MULTI-SCALE/MULTI-DATA INVERSION FOR ELASTIC EARTH STRUCTURE - A CONCEPT

Andreas Fichtner¹, Michael Afanasiev¹, Korbinian Sager¹ and Laura Ermert¹

¹Department of Earth Sciences, ETH Zurich
Sonneggstrasse 5, 8092 Zurich, Switzerland
e-mail: andreas.fichtner@erdw.ethz.ch

Keywords: Seismic tomography, computational seismology, seismic wave propagation, inverse theory

Abstract. Complex interactions of small- and large-scale processes are characteristic for the physics of the Earth, and their proper quantification is key to the integration of interdependent geophysical systems that are today mostly treated as isolated.

Inferring Earth structure over a wide range of scales is the long-standing goal of seismic tomography. While much progress has been made in recent years, tomographic resolution remains limited by our inability to model and invert seismic wave propagation across the complete observable frequency band with the currently available computational resources.

Here we propose a new concept for multi-scale seismic tomography intended to resolve Earth structure from local to global scales, including mantle as well as detailed crustal features. For this we develop a multi-scale full waveform inversion technique that assimilates complete teleseismic and regional seismograms in a broad frequency band. Being based on spectral-element modelling and adjoint techniques, our method simultaneously solves multiple regional- and continental-scale inverse problems in order to jointly resolve Earth structure with resolving lengths ranging from around 20 to more than 5000 km. To further increase the exploitable frequency band beyond what can be modelled numerically, we combine full waveform inversion with classical ray tomography that assimilates the arrival times of high-frequency body waves into the tomographic model. This combination results in an improved resolution of Earth structure, especially below 300 km depth.

We apply our method to Europe and Western Asia, where resolution is particularly high beneath the North Atlantic, the Western Mediterranean and Anatolia. Quantitative resolution analysis based on second-order adjoints, as well as comparisons with observed ambient noise correlations, allow us to assess the quality and predictive power of the final model.
1 INTRODUCTION

1.1 Scaling in waveform inversion

During the past decade, full waveform inversion in seismology has developed from a research topic into a mature technique with numerous applications [1, 2, 3, 4]. Factors contributing to this development include the design of misfit functionals that explicitly introduce traveltime information [5, 6, 7], the optimisation of numerical forward modelling tools [8, 9], and advanced data processing and window selection techniques [10, 11]. While growing compute power has contributed as well, the availability of computational resources is still the most significant limitation. Even with the unrealistic assumption of perfect scaling, doubling the maximum frequency in a waveform inversion multiplies the computational requirements by a factor of $2^5 = 32$. A factor $2^3 = 8$ is caused by the necessary refinement of the numerical mesh in three dimensions. An additional factor 2 is imposed by the CFL stability condition that enforces a reduction of the numerical time step. Finally, the volume covered by sensitivity kernels is anti-proportional to frequency, meaning that at least twice as many data - and therefore simulations - are needed to cover the same volume and avoid a loss of resolution. Current full waveform inversions assimilate seismic waves that have propagated on the order of 100 wavelengths [1, 2, 3, 4]. This is in contrast to the $O(1000)$ wavelengths that a 1 Hz P wave travels through the Earth. The factor 10 in propagating wavelength translates to a factor of $10^5$ by which computational resources would have to increase at least, in order to allow seismologists to assimilate recordings with the complete observable bandwidth into a full waveform inversion.

This simple scaling argument, where storage requirements have not even been considered, illustrates the severity of the problem, and it calls for new methodological developments that allow us to broaden the exploitable bandwidth in tomographic inversions.

1.2 Objectives

The purpose of this contribution is to describe the concept of a new tomographic technique intended to assimilate data in a sufficiently broad frequency range to allow for the resolution of both crustal and mantle structure. The backbone of our method is the coupling of tomographic inversions on different scales and with different forward problem solvers. This strategy includes the combination of full waveform inversion of shorter-period data on regional scales with full waveform inversion of longer-period data on continental scales, but also the coupling with ray tomography that inverts the traveltimes of high-frequency waves that cannot be modelled numerically.

1.3 Outline

This paper is organised as follows: In section 2 we provide a summary of multi-scale full waveform inversion and the coupling with ray tomography. The first application of our approach to European upper-mantle and crustal structure is presented in section 3. Finally, we discuss limitations of our method, as well as future directions in section 4. Since our main goal is to introduce a concept, we will keep the description of mathematical and data processing details at a minimum level, though providing numerous references.
2 THEORY

In order to broaden the exploitable bandwidth of seismic data, we combine tomographic inversions that operate on different scales and are based on different forward problem solvers. Our approach includes (i) multi-scale full waveform inversion through coupled inversions from local to continental scales, and (ii) the coupling of full waveform inversion with classical ray tomography. We describe these components of our method separately in the following paragraphs.

2.1 Multi-scale full waveform inversion

The basic version of our full waveform inversion rests on spectral-element simulations of seismic wave propagation through spherical sections of anisotropic, visco-elastic Earth models [12]. With the help of a semi-automatic window selection algorithm [11], we choose time windows where observed and computed seismograms are sufficiently similar to allow for physically meaningful measurements of time- and frequency-dependent phase differences [7]. From these measurements, we derive adjoint sources used in adjoint simulations to produce Fréchet kernels for anisotropic Earth structure [13, 14, 15]. The Fréchet kernels are then used to iteratively improve the Earth model. To ensure that as many data as possible enter the inversion, the window selection is repeated every few iterations, typically leading to an increase in the number and length of time windows where measurements can be performed. Furthermore, we re-invert for the source location, timing and mechanism of the earthquakes in order to avoid contamination of the Earth model from inaccurate source information.

To jointly assimilate longer-period data on larger scales and shorter-period data on smaller scales, we employ the following inversion scheme [16]: From a multi-scale master model we extract a smaller subvolume that contains both sources and receivers. The limited size of the subvolume allows us to model and invert seismic data at higher frequencies. The updated sub-model is then re-introduced into the master model. To compute an update on larger scales, we first employ non-periodic homogenisation in order to construct a smoothed long-wavelength equivalent of the master model [17, 18]. Being smooth, the long-wavelength equivalent can be meshed more coarsely, which reduces the computational requirements to a level where iterative updating with the longer-period data becomes feasible. The long-period update is then interpolated onto the master model, and the procedure is repeated until data are fit to within their errors. Depending on the detail of Earth structure, the homogenisation step may be omitted and replaced by directly meshing with a coarser grid spacing.

The advantage of this approach lies in its inherent scalability. There is no limit on the size and number of the submodels, meaning that high-frequency data for small local domains can be consistently assimilated into the master model. The major limitation is the requirement that sufficiently many sources and receivers must be located within the submodels, though this may be overcome using interferometric techniques [19, 20].

2.2 Ray tomography combined with full waveform inversion

Ray tomography uses traveltimes of distinct seismic phases to infer Earth structure. It relies on the simplifying assumption that seismic waves can be represented by rays. The intensive use of ray theory, which is closely related to geometrical optics, stems from its simplicity, its intuitive interpretation as well as its computational efficiency. Indeed, most seismological inferences concerning the structure of the Earth are based on this assumption.

Within the ray theoretical framework the arrival time of seismic waves solely depends on the
velocity along the geometrical ray path between source and receiver. Instead of simulating the full seismic wavefield, conventional approaches for computing seismic traveltimes try to find the path taken by seismic energy between source and receiver based on Snell’s law by shooting (e.g. [21, 22, 23, 24]) or pseudo/bending methods (e.g. [21, 25, 26]). These schemes may, however, fail to converge to the true two-point path even in mildly heterogeneous media and do not guarantee that the computed ray indeed corresponds to a first arrival. To solve these problems grid based solutions of the infinite frequency approximation of the wave equation, i.e. the eikonal equation, were and are still developed as an alternative approach (e.g. [27, 28, 29]). Most grid-based traveltime solvers can, however, only track first arrivals in continuous media, but it is possible to adapt grid-based schemes to track later arriving reflection and refraction phases in layered media (e.g. [30, 31, 32, 33, 34, 35]). Together with their robustness and advancing efficiency, these schemes are nowadays established as coequal, if not preferable, alternatives to conventional ray tracing methods.

In our application, we use the software package FMTOMO [38] for ray tomography. It computes the synthetic traveltimes with a grid based eikonal solver and adjusts the model parameters using a subspace inversion scheme [39]. On continental scales, limitations regarding the maximum amount of grid points for the propagation of information through the model prohibit a detailed representation of the crust and upper mantle. Besides crustal phases we also exclude the highly non-linear transition zone from crustal to upper/lower mantle phases and confine ourselves for now to direct waves for epicentral distances larger than 25°.

Ray tracing as well as grid-based eikonal solvers are only valid in the limit of a hypothetical infinite frequency wave. Its applicability relies on the assumption of weak and relatively large-scale lateral heterogeneities, which is not valid for structures with sizes that are comparable to or less than the dominant wavelength. At finite frequencies seismic waves are sensitive to 3D structure off the ray. Scattering, wave-front healing and other diffraction effects contribute to their traveltimes (e.g. [36]). Neglecting these effects, ray tomography is limited to high frequency information and especially long period waves cannot be inverted for.

Due to computational limitations full waveform inversion is limited to an intermediate period band, typically 10–200 s on regional to continental scales (e.g. [37, 4, 16]). At periods above 200 s, the self-gravitation of the Earth becomes relevant. Since this cannot be modelled with sufficient accuracy by numerical schemes, longer periods are omitted. More severely, the information contained in high frequency P- and S-wave traveltimes is therefore missing. Consequently, full waveform inversion yields promising results within the upper 300 km, where surface wave sensitivity is large. P- and S-velocity heterogeneities at greater depth are, however, less well resolved. Ray tomography explicitly incorporates information from arrivals of high frequency P- and S-waves. By alternating between full waveform inversion and ray tomography in an inversion, we intend to combine their complementary merits and produce 3D Earth models explaining both waveform and traveltime data sets to constrain Earth structure in more detail.

3 APPLICATION

We illustrate our approach with preliminary results from an ongoing tomographic inversion for crustal and upper-mantle structure beneath Europe and western Asia. In order to emphasise the concept and the method, we will mostly focus on the combination of waveform inversion and ray tomography, and on the Western Mediterranean as a smaller-scale subregion. Other subregions such as Anatolia or the North Atlantic will receive less attention.
3.1 Waveform data

The seismic waveform data used in the construction of our model fall into three categories, depending on scale and subregion: (i) On the continental (10,000 km) scale we use 14,525 recordings from 84 earthquakes in the period range from 30 to 200 s, meaning that these data mostly constrain large-scale upper-mantle structure. (ii) Within the Anatolian region, that is on a 10 times smaller scale, we complement the data set with 2,312 regional recordings from 29 earthquakes in Turkey with periods from 8 - 50 s that provide additional information on crustal heterogeneities [41]. (iii) For the Western Mediterranean region, we added 13,089 recordings from 52 events in the period range between 12 and 120 s, also in order to constrain the uppermost mantle and the crust. The data coverage for the Western Mediterranean is summarised in figure 1. The waveform data used in the alternating traveltime - full waveform inversion was a sub-sampled version of the continental scale dataset, using 52 events and 13,086 seismograms, bandpass filtered between 40 and 100 seconds.

![Figure 1: Source-receiver distribution in the Western Mediterranean region. Red-white beachballs indicate the source mechanisms and locations of the 52 earthquakes used in the full waveform inversion for this subregion. The great circles connecting sources and receivers are plotted in lighter colour when coverage is high.](image)

3.2 Traveltime data

The traveltime updates are based on measurements from the International Seismological Center (www.isc.ac.uk). The data set comprises 710 events with a total of 19,010 S wave traveltimes, measured at a frequency of 1 Hz. Since the measurements are on the transverse component, and are therefore primarily sensitive to SH-velocity, we restrict our velocity updates to the SH velocity component.

3.3 Inversion sequence

We started the inversion sequence with the construction of a long-wavelength model of the European upper mantle using full waveforms with a minimum period of 40 s. Subsequently, smaller-scale refinements within the Anatolian region and the North Atlantic were performed. This was followed by the combined ray and full waveform inversion that primarily updated...
structure below 200 km depth. Avoiding difficulties in modeling the complexities in the crust and upper mantle, the traveltime update primarily introduces information in deeper parts of the model. Therefore, the traveltime update is nearly invisible to the waveforms, and the waveform update does not influence the traveltime misfit. Finally, we refined the model in the Western Mediterranean using full waveform inversion.

This sequence is the result of continuing methodological developments rather than a long-term schedule. More systematic updates are planned for the near future, but require further automation and improved bookkeeping.

3.4 Multi-scale model of the European continent

Figures 2 and 3 summarise the current state of our model. The model is described in terms of the 3D distributions of SH velocity $v_{SH}$, SV velocity $v_{SV}$, P velocity $v_P$, and density $\rho$. Since most of the data are primarily sensitive to S velocity structure, $v_{SH}$ and $v_{SV}$ are resolved best. The European model embedded within longer-wavelength global structure [42, 43] is shown in figure 2. Differences between oceans and continents dominate the 25 km slice. At 100 km, the low-velocity imprints of the Iceland-Jan Mayen plume system [4] and the Alpine orogeny [41] are the dominant features. Around 300 km depth, the heterogeneities are weaker and mostly related to the deep structure of the North Atlantic plumes and the deep continental roots of Eurasia.

Close-up views of the regional submodels for Anatolia, the North Atlantic and the Western Mediterranean can be seen in figure 3. The uppermost mantle beneath Anatolia carries the imprint of the North Anatolian Fault Zone and various volcanic fields, all expressed as low-velocity anomalies [41]. The North Atlantic around 250 km depth is marked by the wide-spread plume head of the Iceland-Jan Mayen plume system, also characterised by lower than average seismic velocities [4]. Higher than average velocities associate with the subduction oceanic lithosphere are dominant in a belt stretching from the Strait of Gibraltar though the Western Mediterranean and into Italy.

Figure 2: Horizontal slices through the SH velocity structure beneath Europe at depth of 25 km, 100 km and 300 km.
Figure 3: Close-up views of SH velocity structure in the mantle beneath Anatolia, the North Atlantic, and the Western Mediterranean.

3.5 Model validation

The assessment of model quality and uncertainty is among the most challenging aspects of seismic tomography. With the help of second-order adjoints we can compute resolution lengths, that is the width of point-spread functions as a function of position and direction [44]. Maps of resolution length for the Western Mediterranean are shown in figure 4. The maps indicate that resolution is generally better in vertical direction because surface waves dominate the waveform data set. With increasing depth, the resolved volume shrinks, reflecting the decreasing body wave coverage in this regional setting.

In addition to formal resolution analysis, we study the fit to data that have not been used in the inversion. Ambient noise correlations are particularly useful for this purpose as they do not suffer from incorrect earthquake locations and source mechanisms. Furthermore, they probe the model in regions that may be less well covered by earthquake data. To compute noise correlations, we obtained continuous seismic data for the year 2012 from the IRIS DMC (www.iris.edu). The instrument response was removed for the recordings, after which data were cut into overlapping nine-hour-segments, decimated and bandpass-filtered between 4 and 20 s. The normalised cross-correlation was then calculated without further preprocessing. Geometric normalisation of each correlation window ensures the downweighting of high energy signals e.g. from earthquakes.

Figure 5 shows a selection of correlation functions compared to numerical Green functions. Provided that a surface wave clearly emerges from the noise correlation data, it is matched by the Green function to within a few seconds. A remarkable example is shown in the third row of figure 5 where surface waves are not explained but the rare body wave arrivals in the correlation function are explained by the numerical Green function. The above comparison is currently limited to a visual analysis that adds subjective plausibility to the model. A quantitative comparison would require more detailed information on the distribution of noise sources.

4 DISCUSSION AND CONCLUSIONS

In the previous sections we introduced a concept for the construction of multi-scale Earth models based on the solution of tomographic inverse problems on different scales and with forward problem solvers that operate in complementary frequency bands. Specifically, we combined full waveform inversion on regional to continental scales with ray tomography that assimilates high-frequency traveltime data that cannot be modelled numerically due to limited
computational power.

While the focus of this contribution is on the European continent, our developments are ultimately designed to feed into the construction of the Comprehensive Seismic Earth Model (CSEM), a 3D global multi-scale Earth model. As illustrated in Figure 2, the goal of the CSEM is to represent the Earth on all seismically accessible scales, containing high resolution sub-regions where data and computational limitations allow, and showcasing a smooth 3D Earth in regions where detailed inversions have yet to be performed. By building a high resolution Earth model in this manner, we sidestep the scaling issues outlined in the introduction, which become even more severe when inversions move to the global scale. In addition to the multiscale model of Europe, the CSEM also includes high-resolution images of Japan, Australia, and the South Atlantic, with inversions of Africa and North America currently underway.

With such a variety of models across the scales, analysing the globally-varying resolution becomes a particular challenge. Methods based on second-order adjoints can provide an approximation of resolution in full-waveform inverse problems [40], but the theory extending this method to general multi-scale multi-solver problems remains to be developed.

The incorporation of ambient noise correlations into full waveform inversion constitutes another future challenge. While travel times are only weakly affected by heterogeneous and varying noise sources [45], complete waveforms that feed into full waveform inversion are strongly
Figure 5: Comparison of ambient noise correlations (black) and numerical Green functions (red) for three different station pairs. Rarely observed matching P and S wave arrivals for the station pair TIR-BFO are marked by dashed lines.

distorted when the noise source distribution is not isotropic [46, 47]. Properly accounting for the distribution of noise sources is therefore essential before correlations can be used to improve earthquake-based full waveform inversion.

Acknowledgements: This research was supported by the Swiss National Supercomputing Center (CSCS) in the form of the GeoScale and CH1 projects, by the Swiss National Science Foundation (SNF) under grant 200021_149143, and by the Netherlands Organisation for Scientific Research (VIDI grant 864.11.008).

REFERENCES


