

Surface-wave tomography using SeisLib: a Python package for multiscale seismic imaging

Fabrizio Magrini¹, Sebastian Lauro², Emanuel Kästle³ and Lapo Boschi^{4,5,6}

¹*Institute of Geosciences, Johannes Gutenberg University, 55099 Mainz, Germany. E-mail: fmagrini@uni-mainz.de*

²*Department of Mathematics and Physics, Università degli Studi Roma, 00146 Roma RM, Italy*

³*Institute for Geological Sciences, Freie Universität, 12249 Berlin, Germany*

⁴*Dipartimento di Geoscienze, Università degli Studi di Padova, 35131 Padova PD, Italy*

⁵*CNRS, INSU, Institut des Sciences de la Terre de Paris, Sorbonne Université, 75252 Paris, France*

⁶*Istituto Nazionale di Geofisica e Vulcanologia, 40100 Bologna BO, Italy*

Accepted 2022 June 20. Received 2022 June 16; in original form 2022 February 19

SUMMARY

To improve our understanding of the Earth's interior, seismologists often have to deal with enormous amounts of data, requiring automatic tools for their analyses. It is the purpose of this study to present SeisLib, an open-source Python package for multiscale seismic imaging. At present, SeisLib includes routines for carrying out surface-wave tomography tasks based on seismic ambient noise and teleseismic earthquakes. We illustrate here these functionalities, both from the theoretical and algorithmic point of view and by application of our library to seismic data from North America. We first show how SeisLib retrieves surface-wave phase velocities from the ambient noise recorded at pairs of receivers, based on the zero crossings of their normalized cross-spectrum. We then present our implementation of the two-station method, to measure phase velocities from pairs of receivers approximately lying on the same great-circle path as the epicentre of distant earthquakes. We apply these methods to calculate dispersion curves across the conterminous United States, using continuous seismograms from the transportable component of USArray and earthquake recordings from the permanent networks. Overall, we measure 144 272 ambient-noise and 2055 earthquake-based dispersion curves, that we invert for Rayleigh-wave phase-velocity maps. To map the lateral variations in surface-wave velocity, SeisLib exploits a least-squares inversion algorithm based on ray theory. Our implementation supports both equal-area and adaptive parametrizations, with the latter allowing for a finer resolution in the areas characterized by high density of measurements. In the broad period range 4–100 s, the retrieved velocity maps of North America are highly correlated (on average, 96 per cent) and present very small average differences (0.14 ± 0.1 per cent) with those reported in the literature. This points to the robustness of our algorithms. We also produce a global phase-velocity map at the period of 40 s, combining our dispersion measurements with those collected at global scale in previous studies. This allows us to demonstrate the reliability and optimized computational speed of SeisLib, even in presence of very large seismic inverse problems and strong variability in the data coverage. The last part of the manuscript deals with the attenuation of Rayleigh waves, which can be estimated through SeisLib based on the seismic ambient noise recorded at dense arrays of receivers. We apply our algorithm to produce an attenuation map of the United States at the period of 4 s, which we find consistent with the relevant literature.

Key words: Inverse theory; Seismic tomography; Surface waves and free oscillations.

1 INTRODUCTION

The wealth of publicly available seismic data has been steadily increasing for the last few decades, and with it, our ability to

image the Earth's interior at increasingly high resolution. Every year, hundreds of seismometers are installed or redeployed across our planet, with the precise purpose of better resolving features of the Earth that can only be probed indirectly. As a result,

seismologists have to deal with an amount of data that could be considered overwhelming, if it were not for the possibility of using automated tools in their analysis.

In recent years, Python has established itself as one of the most popular programming languages among scientific researchers (e.g., Millman & Aivazis 2011). Besides being open source and characterized by a relatively shallow learning curve (e.g., Nagar 2017), its popularity can be ascribed to its very active community, as testified by the wide variety of packages that continuously become available for installation.

In the field of seismology, the rapid increase in the popularity of Python can probably be traced back to the publication of *ObsPy* (Beyreuther *et al.* 2010), which has provided its users with an all-round framework for seismic data analysis. Since then, and especially in the last few years, several Python packages have been designed to tackle more or less specific tasks in seismic imaging. For example, *rf* (Eulendorf 2020) allows for calculating receiver functions. *MSNoise* (Lecocq *et al.* 2014) and *NoisePy* (Jiang & Denolle 2020) provide routines for monitoring and extracting group velocities of surface waves generated by seismic ambient noise; *PyKonal* (White *et al.* 2020) and *scikit-fmm* (<https://github.com/scikit-fmm>) implement the fast-marching method (Sethian 1996) to track wavefronts propagating in heterogeneous media; and this list could be further extended.

It is the purpose of this study to present *SeisLib*, an open source Python package for multiscale seismic imaging (<https://pypi.org/project/seislib/>). *SeisLib* is the result of a long-term effort of our team to gather, optimize and make open source some of the Python codes behind our seismological publications over the last few years. As opposed to most of the seismological Python packages available to date, *SeisLib* has been conceived to encompass a wide variety of algorithms to obtain seismic images of the Earth's interior.

At the time of writing, our library is in rapid expansion, and includes routines to (i) download and process continuous seismic-ambient-noise recordings to extract phase velocities of Rayleigh and Love waves (e.g., Boschi *et al.* 2013; Kästle *et al.* 2016) as well as Rayleigh-wave attenuation as a function of frequency (e.g., Boschi *et al.* 2019; Magrini & Boschi 2021; Magrini *et al.* 2021); (ii) download and process surface-wave recordings generated by teleseismic earthquakes, to retrieve Rayleigh and Love phase velocities based on pairs of receivers lying on the same great-circle path as the epicentre (e.g., Meier *et al.* 2004; Magrini *et al.* 2020a) and (iii) least-squares imaging of lateral variations in surface-wave velocity, based on both equal-area and adaptive parametrizations (e.g., Boschi & Dziewonski 1999; Schaefer *et al.* 2011). A diagram of the main modules included in *SeisLib* is shown in Fig. 1.

The manuscript is structured so as to illustrate the background theory and our algorithmic implementation inherent to each of the above modules. Sections 2 and 3 deal with the calculation of surface-wave dispersion curves from seismic ambient noise and teleseismic earthquakes, respectively. In this paper, examples are restricted to Rayleigh waves, but Love-wave velocities as well can be measured via *SeisLib* routines. Section 4 concerns the mapping of spatial variations in surface-wave velocity from the local to the global scale. Our strategy to retrieve Rayleigh-wave attenuation is presented in Section 5. At the end of each of these sections, an application of our library to a real-world seismological task is presented. This allows us to both illustrate the potential of *SeisLib* in seismic imaging and cross-validate our results against previous studies.

Upon reading this paper, some readers wishing to make practical use of *SeisLib* might want to learn more about technical aspects of the package. They should be aware that the software comes with

a detailed documentation (<https://seislib.readthedocs.io/en/latest/>) that will hopefully answer most of their questions.

2 PHASE VELOCITY FROM SEISMIC AMBIENT NOISE

2.1 Empirical Green's function

As shown in several theoretical studies, the cross correlation of seismic ambient noise can be related to the surface-wave Green's function between the two points of observation (for a review see, e.g., Campillo *et al.* 2014; Boschi & Weemstra 2015). In case of a *diffuse* ambient wavefield recorded at two receivers on the vertical component, the empirical Rayleigh-wave Green's function has a relatively simple expression, that can be used to constrain the velocity structure of the Earth's interior. In the frequency domain, this is proportional to a zeroth order Bessel function of the first kind (J_0), and reads

$$\Re\{\rho(\mathbf{x}_A, \mathbf{x}_B, \omega)\} \approx J_0\left(\frac{\omega\Delta}{c}\right) e^{-\alpha\Delta} \quad (1)$$

(e.g., Nakahara 2012; Magrini & Boschi 2021), where Δ , ω and c denote interstation distance, angular frequency and phase velocity, respectively, ρ the statistical expectation of the normalized cross-spectrum associated with the two ambient-noise recordings, and $\Re\{\dots\}$ maps a complex number into its real part. Consistent with Magrini & Boschi (2021), the exponential damping term in eq. (1) accounts for the (possibly frequency-dependent) attenuation of the Rayleigh waves propagating between the two receivers $\mathbf{x}_A$ and $\mathbf{x}_B$, through the coefficient α .

In ambient-noise seismology, where continuous seismograms of relatively long duration (months or years) are used, the statistical expectation of ρ is replaced by an ensemble average of the cross-spectra calculated over a relatively large number of time windows. This contributes to approximating the condition of a diffuse ambient wavefield (e.g., Boschi & Weemstra 2015), allowing the use of the above equation to measure the (average) interstation phase velocity. In practice, since phase velocity is related to the *phase* of the empirical Green's function, but not to its amplitude, the exponential damping term in eq. (1) is often neglected (e.g., Kästle *et al.* 2018; Magrini *et al.* 2020b), simplifying the problem of retrieving c from the data. This approach resulted in numerous successful applications of velocity imaging and monitoring, and can nowadays be considered standard in ambient-noise tomography (e.g., Nakata *et al.* 2019, and references therein).

2.2 Velocity ambiguity

To retrieve fundamental-mode Rayleigh-wave phase velocities, *SeisLib* implements an automated algorithm. Given a pair of receivers, it first subdivides their simultaneous recordings into possibly overlapping time windows (e.g., Seats *et al.* 2012). Cross-correlation is then performed in the frequency domain for each time window, and the resulting cross-correlations are normalized by spectral whitening (e.g., Bensen *et al.* 2007); finally, ensemble averaging the normalized cross-correlations ensures the retrieval of a unique, robust cross-spectrum, corresponding to ρ in the left-hand side (LHS) of eq. (1) (Fig. 2).

To better constrain the phase of the cross-spectrum, *SeisLib* pre-processes it following a three-steps procedure (Sadeghisorkhani

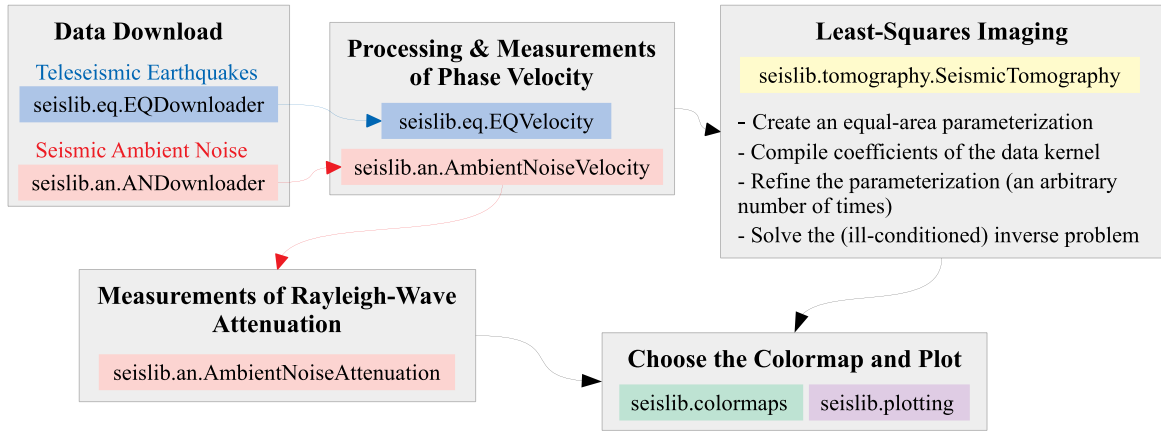


Figure 1. Workflow diagram showing all the main features of SeisLib. The modules *seislib.an* and *seislib.eq* contain Python classes to download, process and extract dispersion curves from recordings of seismic ambient noise (Section 2) and teleseismic earthquakes (Section 3), respectively. The retrieved dispersion curves can then be inverted for spatial variations in surface-wave phase velocity at different periods (Section 4), through the *SeismicTomography* class of the *seislib.tomography* module. The *seislib.an* module also allows for measuring Rayleigh-wave attenuation based on dense seismic arrays (Section 5). Spatial variations in velocity and attenuation can be visualized using the routines included in *seislib.plotting*. SeisLib also comes packed with the scientific colourmaps of Fabio Cramer (Cramer 2018), which can be accessed through *seislib.colormaps*.

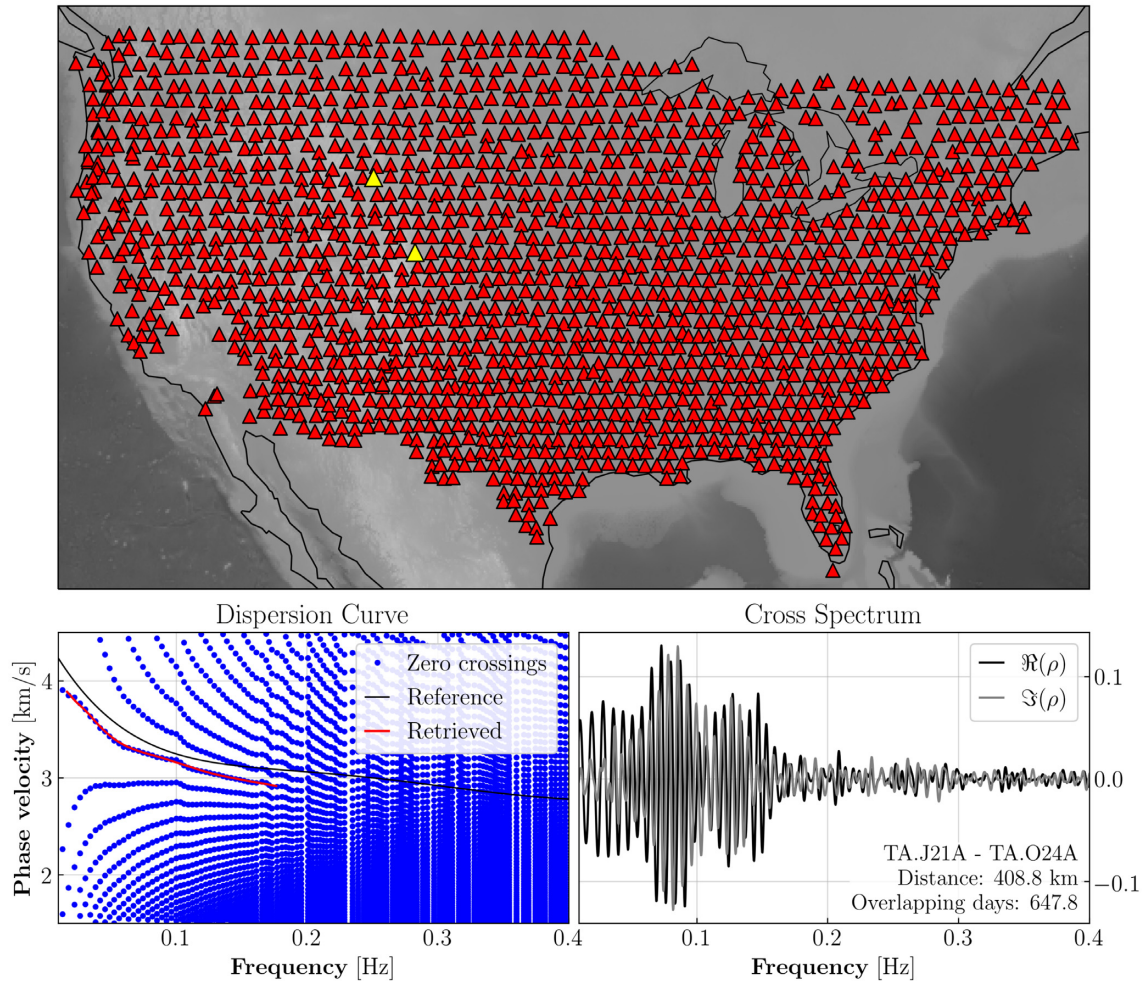


Figure 2. Top panel: location of the receivers (triangles) of the transportable component of USArray used in this study. The yellow triangles refer to the two receivers used to obtain the dispersion curve shown in the bottom-left-hand panel (red line). The dispersion curve was retrieved from the zero crossings (blue dots) of the real part of the cross-spectrum (black line in the bottom-right-hand panel) of their continuous seismograms, as explained in Section 2.3. For illustration purposes, the imaginary part of the cross-spectrum is shown in grey. Station codes, interstation distance, and amount simultaneous recordings (in days) available for the two receivers are reported in the bottom right-hand panel.

et al. 2018; Magrini & Boschi 2021). (i) It inverse-Fourier transforms the cross-spectrum; (ii) it zero-pads the resulting signal (in the time-domain) at times corresponding to interstation phase velocities that are much slower or much faster than those typical of a Rayleigh-wave fundamental mode in the microseism frequency band (i.e. $\sim 1.5\text{--}4.5\text{ km s}^{-1}$) and (iii) the padded signal is finally brought back into the frequency domain via forward Fourier transform. The above operations have the effect of both smoothing the spectrum and filtering out most of the arrivals that are not associated with fundamental-mode Rayleigh waves, overall benefiting the subsequent analysis.

Once the cross-spectrum has been computed, its real part is taken and possible values of phase velocity can be found at discrete angular frequencies ω_n , associated with the zero crossings of $\Re\{\rho\}$ (Fig. 2). In fact, if we consider the discrete set of values z_m ($m = 1, 2, \dots, M$) such that $J_0(z_m) = 0$, found by estimating a realistic minimum value for $c(\omega)$, then

$$c(\omega_n) = \frac{\omega_n \Delta}{z_m} \quad (2)$$

(e.g., Ekström *et al.* 2009; Boschi *et al.* 2013) identifies all the possible (frequency-dependent) phase velocities for which the zero crossings of $\Re\{\rho\}$ coincide with those of J_0 . Eq. (2) states that, at a given angular frequency ω_n , M values of c are in principle possible, resulting in a velocity ambiguity that needs to be overcome algorithmically. Our implementation of such algorithm relies on the works of Kästle *et al.* (2016, 2021), and is briefly summarized below.

2.3 Resolution of the ambiguity

Kästle *et al.* (2016) noted that the number of possible phase velocities can be conveniently halved by accounting for only those zero crossings where the derivatives of J_0 and $\Re\{\rho\}$ have the same sign. The problem can be further simplified introducing a reference dispersion curve, that allows for eliminating all values of c that are too far away from the reference. This is especially helpful at relatively low frequencies, where the velocity ambiguity in question is relatively small and consists of only few options largely spaced apart from each other (Fig. 3b). Our algorithm starts picking the dispersion curve exactly at such relatively low frequencies, and extends the dispersion curve iterating over progressively larger ones.

As in Kästle *et al.* (2021), instead of picking the phase velocities at each zero crossing directly, we first create a heatmap from cosine kernels around each crossing (see Fig. 3). The approach is similar to a 2-D kernel-density estimation (KDE) (Parzen 1962). In practice, for each zero crossing, we define an intensity that equals one at its origin (i.e. at the value of frequency and velocity corresponding to the zero crossing itself) and that decays with distance from it as a cosine, so as to be zero at the border of the kernel (Fig. 3d). This yields elliptical-shaped kernels around each zero crossing, that overlap with each other. The final heatmap simply results from the summation over all the overlapping kernels (Fig. 3b). Low-quality crossings (i.e. those that are too close to each other or where the amplitude of the adjacent maximum/minimum in the cross-spectrum is too low (Fig. 3a), are ignored in the heatmap creation.

As anticipated, the picking of the dispersion curve starts at relatively low frequencies (where the velocity ambiguity is not pronounced) and continues iteratively towards larger ones. Successive picks are taken where the intensity of the heatmap is maximized, and the research for such maxima is guided by the derivative with frequency of the reference curve together with the velocity value picked at the last iteration. Whenever the heatmap intensity is too

low, or the slope of the thus retrieved dispersion curve deviates too much from the reference curve, the iteration is skipped and no velocity value is taken. The algorithm stops when the frequency range spanned by the skipped iterations is larger than three times the frequency spacing between adjacent zero crossings.

Finally, it is worth mentioning that SeisLib also allows for retrieving dispersion curves of Love and radially-polarized Rayleigh waves, albeit not discussed in this study. These are calculated in a very similar fashion as above, with the only differences that (i) cross-correlations are performed on transverse- and radial-component recordings and (ii) the zero crossings of $J_0(z) - J_2(z)$ (with J_2 denoting the second order Bessel function of the first kind) replace those of J_0 in the determination of the ambiguous velocities via eq. (2) (Kästle *et al.* 2016).

2.4 Application to USArray data

To test the above implementation, we download vertical-component recordings from the transportable component of USArray, consisting of over 400 broad-band seismometers deployed in 1600 different locations across the conterminous United States (Fig. 2). We pre-process each seismogram so as to remove mean and linear trends, before tapering (5 per cent) and bandpass filtering between 0.01 and 0.5 Hz; the pre-processed seismograms were also deconvolved with the instrument response to get displacement, and downsampled to 1 Hz. (All the above operations are carried out automatically in SeisLib.) The thus acquired data set covers a period range between May 2004 and September 2019.

Similar to Ekström (2014), owing to the very large size of the data set, we only cross-correlate station pairs characterized by interstation distances of 600 km or less, and for which at least six months of simultaneous recordings are available. We carry out the cross-correlations over one-hour-long time windows, overlapping with each other by 50 per cent. Application of our algorithm resulted in, overall, 144 272 dispersion curves spanning surface-wave periods between 3 and 40 s. These have been selected so as to be characterized by relatively large ratios between inter-station distance and wavelength (>2) (e.g., Bensen *et al.* 2007).

Examples of dispersion curves, together with average, standard deviation and number of measurements as a function of period are illustrated in Fig. 4.

3 PHASE VELOCITY FROM TELESEISMIC EARTHQUAKES

3.1 Two-station method

To retrieve earthquake-based phase-velocity measurements, SeisLib implements a *two-station method* (e.g., Meier *et al.* 2004; Soomro *et al.* 2016; Magrini *et al.* 2020a). The two-station method builds on the assumption that a fundamental-mode surface wave that has travelled large distances ($\gtrsim 20^\circ$) from the epicentre of a strong earthquake (magnitudes $\gtrsim 5.5$) can be decomposed into a sum of monochromatics, or plane waves (e.g., Ekström *et al.* 1997). This assumption is particularly convenient if we consider two receivers approximately lying on the same great-circle path as the epicentre. As the wave front propagates from the first receiver to the second, its amplitude gets modified (by interstation attenuation and/or site effects) and its phase gets shifted according to interstation distance Δ and average velocity of propagation c . If attenuation is small, the phase shift can be expressed, in the frequency domain, as

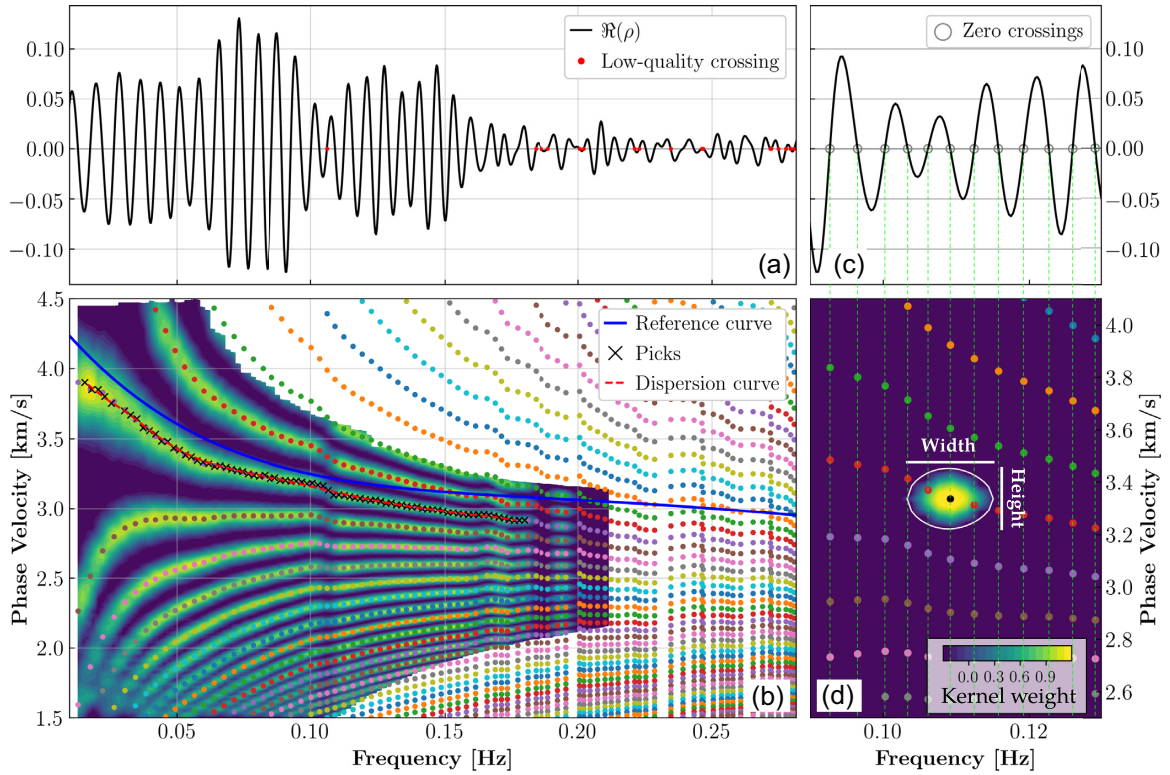


Figure 3. Illustration of SeisLib's algorithm (2.3) to extract dispersion curves from seismic ambient noise, applied to the receiver pair shown in Fig. 2. The real part of the cross-spectrum (a) is first used to retrieve all possible velocities associated with its zero crossings (dots in b, coloured to better distinguish the different branches). From the zero crossings (in the frequency–velocity plane), a heatmap is then created (background colour in b) using a KDE-like method, through cosine kernels centred on them. This is illustrated in (d) for a single zero crossing, while (c) highlights the relation between the zero crossings seen in (d) and the cross-correlation amplitude. The width and height of the kernels are tunable parameters in our algorithm, and are expressed in units relative to the frequency spacing between adjacent crossings and the velocity spacing between parallel branches, respectively. In this study, we set them to 3 and 1.2. As seen in (b), heatmaps are only computed in the most relevant area of the dispersion diagram, to optimize computational speed.

$\phi_2(\omega) - \phi_1(\omega) = \omega \Delta / c(\omega)$ (e.g., Magrini *et al.* 2020a), where the subscripts denote the receiver and ϕ the phase.

It is understood that the observed phase delay is invariant under 2π translations, giving rise to an ambiguity in phase velocity

$$c(\omega) = \frac{\omega \Delta}{\phi_2(\omega) - \phi_1(\omega) + 2n\pi}, \quad (3)$$

where n is integer. In other words, the delay in the arrival time of the wave front at the second receiver can be explained by different average interstation phase velocities (Fig. 5). And similar to Section 2.1, such ambiguity needs to be solved algorithmically.

Retrieving (ambiguous) dispersion measurements is relatively simple, and SeisLib's implementation follows closely the iterative procedure described, for example in Soomro *et al.* (2016) and Magrini *et al.* (2020a). Consider two seismograms recording a strong teleseismic earthquake, and assume the two receivers lie on the same great-circle path as the epicentre. At a given frequency, (i) we first narrow bandpass filter the two seismogram about this frequency; then (ii) taper them, in the time domain, about (here, ± 30 per cent) the arrival time predicted by a reference model; finally, (iii) cross correlate the resulting signals in the frequency domain to obtain the phase delay in eq. (3). Having computed the phase delay, (iv) we calculate $c(\omega)$ by simply replacing n with a finite set of integer numbers (in our code, between -10 and 10), and discard all the resulting values of c that fall outside of a plausible velocity range. The dispersion measurements in Fig. 5 (and throughout the rest of

the section) were obtained following the same procedure, where we discarded all velocities smaller than 2.5 or larger than 5 km s^{-1} .

The above analysis is carried out on vertical-component recordings to measure the dispersion of vertically polarized Rayleigh waves, and on horizontal-component recordings to measure that of Love and radially polarized Rayleigh waves (transverse and radial components, respectively). It has been pointed out that Love-wave observations made by the two-station method are likely to be less accurate than Rayleigh-wave ones, owing to overtone interference (Foster *et al.* 2014b; Hariharan *et al.* 2022).

3.2 The issue of quality selection

To obtain an interstation dispersion curve based on the two-station method, what can be considered as the 'standard' procedure consists of collecting, for a given pair of receivers, dispersion measurements from a relatively large number of teleseismic earthquakes (e.g., Meier *et al.* 2004; Darbyshire *et al.* 2004; Darbyshire & Lebedev 2009). For each earthquake, one dispersion curve is typically extracted based on arbitrary quality-selection criteria; these can involve, for example, the distance of the picked velocities from a reference model, the smoothness of the considered picks, or the length of the retrieved dispersion curve (Soomro *et al.* 2016). Such criteria allow for assessing whether or not the dispersion curve should be accepted, or what parts of it should be considered reliable, and are motivated by the not uncommon occurrence of noisy dispersion

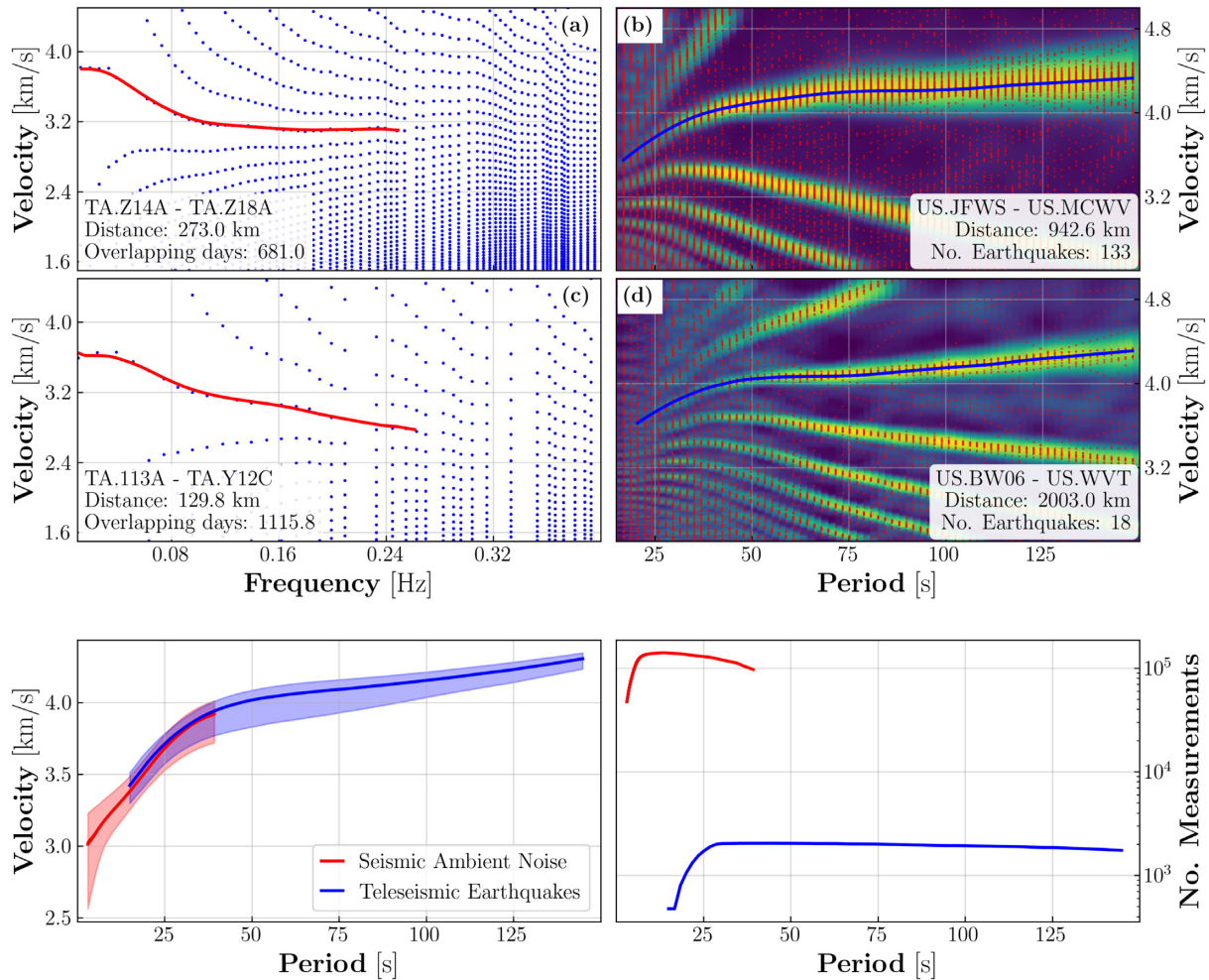


Figure 4. (a) and (c): zero crossings (blue dots) and dispersion curves (red lines) obtained from continuous recordings of seismic ambient noise at different pairs of receivers, as explained in Section 2. (b) and (d): ensemble of dispersion measurements (red dots) and dispersion curves obtained from earthquake recordings at two different station pairs, as explained in Section 3. The background colour indicates the density of dispersion measurements (P_d). The bottom panels report some statistics on the compiled data sets of dispersion curves. On the left, the median phase velocities (thick lines) and 10th and 90th percentiles (colour shades) measured in the conterminous United States from seismic ambient noise (red) and teleseismic earthquakes (blue). The number of measurements obtained for the two data sets as a function of period is reported on the right.

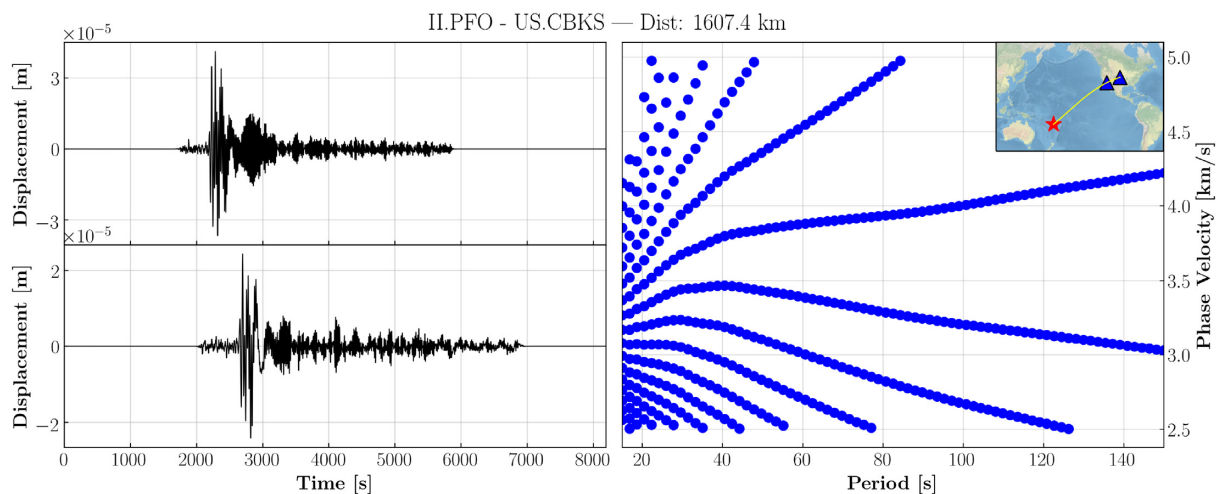


Figure 5. Left-hand panel: teleseismic earthquake (magnitude 6.5) recorded on the vertical component of a pair of receivers lying on the same great-circle path as the epicentre. Right-hand panel: ambiguous dispersion measurements obtained from the recordings. Station codes and interstation distance are reported in the title. The time associated with the recordings is expressed in seconds after the origin time of the earthquake.

measurements. Once all earthquakes have been processed, and a minimum number of dispersion curves has been retrieved for the station pair considered, their average over frequency is generally taken as the final, ‘robust’ interstation dispersion curve. In addition, those frequencies characterized by relatively large standard deviations are optionally discarded (e.g., Soomro *et al.* 2016).

As a result of the above quality selections, a cumbersome research for the optimal compromise between the quality of the retrieved dispersion curves and the number of station pairs rejected in the procedure is often demanded by the data (e.g., Magrini *et al.* 2020a). Not finding the optimal compromise would therefore result either in low-quality measurements or in a scarcity of them, with implications, in both circumstances, for the subsequent velocity imaging of the area of interest.

3.3 SeisLib’s approach

To overcome the above issue, we have designed an innovative procedure. Instead of extracting one dispersion curve for each event and then taking their average over frequency, we compute the final dispersion curve directly from all the available (ambiguous) dispersion measurements, using them as an ensemble. This is illustrated in Fig. 6, where the dispersion measurements collected from 51 teleseisms have been gathered together and used to measure phase velocity as a function of surface-wave period. The main ideas behind our approach can be enumerated as follows.

- (i) A reference model can be seen as the prior probability $\mathbf{P}_m$ that a wave-train propagates, at a given period (or frequency), with a certain speed (Fig. 6a).
- (ii) The dispersion measurements collected from several events are likely to cluster into different branches in the period-velocity plane; these branches are expected to be more or less coherent depending on the noisiness of the individual measurements, and one of those will correspond to the average, true dispersion curve (Fig. 6b).
- (iii) Similar to our Section 2.3, the density of dispersion measurements can be mapped via a KDE-like approach; in turn, the resulting map can be interpreted as the probability of observation $\mathbf{P}_d$ (Fig. 6b).
- (iv) Such mapping allows for reducing the intrinsic noise of the individual dispersion measurements *without* discarding any measurement through arbitrary quality-selection criteria. In principle, this favours reproducibility.
- (v) Multiplying the probability of observation by the prior probability should contribute to identifying the correct branch, by down-weighting those regions in the period-velocity plane that are less likely to describe the wave-train propagation according to the reference model (Fig. 6c).

In our codes, we define $\mathbf{P}_m$ as a $m \times n$ matrix, where m and n denote the number of discrete velocities and periods that finely sample the period-velocity plane (Fig. 6a). Each j th column of $\mathbf{P}_m$ is initialized to a Gaussian function evaluated at each i th velocity v (associated with the rows of $\mathbf{P}_m$) and centred onto the reference velocity μ (associated with the j th period), that is $\frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(v_i-\mu)^2}{2\sigma^2})$. (The possibly period-dependent standard deviation σ should be chosen *a priori*.) The column is then normalized so that its maximum is equal to one.

Note that, since each column of $\mathbf{P}_m$ is normalized by its maximum (rather than its sum), referring to the j th column of $\mathbf{P}_m$ as the prior probability that a wave train propagates at speed μ is technically incorrect. In the following, however, we will stick to such

terminology, as the normalization criterion should affect neither the effectiveness of our method nor its probabilistic interpretation.

The probability of observation $\mathbf{P}_d$ is defined in a very similar fashion to the above, with the only difference that its entries are initialized to the density of dispersion measurements; this is calculated, as in Section 2.3, through a KDE approach. Each column of $\mathbf{P}_d$ is then normalized by its maximum (Fig. 6b).

Finally, we define the weighted probability of observation $\mathbf{P}_w = \mathbf{P}_m \odot \mathbf{P}_d$ ($\odot$ denotes element-wise multiplication), which incorporates the information associated with both the data and the reference model (Fig. 6c). This is used to pick the inter-station dispersion curve, as explained in the following.

3.4 Retrieval of the dispersion curve

The algorithm for picking the dispersion curve, based on the values of $\mathbf{P}_w$, can be described as a two-steps procedure. In both steps, the key aspect considered by the algorithm (at a given period) is the probability that the dispersion curve passes through a given velocity. Such probability, as a function of velocity, typically resembles a series of peaks and valleys (Fig. 6), where only one maximum corresponds to the true phase velocity.

The first step begins with the identification of the correct dispersion branch in the phase-velocity ambiguity. This is done iterating over the j th index of $\mathbf{P}_w$, from longer periods (where the ambiguity is less pronounced) to shorter periods. Whenever the probability profile associated with the j th period presents a global maximum that (i) exceeds a minimum value $P_{\min} = 0.6$ and (ii) is at least 100 per cent higher than all the other maxima, the corresponding phase velocity is included in the dispersion curve (Fig. 6d). This stage ensures that only the most robust velocity branch in $\mathbf{P}_w$ is taken, and terminates when for three consecutive iterations the above conditions are not verified (Fig. 6e).

The second step involves picking the dispersion curve at shorter periods. The logic is similar to the above, with only one main difference. At the j th period, the phase velocity of the following (shorter) period is predicted by the slope of the reference curve. The predicted value c_{pred} is then used to further ‘filter’ the probability profile corresponding to the $(j+1)$ th column of $\mathbf{P}_w$, multiplying it by a skewed Gaussian whose maximum is centred onto c_{pred} (Figs. 6f and e). The use of a skewed Gaussian (rather than a Gaussian) has the precise purpose of downweighting all probability peaks associated with faster velocities than c_{pred} . This is justified by the assumption that phase velocity decreases (approximately) linearly with decreasing period, which is especially reasonable in the period range investigated by a two-station method (~ 15 – 150 s). In the resulting (‘filtered’) probability profile, the phase velocity associated with the global maximum is included in the dispersion curve only if the peak in question (i) exceeds a minimum value of $P_{\min}$ chosen by the user (we found optimal values in the range 0.15–0.35) and (ii) is at least 30 per cent higher than all the other peaks. If, at any point of this second step, a value of phase velocity is rejected, the iterations terminate.

3.5 Application to permanent stations from North America

To verify the effectiveness of our approach, we download vertical-component recordings of strong (magnitudes between 6 and 8.5) teleseismic earthquakes from 90 permanent stations deployed in North America (Fig. 7). We use all publicly-available seismic data

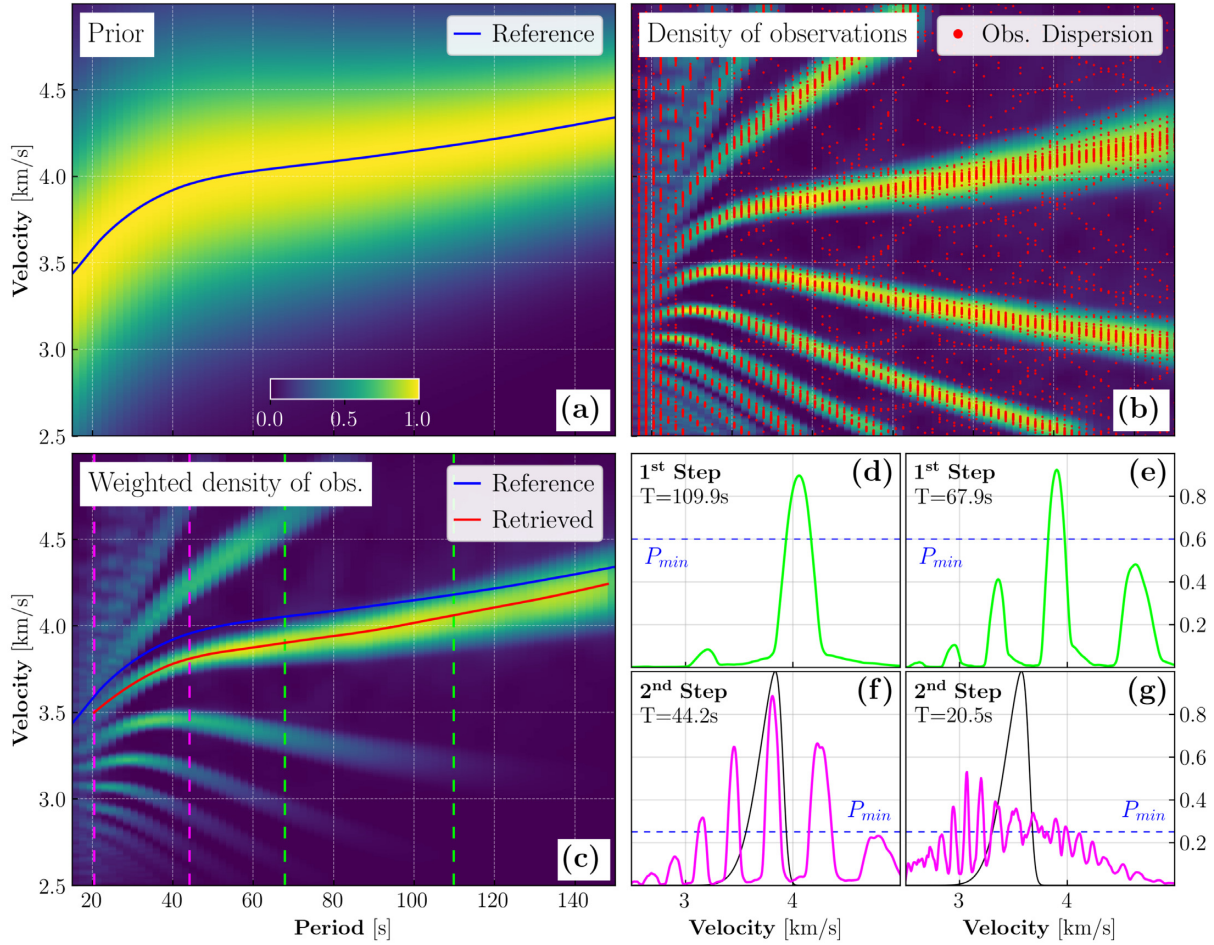


Figure 6. Illustration of our algorithm for retrieving earthquake-based phase velocities. (a) Prior probability $\mathbf{P}_m$ corresponding to a reference curve (blue line). Note the dependency of the standard deviation of the prior model as a function of period: in this study, we chose it to decrease geometrically from 0.7 at 15 s to 0.4 at 150 s. (b) Density (or probability) of observations $\mathbf{P}_d$ associated with dispersion measurements (red dots) from 51 teleseismic earthquakes. (c) Weighted probability of observation $\mathbf{P}_w = \mathbf{P}_m \odot \mathbf{P}_d$ used to retrieve a dispersion curve (red). Cross-sections of (c), corresponding to the coloured vertical lines, are shown in (d), (e), (f) and (g). Panel (e) corresponds to the ending of the first step of the procedure described in Section 3.4 (the height of the global maximum is slightly smaller than twice the second highest peak). The second step of such procedure is illustrated in (f) and (g). Multiplication of the probability profile (pink line) by the skewed Gaussian (black line) is not shown for simplicity. When the resulting (‘filtered’) probability profile does not present a global maximum which is at least 30 per cent larger than the other peaks, or that does not exceed $P_{\min}$ (0.25 in this study) the iterations terminate. This circumstance is shown in (g).

from January 2001 to June 2021, and selected shallow (<50 km) earthquakes characterized by epicentral distances between 20° and 140° from a given receiver. We pre-process each seismogram so as to remove mean and linear trends, before tapering (5 per cent) and bandpass filtering between 0.01 and 0.5 Hz. We remove the instrumental response to convert the downloaded seismograms to displacement, and downsample to 1 Hz.

We first identify triplets of epicentres-receivers approximately lying on the same great-circle path (maximum azimuthal deviation of 7°). Overall, we find 2092 station pairs with interstation distance of 3000 km or less and aligned with at least 8 epicentres, to be used in the calculation of dispersion curves. For each of such triplets, we then collect dispersion measurements as explained in Section 3.1, by analysing 75 discrete surface-wave periods linearly spaced in the range 15–150 s. Finally, application of our two-station algorithm on the dispersion measurements resulted in 2055 Rayleigh-wave interstation dispersion curves, one for each pair of receivers.

Visual inspection of the thus retrieved dispersion curves points to their robustness, and hence of our method. These are smooth and faithfully reproduce the average trend with period of the density of observations ($\mathbf{P}_d$) corresponding to the correct branch in the ambiguous dispersion measurements. This is illustrated in Fig. 4, together with the average dispersion curve and the number of measurements as a function of period.

3.6 Reliability of the two-station method

In previous studies, it has been noticed that the phase velocities measured through a two-station method tend to be faster (on average, by ~ 1 per cent) than those obtained, based on the same station pairs, from the analysis of seismic ambient noise (e.g., Yao *et al.* 2006; Foster *et al.* 2014a; Kästle *et al.* 2016). Such a systematic bias can be attributed to the fact that, in the presence of structural heterogeneity, the true propagation path of a surface wave between two receivers

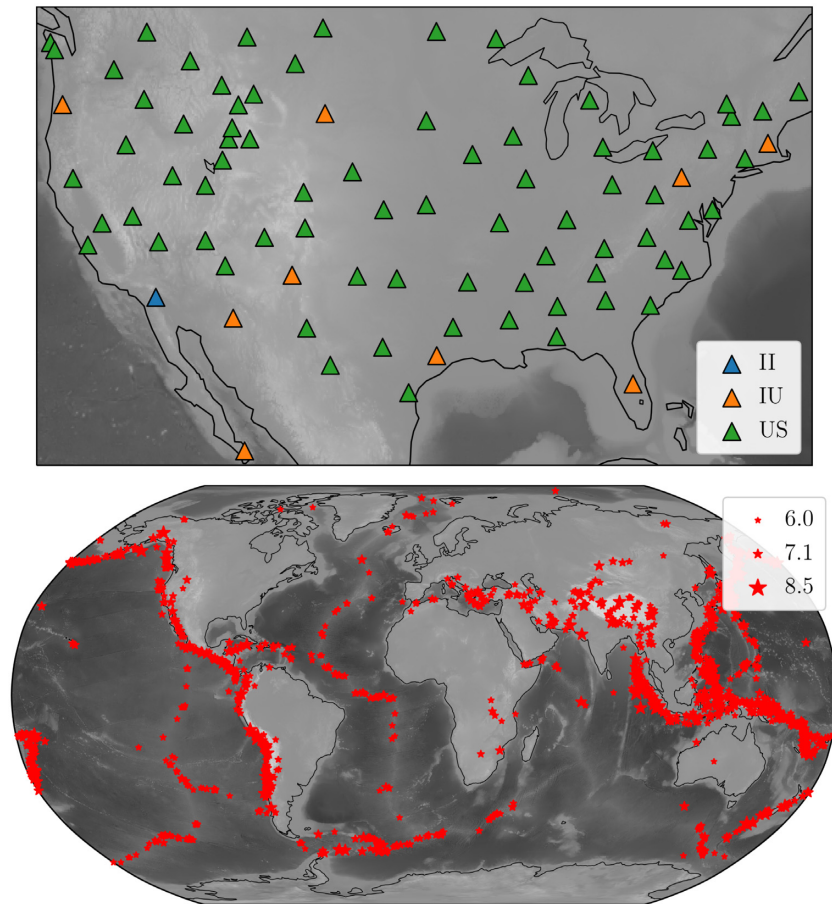


Figure 7. Top panel: location of the 90 receivers (triangles) used to measure Rayleigh-wave dispersion curves from teleseismic earthquakes. These belong to three different permanent networks, as indicated by the colours. Bottom panel: epicentres (stars) of the earthquakes used in the analysis.

is necessarily longer than the corresponding great-circle path. Because of the stronger spatial variations in surface-wave velocity at relatively short periods (corresponding to larger deviations from the predicted propagation path), the reported bias generally tends to decrease with period (Chen *et al.* 2018; Magrini *et al.* 2020a). A similar pattern is observed in our average dispersion curves in Fig. 4, with slightly faster earthquake-based phase velocities at all periods, from ~ 1.2 per cent at 15 s to ~ 0.5 per cent at 40 s, with a minimum of ~ 0.4 per cent at 32 s.

In principle, the above discrepancy could be addressed by estimating the impinging angle, as a function of period, of the earthquake-generated wavefronts at the considered station pairs, so as to convert the (measured) apparent phase velocities into ‘true’ ones. In Foster *et al.* (2014a) and Chen *et al.* (2018), for example, the impinging angles were sought using arrays of receivers, through a sort of grid-search based on a beamforming-like analysis. To decrease the computational burden inherent to that approach, Magrini *et al.* (2020a) defined a cost function based on individual pairs of receivers, to minimize the misfit between the vertical-component seismograms and the Hilbert-transformed radial components rotated at different angles.

Arrival-angle corrections, however, typically lead to discard relatively large amounts of seismograms (for technical details, see Foster *et al.* 2014a; Magrini *et al.* 2020a); as previously mentioned, this is likely to negatively impact the outcome of a tomographic study, especially when only few years of recordings are available. Moreover, neither of the above approaches allows for estimating the

impinging angle of Love waves at the receivers; in the former case, mainly due to overtone-interference effects (Foster *et al.* 2014a), in the latter, simply because Love waves do not cause displacement on the vertical components. Considering that the calculation of Rayleigh- and Love-wave phase velocities often represents a preliminary step before a joint inversion for, e.g., shear-wave velocity and/or radial anisotropy, the benefits of correcting only one kind of surface-wave velocity are to be carefully evaluated. Because of these limitations, at the time of writing SeisLib does not natively allow for estimating the impinging angles. We plan to work on this matter in the future. In the meanwhile, we invite the interested readers to consider the work of Magrini *et al.* (2022), where a simple *a posteriori* correction was applied to the inter-station dispersion curves at different periods; this significantly reduced the discrepancy between ambient-noise and earthquake-based phase velocities (for both Rayleigh and Love waves), without discarding potentially important seismograms.

4 LEAST-SQUARE IMAGING OF LATERAL VARIATIONS IN SURFACE-WAVE VELOCITY

Having calculated Rayleigh-wave dispersion curves, we show in this section how such measurements can be used to map spatial variations in surface-wave velocity. After giving a theoretical background on the inversion inherent to this task, we invert our data set of phase velocities at different periods. In doing so, we show

two different applications: the first focusing on the conterminous United States, the second involving a much larger data set including measurements collected at global scale.

4.1 An ill-conditioned problem

To map lateral variations in surface-wave velocity, SeisLib implements a least-square inversion scheme based on ray theory (e.g., Boschi & Dziewonski 1999). This method has been used in several seismological studies (e.g., Zhang & Tanimoto 1991; Ekström 2014; Lu *et al.* 2018), and rests on the assumption that surface waves propagate, from a given point on the Earth's surface to another, without deviating from the great-circle path connecting them. Under this assumption, the traveltime along the great-circle path can be written $t = \int_{\text{path}} s(\phi(l), \theta(l)) dl$, where ϕ and θ denote longitude and latitude, and s the sought Earth's slowness.

Let us consider a discrete parametrization of the Earth's surface, and assume each block (or grid cell) of such a parametrization has constant slowness. The above integral expression can then be reformulated in the discrete form

$$s = \frac{1}{L} \sum_n s_n l_n, \quad (4)$$

where L is the length of the great-circle path and l the distance travelled by the surface wave through the n th block. Eq. (4) represents the *forward* calculation that allows for retrieving the average velocity of propagation between two points on the Earth's surface (that is what we measured in Sections 2 and 3), provided that the (discrete) spatial variations in velocity (or slowness) are known. If we now define the $m \times n$ matrix such that $A_{ij} = l_j/L_i$, where L_i is the length of the great circle associated with i th observation, we can switch to matrix notation and write

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{d}, \quad (5)$$

where $\mathbf{d}$ is an m -vector whose k th element corresponds to the measured slowness, and $\mathbf{x}$ the sought n -vector whose k th element corresponds to the model coefficient s_k .

Matrix $\mathbf{A}$ can be computed numerically in a relatively simple fashion. For each pair of receivers for which a velocity measurement is available, its i th entries can be found by calculating the fraction of great-circle path connecting them through each of the n blocks associated with the parametrization. In real-world seismological applications, however, the system of equation in (5) is often strongly overdetermined, i.e. the number of data points is much larger than the number of model parameters ($m \gg n$). This implies that, although $\mathbf{A}$ is known, it is not invertible, and therefore (5) cannot be solved exactly for $\mathbf{x}$.

As pointed out by Boschi & Dziewonski (1999), the standard procedure for overcoming the above issue would involve finding a solution of (5) in least-squares sense, i.e. minimizing $\|\mathbf{A} \cdot \mathbf{x} - \mathbf{d}\|$, where $\|\mathbf{v}\| = (\mathbf{v}^T \cdot \mathbf{v})^{1/2}$ is the Euclidean norm of a vector $\mathbf{v}$. The minimum $\mathbf{x} = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \mathbf{A}^T \cdot \mathbf{d}$ can, indeed, be found analytically (e.g., Menke 2012), but a further complication arises from the fact that also $(\mathbf{A}^T \cdot \mathbf{A})$ is generally very close to being singular. In other words, small perturbations in $\mathbf{A}$ might result in large changes in $(\mathbf{A}^T \cdot \mathbf{A})^{-1}$, hence in the solution $\mathbf{x}$.

4.2 Regularized least-squares solution

To overcome the above issues, we regularize the seismic inverse problem by imposing a constraint on the smoothness of the model

parameters. This corresponds to the condition that the roughness of the solution $s(\phi, \theta)$ be minimum, i.e. $\int_{\Omega} \|\nabla s(\phi, \theta)\|^2 d\Omega = \min$ (Boschi & Dziewonski 1999), where the surface gradient $\nabla = \frac{1}{\cos \theta} \hat{\phi} \frac{\partial}{\partial \phi} + \hat{\theta} \frac{\partial}{\partial \theta}$, and $\hat{\phi}$ and $\hat{\theta}$ are the versors corresponding to the directions of increasing longitude and latitude, respectively. In discrete form, the above constraint can be written

$$\|\mathbf{R} \cdot \mathbf{x}\|^2 = \mathbf{x}^T \cdot \mathbf{R}^T \cdot \mathbf{R} \cdot \mathbf{x} = \min \quad (6)$$

(e.g., Sambridge 1990), where the roughness operator $\mathbf{R}$ is dependent on the parametrization. If we now assume that a solution $\mathbf{x}_0 \sim \mathbf{x}$ is approximately known (which is generally the case at seismic frequencies), the regularized least-squares solution

$$\mathbf{x} = \mathbf{x}_0 + (\mathbf{A}^T \cdot \mathbf{A} + \mu^2 \mathbf{R}^T \cdot \mathbf{R})^{-1} \cdot \mathbf{A}^T \cdot (\mathbf{d} - \mathbf{A} \cdot \mathbf{x}_0) \quad (7)$$

(e.g., Aster *et al.* 2018) allows for retrieving the sought spatial variations in velocity, where the scalar weight μ should be chosen via L-curve analysis (Hansen 1999).

In our implementation of the imaging described thus far, the Earth's surface is parametrized by means of equal-area grids. These prevent from artificially increasing the resolution of the resulting maps at latitudes different than zero (the effect is more prominent nearby the poles), and should therefore be preferred to Cartesian grids when investigating relatively large areas. Similar to Bijwaard *et al.* (1998) and Schaefer *et al.* (2011), SeisLib also allows for adaptive parametrizations, with finer resolution in the areas characterized by relatively high density of measurements. If we consider a given block intersected by more than a certain number of interstation great-circle paths, the finer resolution is achieved by splitting it in four sub-blocks, at the midpoint along both latitude and longitude. (In SeisLib, the operation can be performed an arbitrary number of times.) It is the purpose of the following sections to provide a robust definition of the roughness operator $\mathbf{R}$ for both kinds of parametrizations.

4.3 Roughness operator: equal-area parametrizations

For an equal-area grid consisting of n blocks spaced by Δ along both $\hat{\phi}$ and $\hat{\theta}$, the components of the surface gradient of a generic model m evaluated at block i can be approximated as

$$\begin{aligned} \frac{1}{\cos \theta} \left(\frac{\partial m}{\partial \phi} \right)_i &\approx \frac{m_i - m_{iE}}{\Delta} \\ \left(\frac{\partial m}{\partial \theta} \right)_i &\approx \frac{m_i - m_{iS}}{\Delta} \end{aligned} \quad (8)$$

(e.g., Boschi & Dziewonski 1999), where the subscripts E and S denote the blocks bounding i in the eastern and southern directions, respectively (Fig. 8). Plugging (8) into the integral expression for the roughness of m , we can write

$$\begin{aligned} \int_{\Omega} \|\nabla m(\theta, \phi)\|^2 d\Omega &\approx \sum_{i=1}^n \left[\left(\frac{m_i - m_{iE}}{\Delta} \right)^2 + \left(\frac{m_i - m_{iS}}{\Delta} \right)^2 \right] \Delta^2 \\ &\approx \sum_{i=1}^n [(m_i - m_{iE})^2 + (m_i - m_{iS})^2]. \end{aligned} \quad (9)$$

To constrain the regularized least-squares solution (7) through eq. (9), we define the $n \times n$ matrixes $\mathbf{R}_E$ and $\mathbf{R}_S$ (n is the number of model parameters) incorporating the coefficients associated with the partial derivatives (8) along the eastern and southern directions,

respectively. Since the minimum of (9) does not depend on Δ , $\mathbf{R}_E$ and $\mathbf{R}_S$ can be computed such that the only non-zero elements of the k th row of $\mathbf{R}_E$ are the k th and kE th ones, and the only non-zero elements of $\mathbf{R}_S$ are the k th and kS th ones, with their values being 1 and -1 , respectively. Then

$$\mathbf{R}^T \cdot \mathbf{R} = \mathbf{R}_E^T \cdot \mathbf{R}_E + \mathbf{R}_S^T \cdot \mathbf{R}_S \quad (10)$$

can substituted into eq. (7) to constrain the smoothness of the final solution.

4.4 Roughness operator: adaptive parametrizations

We now consider an adaptive parametrization, characterized by a non-constant spacing of the grid cells along $\hat{\theta}$ and $\hat{\phi}$, depending on the density of measurements (Fig. 8). The area of each block is therefore a function of location, and the right-hand side (RHS) of eq. (9) should be modified accordingly.

To tackle this problem, first note that the number of blocks in the parametrization could be increased by iteratively splitting the larger ones into four sub-blocks (as explained in Section 4.2), until a constant spacing Δ remains across the whole parametrization. This would correspond to translating the adaptive grid into an equal-area one, whose roughness can be computed as in eq. (9). Importantly, this procedure does not need to be carried out in practice, but the inherent idea leads to useful algebraic simplifications.

Consider, for example, the case of an adaptive parametrization defined by square blocks of only two different dimensions, where the area of the smaller ones is Δ^2 . By definition, the larger blocks in such grid will have areas equal to $4\Delta^2$, i.e., can be thought of as to be constituted by four ‘fundamental blocks’ (Fig. 8). In general, if we consider an adaptive grid that has been refined an arbitrary number of times, the area of the i th block $\Delta_i^2 = k_i \Delta^2$, where k_i denotes the number of fundamental blocks belonging to i and is either 1 (if i is a fundamental block itself), 4, or a multiple of 4. The components of the surface gradient of m evaluated at block i can then be expressed as the *average* partial derivatives

$$\begin{aligned} \left\langle \frac{1}{\cos \theta} \left(\frac{\partial m}{\partial \phi} \right)_i \right\rangle &\approx \frac{1}{k_i} \sum_{j=1}^{k_i} \frac{m_{ij} - m_{iE}}{\Delta} \\ \left\langle \left(\frac{\partial m}{\partial \theta} \right)_i \right\rangle &\approx \frac{1}{k_i} \sum_{j=1}^{k_i} \frac{m_{ij} - m_{iS}}{\Delta}, \end{aligned} \quad (11)$$

where the subscript j identifies the fundamental blocks in i (Fig. 8) and, if $k_i = 1$, the two components correspond to those in (8).

It follows that

$$\begin{aligned} \|\nabla m(\theta, \phi)_i\|^2 &\approx \frac{1}{k_i^2} \left[\left(\sum_{j=1}^{k_i} \frac{m_{ij} - m_{iE}}{\Delta} \right)^2 \right. \\ &\quad \left. + \left(\sum_{j=1}^{k_i} \frac{m_{ij} - m_{iS}}{\Delta} \right)^2 \right], \end{aligned} \quad (12)$$

and the LHS of eq. (9) can be rewritten

$$\begin{aligned} \int_{\Omega} \|\nabla m(\theta, \phi)\|^2 d\Omega &\approx \sum_{i=1}^n \frac{1}{k_i^2} \left[\left(\sum_{j=1}^{k_i} \frac{m_{ij} - m_{iE}}{\Delta} \right)^2 \right. \\ &\quad \left. + \left(\sum_{j=1}^{k_i} \frac{m_{ij} - m_{iS}}{\Delta} \right)^2 \right] \Delta_i^2 \\ &\approx \sum_{i=1}^n \frac{1}{k_i} \left[\left(\sum_{j=1}^{k_i} m_{ij} - m_{iE} \right)^2 \right. \\ &\quad \left. + \left(\sum_{j=1}^{k_i} m_{ij} - m_{iS} \right)^2 \right]. \end{aligned} \quad (13)$$

Note that, if $k_i = 1$ for all blocks i , expression (13) coincides with (9), as expected for an equal-area parametrization.

Importantly, this allows us to regularize the least-squares solution (7) based on adaptive parametrizations that have been refined an arbitrary number of times. Similar to the previous section, the minimum of (13) does not depend on Δ , and computing the coefficients of $\mathbf{R}_E$ and $\mathbf{R}_S$ from (11) is straightforward. To fix ideas, we might consider the extremely simple case of a fundamental grid consisting of only 9 blocks, 4 of which are merged to form a larger adaptive cell (Fig. 8). According to our convention, in this case the roughness operators are

$$\mathbf{R}_E = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & -0.25 & -0.25 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (14)$$

and

$$\mathbf{R}_S = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -0.25 & -0.25 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}. \quad (15)$$

In practice, the definition of $\mathbf{R}_S$ is slightly complicated by the latitudinal changes in the longitudinal extent of blocks within any equal-area grid. This results in many pairs of east–west strings of blocks, sharing their southern–northern boundary, being offset with respect to one another, as apparent from Fig. 8. Whenever this is the case, the j th term in the sum $\frac{1}{k_i} \sum_{j=1}^{k_i} (m_{ij} - m_{iS})^2$ in eq. (13) are weighted by the length of south–north boundary shared by blocks i_j and i_jS .

4.5 Numerical validation

To verify the validity of the above derivation, we carried out a numerical test. In practice, we first defined an equal-area parametrization consisting of $5^\circ \times 5^\circ$ blocks at the equator (the increment in longitude between neighbouring blocks is gradually increased from the equator to the poles to preserve the area). We then refined all blocks of longitudes $-120^\circ < \phi < 120^\circ$ and latitudes $-60^\circ < \theta < 60^\circ$, splitting them in four sub-blocks. The same operation was repeated so as to refine a second time all blocks of longitudes $-60^\circ < \phi <$

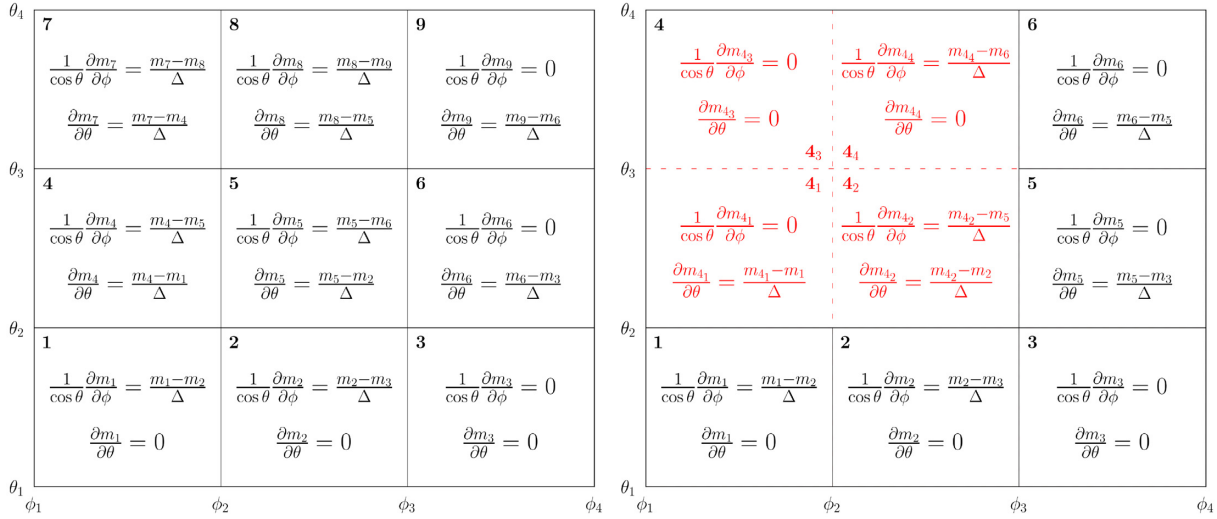


Figure 8. Our definition of the partial derivatives of a generic model m with respect to ϕ and θ , in the cases of equal-area (left-hand panel) and adaptive (right-hand panel) parametrizations. Block indexes are reported, in bold, on their upper left. In general, an adaptive parametrization is characterized by the presence of blocks of different size. The smallest ones can be thought of as the fundamental blocks of larger ones (as shown by red indexes and dashed lines). The idea can be exploited to approximate the roughness of the model, via eq. (13), and to compute $\mathbf{R}_E$ and $\mathbf{R}_S$ via eqs. (14) and (15). On the right-hand side, to decrease the notational burden, the subscript j in eq. (11) has been omitted in the model parameters m_1, m_2, m_3, m_5 and m_6 , where it has a constant value of 1.

60° and latitudes $-30^\circ < \theta < 30^\circ$. The thus obtained parametrization, illustrated in Fig. 9, consists of 7822 blocks: 694 of $5^\circ \times 5^\circ$, 2744 of $2.5^\circ \times 2.5^\circ$ and 4384 of $\Delta^2 = 1.25^\circ \times 1.25^\circ$.

We then defined the arbitrary function

$$f(\phi, \theta) = [\sin(a\phi) + \sin(b\theta)] \cos\left(\frac{\pi\theta}{180}\right) \sin\left(\frac{\pi\phi}{180}\right), \quad (16)$$

where $a = \pi/60$ and $b = \pi/30$ (Fig. 9), and found analytical formulae for its partial derivatives

$$\begin{aligned} \frac{\partial f}{\partial \phi} &= \frac{1}{180} \cos\left(\frac{\pi\theta}{180}\right) \left(\pi \cos\left(\frac{\pi\phi}{180}\right) [\sin(a\phi) + \sin(b\theta)] \right. \\ &\quad \left. + 180a \sin\left(\frac{\pi\phi}{180}\right) \cos(a\phi) \right) \\ \frac{\partial f}{\partial \theta} &= \frac{1}{180} \sin\left(\frac{\pi\phi}{180}\right) \left(180b \cos\left(\frac{\pi\theta}{180}\right) \cos(b\theta) \right. \\ &\quad \left. - \pi \sin\left(\frac{\pi\theta}{180}\right) [\sin(a\phi) + \sin(b\theta)] \right). \end{aligned} \quad (17)$$

Finally, we used (17) to calculate the norm of the gradient of f for each block of the parametrization. We did so both analytically,

$$\|\nabla m(\theta_i, \phi_i)\| = \sqrt{\left(\frac{\partial f}{\partial \phi}(\theta_i, \phi_i)\right)^2 + \left(\frac{\partial f}{\partial \theta}(\theta_i, \phi_i)\right)^2}, \quad (18)$$

and numerically,

$$\|\nabla m(\theta_i, \phi_i)\| \approx \sqrt{\left(\frac{\mathbf{R}_E}{\Delta} \cdot \mathbf{m}\right)_i^2 + \left(\frac{\mathbf{R}_S}{\Delta} \cdot \mathbf{m}\right)_i^2}, \quad (19)$$

where $\mathbf{m}$ is the vector of 7822 elements whose i th element consists of $f(\phi_i, \theta_i)$.

By comparison of the lateral variations of (18) and (19), shown in Fig. 10, we conclude that the approximate expression (13) is essentially correct. Discrepancies are limited to the boundaries between regions of different adaptive-grid density, and are comparable with

the errors inherent in discretizing a relatively rapidly varying function on a relatively coarse grid. We conclude that expression (13) is a valid measure of adaptive-grid model roughness.

4.6 Phase-velocity maps of the conterminous United States

We apply in this section our least-squares inversion algorithm to the previously calculated dispersion curves, to obtain phase-velocity maps of the conterminous United States. At surface-wave periods less or equal to 40 s, corresponding to the period range spanned by our ambient-noise phase velocities, we initially parametrize the study area with equal-area blocks of $2^\circ \times 2^\circ$ (Figs. 11a and d). We then refine, up to a maximum number of three times, all blocks intersected by at least 500 interstation great-circle paths, so as to reach a fine resolution of 0.25° in the areas characterized by a relatively high density of measurements (Figs. 11b and e). At longer periods (>40 s), where we have a relatively small number of (earthquake-based) measurements (see Fig. 4), we initialize the parametrization with $1.5^\circ \times 1.5^\circ$ blocks, and refine all those intersected by more than 20 rays up to two times, to obtain a maximum resolution of 0.35° . (Note that the thresholds involved in the refinement of the parametrization, e.g., 500 or 20, are necessarily arbitrary: in SeisLib, they can be tuned by the user in a very simple fashion, according to the details of the considered tomographic application.)

At all investigated periods, we first compile the coefficients of the data kernel $\mathbf{A}$ and of the roughness operator $\mathbf{R}^T \cdot \mathbf{R}$, as explained in the previous sections. We then solve eq. (7) for $\mathbf{x}$, using a constant reference model $\mathbf{x}_0$ corresponding to the average interstation slowness at the considered period, and an appropriate roughness damping μ based on L-curve analysis. Finally, we retrieve the sought phase-velocity structure by taking the inverse of $\mathbf{x}$. (In SeisLib, all these operations are carried out automatically.) Phase-velocity maps at the periods of 8 and 30 s are shown in Fig. 11 (panels c and f). Several other maps are shown in the supplementary information associated with this paper.

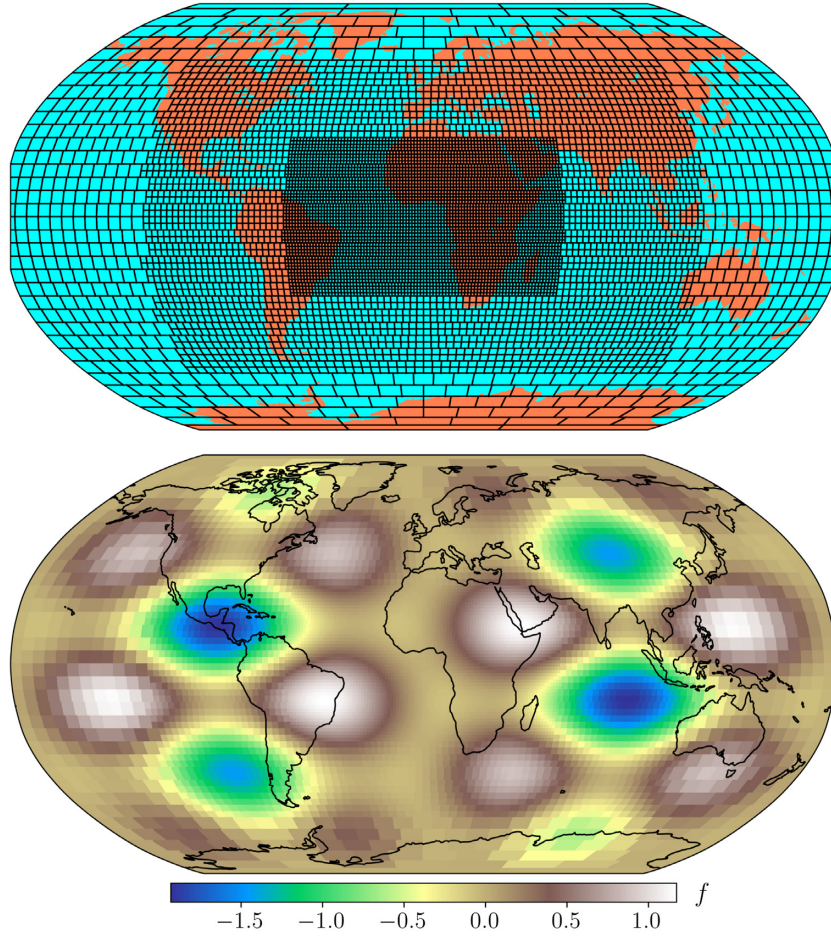


Figure 9. Adaptive grid and lateral variations of the function f defined in eq. (16), used to verify the validity of the approximate expression (13).

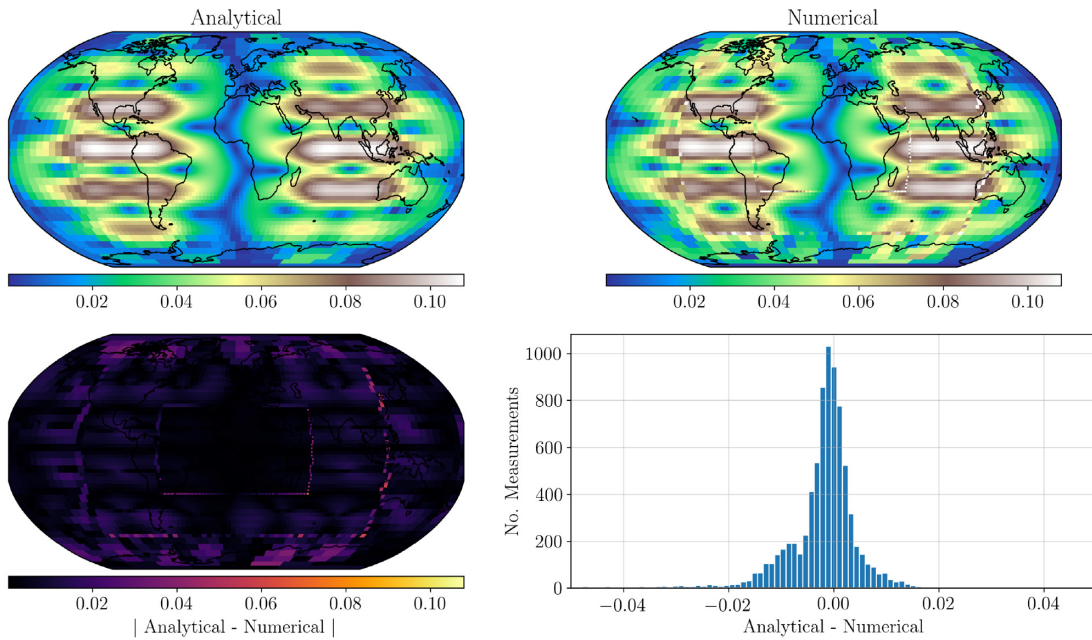


Figure 10. Results of the numerical test to verify the validity of the approximate expression (13). Top panel: lateral variations of the norm of the gradient of f , defined in eq. (16), as calculated through the direct implementation of eq. (18) (left-hand panel) and eq. (19) (right-hand panel). Bottom panel: discrepancy between the above maps calculated at each parametrization block, shown as a function of latitude and longitude (left-hand panel) or as a histogram (right-hand panel). The discrepancies are generally close to zero compared to the values of the functions that are compared, confirming that our approach is valid.

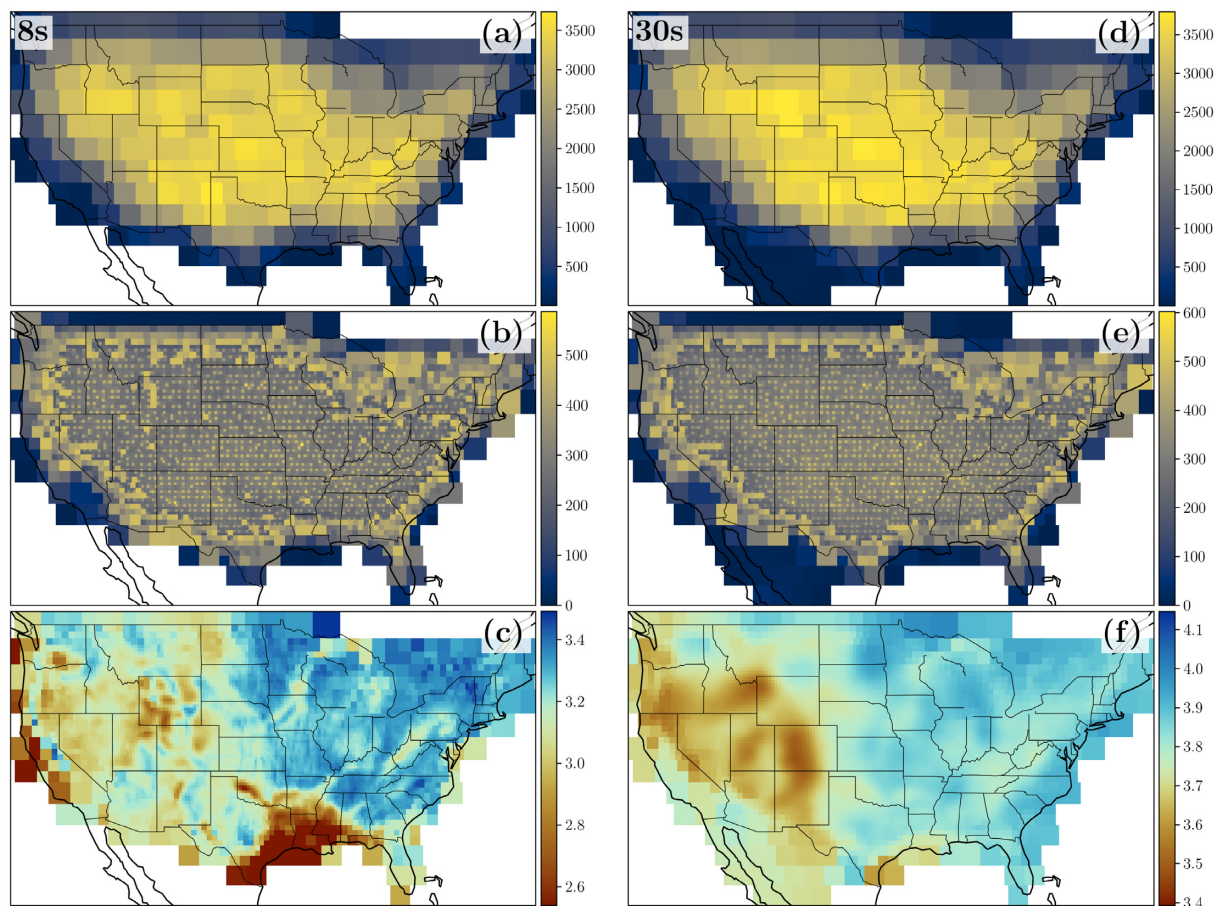


Figure 11. Illustration of the procedure followed to obtain phase-velocity maps (panels c and f, in km s^{-1}) at surface-wave periods of 8 s (left-hand panel) and 30 s (right-hand panel). At both periods, we defined an initial, equal-area parametrization of $2^\circ \times 2^\circ$ (a and d). We then refined each block iteratively up to a maximum number of three times, if intersected by at least 500 interstation great-circle paths. The resulting adaptive grids are shown in panels (b) and (e), where the background colour indicates the ray-count per block. In general, a higher resolution is achieved in the central part of the study area, as constrained by a larger number of dispersion measurements.

Visual inspection of the thus retrieved images suggests that the spatial variations in phase-velocity reflect actual Earth's features at all periods. At relatively long periods (≥ 20 s), the maps show a large-scale division of the continent into a slower west and a faster east. Such division is well documented (e.g., Nettles & Dziewoński 2008; Ekström 2014, 2017), and ascribed to the transition between the tectonically deformed west and the stable, cratonic part of North America. At shorter periods, we observe relatively strong lateral variations in velocity, corresponding to known geologic domains. Among the most prominent ones, we find relatively low velocities in regions characterized by thick sedimentary strata (e.g., Gulf of Mexico) and/or undergoing active tectonics processes (i.e. most of the Pacific Mountain system). We also find evidence of the surface expression of the mid-continent rift, in the form of lower-than-average velocities extending from Lake Superior to Kansas.

To further substantiate the robustness of our results, we compare them quantitatively with the Rayleigh-wave velocity maps obtained by Jin & Gaherty (2015) and Ekström (2017). To do so, we interpolate their maps on our grid and calculate, at a given period, (i) the Pearson correlation coefficient, to quantify the extent of spatial correlation between our maps and theirs and (ii) the relative differences in velocity at each block of the parametrization. The outcome of this analysis is illustrated in Fig. 12. At all periods, we find that our maps are highly correlated both with those of Jin & Gaherty

(2015) and with those of Ekström (2017), with an overall average correlation coefficient of 0.96. Moreover, we observe very small mean relative differences, with average absolute values of 0.14 ± 0.1 per cent. These results not only point to the robustness of our inversion scheme, but also to the goodness of the algorithms, explained in Sections 2 and 3, used to calculate ambient-noise and earthquake-based phase velocities.

4.7 Global-scale phase-velocity maps

SeisLib was designed to allow for an effortless joint use of multiple data sets, with in mind those areas for which independent measurements (possibly from different studies) are available. To conclude this section, we show an application of our library to a similar circumstance, and in particular to the least-squares imaging of Rayleigh-wave velocity at global scale.

To do this, we join our North American phase-velocity measurements at 40 s (involving both our ambient-noise and earthquake-based dispersion curves) with two different, independent data sets: one collected by Kästle *et al.* (2018) in a large region covering most of central Europe; the other consisting of one-station dispersion measurements collected at global scale by Ekström *et al.* (1997). Overall, this joint data set consists of 289 706

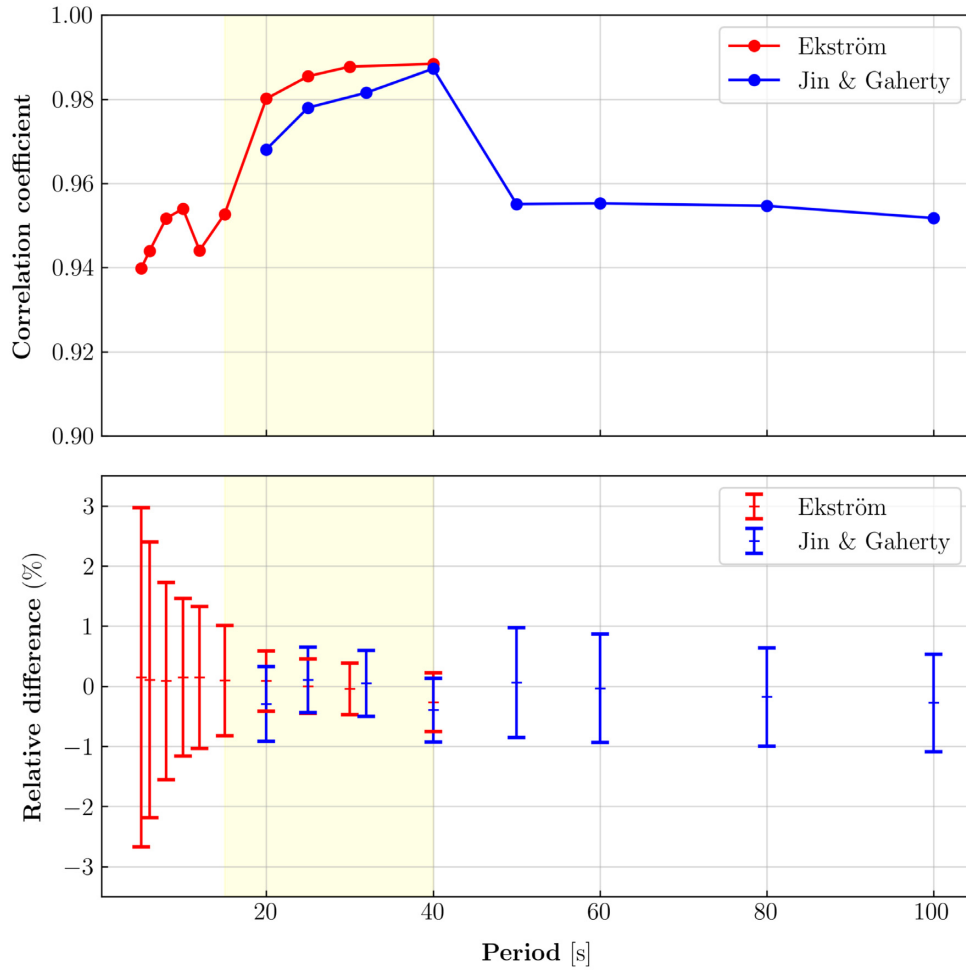


Figure 12. Pearson correlation coefficient (top panel) and relative differences (bottom panel) between our phase-velocity maps and those of Jin & Gaherty (2015) (blue) and Ekström (2017) (red). In the bottom panel, the extent of the vertical bars defines the standard deviation of the relative differences, with the mean reported in the middle as a thin horizontal line. Positive percentage values correspond to larger values in our maps. In the period range 15–40 s, indicated by the yellow shades, we derived our maps from a joint inversion of earthquake-based and ambient-noise dispersion curves.

measurements: 179 657 from Ekström *et al.* (1997) and 19 194 from Kästle *et al.* (2018).

Similar to the previous section, we initially adopt an equal-area parametrization defined by 1,654 relatively large blocks of $5^\circ \times 5^\circ$ (Fig. 13a). In three steps, we then refine all blocks intersected by at least 5000 (first step), 4000 (second step) and 3500 (third step) great-circle paths, each associated with a velocity measurement. The resulting adaptive grid consists of 3205 blocks, and is characterized by a finest resolution of 0.625° (Fig. 13b). Finally, we solve the inverse problem in the same fashion as above.

The retrieved lateral variations, shown in Fig. 13, should be compared with global-scale surface-wave velocity images obtained in previous studies (e.g., Ekström *et al.* 1997; Larson & Ekström 2001). Similar to these studies, we find relatively low velocities associated with active plate boundaries, while higher velocities characterize tectonically stable regions. Such structures can be seen at a relatively coarse resolution (i.e. 5°) across most of the globe. In North America and in central Europe, however, our maps are characterized by a relatively high resolution, owing to the dense number of measurements available in these regions. In North America, the enhanced resolution allows us to discriminate most of the features already discussed in the previous section. In Europe, we can clearly

identify the relatively low velocities associated with the Alpine chain contrasting with the higher ones in northern Europe, as in Kästle *et al.* (2018).

Based on these results, we conclude that SeisLib allows for robustly retrieving lateral variations in surface-wave velocity at global scale, even in presence of strong inhomogeneities in the data coverage. Overall, loading the data into memory, defining and refining the parametrization, and solving such a large inverse problem took 51.5 s on the first author's machine, a 6-core (AMD Ryzen 5 5600H) laptop with 16 Gb of RAM. This testifies the computational efficiency of SeisLib, which exploits the Cython compiler framework to compute the coefficients of the **A** matrix at C-like speed. The high resolution achieved in North America and Europe, in general, also points to the effectiveness of adaptive grids in seismic imaging. Adaptive tomography is the best way to maximize resolution in *global* models, wherever data coverage makes it possible; global models parametrized with *uniformly* distributed, local basis functions (or global basis functions such as spherical harmonics) would require a much more significant computational effort to achieve the same resolution in well sampled areas—and resolution would still be poor in undersampled ones.

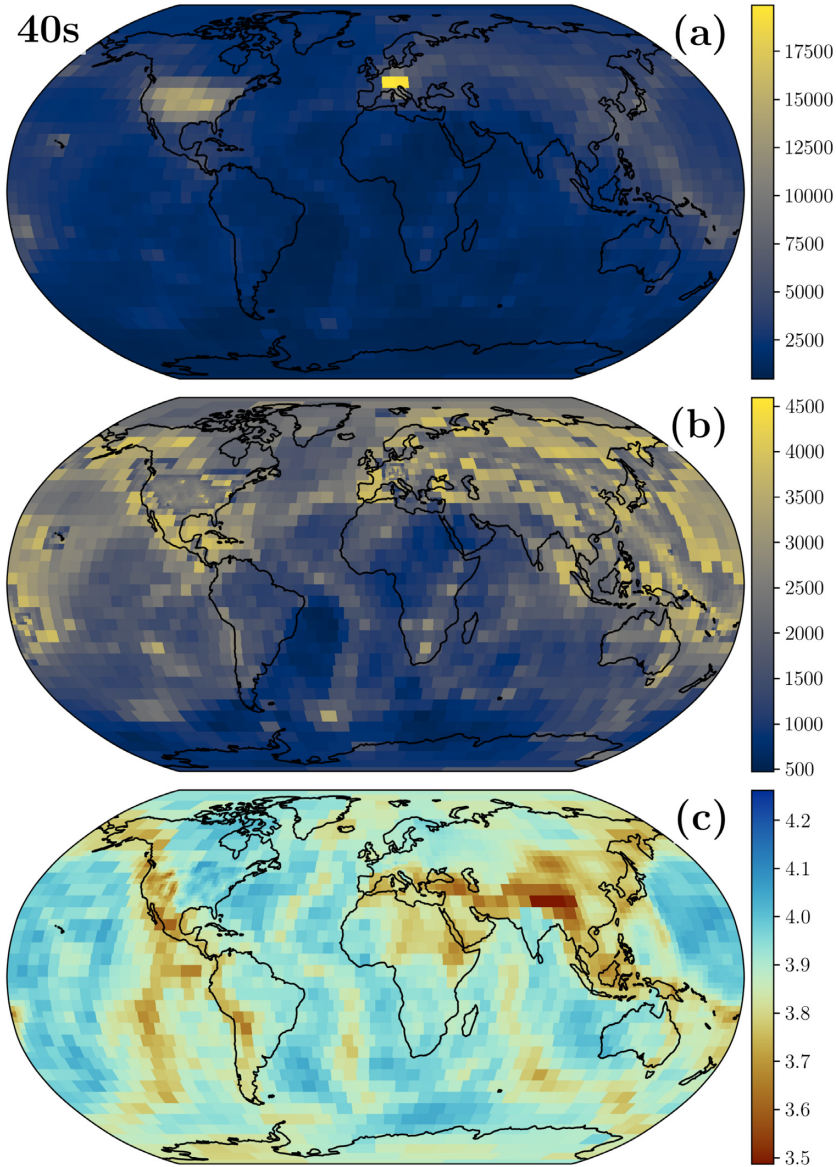


Figure 13. Same as Fig. 11, but obtained at global scale combining our North American phase velocities at 40 s with those of Ekström *et al.* (1997) (from a global distribution of earthquakes and stations) and Kästle *et al.* (2018) (collected in central Europe).

5 RAYLEIGH-WAVE ATTENUATION FROM SEISMIC AMBIENT NOISE

As anticipated, SeisLib also allows for estimating Rayleigh-wave attenuation as a function of frequency, based on seismic ambient noise recordings. As opposed to the problem of calculating phase velocities, which are only related to the phase of the empirical Green's function, estimates of attenuation should rely on its *amplitude*. Previous studies showed how the amplitude of the empirical Green's function should be treated with caution, because of its sensitivity to parameters such as the distribution of noise sources (e.g., Tsai 2011). Accordingly, Magrini & Boschi (2021) designed a procedure that should contribute to ‘regularizing’ the subsequent inversion for the Rayleigh-wave attenuation coefficient α , as verified by a suite of numerical tests. Such procedure builds on the previous work of Boschi *et al.* (2019, 2020), and is summarized in the next section.

5.1 Grid-search for the attenuation coefficient

Let us consider a relatively dense array of receivers, and focus on two of them located at $\mathbf{x}_A$ and $\mathbf{x}_B$. Following Magrini & Boschi (2021), we can rewrite the Rayleigh-wave empirical Green's function

$$\Re \left\{ \frac{u(\mathbf{x}_A, \omega) u^*(\mathbf{x}_B, \omega)}{\langle |u(\mathbf{x}, \omega)|^2 \rangle_{\mathbf{x}}} \right\} \approx J_0 \left(\frac{\omega \Delta}{c} \right) e^{-\alpha(\omega) \Delta}, \quad (20)$$

where u denotes the ambient-noise recording in the frequency domain, $*$ complex conjugation and Δ interstation distance. As in eq. (1), the LHS of (20) is calculated by ensemble averaging the normalized cross-spectra associated with the receiver pair in question, over a sufficiently large number of time windows. The normalization term $\langle |u(\mathbf{x}, \omega)|^2 \rangle_{\mathbf{x}}$ corresponds to the average power spectral density (PSD) recorded by the seismic array. Similar to empirical

normalization terms such as whitening (used in Section 2) or one-bit (e.g., Bensen *et al.* 2007), PSD-normalization serves to mitigate the effect of anomalous signals such as large or nearby earthquakes, but should be considered preferable to more robustly constrain the amplitude of the empirical Green's function (Boschi *et al.* 2019).

Eq. (20) allows to formulate an inverse problem to constrain α from the data. The strategy proposed by Magrini & Boschi (2021), implemented in SeisLib, consists of minimizing the cost function

$$C(\alpha, \omega) = \sum_{i,j} \Delta_{ij}^2 \left| \text{env} \left[\Re \left\{ \frac{u(\mathbf{x}_i, \omega) u^*(\mathbf{x}_j, \omega)}{|u(\mathbf{x}, \omega)|^2_{\mathbf{x}}} \right\} \right] - \text{env} \left[J_0 \left(\frac{\omega \Delta_{ij}}{c_{ij}(\omega)} \right) e^{-\alpha(\omega) \Delta_{ij}} \right] \right|^2, \quad (21)$$

where the subscripts i and j denote the pair of receivers and the weight Δ_{ij}^2 is used to compensate for the decrease in the amplitude of J_0 with interstation distance, due to geometrical spreading. The envelope function *env* is implemented by fitting a combination of cubic splines to the maxima of the absolute value of its argument, and has beneficial effects on the stability of the inversion (e.g., Boschi *et al.* 2019).

In eq. (21), phase velocity c is assumed to be known, i.e., it should be calculated in a preliminary step as explained in Section 2. The minimum of $C(\alpha, \omega)$ can then be found via ‘grid-search’ over α , for a discrete set of values of ω . The alert reader might notice at this point that, since the sum is carried out over all possible combinations of station pairs belonging to the considered array, only one attenuation curve can be extracted from such minimization. This strategy, albeit in a sense restrictive, has been shown to yield robust estimates of the frequency dependency of α even in presence of a heterogeneous distribution of noise sources (Magrini & Boschi 2021). If the array has good azimuthal coverage, using all station pairs as an ensemble in the minimization of $C(\alpha, \omega)$ allows for sampling most azimuths of wave propagation. In turn, this should ‘regularize’ the inversion by decreasing unwanted effects due to inhomogeneities in the distribution of the noise sources, or compensating for a non-perfectly diffuse ambient seismic field.

5.2 Lateral variations of the attenuation coefficient: conterminous United States

To illustrate the potential of SeisLib in estimating the lateral variations in the Rayleigh-wave attenuation coefficient, we apply it here to the continuous seismograms from the transportable component of USArray, previously used in Section 2. For this purpose, USArray represents an ideal playground: its very dense distribution of receivers can be exploited to subdivide the study area in many sub-arrays, so as to retrieve one attenuation curve for each of them.

To do so, we subdivide the study area in relatively large, equal-area blocks of $2.5^\circ \times 2.5^\circ$, overlapping along both longitude and latitude by 50 per cent. Each of these overlapping blocks identifies a subset of receivers of the whole USArray (Fig. 14), which is then processed by SeisLib to (i) subdivide their simultaneous recordings into relatively small time windows (here, 6-hour long); for each of these time windows, SeisLib (ii) cross-correlates the recordings from all pairs of receivers and (iii) normalizes the cross-correlations by the average PSD of the subarray. Having computed the PSD-normalized cross-correlations, these are used to (iv) retrieve the LHS of eq. (20), by ensemble average. For the same station pairs

used to obtain the cross-spectra, (v) their corresponding Bessel functions $J_0 \left(\frac{\omega \Delta}{c(\omega)} \right)$ are then computed, where phase velocity c should be retrieved in a preliminary step, as in Section 2. Finally, (vi) the envelopes of the normalized cross-spectra and of the corresponding Bessel functions are taken, and eq. (21) is used to retrieve α as a function of frequency (Fig. 14). Here, we search for the minimum of $C(\alpha, \omega)$ among 350 values of α , geometrically spaced between 5×10^{-8} and 10^{-4} m^{-1} .

Having retrieved one attenuation curve per subarray, these can be used to map the lateral variations of α at discrete frequencies. In SeisLib, this is done by first creating a new equal-area grid, whose blocks should be smaller than those used to identify the overlapping subarrays described above. Then, the attenuation value attributed to the i th block of such parametrization is

$$\alpha'_i(\omega) = \frac{\sum_k \alpha_k(\omega) N_k}{\sum_k N_k}, \quad (22)$$

where the subscript k denotes each block of the coarsest parametrization (corresponding to the k th subarray) that intersects (or contains) the block i , and the weight N_k the number of cross-spectra used in the inversion for α_k . Incidentally, Magrini *et al.* (2021) weighted the above average by the lengths of ray-path segments crossing the block i associated with each receiver pair, resulting in slightly different results.

An example of a Rayleigh-wave attenuation map, obtained at the surface-wave period of 4 s through eq. (22) and using a parametrization of $1^\circ \times 1^\circ$, is shown in Fig. 14. A detailed interpretation of the map in question is beyond the scope of this study (the interested reader will find a thorough discussion in Magrini *et al.* 2021). In general, the features visible in our image are analogous to those in Magrini *et al.* (2021). We observe relatively high attenuation in the western North America and in large sedimentary basins (e.g., Gulf of Mexico), as opposed to the lower attenuation associated with the (tectonically more stable) east.

6 CONCLUSIONS

The purpose of this paper was to illustrate the functionalities of SeisLib, an open-source Python package for seismic imaging from the local to the global scale. As opposed to several other packages available for download (e.g., Jiang & Denolle 2020; White *et al.* 2020), SeisLib has been designed to encompass a wide variety of tools to obtain seismic images of the subsurface. At the time of writing, these involve routines to download, process, and invert surface-wave velocity measurements, and we plan to implement additional features for body-wave analysis in the near future.

We structured this article so as to emulate a real-world application of surface-wave tomography, based on the calculation of both seismic-ambient-noise and earthquake-based Rayleigh-wave phase velocities. In doing so, we first illustrated the theoretical and algorithmic background of each SeisLib module, and then applied them to seismic data (downloaded and pre-processed through SeisLib itself) from North America. Earthquake-based dispersion curves are calculated in SeisLib using the two-station method (e.g., Meier *et al.* 2004), which we implemented by revisiting and improving the earlier work of Magrini *et al.* (2020a). The algorithm for retrieving ambient-noise dispersion curves builds on the study of K  stle *et al.* (2016) and has been optimized to process very large data sets such as that of USArray.

To map lateral variations in surface-wave velocity, SeisLib relies on a least-squares inversion scheme based on ray theory

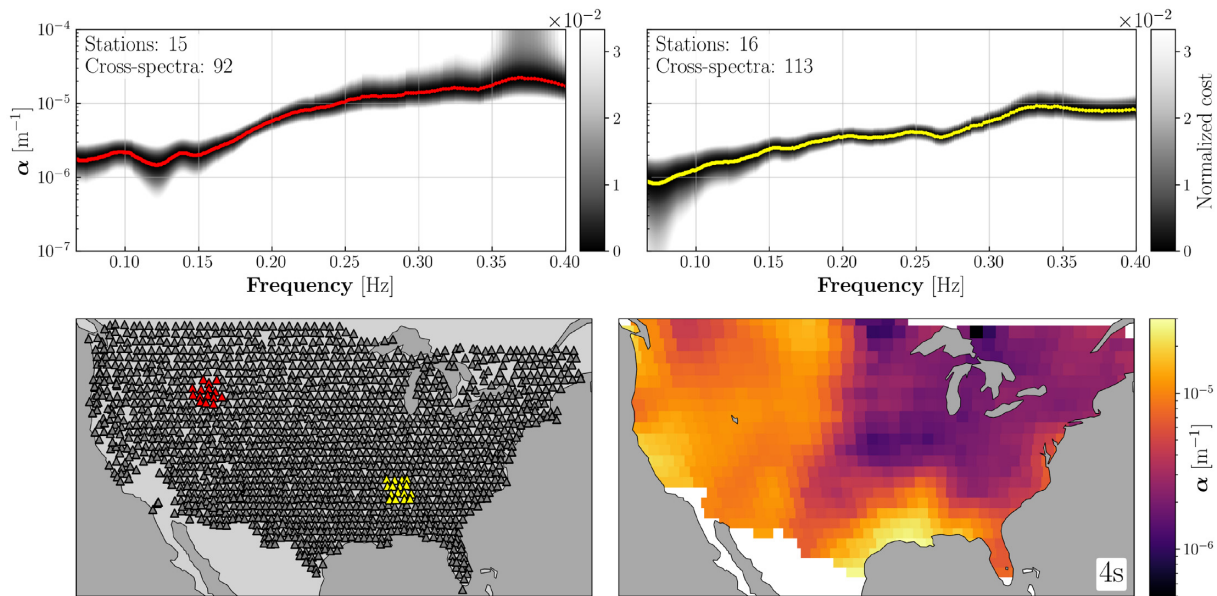


Figure 14. Normalized cost function $C(\alpha, \omega)$ (top panel) obtained from the inversion of two different subarrays (bottom left-hand panel) of the USArray. The red and yellow dots in the top panels indicate the minima of the cost function, i.e., the retrieved attenuation curve for the two seismic arrays (which are coloured accordingly). The bottom right-hand panel shows the attenuation map retrieved at the period of 4 s through eq. (22), based on an equal-area parametrization of $1^\circ \times 1^\circ$ (Section 5.2).

(e.g., Boschi & Dziewonski 1999). Our implementation allows for both equal-area and adaptive parametrizations, with the latter being characterized by a finer resolution in the regions of relatively high density of measurements. We defined rigorously the roughness operator (used to regularize the seismic inverse problem) for both circumstances, and verified our findings numerically. The roughness of a generic model is given, in discrete form, by eq. (13). Importantly, this equation holds for both equal-area grids and adaptive grids that have been refined an arbitrary number of times, as far as the area of the larger blocks in the considered parametrization is a multiple of that of the finest blocks. To our knowledge, this general result does not appear in the earlier seismological literature.

We applied the above inversion scheme to our phase-velocity measurements, amounting to 144 272 ambient-noise and 2055 earthquake-based dispersion curves. In the broad period range 4–100 s, our maps of North America are highly correlated (on average, 96 per cent) with those of Ekström (2014, 2017) and Jin & Gaherty (2015), obtained with data and methods that are entirely independent from ours. Very small mean relative differences (0.14 ± 0.1 per cent) are also found between our maps and those of the same authors, confirming the robustness of the algorithms used to calculate the dispersion curves and to map the lateral variations.

In the last section of the manuscript, we briefly illustrated the potential of SeisLib in retrieving Rayleigh-wave attenuation as a function of frequency, from seismic ambient noise. Our implementation builds on the strategy of Magrini & Boschi (2021), which exploits a dense seismic array to obtain one attenuation curve from the inversion of all available cross-spectra, using them as an ensemble of measurements. This approach can be used to map Rayleigh-wave attenuation, by identifying several (spatially overlapping) subarrays in the area of interest so as to calculate one attenuation curve for each of them. We followed this logic to obtain a Rayleigh-wave attenuation map of the conterminous United States, at the period of 4 s, which is consistent with the existing literature.

Our results suggest that SeisLib can be used reliably in seismic imaging. Importantly for novice programmers, its modular

architecture makes it easy to carry out long processing tasks in an automated fashion, and its optimized speed allows to handle very large data sets even in presence of limited computational resources. The latest release of our library can be downloaded at <https://pypi.org/project/seislib/>.

ACKNOWLEDGMENTS

We are grateful to Fabio Cramer for having allowed us to use his scientific colour maps (<https://www.fabiocramer.ch/colourmaps/>), which are the default in several functions of our library. Funded by the Deutsche Forschungsgemeinschaft (DFG - German Research Foundation) under the Individual Research Project: SI 1748/4-1. EK has received funding from the German Science Foundation: Deutsche Forschungsgemeinschaft (www.dfg.de; SPP-2017, Project Ha 2403/21-1).

DATA AVAILABILITY

We obtained seismic waveforms, related metadata, and derived products used in this study from the IRIS Data Management Center (<https://ds.iris.edu/ds/nodes/dmc/>). This study is based on publicly available seismic data from the Transportable Array network (<https://doi.org/10.7914/SN/TA>), the United States National Seismic network (<https://doi.org/10.7914/SN/US>) and the Global Seismograph networks (<https://doi.org/10.7914/SN/II> and <https://doi.org/10.7914/SN/IU>). Digital elevation data were downloaded from <https://earthexplorer.usgs.gov/>.

REFERENCES

- Aster, R.C., Borchers, B. & Thurber, C.H., 2018. *Parameter Estimation and Inverse Problems*, Elsevier.
- Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A.L., Lin, F., Moschetti, M., Shapiro, N. & Yang, Y., 2007. Processing seismic ambient noise

- data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.*, **169**(3), 1239–1260.
- Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y. & Wassermann, J., 2010. ObsPy: A Python toolbox for seismology, *Seismol. Res. Lett.*, **81**(3), 530–533.
- Bijwaard, H., Spakman, W. & Engdahl, E.R., 1998. Closing the gap between regional and global travel time tomography, *J. geophys. Res.*, **103**(B12), 30 055–30 078.
- Boschi, L. & Dziewonski, A.M., 1999. High-and low-resolution images of the Earth's mantle: Implications of different approaches to tomographic modeling, *J. geophys. Res.*, **104**(B11), 25 567–25 594.
- Boschi, L. & Weemstra, C., 2015. Stationary-phase integrals in the cross correlation of ambient noise, *Rev. Geophys.*, **53**(2), 411–451.
- Boschi, L., Weemstra, C., Verbeke, J., Ekström, G., Zunino, A. & Giardini, D., 2013. On measuring surface wave phase velocity from station–station cross-correlation of ambient signal, *Geophys. J. Int.*, **192**(1), 346–358.
- Boschi, L., Magrini, F., Cammarano, F. & van der Meijde, M., 2019. On seismic ambient noise cross-correlation and surface-wave attenuation, *Geophys. J. Int.*, **219**(3), 1568–1589.
- Boschi, L., Magrini, F., Cammarano, F. & van der Meijde, M., 2020. Erratum: On seismic ambient noise cross-correlation and surface-wave attenuation, *Geophys. J. Int.*, **222**(2), 1090–1092.
- Campillo, M., Roux, P., Romanowicz, B. & Dziewonski, A., 2014. Seismic imaging and monitoring with ambient noise correlations, in *Treatise on Geophysics*, Vol. 1, 2nd edn, pp. 256–271, eds Romanowicz, B. & Dziewonski, A., Elsevier.
- Chen, H., Ni, S., Chu, R., Chong, J., Liu, Z. & Zhu, L., 2018. Influence of the off-great-circle propagation of Rayleigh waves on event-based surface wave tomography in Northeast China, *Geophys. J. Int.*, **214**(2), 1105–1124.
- Cramer, F., 2018. Scientific colour-maps, *Zenodo*, 10.
- Darbyshire, F.A. & Lebedev, S., 2009. Rayleigh wave phase-velocity heterogeneity and multilayered azimuthal anisotropy of the Superior Craton, Ontario, *Geophys. J. Int.*, **176**(1), 215–234.
- Darbyshire, F.A. *et al.*, 2004. A first detailed look at the Greenland lithosphere and upper mantle, using Rayleigh wave tomography, *Geophys. J. Int.*, **158**(1), 267–286.
- Ekström, G., 2014. Love and Rayleigh phase-velocity maps, 5–40 s, of the western and central USA from USArray data, *Earth planet. Sci. Lett.*, **402**, 42–49.
- Ekström, G., 2017. Short-period surface-wave phase velocities across the conterminous United States, *Phys. Earth planet. Inter.*, **270**, 168–175.
- Ekström, G., Tromp, J. & Larson, E. W.F., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.*, **102**(B4), 8137–8157.
- 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**, doi:10.1029/2009GL039131.
- Eulenfeld, T., 2020. rf: Receiver function calculation in seismology, *J. Open Sour. Software*, **5**(48), 1808.
- Foster, A., Ekström, G. & Nettles, M., 2014a. Surface wave phase velocities of the Western United States from a two-station method, *Geophys. J. Int.*, **196**(2), 1189–1206.
- Foster, A., Nettles, M. & Ekström, G., 2014b. Overtone interference in array-based love-wave phase measurements, *Bull. seism. Soc. Am.*, **104**(5), 2266–2277.
- Hansen, P.C., 1999. The L-curve and its use in the numerical treatment of inverse problems, *Computational Inverse Problems in Electrocardiology*, Advances in Computational Bioengineering, pp. 19–142, Johnston, P.R., WIT Press.
- Hariharan, A., Dalton, C.A., Babikoff, J. & Ekström, G., 2022. Controls on surface wave overtone interference, *Geophys. J. Int.*, **228**(3), 1665–1683.
- Jiang, C. & Denolle, M.A., 2020. NoisePy: a new high-performance python tool for ambient-noise seismology, *Seismol. Res. Lett.*, **91**(3), 1853–1866.
- Jin, G. & Gaherty, J.B., 2015. Surface wave phase-velocity tomography based on multichannel cross-correlation, *Geophys. J. Int.*, **201**(3), 1383–1398.
- Kästle, E., 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**, 1493–1512.
- Kästle, E.D., El-Sharkawy, A., Boschi, L., Meier, T., Rosenberg, C., Bellahsen, N., Cristiano, L. & Weidle, C., 2018. Surface wave tomography of the Alps using ambient-noise and earthquake phase velocity measurements, *J. geophys. Res.*, **123**(2), 1770–1792.
- Kästle, E.D., Molinari, I., Boschi, L. & Kissling, E., the AlpArray Working Group, 2021. Azimuthal anisotropy from eikonal tomography: example from ambient-noise measurements in the AlpArray network, *Geophys. J. Int.*, **229**(1), 151–170.
- Larson, E.W. & Ekström, G., 2001. Global models of surface wave group velocity, in *Monitoring the Comprehensive Nuclear-Test-Ban Treaty: Surface Waves*, pp. 1377–1399, Springer.
- Lecocq, T., Caudron, C. & Brenguier, F., 2014. MSNoise, a python package for monitoring seismic velocity changes using ambient seismic noise, *Seismol. Res. Lett.*, **85**(3), 715–726.
- Lu, Y., Stehly, L., Paul, A. & Group, A.W., 2018. High-resolution surface wave tomography of the European crust and uppermost mantle from ambient seismic noise, *Geophys. J. Int.*, **214**(2), 1136–1150.
- Magrini, F. & Boschi, L., 2021. Surface-wave attenuation from seismic ambient noise: numerical validation and application, *J. geophys. Res.*, **126**(1), e2020JB019865.
- Magrini, F., Diaferia, G., Boschi, L. & Cammarano, F., 2020a. Arrival-angle effects on two-receiver measurements of phase velocity, *Geophys. J. Int.*, **220**(3), 1838–1844.
- Magrini, F., Diaferia, G., Fadel, I., Cammarano, F., van der Meijde, M. & Boschi, L., 2020b. 3-D shear wave velocity model of the lithosphere below the Sardinia–Corsica continental block based on Rayleigh-wave phase velocities, *Geophys. J. Int.*, **220**(3), 2119–2130.
- Magrini, F., Boschi, L., Gualtieri, L., Lekić, V. & Cammarano, F., 2021. Rayleigh-wave attenuation across the conterminous United States in the microseism frequency band, *Sci. Rep.*, **11**(1), 1–9.
- Magrini, F., Diaferia, G., El-Sharkawy, A., Cammarano, F., van der Meijde, M., Meier, T. & Boschi, L., 2022. Surface-wave tomography of the central-western Mediterranean: new insights into the Liguro-Provençal and Tyrrhenian basins, *J. geophys. Res.*, **127**(3), e2021JB023267.
- Meier, T., Dietrich, K., Stöckhert, B. & Harjes, H.-P., 2004. One-dimensional models of shear wave velocity for the eastern Mediterranean obtained from the inversion of Rayleigh wave phase velocities and tectonic implications, *Geophys. J. Int.*, **156**(1), 45–58.
- Menke, W., 2012. *Geophysical Data Analysis: Discrete Inverse Theory*, MATLAB edition, Vol. 45, Academic Press.
- Millman, K.J. & Aivazis, M., 2011. Python for scientists and engineers, *Comp. Sci. Eng.*, **13**(2), 9–12.
- Nagar, S., 2017. *Introduction to Python for Engineers and Scientists: Open Source Solutions for Numerical Computation*, Apress.
- Nakahara, H., 2012. Formulation of the spatial autocorrelation (SPAC) method in dissipative media, *Geophys. J. Int.*, **190**(3), 1777–1783.
- Nakata, N., Gualtieri, L. & Fichtner, A., 2019. *Seismic Ambient Noise*, Cambridge Univ. Press.
- Nettles, M. & Dziewoński, A.M., 2008. Radially anisotropic shear velocity structure of the upper mantle globally and beneath North America, *J. geophys. Res.*, **113**(B2), doi:10.1029/2006JB004819.
- Parzen, E., 1962. On estimation of a probability density function and mode, *Ann. Math. Stat.*, **33**(3), 1065–1076.
- Sadeghisorkhani, H., Gudmundsson, Ó. & Tryggvason, A., 2018. Gspecdisp: a Matlab GUI package for phase-velocity dispersion measurements from ambient-noise correlations, *Comp. Geosci.*, **110**, 41–53.
- Sambridge, M., 1990. Non-linear arrival time inversion: constraining velocity anomalies by seeking smooth models in 3-D, *Geophys. J. Int.*, **102**(3), 653–677.
- Schaefer, J., Boschi, L. & Kissling, E., 2011. Adaptively parametrized surface wave tomography: methodology and a new model of the European upper mantle, *Geophys. J. Int.*, **186**(3), 1431–1453.

- Seats, K.J., Lawrence, J.F. & Prieto, G.A., 2012. Improved ambient noise correlation functions using Welch's method, *Geophys. J. Int.*, **188**(2), 513–523.
- Sethian, J.A., 1996. A fast marching level set method for monotonically advancing fronts, *Proc. Natl. Acad. Sci.*, **93**(4), 1591–1595.
- Soomro, R., Weidle, C., Cristiano, L., Lebedev, S., Meier, T. & Group, P.W., 2016. Phase velocities of Rayleigh and Love waves in central and northern Europe from automated, broad-band, interstation measurements, *Geophys. J. Int.*, **204**(1), 517–534.
- Tsai, V.C., 2011. Understanding the amplitudes of noise correlation measurements, *J. geophys. Res.*, **116**(B9), doi:10.1029/2011JB008483.
- White, M.C., Fang, H., Nakata, N. & Ben-Zion, Y., 2020. PyKonal: a Python package for solving the eikonal equation in spherical and Cartesian coordinates using the fast marching method, *Seismol. Res. Lett.*, **91**(4), 2378–2389.
- Yao, H., van Der Hilst, R.D. & De Hoop, M.V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis-I. Phase velocity maps, *Geophys. J. Int.*, **166**(2), 732–744.
- Zhang, Y.-S. & Tanimoto, T., 1991. Global Love wave phase velocity variation and its significance to plate tectonics, *Phys. Earth planet. Inter.*, **66**(3–4), 160–202.

SUPPORTING INFORMATION

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

Figure S1. Phase velocity maps of the conterminous United States at different periods. Note the size of the different blocks constituting the parametrization, which varies with location and period depending on the density of measurements. In general, a higher resolution is achieved in the central part of the study area, as constrained by a larger number of interstation dispersion curves.

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.