Full-Waveform Tomography of the African Plate Using Dynamic Mini-Batches

Author(s):
van Herwaarden, Dirk-Philip; Thrastarson, Sölvi; Hapla, Václav; Afanasiev, Michael; Trampert, Jeannot; Fichtner, Andreas

Publication date:
2023-06

Permanent link:
https://doi.org/10.3929/ethz-b-000618529

Originally published in:
Journal of Geophysical Research: Solid Earth 128(6), https://doi.org/10.1029/2022JB026023

JGR Solid Earth

RESEARCH ARTICLE
10.1029/2022JB026023

Full-Waveform Tomography of the African Plate Using Dynamic Mini-Batches

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, Zürich, Switzerland, [2]Now at Mondaic Ltd, Zürich, Switzerland, [3]Department of Earth Sciences, Utrecht University, Utrecht, The Netherlands

Key Points:

- We invert for crustal and mantle structure of the African Plate using recordings of 397 earthquakes down to periods of 35 s
- 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 final model and continue the inversion

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D.-P. van Herwaarden,
dirkphilip.vanherwaarden@erdw.ethz.ch

Citation:

van Herwaarden, D.-P., Thrastarson, S., Hapla, V., Afanasiev, M., Trampert, J., & Fichtner, A. (2023). Full-waveform tomography of the African Plate using dynamic mini-batches. *Journal of Geophysical Research: Solid Earth*, *128*, e2022JB026023. https://doi.org/10.1029/2022JB026023

Received 10 NOV 2022
Accepted 21 MAY 2023

Author Contributions:

Conceptualization: Dirk-Philip van Herwaarden, Andreas Fichtner
Data curation: Dirk-Philip van Herwaarden, Jeannot Trampert
Funding acquisition: Andreas Fichtner
Methodology: Dirk-Philip van Herwaarden
Project Administration: Andreas Fichtner
Resources: Dirk-Philip van Herwaarden
Software: Dirk-Philip van Herwaarden, Solvi Thrastarson, Vaclav Hapla, Michael Afanasiev
Supervision: Andreas Fichtner
Validation: Dirk-Philip van Herwaarden

Abstract We present a seismic model of the Africa Plate, constructed with the technique of full-waveform inversion. The purpose of our model is to serve as a foundation for quantitative geodynamic and geochemical interpretation, earthquake-induced ground motion predictions, and earthquake source inversion. Starting from the first-generation Collaborative Seismic Earth Model, 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. We use dynamically changing mini-batches of the complete data set to compute approximate gradients at each iteration. This approach has three significant advantages: (a) it reduces computational costs for model updates and the inversion, (b) it enables the use of larger datasets without increasing iteration costs, and (c) it makes it trivial to assimilate new data since we can extend the data set without changing the misfit function. We invert data from 397 unique earthquakes and 184,356 unique source-receiver pairs. We clearly image tectonic features such as the Afar triple junction and 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 In this paper, we contribute to the knowledge about geological structures under the African continent. For that, we use publicly available seismograms (records of the ground motion at seismic stations) measured in Africa and the regions around it. These we process with an advanced imaging method called full-waveform inversion (FWI). The result is a model of the speed of seismic waves propagating through the crust and upper mantle under Africa. We can use the model to generate maps of Earth's interior temperature anomalies in the given region. We can, for instance, clearly see hot structures beneath mountain ranges and volcanoes. Our FWI implementation benefits from several recent advances, allowing us to include a much higher number of seismograms than before, resulting in models with unprecedented resolution. We update the publicly available Collaborative Seismic Earth Model with those outputs. Finally, we propose a new strategy to assess the resolution and reliability of our resulting models.

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 challenging to assess mainly due to the relatively sparse station coverage (International Seismological Centre, 2023). This situation has been substantially improved by the AfricaArray (Nyblade et al., 2011) and our study benefits from it greatly.

In addition to a relative lack of stations in Africa compared to other continents, earthquakes providing a sufficient signal-to-noise ratio for waveform tomography (with magnitudes of around five and more) are uncommon here. These factors make it challenging to perform high-quality seismic tomography 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., Celli et al., 2020; Fishwick, 2010; Pasyanos & Nyblade, 2007; Priestley et al., 2008; Ritsema & van Heijst, 2000; Sebai et al., 2006) 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

Visualization: Dirk-Philip van Herwaarden
Writing – original draft: Dirk-Philip van Herwaarden
Writing – review & editing: Dirk-Philip van Herwaarden, Solvi Thrastarson, Vaclav Hapla, Jeannot Trampert, Andreas Fichtner

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 Western and Central Europe, the U.S.A., 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., Fichtner, 2010; Liu & Gu, 2012; Virieux & Operto, 2009). 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., Fichtner et al., 2006; Tarantola, 1984; Tarantola, 1988; Tromp et al., 2005). 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.

The purpose of this manuscript is to describe the construction of the model and its quality assessment that enables its use in applications that go beyond the standard qualitative interpretations of slow and fast regions. A primary motivation for constructing this model is to update the Collaborative Seismic Earth Model (CSEM) (Fichtner et al., 2018), which serves as a framework to integrate information from various tomographic studies at different scales. This model can then serve as a reference for other researchers, as a base for further, more interpretative studies, and as an input for applications like geodynamic data assimilation, earthquake-induced ground motion predictions, and earthquake source inversion.

The manuscript is structured as follows. In Section 2, we introduce the geologic background of the study region; Section 3 shows the used data set; 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

This section provides a brief overview of tectonic features that may be considered of broader interest and will appear again later when we roughly interpret the model in terms of regional geology. Figure 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 (e.g., Begg et al., 2009; De Waele et al., 2008; Jessell et al., 2016). On the other hand, Africa also contains more volcanically and tectonically active regions.

Significant hotspot areas in the Atlantic Ocean include Cape Verde or the Walvis Ridge (Duncan & Richards, 1991) and the Comoro Islands in the Indian Ocean (Nougier et al., 1986). In the northwest of the continent itself, the Atlas Mountains are located, separating the Sahara Desert from the Mediterranean Sea (Gomez et al., 2000). 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 (e.g., Franz et al., 1994; J. Liégeois et al., 2005). Eastern Africa is home to the Afar Triple junction (Keir et al., 2013) and the EARS (Chorowicz, 2005). 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, for example, São Tomé and Principe (Dedzo et al., 2013; Fitton, 1987). In the south, we find a large region of significant topography at the Southern African Plateau, surrounded by the Great Escarpment (ten Brink & Stern, 1992). 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, for example, by Fishwick and Bastow (2011) suggest that parts of Northern Africa have lower than

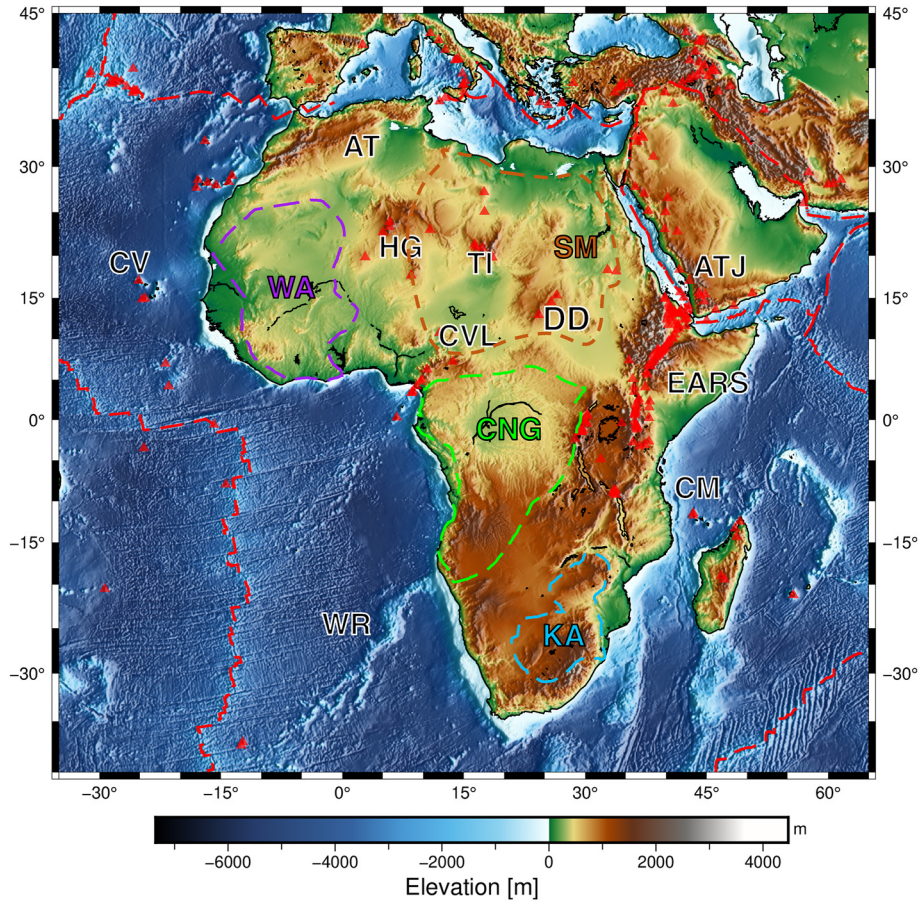

Figure 1. Elevation plot with volcanoes marked as red triangles. Major plate boundaries (Bird, 2003) are marked with a red dashed line. The main cratons (Begg et al., 2009) are indicated on the map: WA: West African Craton, KA: Kalahari Craton, CNG: Congo Craton. SM: Saharan Metacraton. Other features: 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).

average mantle velocities and that several high-velocity regions may mark the locations of cratonic fragments within the Saharan Metacraton (Abdelsalam et al., 2002; J.-P. 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; for example, 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

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 5.5–6.7. Empirically, this range provides a good signal while minimizing the finite-source effects associated with large earthquakes (Vallée, 2013).

To deal with the sparsity of seismic data on the African continent, we extend the modeled domain to use earthquakes around the entire African Plate and exploit waveform recordings in the much more densely instrumented European countries. Our data set comprises 184,356 three-component seismograms, which were recorded by 6,088 distinct receivers. Figure 2a shows the source-receiver path coverage plot of the entire data set.

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 data set, we also consider a separate one that we refer to as the validation data set. 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 data set is shown in Figure 2b and discussed in more detail in the next section.

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., Chaljub et al., 2003; Faccioli et al., 1996; Komatitsch & Tromp, 1999). The implementation allows for ocean loading (Komatitsch & Tromp, 2002), viscoelasticity (e.g., Robertsson et al., 1994), the ellipticity of the Earth, and topography and bathymetry. We include all these effects in our modeling. Simply applying a load for the oceans, rather than modeling the liquid ocean layer has been shown to sufficiently approximate the effect of the oceans at periods below 20 s (Wehner, Rawlinson, et al., 2022).

Our mesh extends down to the outer core, as we are primarily interested in surface and body waves from source-receiver pairs with limited epicentral distances. We add sponge layers of 500 km thickness at the edges of the domain, except for the surface, where we impose a free-surface boundary condition. These serve as absorbing boundaries. 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., Fichtner et al., 2006; Tarantola, 1984; Tarantola, 1988; Tromp et al., 2005).

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 data set. The gradient associated with each iteration is then only computed for this subset. The subset size may change between iterations, depending on the estimated redundancy in the mini-batches. This redundancy is estimated by measuring the difference between the mini-batch gradient and a smaller subset of events in the same mini-batch. 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 upper bound 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 data set 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 (Nocedal & Wright, 2006) 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.

4.1. Starting Model

The starting model is the first generation of the transverse isotropic CSEM (Fichtner et al., 2018). The CSEM is a framework to sequentially update an initial global background model with models of increasingly more

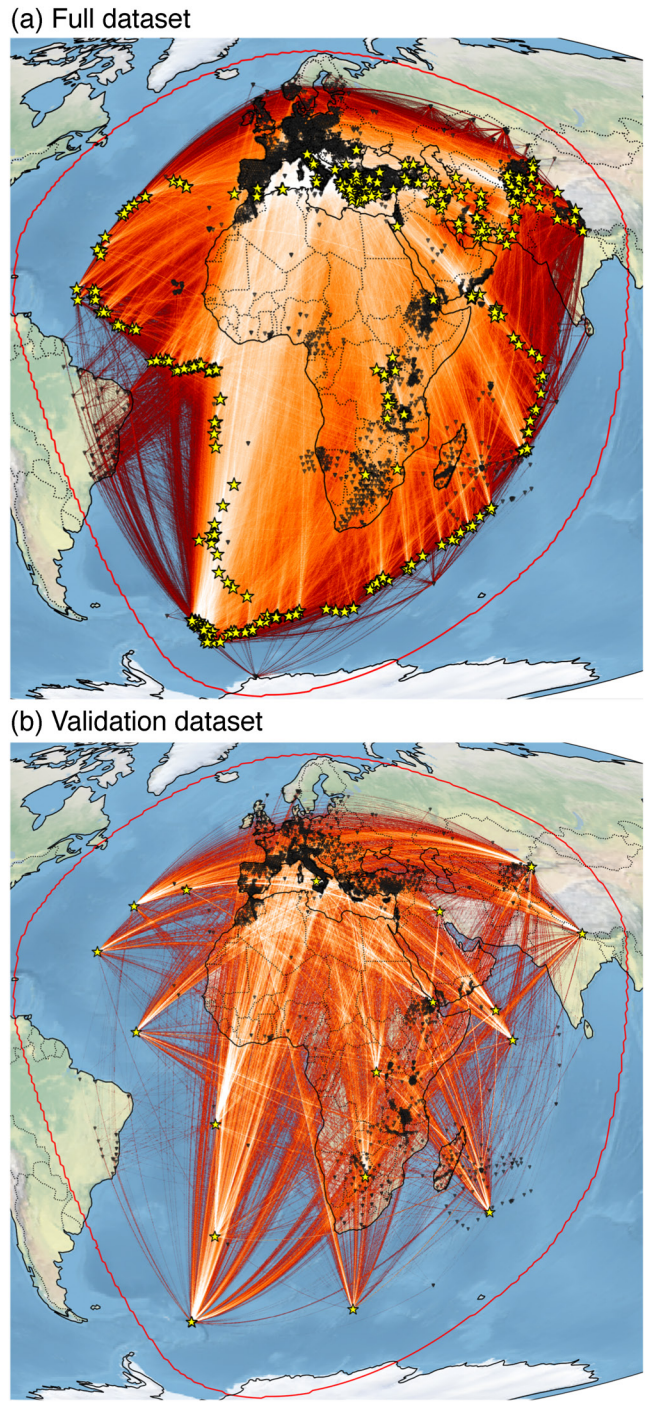

Figure 2. Source-receiver path coverage plots of the datasets used in this study. Bright colors represent a higher density of paths. Yellow stars represent earthquake locations, and black marks denote station locations. The red line represents the edge of the modeled domain. (a) The full data set contains 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 data set 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 data set to measure improvements in waveform fit. Note that although we show source-receiver paths, this plot is only a visual representation of data coverage and does not account for the finite extent of sensitivity kernels or the use of multiple wave types per source-receiver pair.

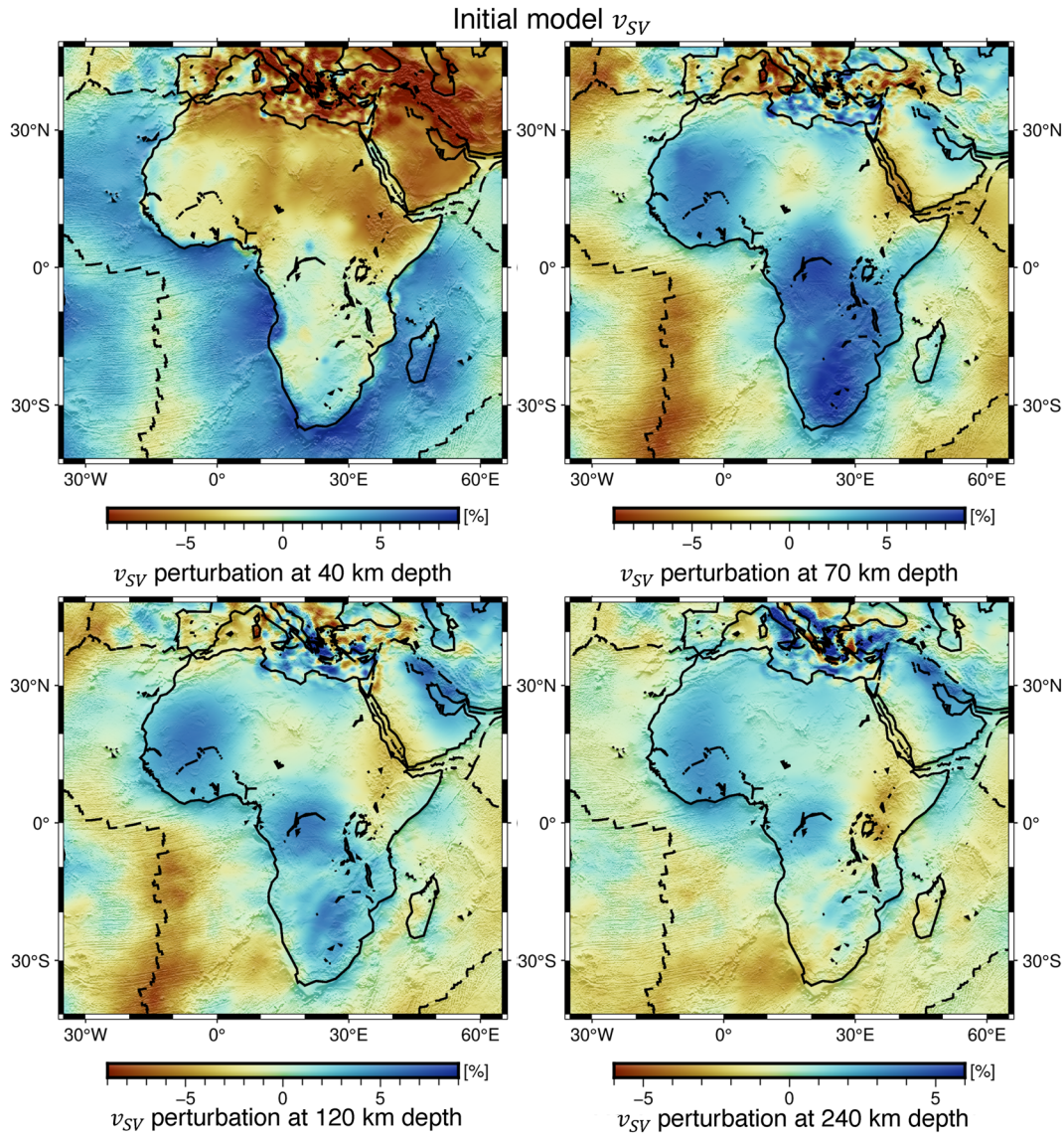

Figure 3. The starting model of the v_{SV} parameter at 40, 70, 120, and 240 km depth. The main plate boundaries (Bird, 2003) are plotted on top with dashed black lines. Perturbations are plotted relative to the lateral means.

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 simulation domain that we use, regional updates have been made to the CSEM 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 Figure 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, that is, 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) give a more detailed explanation of station weighting methods in regional to global scale FWI.

In addition to the misfits and gradients used for the optimization process, we also compute misfits for the validation data set shown in Figure 2b. We calculate no gradients for this data set 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. This 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 data set serves multiple purposes. First, it enables us to have a measure for convergence since a vanishing misfit decrease suggests we approach a minimum. Second, a strategy like this may help prevent us from overfitting our model to the data; if the misfit decreases for the full data set 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 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

We split the inversion into two phases of varying period bands. In the first phase of the inversion, we use a period band of 55–130. After convergence, we move on to the second phase where we use a broader period band of 35–130 s and iterate until convergence again. 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. This is a standard procedure used also in for example, Bunks et al. (1995), Krischer et al. (2018), and Blom et al. (2020).

We regularize the model by smoothing the gradients using the diffusion equation. This technique was popularized 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 the gradient 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.

We then compute the model updates based on these smoothed gradients and apply smooth updates to both the crust and the mantle. In contrast to classical approaches where static crustal corrections have been applied, we treat shallow structures the same as the mantle and allow them to change as requested by the data. Though the resulting shallow structure may not be highly resolved, this approach importantly ensures that inaccuracies of the initial crustal structure do not leave artifacts in the mantle structure.

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, for example, station- and event information, we need to keep track of heavy data such as synthetic traces, measurement windows, adjoint sources, and forward wavefield snapshots. 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). We show the v_{SV} parameter as it is 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 this parameter, which we confirm in later model uncertainty tests. Figure 4 show how the v_{SV} model changes as a function of the iteration number. We can see that velocity anomalies intensify throughout the inversion. The data is plotted as a perturbation to the lateral mean velocities. For reference, Figure 5 shows the lateral mean velocities for the top 250 km for the initial and final model. The evolution and the final model of the v_{SH} parameter are shown in Figures S3, S4, and S5 in Supporting Information S1 because the results are visually very similar to v_{SV}.

5.1. Final Model

Figure 6 shows the final v_{SV} velocity model at a selection of depths (40, 70, 120, and 240 km). Figure 7 focuses on the final v_{SV} model at 70 km depth, with the location of several tectonic features plotted on top.

Several mountain ranges, such as the Hoggar, Aïr, and Tibesti Mountains, believed to be of volcanic origin (Permenter & Oppenheimer, 2007; Schlüter, 2008), correspond clearly to low-velocity anomalies, especially at 70 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 (Figure 3). Furthermore, we observe relatively high velocities in the region of the Saharan Metacraton that Abdelsalam et al. (2002) hypothesized to be home to cratonic fragments. The West African, Congo, and Kalahari Craton areas also show high-velocity anomalies.

In Figure 8, we show cross-sections through the East African Rift System and the West African Craton. 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., Boyce et al., 2021; Ebinger & Sleep, 1998; Moucha & Forte, 2011; Ritsema et al., 1999; Thrastarson et al., 2022), and our model aligns with those findings.

Note that our model improves the CSEM beyond the African continent itself even in areas already processed in earlier works. For instance, the CSEM in the Eastern Mediterranean has already been updated with an inversion by Blom et al. (2020). Our model updates deeper sections and improves the definition of the subduction zone. This can be seen in Figure S6 in Supporting Information S1.

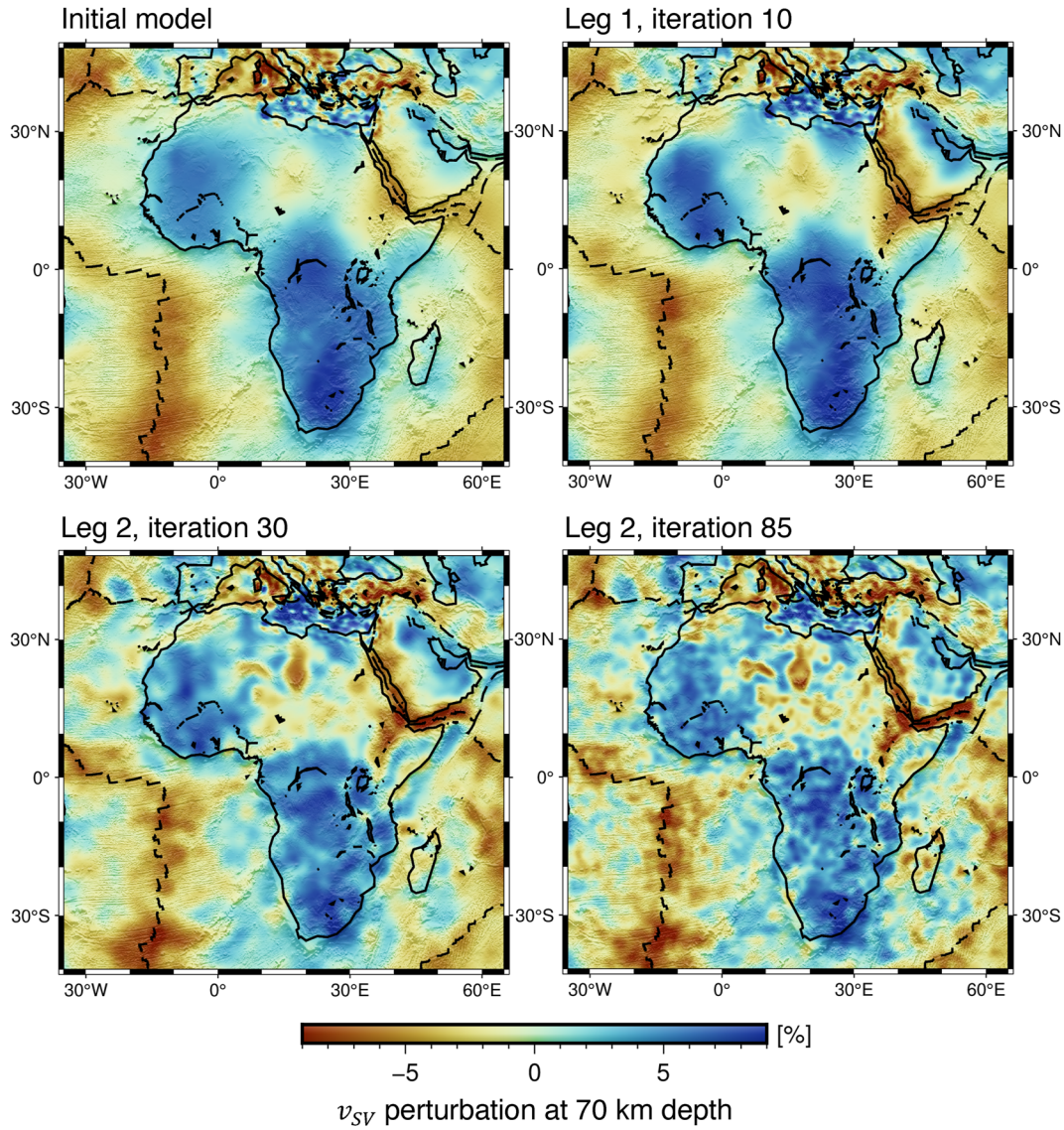

Figure 4. Progression of the v_{SV} model at 70 km depth. We show the state at selected iterations. Leg 1 refers to the period band of 55–130 s; Leg 2 refers to the period band of 35–130 s. Perturbations are plotted relative to the lateral mean. Notice that scale lengths become smaller as we increase the frequency, and regions such as the Mid-Atlantic Ridge and the Afar Triple junction become well-defined as low-velocity anomalies.

5.2. Validation Misfits and Waveform Fit

As mentioned in Section 4, we assess convergence by evaluating waveform fits of an independent data set. 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 data set. While not commonly done in the Earth science community, it is standard practice in machine learning (e.g., Tariq et al., 2018; Zhang & Sabuncu, 2018) to estimate the quality of predictions on independent data, as it helps to estimate overfitting. Figure 9 shows the evolution of the L_2 misfit throughout the inversion in the two period bands. Information about the batch sizes during the inversion can be found in Figures S1 and S2 in Supporting Information S1.

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

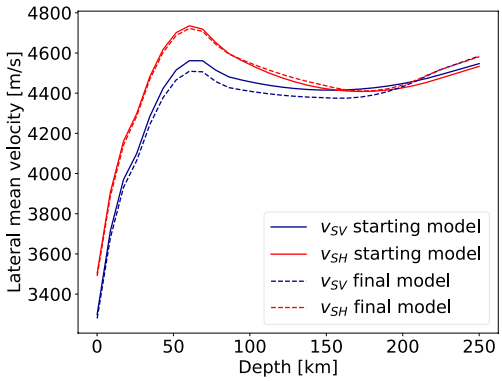

Figure 5. Depth profile of the lateral mean velocities for the v_{SV} and v_{SH} parameters of the initial and final model.

One can observe that the misfit curves do not decrease monotonically. This is not surprising, as the objective function we aim to minimize 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. Figure 10 shows a selection of three observed and synthetic seismograms computed from the validation data set.

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 itself is not tested; it evaluates only the FWI implementation and resolution achievable with the given data. Furthermore, checkerboard tests disregard the noise and hence provide limited insight into the solution's intrinsic non-uniqueness (Deal & Nolet, 1996; Rawlinson et al., 2014).

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., Gao et al., 2021; Nocedal & Wright, 2006),

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

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 $\boldsymbol{h}$ tending toward 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, Blom, 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

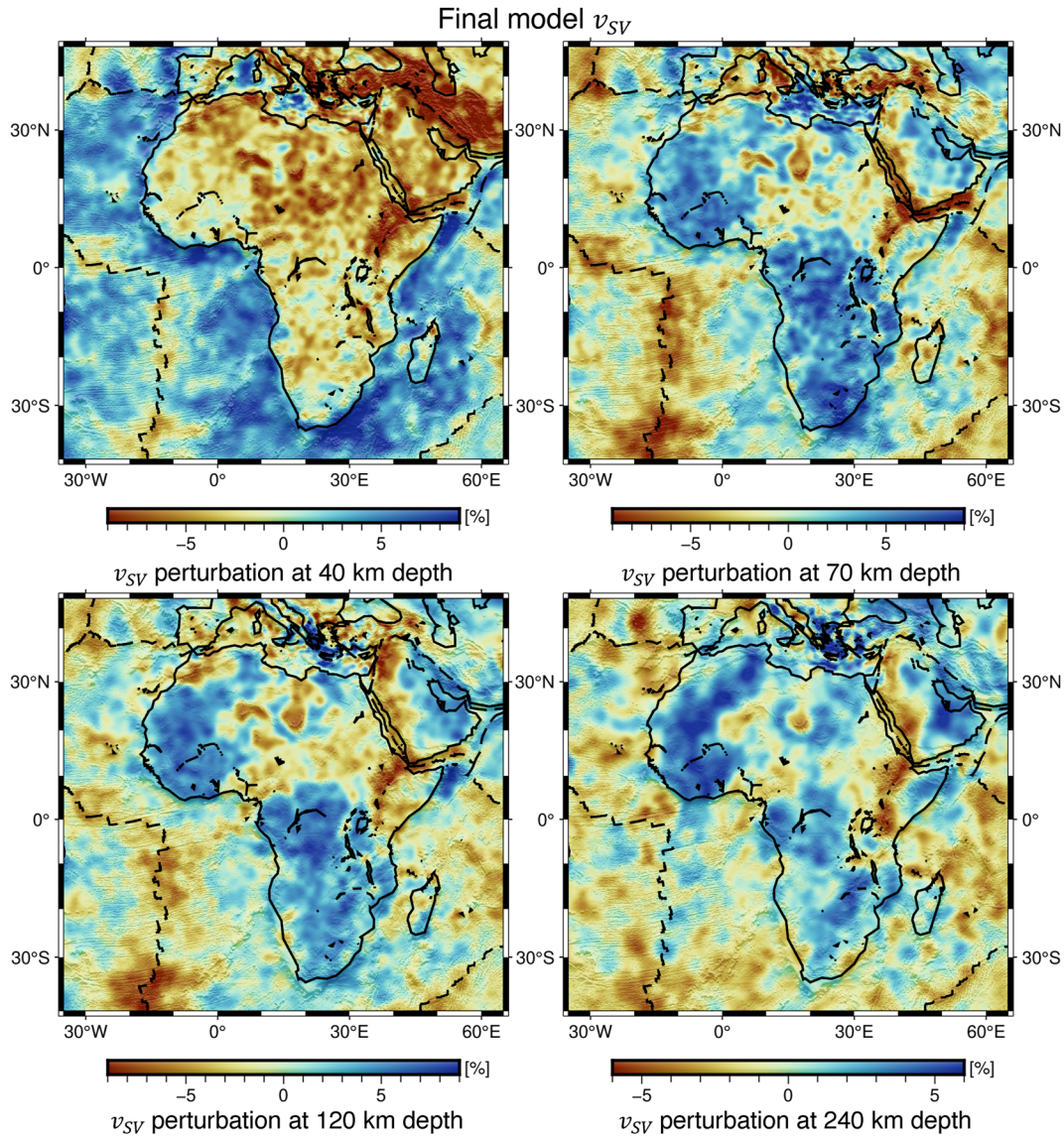

Figure 6. The final model of the v_{SV} parameter at 40, 70, 120, and 240 km depth. The main plate boundaries (Bird, 2003) are plotted on top with a black dashed line. Perturbations are plotted relative to the lateral mean. Notice how the Afar triple junction seems to morph into an elongated low-velocity anomaly underneath the entire EARS at greater depths, with especially low velocities imaged east of Lake Victoria and the Afar region.

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 data set with all its natural imperfections.
3. There is no need to create a costly separate synthetic data set.
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 data set.
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

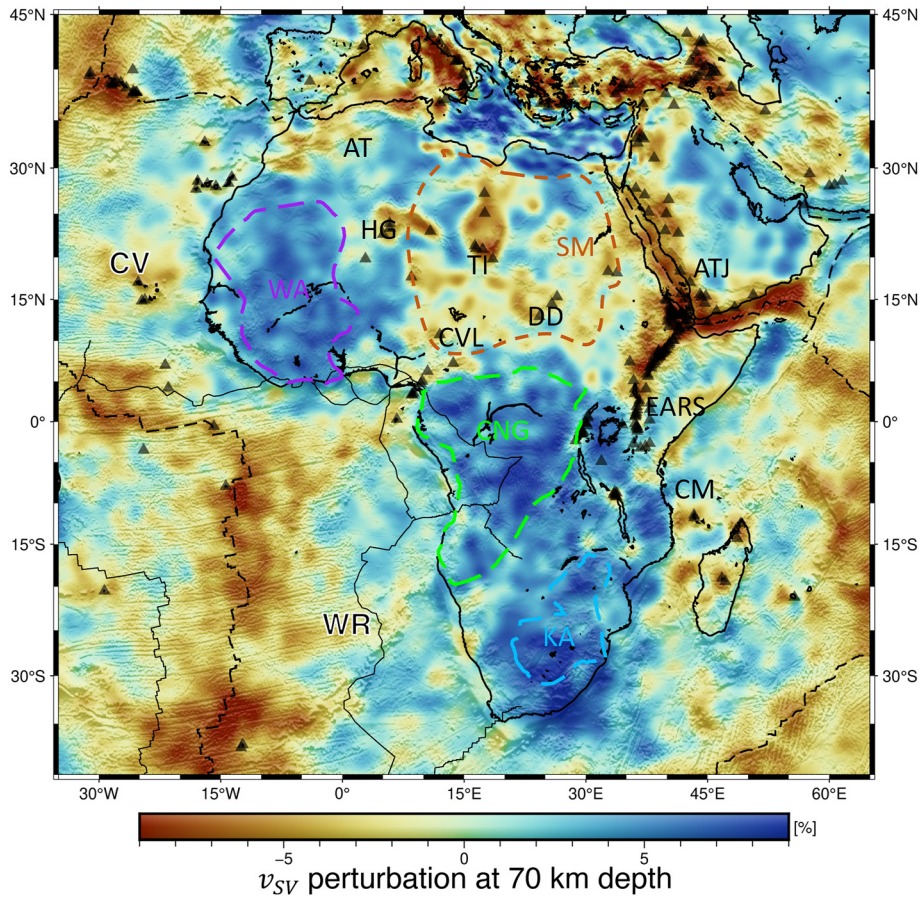

Figure 7. The final model of the v_{SV} parameter at 70 km depth. Perturbations are plotted with respect to the lateral mean. Volcano locations (NCEI, 2022) are indicated with black triangles. These correspond well to low-velocity regions in the model. The main cratons (Begg et al., 2009) are indicated on the map: WA: West African Craton, KA: Kalahari Craton, CNG: Congo Craton. SM: Saharan Metacraton. 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.

previously shown final model, see Figure 11. The test results for the v_{SV} parameter are summarized in Figures 12 and 13 at 150 and 400 km depth, respectively.

The restitution test figures (Figures 12 and 13) 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 source-receiver path crossings translates into the inability to restore the model.

At 150 km depth, our restitution test revealed small errors beneath the African continent, but larger errors were observed in oceanic regions and at the edges of the modeled domain. More specifically, in onshore Africa, we observed lateral smearing of approximately 100 km and amplitude errors mostly below 50 m/s, corresponding to a 1% velocity perturbation. Note that these values are relative to a perturbation scale length of about 500 km. 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. The nullspace is larger, and it is possible to vary the model to a greater degree while retaining similar waveform misfits. Another slice at 220 km depth is shown in Supporting Information S1. We have the greatest confidence in the upper few hundred kilometers of our model, particularly beneath the continents.

7. Discussion

We present a technical description of our full-waveform African Plate tomography model and its construction, complemented with proxies for resolution. Translating this into inferences of, for example, geodynamic

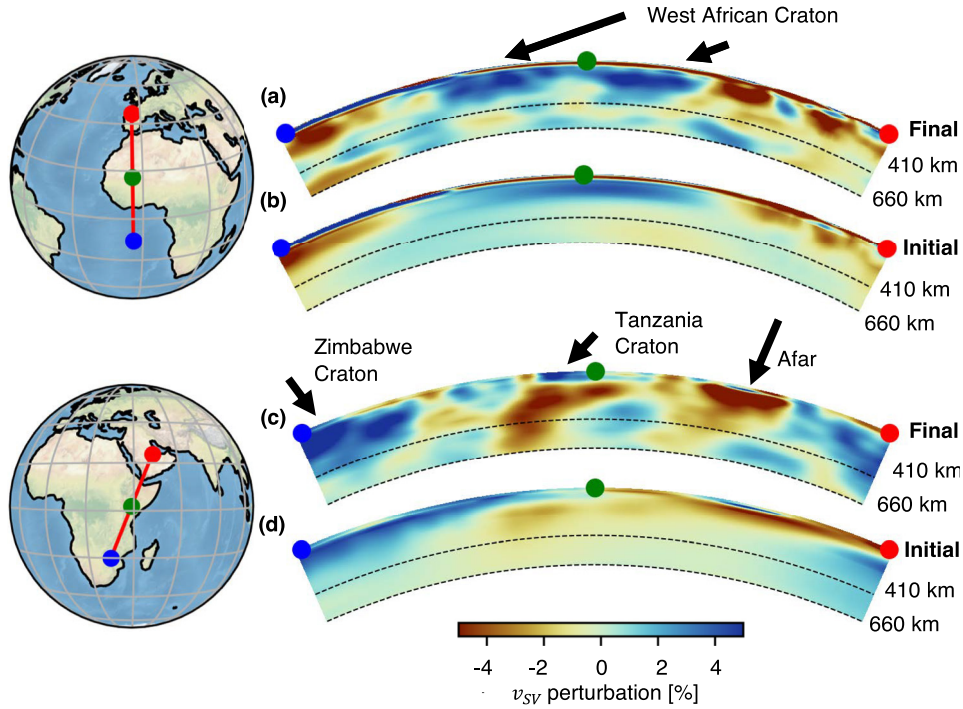

Figure 8. Selected cross-sections through the initial and final model. Perturbations of the v_{SV} parameter are plotted with respect to PREM (Dziewoński & Anderson, 1981). Subfigures (a and b) show cross-sections through the West African Craton for the final and initial model, respectively. The craton appears to consist of two parts that extend to greater depths. This corresponds well with the model by Celli et al. (2020). Subfigures (c and d) show cross-sections through the East African Rift System for the final and initial model, respectively. Low-velocity anomalies can be observed beneath the Tanzania Craton and the Afar region. At the very left of the plot, we can see high velocities in the Zimbabwe Craton.

processes, chemical composition, and temperature, and carefully propagating our uncertainties into uncertainties of these various inferences will require collaboration with experts, for which the authors will openly share data and models.

There are external factors limiting model resolution 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 a reliable model down to mid-upper mantle depths even in southern Africa with comparatively little data.

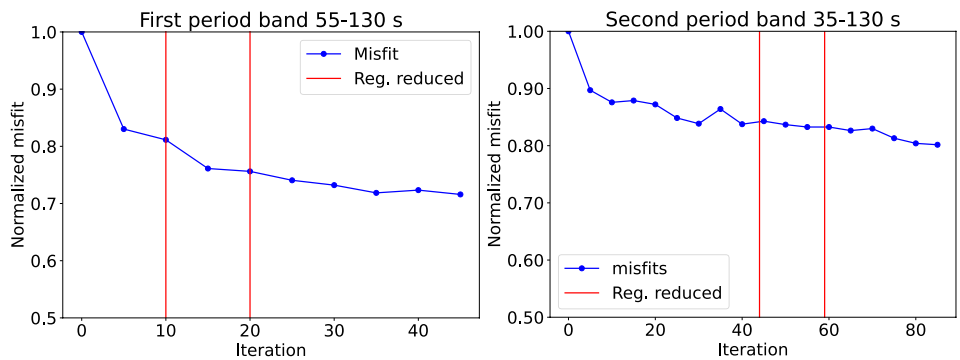

Figure 9. Normalized misfit reductions for the validation data set over the course of the iterations in the first (left) and second (right) leg of the inversion. Misfits are computed for the entire trace length. Misfits do not decrease monotonically, as this misfit is computed separately from the optimization, with L_2 misfits instead of time-frequency phase misfits, and for different data. The smoothing lengths were relative to wavelengths, based on velocities in the starting model of each period band, and defined in [radial, lateral] directions. From left to right the smoothing lengths evolved as follows: [0.5, 1.0], [0.4, 0.8], [0.3, 0.6] in the first frequency band and [0.4, 0.8], [0.3, 0.6] in the second frequency band.

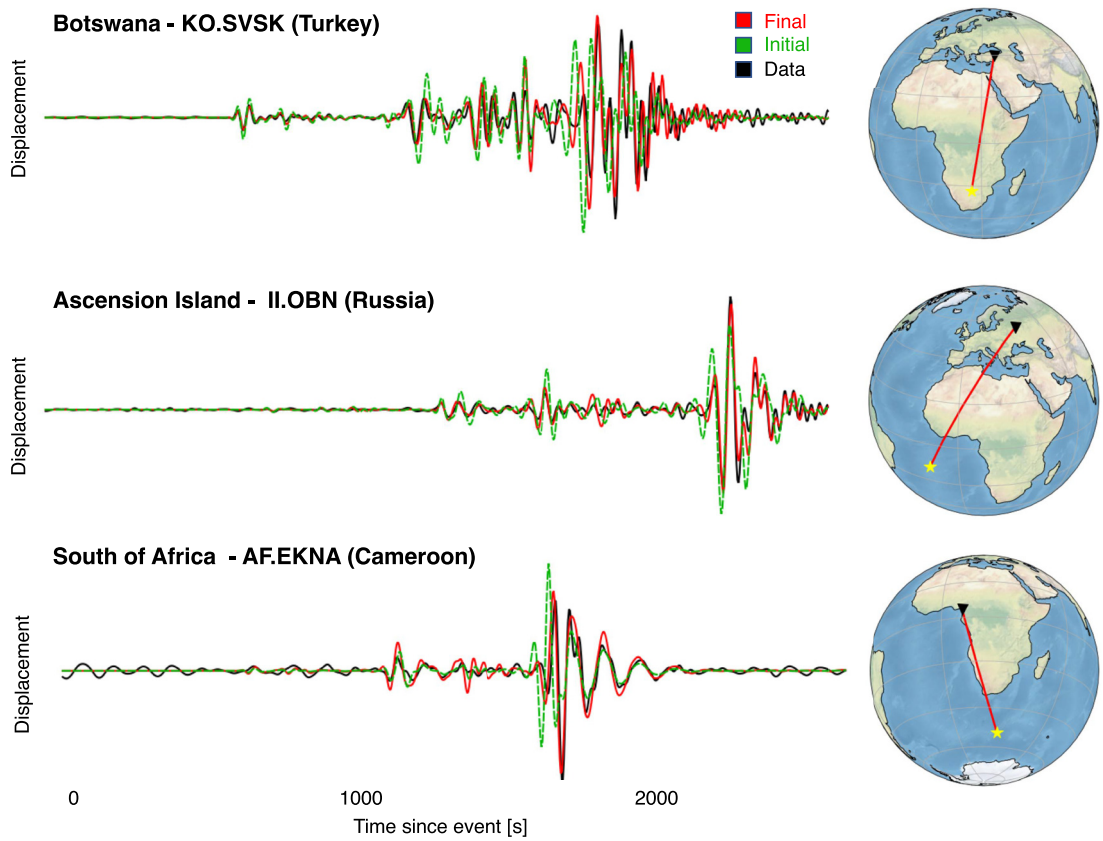

Figure 10. A selection of vertical-component seismograms from the validation data set for source-receiver pairs sensitive to the crustal and mantle structure of the African Plate. Top: a magnitude 6.5 earthquake occurring on 3 April 2017, recorded in Turkey; Middle: a magnitude 6.0 earthquake occurring on 2 February 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.

In this study, we did not invert for the source parameters but assumed the errors in the parameters provided by the GCMT Catalog (Ekström et al., 2012) to be negligible, since, for example, Hjörleifsdóttir and Ekström (2010), Bozdağ et al. (2016) only found minor updates when doing a source inversion. Most earthquakes in our data set 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. Nevertheless, the mini-batch approach allows us to include many more events, so their location inaccuracies will likely average out more efficiently.

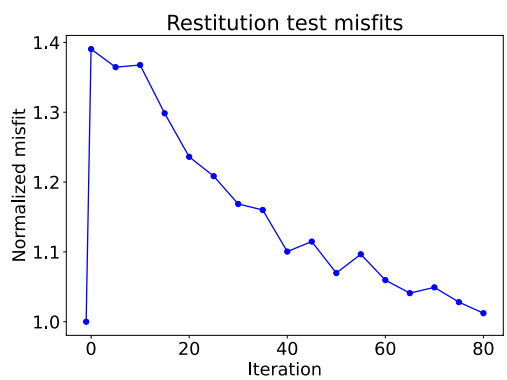

Figure 11. Validation event misfit for the 80 iterations within the restitution test. Initially, the misfit increases when we introduce errors by perturbing the model. Then we iterate until the misfits are approximately the same as for the final model.

Several other studies focusing on the African upper mantle have been published before (e.g., Celli et al., 2020; Emry et al., 2019; Fishwick, 2010; Pasyanos & Nyblade, 2007; Priestley et al., 2008; Ritsema & van Heijst, 2000; Sebai et al., 2006). 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 the recent models by Emry et al. (2019) and Celli et al. (2020). Figure 14 shows the isotropic shear wave velocity perturbations of the three models. For our model, we computed the Voigt averaged shear wave velocity

$$v_S = \sqrt{\left(2v_{SV}^2 + v_{SH}^2\right)/3}$$

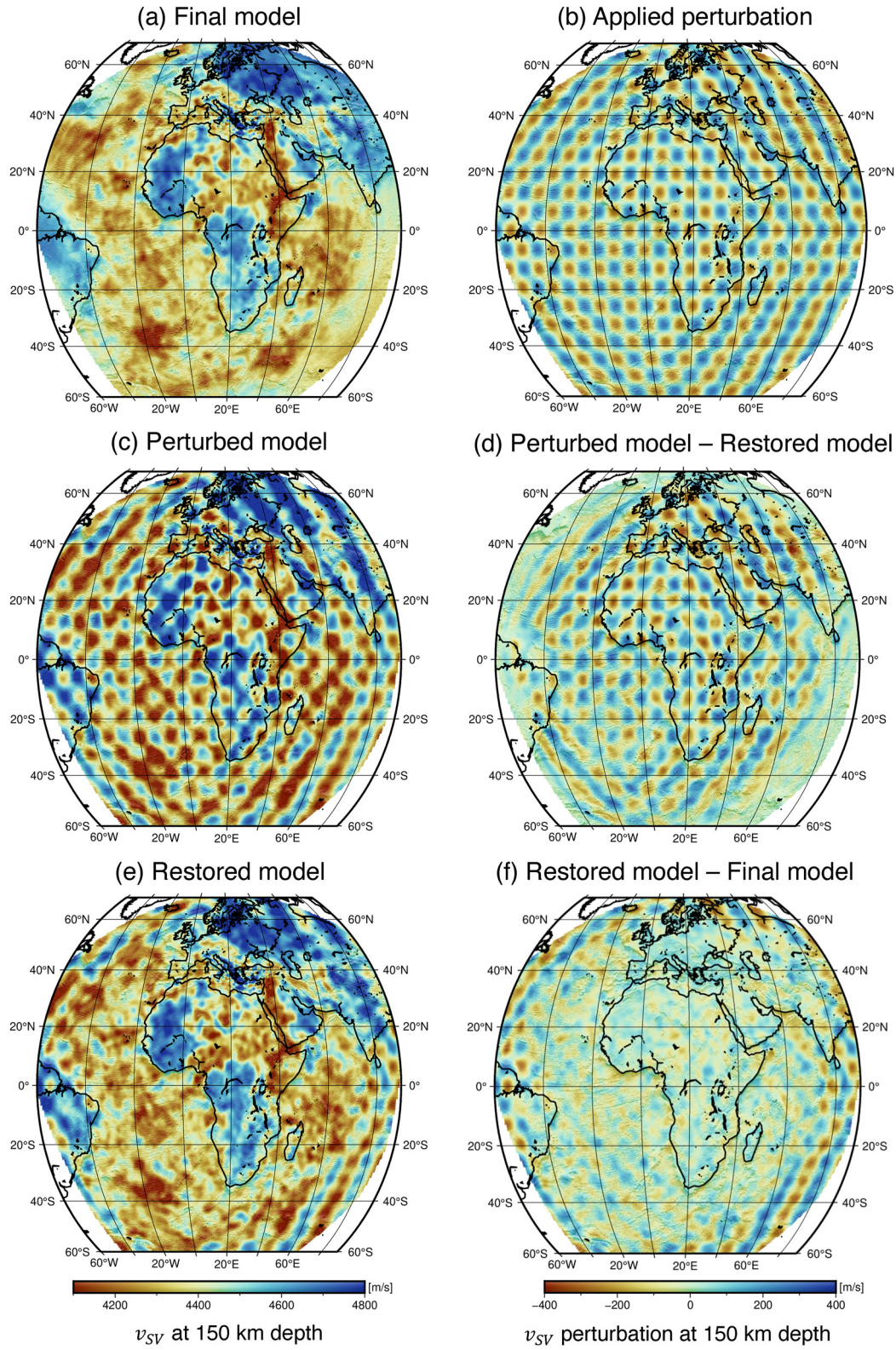

Figure 12. Restitution test for the v_{SV} parameter at 150 km depth. Panel (a) shows the final model after 130 mini-batch iterations, (b) shows the applied perturbation, (c) the model with the perturbation, (e) the model after the 80 repairing iterations, and (d) the difference between the perturbed and restored model. Finally, (f) shows the difference between the restored and the final model; this difference is relatively low beneath the African continent (at most 160 m/s).

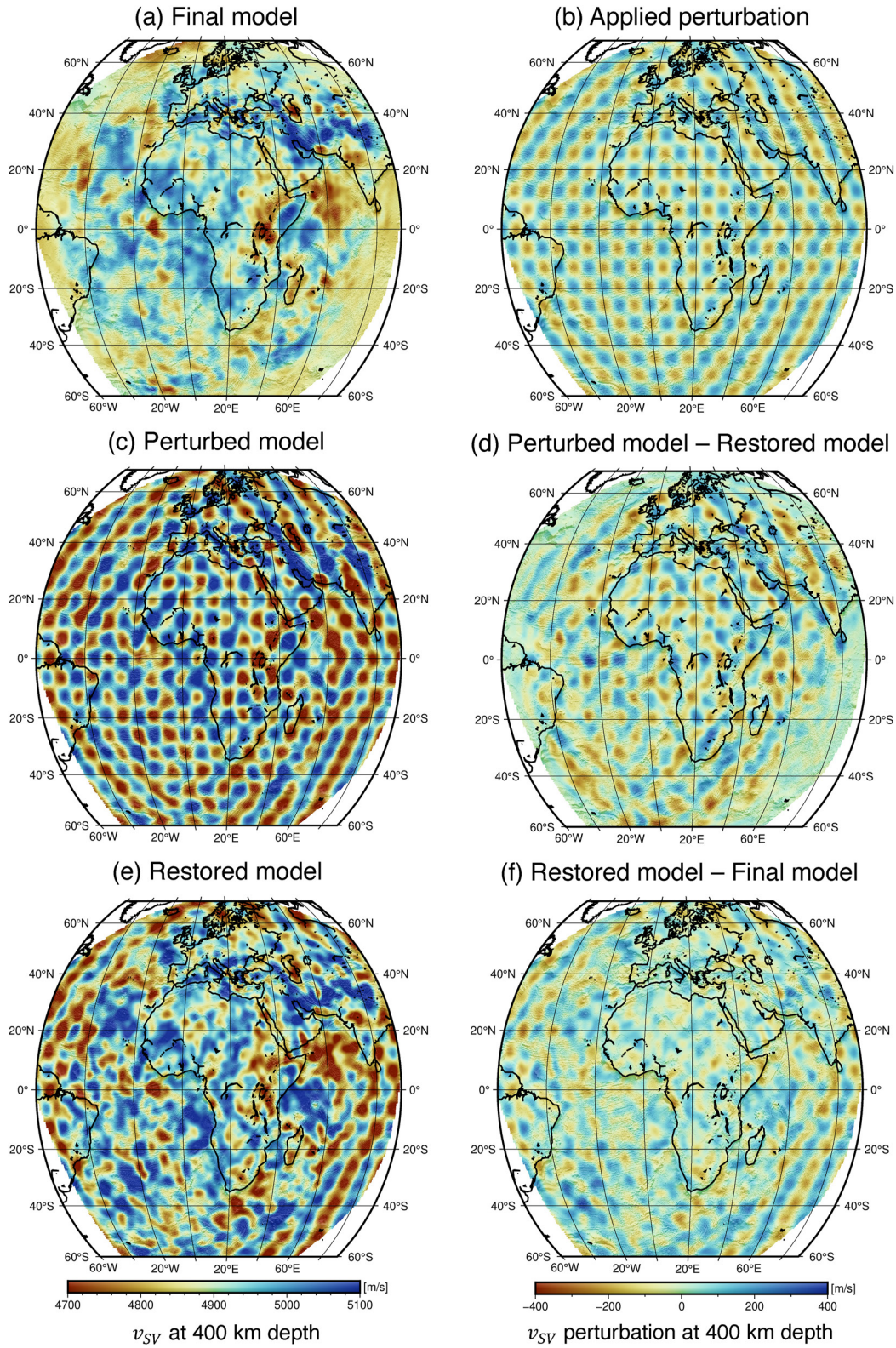

Figure 13. Restitution test for the v_{SV} parameter at 400 km depth. The description of the panels remains the same as in Figure 12. Notice the restoration errors remain larger here than in Figure 12. Panel (d) indicates that we can retrieve anomalies at the scale of the perturbations; however, the 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 (Figure 12) suggests velocities are less tightly constrained and thus imaged with less certainty.

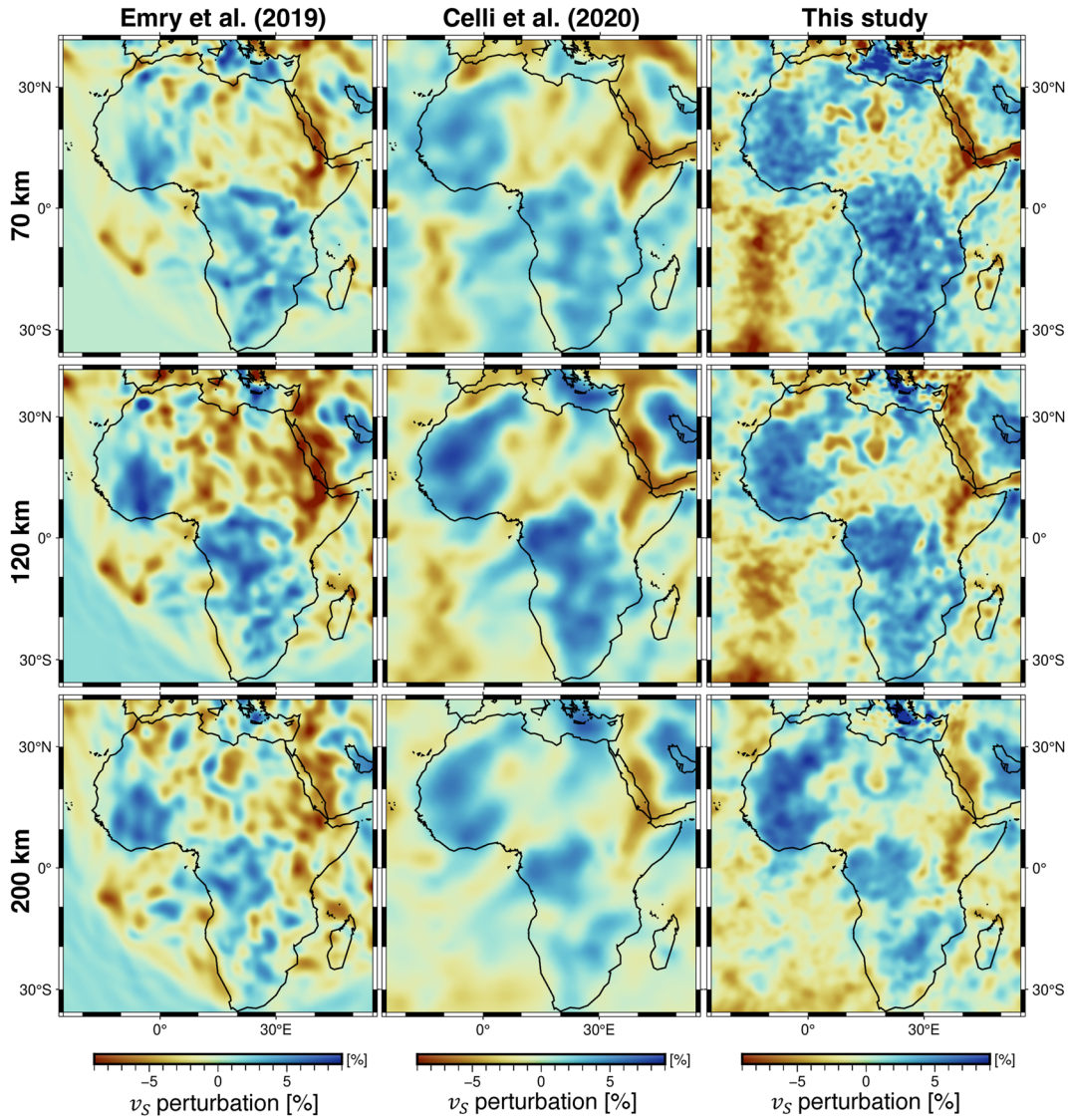

Figure 14. A comparison between the shear velocity (v_S) models of Emry et al. (2019), Celli et al. (2020) and this study at 70, 120, and 200 km depth. Perturbations are plotted relative to the lateral mean of the respective model.

(Babuska & Cara, 1991). Below, we focus only on a few selected features; as explained earlier, exhaustive comparisons and interpretations are beyond the scope of this paper.

Similar to our study, the model by Celli et al. (2020) was produced using a large data set of earthquake recordings and through waveform tomography. However, their synthetic waveforms and sensitivity kernels were computed from normal mode summations. A 1D reference Earth model was used, which limits taking nonlinearity and higher-order scattering into account.

Like our study, Emry et al. (2019) solved the seismic wave equation in 3D to obtain synthetic seismograms. In contrast, however, they used empirical Green's functions extracted from ambient noise instead of earthquake recordings. The common assumption that ambient noise cross-correlations are equivalent to Green's functions, not considering the noise source distribution, can potentially lead to imaging artifacts (e.g., Fichtner et al., 2016; Sager et al., 2018; Wapenaar, 2004). While it is, in principle, possible to consider the noise source distribution, 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.

Our model seems to be in accordance with Celli et al. (2020) to a great extent. It appears that some of the features in their model show more resemblance to our model at an earlier stage of the inversion and longer periods (see

Figure 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. 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).

Comparison with the model by Emry et al. (2019) is more complicated, likely due to the different methodology, assumptions, and data. Particularly at 120 and 200 km depth, that study shows low-velocity anomalies at similar locations as our study. These include the overall structure of the EARS, the high-velocity anomaly under Lake Victoria, contrasting with the nearby low-velocity anomalies beneath Lake Turkana and the Albertine Rift, especially at 120 km depth under the volcanic area between Lake Kivu and Lake Albert and northward of Lake Malawi (not visible in Celli et al. (2020)); the low velocities of the Angola Dome; and others. Our model also seems to confirm the suggestions that the distribution of low velocities in the "African Superplume" is more complex than previously thought. More specifically, our Figure 8 suggests at least two distinct low-velocity anomalies beneath the Tanzania Craton and the Afar region.

Structures of the Sahara Metacraton are relatively aligned between the model by Emry et al. (2019) and our model at the 200 km depth but noticeably different at shallower depths. More specifically, the low-velocity area beneath Tibesti, stretching northwards to the Gulf of Sidra manifests itself consistently at all imaged depths in both models. The models agree that at the 200 km depth, this area is surrounded by high-velocity anomalies. However, in our model, this structure extends to shallower, crustal depths, whereas the other model includes more low-velocity occurrences there. Hence, our model does not seem to confirm the statement in 5.1.2 of Emry et al. (2019) that high-velocity features are subdued at shallow upper mantle depths beneath the Hoggar, Tibesti, and Darfur Domes.

At the edges of the domain, in the oceans, there are more significant differences. This is likely caused by the fact that the data used by Emry et al. (2019) were primarily sensitive to structure between land-based stations. Other locations where models disagree include the strikingly contrasting high-velocity area beneath the High Atlas and the low velocity eastwards or the separation between high-velocity blocks in the Congo Craton.

With the proposed restitution test in hand, it is possible to see where we can reliably recover structure. This test is cheaper to compute than a traditional checkerboard test because we do not need to create a synthetic data set first. No approximate versions of the real data set 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 extremely costly. Doing so, however, might make it possible to obtain more quantitative estimates of model uncertainty.

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, even in places with limited station availability suggests that the imaged features are reliable. However, our novel model uncertainty analysis indicates that we should 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.

Data Availability Statement

All seismic data is publicly 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; Krischer, Megies, et al., 2015; Megies et al., 2011) to download the seismic data. The

model will be available at https://cos.ethz.ch/models.html and on IRIS EMC (IRIS, 2011). The majority of the figures were created with GMT (Wessel et al., 2019). The scientific color maps that we used were created by Crameri (2021) and are available online.

Acknowledgments
We would like to extend our sincere gratitude to Debby Wehner and an anonymous reviewer for their invaluable feedback and insightful comments on our manuscript. Their expertise and constructive suggestions significantly contributed to the improvement of this work. We also gratefully acknowledge the support of the European Research Council (ERC) through the EU's Horizon 2020 program (Grant 714069), as well as the Platform for Advanced Scientific Computing (PASC) program for funding the projects "Salvus" (2017–2020) and "Bayesian Waveform Inversion" (2021–2024). Furthermore, we gratefully acknowledge support from the Swiss National Supercomputing Centre (CSCS) in the form of computational resources Grant s1040. Also, we would like to thank Nicolas Celli for providing us with polygons of the Cratonic regions and scripts to visualize his model (Celli et al., 2020). In addition, we would like to thank IRIS DMC (IRIS, 2011), whose FDSN web services provided the majority of the data. Furthermore, we would like to thank everyone that contributed to the seismic data collection of the used seismic networks. We used data and are grateful for all the following networks: 1B, 1C, 1E, 1N, 1T, 2H, 3C, 3D, 3H, 4A, 4C, 4H, 5H, 5J, 6A, 6C, 6G, 7B, 7C, 7E, 8A, 8C, 8E, 9A, 9C, AB, AC, AF, AW, BA, BE, BL, BN, BR, BS, BW, BX, C4, CA, CB, CH, CK, CL, CQ, CR, CZ, DK, DZ, EB, EE, EI, ES, FN, FR, G, GB, GE, GH, GO, GR, GT, GU, GX, HA, HC, HE, HF, HL, HP, HT, HU, IB, IC, II, IM, IN, IO, IP, IS, IU, IV, IX, JS, KC, KN, KO, KR, KW, KZ, LC, LX, M1, MD, ME, MN, MP, MR, MT, NI, NJ, NL, NO, NR, NS, OE, ON, OT, OX, PF, PL, PM, PZ, RD, RF, RO, S, SI, SJ, SK, SL, SS, ST, SX, TH, TJ, TT, TU, TV, UD, UK, UP, WM, X5, X6, X7, XA, XB, XC, XE, XF, XG, XH, XI, XJ, XK, XM, XO, XP, XQ, XS, XV, XW, XY, XZ, Y1, Y2, Y4, YA, YB, YD, YF, YG, YH, YI, YJ, YK, YL, YP, YQ, YR, YT, YU, YV, YW, YY, YZ, Z2, Z3, Z4, ZA, ZB, ZC, ZD, ZE, ZF, ZH, ZI, ZK, ZM, ZO, ZP, ZQ, ZR, ZS, ZT, ZU, ZV, ZW, ZX, ZZ. Open access funding provided by Eidgenossische Technische Hochschule Zurich.

References

Abdelsalam, M. G., Liégeois, J.-P., & Stern, R. J. (2002). The Saharan Metacraton. *Journal of African Earth Sciences*, *34*(3–4), 119–136. https://doi.org/10.1016/s0899-5362(02)00013-1

Afanasiev, M., Boehm, C., van Driel, M., Krischer, L., Rietmann, M., May, D. A., et al. (2019). Modular and flexible spectral-element waveform modelling in two and three dimensions. *Geophysical Journal International*, *216*(3), 1675–1692. https://doi.org/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. https://doi.org/10.1002/2016gl068069

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

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. https://doi.org/10.1109/tpami.2002.1008390

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), Q12022. https://doi.org/10.1029/2008gc002107

Begg, G., Griffin, W., Natapov, L., O'Reilly, S. Y., Grand, S., O'Neill, C., et al. (2009). The lithospheric architecture of Africa: Seismic tomography, mantle petrology, and tectonic evolution. *Geosphere*, *5*(1), 23–50. https://doi.org/10.1130/ges00179.1

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), Q11013. https://doi.org/10.1029/2006gc001398

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., & Behr, Y. (2010). ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, *81*(3), 47–58. https://doi.org/10.1785/gssrl.81.3.530

Bird, P. (2003). An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems*, *4*(3), 1027. https://doi.org/10.1029/2001gc000252

Blom, N., Gokhberg, A., & Fichtner, A. (2020). Seismic waveform tomography of the central and eastern Mediterranean upper mantle. *Solid Earth*, *11*(2), 669–690. https://doi.org/10.5194/se-11-669-2020

Boyce, A., Bastow, I., Cottaar, S., Kounoudis, R., Guilloud De Courbeville, J., Caunt, E., & Desai, S. (2021). Afrp20: New p-wavespeed model for the African mantle reveals two whole-mantle plumes below east Africa and neoproterozoic modification of the Tanzania craton. *Geochemistry, Geophysics, Geosystems*, *22*(3), e2020GC009302. https://doi.org/10.1029/2020gc009302

Bozdağ, E., Peter, D., Lefebvre, M., Tromp, J., Komatitsch, D., Pugmire, D., & Podhorszki, N. (2016). Global adjoint tomography: First-generation model. *Geophysical Journal International*, *207*(3), 1739–1766. https://doi.org/10.1093/gji/ggw356

Bunks, C., Saleck, F. M., Zaleski, S., & Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, *60*(5), 1457–1473. https://doi.org/10.1190/1.1443880

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. https://doi.org/10.1038/s41467-019-13871-2

Chaljub, E., Capdeville, Y., & Vilotte, J. (2003). Solving elastodynamics in a solid heterogeneous 3-sphere: A spectral element approximation on geometrically non-conforming grids. *Journal of Computational Physics*, *183*(2), 457–491. https://doi.org/10.1016/s0021-9991(03)00119-0

Chorowicz, J. (2005). The East African Rift System. *Journal of African Earth Sciences*, *43*(1–3), 379–410. https://doi.org/10.1016/j.jafrearsci.2005.07.019

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. https://doi.org/10.1016/j.tecto.2013.06.015

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

Deal, M. M., & Nolet, G. (1996). Nullspace shuttles. *Geophysical Journal International*, *124*(2), 372–380. https://doi.org/10.1111/j.1365-246x.1996.tb07027.x

Dedzo, M. G., Kamgang, P., Njonfang, E., Tefogoum, G. Z., Dongmo, A. K., & Nkouathio, D. (2013). Mapping and assessment of volcanic hazards related to the ignimbritic eruption by AMS in Bambouto volcano (Cameroon volcanic line). *The Open Geology Journal*, *7*(1), 1–13. https://doi.org/10.2174/1874262901307010001

De Waele, B., Johnson, S., & Pisarevsky, S. (2008). Palaeoproterozoic to neoproterozoic growth and evolution of the eastern Congo craton: Its role in the Rodina puzzle. *Precambrian Research*, *160*(1–2), 127–141. https://doi.org/10.1016/j.precamres.2007.04.020

Duncan, R. A., & Richards, M. (1991). Hotspots, mantle plumes, flood basalts, and true polar wander. *Reviews of Geophysics*, *29*(1), 31–50. https://doi.org/10.1029/90rg02372

Dziewoński, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, *25*(4), 297–356. https://doi.org/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. https://doi.org/10.1038/27417

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. https://doi.org/10.1016/j.pepi.2012.04.002

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. https://doi.org/10.1029/2018gc007804

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. https://doi.org/10.1190/1.1444036

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*(1–2), 86–104. https://doi.org/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. *Geophysical Journal International*, *175*(2), 665–685. https://doi.org/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. *Geophysical Journal International*, *179*(3), 1703–1725. https://doi.org/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 and Planetary Science Letters*, *373*, 109–117. https://doi.org/10.1016/j.epsl.2013.04.027
Fichtner, A., Stehly, L., Ermert, L., & Boehm, C. (2016). Generalized interferometry – I: Theory for interstation correlations. *Geophysical Journal International*, *208*(2), 603–638. https://doi.org/10.1093/gji/ggw420
Fichtner, A., & Trampert, J. (2011). Resolution analysis in full waveform inversion. *Geophysical Journal International*, *187*(3), 1604–1624. https://doi.org/10.1111/j.1365-246x.2011.05218.x
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. https://doi.org/10.1093/gji/ggt118
Fichtner, A., van Herwaarden, D.-P., Afanasiev, M., Simutė, S., Krischer, L., Çubuk Sabuncu, Y., et al. (2018). The collaborative seismic Earth model: Generation 1. *Geophysical Research Letters*, *45*(9), 4007–4016. https://doi.org/10.1029/2018GL077338
Fichtner, A., & Villaseñor, A. (2015). Crust and upper mantle of the western Mediterranean – Constraints from full-waveform inversion. *Earth and Planetary Science Letters*, *428*, 52–62. https://doi.org/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. https://doi.org/10.1016/j.lithos.2010.05.011
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. https://doi.org/10.1144/SP357.19
Fitton, J. (1987). The Cameroon line, West Africa: A comparison between oceanic and continental alkaline volcanism. *Geological Society, London, Special Publications*, *30*(1), 273–291. https://doi.org/10.1144/gsl.sp.1987.030.01.13
Franz, G., Pudlo, D., Urlacher, G., Haussmann, U., Boven, A., & Wemmer, K. (1994). The Darfur dome, western Sudan: The product of a subcontinental mantle plume. *Geologische Rundschau*, *83*(3), 614–623. https://doi.org/10.1007/bf00194166
Furman, T. (2007). Geochemistry of East African rift basalts: An overview. *Journal of African Earth Sciences*, *48*(2–3), 147–160. https://doi.org/10.1016/j.jafrearsci.2006.06.009
Gao, Y., Tilmann, F., van Herwaarden, D.-P., Thrastarson, S., Fichtner, A., Heit, B., et al. (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. https://doi.org/10.1029/2021jb021984
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. https://doi.org/10.1029/2019JB018428
Gomez, F., Beauchamp, W., & Barazangi, M. (2000). Role of the Atlas mountains (northwest Africa) within the African-Eurasian plate-boundary zone. *Geology*, *28*(9), 775–778. https://doi.org/10.1130/0091-7613(2000)028<0775:rotamn>2.3.co;2
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. https://doi.org/10.1016/j.pepi.2009.11.003
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*.
International Seismological Centre. (2023). International seismograph station registry (IR). https://doi.org/10.31905/EL3FQQ40
IRIS, D. (2011). Data services products: EMC, a repository of Earth models. https://doi.org/10.17611/DP/EMC.1
Jessell, M. W., Begg, G. C., & Miller, M. S. (2016). The geophysical signatures of the west African craton. *Precambrian Research*, *274*, 3–24. https://doi.org/10.1016/j.precamres.2015.08.010
Keir, D., Bastow, I. D., Pagli, C., & Chambers, E. L. (2013). The development of extension and magmatism in the red sea rift of afar. *Tectonophysics*, *607*, 98–114. https://doi.org/10.1016/j.tecto.2012.10.015
Komatitsch, D., & Tromp, J. (1999). Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical Journal International*, *139*(3), 806–822. https://doi.org/10.1046/j.1365-246x.1999.00967.x
Komatitsch, D., & Tromp, J. (2002). Spectral-element simulations of global seismic wave propagation, Part II: 3-D models, oceans, rotation, and gravity. *Geophysical Journal International*, *150*(1), 303–318. https://doi.org/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. *Journal of Geophysical Research: Solid Earth*, *123*(7), 5902–5928. https://doi.org/10.1029/2017jb015289
Krischer, L., Fichtner, A., Zukauskaite, S., & Igel, H. (2015). Large-scale seismic inversion framework. *Seismological Research Letters*, *86*(4), 1198–1207. https://doi.org/10.1785/0220140248
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. https://doi.org/10.1088/1749-4699/8/1/014003
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. https://doi.org/10.1111/j.1365-246x.2005.02708.x
Liégeois, J., Benhallou, A., Azzouni-Sekkal, A., Yahiaoui, R., & Bonin, B. (2005). The Hoggar swell and volcanism: Reactivation of the Precambrian Tuareg shield during alpine convergence and West African Cenozoic volcanism. *Special Papers - Geological Society of America*, *388*, 379.
Liégeois, J.-P., Abdelsalam, M. G., Ennih, N., & Ouabadi, A. (2013). Metacraton: Nature, genesis and behavior. *Gondwana Research*, *23*(1), 220–237. https://doi.org/10.1016/j.gr.2012.02.016
Liu, Q., & Gu, Y. (2012). Seismic imaging: From classical to adjoint tomography. *Tectonophysics*, *566–567*, 31–66. 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 of Geophysics*, *54*, 47–58.
Meier, U., Curtis, A., & Trampert, J. (2007). Fully nonlinear inversion of fundamental mode surface waves for a global crustal model. *Geophysical Research Letters*, *34*(16), L16304. https://doi.org/10.1029/2007GL030989
Moucha, R., & Forte, A. M. (2011). Changes in African topography driven by mantle convection. *Nature Geoscience*, *4*(10), 707–712. https://doi.org/10.1038/ngeo1235
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. https://doi.org/10.1002/ggge.20150
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. https://doi.org/10.1007/978-0-387-40065-5

Nougier, J., Cantagrel, J., & Karche, J. (1986). The comores archipelago in the western Indian ocean: Volcanology, geochronology and geodynamic setting. *Journal of African Earth Sciences*, *5*(2), 135–145. https://doi.org/10.1016/0899-5362(86)90003-5

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. https://doi.org/10.1029/2011eo190002

Pasyanos, M. E., & Nyblade, A. A. (2007). A top to bottom lithospheric study of Africa and Arabia. *Tectonophysics*, *444*(1–4), 27–44. https://doi.org/10.1016/j.tecto.2007.07.008

Patera, A. T. (1984). A spectral element method for fluid dynamics: Laminar flow in a channel expansion. *Journal of Computational Physics*, *54*(3), 468–488. https://doi.org/10.1016/0021-9991(84)90128-1

Permenter, J. L., & Oppenheimer, C. (2007). Volcanoes of the Tibesti massif (Chad, northern Africa). *Bulletin of Volcanology*, *69*(6), 609–626. https://doi.org/10.1007/s00445-006-0098-x

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. https://doi.org/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. https://doi.org/10.1111/j.1365-246x.2008.03951.x

Rawlinson, N., Fichtner, A., Sambridge, M., & Young, M. K. (2014). *Chapter one - Seismic tomography and the assessment of uncertainty* (Vol. 55, pp. 1–76). Elsevier. https://doi.org/10.1016/bs.agph.2014.08.001

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 and Planetary Science Letters*, *367*, 39–51. https://doi.org/10.1016/j.epsl.2013.02.022

Ritsema, J., & van Heijst, H. (2000). New seismic model of the upper mantle beneath Africa. *Geology*, *28*(1), 63–66. https://doi.org/10.1130/0091-7613(2000)028<0063:nsmotu>2.3.co;2

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. *Geophysical Journal International*, *149*(2), 482–489. https://doi.org/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. https://doi.org/10.1126/science.286.5446.1925

Robertsson, J. O. A., Blanch, J. O., & Symes, W. W. (1994). Viscoelastic finite-difference modelling. *Geophysics*, *59*(9), 1444–1456. https://doi.org/10.1190/1.1443701

Romanowicz, B. A., & Dziewonski, A. M. (1986). Toward a federation of broadband seismic networks. *Eos, Transactions American Geophysical Union*, *67*(25), 541–542. https://doi.org/10.1029/EO067i025p00541

Ruan, Y., Lei, W., Modrak, R., Örsvuran, R., Bozdag, E., & Tromp, J. (2019). Balancing unevenly distributed data in seismic tomography: A global adjoint tomography example. *Geophysical Journal International*, *219*(2), 1225–1236. https://doi.org/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. https://doi.org/10.1029/2019jb018644

Sager, K., Ermert, L., Boehm, C., & Fichtner, A. (2018). Towards full waveform ambient noise inversion. *Geophysical Journal International*, *212*(1), 566–590. https://doi.org/10.1093/gji/ggx429

Schlüter, T. (2008). *Geological atlas of Africa* (Vol. 307). Springer.

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. https://doi.org/10.1016/j.pepi.2005.09.009

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. https://doi.org/10.1126/science.1175298

Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, *49*(8), 1259–1266. https://doi.org/10.1190/1.1441754

Tarantola, A. (1988). Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation. *Pure and Applied Geophysics*, *128*, 365–399. https://doi.org/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. https://doi.org/10.1371/journal.pmed.1002705

ten Brink, U., & Stern, T. (1992). Rift flank uplifts and hinterland basins: Comparison of the transantarctic mountains with the great escarpment of southern Africa. *Journal of Geophysical Research*, *97*(B1), 569–585. https://doi.org/10.1029/91jb02231

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

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

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

Tromp, J., Tape, C., & Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, *160*(1), 195–216. https://doi.org/10.1111/j.1365-246x.2004.02453.x

Vallée, M. (2013). Source time function properties indicate a strain drop independent of earthquake depth and magnitude. *Nature Communications*, *4*(1), 2606. https://doi.org/10.1038/ncomms3606

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. https://doi.org/10.1093/gji/ggaa079

Virieux, J., & Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, *74*(6), WCC127–WCC152. https://doi.org/10.1190/1.3238367

Wapenaar, K. (2004). Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation. *Physical Review Letters*, *93*(25), 254301. https://doi.org/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*, *127*(3), e2021JB022930. https://doi.org/10.1029/2021jb022930

Wehner, D., Rawlinson, N., Greenfield, T., 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*, *23*(11), e2022GC010563. https://doi.org/10.1029/2022gc010563

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. https://doi.org/10.1029/2019gc008515

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