Full-waveform tomography of the African Plate

Dirk-Philip van Herwaarden[1], Solvi Thrastarson[1], Vaclav Hapla[1], Michael Afanasiev[1,2], Jeannot Trampert[3], and Andreas Fichtner[1]

[1]Department of Earth Sciences, Institute of Geophysics, ETH Zürich, 8092 Zürich, Switzerland
[2]Now at Mondaic Ltd, Zürich, Switzerland
[3]Department of Earth Sciences, Utrecht University, Princetonlaan 8a, 3584 CB Utrecht, The Netherlands

Key Points:

- We invert waveform recordings of 397 earthquakes down to periods of 35 s to recover crustal and mantle structure of the African Plate.
- We apply the dynamic mini-batch approach to reduce computational costs and enable inclusion of more data.
- We present a new approach to uncertainty analysis in which we perturb the model and continue the inversion.

Corresponding author: Dirk-Philip van Herwaarden, `dirkphilip.vanherwaarden@erdw.ethz.ch`

Abstract

We present a full-waveform inversion (FWI) study of the African plate. Starting from the first-generation Collaborative Seismic Earth Model (CSEM), we invert seismograms filtered to a minimum period of 35 s and compute gradients of the misfit function with respect to the model parameters using the adjoint state method. In contrast to the conventional FWI approach, we use dynamically changing data subsets (mini-batches) of the complete dataset to compute approximate gradients at each iteration. This approach has three significant advantages: 1) it reduces computational costs for model updates and the inversion, 2) it enables the use of larger datasets without increasing iteration costs, and 3) it makes it trivial to assimilate new data since we can add it to the complete dataset without changing the misfit function, thereby enabling "evolutionary FWI". We perform 130 mini-batch iterations and invert data from 397 unique earthquakes and 184,356 unique source-receiver pairs at the cost of approximately 10 full-data iterations. We clearly image tectonic features such as the Afar triple junction. Particularly interesting are the low-velocity zones below the Hoggar, Aïr, and Tibesti Mountains, pronounced more than in earlier works. Finally, we introduce a new strategy to assess model uncertainty. We deliberately perturb the final model, perform additional mini-batch iterations, and compare the result with the original final model. This test uses actual seismic data instead of artificially generated synthetic data and requires no assumptions about the linearity of the inverse problem.

Plain Language Summary

We present a seismic tomography of the African plate. We use the full-waveform inversion (FWI) technique to update the velocity structure of a global Earth model, called the Collaborative Seismic Earth Model. FWI promises to create more detailed images of the subsurface, compared to other approaches that rely on less complete physics. This technique, however, requires a lot of computing power. We use a stochastic optimization method that has been shown to reduce the need for computational resources. It also allows us to include more seismic data to update our model given a certain budget. In our final model, we see low-velocity structures beneath mountain ranges in the central Sahara, such as Hoggar and Tibesti, that have not been seen so clearly in earlier works, and many other tectonic features. Finally, we propose a new strategy to test the quality of the resulting model. We introduce errors into our final velocity model and continue to update the seismic structure. We then test if we can remove the introduced errors from our model. The outcome gives us an idea about the resolution as well as the reliability of our solution.

1 Introduction

The African continent is home to several unique surface anomalies, such as the East African Rift System (EARS), the Afar Depression, and significant topography in the southern part of the continent that cannot be explained by plate tectonics alone (Fishwick & Bastow, 2011). At the same time, the seismic structure, which may help to explain some of these distinctive features, has been notoriously difficult to assess for decades. Relative to other continents, there is a lack of seismic instrumentation, primarily due to the economic and political situation across much of the continent. The station coverage in Africa has been substantially improved by the AfricaArray (Nyblade et al., 2011). This study takes advantage of the resulting substantial extension of the seismic data collection.

In addition to a relative lack of stations in Africa, high-magnitude earthquakes are uncommon there. Emry et al. (2019) performed tomographic imaging using ambient noise to mitigate this issue. This approach has the benefit of providing additional data over just earthquake data. However, the common assumption that ambient noise cross-correlations

are equivalent to Green's functions, not considering the noise source distribution, may lead to imaging artefacts (e.g. Wapenaar, 2004; Fichtner et al., 2016; Sager et al., 2018). While it is, in principle, possible to take the noise source distribution into account, it has only been done on the global scale with long-period data (Sager et al., 2020) due to the implementation complexity and high computational costs.

These factors make it challenging to obtain high-quality tomographic images that might shed light on the origins of Africa's unique tectonic features. During the past two decades, there have been a number of large-scale surface wave studies (e.g. Ritsema & van Heijst, 2000; Sebai et al., 2006; Pasyanos & Nyblade, 2007; Priestley et al., 2008; Fishwick, 2010; Celli et al., 2020) with a focus on the African continent. Generally, these studies agreed well with each other and find low-velocity anomalies beneath the EARS and high-velocity anomalies beneath the major cratons. With more data becoming available and ever improving methodology, the seismic structure in these models is becoming more and more detailed. However, the resolution is still relatively low compared to well instrumented regions such as Europe, North America, Australia, China and Japan.

In this contribution, we aim to improve upon the state-of-the-art tomographic models by using the technique of full-waveform inversion (FWI) (e.g. Virieux & Operto, 2009; Fichtner, 2010; Q. Liu & Gu, 2012). In contrast to the traditional traveltime tomography methods, FWI extracts information from the entire waveform, and the sensitivity of each recording can be accurately computed through the adjoint state method (e.g. Tarantola, 1984, 1988; Tromp et al., 2005; Fichtner et al., 2006). This enables us to account for finite-frequency and wavefield scattering effects, and recover a more detailed structure given the same number of recordings (e.g. Pratt, 1999; Virieux & Operto, 2009). We compute model updates with the dynamic mini-batch approach introduced in van Herwaarden et al. (2020), and described more in detail later in Section 4.

To deal with the sparsity of seismic data on the African continent, we extend the modelled domain to use earthquakes around the entire African Plate and exploit waveform recordings in the much more densely instrumented European countries. Celli et al. (2020) took a similar approach by applying automated multimode inversion (AMI) (Lebedev et al., 2005) to surface and S waves. Eventually, our model will become part of the CSEM (Fichtner, van Herwaarden, Afanasiev, Simutė, et al., 2018), which serves as a framework to integrate information from various tomographic studies at different scales.

This manuscript is structured as follows. In Section 2, we introduce the geologic background of the study region; Section 3 shows the used dataset; in Section 4, we describe the methodology; Section 5 shows the results of the waveform inversion and the final model; and Section 6 the results of the uncertainty analysis. Finally, we conclude the manuscript with a discussion and conclusions in Sections 7 and 8.

2 Geologic background

Africa is geologically rich, and any section on its geologic background can only be incomplete. Hence, the following paragraphs are limited to the description of a few features that may be considered of broader interest and will appear again later when we roughly interpret the model in terms of the regional geology. Fig. 1 shows an elevation plot of the continent, the most significant plate boundaries, and volcanoes that we marked with red triangles, as well as the location of major cratons and tectonic features.

The continent includes several large cratons. While often referred to as a single craton, these regions can also be considered collections of smaller cratonic fragments. The most significant ones are the West African Craton, the Congo Craton, and the Kalahari Craton, composed of the Kaapvaal and Zimbabwe Craton (Begg et al., 2009).

On the other hand, Africa also contains much more volcanically and tectonically active regions. Significant hotspot areas in the Atlantic Ocean include Cape Verde or the Walvis Ridge and the Comoro Islands in the Indian Ocean. In the northwest of the continent itself, the Atlas Mountains are located, separating the Sahara Desert from the Mediterranean Sea. Continuing in the southeast direction, we can see a distinct mountain range group in the central Sahara, including the Hoggar, Aïr, and Tibesti Mountains, followed by the Darfur Dome in the southeast Sahara corner. Eastern Africa is home to the Afar Triple junction and the EARS. On the opposite side of the continent, the Cameroon volcanic line starts near Lake Chad and stretches in the southwest direction into the Atlantic Ocean, forming an island chain including, e.g., São Tomé and Principe. In the south, we find a large region of significant topography at the Southern African Plateau, surrounded by the Great Escarpment. Here, bands of a highly elevated surface fall steeply toward the coasts.

The geology of Northern Africa is hard to observe from the surface, as most of it is covered beneath the Sahara desert. Several regions of elevated topography exist throughout Northern Africa, some of which expose altered Proterozoic or Archean rocks, and Cenozoic volcanic deposits cap (Abdelsalam et al., 2002). The uplift in these areas has been proposed to have a dynamic origin (Burke & Gunnell, 2008), as no known plate boundaries exist there. In particular, very few seismic stations are installed in the Sahara. Despite the relative lack of information, earlier studies, e.g., by Fishwick and Bastow (2011) suggest that parts of Northern Africa have slower than average mantle velocities and that several high-velocity regions may mark the locations of cratonic fragments within the Saharan Metacraton (Liégeois et al., 2013).

The EARS is an area that is composed of multiple rift branches. It extends to the Red Sea in the northeast and along Tanzania to Mozambique and Zambia in the south (Chorowicz, 2005). The Afar region and the Red Sea in the north are home to the only known area on Earth where incipient seafloor spreading is sub-aerially exposed (Almalki et al., 2016). Most of the EARS follows the same trends as the mobile belts, especially in the south where several cratons are located. The northern and eastern branches of the EARS are more magmatic than the western and southern branches (Furman, 2007). Tomographic studies have confirmed this; e.g. Benoit et al. (2006); Bastow et al. (2008); Fishwick (2010) find lower velocities in the north relative to the south. Large parts of the EARS also coincide with regions of high elevations that may be associated with dynamic uplift due to mantle upwellings (e.g. Mulibo & Nyblade, 2013).

3 Data

Africa's two essential station networks are the AfricaArray (Nyblade et al., 2011) and the NARS array (Utrecht University, 2013). We downloaded recordings from 397 earthquakes through the FDSN web services (Romanowicz & Dziewonski, 1986). The earthquake source information was extracted from the GCMT Catalog (Ekström et al., 2012) and selected to fall within a magnitude range of 5.5 to 6.7. Empirically, this range provides a good signal while minimizing the finite-source effects associated with large earthquakes (Vallée, 2013). We use 184,356 unique source-receiver pairs recorded by 6,088 three-component receivers. Fig. 2a shows a ray coverage plot of the entire dataset.

Before inverting, we preprocess the data as follows. First, the data is detrended, and we subtract the mean. Next, we remove the instrument response and filter the seismogram to the period band of interest.

In addition to the entire dataset, we also consider a separate one that we refer to as the validation dataset. We do not use these data to update the model directly but as a reference throughout the inversion to assess improvements in waveform fit. The validation dataset is shown in Fig. 2b and discussed in more detail in the next section.

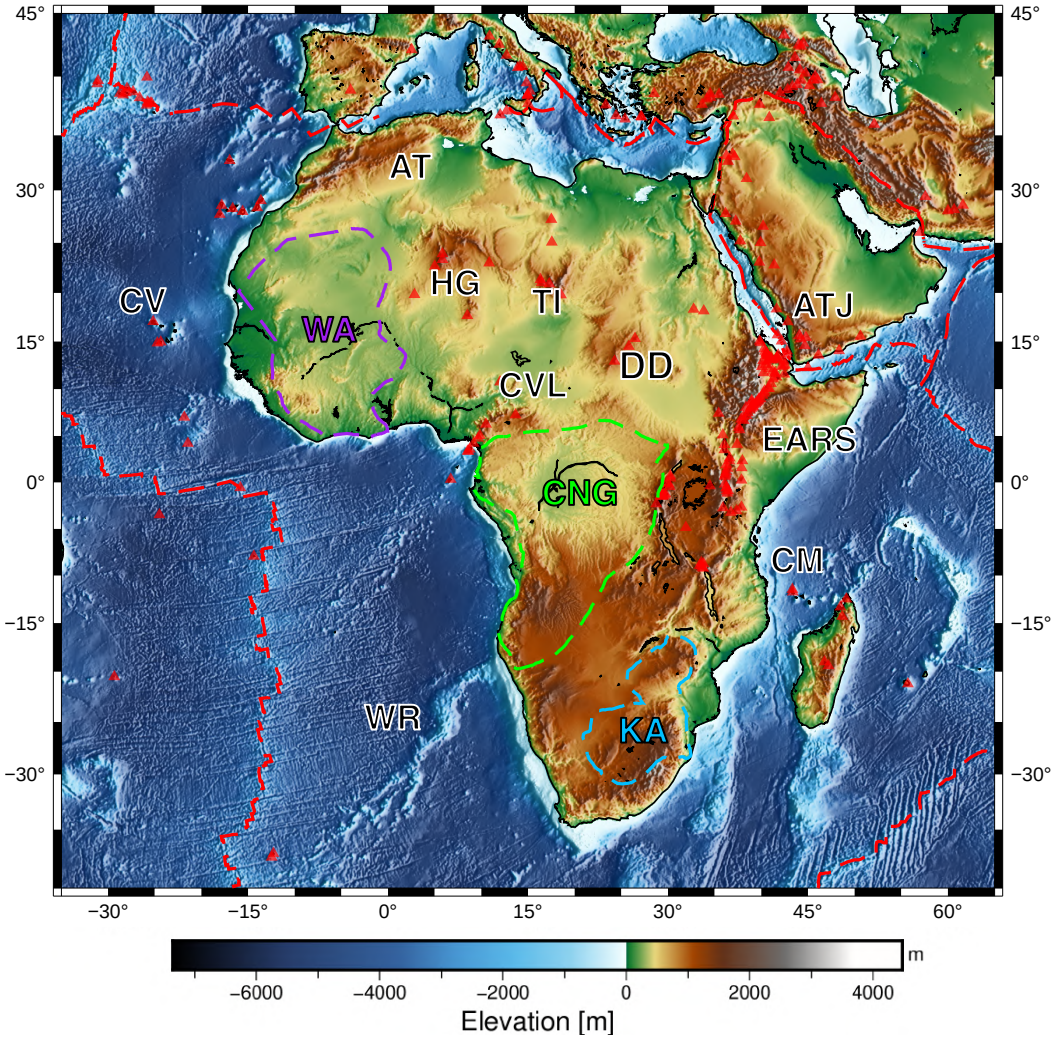

Figure 1. Elevation plot with volcanoes marked as red triangles. Major plate boundaries (Bird, 2003) are marked with a red dashed line. The main cratons (Begg et al., 2009) are indicated on the map: **WA:** West African Craton, **KA:** Kalahari Craton, **CNG:** Congo Craton. Other features: **EARS:** East African Rift System, **CVL:** Cameroon Volcanic Line, **AT:** Atlas Mountains, **HG:** Hoggar mountains, **TI:** Tibesti Mountains, **DD:** Darfur Dome, **ATJ:** Afar Triple Junction, **CV:** Cape Verde, **WR:** Walvis Ridge, **CM:** Comoro Volcanic Islands. Volcanoes were taken from the online database provided by NCEI (2022), and elevation data was supplied by GMT (Wessel et al., 2019).

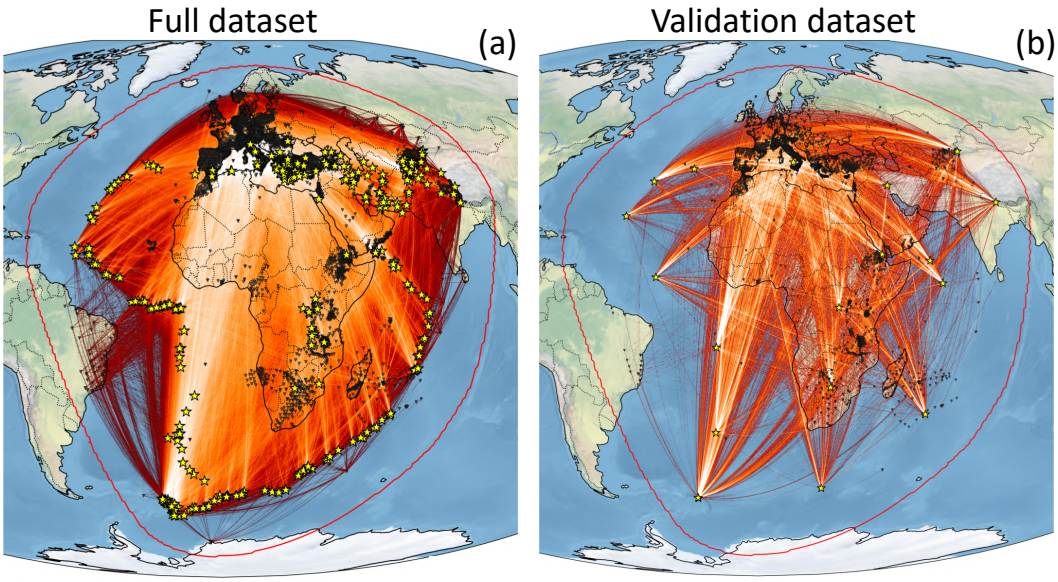

Figure 2. Raydensity plots of the datasets used in this study. Bright colors represent a higher density of rays. Yellow stars represent earthquake locations, and black marks denote station locations. The red line represents the edge of the modeled domain. **(a)** The full dataset contains data from 397 earthquakes. All events were recorded between 1995 and 2020 by 6,088 stations, resulting in 184,356 unique source-receiver pairs. **(b)** The validation dataset consists of 19 events and 9,525 unique source-receiver pairs. We do not use these data to update the model, or in other words, we do not specifically try to fit these data. At several stages in the inversion, the misfit is computed for this independent dataset to measure improvements in waveform fit.

4 Forward and inverse modeling

We perform forward and adjoint modeling with the wave propagation and inversion package Salvus (Afanasiev et al., 2019). Salvus utilizes the spectral-element method. The method was first introduced in the fluid dynamics community (Patera, 1984) and later adapted to solve the seismic wave equation (e.g. Faccioli et al., 1996; Komatitsch & Tromp, 1999; Chaljub et al., 2003). The implementation allows for ocean loading (Komatitsch & Tromp, 2002), viscoelasticity (e.g. Robertsson et al., 1994), ellipticity of the Earth, and topography.

Our mesh extends down to the outer core, as we are primarily interested in surface waves and first arrivals from source-receiver pairs with limited epicentral distances. We place absorbing boundaries on all outer edges except the surface, where we impose a free-surface boundary condition. Gradients of the misfit function with respect to the model parameters (or just gradients from hereon) are computed with the adjoint state method (e.g. Tarantola, 1984, 1988; Tromp et al., 2005; Fichtner et al., 2006).

We largely follow the workflow established in Fichtner et al. (2009) and Tape et al. (2009), enhanced with the dynamic mini-batch approach introduced in van Herwaarden et al. (2020). In contrast to conventional FWI, this technique operates on subsets of the full dataset. The gradient associated with each iteration is then only computed for this subset. The subset size may change between iterations, depending on the estimated difference of the mini-batch gradient approximation with respect to the full dataset gradient. This concept has also been described as adaptive gradient optimization (AGO) (Bernal-Romero & Iturrarán-Viveros, 2021).

The method has three significant benefits. First, it enables the use of significantly larger datasets while keeping similar iteration costs. Therefore, we can incorporate more data, which provides the inversion with more information and possibly leads to a better image. Second, the approach has been shown in van Herwaarden et al. (2020) to require fewer simulations than conventional FWI to obtain a model of the same quality, thereby extending the frequency range for a given computational budget. Third, it becomes trivial to assimilate new data on the fly during the inversion process. We can easily add data to the dataset from which we sample the mini-batches. This fact leads to an "evolutionary" mode of FWI (van Herwaarden et al., 2021), where the model can evolve as new data becomes available. In addition, the use of mini-batches may help to prevent data overfitting (e.g. Hoffer et al., 2017). Since we compute the gradient for each subset, information unique to each subset becomes harder to fit.

Like in van Herwaarden et al. (2020), we use trust-region L-BFGS (D. C. Liu & Nocedal, 1989) as the optimization algorithm. L-BFGS is a quasi-Newton method, where we calculate the model updates by scaling the gradients with the L-BFGS approximation of the inverse Hessian. We compute this approximation from a history of gradients and models from earlier iterations.

We then invert in two separate stages. We first filter the data to a period range of 55–130 s. After convergence in this period band, we filter the data to 35–130 s. We do this for two reasons. First, by starting at longer periods, we reduce the computational requirements in the wavefield simulations and, thus, the overall cost of the inversion. Second, this mitigates the risk of possible cycle-skipping issues when the starting model is not good enough.

4.1 Starting model

The starting model is the first generation of the transverse isotropic Collaborative Seismic Earth Model (CSEM) (Fichtner, van Herwaarden, Afanasiev, Simute, et al., 2018). The CSEM is a framework to sequentially update an initial global background model with

models of increasingly more detail. The global background model consists of the Preliminary Reference Earth Model (PREM) (Dziewoński & Anderson, 1981), where a linear gradient replaces the 220 km discontinuity. The mantle is perturbed by velocities from S20RTS (Ritsema et al., 1999), to which P-velocity variations are scaled (Ritsema & van Heijst, 2002). Finally, the crust of PREM is overwritten by the crustal model of (Meier et al., 2007), derived from a surface wave inversion, to complete the global background. Within the domain that we consider, regional updates have been made to this model in Europe (Fichtner, Trampert, et al., 2013), the South Atlantic (Colli et al., 2013), the North Atlantic (Rickers et al., 2013), the Western Mediterranean (Fichtner & Villaseñor, 2015), Turkey (Fichtner, Saygin, et al., 2013) and the Central and Eastern Mediterranean (Blom et al., 2020).

The model is parameterized in vertically and horizontally propagating/polarized P/S velocities, v_{PH}, v_{PV}, v_{SH}, v_{SV}, density ρ, and the dimensionless parameter η. Independently constraining the entire set of parameters is unrealistic, given the limited data coverage. For this reason, we only invert for isotropic P-wave velocity v_P, and transverse isotropic S-wave velocities v_{SH} and v_{SV}. To reduce the parameter space, we set $\eta = 1$ and set $v_P = v_{PV} = v_{PH}$. Fichtner, Trampert, et al. (2013) give a more detailed explanation of this decision. We show several depth slices through the initial distributions of the v_{SV} parameter in Fig. 3.

4.2 Misfit functional and validation misfits

The misfit functional that we optimize is the time- and frequency-dependent phase misfit (Fichtner et al., 2008). It does not require the isolation of specific phases, and eliminates uncertainties related to imprecise earthquake magnitudes by ignoring constant amplitude scalings. However, it keeps the waveform information, i.e., the information on relative amplitudes of neighboring wiggles.

We also employ a station weighting scheme that empirically leads to faster convergence by down-weighting clusters of stations. For this purpose, misfits at station location $\mathbf{x}_r$ are multiplied by the factor

$$W_r = c \left(\sum_{i=1, i \neq r}^{n} \frac{1}{||\mathbf{x}_i - \mathbf{x}_r||} \right)^{-1} , \tag{1}$$

where n is the total number of all other station locations $\mathbf{x}_i$ for the respective event. The events are then normalized by factor c, so the average weight per station is constant across events. Ruan et al. (2019) gives a more detailed explanation of station weighting methods in regional to global scale FWI in 2D.

In addition to the misfits and gradients used for the optimization process, we also compute misfits for the validation dataset shown in Fig. 2b. We calculate no gradients for this dataset and do not use them to compute model updates directly. We compute the L_2 norm of the difference between these data and synthetics over the full trace every five iterations. We use full traces to avoid the bias that selecting measurement windows would introduce. On the other hand, it means that we include all the noise in this independent measurement, and therefore misfit decrease will be smaller than it would be for only high-quality data.

The use of a validation dataset serves multiple purposes. First, it enables us to have a measure for convergence since a vanishing misfit decrease suggests we approach a minimum. Secondly, a strategy like this may help prevent us from overfitting our model to the data; if the misfit decreases for the full dataset but not for the validation one, we likely fit noise at that stage. The third benefit of this approach is that it guides us on the choice of the regularization parameters. Suppose independent data misfits become smaller when we reduce the gradient smoothing. In such a case, the features we introduce will likely

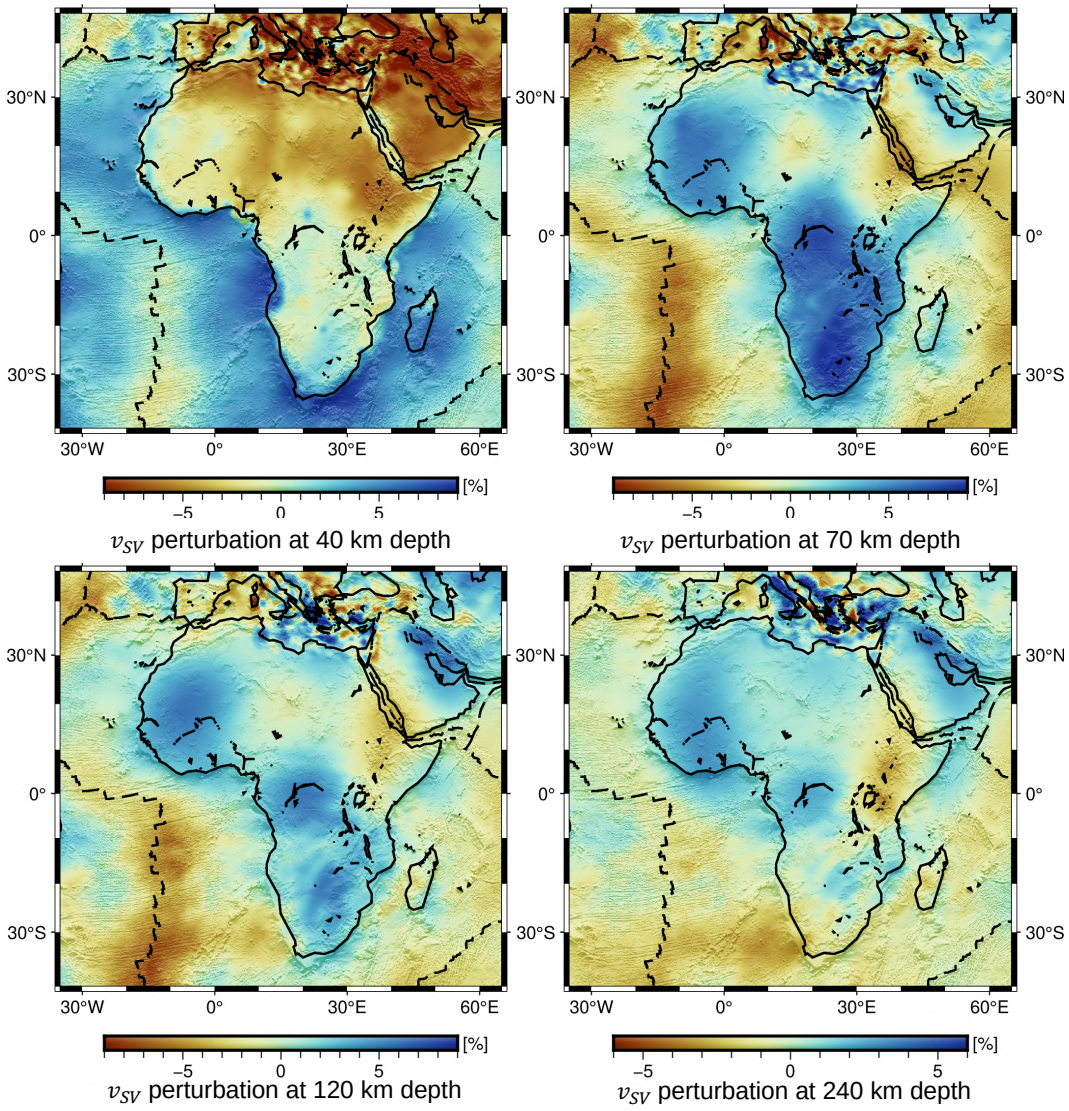

Figure 3. The starting model of the v_{SV} parameter at 40, 70, 120, and 240 km depth. The main plate boundaries (Bird, 2003) are plotted on top with dashed black lines. Perturbations are plotted relative to the lateral mean.

generalize to unseen data, and we are probably further improving the model rather than just fitting noise. Thrastarson et al. (2022) have introduced and explained this concept in more detail.

4.3 Multiscale inversion and regularization

To avoid trapping in a local minimum and to reduce computational costs, we split the inversion into different phases of varying period bands. This is a standard procedure to prevent cycle-skipping problems (e.g. Bunks et al., 1995; Krischer et al., 2018; Blom et al., 2020).

In the first phase of the inversion, we use the period band of 55–130 s and move on to a broader period band of 35–130 s. We regularize the model by smoothing the gradients using the diffusion equation. This technique was popularised first in image processing (e.g. Barash, 2002) and adopted into seismic imaging within Salvus (Afanasiev et al., 2019). Effectively, this smoothing operator is equivalent to convolving the gradient with a Gaussian kernel. A great benefit of this technique is that it enables velocity-dependent and spatially varying smoothing as well as anisotropic smoothing. We can express the smoothing length relative to the seismic wavelength at the given place in the model and smooth more in areas with higher velocities, where we expect to retrieve less detail. Furthermore, we mitigate the problem of source imprint formation by cutting out a region with a radius of 250 km around each source location.

Initially, we start conservatively by convolving with a Gaussian with a standard deviation of one minimum wavelength in the lateral direction and half a wavelength radially. We expect the radial resolution to be higher as the surface wave sensitivity kernels have a shorter cross-section in the radial direction (Takeuchi & Saito, 1972). We incrementally reduce smoothing and eventually use a standard deviation of 0.3 minimum wavelengths in the radial direction and 0.6 minimum wavelengths in the lateral direction.

4.4 Workflow management

Workflow management is often an underappreciated part of an FWI. A large-scale FWI can involve millions of different files and terabytes of storage. In addition to metadata, e.g., station- and event information, we need to keep track of heavy data such as synthetic traces, measurement windows, and adjoint sources. All these files must be tracked and used for the appropriate computing processes in each iteration. Without any form of automation, this becomes intractable for large projects.

To facilitate the management of these files and the inversion, we use LASIF (Krischer, Fichtner, et al., 2015; Thrastarson, van Herwaarden, Krischer, & Fichtner, 2021). LASIF is a software package that automates many previously time-consuming tasks associated with FWI, such as data acquisition, processing, organization, and window picking. Once set up, we use the automation package Inversionson (Thrastarson, van Herwaarden, & Fichtner, 2021), which was developed to automatically perform iterations and submit waveform simulation jobs to a high-performance computing (HPC) cluster.

5 Full-waveform tomography of the African Plate

In this section, we show the model evolution, the misfits of the independent validation data, and the final model and its significant features. Starting from the first generation CSEM, we perform 45 iterations in the first period band (55–130 s) and 85 in the second period band (35–130 s). Figs. 4 and 5 show how the v_{SV} and v_{SH} models change as a function of the iteration number. We can see that areas, such as the Afar Triple junction, start sharpening throughout the inversion. We show these two param-

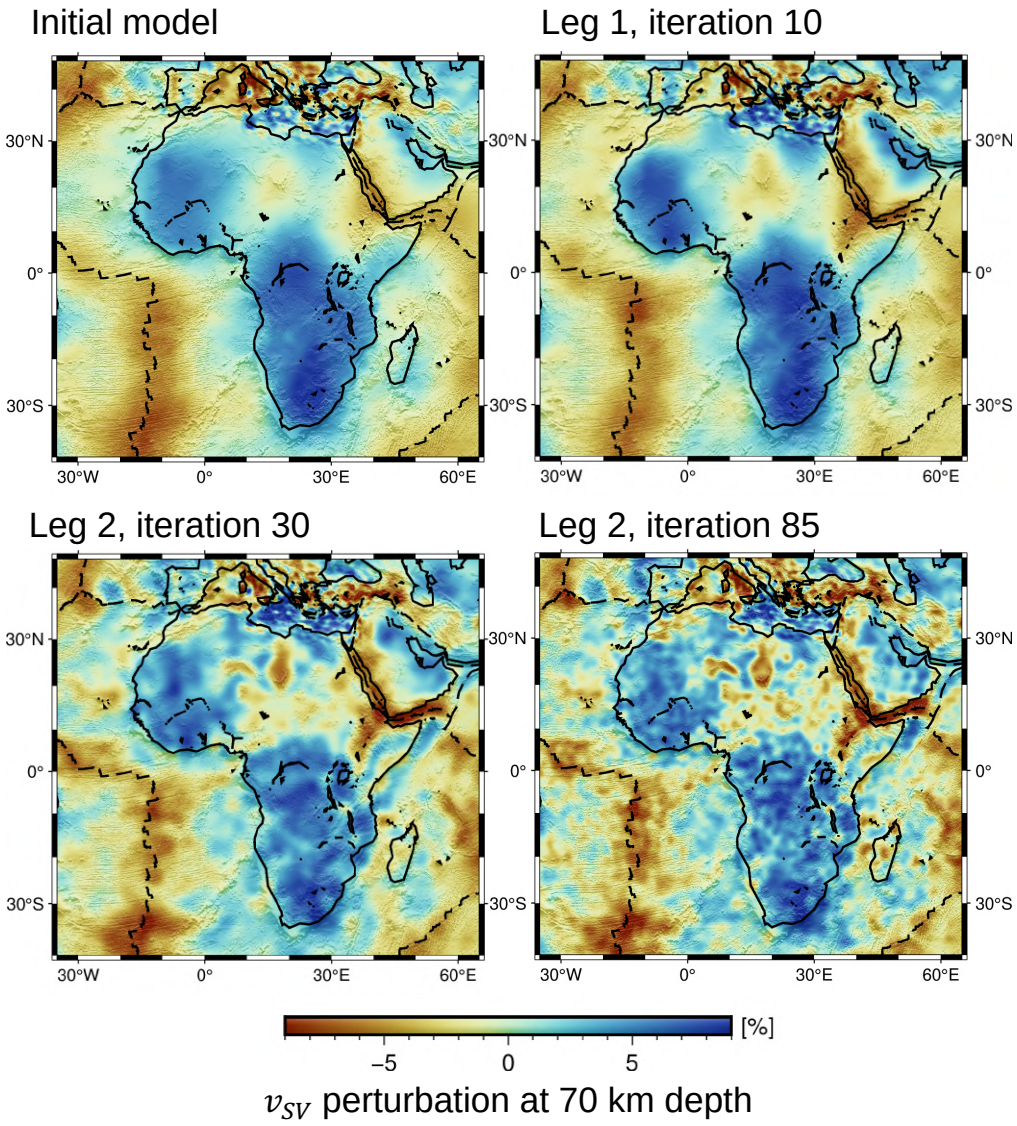

Figure 4. The v_{SV} model plotted at a selection of iterations. Perturbations are plotted relative to the lateral mean. Notice that scale lengths become smaller as we increase the frequency, and regions such as the Mid-Atlantic Ridge and the Afar Triple junction become well-defined in the form of low-velocity anomalies.

eters as they are most sensitive to the surface waves that dominate the signal within the band-pass-filtered data. For the same reason, we expect the best results for these two parameters, which we confirm in later model uncertainty tests. The data is plotted as a perturbation to the lateral mean velocities. For reference, Fig. 6 shows the lateral mean velocities for the top 250 km for the initial and final model.

5.1 Final model

Figs. 7 and 8 show the final v_{SV} and v_{SH} velocity models at a selection of depths (40, 70, 120, and 240 km). Fig. 9 focuses on the final v_{SV} model at 70 km depth, with the location of several tectonic features plotted on top.

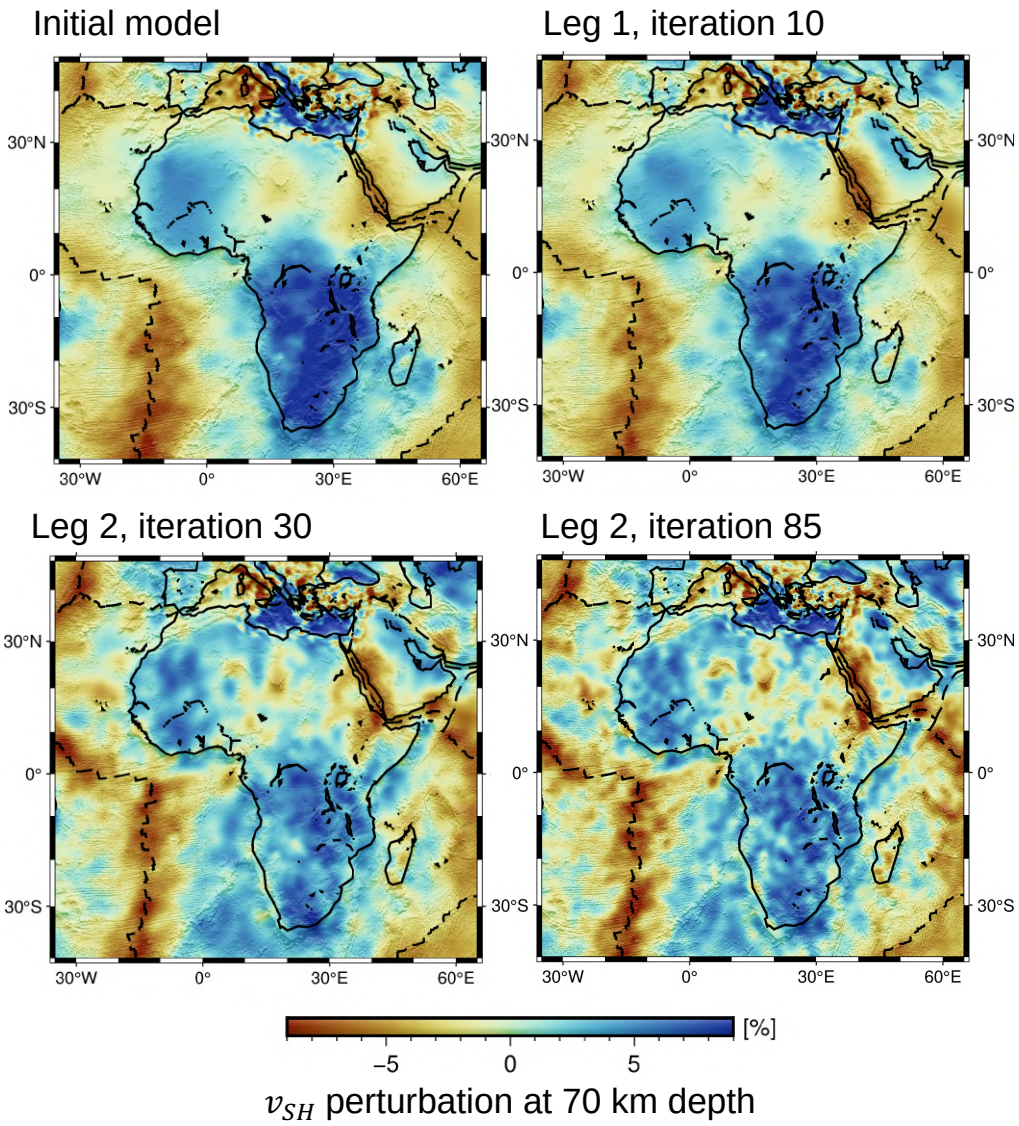

Figure 5. The v_{SH} model plotted at a selection of iterations. The description remains the same as in Fig. 4.

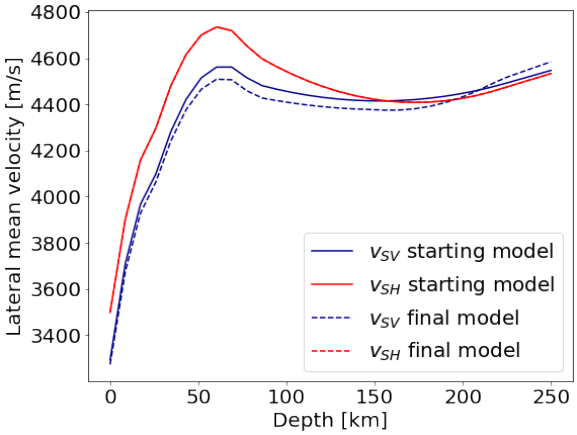

Figure 6. Depth profile of the lateral mean velocities for the v_{SV} and v_{SH} parameters of the initial and final model.

Several mountain ranges believed to be of a dynamic origin, such as the Hoggar, Aïr, and Tibesti Mountains, correspond clearly to low-velocity anomalies, especially at 70 km and 120 km depth. The most volcanically active areas, such as the Afar triple junction, the Cameroon Volcanic line, Cape Verde, and the Walvis Ridge, and even solitary volcanoes overlap favorably with low-velocity anomalies in the model, too. The Mid-Atlantic Ridge sharpened significantly compared to the initial model (Fig. 3). Furthermore, we observe relatively high velocities in Northern Africa in regions that Abdelsalam et al. (2002) hypothesized to be home to cratonic fragments. The West African, Congo, and Kalahari Craton areas also appear as high-velocity anomalies.

In Fig. 10, we show a cross-section through the East African Rift System. Low velocities extend deep downwards beneath the Afar region and Tanzania. The area beneath Tanzania and Kenya is thought to be home to a mantle superplume (e.g. Ebinger & Sleep, 1998; Thrastarson et al., 2022), and our model aligns with those findings. The model suggests that the low-velocity anomaly beneath the Tanzania region may become less pronounced below the 660 km discontinuity. Due to the dominance of the surface waves in our data, we likely have our highest sensitivity in the top several hundred kilometers of the Earth. Therefore, the findings at these deeper depths need to be interpreted with additional caution. Fig. 11 shows a vertical slice through the v_{SV} model in the Eastern Mediterranean. The CSEM in this region has already been updated with an inversion by (Blom et al., 2020). We see updates in the deeper sections and the southern part of the vertical slice that appear to improve the definition of the subduction zone.

5.2 Validation misfits and waveform fit

As mentioned in Section 4, we assess convergence by evaluating waveform fits of an independent dataset. The dynamic mini-batch approach (van Herwaarden et al., 2020) that we follow here does not give us a standard misfit curve for the whole dataset. While not commonly done in the Earth science community, it is standard practice in machine learning (e.g. Zhang & Sabuncu, 2018; Tariq et al., 2018) to estimate the quality of predictions on independent data, as it helps to estimate overfitting. Fig. 12 shows the evolution of the L_2 misfit throughout the inversion in the two period bands.

To eliminate the bias towards the windows that showed a good fit for the initial model, we computed the L_2 misfits over the whole trace. This strategy also implies that

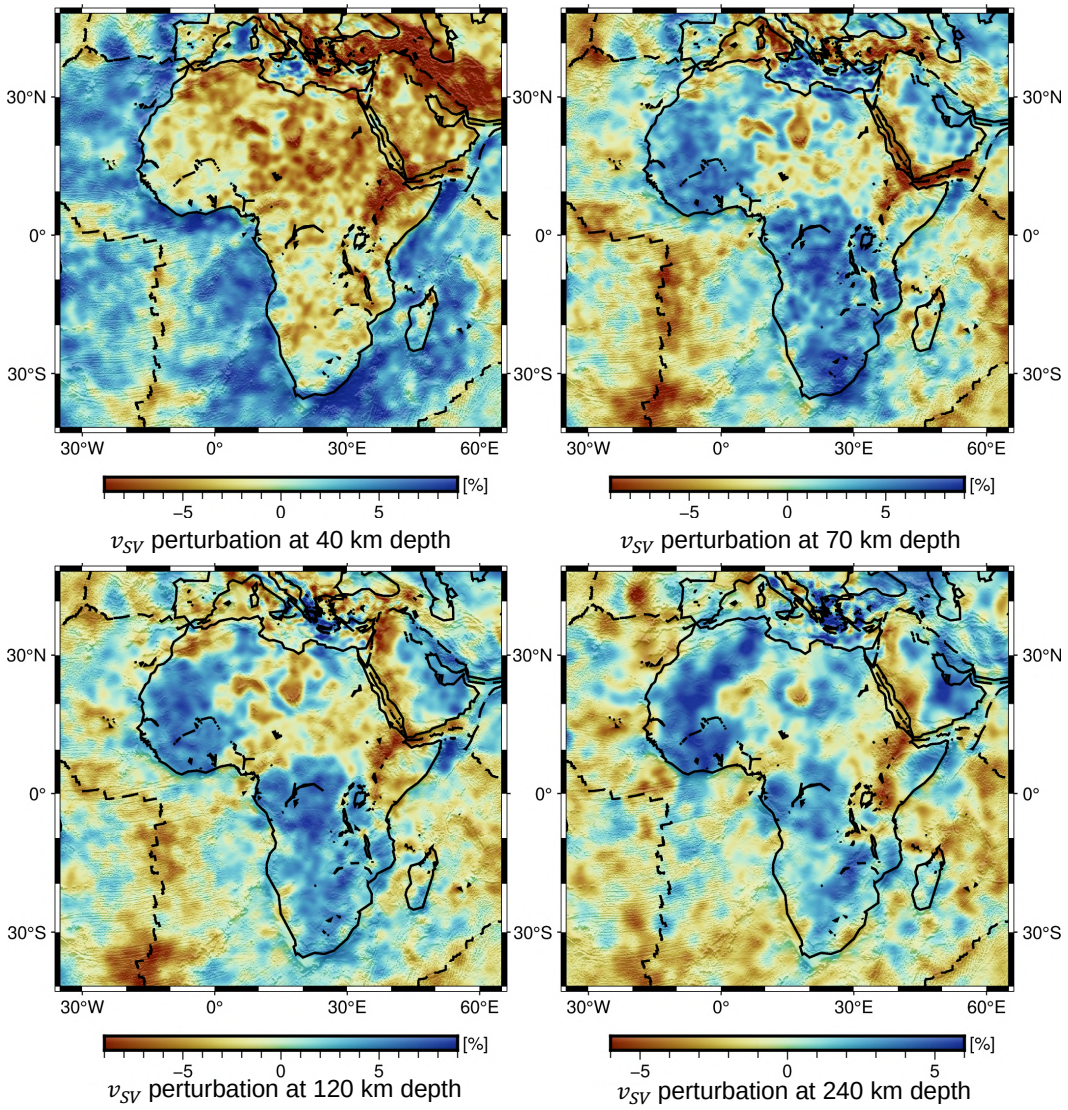

Figure 7. The final model of the v_{SV} parameter at 40, 70, 120, and 240 km depth. The main plate boundaries (Bird, 2003) are plotted on top with a black dashed line. Perturbations are plotted relative to the lateral mean. Notice how the Afar triple junction seems to morph into an elongated low-velocity anomaly underneath the entire EARS at greater depths, with especially low velocities imaged east of Lake Victoria and the Afar region.

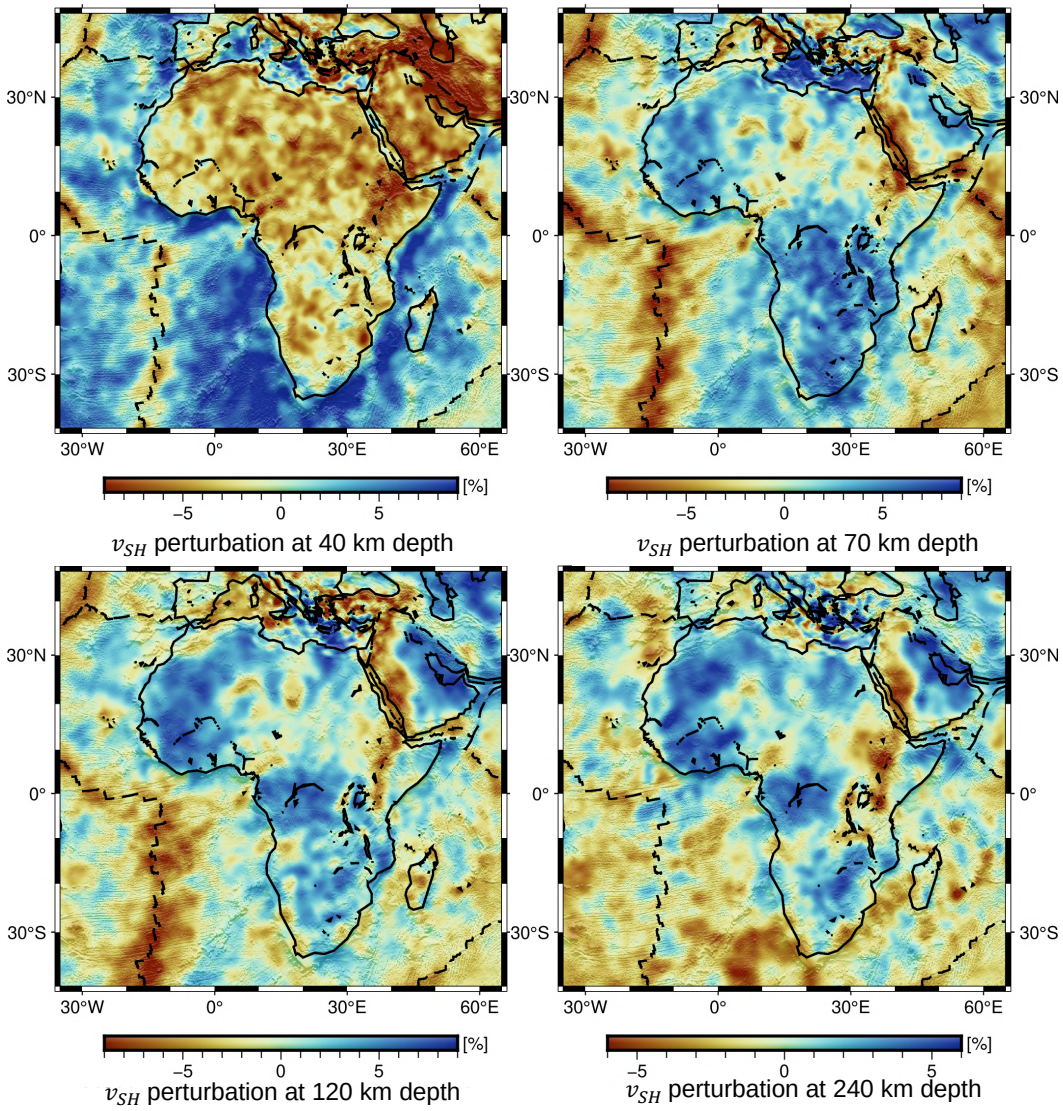

Figure 8. The final model of the v_{SH} parameter at 40, 70, 120, and 240 km depth. The main plate boundaries (Bird, 2003) are plotted on top with black dashed line. We see similar patterns as in Fig. 7.

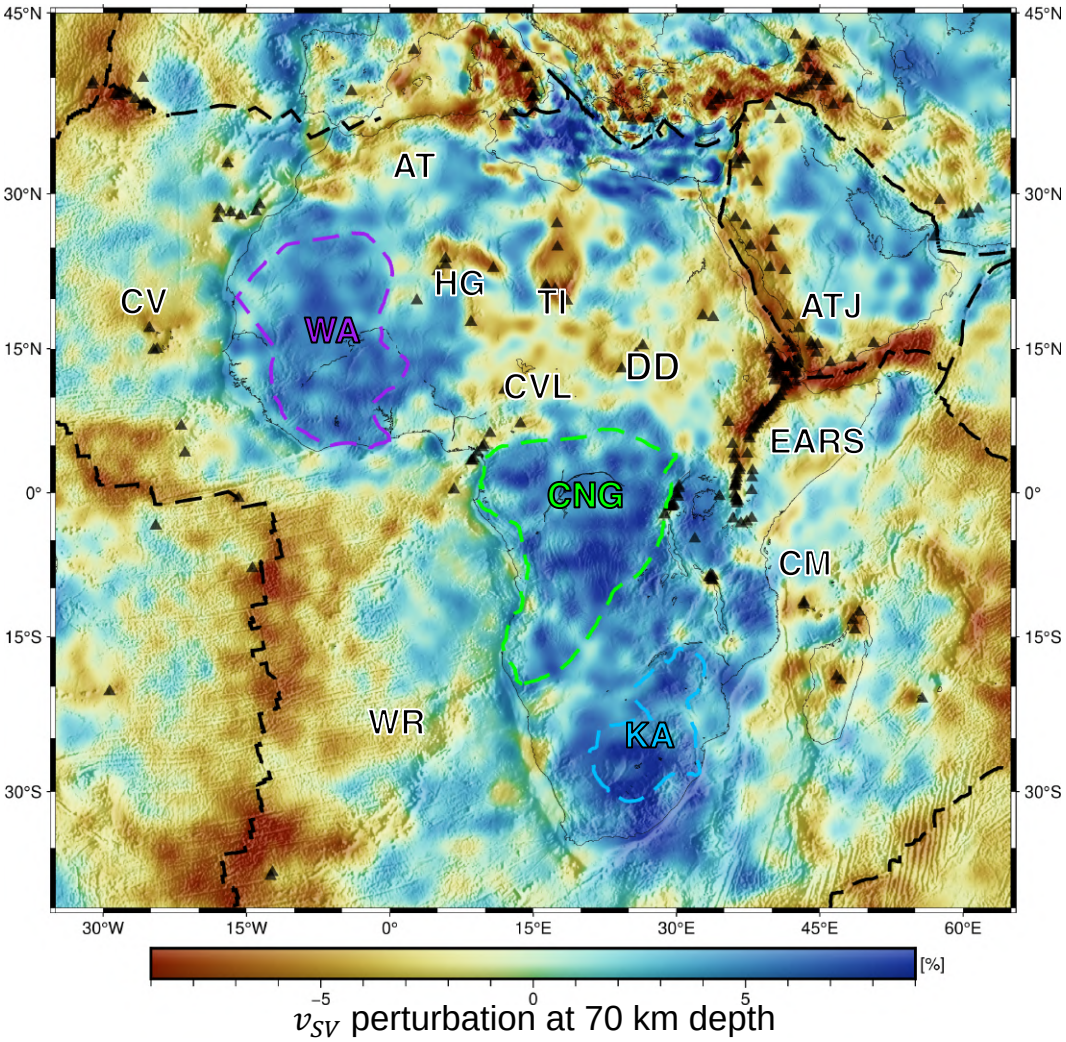

Figure 9. The final model of the v_{SV} parameter at 70 km depth. Perturbations are plotted with respect to the lateral mean. The main cratons (Begg et al., 2009) are indicated on the map: **WA:** West African Craton, **KA:** Kalahari Craton, **CNG:** Congo Craton. Other features: **EARS:** East African Rift System, **CVL:** Cameroon Volcanic Line, **AT:** Atlas Mountains, **HG:** Hoggar mountains, **TI:** Tibesti Mountains, **DD:** Darfur Dome, **ATJ:** Afar Triple Junction, **CV:** Cape Verde, **WR:** Walvis Ridge, **CM:** Comoro Volcanic Islands. In addition, volcano locations (NCEI, 2022) are indicated with black triangles. These locations correspond well to low-velocity regions in the model.

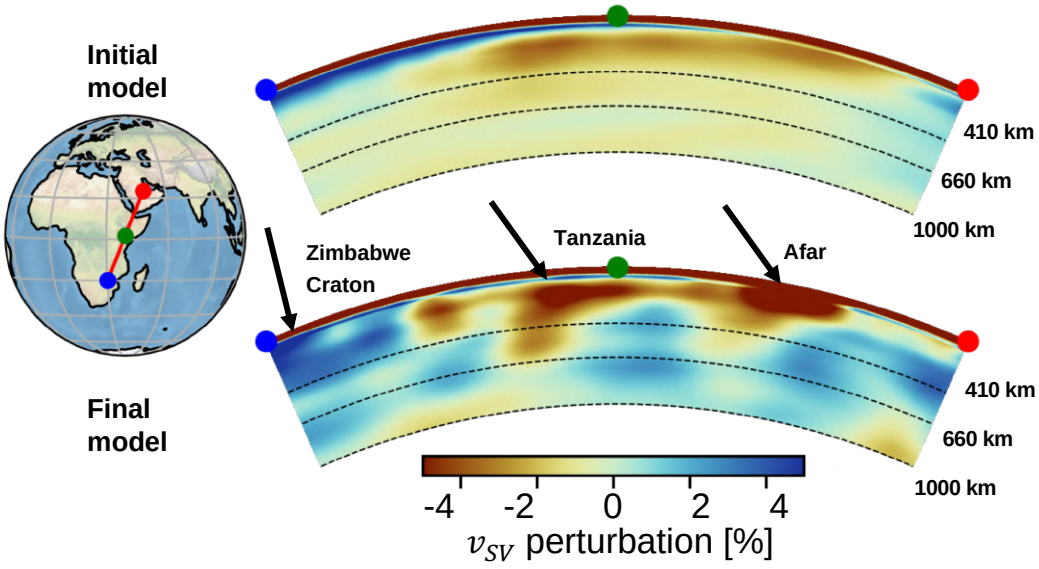

Figure 10. A cross-section through the East African Rift System. The top figure shows the initial model, and the bottom shows the final model. Low-velocity anomalies can be observed beneath Tanzania and the Afar region. At the very left of the plot, a high-velocity anomaly can be seen underneath the Zimbabwe Craton. Perturbations of the v_{SV} parameter are plotted with respect to PREM (Dziewoński & Anderson, 1981).

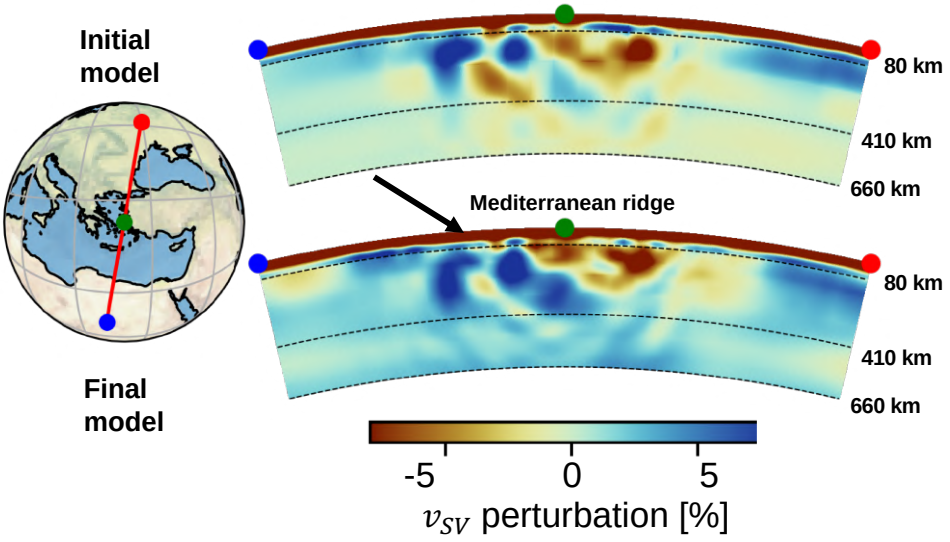

Figure 11. Vertical slice through v_{SV} model in the subduction zone of the Eastern Mediterranean. This region was previously inverted by Blom et al. (2020) and part of the initial CSEM model. Compared to the initial model, we observe a better continuation of the subduction zone from the surface to roughly 400 km depth. This likely owes to sensitivity across the entire domain of the subduction zone. Perturbations of the v_{SV} parameter are plotted with respect to PREM (Dziewoński & Anderson, 1981).

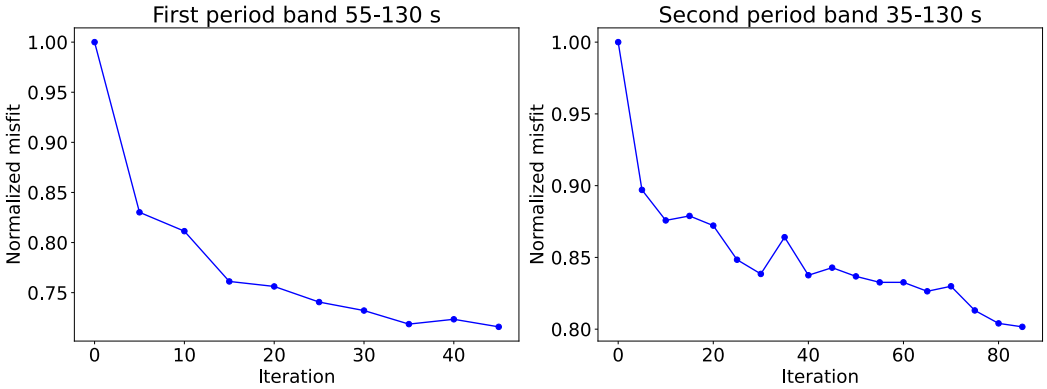

Figure 12. Normalized misfit reductions for the validation dataset over the course of the iterations in the first- (**left**) and second (**right**) leg of the inversion. Misfits are computed for the entire trace length. Misfits do not decrease monotonically, as this misfit is computed separately from the optimization, with L_2 misfits instead of time-frequency phase misfits, and for different data.

noisy (parts of) seismograms are included, making misfit reductions smaller than they would be if we only considered high-quality data.

One can observe that the misfit curves do not decrease monotonically. This is not surprising as the objective function we decrease during the optimization is different and computed for other data.

We can see a significantly improved fit, although we did not invert for these seismograms. Hence, we can assume that the improvements in waveform fit will generalize to unseen data. Fig. 13 shows a selection of three observed and synthetic seismograms computed from the validation dataset.

6 Model uncertainty analysis

Several uncertainty-assessment strategies are commonly employed to see if we can trust the retrieved models. In this section, we first briefly introduce commonly applied options. Then we propose an alternative, particularly well, but not only suited to test models created with mini-batch inversion workflows. Fully probabilistic approaches, including those that use Hamiltonian Monte Carlo sampling (e.g. Gebraad et al., 2020), are still not feasible for problems where even the deterministic methods require significant resources on modern HPC clusters. We, therefore, limit the discussion to rather qualitative assessments of uncertainty.

The most commonly applied test to assess model quality may be the checkerboard test. A model with a checkerboard pattern is created to produce synthetic data, which are then used as if they were the observed data and inverted for. Several checkerboard patterns can be tested to see what resolution might be achieved. An example of such a test on a continental scale can be found in van Herwaarden et al. (2020). The main downside of this approach is that the resulting model is not a direct input to the checkerboard test itself. Thus we can evaluate only our FWI implementation and resolution achievable with the given set of sources and receivers, but not the quality of the actual final model. Furthermore, the synthetic data do not contain natural imperfections present in real data, such as noise, outliers, timing errors, and gaps in recordings. The intrinsic non-uniqueness of the original inverse problem is thus not considered. In the context of trav-

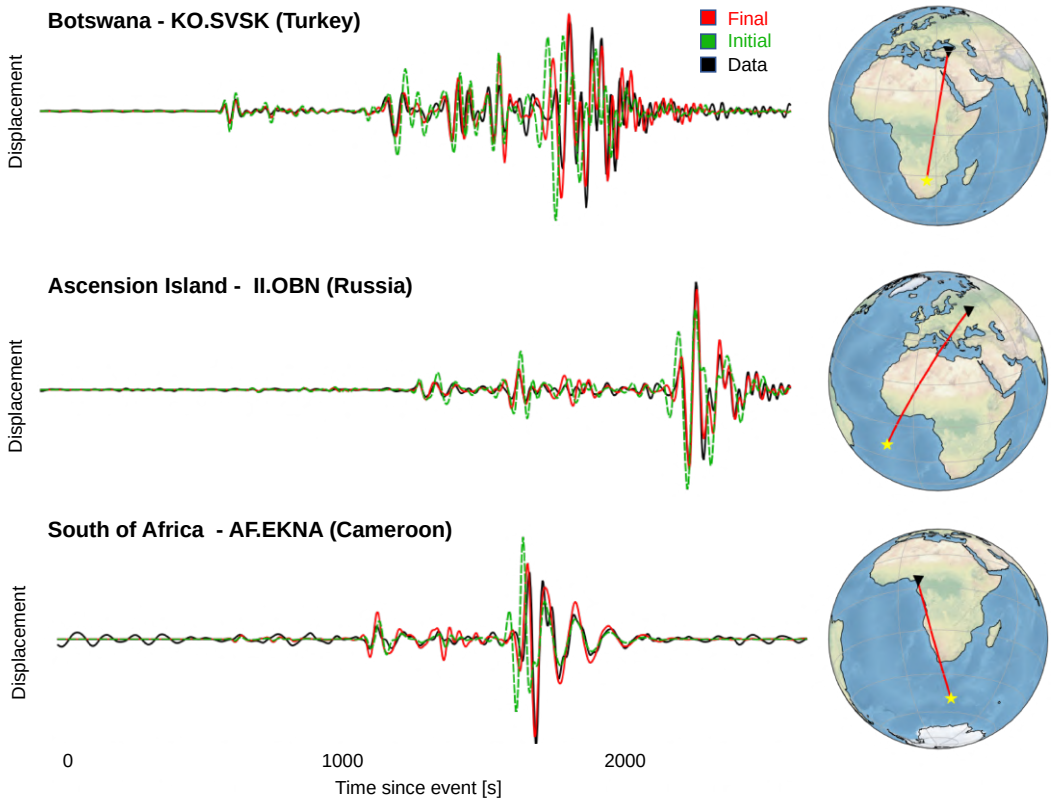

Figure 13. A selection of vertical-component seismograms from the validation dataset for source-receiver pairs sensitive to the crustal and mantle structure of the African Plate. **Top:** a magnitude 6.5 earthquake occurring on April 3rd, 2017, recorded in Turkey; **Middle:** a magnitude 6.0 earthquake occurring on February 2nd, 2017, recorded in Obninsk, Russia; **Bottom:** a magnitude 5.9 earthquake occurring south of the South Africa coast, recorded in Cameroon. Synthetic waveforms from the final model are shown in red, synthetic waveforms from the initial model are shown in green, and recorded data is shown in black. Source locations are indicated with yellow stars, and receiver locations are marked with black triangles.

eltime inversion, the impact of noise on the non-uniqueness of the inversion problem is nicely illustrated by Deal and Nolet (1996).

Another strategy consists in the use of point spread functions (PSFs) (e.g. Fichtner & Trampert, 2011). Here, the Hessian near the objective function's minimum is interpreted as the inverse of the covariance matrix, and hence as a conservative estimate of the PSF. The Hessian-vector multiplication can be approximated through finite differences of two gradients (e.g. Nocedal & Wright, 2006; Gao et al., 2021),

$$\mathbf{H}\delta\mathbf{m} = \frac{\mathbf{g}(\mathbf{m} + h\delta\mathbf{m}) - \mathbf{g}(\mathbf{m})}{h}, \tag{2}$$

where $\mathbf{H}$ is the Hessian, $\delta\mathbf{m}$ the model- or point perturbation, and $\mathbf{g}(\mathbf{m})$ is the gradient of the misfit function with respect to the model $\mathbf{m}$. This approximation becomes exact in the limit of h tending towards 0. Notice that at a minimum of the misfit function, the gradient becomes 0, and the Hessian-vector product gets simplified to the gradient with respect to the perturbation $\delta\mathbf{m}$.

Alternatively, one can utilize the L-BFGS approximation of the Hessian, based on the history of the model and gradient updates (e.g. Wehner et al., 2022). This approach requires no additional simulations at the cost of a loss in accuracy. In areas with no or only minor gradient information, the Hessian approximation remains identical to the initial guess, typically an identity matrix. As a result, the product of the approximate Hessian with the point perturbation becomes the point perturbation itself in such areas. This makes the interpretation rather difficult; the seemingly best results are achieved in the areas with the worst data coverage.

In this study, we use mini-batches, as described in Section 4. It brings up the additional question of how to compute the Hessian-vector products. On the one hand, using a constant data subset would mean that for the uncertainty analysis, we use different data than for the inversion itself. On the other hand, using all the data would be very expensive and not accurately represent the stochastic nature of the mini-batches.

6.1 Restitution test

To address some of the disadvantages of the approaches above, we propose another strategy. We deliberately perturb the final model or, in other words, introduce known errors into the model that we imaged with the inversion. We then continue the inversion as usual, performing iterations to reduce the misfit and correct the introduced errors. If the final model converged to a stable solution, and we can image the introduced perturbations, it should be possible to restore the model and return to the original unperturbed model. This approach has several advantages:

1) The final model becomes an integral part of the test.
2) We use the actual dataset with all its natural imperfections.
3) There is no need to create a costly separate synthetic dataset.
4) The dynamic mini-batches can be used in the same way as during the inversion; we do not need a special data subset or the entire dataset.
5) The non-linear nature of the full-waveform inversion problem is naturally considered.

For the tests presented in this paper, we applied Gaussian perturbations to all model parameters. The Gaussians had an amplitude of 7%, a standard deviation of 225 km, peaks positioned onto a grid with a step of approximately 6 standard deviations, based at 225 km depth, and summed together, forming a checkerboard pattern. We then performed additional 80 mini-batch iterations until validation misfits were approximately the same as for the previously shown final model, see Fig. 14. The test results for the v_{SV} parameter are summarized in Figs. 15 and 16 at 150 and 400 km depth, respectively.

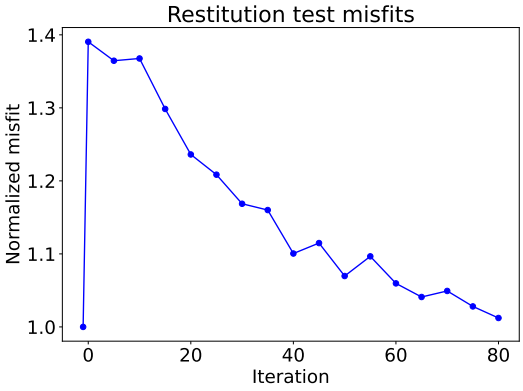

Figure 14. Validation event misfit for the 80 iterations within the restitution test. Initially, the misfit increases when we introduce errors by perturbing the model. Then we iterate until the misfits are approximately the same as for the final model.

The restitution test figures (Figs. 15 and 16) show the whole inverted domain, unlike the previous model figures. We also include the edges of the domain, where a lower quality of recovery can be expected. One can observe how the lack of ray coverage (crossings) translates into the inability to restore the model.

In the perturbation test at 150 km depth, the perturbed model restores almost perfectly to the original final model. That indicates the velocity structures are imaged with a high degree of certainty. At 400 km depth, however, we can see that the mismatch between the final and restored model is considerable. Hence, the model uncertainty at this depth is relatively higher, and one should interpret the model parameters with greater caution. In other words, the nullspace is larger, and it is possible to vary the model more while keeping very similar waveform misfits.

7 Discussion

We presented a full-waveform tomography of the African Plate. We believe these results are meaningful as many expected tectonic features are visible in the image. There are, however, external limiting factors that we can hardly influence. The most important one is the relatively small number of stations in the domain of interest. We must rely heavily on stations that are placed in Europe to make these results possible. Despite this fact, the results of the restitution test give us confidence that we get decent model quality even in the south with comparatively little data.

Note that we present a rather technical view of the model and its construction, together with some proxies for resolution. Translating this into inferences of, e.g., geodynamic processes, chemical composition, and temperature, and carefully propagating our uncertainties into uncertainties of these various inferences, will require collaboration with experts in these fields. For this, we will happily share our data.

In this study, we did not invert for the source parameters but assumed the parameters provided by the GCMT Catalog (Ekström et al., 2012) to be correct, since, e.g., Hjörleifsdóttir and Ekström (2010); Bozdağ et al. (2016) only found minor updates when doing a source inversion. Most earthquakes in our dataset occurred on the edge of the domain of interest, with most stations on one side of the event. This lack of azimuthal coverage makes it virtually impossible to find more accurate source locations. Neverthe-

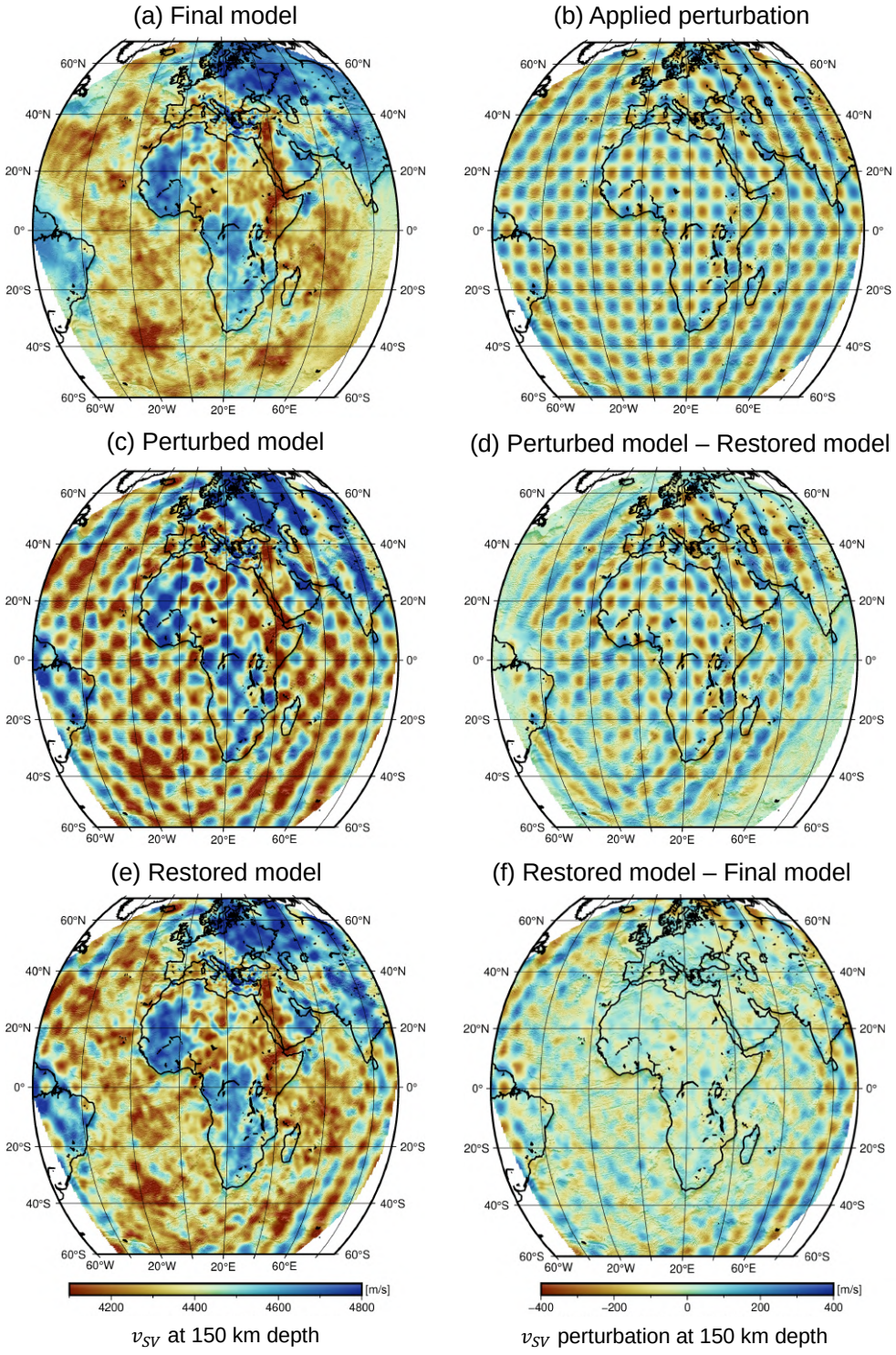

Figure 15. Restitution test for the v_{SV} parameter at 150 km depth. Panel **(a)** shows the final model after 130 mini-batch iterations, **(b)** shows the applied perturbation, **(c)** the model with the perturbation, **(e)** the model after the 80 repairing iterations, and **(d)** the difference between the perturbed and restored model. Finally, **(f)** shows the difference between the restored and the final model; this difference is relatively low in the domain of interest, implying the final model is reliable there.

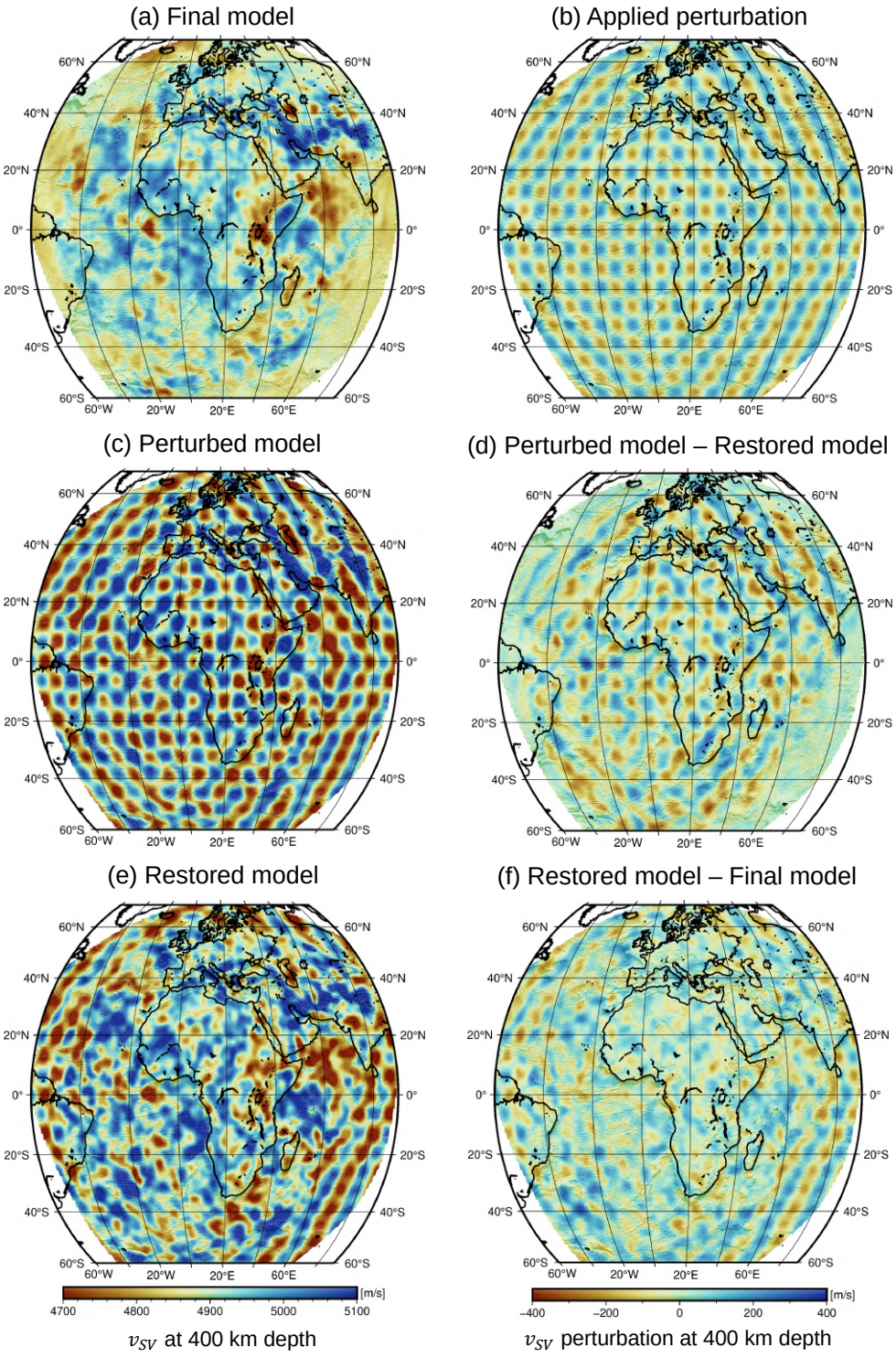

Figure 16. Restitution test for the v_{SV} parameter at 400 km depth. The description of the panels remains the same as in Fig. 15. Notice the restoration errors remain larger here than in Fig. 15. Panel **(d)** indicates that we can retrieve anomalies at the scale of the perturbations; however, the errors remain relatively large as seen in **(f)**. The misfits for models **(a)** and **(e)** are very similar. Therefore, the greater difference between models **(a)** and **(e)** at 400 km depth than that at 150 km depth (Fig. 15) suggests velocities are less tightly constrained and thus imaged with less certainty.

less, the mini-batch approach allows us to include many more events, so their location inaccuracies will likely average out more efficiently.

Several other studies focusing on the African upper mantle have been published before (e.g. Ritsema & van Heijst, 2000; Sebai et al., 2006; Pasyanos & Nyblade, 2007; Priestley et al., 2008; Fishwick, 2010; Emry et al., 2019; Celli et al., 2020). These models generally agree on large-scale features, such as low velocities beneath the EARS and high velocities beneath the major cratons. We compare our model to a recent one by Celli et al. (2020). Similar to our study, that model was produced using extensive data from the AfricaArray (Nyblade et al., 2011), and synthetic waveforms were used. In contrast to our study, waveforms were computed from normal mode summations rather than by solving the wave equation. Figs. 17 and 18 compare the isotropic shear wave velocity perturbations between the two models. For this purpose, we computed the Voigt averaged shear wave velocity, where $v_S = \sqrt{(2v_{SV}^2 + v_{SH}^2)/3}$.

Broadly, there is a notable similarity between the two models. It appears that some of the features in the model by Celli et al. (2020) show more resemblance to our model at an earlier stage of the inversion and longer periods (see Fig. 4). However, we retrieve more localized structures after additional iterations at shorter periods. We observe distinctive low-velocity anomalies beneath the Hoggar, Tibesti, and Aïr Mountains. These anomalies are not typically seen in other models of the African upper mantle, although they are expected beneath regions with dynamic topography. We also see the Cameroon Volcanic Line as a thinner, more defined low-velocity anomaly. This increased level of detail likely owes to the fact that we compute gradients with respect to an evolving 3D model rather than a constant 1D referential one, as is the case in the Automated Multimode Inversion (Lebedev et al., 2005) method used by Celli et al. (2020).

With the proposed restitution test in hand, it is possible to see where we can recover structure. At the same time, the lack of precision in recovering the final model after applying perturbations gives us an idea about the model uncertainty. This test is cheaper to compute than a traditional checkerboard test because we do not need to create a synthetic dataset first. No approximate versions of the real dataset are used, and we directly test the final model itself. In contrast to the Hessian-vector product approach, we can incorporate the non-linearity of the FWI imaging process.

The proposed restitution test is approximately equally expensive as running the inversion itself. While testing different model perturbations would be potentially interesting, it would also make the uncertainty analysis more costly. Doing so, however, might make it possible to obtain more quantitative estimates of model uncertainty. Perhaps, this could be an interesting subject of future research.

8 Conclusions

We presented a full-waveform tomography of the African continent using the dynamic mini-batch technique. To our knowledge, this is the most extensive continental-scale full-waveform inversion study performed for this area to date. We incorporated data from 184,356 unique source-receiver pairs and achieved significant improvements in waveform fits at 35 s, even for data we did not use within the inversion process.

Many significant features, such as the Afar Triple Junction, the Cameroon Volcanic Line, and mountain ranges of dynamic origin, are visible in the images and align well with the surface anomalies. Additionally, we imaged a low-velocity region beneath the EARS with low-velocity anomalies extending deeper down beneath Ethiopia and the Tanzania-Kenya border region.

The alignment with these tectonic features suggests that the imaged seismic structure is meaningful. However, our novel model uncertainty analysis indicates that we should

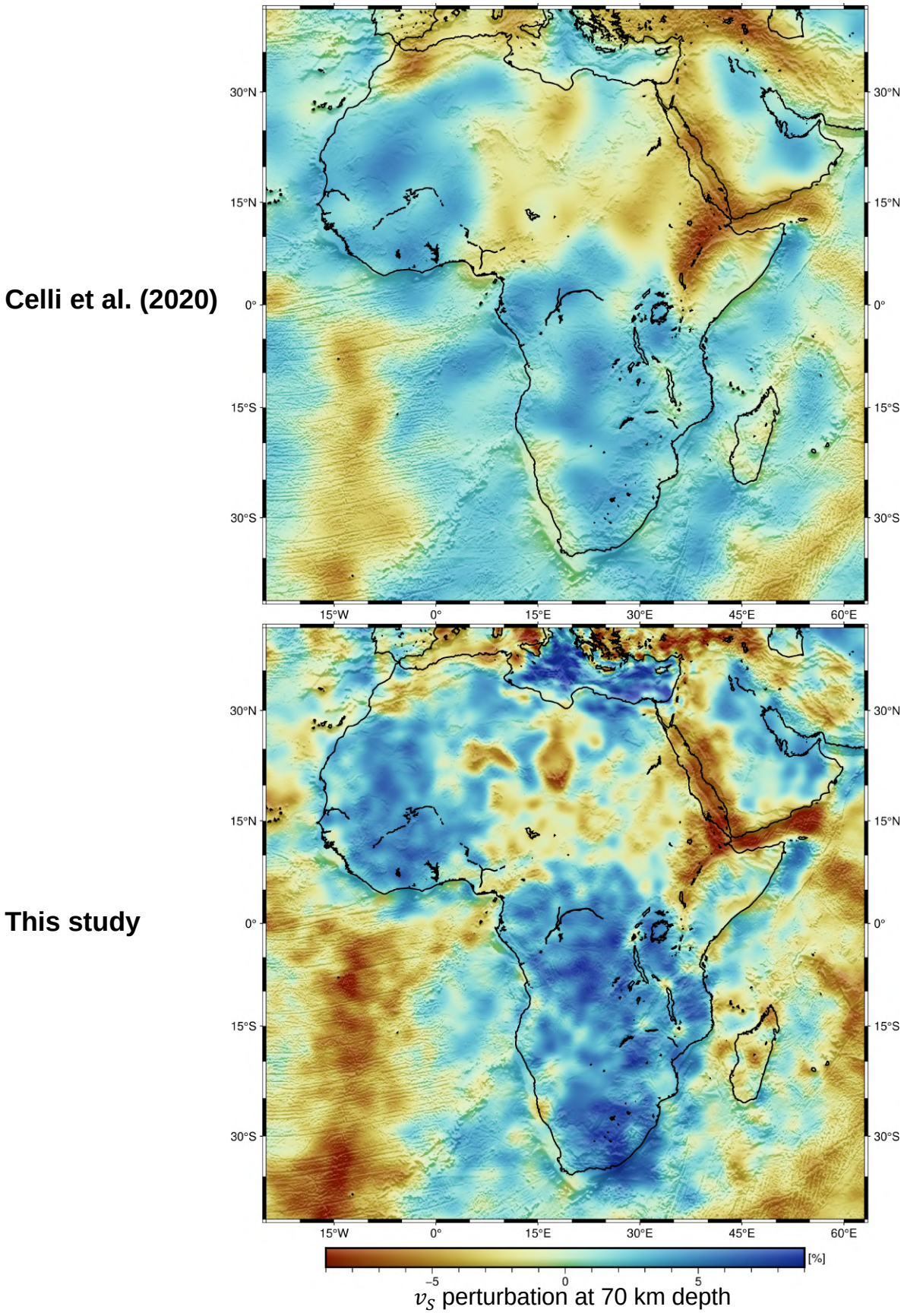

Figure 17. A comparison between the shear velocity model of Celli et al. (2020) and this study at 70 km depth. Perturbations are plotted relative to the lateral mean.

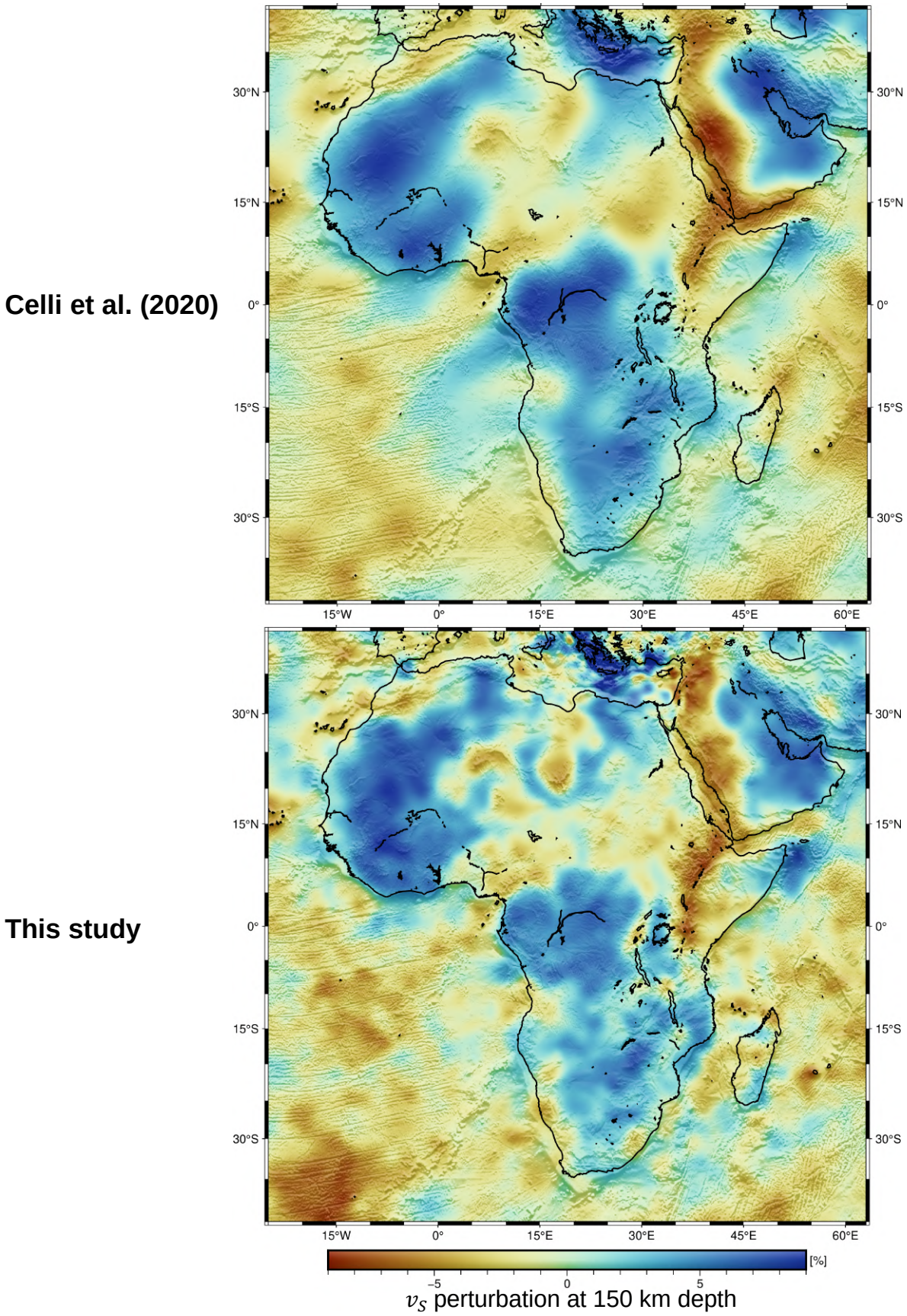

Figure 18. A comparison between the shear velocity model of Celli et al. (2020) and this study at 150 km depth. Perturbations are plotted relative to the lateral mean.

be careful when interpreting deeper-lying model attributes. We would gladly share the model and collaborate to get more in-depth interpretations of the tomographic results.

Acknowledgments

This work was supported by the European Research Council (ERC) under the EU's Horizon 2020 programme (grant no. 714069). We also gratefully acknowledge support from the Swiss National Supercomputing Centre (CSCS) in the form of computing time grant s1040. In addition, we would like to thank IRIS DMC, whose FDSN web services provided the majority of the data, as well as the operators of the NARS-Botswana deployment and AfricaArray. Also, we would like to thank Nicolas Celli for providing us with polygons of the Cratonic regions and scripts to visualize his model.

Open Research

Data Availability Statement

All seismic data is publically available and can be obtained through the FDSN web services (Romanowicz & Dziewonski, 1986). We used LASIF (Thrastarson, van Herwaarden, Krischer, & Fichtner, 2021) and ObsPy (Beyreuther et al., 2010; Megies et al., 2011; Krischer, Megies, et al., 2015) to download the seismic data. The model, which we refer to as AFR2022 will be made available on IRIS EMC (IRIS, 2011) upon publication. The majority of the figures were created with GMT (Wessel et al., 2019). The scientific colour maps that we used were created by Crameri (2021) and are available online.

References

Abdelsalam, M. G., Liégeois, J.-P., & Stern, R. J. (2002). The saharan metacraton. *Journal of African Earth Sciences*, *34*(3-4), 119–136.

Afanasiev, M., Boehm, C., van Driel, M., Krischer, L., Rietmann, M., May, D. A., ... Fichtner, A. (2019). Modular and flexible spectral-element waveform modelling in two and three dimensions. *Geophys. J. Int.*, *216*(3), 1675-1692. doi: 10.1093/gji/ggy469

Almalki, K. A., Betts, P. G., & Ailleres, L. (2016). Incipient seafloor spreading segments: insights from the red sea. *Geophysical Research Letters*, *43*(6), 2709–2715.

Barash, D. (2002). Fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(6), 844–847.

Bastow, I., Nyblade, A., Stuart, G., Rooney, T., & Benoit, M. (2008). Upper mantle seismic structure beneath the ethiopian hot spot: Rifting at the edge of the african low-velocity anomaly. *Geochemistry, Geophysics, Geosystems*, *9*(12).

Begg, G., Griffin, W., Natapov, L., O'Reilly, S. Y., Grand, S., O'Neill, C., ... others (2009). The lithospheric architecture of africa: Seismic tomography, mantle petrology, and tectonic evolution. *Geosphere*, *5*(1), 23–50.

Benoit, M. H., Nyblade, A. A., Owens, T. J., & Stuart, G. (2006). Mantle transition zone structure and upper mantle s velocity variations beneath ethiopia: Evidence for a broad, deep-seated thermal anomaly. *Geochemistry, Geophysics, Geosystems*, *7*(11).

Bernal-Romero, M., & Iturrarán-Viveros, U. (2021). Accelerating full-waveform inversion through adaptive gradient optimization methods and dynamic simultaneous sources. *Geophysical Journal International*, *225*(1), 97–126.

Beyreuther, M., Barsch, R., Krischer, L., & Wassermann, J. (2010). ObsPy: A Python toolbox for seismology. *Seis. Res. Lett.*, *81*, 47-58.

Bird, P. (2003). An updated digital model of plate boundaries. *Geochemistry, Geo-*

physics, Geosystems, *4*(3).

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

Bozdağ, E., Peter, D., Lefebvre, M., Tromp, J., Komatitsch, D., Pugmire, D., ... Podhorszki, N. (2016). Global adjoint tomography: First-generation model. *Geophys. J. Int.*, *207*(3), 1739-1766. doi: 10.1093/gji/ggw356

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

Burke, K., & Gunnell, Y. (2008). *The african erosion surface: a continental-scale synthesis of geomorphology, tectonics, and environmental change over the past 180 million years* (Vol. 201). Geological Society of America.

Celli, N. L., Lebedev, S., Schaeffer, A. J., & Gaina, C. (2020). African cratonic lithosphere carved by mantle plumes. *Nature communications*, *11*(1), 1–10.

Chaljub, E., Capdeville, Y., & Vilotte, J. (2003). Solving elastodynamics in a solid heterogeneous 3-Sphere: a spectral element approximation on geometrically non-conforming grids. *J. Comp. Physics*, *183*, 457-491.

Chorowicz, J. (2005). The east african rift system. *Journal of African Earth Sciences*, *43*(1-3), 379–410.

Colli, L., Fichtner, A., & Bunge, H.-P. (2013). Full waveform tomography of the upper mantle in the South Atlantic region: Imaging westward fluxing shallow asthenosphere? *Tectonophysics*, *604*, 26-40.

Crameri, F. (2021, February). *Scientific colour maps.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.4491293` doi: 10.5281/zenodo.4491293

Deal, M. M., & Nolet, G. (1996). Nullspace shuttles. *Geophysical Journal International*, *124*(2), 372–380.

Dziewoński, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. *Phys. Earth Planet. Inter.*, *25*(4), 297–356. doi: 10.1016/0031-9201(81)90046-7

Ebinger, C. J., & Sleep, N. (1998). Cenozoic magmatism throughout east africa resulting from impact of a single plume. *Nature*, *395*(6704), 788–791.

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

Emry, E. L., Shen, Y., Nyblade, A. A., Flinders, A., & Bao, X. (2019). Upper mantle earth structure in africa from full-wave ambient noise tomography. *Geochemistry, Geophysics, Geosystems*, *20*(1), 120–147.

Faccioli, E., Maggio, F., Quarteroni, A., & Tagliani, A. (1996). Spectral-domain decomposition methods for the solution of acoustic and elastic wave equations. *Geophysics*, *61:4*, 1160-1174.

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

Fichtner, A., Bunge, H.-P., & Igel, H. (2006). The adjoint method in seismology - I. Theory. *Phys. Earth Planet. Inter.*, *157*, 86-104. doi: 10.1016/j.pepi.2006.03.016

Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H.-P. (2008). Theoretical background for continental- and global-scale full-waveform inversion in the time-frequency domain. *Geophys. J. Int.*, *175*(2), 665-685. doi: 10.1111/j.1365-246X.2008.03923.x

Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H.-P. (2009). Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophys. J. Int.*, *179*(3), 1703-1725. doi: 10.1111/j.1365-246X.2009.04368.x

Fichtner, A., Saygin, E., Taymaz, T., Cupillard, P., Capdeville, Y., & Trampert, J. (2013). The deep structure of the North Anatolian Fault Zone. *Earth Planet.*

Sci. Lett., *373*, 109-117. doi: 10.1016/j.epsl.2013.04.027

Fichtner, A., Stehly, L., Ermert, L., & Boehm, C. (2016, 11). Generalized interferometry – I: theory for interstation correlations. *Geophysical Journal International*, *208*(2), 603-638. Retrieved from `https://doi.org/10.1093/gji/ggw420` doi: 10.1093/gji/ggw420

Fichtner, A., & Trampert, J. (2011). Resolution analysis in full waveform inversion. *Geophys. J. Int.*, *187*, 1604-1624.

Fichtner, A., Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y., & Villasenor, A. (2013). Multiscale full waveform inversion. *Geophysical Journal International*, *194*(1), 534–556.

Fichtner, A., van Herwaarden, D.-P., Afanasiev, M., Simutė, S., Krischer, L., Çubuk Sabuncu, Y., ... Igel, H. (2018). The Collaborative Seismic Earth Model: Generation 1. *Geophys. Res. Lett.*, *45*(9), 4007-4016. doi: 10.1029/2018GL077338

Fichtner, A., van Herwaarden, D.-P., Afanasiev, M., Simute, S., Krischer, L., Ã‡ubuk Sabuncu, Y., ... Igel, H. (2018). The Collaborative Seismic Earth Model: Generation 1. *Geophys. Res. Lett.*, *45*(9), 4007-4016. doi: 10.1029/2018GL077338

Fichtner, A., & Villaseñor, A. (2015). Crust and upper mantle of the western Mediterranean – Constraints from full-waveform inversion. *Earth Planet. Sci. Lett.*, *428*, 52-62. doi: 10.1016/j.epsl.2015.07.038

Fishwick, S. (2010). Surface wave tomography: imaging of the lithosphere–asthenosphere boundary beneath central and southern africa? *Lithos*, *120*(1-2), 63–73.

Fishwick, S., & Bastow, I. D. (2011). Towards a better understanding of african topography: a review of passive-source seismic studies of the african crust and upper mantle. *Geological Society, London, Special Publications*, *357*(1), 343–371. doi: 10.1144/SP357.19

Furman, T. (2007). Geochemistry of east african rift basalts: an overview. *Journal of African Earth Sciences*, *48*(2-3), 147–160.

Gao, Y., Tilmann, F., van Herwaarden, D.-P., Thrastarson, S., Fichtner, A., Heit, B., ... Schurr, B. (2021). Full waveform inversion beneath the central andes: Insight into the dehydration of the nazca slab and delamination of the back-arc lithosphere. *JGR: Solid Earth*, *126*.

Gebraad, L., Boehm, C., & Fichtner, A. (2020). Bayesian elastic full-waveform inversion using hamiltonian monte carlo. *Journal of Geophysical Research: Solid Earth*, *125*(3), e2019JB018428. doi: 10.1029/2019JB018428

Hjörleifsdóttir, V., & Ekström, G. (2010). Effects of three-dimensional earth structure on cmt earthquake parameters. *Physics of the Earth and Planetary Interiors*, *179*(3-4), 178–190.

Hoffer, E., Hubara, I., & Soudry, D. (2017). Train longer, generalize better: closing the generalization gap in large batch training of neural networks. *Advances in neural information processing systems*, *30*.

IRIS, D. (2011). *Data services products: Emc, a repository of earth models.* doi: https://doi.org/10.17611/DP/EMC.1.

Komatitsch, D., & Tromp, J. (1999). Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophys. J. Int.*, *139*, 806-822.

Komatitsch, D., & Tromp, J. (2002). Spectral-element simulations of global seismic wave propagation, Part II: 3-D models, oceans, rotation, and gravity. *Geophys. J. Int.*, *150*(1), 303-318. doi: 10.1046/j.1365-246X.2002.01716.x

Krischer, L., Fichtner, A., Boehm, C., & Igel, H. (2018). Automated Large-Scale Full Seismic Waveform Inversion for North America and the North Atlantic. *J. Geophys. Res.: Solid Earth*, *123*(7), 5902-5928.

Krischer, L., Fichtner, A., Zukauskaite, S., & Igel, H. (2015). Large-scale seismic in-

version framework. *Seismological Research Letters*, *86*(4), 1198–1207.
Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., & Wassermann, J. (2015). ObsPy: a bridge for seismology into the scientific Python ecosystem. *Computational Science & Discovery*, *8*(1), 014003.
Lebedev, S., Nolet, G., Meier, T., & Van Der Hilst, R. D. (2005). Automated multimode inversion of surface and s waveforms. *Geophysical Journal International*, *162*(3), 951–964.
Liégeois, J.-P., Abdelsalam, M. G., Ennih, N., & Ouabadi, A. (2013). Metacraton: nature, genesis and behavior. *Gondwana Research*, *23*(1), 220–237.
Liu, D. C., & Nocedal, J. (1989). On the limited-memory BFGS method for large-scale optimisation. *Mathematical Programming*, *45*, 503-528.
Liu, Q., & Gu, Y. (2012). Seismic imaging: From classical to adjoint tomography. *Tectonophysics*, *566-567*, 31-66. doi: https://doi.org/10.1016/j.tecto.2012.07.006
Megies, T., Beyreuther, M., Barsch, R., Krischer, L., & Wassermann, J. (2011). Obspy - what can it do for data centers and observatories? *Annals Geophys.*, *54*, 47-58.
Meier, U., Curtis, A., & Trampert, J. (2007). Fully nonlinear inversion of fundamental mode surface waves for a global crustal model. *Geophys. Res. Lett.*, *34*, doi:10.1029/2007GL030989.
Mulibo, G. D., & Nyblade, A. A. (2013). The p and s wave velocity structure of the mantle beneath eastern africa and the african superplume anomaly. *Geochemistry, Geophysics, Geosystems*, *14*(8), 2696–2715.
NCEI. (2022). *NCEI Global Historical Hazard Database.* Retrieved from `https://www.ngdc.noaa.gov/hazel/view/hazards/volcano/`
Nocedal, J., & Wright, S. (2006). *Numerical optimization.* Springer Science & Business Media. doi: 10.1007/978-0-387-40065-5
Nyblade, A. A., Durrheim, R., Dirks, P., Graham, G., Gibson, R., & Webb, S. (2011). Geoscience initiative develops sustainable science in africa. *Eos, Transactions American Geophysical Union*, *92*(19), 161–162.
Pasyanos, M. E., & Nyblade, A. A. (2007). A top to bottom lithospheric study of africa and arabia. *Tectonophysics*, *444*(1-4), 27–44.
Patera, A. T. (1984). A spectral element method for fluid dynamics: laminar flow in a channel expansion. *J. Comput. Phys.*, *54*, 468-488.
Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model. *Geophysics*, *64*(3), 888-901. doi: 10.1190/1.1444597
Priestley, K., McKenzie, D., Debayle, E., & Pilidou, S. (2008). The african upper mantle and its relationship to tectonics and surface geology. *Geophysical Journal International*, *175*(3), 1108–1126.
Rickers, F., Fichtner, A., & 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 Planet. Sci. Lett.*, *367*, 39-51.
Ritsema, J., & van Heijst, H. J. (2002). Constraints on the correlation of P- and S-wave velocity heterogeneity in the mantle from P, PP, PPP and PKPab traveltimes. *Geophys. J. Int.*, *149*(2), 482-489. doi: 10.1046/j.1365-246X.2002.01631.x
Ritsema, J., van Heijst, H. J., & Woodhouse, J. H. (1999). Complex shear wave velocity structure imaged beneath Africa and Iceland. *Science*, *286*(5446), 1925-1928. doi: 10.1126/science.286.5446.1925
Ritsema, J., & van Heijst, H. (2000). New seismic model of the upper mantle beneath africa. *Geology*, *28*(1), 63–66.
Robertsson, J. O. A., Blanch, J. O., & Symes, W. W. (1994). Viscoelastic finite-difference modelling. *Geophysics*, *59*, 1444-1456.
Romanowicz, B. A., & Dziewonski, A. M. (1986). Toward a federation of broadband

seismic networks. *Eos, Transactions American Geophysical Union*, *67*(25), 541-542. doi: 10.1029/EO067i025p00541

Ruan, Y., Lei, W., Modrak, R., Ã–rsvuran, R., BozdaÄŸ, E., & Tromp, J. (2019, 08). Balancing unevenly distributed data in seismic tomography: a global adjoint tomography example. *Geophysical Journal International*, *219*(2), 1225-1236. doi: 10.1093/gji/ggz356

Sager, K., Boehm, C., Ermert, L., Krischer, L., & Fichtner, A. (2020). Global-scale full-waveform ambient noise inversion. *Journal of Geophysical Research: Solid Earth*, *125*(4), e2019JB018644.

Sager, K., Ermert, L., Boehm, C., & Fichtner, A. (2018). Towards full waveform ambient noise inversion. *Geophys. J. Int.*, *212*(1), 566-590. doi: 10.1093/gji/ggx429

Sebai, A., Stutzmann, E., Montagner, J.-P., Sicilia, D., & Beucler, E. (2006). Anisotropic structure of the african upper mantle from rayleigh and love wave tomography. *Physics of the Earth and Planetary Interiors*, *155*(1-2), 48–62.

Takeuchi, H., & Saito, M. (1972). Seismic surface waves. *Methods in computational physics*, *11*, 217–295.

Tape, C., Liu, Q., Maggi, A., & Tromp, J. (2009). Adjoint tomography of the southern California crust. *Science*, *325*(5943), 988-992. doi: 10.1126/science.1175298

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

Tarantola, A. (1988). Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation. *Pure Appl. Geophys.*, *128*, 365-399. doi: 10.1007/978-3-0348-7722-0_19

Tariq, Q., Daniels, J., Schwartz, J. N., Washington, P., Kalantarian, H., & Wall, D. P. (2018). Mobile detection of autism through machine learning on home video: A development and prospective validation study. *PLoS medicine*, *15*(11), e1002705.

Thrastarson, S., van Herwaarden, D. P., & Fichtner, A. (2021). Inversionson: Fully automated seismic waveform inversions. *EarthArxiv*.

Thrastarson, S., van Herwaarden, D.-P., Krischer, L., & Fichtner, A. (2021). Lasif: Large-scale seismic inversion framework, an updated version. *EarthArXiv*.

Thrastarson, S., van Herwaarden, D.-P., Krischer, L., Boehm, C., van Driel, M., Afanasiev, M., & Fichtner, A. (2022, 03). Data-adaptive global full-waveform inversion. *Geophysical Journal International*, *230*(2), 1374-1393. doi: 10.1093/gji/ggac122

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

Utrecht University. (2013). *Nars.* International Federation of Digital Seismograph Networks. Retrieved from `http://www.fdsn.org/doi/10.7914/SN/NR` doi: 10.7914/SN/NR

Vallée, M. (2013). Source time function properties indicate a strain drop independent of earthquake depth and magnitude. *Nature Communications*, *4*(1), 2606.

van Herwaarden, D. P., Afanasiev, M., Thrastarson, S., & Fichtner, A. (2021). Evolutionary full-waveform inversion. *Geophysical Journal International*, *224*(1), 306–311.

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

Virieux, J., & Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, *74*, WCC127-WCC152.

Wapenaar, K. (2004, Dec). Retrieving the elastodynamic green's function of an

arbitrary inhomogeneous medium by cross correlation. *Phys. Rev. Lett.*, *93*, 254301. doi: 10.1103/PhysRevLett.93.254301

Wehner, D., Blom, N., Rawlinson, N., 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*, e2021JB022930.

Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5556–5564.

Zhang, Z., & Sabuncu, M. R. (2018). Generalized cross entropy loss for training deep neural networks with noisy labels. *arXiv preprint arXiv:1805.07836*.