Iterative Helmholtz Surface Wave Tomography Using Generalized Wave Equation Smoothing Splines - Application to Western Europe

A Sergeant, K Delage, Sébastien Chevrot

JGR Solid Earth

RESEARCH ARTICLE
10.1029/2024JB030087

Iterative Helmholtz Surface Wave Tomography Using Generalized Wave Equation Smoothing Splines—Application to Western Europe

A. Sergeant[1], **K. Delage**[1], **and S. Chevrot**[1]

[1]Geosciences Environment Toulouse, UMR 5563, CNRS, Observatoire Midi-Pyrénées, Université Paul Sabatier, Toulouse, France

Key Points:

- We define generalized Helmholtz and transport equation smoothing splines to improve Helmholtz tomography
- We use an iterative Helmholtz tomography approach to obtained phase velocity maps for Rayleigh waves at periods between 25 and 120 s
- Compared with eikonal tomography, velocity anomalies obtained by Helmholtz tomography are stronger and more sharply defined

Supporting Information:

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

Correspondence to:

S. Chevrot,
sebastien.chevrot@get.omp.eu

Citation:

Sergeant, A., Delage, K., & Chevrot, S. (2025). Iterative Helmholtz surface wave tomography using generalized wave equation smoothing splines—Application to Western Europe. *Journal of Geophysical Research: Solid Earth*, *130*, e2024JB030087. https://doi.org/10.1029/2024JB030087

Received 21 AUG 2024
Accepted 1 FEB 2025

Author Contributions:

Conceptualization: S. Chevrot
Data curation: A. Sergeant, K. Delage
Formal analysis: A. Sergeant
Funding acquisition: S. Chevrot
Investigation: A. Sergeant, S. Chevrot
Methodology: A. Sergeant, K. Delage, S. Chevrot
Project administration: S. Chevrot
Resources: S. Chevrot
Software: A. Sergeant, K. Delage, S. Chevrot
Supervision: S. Chevrot
Validation: A. Sergeant, K. Delage, S. Chevrot
Visualization: A. Sergeant

Abstract Over the last decade, the expansion and densification of broadband seismic networks has led to an increased interest in tomographic approaches based on gradiometry, which relies on the spatial derivatives of the phase and amplitude of surface wavefronts. Nevertheless, the use of amplitude information remains challenging. In this study, we present a novel approach to Helmholtz tomography based on the use of generalized wave equation smoothing splines. The application of these generalized smoothing splines results in the generation of smooth amplitude and phase fields that satisfy the Helmholtz and transport equations, and hence the wave equation. This enables the direct derivation of phase velocity maps using the Helmholtz equation. We apply this new Helmholtz tomography approach to Rayleigh waves recorded by permanent and temporary networks in Western Europe from 2008 to 2022 to obtain phase velocity maps for periods between 25 and 120 s. These maps reveal the detailed structure of the crust and upper mantle in this region. Compared to eikonal tomography, the phase velocity anomalies obtained by Helmholtz tomography are stronger in amplitude and more sharply defined.

Plain Language Summary Over the past decade, the densification of seismic networks has had a profound impact on the way tomography of the Earth's interior is performed. In particular, methods based on the direct exploitation of the phase and amplitude of surface waves recorded by dense seismic networks have led to significant improvements in the spatial resolution of tomographic models. This tomographic approach is also easier to implement because it simplifies to a 2D interpolation problem. However, exploiting the amplitude information remains challenging. Here, we introduce a new method to interpolate phase and amplitude measurements on surface waves using the equations that describe wave propagation as a constraint. From the interpolated phase and amplitude fields, we can easily and efficiently derive 2D maps of the phase velocity over the region of interest. We have applied this new method to data from Western Europe collected between 2008 and 2022, focusing on Rayleigh waves with periods between 25 and 120 s. The phase-velocity maps show the structures of the Earth's crust and upper mantle in this region in great detail. Compared to classical tomographic methods, the velocity anomalies are stronger and more sharply defined.

1. Introduction

Modern seismology is characterized by a rapid and massive densification of temporary and permanent seismic networks, which has significantly impacted the practice of seismic tomography. Eikonal tomography, in particular, has regained interest, driven by the remarkable results obtained by exploiting the long period surface waves recorded by the USArray transportable array (Jin & Gaherty, 2015; Lin & Ritzwoller, 2011; Lin et al., 2009; Liu & Holt, 2015). Eikonal tomography belongs to a new class of wavefield-based tomography approaches that take advantages of seismic wavefields recorded by dense regional arrays (Zhan et al., 2018). The basic idea of eikonal tomography is that if we can characterize the phase of surface wavefronts, we can obtain a phase velocity model from the phase gradient, without having to compute synthetic seismograms or solving a large inverse tomographic problem (Lehujeur & Chevrot, 2020). The model is simply derived from the observed variations of the phase across the array, hence the name gradiometry sometimes given to this approach. Furthermore, eikonal tomography does not require a precise description of the complexity of the incoming wavefields, since they can be approximately reconstructed using appropriate boundary conditions when interpolating the phase measurements with smoothing splines (Chevrot & Lehujeur, 2022).

Writing – original draft: A. Sergeant, S. Chevrot
Writing – review & editing: A. Sergeant, S. Chevrot

Initial concerns on eikonal tomography were raised by Wielandt (1993), who argued that amplitude corrections, which involve computing the Laplacian of the amplitude field, are necessary to ensure accurate reconstruction of phase velocities. However, Lin et al. (2009) have shown that eikonal tomography can yield phase velocity models comparable to those obtained by Helmholtz tomography. This apparent contradiction was resolved by Lehujeur and Chevrot (2020), who demonstrated that the azimuthal average of amplitude corrections tends to zero. Therefore, provided that a sufficient number of surface wavefronts coming from different azimuths are considered, eikonal tomography leads to a good approximate but smooth reconstruction of the phase velocity model. Lehujeur and Chevrot (2020) also demonstrated that, using the complete Helmholtz equation, a perfect reconstruction of the model is achieved with a single source. However, in their synthetic experiment, the phase and amplitude were perfectly known. It therefore remains to be determined whether such a result can be reproduced with real data, from incomplete and irregular sampling of the phase and amplitude fields at station positions, and with measurements potentially contaminated by noise.

This study aims to introduce a novel iterative approach to Helmholtz tomography, utilizing a new class of smoothing splines satisfying the transport and Helmholtz equations for interpolating the amplitude and phase fields. With this new tomographic approach, we perform a Helmholtz tomography of Western Europe using teleseismic surface waves recorded by permanent and temporary networks from 2008 to 2022. To deal with the spatially and temporally variable spatial distribution of stations recording each event, we implement a weighting scheme based on the local station density, which allows the determination of phase velocity maps with fine spatial resolution for periods between 25 and 120 s. We compare the results of eikonal tomography to those of Helmholtz tomography, and discuss the benefits of using the amplitude information. Finally, we provide preliminary interpretations of our phase velocity maps for Western Europe.

2. Helmholtz Tomography: Theoretical Background

In a smoothly varying Earth, Love and Rayleigh fundamental mode displacement wavefields for angular frequency ω are given by a 2D field $\mathbf{u}$ that satisfies the wave equation (Tanimoto, 1990; Tromp & Dahlen, 1993)

$$\Delta u(\mathbf{r}) + \frac{\omega^2}{c^2} u(\mathbf{r}) = 0, \tag{1}$$

where Δ is the Laplacian operator, $\mathbf{r}$ the position, and c the local Love or Rayleigh phase velocity.

A simple proxy solution of Equation 1 is given by

$$u(\mathbf{r}) = A(\mathbf{r}) e^{i\omega T(\mathbf{r})}, \tag{2}$$

where $A(\mathbf{r})$ and $T(\mathbf{r})$ are the amplitude and phase of u at location $\mathbf{r}$. Substituting 2 into 1 leads to (e.g., Červený, 2001)

$$\Delta \mathbf{A} - \omega^2 \mathbf{A} \|\nabla \mathbf{T}\|^2 + \frac{\omega^2}{c^2} \mathbf{A} + i\omega (2\nabla \mathbf{A}.\nabla \mathbf{T} + \mathbf{A}\Delta \mathbf{T}) = 0, \tag{3}$$

where ∇ is the gradient operator, and the spatial dependence of $\mathbf{A}$ and $\mathbf{T}$ have been omitted.

From the imaginary part of 3, we derive the transport equation

$$2\nabla \mathbf{T} \cdot \nabla \mathbf{A} + \mathbf{A}\Delta \mathbf{T} = 0, \tag{4}$$

and from the real part, the Helmholtz equation

$$\mathbf{s}^2 = \|\nabla \mathbf{T}\|^2 - \frac{\Delta \mathbf{A}}{\omega^2 \mathbf{A}}, \tag{5}$$

where $s = 1/c$ is the surface wave slowness. Following Lehujeur and Chevrot (2020), we define $\mathbf{a} = \ln \mathbf{A}$. Using this definition

$$\frac{\Delta \mathbf{A}}{\mathbf{A}} = \|\nabla \mathbf{a}\|^2 + \Delta \mathbf{a} \approx \Delta \mathbf{a}, \tag{6}$$

because to first order the squared gradient is much smaller than the Laplacian (Lehujeur & Chevrot, 2020). We can therefore rewrite the Helmholtz equation as

$$\mathbf{s}^2 = \|\nabla \mathbf{T}\|^2 - \frac{\Delta \mathbf{a}}{\omega^2}. \tag{7}$$

Similarly, since

$$\frac{\nabla \mathbf{A}}{\mathbf{A}} = \nabla \mathbf{a}, \tag{8}$$

we can rewrite the transport equation as

$$2\nabla \mathbf{T} \cdot \nabla \mathbf{a} + \Delta \mathbf{T} = 0. \tag{9}$$

The main idea of Helmholtz tomography is to derive phase velocities directly from the Helmholtz equation, that is, from the spatial variations of the phase and amplitude of surface waves, thus avoiding the computation of synthetic seismograms and solving a large tomographic inverse problem as in classical surface wave tomography approaches.

3. Reconstruction of Phase and Amplitude Surfaces With Smoothing Splines

Helmholtz tomography requires the reconstruction of travel time and amplitude fields by interpolating irregularly distributed measurements at the position of the seismic stations onto a regular grid. The idea of using a regular grid is that spatial derivations can be performed with a simple finite-difference scheme. The algorithm described below can be implemented on either a Cartesian or a spherical grid, depending on the size of the domain to be imaged. So far, interpolation relied mainly on splines, which ensure minimum curvature of the derived surface. Lin et al. (2009) and Lin and Ritzwoller (2011) were the first to use splines in tension to reconstruct travel time surfaces. The main issue with splines in tension is that the interpolated field is imposed to fit the data exactly at the station locations. Since phase and amplitude measurements are noisy, this interpolation approach will produce small-scale oscillations of the interpolated fields, which will be amplified when computing spatial derivatives. A better choice is to use smoothing splines, which search for an optimal compromise between fitting the measurements and minimizing the curvature of the interpolated surface (Chevrot & Lehujeur, 2022; Hastie et al., 2017). Compared with splines in tension, the interpolated surfaces obtained with smoothing splines are smoother and less sensitive to noise measurements.

3.1. Definition of Smoothing Splines

Let us consider N samples of variable y sampled at positions $r_i, i = 1, \ldots, N$. A smoothing spline $f(\mathbf{r})$ is a function with two continuous derivatives that minimizes

$$\chi(f, \lambda) = \sum_{i=1}^{N} \left[y_i - f(r_i)\right]^2 + \mu \int \|\Delta f(\mathbf{r})\|^2 dr^2 \tag{10}$$

where μ is the smoothing parameter, which can be determined for example, by generalized cross-validation (Chevrot & Lehujeur, 2022; Hastie et al., 2017).

3.2. Interpolation of Travel Times With Smoothing Splines

Lehujeur and Chevrot (2020) have demonstrated that it is possible to approximately reconstruct phase velocity models from the eikonal equation, using only the gradient of the phase, if a sufficient number of wavefronts coming from different azimuths are used.

Let us formulate the smoothing interpolation problem for the reconstruction of travel time fields of surface waves. We search for the travel time field $\mathbf{T}$ that represents an optimal compromise between fitting the observed travel times $\tilde{\mathbf{T}}$ and minimizing the curvature of the interpolated field, quantified by the norm of the Laplacian of $\mathbf{T}$. At the edges of the grid we impose Neumann boundary conditions such that the gradient of the phase along the normal direction is equal to the apparent slowness along that direction. We introduce the derivative boundary operator $\mathbf{B}$ such that $\mathbf{B}\cdot\mathbf{T} = s_x$ at the left and right boundaries, and $\mathbf{B}\cdot\mathbf{T} = s_y$ at the bottom and top boundaries. Using these different notations, the smoothing spline problem for $\mathbf{T}$ can be written

$$\chi(\mathbf{T}) = \|\mathbf{P}\cdot\mathbf{T} - \tilde{\mathbf{T}}\|^2 + \|\mathbf{B}\cdot\mathbf{T} - \mathbf{b}\|^2 + \mu\|\mathbf{L}\cdot\mathbf{T}\|^2, \tag{11}$$

where $\mathbf{L}$ is the discretized Laplacian operator, $\mathbf{P}$ the operator that samples $\mathbf{T}$ at the station positions, and $\mathbf{b}$ a vector containing the values of the boundary conditions (s_x or s_y).

The solution of the smoothing spline for $\mathbf{T}$ is then given by

$$\mathbf{T} = \left[\mathbf{P}^\mathbf{T}\mathbf{P} + \mathbf{B}^\mathbf{T}\mathbf{B} + \mu\mathbf{L}^\mathbf{T}\mathbf{L}\right]^{-1}\left[\mathbf{P}^\mathbf{T}\tilde{\mathbf{T}} + \mathbf{B}^\mathbf{T}\mathbf{b}\right], \tag{12}$$

where $^\mathrm{T}$ denotes the transpose operator.

In practice, instead of interpolating the phase directly, it is more convenient to estimate the phase field residuals $\boldsymbol{\delta}\mathbf{T}$, which represent the deviations of the measured phase from the phase of a plane wave propagating in a homogeneous model with slowness $\mathbf{s}_0$. $\mathbf{T}$ is then replaced by $\boldsymbol{\delta}\mathbf{T}$ in 11, with Neumann boundary conditions set to zero on the edges of the grid ($\mathbf{b} = \mathbf{0}$). In that case, the solution of the smoothing spline for $\boldsymbol{\delta}\mathbf{T}$ is

$$\boldsymbol{\delta}\mathbf{T} = \left[\mathbf{P}^\mathbf{T}\mathbf{P} + \mathbf{B}^\mathbf{T}\mathbf{B} + \mu\mathbf{L}^\mathbf{T}\mathbf{L}\right]^{-1}\mathbf{P}^\mathbf{T}\boldsymbol{\delta}\tilde{\mathbf{T}}. \tag{13}$$

We estimate the optimal value of the regularization coefficient μ by generalized cross-validation (GCV) (Craven & Wahba, 1978; Hastie et al., 2017; Tenorio, 2001). For each tested value of μ, we compute the GCV prediction error (Chevrot & Lehujeur, 2022; Hastie et al., 2017; Wahba, 1990)

$$\mathrm{Error}_{\mathrm{gcvd}}(\mu) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{T}_i^\mu - T_i}{1 - \mathrm{trace}(\mathbf{S})/N}\right)^2 \tag{14}$$

where N is the number of stations, $\hat{T}_i^\mu$ the predicted phase value at station i using smoothing parameter μ, T_i the phase measured at station i, and

$$\mathbf{S} = \mathbf{P}\left(\mathbf{P}^T\mathbf{P} + \lambda\mathbf{L}^T\mathbf{L}\right)^{-1}\mathbf{P}^T \tag{15}$$

a symmetric, positive definite matrix, often referred to as the influence (Wahba, 1990) or smoother (Hastie et al., 2017) matrix. The value of the smoothing parameter μ which minimizes the GCV prediction error represents a compromise between fitting the data and minimizing the curvature of the reconstructed surface. It is mainly controlled by data distribution (i.e., number of data points, station coverage and density, and noise in travel time measurements) and by the grid size.

3.3. Interpolation of Amplitudes With Smoothing Splines

Isotropic phase velocity maps can also be derived with Helmholtz tomography, using the full Helmholtz equation (e.g., Lin et al., 2009). This requires the computation of amplitude corrections and therefore an interpolation of the amplitude measurements $\tilde{\mathbf{a}}$. Since it is more convenient to work with relative amplitudes (or amplitude residuals), we consider $\mathbf{a} = \ln\mathbf{A} - \ln A_0$, where A_0 is the average absolute amplitude across the network. Following an approach similar to the one described above for the phase, we search for the amplitude field $\mathbf{a}$ that represents an optimal compromise between fitting the observed amplitudes $\tilde{\mathbf{a}}$ and minimizing the norm of the Laplacian of $\mathbf{a}$. At the edges of the grid we impose Neumann boundary conditions such that the gradient of the amplitude along the

normal direction is zero. We introduce the derivative boundary operator $\mathbf{B}$ so that $\mathbf{B}\cdot\mathbf{a} = 0$ at the four edges. Using these different notations, the smoothing spline problem for $\mathbf{a}$ can be written

$$\chi(\mathbf{a}) = \|\mathbf{P}\cdot\mathbf{a} - \tilde{\mathbf{a}}\|^2 + \|\mathbf{B}\cdot\mathbf{a}\|^2 + \mu\|\mathbf{L}\cdot\mathbf{a}\|^2. \tag{16}$$

The solution of the smoothing spline problem for $\mathbf{a}$ is given by

$$\mathbf{a} = \left[\mathbf{P}^\mathbf{T}\mathbf{P} + \mathbf{B}^\mathbf{T}\mathbf{B} + \mu\mathbf{L}^\mathbf{T}\mathbf{L}\right]^{-1}\mathbf{P}^\mathbf{T}\tilde{\mathbf{a}}. \tag{17}$$

4. Generalized Wave Equation Smoothing Splines

In this section, we generalize the definition of smoothing splines, in order to reconstruct phase and amplitude fields that satisfy the Helmholtz and transport equations, and consequently the wave equation.

4.1. Helmholtz Smoothing Splines

As we just saw in the previous section, smoothing splines for the amplitude are obtained by imposing the constraint

$$\Delta\mathbf{a} = 0. \tag{18}$$

Consequently, the interpolated amplitude field $\mathbf{a}$ will approximately fit the observed amplitudes and minimize 18. Classical smoothing splines as defined in 17 therefore search for an amplitude field that minimizes the amplitude correction in the Helmholtz equation, which will lead to underpredicted amplitude corrections. This simply explains the rather small differences that have been observed so far between eikonal and Helmholtz tomography (e.g., Lin et al., 2009). To impose that the amplitude field satisfies the Helmholtz equation, we only need to replace the constraint 18 by

$$\Delta\mathbf{a} = \omega^2\left(\|\nabla\mathbf{T}\|^2 - \mathbf{s}^2\right). \tag{19}$$

This expression defines a new type of smoothing splines, hereafter called Helmholtz smoothing splines, which give smooth amplitude fields satisfying the Helmholtz equation (as they should). We note that to compute Helmholtz smoothing splines, we need the slowness model $\mathbf{s}^2$. This suggests a simple algorithm for the implementation of Helmholtz tomography, which consists of first performing eikonal tomography with travel time and amplitude fields reconstructed with classical smoothing splines. The average squared slowness model $\bar{\mathbf{s}}^2$ derived from all events can then be used to compute Helmholtz smoothing splines for the amplitude fields. This provides amplitude corrections that can be used to derive a new slowness model from the complete Helmholtz equation.

Helmholtz smoothing splines minimize the cost function

$$\chi(\mathbf{a}) = \|\mathbf{P}\cdot\mathbf{a} - \tilde{\mathbf{a}}\|^2 + \|\mathbf{B}\cdot\mathbf{a} - \mathbf{b}\|^2 + \mu\|\mathbf{L}\cdot\mathbf{a} - \omega^2\left(\|\nabla\mathbf{T}\|^2 - \bar{\mathbf{s}}^2\right)\|^2, \tag{20}$$

where $\|\nabla\mathbf{T}\|^2$ is the norm of the phase gradient computed in the eikonal approximation (i.e., using the travel time smoothing spline as defined earlier), and $\bar{\mathbf{s}}^2$ is the squared-slowness model obtained by eikonal tomography. At the edges of the grid, we apply Neumann boundary conditions set to 0, so that the amplitude field slowly decreases toward the average amplitude far from the sensor array.

The solution for the amplitude field computed with Helmholtz smoothing splines is given by

$$\mathbf{a} = \left[\mathbf{P}^\mathbf{T}\mathbf{P} + \mathbf{B}^\mathbf{T}\mathbf{B} + \mu\mathbf{L}^\mathbf{T}\mathbf{L}\right]^{-1}\left[\mathbf{P}^\mathbf{T}\tilde{\mathbf{a}} - \mu\mathbf{L}^\mathbf{T}\omega^2\left(\|\nabla\mathbf{T}\|^2 - \bar{\mathbf{s}}^2\right)\right]. \tag{21}$$

4.2. Transport Smoothing Splines

We can follow a similar approach to define a new type of smoothing splines to interpolate phase measurements so that the phase field satisfy the transport Equation 9. To do so, we just need to replace the constraint

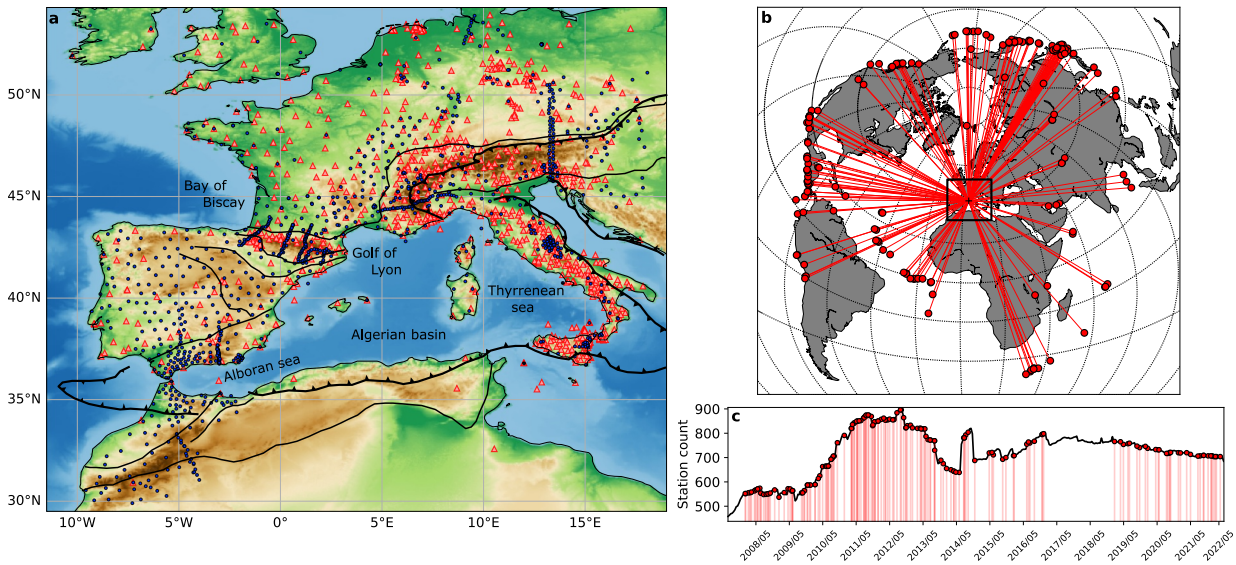

Figure 1. (a) Map showing the study region, surface topography, main tectonic features, and seismic station locations. Stations from permanent and temporary networks are indicated by red triangles and blue circles, respectively. Black lines indicate sutures and subduction trenches, thin black lines indicate deformation fronts (modified after Lanari et al. (2023)). (b) Epicenters of events used in this study. (c) Number of recording stations as a function of time. The earthquakes selected are indicated by red vertical lines.

$$\Delta \mathbf{T} = 0 \tag{22}$$

by

$$2\nabla \mathbf{a} \cdot \nabla \mathbf{T} + \Delta \mathbf{T} = 0, \tag{23}$$

that is, changing the Laplacian operator $\mathbf{L}$ by a new operator $\mathbf{K} = 2\nabla \mathbf{a} \cdot \nabla + \mathbf{L}$ in 11. The solution for the phase field computed with transport smoothing splines then becomes

$$\mathbf{T} = \left[\mathbf{P}^{\mathbf{T}}\mathbf{P} + \mathbf{B}^{\mathbf{T}}\mathbf{B} + \mu \mathbf{K}^{\mathbf{T}}\mathbf{K}\right]^{-1}\left[\mathbf{P}^{\mathbf{T}}\tilde{\mathbf{T}} + \mathbf{B}^{\mathbf{T}}\mathbf{b}\right]. \tag{24}$$

Note that to determine transport smoothing splines the amplitude field is needed. This naturally suggests an iterative approach during which the phase and amplitude fields will be updated alternatively, in order to finally converge toward phase and amplitude fields that will satisfy both the Helmholtz and the transport equations.

5. Application: Eikonal and Helmholtz Tomography of Western Europe

5.1. Seismic Data Set

We use all publicly available (last access: November 2022) medium-band (cut-off period of 30, 40, and 60 s), broad-band (cut-off period of 120 s) and ultra broad-band stations (cut-off period of 360 s) recording in Western Europe in the time period 2008–2022 (Figure 1). Note that we discarded the year 2017 and part of 2018, when the French permanent network and the AlpArray deployment were not completed. In total, we collected data from 2033 stations (938 stations from permanent national networks and 1,095 stations from temporary deployments). The data set includes records from 78 broadband permanent stations in Spain, operated by the Instituto Geográfico Nacional. In Northern Marocco, Spain and across the Gibraltar arc, we use temporary stations from the IberArray (IB network, 2008–2013) (Institute Earth Sciences "Jaume Almera" CSIC (ICTJA Spain), 2007) and PICASSO (XB network, 2010–2012) (Levander et al., 2009) deployments. Over the Pyrénées and southern France, we use stations from PYROPE (X7 network, 2010–2013) (Chevrot et al., 2017) and Orogen-X (ZU network, 2015–2016) (Chevrot, Diaz, et al., 2018) experiments. Across the Alps, we use stations from AlpArray (Z3 network, 2016–2020) (AlpArray Seismic Network, 2015), CIFALPS (YP network, 2012–2013) (Zhao, Paul,

Solarino, & RESIF, 2016) and EASI experiments (XT network, 2014–2015) (AlpArray Seismic Network, 2014). Most of the temporary networks were not deployed simultaneously. The IberArray, PYROPE, and Orogen-X profiles were deployed in successive stages covering different areas of the Iberian peninsula and the Pyrénées (see Chevrot et al. (2014) and Chevrot, Sylvander, et al. (2018) for a complete description). Consequently, the station distribution is uneven and varies over time. The average station spacing is ∼60 km.

We select all teleseismic shallow (hypocentral depth ≤50 km) events with magnitude greater than 6.0 in the epicentral distance range $25° - 95°$ from the center of the region. We discard closer events to avoid near-source effects and interference with higher modes (Hariharan et al., 2020, 2022). We extract the vertical component of all the waveforms recorded by the stations within the region shown in Figure 1a. The seismograms are then deconvolved from their station response and resampled at 1 Hz.

5.2. Data Analysis

To perform the event selection, we first isolate the fundamental mode of Rayleigh waves using the method of Chevrot and Lehujeur (2022). The waveforms are bandpass filtered between 20 and 50 s and aligned using a time-shift corresponding to the apparent slowness determined by beamforming. The fundamental mode is then isolated using a taper obtained by taking the envelope of the average trace. We compute the correlation coefficient between the extracted waveforms and their average, and keep events with an average correlation coefficient greater than 0.8. We then visually inspect the record sections and keep the events for which a clear and simple wavefront is observed. Our final selection contains 174 teleseismic events (Figure 1b). The azimuthal distribution is quite homogeneous, so we do not expect a significant contribution from azimuthal anisotropy in our final average phase velocity maps.

5.2.1. Extraction of Fundamental Mode

The extraction of Rayleigh wave traveltimes and amplitudes at each period consists of two steps. First, we isolate the fundamental mode using a phase-matching filtering approach based on Rayleigh group velocity dispersion curves. We then decompose the fundamental mode Rayleigh wave into a set of monochromatic waves for each investigated period to measure their arrival time and amplitude. To isolate the fundamental mode, we use the frequency–time analysis (FTAN) of Levshin et al. (1992). FTAN involves applying a series of narrow-band Gaussian filters at multiple frequencies ω_i to obtain a time-frequency dispersion diagram. The width of the Gaussian filters varies as a function of frequency, according to the function $\alpha(\omega_i)$ from Levshin et al. (1992). To mitigate interference effects at short period, we use high values of α which provides very narrow-band filters. We set $\alpha = 80$ at period 15 s and then impose a logarithmic decrease with increasing periods so that α reaches 20 at 80 s period, above which group velocity dispersion curves are nearly flat. Using larger narrow-band filters at long periods gives better temporal resolution. The group velocity dispersion curve is then given by the local maximum amplitudes of the dispersion diagram (Figure 2b). It can sometimes be difficult to automatically pick the continuous dispersion curve on individual records. Therefore, for each record, we use an average group velocity curve determined by averaging the dispersion diagrams from the 20 nearest stations within a 200 km radius, following P. Zhou et al. (2022). Finally, to extract the fundamental mode, we apply a time-frequency phase-matching filter based on the average dispersion curve to the raw waveforms. By doing so, the resulting waveforms are cleaned of contamination from other phases (Figure 2a).

5.2.2. Phase and Amplitude Measurements

To measure the travel time and amplitude of the fundamental mode of Rayleigh waves, we use a method similar to Kolínský and Bokelmann (2019). The recordings from each station are processed separately. To extract the surface wave train at each period, we use narrow-band Gaussian filters in the frequency domain. As mentioned in Section 5.2.1, the width of the Gaussian filters can be tuned to the frequency. Kolínský (2004) argues that they should be adjusted according to the properties of the signal being analyzed. For simplicity, we use the same filter parameters for all the events, with $\alpha = 50$ for all frequencies. After the Gaussian filtering, each waveform spectrum is transformed back into the time domain. This results in a collection of quasi-monochromatic signals. The frequency of the extracted signals is not constant due to the Gaussian filtering (Kolínský et al., 2014). Therefore, in the remainder of the analysis, instead of the center frequencies of the filters, we use the

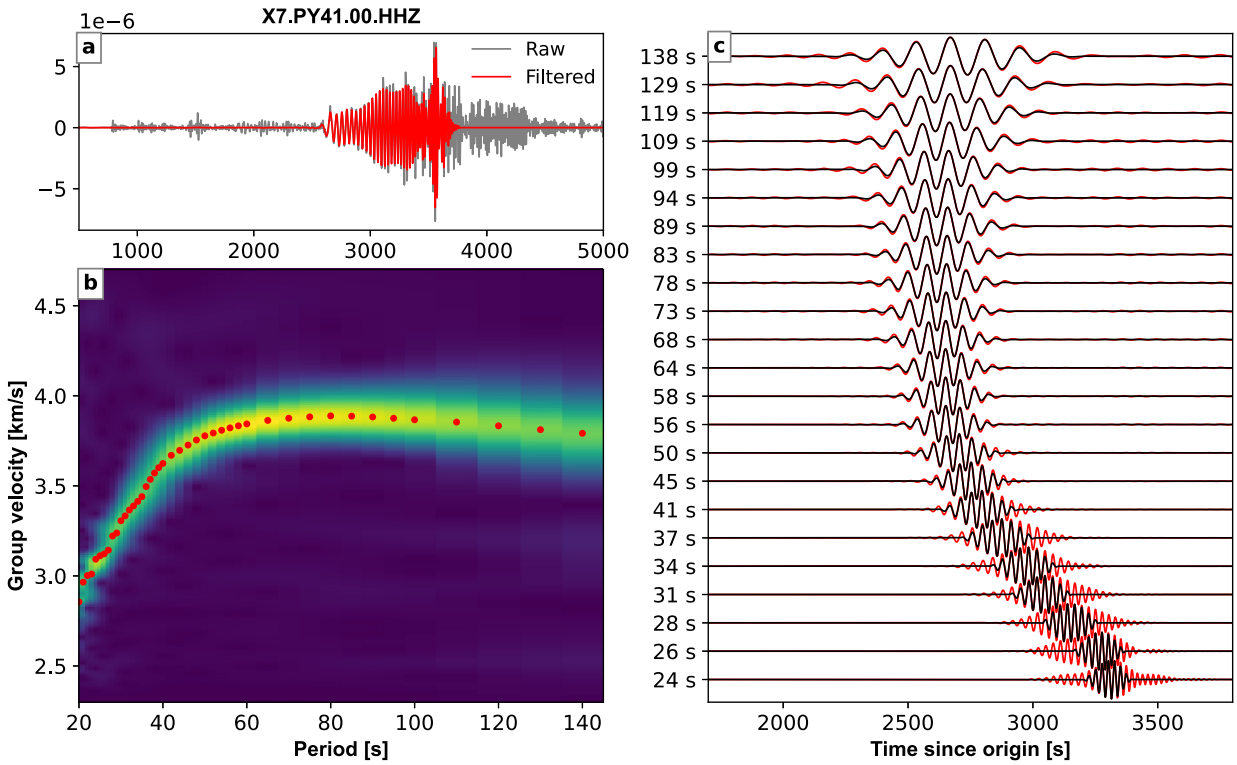

Figure 2. Time-frequency filtering to isolate the fundamental mode wave train. Results are shown for the M_w = 6.2, 2011 August 17, earthquake near the island of Honshu, Japan, recorded at station X7.PY41. (a) Vertical component seismograms before (gray) and after (red) application of the isolation filter. (b) FTAN diagram of the raw waveform. Red points indicate the picked group velocity dispersion curve. (c) Quasi-monochromatic signals (red) extracted from the isolated waveform, compared to tapered signals (black), as a function of time and period. To taper the signals, we use cosine windows of four period-length, centered on the envelope maximum of the analytic signal. The tapered waveforms are used to measure the travel times and amplitudes from their Fourier spectra at each period.

instantaneous frequencies computed from the analytical signal of each extracted quasi-monochromatic waveform, following Levshin et al. (1992).

To extract the phase arrivals, we taper the signal twice and look at the waveform maxima. First, we use a large taper window of eight periods length, centered on the arrivals predicted by the averaged group velocity curve computed for the region in the previous step. This gives us a rough isolation of the fundamental mode at this frequency (Figure 2c, red lines), which still needs to be explored to accurately pick the wavefront arrival. For periods longer than 60 s, the fundamental mode is characterized by a flat dispersion curve and we usually observe a single envelope maximum in the time window. At higher frequencies, we sometimes observe several groups of waves that are not easy to separate. To pick the arrivals, we calculate the envelope of the extracted signals from the modulus of the analytical signals. This envelope is then used to further reduce the signals around the identified envelope maximum using a four-period length cosine window, similar to Kolínský and Bokelmann (2019) (Figure 2c, black lines). To accurately select the maximum corresponding to the fundamental mode at all frequencies, we use a continuous ridge selection method. We start by windowing the signal at the lowest frequency and then progressively consider higher frequencies. As we move to a higher frequency, we always select the maximum that is closest in time to the maximum selected at the previous frequency. The set of quasi-monochromatic signals previously extracted is then limited to the fundamental mode wave train at each period (see example of a record section before and after cleaning for the 40 s period in Figures 3a and 3b). The phase and amplitude of the Rayleigh waves are then measured from the phase and amplitude of the Fourier transform of the recordings of the isolated fundamental mode wave trains.

In the last step, we unwrap the phase (Figure 3d). To do so, the phase measurements are sorted by increasing epicentral distance. We start with the station closest to the source and successively add the phase difference between neighboring stations. The phase measurements are then converted into arrival times by dividing by the angular frequency. In the following, when we use the term "phase," we will refer specifically to the travel time of

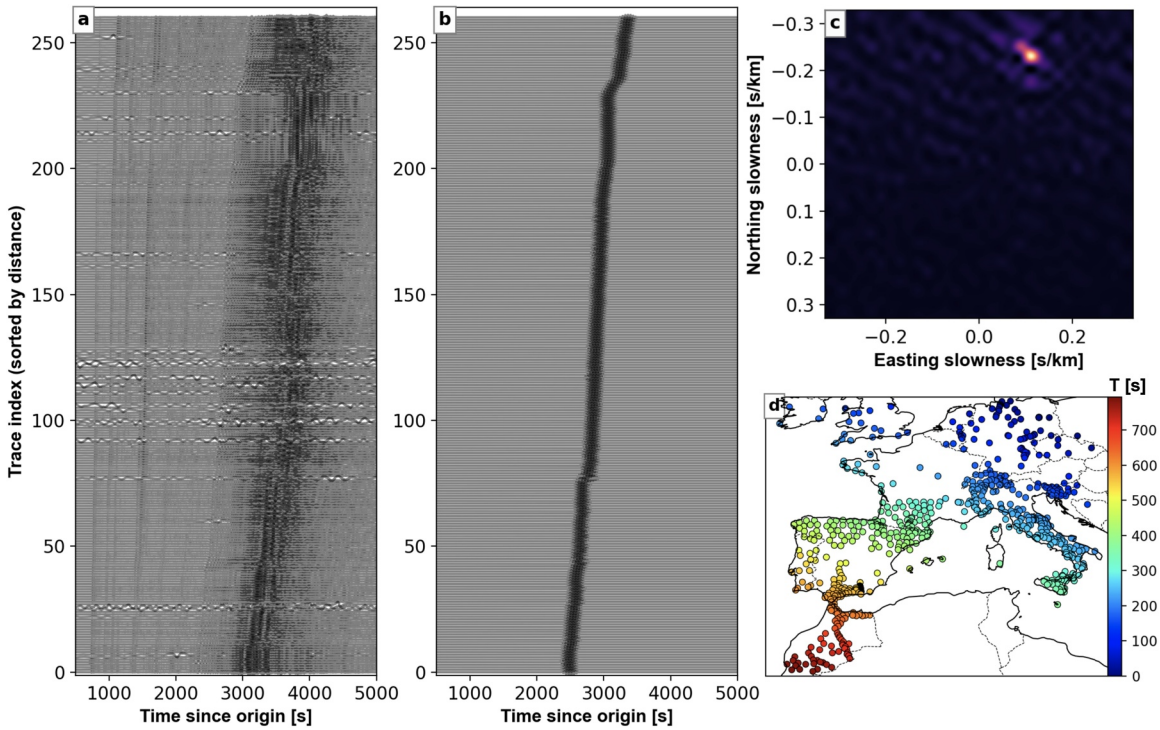

Figure 3. Filtering of Rayleigh waves and travel time measurements for the 2011 Japan event at period 40 s. (a) Raw vertical component seismic records, low-pass filtered at 10 s. (b) Tapered signals extracted for period 40 s, after time–frequency filtering. (c) Results of the beamforming for the record section shown in panel (b). The main beam gives a slowness of 0.256 s/km, in good agreement with the expected slowness of a fundamental-mode Rayleigh wave at 40 s. The propagation direction points to the source region in Japan. (d) Travel time measurements relative to the station closest to the source, after phase unwrapping.

the wavefront. We compute catalogs of amplitudes and arrival times for each event and for the periods 25 s, 30 s, 40 s, 50 s, 60 s, 70 s, 80 s, 90 s, 100 s, and 120 s.

Because Helmholtz tomography involves the computation of gradients, it is particularly sensitive to the presence of outliers. It is therefore necessary to identify and remove these outliers in order to obtain clean data sets of phase and amplitude measurements. To this end, we have developed an approach based on a statistical description of the spatial variations of phase and amplitude over a regional seismic network, which is described in Appendix A.

5.3. The Iterative Helmholtz Tomography Approach

Since our study region is quite large, we use a spherical grid for interpolating the phase and amplitude measurements, as well as gradient and Laplacian operators expressed in spherical coordinates. We use a spatial discretization such that the dominant wavelength is sampled by a minimum of 15 grid points. For the 25 s period, we use a discretization of 0.09° and 0.06° in longitude and latitude, respectively, corresponding to an average grid cell size of 7 km. For the 30 s period, the grid cell dimensions are set to 0.1° and 0.07° in longitude and latitude, respectively (cell size ∼8 km). For periods 40–120 s, we use a discretization of 0.14° and 0.1° in longitude and latitude, respectively, corresponding to an average grid cell size of 11 km.

As seen in Section 5, Helmholtz smoothing splines require a slowness model, which makes it necessary to use an iterative approach. In the first iteration, the phase and amplitude fields are interpolated with regular smoothing splines, using 13 and 17. We interpolate the phase residuals $\delta\mathbf{T}$, obtained by removing from the observed travel times those predicted by the average plane wave velocity and azimuth measured by beamforming. This means that outside the recording network, the phase gradient gradually reaches the apparent plane wave slowness s_0 toward the edges of the tomographic grid. After interpolation, we compute the gradient of the phase and the Laplacian of the amplitude, from which we derive the slowness map for each event and period according to Equation 7. We then determine the average slowness map $\bar{\mathbf{s}}_l$ for each period by averaging all the individual slowness maps (see details for the stacking procedure in Section 5.5). The results of this first iteration are equivalent to those of

eikonal tomography, as the amplitude corrections derived from smoothing splines, that is, from 13 and 17, have only a small effect on the average slowness model $\bar{\mathbf{s}}_1$.

In the second iteration, the phase and amplitude measurements are interpolated with generalized smoothing splines, using the slowness model $\bar{\mathbf{s}}_1$ obtained in the first iteration. We first interpolate the travel time residuals $\delta\mathbf{T} = \tilde{\mathbf{T}} - \mathbf{T}_{ref}$, where the reference travel times $\mathbf{T}_{ref}$ are computed in the slowness model $\bar{\mathbf{s}}_1$ using a 2D fast marching method (Sethian & Popovici, 1999). The computations are performed on the surface of a spherical Earth (Chen et al., 2023), with the velocities outside the tomographic grid set to the average phase velocity inside. We also update the boundary conditions so that the phase gradient gradually reaches the slowness model $\bar{\mathbf{s}}_1$ at the edges of the grid. Once we have the new phase field $\mathbf{T}$, we can interpolate the amplitudes with Helmholtz smoothing splines, using the gradient of the phase just computed and the slowness model $\bar{\mathbf{s}}_1$. From the gradient of the phase and the Laplacian of the amplitude, we can then derive a new slowness map for each event and period, using the Helmholtz equation. After averaging these maps, we obtain a new average slowness model $\bar{\mathbf{s}}_2$. This whole procedure can be repeated to further refine the average slowness model.

5.4. Examples of Phase and Amplitude Interpolation With Smoothing Splines

Figure 4 shows the interpolated fields $\delta\mathbf{T}$ in (d) and $\mathbf{a}$ in (e), the eikonal squared slowness map in (g), the amplitude correction map $-\Delta\mathbf{a}/\omega^2$ in (h) and the Helmholtz squared slowness map in (i) obtained using classical smoothing splines for interpolating $\delta\mathbf{T}$ and $\mathbf{a}$. These results are for the 40 s Rayleigh wave from a Japan event arriving from the northeast. At this period, Rayleigh waves are mostly sensitive to the 30–80 km depth range. The phase and amplitude measurements, indicated by colored circles, are highly irregularly distributed in space. In regions where dense measurements are available, clear correlations are observed between the residual phase and amplitude fields (Figures 4d and 4e) and the amplitude correction map (Figure 4h). In the western Alps, the amplitude map shows prominent stripes elongated along the propagation direction. This is a simple consequence of the transport equation which predicts that far from the source, where the curvature of the wavefront is small, and hence the Laplacian of $\mathbf{T}$ is small, the gradient of $\mathbf{a}$ is perpendicular to the gradient of $\mathbf{T}$, that is, to the propagation direction. Therefore, the amplitude correction term in the Helmholtz equation takes into account (and corrects for) the scattering effects that can bias phase measurements, even at short period. Indeed, when the amplitude term is taken into account, the stripe interference pattern present in the eikonal slowness map (Figure 4g) is no longer observed, and a well-defined slow anomaly appears, highlighting the deeper crustal roots of the Alpine arc (Figure 4i).

Figure 5 shows the results of the interpolation in the second iteration, using the updated travel time residuals and the average phase slowness map $\bar{\mathbf{s}}_1$ derived at the end of iteration 1 (Figure 5f). The travel time residuals $\delta\mathbf{T}$ in Figure 5d, obtained using transport smoothing splines, show a different pattern compared to those obtained in the first iteration (Figure 4d). As expected, these travel time residuals are smaller as they are computed in a more realistic slowness model. The slowness map derived from the phase gradient (i.e., the eikonal solution in Figure 5g) is similar to the one obtained at the first iteration, except toward the edges of the grid far from any recording station (e.g., from Algeria to Libya, and near the western shores of the Bay of Biscay), where the computed squared slowness tends toward $\mathbf{s}_0^2$. The residual amplitude map shows a large low anomaly extending from the Algerian basin to the southern edge of the grid (Figure 5e). This anomaly, caused by poor station coverage, produces stripes in the amplitude correction map around the Mediterranean islands (Figure 5h). With the amplitude field reconstructed using Helmholtz smoothing splines, such spurious anomalies are now attenuated in the updated phase slowness map (Figure 4i). In summary, using the a priori model $\bar{\mathbf{s}}_1^2$ to reconstruct the amplitude field with Helmholtz smoothing splines allowed us to mitigate artifacts in both the phase gradient and the amplitude Laplacian resulting from the sparse and irregular distribution of the measurements, resulting in more consistent and robust phase velocity maps.

To further illustrate the effect of using Helmholtz smoothing splines to determine the amplitude corrections, Figures 6a–6c show the interpolated amplitude fields obtained at iteration 1 (with classical smoothing splines) and iteration 2 with Helmholtz smoothing splines, using the slowness model $\bar{\mathbf{s}}_1$ obtained by eikonal tomography in the first iteration, for a 2020 Mid-Atlantic Ridge even. At that time, the station coverage was quite good in France, but very poor in Spain. In both cases, the values of the interpolated fields are close to the measurements, which means that both interpolation methods provide similar fits to the amplitude measurements. However, the interpolated field with Helmholtz smoothing splines shows significant differences in regions where seismic stations are absent

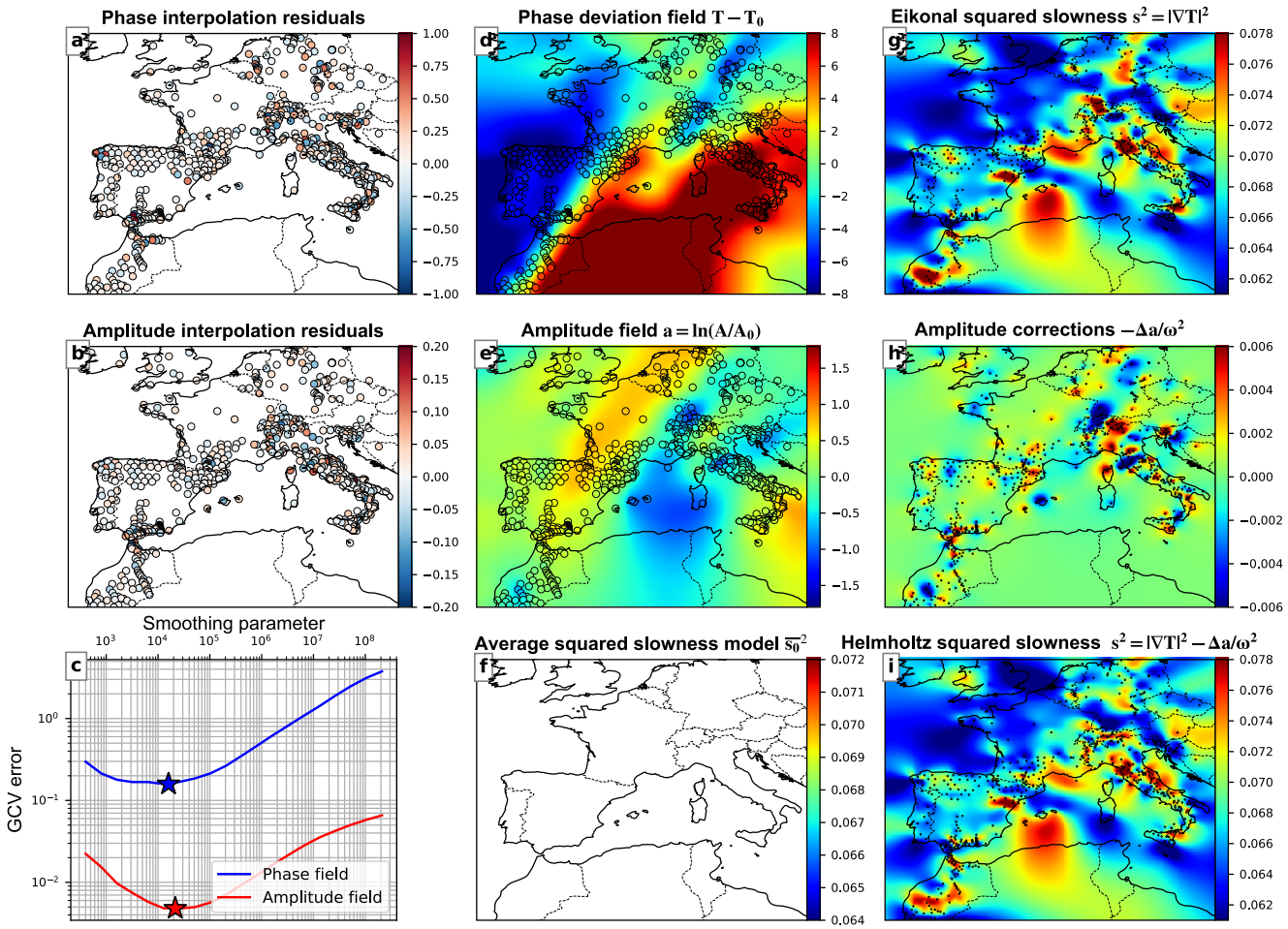

Figure 4. Interpolation results for the first iteration of Helmholtz tomography using classical smoothing splines. Results are for a 40 s Rayleigh wave generated by the 2011 Japan event. (a) Differences between the observed and interpolated travel time field at station locations. (b) Same as (a) but for the amplitude. (c) GCV error as a function of smoothing parameters for the travel times (blue line) and amplitudes (red line). Optimal values of smoothing parameters are indicated by stars. (d) Interpolated phase residual field (deviations from the average plane wave). Observed travel time residuals are indicated by colored circles. (e) Same as (d) but for the interpolated relative amplitude field. (f) Null a priori slowness model (not considered in the first iteration). (g) Eikonal squared apparent slowness map derived from the squared gradient of the phase field. Seismic stations are indicated with black dots. (h) Amplitude correction map derived from the Laplacian of the amplitude field. (i) Helmholtz squared apparent slowness map derived from the squared gradient of the phase (g) plus the amplitude correction term (h).

or very coarsely distributed, such as in northern Spain, the Mediterranean Sea, or the Atlantic Ocean, for example, This is easily explained by the constraint 18, which tends to minimize the curvature and thus the absolute value of the amplitude field between stations. These differences are amplified in the amplitude corrections, since they are derived from the Laplacian of the interpolated fields. The amplitude corrections derived from classical smoothing splines (Figure 6b) are much weaker than those derived from Helmholtz smoothing splines (Figure 6d). For the latter, the information provided by the phase gradient and by the structural model have at least partly compensated for the irregular station distribution to improve the reconstruction of the amplitude field. Simply minimizing the curvature of the amplitude field (i.e., using regular smoothing splines) leads to highly biased and underestimated amplitude corrections.

5.5. Stacking Scheme

In this section we detail the procedure for stacking the apparent slowness maps from all individual events to obtain robust average slowness (or phase velocity) models at the end of each iteration and for each period.

Prior to averaging, we assess the quality of the individual slowness maps and exclude events or individual grid cells that show patterns of unusually high variability in slowness values. Event maps with a mean phase slowness

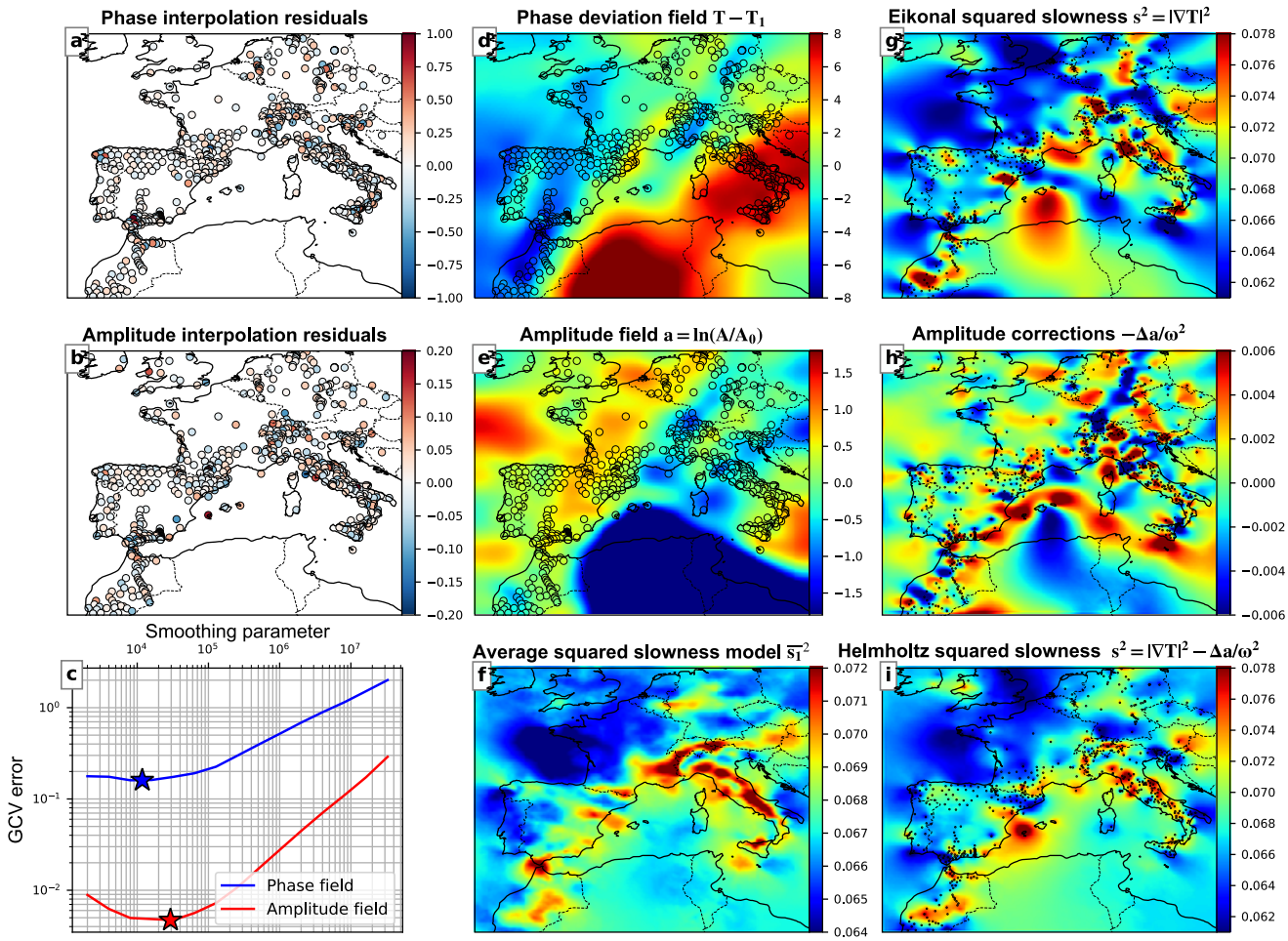

Figure 5. Same as Figure 4 but for the interpolations performed in the second iteration of Helmholtz tomography using generalized wave smoothing splines. Iteration 2 uses the average phase slowness model $\bar{s}_1$ (f) to (1) update the traveltime residuals in panel (d) that are interpolated with transport smoothing splines, and (2) to constrain the Helmholtz smoothing splines used for amplitude interpolation in panel (e).

that differs by more than 4% from the raw stacked result are removed. For our data set, this condition typically eliminates between 2 and 15 events, depending on the period considered. When a small number of maps are averaged, the average phase velocity model may show significant artifacts caused by scattering effects (Chevrot & Lehujeur, 2022; Lehujeur & Chevrot, 2020). These scattering effects may not be properly accounted for by the amplitude corrections, in particular in regions where station coverage is sparse. These artifacts gradually disappear by increasing the number of sources, and coherent patterns in the lateral variations of phase slowness emerge and become increasingly stable. Chevrot and Lehujeur (2022) and P. Zhou et al. (2022) used between 20 and 30 events to obtain stable results on permanent seismic networks in California and China. In our case of temporally evolving networks, the average maps become stable after stacking the results from ∼80 sources (Movie S1).

We design two period and event dependent weighting procedures. The first weighting function **W** is based on the quality of the slowness measurements at individual grid cells. We assign a zero weight to grid cells whose value differs by more than three standard deviation from the average value over all event maps. Since the standard deviation is still influenced by outlying values, we perform a bootstrap analysis on the final data sets to explore the variability of slowness measurements in each grid cell and obtain a more robust averaging scheme. To construct a confidence interval for slowness at each grid cell, we take 200 random realizations and find their 10th and 90th percentiles. We then assign a zero weight to grid cells on individual event maps if the local slowness is below the fifth percentile of the tenth percentile distribution from all bootstrap samples. Conversely, a grid cell value is discarded if the local slowness is above the 95th percentile of the 90th percentile distribution.

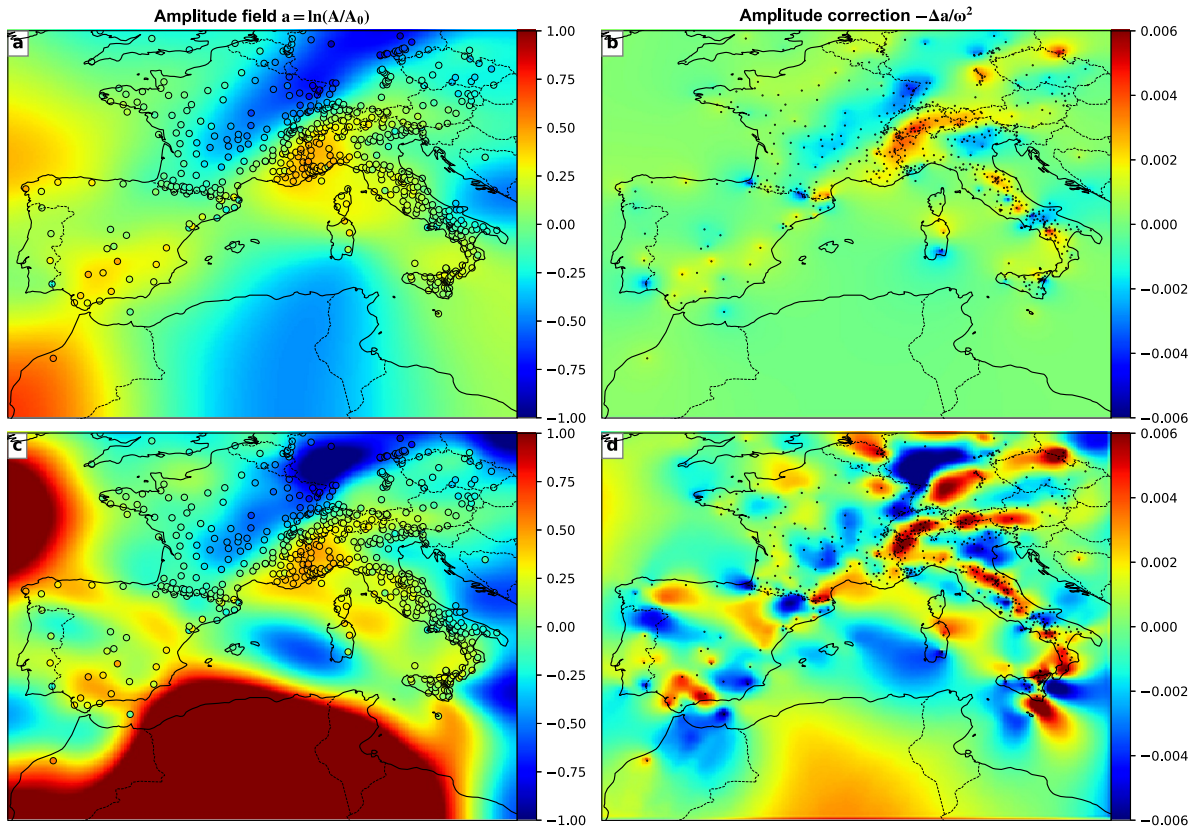

Figure 6. Measured (colored circles) and interpolated amplitude of Rayleigh waves at 50 s period for the Mw = 6.9 2020-9-18 Mid-Atlantic Ridge earthquake. Interpolated amplitude with (a) classical smoothing splines (iteration 1) and with (c) Helmholtz smoothing splines (iteration 2). Amplitude corrections derived from classical smoothing splines (b) and Helmholtz smoothing splines (d).

The second weighting function **w** deals with the spatial variability of seismic measurements that constrain the individual apparent slowness maps. For seismic networks whose geometry is subject to change over time, which is inevitable when utilizing temporary deployments, the stacking results may be influenced by the temporal variation in the spatial distribution of stations. While anomalies may be well resolved when considering dense measurements from a temporary network, our ability to image the same anomalies with a sparser station distribution at another time is subsequently reduced. To compensate for this effect, we introduce additional weights that account for the distribution of receivers at the time of an event, which are applied to each individual map before azimuthal averaging. The purpose of this weighting scheme is to down-weight sparsely sampled regions and give more weight to geographic zones in individual event maps that were covered by dense regional arrays. Our weighting scheme is derived from the geographical weighting function of Ruan et al. (2019), which aims to compensate for the uneven source and station distributions in global adjoint tomography. Given a set of N receivers, we compute the Euclidian distance d_{ij} for each receiver-cell pair. The weight w_i associated to each grid cell i for an event slowness map is computed as follows:

$$w_i = \sum_{j=1}^{N} \exp\left[-\left(\frac{d_{ij}}{d_0}\right)^2\right] \tag{25}$$

where d_0 is a reference distance set to twice the seismic wavelength. Examples of weighting functions **w** are illustrated in Figure S1 in Supporting Information S1 for two station distributions and two periods.

Last but not least, isolated stations often produce focal spots with large amplitude corrections that may persist after averaging. Such spots can be caused for example, by site effects, inaccurate station response, or isolated stations around which the interpolation is poorly constrained. To minimize such artifacts, we apply a median filter

to each individual slowness map, over a quarter wavelength radius prior to azimuthal averaging. Given a set of M events, the final slowness map is then computed as:

$$\mathbf{s} = \left(\sum_{k=1}^{M} \mathbf{w_k} \mathbf{W_k} \hat{\mathbf{s}}_\mathbf{k} \right) \left(\sum_{k=1}^{M} \mathbf{w_k} \mathbf{W_k} \right)^{-1}, \tag{26}$$

where $\mathbf{W_k}$ is the weighting function based on the quality check of grid cell measurements and associated to the event k, $\mathbf{w_k}$ is the weighting function associated to the station distribution at time k, and $\hat{\mathbf{s}}_\mathbf{K}$ is the median-smoothed slowness map for event k.

5.6. Resolution Tests

The depth sensitivity of Rayleigh waves depends on the period. At 100 s period, Rayleigh waves are sensitive to lithospheric and asthenospheric structures, while at short periods ($T \leq 30$ s) they are mostly sensitive to crustal structures (Lebedev et al., 2013). The lateral resolution of surface waves is also period dependent. The smallest anomaly we can resolve is determined by the wavelength and the station distribution. For Helmholtz tomography, which is based on travel time gradiometry, the velocity anomalies might not be well resolved everywhere if the seismic network is heterogeneous. To estimate the lateral resolution of the phase velocity maps, we perform checkerboard resolution tests.

The checkerboard pattern consists of blocks of alternating positive and negative velocity anomalies, each set at $\pm 10\%$ with respect to an average phase velocity of 4 km/s. We consider block sizes of 100, 200, and 300 km (Figures 7a, 7e, and 7i). We design three types of resolution tests to be performed for each checkerboard pattern. Each test considers different seismic network configurations. For each of the 173 events selected in our study, we compute synthetic travel times at station locations with the Fast Marching Method, using the velocity model defined by the checkerboard pattern. Velocity maps are then computed for each event from the gradient of the phase previously interpolated using Equation 12. We then stack all the individual maps to obtain the final average velocity map. The retrieved velocities are shown for 40 s Rayleigh waves in Figure 7 for each checkerboard pattern.

In the first resolution test we assume that each event was recorded by all stations assigned to the region during the whole investigated time period (black dots in Figure 7i). Even if the station distribution is still heterogeneous, the contribution of each event is even in terms of spatial resolution. The results of this synthetic test are shown in Figures 7b, 7f, and 7j for velocity anomalies of size 100, 200, and 300 km, respectively. The velocity anomalies are generally well resolved on the continent where the station distribution is dense enough. Note that in northern France, the retrieved velocity anomalies of size 100 km are slightly smeared to the west because of a poor station coverage in this region.

The second and third resolution tests are close to the real case in terms of station distributions. This time we consider only the travel times measured at the stations operating at the time of the event. Each source then has a different contribution in terms of its ability to resolve the anomalies, due to the irregular spatial as well as temporal station distribution. The second resolution test uses a classical stacking scheme, while the third resolution test uses the weighting scheme based on Equation 25. Here we use $d_0 = 300$ km, which is twice the wavelength of a 40 s Rayleigh wave. This weighting scheme will give more weights to grid cells where dense measurements are available (Figure S1 in Supporting Information S1).

The comparison of the second resolution test (Figures 7c, 7g, and 7k) with the first one (Figures 7b, 7f, and 7j) shows that the station coverage strongly impacts our ability to retrieve the strength of the anomalies. For velocity anomalies of size $\leq$200 km, the spatial resolution is degraded in regions that are not always covered by permanent networks (e.g., in Morocco, Iberia, or France, especially in its northern part). In comparison, in the third resolution test (Figures 7d, 7h, and 7l) the anomalies are better recovered for all block sizes. With the weighting scheme, the contribution of each source is more evenly balanced in the final stack.

The synthetic resolution tests suggest that velocity anomalies of size >100 km are expected to be well resolved using the weighting scheme. As an indication, we estimate a confidence region indicated by the black contour in Figure 7l. Within this region, the resolution tests show that we are able to retrieve accurate velocity anomalies of size $\sim$300 km whose amplitudes differ by less than 5% from the synthetic case.

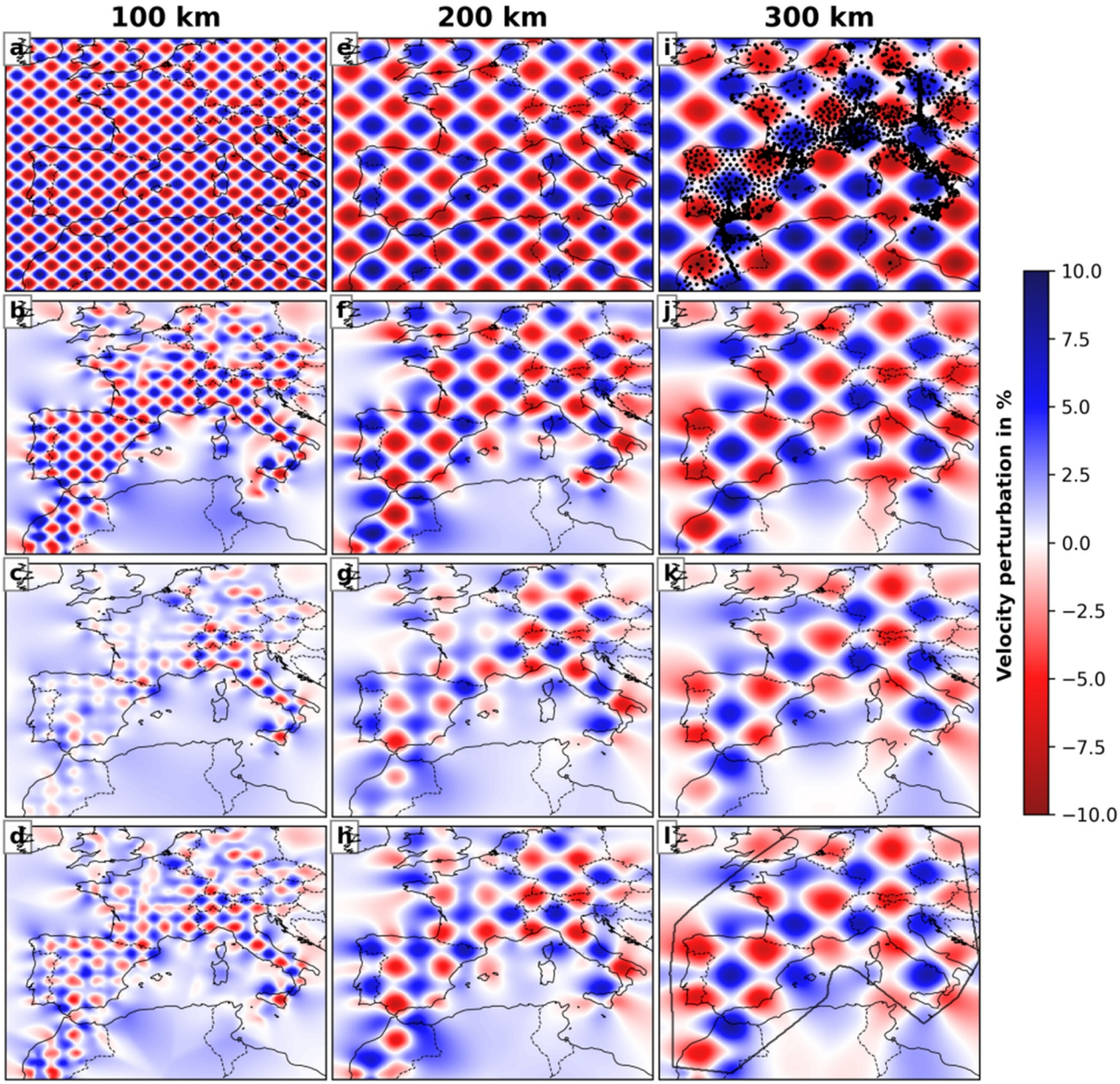

Figure 7. Synthetic checkerboard resolution tests for the Rayleigh-wave phase-velocity. The checkerboard consists of blocks of alternative positive and negative velocity anomalies (±10% of deviation from 4 km/s) discretized on block sizes of (a) 100, (e) 200 and (i) 300 km. The results of the resolution tests at period 40 s are shown in rows 2, 3 and 4 for different station distributions and weighting stacking schemes. Each map shows the velocity pattern recovered after stacking individual slowness maps for the 173 events included in our study. For each event, the derived slowness map is the eikonal solution for a given distribution of sampling stations. The second row (b, f, j) shows the velocity maps derived when assigning all stations that were available in the region during the time period 2008–2022 (black dots in i) to each event. The third row (c, g, k) shows the derived velocities with real case conditions, that is, only the stations available at the time of the event contributed to the event. The fourth row (d, h, l) shows the derived velocities with real-case conditions as in panels (c, g, k), but computed with appropriate weighting stacking scheme. Here, we assigned to each grid cell i and each event, the weighting function w_i based on station locations and density, before azimuthal averaging. The black contour in panel (l) indicates the confidence region where the retrieved velocities for block size of 300 km differ from the original velocities by less than 5%.

5.7. Phase Velocity Maps

We performed six iterations of the previously described Helmholtz tomography algorithm to determine the phase velocity maps. However, we observed that the changes in the average slowness maps become negligible after the third iteration. Therefore, we will only consider the results of the first three iterations. Figure 8 shows the phase velocity maps for period 40 s obtained after the first, second and third iterations. The first iteration uses classical

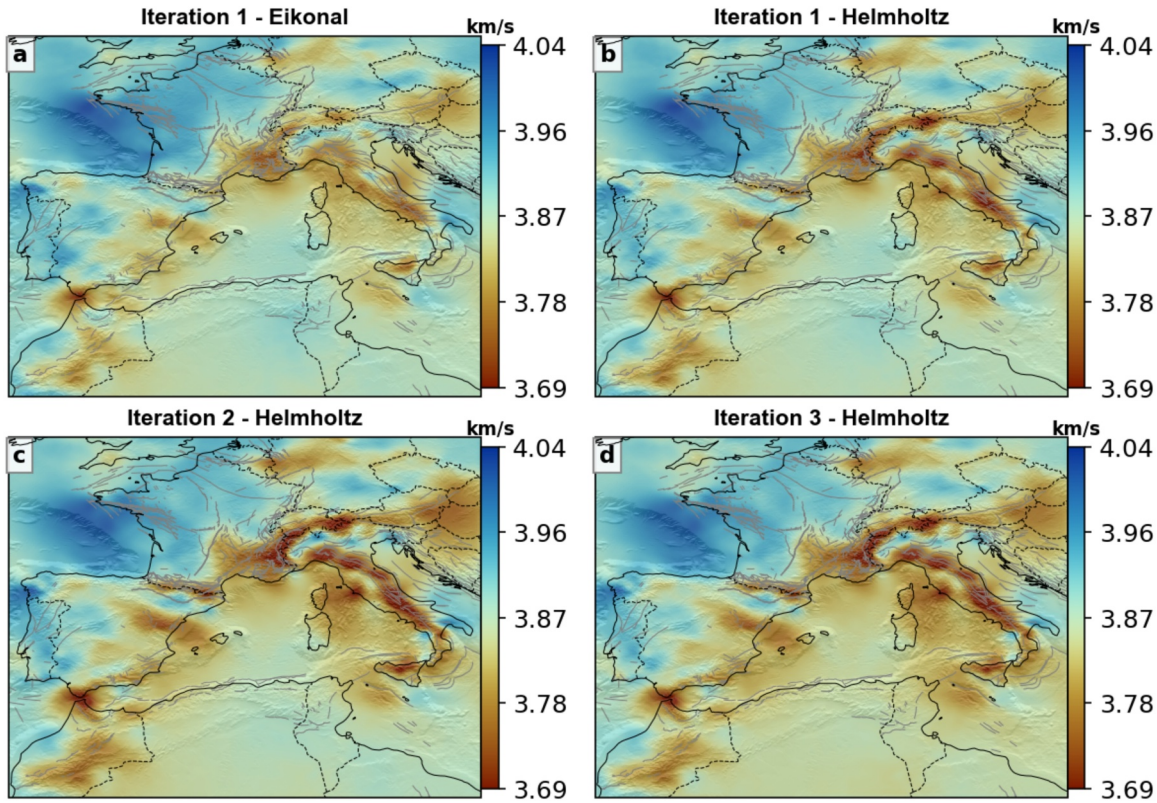

Figure 8. Average phase velocity maps for period 40 s obtained after stacking (a) eikonal velocities at iteration 1, and Helmholtz velocities at iteration (b) 1, (c) 2, and (d) 3. For comparison, we use the same velocity colormap for all figures.

smoothing splines, while the second and third iterations use Helmholtz and transport smoothing splines. Here we show the eikonal solution for the phase velocities obtained in the first iteration (Figure 8a), that is, the phase velocities determined from the phase gradient only. For the first and later iterations, we show the Helmholtz solution (Figures 8b–8d), which includes the amplitude correction term. The amplitude information improves the spatial resolution of the structural anomalies. The velocity anomalies appear stronger in amplitude and their contours appear sharper. Taking into account the amplitude of the seismic wavefield reveals hidden structures that could hardly be resolved using only the phase information, and better recovers the amplitudes of the phase velocity anomalies.

Figures 9 and 10 show the final phase velocity maps obtained at the end of the third iteration for periods 25, 30, 40, 60, 80, 90, and 120 s. At 25 s, the phase velocities are well correlated with the topography, as the thick crustal roots produce pronounced low velocity anomalies beneath mountain ranges (e.g., beneath the Appenines, Alps, Pyrénées, Betics). At 30 s, a slow anomaly starts to emerge beneath the Massif Central. At periods longer than 60 s, fast anomalies are resolved at the locations of downgoing slabs such as the Calabrian slab, Gibraltar, the Alpine arc, and the Dinarides. A more detailed interpretation of the velocity maps is given in Section 6.3.

6. Discussion

6.1. Eikonal Versus Helmholtz Tomography

Conventional surface wave tomography uses only the phase information to invert for phase velocities. However, neglecting the amplitude has important consequences for the resolution limit of tomography, which have been investigated by Lehujeur and Chevrot (2020). Their numerical experiments and theoretical developments demonstrated that even for a perfect knowledge of the phase and ideal azimuthal coverage, the recovered model is the true model, but convolved by Bessel functions, which have a main radial beam of about half the wavelength. This tomographic filter represents the fundamental resolution limit of travel time tomography. They have also shown that if both the phase and amplitude are perfectly known, a perfect reconstruction can then be achieved,

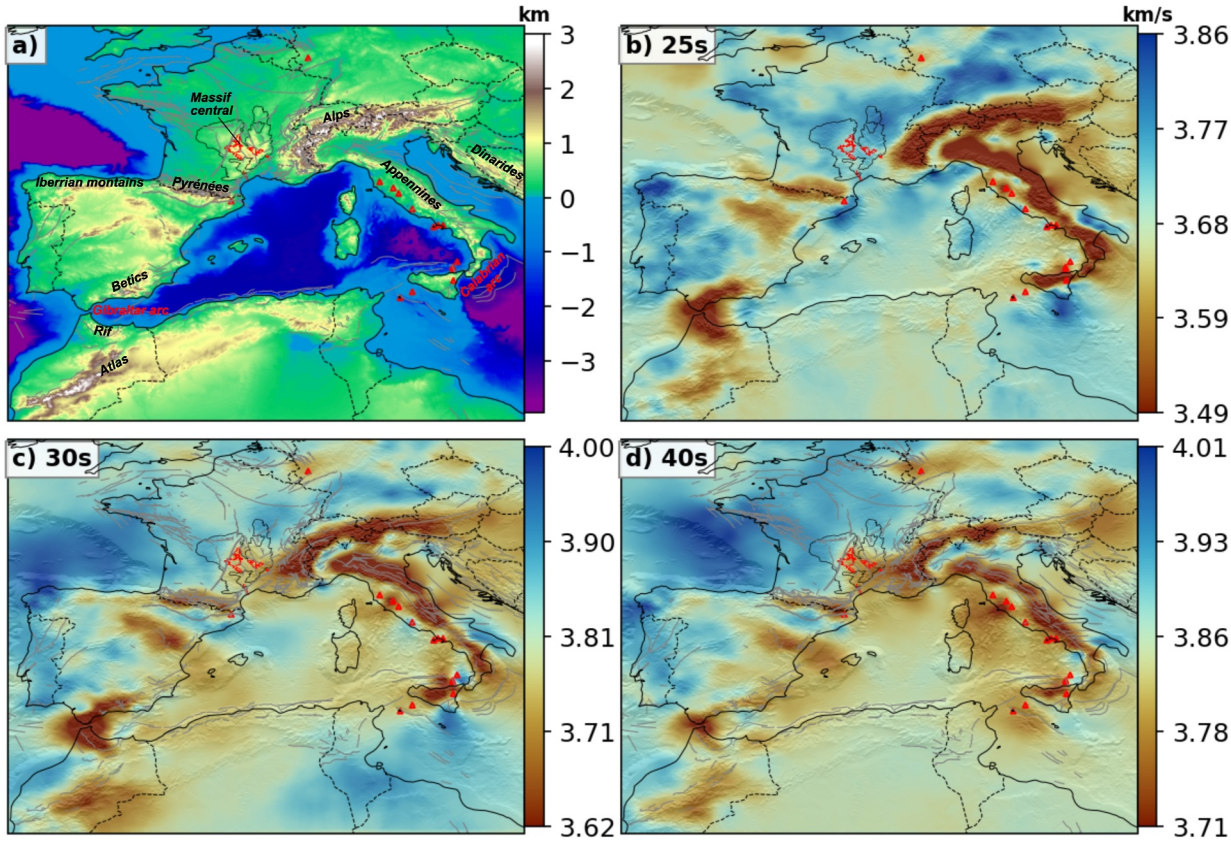

Figure 9. Topography (a) and average phase velocity maps for periods (b) 25, (c) 30, and (d) 40 s. Final results are those obtained at iteration 3. In panel (a), mountain ranges are labeled in black. Subduction zones of Gibraltar and Calabrian arcs are labeled in red. Faults and major tectonic boundaries are indicated by gray lines. Volcanoes are indicated by red triangles. Red contours in Massif Central indicate the Cenozoic volcanic rocks, whereas the black lines outline the Hercynian orogenic belt.

even with a single source. Using a different approach, Yang and Forsyth (2006) also demonstrated the importance of using the amplitude information in surface wave tomography. They found that 2D phase and amplitude sensitivity kernels computed by normal-mode summation (Y. Zhou et al., 2004) can accurately predict the perturbation of the wavefield even when the size of the anomaly is smaller than the wavelength. They also used these 2D kernels to invert synthetic data obtained by numerical modeling of the propagation of Rayleigh wave through a checkerboard model. They found that the checkerboard pattern can be recovered by inverting either amplitude or phase data, but the strength of the anomalies were significantly underestimated in both inversions. In contrast, jointly inverting both the amplitude and phase data, the anomalies were almost perfectly recovered. Even though these results clearly suggest that surface wave amplitude contains important information about Earth structure and should be inverted together with phase data in surface wave tomography, tomographic studies were so far exclusively relying on phase inversions, with only a few exceptions (e.g., Lin et al., 2009).

Our Helmholtz tomography results confirm that indeed the amplitude information improves the spatial resolution of surface wave tomography, with phase velocity anomalies that are stronger in amplitude and more sharply defined. Better constraining the absolute amplitude of phase velocity anomalies is obviously important to quantify the chemical and thermal anomalies of the lithosphere.

6.2. Helmholtz Tomography Versus Wavefield Reconstruction Imaging

The wave equation smoothing splines introduced here share some similarities with the wavefield reconstruction inversion (WRI) method, introduced by van Leeuwen and Herrmann (2013). WRI is a novel formulation of full waveform inversion, which involves minimizing the functional

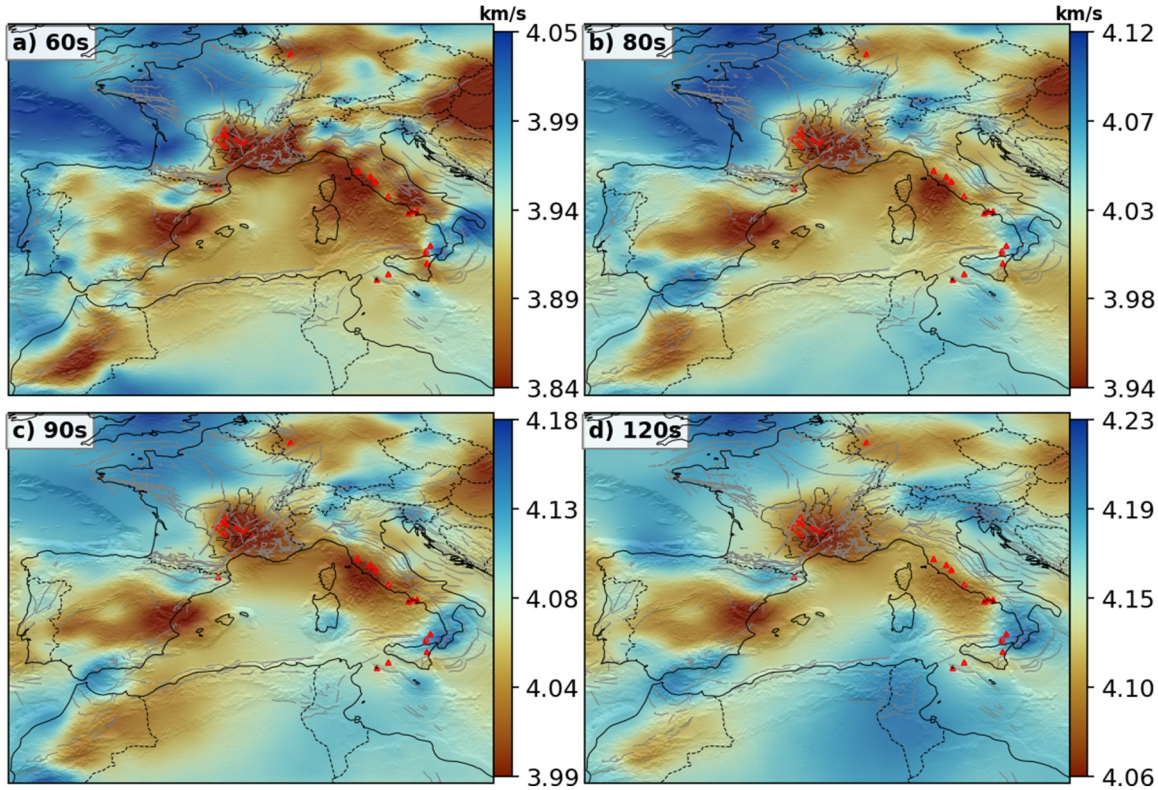

Figure 10. Average phase velocity maps for periods (a) 60, (b) 80, (c) 90, and (d) 120 s obtained at iteration 3.

$$\chi(\mathbf{m},\mathbf{u}) = ||\mathbf{P}\cdot\mathbf{u} - \tilde{\mathbf{u}}||^2 + \mu||\mathbf{A}(\mathbf{m})\mathbf{u}||^2, \tag{27}$$

where $\mathbf{A}(\mathbf{m}) = \omega^2\text{diag}(\mathbf{m}) + \mathbf{L}$ is the discretized scalar wave equation operator, and $\tilde{\mathbf{u}}$ the wavefields observed at the stations. The first term in Equation 27 quantifies the data misfit whereas the second term enforces the wave equation. Compared with the classical formulation of full waveform inversion, the wave equation is not matched exactly. van Leeuwen and Herrmann (2013) formulated an inverse problem that simultaneously reconstructs the seismic velocities and the wavefield. A variant of WRI has been proposed by de Ridder and Maddison (2018) to invert ambient seismic noise wavefields recorded by dense arrays.

The main difference between WRI and our Helmholtz tomography approach is that the latter exploits the observed wavefield whereas the former exploits its phase and amplitude. While WRI can be considered as a kind of smoothing spline, interpolating the wavefield is more difficult because it is highly oscillatory, unlike the amplitude and phase fields, which are smoother and more regular. In addition, WRI involves solving a non linear inverse problem, with the pitfall of converging toward a local instead of a global minimum of the cost function. Instead, in Helmholtz tomography, the non-linearity is hidden in phase unwrapping and in the iterative Helmholtz tomography algorithm, which progressively reconstructs accurate amplitude corrections. The average phase velocity can be robustly determined by beamforming, for example, and it is therefore easy to correct phase measurements for cycle skipping effects. Consequently, whereas these two approaches have many similarities, Helmholtz tomography avoids the need for a sufficiently accurate starting model for the FWI to converge to the true velocity structure (Virieux & Operto, 2009). Since the final amplitude and phase fields reconstructed by generalized smoothing splines satisfy the Helmholtz and transport equations, respectively, it should be equivalent to FWI.

Note that although generalized smoothing splines combined with cross-validation are here considered for the first time for the Helmholtz tomography problem, they have been used for the assimilation of meteorological data for several decades (e.g., Wahba & Wendelberger, 1980).

6.3. Preliminary Interpretations of the Final Helmholtz Phase Velocity Maps

We now present some preliminary interpretations of our phase velocity models, as well as comparisons with the results of previous tomographic studies. The 3D velocity model derived from these phase velocity maps will be presented in a future contribution.

At 25 s, Rayleigh waves have a maximum sensitivity to shear velocities at a depth of about 30 km. Consequently, at this period, the phase velocity of Rayleigh waves will be mainly controlled by variations of crustal thickness. Indeed, the 25 s phase velocity map (Figure 9b) shows pronounced low velocity anomalies that are spatially correlated with the high topography of Cenozoic mountain ranges such as the Alps, the Appennines, the Calabrian Arc, the Pyrénées, the Iberian Mountains, the Betics, the Rif, and the Atlas (Figure 9a).

The AlpArray experiment has provided new opportunities to study the Alpine domain, and a number of surface wave models obtained from earthquake recordings and/or ambient noise have been published recently (Alder et al., 2021; El-Sharkawy et al., 2020; Kästle et al., 2018; Lu et al., 2018; Nouibat et al., 2023). These different models are generally in good agreement with our results. All these models show low velocity anomalies under the Alps and Apennines, spatially correlated with the surface topography (Figure 9a). These results confirm the presence of thick crustal roots that support the topography. At longer periods, positive velocity anomalies are observed beneath the Western Alps, the Northern Alps and the Dinarides, which are probably associated with downgoing slabs. Uppermost mantle structures in the Alpine region imaged with regional travel time tomography remain controversial, particularly with regard to the geometry and continuity of these downgoing slabs (e.g., Kästle et al., 2020). A significant part of the discrepancies observed between the different models (e.g., Giacomuzzi et al., 2011; Koulakov et al., 2009; Lippitsch et al., 2003; Zhao, Paul, Malusà, et al., 2016) probably stems from the effects of crustal structures, which can produce severe smearing artifacts down to 150–200 km depth (Chevrot et al., 2014). The phase velocity maps obtained in this study could therefore provide key constraints for the construction of a 3D reference crustal model that could be used in regional travel time tomography to reduce these smearing artifacts. They could also be used in joint inversions with body travel times or even complete waveforms (Beller et al., 2018; Kan et al., 2023; Wang et al., 2016).

Compared to the Alps, the Pyrénées are characterized by a much less pronounced reduction in phase velocities at 25 s period, which could be attributed to a smaller amount of crustal shortening and thickening. We also note the absence of a low velocity anomaly beneath the eastern Pyrénées, suggesting a near-normal crustal thickness in that part of the mountain range, in good agreement with the results of receiver function studies (Chevrot, Sylvander, et al., 2018; Diaz et al., 2018). The signature of the underthrusted Iberian lithosphere underlying the western and central Pyrénées is observed for periods between 40 and 70 s, in good agreement with the migrated receiver function sections along the Pyrenan transects (Chevrot, Sylvander, et al., 2018). Beneath the western Pyrénées, we also observe lower than average phase velocities over the same range of periods, suggesting a hot and thinned lithosphere in this region.

Beneath the Betics and Rif, a strong horseshoe-shape anomaly is clearly imaged at periods between 25 and 40 s, in good agreement with the recent results of Feng and Diaz (2023). This anomaly is particularly strong beneath Gibraltar, where the crust is dragged down by the descending lithosphere (Villaseñor et al., 2015). In this region, we therefore evidence a very thick crust underlying a moderate topography. At longer periods ($>$ 70 s), a high velocity anomaly is imaged beneath the Betics, a feature that has been consistently evidenced by previous regional travel-time tomographic studies (e.g., Bezada et al., 2013; Bonnin et al., 2014; Villaseñor et al., 2015). We note that this anomaly is barely visible in the results obtained after the first iteration (eikonal tomography). Taking into account the amplitude information is therefore essential to obtain a well-resolved image of small-scale structures (compared to the dominant wavelength).

Beneath the Massif Central the phase velocities at 25 s period are close to normal, and therefore the topography does not seem to be supported by thick crustal roots. On the other hand, we image below-average phase velocities for periods greater than 40 s, indicating a thin, hot lithosphere, overlying a hot, buoyant asthenosphere. From the joint interpretation of geothermobarometric data from xenoliths, gravimetric data, and the tomographic model obtained by Granet et al. (1995), Sobolev et al. (1997) point to the presence of a hot asthenosphere at 70–120 km depth beneath the Massif Central, with a temperature excess of the order of 150–200° K. They attributed this anomaly to a mantle plume extending to a depth of at least 300 km, with a diameter of 200 km. Our tomographic images do not support this latter interpretation, as the anomaly at depth beneath the Massif Central is triangular

rather than circular, and is not centered on the volcanic fields, but instead extends southeast toward the Mediterranean Sea.

Beneath the High Atlas, we document a moderate low velocity anomaly at 25 s period, compatible with a moderate crustal thickening. As in the Massif Central, the topography, which can reach up to 4,000 m above sea level, is only partially supported by the crust, and a buoyancy contribution from subcrustal levels must be invoked. The regional P wave tomography by Seber et al. (1996) has revealed a low velocity anomaly in the upper mantle under the High Atlas, suggesting a thin, hot lithosphere. Gravity, heat flow, and topography data can be explained by a an elongated band of thinned lithosphere extending from the Anti-Atlas to the Middle Atlas and crossing the High Atlas (Teixell et al., 2005). Its NE–SW orientation, oblique to the mean trend of the Atlas system, is very similar to the geometry of the low velocity anomaly in our phase velocity maps for periods between 40 and 90 s. However, the origin of this deep mantle anomaly remains debated (Missenard et al., 2006; Teixell et al., 2005).

7. Conclusions

Eikonal tomography is an efficient approach for surface wave tomography which can provide robust, finely resolved, phase velocity maps, provided a sufficient number of wavefronts coming from varying azimuths can be exploited. However, even in the case of homogenous sampling of the phase and perfect azimuthal coverage, the recovered model is a filtered and smoothed version of the true model. The absolute values of velocity perturbations are thus generally underestimated, which is a well-known limitation of classical tomographic approaches which can be overcome by using the amplitude information. We have shown that this can be achieved by introducing a new family of generalized smoothing splines for the amplitude, which satisfy the Helmholtz equation. The method starts with a classical eikonal tomography approach, based on phase measurements only. The resulting phase velocity models are then used to interpolate the amplitude measurements, using generalized Helmholtz smoothing splines. The comparison between the results of eikonal and Helmholtz tomography in Western Europe demonstrates the importance of the amplitude information in obtaining tomographic models with a sharper spatial resolution. This new implementation of Helmholtz tomography is equivalent to a complete waveform inversion, but for a computational cost that is reduced by several orders of magnitude.

Appendix A: A Kriging-Based Method to Remove Outliers in Phase and Amplitude Measurements

Because Helmholtz tomography involves the computation of gradients, it is particularly sensitive to the presence of outliers, especially for irregular and sparse arrays. Anomalous values in the phase and amplitude data sets can arise, for example, from problems with sensors or metadata (e.g., clock problem, incorrect instrument response). Whereas the identification of strongly anomalous phase or amplitude measurements is straightforward, it is often necessary to compare these measurements with those at neighboring stations to identify outliers. Since the spatial variations of phase and amplitude produced by wave propagation effects in a heterogeneous medium are smooth and coherent, this suggests to use an approach that relies on a statistical description of the spatial variations of the phase and amplitude to detect potential outlier measurements. The basic idea is to compare the value predicted by this statistical model at a given station using the measurements from the other stations, and infer whether the deviation from the observation exceeds some expected statistical error.

In practice, we use a hierarchical cross-validation approach to assess the impact of each measurement (phase or amplitude) on a set of regression models. The regression models are built using the kriging interpolation method on multiple training data sets that undersample the actual catalog of measurements. Each event and each period is processed separately. Kriging allows us to compute predictions at a given point from a weighted average of the measurements in its vicinity. It is based on the assumption that the spatial variations of the field can be described by a statistical covariance function, which can be approximated by a variogram (Cressie, 1993). This method provides an unbiased estimator with minimum variance. We exploit the property of kriging, which gives the best linear unbiased prediction at unsampled locations.

We work with data sets of phase and amplitude residuals, which are considered as samples of a normal distribution (Figures A1a and A1b). Phase residuals δT are defined as the difference between the travel time measurements T minus the predicted travel times T_0 of the average plane wave determined by beamforming

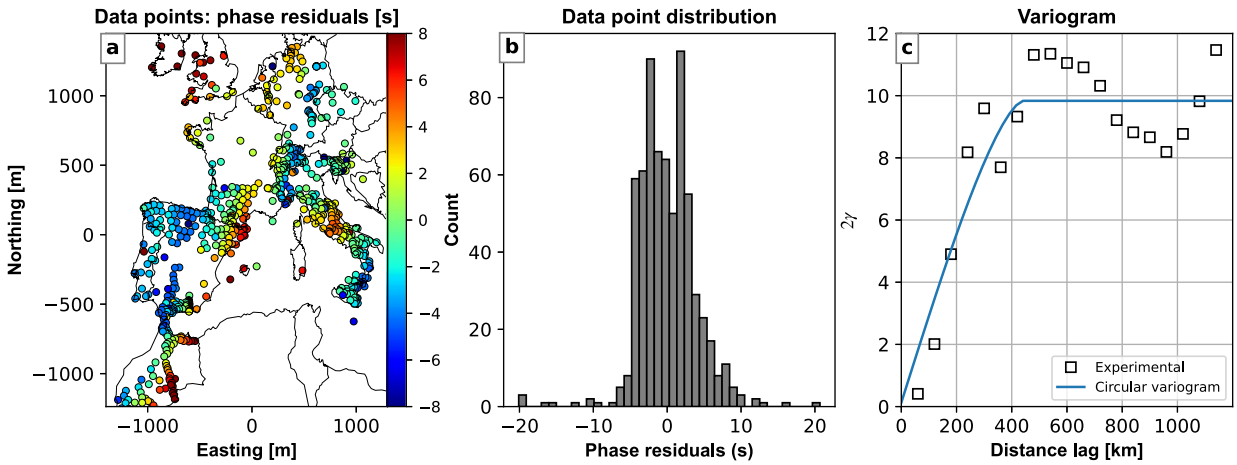

Figure A1. Distribution and spatial structure of traveltime residuals for the 2011 Japan event at period 40 s. (a) Observed travel time residuals δT. They correspond to the difference between the Rayleigh wave travel times measured at 40 s and the travel time of the average plane wave determined by beamforming. (b) Distribution of δT. (c) Experimental variogram of δT observations (squares) and best-fitting variogram model.

(Figure 3c). Amplitude residuals are defined as the logarithm of the relative amplitudes with respect to the average over the stations, $\delta a = \ln(A/A_0)$. For kriging, we use Cartesian coordinates in a transverse Mercator projection, a cylindrical projection with a line of tangency of the cylinder corresponding to the central meridian of the regional domain.

The identification of outliers is based on the difference between the observed data and the data predicted by the regression model constructed by kriging. To predict spatially dependent data, it is necessary to first model the spatial correlation of the data set with an empirical variogram model. Robust estimates of variograms for δT or δa can be derived from Cressie and Hawkins (1980):

$$2\gamma(h) = \frac{\left(|N(h)|^{-1}\sum_{N(h)}\sqrt{|Z_i - Z_j|}\right)^4}{0.487 + \frac{0.494}{|N(h)|} + \frac{0.045}{|N(h)|^2}}, \tag{A1}$$

where $\gamma(h)$ is the semivariogram, $N(h)$ the number of data pairs in the spatial offset bin h, and $Z_i - Z_j$ the difference between the values of δT or δa at the locations i and j. The covariance function of the regression model is then prescribed by a continuous function $\hat{\gamma}(h)$ that best fits the empirical variogram γ. Figure A1c shows the variogram associated with the travel time residuals from Figure A1a. At distances $h < 300$ km, corresponding to two wavelengths at 40 s period, the values of δT are strongly correlated. At distances $h > 400$ km, the variogram reaches a plateau and this correlation disappears. The critical distance at which the data are no longer correlated is called the variogram range, which increases with the period.

When constructed on a sample of data S, the kriging predictor $\hat{Z}$ provides a perfect match to the observed data: $\hat{Z}(s_0|S) = Z(s_0)$, if $s_0 \in S$. To identify observations that are outliers with respect to the null model of no outliers, we use a cross-validation approach. The idea is to build several kriging predictors, each based on a subsample of data S_{-i} in which we have omitted an observation at location i. We then evaluate the normalized prediction error at the observation point i that was not included in the training data set:

$$\hat{e}_{i(S_{-i})} = \frac{Z_i - \hat{Z}_{i(S_{-i})}}{\sigma_{i(S_{-i})}}, \tag{A2}$$

where $\sigma^2_{i(S_{-i})}$ is the kriging variance (Cerioli & Riani, 1999). If an observation point has a prediction error $e_i > 2.5$, it is considered an outlier.

A standard cross-validation approach to testing for the presence of outliers in regression models would not deal with swamping (outlier contaminating the estimated value of its neighbors) or masking (a cluster of outliers merging into another cluster) effects. Our approach to identifying outliers in spatial prediction models is based on a forward search (Atkinson, 1994; Cerioli & Riani, 1999). The forward search algorithm consists of ordering the observations from those that are most consistent with a specified regression model to those that are least consistent.

We start by identifying a subset S_p of p data points that are free of outliers. To construct this initial data set, we use a clustering approach. We first perform a standard cross-validation of the kriging predictors with all data points. If we have N observations, we build N regression models for which we have removed one observation (i.e., S_{-i} training sample). For each data point i, we compute the prediction error $\hat{e}_{i(S_{-i})}$ from the regression model based on S_{-i}. We keep only the observations for which $\hat{e}_{i(S_{-i})} < 3$. We then use the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm (Ester et al., 1996) on all the selected data points. The DBSCAN attributes are the spatial coordinates of each data point and the variance of the δT (or δa) values for all data points within a 150 km radius of the station. Clustering is used to bring together similar observations, usually in the same geographical regions. It excludes observations with unusual extreme values (and therefore high variance), and observations from isolated stations for which it is difficult to judge whether they are outliers. The p data points in the initial subset S_p are selected from the data points that were successfully clustered into different families. The number of points to include in the initial subset depends on the total number of observations available and must be at least $p \geq 2$.

Once we have the initial outlier-free subset of size p, we run a regression model to estimate the kriging values at the remaining $m = N - p$ points. We then iteratively build $N - p$ regression models, each time incrementing the current subset of samples S_m by the additional data point that has the smallest prediction errors relative to the current data set. Potential outliers are then pushed to the end of the search and will not contaminate the earlier models and the predictions at neighboring sample points. We can monitor prediction errors throughout the forward search. Data points with standardized prediction errors greater than 2.5 during 70% of the iterations are classified as outliers. They are removed from further analysis.

Figure A2 illustrates the preliminary outlier detection procedure with DBSCAN for the travel time residuals δT measured at 40 s from a $M_w = 6.2$ earthquake in Japan. Figure A2a shows the results of DBSCAN, which identified 7 different clusters. Black dots indicate stations which could not be included in any of these clusters. In total we count 85 unclustered data points. As the clustering attributes are the spatial coordinates and the δT variance, data points identified as outliers at this stage include isolated stations (e.g., stations in Corsica, center of France, and Balearic islands), and stations with extreme δT values or outlying measurements with respect to their neighbors (e.g., stations in Ireland, one dark blue station in Sicily, 3 dark blue stations in Slovenia, Figure A2c). DBSCAN also classifies as outliers stations that are spatially close to outlying measurements that contaminate the δT variance at that location (e.g., blue station in Portugal close to a dark red outlying station). In the presence of a non-stationary group of extreme δT values, that is, a group of potential outliers, the DBSCAN algorithm tends to cluster these data points into a family and is unable to discriminate outliers (e.g., dark red stations in northern England, dark red stations in the southern part of the Morocco profile).

For the forward search, we choose $p = 20$ data points from the outlier-free subset of measurements identified by DBSCAN. Since the ordering of the data in the forward search is somewhat dependent on the distribution of data values, we choose these 20 samples so that they follow approximately the same normal distribution as the entire data set. Figure A2d shows the evolution of the prediction error $\hat{e}_i$ of each data point i during the forward search. Having an initial subset of 20 data points and a total number of 661 stations, we compute 641 regression models with training subset S_m of evolving size m. Each spatial location is tracked until it joins the subset S_m. This plot typically has a scissor shape, with stations with small (large) $\hat{e}_i$ joining the subset in the first (last) stages of the forward search. Curves corresponding to unusual observations are highlighted. The outliers are identified by errors greater than 2.5 times the standard deviation. Figure A2e shows the evolution of the number of data points for which $\hat{e}_i > 2.5$ through the forward search iterations. Around iteration 300, half of the available observations are used to build the regression model, and the total number of data points with prediction errors exceeding 2.5 is maximum around 70. Around iteration 400, the number of outliers begins to gradually decrease, as a certain amount of extreme values of δT (e.g., dark red curves and blue curves in Figure A2d) have joined the kriging subset. The decrease in the number of outliers accelerates toward the end of the forward search, as atypical

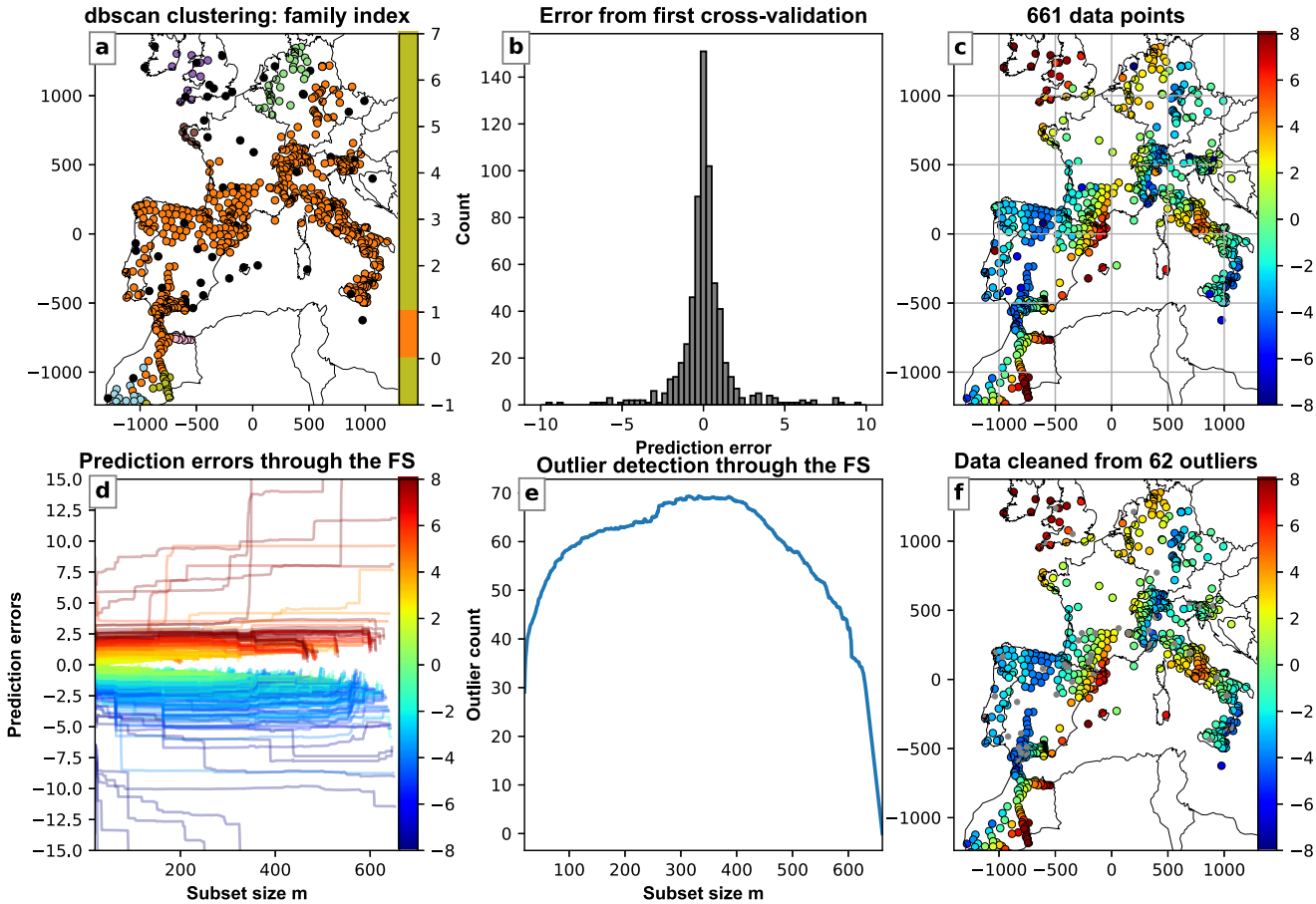

Figure A2. Outlier detection procedure applied to the phase residuals δT for the 2011 Japan event at period 40 s. (a) Results of DBSCAN clustering. Each color corresponds to one cluster based on data point density and the variance of their values in a small surrounding region. Data points indicated by black dots are not clustered in any of the 7 groups and may indicate outlying values of δT. (b) Distribution of kriging prediction errors $\hat{e}_{i(S_{-i})}$ obtained for the regression models. Most data points have a prediction error $\hat{e}_{i(S_{-i})} < 2.5$. Data points with $\hat{e}_{i(S_{-i})} > 3$ are labeled as potential outliers and are discarded for the construction of the data subset S_p to start the forward search. (c) Map of the 661 original data points and associated δT values. (d) Prediction errors during the forward search. Each line corresponds to one data point, and its color indicates the δT value as in panel (c). Each location is monitored until it joins the subset S_m. (e) Number of data points with prediction errors $\hat{e}_i > 2.5$ along the iterations of the forward search. (f) Observed δT field after removing 62 outliers, indicated by gray dots.

observations have joined the kriging subset. From the forward search, we classify observations as outliers if they have standardized prediction errors greater than 2.5 in 70% of the iterations.

The most critical step of this procedure is the construction of the initial subset. There is a trade-off between the size of the initial subset S_p, the speed of the forward search, and its efficiency in pushing forward the pocket of true outliers. Since it is difficult to (a) construct an outlier-free subsample of data (e.g., clusters of outliers are difficult to identify) and (b) subsample the entire data distribution from a reduced subset S_p of size 20, the forward search is repeated several times. Here we use 40 realizations of the forward search, with a different initial subset S_{20} each time. An observation is definitively classified as outlier if it is present in 28 out of 40 forward searches. In the end, we remove 62 observations from the initial catalog of 661 phase measurements (Figure A2f).

The accuracy of the kriging results depends on our ability to accurately describe the 2D spatial structure of the phase and amplitude fields with a 1D covariance function. This method is very sensitive to station sampling, and the geometry of structural anomalies, which control the spatial structure of the seismic wavefield. The cleaning algorithm is therefore most effective for dense and regular seismic arrays deployed in regions where structural anomalies produce smooth variations of the phase and amplitude. Due to the large area covered by our study and the sparse station distribution in some regions of our tomographic domain, the detection of isolated outliers remains challenging.

Data Availability Statement

The seismic data used in this study are publicly available from ORFEUS Data Center (https://www.orfeus-eu.org/data/odc/). The smoothing splines python package is available at (Sergeant et al., 2024).

Acknowledgments
We thank the associate editor Fenglin Niu, Petr Kolínský, and an anonymous reviewer for their thoughtful and constructive comments. This work was supported by the ANR project FRLITHO3D (ANR-21-CE49-0013-03). Data from the AlpArray Seismic Network (AASN) were made freely available to the public as part of the AlpArray initiative (www.alparray.ethz.ch). We acknowledge the operation of the temporary AlpArray Seismic Network under FDSN network code Z3, see Hetényi et al. (2018) and AlpArray Seismic Network (2015), used in this study.

References

Alder, C., Debayle, E., Bodin, T., Paul, A., Stehly, L., & Pedersen, H. (2021). Evidence for radial anisotropy in the lower crust of the Apennines from Bayesian ambient noise tomography in Europe. *Geophysical Journal International*, *226*(2), 941–967. https://doi.org/10.1093/gji/ggab066

AlpArray Seismic Network. (2014). Eastern Alpine Seismic Investigation (EASI) - AlpArray Complimentary Experiment [dataset]. AlpArray Working Group. Retrieved from http://networks.seismo.ethz.ch/networks/xt/doi:10.12686/ALPARRAY/XT_2014

AlpArray Seismic Network. (2015). Alparray Seismic Network (AASN) temporary component [dataset]. AlpArray Working Group. Retrieved from http://networks.seismo.ethz.ch/networks/z3/doi:10.12686/ALPARRAY/Z3_2015

Atkinson, A. C. (1994). Fast very robust methods for the detection of multiple outliers. *Journal of the American Statistical Association*, *89*(428), 1329–1339. https://doi.org/10.1080/01621459.1994.10476872

Beller, S., Monteiller, V., Operto, S., Nolet, G., Paul, A., & Zhao, L. (2018). Lithospheric architecture of the South-Western Alps revealed by multiparameter teleseismic full-waveform inversion. *Geophysical Journal International*, *212*(2), 1369–1388. https://doi.org/10.1093/gji/ggx216

Bezada, M., Humphreys, E., Toomey, D., Harnafi, M., Dávila, J., & Gallart, J. (2013). Evidence for slab rollback in westernmost Mediterranean from improved upper mantle imaging. *Earth and Planetary Science Letters*, *368*, 51–60. https://doi.org/10.1016/j.epsl.2013.02.024

Bonnin, M., Nolet, G., Villasenor, A., Gallart, J., & Thomas, C. (2014). Multiple-frequency tomography of the upper mantle beneath the African/Iberian collision zone. *Geophysical Journal International*, *198*(3), 1458–1473. https://doi.org/10.1093/gji/ggu214

Cerioli, A., & Riani, M. (1999). The ordering of spatial data and the detection of multiple outliers. *Journal of Computational & Graphical Statistics*, *8*(2), 239–258. https://doi.org/10.1080/10618600.1999.10474812

Chen, Y., Chen, Y., Fomel, S., Savvaidis, A., Saad, O. M., & Oboué, Y. A. S. I. (2023). Pyekfmm: A python package for 3D fast-marching-based travel-time calculation and its applications in seismology. *Seismological Research Letters*. https://doi.org/10.1785/0220230042

Chevrot, S., Diaz, J., Sylvander, M., & Ruiz, M., & RESIF. (2018a). Seismic network Zu: Orogen-X temporary experiment (RESIF-SISMOB) [dataset]. *RESIF - Réseau Sismologique et géodésique Français*. https://doi.org/10.15778/RESIF.ZU2015

Chevrot, S., & Lehujeur, M. (2022). Eikonal surface wave tomography with smoothing splines—Application to Southern California. *Geophysical Journal International*, *229*(3), 1927–1941. https://doi.org/10.1093/gji/ggac034

Chevrot, S., & Sylvander, M., & RESIF. (2017). Seismic network X7:PYROPE PYRenean Observational Portable Experiment (RESIF-SISMOB) [dataset]. *RESIF - Réseau Sismologique et Géodésique Français*. https://doi.org/10.15778/RESIF.X72010

Chevrot, S., Sylvander, M., Diaz, J., Martin, R., Mouthereau, F., Manatschal, G., et al. (2018). The non-cylindrical crustal architecture of the Pyrenees. *Scientific Reports*, *8*(1), 9591. https://doi.org/10.1038/s41598-018-27889-x

Chevrot, S., Villaseñor, A., Sylvander, M., Benahmed, S., Beucler, E., Cougoulat, G., et al. (2014). High-resolution imaging of the Pyrenees and Massif Central from the data of the PYROPE and IBERARRAY portable array deployments. *Journal of Geophysical Research: Solid Earth*, *119*(8), 6399–6420. https://doi.org/10.1002/2014jb010953

Craven, P., & Wahba, G. (1978). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, *31*(4), 377–403. https://doi.org/10.1007/bf01404567

Cressie, N. (1993). Statistics for spatial data. In *Statistics for spatial data* (p. 900).

Cressie, N., & Hawkins, D. (1980). Robust estimation of the variogram: I. *Journal of the International Association for Mathematical Geology*, *12*(2), 115–125. https://doi.org/10.1007/bf01035243

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

de Ridder, S. A. L., & Maddison, J. R. (2018). Full wavefield inversion of ambient seismic noise. *Geophysical Journal International*, *215*(2), 1215–1230. https://doi.org/10.1093/gji/ggy328

Diaz, J., Vergés, J., Chevrot, S., Antonio-Vigil, A., Ruiz, M., Sylvander, M., & Gallart, J. (2018). Mapping the crustal structure beneath the eastern Pyrenees. *Tectonophysics*, *744*, 296–309. https://doi.org/10.1016/j.tecto.2018.07.011

El-Sharkawy, A., Meier, T., Lebedev, S., Behrmann, J. H., Hamada, M., Cristiano, L., et al. (2020). The slab puzzle of the Alpine-Mediterranean region: Insights from a new, high-resolution, shear wave velocity model of the upper mantle. *Geochemistry, Geophysics, Geosystems*, *21*(8). https://doi.org/10.1029/2020gc008993

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd* (Vol. 96, pp. 226–231).

Feng, L., & Diaz, J. (2023). A high-resolution shear velocity model of the crust and uppermost mantle beneath westernmost Mediterranean including radial anisotropy. *Journal of Geophysical Research: Solid Earth*, *128*(9). https://doi.org/10.1029/2023jb026868

Giacomuzzi, G., Chiarabba, C., & De Gori, P. (2011). Linking the Alps and Apennines subduction systems: New constraints revealed by high-resolution teleseismic tomography. *Earth and Planetary Science Letters*, *301*(3–4), 531–543. https://doi.org/10.1016/j.epsl.2010.11.033

Granet, M., Wilson, M., & Achauer, U. (1995). Imaging a mantle plume beneath the French Massif Central. *Earth and Planetary Science Letters*, *136*(3–4), 281–296. https://doi.org/10.1016/0012-821x(95)00174-b

Hariharan, A., Dalton, C. A., Babikoff, J., & Ekström, G. (2022). Controls on surface wave overtone interference. *Geophysical Journal International*, *228*(3), 1665–1683. https://doi.org/10.1093/gji/ggab424

Hariharan, A., Dalton, C. A., Ma, Z., & Ekström, G. (2020). Evidence of overtone interference in fundamental-mode Rayleigh wave phase and amplitude measurements. *Journal of Geophysical Research: Solid Earth*, *125*(1). https://doi.org/10.1029/2019jb018540

Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning*. Springer.

Hetényi, G., Molinari, I., Clinton, J., Bokelmann, G., Bondár, I., Crawford, W. C., et al. (2018). The Alparray seismic network: A large-scale European experiment to image the Alpine orogen. *Surveys in Geophysics*, *39*(5), 1009–1033. https://doi.org/10.1007/s10712-018-9472-4

Institute Earth Sciences "Jaume Almera" CSIC (ICTJA Spain). (2007). IberArray [dataset]. *International Federation of Digital Seismograph Networks*. Retrieved from https://www.fdsn.org/networks/detail/IB/doi:10.7914/SN/IB

Jin, G., & Gaherty, J. B. (2015). Surface wave phase-velocity tomography based on multichannel cross-correlation. *Geophysical Journal International*, *201*(3), 1383–1398. https://doi.org/10.1093/gji/ggv079

Kan, L., Chevrot, S., & Monteiller, V. (2023). Dehydration of the Subducting Juan de Fuca Plate and Fluid Pathways Revealed by Full Waveform Inversion of Teleseismic P and SH waves in Central Oregon. *Journal of Geophysical Research: Solid Earth*, *128*(4). https://doi.org/10.1029/2022jb025506

Kästle, E. D., El-Sharkawy, A., Boschi, L., Meier, T., Rosenberg, C., Bellahsen, N., et al. (2018). Surface wave tomography of the Alps using ambient-noise and earthquake phase velocity measurements. *Journal of Geophysical Research: Solid Earth*, *123*(2), 1770–1792. https://doi.org/10.1002/2017jb014698

Kästle, E. D., Rosenberg, C., Boschi, L., Bellahsen, N., Meier, T., & El-Sharkawy, A. (2020). Slab break-offs in the Alpine subduction zone. *International Journal of Earth Sciences*, *109*(2), 587–603. https://doi.org/10.1007/s00531-020-01821-z

Kolínský, P. (2004). Surface waves dispersion curves of Eurasian earthquakes: The SVAL program. *Acta Geodynamica et Geromaterialia*, *1*(2), 165–183.

Kolínský, P., & Bokelmann, G. (2019). Arrival angles of teleseismic fundamental mode Rayleigh waves across the AlpArray. *Geophysical Journal International*, *218*(1), 115–144. https://doi.org/10.1093/gji/ggz081

Kolínský, P., Valenta, J., & Málek, J. (2014). Velocity model of the Hronov-poříčíFault Zone from Rayleigh wave dispersion. *Journal of Seismology*, *18*(3), 617–635. https://doi.org/10.1007/s10950-014-9433-4

Koulakov, I., Kaban, M. K., Tesauro, M., & Cloetingh, S. (2009). P- and S-velocity anomalies in the upper mantle beneath Europe from tomographic inversion of ISC data. *Geophysical Journal International*, *179*(1), 345–366. https://doi.org/10.1111/j.1365-246x.2009.04279.x

Lanari, R., Boutoux, A., Faccenna, C., Herman, F., Willett, S. D., & Ballato, P. (2023). Cenozoic exhumation in the Mediterranean and the Middle East. *Earth-Science Reviews*, *237*, 104328. https://doi.org/10.1016/j.earscirev.2023.104328

Lebedev, S., Adam, J. M.-C., & Meier, T. (2013). Mapping the Moho with seismic surface waves: A review, resolution analysis, and recommended inversion strategies. *Tectonophysics*, *609*, 377–394. https://doi.org/10.1016/j.tecto.2012.12.030

Lehujeur, M., & Chevrot, S. (2020). On the validity of the Eikonal equation for surface-wave phase-velocity tomography. *Geophysical Journal International*, *223*(2), 908–914. https://doi.org/10.1093/gji/ggaa352

Levander, A., Humphreys, G., & Pat, R. (2009). Program to investigate convective Alboran sea system overturn [dataset]. *International Federation of Digital Seismograph Networks*. Retrieved from https://www.fdsn.org/networks/detail/XB_2009/doi:10.7914/SN/XB_2009

Levshin, A., Ratnikova, L., & Berger, J. (1992). Peculiarities of surface-wave propagation across central Eurasia. *Bulletin of the Seismological Society of America*, *82*(6), 2464–2493. https://doi.org/10.1785/bssa0820062464

Lin, F.-C., & Ritzwoller, M. H. (2011). Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure: Helmholtz surface wave tomography. *Geophysical Journal International*, *186*(3), 1104–1120. https://doi.org/10.1111/j.1365-246x.2011.05070.x

Lin, F.-C., Ritzwoller, M. H., & Snieder, R. (2009). Eikonal tomography: Surface wave tomography by phase front tracking across a regional broad-band seismic array. *Geophysical Journal International*, *177*(3), 1091–1110. https://doi.org/10.1111/j.1365-246x.2009.04105.x

Lippitsch, R., Kissling, E., & Ansorge, J. (2003). Upper mantle structure beneath the Alpine orogen from high-resolution teleseismic tomography. *Journal of Geophysical Research*, *108*(B8). https://doi.org/10.1029/2002jb002016

Liu, Y., & Holt, W. E. (2015). Wave gradiometry and its link with Helmholtz equation solutions applied to USArray in the eastern U.S. *Journal of Geophysical Research: Solid Earth*, *120*(8), 5717–5746. https://doi.org/10.1002/2015jb011982

Lu, Y., Stehly, L., & Paul, A. (2018). High-resolution surface wave tomography of the European crust and uppermost mantle from ambient seismic noise. *Geophysical Journal International*, *214*(2), 1136–1150. https://doi.org/10.1093/gji/ggy188

Missenard, Y., Zeyen, H., Frizon de Lamotte, D., Leturmy, P., Petit, C., Sébrier, M., & Saddiqi, O. (2006). Crustal versus asthenospheric origin of relief of the Atlas Mountains of Morocco. *Journal of Geophysical Research*, *111*(B3). https://doi.org/10.1029/2005jb003708

Nouibat, A., Brossier, R., Stehly, L., Cao, J., & Paul, A. (2023). Ambient-noise wave-equation tomography of the Alps and Ligurian-Provence Basin. *Journal of Geophysical Research: Solid Earth*, *128*(10). https://doi.org/10.1029/2023jb026776

Ruan, Y., Lei, W., Modrak, R., Örsvuran, R., Bozdağ, E., & Tromp, J. (2019). Balancing unevenly distributed data in seismic tomography: A global adjoint tomography example. *Geophysical Journal International*, *219*(2), 1225–1236. https://doi.org/10.1093/gji/ggz356

Seber, D., Barazangi, M., Tadili, B. A., Ramdani, M., Ibenbrahim, A., & Ben Sari, D. (1996). Three-dimensional upper mantle structure beneath the intraplate Atlas and interplate Rif mountains of Morocco. *Journal of Geophysical Research*, *101*(B2), 3125–3138. https://doi.org/10.1029/95jb03112

Sergeant, A., Delage, K., & Chevrot, S. (2024). pyHelmholtzTomo Python library. Retrieved from https://hal.science/hal-04757812v1

Sethian, J. A., & Popovici, A. M. (1999). 3-D traveltime computation using the fast marching method. *Geophysics*, *64*(2), 516–523. https://doi.org/10.1190/1.1444558

Sobolev, S. V., Zeyen, H., Granet, M., Achauer, U., Bauer, C., Werling, F., et al. (1997). Upper mantle temperatures and lithosphere-asthenosphere system beneath the French Massif Central constrained by seismic, gravity, petrologic and thermal observations. *Tectonophysics*, *275*(1–3), 143–164. https://doi.org/10.1016/s0040-1951(97)00019-x

Tanimoto, T. (1990). Modelling curved surface wave paths: Membrane surface wave synthetics. *Geophysical Journal International*, *102*(1), 89–100. https://doi.org/10.1111/j.1365-246x.1990.tb00532.x

Teixell, A., Ayarza, P., Zeyen, H., Fernàndez, M., & Arboleya, M. (2005). Effects of mantle upwelling in a compressional setting: The Atlas Mountains of Morocco. *Terra Nova*, *17*(5), 456–461. https://doi.org/10.1111/j.1365-3121.2005.00633.x

Tenorio, L. (2001). Statistical regularization of inverse problems. *SIAM Review*, *43*(2), 347–366. https://doi.org/10.1137/s0036144500358232

Tromp, J., & Dahlen, F. A. (1993). Variational principles for surface wave propagation on a laterally heterogeneous Earth-III. Potential representation. *Geophysical Journal International*, *112*(2), 195–209. https://doi.org/10.1111/j.1365-246x.1993.tb01449.x

van Leeuwen, T., & Herrmann, F. J. (2013). Mitigating local minima in full-waveform inversion by expanding the search space. *Geophysical Journal International*, *195*(1), 661–667. https://doi.org/10.1093/gji/ggt258

Villaseñor, A., Chevrot, S., Harnafi, M., Gallart, J., Pazos, A., Serrano, I., et al. (2015). Subduction and volcanism in the Iberia–North Africa collision zone from tomographic images of the upper mantle. *Tectonophysics*, *663*, 238–249. https://doi.org/10.1016/j.tecto.2015.08.042

Virieux, J., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, *74*(6), WCC1–WCC26. https://doi.org/10.1190/1.3238367

Wahba, G. (1990). *Spline models for observational data*. SIAM.

Wahba, G., & Wendelberger, J. (1980). Some new mathematical methods for variational objective analysis using splines and cross validation. *Monthly Weather Review*, *108*(8), 1122–1143. https://doi.org/10.1175/1520-0493(1980)108⟨1122:snmmfv⟩2.0.co;2

Wang, Y., Chevrot, S., Monteiller, V., Komatitsch, D., Mouthereau, F., Manatschal, G., et al. (2016). The deep roots of the western Pyrenees revealed by full waveform inversion of teleseismic P waves. *Geology*, *44*(6), 475–478. https://doi.org/10.1130/g37812.1

Wielandt, E. (1993). Propagation and structural interpretation of non-plane waves. *Geophysical Journal International*, *113*(1), 45–53. https://doi.org/10.1111/j.1365-246x.1993.tb02527.x

Yang, Y., & Forsyth, D. W. (2006). Regional tomographic inversion of the amplitude and phase of Rayleigh waves with 2-D sensitivity kernels. *Geophysical Journal International*, *166*(3), 1148–1160. https://doi.org/10.1111/j.1365-246x.2006.02972.x

Zhan, Z., Li, Q., & Huang, J. (2018). Application of wavefield compressive sensing in surface wave tomography. *Geophysical Journal International*, *213*(3), 1731–1743. https://doi.org/10.1093/gji/ggy082

Zhao, L., Paul, A., Malusà, M. G., Xu, X., Zheng, T., Solarino, S., et al. (2016). Continuity of the Alpine slab unraveled by high-resolution P wave tomography. *Journal of Geophysical Research: Solid Earth*, *121*(12), 8720–8737. https://doi.org/10.1002/2016jb013310

Zhao, L., Paul, A., & Solarino, S., & RESIF. (2016). Seismic network Yp: CIFALPS temporary experiment (China-Italy-France Alps seismic transect) [dataset]. *RESIF - Réseau Sismologique et Géodésique Français*. https://doi.org/10.15778/RESIF.YP2012

Zhou, P., Chevrot, S., Lehujeur, M., Xia, S., & Yu, C. (2022). Eikonal surface wave tomography of central and eastern China. *Geophysical Journal International*, *231*(3), 1865–1879. https://doi.org/10.1093/gji/ggac296

Zhou, Y., Dahlen, F. A., & Nolet, G. (2004). Three-dimensional sensitivity kernels for surface wave observables. *Geophysical Journal International*, *158*(1), 142–168. https://doi.org/10.1111/j.1365-246x.2004.02324.x