

Imaging Alpine crust using ambient noise wave-equation tomography

Y. Lu¹, L. Stehly, R. Brossier, A. Paul and AlpArray Working Group*

Université Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISTerre, 38000 Grenoble, France. E-mail: yang.lu@univ-grenoble-alpes.fr

Accepted 2020 March 18. Received 2020 March 16; in original form 2019 October 10

SUMMARY

We present an improved crustal Vs model and Moho depth map using ambient noise wave-equation tomography. The so-called ‘ambient noise wave-equation tomography’ is a method to invert seismic ambient noise phase dispersion data based on elastic waveform simulation, which accounts for 3-D and finite-frequency effects. We use cross-correlations of up to 4 yr of continuous vertical-component ambient seismic noise recordings from 304 high-quality broad-band stations in the Alpine region. We use model LSP_Eucrust1.0 obtained from traditional ambient noise tomography as initial model, and we iteratively improve the initial model by minimizing frequency-dependent phase traveltimes between the observed and synthetic waveforms of Rayleigh waves in the period range 10–50 s. We obtain the final model after 15 iterations with ~65 per cent total misfit reduction compared to the initial model. At crustal depth, the final model significantly enhances the amplitudes and adjusts the shapes of velocity anomalies. At Moho and upper-mantle depth, the final model corrects an obvious systematic velocity shift of the initial model. The resulting isovelocity Moho map confirms a Moho step along the external side of the external crystalline massifs of the northwestern Alps and reveals underplated gabbroic plutons in the lower most crust of the central and eastern Alps. Ambient noise wave-equation tomography turns out to be a useful tool to refine shear wave velocity models obtained by traditional ambient noise tomography based on ray theory.

Key words: Seismic noise; Tomography; Crustal structure; Moho discontinuity.

1 INTRODUCTION

Following Shapiro *et al.* (2005), surface waves reconstructed from cross-correlations of ambient seismic noise have been widely used to image Earth structures at all scales, giving rise to a class of methods referred to as ‘ambient noise tomography’ (at local scale: e.g. Brenguier *et al.* 2007; Picozzi *et al.* 2009, at regional scale: e.g. Bensen *et al.* 2009; Stehly *et al.* 2009 and at global scale: e.g. Nishida *et al.* 2009; Haned *et al.* 2016). Traditional ambient noise tomography exploits the dispersive characteristics of surface waves using a two-step inversion approach: (i) period-dependent group or phase velocity measurements are inverted to estimate 2-D velocity maps based on the ray-theory assumption; (ii) local dispersion curves extracted from 2-D velocity maps at grid points are inverted to obtain 1-D shear wave velocity models, which together construct a final pseudo 3-D model.

Traditional ambient noise tomography has been used to resolve structures of the Alpine crust and uppermost mantle in a variety of studies (Yang *et al.* 2007; Li *et al.* 2009; Stehly *et al.* 2009; Verbeke *et al.* 2012; Molinari *et al.* 2015; Kästle *et al.* 2016; Lu

et al. 2018). For instance, Stehly *et al.* (2009) constructed a crustal shear wave velocity model and Moho map for the greater Alpine region using cross-correlations computed for 150 stations in the period band 5–80 s. The most up-to-date study on this region is by Lu *et al.* (2018), who made use of a large cross-correlation data set computed for 1293 broad-band stations across much of Europe. Rayleigh wave group velocity measurements were used to estimate 2-D group velocity maps in the period band 5–150 s. A Bayesian probabilistic depth inversion was applied to construct the 3-D model LSP_Eucrust1.0 for the European crust and uppermost mantle, including an isotropic shear wave velocity model and a probabilistic model of interfaces, which showed good agreement with previous regional geophysical studies.

However, the two-step inversion approach employed by traditional ambient noise tomography has two basic limitations: (i) the use of ray theory, which is simplified to straight ray assumption in most cases, essentially states that the surface wave traveltimes is only sensitive to the structure along the zero-width source–receiver path. It only holds under high-frequency assumption where the scale length of medium heterogeneities is much larger than the seismic wavelength (Snieder 1986; Yomogida & Aki 1987; Spetzler *et al.* 2002; Zhou *et al.* 2004); (ii) the point-by-point 1-D depth inversion

* AlpArray Working Group (<http://www.alparray.ethz.ch/home/>)

in the second step is only valid in a laterally homogeneous media. Given the strong heterogeneity of the Alpine crust, these two basic limitations may lead to non-negligible misfits that should be properly accounted for.

Nevertheless, methodological efforts to improve ambient noise tomography have been mainly dedicated to improvements of technical aspects of the two-step workframe, such as traveltimes measurement and correction (e.g. Ekström *et al.* 2009; Yao & van der Hilst 2009), 2-D Eikonal and Helmholtz tomography (e.g. Lin *et al.* 2009; Lin & Ritzwoller 2011), 2-D ray-tracing tomography (e.g. Nicolson *et al.* 2012; Saygin & Kennett 2012), 2-D Bayesian tomography (e.g. Bodin *et al.* 2012a; Young *et al.* 2013) and 1-D nonlinear depth inversion (e.g. Bodin *et al.* 2012b; Lu *et al.* 2018).

Following rapid advances in computational facilities and numerical methods, tomography methods based on solving the wave equation became feasible in recent years. These methods consist in minimizing of a misfit function defined by traveltimes or waveform differences between the observed and synthetic waveforms. The synthetic waveforms are computed from an initial model using full 3-D wavefield modelling techniques, such as the finite-difference method (Graves 1996; Operto *et al.* 2007) or the spectral-element method (SEM; Komatitsch & Vilotte 1998; Komatitsch & Tromp 1999). Sensitivity kernels or gradients of the misfit function with respect to physical parameters are usually calculated by the adjoint-state approach, given by the interaction between the forward and adjoint wavefields (Tromp *et al.* 2004; Liu & Tromp 2006; Plessix 2006). The initial model is iteratively improved using optimization methods, like the conjugate-gradient method (Mora 1987) or the quasi-Newton method (Nocedal 1980). Depending on the choice of the misfit function, these methods roughly fall into two categories: wave-equation tomography (WET) when the misfit function is defined by traveltimes difference (Luo & Schuster 1991), and full waveform inversion (FWI) when the misfit function involves waveform difference (Virieux & Operto 2009). Here, we clarify that the ‘adjoint tomography’ refers specifically to FWI applied to large-scale objects (regional to global) using the adjoint-state approach for the computation of sensitivity kernels (Liu & Gu 2012).

Applications of wave-equation-based tomographic methods are primarily conducted using earthquake data at large-scale and artificial sources for smaller-scale objectives. These applications take advantage of accurate 3-D wavefield modelling, providing improved images of the Earth beyond those obtained by classical tomographic techniques (e.g. applications at regional and continental scales: Fichtner *et al.* 2009; Tape *et al.* 2009; Zhu *et al.* 2012; Yuan *et al.* 2014; Fichtner & Villaseñor 2015; Beller *et al.* 2018).

However, FWI of ambient noise cross-correlation is still challenging. Indeed, while the phase (traveltimes) of the Green’s function can be accurately measured from cross-correlation, its amplitude is poorly estimated due to its high sensitivity to the anisotropic distribution of noise sources, the medium heterogeneity and the pre-processing procedure (Stehly & Boué 2017; Sager *et al.* 2018). Fichtner (2014) stated that FWI of ambient noise cross-correlations is not achievable unless the noise source distribution and pre-processing procedure are properly accounted for in the inversion. Tromp *et al.* (2010) attempted to drop the concept of Green’s function retrieval and to establish cross-correlations as self-consistent observables. The basic idea is to compute the synthetic cross-correlation for a station pair by using simulated noise fields in the forward simulation, rather than by treating the station pair as a source–receiver pair. With synthetic data, Sager *et al.* (2018) showed improved accuracy and resolution of images with respect to those

obtained from WET. Nonetheless, the method requires a good estimate of noise source fields, which are temporally and spatially non-stationary. Thus, FWI of ambient noise cross-correlation has not been applied in practice so far.

Meanwhile, some recent tomographic applications reveal the potential of providing improved velocity models by inverting ambient noise cross-correlation data using WET. For instance, Chen *et al.* (2014) applied ambient noise WET to refine a crustal model of southeastern Tibet using ambient noise recorded at 25 stations in the period band 10–40 s. The new method addresses the limitations of traditional ambient noise tomography as it accounts for 3-D and finite frequency effects. Unlike earthquake-based WET, observed waveforms in ambient noise WET are Green’s function reconstructed from cross-correlations, primarily for the vertical component of Rayleigh waves. Synthetic waveforms are generated by applying a single force on the surface or near-surface. The misfit function can be defined by cross-correlation type traveltimes differences as in classical WET (Luo & Schuster 1991), or frequency-dependent phase traveltimes differences (Tape *et al.* 2009). To avoid any confusion of notation, we combine the two names ‘full-wave ambient noise tomography’ (Gao & Shen 2014) and ‘ambient noise adjoint tomography’ (Chen *et al.* 2014; Liu *et al.* 2017) in ‘ambient noise WET’, since these methods define the misfit function using traveltimes differences as in WET.

In this study, we apply ambient noise WET to the Alpine region. The primary goal is to refine the shear wave velocity model *LSP_Eucrust1.0* obtained from traditional ambient noise tomography. To that end, we select ~10 000 high-quality cross-correlations computed for 304 broad-band stations (Fig. 1b). Following a similar approach as Chen *et al.* (2014) and Liu *et al.* (2017), we progressively improve the initial model *LSP_Eucrust1.0* by minimizing frequency-dependent phase traveltimes differences between the observed and synthetic Rayleigh waves in the period range 10–50 s. The inversion is performed using the SEM46 code package developed within the SEISCOPE consortium initially for exploration targets (Trinh *et al.* 2019). The SEM46 code package employs SEM for forward simulation, adjoint-state method for misfit kernel computation and the SEISCOPE optimization toolbox for model update (Métivier & Brossier 2016).

The paper is organized as follows. First, we briefly introduce the ambient noise cross-correlation data and initial model in Section 2. Then, we present our tomographic workflow in Section 3, and we show the misfit evolution (total misfit, misfit histograms, misfit spatial distribution) and resolution tests in Section 4. We display the tomographic results (V_s depth slices and Moho depth map) in Section 5. Finally, we discuss some important aspects of the methodology in Section 6.

2 DATA AND INITIAL MODEL

This study uses the data set of Lu *et al.* (2018), which consists in cross-correlations computed from 4 yr (from July 2012 to June 2016) of vertical-component seismic noise recorded by 1293 broad-band stations across Europe. We select cross-correlations related to 304 stations in the Alpine region (Fig. 1, 3.5°E–18.5°E; 43°N–49°N). All 304 stations serve as virtual receivers, from which we select 64 stations as virtual sources to reduce computational cost and create an uniform distribution of virtual sources, providing us with ~18 000 cross-correlations. Instead of using the wide period band 5–150 s as in Lu *et al.* (2018), we apply a bandpass filter 10–50 s to cross-correlations to further reduce the high computational cost of

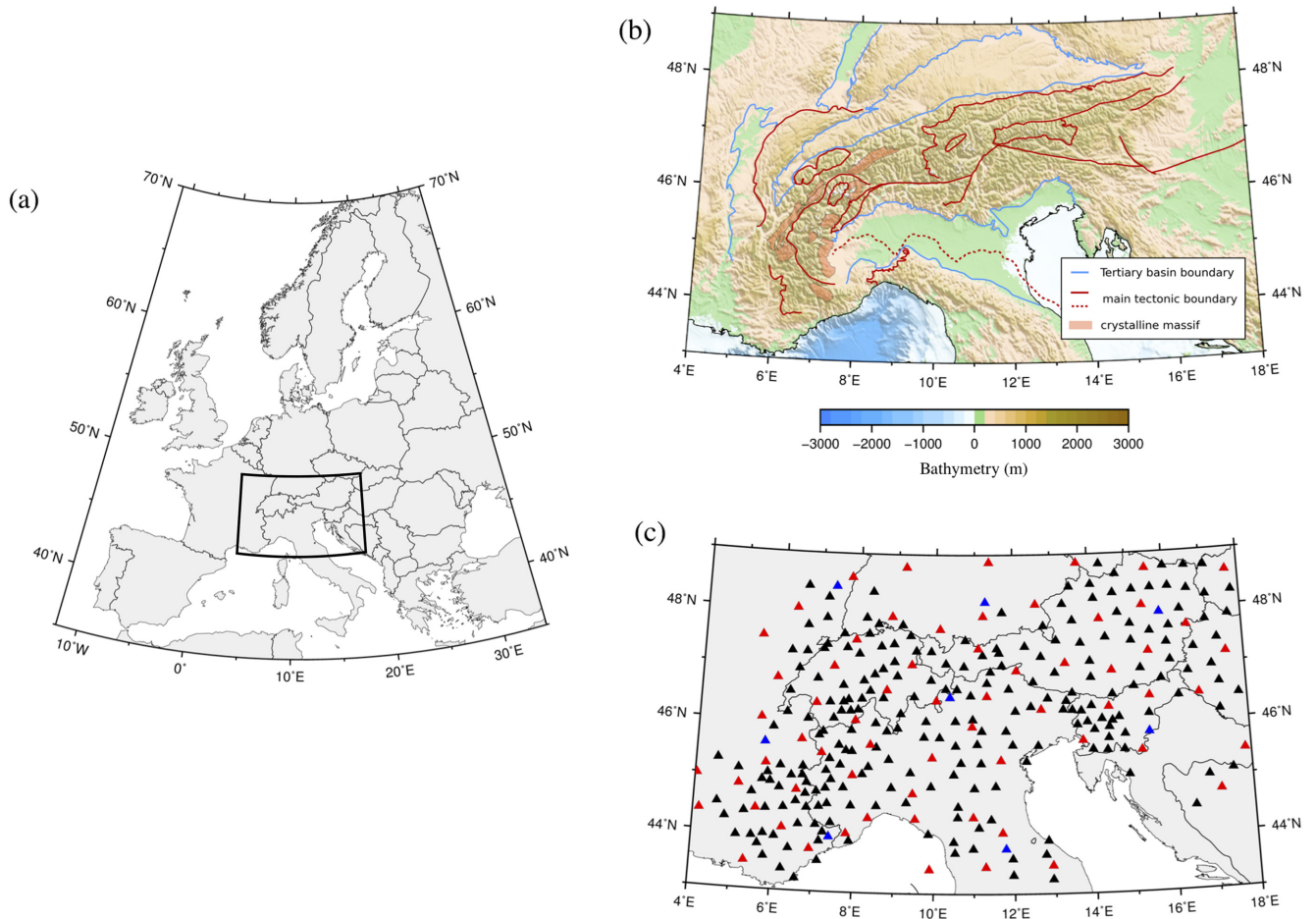


Figure 1. Tectonic map and seismic station map of the study region. (a) The black box indicates the study region of this work; (b) generalized tectonic map of the Alps from Schmid *et al.* (2004) and Handy *et al.* (2010); (c) location map of the 304 seismic stations used in this tomographic study (triangles). Red triangles: 64 selected virtual source stations for inversion. Blue triangles: 8 selected virtual source stations for validation.

3-D numerical simulation. We keep only reliable cross-correlations by applying the following criteria: (1) we choose cross-correlations with signal-to-noise ratio (SNR) greater than 5, while the SNR is defined as the ratio between the envelope maximum of surface waves and the standard deviation of late coda waves. Station pairs with SNR smaller than 5 show low similarity between the observed and synthetic waveforms due to the high noise level in the observed waveforms (see Supporting Information Fig. S1); (2) we require the inter-station distance to be larger than one wavelength for the maximum period considered (Luo *et al.* 2015). Accordingly, we keep $\sim 10\,000$ high-quality cross-correlations in our inversion data set.

Besides, we build a validation data set to avoid over-interpreting the inversion data set. The validation data set is independent of the inversion data set, and it is not involved in the inversion procedure. It consists in ~ 1400 high-quality cross-correlations related to 8 virtual source stations that are selected using the same criteria as for the inversion data set (Fig. 1b). The basic motivation for a validation data set is that an improved model should provide better fit to both inversion and validation data sets. If not, the inversion procedure would start mapping errors of the inversion data set into the model, and the iterations should stop.

We choose model LSP_Eucrust1.0 derived from traditional ambient noise tomography as initial model (Fig. 2). On one hand, LSP_Eucrust1.0 is so far the best resolved shear wave velocity

model at the scale of the whole Alpine crust. It was validated by comparison with different geophysical studies, in particular controlled source seismic experiments at various locations along the Alpine arc. On the other hand, an important goal of the study is to test how ambient noise WET improves traditional ambient noise tomography results.

3 METHOD AND IMPLEMENTATION

We present the ambient noise WET method and implementation in this section. The main characteristics of our tomography are:

- (i) observations are vertical components of Rayleigh waves reconstructed from noise cross-correlations;
- (ii) forward simulation and misfit gradient are computed using 3-D full elastic wavefield simulation;
- (iii) misfit function is defined by frequency-dependent phase traveltimes differences between observed and synthetic waveforms.

We follow the workflow shown in Fig. 3, while details of the main steps are given below. Since the theory and concepts of the wave-equation-based tomography have been well documented in many studies (e.g. Luo & Schuster 1991; Tromp *et al.* 2004; Virieux & Operto 2009; Liu & Gu 2012), we outline our implementation with a focus on the specific features related to noise cross-correlation data.

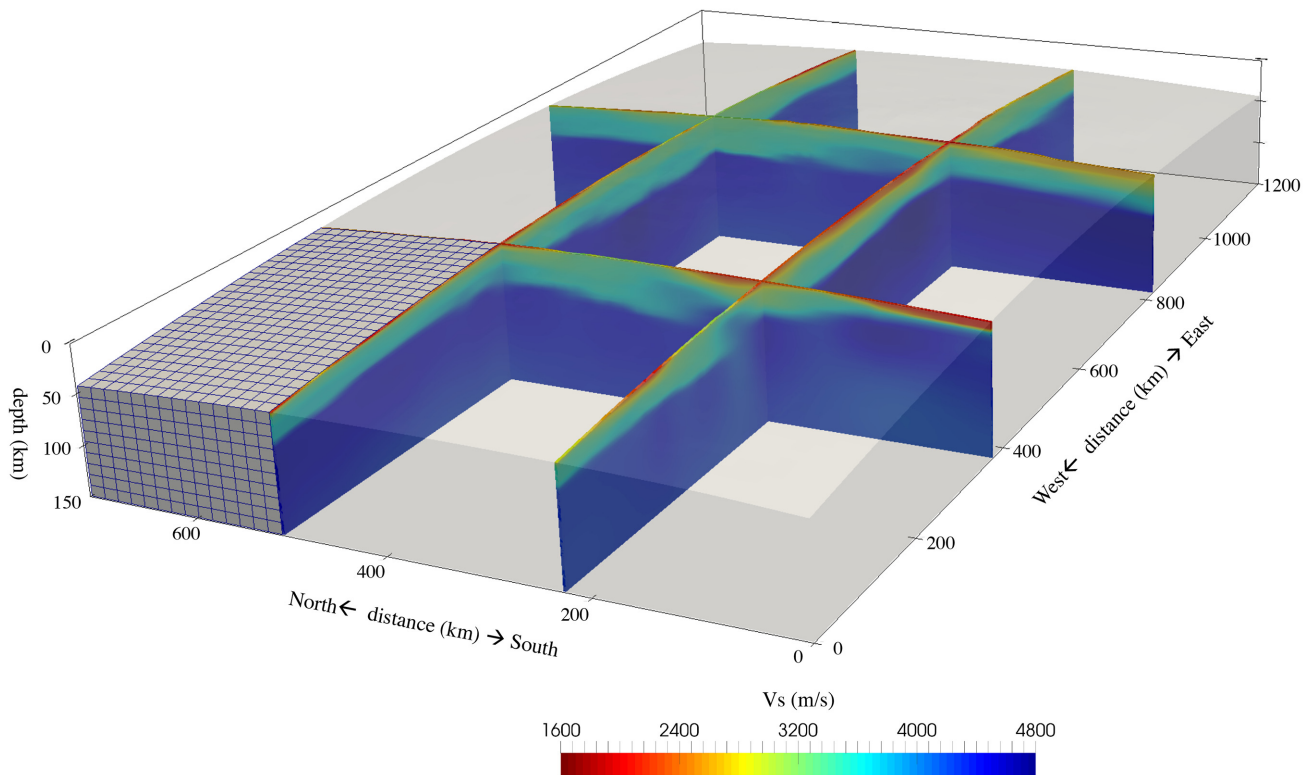


Figure 2. Mesh and shear wave velocity of the initial model. The grey volume indicates the scope of the initial model. It is centred at $[11^\circ\text{E}, 46^\circ\text{N}]$ with lateral dimensions 1200 km in the east–west direction and 720 km in the north–south direction. We show shear wave velocity values of the initial model in four representative depth-sections. We show a chunk of the mesh for the northwest portion of the initial model. The topography and curvature of the Earth are accounted by vertically deforming the mesh grid.

3.1 Forward simulation

The SEM has proven to be an accurate and efficient method of wavefield simulation in a 3-D heterogeneous media (Komatitsch & Vilotte 1998; Komatitsch & Tromp 1999, 2002a,b). It benefits from the geometrical flexibility of finite-element method and the exponential convergence rate of spectral methods. Given these advantages, the SEM is recognized as the most popular forward simulation method in wave-equation-based tomographic studies at regional to global scales.

We compute synthetic waveforms using the 3-D elastic wave-equation solver of SEM46 package, which is developed in a hexahedra-based SEM frame (Trinh *et al.* 2019). Our initial model is discretized using a mesh built in Cartesian coordinates (Fig. 2). It is centred at $[11^\circ\text{E}, 46^\circ\text{N}]$ with lateral dimensions 1200 km in the east–west direction and 720 km in the north–south direction. We handle the topography and the curvature of the Earth by vertically deforming the mesh. Thus, the vertical dimension ranges approximately from 150 km in the centre to 110 km at the edges due to the vertical compression of element size in our mesh. Using Lagrange polynomials of order 4, we choose 15 km as the element size in the horizontal direction, and 10–15 km in the vertical direction with respect to the volume condition to ensure accuracy of wavefield simulation (Komatitsch & Tromp 1999). As a consequence, the average spacing between Gauss–Lobatto–Legendre (GLL) points is around 3.5 km. Since we do not have access to the quality factor (Q) model of the Alpine region, we ignore the attenuation and parametrize each GLL point of the initial model with isotropic S -wave velocity (V_s), P -wave velocity (V_p) and density (ρ). The V_s values of the initial model are taken from model LSP_Eucrust1.0, while V_p and

ρ are converted from V_s using empirical formulas (Ludwig *et al.* 1970; Brocher 2005).

In order to simulate the vertical component of Rayleigh waves, we apply a vertical single force at each virtual source position on the free surface, and a Dirac delta function filtered in the desired period band 10–50 s is used as source time function. In this way, the record of synthetic waveform u_{syn} at a receiver position for model m is the convolution product between the source time function $s(t)$ and the synthetic Green’s function $G_{\text{syn}}(m, t)$ for the source–receiver pair:

$$u_{\text{syn}} = s(t) \otimes G_{\text{syn}}(m, t), \quad (1)$$

where \otimes denotes the convolution operator. The time step of simulation is 0.004 s, meeting the Courant–Friedrichs–Lewy stability condition (Komatitsch & Tromp 1999).

3.2 Misfit function

The Green’s function $G_{\text{obs}}(t)$ for a station pair can be estimated from the time derivative of cross-correlation $C(t)$ of seismic noise recorded at the two stations (Lobkis & Weaver 2001; Weaver & Lobkis 2001; Derode *et al.* 2003a,b; Snieder 2004; Wapenaar 2004):

$$G_{\text{obs}}(t) \approx -A \frac{\partial C(t)}{\partial t}, \quad (2)$$

where A is the amplitude term related to noise source energy. To compare with synthetic waveforms, we convolve the source time function $s(t)$ used for the simulations with the opposite of the time

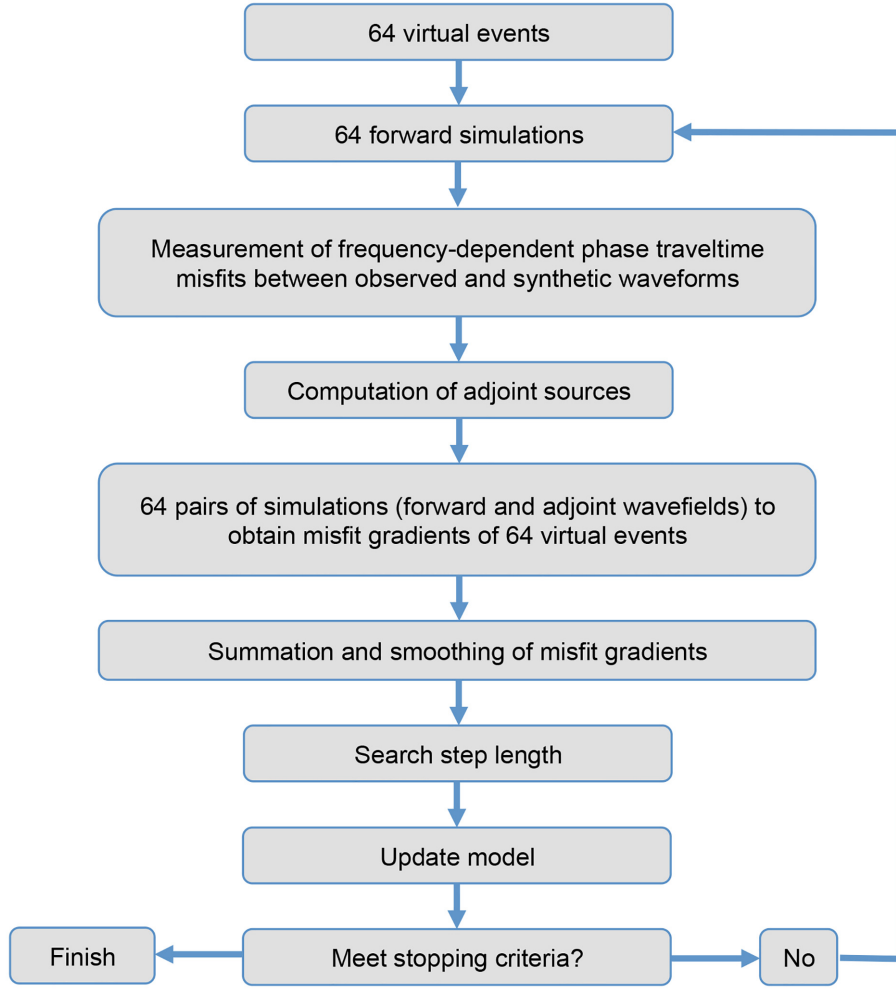


Figure 3. General ambient noise wave-equation tomography workflow.

derivative of the cross-correlation:

$$u_{obs} = s(t) \otimes -\frac{\partial C(t)}{\partial t} \approx \frac{1}{A} s(t) \otimes G_{obs}(t). \quad (3)$$

This convolution product serves as observed waveform at each receiver. Both observed and synthetic waveforms are tapered to keep only Rayleigh waves. For this purpose, we use a time window with $\min[D/5, T_{\text{umax}} - 50]$ and $\max[D/2, T_{\text{umax}} + 50]$ as lower and upper limits, where D is the interstation distance expressed in km and T_{umax} refers to the arrival time of the maximum of the envelope of synthetic waveform. $D/5$ and $D/2$ correspond to the range of typical phase arrival times of Rayleigh waves in the 10–55 s period band. $T_{\text{umax}} - 50$ and $T_{\text{umax}} + 50$ refer to 50 s (one maximum period) before and after the maximum of the envelope of the synthetic waveform, which are complementary criteria making it possible to include the entire Rayleigh wave train.

We define the misfit function using frequency-dependent differences in phase traveltimes (Tape *et al.* 2009), rather than the cross-correlation type misfit function used in classical WET (Luo & Schuster 1991). Because the cross-correlation type misfit function involves the amplitude information of waveforms that is not reliable for noise cross-correlations, and it is only suitable for identical waveforms when the difference between the two waveforms is mostly a shift in time than a distortion of waveforms (Fichtner *et al.* 2008).

We measure the phase traveltimes differences using the multitaper method, which is efficient analysing phase spectra for time- and band-limited signals (Thomson 1982). Taking advantage of a series of prolate spheroidal eigentapers (Slepian & Pollak 1961), the method provides independent measurements of phase traveltimes differences between observed and synthetic waveforms. The final measurement is determined as the average of the ensemble of independent measurements. For model m , the misfit function is expressed as

$$F(m) = \frac{1}{2} \sum_i \sum_w \Delta T_i(w, m)^2, \quad (4)$$

where w is the frequency and $\Delta T_i(w, m) = T_i^{\text{syn}}(w, m) - T_i^{\text{obs}}(w)$ denotes the frequency-dependent phase traveltimes differences for the i th station pair. Since we only use phase traveltimes information, the amplitude term A in eq. (2) can be ignored. The phase traveltimes differences between the observed and synthetic waveforms (u_{obs} and u_{syn}) are indeed the differences between the observed and synthetic Green's function computed from model m ($G_{\text{obs}}(t)$ and $G_{\text{syn}}(t, m)$). Unlike earthquake-based tomography, the estimate of source time functions is not required.

3.3 Misfit gradient

We compute the misfit gradients (also referred to as Fréchet kernels or misfit kernels) using the adjoint-state approach (Tromp *et al.* 2004; Liu & Tromp 2006; Plessix 2006). Although Rayleigh waves show much weaker sensitivity to V_p than V_s , the sensitivity kernel with respect to V_p shows quite different shape from that of V_s (Zhou *et al.* 2004). To prevent V_p anomalies to be mapped into the V_s model, we invert the Rayleigh waves traveltime for both V_p and V_s (Chen *et al.* 2014; Gao & Shen 2014; Wang *et al.* 2018). We express the perturbation of the misfit function for model m by the following linear relation:

$$\delta F(m) = \int_V [K_{V_s}(m)\delta \ln m_{V_s} + K_{V_p}(m)\delta \ln m_{V_p}], \quad (5)$$

where $K_{V_s}(m)$ and $K_{V_p}(m)$ denote the misfit gradients with respect to V_s and V_p , while $\delta \ln m_{V_s}$ and $\delta \ln m_{V_p}$ are relative perturbations of the two parameters.

The misfit gradient for each single station pair is given by the interaction of the forward and adjoint wavefields. Regarding our choice of misfit function, the adjoint source is related to the derivative of synthetic waveforms weighted by the frequency-dependent phase difference measurements (Tape 2009). We sum gradients computed for all source stations, and we smooth the summed gradient using an approximate Laplace filter to remove artificial high-wavenumber gradient artefacts (Trinh *et al.* 2017). The filter width decreases with increasing iterations, from 8 to 2 km in the horizontal direction and from 4 to 1 km in the vertical direction, so that the tomography resolves heterogeneities of smaller size.

3.4 Iterative inversion procedure

The inversion corresponds to a local optimization problem based on the second-order Taylor expansion of the misfit function:

$$F(m + \delta m) \approx F(m) + g(m)^T \delta m + \frac{1}{2} \delta m^T H(m) \delta m, \quad (6)$$

where δm is the model update, $g(m)$ is the gradient of the misfit function, and $H(m)$ is the Hessian as a second-order derivative of the misfit function with respect to model parameters. The misfit $F(m + \delta m)$ can be reduced by updating the model via

$$\delta m = -\alpha H(m)^{-1} g(m), \quad (7)$$

where α and $-H(m)^{-1} g(m)$ are the step length and optimal direction of the model update. However, the optimum update direction is usually not available in practice because accessing the full Hessian matrix is unfeasible. In our case, we address the problem using the limited memory quasi-Newton method regarding its simplicity and efficiency (Nocedal 1980). The method approximates $-H(m)^{-1} g(m)$ as a whole using gradient values of a few previous iterations. It shows faster convergence than the nonlinear conjugate gradient and steepest-descent algorithms in the numerical tests conducted by Métivier & Brossier (2016). A line search strategy is employed to find the appropriate step length of the model update.

3.5 Implementation

Our implementation consists in 64 forward simulations for computing synthetic waveforms, and 64 pairs of simulations for computing the misfit gradient at each iteration. The model is divided into 4 sub-domains, and simulations are performed in parallel using 256 cores. For each iteration, the computational cost is approximately 1000

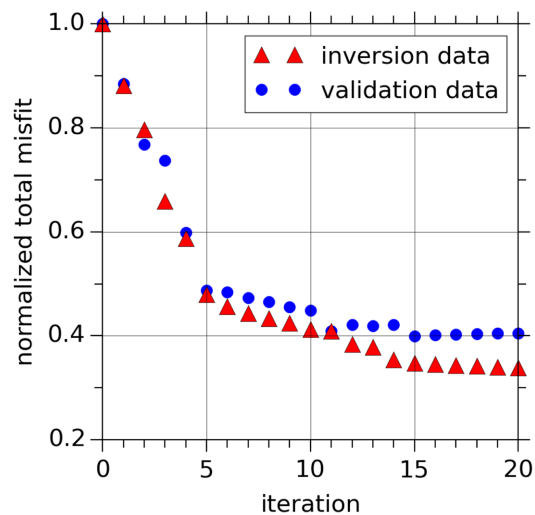


Figure 4. Total misfit reduction over iterations. Red triangles and blue dots represent the evolution of the normalized total misfit for the inversion and validation data sets, respectively.

core hours on Froggy architecture (Intel Sandy Bridge EP E5-2670, 8c/2.6 GHz/20M/8 GT/s, <https://ciment.ujf-grenoble.fr/wiki>).

4 MISFIT REDUCTION AND RESOLUTION TESTS

4.1 Total misfit

Fig. 4 shows the reduction of total misfit over iterations (0–20) for the inversion and validation data sets. The total misfit of the inversion data set shows a quick drop in the first five iterations and a gradual decrease in the following iterations. In general, the total misfit of the validation data set shows a similar decrease as the inversion data set, except for a slow and steady increase after 15 iterations. As mentioned before, an improved model should provide better fits for both inversion and validation data sets. The increase in the misfit of the validation data set after 15 iterations indicates that the inversion procedure starts to over-interpret the data. Thus, we choose to stop at the 15th iteration. For convenience, we name the updated model according to the iteration number, with m_0 being the initial model and m_{15} the final model. The final model (m_{15}) shows a significant reduction of misfit with respect to the initial model (m_0), which amounts to ~ 65 per cent for both inversion and validation data sets. Suporging Information Fig. S2 shows example of observed waveforms and synthetics computed from the initial (m_0) and final (m_{15}) models, which displays obvious improvement of the traveltime fits.

4.2 Histograms of misfit

In Fig. 5, we compare histograms of traveltime misfits (inversion data set) for the initial (m_0) and final models (m_{15}) at 10, 15, 25 and 50 s. The pure traveltime misfits vary according to interstation distance, as station pairs with large interstation distance naturally have large misfits. Thus, in this section and the following, we express the misfit using a distance weighted traveltime misfit $\delta T_i(w, m)/D \times 100$, which corresponds to the traveltime misfit for waves propagating 100 km.

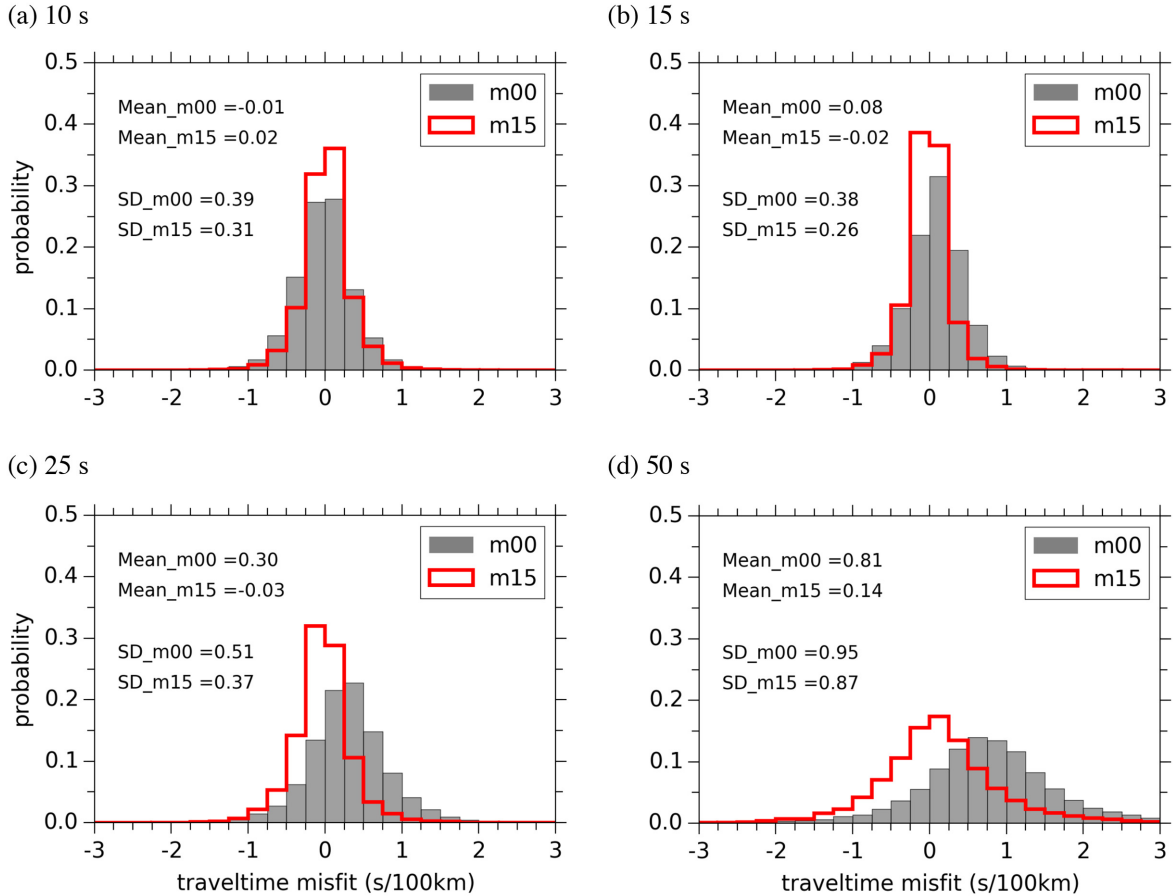


Figure 5. Comparison of histograms of misfits (inversion data set) for the initial (m00) and final models (m15) at 10, 15, 25 and 50 s. The misfit refers to the travelt ime misfit for waves propagating 100 km. Mean refers to the mean misfit, and SD refers to the standard deviation.

The histograms of misfits for the initial model have two main characteristics: (i) the standard deviation of misfits increases with period; (ii) the mean value of misfits shows a positive shift at 25 and 50 s, which means that the velocity values in the initial model are lower than those of the true model. At present, we found no explanation for this bias of the ANT model towards low velocities. In general, these misfits have a limited range, indicating that the initial model is well resolved using traditional ambient noise tomography. For instance, at 25 s, the misfits are in the range 0.30 ± 1.02 s/100 km considering two standard deviations.

Nevertheless, the final model shows an obvious improvement with respect to the initial model in terms of both the standard deviation and mean value of misfits. The misfits of the final model are centred around zero with more concentrated distribution. At 25 s, the misfits of the final model are mostly in the range -0.03 ± 0.74 s/100 km. Assuming an average V_s of 3.8 km s^{-1} at this period, we can roughly estimate the V_s misfits to be $-0.00 \pm 0.10 \text{ km s}^{-1}$ for the final model.

The misfits in both the initial and final models are much larger than potential biases of phase velocity measurements due to the uneven noise source distribution. They are estimated to be smaller than 0.5 per cent (e.g. <0.15 s/100 km at 25 s) by the numerical tests conducted by Yao & van der Hilst (2009).

4.3 Spatial distribution of misfit

In addition to the histograms of Fig. 5, the spatial distribution of misfit can also be used to document the quality of the final model.

At each period, we discretize the study region with a mesh of square cells of size 0.2° . For each cell, we calculate the mean value of misfits for all paths crossing the cell assuming ray paths are great circles. Misfits due to imperfection in the isotropic velocity model should display misfit anomaly patterns coinciding with geological structures, while misfits due to data errors are expected to display random patterns.

Fig. 6 shows the spatial distribution of average misfit for the initial (m00) and final (m15) models at representative periods. For the initial model (a1 and b1), the most striking features are low misfits along the western Alps and high misfits in the north Apennines and the north Adriatic sea. These regional misfit patterns disappear in the final model (a2 and b2). Large misfits in the final model concentrate along the boundary of the study region, where data coverage is poor. Besides, positive misfit shifts are observed for the initial model at long period, which are corrected in the final model.

4.4 Resolution tests

We performed 3-D spike tests to evaluate the resolution of the final V_s model. The spike is a vertical cylinder with radius ~ 25 km, thickness ~ 20 km, and 10 per cent positive velocity perturbation with respect to the initial model (m00). The background model is the initial m00 model to ensure similar resolution characteristics. Using the method described in Section 3, synthetic waveforms are computed and inverted for the same station pair coverage as observed data. Fig. 7 shows the result of the spike tests for the centre of

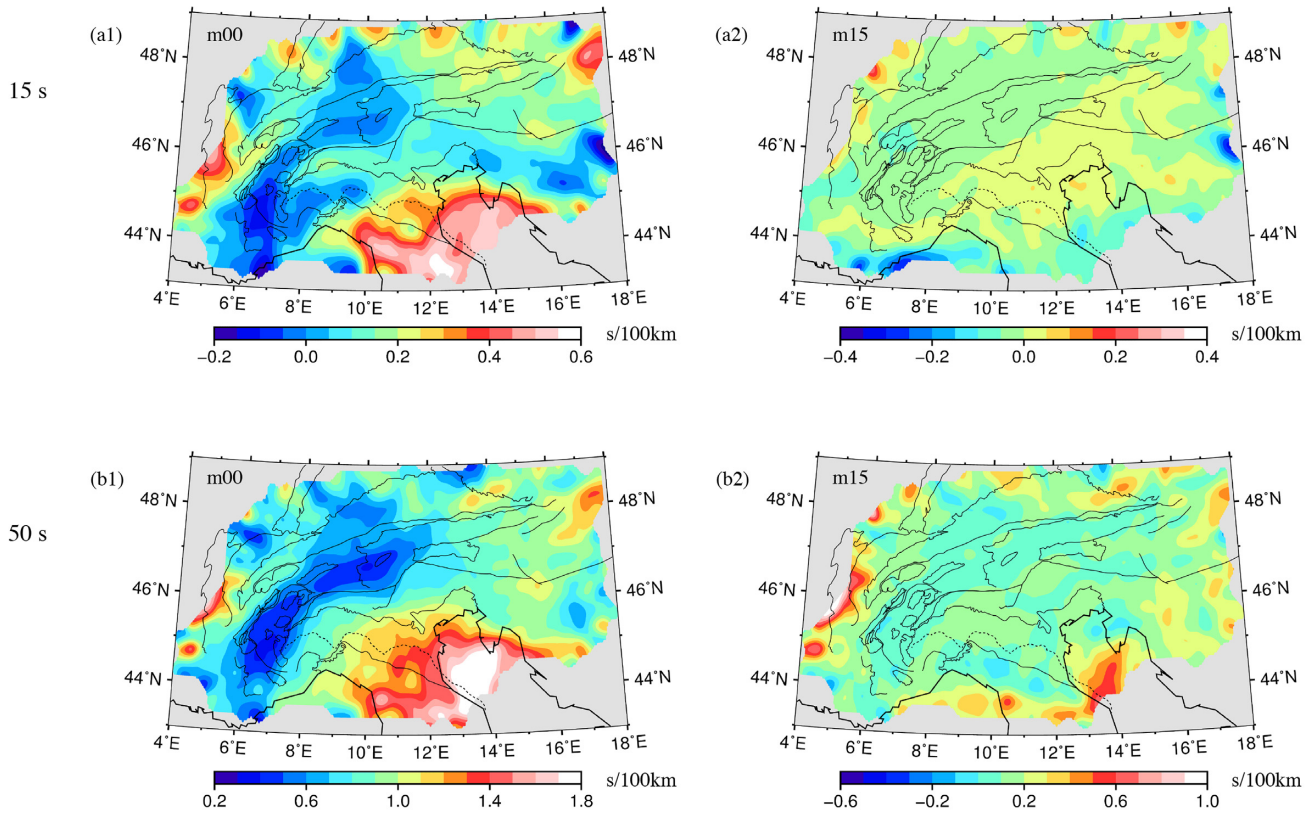


Figure 6. Comparison of spatial distributions of average misfit for the initial model (m00; a1 and b1) and final model (m15; a2 and b2) at 15 and 50 s. The colour scale displays the mean value of misfits for all paths crossing each 0.2° cell assuming ray paths are great circles. We only display cells crossed by more than 10 paths. The black lines in each map correspond to the geological and tectonic boundaries of the generalized tectonic map of the Alps shown in Fig. 1(b).

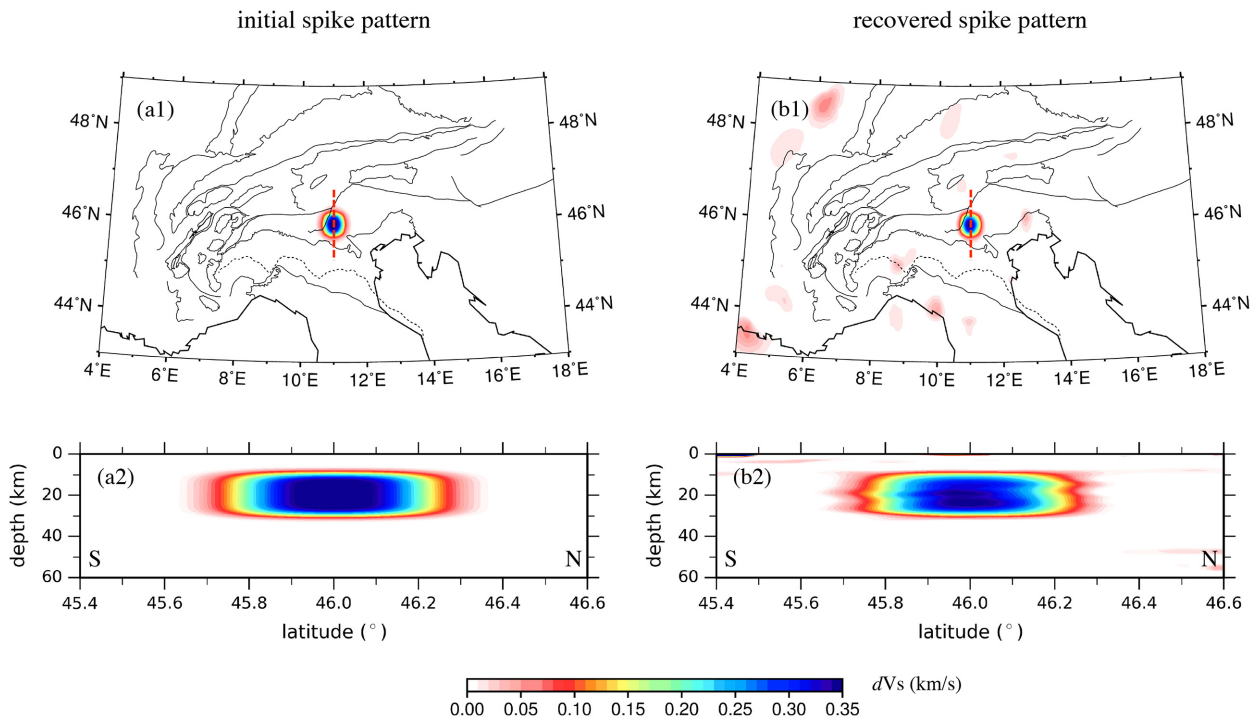


Figure 7. Resolution assessment of the final Vs model using a spike test. We display 20 km depth slices and cross-sections of the initial (a1 and a2) and recovered (b1 and b2) spike by removing the background model (m15). The red dashed line in (a1) and (b1) indicates the location of the cross-section.

the study region. The velocity perturbation is well recovered except for some slight horizontal smearings in the direction of Rayleigh-wave propagation. This result demonstrates that a structure of size ~ 50 km is well resolved at middle and lower crustal depths.

5 RESULTS

5.1 Depth slices in the shear wave velocity model

In Fig. 8, we show the comparison of shear wave velocity depth slices at 10, 30 and 50 km in the initial (m00) and final (m15) models. At 10 km depth (Figs 8a1 and a2), the strong low-velocity anomaly is related to the thick sedimentary basin of the Po basin (PB in Figs 8a1 and a2), while the high-velocity anomaly along the alpine belt is mainly due to the crystalline core of the mountain range. At this depth, the final model shows striking differences with the initial model. First, the final model displays velocity anomalies with more details. Some new features emerge in the final model, as for instance the Ivrea body (IB in Figs 8a1 and a2), a high-velocity, high-density body interpreted as a wedge of Adriatic upper-mantle roofing at 10 km depth in the crust of the western Alps (Closs & Labrousse 1963; Nicolas *et al.* 1990). Fig. 8(d) shows the Bouguer anomaly map of the study area. It documents a better match in shape and location between the 3.6 km s^{-1} isoline of model m15 (white dashed line in Figs 8a1 and a2) and the 100 mGal contour of the Ivrea Bouguer anomaly (white dashed line in Fig. 8d). Second, the final model shows much stronger velocity contrasts. For example, the low-velocity anomaly beneath the Po basin (PB) is 10 per cent lower in the final model than in the initial model. At 30 km depth, the most striking features are low-velocity anomalies related to the deep crustal roots of the Alps (AL in Figs 8b1 and b2) and the Apennines (AP in Figs 8b1 and b2). As in the 10 km depth slice, the final model shows higher resolution and stronger velocity contrasts with respect to the initial model. Besides, we observe an important shift of mean V_s values towards higher values from 3.91 to 3.99 km s^{-1} . This difference in average V_s is related to the shift in the mean of traveltimes misfits observed at long periods for model m00 that does not exist anymore in model m15 (see Section 4.2 and Fig. 5). Moreover, the final model shows a change in the shape of the low-velocity anomalies due to the Alpine and Apenninic crustal roots ($dV_s/V_s < 8$ per cent), which are narrower than in the initial model. At 50 km, the final model shows similar patterns as the initial model, and the strongest change is the shift in mean velocity from 4.41 to 4.47 km s^{-1} .

5.2 Moho depth map

The depth maps of isovelocity surfaces $V_s = 4.2 \text{ km s}^{-1}$ and $V_s = 4.0 \text{ km s}^{-1}$ are displayed in Fig. 9. The depth map of $V_s = 4.2 \text{ km s}^{-1}$ shown in Fig. 9(a) may be considered as a Moho depth map since 4.2 km s^{-1} is the typical velocity limit between crustal and mantle material, while the depth map of the isovelocity surface $V_s = 4.0 \text{ km s}^{-1}$ corresponds to the upper boundary of either anomalously high-velocity lower crustal material, or anomalously low-velocity upper mantle.

In the depth map of $V_s = 4.2 \text{ km s}^{-1}$ (Fig. 9a), the colour scale emphasizes the transition from normal crustal thickness for Phanerozoic western Europe (≤ 35 km, red colour) to thickened crust (> 35 km, blue colour). The crustal roots of the mountain ranges are clearly visible beneath the Alpine arc, the northern Apennines in the south and the Dinarides in the southeast. The boundary

between the European Moho (lower plate) and the Adriatic Moho (upper plate) is marked by an abrupt depth change that roughly follows the western and northern boundaries of the Po basin. This boundary is shifted towards the west and the north on the depth map of isovelocity 4.0 km s^{-1} (Fig. 9b) as a consequence of the eastward to southward dip of the subduction of Europe beneath Adria. The transition to the very shallow Moho of the stretched crust of the Ligurian sea is also clear. At smaller scale, Fig. 9(a) shows a clear Moho step of ~ 8 km that is almost linear in shape and closely follows the northwestern boundary of the external crystalline massifs of the western Alps (yellow dashed line in Fig. 9a). This sudden Moho depth change in the northwestern Alps contrasts with the more gradual increase in crustal thickness from the external to the internal parts of the central and eastern Alps. The rectilinear character of the Moho depth change, which contrasts with the curved shape of western Alpine structures such as the Penninic front, and its SW–NE orientation suggests that the Moho step may be controlled by a Variscan lithospheric discontinuity. A second remarkable small-scale feature of the Moho map is the thinner crust beneath the Lepontine dome of the central Alps, which is thinner than in any other parts along the strike of the belt. This feature was already visible in the Moho maps of Lu *et al.* (2018), but with less detail, while it was absent from the previous Moho depth maps of the Alps (Molinari & Morelli 2011; Spada *et al.* 2013). In the northern Apennines, the northern boundary of the thick crust closely follows the two eastern lobes of the Apenninic front (thin black dashed line in Fig. 9a), but not the westernmost lobe. This difference could be related to the gap observed from teleseismic tomography in-between the Alpine (or European) and Apenninic subduction slabs beneath the westernmost Po basin and the Ligurian Alps (Zhao *et al.* 2016).

The depth map of $V_s = 4.0 \text{ km s}^{-1}$ shown in Fig. 9(b) displays the Ivrea body, a wedge of serpentized Adriatic upper mantle at less than 20 km depth beneath the border between the Po basin and the western Alps (white dashed line in Fig. 9b). Its location well corresponds to the Bouguer anomaly high of Fig. 8(d), and its along-strike changes in depth even mimic the along-strike changes in the amplitude of the Bouguer anomaly. More unexpectedly, Fig. 9(b) displays two small-size highs of the isovelocity surface 4.0 km s^{-1} that are located along the northern boundary of the Adriatic plate in the vicinity of the Giudicarie line (yellow dashed lines in Fig. 9b). The two anomalies have a size of ~ 50 km, close to the lower resolution limit of our WET. Such a small size may explain why they have never been imaged in previous tomographies. The anomalies coincide with the two main provinces of Permian magmatism in northern Italy: the Val Trompia–Val Caffaro province for the western one, and the Athesian Volcanic Group for the eastern one (M. Malusá personal communication 2019). This spatial coincidence may suggest that the anomalies are due to gabbros underplated in the lowermost crust.

6 DISCUSSION

Here, we discuss some essential technical aspects which might affect the robustness of our tomography. We first test the influence of the initial model in Section 6.1. As it uses rather long-period surface waves (> 10 s), our tomography is less affected by cycle-skipping problems as other wave-equation-based tomographies. Therefore, Section 6.1 focuses on the impact of the initial model on the final model. We check the validity of our choice of misfit function in Section 6.2. Since we define the misfit function by the ‘overall’ phase shift between observed and simulated Rayleigh waves, we assume

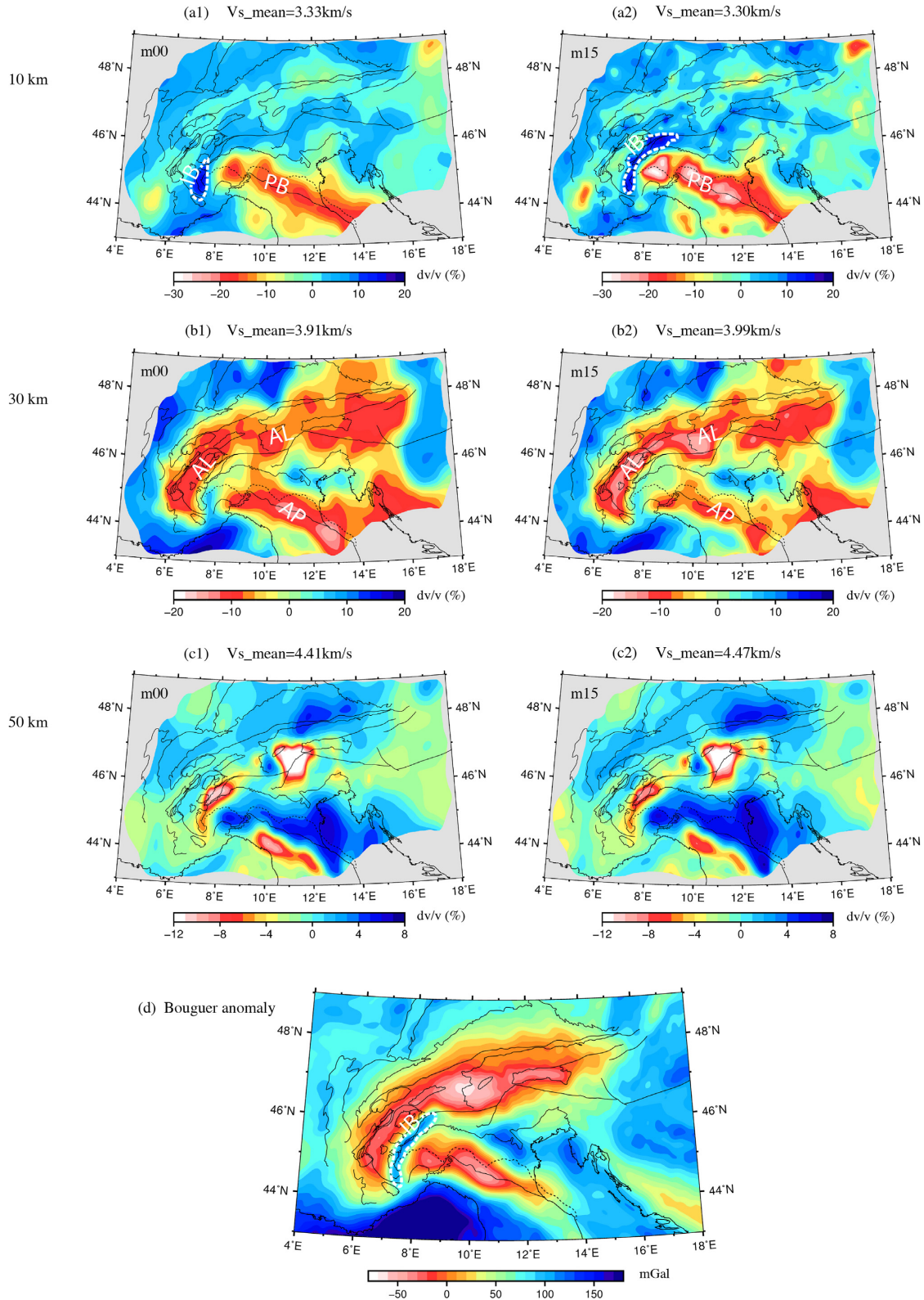


Figure 8. Comparison of Vs depth slices at 10, 30 and 50 km in the initial model (m00, left-hand column) and final model (m15, right-hand column), and Bouguer anomaly map (WGM2012; Bonvalot *et al.* 2012). The relative variations of velocity to the mean velocity of each slice (written above the maps) are displayed to emphasize the changes in the shape and amplitude of anomalies. The white dashed lines in (a1) and (a2) are the 3.6 km s^{-1} Vs contour, which should be compared with the 100 mGal contour of the Ivrea Body gravity anomaly shown by the white dashed line in (d). AL: Alps; AP: Apennines; IB: Ivrea Body; PB: Po basin.

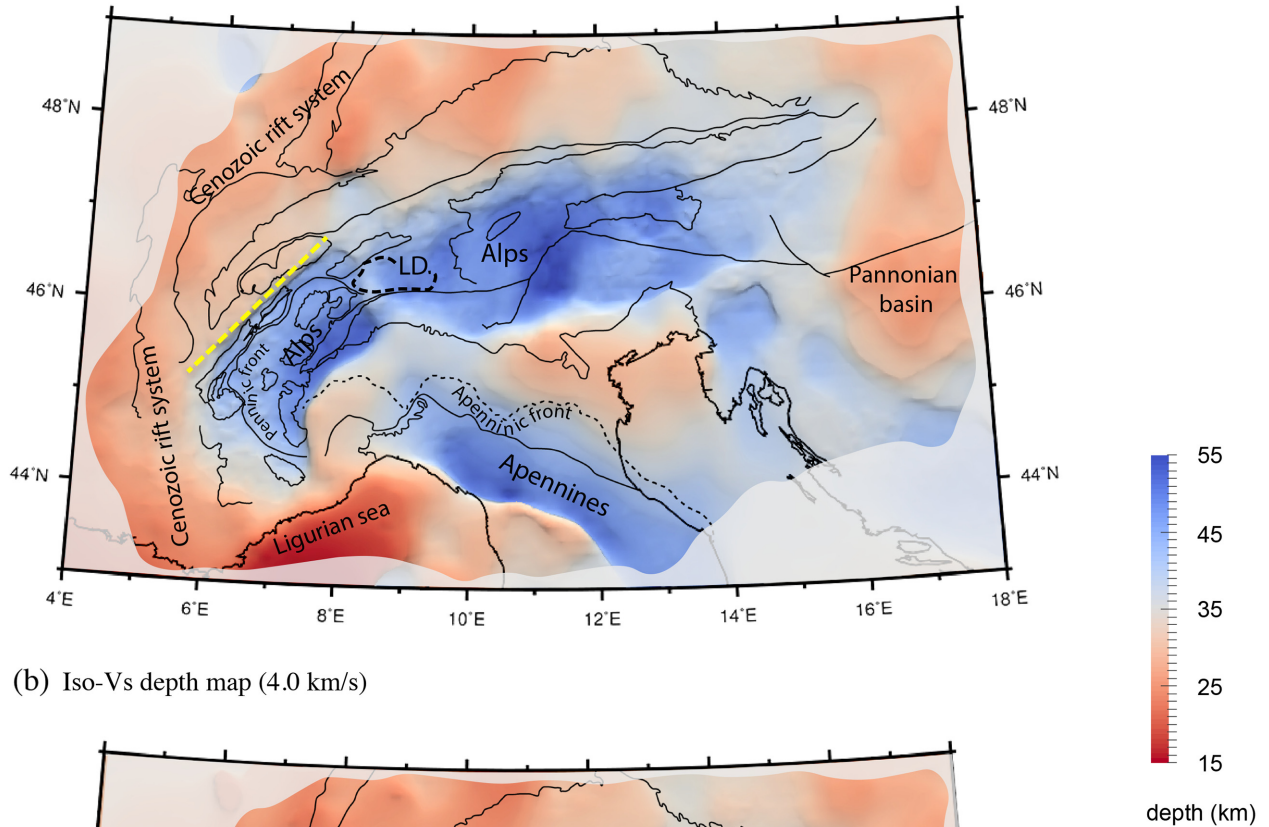
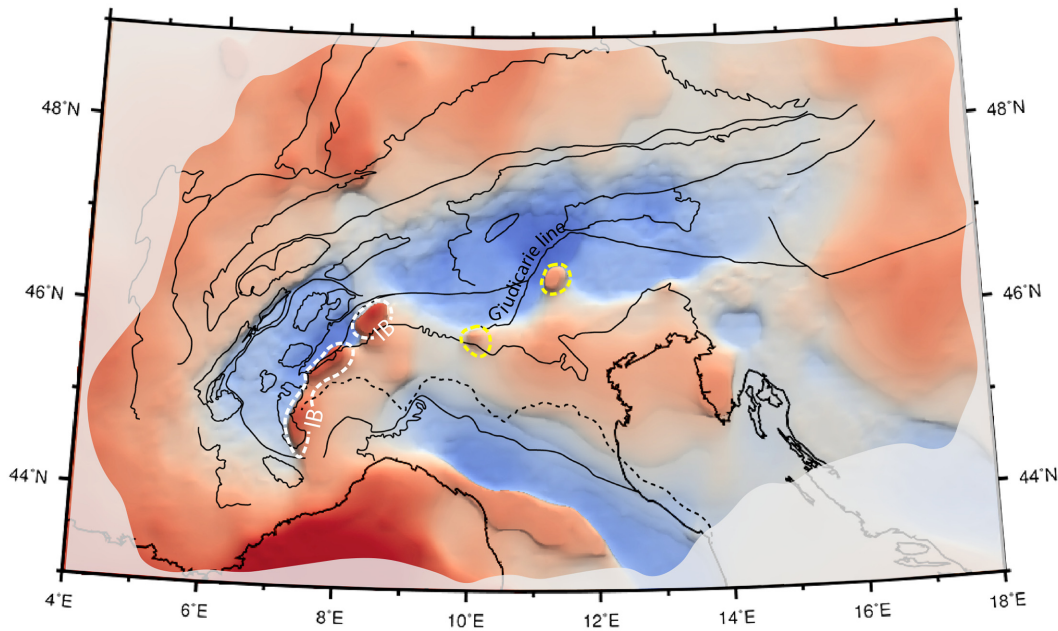
(a) Iso- V_s depth map (4.2 km/s)(b) Iso- V_s depth map (4.0 km/s)

Figure 9. Moho depth maps estimated from the depth of the isovelocity surfaces $V_s = 4.2$ and 4.0 km s^{-1} . In (a), the yellow dashed straight line outlines the Moho step along the outer boundary of the external crystalline massifs of the northwestern Alps. The thick black dashed line shows the Lepontine Dome (LD). In (b), the white dashed line outlines the shallow Moho ($\leq 20 \text{ km}$) on top of the Ivrea body (IB). The yellow dashed line indicates the location of small-scale bodies of anomalously high shear wave velocity ($V_s > 4.0 \text{ km s}^{-1}$) at lower crustal depths. Suporging Information Fig. S3 displays uncertainties of the isovelocity surface $V_s = 4.20 \text{ km s}^{-1}$. Suporging Information Fig. S4 shows the comparison of isovelocity surfaces $V_s = 4.00 \text{ km s}^{-1}$ for the initial (m00) and final (m15) models.

that we deal with a single event (wave train). This assumption is valid only in a medium devoid of strong scattering. In Section 6.3, we test the validity of the assumption of isotropy using the azimuthal distribution of misfits in the final model.

6.1 Influence of the initial model

Fig. 10 displays the initial and final models of the WET for a smoothed initial model m00_smooth, which is built by applying a Gaussian filter of length $\sim 100 \text{ km}$ to our previous initial model

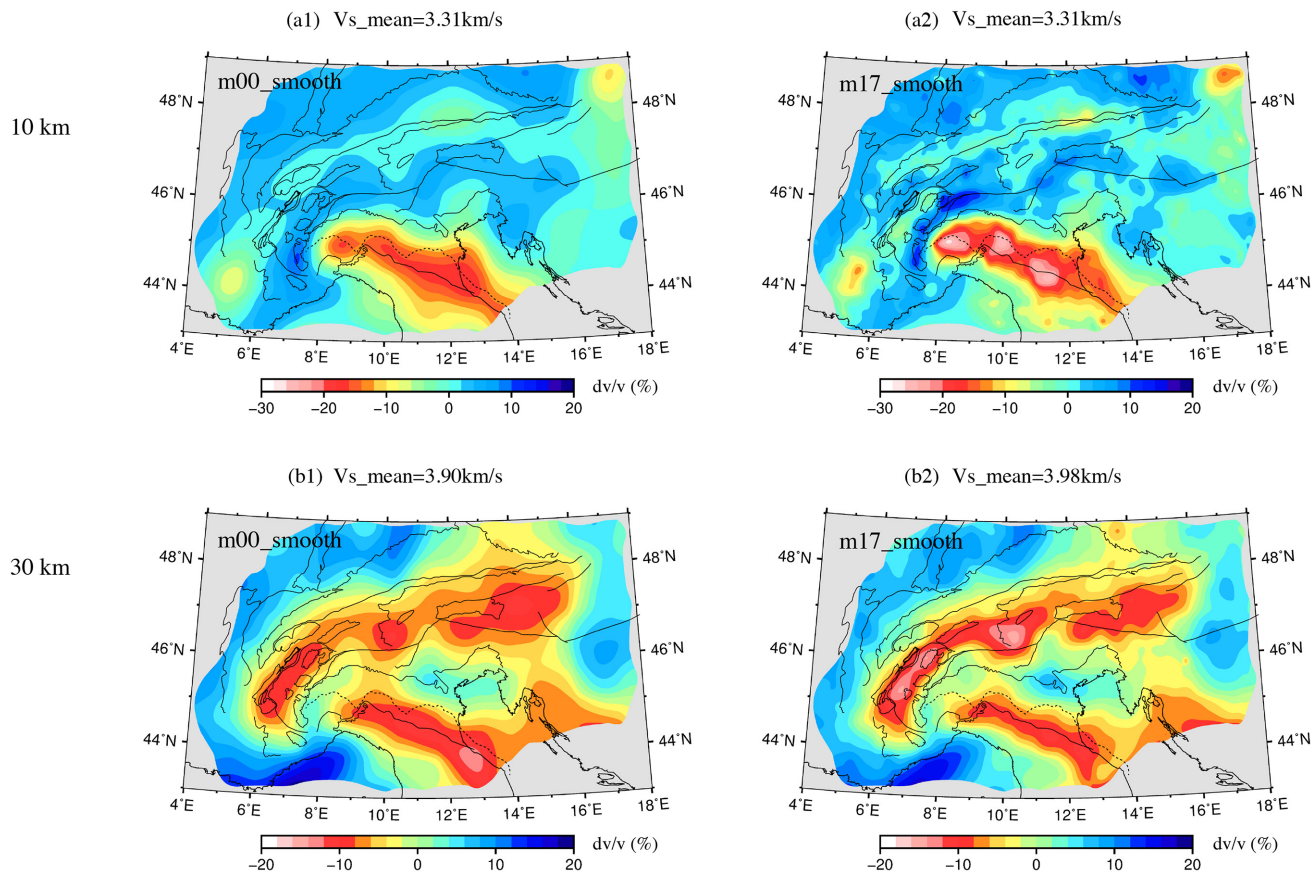


Figure 10. Vs depth slices at 10 and 30 km in the smooth initial model (m00_smooth; a1 and b1) and the corresponding final model (m17_smooth; a2 and b2). As in Fig. 8, we display the relative variations of velocity with respect to the average velocity in the depth slice, and we discard areas with poor data coverage.

m00. The final model m17_smooth is obtained after 17 iterations with ~ 66 per cent of total misfit reduction compared to the initial model m00_smooth. In the 10 and 30 km depth slices, the final model m17_smooth has similar values of mean Vs to m15 and it displays similar main features as discussed in Figs 8(a2) and (b2) and Section 5.2.

However, the remaining total misfit in the final model m17_smooth is ~ 13.5 per cent higher than in m15. A comparison of the depth slices in m17_smooth (Figs 10a2 and b2) and m15 (Figs 8a2 and b2) shows that the two final models still differ in a few small-scale structures. For instance, at 30 km depth, the low-velocity body corresponding to the crustal roots of the Apennines has a much more continuous shape in model m17_smooth than in m15. This implies the need for an accurate initial model for resolving small-scale structures, while large-scale structures can be recovered even where starting from a smoothed model.

6.2 Assumption of single event for the misfit function

It is well known that the heterogeneity of the lithosphere leads to complex propagation of Rayleigh waves due to scattering (e.g. Bungum & Capon 1974; Levshin & Berteussen 1979). We measure phase shift assuming that there is a single wave arrival in the time window of our measurements. Our phase shift could be biased if two wave arrivals are present in our measurement window, for instance because of multipathing or scattering (Bozdağ *et al.* 2011; Rickers *et al.* 2012). In Fig. 11, we show snapshots of the simulated

vertical motions at the surface for source station CH.WOLEN. In the 10–20 s period band, we observe a slightly scattered Rayleigh wave packet (black dashed frames in Fig. 11a2), which is generated by the complex crustal structure at the intersection of the western Alps, the northern Apennines and the Ligurian sea (black dashed frames in Fig. 11a1). As a consequence, the phase shift measurements for station pairs with interstation path crossing this region would be slightly biased by scattered waves. In the longer period band 20–50 s, there is no clear scattering since the wavelengths are mostly larger than the sizes of medium heterogeneities. We conclude that the effect of scattering is limited in our study. However, this effect should be properly taken into account when using Rayleigh waves at periods shorter than 10 s, when the wavelengths are comparable to or smaller than the sizes of medium heterogeneities.

6.3 Assumption of isotropic medium

Fig. 12 displays the azimuthal distribution of misfit for the initial (m00) and final (m15) models at 10, 15, 25 and 50 s. For each period, we observe obvious azimuthal variations of misfits with similar shapes for the two models, while the main difference is an overall correction towards zero for the final model in particular at long periods (25, 50 s).

The remaining azimuthal variations of misfits in the final model indicate the importance of azimuthal effects, which mainly concerns two aspects: (i) the potential azimuthal anisotropy in the crust

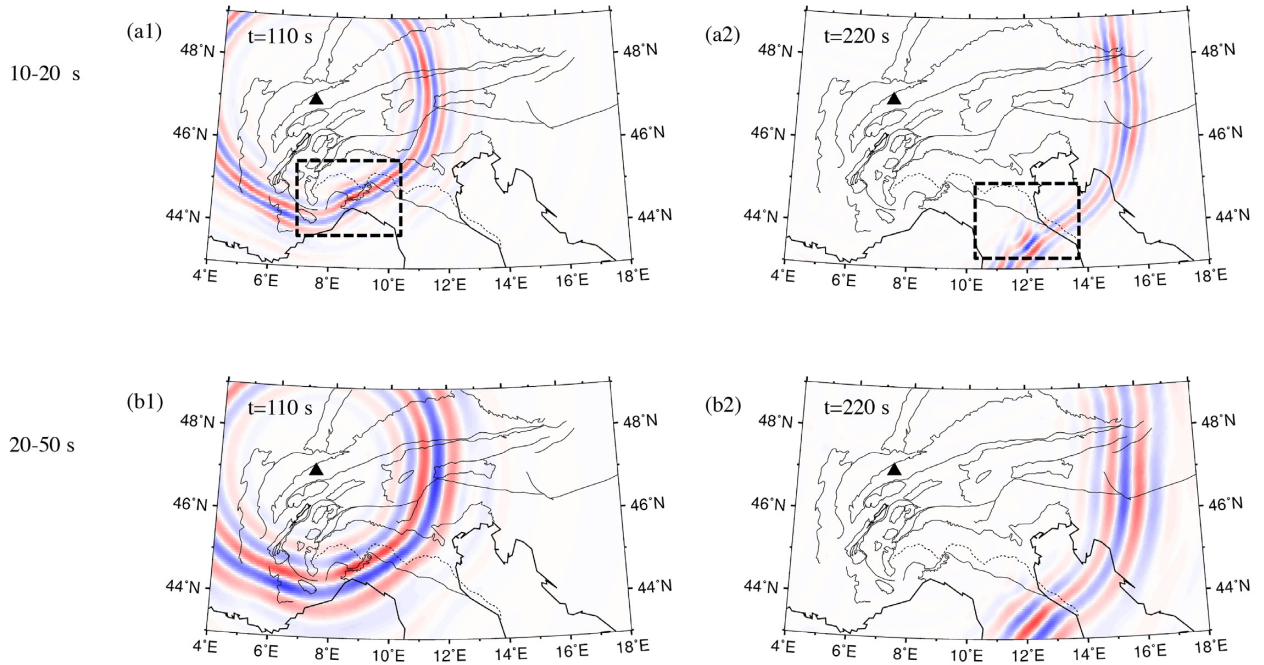


Figure 11. Snapshots of simulated wavefields at the surface in the period bands 10–20 s and 20–50 s at times 110 and 220 s. The black triangle indicates the source station CH.WOLEN. The two black dashed frames in (a1) and (a2) highlight scattered Rayleigh wave packets in the period band 10–20 s.

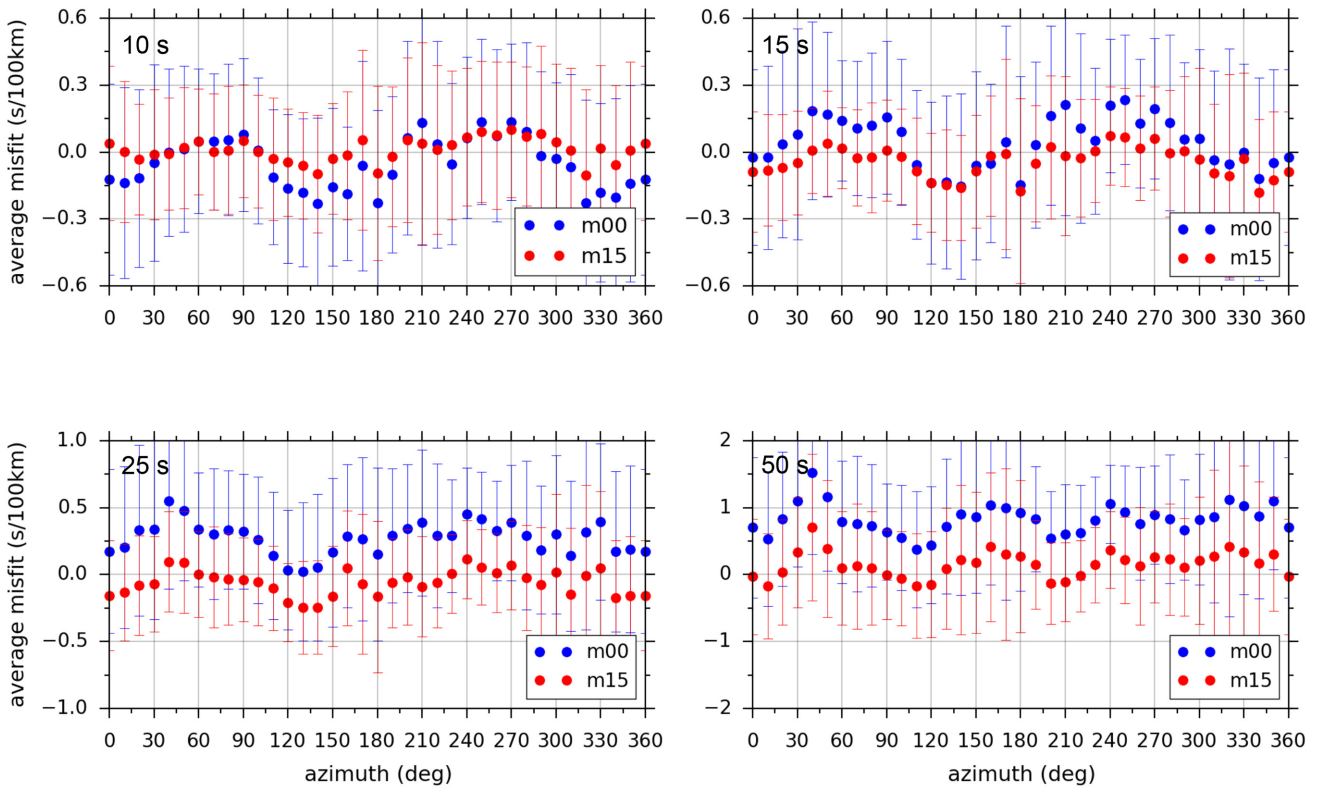


Figure 12. Azimuthal distributions of misfit at 10, 15, 25 and 50 s. Blue and red dots display the average misfits over azimuthal bins of 10° for the initial (m00) and final models (m15), respectively. Error bar refers to the one standard deviation of misfits in each azimuthal bin.

that is ignored in our inversion (Fry *et al.* 2010); (ii) the bias in phase velocity estimates from noise cross-correlation due to the uneven distribution of noise sources (Yao & van der Hilst 2009; Froment *et al.* 2010). The first aspect can be addressed by including

anisotropy in the inversion procedure. The impact of the second aspect can be reduced by using higher-order noise cross-correlations, such as correlations of the coda of noise cross-correlations, the so-called C3 (Stehly *et al.* 2008). The coda of noise correlations

provides a diffuse wavefield, so one could have a better phase velocity estimate than that obtained from the original directional wavefield.

There is possible influence of the shape of the Alps (or of lateral heterogeneities), which might be misinterpreted as azimuthal variation. But the influence should be limited for the following two reasons: (i) The azimuthal variations due to lateral heterogeneities should show different patterns in the initial and final models since the lateral heterogeneities are not equally resolved in different directions by the inversion. On the contrary, the azimuthal variation due to anisotropy and uneven noise source effects are unlikely to be modified by the inversion; (ii) The average misfit plot includes the whole data set and not only the Alps, or the E–W elongated part of the central and eastern Alps. So it is more influenced by general features of the whole region, rather than by the E–W trend of part of the region.

7 CONCLUSIONS

In this study, we performed ambient noise WET to build a truly 3-D high-resolution shear wave velocity model for the Alpine crust and uppermost mantle. The ambient noise WET overcomes the assumptions of high-frequency and lateral-homogeneity in the two-step inversion approach employed by traditional ambient noise tomography. With this method, we iteratively improved the initial model *LSP_Eucrust* 1.0 by minimizing the phase differences between observed and synthetic waveforms in the period range 10–50 s. The observed waveforms are estimated from noise correlations, while the synthetic waveforms are calculated numerically with the SEM. The gradient of the misfit function is computed with the adjoint-state approach, and the model is updated using the quasi-Newton l-LBFGS optimization method.

We obtained the final model after 15 iterations with ~65 per cent total misfit reduction compared to the initial model. At longer periods (25 and 50 s), the inversion corrected a strong positive shift of the histogram of misfits in the initial model in particular at periods ≥ 25 s. The histograms of misfits for the final model are symmetric around zero with more concentrated distributions. We validated the inversion mainly in the following two aspects: (i) we used a validation data set to avoid overinterpreting noise in the data; (ii) by analysing the spatial distributions of misfits for the initial and final models, we confirmed that the reduced misfits are primarily meaningful structure-induced misfits, rather than data errors.

At intracrustal and Moho depths, the final model shows higher resolution and significantly stronger velocity contrasts than the initial model. It reveals new features and precise features that existed in the initial model. The isovelocity Moho map derived from the final model confirms the Moho step in the northwest Alps and a local Moho high in the central Alps. It also reveals small-size high-velocity anomalies in the lower crust of northern Italy that may be interpreted as gabbroic plutons of Permian age underplated in the lower crust of the Adriatic plate.

The resulting model is up-to-now the highest resolution shear wave velocity model of the Alpine crust that has a great potential for deciphering a 3-D geological model of the crust of the study area. The ambient noise WET used in this study is an important complementary tool to refine the Vs model resulting from traditional ambient noise tomography. Further improvements include the consideration of anisotropy in the tomography scheme. Our final model can now be used as the initial model in earthquake-based FWI to better constrain the *P*-wave velocity. The resulting

Vs model and Moho depth map will be distributed on the authors' website <https://sites.google.com/view/seismology-yanglu>.

8 ORIGIN of DATA

Waveform data used in this paper belong to the permanent networks with codes BW (Department of Earth and Environmental Sciences, Geophysical Observatory, University of Munchen 2001), CH (Swiss Seismological Service at ETH Zürich 1983), FR (RESIF 1995), G (Institut De Physique Du Globe De Paris, & Ecole Et Observatoire Des Sciences De La Terre De Strasbourg 1982), GE (GEOFON Data Centre 1993), GR, GU (University of Genova 1967), HU (Kövesligethy Radó Seismological Observatory 1992), IV (INGV Seismological Data Centre, 2006), MN (MedNet Project Partner Institutions 1990), MT, NI [OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) and University of Trieste 2002], OE, RD, SI, SK (ESI SAS 2004), SL (Slovenian Environment Agency 2001) and ST (Geological Survey-Provincia Autonoma di Trento 1981). We also used data of temporary experiments, namely, AlpArray [network code Z3 (2015–2020), Hetényi *et al.* (2017)], CIFALPS [network code YP (2012–2013), Zhao *et al.* (2016)] and PYROPE [network code X7 (2010–2014), Chevrot *et al.* (2017)].

ACKNOWLEDGEMENTS

We warmly thank Ludovic Métivier, Weiguang He and Irnaka Theodosius Marwan from the SEISCOPE group for valuable discussions on FWI. The modelling and inversion code SEM46 have been developed in the frame of the SEISCOPE consortium (<http://seiscope2.osug.fr>), and we would like to thank the associated industrial sponsors. We appreciate the constructive feedback on our inversion methodology from Andreas Fichtner. We also acknowledge Marco Giovanni Malusà, Stéphane Guillot, Stéphane Schwartz and Thierry Dumont for their help with the preliminary interpretation of our results and the geological map. This study was granted access to the HPC resources of the Froggy platform of the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>), which is supported by the Rhône-Alpes region (GRANT CPER07_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56) and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the programme 'Investissements d'Avenir' supervised by the Agence Nationale pour la Recherche. We gratefully thank the operators of European permanent seismic networks who make their data available through EIDA (<http://www.orfeus-eu.org/eida>). The Z3 network is operated by the AlpArray Seismic Network Team: György Hetényi, Rafael Abreu, Ivo Allegretti, Maria-Theresia Apoloner, Coralie Aubert, Maxime Bes De Berc, Götz Bokelmann, Didier Brunel, Marco Capello, Martina Čarman, Adriano Cavaliere, Jérôme Chèze, Claudio Chiarabba, John Clinton, Glenn Cougoulat, Wayne Crawford, Luigia Cristiano, Tibor Czifra, Ezio D'alema, Stefania Danesi, Romuald Daniel, Iva Dasović, Anne Deschamps, Jean-Xavier Dessa, Cécile Doubre, Sven Egdorf, ETHZ-SED Electronics Lab, Tomislav Fiket, Kasper Fischer, Wolfgang Friederich, Florian Fuchs, Sigward Funke, Domenico Giardini, Aladino Govoni, Zoltán Gráczer, Gidera Gröschl, Stefan Heimers, Ben Heit, Davorka Herak, Marijan Herak, Johann Huber, Dejan Jarić, Petr Jedlička, Yan Jia, Hélène Jund, Edi Kissling, Stefan Klingen, Bernhard Klotz, Petr Kolínský, Michael Korn, Josef Kotek, Lothar Kühne, Krešo Kuk, Jürgen Loos, Deny Malengros, Lucia Margheriti, Christophe Maron, Xavier Martin, Marco Massa, Francesco Mazzarini, Thomas

Meier, Laurent Métral, Irene Molinari, Milena Moretti, Helena Žlebčíková, Anna Nardi, Jurij Pahor, Anne Paul, Catherine Péquegnat, Damiano Pesaresi, Davide Piccinini, Claudia Piro-mallo, Thomas Plenefisch, Jaroslava Plomerová, Silvia Pondrelli, Snježan Prevolnik, Roman Racine, Marc Régnier, Miriam Reiss, Joachim Ritter, Georg Rümpker, Simone Salimbeni, Detlef Schulte-kortnack, Werner Scherer, Sven Schippkus, Vesna Šipka, Daniele Spallarossa, Kathrin Spieker, Josip Stipčević, Angelo Strollo, Bálint Süle, Gyöngyvér Szanyi, Eszter Szűcs, Christine Thomas, Frederik Tilmann, Stefan Ueding, Massimiliano Vallocchia, Luděk Vecsey, René Voigt, Joachim Wassermann, Zoltán Wéber, Christian Weidle, Viktor Westergom, Gauthier Weyland, Stefan Wiemer, David Wolyniec, Thomas Zieke and Mladen Živčić. This work is funded by Agence Nationale de la Recherche (contract ANR-15-CE31-0015) and Labex OSUG@2020 (Investissement d'Avenir, ANR-10-LABX-56).

REFERENCES

- Beller, S., Monteiller, V., Operto, S., Nolet, G., Paul, A. & Zhao, L., 2018. Lithospheric architecture of the South-Western Alps revealed by multi-parameter teleseismic full-waveform inversion, *Geophys. J. Int.*, **212**(2), 1369–1388.
- Bensen, G.D., Ritzwoller, M.H. & Yang, Y., 2009. A 3-D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise, *Geophys. J. Int.*, **177**(3), 1177–1196.
- Bodin, T., Sambridge, M., Rawlinson, N. & Arroucau, P., 2012a. Transdimensional tomography with unknown data noise, *Geophys. J. Int.*, **189**(3), 1536–1556.
- Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K. & Rawlinson, N., 2012b. Transdimensional inversion of receiver functions and surface wave dispersion, *J. geophys. Res.*, **117**(B2), B02301.
- Bonvalot, S. *et al.*, 2012. *World Gravity Map*, Commission for the Geological Map World. Eds. BGI-CGMW-CNES-IRD, Paris.
- Bozdağ, E., Trampert, J. & Tromp, J., 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.*, **185**, 845–870.
- Brenguier, F., Shapiro, N.M., Campillo, M., Nercessian, A. & Ferrazzini, V., 2007. 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlations, *Geophys. Res. Lett.*, **34**(2), L02305.
- Brocher, T.M., 2005. Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. seism. Soc. Am.*, **95**(6), 2081–2092.
- Bungum, H. & Capon, J., 1974. Coda pattern and multipath propagation of Rayleigh waves at NORSAR, *Phys. Earth planet. Inter.*, **9**(2), 111–127.
- Chen, M., Huang, H., Yao, H., van der Hilst, R. & Niu, F., 2014. Low wave speed zones in the crust beneath SE Tibet revealed by ambient noise adjoint tomography, *Geophys. Res. Lett.*, **41**(2), 334–340.
- Chevrot, S., Sylvander, M. & RESIF, 2017. Seismic network X7: PYROPE PYrenean Observational Portable Experiment (RESIF-SISMOB), RESIF - Réseau Sismologique et géodésique Français, doi:10.15778/resif.x72010.
- Closs, H. & Labrouste, Y., 1963. Recherches séismologiques dans les Alpes occidentales au moyen de grandes explosions en 1956, 1958 et 1960, *Mémoire Collectif, Année Géophysique Internationale, C.N.R.S., Sér.III, Fasc.2*, 236.
- Department of Earth and Environmental Sciences, Geophysical Observatory, University of Munchen, 2001. BayernNetz, doi:10.7914/SN/BW .
- Derode, A., Larose, E., Campillo, M. & Fink, M., 2003a. How to estimate the Green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves, *Appl. Phys. Lett.*, **83**(15), 3054–3056.
- Derode, A., Larose, E., Tanter, M., de Rosny, J., Tourin, A., Campillo, M. & Fink, M., 2003b. Recovering the Green's function from field-field correlations in an open scattering medium (L), *J. acoust. Soc. Am.*, **113**(6), 2973–2976.
- Ekström, G., Abers, G.A. & Webb, S.C., 2009. Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation, *Geophys. Res. Lett.*, **36**(18), L18301.
- ESI SAS, 2004. *National Network of Seismic Stations of Slovakia*, doi:10.14470/fx099882 .
- Fichtner, A., 2014. Source and processing effects on noise correlations, *Geophys. J. Int.*, **197**(3), 1527–1531.
- Fichtner, A. & Villasenör, A., 2015. Crust and upper mantle of the western Mediterranean—constraints from full-waveform inversion, *Earth Planet. Sci. Lett.*, **428**, 52–62.
- Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2008. Theoretical background for continental- and global-scale full-waveform inversion in the time-frequency domain, *Geophys. J. Int.*, **175**(2), 665–685.
- Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2009. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.*, **179**(2), 1703–1725.
- Froment, B., Campillo, M., Roux, P., Gouédard, P., Verdel, A. & Weaver, R.L., 2010. Estimation of the effect of nonisotropically distributed energy on the apparent arrival time in correlations, *Geophysics*, **75**(5), SA85–SA93.
- Fry, B., Deschamps, F., Kissling, E., Stehly, L. & Giardini, D., 2010. Layered azimuthal anisotropy of Rayleigh wave phase velocities in the European Alpine lithosphere inferred from ambient noise, *Earth planet. Sci. Lett.*, **297**(1–2), 95–102.
- Gao, H. & Shen, Y., 2014. Upper mantle structure of the Cascades from full-wave ambient noise tomography: Evidence for 3D mantle upwelling in the back-arc, *Earth planet. Sci. Lett.*, **390**, 222–233.
- GEOFON Data Centre, 1993. GEOFON Seismic Network. Deutsches Geoforschungszentrum GFZ, doi:10.14470/tr560404 .
- Geological Survey-Provincia Autonoma di Trento, 1981. Trentino Seismic Network, doi:10.7914/SN/ST .
- Graves, R.W., 1996. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bull. seism. Soc. Am.*, **86**(4), 1091–1106.
- Handy, M.R., M. Schmid, S., Bousquet, R., Kissling, E. & Bernoulli, D., 2010. Reconciling plate-tectonic reconstructions of Alpine Tethys with the geological-geophysical record of spreading and subduction in the Alps, *Earth-Sci. Rev.*, **102**(3–4), 121–158.
- Haned, A., Stutzmann, E., Schimmel, M., Kiselev, S., Davaile, A. & Yelles-Chaouche, A., 2016. Global tomography using seismic hum, *Geophys. J. Int.*, **204**(2), 1222–1236.
- Hetényi, G., Molinari, I., Clinton, J. & Bokelmann, G., 2017. The AlpArray seismic network: a large-scale European experiment to image the Alpine orogen, *Surv. Geophys.*, **39**(5), 1009–1033.
- INGV Seismological Data Centre, 2006. *Rete Sismica Nazionale (RSN)*, Istituto Nazionale di Geofisica e Vulcanologia (INGV).
- Institut De Physique Du Globe De Paris, Ecole Et Observatoire Des Sciences De La Terre De Strasbourg, 1982. GEOSCOPE, French Global Network of broad band seismic stations, doi:10.18715/geoscope.g .
- Kästle, E.D., Soomro, R., Weemstra, C., Boschi, L. & Meier, T., 2016. Two-receiver measurements of phase velocity: cross-validation of ambient-noise and earthquake-based observations, *Geophys. J. Int.*, **207**(3), 1493–1512.
- Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Int.*, **139**(3), 806–822.
- Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.*, **149**(2), 390–412.
- Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.*, **150**(1), 303–318.
- Komatitsch, D. & Vilotte, J.P., 1998. The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. seism. Soc. Am.*, **88**(2), 368–392.
- Kövesligethy Radó Seismological Observatory, 1992. Hungarian National Seismological Network, doi:10.14470/uh028726 .

- Levshin, A. & Berteussen, K.A., 1979. Anomalous propagation of surface waves in the Barents Sea as inferred from NORSAR recordings, *Geophys. J. Int.*, **56**(1), 97–118.
- Li, H., Su, W., Wang, C.-Y. & Huang, Z., 2009. Ambient noise Rayleigh wave tomography in western Sichuan and eastern Tibet, *Earth planet. Sci. Lett.*, **282**(1–4), 201–211.
- Lin, F.-C. & Ritzwoller, M.H., 2011. Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure, *Geophys. J. Int.*, **186**(3), 1104–1120.
- Lin, F.-C., Ritzwoller, M.H. & Snieder, R., 2009. Eikonal tomography: surface wave tomography by phase front tracking across a regional broadband seismic array, *Geophys. J. Int.*, **177**(3), 1091–1110.
- Liu, Q. & Gu, Y., 2012. Seismic imaging: from classical to adjoint tomography, *Tectonophysics*, **566–567**, 31–66.
- Liu, Q. & Tromp, J., 2006. Finite-frequency kernels based on adjoint methods, *Bull. seism. Soc. Am.*, **96**(6), 2383–2397.
- Liu, Y., Niu, F., Chen, M. & Yang, W., 2017. 3-D crustal and uppermost mantle structure beneath NE China revealed by ambient noise adjoint tomography, *Earth planet. Sci. Lett.*, **461**, 20–29.
- Lobkis, O.I. & Weaver, R.L., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *J. acoust. Soc. Am.*, **110**(6), 3011–3017.
- Lu, Y., Stehly, L., Paul, A. & AlpArray Working Group, 2018. High-resolution surface wave tomography of the European crust and uppermost mantle from ambient seismic noise, *Geophys. J. Int.*, **214**(2), 1136–1150.
- Ludwig, W., Nafe, J. & Drake, C., 1970. Seismic refraction, in *The Sear*, Vol. 4, pp. 53–84, ed. Maxwell, A.E., Wiley-Interscience.
- Luo, Y. & Schuster, G.T., 1991. Wave equation travelttime inversion, *Geophysics*, **56**(5), 645–653.
- Luo, Y., Yang, Y., Xu, Y., Xu, H., Zhao, K. & Wang, K., 2015. On the limitations of interstation distances in ambient noise tomography, *Geophys. J. Int.*, **201**, 652–661.
- MedNet Project Partner Institutions, 1990. Mediterranean Very Broadband Seismographic Network (MedNet), doi:10.13127/sd/fbbtdtd6q .
- Métivier, L. & Brossier, R., 2016. The SEISCOPE optimization toolbox: a large-scale nonlinear optimization library based on reverse communication, *Geophysics*, **81**(2), F1–F15.
- Molinari, I. & Morelli, A., 2011. EPcrust: a reference crustal model for the European Plate, *Geophys. J. Int.*, **185**(1), 352–364.
- Molinari, I., Verbeke, J., Boschi, L., Kissling, E. & Morelli, A., 2015. Italian and Alpine three-dimensional crustal structure imaged by ambient-noise surface-wave dispersion, *Geochem. Geophys. Geosyst.*, **16**(12), 4405–4421.
- Mora, P., 1987. Nonlinear two dimensional elastic inversion of multioffset seismic data, *Geophysics*, **52**(9), 1211–1228.
- Nicolas, A., Hirn, A., Nicolich, R. & Polino, R., 1990. Lithospheric wedging in the western Alps inferred from the ECORS-CROP traverse, *Geology*, **18**(7), 587.
- Nicolson, H., Curtis, A., Baptie, B. & Galetti, E., 2012. Seismic interferometry and ambient noise tomography in the British Isles, *Proc. Geol. Assoc.*, **123**(1), 74–86.
- Nishida, K., Montagner, J.-P. & Kawakatsu, H., 2009. Global surface wave tomography using seismic hum, *Science*, **326**(5949), 112–112.
- Nocedal, J., 1980. Updating quasi-Newton matrices with limited storage, *Math. Comput.*, **35**(151), 773–782.
- OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) and University of Trieste, 2002. North-East Italy Broadband Network, doi:10.7914/SN/NI .
- Operto, S., Virieux, J., Amestoy, P., L'Excellent, J.-Y., Giraud, L. & Ali, H.B.H., 2007. 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: a feasibility study, *Geophysics*, **72**(5), SM195–SM211.
- Picozzi, M., Parolai, S., Bindi, D. & Strollo, A., 2009. Characterization of shallow geology by high-frequency seismic noise tomography, *Geophys. J. Int.*, **176**(1), 164–174.
- Plessix, R.-E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.*, **167**(2), 495–503.
- RESIF, 1995. RESIF-RLBP French Broad-band network, RESIF-RAP strong motion network and other seismic stations in metropolitan France. RESIF - Réseau Sismologique et géodésique Français, doi:10.15778/resif.fr .
- Rickers, F., Fichtner, A. & Trampert, J., 2012. Imaging mantle plumes with instantaneous phase measurements of diffracted waves, *Geophys. J. Int.*, **190**, 650–664.
- Sager, K., Ermert, L., Boehm, C. & Fichtner, A., 2018. Towards full waveform ambient noise inversion, *Geophys. J. Int.*, **212**(1), 566–590.
- Saygin, E. & Kennett, B.L.N., 2012. Crustal structure of Australia from ambient seismic noise tomography, *J. geophys. Res.*, **117**(B1), 15 303–15 318.
- Schmid, S. M., Fügenschuh, B., Kissling, E. & Schuster, R., 2004. Tectonic map and overall architecture of the Alpine orogen, *Eclogae Geol. Helv.*, **97**(1), 93–117.
- Shapiro, N. M., Campillo, M., Stehly, L. & Ritzwoller, M. H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science*, **307**(5715), 1615–1618.
- Slepian, D. & Pollak, H. O., 1961. Prolate spheroidal wave functions, Fourier analysis and uncertainty—I, *Bell Syst. Tech. J.*, **40**(1), 43–63.
- Slovenian Environment Agency, 2001. Seismic Network of the Republic of Slovenia, doi:10.7914/SN/SL .
- Snieder, R., 1986. 3D linearized scattering of surface waves and a formalism for surface wave holography, *Geophys. J. Int.*, **84**(3), 581–605.
- Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E*, **69**(4), 046610.
- Spada, M., Bianchi, I., Kissling, E., Agostinetti, N. P. & Wiemer, S., 2013. Combining controlled-source seismology and receiver function information to derive 3-D moho topography for Ttaly, *Geophys. J. Int.*, **194**(2), 1050–1068.
- Spetzler, J., Trampert, J. & Snieder, R., 2002. The effect of scattering in surface wave tomography, *Geophys. J. Int.*, **149**(3), 755–767.
- Stehly, A. & Boué, P., 2017. On the interpretation of the amplitude decay of noise correlations computed along a line of receivers, *Geophys. J. Int.*, **209**, 358–372.
- Stehly, L., Campillo, M., Froment, B. & Weaver, R. L., 2008. Reconstructing Green's function by correlation of the coda of the correlation (C3) of ambient seismic noise, *J. geophys. Res.*, **113**(B11), B11306.
- Stehly, L., Fry, B., Campillo, M., Shapiro, N. M., Guilbert, J., Boschi, L. & Giardini, D., 2009. Tomography of the Alpine region from observations of seismic ambient noise, *Geophys. J. Int.*, **178**(1), 338–350.
- Swiss Seismological Service ETH Zürich, 1983. National Seismic Networks of Switzerland, doi:10.7914/SN-CA .
- Tape, C., 2009. Seismic tomography of southern California using adjoint methods, *PhD thesis*, California Institute of Technology, Pasadena, CA.
- Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the southern California crust, *Science*, **325**(5943), 988–992.
- Thomson, D. J., 1982. Spectrum estimation and harmonic analysis, *Proc. IEEE*, **70**(9), 1055–1096.
- Trinh, P. T., Brossier, R., Métivier, L., Virieux, J. & Wellington, P., 2017. Bessel smoothing filter for spectral element mesh, *Geophys. J. Int.*, **209**(3), 1489–1512.
- Trinh, P. T., Brossier, R., Métivier, L., Tavad, L. & Virieux, J., 2019. Efficient time-domain 3D elastic and viscoelastic full-waveform inversion using a spectral-element method on flexible Cartesian-based mesh, *Geophysics*, **84**(1), R61–R83.
- Tromp, J., Tape, C. & Liu, Q., 2004. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**(1), 195–216.
- Tromp, J., Luo, Y., Hanasoge, S. & Peter, D., 2010. Noise cross-correlation sensitivity kernels, *Geophys. J. Int.*, **183**(2), 791–819.
- University of Genova, 1967. Regional Seismic Network of North Western Italy, doi:10.7914/SN/GU .
- Verbeke, J., Boschi, L., Stehly, L., Kissling, E. & Michelini, A., 2012. High-resolution Rayleigh-wave velocity maps of central Europe from a dense ambient-noise data set, *Geophys. J. Int.*, **188**(3), 1173–1187.

- Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics*, **74**(6), WCC1–WCC26.
- Wang, K., Yang, Y., Basini, P., Tong, P., Tape, C. & Liu, Q., 2018. Refined crustal and uppermost mantle structure of southern California by ambient noise adjoint tomography, *Geophys. J. Int.*, **215**, 844–863.
- Wapenaar, K., 2004. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, *Phys. Rev. Lett.*, **93**(25), 1–4.
- Weaver, R. L. & Lobkis, O. I., 2001. Ultrasonics without a source: thermal fluctuation correlations at MHz Frequencies, *Phys. Rev. Lett.*, **87**(13), 134301.
- Yang, Y., Ritzwoller, M. H., Levshin, A. L. & Shapiro, N. M., 2007. Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.*, **168**(1), 259–274.
- Yao, H. & van der Hilst, R. D., 2009. Analysis of ambient noise energy distribution and phase velocity bias in ambient noise tomography, with application to SE Tibet, *Geophys. J. Int.*, **179**(2), 1113–1132.
- Yomogida, K. & Aki, K., 1987. Amplitude and phase data inversions for phase velocity anomalies in the Pacific Ocean basin, *Geophys. J. Int.*, **88**(1), 161–204.
- Young, M. K., Rawlinson, N. & Bodin, T., 2013. Transdimensional inversion of ambient seismic noise for 3D shear velocity structure of the Tasmanian crust, *Geophysics*, **78**(3), WB49–WB62.
- Yuan, H., French, S., Cupillard, P. & Romanowicz, B., 2014. Lithospheric expression of geological units in central and eastern North America from full waveform tomography, *Earth planet. Sci. Lett.*, **402**, 176–186.
- Zhao, L. *et al.*, 2016. Continuity of the Alpine slab unraveled by high-resolution *P* wave tomography, *J. geophys. Res.*, **121**(12), 8720–8737.
- Zhou, Y., Dahlen, F. A. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.*, **158**(1), 142–168.
- Zhu, H., Bozdağ, E., Peter, D. & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography, *Nat. Geosci.*, **5**(7), 493–498.

SUPPORTING INFORMATION

Supplementary data are available at [GJI](#) online.

Figure S1. Evolution of correlation coefficient with respect to SNR. The correlation coefficients are calculated for all station pairs between the observed and synthetic waveforms computed for the initial model m00.

Figure S2. Comparison of the observed waveforms (orange line) with synthetics computed from the initial (grey line) and final models (orange line). We show the comparison in three period bands for four station pairs WOLEN-A206A, WOLEN-ASQU, WOLEN-BOJS and WOLEN-MPLH. The waveforms are windowed to keep only Rayleigh waves and normalized by their maximum absolute amplitudes. Comparison of frequency-dependent phase traveltimes misfits for the initial (grey dots) and final models (orange dots) is attached below the waveform comparison.

Figure S3. Uncertainties of the isovelocity surface $V_s = 4.2 \text{ km s}^{-1}$ shown in Fig. 9(a). The uncertainties are evaluated by the vertical distance between 4.17 and 4.23 km s^{-1} isovelocity surfaces.

Figure S4. 3-D views of the shaded Moho depth maps estimated from the depth of the isovelocity surface $V_s = 4.0 \text{ km s}^{-1}$ for the initial (a, m00) and final models (b, m15). In (a) and (b), the black arrow refers to the Moho step along the external side of the north-western Alps (see Section 5.2). In (b), the orange arrow indicates the location of anomalously high shear wave velocity ($V_s > 4.0 \text{ km s}^{-1}$) bodies, which are interpreted as gabbros underplated in the lowermost crust of the Adriatic plate (see Section 5.2). IB: Ivrea body; LD: Lepontine Dome.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.