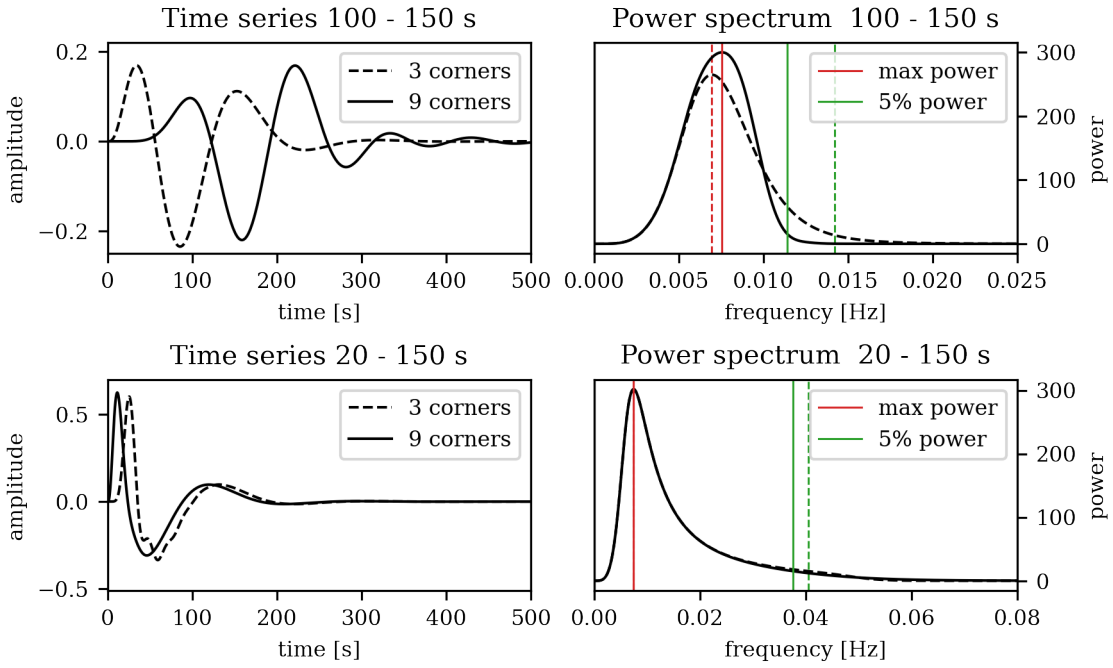


Figure 4.5: Comparison using three (dashed lines) vs nine (solid lines) lowpass filter corners for the Butterworth bandpass filtered source time function. *Left column:* The source time function at (top row) 100 –150 s and (bottom row) 20 –150 s. *Right column:* The corresponding power spectrum.

[5]This is the average location in space and time of the energy release and thus differs from the hypocentre, which is where the rupture started.

Figure 4.5 shows the source time function and its corresponding power spectrum for the first and last iteration bands at 100 – 150 s and 20 – 150 s periods, respectively. Here, I use nine corners for the lowpass filter to strongly damp higher frequencies. At 100 – 150 s, five per cent of the energy is below 90 s compared to 70 s when using three corners (which is the *ObsPy* default). This is not an issue per se, provided that the mesh properly accounts for these high frequencies (see Section 3.2.1). Note that the difference is more pronounced at longer periods.

4.2.4 Event selection

While frequent earthquakes with depths of up to 700 km generated in Southeast Asia (see Chapter 2) provide an excellent dataset for regional tomography, the high computational cost of FWI only permits the use of a small event dataset compared to ray tomographic studies. In fact, it is challenging to identify an optimal subset of events from the vast number of earthquakes within the region, for which events yield high-quality waveforms and are spatially distributed as evenly as possible.

There is a trade-off between earthquakes being small enough to be reasonably approximated by a point source and emitting enough energy at the relatively long periods currently considered. Record sections show that events with $M_w \geq 5.5$ have sufficient energy to generate high signal-to-noise-ratio waveforms at distant receivers within the domain. Hence, the event selection takes place using a process of elimination, where all events with magnitudes $M_w \geq 5.5$ within the domain are selected first. However, most large-magnitude earthquakes are characterised by rupture durations of several seconds and rupture lengths of tens of kilometres (e.g. Wells and Coppersmith, 1994; Leonard, 2014). Thus, earthquakes with $M_w > 7.5$ are disregarded since the point source assumption is strictly not valid. To avoid interference, events are eliminated if another event of $M_w \geq 7.0$ occurred elsewhere in the world in the past 12 hours, or a $M_w \geq 5.0$ earthquake occurred in an extensive area encompassing the domain either 60 minutes before or 30 minutes[6] after event time.

Earthquakes with a large number of recordings are preferred to enhance the efficiency of the adjoint-based inversion (see Section 3.3.2). Hence, I optimise the event catalogue for data availability, including the deployment of temporary networks, as described in Section 4.2.1. However, Southeast Asia is a seismically very active region and a wealth of earthquakes remain, even after this elimination process. Thus, one may wonder how to identify suitable events that yield waveforms with a high signal-to-noise ratio from a "prefiltered" event catalogue (negligible finite source effects, optimised for the number of recorded stations, etc).

I find that evaluating events based on the observed data from a set of high-quality stations (e.g. vertical components of permanent stations) is an effective way of reducing the number of events. Permanent stations tend to yield waveforms with a high signal-to-noise ratio because they are usually deployed with great care, e.g. by trained technicians in vaults, at sites that have been carefully selected for their low noise characteristics[7], and typically use high-quality instruments like *Streckeisen STS-2s*. However, it should be noted that events may

[6] This roughly translates to the maximum wavefield simulation time used in this study (see Section 4.1.3).
[7] Although this is sometimes trumped by other factors like accessibility and cost.

become suitable at later stages of the inversion, e.g. lower magnitude earthquakes exhibit sufficient energy at shorter periods.

In the final step, I review the source time functions (see Section 4.2.3) for the remaining events using *SCARDEC* (Vallée et al., 2011). Events are removed if the source time functions have durations comparable to the minimum period considered or appear generally suspicious. I find that these source time functions correlate with large event misfits throughout the inversion (see Figure 4.6). Moreover, using multiple earthquakes that occur in a similar location does not improve the inversion result, but will significantly increase compute time (see Section 3.3.2), so I am careful to include only those events that are likely to contribute meaningfully to the final model. For example, 15 events with depths > 300 km and several events in unique locations (e.g. around northern Borneo and the Indonesian back-arc basin) are selected to help improve data coverage and to ensure a spatial distribution that is as uniform as possible.

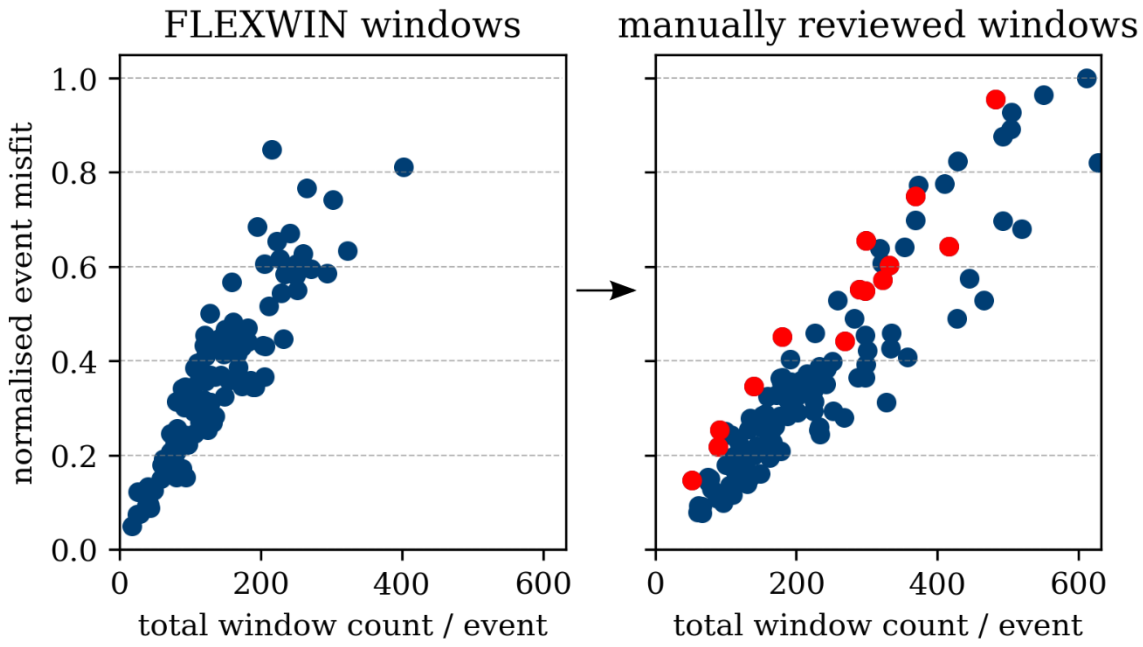


Figure 4.6: Effect of reselecting windows for the period band 65 – 150 s. Blue dots indicate normalised event misfits computed using the time-frequency phase misfits as presented in Section 3.3.1. Red dots indicate events that have suspicious source time functions according to Vallée (2013) and were removed before decreasing the period to 50 – 150 s.

In summary, the final event catalogue contains 143 earthquakes that range in size between $5.5 \leq M_w \leq 7.5$ and occurred between 2008 and 2020. A core event catalogue – events used at every period band – consists of 50 events of size $5.8 \leq M_w \leq 6.9$. The remaining 93 events are added or removed as the iterations progress, e.g. events with $M_w > 7.0$ are discarded from 50 s onwards to mitigate finite-source effects contaminating the tomography and lower-magnitude events are added at shorter periods. Figure 4.7 presents the focal mechanisms of the event catalogue. Table 4.2 presents the number of events used per period band, which decreases as the iterations progress due to the increase in computational cost

and careful event monitoring as shorter periods are considered. A detailed overview of the events, including their locations and the period bands they were used at can be found in Section A1 of the Appendix.

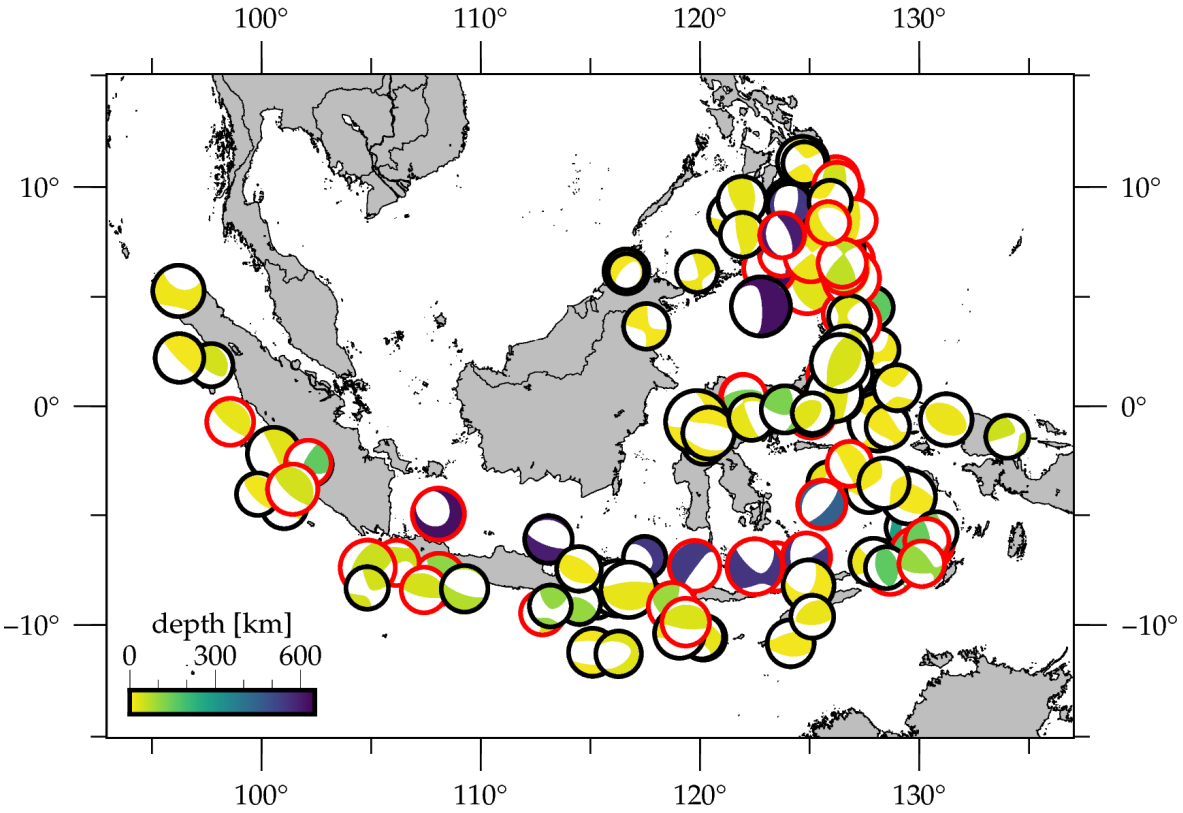


Figure 4.7: Focal mechanisms of the 143 earthquakes ($5.5 \leq M_w \leq 7.5$) used throughout this study, coloured by depth. Red circled beachballs indicate events used in every period band, while black circled beachballs indicate events which are added or removed as the iterations progress.

4.2.5 Data selection

Ideally, we would like to invert for the entire waveform (hence the term *full-waveform* inversion). However, noisy waveform portions and cycle skips can contaminate the tomography and need to be removed. Furthermore, depending on the chosen misfit functional (see Section 3.3.1), we need to balance arrivals of different amplitudes since large-amplitudes will otherwise dominate the misfit per trace. As a result, depth sensitivity largely derived from small-amplitude body wave arrivals ends up being lost. Thus, the commonly heard remark that waveform tomographic methods are predominantly sensitive to surface waves is partially a result of the data selection[8] and can be somewhat mitigated by specifically accounting for body wave signals separately. Both can be achieved by using measurement "windows", which define the parts of a seismogram suitable for waveform comparison.

An obvious problem with window picking is that most studies incorporate tens of thousands of seismograms, with numbers continuously increasing. Manual window picking is

[8] as well as the long periods currently considered in FWI

time-consuming, in particular because the window sets should be reviewed at each period band. Several toolboxes have been published that aim to automate the selection of suitable measurement windows. The following two window selection algorithms are commonly used: *LASIF* (*LArge-scale Seismic Inversion Framework*, Krischer et al., 2015) and *FLEXWIN* (*FLEXibility in picking time WINdows*, Maggi et al., 2009) with its Python port *pyflex* (Krischer and Casarotti, 2015). However, there is still a trade-off between including as much signal as possible and avoiding noisy data. Furthermore, an automated separation of body and surface wave arrivals is particularly challenging.

I compared the resulting windows for our dataset using the two above mentioned algorithms and found that *FLEXWIN* naturally separates different amplitude arrivals more effectively. *FLEXWIN* operates on pairs of observed and synthetic single component seismograms and applies a short-term average over long-term average (STA/LTA) to the envelope of the synthetic. Then, a multi-stage selection procedure examines whether there is an adequate match to the observed waveform by evaluating criteria such as the signal-to-noise ratio, time shift, shape comparison and amplitude ratio.

Windows for the first period band (100 – 150 s) are suggested using *pyflex*. However, more than 30 % of the suggested windows require adjustments and many suitable waveform portions were not detected by the algorithm. Thus, all windows are reviewed for each period band, employing the *LASIF* graphical user interface for the manual inspection of the *pyflex* windows. This allows accurate manual window picking since start and end time of windows can directly be picked per mouse click on the seismograms. Over the course of my PhD, I reviewed $>$ 500,000 seismograms, which is by far the most time-consuming part of the inversion setup, but it triples the analysed window lengths compared to the tuned *FLEXWIN* algorithm.

Figure 4.8 shows a comparison of *FLEXWIN* (blue) and manually selected (grey) windows for a three-component seismogram of a station in Sumatra (Indonesia), which recorded a M_w6.0 event in the Banda Arc region. This shows that the manual data selection incorporates more data; *FLEXWIN* suggests two windows comprising a length of $\sim$250 s, while I select three windows adding up to $\sim$750 s for this station. However, this example also shows that the data selection process is unavoidably subjective (e.g. see windows on horizontal components). Figure 4.8 also provides a basic demonstration of the data selection procedure, which can be summarised as follows:

- Windows are selected on both horizontals and the vertical component.
- Windows are only selected for seismogram portions suitable for waveform comparison (e.g. no cycle skips, clear signal, high-quality waveforms). Thus, no windows are selected on traces that do not yield a sufficient signal-to-noise ratio for waveform tomography (see Table 4.2) and windows can significantly vary in length.
- I aim to specifically account for body wave arrivals in order to enhance depth sensitivity. Thus, more than one window can be selected per trace (typically one shorter window for body wave arrivals and one longer window for surface wave arrivals).

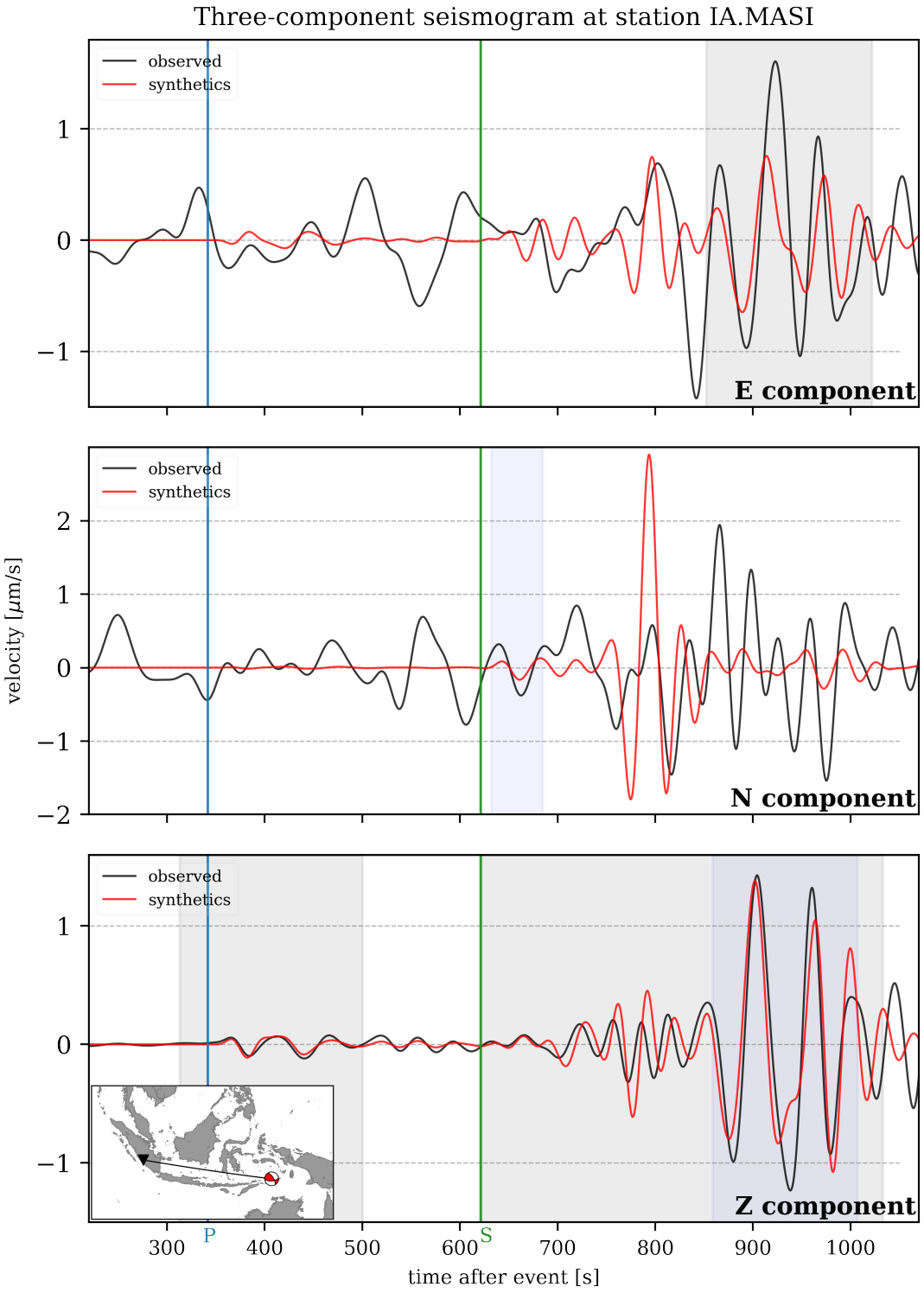


Figure 4.8: Comparison of synthetic (red) and observed (black) waveforms at 30 – 150 s for a three-component seismogram from a station in Sumatra (Indonesia), which recorded a M_w6.0 event in the Banda Arc region in November 2015. Blue windows were suggested by *FLEXWIN* (Maggi et al., 2009), while grey windows were manually selected and used throughout the inversion. Body wave arrivals are accounted for separately (see first grey window on vertical component) to enhance depth sensitivity. Vertical lines indicate predicted P- (blue) and S-wave (green) first arrival times obtained from the *TauP* toolkit (Crotwell et al., 1999) for *PREM* (Dziewoński and Anderson, 1981).

Note that only shaded grey areas in Figure 4.8 define the parts of the seismogram used in the inversion, and a time-frequency phase misfit value (see Equation 3.9) is computed for each of these windows. The misfit between synthetic and observed waveforms is computed before determining the adjoint wavefield to obtain the gradient (see Section 3.3.2), which allows for further quality control. Besides evaluating misfit maps in order to identify suspicious receiver misfits, I find it useful to plot the event misfit vs the number of windows selected per event in order to identify suspicious events, since we would expect a somewhat linear behaviour. Figure 4.6 presents an example for *FLEXWIN* vs manually reviewed windows, which demonstrates that 1) I select more windows per event than the *FLEXWIN* algorithm, 2) the linear relationship is better and 3) the slope for *FLEXWIN* windows is steeper, which means the overall event misfit for the same number of windows is larger. It should be noted that while the misfit calculation is non-linear, one can usually observe a cumulative effect of station misfit with epicentral distance in real-world applications. Thus, an event that has many windows for a dense network at a large epicentral distance will naturally have a large event misfit. The effect of such dense networks can be somewhat mitigated by the geographical station weighting described in Section 4.3.

Figure 4.9 shows the waveform match for a single source-receiver pair across four period bands; from 50 s downwards, body wave signals become clearly identifiable and are selected in a separate window to maximise depth sensitivity. In the final period band (20 – 150 s), windows around small-amplitude arrivals contribute approximately 30 % of the total number of windows. 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 analysed window length per event decreases despite the increasing number of windows. From 30 s downwards, the 3-D wavefield becomes increasingly complex (e.g. due to crustal scattering), which in turn translates to a smaller number of events used and thus, fewer windows.

In summary, $\sim$13,000 unique source-receiver pairs and a total analysed time window length of $\sim$1,000 – 3,000 h per period band (see Table 4.2) are used throughout this study, compared to $<$ 100 h in the continental-scale Australian FWI study by Fichtner et al. (2009), which emphasises the increase in computational resources and advances in FWI over the past decade. While the majority of windows are selected on the vertical component, 33 – 42 % of the windows per period band are selected on horizontal components. The final window sets for each period band can be found on a *Zenodo* repository at `https://doi.org/10.5281/zenodo.5573139`.

The station map in Figure 4.10 presents the number of events for which a given station contributes waveforms to the inversion, which provides an impression of the data coverage used in this study. This shows that stations in unique locations contribute many waveforms (e.g. Australia, South China Sea and south of the Indonesian volcanic arc). Furthermore, it shows that the *nBOSS* network in northern Borneo contributes relatively few waveforms, which is a result of the short deployment time frame and lack of instrument sensitivity (see Section 4.2.1). The maximum source-receiver distance in this study is $\sim$5,600 km.

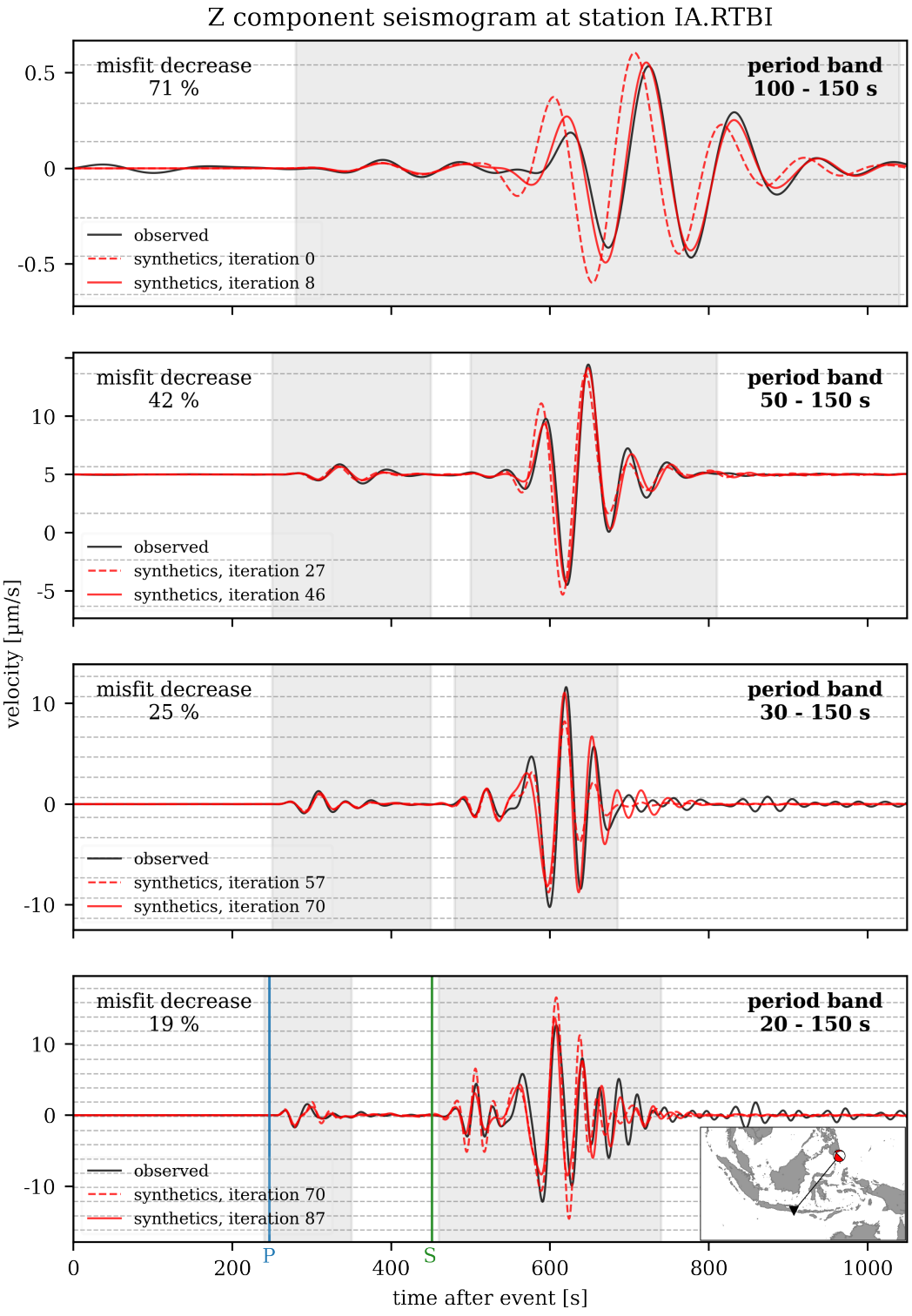


Figure 4.9: Waveform match improvement across four of the seven period bands for the vertical component of a station on Bali, Indonesia, which recorded a M_w6.2 event south of the Philippines. For each period band, the final synthetics (solid red) match the observed waveforms (black) better than the synthetics from the initial iteration (dashed red). From 50 s onwards, an additional window (grey box) 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* (Dziewoński and Anderson, 1981).

Table 4.2: Summary of the data selection displaying the number of events, seismogram traces, selected windows, average number of windows per event, percentage of traces with windows, total window length in hours, 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	∑ window length [h]	avg. window length per event [h]	# unique s-r pairs
100 – 150 s (I)	118	67,401	20,594	175	22.4	2,306	19.5	10,312
80 – 150 s (II)	118	67,317	25,614	217	27.5	2,995	25.4	11,604
65 – 150 s (III)	118	68,460	26,988	229	28.6	3,103	26.3	12,269
50 – 150 s (IV)	117	64,449	25,583	219	28.7	2,711	23.2	12,060
40 – 150 s (V)	106	58,464	32,081	302	38.1	2,586	24.4	12,960
30 – 150 s (VI)	83	44,787	26,679	321	40.9	1,519	18.3	10,279
20 – 150 s (VII)	71	38,727	22,683	319	40.6	1,064	15.0	8,656

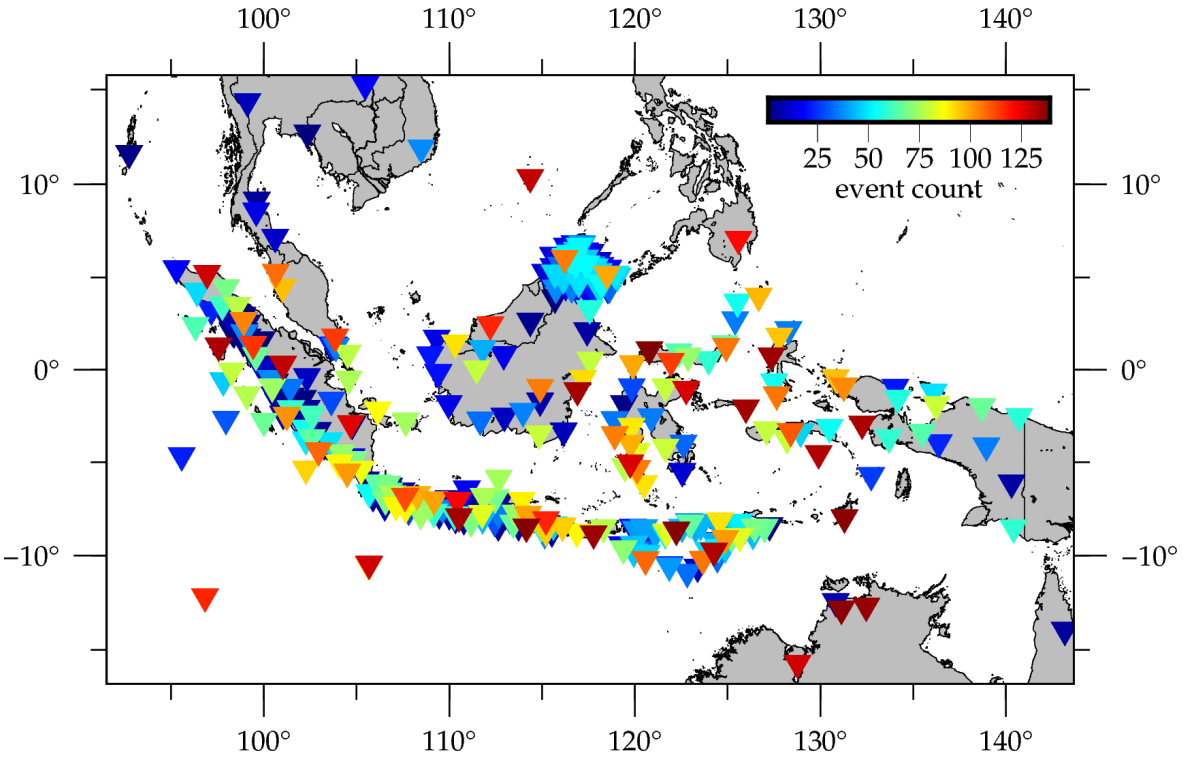


Figure 4.10: Distribution of the 440 seismic stations used in this study (inverted triangles). Colours denote the number of events for which a given station contributes waveforms to the inversion. Note that stations with a number of events < 30 are plotted in the background. These are usually temporary arrays deployed over a short period of time.

A note on the importance of data selection

Blom et al. (2020) demonstrate the importance of separating small- and large-amplitude arrivals in order to enhance depth sensitivity. As a result, they also suggest that a manual revision of selected windows is advisable. Here, I investigate the effect on the gradient of a manual window revision compared to the automatically suggested *FLEXWIN* windows. However, the results emphasise the importance of the data selection procedure in general.

I compute the initial gradient at 50 s for a subset of five events[9], which are spatially distributed as uniformly as possible and yield many recordings with high-quality waveforms. The 50 s period band is chosen because body wave arrivals become clearly identifiable in this period band, yet it is still computationally feasible to run test simulations. Figure 4.11 presents a comparison of the smoothed v_{SV} gradients for the *FLEXWIN* and manually reviewed windows for a vertical cross-section through Sulawesi, obtained using the time-frequency phase misfit function employed in this study (see Equation 3.9). The smoothed gradient is presented since it forms the basis for the model update (see Section 3.3.3). It is based on 902 *FLEXWIN* windows and 2,579 manually selected windows, resulting in a 2.5 times longer analysed window length in the latter case. The cross-section demonstrates that for the manual window selection 1) the gradient is stronger as more measurements are taken into account, and 2) finer-scale structure can be recovered, in particular at greater depths. Note that negative gradient values translate to an increase in the model parameter, and that the gradient is relative to the previous iteration only. However, the cross-section shows evidence of the subducting slab along the Indonesian volcanic arc in the south, which becomes clearer for manually selected windows.

[9]Event numbers #27, #39, #46, #91 and #96 in Section A1 of the Appendix. Figure 4.12 shows the event locations.

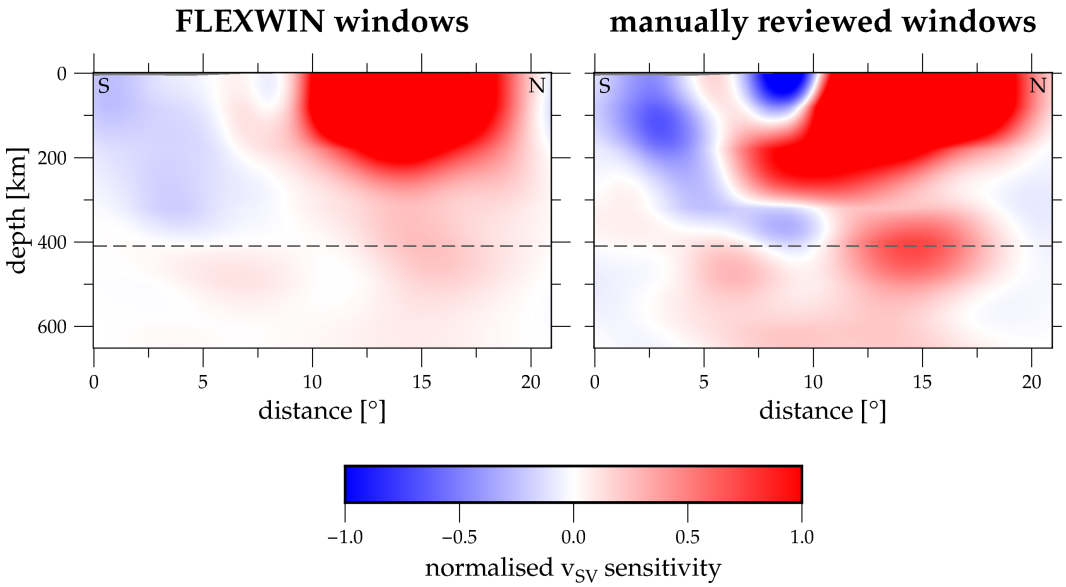


Figure 4.11: North-south cross-section of the smoothed v_{SV} gradients at 50 – 150 s for (*left*) automatically suggested *FLEXWIN* measurement windows (Maggi et al., 2009) and (*right*) manually reviewed windows. v_{SV} sensitivities are normalised to values in the right plot. The section's location corresponds to the green dotted section in Figure 5.11.

4.3 Geographical weighting

A major problem in seismic tomography is the highly uneven geographical distribution of earthquakes and seismic stations. For closely located stations, measurements (including errors) are correlated (e.g. Li and Romanowicz, 1996), which also means large portions of the data are redundant. This leads to slower convergence and oscillatory behaviour in model updates, as well as spatial bias in inversion results (e.g. Bozdağ et al., 2016). Thus, minimising the effect of dense regional networks is of particular importance in adjoint waveform tomography due to the high computational cost of the method.

In the data space, two primary strategies have been introduced to balance uneven station distributions in seismic tomography studies: Merging correlated data (e.g. Spakman and Nolet, 1988) and down-weighting correlated data (e.g. Li and Romanowicz, 1996). Here, I implement a geographical weighting based on the weighting scheme for regional to global scale adjoint waveform tomography introduced by Ruan et al. (2019). The weight w for each receiver i is computed as follows:

$$w_i^{-1} = \sum_{j=1}^{N} \exp\left[-\left(\frac{\Delta_{ij}}{\Delta_0}\right)^2\right] \tag{4.1}$$

where Δ_{ij} is the distance in degrees between each receiver pair and Δ_0 is a reference distance parameter governing the ratio of maximum to minimum weights. As suggested by Ruan et al. (2019), it is chosen such that the ratio is about one third of the largest possible ratio for all choices of Δ_0. From this equation, it follows that a station is assigned a larger weight if it has few nearby stations, and vice versa.

As described in Section 4.2.1, this study incorporates data from several dense regional networks in Southeast Asia such as the *north Borneo Orogeny Seismic Survey* (Rawlinson, 2018) and the *Banda Arc Project* (Miller, 2014). Stations from these networks should be down-weighted and carry roughly equal weights to obtain a more uniform spatial sampling and to avoid convergence problems. Following Equation 4.1, closely located stations, (e.g. in northern Borneo and on Java), have a five to ten times lower weight than isolated stations, e.g. those located on remote islands (see Figure 4.12).

To investigate the effect of the Ruan et al. (2019) station weighting, I compute the initial gradient at 50 s for the same subset of five events used in Section 4.2.5, which were recorded by the *nBOSS* and *Banda Arc Project* networks (see Section 4.2.1). Figure 4.12 demonstrates that the station weighting has a significant effect on the gradient. Again, the gradient is relative to the previous iteration only and thus, one should be careful with an analysis of any features. However, the gradient for Ruan et al. (2019) looks more balanced and geologically more reasonable, e.g. around the Banda Arc curvature. In particular, it appears that one can observe an imprint of the path coverage[10] on the gradient, if no station weighting is applied. Furthermore, implementing a geographical station weighting results in a $\sim$0.5 % larger misfit

[10]Note that this is rather an impression of the path coverage – especially at the long periods considered here (50 – 150 s) – since we do not make the assumption of ray paths.

decrease[11], confirming that it leads to faster convergence. Here, a clear effect of the station weighting on the gradient down to ∼300 km depth is observed.

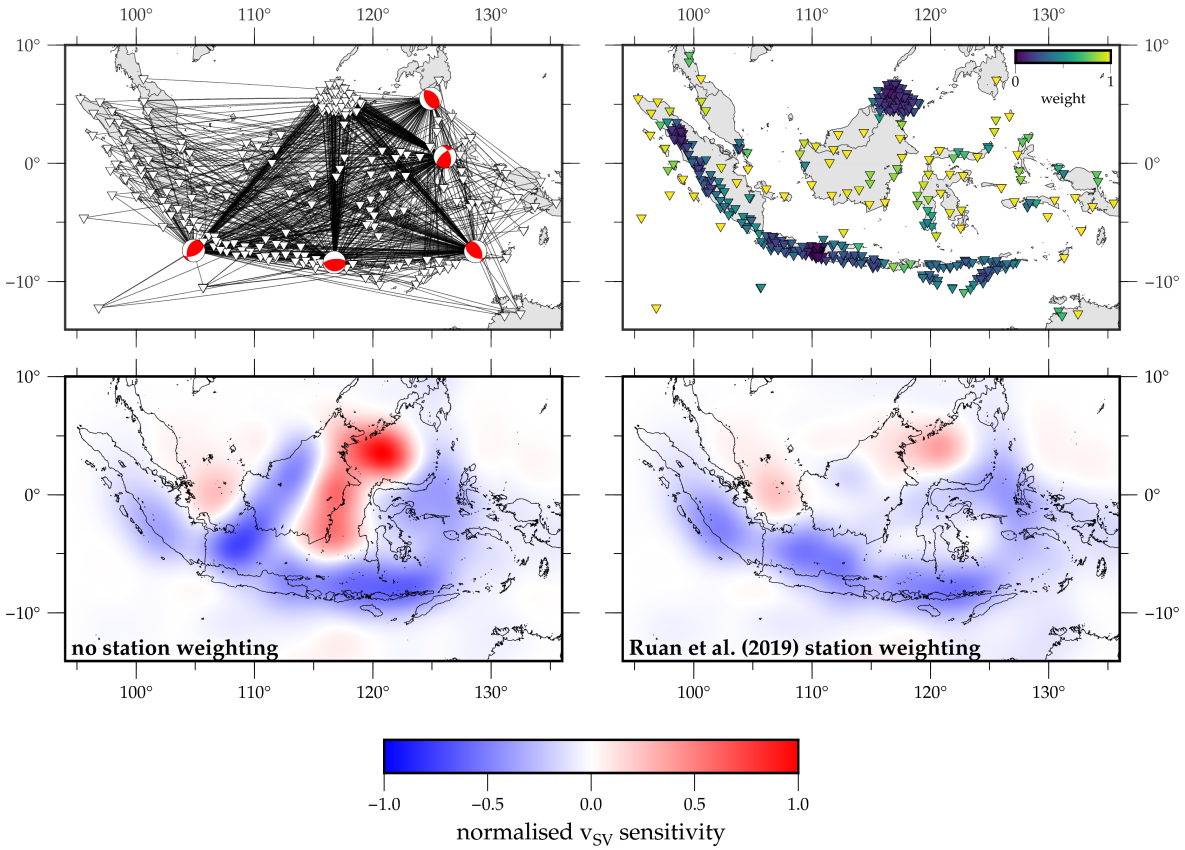


Figure 4.12: A demonstration of the effect of the geographical station weighting used throughout this study. *Top left:* Impression of the data coverage via great circle paths connecting the five events used here (event numbers #27, #39, #46, #91 and #96 in Section A1 of the Appendix) with all receivers that recorded the corresponding waveform data. *Top right:* Normalised station weights for all stations used throughout the inversion. Note that station weights are computed for each event separately and that no event was recorded by all stations. *Bottom row:* Smoothed v_{SV} gradient at 50 km depth for (*left*) no vs (*right*) Ruan et al. (2019) station weighting at 50 – 150 s. The source imprints were removed before the smoothing operator was applied (see Section 4.5).

Throughout this study, only a geographical weighting for receivers is adopted since events were already selected spatially as uniformly as possible (see Section 4.2.4). The station weights are computed for each event individually after the window picking has been carried out (see Section 4.2.5) to first identify stations that are actually used during the inversion. Another avenue to pursue could be to assign different weights to components or windows, e.g. to enhance depth sensitivity by increasing the weight of body wave windows.

4.4 Accounting for surface elevation and the ocean

The success of FWI is strongly dependent on being able to produce realistic synthetics from potentially complex models of the Earth's interior. In particular, surface topography,

[11] I observe a total misfit decrease of 3.6 and 4.1 % for no station weighting and Ruan et al. (2019) station weighting, respectively.

bathymetry and the ocean can have a significant effect on ground motion; for example, surface topography can lead to seismic wave scattering (e.g. Lee et al., 2008) and the fluid ocean elongates the Rayleigh surface wave train (Todoriki et al., 2016; Fernando et al., 2020). However, these effects are not routinely accounted for in adjoint waveform tomography, largely due to the need for sophisticated meshing techniques, plus the added computational burden associated with simulating waves in the relatively low-velocity fluid ocean.

In this study, large topographic contrasts are present; several mountain ranges with peaks over 4,000 m high can be found within the region, the tallest being Puncak Jaya (4,884 metres) on Papua, Indonesia. At the opposite end of the scale, the deepest point within the region – the so-called Galathea Depth – is located along the Philippine Trench at 10,540 metres below sea level. In the following, the implementation of surface elevation (topography and bathymetry) and the ocean in this study is discussed. Furthermore, shorter period seismic waves are expected to be particularly sensitive to variations in surface elevation and the fluid ocean due to increased sensitivity to the upper few kilometers. Consequently, I investigate their frequency-dependent effects on synthetic waveforms.

4.4.1 Surface topography and bathymetry

Recent studies have concluded that neglecting surface topography can result in severe velocity artefacts if topographic fluctuations have wavelengths on the order of the minimum seismic wavelength (e.g. Nuber et al., 2016; Espindola-Carmona and Peter, 2018). However, realistic topography and bathymetry are not routinely accounted for, mainly due to the difficulty of incorporating them in finite-difference methods. As described in Section 3.2.4, the spectral-element method, on the other hand, provides a greater flexibility in terms of meshing and thus, includes a natural ability to handle complex topographic media (Komatitsch and Tromp, 1999).

Topography and bathymetry are implemented by distorting the upper layer of elements, which requires advanced meshing techniques, but it does not lead to a considerable increase in computational cost. Thus, this study implements the global relief model *Earth2014* (Hirt and Rexer, 2015) at all period bands, which yields the surface elevation and ocean depth relative to the *WGS84* ellipsoid with an approximate spatial resolution of ten kilometres. However, it should be noted that topography is only meshed to accommodate the maximum frequency considered within each period band.

Figure 4.13 presents the frequency-dependent effect of surface topography and bathymetry on synthetic waveforms. For this simulation setup, implementing surface elevation can result in a considerable phase advance and change in amplitude of the surface wave train at periods $\leq$ 20 s for this simulation setup. Further three-component waveforms at 15 s and synthetic tests investigating path-dependent effects are presented in Section 6.2.

An L_2 gradient for 20 – 150 s period shows a clear effect of topography confined mostly to the near surface but with notable sensitivity down to depths of about 200 – 300 km. Thus, neglecting topography can lead to artefacts at greater depths, which was previously suggested in a synthetic near-surface study by Nuber et al. (2016).

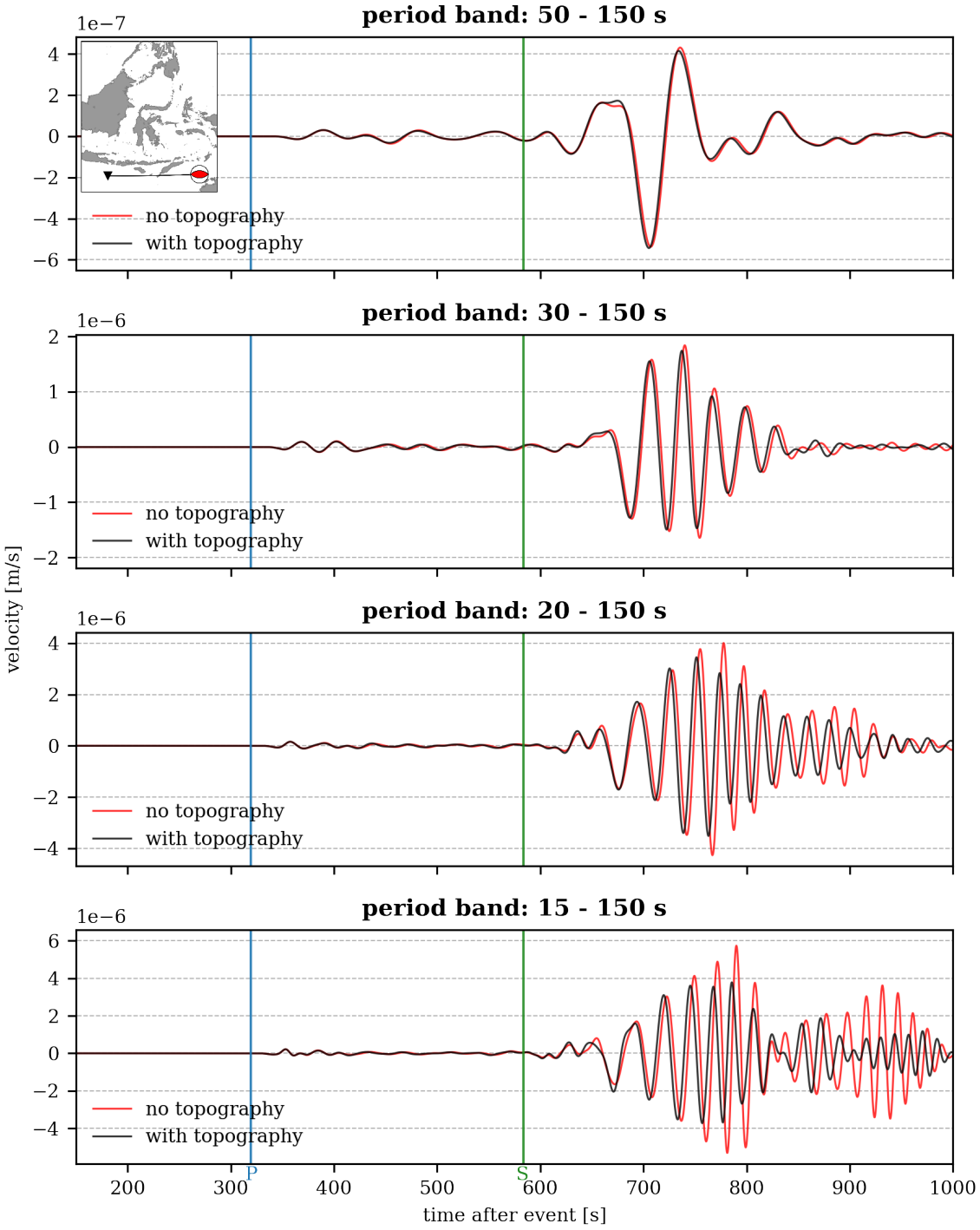


Figure 4.13: North component of a single source-receiver pair (Δ 25°) illustrating the effect of topography via synthetic waveforms at 50, 30, 20 and 15 s period for a synthetic M6.0 event in the Banda Arc region. Background model: *PREM* (Dziewoński and Anderson, 1981) with *CRUST1.0* (Laske et al., 2013). Vertical lines indicate predicted P- (blue) and S-wave (green) first arrival times obtained from the *TauP* toolkit (Crotwell et al., 1999) for *PREM* (Dziewoński and Anderson, 1981).

4.4.2 Ocean effect

Simulating the fluid ocean is a greater challenge and its effect on 3-D seismic wave propagation has been the subject of several recent studies[12]. Todoriki et al. (2016) conclude that conversions at the sea floor and attenuation in the sea water resemble a low-pass filter and extend the duration of Rayleigh waves. An et al. (2017), Yue et al. (2017) and Fernando et al. (2020) observe water-reverberation phases at characteristic periods shorter than 20 s. An et al. (2017) further demonstrate that such phases are more efficiently excited by near-trench earthquake ruptures due to sharp bathymetry gradients. Fernando et al. (2020) include a global ocean layer of constant thickness in *AxiSEM3D* (Nissen-Meyer et al., 2014; Leng et al., 2016; Leng et al., 2019) in order to simulate water-depth phases at shorter periods and show that this improves the waveform match for island stations. However, they emphasise that *AxiSEM3D* currently does not support arbitrarily patched oceans and thus, path-dependent effects remain unclear.

For our purposes, two different methods are available to account for the fluid ocean:

1. The ocean can be approximated by the weight of its water column, which means the ocean does not need to be meshed explicitly. This reproduces the ocean with good fidelity, provided the thickness of the ocean layer is small compared to the wavelength (Komatitsch and Tromp, 2002). This method is referred to as "ocean loading" (or "water column approximation") and is implemented via a modification of the mass matrix.
2. The volume of the ocean can be replaced by acoustic elements to represent it as a fluid medium. This allows reverberations to be simulated, but it requires solving a coupled system of the elastic (see Equation 3.4) and acoustic (see Equation 3.6) wave equations, which is computationally more expensive.

In fact, a combination of the methods above is implemented in *Salvus* when accounting for the fluid ocean, where deep ocean areas are modelled with acoustic elements, while the fluid ocean is replaced by an equivalent load when the ocean becomes too shallow (here: < 1.5 km). The seismic properties of the fluid ocean are set to $v_P = 1{,}450$ m/s, $v_{SH} = v_{SV} = 0$ and $\rho = 1{,}020$ kg/m^3.

Figure 4.14 shows that approximating the fluid ocean by its weight is a valid assumption at the periods considered in this study. At shorter periods (≤ 20 s), seismic waves within the water layer play a crucial role and these reverberated phases should not be neglected. This will be discussed in more detail in Chapter 6, where the fluid ocean is simulated explicitly. Further three-component waveforms at 15 s and synthetic tests investigating path-dependent effects are presented in Section 6.2.

[12]For a more detailed introduction to this topic, I refer the interested reader to Fernando et al. (2020).

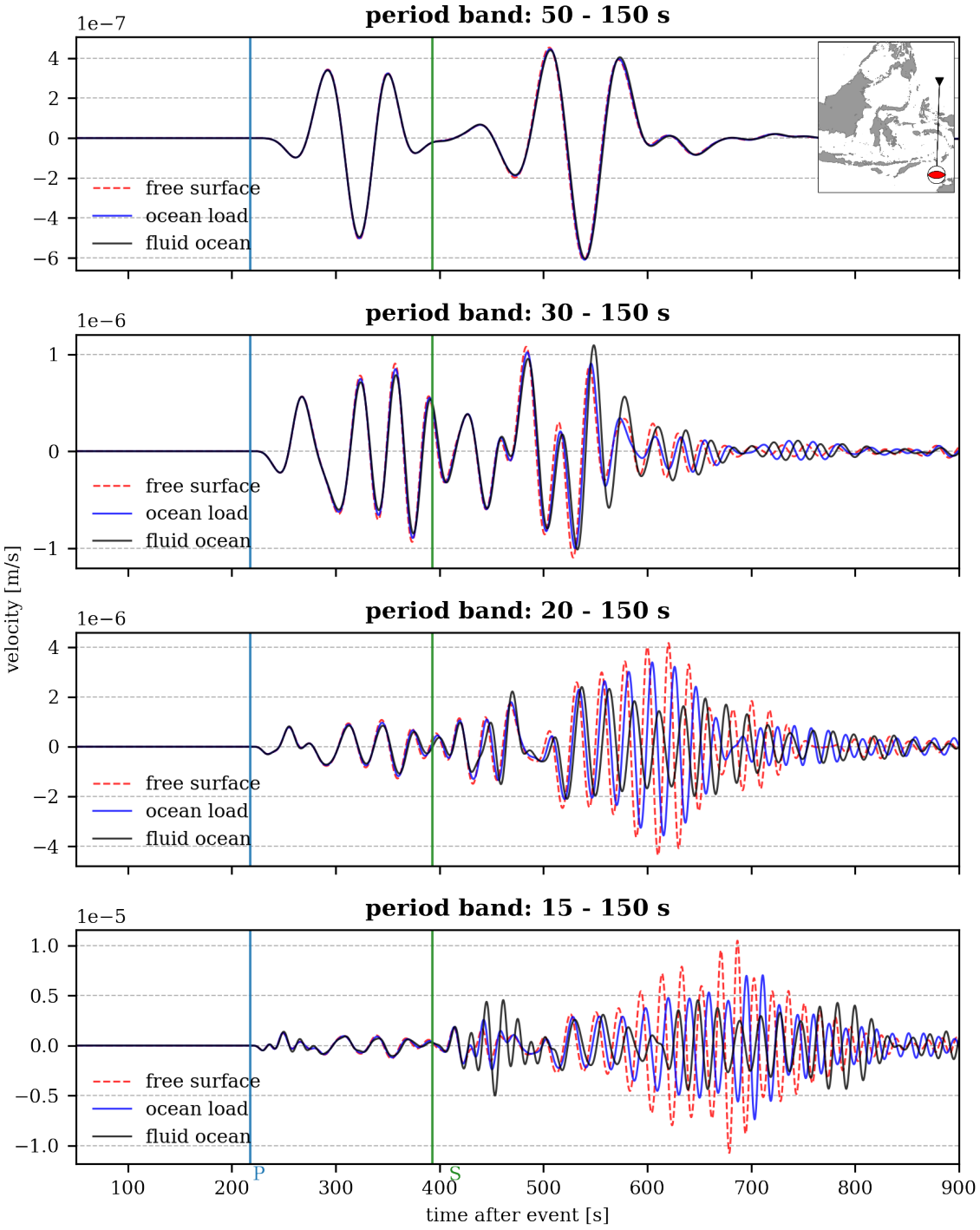


Figure 4.14: North component (here: radial) of a single source-receiver pair (Δ 15°) illustrating the effect of the fluid ocean via synthetic waveforms at 50, 30, 20 and 15 s period for a synthetic M6.0 event in the Banda Arc region. Background model: *PREM* (Dziewoński and Anderson, 1981) with *CRUST1.0* (Laske et al., 2013). Vertical lines indicate predicted P- (blue) and S-wave (green) first arrival times obtained from the *TauP* toolkit (Crotwell et al., 1999) for *PREM* (Dziewoński and Anderson, 1981).

4.5 Gradient preconditioning

The following two processing steps are applied to the raw gradient(s) before a model update is computed. These are regularisation techniques to mitigate the ill-posedness of the inverse problem (see Chapter 3.3). Furthermore, they can provide significant overall computational savings by accelerating the convergence of the optimisation algorithm (e.g. Modrak and Tromp, 2016; Liu et al., 2022). Note that the preconditioner does not affect the location of the minimum we seek (see Section 3.3.3), but the path the iterations take in the model space.

4.5.1 Source imprint removal

Event kernels usually show large sensitivities around the source region, with values several times higher than the surrounding area in this study. The focused sensitivity is a result of all energy coming together in this zone and the fact that an update in this region will have an effect on all waveforms. Modrak and Tromp (2016) argue that these near field features result from inaccuracies in the numerical treatment of wave propagation in the close vicinity of sources and receivers. Furthermore, errors in the source parameters have concentrated effects in the source region (e.g. Blom et al., 2022). Thus, these imprints have to be removed in order to avoid a strong localisation of model updates. Some studies employ clipping of gradients at a certain percentile (e.g. Blom et al., 2020), but smoothing is regarded as a better alternative since it can mathematically be implemented in optimisation theory, while clipping cannot (Christian Böhm, pers. comm., March 2020).

The size of the source imprint is event- and frequency-dependent and can vary between parameters. Furthermore, it increases with the number of measurements made (see Section 4.2.5). For each period band, I manually investigate several event kernels at hypocentral depth to quantify the volume that needs to be removed. The maximum radius is determined, and a sphere with this radius is removed for all events and all inversion parameters (as illustrated in Figure 4.15). I favour 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. If the imprints are removed before summing the event gradients, sensitivity around source regions can be obtained from other events.

In this study, the chosen source imprint radius decreases from 500 km to 300 km across all period bands (100 – 150 s to 20 – 150 s) as shown in Table 4.3. Furthermore, I remove a sphere with a radius of 50 km for each receiver (e.g. the 20 s gradient in Figure 4.15 shows receiver imprints around Java, Indonesia). However, receiver imprints are usually small (even at the surface) and likely removed by the smoothing operator described in the next section.

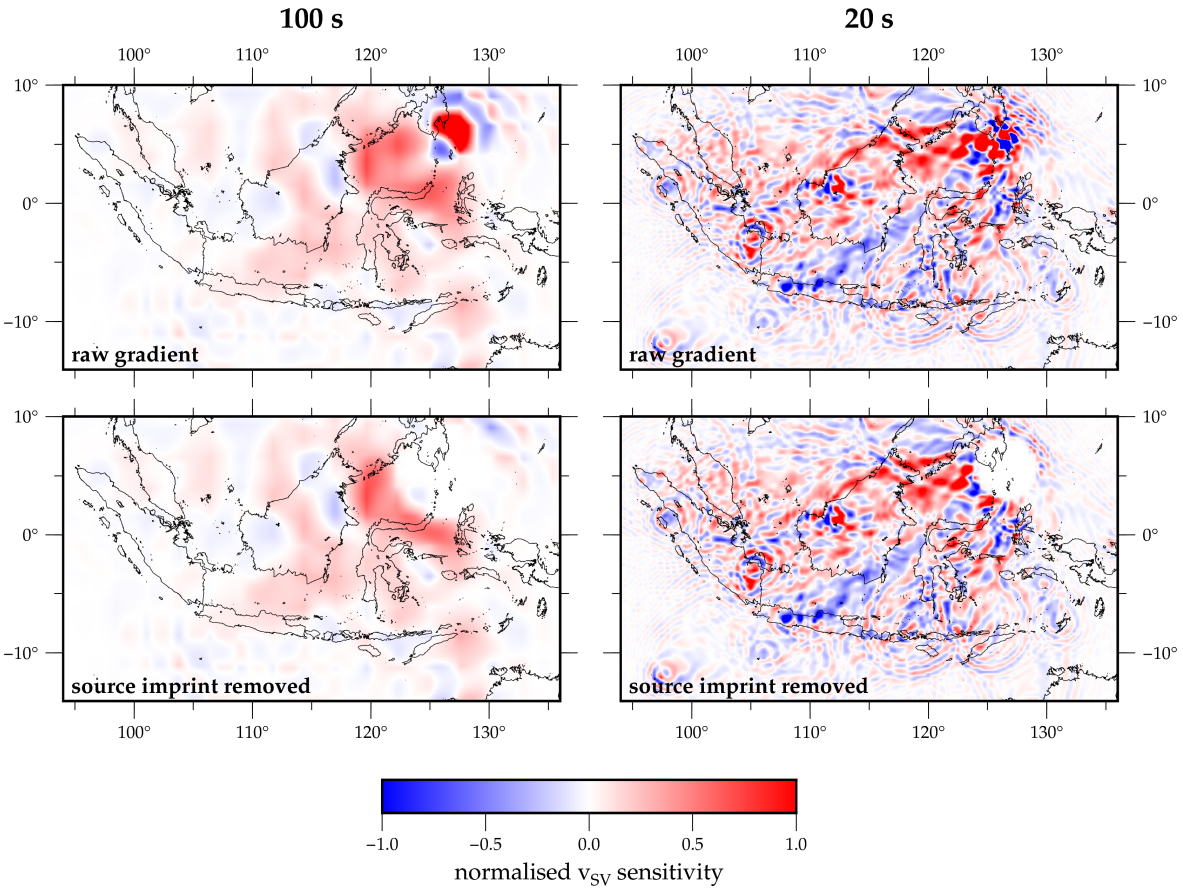


Figure 4.15: Source imprint removal for a M_W6.3 event southeast of the Philippines at (*left*) 100 – 150 s and (*right*) 20 – 150 s for a depth slice at 75 km, which is the event's hypocentral depth taken from *GCMT* (Ekström et al., 2012). The upper row shows the raw v_{SV} kernel, and the bottom row shows the event kernel after the source imprint has been removed. Note the radius decrease of the source imprint removal and the overall smaller scale structure as shorter periods are considered. The sensitivities are normalised 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.

4.5.2 Smoothing

After the source imprint for each event has been removed, the event kernels are summed to produce the misfit gradient. To avoid the appearance of small-scale artefacts, a smoothing operator is applied to the summed gradient. Note that the gradient is smoothed and not the model itself. In this study, I apply *Salvus*' anisotropic, depth-dependent, diffusion-based smoothing operator following Böhm et al. (2019), which is briefly described below:

Anisotropic smoothing Here, anistotropic smoothing means it is smoothed less in the vertical direction because it is assumed that the model varies significantly more in this direction and there are usually better constraints in the vertical direction.

Depth-dependent smoothing Elastic parameters vary with depth and so do wavelengths, e.g. seismic velocities are usually higher at greater depths. Thus, smoothing lengths can be adapted to the local wavelengths of the model, i.e. it is smoothed less at shallower depths.

Diffusion-based smoothing Common smoothing methods involve Gaussian filters, but they are incompatible with many spectral-element solvers that are currently used (e.g. Böhm et al., 2019). Thus, *Salvus* implements a diffusion-based smoothing operator, which employs the smoothing effect of the heat equation[13].

Initial model updates (100 – 65 s) use an anisotropic smoothing operator (no depth-dependent smoothing; horizontal and vertical smoothing lengths are fixed across the model). The smoothing lengths decrease from 450 to 300 km horizontally, and from 100 to 65 km vertically, which are comparable values to those employed in other studies (e.g. Krischer et al., 2015). From 50 s downwards, depth-dependent smoothing is applied in order to account for the local wavelengths of the model. The smoothing lengths range from 1.0 – 0.5 λ_{min} in the horizontal and 0.2 λ_{min} in the vertical direction. The respective wavelengths are based on the shear-wave velocity of the initial model in each period band. An overview of the smoothing lengths per period band can be found in Table 4.3 and an example of a smoothed gradient is presented in Figure 4.16. Note that some period bands are divided into two legs with decreasing smoothing lengths to mitigate the drawback of losing too much structural detail. However, decreasing the smoothing length is associated with a reset of the inverse Hessian approximation since it is associated with a change in the inverse problem.

Table 4.3: 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 downwards, a depth-dependent, anisotropic, diffusion-based smoothing (DD) is used.

period band	smoothing type	smoothing lengths – horizontal, vertical	source imprint removal [km]
100 – 150 s (Ia)	PA	450, 100 km	500
100 – 150 s (Ib)	PA	375, 100 km	500
80 – 150 s (IIa)	PA	375, 80 km	450
80 – 150 s (IIb)	PA	300, 80 km	450
65 – 150 s (III)	PA	300, 65 km	400
50 – 150 s (IVa)	DD	1.0, 0.2 λ_{min}	350
50 – 150 s (IVb)	DD	0.75, 0.2 λ_{min}	350
40 – 150 s (V)	DD	0.5, 0.2 λ_{min}	300
30 – 150 s (VI)	DD	0.5, 0.2 λ_{min}	300
20 – 150 s (VII)	DD	0.5, 0.2 λ_{min}	300

A number of best practices have been established on how to choose smoothing lengths, which I will elaborate on in the following. Stronger smoothing tends to avoid local minima and small-scale artefacts, but it slows down the reduction of the misfit since structural detail is lost. A strategy that seems to work well in practice is to start with stronger smoothing and to reduce it when necessary (Barnier et al., 2018). This further implies that the longest absolute smoothing length within a period band should not be larger than the previous ones, e.g. one wavelength at 20 s is larger than half a wavelength at 30 s. This will likely result in rejected trial models since longer-wavelength structure was already obtained during previous

[13] An illustrative example can be found in *Why Blurring an Image is Similar to Warming Your Coffee* (Rome, 2016).

iterations. A rule of thumb to estimate a suitable smoothing length s_{II} for the subsequent period band T_{II} can be expressed as

$$s_{\mathrm{II}} \leq \frac{s_{\mathrm{I}} T_{\mathrm{I}}}{T_{\mathrm{II}}} \tag{4.2}$$

where s represents the smoothing length (e.g. 400 km or $0.5\lambda_{\mathrm{min}}$) at the period band T in seconds. I and II represent the longer and shorter period bands, respectively.

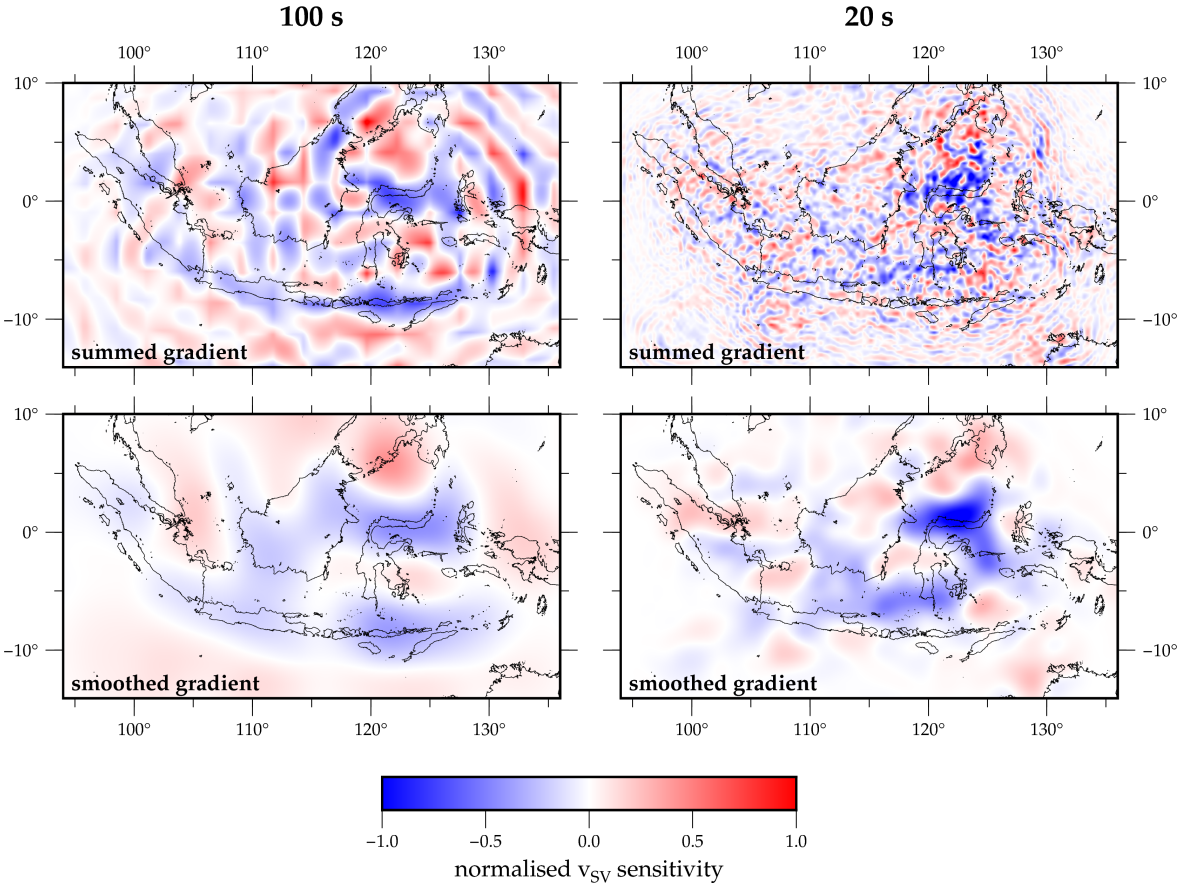


Figure 4.16: An example showing the effect of smoothing the misfit gradient, shown for (*left*) 100 – 150 s and (*right*) 20 – 150 s for a depth slice at 200 km. The upper row shows the summed v_{SV} gradient after the source imprint has been removed (see Section 4.5.1), while the bottom row shows the smoothed gradient. The sensitivities are normalised per period band since the gradients at 100 and 20 s vary by two orders of magnitude. Note the increased sensitivity to smaller scale structure as the period is decreased.

4.6 Inversion parameters

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) and radially anisotropic S-wave velocity (v_{SH} and v_{SV}). Furthermore, it is advisable to invert for 3-D density variations in order to avoid artefacts (Blom et al., 2017). For P-wave structure, I invert for an isotropic parameter since the v_{PH} gradient is one order less than v_{PV} and thus, it has a second-order effect. This means the P-wave anisotropy parameter η is kept constant at one.

It is possible to explore additional inversion parameters (such as attenuation, e.g. Xing and Zhu, 2021), but this holds the risk of overfitting, i.e. starting to fit noise. In Southeast Asia, this would not have been feasible due to the limited (azimuthal) data coverage (see Section 4.2.1), which would be unable to resolve the increased trade-offs caused by the introduction of another class of free parameter.

It should be noted that the simulation parameters usually differ from the inversion parameters, e.g. the forward problem is solved for a different set of parameters (e.g. v_{PV}, v_{PH}, ...) than the ones inverted for (e.g. v_P, ...). Furthermore, it is also important how the inversion parameters are updated, e.g. as absolute or relative deviations from a background model. Here, I use relative deviations from the initial model within each period band, which circumvents issues with differences between magnitudes and units of different parameters (e.g. P- vs S-wave velocity and velocity vs density, respectively) and can lead to faster updates at greater depths (Christian Böhm, pers. comm., May 2020).

4.7 Synthetic tests

Synthetic tests can be an intuitive way of assessing the ability of a given source-receiver distribution to recover seismic structure. In this section, I will describe how to set up an efficient test simulation, run a synthetic recovery test for the Southeast Asian study region and present preliminary results from exploring cross-talk between inversion parameters for the simulation setup described in this chapter.

4.7.1 How to reduce the computational cost for test simulations

Reducing the computational cost of test simulations in the context of *Salvus* (release 0.11.44) could involve one or more of the following:

- 2-D instead of 3-D.
- Consider longer period data and a small event catalogue.
- Do not store adjoint checkpoints if the gradient is not computed.
- Decrease the number of mesh elements per λ_{min} (e.g. to 1.0), which reduces the total number of elements.
- Neglect attenuation for purely synthetic studies.
- Use a smaller domain extension, incl. mesh depth.
- A shorter wavefield simulation length reduces the total number of time steps computed.
- Reduce the absorbing boundary width (see Section 4.1.2), which reduces the total number of mesh elements.
- Some misfit calculations are computationally cheaper than others, e.g. a cross-correlation time shift or L_2 misfit evaluation is faster than computing time-frequency phase misfits (see Section 3.3.1).

4.7.2 Checkerboard recovery test

In this section, I explore whether we can recover a 3-D checkerboard pattern of positive and negative Gaussian anomalies ($\pm$ 10 %) with a radius of 150 km in v_{SV}, using the Southeast Asian station distribution shown in Figure 4.4 and the inversion setup described throughout this chapter. Synthetic waveforms obtained from the model with the checkerboard pattern serve as observed data, with 5 % normally distributed noise added to it. Due to the high computational cost of FWI, I only use a set of five synthetic sources (see Figure 4.17), which are spatially evenly distributed, while reflecting a realistic event distribution for the region (see Figure 4.7). On the one hand this means the recovery test is in some sense a worst-case scenario, on the other hand I assume that all stations in Southeast Asia contribute useful waveforms to the inversion. Starting off from the 1-D *CSEM* (see Figure 4.2), I run ten iterations at 50 – 150 s and ten iterations at 20 – 150 s period ranges to emulate the multi-scale approach applied throughout this study (see Section 3.3.3).

Figure 4.17 presents the results of the synthetic recovery test at depths of 100, 300 and 500 km. This shows that we are able to recover high- and low-velocity anomalies, and have good constraints around the Sundaland block, Sulawesi, Banda Arc and south of the Philippines, keeping in mind the small event catalogue and few iterations carried out. However, there is little constraint north of $\sim$5° latitude and south of $\sim$-10° latitude but we are able to recover the sign of the input perturbation around northern Australia in the southeast. It is unlikely that the region of good recovery in the north-south direction would substantially increase if the full set of earthquakes were used (see Figure 4.7) in the synthetic test. Note that sensitivity and hence the ability to recover features decreases with depth, as expected from a surface wave dominated method (and because measurements are made at the Earth's surface). Again, if the full set of earthquakes were used to generate the synthetic dataset, it is unlikely that the depth resolution would be significantly increased. The results for input perturbations in the other inversion parameters are presented in Section A2 of the Appendix, which show that input perturbations in v_{SH} can be recovered but v_P and ρ suffer from significant cross-talk as described in the next section.

4.7.3 Parameter cross-talk

The tests described in Section 4.7.2 allow me to explore how a perturbation in one parameter trades off with other parameters throughout the inversion. For example, Figure 4.17 shows that a v_{SV} anomaly also results in updates in v_{SH} and density. The results of parameter trade-offs are summarised in Table 4.4; for perturbations in v_{SV} and v_{SH}, the strongest update is in the same parameters, as desired. However, P-wave and density structure get mapped into S-wave structure, and result in an overall misfit decrease of $\sim$20 – 25 % compared to $\sim$40 % for shear-wave structure. Thus, I conclude that we can only constrain shear-wave structure reliably with the given dataset and at the periods considered throughout this study.

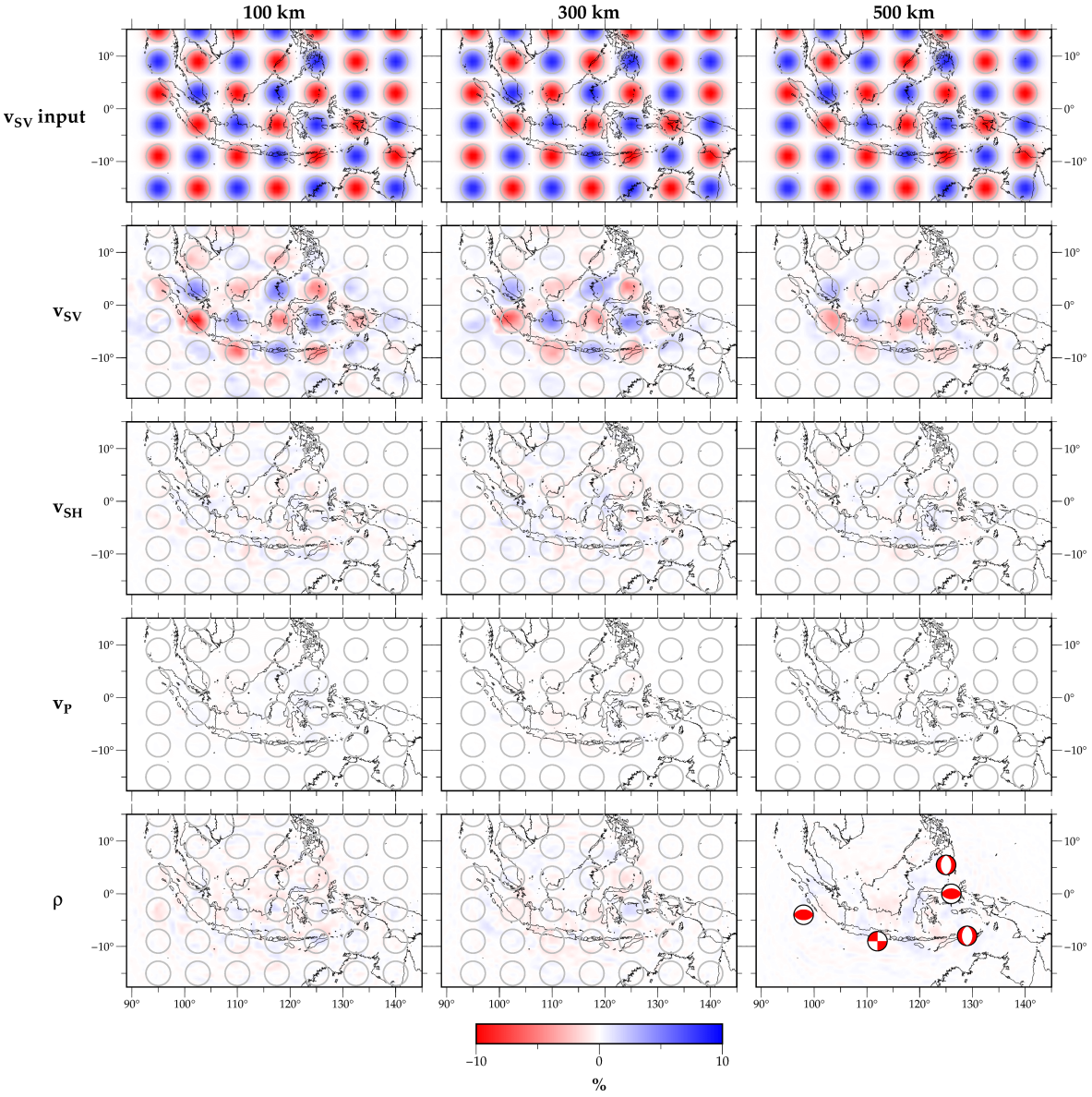


Figure 4.17: Depth slices for a checkerboard recovery test. *Top panel:* Input perturbation of $\pm$ 10 % in v_{SV}. *Panels below:* The final model after 20 iterations for all inversion parameters (v_{SV}, v_{SH}, v_P and density ρ). Perturbations are in % and relative to the 1-D starting model (*CSEM*). Grey circles indicate the contours of the input perturbation. The columns indicate depths at (*left*) 100 km, (*middle*) 300 km and (*right*) 500 km. The bottom right panel shows the focal mechanisms for the five events (M_w6.0) used for this test.

Table 4.4: Summary of cross-talk in anisotropic synthetic tests. Moderate and weak updates correspond to values of ∼50 % and ∼10 % of the strongest update or input perturbation, respectively.

perturbed parameter	**strongest update**	**moderate update**	**weak update**
$\boldsymbol{v_{SV}}$	v_{SV}	v_{SH}, ρ	v_P
$\boldsymbol{v_{SH}}$	v_{SH}	v_{SV}, ρ	v_P
$\boldsymbol{v_P}$	–	v_{SV}, v_{SH}, v_P, ρ	–
ρ	–	v_{SV}, v_{SH}, ρ	v_P

5 Southeast Asian waveform tomography: Results – *SASSY21*

In Chapter 4, I presented the inversion setup for my Southeast Asian adjoint waveform tomography study. The final model was obtained after 87 trust-region based L-BFGS model updates and reveals detailed anomalies down to the mantle transition zone. In this chapter, I will discuss the inversion performance and assess the robustness of the final model. Furthermore, I highlight and discuss some of the key features of the final model, including a detailed comparison with other recent tomographic models of the region. Henceforth, I will refer to the final model of this study as *SASSY21*. Many of the results presented in this chapter are also contained in the publication by Wehner et al. (2022), which was recently published in the Journal of Geophysical Research: Solid Earth.

5.1 Inversion performance

Beginning from a 1-D reference model (see Section 4.1.1), a total of 87 L-BFGS inversion iterations divided over seven period bands that range from 100 – 150 s down to 20 – 150 s were carried out following a multi-scale approach, where higher frequency content is successively added (see Section 3.3.3). Thus, I take advantage of the relatively low computational cost and the broad sensitivity of seismic signals at these long periods to obtain long-wavelength structure. The inversion process was performed on the University of Cambridge Research Computing Services' supercomputer[1] and required $>$ 50,000 CPU hours, half of which were used during the final period band (20 – 150 s), which can be attributed to the denser wavefield sampling at shorter periods (see Section 3.2.1).

The criteria for adding shorter period data are based on when the misfit decrease stagnates or the number of events that decrease their misfit significantly drops (below $\sim$70 %). Each broadening of the period band is accompanied by a mesh interpolation (see Section 4.1.4) and review of events and measurement windows (see Section 4.2.4 and Section 4.2.5, respectively). It should be noted that some studies prefer to run many iterations (e.g. 256 iterations over five period bands for the western United States by Rodgers et al., 2022). I have run several additional iterations beyond the 87 iterations presented here and the misfit decrease is not significant ($<$ 0.5 %). However, suspicious smaller-scale features start to appear and it is suspected that this is when overfitting becomes apparent.

[1] See *Acknowledgements* for more details.

5.1.1 Misfit development

The misfit development for all seven period bands used in this study is presented in Figure 5.1. For comparison, a decade ago, it was common to a run a few conjugate gradient iterations (see Section 3.3.3) resulting in a misfit decrease of a few per cent (e.g. < 5 % across eleven iterations in Fichtner et al., 2009), which emphasises the increase in computational resources and advances in FWI over the past decade. In this study, the overall misfit decrease is remarkable, which I mainly attribute to the 1-D starting model leaving a lot of room for improvement. The first period band yields the greatest misfit decrease of 42 %. 29 % 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).

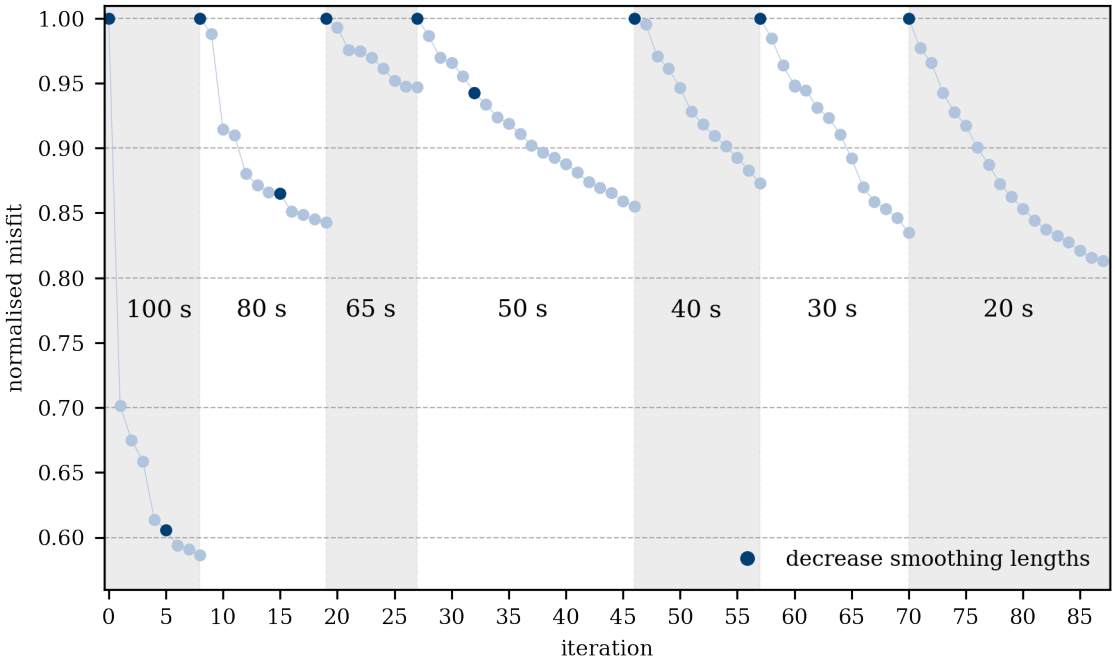


Figure 5.1: Misfit development across 87 iterations, normalised by the initial misfit within each period band. Throughout this study, time-frequency phase misfits are used (see Section 3.3.1). Dark blue dots indicate a smoothing length decrease (see Section 4.5.2). Each broadening of the period content is accompanied by a mesh interpolation and data review (events and windows).

For many period bands (e.g. 80 – 150 s, 40 – 150 s), a strong misfit decrease for the second iteration is observed, which may be the result of the trust-region based L-BFGS optimisation scheme used in this study (see Section 3.3.3). 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. This tends to speed up convergence and spare computational resources since repeated forward evaluations as a result of small step lengths can be avoided. Here, the Hessian approximation

is based on the history of the past ten gradients since FWIs using these scales and period bands are relatively convex and thus, the change in curvature between iterations is small. Note that this is not the case at e.g. exploration scales, where the problem is highly non-convex (Christian Böhm, pers. comm., November 2021).

I only encountered rejected trial models (that is, a model resulting in a higher overall waveform misfit than the previous model) during the first iteration within a period band when the smoothing lengths (see Section 4.5.2) and/or the initial step length were too large. Equation 4.2 can be used to compute suitable smoothing lengths and I find that decreasing the initial step length once the wavefield becomes significantly more complex (typically at 30 s) helps to avoid overshooting. However, Modrak and Tromp (2016) mention that rejected trial models are not unexpected even in between period bands, especially in applications involving multi-parameter inversion or noisy data.

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 Figure 5.2). In the following section, I will validate the final model in more detail.

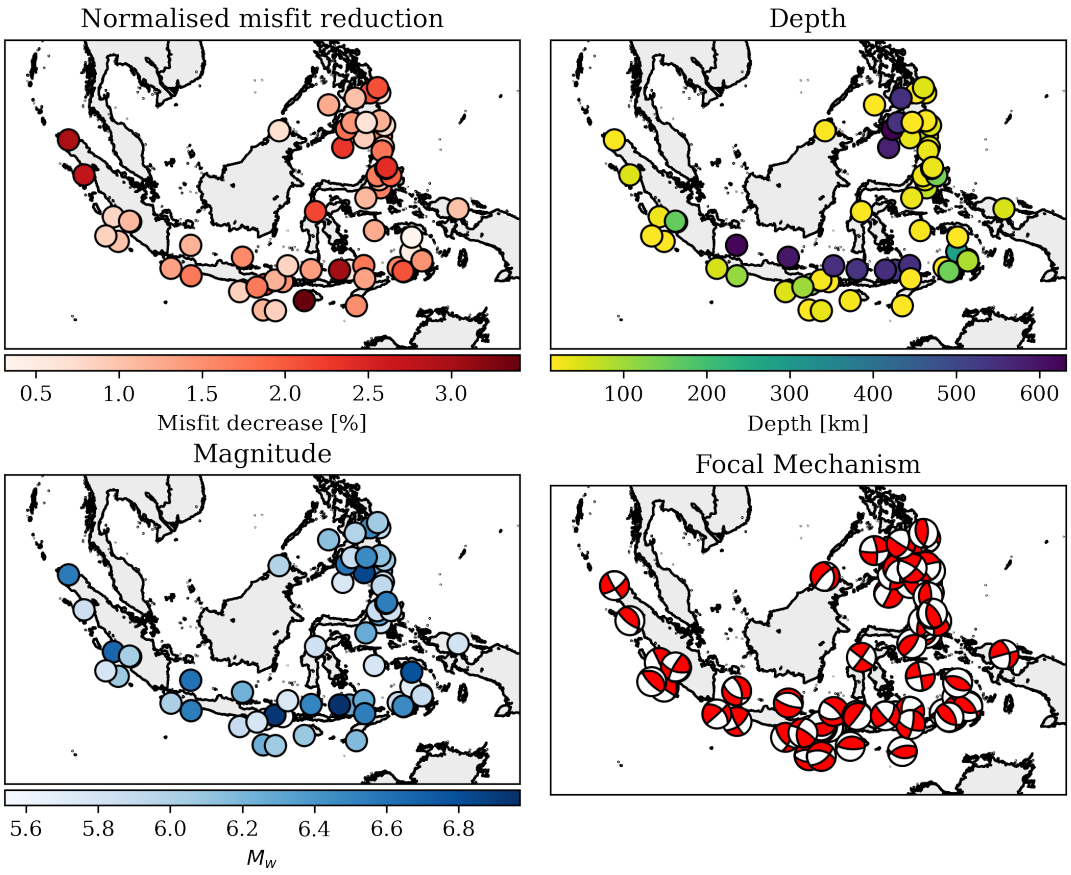


Figure 5.2: Overview of the 71 events used in the final period band (20 – 150 s). *Top left:* The normalised event misfit reduction for each event between the starting and final model. *Top right:* Events coloured by depth. *Bottom left:* Events coloured by magnitude. *Bottom right:* Focal mechanisms taken from *GCMT* (Ekström et al., 2012).

5.2 Model validation

In the following, I will assess the robustness of *SASSY21*. As described in Section 3.4, the checkerboard test is popular in traditional ray-based tomography (e.g. Rawlinson and Spakman, 2016), but it is currently 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., 2022). However, one can pursue more data-driven approaches towards validating the final model, e.g. by evaluating its ability to fit independent data and investigating trade-offs among parameters by analysing the Hessian-vector product for a test function, as described below.

5.2.1 Waveform match improvement

The misfit development described in the previous section is entirely driven by an improvement in the match between observed and synthetic waveforms. Figure 4.9 presents the waveform comparison across four of the seven period bands for the vertical component of a single source-receiver pair. The long-period seismograms show a similar pattern across all receivers, which is expected since they are sensitive to the overall structure of the region due to the broad sensitivity of seismic signals at these periods. Here, 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. However, *CSEM* (see Section 4.1.1, Fichtner et al., 2018) is deemed to be an acceptable starting model for Southeast Asia since it still matches the longest-period data (100 – 150 s) to within half a cycle.

Several three-component waveform fits for the initial synthetics, final synthetics and observed waveforms for a variety of events are presented in Figure 5.3. The waveform match improvement is associated with an overall misfit decrease of $\sim$54 %. However, this is indicative since the 1-D starting model is not able to properly explain the observed waveforms – in particular the elongated surface wave train – at 20 s. The initial synthetics show a spike but no elongated surface wave train. I suspect this to be the Airy phase for Rayleigh waves (e.g. Pekeris, 1948; Bormann et al., 2012), which is typical for 1-D models, while 3-D crustal models can result in a longer surface wave train. This emphasises the importance of the multi-scale approach (see Section 3.3.3) but makes a meaningful misfit comparison to observed waveforms difficult, which is further complicated by the varying windows and number of events per period band.

In order to avoid bias related to windowing of the data (e.g. Thrastarson et al., 2022), I evaluate the L_2 (see Section 3.3.1) misfit for the starting model (*CSEM*) and the final model (*SASSY21*) at the final period band (20 – 150 s) on the full trace for all traces used throughout the inversion[2]. This results in a misfit decrease of $\sim$18 % for all 71 events used in the final period band. However, the full-trace L_2 misfit decrease highly depends on the noise level, and can vary between 9 % (for all available traces) and 86 % (for a handful of traces with an exceptional signal-to-noise ratio). Thus, the absolute misfit decrease is indicative but it is reassuring that a misfit decrease is observed, even for the worst-case scenario.

[2]That is, one or more windows were selected for this trace.

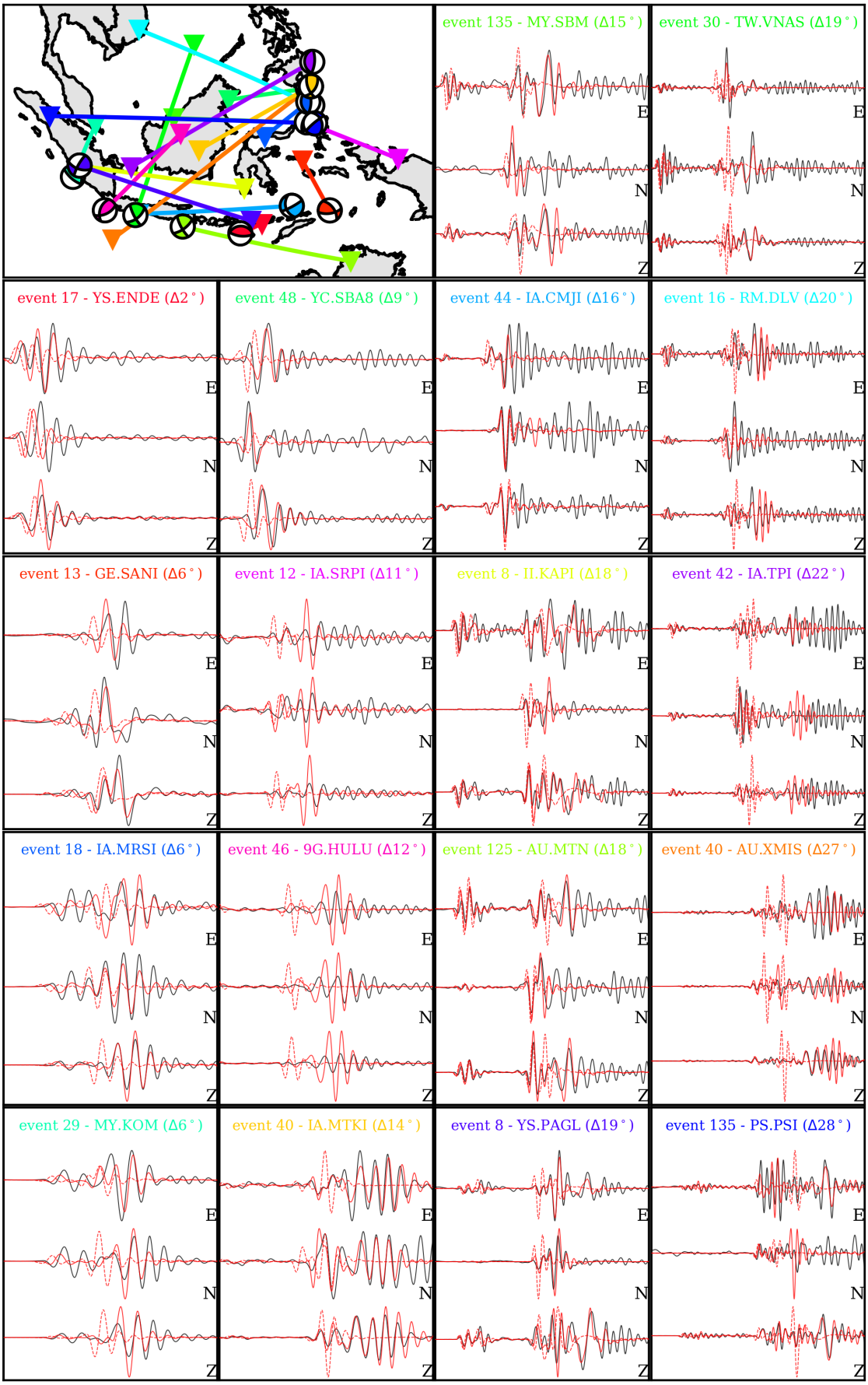


Figure 5.3: Three-component waveform match between the initial model (iteration 0, dashed red), the final synthetics (iteration 87, solid red) and observed waveforms (black) for 18 source-receiver pairs. The event numbers are taken from Section A1 of the Appendix.

Figure 5.3 demonstrates that the final model explains observed waveforms significantly better than the starting model. This is reinforced by computing *FLEXWIN* windows[3] for the starting and final model at 20 s period for the 71 events used in the final period band. This results in a doubling of window lengths in the latter case (192 h to 406 h), thus indicating that *SASSY21* satisfies the data significantly better than the starting model. Note that *SASSY21* is able to explain true-amplitudes despite only utilising relative amplitude information throughout the inversion (see Section 3.3.1).

5.2.2 Misfit contribution by parameter

Some parameters are more difficult to constrain than others (see Section 4.7.2) due to variations in seismic wave sensitivity. As a result, the interpretation of the final model will largely be based on shear-wave structure as further explained in Section 5.3. While I do not interpret density and P-wave velocity structure, it can be shown that the updates to these parameters are required to decrease the waveform misfit as described below.

I test each inversion parameter separately by replacing the corresponding parameter in the final model by the values of the starting model. Then, I simulate the wave propagation through the modified model and evaluate the misfit. A summary of the resulting misfit increase compared to *SASSY21* is provided in Table 5.1, which shows that v_{SV} and v_{SH} drive the inversion because they have the largest effect on the misfit, but the updates in v_P and density are also required to explain observed waveforms.

Table 5.1: Time-frequency phase misfit increase per inversion parameter class when replacing the corresponding inversion parameters in *SASSY21* by the values of the starting model. Note that the individual event misfits increased for all parameters.

	v_{SV}	v_{SH}	v_P	ρ
misfit increase	+60 %	+42 %	+12 %	+5 %

5.2.3 Ability to satisfy unused data

I tested the validity of *SASSY21* 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. A detailed overview of the events can be found in Section A1 of the Appendix.

The 3-D synthetics in the presence of the final model result in an event misfit decrease that is only 3 % lower compared to data used in the actual inversion. Figure 5.4 shows that synthetics obtained from the final model are able to explain true-amplitude data from horizontal and vertical components as well as body and surface wave phase arrivals. For comparison, Figure 5.4 also shows synthetics obtained from the 1-D starting model *CSEM* at

[3]Note that *SASSY21* is obtained using manually selected windows, which yield an analysed window length between ∼1,000 – 3,000 h per period band (see Table 4.2). Here, *FLEXWIN* windows are solely computed for the purpose of assessing the final model in order to avoid time-consuming manual window picking for the starting model.

this period. Note that the synthetics from the starting model show a significant advance for body wave arrivals.

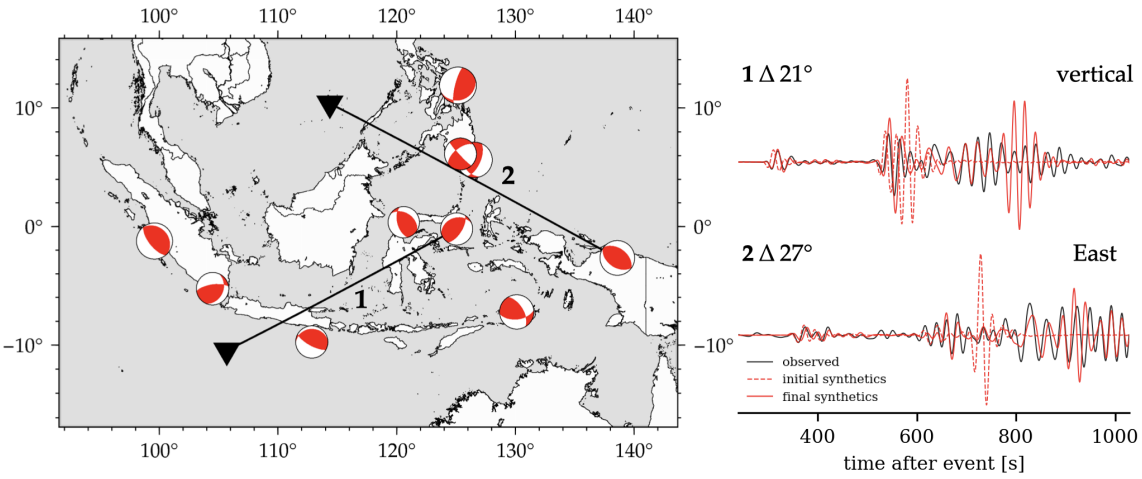


Figure 5.4: *Left:* Map of the validation dataset consisting of ten earthquakes of M_w5.5 – 6.5 with a relatively even spatial distribution. *Right:* Horizontal and vertical component seismograms at two different stations with epicentral distances of 21° and 27°.

5.2.4 Hessian-vector product analysis

Uncertainty quantification based on exploiting the inverse Hessian H^{-1} is currently prohibitively expensive to handle in FWI (see Section 3.4). Consequently, several studies have analysed the Hessian-vector product $H\delta m$ for a test function δm (e.g. Fichtner and van 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 an approximation of the Hessian with L-BFGS during the inversion has already been constructed (see Section 3.3.3), this can be directly applied to a model perturbation in order to obtain a qualitative analysis of inter-parameter trade-offs.

Figure 5.5 presents a visualisation of the Hessian-vector product for a v_{SV} input perturbation. Note that $H\delta m$ is shown relative to the input and indicates the sensitivity of the misfit with respect to perturbations of the final model. This reveals that the model is most sensitive to changes beneath the Sundaland block, as expected from the data coverage (see Figure 4.10), and less sensitive to changes at 300 km depth than 100 km depth as previously demonstrated in Figure 4.17. Furthermore, the inversion appears to suffer from some cross-talk between parameters, which is more pronounced for v_{SH} and density than for v_P, as reinforced by the synthetic tests in Section 4.7.3.

A similar approach is presented in Rodgers et al. (2022) with the difference that a small model perturbation ($\pm$ 0.2 %) is added to the final model, and the gradient for the entire dataset is computed for the perturbed model, which allows to capture the curvature in the direction of δm more accurately.

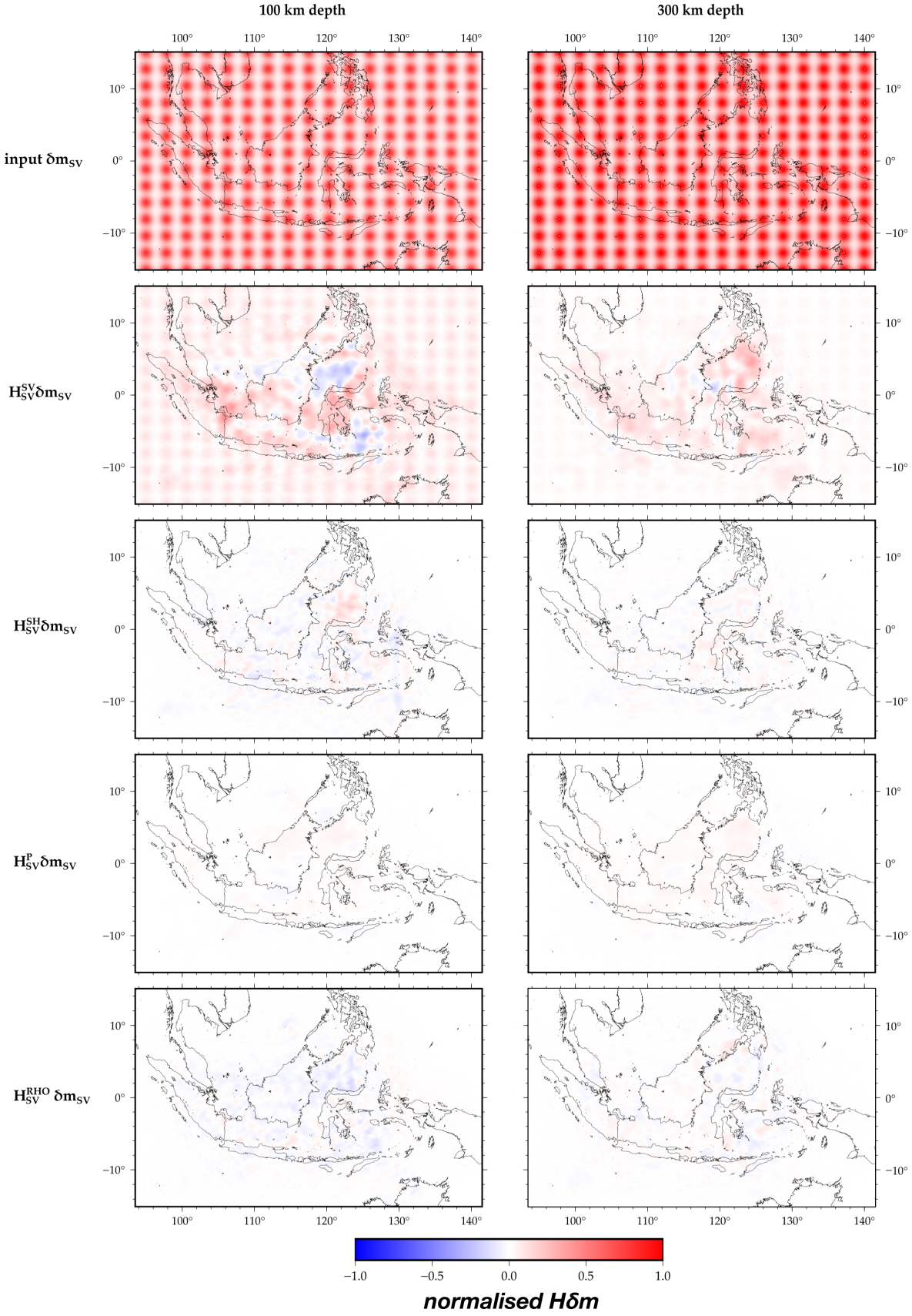


Figure 5.5: Visualisation of $H\delta m_{SV}$ for all inversion parameters at 100 and 300 km depth. *Top panel:* Depth slices of the input perturbation: a 3-D checkerboard pattern of Gaussian v_{SV} spheres with a standard deviation of 70 km. *Panels below:* $H\delta m$ for all inversion parameters (v_{SV}, v_{SH}, v_P and ρ) relative to the input and normalised to $H^{SV}_{SV}\delta m$.

5.3 Final model: *SASSY21*

After 87 L-BFGS iterations, the model is substantially modified from the crust down to the transition zone. Figure 5.6 shows the depth-averaged perturbations, which reveal dominantly 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 behaviour in their updates. The lack of suspicious behaviour is reassuring, because both parameters are difficult to constrain during the inversion (see Section 4.7.2), since they are less sensitive to the data than v_{SV}.

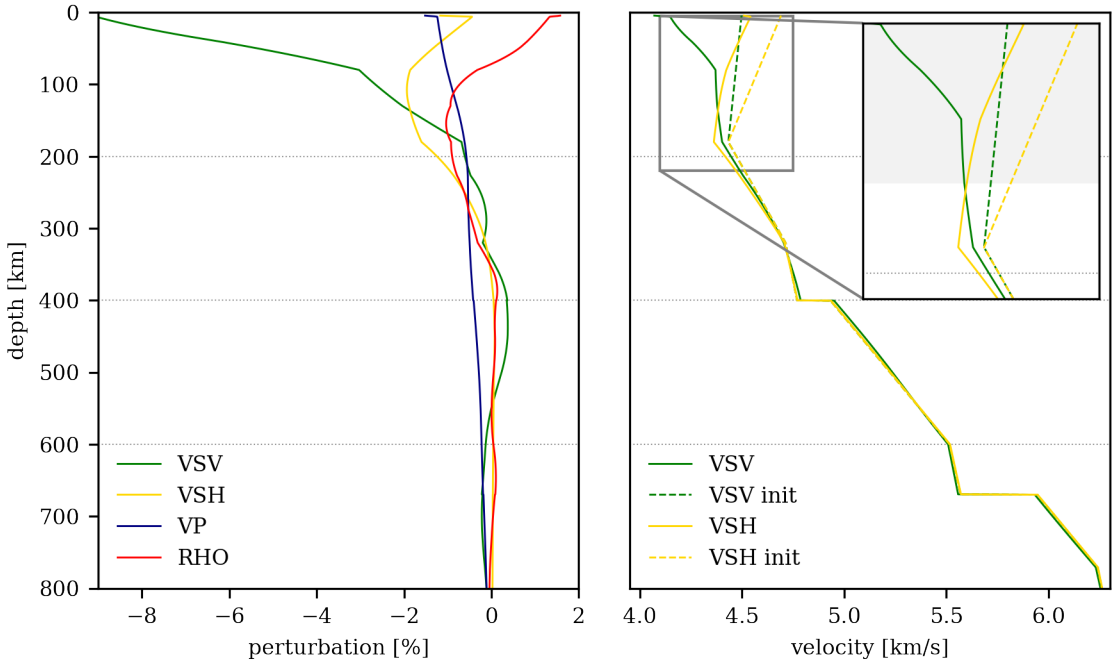


Figure 5.6: *Left:* Depth average of the magnitude of the relative difference between the initial and final model for all inversion parameter classes. *Right:* The depth-averaged absolute v_{SV} and v_{SH} values for the initial and final model, including a zoom-in of the upper 220 km. The grey highlighted area denotes depth values with positive radial anisotropy ($v_{SH} > v_{SV}$) in the final model.

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. Furthermore, the starting model does not include crustal structure for reasons outlined in Section 5.4.3 and thus, strong updates are required to deal with the large corrections that are needed. In addition, the lithosphere tends to be the most heterogeneous part of the upper mantle. I attribute the somewhat linear variation in elastic parameters with depth in the upper ∼70 km (see Figure 5.6) to the wavelength of the short-period waves (∼20 – 30 s), that is, seismic waves at this period are sensitive to the bulk crustal structure (e.g. Capdeville et al., 2010a). The kink at 70 km depth does not coincide with a mesh element boundary.

The final model reveals 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: $v_S = \sqrt{(2{v_{SV}}^2 + {v_{SH}}^2)/3}$ (e.g. Babuska and Cara, 1991; Panning and Romanowicz, 2006). Furthermore, the following analysis is based on the consensus that cold, subducted lithosphere produces strong negative temperature anomalies, which are reflected by relatively high seismic velocities in tomographic images (e.g. Grand, 2002).

Figure 5.7 shows v_s depth slices from 50 to 700 km through the final model. The results for all inversion parameters are presented in Section A3 of the Appendix. In the upper ∼100 km, the oceanic lithosphere beneath 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 as expected from the large, thick Cenozoic sedimentary basins in this region (see Chapter 2). 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, I will discuss some of the key features of the final model in more detail.

5.3.1 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 5.6). The need for a regional low-velocity zone was already apparent from strongly delayed observed waveforms at long periods (see Figure 4.9). The low lithospheric velocities are consistent with previous tomography studies (e.g. van der Hilst et al., 1997; Lebedev and Nolet, 2003; Zenonos et al., 2019) and other measurements such as high heat flow (e.g. Artemieva and Mooney, 2001). This suggests a thin, warm and weak lithosphere, which may be the result of long-term subduction beneath the Sundaland block (see Section 2.2). While I focus on the interpretation of S-wave structure, it is worth mentioning that P-wave structure shows local updates of -6 % beneath the Sundaland block, with the misfit analysis in Section 5.2.2 suggesting that this low-velocity zone is necessary to explain observed waveforms.

Seismic anisotropy – the directional dependence of seismic velocities – has long been linked to past and present deformation in the Earth's interior, e.g. radial anisotropy is frequently used as an indicator for the record of material flow (e.g. Hess, 1964; Montagner and Tanimoto, 1991). The regional low-velocity zone is characterised by strong radially anisotropic values of up to 18 %. For the upper 130 km, an overall positive radial anisotropy ($v_{SH} > v_{SV}$) is observed, 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 5.6, while Figure 5.8 presents lateral variations at 150 and 250 km depth. These results reveal negative radial anisotropy along the slabs (which is in good agreement with Sturgeon et al., 2019) and beneath Sundaland in the west, and positive radial anisotropy around the Celebes Sea, Sulawesi and the Banda Sea. This is believed to be the result of two different mechanisms:

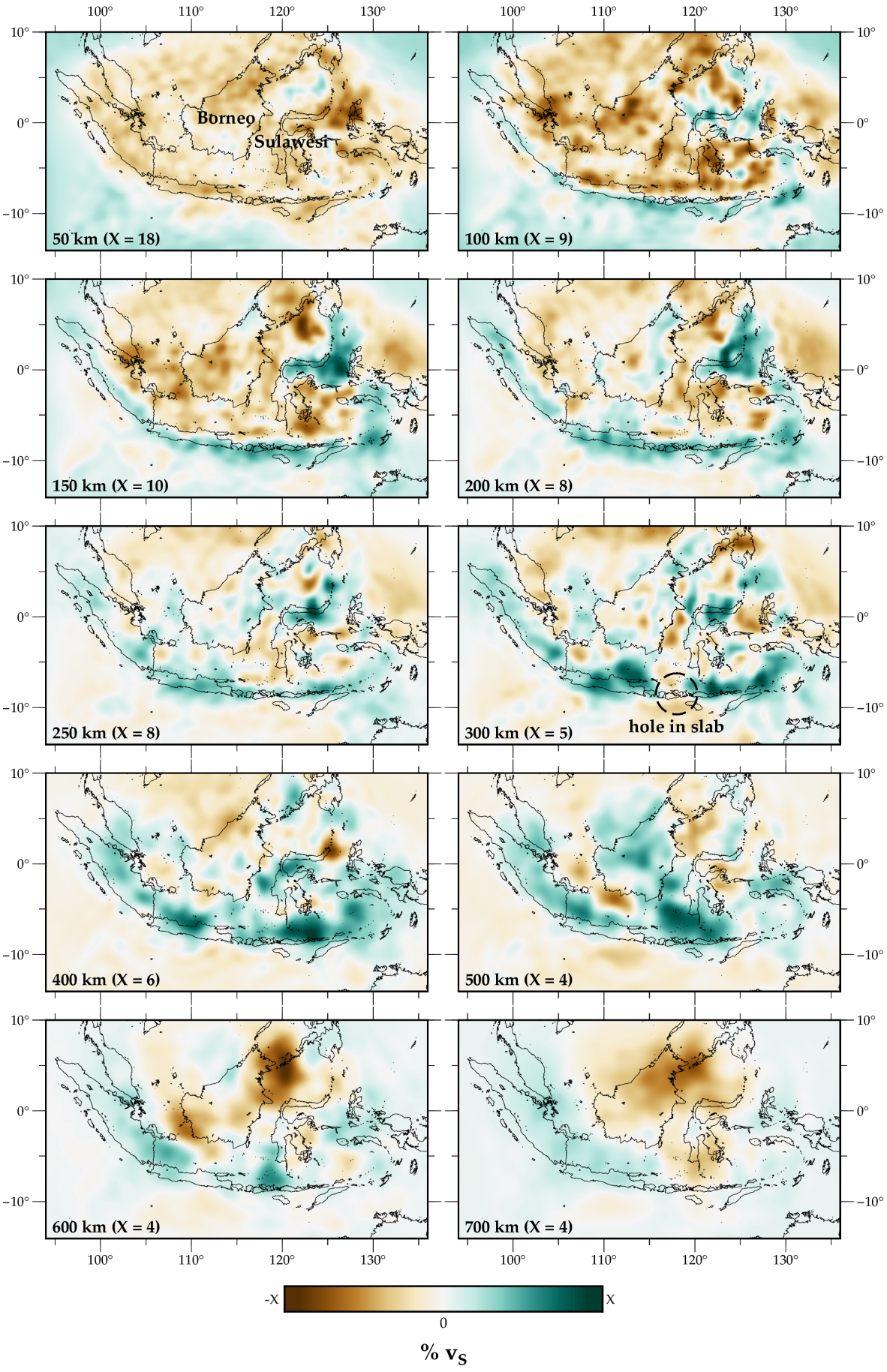


Figure 5.7: Shear-wave (v_S) depth slices between the range 50 and 700 km. Perturbations are in % relative to the depth-average. The limits of the colourscale X are shown in the lower left corner of each plot.

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 and Kawakatsu, 2012).
2. Negative frozen-in anisotropy of continental-lithosphere roots during formation (Priestley et al., 2021), thus explaining negative values beneath the Sundaland block.

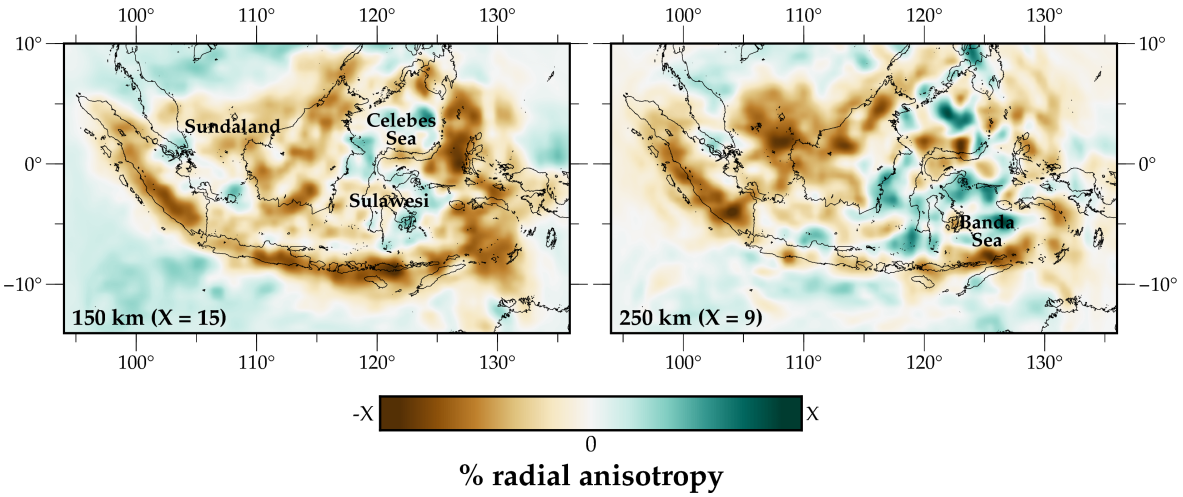


Figure 5.8: Radial anisotropy ($\frac{v_{SH}-v_{SV}}{v_S}$) in % for the final model at (*left*) 150 km and (*right*) 250 km depth. The limits of the colourscale *X* are shown in the lower left corner of each plot.

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, I refrain from a more detailed geological interpretation of the radial anisotropy. Nevertheless, radial anisotropy depth slices from 50 to 700 km are presented in Section A3 of the Appendix for completeness.

The following interpretation is built upon SASSY21 and was done by Conor Andrew Bacon, postdoctoral researcher at the University of Cambridge.

Bacon et al. (submitted) study seismic anisotropy across the Malaysian state of Sabah, northern Borneo, using shear-wave splitting analysis. However, it is difficult to constrain the depth of an anisotropy source with shear-wave splitting measurements, and thus, we extract 1-D v_{SV} and v_{SH} profiles for northern Borneo from *SASSY21* to analyse radial anisotropy, which provides better vertical resolution. This reveals positive radial anisotropy in the upper 50 – 75 km for northern Borneo (compared to ∼130 km for Southeast Asia), which is roughly consistent with the thickness of an anisotropic layer observed in Greenfield et al. (2022) using two-plane wave tomography. In summary, we conclude that the observed seismic anisotropy is principally attributable to the lithosphere and corresponding mechanisms of deformation.

Lithosphere–asthenosphere boundary (LAB)

The lithosphere–asthenosphere boundary (LAB) is defined as the thermal boundary between the Earth's rigid outer layer (including both the crust and non-convecting uppermost mantle) and the warmer, ductile asthenosphere, and thus has significant implications for plate tectonics. In order to determine the lithospheric thickness, Priestley and McKenzie (2006) suggest an approach to convert regional variations in absolute shear-wave velocities to temperature based on an empirical parameterisation. Theoretically, the base of the lithosphere can be marked by a rapid decrease in the temperature gradient with depth, which is usually observed at around 1,333°C. Following this approach, Greenfield et al. (2022) determine a lithospheric thickness of ∼100 km beneath Sabah, northern Borneo, based on a tomography model obtained from two plane wave tomography. This is much thicker than the ∼50 km estimated from global tomography models (Priestley et al., 2018; Roberts et al., 2018; Hoggard et al., 2020), which indicate that there is essentially no lithospheric mantle present.

Here, I follow the same approach – with parameter values taken from Table 1 in Roberts et al. (2018) – to convert shear-wave velocity values of *SASSY21* to temperature (see look-up chart in Figure 8e in Roberts et al., 2018). For comparison, I also convert the global shear-wave tomography model *SL2013* (Schaeffer and Lebedev, 2013) to temperature. For both models, the temperature values are too high to select a 1,333°C contour surface for many locations in Southeast Asia as a result of the anomalously low velocities within the region. The 1-D temperature and temperature gradient profiles for both models are similar; while the *SL2013* temperature profiles are smoother (as one would expect from a global model), the resulting temperature profiles and gradients show strong variations for both models, which makes the contouring of the LAB a difficult exercise. Thus, I compute the average temperature gradient for four different subregions as shown in Figure 5.9 to get a smoother, representative profile. This shows that the LAB is around ∼175 km for the North Australian craton and ∼70, 80 km for northern Borneo, the Sundaland block and the Celebes Sea. Note that *SL2013* tends to yield shallower LAB values than *SASSY21*.

It should be noted that the approach described above makes the assumption that any variation in v_S arises from temperature variations only. However, v_S is also affected by composition, the presence of melt and variations in grain size (e.g. Dalton et al., 2014). It is likely, therefore, that the strong 3-D perturbations in seismic structure recovered by FWI have a significant contribution from these other parameters, which makes a more detailed analysis of the LAB structure using the above approach of questionable value.

5.3.2 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. The depth slices in Figure 5.7 show the Sunda slab descending at depths $\geq$ 100 km down to the mantle transition zone. It is assumed that 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

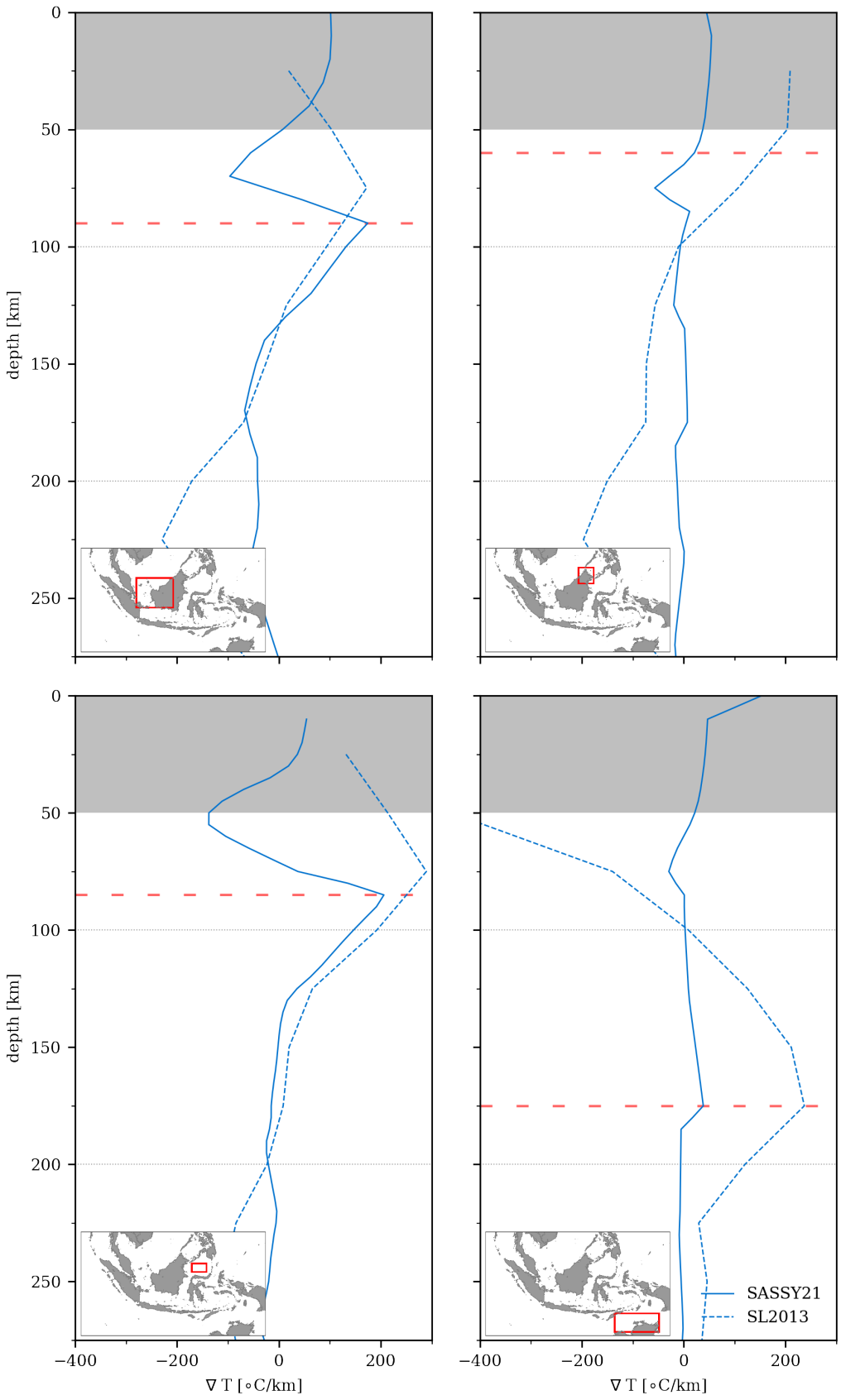


Figure 5.9: Temperature gradient profiles for *SL2013* (dashed line, Schaeffer and Lebedev, 2013) and this study (solid) for four different regions within Southeast Asia. The red rectangle in the inset map in each plot shows the region used for averaging shear-wave velocities before computing the corresponding temperature using the approach of Priestley and McKenzie (2006). The red dashed line in each plot indicates the manually selected LAB value. The grey highlighted area denotes the upper 50 km, which is not used for mapping the LAB.

in Section 4.5. Further east, the bending of the Banda Arc is imaged as one continuous slab from ∼100 – 300 km depth. A geodynamic modelling study by Moresi et al. (2014) potentially supports the interpretation of a single bent and deformed slab by modelling how the curvature of this system could have developed from northward motion of the (Indo-) Australian plate (see Section 2.4.2). 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 synthetic recovery test presented in Section 4.7.2 has shown that the given simulation setup yields sufficient sensitivity in the southeast to recover the sign of an anomaly around northern Australia.

As the iterations progress, shorter-period data is added and thus, smaller-scale features are resolved, which is illustrated in Figure 5.10. Subduction zones first become apparent at 50 s[4], which I largely attribute to body wave arrivals becoming clearly identifiable at this period (see Section 4.2.5). Furthermore, the smoothing lengths are decreased based on the minimum wavelength considered (see Section 4.5.2) and thus, features become sharper as the dominant period is decreased. The bottom panel of Figure 5.10 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.

Figure 5.11 presents several cross-sections through *SASSY21*. Figure 5.11a shows a south-north cross-section through Sumatra, revealing a steeply dipping Sunda slab, which appears to penetrate through the mantle transition zone. Furthermore, one can distinguish between high velocities of the oceanic crust of the subducting Australian plate in the south and low velocities beneath the continental shelf of the Sundaland block in the north (see Section 5.3.1). Further east, Figure 5.11d 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. Subduction of the Philippine Sea plate is associated with a high-velocity zone down to depths of ∼200 – 400 km, which is significantly deeper than the previously suggested ∼100 km (see Section 2.5). Figure 5.11b and Figure 5.11c will be discussed in the subsequent sections.

[4] Although the 100 s gradient already shows evidence of subduction (see Figure 4.16).

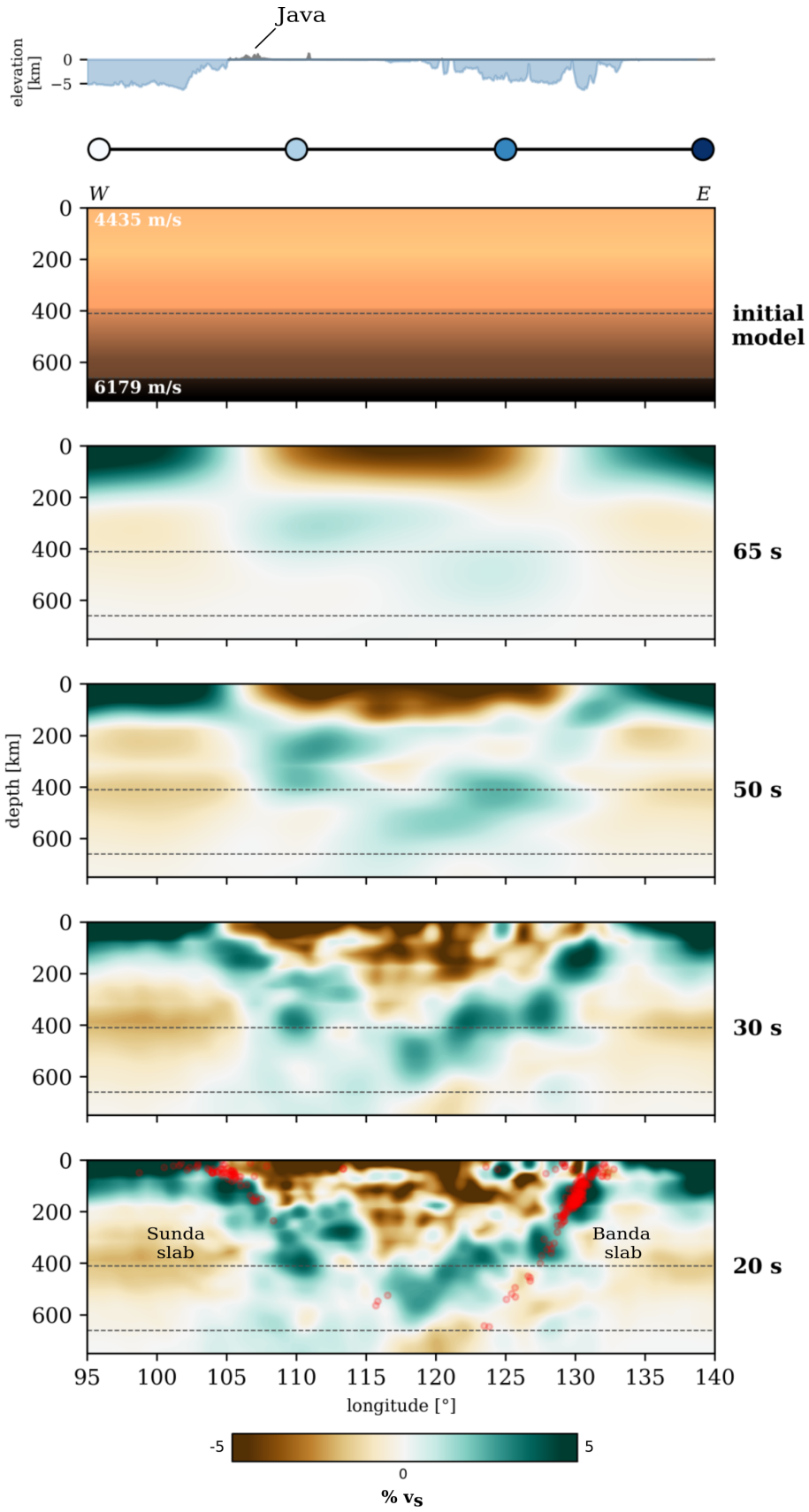


Figure 5.10: West-east v_S cross-section across different period bands. The section's location corresponds to the blue dotted section in Figure 5.11. The top plot shows the absolute values of the initial model, while the other plots show perturbations from the depth-average in % for the final iteration within the respective period band. Earthquake locations (red dots in bottom plot) are taken from the *ISC* catalogue (International Seismological Centre, 2016) and are within 25 km from the cross-section slice. The vertical-horizontal ratio is 1:2.

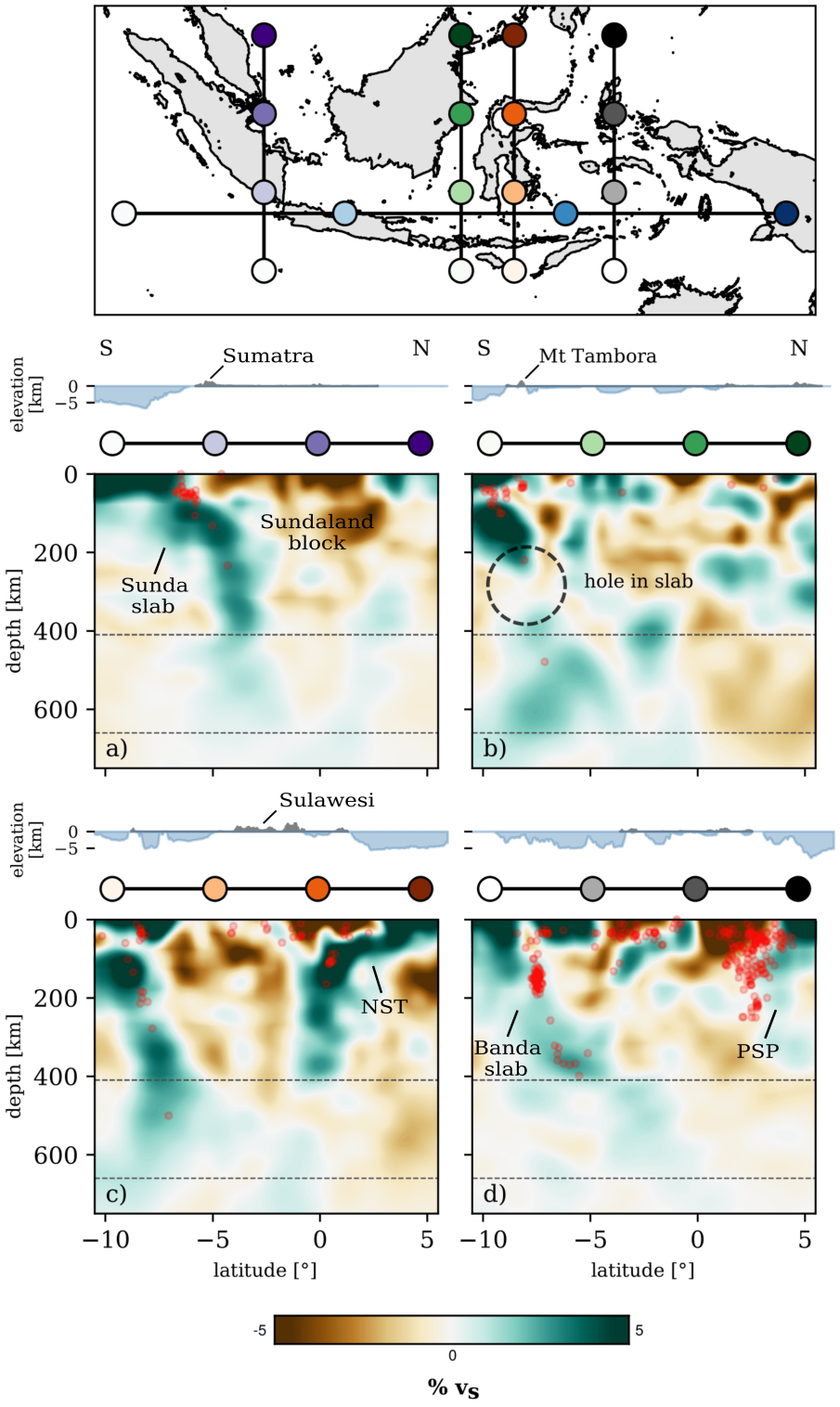


Figure 5.11: South-north v_S cross-sections illustrating subduction around the Philippine Sea Plate (PSP) and North Sulawesi Trench (NST), and a map showing their locations. The blue dotted section corresponds to the west-east section in Figure 5.10. Perturbations are in % relative to the depth-average. Earthquake locations (red dots) are taken from the *ISC* catalogue (International Seismological Centre, 2016) and are within 25 km from the cross-section slice. The vertical-horizontal ratio is 1:2.

5.3.3 Hole in slab beneath Mount Tambora

The 300 km v_s depth slice in Figure 5.7 reveals a hole in the slab east of Java, roughly beneath Mount Tambora. The existence of this hole was previously suggested based on ray tomography studies (e.g. Widiyantoro et al., 2011; Hall and 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 tomography model, Harris et al. (2020) conclude 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 and 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 5.11b shows the hole in a cross-section, which has dimensions of ∼300 x 100 km.

5.3.4 High-velocity zone(s) beneath Borneo

SASSY21 shows a high-velocity zone beneath northern Borneo, which extends from ∼100 to 300 km depth (see Figure 5.7). A similar anomaly was imaged previously in ray tomography studies (Hall and Spakman, 2015; Zenonos et al., 2019), but was regarded as suspicious owing to the poor data coverage. However, this study uses data from a dense, regional network in this region (see Figure 4.10) and I thus argue that this feature is likely not an artefact and may be associated with the post-subduction setting of the study area. This is consistent with the absence of seismicity and the known Neogene history (∼23 – 2 Myr) of northern Borneo (see Section 2.2.1)

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 5.7). The anomaly appears connected with the one identified beneath northern Borneo but this may simply be a result of the long periods considered. This anomaly may be associated with underthrusting from the accretion of Sulawesi in the east during the Miocene (∼23 – 5 Myr, see Section 2.3).

5.3.5 Comparison with other models

Figure 5.12 and Figure 5.13 show a comparison between *SASSY21* and other shear-wave tomography models. A summary of the tomography studies presented in this comparison is provided in Table 5.2. Note that Zenonos et al. (2020) and *SASSY21* are continental-scale models for Southeast Asia, while the other models are from global studies. *SASSY21* and *GLAD-M25* are obtained using full-waveform inversion.

Figure 5.12 shows a comparison of selected models at 100 and 200 km depth. For all models, the high-velocity zone along the Indonesian volcanic 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 (see Section 2.4.2). However, *SASSY21* is able to resolve smaller-scale features,

Table 5.2: Overview of models used for comparison throughout this study, adapted from Table 1 in Hosseini et al. (2018). Note that *SP12RTS* is a successor to *S20RTS* (Ritsema et al., 1999; Ritsema et al., 2009) and *S40RTS* (Ritsema et al., 2011).

model	data type	reference model	reference	source
CAM2016	surface waves	*SPREM*	Ho et al. (2016)	IRIS EMC website (Hutko et al., 2017)
GAP-P4	body waves	*GAP*	Obayashi et al. (2013)	Masayuki Obayashi, pers. communication
GLAD-M25	body and surface waveforms	*STW105*	Lei et al. (2020)	IRIS EMC website (Hutko et al., 2017)
SEISGLOB2	surface waves, body waves, normal modes	*PREM*	Durand et al. (2017)	IRIS EMC website (Hutko et al., 2017)
SL2013	surface waves	*AK135*	Schaeffer and Lebedev (2013)	Andrew J. Schaeffer, pers. website
SP12RTS	surface waves, body waves, normal modes	*PREM*	Koelemeijer et al. (2016)	Paula Koelemeijer, pers. website
UUP07	body waves	*AK135*	Amaru (2007)	`atlas-of-the-underworld.org`
Zenonos	body waves	*AK135*	Zenonos et al. (2019)	Aristides Zenonos, pers. communication

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 this study (see Section 4.2.1).

Figure 5.13 presents a comparison between *SASSY21* and other continental- and global-scale tomography models for a north-south cross-section through Sulawesi and the Indonesian volcanic arc. While several other models show hints of descending slabs along this section, the comparison shows that *SASSY21* is of unprecedented detail. The slab associated with the North Sulawesi Trench extends down to ∼410 km depth, while *Slab2* (Hayes et al., 2018) tracks it down to only ∼200 km (see Figure 2.5). The slab appears to have a small dip angle at shallower depths before subducting steeply from ∼100 km onwards, similar to a deep aseismic slab observed beneath Alaska (Gou et al., 2019). Aseismic slabs have been observed worldwide (e.g. Huang et al., 2013; Gou et al., 2019; Fan and Zhao, 2018) and have been associated with high temperatures, which facilitate a slab descent generating little seismicity (Huang et al., 2013). The deep subduction along the North Sulawesi Trench is observed from ∼120 – 123°E longitude.

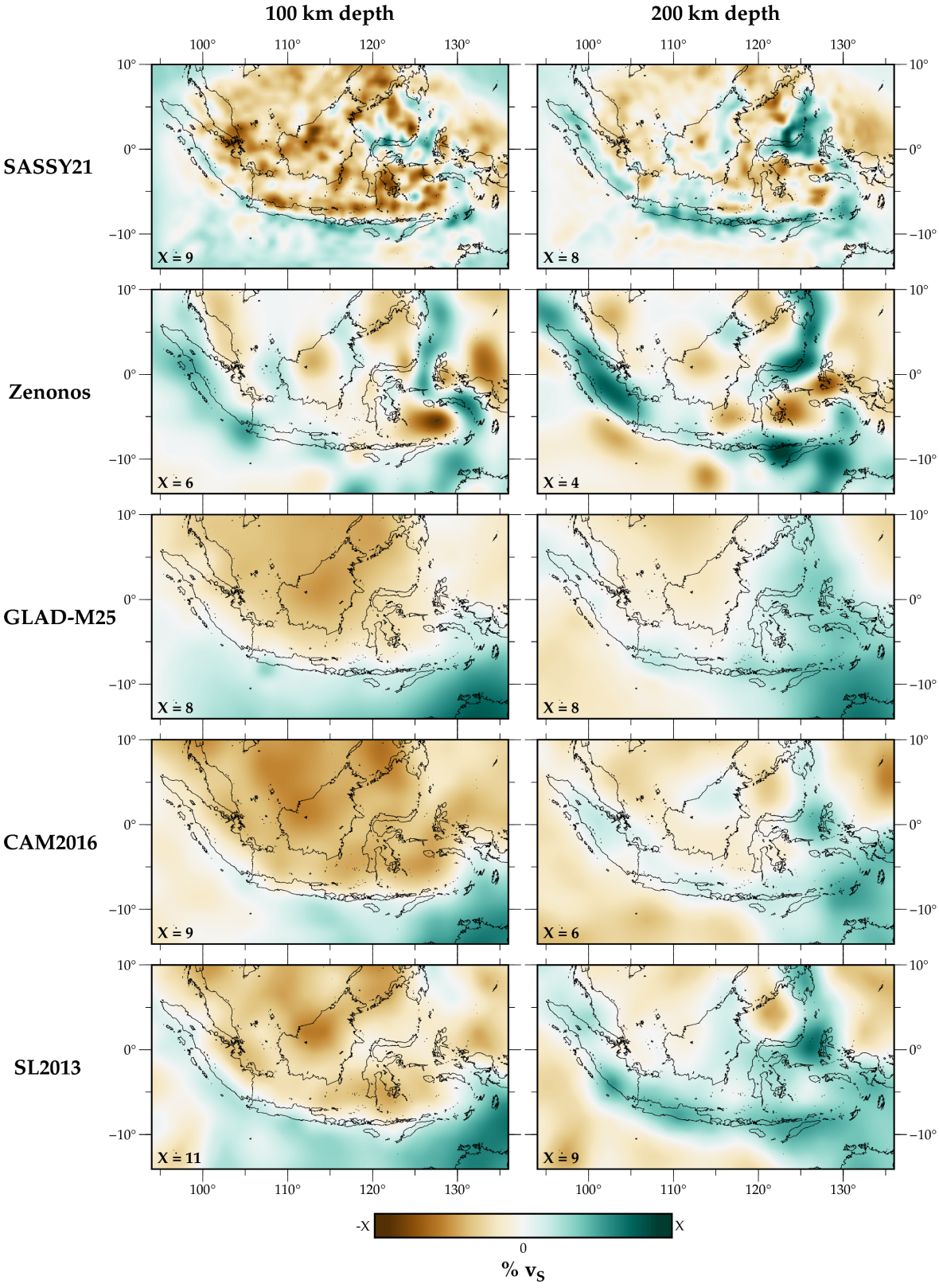


Figure 5.12: Shear-wave depth slices at (*left*) 100 km and (*right*) 200 km depth for five different models (see overview in Table 5.2). Perturbations are in % relative to the depth-average within the region. The limits of the colourscale *X* are shown in the lower left corner of each plot.

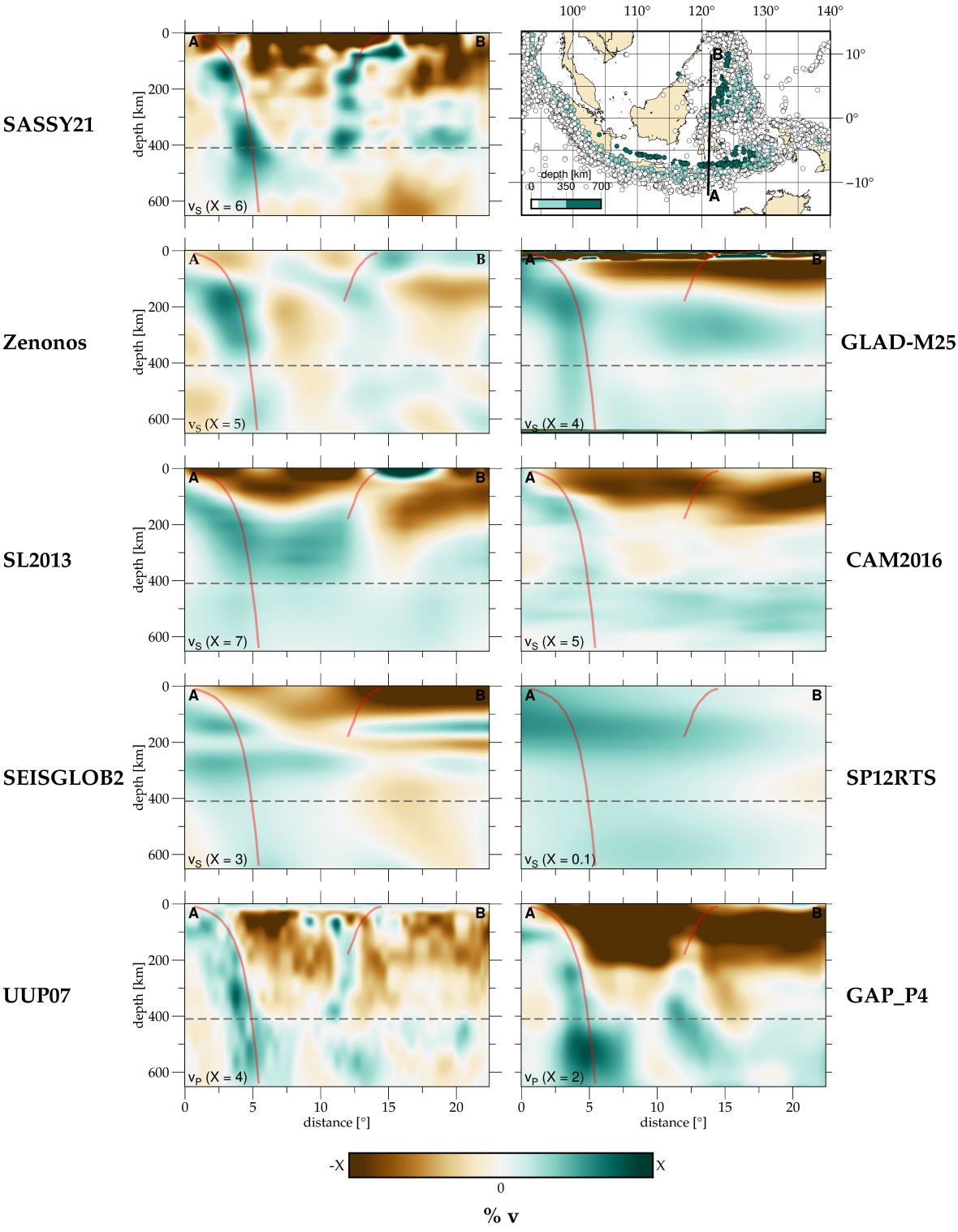


Figure 5.13: Comparison between *SASSY21* and various regional and global tomography models (see overview in Table 5.2) for a north-south cross-section through Sulawesi and the Indonesian volcanic arc. Perturbations are in % relative to the depth-average within the region. The limits of the colorscale *X* are shown in the lower left corner of each plot. Red lines in the sections indicate *Slab2* values (Hayes et al., 2018). The map in the top right corner shows the section's location and seismicity ($M_w > 5$) taken from the *ISC* catalogue (International Seismological Centre, 2016), coloured by depth.

5.4 Limitations

An obvious limitation of FWI is the high computational cost of the forward problem, which translates to the use of relatively long periods and smaller datasets compared to ray tomography studies (see Section 3.5). 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. As mentioned in Section 3.2.2, seismic waves have measurable sensitivity to only a few parameters in many cases (Sieminski et al., 2009) and 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 such as density (e.g. Blom et al., 2017) and attenuation (e.g. Yang et al., 2020). Based on synthetic tests, they conclude that these parameters can be imaged in theory but this has proven to be difficult using real data, mainly because of a lack of dense data coverage and the difficulty in determining the optimal observables for constraining a specific parameter.

Further limitations and ongoing developments of FWI are discussed in Section 3.5. For example, lower mantle and core phases (e.g. "PcP") are contained in observed waveforms but are not modelled since the wavefield propagation is only computed through the target region. However, many of these phases occur at shorter periods of $\sim$1 – 10 s, greater epicentral distances, have low amplitude and/or will likely be removed by the data selection procedure (Sanne Cottaar, pers. comm., February 2021). Furthermore, great care was taken to build an event catalogue that avoids interference with other events (see Section 4.2.4), which prevents core phases from teleseismic events interfering with the seismograms.

5.4.1 Source inversion

As a result of the limited data availability and lack of azimuthal coverage, I chose not to invert for source parameters, while acknowledging the potential for source errors to map as artefacts in the tomography model. It is worth noting, however, that in addition to the need for sufficient data coverage, seismic source inversion is also a challenging task due to the non-linearity of the problem. Furthermore, two additional simulations would be needed in order to construct the gradients to jointly invert for structure and sources in waveform tomography (Liu et al., 2004), which would result in a significant increase in computational cost. As a result of these factors, source parameters are either not inverted for (e.g. van Herwaarden et al., 2020; Gao et al., 2021) or they are updated in consecutive iterations (rather than jointly) with the seismic structure if the data coverage is sufficient (e.g. Tape et al., 2010; Tao et al., 2018).

Blom et al. (2022) investigate the issue of source parameter errors in adjoint waveform tomography, and demonstrate that plausible source errors can result in artefacts appearing several wavelengths away from the hypocenter. They also suggest that strong, local source effects can be mitigated by removing a near-source volume from the event gradient, as employed in this study (see Section 4.5). I believe that the effects of source errors are further suppressed through the careful event selection (see Section 4.2) and monitoring throughout the inversion, e.g. I remove events that develop a suspicious waveform misfit as the period

decreases and events with long or complex source time functions. Nevertheless, the ideal tomographic approach is a joint inversion of structure and source parameters (Valentine and Woodhouse, 2010), something that is the subject of ongoing developments by the waveform tomography community (e.g. Zhao et al., 2006; Hejrani et al., 2017; Fichtner and Simutė, 2018).

5.4.2 Mesh refinement

As described in Section 4.1.4, the final mesh of each period band is interpolated onto an empty *CSEM* mesh (Fichtner et al., 2018) at the new period band using *MultiMesh* (Thrastarson et al., 2021). Thus, the mesh is finer, but it is still based on the 1-D elastic parameters of *CSEM* and does not account for the 3-D variations introduced throughout the inversion. In particular, *CSEM* is too fast for Southeast Asia, which means the elements are too large. It is hoped that the mesh generation based on 1.5 elements per $\lambda_{\min}$ (see Section 4.1.3) somewhat mitigates this issue but the final model is up to 20 % slower than the starting model (see Figure 5.6 and Figure 5.7) and thus, the wavefield is likely to be undersampled. Note that any high-velocity features should be sampled correctly since the wavelengths are longer. Furthermore, the discontinuities at 410- and 660 km in the mantle transition zone are still properly implemented since they are explicitly meshed as element boundaries in the starting model (see Section 4.1.1).

A new *Salvus* functionality (as of October 2021) enables mesh refinements that account for local variations in parameters and thus allows me to investigate its effect on synthetic waveforms. Here, I refine the final model *SASSY21* and run a forward simulation for two events used throughout the inversion. Figure 5.14 shows a comparison of the synthetics for the source-receiver pair with the highest L_2 misfit (see Section 3.3.1) for each event. This demonstrates that the waveform differences are minor (likely below noise level) but there is evidence of numerical dispersion (see Section 3.2.1) in the later part of the seismograms for the original mesh interpolation type used throughout the inversion. However, Table 5.3 demonstrates the significant increase in computational cost associated with the refinement of *SASSY21* for the final period band (20 – 150 s). Thus, I argue that a proper refinement of the mesh is advisable but does not result in significant waveform differences and was not feasible in this study from a computational point of view.

Table 5.3: A comparison of the computational cost at 20 – 150 s for *SASSY21* built upon the starting model *CSEM* vs a refined *SASSY21* mesh that accounts for 3-D variations introduced throughout the inversion. The computation time is based on a wavefield simulation time of 1,100 s on 56 CPU cores. This is because the Cascade Lake nodes used have 56 CPUs (`https://www.hpc.cam.ac.uk`).

mesh	# mesh elements	time step [s]	comp. time [min]
SASSY21	207,636	0.28	∼6
***SASSY21* refined**	557,960	0.13	∼34

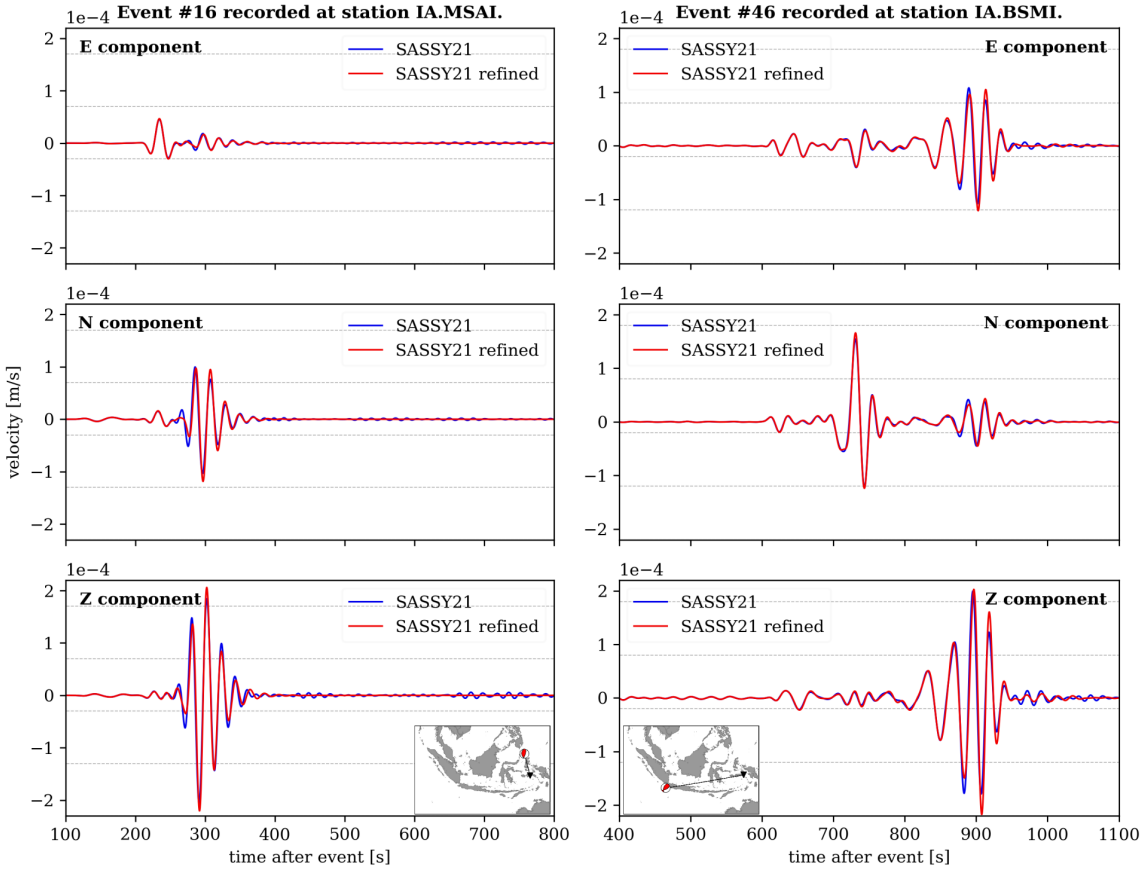


Figure 5.14: Three-component seismograms of two events for the original *SASSY21* mesh (blue) vs a refined *SASSY21* mesh (red) that accounts for 3-D variations introduced throughout the inversion. The maps in the bottom panels show the location of the corresponding source and receiver. The event numbers (#) are taken from Section A1 of the Appendix.

5.4.3 Crustal model

Many ray-based tomography studies employ crustal corrections to account for the behaviour of seismic waves in the upper few kilometres. Bozdağ and Trampert (2008) propose that this can be circumvented in FWI since crust and mantle are jointly imaged as a result of the computation of 3-D sensitivity kernels. In this study, I did not explicitly include crustal structure (e.g. *CRUST1.0*, Laske et al., 2013) in the prior model because:

1. Accommodating the small mesh elements that are required to properly account for the low velocities and small-scale 3-D variations in crustal seismic structure requires advanced meshing techniques and results in a significant increase in computational resources, which were not available when the initial iterations were run.
2. While a crustal model can provide useful prior information, erroneous structure is removed slowly – if at all –, and *CRUST1.0* is potentially not accurate in Southeast Asia. A comparison of *CRUST1.0* model predictions against 25 s Rayleigh wave group velocity maps by Laske et al. (2013) indicate that Southeast Asia has some of the largest misfits on the globe, although this may in part be due to their tests not accounting for

upper mantle structure.

3. The crust is relatively thin in Southeast Asia (∼20 km on average for *CRUST1.0*), with much of it comprising a mix of oceanic and extended continental crust, which likely has less of an effect on the longer period waveforms ($>$ 20 s) compared to normal continental crust of average thickness ∼35 km.

Hence, I decided not to add any prior information about the crustal structure. While the absolute shallow velocities of *SASSY21* are slower than the starting model, they are still too fast for crustal structure (see Figure 5.6) and thus, I cannot exclude that this may have introduced artefacts into the final model, although I do take care to only interpret longer wavelength features that are also seen in the models constrained by longer period data obtained in earlier iterations and that are consistent with the geology and tectonics of the region.

In order to investigate the effect of a crustal model on synthetic waveforms, I generate a set of synthetic waveforms for the starting model *CSEM* and final model *SASSY21* with *CRUST1.0* inserted. *CRUST1.0* is implemented for v_P, v_S and density, and extends down to ∼50 km depth in the southeast (North Australian craton) with a maximum depth of ∼25 km for the rest of Southeast Asia (see Figure 5.16). Below the Mohorovičić discontinuity, the elastic parameters are taken from the background model (*CSEM* or *SASSY21*). The results are presented below.

Starting model *CSEM* with *CRUST1.0*

Implementing a crustal model requires sophisticated meshing tools and results in a significant increase in computational cost since the mesh is refined to account for the low velocities and 3-D variations introduced by *CRUST1.0*. Table 5.4 provides an overview of the increase in computational cost at 100, 50 and 20 s, which is considerable even at longer periods. This confirms that implementing a crustal model would not have been computationally feasible for this study. However, this could be somewhat circumvented by implementing a longer-wavelength equivalent of the crust (e.g. Thrastarson et al., 2022).

Figure 5.15 presents a comparison of synthetics for the 1-D starting model *CSEM* and *CSEM* with *CRUST1.0* at 100, 50 and 20 s. The effect of a crustal model becomes increasingly noticeable at shorter periods but does not have a significant effect on body wave arrivals. Note that the initial period band for this study is 100 – 150 s.

SASSY21 with *CRUST1.0*

I have superimposed *CRUST1.0* (Laske et al., 2013) onto the final model and generated synthetic waveforms for two real events. It should be noted that this test is merely an indicator of how a crustal model affects synthetics and that *CRUST1.0* would have been updated during an actual inversion. Figure 5.16 confirms that the final model explains observed waveforms better than *SASSY21* with *CRUST1.0*. This is likely due to some combination of 1) *CRUST1.0* potentially not being accurate in Southeast Asia (see above) and 2) the vertical resolution that can be achieved using the periods I exploit ($>$ 20 s) being limited, such that crustal

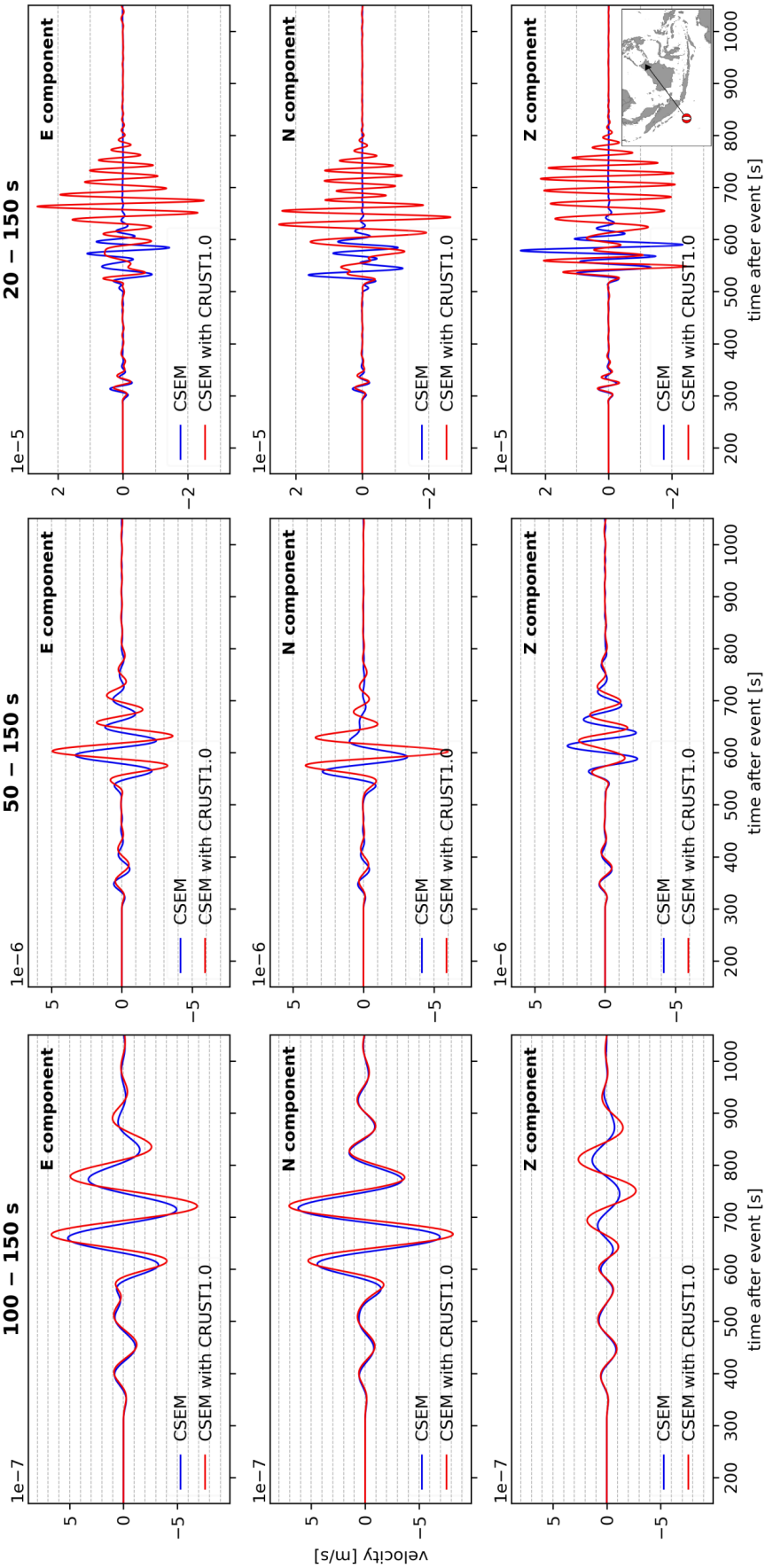


Figure 5.15: Three-component synthetic waveforms for the 1-D *CSEM* and *CSEM* with *CRUST1.0*, recorded at a station in northern Borneo in the period bands 100 – 150 s, 50 – 150 s and 20 – 150 s for a synthetic M_w6.5 event at a depth of 40 km. The inset map shows the location of the source and receiver.

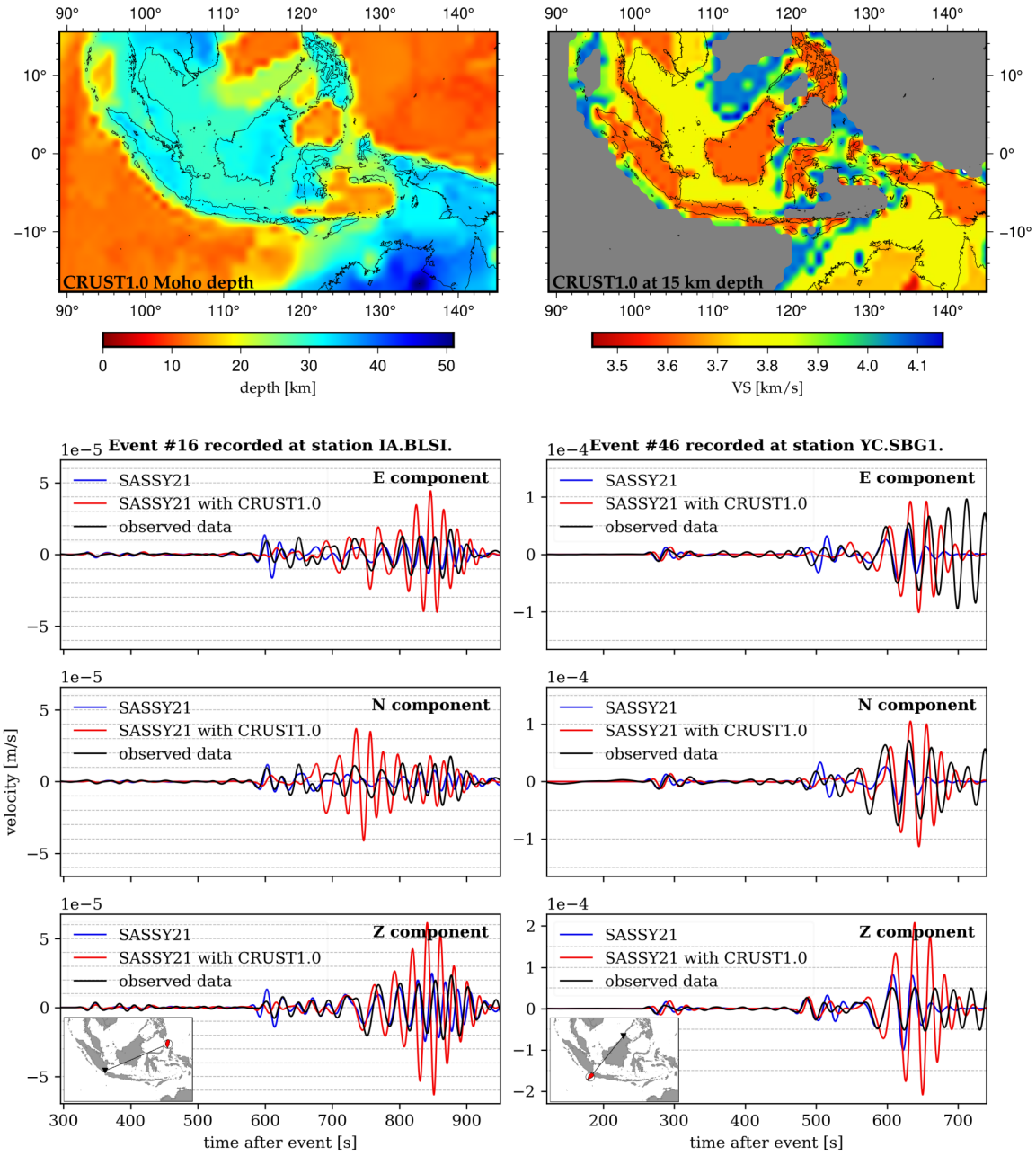


Figure 5.16: *Top left:* Depth of the Mohorovičić discontinuity (Moho) taken from *CRUST1.0* (Laske et al., 2013). *Top right:* S-wave velocities of *CRUST1.0* at 15 km depth. The grey area indicates parts of the model below the Moho, which are replaced by the background model (*SASSY21* in this case). *Bottom:* Two three-component seismograms in the period band 20 – 150 s. The event numbers (#) are taken from Section A1 of the Appendix. The inset map in the bottom left corner of each plot shows the location of the corresponding source and receiver.

Table 5.4: Overview of the number of mesh elements, time step and compute time for the *CSEM* vs *CSEM* with *CRUST1.0* mesh at 100, 50 and 20 s. The domain extension for Southeast Asia is the same as in the main study but the mesh depth is decreased to 500 km. The compute time per core is based on a wavefield simulation time of 1,050 s.

period band	# mesh elements	time step [s]	comp. time on 1 core [min]
100 – 150 s			
CSEM	12,540	0.48	∼8
CSEM with *CRUST1.0*	31,894	0.41	∼25
50 – 150 s			
CSEM	26,609	0.39	∼19
CSEM with *CRUST1.0*	61,290	0.21	∼190
20 – 150 s			
CSEM	160,446	0.31	∼150
CSEM with *CRUST1.0*	433,455	0.14	∼878

anomalies smear into the upper mantle structure of *SASSY21* in a way that is inconsistent with *CRUST1.0*, since *SASSY21* has already retrieved crustal structure at the scale-length to which the minimum periods are sensitive. If shorter period information could be used, then the crust would be better resolved both vertically and horizontally but this would come with a significantly increased computational burden. However, this test demonstrates that it is not meaningful to add a crustal model at later stages of the inversion.

5.5 Model and data availability

The final model is available as *NetCDF* and *HDF5* files, with the former being readable by e.g. *xarray* (Hoyer and Hamman, 2017) and the latter suitable for viewing with *ParaView* (Ahrens et al., 2005) and interaction with *Salvus* (Afanasiev et al., 2019). I further provide *SASSY21* in *CSV* format, a Jupyter Notebook (Kluyver et al., 2016) demonstrating how to interact with the different file formats, an example that demonstrates the processing steps for the observed waveforms and the filtered, windowed data to allow reproduction of the results. The data, final model and a 3-D model fly-through can be found on a *Zenodo* repository at `https://doi.org/10.5281/zenodo.5573139` (Wehner et al., 2021). Furthermore, the pre-gridded model in *NetCDF Classic* format is available on the *IRIS EMC* website (Trabant et al., 2012; Hutko et al., 2017) at `https://doi.org/10.17611/dp/emc.2022.sassy21.1`.

The implementation of SASSY21 in the Collaborative Seismic Earth Model presented below was carried out by Sebastian Noe, PhD student at ETH Zürich.

SASSY21 is implemented in the *Collaborative Seismic Earth model (CSEM): Generation 2* (Noe et al., in prep.). *CSEM* is a multi-scale global tomographic Earth model that evolves continuously through successive regional- and global-scale refinements. It is a group effort initiated by the *Seismology and Wave Physics Group* at *ETH Zürich*, and gathers individual

tomography models to create a global model. The current *CSEM* is successively refined by adding finer-scale structure constrained by newly added data. It intends to bridge the gap between regional and global tomography, and aims to deliver the first multi-scale Earth model, where crust and mantle are jointly resolved. Consequently, it has the potential to serve as prior information for many seismic studies around the globe. Figure 5.17 presents a depth slice through *CSEM2* and shows how *SASSY21* was tapered into surrounding regions.

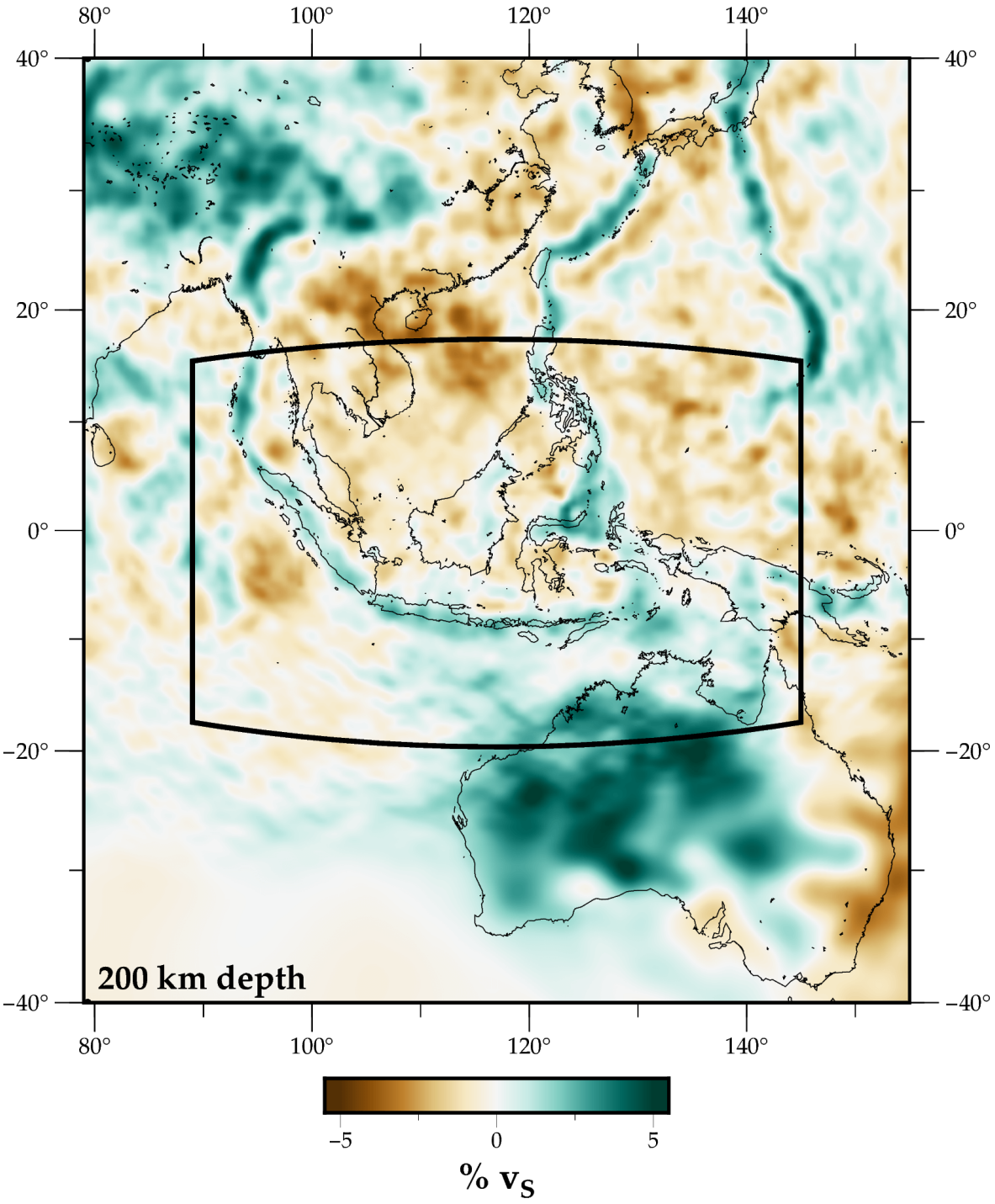


Figure 5.17: v_S depth slice at 200 km through *CSEM: Generation 2* (Noe et al., in prep.) at 30 s. The black box indicates the *SASSY21* study area.

6 Eastern Indonesian waveform tomography – *SASSIER22*

In Chapter 4 and 5, I presented my continental-scale Southeast Asian waveform tomography study at periods down to 20 s. In this chapter, a new 3-D seismic structural model of the lithosphere and underlying mantle beneath eastern Indonesia and its surroundings is presented, which exploits seismic data filtered at periods between 15 – 150 s period. *SASSY21* (see Chapter 5) is used as a starting model, and the study region is characterised by particularly good data coverage, which facilitates a more refined image. Results presented in this chapter can also be found in Wehner et al. (submitted).

In Section 4.4, I describe how surface elevation (topography and bathymetry) and the fluid ocean are implemented, and demonstrate that their effects on synthetic waveforms become pronounced at periods $\leq$ 20 s. In this chapter, I present results from investigating path-dependent effects of surface elevation and the ocean on synthetics, and I compare the final model generated from a fluid ocean inversion to the tomographic result obtained with the frequently used ocean loading approximation (see Section 4.4.2). Furthermore, I highlight some of the key features of the final model, *SASSIER22*, including a detailed comparison with other recent tomographic models of the region.

6.1 Inversion setup

In this study, I focus on a subregion of Southeast Asia with good data coverage, which encompasses the regions around the Banda Arc, Borneo, Sulawesi and the southern segment of the Philippine Trench. The domain extension is shown in Figure 6.1 and comprises an area of approximately 2,800 km in the north-south direction and 2,500 km in the east-west direction, which represents a geographic coverage of $\sim$30 % of the continental-scale Southeast Asian study area presented in Chapter 5. Furthermore, the vertical mesh extension is decreased from 800 to 650 km due to the shallower depth sensitivity that arises from the use of shorter source-receiver distances.

Throughout this study, I closely follow the inversion setup described in Chapter 4 as summarised in Table 6.1. I use seismic data filtered at periods from 15 to 150 s and run two inversions: one accounting for the fluid ocean explicitly and one with the commonly used ocean loading approximation.

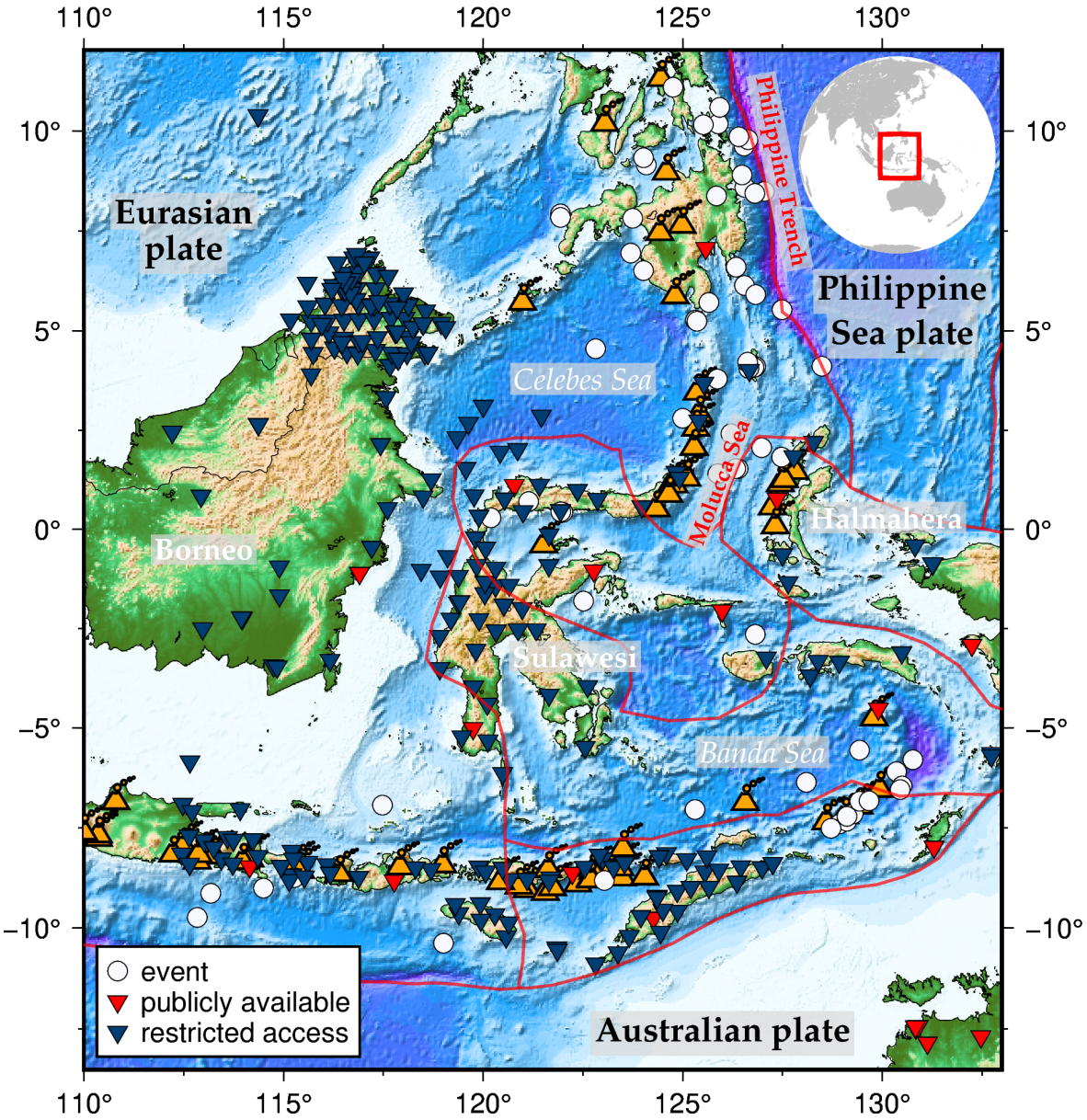


Figure 6.1: Map showing the study region (∼2,800 x 2,500 km), including the source-receiver distribution, for the eastern Indonesian full-waveform tomography. Publicly available stations are denoted by red triangles, stations with restricted access by blue triangles and events by white circles. Note that only stations used throughout the inversion are shown. The maximum source-receiver distance is ∼2,800 km. Orange symbols indicate volcanoes with eruptions during the Holocene (the past ∼12,000 years) as taken from the *Global Volcanism Program* (Venzke, 2013). Plate tectonic boundaries (red lines) are taken from Bird (2003). Topographic variations are extracted from *ETOPO1* (Amante and Eakins, 2009).

Table 6.1: Summary of the inversion settings for the Sulawesi waveform tomography study.

starting model	*SASSY21*	see Chapter 5
frequency-dependent attenuation	approximated with five linear solids (e.g. van Driel and Nissen-Meyer, 2014)	see Section 4.1.3
misfit function	time-frequency phase misfits (Fichtner et al., 2008)	see Section 3.3.1
data selection	manual, accounting for body-wave arrivals separately	see Section 4.2.5
station weighting	Ruan et al. (2019)	see Section 4.3
gradient preconditioning	1. source imprint removal 2. smoothing operator	see Section 4.5.1 see Section 4.5.2
optimisation scheme	trust-region based L-BFGS (e.g. Conn et al., 2000)	see Section 3.3.3
inversion parameters	v_P, v_{SH}, v_{SV} and ρ	see Section 4.6

Mesh generation

Table 6.2 presents an overview of the technical parameters used in the multi-scale approach (Bunks et al., 1995) employed in this study (see Section 3.3.3), which involves a gradual increase in the shorter-period content as the iterations progress (from 20 – 150 s to 15 – 150 s). I use 1.2 elements per minimum wavelength for all simulations and the mesh interpolation between period bands accounts for 3-D variations introduced throughout the inversion (see Section 5.4.2).

Table 6.2: Summary of the mesh generation displaying the number of iterations, number of mesh elements, time step for wavefield simulation in seconds and the duration of each earthquake simulation in seconds. The compute time per event in minutes in the last column is obtained by using the values in this table, but does not include storing wavefield snapshots for the adjoint simulation (see Section 3.3.2). Note that absolute run times are indicative since a simulation with exactly the same settings can have very different run times depending on the specifications of the computer it was executed on. The wavefield simulation was run in parallel on up to 56 cores for the actual inversion.

period band	**# iterations**	**# mesh elements**	**time step [s]**	**simulation length [s]**	**run time on 1 core [min]**
ocean layer:					
20 – 150 s	15	62,006	0.09	1,000	∼201
15 – 150 s	19	161,855	0.09	900	∼403
ocean load:					
15 – 150 s	19	153,223	0.14	900	∼275

I use *Salvus* (Afanasiev et al., 2019) to mesh the fluid ocean explicitly (see Section 4.4.2) for this study. Figure 6.2 shows the distribution of elastic and acoustic mesh elements for the fluid ocean inversion and visualises which parts of the domain are approximated by an equivalent load when the ocean becomes too shallow (here: < 1.5 km). For the ocean load inversion,

no acoustic elements are implemented but a boundary condition is used to approximate the ocean by the weight of its water column. Table 6.2 illustrates the computational cost associated with implementing an ocean load vs a fluid ocean for this simulation setup. Note how the number of mesh elements increases by < 10 % but the time step for the wavefield simulation is significantly reduced as a result of the low velocities introduced in the fluid ocean, thus leading to a considerable increase in computational cost.

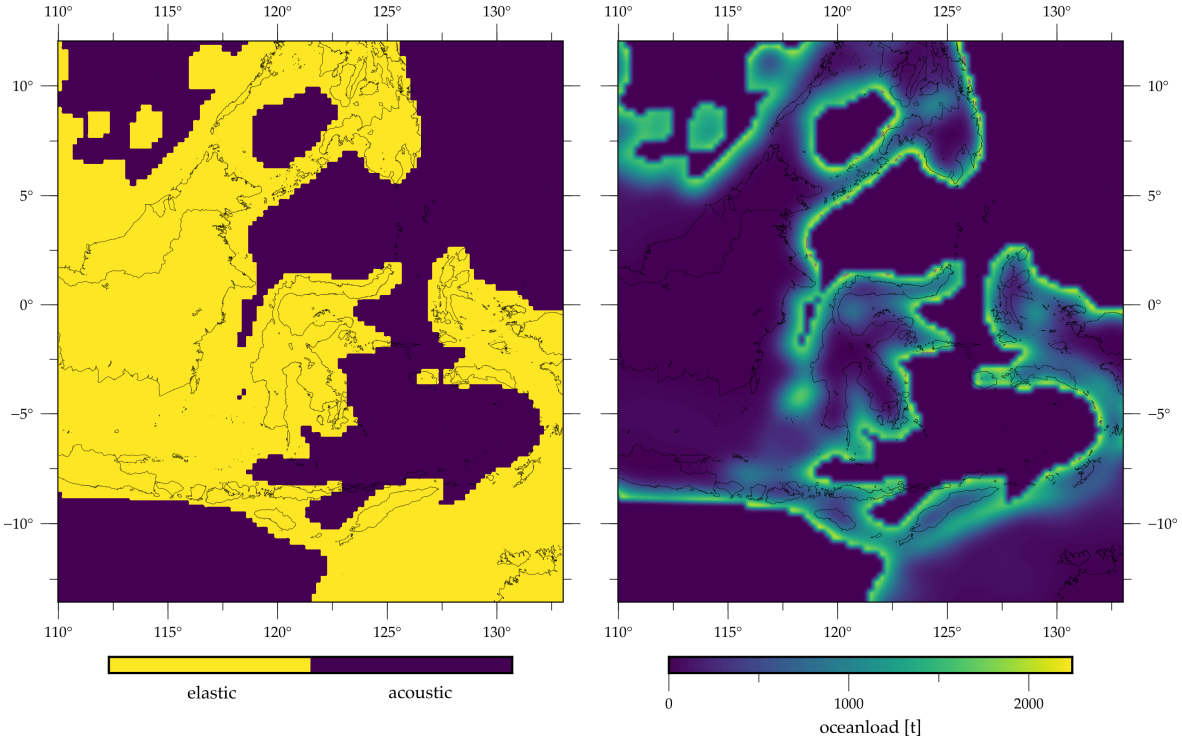


Figure 6.2: An impression of the 15 s mesh that explicitly meshes the fluid ocean. *Left:* Purple areas are acoustic (fluid), yellow areas are elastic (solid). *Right:* Area for which the ocean load (water column thickness times density) approximation is used, where the fluid ocean is < 1.5 km deep.

Event and data selection

To date, data from only a relatively small proportion of permanent network stations in Southeast Asia have been made publicly available, which are downloaded using *obspyDMT* (Hosseini and Sigloch, 2017). However, I have been able to include data from several networks with restricted access, resulting in an unprecedented dataset that comprises recordings from 255 stations within the study region (see Figure 6.1). In addition to the stations used to obtain the starting model *SASSY21* (see Section 4.2.1), data from two recent deployments with restricted access are added: 20 stations on Sulawesi (network code: *SL*) and 11 ocean bottom seismometers (network code: *5R*, Rawlinson et al., 2020) in the Celebes Sea. A detailed overview of the available stations is available in Section A4 of the Appendix.

The event selection process follows Section 4.2.4 and aims to produce an as uniform distribution of earthquakes throughout the study region as possible. At the same time, I avoid events potentially affected by interference with other events, mitigate finite-source effects by

removing large-magnitude earthquakes and review source time functions using *SCARDEC* by Vallée et al. (2011). Furthermore, recent events are preferred due to the deployment of temporary networks within the region, which increases the number of recordings and thus, enhances the efficiency of the adjoint-based inversion (see Section 3.3.2).

The final event catalogue for this study contains 67 earthquakes (see Figure 6.1) of magnitude $5.4 \leq M_w \leq 6.0$, which occurred between April 2014 and December 2020. It should be noted that only 20 out of the 67 events were used in the construction of *SASSY21*. Event locations and moment tensors are taken from the *GCMT* catalogue (Ekström et al., 2012) and remain constant throughout the inversion for reasons discussed in Section 5.4.1. A detailed overview of the events can be found in Section A1 of the Appendix.

In FWI, we aim to invert three-component seismograms, but noisy waveform segments and cycle skips can contaminate the tomography and need to be removed as discussed in more detail in Section 4.2.5. Here, I manually select seismogram portions ("windows") that are suitable for waveform comparison in order to

1. Avoid noisy portions of the seismogram.
2. Exploit as many waveforms as possible.
3. Optimise for depth sensitivity by accounting for small-amplitude body-wave arrivals separately.

In summary, this study uses ∼6,500 windows gathered from ∼7,100 unique source-receiver pairs resulting in $>$ 200 h of analysed waveform data per period band for the fluid ocean inversion (see Table 6.3). For the ocean load inversion, the same source-receiver pairs are used but the analysed window length is ∼22 % shorter due to the match between predicted and observed surface wave trains being of shorter duration on many components.

Table 6.3: Summary of the event and data selection that includes the number of events, number of windows, average number of windows per event, average window length per event in hours and number of unique source-receiver pairs per period band for the fluid ocean inversion. The same source-receiver pairs are used for the ocean load inversion.

period band	# events	# windows	avg. # windows per event	avg. window length per event [h]	unique source-receivers pairs
20 – 150 s	67	6,459	96	3.4	7,074
15 – 150 s	55	5,758	105	3.7	5,997

6.2 Effects arising from surface elevation and the ocean

In Section 4.4.2, I demonstrate frequency-dependent effects arising from surface elevation (topography and bathymetry) and the fluid ocean on synthetic waveforms, which shows that they become pronounced at periods $\leq$ 20 s. In this section, I undertake additional synthetic tests in order to illustrate path-dependent effects. I present three-component waveforms and misfit maps at 15 s for a synthetic grid of stations, which record three synthetic events with a diverse range of locations (on-shore and off-shore), depths and focal mechanisms (see Table 6.4). The effects of surface elevation and the fluid ocean are assessed separately.

Table 6.4: Source information used for synthetic tests. All events are M6.0, which generates sufficient energy across the domain. The last column indicates the misfit function used to produce the misfit maps presented in this section, which are explained in more detail in Section 3.3.1.

#	location	longitude	latitude	depth [km]	focal mechanism	misfit function
1	Australian continental shelf	130	-10.5	30	reverse fault	TF
2	Borneo (Sundaland shelf)	113	0	100	normal fault	L_2
3	Philippine Trench	127	9	10	strike-slip	CC

Synthetics waveforms are computed for a wavefield simulation time of 1,000 s, which allows us to capture the entire surface wave train for the largest possible source-receiver distance for this simulation setup. I use the anisotropic *PREM* (Dziewoński and Anderson, 1981) with the crust represented by *CRUST1.0* (Laske et al., 2013) as a background model since this introduces realistic 3-D structure with low velocities near the surface.

Surface topography and bathymetry

As described in Section 4.4.1, implementing surface elevation requires sophisticated meshing techniques but does not result in a notable increase in computational cost. For this reason, surface topography and bathymetry are implemented across all period bands to obtain the starting model *SASSY21* (see Section 4.1.1) and are also included in this study.

Figure 6.3 shows synthetic three-component seismograms for two stations at similar epicentral distances for each event in Table 6.4. For this simulation setup, implementing surface topography and bathymetry can result in a considerable phase advance and change in amplitude of the surface wave train at periods $\leq$ 20 s. It can have an effect on both horizontal (radial-transverse as well as east-north) and the vertical component. This is reinforced by Table 6.5, which shows the average misfit contribution per component for each event. In order to investigate path-dependent effects of surface elevation, I compute full-trace misfits as visualised via misfit maps in Figure 6.4. This shows that the effect is stronger for source-receiver paths passing through oceanic regions for this simulation setup, which is not surprising as bathymetry exhibits larger variations than topography in Southeast Asia. Thus, a stronger effect from oceans than mountain ranges can be observed.

Comparing the total event misfits shows that event 3 exhibits by far the largest misfit, with differences of 1 – 2 orders compared to event 1 and event 2. I suspect that this is caused by the shallow hypocentral depth of 10 km (see Table 6.4) and test this by placing the event at 100 km depth. This significantly reduces the event misfit to values of the same order as event 2 and thus, I conclude that surface topography and bathymetry have a larger effect for shallow events. Note that – without normalising the waveforms – an L_2 misfit is not diagnostic for this assessment since it utilises absolute amplitudes, which are naturally lower for deeper events. More detailed tests would be required to evaluate the effect of focal mechanisms and source locations (e.g. near trenches).

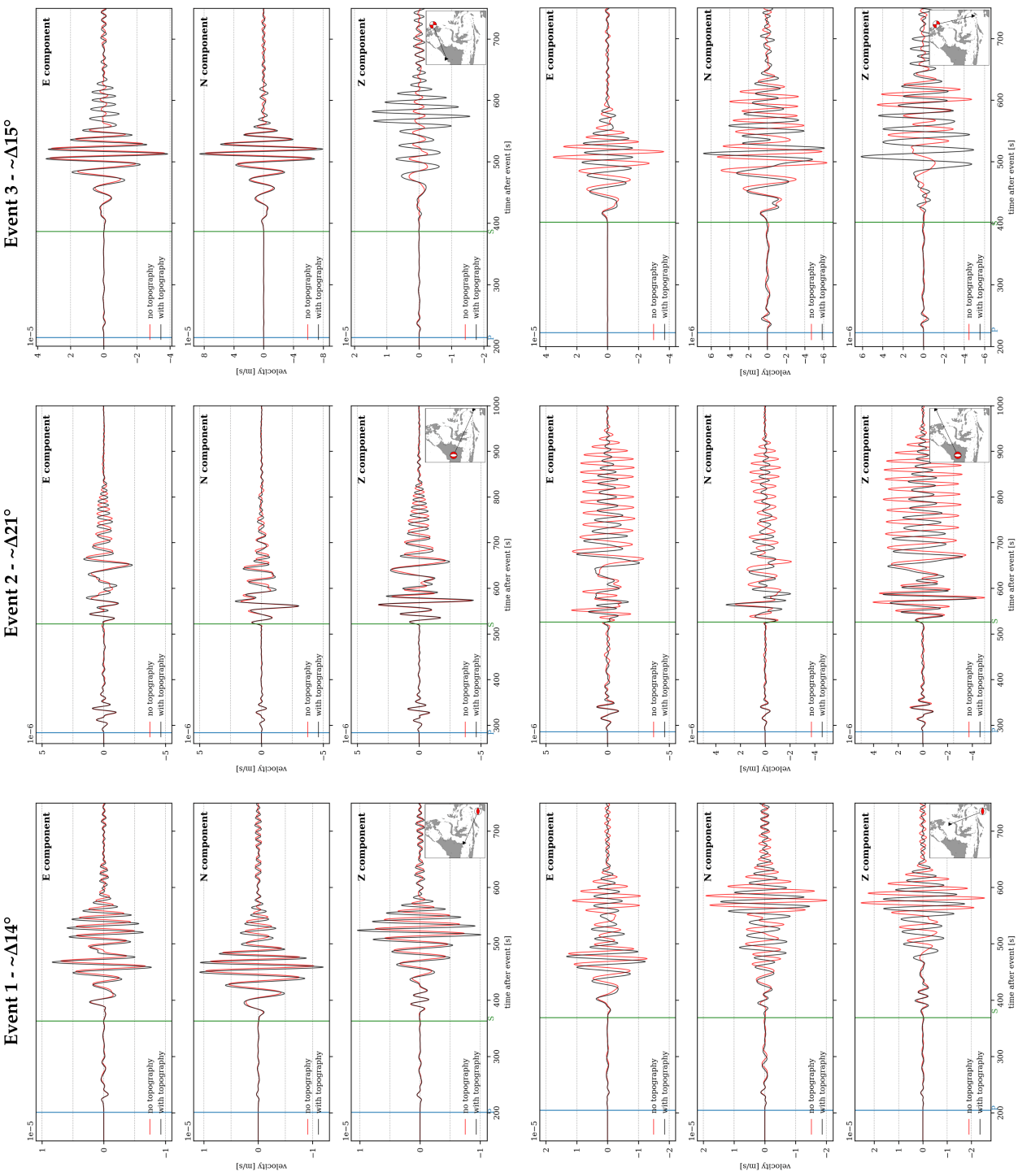


Figure 6.3: The effect of surface topography and bathymetry on three-component synthetic waveforms at 15 s. Two stations at similar epicentral distances are presented for each event in Table 6.4. The top row shows source-receivers pair with comparably small waveform differences, the bottom row shows source-receiver pairs with a significant effect. (*Left column*) event 1, (*middle column*) event 2 and (*right column*) event 3. Vertical lines indicate predicted P- (blue) and S-wave (green) first arrival times obtained from the *TauP* toolkit (Crotwell et al., 1999) for *PREM* (Dziewoński and Anderson, 1981).

Table 6.5: Surface topography and bathymetry synthetic tests: Misfit contribution per component for each event averaged over all stations shown in Figure 6.4.

component	event 1	event 2	event 3
E	26 %	34 %	40 %
N	28 %	33 %	34 %
Z	46 %	33 %	26 %

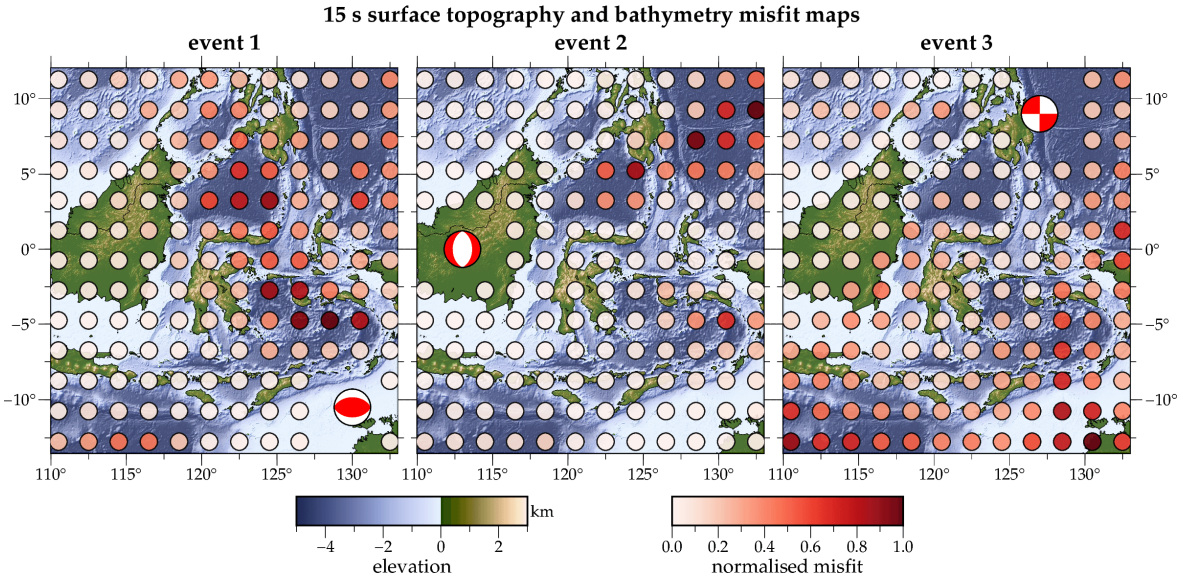


Figure 6.4: The effect of surface topography and bathymetry on synthetic waveforms at 15 s illustrated via misfit maps for the three events in Table 6.4. Misfits are computed on the full trace using the misfit function indicated by the rightmost column. Topographic variations are taken from *ETOPO1* (Amante and Eakins, 2009).

The fluid ocean

The ocean was approximated by an equivalent ocean load to obtain the starting model *SASSY21* (see Section 4.1.1). Figure 4.14 demonstrates that this a valid approximation at the periods considered for *SASSY21* (20 – 150 s) compared to explicitly modelling the fluid ocean, which is computationally more expensive. At shorter periods, however, simulating the fluid ocean has a noticeable effect on the later parts of the seismogram (phase delay, amplitudes and duration of the surface wave train).

I repeat the same tests presented in the previous section for the fluid ocean in order to investigate path-dependent effects. Surface elevation is implemented for both simulations with a free-surface condition (neglecting the ocean entirely) and simulations with a fluid ocean (the ocean is modelled as a fluid medium with acoustic elements). Figure 6.5 shows synthetic three-component seismograms for two stations at similar epicentral distances for each event in Table 6.4. The strongest effects are observed on the radial and vertical components (see Table 6.6), which suggests that the ocean's effects are confined to the source–receiver plane

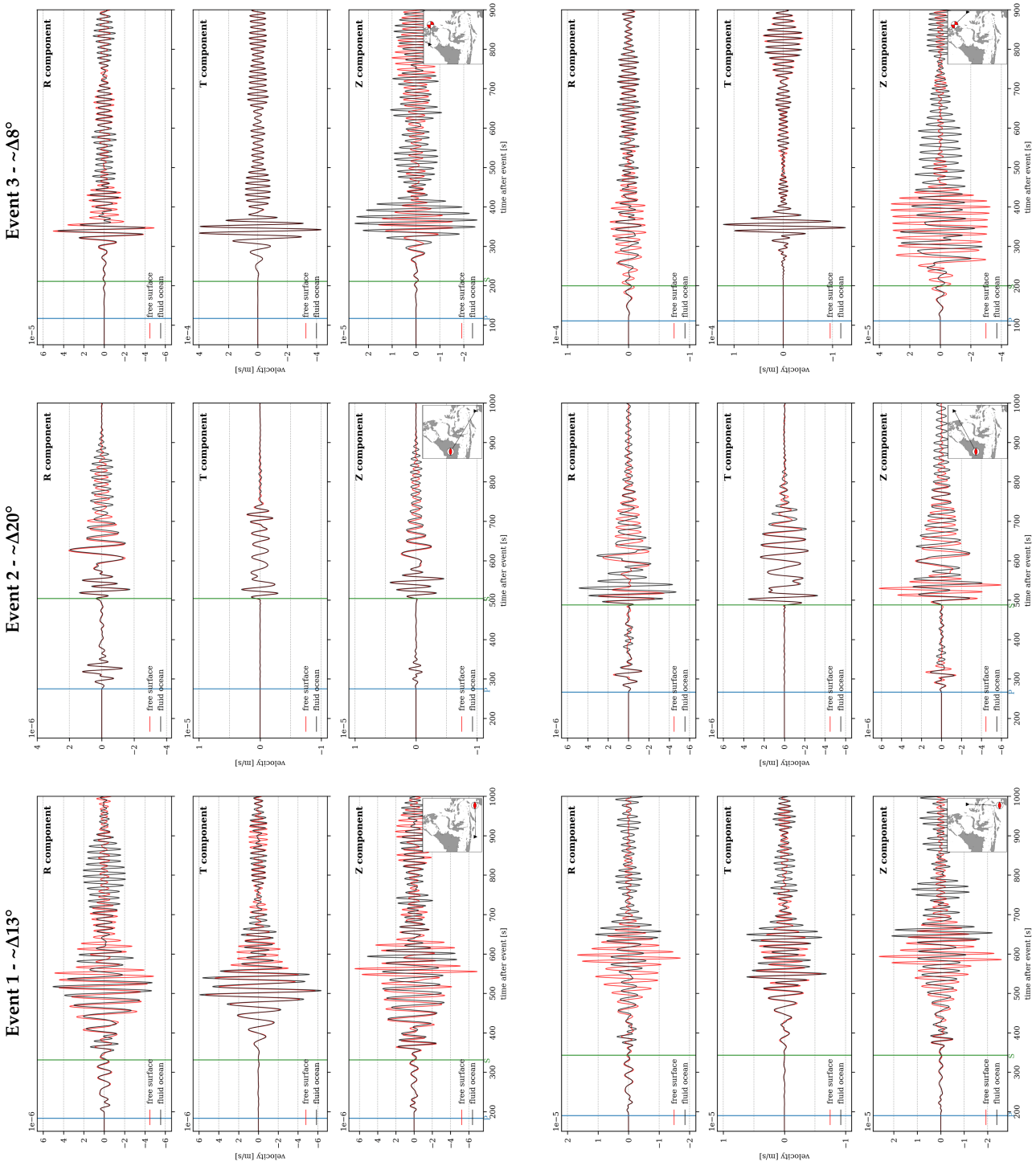


Figure 6.5: The effect of the fluid ocean on three-component synthetic waveforms at 15 s. Two stations at similar epicentral distances are presented for each event in Table 6.4. Note that the seismogram components are rotated from a North-East to a radial-transverse system. (*Left column*) event 1, (*middle column*) event 2 and (*right column*) event 3. Vertical lines indicate predicted P- (blue) and S-wave (green) first arrival times obtained from the *TauP* toolkit (Crotwell et al., 1999) for *PREM* (Dziewoński and Anderson, 1981).

(Fernando et al., 2020). Note that this was not the case for the surface elevation tests presented in the previous section, where both radial and transverse components can be likewise affected. Figure 6.6 illustrates the path-dependent effect of the fluid ocean on synthetic waveforms via misfit maps. While the misfit calculation is non-linear, misfit values tend to be higher at larger epicentral distances and near the Philippines – the deepest location within the study region – but also the Banda Sea and Celebes Sea. Preliminary tests do not show an obvious correlation of the event misfits with respect to their location or depth.

Table 6.6: Fluid ocean synthetic tests: Misfit contribution per component averaged over all stations for each event shown in Figure 6.4.

component	event 1	event 2	event 3
R	58 %	32 %	28 %
T	1 %	2 %	4 %
Z	41 %	66 %	68 %

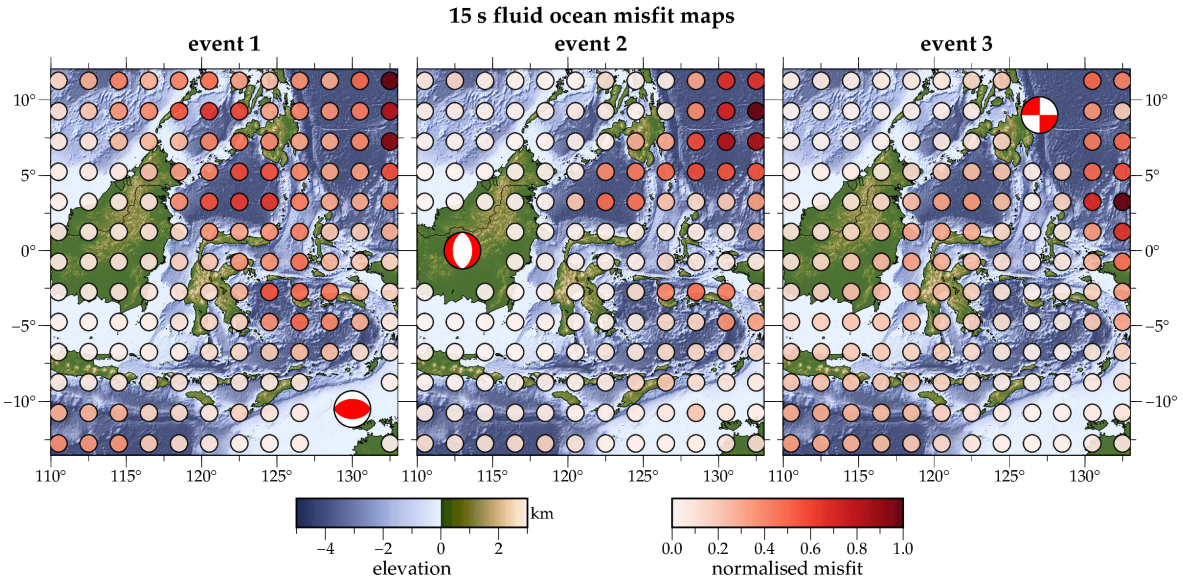


Figure 6.6: The effect of the fluid ocean on synthetic waveforms at 15 s illustrated via misfit maps for the three events in Table 6.4. Misfits are computed on the full trace using the misfit function indicated by Table 6.4's last column. Topographic variations are taken from *ETOPO1* (Amante and Eakins, 2009).

Note that I test the misfit maps, event misfits and component-dependent values presented for three misfit functions (time-frequency phase misfits, L_2 and cross-correlation time shifts – see Section 3.3.1), and for different minimum periods (50 s, 30 s, 20 s and 15 s). The overall pattern is very similar, but individual station misfits can vary because the misfit calculation is non-linear. Furthermore, a cross-correlation time shift is not able to capture the full complexity of the later parts of the surface wave train and cycle skips on some components are observed at shorter periods, which can result in misleading misfit values.

For comparison, Figure 6.7 shows the misfit maps at 50 s using the same parameters to obtain the misfit maps in Figure 6.6. Note that a path-dependent effect for the deep ocean near the Philippines can already be observed at these long periods, even though there is only a minor difference in synthetic waveforms (see Figure 4.14), which would be below the noise level for any real-world application. At 15 s, an increased sensitivity to other oceanic regions such as the Banda Sea and Celebes Sea is observed (see Figure 6.6).

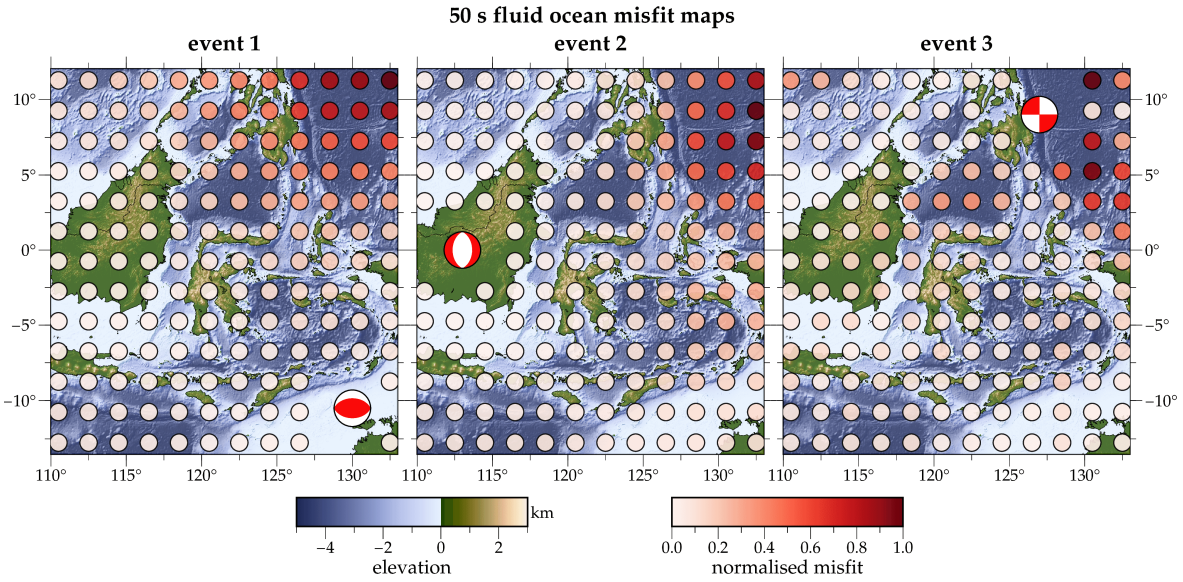


Figure 6.7: The effect of the fluid ocean on synthetic waveforms at 50 s illustrated via misfit maps for the three events in Table 6.4. Misfits are computed on the full trace using the misfit function indicated by Table 6.4's last column. Topographic variations are taken from *ETOPO1* (Amante and Eakins, 2009).

6.3 Results and Discussion

For the inversion setup described in Section 6.1, a total of 34 L-BFGS iterations divided over two period bands (15 iterations at 20 – 150 s and 19 iterations at 15 – 150 s) are carried out, accounting for the fluid ocean explicitly. In the following, I will refer to this model as *SASSIER22*. Furthermore, I run 19 iterations at 15 – 150 s with the fluid ocean approximated by its weight ("ocean load") using the same source-receiver pairs, which allows for a direct comparison of the tomographic results. In this section, I present the inversion results, including a model validation, some key highlights of the final model *SASSIER22*, and a comparison with the model obtained from ocean loading and with other recent tomographic models of the region.

6.3.1 Model validation

Once a model has been determined, we would like to assess the robustness of the solution but this is non-trivial in FWI as discussed in Section 3.4. However, one can pursue more data-driven approaches towards validating the final model as presented below.

Misfit development

Figure 6.8 presents the misfit decrease per period band, which indicates that the waveform difference is successively reduced throughout the inversion, for both the fluid ocean and ocean loading. The overall misfit, based on the entire dataset, decreases by ∼40 % and ∼34 % for the fluid ocean and ocean load, respectively. However, ocean loading shows a ∼10 % higher misfit value, even though the analysed window length is ∼22 % less compared to the fluid ocean inversion due to the match between predicted and observed surface wave trains being of shorter duration on many components. Furthermore, a slightly stronger misfit decrease for ocean load can be observed, which I attribute to a combination of the following:

- I do not run any simulations at 20 s for ocean loading since *SASSY21* already ran 17 iterations at this period band for continental-scale Southeast Asia (see Figure 5.1). Thus, some of the misfit decrease achieved during the additional 15 iterations for ocean layering may have been accounted for in the 15 s ocean load period band.
- During the fluid ocean inversion, the ocean layer and the mesh element layer below are not updated to avoid potential artefacts in the gradient from the smoothing preconditioner (see Table 6.1) and the interface condition. However, this only affects the upper few kilometers.

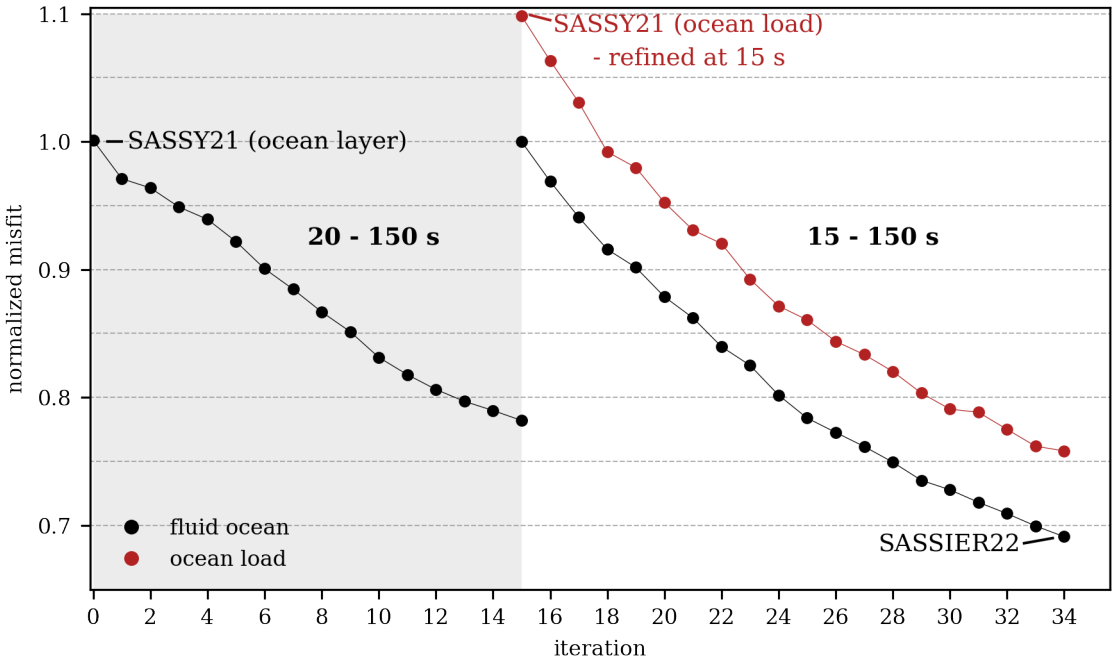


Figure 6.8: Misfit development across 34 iterations, normalised by the initial misfit within each period band. Black and red dots are normalised misfit values for the fluid ocean and ocean load inversion, respectively. The change to the 15 – 150 s period band is accompanied by a mesh interpolation and data review.

Analysed window length

Table 6.3 shows an increase in the window length per event as the period band is broadened, thus indicating that more data per event is successively included as the iterations progress. For comparison, I use the automated data selection algorithm *FLEXWIN* (Maggi et al., 2009) to suggest windows for the starting and final model for both ocean load and ocean layer; *FLEXWIN* selects a similar total window length for the starting model for both ocean types, which reinforces that ocean loading is a valid approximation at 20 s. However, the window length for ocean layer increases by 63 %, while the increase for ocean load is only 39 % for the final model, indicating that the model with an ocean layer matches observed waveforms better at 15 s.

Validation dataset

I test the validity of *SASSIER22* for three events not used in the inversion as shown in Figure 6.9. This also shows that *SASSIER22* is able to explain true-amplitude waveforms from network *7G* (Passarelli et al., 2016) – see source-receiver pair #1 –, which only recently became available and was thus not used to obtain the starting model *SASSY21* or *SASSIER22*. Note that the wavefield gets significantly more complex at these periods (compared to ∼30, 40 s) and that the minimum period only decreased from 20 to 15 s.

Misfit increase per parameter

For *SASSIER22*, I test each inversion parameter separately by replacing the corresponding parameter in the final model by the values of the starting model and evaluate the misfit increase. This shows that shear-wave velocity drives the inversion (v_{SV}: + 24 %, v_{SH}: + 18 %) but the updates in v_P (+ 7 %) and density (+ 4 %) are also required to explain observed waveforms.

6.3.2 Final model: *SASSIER22*

The final tomographic model is obtained in a period band for which the wavefield is surface-wave dominated. However, P-wave velocity shows meaningful updates down to the mantle transition zone; for example, higher velocities beneath the Celebes Sea and hints of subduction along the Indonesian volcanic arc, North Sulawesi Trench and Philippine Trench can be observed. Nonetheless, the P-wave model is not as clear as the S-wave model, which I largely attribute to the broader sensitivity kernels of P-waves. As a consequence, the following interpretation is based on shear-wave velocity v_S, which is computed using the Voigt average: $v_S = \sqrt{(2{v_{SV}}^2 + {v_{SH}}^2)/3}$ (Babuska and Cara, 1991; Panning and Romanowicz, 2006). The results for all inversion parameters are presented in Section A5 of the Appendix.

Figure 6.10 shows v_S depth slices for the starting model *SASSY21* and the final model *SASSIER22* from 25 to 300 km. The depth slices at 25 and 50 km show that *SASSIER22* is significantly more sensitive to shallower structure as a result of the shorter-period data considered. For instance, one can distinguish between high velocities arising from oceanic

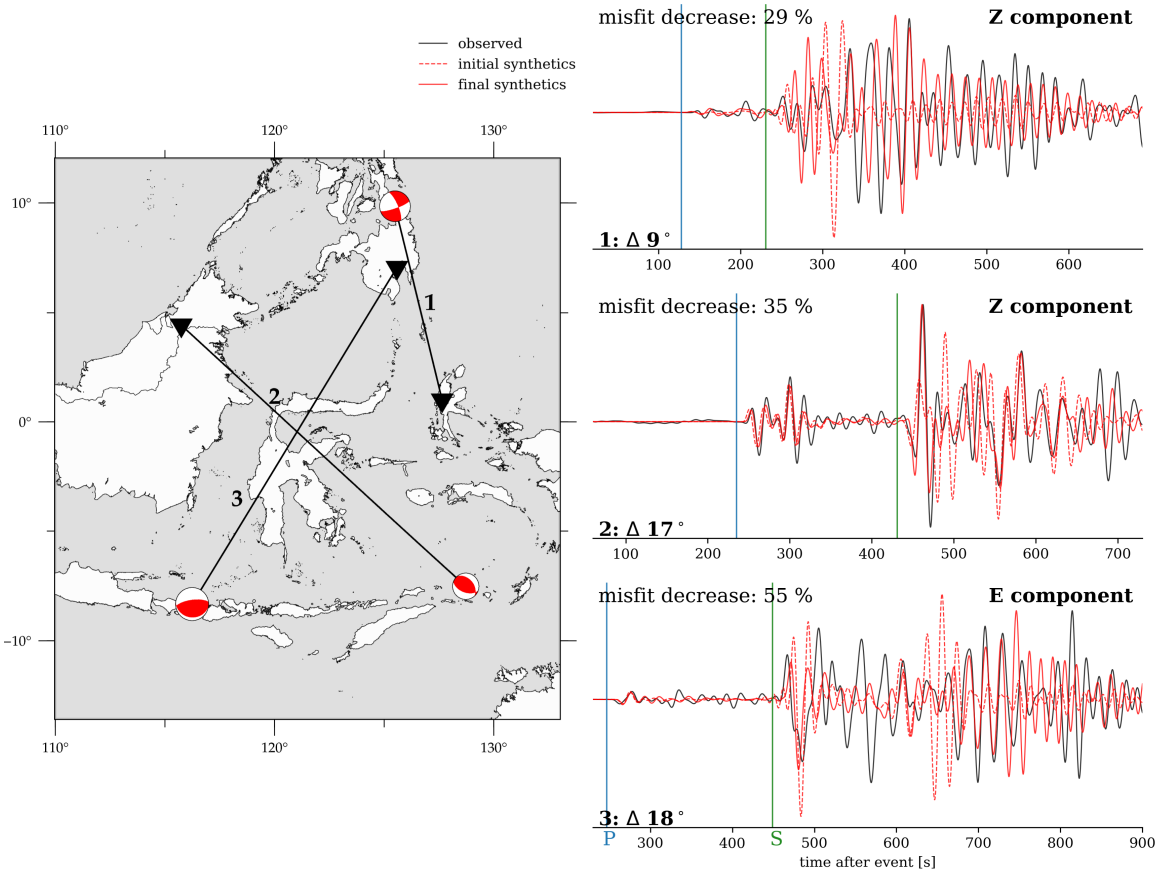


Figure 6.9: *Left:* Map of three validation earthquakes (M_w5.5 – 6.0). *Right:* Vertical and horizontal seismogram traces showing the waveform match improvement between the initial model (*SASSY21*) and the final model (*SASSIER22*) for three source-receiver pairs. Epicentral distances are given in the bottom left corner of each trace. The final model *SASSIER22* (iteration 34) satisfies the data better than the starting model (iteration 0) as indicated by the full-trace time-frequency phase misfit (Fichtner et al., 2008) decrease in the upper left corner of each trace. Vertical lines indicate predicted P- (blue) and S-wave (green) first arrival times obtained from the *TauP* toolkit (Crotwell et al., 1999) for *PREM* (Dziewoński and Anderson, 1981).

lithosphere (e.g. beneath the Celebes Sea) and low velocities arising from continental lithosphere (e.g. beneath Borneo). At greater depths ($\geq$ 100 km), significant updates around the Philippine Trench are observed. Furthermore, the starting and final model agree on high velocities along the northern arm of Sulawesi and the Indonesian volcanic arc, including a hole at ~300 km depth. The hole may be associated with the pertusion of continental lithosphere from the northward moving Australian plate as discussed in Section 5.3.3. The 180° curvature around the Banda Arc is still imaged as a single bent and deformed slab (see Section 5.3.2), and the steeply dipping slab associated with the North Sulawesi Trench is still tracked down to the mantle transition zone (see Section 5.3.5).

Figure 6.11 presents a comparison between the starting and final models of this study, and other continental- and global-scale tomography models for a west-east cross-section through the southern end of the Philippine Trench. For *SASSIER22*, a convergent double subduction in this region is visible (see Section 2.5), which extends from ~1.5 – 6°N latitude

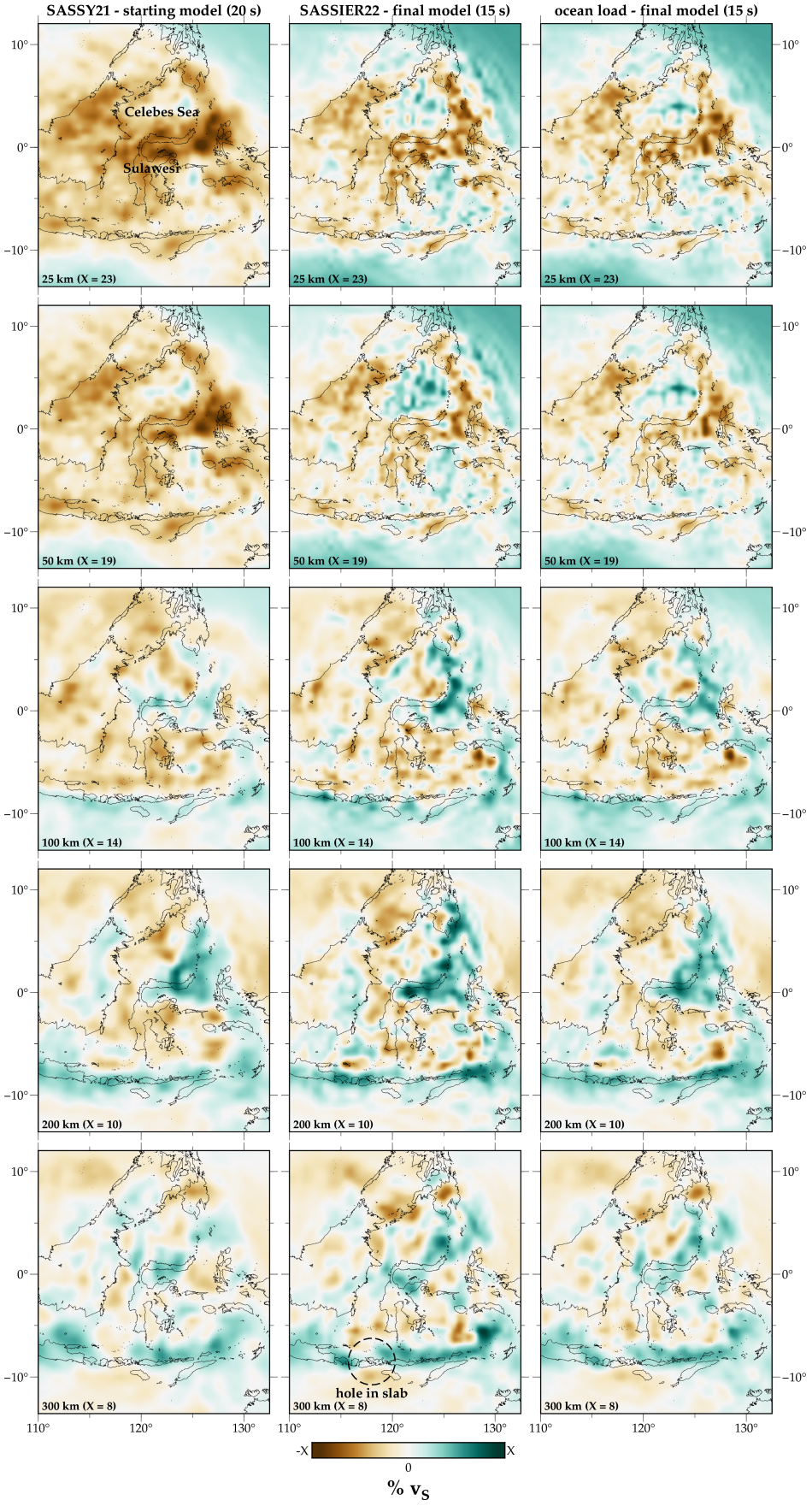


Figure 6.10: Shear-wave velocity v_S depth slices from 25 to 300 km for (*left column*) the starting model *SASSY21*, (*middle column*) the final model *SASSIER22* obtained by simulating the fluid ocean explicitly and (*right column*) the final model obtained using ocean loading. Perturbations are in % relative to the depth-average. The limits of the colorscale X are shown in the lower left corner of each plot.

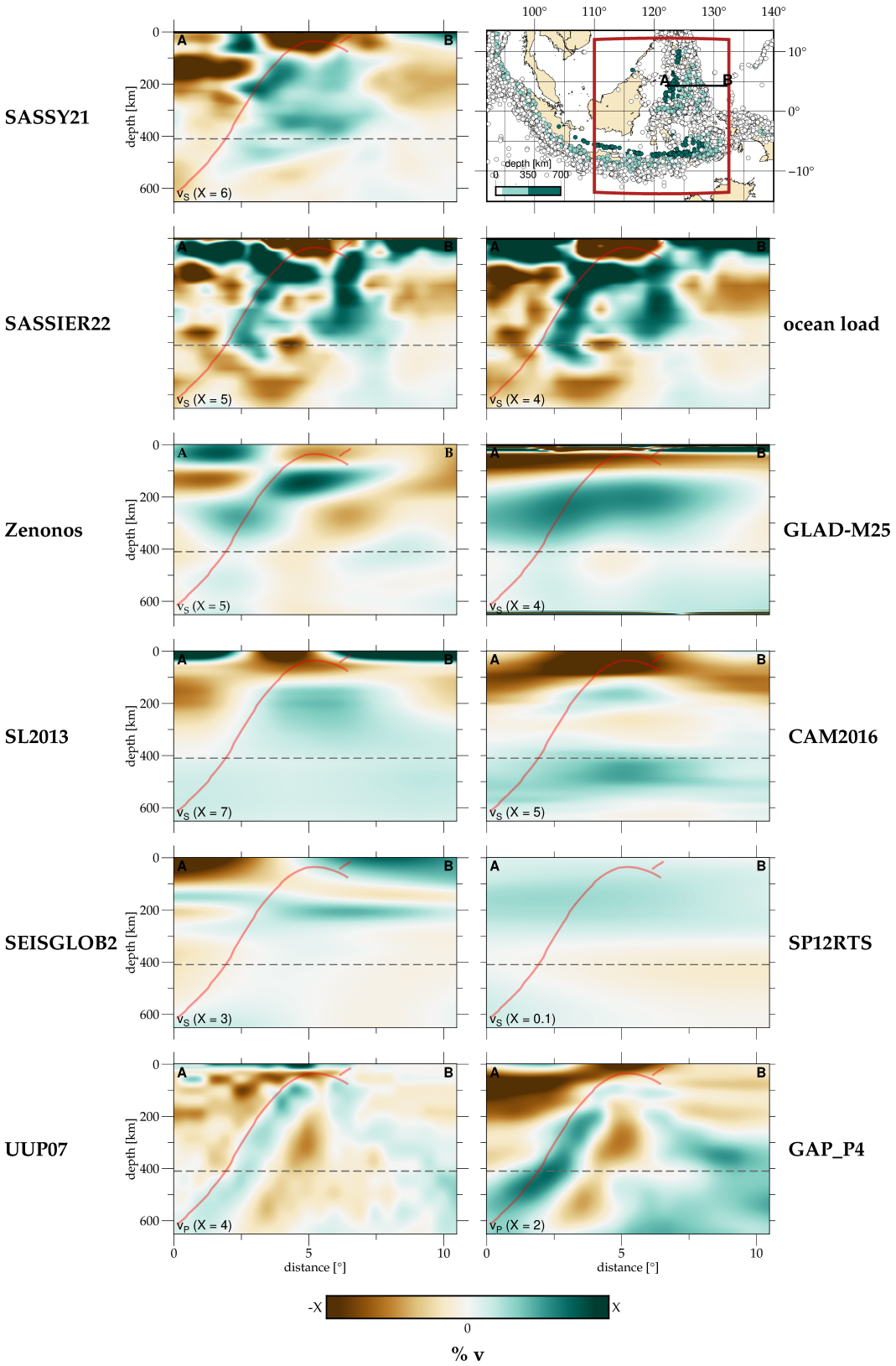


Figure 6.11: Comparison between the starting and final models of this study, and various regional and global tomographic models for an east-west cross-section through the Philippine Trench. An overview of the different models can be found in Table 5.2. Note that some models are v_P (bottom two panels) and some are v_S and thus, a direct comparison is not always warranted. 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. Red lines in the sections indicate slab depths taken from *Slab2* (Hayes et al., 2018), which is derived from the distribution of earthquake locations in this region. The map in the top right corner shows the section's location and seismicity ($M_W > 5$) taken from the *ISC* catalogue (International Seismological Centre, 2016), coloured by depth. The dark red box indicates the study area for *SASSIER22*.

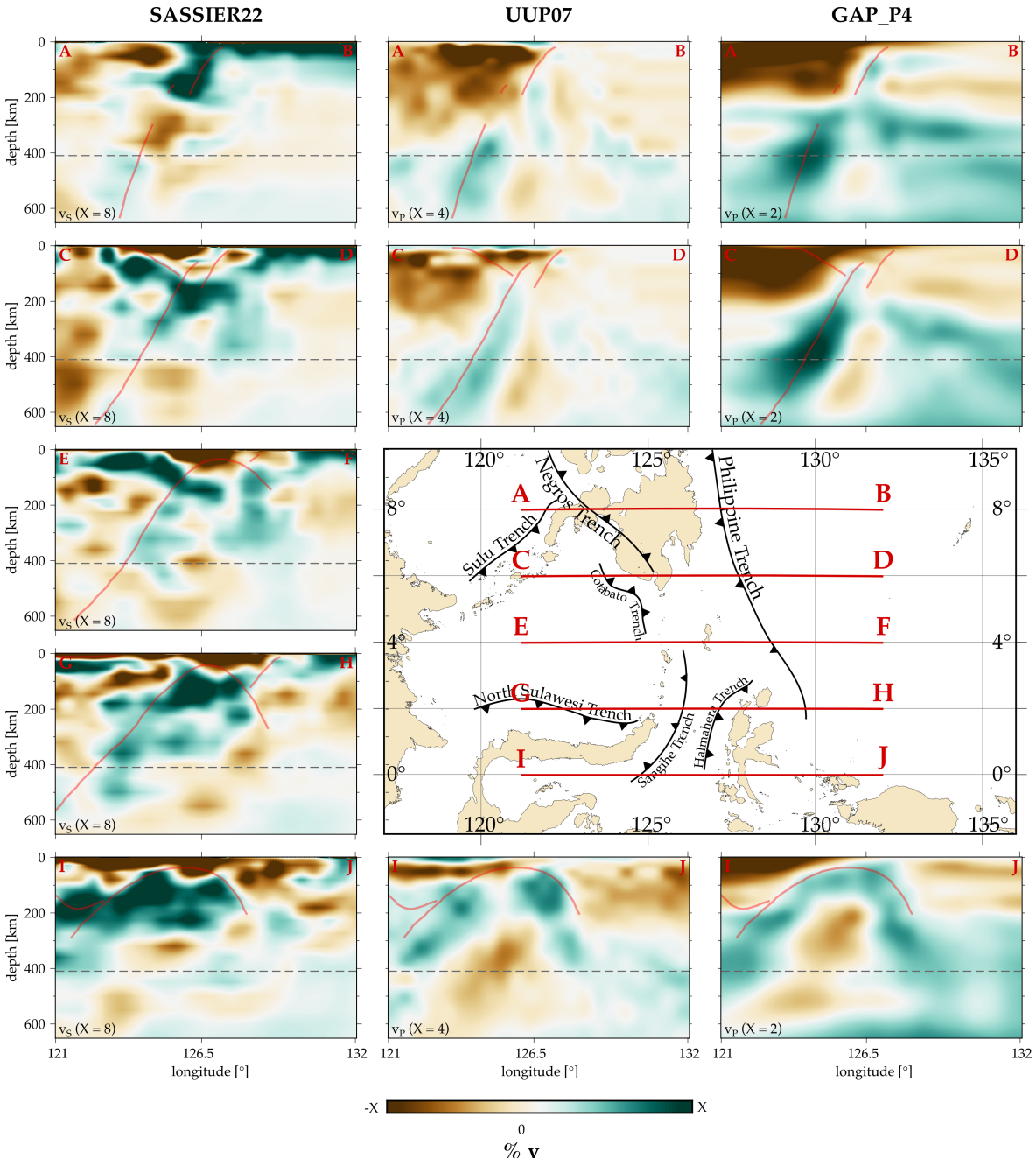


Figure 6.12: Cross-section from 0 – 8°N latitude through *SASSIER22*, *UU-P07* and *GAP_P4*. Perturbations are in % relative to the depth-average. The limits of the colorscale *X* are shown in the lower left corner of each plot. Note that *SASSIER22* is v_S but *UUP07* (Amaru, 2007) and *GAP_P4* (Obayashi et al., 2013) are P-wave models and thus, a direct comparison is not always warranted. Red lines in the sections indicate slab depths taken from *Slab2* (Hayes et al., 2018). The tectonic map on the right is modified from Aurelio (2000).

and was not apparent in the starting model *SASSY21*. In particular, a westward dipping slab beneath the southern end of the Philippine Trench and an eastward dipping slab that exhibits a reversal in subduction polarity from 100 km down to the mantle transition zone are observed. Such a subduction polarity flip has previously been suggested for settings where two oceanic plates collide (Su et al., 2019; Gasser et al., 2021). Chen et al. (2020) suggest an east-dipping slab from 80 to 300 km depth at 1 – 6°N latitude based on the *ISC-EHB* (Engdahl et al., 2020) earthquake distribution for this region. Other shear-wave models do not show a clear high-velocity anomaly but a divergent double subduction is visible in P-wave models (Amaru, 2007; Obayashi et al., 2013) further south. Hints of this inverted U-shaped subduction zone from -0.5 – 1.5°N latitude can be observed in *SASSIER22*, which is consistent with *UUP07* (Amaru, 2007) and *GAP-P4* (Obayashi et al., 2013).

Figure 6.12 presents further cross-sections from 0 – 8°N latitude showing the transition from a divergent subduction system in the Molucca Sea in the south to a convergent subduction system around the Philippine Trench further north. Note that the convergent double subduction system is further complicated by the eastward subduction of two plates – along the Cotabato Trench and Negros Trench – as shown in the map in Figure 6.12. However, this complexity is not evident in the slab model *Slab2* (see Figure 2.5, Hayes et al., 2018) nor the plate tectonic boundary model by Bird (2003) as shown in Figure 6.1.

6.3.3 The effect of the ocean on the tomographic model: Ocean load vs fluid ocean

Figure 6.10 presents a comparison of the tomographic models obtained from ocean loading vs ocean layer at iteration 19 and 34, respectively, for depth slices from 25 to 300 km. The main differences appear to be in the upper ∼50 km, with higher velocities obtained beneath the Celebes and Banda Sea for the fluid ocean model, as expected for oceanic lithosphere. However, Figure 6.11 shows that there are also significant differences at greater depths, as illustrated via a comparison of cross-sections through the double subduction around the southern segment of the Philippine Trench, which was not apparent in the starting model. Both ocean types are able to recover the double subduction, but the slab associated with subduction along the Philippine Trench is clearer in *SASSIER22*.

I use *Salvus* to mesh irregularly patched oceans (rather than a global ocean layer of constant thickness as in Fernando et al., 2020), which allows me to investigate waveforms along different source-receiver paths. Figure 6.13 presents several waveform matches for observed data with synthetics obtained from both ocean types. This shows that there is a minor difference between synthetic waveforms for a source-receiver path that mostly passes through a continental shelf (#3), while the synthetics from the final fluid ocean model are able to explain true-amplitude observed surface wave trains significantly better than synthetics from the final ocean loading model for source-receiver paths passing partially (#2) or entirely (#1) through oceanic regions. Note that the receiver for source-receiver pair #1 is an ocean-bottom seismometer.

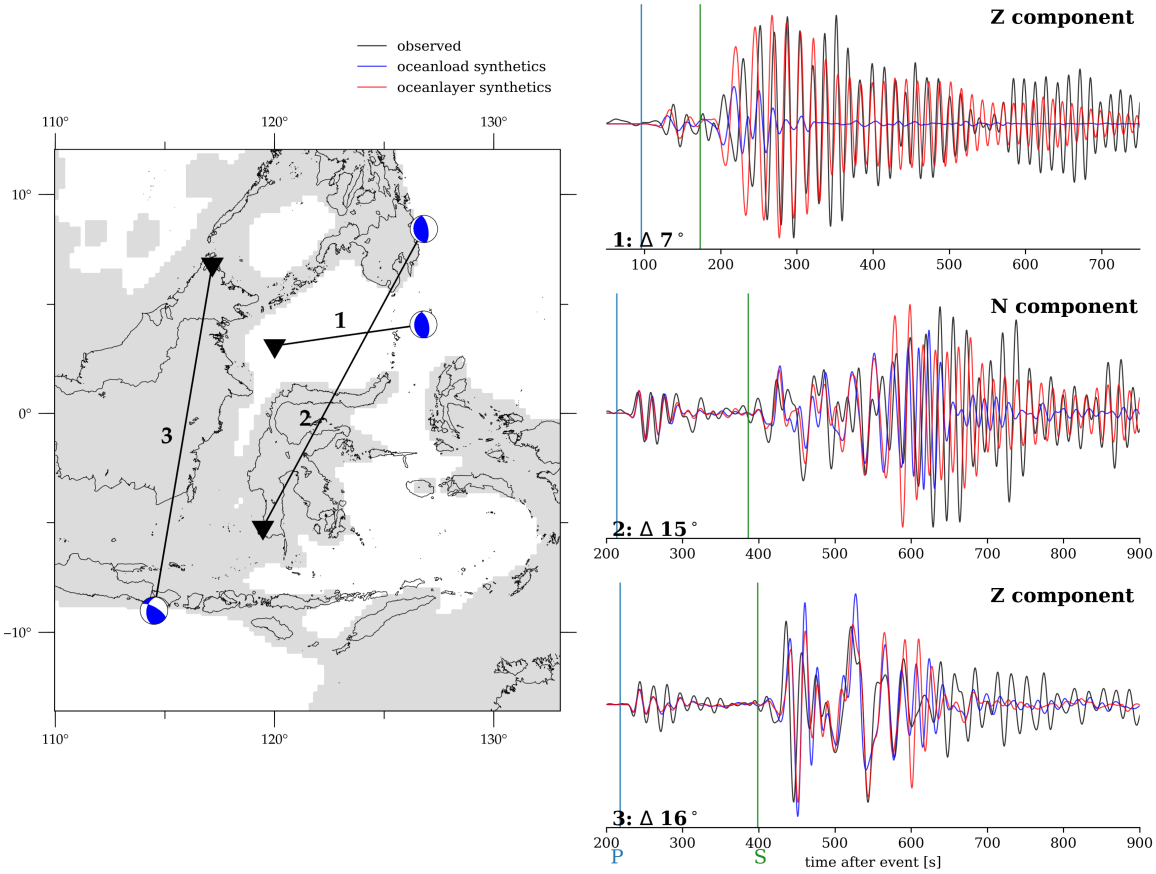


Figure 6.13: *Left:* Map of three earthquakes (M_w5.6 – 5.8). The white area indicates where the fluid ocean is $>$ 1.5 km deep and thus, meshed explicitly to obtain *SASSIER22*. *Right:* Vertical and horizontal seismogram traces showing the waveform match at 15 s between the observed waveforms (black), synthetics from the final model obtained using the ocean load approximation (blue) and synthetics from the final model that meshes the fluid ocean explicitly (*SASSIER22*, red) for the three source-receiver pairs shown on the left. Epicentral distances are given in the bottom left corner of each trace. Vertical lines indicate predicted P- (blue) and S-wave (green) first arrival times obtained from the *TauP* toolkit (Crotwell et al., 1999) for *PREM* (Dziewoński and Anderson, 1981).

6.3.4 Limitations

For reasons discussed in Section 5.4, I do not invert for source parameters and any parameters additional to v_P, v_{SH}, v_{SV} and ρ. Furthermore, I only investigate the effects of surface topography, bathymetry and the fluid ocean for a Southeast Asian simulation setup at 15 s at a regional scale. Thus, the applicability of my results for other regions and different scales remains limited, but the results are in good agreement with Fernando et al. (2020), who consider epicentral distances of up to 90° across the Pacific Ocean. While the minimum period band is only decreased from 20 to 15 s compared to *SASSY21*, I expect that the effects of surface elevation and the fluid ocean become even more significant at shorter periods and when running additional iterations. However, this would require an improved data coverage, e.g. around the Philippines.

7 Future work: Icelandic waveform tomography

The aim of this project is to exploit the wealth of seismic data collected in Iceland over the past 25 years to constrain a regional 3-D seismic structural model of the crust and upper mantle beneath Iceland using full-waveform inversion down to periods of a few seconds[1]. The resulting 3-D model will overcome the inherent limitations of tomographic inversions based on ray theory (see Section 3.1.2), allowing us to better recover the strong velocity contrasts that characterise the Icelandic lithosphere. This will provide new insights into subsurface structure, improve reference models for microseismic analysis and future tomographic studies, and permit significantly more accurate Green's function calculations for waveform-based seismic source inversion.

It should be noted that the Icelandic study area is significantly smaller than the Southeast Asian studies described in the previous chapters; for example, the geographic coverage is < 2 % of the continental-scale Southeast Asian study region presented in Chapter 5. The smaller domain size makes it computationally feasible to consider shorter-period data, which translates to the recovery of smaller-scale features. Moreover, one can expect increased sensitivity to crustal structure, topography and source parameters and thus, initial input needs to be chosen with great care. Compared to Southeast Asia, Iceland is relatively well-studied and thus, a wealth of prior information such as robust earthquake locations and crustal models are available.

Overall, Iceland represents a very contrasting case study to Southeast Asia, both in terms of its tectonic setting and the challenges of implementing adjoint waveform tomography. As such, it helps to extend the breadth of my experience and understanding of this field, and arguably broadens the interest of my dissertation to the readership. In the following, I will present preliminary results for my Icelandic adjoint waveform tomography study, which includes a seismic moment tensor inversion and a test inversion for seismic structure. I will also discuss data coverage and availability.

[1]For comparison, Chow et al. (2022) apply FWI to the northern Island of New Zealand – which has a very similar domain size – and they use ∼60 events, with a minimum period of 4 s.

7.1 Tectonic setting and previous work

Iceland lies in the North Atlantic Ocean and straddles the mid-ocean ridge (MOR) that separates the North American and Eurasian tectonic plates (see Figure 7.1). It is widely accepted that an underlying mantle plume produces positive dynamic topography and a thicker crust (via enhanced melt production), which exposes Iceland above sea level (Morgan, 1971; White and McKenzie, 1989). On shore, the MOR broadly separates into three volcanic zones (northern, eastern and western, e.g. Thordarson and Larsen, 2007) along which frequent seismic and volcanic activity occurs (such as the 2021 Fagradalsfjall eruption on the Reykjanes Peninsula, see Figure 7.2).

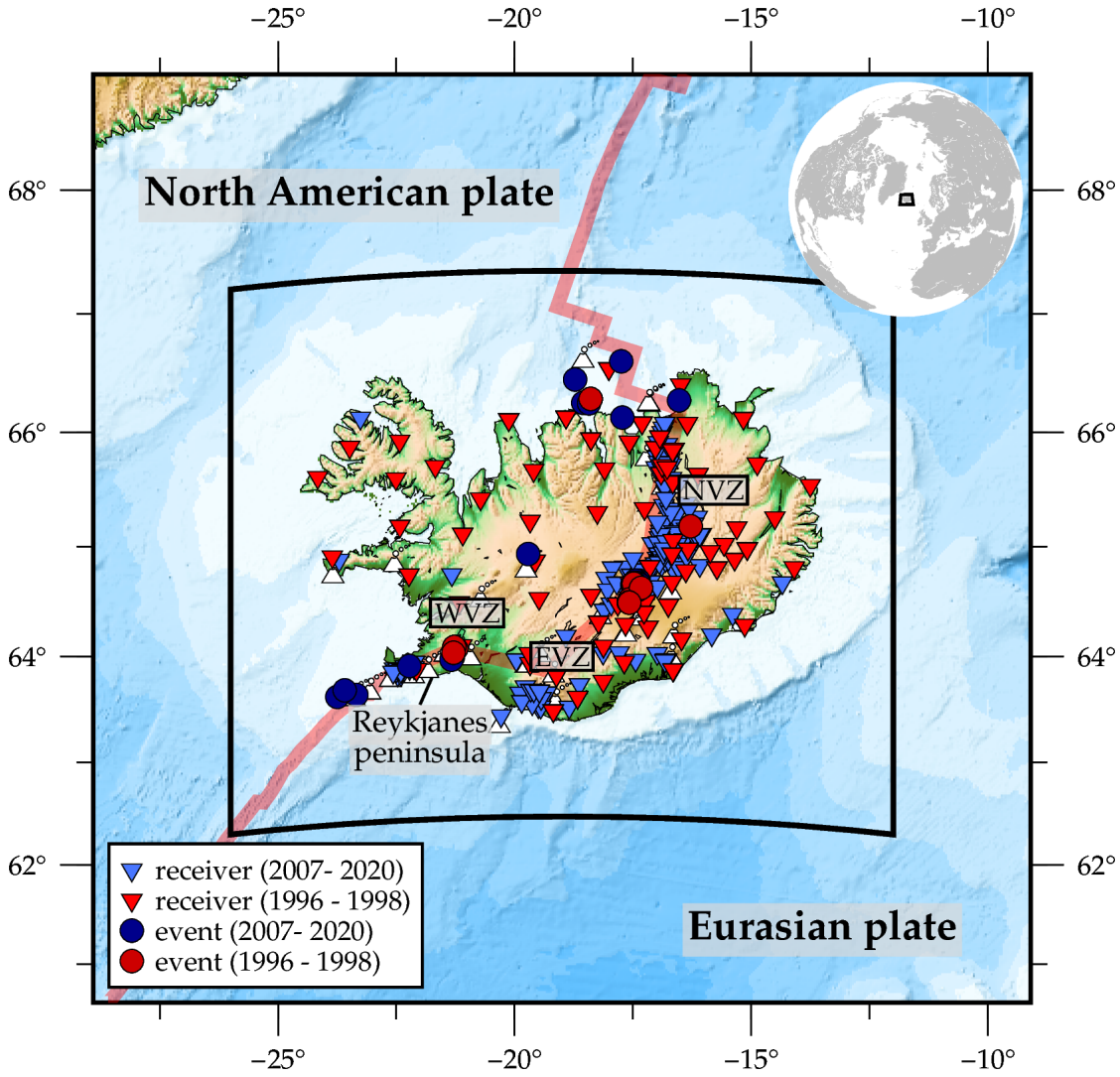


Figure 7.1: Map showing the study region (black rectangle) including the source-receiver distribution for the Icelandic adjoint waveform tomography: Stations are denoted by inverted triangles and events by circles. Note that only stations used throughout the test inversion (see Section 7.2) are shown. *EVZ*, *NVZ* and *WVZ* indicate the eastern, northern and western volcanic zone, respectively. The plate tectonic boundary (red line) is taken from Bird (2003). Topographic variations are extracted from *ETOPO1* (Amante and Eakins, 2009). White symbols in the background indicate volcanoes that erupted during the Holocene (the past ∼12,000 years) as taken from the *Global Volcanism Program* (Venzke, 2013).

So far, several active and passive seismic tomography studies have imaged the upper ∼25 km beneath Iceland from a local (e.g. Greenfield et al., 2016; Obermann et al., 2016) to a regional (e.g. Tryggvason et al., 2002; Green et al., 2017; Volk, 2021) scale. Green et al. (2017) and Volk (2021) reveal a spatial correlation between low seismic velocities and active volcanic rift zones, and faster velocities with the older, cooler, less fractured, non-volcanically-active crust. Active source experiments reveal a highly variable crustal thickness of between 20 and 40 km across the island and anomalously high P-wave velocities in the lower crust (Staples et al., 1997; Brandsdóttir et al., 1997; Darbyshire et al., 1998; Darbyshire et al., 2000). A regional P-wave tomography model by Yang and Shen (2005) reveals a broad low-velocity anomaly in the middle and lower crust underlying a high-velocity body in the shallow crust in central Iceland. Several tomographic studies at different scales image a large low-velocity anomaly in the upper mantle, which is associated with the upwelling mantle plume (e.g. Wolfe et al., 1997; Bijwaard and Spakman, 1999; Foulger et al., 2001; Allen et al., 2002). However, recent tomographic models suggest that the plume extends into the lower mantle (e.g. Rickers et al., 2013; French and Romanowicz, 2014; French and Romanowicz, 2015) but its deep part may be located beneath closer to Greenland and not directly beneath Iceland (Celli et al., 2021).

7.2 Test inversion

One of the main challenges for regional tomography beneath Iceland is the heterogeneous source-receiver distribution, which will be discussed in more detail in Section 7.3. Therefore, as a first step, I explore the feasibility of a regional-scale Icelandic FWI by investigating the available dataset and its associated depth sensitivity.

I run a test inversion using the *Salvus* software package (release 0.11.44, Afanasiev et al., 2019) for six events of size $5.2 \leq M_W \leq 6.0$ that are available in the *GCMT* catalogue (Ekström et al., 2012), which occurred between September 1996 and June 2020 and are spatially as evenly distributed as possible (see Figure 7.3). The station distribution is presented in Figure 7.3. I use the anisotropic *PREM* (Dziewoński and Anderson, 1981) as a starting model for reasons discussed in Section 4.1.1. However, I intend to implement a crustal model for the actual inversion (e.g. Meier et al., 2007; Jenkins et al., 2018) since we have better constraints of Iceland's crustal structure compared to Southeast Asia (see Section 5.4.3). Topography and bathymetry are implemented using *Earth2014* (resolution: ∼1.8 km, Hirt and Rexer, 2015) for the test inversion, but a higher resolution topography model may be desirable. The fluid ocean is approximated by the weight of its water column (see Section 4.4.2), which will likely be a valid approximation since all events and receivers are located on the Icelandic continental shelf. However, I will further investigate this before running the actual inversion. Measurement windows are selected for three-component waveforms as described in Section 4.2.5, with a particular focus on accounting for body wave arrivals in order to enhance depth sensitivity.

Figure 7.2: *From left to right:* Conor Andrew Bacon, Thomas Edmund Barnaby Winder, Joseph William Fone and Timothy Stephen Greenfield at the 2021 Fagradalsfjall eruption on the Reykjanes Peninsula in the southwest of Iceland (see Figure 7.1). Photo taken by author in August 2021.

A total of 32 iterations across four period bands (40 – 100 s, 30 – 100 s, 20 – 100 s and 10 – 100 s) are run and the inversion parameters are restricted to radially anisotropic shear-wave velocity, isotropic P-wave velocity and density (see Section 4.6). The waveform misfit is successively decreased, which is reinforced by the fact that the number of unique source-receiver pairs increases from 136 to 216, and the total analysed window length increases from ∼6 – 10 h from the first to the last period band. I have run two additional iterations at 5 s but suspicious smaller-scale features start to appear and it is suspected that this is caused by the lack of data coverage from such a small dataset.

Figure 7.3 shows two shear-wave depth slices through the final model at 10 s minimum period, which reveals a low-velocity zone in the upper mantle that can likely be associated with the hot mantle plume beneath Iceland. However, its extension outside of Iceland cannot be imaged as expected from the given data coverage (see Figure 7.1). Furthermore, the preliminary model shows faster crustal velocities in eastern Iceland at 10 km depth. Interestingly, a thicker crust in this area has previously been suggested by Bjarnason and Schmeling (2009) and Volk et al. (2021).

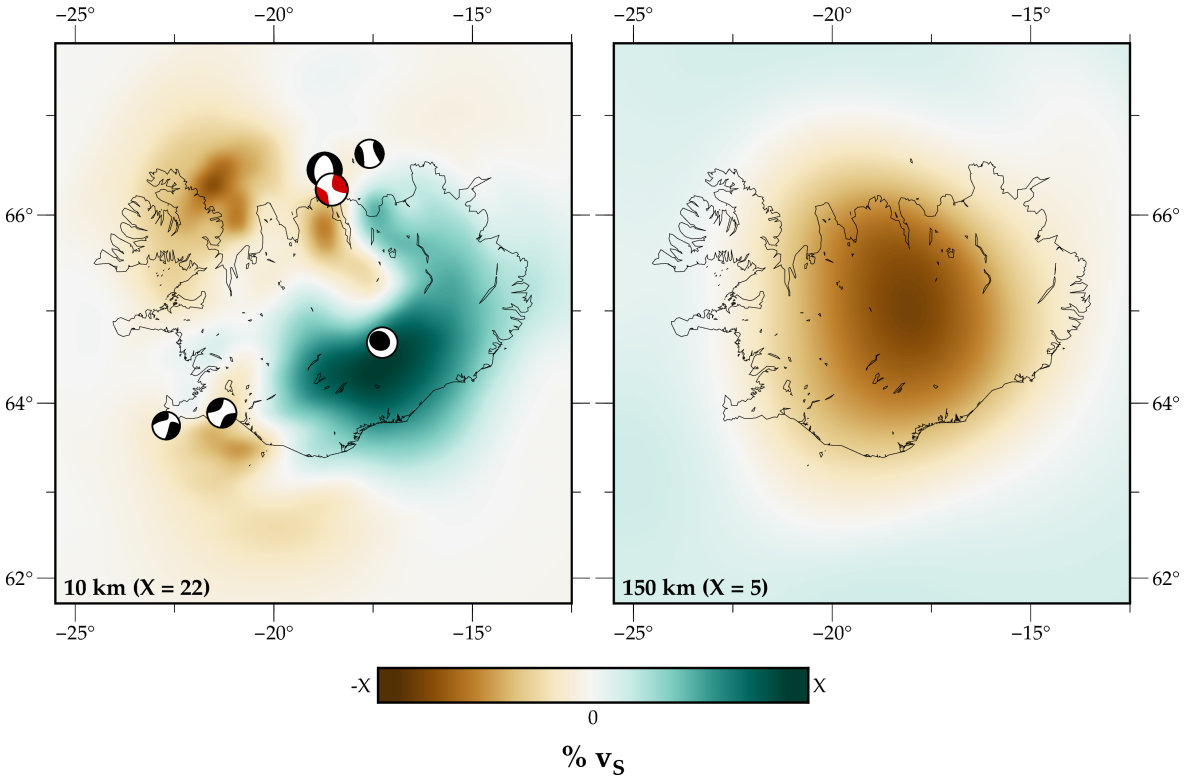


Figure 7.3: Shear-wave depth slices at (*left*) 10 km and (*right*) 150 km through the final model produced by the test inversion at 10 s minimum period. Perturbations are in % relative to the depth-average. The locations of the six events used for the test inversion are plotted in the left panel. Event information is taken from the *GCMT* catalogue (Ekström et al., 2012), which sets the hypocentral depths to 12 – 15 km for these earthquakes. The red beachball corresponds to the moment tensor inversion presented in Figure 7.4.

7.3 Event and data availability

Most earthquakes occur along the MOR, which is significantly offset to the east across most of Iceland, following a similar pattern of volcano locations (see Figure 7.1). Many studies have targeted local areas of interests, such as volcanoes, which explains the distribution of seismic networks. Furthermore, many short-period instruments were deployed, which are less suitable for a regional-scale study due to the relatively long periods currently considered in FWI (see Section 3.5).

7.3.1 Seismic stations

For this study, two time frames are of particular importance:

1. 1996 – 1998, when the *XD* (Nolet, 1996) broadband network was deployed across Iceland.
2. From 2013 onwards, when data from several temporary networks are available (mainly from the *Cambridge Volcano Seismology* research group), which are largely located along the MOR.

Publicly available data including instrument responses are downloaded automatically using *obspyDMT* (Hosseini and Sigloch, 2017, see Section 4.2.1). Restricted data were already available or requested from the *Icelandic Meteorological Office*[2] *(IMO* – pending). Observed waveforms are processed as described in Section 4.2.2 and a geographical station weighting following Ruan et al. (2019) is employed in order to balance the effect of dense seismic networks (see Section 4.3).

7.3.2 Event catalogue

The event selection takes place using a process of elimination, similar to Section 4.2.4 and is summarised below:

1. Seismic record sections show that events with $M \geq 4.0$ generate sufficient energy across the domain at the periods considered ($\sim$5 – 10 s).
2. All events between 1996 – 1998 and 2013 – 2020 are selected to account for the deployment of several temporary networks as described in Section 7.3.1.
3. To avoid interference, events are eliminated if another event occurred a few minutes before or after origin time. This is done using the *Icelandic Meteorological Office's (IMO)* earthquake catalogue (Rögnvaldsson and Slunga, 1993), which contains events with magnitudes of $-1.2 \leq M_{Lw} \leq 5.8$. This eliminates a lot of events as a result of the many earthquake swarms that occur in Iceland (e.g. Greenfield et al., 2020).
4. Events in unique locations are added[3], e.g. a 2007 event in western Iceland (see Figure 7.1).

[2]Previously known as the *Southern Iceland Lowlands (SIL).*

[3]In particular, I aim to include off-shore and on-shore earthquakes outside of the volcanic rift zones. However, the vast majority of these events are excluded in step 5 as they exhibit an insufficient signal-to-noise ratio.

5. Observed waveforms are investigated to evaluate the signal-to-noise ratio. I find permanent stations (here: *II.BORG*) to be a good indicator of the usefulness of an event, as previously suggested in Section 4.2.4.
6. Events are selected in order to achieve as even a spatial distribution as possible.

A total of 44 events with magnitudes[4] $4.1 \leq M_w \leq 5.8$ are selected. Event information (location and focal time) is extracted from Jónasson et al. (2021). For the 44 events used here, all values in Jónasson et al. (2021) are taken from the *IMO* catalogue (Rögnvaldsson and Slunga, 1993). However, several local studies have relocated individual events (e.g. Ágústsdóttir et al., 2019; Konstantinou et al., 2020; Southern et al., 2022) and thus, I adopt the values from Konstantinou et al. (2020) for four events since their epicentral locations vary by $\sim$5 – 20 km compared to the regional catalogue. For other local studies, epicentral distances vary by 1 – 2 km and origin times by $< 1/10$ s, which I argue is negligible at this scale. For comparison, *GCMT* (Ekström et al., 2012) epicentres differ by up to 22 km and 11 km on average compared to the *IMO* catalogue. Note that shallow events are usually fixed at 12 km depth in the *GCMT* catalogue, which would not be a feasible input for Iceland since most earthquakes occur at even shallower depths. An overview of the event information is provided in Section A1 of the Appendix.

7.3.3 Seismic moment tensor inversion

The seismic moment tensor inversion described in this section was performed by Félix Rodrigo Rodríguez Cardozo, Postdoctoral Research Scholar at the University of South Florida.

At the short periods considered, the source parameters become increasingly important. Ideally, we would like to invert for source and seismic structure simultaneously but this is the subject of ongoing development as further elaborated on in Section 5.4.1. As a result, most studies do not invert for source parameters or they are updated in consecutive iterations (rather than jointly) with the seismic structure. Here, we follow the latter approach and invert for seismic moment tensor (MT) solutions before running an inversion for seismic structure.

We use *TDMT_INV* (Dreger, 2003) with a period band of 10 – 50 s for the moment tensor inversion. *TDMT_INV* requires the observed waveforms, the event epicentre, receiver coordinates, a background velocity model and a set of precalculated Green's functions (GF, Morse and Feshbach, 1954) as input. GFs are the response in displacement to an impulse perturbation and thus, their convolution with a seismic source yields synthetic waveforms, which can be used for a source inversion. Here, the GFs are obtained using the frequency-wavenumber integration method proposed by Saikia (1994) with respect to a 1-D velocity model. The velocity model is based on a section from a refraction profile from the southwest of Iceland (Bjarnason et al., 1993) and is used for routine locations of seismicity in Iceland by *IMO*. It should be noted that Hejrani et al. (2017) observe a strong trade-off between centroid location and time with MT components when comparing MT solutions obtained

[4]Moment magnitudes are extracted from Jónasson et al. (2021) who take them from *GCMT* (Ekström et al., 2012) or when not available, compute proxy M_w values using χ^2 regression, normally on the surface-wave magnitude but exceptionally on the body-wave magnitude.

with respect to a 3-D vs a 1-D background model for a continental-scale study. They also observe a sizeable reduction in the MT's non-double component when using a 3-D velocity model for calculating the GFs. This is particularly intriguing for an application in Iceland since there is evidence that a double-couple fault model does not fit some of the observed events (e.g. Nettles and Ekström, 1998). However, Harris et al. (2022) compare MT solutions for a local study and observe a remarkable degree of consistency in terms of source geometry, kinematics, and magnitude, but notice a difference in hypocentral depths. One aim of this project is to recalculate MT solutions with the final 3-D model obtained in this study.

During the seismic moment tensor inversion, the source depth is varied from 1 to 20 km and we select the depth solution that results in the highest variance reduction. The epicentre is kept fixed because the *IMO* catalogue is obtained using a larger dataset, i.e. shorter-period data and waveforms from additional stations. Thus, the *IMO* uncertainties are smaller than the ones we could obtain running inversions at different epicentres with long-period data. Furthermore, off-shore events (see Figure 7.1) likely have a poor azimuthal data coverage.

Figure 7.4 presents the MT inversion result for one of the events used in the test inversion. In order to validate the moment tensor inversion result, we also present a comparison of the *TDMT_INV* output to *GCMT* (Ekström et al., 2012) and *GEOFON* (Quinteros et al., 2021) MT solutions, which shows that they are in very good agreement. The full list of the moment tensor inversion results will be made available upon publication.

7.4 Outlook

Initial tests with a subset of six earthquakes using *GCMT* moment tensors and a regional dataset have shown great promise in jointly imaging the crust and upper mantle beneath Iceland down to $\sim$150 km depth using a minimum period of $\sim$10 s. Furthermore, seismic MT solutions obtained with *TDMT_INV* using the given dataset are in good agreement with those from the *GCMT* and *GEOFON* catalogues. Before running the actual inversion using the inverted MT solutions, I will explore whether the topography and bathymetry model *EARTH2014* has sufficient resolution at the periods considered, and assess different starting models. Furthermore, I will investigate whether ocean loading (see Section 4.4.2) is a valid approximation for this simulation setup. This should allow me to better recover the strong velocity contrasts, which exist in Iceland, and provide an improved reference model for microseismic analysis and future tomographic studies.

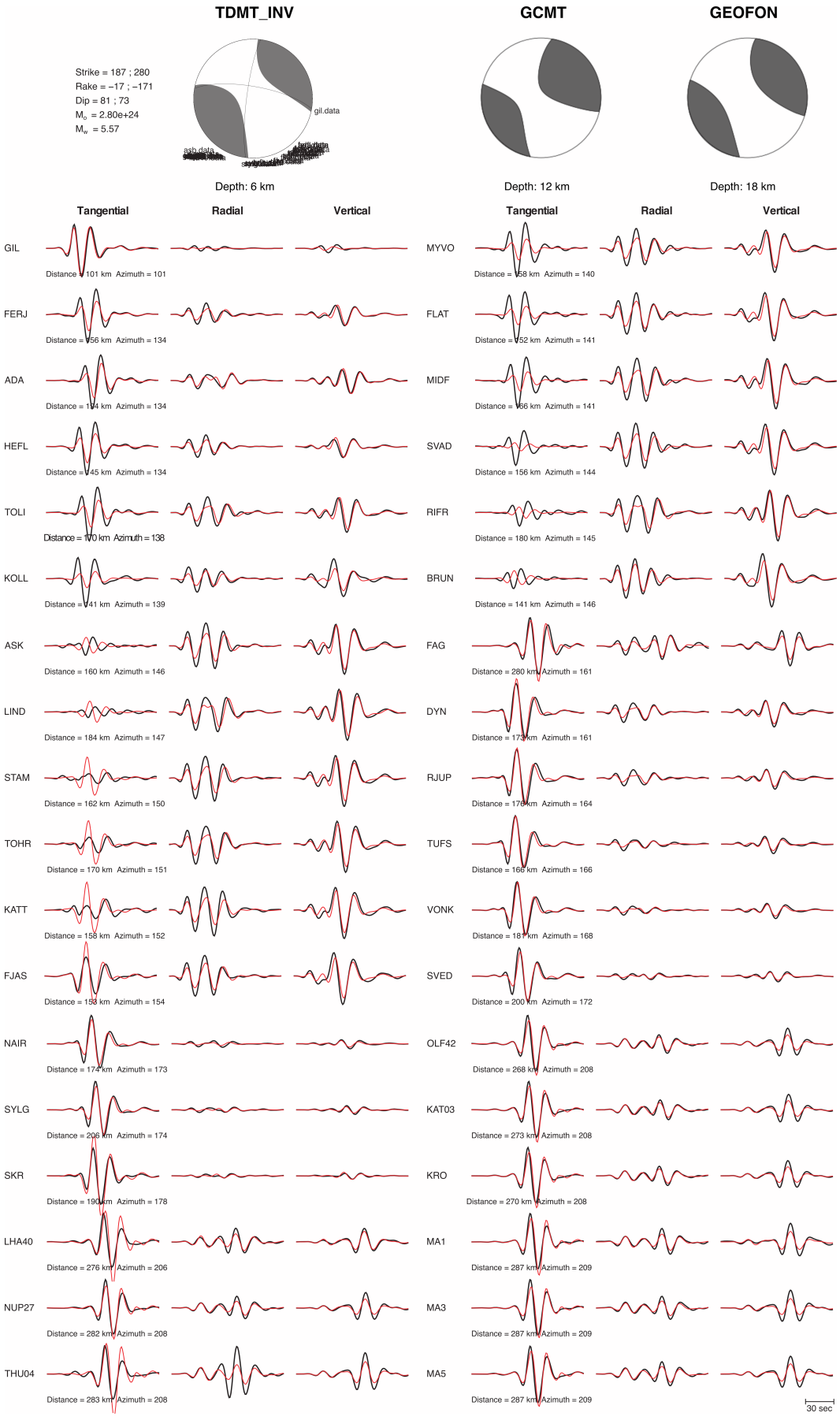


Figure 7.4: Seismic moment tensor inversion result for event #41 in Section A1 of the Appendix and the waveform match (red: synthetics, black: observed) for a subset of stations. The event's location corresponds to the red beachball in Figure 7.3. For comparison, the *GCMT* (Ekström et al., 2012) and *GEOFON* (Quinteros et al., 2021) moment tensor solutions are shown in the top right corner.

8 Conclusions and future directions

8.1 Conclusions

In the first part of this thesis, I imaged the lithosphere and underlying mantle beneath Southeast Asia at periods between 20 – 150 s using an unprecedented dataset that allowed for the application of adjoint waveform tomography to this region for the first time. 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. The final model, *SASSY21*, was reached after 87 L-BFGS iterations and is most reliable for shear-wave velocity due to the natural dominance of surface wave signals in adjoint waveform tomography. The trade-off between inversion parameters was estimated through an analysis of the Hessian-vector product, which showed that the model is most sensitive to changes beneath the Sundaland block, as expected from the data coverage. Furthermore, *SASSY21* is able to explain true-amplitude data from events and receivers not included in the inversion, which also testifies to its robustness.

The final model is able to resolve detailed anomalies down to the transition zone, including multiple subduction zones. The most prominent feature is the (Indo-)Australian plate that descends beneath Indonesia, which can be seen in the form of a steeply dipping Sunda slab in the west. Further east, the Southeast Asia-Australia collision zone is imaged, indicated by high velocities that reflect the presence of the northward moving Australian continental lithosphere. The 180° curve of the Banda Arc is imaged as one continuous slab and *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. Furthermore, a high-velocity zone around northern Borneo is imaged and a previously undiscovered high-velocity feature beneath the east coast of Borneo is revealed. While two subduction systems terminated in the Neogene ($\sim$23 – 2 Myr) 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.

In the second part of this thesis, the minimum period is decreased to 15 s (compared to 20 s for *SASSY21*) and I focus on the eastern Indonesian region, which is characterised by particularly good data coverage that facilitates a more refined image. In this study, *SASSY21* was used as a starting model and the fluid ocean was accounted for explicitly by solving a coupled system of the acoustic and elastic wave equation. This is computationally more expensive but allows seismic waves within the water layer to be simulated, which becomes

important at shorter periods. The final model, *SASSIER22*, after 34 L-BFGS iterations reveals a convergent double subduction zone along the southern segment of the Philippine Trench, which was not evident in the starting model and transitions to a divergent system in the Molucca Sea further south. A more detailed illumination of the slab beneath the North Sulawesi subduction zone reveals a pronounced positive wavespeed anomaly down to 200 km depth, consistent with the maximum depth of seismicity, and a more diffuse but aseismic positive wavespeed anomaly that continues to the 410 km discontinuity.

While the increase in data availability within the region contributes to a better image, a careful inversion setup is also of crucial importance. For example, I have illustrated the importance of data selection and station weighting throughout the inversion, showing that such steps help to recover more refined slabs and remove an imprint of the data coverage from the gradient used for the model update, even at greater depths. Furthermore, I explored the influence of surface elevation (topography and bathymetry) and the fluid ocean, which is usually ignored in other studies. I conclude that their effects on synthetic waveforms become important at periods $\leq$ 20 s. In particular, surface elevation can result in a considerable phase advance and change in amplitude of the surface wave train, and has an effect on both horizontal and the vertical seismogram components in my simulation setup. The fluid ocean results in a phase delay as well as a change in amplitudes and duration of the surface wave train, and affects the radial and vertical components. At periods $\leq$ 20 s, accounting for the fluid ocean explicitly can lead to more realistic lithospheric velocities and a more refined image compared to the frequently used ocean load approximation, even at greater depths. Furthermore, it allows for an improved waveform match for source-receiver paths passing partially or entirely through oceanic regions.

8.2 Future directions

Advances in FWI As computational advances progress and sophisticated, automated workflows continue to be developed, the application of adjoint waveform tomography becomes more established across a wide range of applications. While the inversion setup and performance are still time- and computing-intensive, an increasing number of studies attempt to overcome the current challenges, as described in Section 3.5. In particular, I did not invert for seismic sources during the Southeast Asian tomography studies for reasons discussed in Section 5.4.1. However, we included a seismic moment tensor inversion as a precursor to the Icelandic waveform tomography for which shorter-period data is considered (see Section 7.3.3), which will allow us to investigate how necessary this is in more detail.

Most commonly, adjoint waveform tomography is implemented using a deterministic approach due to the high computational cost of the forward problem, which makes it difficult to investigate what the effect the choice of initial parameters has on the final model, such as prior information about crustal structure (see Section 5.4.3). However, there are ongoing developments towards this; for example, Doody et al. (2022) currently explore tomographic results obtained from different starting models. I find that the

inversion setup described in Chapter 4 translated well across scales and initial tests with a small number of events have shown great promise to jointly image the crust and upper mantle of a contrasting and significantly smaller study region (see Chapter 7). However, this is an ongoing project and more work needs to be done to better recover the strong velocity contrasts beneath Iceland.

Effects arising from surface elevation and the fluid ocean In Chapter 6, I only investigate the effects of surface elevation and the fluid ocean for an eastern Indonesian simulation setup at a minimum period of 15 s. Thus, the applicability of my results for other regions and different scales remains to be studied. Furthermore, more detailed tests would be required to evaluate the effect of focal mechanisms and source locations (e.g. near trenches). However, my results are in good agreement with Fernando et al. (2020), who consider epicentral distances of up to 90° across the Pacific Ocean.

I expect the effects of surface elevation and the fluid ocean to become even more pronounced at shorter periods and when running additional iterations. This would be an avenue to pursue – ideally with an improved data coverage around the Philippines – with the ultimate aim of bridging applications from a continental to an exploration scale.

Anomaly interpretation Adjoint waveform tomography tends to reveal stronger perturbations compared to ray-based tomography studies (Tromp, 2020), indicating that the Earth may be more heterogeneous than we thought. This means some simplified approaches towards interpreting seismological models break down, as encountered during the analysis of the lithosphere–asthenosphere boundary structure in Section 5.3.1. Thus, more advanced theory is needed to explain our observations (e.g. Jackson and Faul, 2010; Wei et al., 2015); for example, temperature variations are not sufficient to explain the strong perturbations observed in shear-wave structure (e.g. Dalton et al., 2014).

Acknowledgements

This research was funded by the Engineering and Physical Sciences Research Council (EPSRC) project reference 2073302.

The *Salvus* software package (release 0.11.23 – 0.11.44, `www.mondaic.com`) was used for the mesh generation, forward and adjoint simulations and non-linear optimisation, 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* (Harris et al., 2019) and *ObsPy* (Beyreuther et al., 2010). Waveform data are handled in the *Adaptable Seismic Data Format* (*ASDF*, Krischer et al., 2016). Visualisations were created using *Inkscape* (Inkscape Project, 2020), *Matplotlib* (Hunter, 2007) and *PyGMT* (Uieda et al., 2021).

Nick I am beyond grateful for the opportunity to pursue a PhD at such an inspiring university. Thank you for your continuous support and excellent feedback. I also can't thank you enough for the incredible fieldwork opportunities during my time here. Hiking through the Bornean jungle and witnessing the 2021 Fagradalsfjall volcanic eruption in Iceland were highlights of my PhD (and they were always accompanied by excellent food).

Nienke Thank you for always having an open ear to all my questions, your cheerful words of encouragement and your reassurance that I am not an impostor. I couldn't have done this without you!

St John's College, Cambridge My beloved home away from home. I never tired of walking through the stunning grounds and ancient buildings of this beautiful place. I am deeply thankful for the wonderful memories with people from all over the world, who are now dear friends. *Souvent me souvient.*

Appendices

A1: Event overview

List of events used to obtain *SASSY21*. Event locations and moment tensors are retrieved from the *GCMT* catalogue and remain constant throughout the inversion. The period bands the events were used in are indicated by roman numerals in the last column, following the notation from Table 4.1. The first 50 events are used across all period bands.

#	origin time	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	13.0	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
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

continued on next page

continued from previous page

#	origin time	M_w	longitude	latitude	depth [km]	period bands
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

continued from previous page

#	origin time	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	15.2	I - V

continued on next page

continued from previous page

#	origin time	M_w	longitude	latitude	depth [km]	period bands
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
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 - III
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	5.23	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

List of events used for the validation dataset described in Section 5.2. Event locations and moment tensors are retrieved from the *GCMT* catalogue.

#	origin time	M_w	longitude	latitude	depth [km]
1	2014-07-14T08:00:03.5	6.25	126.55	5.64	28.6
2	2015-09-27T02:56:15.2	5.56	129.64	-7.29	125.3
3	2016-05-02T04:21:25.1	5.71	104.49	-5.26	125.4
4	2017-01-13T16:39:28.2	5.66	125.04	-0.20	36.2
5	2017-03-31T11:21:01.3	5.52	120.59	0.37	102.4
6	2017-08-31T17:07:01.0	6.34	99.59	-1.23	49.3
7	2019-02-18T19:30:28.7	5.67	112.84	-9.74	35.2
8	2019-03-15T23:07:57.8	5.60	125.37	6.14	82.2
9	2019-04-23T05:37:55.6	6.45	125.16	11.88	53.5
10	2019-06-24T01:05:31.8	6.11	138.58	-2.66	19.8

List of events used to obtain *SASSIER22* (see Chapter 6).

#	origin time	M_w	longitude	latitude	depth [km]	period bands
1	2014-04-17T04:38:20.0	5.76	122.82	4.55	575.0	20 – 150 s, 15 – 150 s
2	2014-10-30T12:11:36.8	5.76	117.48	-6.94	547.4	20 – 150 s, 15 – 150 s
3	2015-03-28T22:28:52.4	5.92	122.00	0.43	130.6	20 – 150 s, 15 – 150 s
4	2015-11-11T23:36:22.0	5.84	128.93	-7.41	137.0	20 – 150 s, 15 – 150 s
5	2015-12-24T23:10:59.7	5.81	129.11	-7.34	132.1	20 – 150 s, 15 – 150 s
6	2016-01-14T13:14:51.3	5.54	126.54	9.62	32.2	20 – 150 s, 15 – 150 s
7	2016-02-28T09:06:25.9	5.55	125.32	5.27	180.6	20 – 150 s, 15 – 150 s
8	2016-03-19T08:51:26.5	5.70	129.43	-5.56	282.0	20 – 150 s, 15 – 150 s
9	2016-03-29T09:18:12.0	5.42	126.21	2.41	61.8	20 – 150 s
10	2016-04-05T08:29:39.2	5.92	126.63	4.21	29.7	20 – 150 s, 15 – 150 s
11	2016-04-06T20:35:15.9	5.49	130.36	-6.10	130.6	20 – 150 s
12	2016-04-13T18:21:55.9	5.97	121.94	7.84	24.2	20 – 150 s
13	2016-04-15T04:50:12.9	5.59	126.98	2.06	108.7	20 – 150 s, 15 – 150 s
14	2016-09-04T02:38:13.9	5.77	125.85	8.38	19.0	20 – 150 s, 15 – 150 s
15	2016-09-16T07:50:48.4	5.49	125.66	5.71	200.1	20 – 150 s, 15 – 150 s
16	2016-10-09T14:46:28.1	5.82	127.48	1.82	141.1	20 – 150 s, 15 – 150 s
17	2016-10-27T08:17:52.2	5.78	125.88	1.40	67.8	20 – 150 s
18	2016-11-16T15:10:13.1	5.71	113.18	-9.14	105.7	20 – 150 s, 15 – 150 s
19	2016-12-19T20:44:43.2	5.42	125.89	10.17	71.7	20 – 150 s, 15 – 150 s
20	2017-01-31T15:55:58.8	5.41	125.31	-7.05	510.7	20 – 150 s, 15 – 150 s
21	2017-03-14T05:55:21.9	5.42	121.14	0.70	87.6	20 – 150 s, 15 – 150 s
22	2017-03-21T23:10:28.1	5.69	115.27	-8.75	130.2	20 – 150 s, 15 – 150 s
23	2017-04-13T16:05:54.6	5.58	127.49	5.53	28.5	20 – 150 s

Continued on next page

continued from previous page

#	origin time	M_w	longitude	latitude	depth [km]	period bands
24	2017-05-20T01:06:16.4	5.98	124.02	9.33	544.6	20 – 150 s, 15 – 150 s
25	2017-07-10T01:41:52.6	5.80	124.76	11.08	13.6	20 – 150 s, 15 – 150 s
26	2017-07-15T12:12:22.5	5.94	121.95	0.44	125.8	20 – 150 s, 15 – 150 s
27	2017-07-23T07:55:55.4	5.40	120.20	0.29	52.6	20 – 150 s
28	2017-09-30T23:15:33.8	5.60	128.49	4.11	18.8	20 – 150 s
29	2017-12-28T17:20:23.4	5.75	126.83	4.10	32.5	20 – 150 s
30	2018-02-26T13:34:58.8	6.00	126.82	-2.65	12.9	20 – 150 s
31	2018-05-10T11:57:00.7	5.68	125.52	10.15	30.7	20 – 150 s, 15 – 150 s
32	2018-05-10T18:02:29.8	5.88	123.70	6.95	543.1	20 – 150 s 15 – 150 s
33	2018-09-07T20:44:11.6	5.42	125.00	2.80	203.1	20 – 150 s, 15 – 150 s
34	2018-10-20T09:47:54.2	5.61	128.10	-6.37	358.5	20 – 150 s, 15 – 150 s
35	2018-11-04T07:55:29.9	5.99	123.75	7.82	599.3	20 – 150 s, 15 – 150 s
36	2018-12-03T14:00:09.3	5.54	128.72	-7.52	142.9	20 – 150 s, 15 – 150 s
37	2019-01-04T06:22:47.0	5.61	130.45	-6.55	137.2	20 – 150 s, 15 – 150 s
38	2019-01-07T03:11:56.4	5.57	121.94	7.93	16.5	20 – 150 s
39	2019-01-15T20:03:23.4	5.76	126.83	5.91	61.9	20 – 150 s, 15 – 150 s
40	2019-01-23T11:39:10.5	5.76	119.01	-10.38	27.3	20 – 150 s, 15 – 150 s
41	2019-02-07T04:15:33.3	5.72	126.39	1.53	40.5	20 – 150 s, 15 – 150 s
42	2019-02-08T11:55:12.6	5.90	126.41	9.85	20.8	20 – 150 s, 15 – 150 s
43	2019-02-18T19:30:28.7	5.67	112.84	-9.74	35.2	20 – 150 s, 15 – 150 s
44	2019-03-06T00:13:04.8	5.86	127.05	8.49	18.1	20 – 150 s, 15 – 150 s
45	2019-04-01T02:16:54.1	5.55	125.93	10.58	45.3	20 – 150 s, 15 – 150 s
46	2019-04-25T21:57:48.8	5.40	122.52	-1.81	17.9	20 – 150 s
47	2019-04-26T08:04:28.9	5.54	126.64	9.78	19.0	20 – 150 s, 15 – 150 s
48	2019-06-01T09:45:30.7	5.46	126.55	6.14	82.4	20 – 150 s, 15 – 150 s
49	2019-06-05T20:41:50.0	5.53	124.02	6.51	465.9	20 – 150 s, 15 – 150 s
50	2019-06-14T20:10:55.2	5.71	130.77	-5.80	129.2	20 – 150 s, 15 – 150 s
51	2019-06-17T05:43:32.7	5.48	123.04	-8.82	117.4	20 – 150 s, 15 – 150 s
52	2019-07-01T16:59:26.1	5.93	124.09	9.15	545.8	20 – 150 s, 15 – 150 s
53	2019-07-16T00:18:38.3	5.78	114.50	-9.01	102.7	20 – 150 s, 15 – 150 s
54	2019-09-10T05:39:25.3	5.74	126.34	6.60	57.2	20 – 150 s
55	2019-09-10T23:32:26.5	5.68	126.79	4.08	29.3	20 – 150 s, 15 – 150 s
56	2019-09-21T19:53:15.4	5.88	130.50	-6.46	87.8	20 – 150 s, 15 – 150 s
57	2020-02-20T20:24:30.1	5.52	129.67	-6.83	172.8	20 – 150 s, 15 – 150 s
58	2020-02-27T19:13:04.0	5.61	125.87	3.79	125.4	20 – 150 s, 15 – 150 s
59	2020-04-05T18:37:14.9	6.02	126.33	1.53	41.3	20 – 150 s, 15 – 150 s
60	2020-05-23T00:13:41.9	5.45	129.12	-7.22	144.5	20 – 150 s, 15 – 150 s
61	2020-05-30T13:06:27.4	5.63	126.66	4.05	41.0	20 – 150 s, 15 – 150 s
62	2020-07-27T17:32:47.3	5.82	126.49	8.90	51.9	20 – 150 s, 15 – 150 s

Continued on next page

continued from previous page

#	origin time	M_w	longitude	latitude	depth [km]	period bands
63	2020-07-31T06:06:44.6	5.74	126.81	8.44	33.5	20 – 150 s, 15 – 150 s
64	2020-09-17T10:32:29.5	5.65	129.44	-6.84	195.1	20 – 150 s, 15 – 150 s
65	2020-11-01T03:43:22.5	5.84	129.27	-7.16	192.5	20 – 150 s, 15 – 150 s
66	2020-11-15T22:37:46.3	5.99	126.56	8.59	54.5	20 – 150 s, 15 – 150 s
67	2020-12-29T16:21:50.1	5.57	125.36	5.24	17.0	20 – 150 s, 15 – 150 s

List of events used for the Icelandic adjoint waveform tomography study described in Chapter 7. Most event information is taken from the *Iceland Meteorological Office (IMO)* catalogue (Rögnvaldsson and Slunga, 1993).

#	origin time	M_w	longitude	latitude	depth [km]	catalogue
1	1996-09-29T10:48:16.17	5.60	-17.49720	64.67128	4.044	Konstantinou et al. (2020)
2	1996-10-01T14:47:47.74	5.45	-17.48367	64.66683	1.840	Konstantinou et al. (2020)
3	1996-10-02T13:17:19.74	4.54	-17.56892	64.52473	4.630	Konstantinou et al. (2020)
4	1996-10-03T13:08:20.33	4.63	-17.33033	64.62819	9.410	Konstantinou et al. (2020)
5	1997-01-03T00:51:18.914	4.50	-17.59151	64.49713	0.144	*IMO*
6	1997-04-12T23:04:44.258	4.23	-21.23817	64.07137	3.593	*IMO*
7	1997-04-19T16:08:18.139	4.74	-17.57657	64.49485	0.062	*IMO*
8	1997-07-22T16:21:40.176	5.12	-18.39388	66.28674	3.500	*IMO*
9	1998-05-31T10:32:35.598	4.27	-16.28294	65.18061	0.042	*IMO*
10	1998-06-04T19:04:44.913	4.64	-21.27744	64.08764	2.126	*IMO*
11	1998-06-04T21:36:53.615	5.43	-21.29378	64.03367	4.975	*IMO*
12	1998-06-04T22:59:57.137	4.75	-21.30797	63.99241	2.386	*IMO*
13	2007-11-26T15:31:29.462	4.81	-19.71449	64.93375	6.087	*IMO*
14	2013-04-02T08:55:54.626	4.95	-17.74955	66.60459	12.914	*IMO*
15	2014-08-24T05:33:41.596	4.91	-17.44975	64.61641	5.995	*IMO*
16	2014-08-27T02:50:36.573	5.32	-17.37260	64.65322	6.220	*IMO*
17	2014-08-29T12:21:46.700	5.27	-17.42961	64.68294	9.207	*IMO*
18	2014-08-30T07:03:02.171	5.36	-17.45727	64.61080	2.864	*IMO*
19	2014-09-03T03:09:54.372	5.35	-17.45095	64.67992	6.377	*IMO*
20	2014-09-06T05:40:50.559	5.06	-17.40545	64.67905	2.840	*IMO*
21	2014-09-11T00:07:38.303	5.22	-17.38777	64.61845	5.850	*IMO*
22	2014-09-16T21:34:12.689	4.99	-17.37817	64.69247	7.711	*IMO*
23	2014-09-21T10:51:45.642	5.38	-17.39161	64.61313	8.880	*IMO*
24	2014-09-23T04:33:55.100	5.11	-17.41786	64.68307	4.851	*IMO*
25	2014-09-25T05:00:03.334	5.12	-17.39366	64.61078	11.733	*IMO*
26	2014-09-26T16:49:28.819	4.99	-17.44700	64.67013	4.354	*IMO*
27	2014-10-21T08:36:35.412	5.36	-17.44208	64.69447	15.574	*IMO*
28	2014-10-31T21:32:22.774	5.02	-17.47631	64.67280	1.853	*IMO*

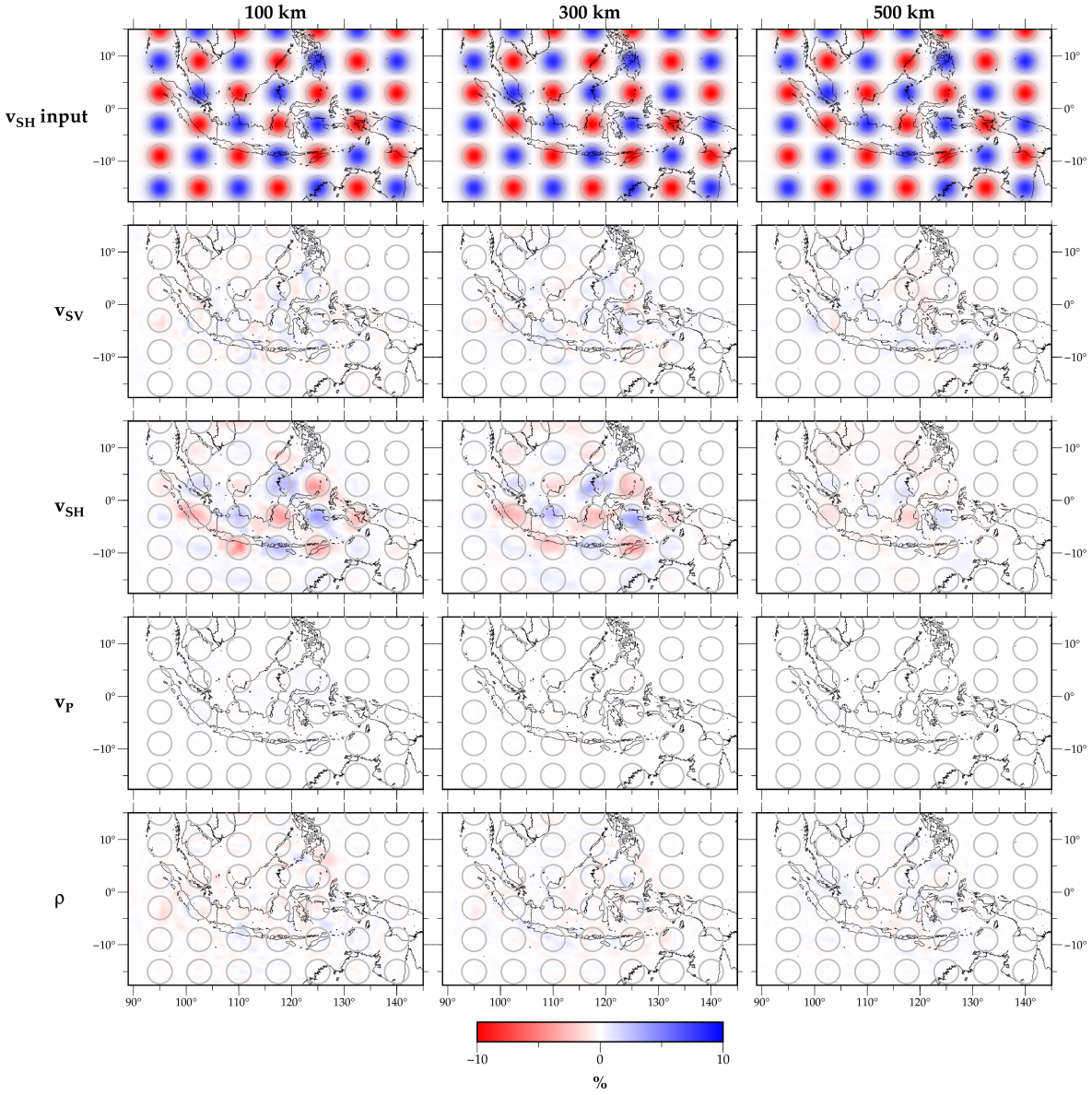
Continued on next page

continued from previous page

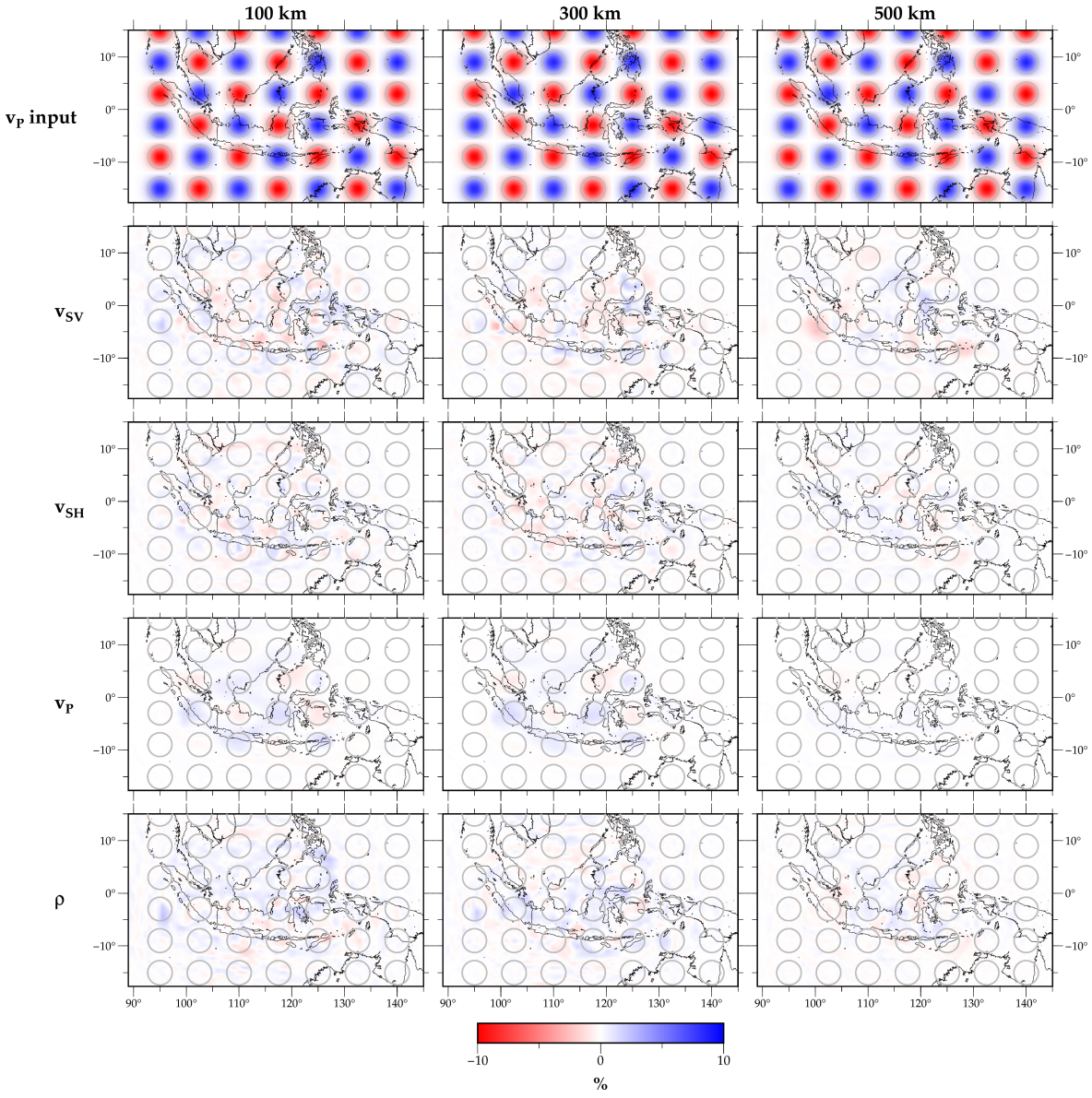
#	origin time	M_w	longitude	latitude	depth [km]	catalogue
29	2014-11-07T07:11:23.082	5.18	-17.51378	64.65782	5.603	*IMO*
30	2014-12-15T09:26:50.109	5.37	-17.40761	64.67323	8.211	*IMO*
31	2015-01-08T18:47:10.063	4.99	-17.49253	64.65169	2.995	*IMO*
32	2015-01-18T07:35:48.997	4.71	-17.47623	64.66704	4.890	*IMO*
33	2015-02-06T03:48:07.167	4.73	-17.36178	64.65706	5.841	*IMO*
34	2015-07-01T02:25:35.937	5.06	-23.74729	63.62218	3.542	*IMO*
35	2017-04-19T12:34:29.496	4.36	-23.34212	63.64071	13.199	*IMO*
36	2017-07-26T13:55:25.917	4.36	-22.22665	63.91145	5.741	*IMO*
37	2018-12-28T01:16:33.684	4.78	-17.45107	64.66445	0.135	*IMO*
38	2018-12-30T02:56:20.073	4.50	-21.33544	63.97004	5.938	*IMO*
39	2019-03-27T20:29:11.469	4.14	-16.52657	66.27478	6.478	*IMO*
40	2019-11-16T13:17:50.987	4.88	-23.59093	63.68489	2.234	*IMO*
41	2020-06-20T19:26:21.534	5.64	-18.55070	66.25351	10.005	*IMO*
42	2020-06-21T19:07:52.978	5.82	-18.72155	66.45129	10.005	*IMO*
43	2020-08-08T03:42:33.019	4.58	-18.42969	66.24963	12.679	*IMO*
44	2020-09-15T14:52:11.782	4.63	-17.72333	66.12726	11.287	*IMO*

A2: Synthetic recovery tests

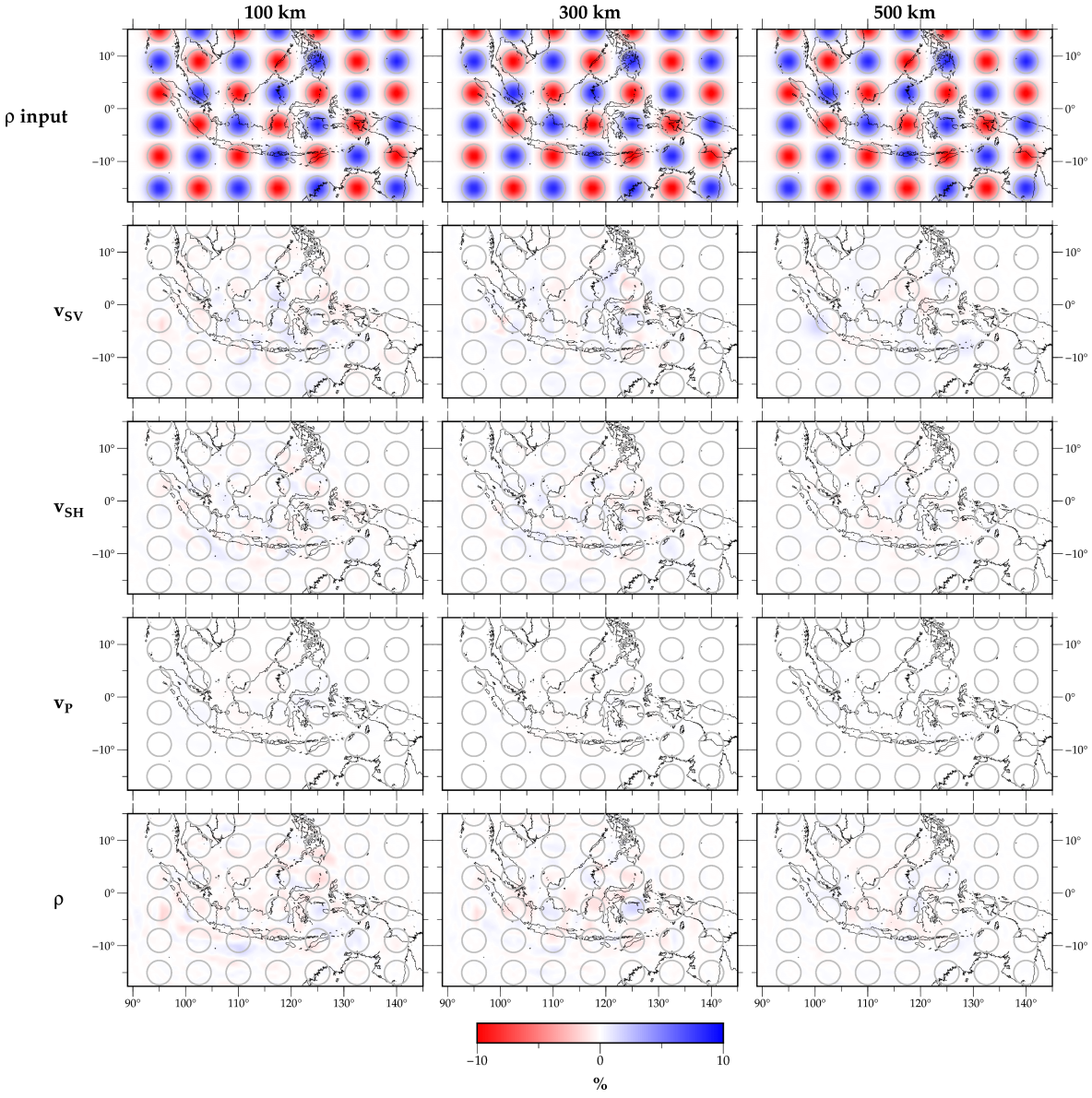
In the following, the results for the other inversion parameters of the synthetic recovery tests described in Section 4.7 are presented.



Depth slices for a checkerboard recovery test. *Top panel:* Input perturbation of ± 10 % in v_{SH}. *Panels below:* The final model after 20 iterations for all inversion parameters (v_{SV}, v_{SH}, v_P and density ρ). Perturbations are in % and relative to the 1-D starting model (*CSEM*). The columns indicate depths at (*left*) 100 km, (*middle*) 300 km and (*right*) 500 km.

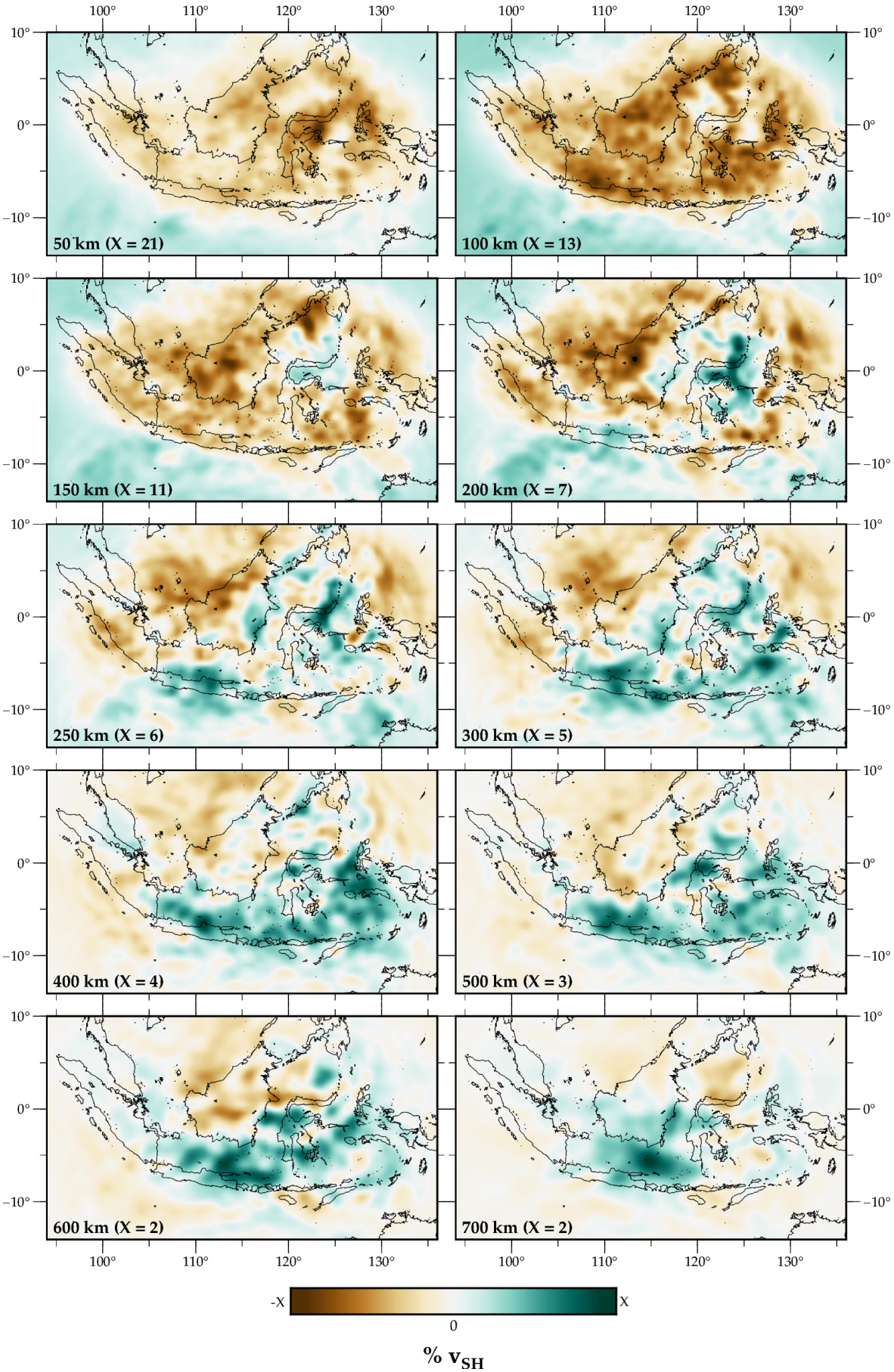


Depth slices for a checkerboard recovery test. *Top panel:* Input perturbation of ± 10 % in v_P. *Panels below:* The final model after 20 iterations for all inversion parameters (v_{SV}, v_{SH}, v_P and density ρ). Perturbations are in % and relative to the 1-D starting model (*CSEM*). The columns indicate depths at (*left*) 100 km, (*middle*) 300 km and (*right*) 500 km.

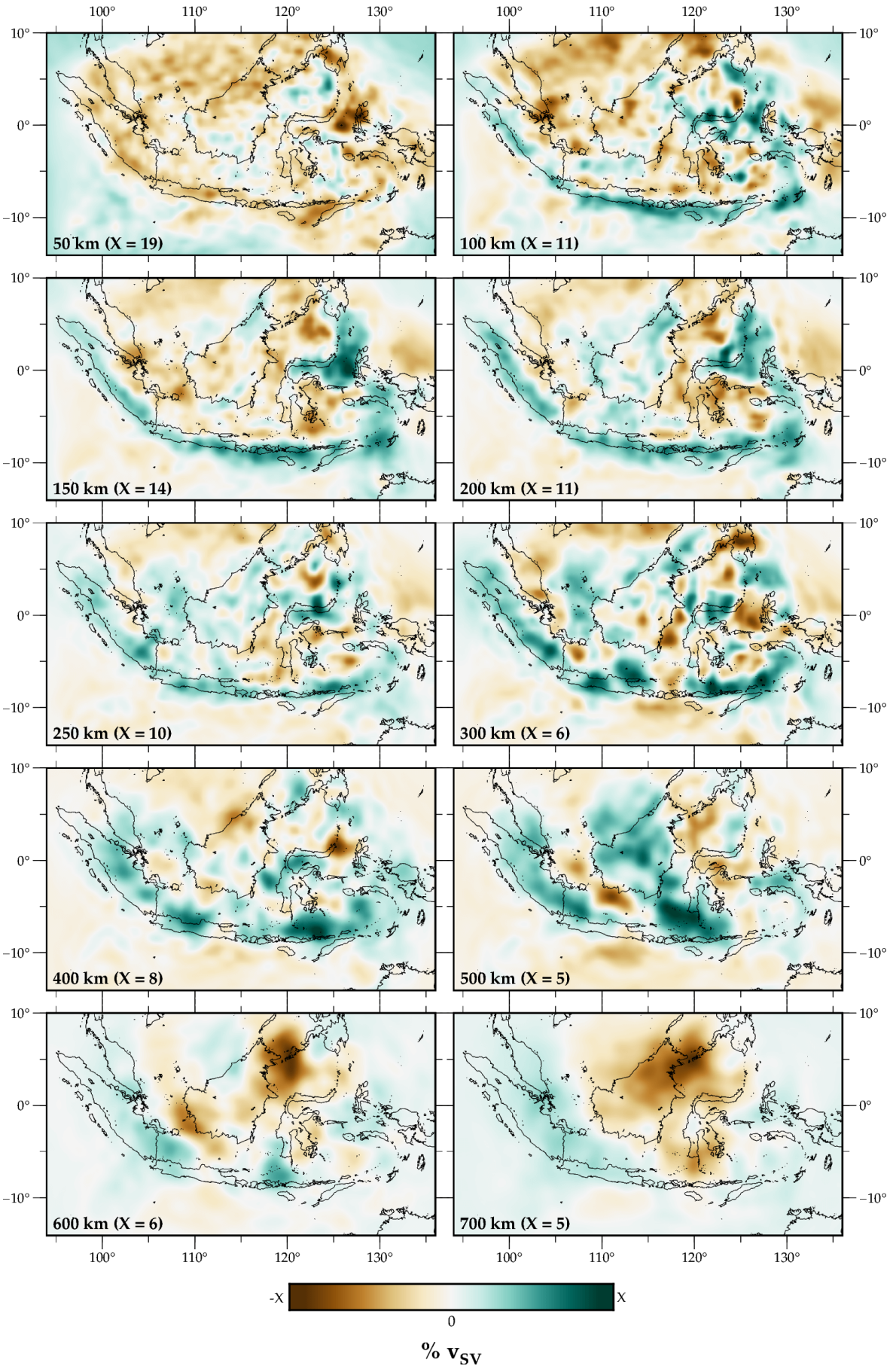


Depth slices for a checkerboard recovery test. *Top panel:* Input perturbation of ± 10 % in ρ. *Panels below:* The final model after 20 iterations for all inversion parameters (v_{SV}, v_{SH}, v_P and density ρ). Perturbations are in % and relative to the 1-D starting model (*CSEM*). The columns indicate depths at (*left*) 100 km, (*middle*) 300 km and (*right*) 500 km.

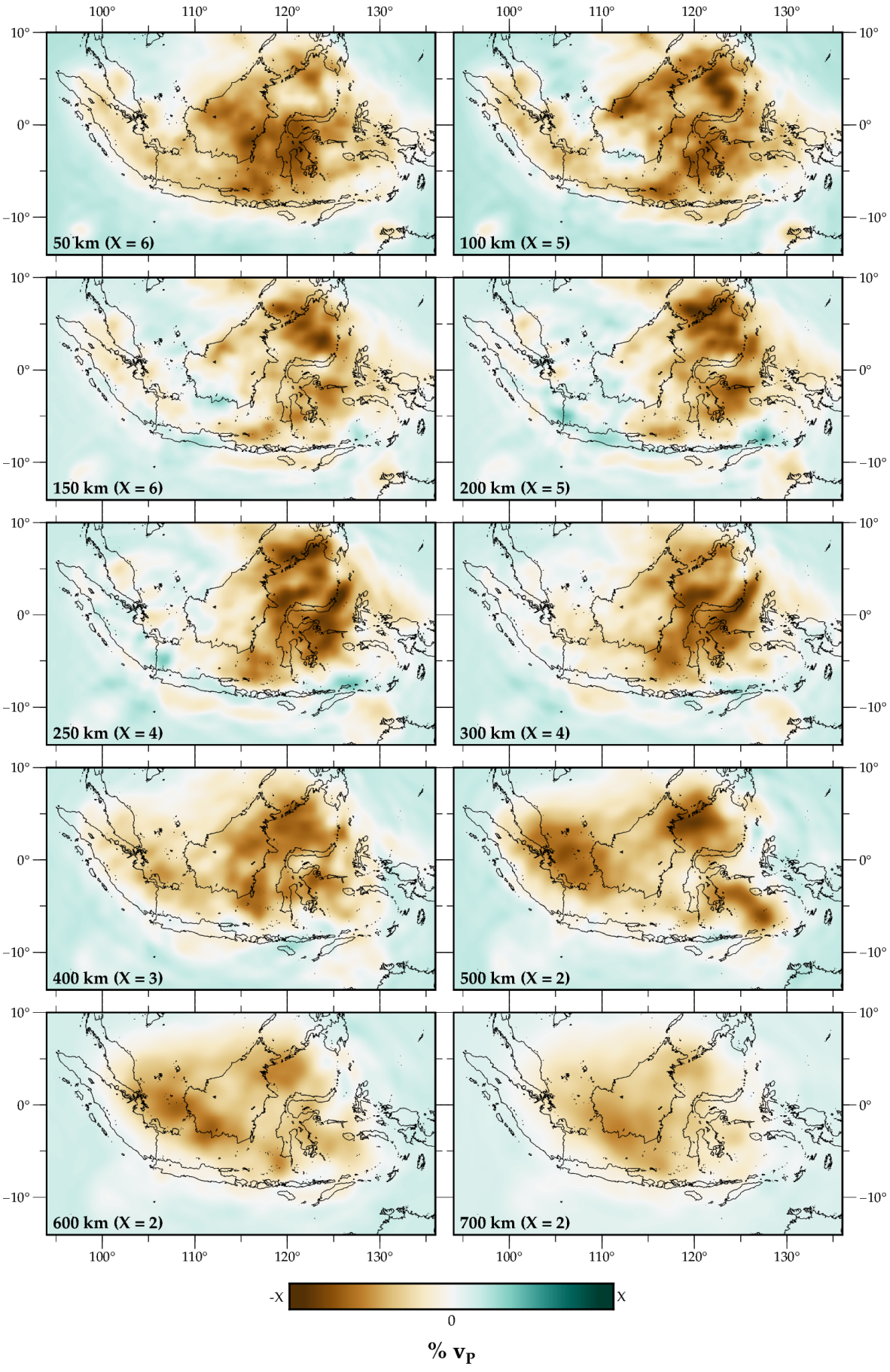
A3: *SASSY21* depth slices



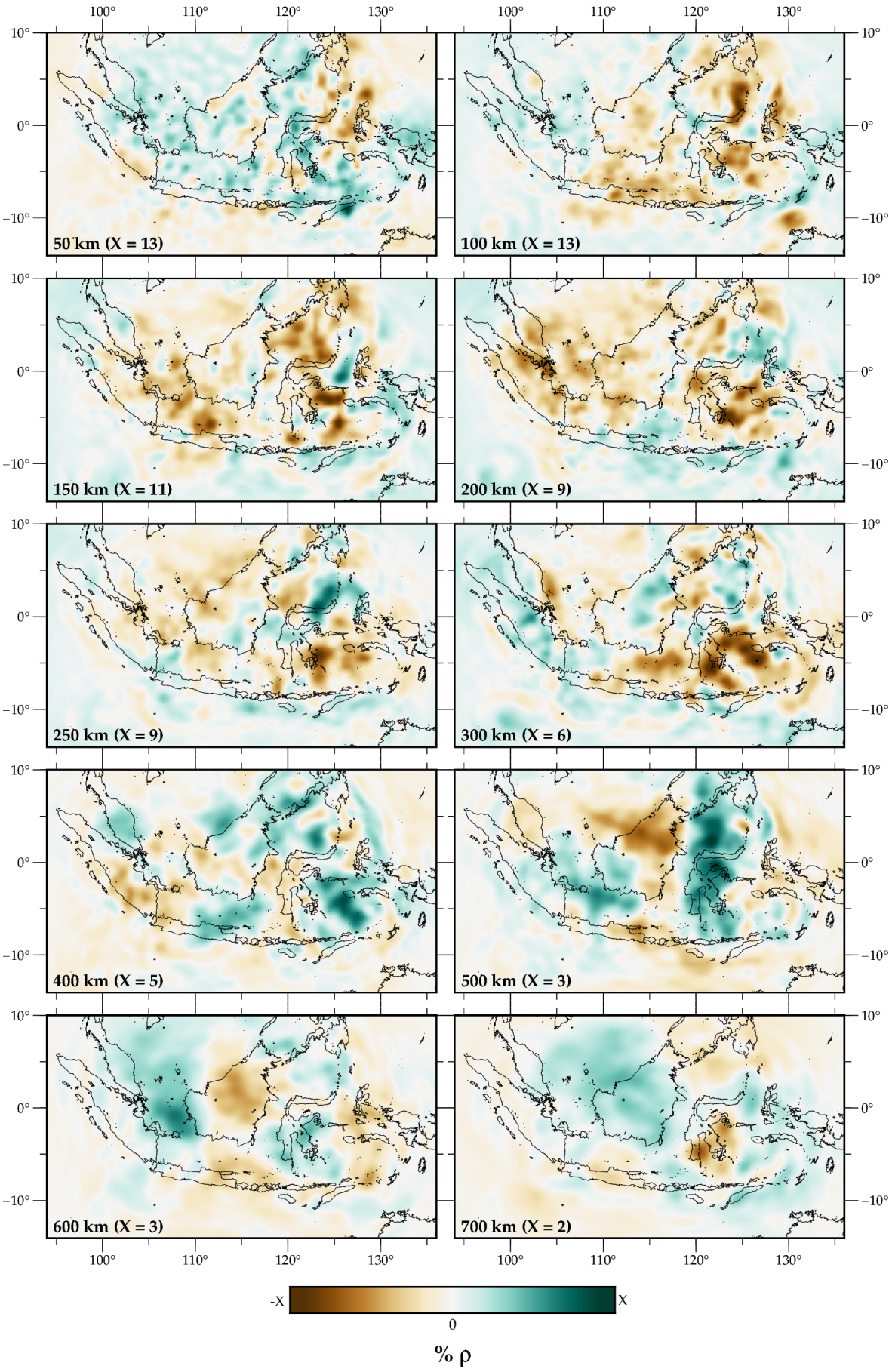
v_{SH} depth slices from 50 to 700 km. Perturbations are in % relative to the depth-average. The limits of the colourscale X are shown in the lower left corner of each plot.



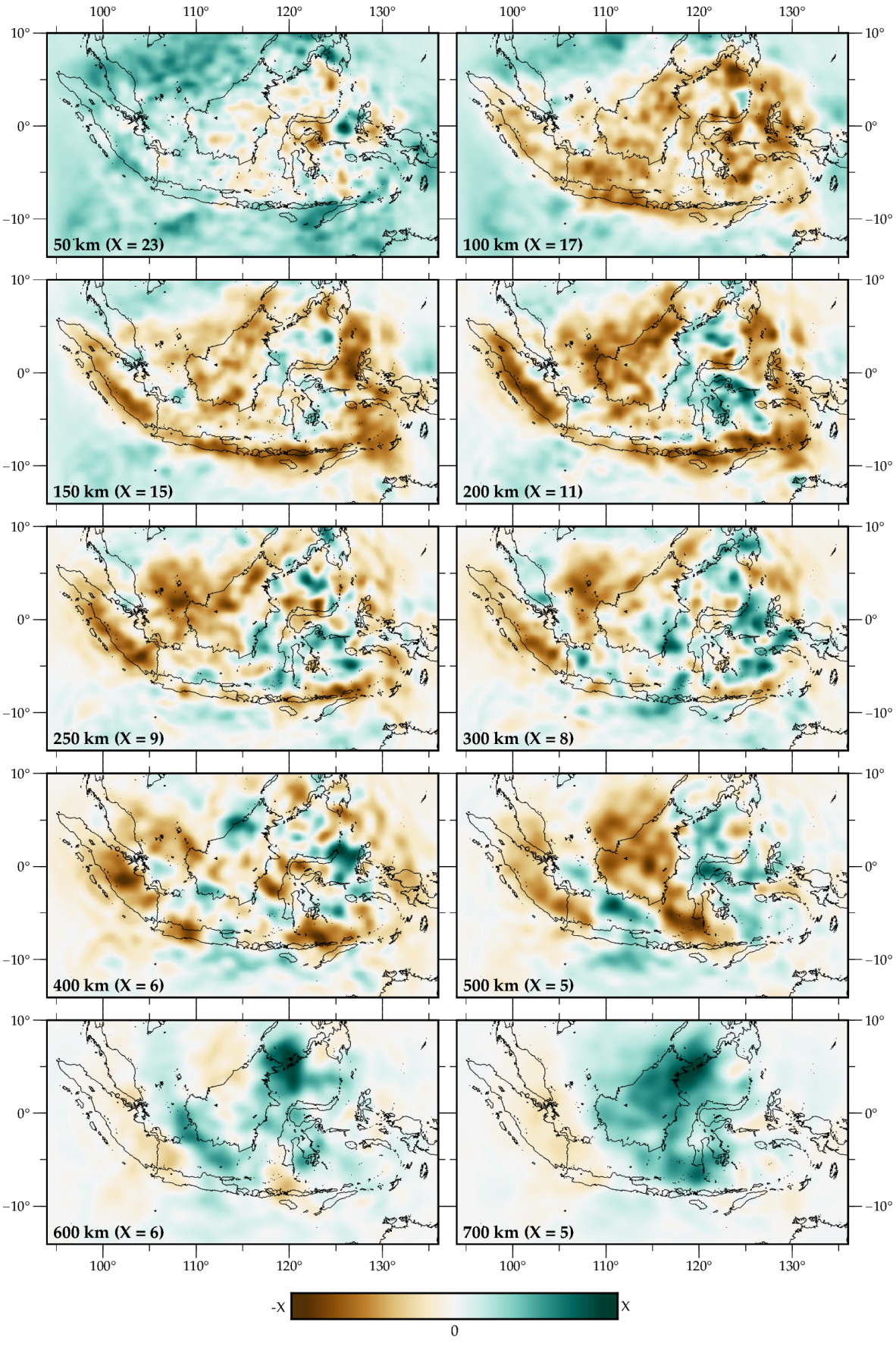
v_{SV} depth slices from 50 to 700 km. Perturbations are in % relative to the depth-average. The limits of the colourscale X are shown in the lower left corner of each plot.



v_P depth slices from 50 to 700 km. Perturbations are in % relative to the depth-average. The limits of the colourscale X are shown in the lower left corner of each plot.



Density (ρ) depth slices from 50 to 700 km. Perturbations are in % relative to the depth-average. The limits of the colourscale X are shown in the lower left corner of each plot.



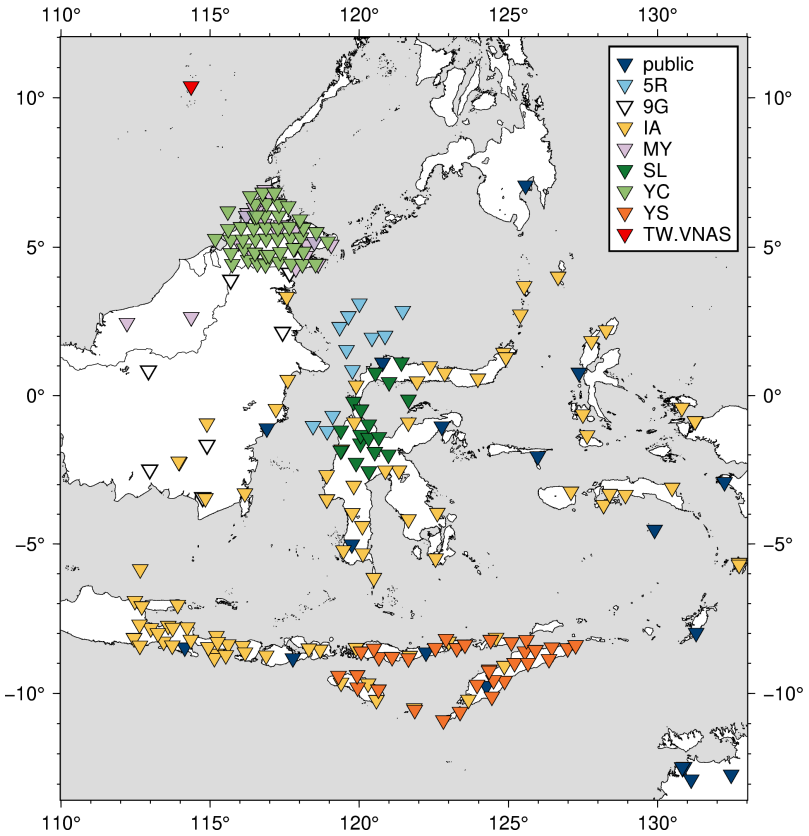
Radial anisotropy ($\frac{v_{SH}-v_{SV}}{v_S}$) depth slices from 50 to 700 km in % for the final model. The limits of the colourscale X are shown in the lower left corner of the plots.

A4: *SASSIER22* station availability

This section provides an overview of the 255 stations used to obtain *SASSIER22* (see Chapter 6). Publicly available waveforms[1] including instrument responses were downloaded automatically using *obspyDMT* (Hosseini and Sigloch, 2017). However, the majority of the 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 et al., 2018)
- *5R* (Rawlinson et al., 2020)
- *SL* (Greenfield et al., in prep.)

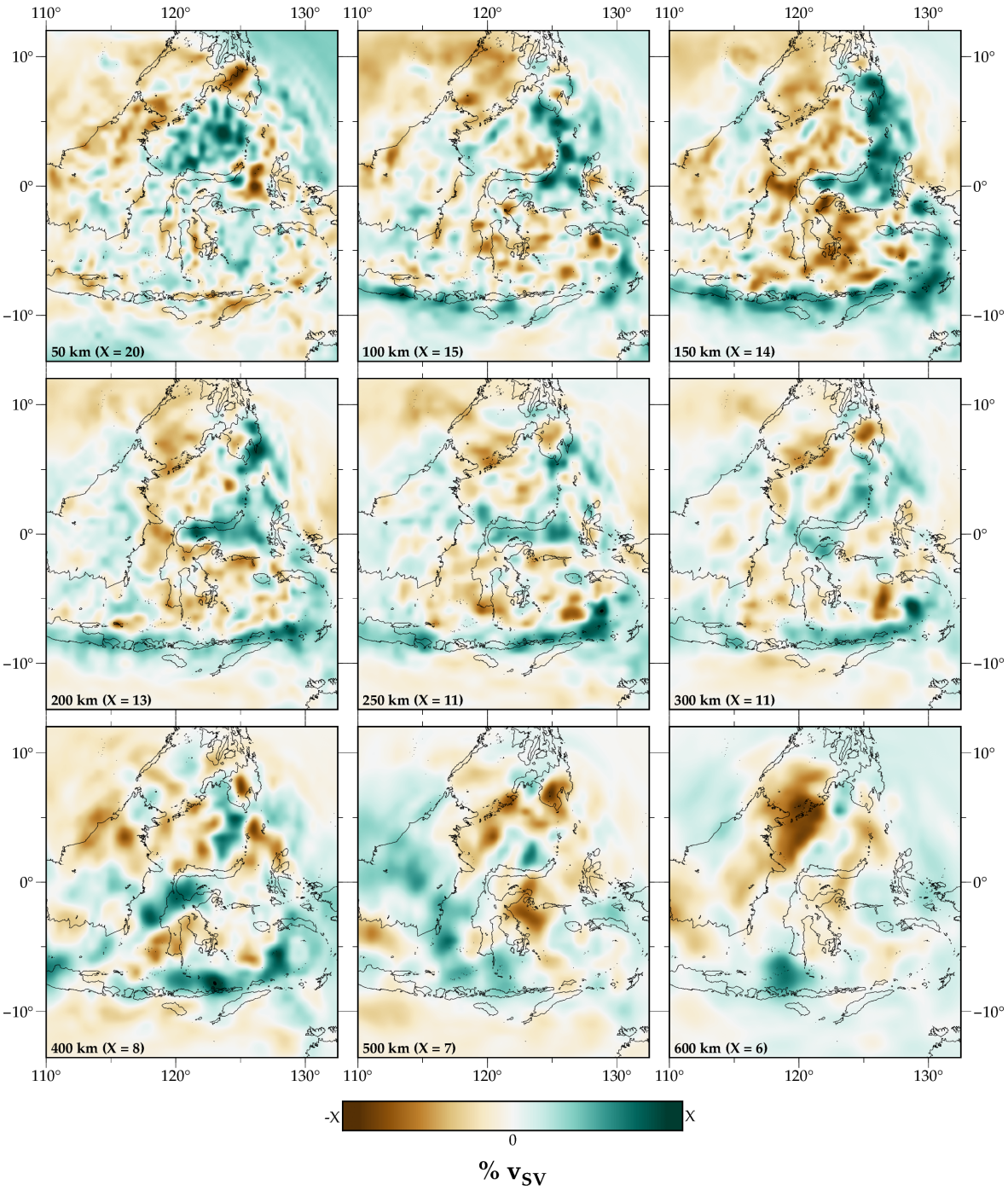
∗ These networks were also used to obtain the starting model *SASSY21* (see Section 4.2.1). Network *5R* and *SL* only became available later.



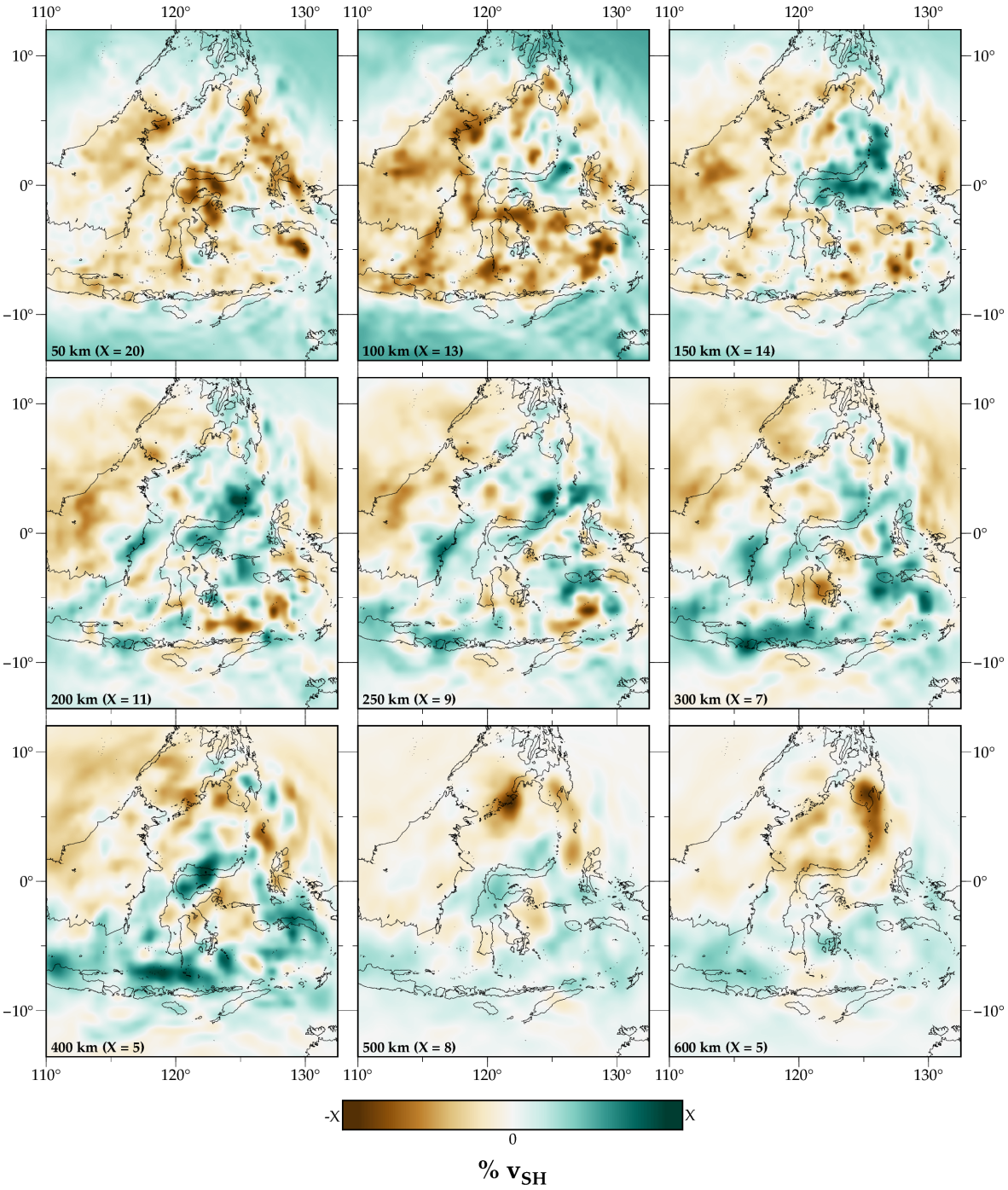
Map showing the locations of the 255 stations used to obtain *SASSIER22* (see Chapter 6).

[1]This includes data from the following networks: *AU* (Glanville and Geoscience Australia, 2021), *GE* (GEOFON Data Centre, 1993), *II* (Scripps Institution of Oceanography, 1986), *IU* (Albuquerque Seismological Laboratory/USGS, 1988) and *MY*.

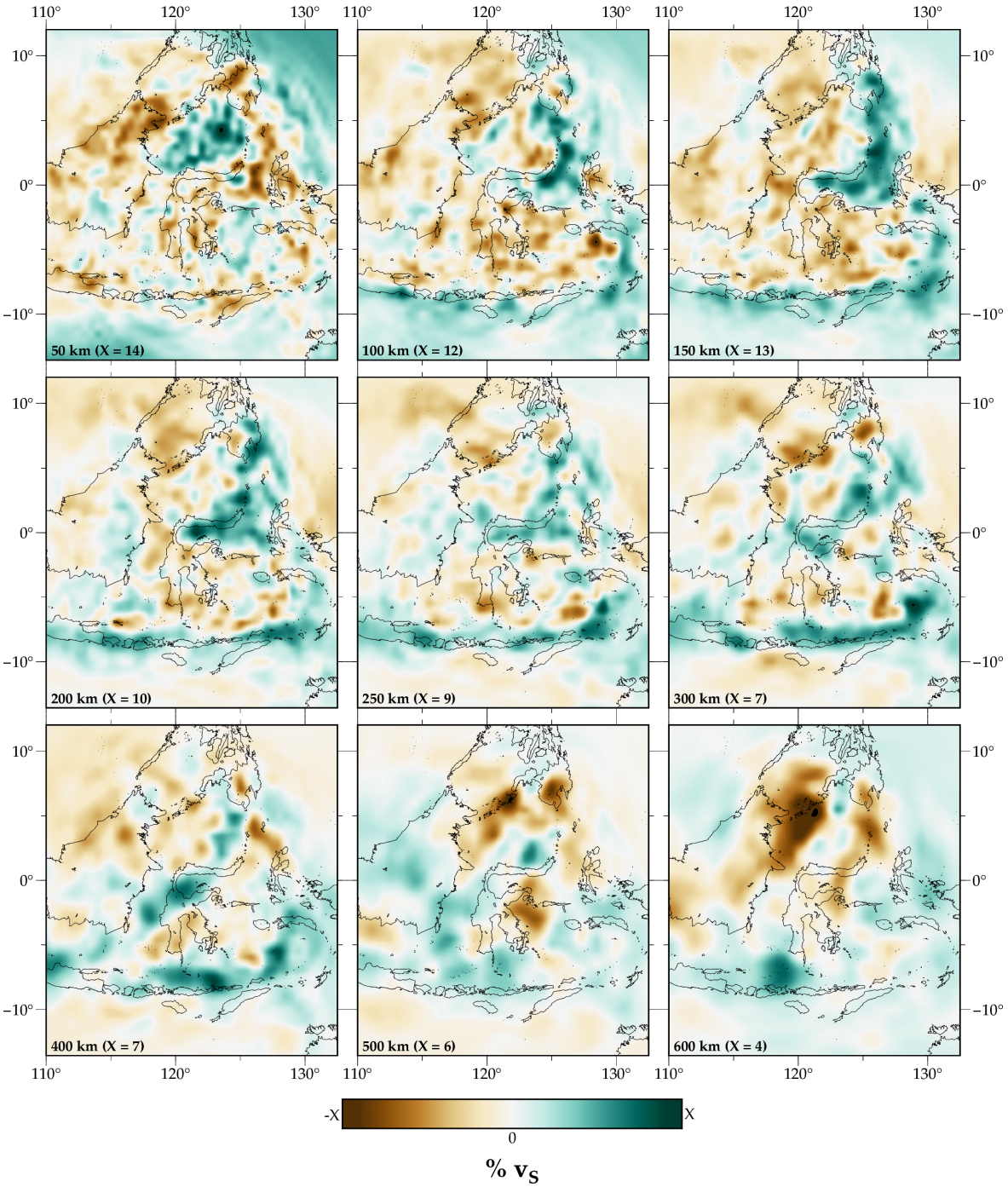
A5: *SASSIER22* depth slices



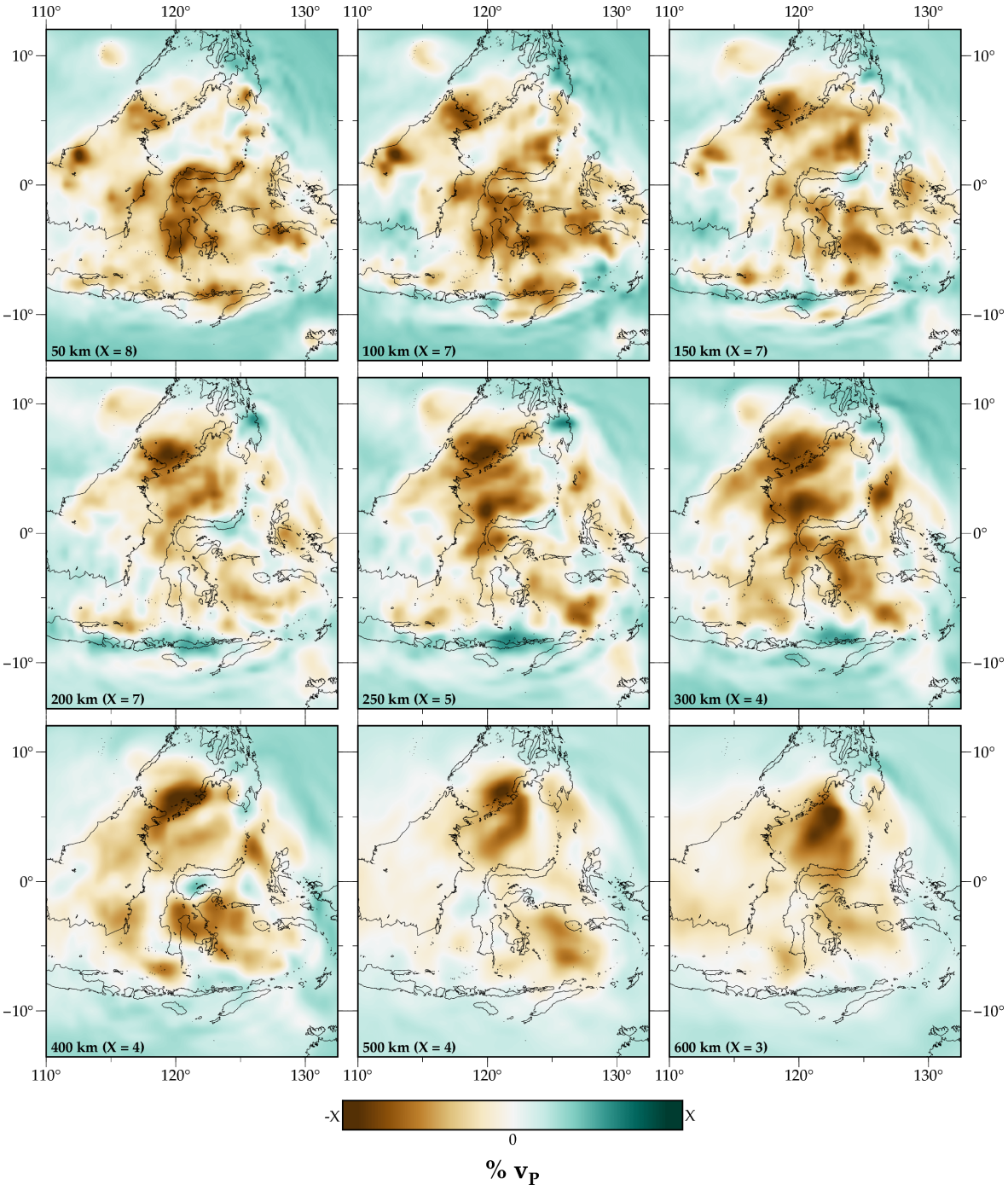
v_{SV} depth slices from 50 to 700 km. Perturbations are in % relative to the depth-average. The limits of the colourscale X are shown in the lower left corner of each plot.



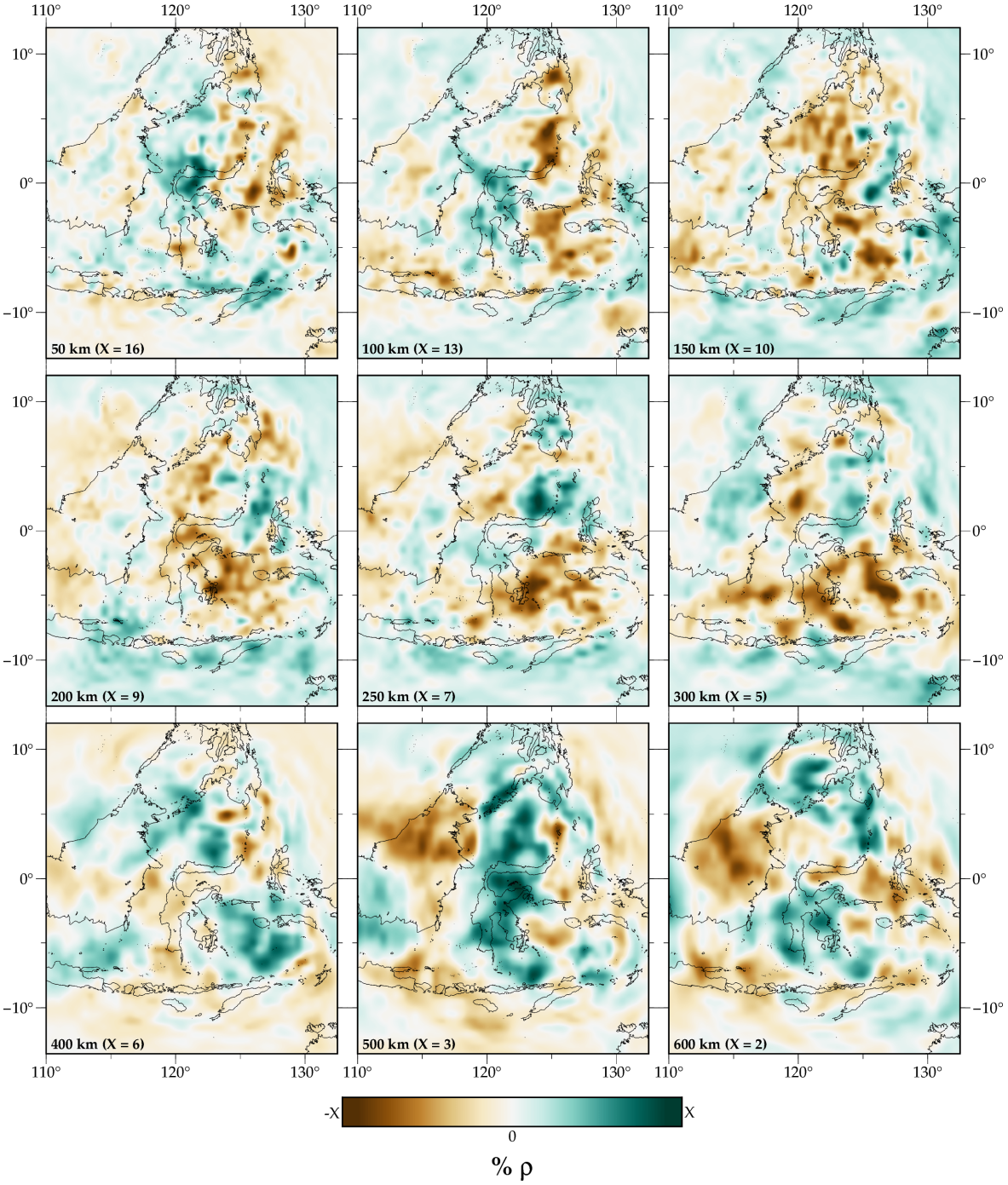
v_{SH} depth slices from 50 to 700 km. Perturbations are in % relative to the depth-average. The limits of the colourscale X are shown in the lower left corner of each plot.



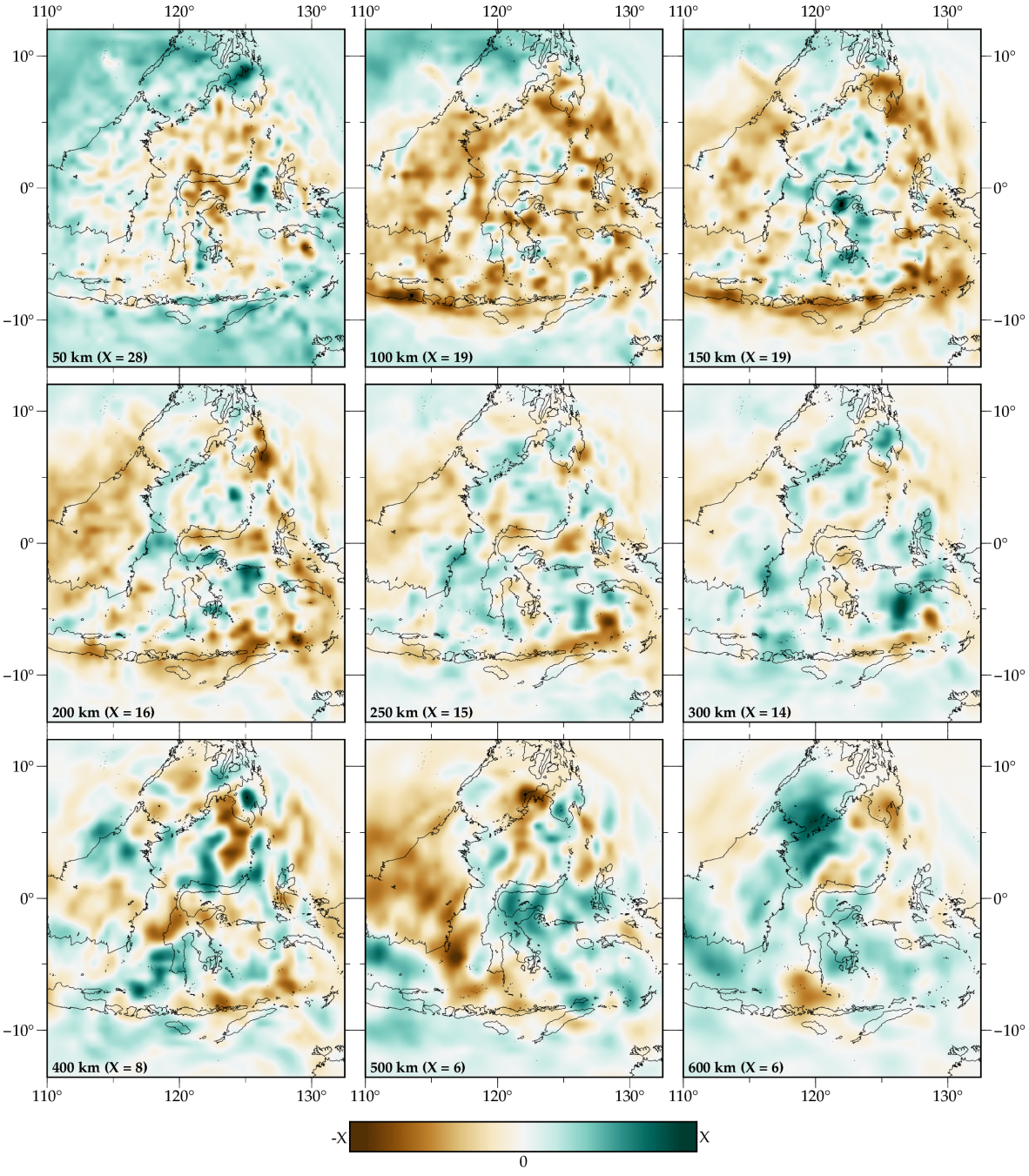
v_S depth slices from 50 to 700 km. v_S is computed using the Voigt average: $v_S = \sqrt{(2{v_{SV}}^2 + {v_{SH}}^2)/3}$ (e.g. Babuska and Cara, 1991; Panning and Romanowicz, 2006. Perturbations are in % relative to the depth-average. The limits of the colorscale *X* are shown in the lower left corner of each plot.



v_P depth slices from 50 to 700 km. Perturbations are in % relative to the depth-average. The limits of the colourscale X are shown in the lower left corner of each plot.



Density (ρ) depth slices from 50 to 700 km. Perturbations are in % relative to the depth-average. The limits of the colourscale X are shown in the lower left corner of each plot.



Radial anisotropy ($\frac{v_{SH}-v_{SV}}{v_S}$) depth slices from 50 to 700 km in % for the final model. The limits of the colourscale X are shown in the lower left corner of the plots.

References

Afanasiev, M., Böhm, C., van Driel, M., Krischer, L., Rietmann, M., May, D. A., Knepley, M. G., and Fichtner, A. (2019). "Modular and flexible spectral-element waveform modelling in two and three dimensions". *Geophysical Journal International* vol. 216.3, pp. 1675–1692.

Afonso, J. C., Rawlinson, N., Yang, Y., Schutt, D. L., Jones, A. G., Fullea, J., and Griffin, W. L. (2016). "3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle: III. Thermochemical tomography in the Western-Central US". *Journal of Geophysical Research: Solid Earth* vol. 121.10, pp. 7337–7370.

Ágústsdóttir, T., Winder, T., Woods, J., White, R. S., Greenfield, T., and Brandsdóttir, B. (2019). "Intense seismicity during the 2014–2015 Bárðarbunga-Holuhraun rifting event, Iceland, reveals the nature of dike-induced earthquakes and caldera collapse mechanisms". *Journal of Geophysical Research: Solid Earth* vol. 124.8, pp. 8331–8357.

Ahrens, J., Geveci, B., and Law, C. (2005). "Paraview: An end-user tool for large data visualization". *The visualization handbook* vol. 717.8.

Aki, K., Christoffersson, A., and Husebye, E. S. (1976). "Three-dimensional seismic structure of the lithosphere under Montana LASA". *Bulletin of the Seismological Society of America* vol. 66.2, pp. 501–524.

Aki, K. and Lee, W. (1976). "Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model". *Journal of Geophysical Research* vol. 81.23, pp. 4381–4399.

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

Albuquerque Seismological Laboratory/USGS (1988). *Global Seismograph Network - IRIS/USGS*. DOI: 10.7914/SN/IU. URL: https://www.fdsn.org/networks/detail/IU/.

Allen, R. M., Nolet, G., Morgan, W. J., Vogfjörd, K., Nettles, M., Ekström, G., Bergsson, B. H., Erlendsson, P., Foulger, G., Jakobsdóttir, S., et al. (2002). "Plume-driven plumbing and crustal formation in Iceland". *Journal of Geophysical Research: Solid Earth* vol. 107.B8, ESE–4.

Altamimi, Z., Rebischung, P., Métivier, L., and Collilieux, X. (2016). "ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions". *Journal of Geophysical Research: Solid Earth* vol. 121.8, pp. 6109–6131.

Amante, C. and Eakins, B. W. (2009). "ETOPO1 arc-minute global relief model: procedures, data sources and analysis".

Amaru, M. (2007). "Global travel time tomography with 3-D reference models". PhD thesis. University Utrecht.

An, C., Yue, H., Sun, J., Meng, L., and Báez, J. C. (Sept. 2017). "The 2015 Mw 8.3 Illapel, Chile, Earthquake: Direction-Reversed Along-Dip Rupture with Localized Water Reverberation".

Bulletin of the Seismological Society of America vol. 107.5, pp. 2416–2426. ISSN: 0037-1106. DOI: 10.1785/0120160393.

Anderson, J. E., Tan, L., and Wang, D. (2012). "Time-reversal checkpointing methods for RTM and FWI". *Geophysics* vol. 77.4, pp. 93–103.

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

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

Aurelio, M. A. (2000). "Shear partitioning in the Philippines: constraints from Philippine Fault and global positioning system data". *Island Arc* vol. 9.4, pp. 584–597.

Australian National University (2011). *Australian Seismometers in Schools*. DOI: 10.7914/SN/S1. URL: http://www.fdsn.org/doi/10.7914/SN/S1.

Babuska, V. and Cara, M. (1991). *Seismic anisotropy in the Earth*.

Bachmann, E. and Tromp, J. (2020). "Source encoding for viscoacoustic ultrasound computed tomography". *The Journal of the Acoustical Society of America* vol. 147.5, pp. 3221–3235.

Backus, G. and Gilbert, F. (1970). "Uniqueness in the inversion of inaccurate gross earth data". *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* vol. 266.1173, pp. 123–192.

Bacon, C. A., Rawlinson, N., Pilia, S., Gilligan, A., Wehner, D., Cornwell, D. G., and Tongkul, F. (submitted). "The signature of lithospheric anisotropy at post-subduction continental margins: new insight from XKS splitting analysis in northern Borneo". *Geochemistry, Geophysics, Geosystems*.

Bamberger, A, Chavent, G, Hemon, C., and Lailly, P (1982). "Inversion of normal incidence seismograms". *Geophysics* vol. 47.5, pp. 757–770.

Bamberger, A. et al. (1977). "Une application de la theorie du controle a un probleme inverse de sismique".

Barnier, G., Biondi, E., and Biondi, B. (2018). "Full waveform inversion by model extension". *2018 SEG International Exposition and Annual Meeting*.

Berkhout, A. (1984). "Multidimensional linearized inversion and seismic migration". *Geophysics* vol. 49.11, pp. 1881–1895.

Bernauer, M., Fichtner, A., and Igel, H. (2014). "Optimal observables for multiparameter seismic tomography". *Geophysical Journal International* vol. 198.2, pp. 1241–1254.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. (2010). "ObsPy: A Python toolbox for seismology". *Seismological Research Letters* vol. 81.3, pp. 530–533.

Bianchi, M., Evans, P., Heinloo, A., and Quinteros, J. (2015). "WebDC3 web interface".

Bijwaard, H. and Spakman, W. (1999). "Tomographic evidence for a narrow whole mantle plume below Iceland". *Earth and Planetary Science Letters* vol. 166.3-4, pp. 121–126.

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

Bird, P. (2003). "An updated digital model of plate boundaries". *Geochemistry, Geophysics, Geosystems* vol. 4.3.

Bjarnason, I. T., Menke, W., Flóvenz, Ó. G., and Caress, D. (1993). "Tomographic image of the mid-Atlantic plate boundary in southwestern Iceland". *Journal of Geophysical Research: Solid Earth* vol. 98.B4, pp. 6607–6622.

Bjarnason, I. T. and Schmeling, H. (2009). "The lithosphere and asthenosphere of the Iceland hotspot from surface waves". *Geophysical Journal International* vol. 178.1, pp. 394–418.

Blom, N. (2018). "Towards imaging density using waveform tomography". PhD thesis. University Utrecht.

Blom, N., Böhm, C., and Fichtner, A. (2017). "Synthetic inversions for density using seismic and gravity data". *Geophysical Journal International* vol. 209.2, pp. 1204–1220.

Blom, N., Gokhberg, A., and Fichtner, A. (2020). "Seismic waveform tomography of the central and eastern Mediterranean upper mantle". *Solid Earth* vol. 11.2, pp. 669–690.

Blom, N., Hardalupas, P.-S., and Rawlinson, N. (2022). "Mitigating the effect of errors in earthquake parameters on seismic (waveform) tomography". *Geophysical Journal International*. DOI: `https://doi.org/10.31223/X5F929`.

Böhm, C., Afanasiev, M., Krischer, L., van Driel, M., and Fichtner, A. (2019). "Anisotropic diffusion-based smoothing filters for full-waveform inversion." vol. 21.

Böhm, C., Hanzich, M., de la Puente, J., and Fichtner, A. (2016). "Wavefield compression for adjoint methods in full-waveform inversion". *Geophysics* vol. 81.6, R385–R397.

Bormann, P., Engdahl, B., and Kind, R. (2012). "Seismic wave propagation and earth models". *New manual of seismological observatory practice 2 (NMSOP2)*. Deutsches GeoForschungsZentrum GFZ, pp. 1–105.

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., and Pugmire, D. (2016). "Global adjoint tomography: first-generation model". *Geophysical Journal International* vol. 207.3, pp. 1739–1766.

Bozdağ, E. and Trampert, J. (2008). "On crustal corrections in surface wave tomography". *Geophysical Journal International* vol. 172.3, pp. 1066–1082.

Bozdağ, E., Trampert, J., and Tromp, J. (2011). "Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements". *Geophysical Journal International* vol. 185.2, pp. 845–870.

Brandsdóttir, B., Menke, W., Einarsson, P., White, R. S., and Staples, R. K. (1997). "Færoe-Iceland ridge experiment 2. Crustal structure of the Krafla central volcano". *Journal of Geophysical Research: Solid Earth* vol. 102.B4, pp. 7867–7886.

Broyden, C. G. (1970). "The convergence of a class of double-rank minimization algorithms 1. general considerations". *IMA Journal of Applied Mathematics* vol. 6.1, pp. 76–90.

Bunks, C., Saleck, F. M., Zaleski, S, and Chavent, G (1995). "Multiscale seismic waveform inversion". *Geophysics* vol. 60.5, pp. 1457–1473.

Capdeville, Y., Guillot, L., and Marigo, J.-J. (2010a). "1-D non-periodic homogenization for the seismic wave equation". *Geophysical Journal International* vol. 181.2, pp. 897–910.

— (2010b). "2-D non-periodic homogenization to upscale elastic media for P–SV waves". *Geophysical Journal International* vol. 182.2, pp. 903–922.

Carter, D. J., Audley-Charles, M. G., and Barber, A. (1976). "Stratigraphical analysis of island arc—continental margin collision in eastern Indonesia". *Journal of the Geological Society* vol. 132.2, pp. 179–198.

Cecchetto, B and Heidrich, W. (2011). "Probabilistic inverse dynamics for blood pattern reconstruction". *16th International Workshop on Vision, Modeling and Visualization, VMV 2011*, pp. 369–376.

Celli, N. L., Lebedev, S., Schaeffer, A. J., and Gaina, C. (2021). "The tilted Iceland Plume and its effect on the North Atlantic evolution and magmatism". *Earth and Planetary Science Letters* vol. 569, p. 117048.

Celli, N. L., Lebedev, S., Schaeffer, A. J., Ravenna, M., and Gaina, C. (2020). "The upper mantle beneath the South Atlantic Ocean, South America and Africa from waveform tomography with massive data sets". *Geophysical Journal International* vol. 221.1, pp. 178–204.

Cerjan, C., Kosloff, D., Kosloff, R., and Reshef, M. (1985). "A nonreflecting boundary condition for discrete acoustic and elastic wave equations". *Geophysics* vol. 50.4, pp. 705–708.

Červený, V (2001). *Seismic Ray Theory*.

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., and Jordan, T. H. (2007). "Full 3D tomography for the crustal structure of the Los Angeles region". *Bulletin of the Seismological Society of America* vol. 97.4, pp. 1094–1120.

Chen, P.-F., Chien, M., Bina, C. R., Yen, H.-Y., and Olavere, E. A. (2020). "Evidence of an east-dipping slab beneath the southern end of the Philippine Trench (1° N–6° N) as revealed by ISC-EHB". *Journal of Asian Earth Sciences* vol. 4, p. 100034.

Chow, B., Kaneko, Y., Tape, C., Modrak, R., Mortimer, N., Bannister, S., and Townend, J. (2022). "Strong upper-plate heterogeneity at the Hikurangi subduction margin (North Island, New Zealand) imaged by adjoint tomography". *Journal of Geophysical Research: Solid Earth* vol. 127.1, e2021JB022865.

Clayton, R. and Engquist, B. (1977). "Absorbing boundary conditions for acoustic and elastic wave equations". *Bulletin of the Seismological Society of America* vol. 67.6, pp. 1529–1540.

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

Courant, R., Friedrichs, K., and Lewy, H. (1928). "Über die partiellen Differenzengleichungen der mathematischen Physik". *Mathematische annalen* vol. 100.1, pp. 32–74.

Crotwell, H. P., Owens, T. J., and Ritsema, J. (1999). "The TauP Toolkit: Flexible seismic travel-time and ray-path utilities". *Seismological Research Letters* vol. 70.2, pp. 154–160.

Dahlen, F., Hung, S.-H., and Nolet, G. (2000). "Fréchet kernels for finite-frequency travel-times—I. Theory". *Geophysical Journal International* vol. 141.1, pp. 157–174.

Dalton, C. A., Langmuir, C. H., and Gale, A. (2014). "Geophysical and geochemical evidence for deep temperature variations beneath mid-ocean ridges". *Science* vol. 344.6179, pp. 80–83.

Darbyshire, F. A., Bjarnason, I. T., White, R. S., and Flóvenz, Ó. G. (1998). "Crustal structure above the Iceland mantle plume imaged by the ICEMELT refraction profile". *Geophysical Journal International* vol. 135.3, pp. 1131–1149.

Darbyshire, F. A., White, R. S., and Priestley, K. F. (2000). "Structure of the crust and uppermost mantle of Iceland from a combined seismic and gravity study". *Earth and Planetary Science Letters* vol. 181.3, pp. 409–428.

Doody, C., Rodgers, A., Böhm, C., Afanasiev, M., Krischer, L., Chiang, A., and Simmons, N. (2022). "Using K-Means Clustering to Compare Adjoint Waveform Tomography Models of California and Nevada". *EGU General Assembly Conference Abstracts*.

Dreger, D. S. (2003). "TDMT_INV: Time domain seismic moment tensor inversion". *Elsevier* vol. 81, p. 1627.

Dreiling, J., Tilmann, F., Yuan, X., Haberland, C., and Seneviratne, S. M. (2020). "Crustal structure of Sri Lanka derived from joint inversion of surface wave dispersion and receiver functions using a Bayesian approach". *Journal of Geophysical Research: Solid Earth* vol. 125.5, e2019JB018688.

van Driel, M., Böhm, C., Krischer, L., and Afanasiev, M. (2020). "Accelerating numerical wave propagation using wavefield adapted meshes. Part I: forward and adjoint modelling". *Geophysical Journal International* vol. 221.3, pp. 1580–1590.

van Driel, M., Kemper, J., and Böhm, C. (2021). "On the modeling of self-gravitation for full 3D global seismic wave propagation". *Geophysical Journal International*.

van Driel, M. and Nissen-Meyer, T. (2014). "Optimized viscoelastic wave propagation for weakly dissipative media". *Geophysical Journal International* vol. 199.2, pp. 1078–1093.

Durand, S, Debayle, E, Ricard, Y, Zaroli, C, and Lambotte, S (2017). "Confirmation of a change in the global shear velocity pattern at around 1000 km depth". *Geophysical Journal International* vol. 211.3, pp. 1628–1639.

Dziewoński, A. (1984). "Mapping the lower mantle: determination of lateral heterogeneity in P velocity up to degree and order 6". *Journal of Geophysical Research: Solid Earth* vol. 89.B7, pp. 5929–5952.

Dziewoński, A. and Anderson, D. L. (1981). "Preliminary reference Earth model". *Physics of the Earth and Planetary Interiors* vol. 25.4, pp. 297–356.

Dziewoński, A., Hager, B. H., and O'Connell, R. J. (1977). "Large-scale heterogeneities in the lower mantle". *Journal of Geophysical Research* vol. 82.2, pp. 239–255.

Eberhart-Phillips, D. and Chadwick, M. (2002). "Three-dimensional attenuation model of the shallow Hikurangi subduction zone in the Raukumara Peninsula, New Zealand". *Journal of Geophysical Research: Solid Earth* vol. 107.B2, ESE–3.

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

Elburg, M., van Bergen, M., and Foden, J. (2004). "Subducted upper and lower continental crust contributes to magmatism in the collision sector of the Sunda-Banda arc, Indonesia". *Geology* vol. 32.1, pp. 41–44.

Engdahl, E. R., Di Giacomo, D., Sakarya, B., Gkarlaouni, C. G., Harris, J., and Storchak, D. A. (2020). "ISC-EHB 1964–2016, an improved data set for studies of Earth structure and global seismicity". *Earth and Space Science* vol. 7.1, e2019EA000897.

Espindola-Carmona, A. and Peter, D. B. (2018). "Impact of surface topography on full-waveform tomography for Central Mexico". *AGU Fall Meeting Abstracts*.

Fan, J. and Zhao, D. (2018). "Evolution of the southern segment of the Philippine Trench: Constraints from seismic tomography". *Geochemistry, Geophysics, Geosystems* vol. 19.11, pp. 4612–4627.

Fernando, B., Leng, K., and Nissen-Meyer, T. (2020). "Oceanic high-frequency global seismic wave propagation with realistic bathymetry". *Geophysical Journal International* vol. 222.2, pp. 1178–1194.

Fichtner, A. (2010). *Full seismic waveform modelling and inversion*.

Fichtner, A., De Wit, M., and 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* vol. 297.3-4, pp. 405–412.

Fichtner, A., van Herwaarden, D.-P., Afanasiev, M., Simutė, S., Krischer, L., Çubuk-Sabuncu, Y., Taymaz, T., Colli, L., Saygin, E., Villaseñor, A., et al. (2018). "The collaborative seismic earth model: generation 1". *Geophysical research letters* vol. 45.9, pp. 4007–4016.

Fichtner, A. and Igel, H. (2008). "Efficient numerical surface wave propagation through the optimization of discrete crustal models—a technique based on non-linear dispersion curve matching (DCM)". *Geophysical Journal International* vol. 173.2, pp. 519–533.

Fichtner, A., Kennett, B. L., Igel, H., and Bunge, H.-P. (2008). "Theoretical background for continental-and global-scale full-waveform inversion in the time–frequency domain". *Geophysical Journal International* vol. 175.2, pp. 665–685.

— (2009). "Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods". *Geophysical Journal International* vol. 179.3, pp. 1703–1725.

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

Fichtner, A. and van Leeuwen, T. (2015). "Resolution analysis by random probing". *Journal of Geophysical Research: Solid Earth* vol. 120.8, pp. 5549–5573.

Fichtner, A. and Simutė, S. (2018). "Hamiltonian Monte Carlo inversion of seismic sources in complex media". *Journal of Geophysical Research: Solid Earth* vol. 123.4, pp. 2984–2999.

Fichtner, A. and Trampert, J. (2011). "Hessian kernels of seismic data functionals based upon adjoint techniques". *Geophysical Journal International* vol. 185.2, pp. 775–798.

Fletcher, R. and Reeves, C. M. (1964). "Function minimization by conjugate gradients". *The Computer Journal* vol. 7.2, pp. 149–154.

Fletcher, R. (1970). "A new approach to variable metric algorithms". *The Computer Journal* vol. 13.3, pp. 317–322.

Foulger, G., Pritchard, M., Julian, B., Evans, J., Allen, R., Nolet, G, Morgan, W., Bergsson, B., Erlendsson, P, Jakobsdottir, S, et al. (2001). "Seismic tomography shows that upwelling beneath Iceland is confined to the upper mantle". *Geophysical Journal International* vol. 146.2, pp. 504–530.

French, S. W. and Romanowicz, B. (2015). "Broad plumes rooted at the base of the Earth's mantle beneath major hotspots". *Nature* vol. 525.7567, pp. 95–99.

French, S. and Romanowicz, B. A. (2014). "Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography". *Geophysical Journal International* vol. 199.3, pp. 1303–1327.

Fukao, Y. and Obayashi, M. (2013). "Subducted slabs stagnant above, penetrating through, and trapped below the 660 km discontinuity". *Journal of Geophysical Research: Solid Earth* vol. 118.11, pp. 5920–5938.

Gao, Y., Tilmann, F., van Herwaarden, D.-P., Thrastarson, S., Fichtner, A., Heit, B., Yuan, X., and 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. DOI: https://doi.org/10.1029/2021JB021984.

Garcia, R., Schardong, L, and Chevrot, S (2013). "A nonlinear method to estimate source parameters, amplitude, and travel times of teleseismic body waves". *Bulletin of the Seismological Society of America* vol. 103.1, pp. 268–282.

Gasser, D., Grenne, T., Corfu, F., Bøe, R., Røhr, T. S., and Slagstad, T. (2021). "Concurrent MORB-type and ultrapotassic volcanism in an extensional basin along the Laurentian Iapetus margin: Tectonomagmatic response to Ordovician arc-continent collision and subduction polarity flip". *GSA Bulletin*.

Gauthier, O., Virieux, J., and Tarantola, A. (1986). "Two-dimensional nonlinear inversion of seismic waveforms: Numerical results". *Geophysics* vol. 51.7, pp. 1387–1403.

Gebraad, L., Böhm, C., and Fichtner, A. (2020). "Bayesian elastic full-waveform inversion using Hamiltonian Monte Carlo". *Journal of Geophysical Research: Solid Earth* vol. 125.3, e2019JB018428.

GEOFON Data Centre (1993). *GEOFON Seismic Network*. en. DOI: 10.14470/TR560404. URL: http://geofon.gfz-potsdam.de/doi/network/GE.

Gilbert, F. and Dziewonski, A. M. (1975). "An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra". *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 278.1280, pp. 187–269.

Gilligan, A, Cornwell, D., Rawlinson, N, Tongkul, F, and Pilia, S (2018). "Post-subduction tectonics in North Borneo: constraints from P receiver functions". *AGU Fall Meeting Abstracts*.

Glanville, H. and Geoscience Australia (2021). *Australian National Seismograph Network Data Collection*. DOI: 10.26186/144675. URL: http://pid.geoscience.gov.au/dataset/ga/144675.

Gokhberg, A, Simutė, S, and Fichtner, A (2016). "SES3D-An Open-source Package for Seismic Waveform Modelling and Inversion". vol. 2016.1, pp. 1–5.

Goldfarb, D. (1970). "A family of variable-metric methods derived by variational means". *Mathematics of Computation* vol. 24.109, pp. 23–26.

Gou, T., Zhao, D., Huang, Z., and Wang, L. (2019). "Aseismic deep slab and mantle flow beneath Alaska: Insight from anisotropic tomography". *Journal of Geophysical Research: Solid Earth* vol. 124.2, pp. 1700–1724.

Grand, S. P. (2002). "Mantle shear–wave tomography and the fate of subducted slabs". *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, pp. 2475–2491.

Green, R. G., Priestley, K. F., and White, R. S. (2017). "Ambient noise tomography reveals upper crustal structure of Icelandic rifts". *Earth and Planetary Science Letters* vol. 466, pp. 20–31.

Greenfield, T., Gilligan, A., Pilia, S., Cornwell, D. G., Tongkul, F., Widiyantoro, S., and Rawlinson, N. (2022). "Post-Subduction Tectonics of Sabah, Northern Borneo, Inferred From Surface Wave Tomography". *Geophysical Research Letters* vol. 49.3, e2021GL096117.

Greenfield, T., White, R. S., and Roecker, S. (2016). "The magmatic plumbing system of the Askja central volcano, Iceland, as imaged by seismic tomography". *Journal of Geophysical Research: Solid Earth* vol. 121.10, pp. 7211–7229.

Greenfield, T., White, R. S., Winder, T., and Ágústsdóttir, T. (2020). "Seismicity of the Askja and Bárðarbunga volcanic systems of Iceland, 2009–2015". *Journal of Volcanology and Geothermal Research* vol. 391, p. 106432.

Greenfield, T., Widiyantoro, S., and Rawlinson, N. (2018). *Kalimantan Temporary Network*. DOI: 10.7914/SN/9G_2018. URL: http://www.fdsn.org/networks/detail/9G_2018/.

Gu, Y. J., Dziewoński, A., and Ekström, G. (2001). "Preferential detection of the Lehmann discontinuity beneath continents". *Geophysical Research Letters* vol. 28.24, pp. 4655–4658.

Guasch, L., Agudo, O. C., Tang, M.-X., Nachev, P., and Warner, M. (2020). "Full-waveform inversion imaging of the human brain". *NPJ Digital Medicine* vol. 3.1, pp. 1–12.

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

Gung, Y and Romanowicz, B (2004). "Q tomography of the upper mantle using three-component long-period waveforms". *Geophysical Journal International* vol. 157.2, pp. 813–830.

Gutenberg, B. (1914). "Über Erdbebenwellen. VII A. Beobachtungen an Registrierungen von Fernbeben in Göttingen und Folgerung über die Konstitution des Erdkörpers (mit Tafel)". *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* vol. 1914, pp. 125–176.

Hadamard, J. (1902). "Sur les problèmes aux dérivées partielles et leur signification physique". *Princeton University Bulletin*, pp. 49–52.

Hall, R and Wilson, M. (2000). "Neogene sutures in eastern Indonesia". *Journal of Asian Earth Sciences* vol. 18.6, pp. 781–808.

Hall, R. (1998). "The plate tectonics of Cenozoic SE Asia and the distribution of land and sea". *Biogeography and geological evolution of SE Asia*, pp. 99–131.

— (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* vol. 20.4, pp. 353–431.

— (2011). "Australia–SE Asia collision: plate tectonics and crustal flow". *Geological Society, London, Special Publications* vol. 355.1, pp. 75–109.

Hall, R. (2012). "Late Jurassic–Cenozoic reconstructions of the Indonesian region and the Indian Ocean". *Tectonophysics* vol. 570, pp. 1–41.

— (2013). "Contraction and extension in northern Borneo driven by subduction rollback". *Journal of Asian Earth Sciences* vol. 76, pp. 399–411.

— (2014). "The origin of Sundaland". *Procedings of Sundaland Resources*.

— (2019). "The subduction initiation stage of the Wilson cycle". *Geological Society, London, Special Publications* vol. 470.1, pp. 415–437.

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

Hall, R. and Spakman, W. (2015). "Mantle structure and tectonic history of SE Asia". *Tectonophysics* vol. 658, pp. 14–45.

Hamilton, W. B. (1979). *Tectonics of the Indonesian region*. Vol. 1078.

Häner, R. and Kriegel, U. (2008). "From GITEWS to InaTEWS: An Architectural Blue Print".

Harig, S., Immerz, A., Griffin, J., Weber, B., Babeyko, A., Rakowsky, N., Hartanto, D., Nurokhim, A., Handayani, T., Weber, R., et al. (2020). "The Tsunami Scenario Database of the Indonesia Tsunami Early Warning System (InaTEWS): evolution of the coverage and the involved modeling approaches". *Pure and Applied Geophysics* vol. 177.3, pp. 1379–1401.

Harris, C., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., et al. (2019). "Array programming with NumPy". *Nature* vol. 585.7825, pp. 357–362.

Harris, C., Miller, M. S., Supendi, P., and Widiyantoro, S. (2020). "Subducted lithospheric boundary tomographically imaged beneath arc-continent collision in eastern Indonesia". *Journal of Geophysical Research: Solid Earth* vol. 125.8, e2019JB018854.

Harris, P. A., Scognamiglio, L, Magnoni, F, Casarotti, E, and Tinti, E (2022). "Centroid Moment Tensor catalog with 3D lithospheric wavespeed model: the 2016–2017 Central Apennines sequence". *Journal of Geophysical Research: Solid Earth*, e2021JB023068.

Harris, R (2011). "The nature of the Banda Arc–continent collision in the Timor region". *Arc-continent collision*. Springer, pp. 163–211.

Havskov, J. and Alguacil, G. (2004). *Instrumentation in earthquake seismology*. Vol. 358. Springer.

Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., and Smoczyk, G. M. (2018). "Slab2, a comprehensive subduction zone geometry model". *Science* vol. 362.6410, pp. 58–61.

Hearn, T. M. and Clayton, R. W. (1986). "Lateral velocity variations in southern California. I. Results for the upper crust from Pg waves". *Bulletin of the Seismological Society of America* vol. 76.2, pp. 495–509.

Hejrani, B., Tkalčić, H., and Fichtner, A. (2017). "Centroid moment tensor catalogue using a 3-D continental scale Earth model: Application to earthquakes in Papua New Guinea and the Solomon Islands". *Journal of Geophysical Research: Solid Earth* vol. 122.7, pp. 5517–5543.

van Herwaarden, D. P., Böhm, C., Afanasiev, M., Thrastarson, S., Krischer, L., Trampert, J., and Fichtner, A. (2020). "Accelerated full-waveform inversion using dynamic mini-batches". *Geophysical Journal International* vol. 221.2, pp. 1427–1438.

Hess, H. (1964). "Seismic anisotropy of the uppermost mantle under oceans". *Nature* vol. 203.4945, pp. 629–631.

Hestenes, M. R. and Stiefel, E. (1952). "Methods of Conjugate Gradients for Solving". *Journal of research of the National Bureau of Standards* vol. 49.6, p. 409.

Higgs, R, Madon, M., and Watts, A. (1999). "Gravity anomalies, subsidence history and the tectonic evolution of the Malay and Penyu Basins (offshore Peninsula Malaysia)". *Basin Research* vol. 11, pp. 285–290.

van der Hilst, R. D., Widiyantoro, S., and Engdahl, E. (1997). "Evidence for deep mantle circulation from global tomography". *Nature* vol. 386.6625, p. 578.

Hirt, C. and 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* vol. 39, pp. 103–112.

Ho, T., Priestley, K., and Debayle, E. (2016). "A global horizontal shear velocity model of the upper mantle from multimode Love wave measurements". *Geophysical Journal International* vol. 207.1, pp. 542–561.

Hoggard, M. J., Czarnota, K., Richards, F. D., Huston, D. L., Jaques, A. L., and Ghelichkhan, S. (2020). "Global distribution of sediment-hosted metals controlled by craton edge stability". *Nature Geoscience* vol. 13.7, pp. 504–510.

Hosseini, K., Matthews, K. J., Sigloch, K., Shephard, G. E., Domeier, M., and Tsekhmistrenko, M. (2018). "SubMachine: Web-based tools for exploring seismic tomography and other models of Earth's deep interior". *Geochemistry, Geophysics, Geosystems* vol. 19.5, pp. 1464–1483.

Hosseini, K. and Sigloch, K. (2017). "ObspyDMT: a Python toolbox for retrieving and processing large seismological data sets". *Solid Earth* 5, pp. 1047–1070.

Hounsfield, G. N. (1980). "Computed medical imaging". *Medical Physics* vol. 7.4, pp. 283–290.

Hoyer, S. and Hamman, J. (2017). "xarray: N-D labeled arrays and datasets in Python". *Journal of Open Research Software* 5.1. DOI: 10.5334/jors.148. URL: http://doi.org/10.5334/jors.148.

Huang, Z., Zhao, D., Hasegawa, A., Umino, N., Park, J.-H., and Kang, I.-B. (2013). "Aseismic deep subduction of the Philippine Sea plate and slab window". *Journal of Asian Earth Sciences* vol. 75, pp. 82–94.

Huang, Z., Zhao, D., and Wang (2015). "P wave tomography and anisotropy beneath Southeast Asia: Insight into mantle dynamics". *Journal of Geophysical Research: Solid Earth* vol. 120.7, pp. 5154–5174.

Hung, S.-H., Dahlen, F., and Nolet, G. (2001). "Wavefront healing: a banana–doughnut perspective". *Geophysical Journal International* vol. 146.2, pp. 289–312.

Hunter, J. D. (2007). "Matplotlib: A 2D graphics environment". *IEEE Annals of the History of Computing* vol. 9.03, pp. 90–95.

Hutko, A. R., Bahavar, M., Trabant, C., Weekly, R. T., Fossen, M. V., and Ahern, T. (2017). "Data products at the IRIS-DMC: Growth and usage". *Seismological Research Letters* vol. 88.3, pp. 892–903.

Igel, H. (n.d.). *Computers, Waves, Simulations: A Practical Introduction to Numerical Methods using Python*. URL: `https://www.coursera.org/learn/computers-waves-simulations` (visited on 01/09/2021).

Igel, H., Djikpéssé, H., and Tarantola, A. (1996). "Waveform inversion of marine reflection seismograms for P impedance and Poisson's ratio". *Geophysical Journal International* vol. 124.2, pp. 363–371.

Inkscape Project (Apr. 16, 2020). *Inkscape*. Version 0.92.5. URL: `https://inkscape.org`.

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

IRIS, D. (2011). *Data services products: EMC, A repository of Earth models*.

Ishizuka, O., Taylor, R. N., Ohara, Y., and Yuasa, M. (2013). "Upwelling, rifting, and age-progressive magmatism from the Oki-Daito mantle plume". *Geology* vol. 41.9, pp. 1011–1014.

Jackson, I. and Faul, U. H. (2010). "Grainsize-sensitive viscoelastic relaxation in olivine: Towards a robust laboratory-based model for seismological application". *Physics of the Earth and Planetary Interiors* vol. 183.1-2, pp. 151–163.

Jalinoos, F., Tran, K. T., Nguyen, T. D., and Agrawal, A. K. (2017). "Evaluation of bridge abutments and bounded wall type structures with ultraseismic waveform tomography". *Journal of Bridge Engineering* vol. 22.12, p. 04017104.

Jeffreys, H and Bullen, K. (1940). "Seismological Tables, London: British Association for the Advancement of Science".

Jeffreys, H. (1926). "The rigidity of the Earth's central core". *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society* vol. 1.7, pp. 371–383.

Jenkins, J., Maclennan, J., Green, R. G., Cottaar, S., Deuss, A., and White, R. S. (2018). "Crustal formation on a spreading ridge above a mantle plume: receiver function imaging of the Icelandic crust". *Journal of Geophysical Research: Solid Earth* vol. 123.6, pp. 5190–5208.

Jónasson, K., Bessason, B., Helgadóttir, Á., Einarsson, P., Guðmundsson, G. B., Brandsdóttir, B., Vogfjörd, K. S., and Jónsdóttir, K. (2021). "A harmonised instrumental earthquake catalogue for Iceland and the northern Mid-Atlantic Ridge". *Natural Hazards and Earth System Sciences* vol. 21.7, pp. 2197–2214.

Katili, J. A. (1978). "Past and present geotectonic position of fi, Indonesia". *Tectonophysics* vol. 45.4, pp. 289–322.

Keep, M. and Haig, D. W. (2010). "Deformation and exhumation in Timor: Distinct stages of a young orogeny". *Tectonophysics* vol. 483.1-2, pp. 93–111.

Kennett, B. L. N. (June 2020). "Radial earth models revisited". *Geophysical Journal International* 222.3, pp. 2189–2204. DOI: `10.1093/gji/ggaa298`.

Kennett, B. and Engdahl, E. (1991). "Traveltimes for global earthquake location and phase identification". *Geophysical Journal International* vol. 105.2, pp. 429–465.

Kennett, B., Engdahl, E., and Buland, R (1995). "Constraints on seismic velocities in the Earth from traveltimes". *Geophysical Journal International* vol. 122.1, pp. 108–124.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., and Willing, C. (2016).

"Jupyter Notebooks – a publishing format for reproducible computational workflows". *Positioning and Power in Academic Publishing: Players, Agents and Agendas*.

Koelemeijer, P, Ritsema, J, Deuss, A, and van Heijst, H.-J. (2016). "SP12RTS: a degree-12 model of shear-and compressional-wave velocity for Earth's mantle". *Geophysical Journal International* vol. 204.2, pp. 1024–1039.

Komatitsch, D. and Tromp, J. (1999). "Introduction to the spectral element method for three-dimensional seismic wave propagation". *Geophysical Journal International* vol. 139.3, pp. 806–822.

— (2002). "Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation". *Geophysical Journal International* vol. 150.1, pp. 303–318.

Komatitsch, D. and 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* vol. 88.2, pp. 368–392.

Konstantinou, K., Utami, I. W., Giannopoulos, D, and Sokos, E (2020). "A Reappraisal of Seismicity Recorded During the 1996 Gjálp Eruption, Iceland, in Light of the 2014–2015 Bárðarbunga–Holuhraun Lateral Dike Intrusion". *Pure and Applied Geophysics* vol. 177.6, pp. 2579–2595.

Kopp, H., Flueh, E. R., Petersen, C. J., Weinrebe, W, Wittwer, A., and Scientists, M. (2006). "The Java margin revisited: Evidence for subduction erosion off Java". *Earth and Planetary Science Letters* vol. 242.1-2, pp. 130–142.

Koroni, M., Borgeaud, A., Fichtner, A., and Deschamps, F. (2021). "Analysis of core-mantle boundary seismic waves using full-waveform modelling and adjoint methods". *arXiv preprint arXiv:2110.11068*.

Kosloff, R and Kosloff, D (1986). "Absorbing boundaries for wave propagation problems". *Journal of Computational Physics* vol. 63.2, pp. 363–376.

Krischer, L. and Casarotti, E. (Sept. 2015). *pyflex: 0.1.4*. Version 0.1.4. DOI: 10.5281/zenodo.31607. URL: https://doi.org/10.5281/zenodo.31607.

Krischer, L., Fichtner, A., Böhm, C., and Igel, H. (2018). "Automated large-scale full seismic waveform inversion for North America and the North Atlantic". *Journal of Geophysical Research: Solid Earth* vol. 123.7, pp. 5902–5928.

Krischer, L., Fichtner, A., Zukauskaite, S., and Igel, H. (2015). "Large-scale seismic inversion framework". *Seismological Research Letters* vol. 86.4, pp. 1198–1207.

Krischer, L., Smith, J., Lei, W., Lefebvre, M., Ruan, Y., Andrade, E. S. de, Podhorszki, N., Bozdağ, E., and Tromp, J. (2016). "An Adaptable Seismic Data Format". *Geophysical Journal International* 207.2, pp. 1003–1011.

Kristeková, M., Kristek, J., and Moczo, P. (2009). "Time-frequency misfit and goodness-of-fit criteria for quantitative comparison of time signals". *Geophysical Journal International* vol. 178.2, pp. 813–825.

Kristeková, M., Kristek, J., Moczo, P., and Day, S. M. (2006). "Misfit criteria for quantitative comparison of seismograms". *Bulletin of the Seismological Society of America* vol. 96.5, pp. 1836–1850.

Kukreja, N., Hückelheim, J., Louboutin, M., Washbourne, J., Kelly, P. H., and Gorman, G. J. (2020). "Lossy checkpoint compression in full waveform inversion". *Geoscientific Model Development Discussions*, pp. 1–26.

Lagmay, A., Tejada, L. G., Pena, R. E., Aurelio, M, Davy, B., David, S, and Billedo, E (2009). "New definition of Philippine Plate boundaries and implications to the Philippine Mobile Belt". *Journal of the Geological Society of the Philippines* vol. 64, pp. 17–30.

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

Laske, G., Masters, G., Ma, Z., and Pasyanos, M. (2013). "Update on CRUST1. 0—A 1-degree global model of Earth's crust". *EGU General Assembly Conference Abstracts* vol. 15, p. 2658.

Lebedev, S. and Nolet, G. (2003). "Upper mantle beneath Southeast Asia from S velocity tomography". *Journal of Geophysical Research: Solid Earth* vol. 108.B1.

Lee, S.-J., Chen, H.-W., Liu, Q., Komatitsch, D., Huang, B.-S., and Tromp, J. (2008). "Three-dimensional simulations of seismic-wave propagation in the Taipei basin with realistic topography based upon the spectral-element method". *Bulletin of the Seismological Society of America* vol. 98.1, pp. 253–264.

Lehmann, I. (1936). "Bur. Cent. Séismologique Int. Sér". *Séismol. Internat. Strasbourg Public. Bur. Centr. Sci.* vol. 14, pp. 87–115.

Lei, W., Ruan, Y., Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., and Pugmire, D. (2020). "Global adjoint tomography—model GLAD-M25". *Geophysical Journal International* vol. 223.1, pp. 1–21.

Leng, K., Nissen-Meyer, T., and Driel, M. van (2016). "Efficient global wave propagation adapted to 3-D structural complexity: a pseudospectral/spectral-element approach". *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society* vol. 207.3, pp. 1700–1721.

Leng, K., Nissen-Meyer, T., Van Driel, M., Hosseini, K., and Al-Attar, D. (2019). "AxiSEM3D: broad-band seismic wavefields in 3-D global earth models with undulating discontinuities". *Geophysical Journal International* vol. 217.3, pp. 2125–2146.

Leonard, M. (2014). "Self-consistent earthquake fault-scaling relations: Update and extension to stable continental strike-slip faults". *Bulletin of the Seismological Society of America* vol. 104.6, pp. 2953–2965.

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

Li, X.-D. and Romanowicz, B. (1996). "Global mantle shear velocity model developed using nonlinear asymptotic coupling theory". *Journal of Geophysical Research: Solid Earth* vol. 101.B10, pp. 22245–22272.

Li, X.-D. and Tanimoto, T. (1993). "Waveforms of long-period body waves in a slightly aspherical Earth model". *Geophysical Journal International* 112.1, pp. 92–102.

Li, X., Hao, T., and Li, Z. (2018). "Upper mantle structure and geodynamics of the Sumatra subduction zone from 3-D teleseismic P-wave tomography". *Journal of Asian Earth Sciences* vol. 161, pp. 25–34.

Liu, D. C. and Nocedal, J. (1989). "On the limited memory BFGS method for large scale optimization". *Mathematical Programming* vol. 45.1, pp. 503–528.

Liu, Q., Beller, S., Lei, W., Peter, D., and Tromp, J. (2022). "Pre-conditioned BFGS-based uncertainty quantification in elastic full-waveform inversion". *Geophysical Journal International* vol. 228.2, pp. 796–815.

Liu, Q., Polet, J., Komatitsch, D., and Tromp, J. (2004). "Spectral-element moment tensor inversions for earthquakes in southern California". *Bulletin of the Seismological Society of America* vol. 94.5, pp. 1748–1761.

Liu, Q. and Tromp, J. (2006). "Finite-frequency kernels based on adjoint methods". *Bulletin of the Seismological Society of America* vol. 96.6, pp. 2383–2397.

Liu, X. and Zhao, D. (2015). "Seismic attenuation tomography of the Southwest Japan arc: new insight into subduction dynamics". *Geophysical Journal International* vol. 201.1, pp. 135–156.

Luo, Y. and Schuster, G. T. (1991). "Wave-equation traveltime inversion". *Geophysics* vol. 56.5, pp. 645–653.

Maggi, A., Tape, C., Chen, M., Chao, D., and Tromp, J. (2009). "An automated time-window selection algorithm for seismic tomography". *Geophysical Journal International* vol. 178.1, pp. 257–281.

Mallet, R. (1862). *Great Neapolitan Earthquake of 1857: The First Principles of Observational Seismology as Developed in the Report to the Royal Society of London of the Expedition Made by Command of the Society Into the Interior of the Kingdom of Naples, to Investigate the Circumstances of the Great Earthquake of Demember 1857*. Vol. 2.

Marano, K. D., Wald, D. J., and Allen, T. I. (2010). "Global earthquake casualties due to secondary effects: a quantitative analysis for improving rapid loss analyses". *Natural Hazards* vol. 52.2, pp. 319–328.

Marquering, H., Dahlen, F., and Nolet, G. (1999). "Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox". *Geophysical Journal International* vol. 137.3, pp. 805–815.

Masson, Y. and Romanowicz, B. (2017). "Box tomography: localized imaging of remote targets buried in an unknown medium, a step forward for understanding key structures in the deep Earth". *Geophysical Journal International* vol. 211.1, pp. 141–163.

McCaffrey, R. (May 2009). "The Tectonic Framework of the Sumatran Subduction Zone". *Earth Planet. Sci. Annu. Rev. Earth Planet. Sci* 3737, pp. 345–66. DOI: 10.1146/annurev.earth.031208.100212.

Meier, U, Curtis, A, and Trampert, J (2007). "Fully nonlinear inversion of fundamental mode surface waves for a global crustal model". *Geophysical Research Letters* 34.16.

Metaxian, J.-P., Surono, Widiyantoro, S., and RESIF (2018). *Seismic network YR:DOMERAPI temporary experiment (RESIF-SISMOB*. en. DOI: 10.15778/RESIF.YR2013.

Metcalfe, I (1990). "Allochthonous terrane processes in Southeast Asia". *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* vol. 331.1620, pp. 625–640.

Métrich, N., Vidal, C. M., Komorowski, J.-C., Pratomo, I., Michel, A., Kartadinata, N., Prambada, O., Rachmat, H., and 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* vol. 58.11, pp. 2257–2284.

Michell, J (1761). "Conjectures concerning the cause and observations upon the phenomena of earthquakes". *Philosophical Transactions of the Royal Society of London* vol. 51, p. 2.

Miller, M. S. (2014). *Transitions in the Banda Arc-Australia continental collision*. DOI: 10.7914/SN/YS_2014. URL: http://www.fdsn.org/doi/10.7914/SN/YS_2014.

Miller, M. S., O'Driscoll, L. J., Roosmawati, N., Harris, C. W., Porritt, R. W., Widiyantoro, S., Costa, L. Teofilo da, Soares, E., Becker, T. W., and Joshua West, A (2016). "Banda arc experiment—Transitions in the Banda arc-Australian continental collision". *Seismological Research Letters* vol. 87.6, pp. 1417–1423.

Mimoun, D., Murdoch, N., Lognonné, P., Hurst, K., Pike, W. T., Hurley, J., Nébut, T., and Banerdt, W. B. (2017). "The noise model of the SEIS seismometer of the InSight mission to Mars". *Space Science Reviews* 211.1, pp. 383–428.

Modrak, R. and Tromp, J. (2016). "Seismic waveform inversion best practices: regional, global and exploration test cases". *Geophysical Journal International* vol. 206.3, pp. 1864–1889.

Mohorovicic, A. (1909). *Potres od 8*.

Möller, T. and Friederich, W. (2021). "A hybrid method to calculate teleseismic body waves in a regional 3D model using GEMINI and SPECFEM." *EGU General Assembly Conference Abstracts*.

Mondaic: Equivalence of Filtering Source Time Functions and Filtering Synthetics. https://mondaic.com/docs/0.11.40/examples/tutorials/sources_and_receivers/equivalence_filtering_stfs_and_filtering_synthetics/tutorial. Accessed: 2021-11-06.

Mondaic: Misfit functionals involving velocity. https://mondaic.com/docs/0.11.44/knowledge_base/velocity_misfits. Accessed: 2021-02-03.

Montagner, J.-P. and Kennett, B. (1996). "How to reconcile body-wave and normal-mode reference Earth models". *Geophysical Journal International* vol. 125.1, pp. 229–248.

Montagner, J.-P. and Tanimoto, T. (1991). "Global upper mantle tomography of seismic velocities and anisotropies". *Journal of Geophysical Research: Solid Earth* vol. 96.B12, pp. 20337–20351.

Monteiller, V., Chevrot, S., Komatitsch, D., and Wang, Y. (2015). "Three-dimensional full waveform inversion of short-period teleseismic wavefields based upon the SEM–DSM hybrid method". *Geophysical Journal International* vol. 202.2, pp. 811–827.

Montelli, R., Nolet, G., Dahlen, F., Masters, G., Engdahl, E. R., and Hung, S.-H. (2004). "Finite-frequency tomography reveals a variety of plumes in the mantle". *Science* vol. 303.5656, pp. 338–343.

Moore, E. H. (1920). "On the reciprocal of the general algebraic matrix". *Bulletin of the American Mathematical Society.* vol. 26, pp. 394–395.

Moresi, L., Betts, P. G., Miller, M. S., and Cayley, R. A. (2014). "Dynamics of continental accretion". *Nature* vol. 508.7495, pp. 245–248.

Morgan, W. J. (1971). "Convection plumes in the lower mantle". *Nature* vol. 230.5288, pp. 42–43.

Morse, P. M. and Feshbach, H. (1954). "Methods of theoretical physics". *American Journal of Physics* vol. 22.6, pp. 410–413.

Mosegaard, K and Sambridge, M (2002). "Monte Carlo methods in geophysical inverse problems". *Reviews of Geophysics*.

Mosegaard, K. and Tarantola, A. (1995). "Monte Carlo sampling of solutions to inverse problems". *Journal of Geophysical Research: Solid Earth* vol. 100.B7, pp. 12431–12447.

Münch, U., Rudloff, A., and Lauterjung, J. (2011). "Postface" The GITEWS Project–results, summary and outlook"". *Natural Hazards and Earth System Sciences* vol. 11.3, pp. 765–769.

Muttarak, R. and Pothisiri, W. (2013). "The role of education on disaster preparedness: case study of 2012 Indian Ocean earthquakes on Thailand's Andaman Coast". *Ecology and Society* vol. 18.4.

Navon, I. (1998). "Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography". *Dynamics of Atmospheres and Oceans* vol. 27.1-4, pp. 55–79.

Nettles, M. and Ekström, G. (1998). "Faulting mechanism of anomalous earthquakes near Bárdarbunga Volcano, Iceland". *Journal of Geophysical Research: Solid Earth* vol. 103.B8, pp. 17973–17983.

Nissen-Meyer, T., Driel, M. van, Stähler, S. C., Hosseini, K., Hempel, S., Auer, L., Colombi, A., and Fournier, A. (2014). "AxiSEM: broadband 3-D seismic wavefields in axisymmetric media". *Solid Earth* vol. 5.1, pp. 425–445.

Nocedal, J. (1980). "Updating quasi-Newton matrices with limited storage". *Mathematics of computation* vol. 35.151, pp. 773–782.

Nocedal, J. and Wright, S. (2006). *Numerical optimization*.

Noe, S., van Herwaarden, D., Thrastarson, S., Masouminia, N., van Driel, M., Böhm, C., Ma, J., Bunge, H.-P., Gao, Y., Tilmann, F, Wehner, D., and Fichtner., A (in prep.). "Collaborative Seismic Earth Model: Generation 2". *Geophysical Research Letters*.

Nolet, G. (1996). *Seismic study of the Iceland hotspot*. DOI: 10.7914/SN/XD_1996. URL: https://www.fdsn.org/networks/detail/XD_1996/.

— (2008). "A breviary of seismic tomography".

Nolet, G., Montelli, R., and Virieux, J. (1999). "Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systems". *Geophysical Journal International* vol. 138.1, pp. 36–44.

Nolet, G. and Zielhuis, A. (1994). "Low S velocities under the Tornquist-Teisseyre zone: Evidence for water injection into the transition zone by subduction". *Journal of Geophysical Research: Solid Earth* vol. 99.B8, pp. 15813–15820.

Nuber, A., Manukyan, E., and Maurer, H. (2016). "Ground topography effects on near-surface elastic full waveform inversion". *Geophysical Journal International* vol. 207.1, pp. 67–71.

Obayashi, M., Yoshimitsu, J., Nolet, G., Fukao, Y., Shiobara, H., Sugioka, H., Miyamachi, H., and Gao, Y. (2013). "Finite frequency whole mantle P wave tomography: Improvement of subducted slab images". *Geophysical Research Letters* vol. 40.21, pp. 5652–5657.

Obermann, A., Lupi, M., Mordret, A., Jakobsdóttir, S. S., and Miller, S. A. (2016). "3D-ambient noise Rayleigh wave tomography of Snæfellsjökull volcano, Iceland". *Journal of Volcanology and Geothermal Research* vol. 317, pp. 42–52.

Oldham, R. D. (1899). *Report of the great earthquake of 12th June, 1897*. Office of the Geological Survey.

— (1906). "The constitution of the interior of the Earth, as revealed by earthquakes". *Quarterly Journal of the Geological Society* vol. 62.1-4, pp. 456–475.

Panning, M. and Romanowicz, B. (2006). "A three-dimensional radially anisotropic model of shear velocity in the whole mantle". *Geophysical Journal International* vol. 167.1, pp. 361–379.

Panning, M. P., Capdeville, Y., and Romanowicz, B. A. (2009). "Seismic waveform modelling in a 3-D Earth using the Born approximation: potential shortcomings and a remedy". *Geophysical Journal International* vol. 177.1, pp. 161–178.

Passarelli, L, Heryandoko, N., Zimmer, M., Cesca, S., Rivalta, E, Rohadi, S, Merdianto, U, Dahm, T, Milkereit, C, et al. (2016). "Jailolo network, West Halmahera-Indonesia 2016/2018".

Patera, A. T. (1984). "A spectral element method for fluid dynamics: laminar flow in a channel expansion". *Journal of Computational Physics* vol. 54.3, pp. 468–488.

Pekeris, C. L. (1948). "Theory of propagation of explosive sound in shallow water".

Penrose, R. (1955). "A generalized inverse for matrices". vol. 51.3, pp. 406–413.

Peter, D., Komatitsch, D., Luo, Y., Martin, R., Le Goff, N., Casarotti, E., Le Loher, P., Magnoni, F., Liu, Q., Blitz, C., et al. (2011). "Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes". *Geophysical Journal International* vol. 186.2, pp. 721–739.

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

Pilia, S., Rawlinson, N, Hall, R, Cornwell, D., Gilligan, A, and Tongkul, F (submitted). "Seismic signature of subduction termination from teleseismic P-and S-wave arrival-time tomography: the case of northern Borneo 2".

Pilia, S., Rawlinson, N., Gilligan, A., and Tongkul, F. (2019). "Deciphering the Fate of Plunging Tectonic Plates in Borneo". *Eos, Transactions American Geophysical Union* vol. 100.10, pp. 18–23.

Plessix, R.-E., Baeten, G., Maag, J. W. de, Klaassen, M., Rujie, Z., and Zhifei, T. (2010). "Application of acoustic full waveform inversion to a low-frequency large-offset land data set". *SEG Technical Program Expanded Abstracts 2010*. Society of Exploration Geophysicists, pp. 930–934.

Pratt, R. G., Shin, C., and Hick, G. (1998). "Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion". *Geophysical Journal International* vol. 133.2, pp. 341–362.

Pratt, R. G. and Worthington, M. H. (1990). "Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave equation method 1". *Geophysical Prospecting* vol. 38.3, pp. 287–310.

Priestley, K., Ho, T., and McKenzie, D. (2021). "The formation of continental roots". *Geology* vol. 49.2, pp. 190–194.

Priestley, K. and McKenzie, D. (2006). "The thermal structure of the lithosphere from shear wave velocities". *Earth and Planetary Science Letters* vol. 244.1-2, pp. 285–301.

Priestley, K., McKenzie, D., and Ho, T. (2018). "A Lithosphere–Asthenosphere Boundary—a Global Model Derived from Multimode Surface-Wave Tomography and Petrology". *Lithospheric Discontinuities*, pp. 111–123.

Priolo, E., Carcione, J. M., and Seriani, G. (1994). "Numerical simulation of interface waves by high-order spectral modeling techniques". *The Journal of the Acoustical Society of America* vol. 95.2, pp. 681–693.

Quinteros, J., Strollo, A., Evans, P. L., Hanka, W., Heinloo, A., Hemmleb, S., Hillmann, L., Jaeckel, K.-H., Kind, R., Saul, J., et al. (2021). "The GEOFON program in 2020". *Seismological Society of America* vol. 92.3, pp. 1610–1622.

Rangin, C. (1991). "The Philippine Mobile Belt: a complex plate boundary". *Journal of Southeast Asian Earth Sciences* vol. 6.3-4, pp. 209–220.

Rawlinson, N and Spakman, W (2016). "On the use of sensitivity tests in seismic tomography". *Geophysical Journal International* vol. 205.2, pp. 1221–1243.

Rawlinson, N., Fichtner, A., Sambridge, M., and Young, M. K. (2014). "Seismic tomography and the assessment of uncertainty". *Advances in Geophysics* vol. 55, pp. 1–76.

Rawlinson, N., Hao, T., Widiyantoro, S., and Lü, C. (2020). *Sulawesi and Celebes Sea Ocean-bottom Seismic Network*. DOI: 10.7914/SN/5R_2019. URL: https://www.fdsn.org/networks/detail/5R_2019/.

Rawlinson, N. and Kennett, B. (2008). "Teleseismic tomography of the upper mantle beneath the southern Lachlan Orogen, Australia". *Physics of the Earth and Planetary Interiors* vol. 167.1-2, pp. 84–97.

Rawlinson, N., Pozgay, S, and Fishwick, S (2010). "Seismic tomography: a window into deep Earth". *Physics of the Earth and Planetary Interiors* vol. 178.3-4, pp. 101–135.

Rawlinson, N. and Sambridge, M. (2003). "Seismic traveltime tomography of the crust and lithosphere". *Advances in Geophysics* vol. 46, pp. 81–199.

Rawlinson, N. (2018). *North Borneo Orogeny Seismic Survey*. DOI: 10.7914/SN/YC_2018. URL: http://www.fdsn.org/doi/10.7914/SN/YC_2018.

Rebeur-Paschwitz, E. von (1889). "The earthquake of Tokio, April 18, 1889". *Nature* vol. 40.1030, pp. 294–295.

Reinarz, A., Charrier, D. E., Bader, M., Bovard, L., Dumbser, M., Duru, K., Fambri, F., Gabriel, A.-A., Gallard, J.-M., Köppel, S., et al. (2020). "ExaHyPE: an engine for parallel dynamically adaptive simulations of wave problems". *Computer Physics Communications* vol. 254, p. 107251.

Richards, S., Lister, G., and Kennett, B. (2007). "A slab in depth: Three-dimensional geometry and evolution of the Indo-Australian plate". *Geochemistry, Geophysics, Geosystems* vol. 8.12.

Rickers, F., Fichtner, A., and Trampert, J. (2012). "Imaging mantle plumes with instantaneous phase measurements of diffracted waves". *Geophysical Journal International* vol. 190.1, pp. 650–664.

Rickers, F., Fichtner, A., and Trampert, J. (2013). "The Iceland–Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: evidence from full-waveform inversion". *Earth and Planetary Science Letters* vol. 367, pp. 39–51.

Ritsema, J., Deuss, A., van Heijst, H., and Woodhouse, J. (2011). "S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements". *Geophysical Journal International* vol. 184.3, pp. 1223–1236.

Ritsema, J., van Heijst, H. J., and Woodhouse, J. H. (1999). "Complex shear wave velocity structure imaged beneath Africa and Iceland". *Science* vol. 286.5446, pp. 1925–1928.

Ritsema, J., Xu, W., Stixrude, L., and Lithgow-Bertelloni, C. (2009). "Estimates of the transition zone temperature in a mechanically mixed upper mantle". *Earth and Planetary Science Letters* vol. 277.1-2, pp. 244–252.

Roberts, G. G., White, N., Hoggard, M. J., Ball, P. W., and Meenan, C. (2018). "A Neogene history of mantle convective support beneath Borneo". *Earth and Planetary Science Letters* vol. 496, pp. 142–158.

Rodgers, A., Krischer, L., Afanasiev, M., Böhm, C., Doody, C., Chiang, A., and Simmons, N. (2022). "WUS256: An Adjoint Waveform Tomography Model of the Crust and Upper Mantle of the Western United States for Improved Waveform Simulations". *Journal of Geophysical Research: Solid Earth* vol. 127.7, e2022JB024549. DOI: https://doi.org/10.1029/2022JB024549.

Roecker, S, Baker, B, and McLaughlin, J (2010). "A finite-difference algorithm for full waveform teleseismic tomography". *Geophysical Journal International* vol. 181.2, pp. 1017–1040.

Rögnvaldsson, S. T. and Slunga, R. (1993). "Routine fault plane solutions for local networks: a test with synthetic data". *Bulletin of the Seismological Society of America* vol. 83.4, pp. 1232–1247.

Romanowicz, B., Chen, L.-W., and French, S. W. (2020). "Accelerating full waveform inversion via source stacking and cross-correlations". *Geophysical Journal International* vol. 220.1, pp. 308–322.

Rome, S. (2016). *Why blurring an image is similar to warming your coffee*. Accessed: 2022-01-15. URL: https://srome.github.io/Why-Blurring-an-Image-is-Similar-to-Warming-Your-Coffee/.

Rosenbrock, H. (1960). "An automatic method for finding the greatest or least value of a function". *The Computer Journal* vol. 3.3, pp. 175–184.

Ruan, Y., Lei, W., Modrak, R., Örsvuran, R., Bozdağ, E., and Tromp, J. (2019). "Balancing unevenly distributed data in seismic tomography: a global adjoint tomography example". *Geophysical Journal International* vol. 219.2, pp. 1225–1236.

Rudloff, A., Lauterjung, J., Münch, U., Tinti, S, et al. (2009). "Preface'The GITEWS Project (German-Indonesian Tsunami Early Warning System)'". *Natural Hazards and Earth System Sciences (NHESS)* vol. 9.4, pp. 1381–1382.

Ryberg, T and Haberland, C (2008). "Lake Toba seismic network, Sumatra, Indonesia".

Saikia, C. K. (1994). "Modified frequency-wavenumber algorithm for regional seismograms using Filon's quadrature: modelling of Lg waves in eastern North America". *Geophysical Journal International* vol. 118.1, pp. 142–158.

Saltus, R. W. and Blakely, R. J. (2011). "Unique geologic insights from "non-unique" gravity and magnetic interpretation". *GSA Today* vol. 21.12, pp. 4–11.

Sambridge, M. (1999). "Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space". *Geophysical journal international* vol. 138.2, pp. 479–494.

Santosa, F. and Symes, W. W. (1988). "Computation of the Hessian for least-squares solutions of inverse problems of reflection seismology". *Inverse Problems* vol. 4.1, p. 211.

Schaeffer, A. and Lebedev, S (2013). "Global shear speed structure of the upper mantle and transition zone". *Geophysical Journal International* vol. 194.1, pp. 417–449.

Schreiman, J., Gisvold, J., Greenleaf, J. F., and Bahn, R. (1984). "Ultrasound transmission computed tomography of the breast." *Radiology* vol. 150.2, pp. 523–530.

Scripps Institution of Oceanography (1986). *Global Seismograph Network - IRIS/IDA*. DOI: 10.7914/SN/II. URL: https://www.fdsn.org/networks/detail/II/.

Seton, M., Müller, R. D., Zahirovic, S, Gaina, C, Torsvik, T, Shephard, G, Talsma, A, Gurnis, M, Turner, M, Maus, S, et al. (2012). "Global continental and ocean basin reconstructions since 200 Ma". *Earth-Science Reviews* vol. 113.3-4, pp. 212–270.

Shanno, D. F. (1970). "Conditioning of quasi-Newton methods for function minimization". *Mathematics of computation* vol. 24.111, pp. 647–656.

Sieminski, A., Lévêque, J.-J., and Debayle, E. (2004). "Can finite-frequency effects be accounted for in ray theory surface wave tomography?" *Geophysical Research Letters* vol. 31.24.

Sieminski, A., Trampert, J., and Tromp, J. (2009). "Principal component analysis of anisotropic finite-frequency sensitivity kernels". *Geophysical Journal International* vol. 179.2, pp. 1186–1198.

Sigloch, K., McQuarrie, N., and Nolet, G. (2008). "Two-stage subduction history under North America inferred from multiple-frequency tomography". *Nature Geoscience* vol. 1.7, pp. 458–462.

Simons, M., Minson, S. E., Sladen, A., Ortega, F., Jiang, J., Owen, S. E., Meng, L., Ampuero, J.-P., Wei, S., Chu, R., et al. (2011). "The 2011 magnitude 9.0 Tohoku-Oki earthquake: Mosaicking the megathrust from seconds to centuries". *Science* vol. 332.6036, pp. 1421–1425.

Simons, W., Socquet, A, Vigny, C, Ambrosius, B., Haji Abu, S, Promthong, C., Subarya, C, Sarsito, D., Matheussen, S, Morgan, P, et al. (2007). "A decade of GPS in Southeast Asia: Resolving Sundaland motion and boundaries". *Journal of Geophysical Research: Solid Earth* vol. 112.B6.

Simutė, S., Steptoe, H., Cobden, L., Gokhberg, A., and Fichtner, A. (2016). "Full-waveform inversion of the Japanese Islands region". *Journal of Geophysical Research: Solid Earth* vol. 121.5, pp. 3722–3741.

Sirgue, L., Barkved, O., Dellinger, J, Etgen, J, Albertin, U, and Kommedal, J. (2010). "Thematic set: Full waveform inversion: The next leap forward in imaging at Valhall". *First Break* vol. 28.4.

Snieder, R. and Trampert, J. (1999). "Inverse problems in geophysics". *Wavefield inversion*. Springer, pp. 119–190.

Socquet, A., Hollingsworth, J., Pathier, E., and Bouchon, M. (2019). "Evidence of supershear during the 2018 magnitude 7.5 Palu earthquake from space geodesy". *Nature Geoscience* vol. 12.3, pp. 192–199.

Soesoo, A., Bons, P. D., Gray, D. R., and Foster, D. A. (1997). "Divergent double subduction: tectonic and petrologic consequences". *Geology* vol. 25.8, pp. 755–758.

Song, T., Hao, T., Zhang, J., Cao, L., and Dong, M. (2022). "Numerical modeling of North Sulawesi subduction zone: Implications for the east–west differential evolution". *Tectonophysics* vol. 822, p. 229172.

Song, T.-R. A. and Kawakatsu, H. (2012). "Subduction of oceanic asthenosphere: Evidence from sub-slab seismic anisotropy". *Geophysical Research Letters* vol. 39.17.

Southern, E. O., Winder, T., White, R. S., and Brandsdóttir, B. (2022). "Ring Fault Slip Reversal at Bárðarbunga Volcano, Iceland: Seismicity during Caldera Collapse and Re-Inflation 2014-2018".

Spakman, W and Nolet, G (1988). "Imaging algorithms, accuracy and resolution in delay time tomography". *Mathematical Geophysics*. Springer, pp. 155–187.

Spakman, W. and Hall, R. (2010). "Surface deformation and slab–mantle interaction during Banda arc subduction rollback". *Nature Geoscience* vol. 3.8, pp. 562–566.

Stähler, S. C. and Sigloch, K. (2014). "Fully probabilistic seismic source inversion–Part 1: Efficient parameterisation". *Solid Earth* vol. 5.2.

Staples, R. K., White, R. S., Brandsdóttir, B., Menke, W., Maguire, P. K., and McBride, J. H. (1997). "Färoe-Iceland ridge experiment 1. Crustal structure of northeastern Iceland". *Journal of Geophysical Research: Solid Earth* vol. 102.B4, pp. 7849–7866.

Steinberger, B., Zhao, D., and Werner, S. C. (2015). "Interior structure of the Moon: constraints from seismic tomography, gravity and topography". *Physics of the Earth and Planetary Interiors* vol. 245, pp. 26–39.

Sturgeon, W., Ferreira, A. M., Faccenda, M., Chang, S.-J., and Schardong, L. (2019). "On the origin of radial anisotropy near subducted slabs in the midmantle". *Geochemistry, Geophysics, Geosystems* vol. 20.11, pp. 5105–5125.

Su, P.-L., Chen, P.-F., and Wang, C.-Y. (2019). "High-Resolution 3-DP Wave Velocity Structures Under NE Taiwan and Their Tectonic Implications". *Journal of Geophysical Research: Solid Earth* vol. 124.11, pp. 11601–11614.

Tao, K., Grand, S. P., and Niu, F. (2018). "Seismic structure of the upper mantle beneath eastern Asia from full waveform seismic tomography". *Geochemistry, Geophysics, Geosystems* vol. 19.8, pp. 2732–2763.

Tape, C., Liu, Q., Maggi, A., and Tromp, J. (2010). "Seismic tomography of the southern California crust based on spectral-element and adjoint methods". *Geophysical Journal International* vol. 180.1, pp. 433–462.

Tape, C., Liu, Q., and Tromp, J. (2007). "Finite-frequency tomography using adjoint methods—Methodology and examples using membrane surface waves". *Geophysical Journal International* vol. 168.3, pp. 1105–1129.

Tarantola, A. (1984). "Inversion of seismic reflection data in the acoustic approximation". *Geophysics* vol. 49.8, pp. 1259–1266.

Thordarson, T. and Larsen, G. (2007). "Volcanism in Iceland in historical time: Volcano types, eruption styles and eruptive history". *Journal of Geodynamics* vol. 43.1, pp. 118–152.

Thrastarson, S., van Driel, M., Krischer, L., Böhm, C., Afanasiev, M., van Herwaarden, D.-P., and Fichtner, A. (2020). "Accelerating numerical wave propagation by wavefield adapted meshes. Part II: full-waveform inversion". *Geophysical Journal International* vol. 221.3, pp. 1591–1604.

Thrastarson, S., van Herwaarden, D.-P., and Fichtner, A. (Feb. 2021). *solvithrastar/MultiMesh: MultiMesh - Python-based interpolations between discretizations*. Version v0.0.2-alpha. DOI: 10.5281/zenodo.4564523.

Thrastarson, S., van Herwaarden, D.-P., Krischer, L., Böhm, C., Driel, M. van, Afanasiev, M., and Fichtner, A. (2022). "Data-adaptive global full-waveform inversion". *Geophysical Journal International*.

Tikhonov, A. N. (1943). "On the stability of inverse problems". vol. 39, pp. 195–198.

Todoriki, M., Furumura, T., and Maeda, T. (2016). "Effects of seawater on elongated duration of ground motion as well as variation in its amplitude for offshore earthquakes". *Geophysical Journal International*, p. 388.

Trabant, C., Hutko, A. R., Bahavar, M., Karstens, R., Ahern, T., and Aster, R. (2012). "Data products at the IRIS DMC: Stepping stones for research and other applications". *Seismological Research Letters* vol. 83.5, pp. 846–854.

Tromp, J. (2020). "Seismic wavefield imaging of earth's interior across scales". *Nature Reviews Earth & Environment* vol. 1.1, pp. 40–53.

Tromp, J. and Bachmann, E. (2019). "Source encoding for adjoint tomography". *Geophysical Journal International* vol. 218.3, pp. 2019–2044.

Tromp, J., Tape, C., and Liu, Q. (2005). "Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels". *Geophysical Journal International* vol. 160.1, pp. 195–216.

Tryggvason, A., Rögnvaldsson, S. T., and Flóvenz, O. G. (2002). "Three-dimensional imaging of the P-and S-wave velocity structure and earthquake locations beneath Southwest Iceland". *Geophysical Journal International* vol. 151.3, pp. 848–866.

Turner, S., Foden, J., George, R., Evans, P., Varne, R., Elburg, M., and Jenner, G. (2003). "Rates and processes of potassic magma evolution beneath Sangeang Api volcano, East Sunda arc, Indonesia". *Journal of Petrology* vol. 44.3, pp. 491–515.

Uieda, L., Tian, D., Leong, W. J., Toney, L., Schlitzer, W., Yao, J., Grund, M., Jones, M., Materna, K., Newton, T., Ziebarth, M., and Wessel, P. (Mar. 2021). *PyGMT: A Python interface for the Generic Mapping Tools*. Version v0.3.1. DOI: 10.5281/zenodo.4592991.

Valentine, A. P. and Woodhouse, J. H. (2010). "Reducing errors in seismic tomography: combined inversion for sources and structure". *Geophysical Journal International* vol. 180.2, pp. 847–857.

Vallée, M. (2013). "Source time function properties indicate a strain drop independent of earthquake depth and magnitude". *Nature communications* vol. 4.1, pp. 1–6.

Vallée, M., Charléty, J., Ferreira, A. M., Delouis, B., and Vergoz, J. (2011). "SCARDEC: a new technique for the rapid determination of seismic moment magnitude, focal mechanism and source time functions for large earthquakes using body-wave deconvolution". *Geophysical Journal International* vol. 184.1, pp. 338–358.

Venzke, E (2013). "Global Volcanism Program. Volcanoes of the World, v. 4.10.6". *Smithsonian Institution*.

Virieux, J. (1984). "SH-wave propagation in heterogeneous media: Velocity-stress finite-difference method". *Geophysics* vol. 49.11, pp. 1933–1942.

Virieux, J. and Operto, S. (2009). "An overview of full-waveform inversion in exploration geophysics". *Geophysics* vol. 74.6, WCC1–WCC26.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". *Nature Methods* 17, pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

Volk, O. (2021). "Velocity and anisotropy structure of the Icelandic crust-An ambient seismic noise analysis". PhD thesis. University of Cambridge.

Volk, O., White, R. S., Pilia, S., Green, R. G., Maclennan, J., and Rawlinson, N. (2021). "Oceanic crustal flow in Iceland observed using seismic anisotropy". *Nature Geoscience* vol. 14.3, pp. 168–173.

Walck, M. C. and Clayton, R. W. (1987). "P wave velocity variations in the Coso region, California, derived from local earthquake travel times". *Journal of Geophysical Research: Solid Earth* vol. 92.B1, pp. 393–405.

Walter, T. R., Haghshenas Haghighi, M., Schneider, F. M., Coppola, D., Motagh, M., Saul, J., Babeyko, A., Dahm, T., Troll, V. R., Tilmann, F., et al. (2019). "Complex hazard cascade culminating in the Anak Krakatau sector collapse". *Nature communications* vol. 10.1, pp. 1–11.

Wang and 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* vol. 21.4, e2019GC008658.

Wang, Y., Chevrot, S., Monteiller, V., Komatitsch, D., Mouthereau, F., Manatschal, G., Sylvander, M., Diaz, J., Ruiz, M., Grimaud, F., et al. (2016). "The deep roots of the western Pyrenees revealed by full waveform inversion of teleseismic P waves". *Geology* vol. 44.6, pp. 475–478.

Wang, Z., Singh, S. C., and Noble, M. (2020). "True-amplitude versus trace-normalized full waveform inversion". *Geophysical Journal International* vol. 220.2, pp. 1421–1435.

Wehner, D., Blom, N., Rawlinson, N., Böhm, C., Miller, M. S., Supendi, P., and Widiyantoro, S. (2022). "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* vol. 127.3, e2021JB022930.

Wehner, D., Blom, N., Rawlinson, N., Daryono, Böhm, C., Miller, M. S., Supendi, P., and Widiyantoro, S. (Oct. 2021). "Supplementary material "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*. DOI: 10.5281/zenodo.5573139. URL: https://doi.org/10.5281/zenodo.5573139.

Wehner, D., Rawlinson, N., Greenfield, T., Miller, M., Supendi, P., Lü, C., and Widiyantoro, S. (submitted). "*SASSIER22*: Full-waveform tomography of the eastern Indonesian region that includes surface topography and the fluid ocean". *Geochemistry, Geophysics, Geosystems*.

Wei, S. S., Wiens, D. A., Zha, Y., Plank, T., Webb, S. C., Blackman, D. K., Dunn, R. A., and Conder, J. A. (2015). "Seismic evidence of effects of water on melt transport in the Lau back-arc mantle". *Nature* vol. 518.7539, pp. 395–398.

Wells, D. L. and Coppersmith, K. J. (Aug. 1994). "New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement". *Bulletin of the Seismological Society of America* vol. 84.4, pp. 974–1002. ISSN: 0037-1106. eprint: https://pubs.geoscienceworld.org/bssa/article-pdf/84/4/974/2707918/BSSA0840040974.pdf.

White, R. and McKenzie, D. (1989). "Magmatism at rift zones: the generation of volcanic continental margins and flood basalts". *Journal of Geophysical Research: Solid Earth* vol. 94.B6, pp. 7685–7729.

Widiyantoro, S. and Hilst, R. van der (1996). "Structure and evolution of lithospheric slab beneath the Sunda arc, Indonesia". *Science* vol. 271.5255, pp. 1566–1570.

Widiyantoro, S., Pesicek, J., and Thurber, C. (2011). "Subducting slab structure below the eastern Sunda arc inferred from non-linear seismic tomographic imaging". *Geological Society, London, Special Publications* vol. 355.1, pp. 139–155.

Wiggins, R. A. (1972). "The general linear inverse problem: Implication of surface waves and free oscillations for earth structure". *Reviews of Geophysics* vol. 10.1, pp. 251–285.

Wolfe, C. J., VanDecar, J. C., Solomon, S. C., et al. (1997). "Seismic structure of the Iceland mantle plume". *Nature* vol. 385.6613, pp. 245–247.

Wolpert, D. H. and Macready, W. G. (1997). "No free lunch theorems for optimization". *IEEE Transactions on Evolutionary Computation* vol. 1.1, pp. 67–82.

Woodhouse, J. H. and Dziewonski, A. M. (1984). "Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms". *Journal of Geophysical Research: Solid Earth* 89.B7, pp. 5953–5986.

Wu, J., Suppe, J., Lu, R., and Kanda, R. (2016). "Philippine Sea and East Asian plate tectonics since 52 Ma constrained by new subducted slab reconstruction methods". *Journal of Geophysical Research: Solid Earth* vol. 121.6, pp. 4670–4741.

Xing, G. and Zhu, T. (2021). "A viscoelastic model for seismic attenuation using fractal mechanical networks". *Geophysical Journal International* vol. 224.3, pp. 1658–1669.

Yamazaki, T., Takahashi, M., Iryu, Y., Sato, T., Oda, M., Takayanagi, H., Chiyonobu, S., Nishimura, A., Nakazawa, T., and Ooka, T. (2010). "Philippine Sea Plate motion since the Eocene estimated from paleomagnetism of seafloor drill cores and gravity cores". *Earth, Planets and Space* vol. 62.6, pp. 495–502.

Yang, J., Zhu, H., Li, X., Ren, L., and Zhang, S. (2020). "Estimating P Wave Velocity and Attenuation Structures Using Full Waveform Inversion Based on a Time Domain Complex-Valued Viscoacoustic Wave Equation: The Method". *Journal of Geophysical Research: Solid Earth* vol. 125.6, e2019JB019129.

Yang, T., Gurnis, M., and Zahirovic, S. (2016). "Mantle-induced subsidence and compression in SE Asia since the early Miocene". *Geophysical Research Letters* vol. 43.5, pp. 1901–1909.

Yang, T. and Shen, Y. (2005). "P-wave velocity structure of the crust and uppermost mantle beneath Iceland from local earthquake tomography". *Earth and Planetary Science Letters* vol. 235.3-4, pp. 597–609.

Yanovskaya, T. (1997). "Resolution estimation in the problems of seismic ray tomography". *Physics of the Solid Earth* vol. 33.9, pp. 762–765.

Yeh, W. (1986). "Review of parameter identification procedures in groundwater hydrology: The inverse problem". *Water Resources Research* vol. 22.2, pp. 95–108.

Yomogida, K. (1992). "Fresnel zone inversion for lateral heterogeneities in the Earth". *Pure and Applied Geophysics* vol. 138.3, pp. 391–406.

Yuan, Y. O., Bozdağ, E., Ciardelli, C., Gao, F., and Simons, F. J. (2020). "The exponentiated phase measurement, and objective-function hybridization for adjoint waveform tomography". *Geophysical Journal International* vol. 221.2, pp. 1145–1164.

Yue, H., Castellanos, J. C., Yu, C., Meng, L., and Zhan, Z. (2017). "Localized water reverberation phases and its impact on backprojection images". *Geophysical Research Letters* vol. 44.19, pp. 9573–9580. DOI: 10.1002/2017GL073254.

Zenonos, A., De Siena, L., Widiyantoro, S., and Rawlinson, N. (2019). "P and S wave travel time tomography of the SE Asia-Australia collision zone". *Physics of the Earth and Planetary Interiors* vol. 293, p. 106267.

— (2020). "Direct inversion of S-P differential arrival times for ratio in SE Asia". *Journal of Geophysical Research: Solid Earth* vol. 125.5, e2019JB019152.

Zhang, X. and Curtis, A. (2020). "Variational full-waveform inversion". *Geophysical Journal International* vol. 222.1, pp. 406–411.

Zhao, D., Lei, J., and Liu, L. (2008). "Seismic tomography of the Moon". *Chinese Science Bulletin* vol. 53.24, pp. 3897–3907.

Zhao, L., Chen, P., and Jordan, T. H. (2006). "Strain Green's tensors, reciprocity, and their applications to seismic source and structure studies". *Bulletin of the Seismological Society of America* vol. 96.5, pp. 1753–1763.

Zhao, L., Jordan, T. H., Olsen, K. B., and Chen, P. (2005). "Fréchet kernels for imaging regional earth structure based on three-dimensional reference models". *Bulletin of the Seismological Society of America* vol. 95.6, pp. 2066–2080.

Zhou, C., Cai, W., Luo, Y., Schuster, G. T., and Hassanzadeh, S. (1995). "Acoustic wave-equation traveltime and waveform inversion of crosshole seismic data". *Geophysics* vol. 60.3, pp. 765–773.

Zhou, Y., Dahlen, F., and Nolet, G. (2004). "Three-dimensional sensitivity kernels for surface wave observables". *Geophysical Journal International* vol. 158.1, pp. 142–168.

Zhou, Y., Ni, S., Chu, R., and Yao, H. (2016). "Accuracy of the water column approximation in numerically simulating propagation of teleseismic PP waves and Rayleigh waves". *Geophysical Journal International* vol. 206.2, pp. 1315–1326.

Zhu, H., Bozdağ, E., Peter, D., and Tromp, J. (2012). "Structure of the European upper mantle revealed by adjoint tomography". *Nature Geoscience* vol. 5.7, pp. 493–498.