A new approach to determine surface-wave attenuation from seismic ambient noise: numerical validation and application

Fabrizio Magrini[1], **Lapo Boschi**[2,3,4]

[1]Department of Sciences, Università degli Studi Roma Tre, Italy

[2]Dipartimento di Geoscienze, Università degli Studi di Padova, Italy

[3]Sorbonne Université, CNRS, INSU, Institut des Sciences de la Terre de Paris, ISTeP UMR 7193, F-75005 Paris, France.

[4]Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

Key Points:

- numerical validation of a new method for measuring ambient surface-wave attenuation
- application to North American data yields attenuation estimates compatible with previous studies
- the method is promising for quantifying ambient surface-wave attenuation at relatively high frequencies

Corresponding author: Fabrizio Magrini, `fabrizio.magrini@uniroma3.it`

Abstract

We validate a new method to determine surface-wave attenuation from seismic ambient noise, both numerically and by application to recordings from a dense broadband array. We generate synthetic recordings of numerically simulated ambient seismic noise in several experimental setups, characterized by different source distributions and different values of attenuation coefficient. We use them to verify that: (I) "cross-terms" cancel out, as predicted by the theory; (II) the source spectrum can be reconstructed from ambient recordings, provided that the density of sources and the attenuation coefficient are known; (III) true attenuation can be retrieved from normalized cross correlations of synthetic signals. We then apply the so validated method to real continuous recordings from 33 broadband receivers distributed within the Colorado Plateau and Great Basin. A preliminary analysis of the signal-to-noise ratio as a function of azimuth reveals a SW-NE preferential directionality of the noise sources within the secondary microseism band (6-8 s), as previously reported by other authors. By nonlinear inversion of noise data we find the attenuation coefficient in the area of interest to range from $\sim 1 \times 10^{-5}$ m^{-1} at 0.3 Hz to $\sim 4.5 \times 10^{-7}$ m^{-1} at 0.065 Hz, and confirm the statistical robustness of this estimate by means of a bootstrap analysis. This result is compatible with previous observations made on the basis of both earthquake-generated and ambient Rayleigh waves. In this regard, the new method proves to be promising in accurately quantifying surface-wave attenuation at relatively high frequencies.

1 Introduction

Over the last century, seismologists have learned to constrain the velocity of seismic waves increasingly well, but its interpretation in terms of temperature, density, viscosity, and composition of the Earth's interior is nonunique and remains problematic. As opposed to their speed of propagation, the *amplitude* of seismograms is directly related to anelastic dissipation; knowing how the Earth attenuates seismic waves, and how such attenuation changes with location within our planet, would tell us much more about its properties than we currently know. But measures of amplitude carry important uncertainty, and the theory relating seismogram amplitude to Earth parameters is cumbersome and occasionally (e.g. Menon et al., 2014; Boschi et al., 2019) controversial.

Several studies have shown that cross correlations of seismic ambient noise approximately coincide with the surface-wave Green's function associated with the two points

of observation. By analysing the *phase* of the empirical Green's function, it is possible to successfully image and monitor the velocity structure of the Earth's interior (see the reviews by, e.g., Campillo & Roux, 2014; Boschi & Weemstra, 2015). The information on the anelastic properties carried by its *amplitude*, on the other hand, is less accurately reconstructed by cross correlation. Initial attempts to constrain surface-wave attenuation from ambient noise (e.g. Prieto et al., 2009; Lawrence & Prieto, 2011) were based on the assumption that attenuation could be accounted for by simply taking the product of the Green's function and an exponential damping term. Tsai (2011) showed that these works omitted a multiplicative factor dependent on source parameters, which, if not accounted for, is likely to introduce a bias in the attenuation estimates; Weemstra et al. (2013) chose to treat that factor as a free parameter in their formulation of the inverse problem. However, Weemstra et al. (2014) showed an additional difficulty associated with the normalization of cross correlations, used in ambient-noise literature to reduce the effects of e.g. strong earthquakes; spectral whitening or other normalization terms affect the amplitude of the empirical Green's function, biasing the measurements of attenuation.

Boschi et al. (2019) recently derived a mathematical expression for the multiplicative factor relating the normalized cross correlations to the Rayleigh-wave Green's function; this factor accounts for the bias introduced by normalization, and incorporates the parameters associated with the source distribution. Based on this theoretical result, they implemented a new method for constraining the Rayleigh-wave attenuation coefficient, *without* prior knowledge of source parameters. This method was tested by Boschi et al. (2019) on an 11-receiver array deployed on the island of Sardinia, Italy. The relatively small size of the array and the lack of literature on seismic attenuation in Sardinia, however, made it difficult to assess the accuracy of the so obtained attenuation estimates. In this study, we validate the method of Boschi et al. (2019) numerically. Finally, we apply the same method to seismograms from a very dense broadband array.

After summarizing the theory (Section 2), we formulate in Section 3 an inverse problem to retrieve the attenuation coefficient from ambient-noise cross correlations. We illustrate in Section 4 two numerical tests of our method. In Section 5 we apply it to a subset of the USArray database, consisting of 33 receivers distributed within the Colorado Plateau and Great Basin. Finally, we note that an algebraic error was found in Boschi et al. (2019), but the corrected equations are employed throughout this study.

2 Theory

2.1 Rayleigh-wave Green's function

Following, e.g., Tsai (2011) and Boschi et al. (2019), we assume that surface-wave attenuation can be accounted for by replacing the equation governing the displacement of a lossless, stretched membrane with that of a damped membrane equation; we define the 2-D Green's function as the membrane response to impulsive initial velocity at the reference-frame origin (e.g. Boschi et al., 2019, App. A),

$$G^d_{2D}(x_1, x_2, \omega) = -\frac{i}{4\sqrt{2\pi}c^2} H_0^{(2)} \left(x \sqrt{\frac{\omega^2}{c^2} - \frac{2i\alpha\omega}{c}} \right), \tag{1}$$

where i, ω, c, and α denote imaginary unit, angular frequency, phase velocity and attenuation coefficient, respectively, $x = \sqrt{x_1^2 + x_2^2}$ is the distance between (x_1, x_2) and the impulsive source, and $H_0^{(2)}$ a zero-order Hankel function of the second kind. Eq. (1) is equivalent to eq. (8) of Boschi et al. (2019), except for a constant factor, dubbed P by Boschi et al. (2019) that served to keep track of the physical dimensions of G_{2D} and that is omitted here for simplicity. As shown by Boschi et al. (2019), provided that attenuation is relatively weak, i.e. $\alpha \ll \omega/c$, and/or the effects of near-field sources are negligible, eq. (1) can be reduced to the more convenient, approximate form

$$G^d_{2D}(x_1, x_2, \omega) \approx -\frac{i}{4\sqrt{2\pi}c^2} H_0^{(2)} \left(\frac{\omega x}{c} \right) \mathrm{e}^{-\alpha x}, \tag{2}$$

employed throughout the rest of the study.

2.2 Cross-correlation of ambient-noise recordings

By the properties of the Green's function, a signal of amplitude $h(\omega)$ and phase ϕ emitted at $\mathbf{x}$ and recorded at $\mathbf{x}_A$ reads $h(\omega) G^d_{2D}(\mathbf{x}_A, \mathbf{x}, \omega) \mathrm{e}^{i\phi}$. The vertical-component, Rayleigh-wave displacement associated with ambient noise can be expressed as a sum over the contributions of individual noise sources,

$$s(\mathbf{x}_A, \omega) = h(\omega) \sum_{j=1}^{N_S} G^d_{2D}(\mathbf{x}_A, \mathbf{x}_j, \omega) \mathrm{e}^{i\phi_j} \tag{3}$$

where N_S denotes the total number of sources, and the index j identifies the source. Eq. (3) is equivalent to (4) of Weemstra et al. (2014). It is also theoretically equivalent to eq. (17) of Boschi et al. (2019), where, however, the argument of the exponential is $i\omega\phi_j$ rather than $i\phi_j$.

Upon the assumption that the amplitude $h(\omega)$ is approximately the same for all noise sources, the cross correlation of seismic ambient noise recorded at two receivers $\mathbf{x}_A$, $\mathbf{x}_B$ can be written

$$\begin{aligned} s(\mathbf{x}_A,\omega)s^*(\mathbf{x}_B,\omega) =& |h(\omega)|^2 \left[\sum_{j=1}^{N_S} G^d_{2D}(\mathbf{x}_A,\mathbf{x}_j,\omega)\mathrm{e}^{i\phi_j}\right]\left[\sum_{k=1}^{N_S} G^{d*}_{2D}(\mathbf{x}_B,\mathbf{x}_k,\omega)\mathrm{e}^{-i\phi_k}\right] \\ =& |h(\omega)|^2 \left[\sum_{j=1}^{N_S} G^d_{2D}(\mathbf{x}_A,\mathbf{x}_j,\omega)G^{d*}_{2D}(\mathbf{x}_B,\mathbf{x}_j,\omega)\right. \\ &\left. + \sum_{j=1}^{N_S}\sum_{k=1,k\neq j}^{N_S} G^d_{2D}(\mathbf{x}_A,\mathbf{x}_j,\omega)G^{d*}_{2D}(\mathbf{x}_B,\mathbf{x}_k,\omega)\mathrm{e}^{i(\phi_j-\phi_k)}\right], \end{aligned} \tag{4}$$

where * denotes complex conjugation and the phases ϕ_1, ϕ_2, ϕ_3, ... are assumed to be random (uniformly distributed between 0 and 2π). Importantly, equation (4) can be simplified considering that if the recordings of noise span a sufficiently long time, or if a sufficiently large amount of uniformly distributed sources are present, the contribution of the "cross-terms" becomes negligible (e.g. Weemstra et al., 2014; Boschi & Weemstra, 2015, App. D), and the second term at the right-hand side (RHS) of (4) cancels out.

As shown by Boschi et al. (2019), the sum at the RHS of eq. (4) can then be replaced by an integral over the entire real plane, and combined with the reciprocity theorem for a lossy membrane (Section 2.2 of Boschi et al., 2019) to yield

$$s(\mathbf{x}_A,\omega)s^*(\mathbf{x}_B,\omega) \approx -\frac{|h(\omega)|^2\rho}{2\sqrt{2\pi}\alpha\omega c}\Im[G^d_{2D}(\mathbf{x}_A,\mathbf{x}_B,\omega)], \tag{5}$$

where the operator $\Im[\ldots]$ maps a complex number into its imaginary part, ρ is the surface density of noise sources, and $1/\sqrt{2\pi}$ arises from the correction of the algebraic error found in Boschi et al. (2019). Equation (5) stipulates that the amplitude of the cross-correlation of ambient-noise recordings carries the information on the surface-wave attenuation coefficient α. This means that α can be retrieved from the data, if inter-station phase velocity and spatial density and power spectral density of noise sources are known.

2.3 Power spectral density as normalization term

As shown by Boschi et al. (2019), the RHS of eq. (5) can be manipulated algebraically, to find an expression for the cross correlation of ambient noise where the source parameters $|h(\omega)|^2$ and ρ conveniently cancel out. In practice, the power spectral density of the signal recorded at any receiver $\mathbf{x}$ is first written as a sum over sources, analogously

to eqs. (3) and (4),

$$
\begin{aligned}
|h(\omega)|^2 &\approx \frac{16c^4}{\rho\, I(\alpha,\omega,c)}\ |s(\mathbf{x},\omega)|^2 \\
&\approx \frac{16c^4}{\rho\, I(\alpha,\omega,c)} < |s(\mathbf{x},\omega)|^2 >_{\mathbf{x}} ,
\end{aligned}
\tag{6}
$$

where $< \ldots >_{\mathbf{x}}$ denotes an average over all available receivers (the function h is assumed to be approximately the same for all noise sources) and

$$
I(\alpha,\omega,c) = \int_0^\infty \mathrm{d}r\ r\ \left| H_0^{(2)}\left(\frac{\omega r}{c}\right)\right|^2 \mathrm{e}^{-2\alpha r} \tag{7}
$$

can be evaluated numerically (our implementation exploits Gaussian quadrature, as provided by the SciPy Python library, Jones et al., 2001). Substituting eq. (6) into (5),

$$
\frac{s(\mathbf{x}_A,\omega)s^*(\mathbf{x}_B,\omega)}{< |s(\mathbf{x},\omega)|^2 >_{\mathbf{x}}} \approx \frac{c}{\omega\,\pi\, I(\alpha,\omega,c)} J_0\left(\frac{\omega|\mathbf{x}_A-\mathbf{x}_B|}{c}\right)\frac{\mathrm{e}^{-\alpha|\mathbf{x}_A-\mathbf{x}_B|}}{\alpha}, \tag{8}
$$

where J_0 denotes the zeroth-order Bessel function of the first kind (e.g. Abramowitz & Stegun, 1964). Eq. (8) is equivalent to (30) of Boschi et al. (2019), except for the mentioned algebraic error (a factor $1/\sqrt{2\pi}$) that here has been corrected. The left-hand side (LHS) of (8) represents the data, i.e. the normalized cross correlation of ambient noise records, while its RHS is our theoretical model. Importantly, as first pointed out by Boschi et al. (2019), $h(\omega)$ and ρ cancel out in the derivation that leads from eq. (5) to (8); it follows that eq. (8) can be used, through an inverse problem, to determine α from the data without prior knowledge of source density and frequency content (as long as both are approximately constant in space). In addition, if the LHS of eq. (8) is calculated as an ensemble-average of relatively small temporal windows with respect to the entire recording time, the normalization term $< |s(\mathbf{x},\omega)|^2 >_{\mathbf{x}}$ mitigates the effect of possible anomalous, ballistic signals like, e.g., large or nearby earthquakes (Boschi et al., 2019). This is often necessary when working with real data and commonly accomplished by one-bit normalization or spectral whitening (e.g. Bensen et al., 2007); these empirical normalization terms, however, albeit useful for retrieving phase or group velocities since they leave the phase of the cross correlations unchanged, are doomed to introduce a bias in their amplitude and therefore in the resulting estimates of α (Weemstra et al., 2014).

We emphasize that eq. (8) only holds if all our theoretical assumptions on the nature of ambient noise and propagation medium are valid, and that in such scenario both its RHS and LHS are purely real. When working with observational data these assumptions are not strictly verified, and the numerical value of the LHS of (8), as obtained from

the data, is only approximately equal to the theoretical model at the RHS; this is why in ambient-noise literature empirical Green's functions commonly show a non-zero imaginary part, and are referred to as "complex coherency" (e.g. Weemstra et al., 2014).

3 Inverse problem

Equation (8) allows to formulate an inverse problem to determine α from cross correlations of recorded ambient signal. Because equation (8) holds for all station pairs, it is desirable that the cost function be related to the weighted sum of the squared differences between LHS and RHS of (8), calculated for each station pair; since the RHS of (8) is an oscillatory function of ω (through the Bessel function J_0), and α only affects its envelope but not its oscillations (e.g. Prieto et al., 2009; Boschi et al., 2019), we introduce the envelope function env to define the cost function

$$C(\alpha,\omega) = \sum_{i,j} |\mathbf{x}_i - \mathbf{x}_j|^2 \sum_k \left| \mathrm{env}\left[\frac{s(\mathbf{x}_i,\omega_k)s^*(\mathbf{x}_j,\omega_k)}{<|s(\mathbf{x},\omega_k)|^2>_{\mathbf{x}}}\right] - \mathrm{env}\left[\frac{c_{ij}(\omega_k)}{\omega_k \pi I[\alpha,\omega_k,c_{ij}(\omega_k)]} J_0\left(\frac{\omega_k|\mathbf{x}_i-\mathbf{x}_j|}{c_{ij}(\omega_k)}\right)\frac{\mathrm{e}^{-\alpha|\mathbf{x}_i-\mathbf{x}_j|}}{\alpha}\right]\right|^2, \tag{9}$$

where the weight $|\mathbf{x}_i-\mathbf{x}_j|^2$ is chosen based on the fact that larger inter-station distances are associated with smaller amplitudes of the cross correlations, due to geometrical spreading, which would result in smaller absolute values of misfit if not weighted accordingly. The minimum of $C(\alpha,\omega)$ can then be found through some form of "grid-search" over α, for a discrete set of values of ω. The formula (9) for $C(\alpha,\omega)$ was selected after experimenting several other options, as partly documented in Boschi et al. (2019).

The summation over receiver pairs i,j at the RHS of (9) involves all the available receivers and, if the array has good azimuthal coverage, most azimuths of wave propagation. Minimizing $C(\alpha,\omega)$ therefore involves finding one function $\alpha(\omega)$ such that a good fit is simultaneously achieved at all azimuths; this has a regularizing effect on the inversion, and should reduce the effects of non-homogeneity in azimuthal source distribution.

Previous studies (e.g. Prieto et al., 2009; Weemstra et al., 2013) formulated inverse problems whose data consisted of azimuthally averaged cross correlations calculated over several station pairs; this was based on the idea that azimuthal averaging is necessary to retrieve a reliable, purely real empirical Green's function (e.g. Asten, 2006; Yokoi & Margaryan, 2008). It has been noticed, however, that this approach might not be equally effective in estimating attenuation. In fact, slightly different inter-station distances or

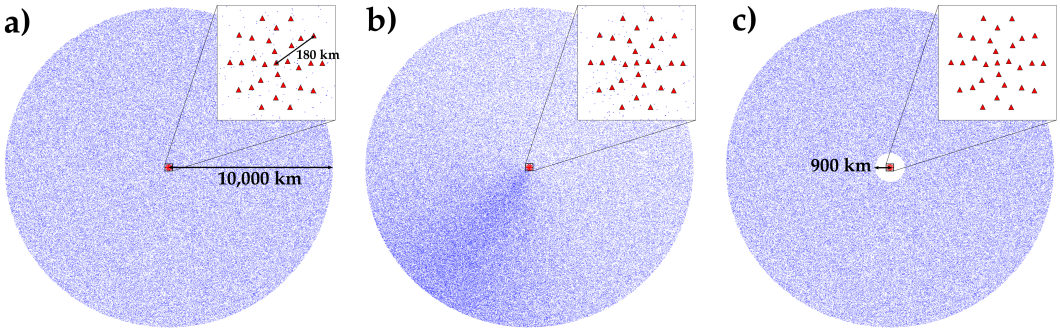

Figure 1. Sources (blue dots) and stations (red triangles) used for simulating seismic noise. (a), (b), and (c) indicate the source distributions used in Sections 4.1 (uniform source distribution), 4.2 (azimuth-dependent source density), and 4.3 (no sources in the near field), respectively.

a laterally inhomogeneous phase velocity would introduce a phase offset of the cross correlations involved in the average; in turn, this would result in a "attenuation-like" effect (Menon et al., 2014), i.e. in a fictitious decrease of the amplitude of the averaged coherency and thus in a bias of the estimates of α.

4 Numerical Validation

We simulate ambient signal via a very large number of randomly distributed, uncorrelated point sources. We next solve an inverse problem, as described above, to retrieve the theoretical value of α; we also verify numerically the emergence of coherent signal in the cross correlations due to cancellation of cross-terms in eq. (4), and the validity of eq. (6), which relates recorded ambient noise and the frequency spectrum of ambient-noise sources. The simulation is carried out in three different experimental setups. First, we present the "ideal" case of a spatially uniform distribution of sources. Since real-world ambient sources are not distributed uniformly (e.g. Hillers et al., 2012), we next discuss the case of an azimuthally heterogeneous source distribution. Finally, we show the results obtained through a source distribution characterized by absence of noise sources in the vicinity of the receivers.

4.1 Uniform source distribution

200,000 point sources are randomly distributed both in the near and far field of 29 receivers, within a circle of radius $R = 1 \times 10^7$ m (Fig. 1a); source locations are de-

fined by their polar coordinates θ, r with respect to one station located at the center of the array; random values of θ between 0 and 2π, and of n between 0 and 1 are generated, and $r = R\sqrt{n}$ (the square root results in a linear growth of the number of sources with increasing distance from the center of the circle, hence constant source density in space). The receivers are randomly deployed in the central part of such distribution on 4 concentric circles, with radii of 45, 90, 135, and 180×10^3 m.

4.1.1 Simulation of seismic ambient noise

Synthetic data are generated using a constant attenuation coefficient, different in each of two experiments (corresponding to different models of attenuation, i.e. $\alpha = 5\times 10^{-7}$ and $\alpha = 1 \times 10^{-6}$ m^{-1}), and a fixed, frequency-dependent phase velocity $c = c(\omega)$; the phase velocity decreases monotonously (and almost linearly) between 0.05 Hz (where $c = 3526$ ms^{-1}) and 0.25 Hz (2851 ms^{-1}), with a slight kink around 0.07 Hz where its derivative with respect to time decreases with increasing frequency. We consider these values to be realistic, based, e.g., on Mitchell (1995) and Ekström (2014).

Each numerical test consisted of 25,000 *realizations* (Cupillard & Capdeville, 2010; Weemstra et al., 2015). At each realization every source emits an independent signal of constant amplitude $h(\omega) = 1$ and random phase ϕ between 0 and 2π. The displacement at the receivers due to the impulsive sources is computed, at each realization, via eq. (3); the LHS of eq. (8) is then implemented for a pair of stations $\mathbf{x}_A$, $\mathbf{x}_B$ by ensemble-averaging the normalized cross-correlations (calculated for each realization k) over N_R realizations,

$$\frac{s(\mathbf{x}_A,\omega)s^*(\mathbf{x}_B,\omega)}{<|s(\mathbf{x},\omega)|^2>_{\mathbf{x}}} = \frac{1}{N_R}\sum_{k=1}^{N_R}\frac{s_k(\mathbf{x}_A,\omega)\,s_k^*(\mathbf{x}_B,\omega)}{<|s_k(\mathbf{x},\omega)|^2>_{\mathbf{x}}}. \tag{10}$$

4.1.2 Cancellation of cross-terms and source spectrum

Real and imaginary parts of normalized cross-correlations, calculated by ensemble-averaging over an increasing number of realizations as in the RHS of eq. (10), are shown in Fig. 2 for a pair of receivers with inter-station distance of 67,600 m. For both chosen values of α, the increase in the smoothness of the real parts and the decrease in the amplitude of the imaginary parts with the number of realizations brings evidence of the cancellation of cross-terms of eq. (4). Fig. 2 also shows that the real coherency obtained when $\alpha = 5 \times 10^{-7}$ m^{-1} is slightly larger than when $\alpha = 1 \times 10^{-6}$ m^{-1}, as expected for a less attenuating medium.

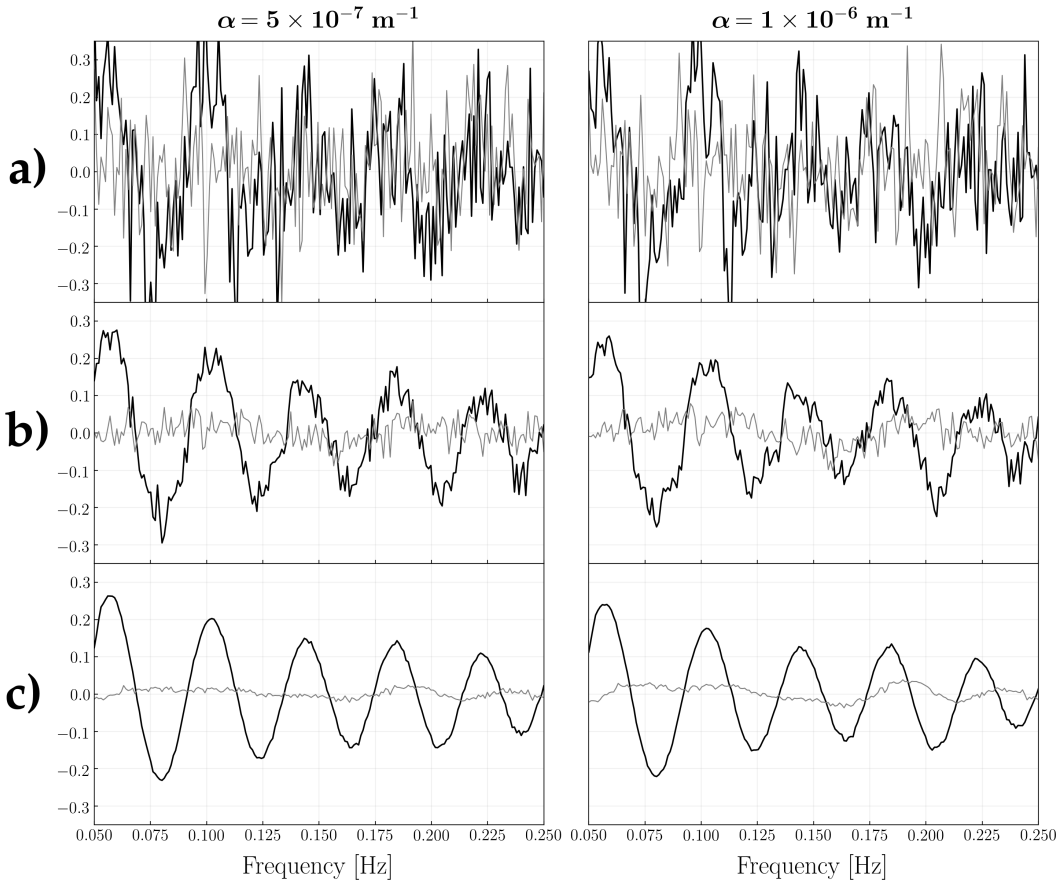

Figure 2. Real (black) and imaginary (gray) parts of LHS of eq. (10), obtained for a pair of receivers with inter-station distance of 67,600 m by ensemble-averaging over (a) 25, (b) 500, and (c) 25,000 realizations. Results are shown for both values of α used in the experimental setup of Section 4.1 (uniform source distribution).

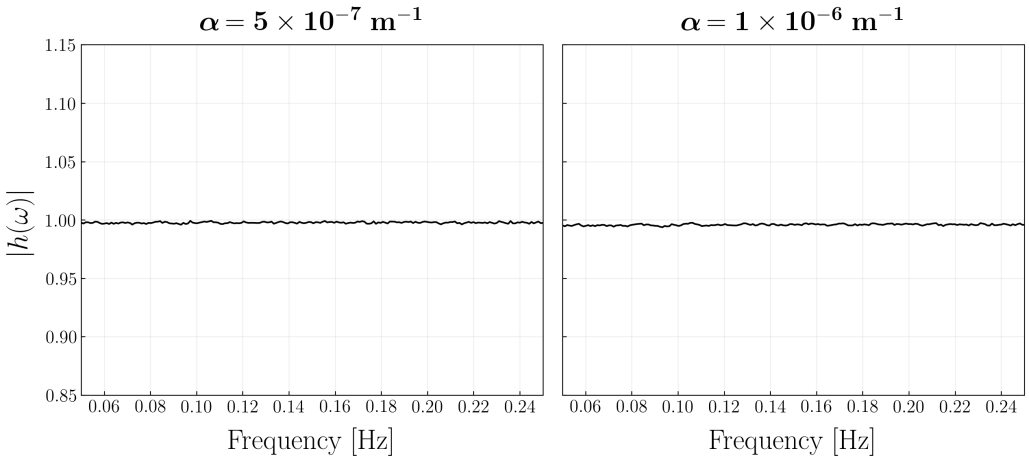

Figure 3. Absolute value of source amplitude $|h(\omega)|$, retrieved from synthetic data for both values of α used in the experimental setup of Section 4.1 (uniform source distribution). $|h(\omega)|$ is calculated by taking the square root of the RHS of eq. (6). Note that, in both numerical tests, noise has been simulated using a constant $h(\omega) = 1$.

Equation (6) indicates that it is possible to retrieve the source spectrum $h(\omega)$ if source density ρ and attenuation coefficient are known, provided that h is approximately the same for all sources (Section 2.2); we show in Fig. 3 that, implementing eq. (6), $h(\omega) = 1$ is retrieved correctly, at least to the second decimal digit, for both values of α. This result validates numerically the derivation of eq. (6), first shown by Boschi et al. (2019) and summarized here in Section 2.

4.1.3 Retrieval of the attenuation coefficient

After a suite of preliminary tests, we chose to implement the envelope function by fitting a combination of cubic splines (De Boor et al., 1978) to the maxima of the absolute value of their arguments, and then smoothing them by means of a running average performed with a Savitzky-Golay filter (Savitzky & Golay, 1964). Smoothing is motivated by the fact that, if the anelastic properties of the Earth are assumed to be smoothly varying with depth, the same behavior is expected for the amplitude of adjacent peaks of the real coherency; abrupt amplitude variations are ascribed to a non-perfectly diffuse wavefield or, in the case of real recordings, simply to noisiness of the empirical Green's function.

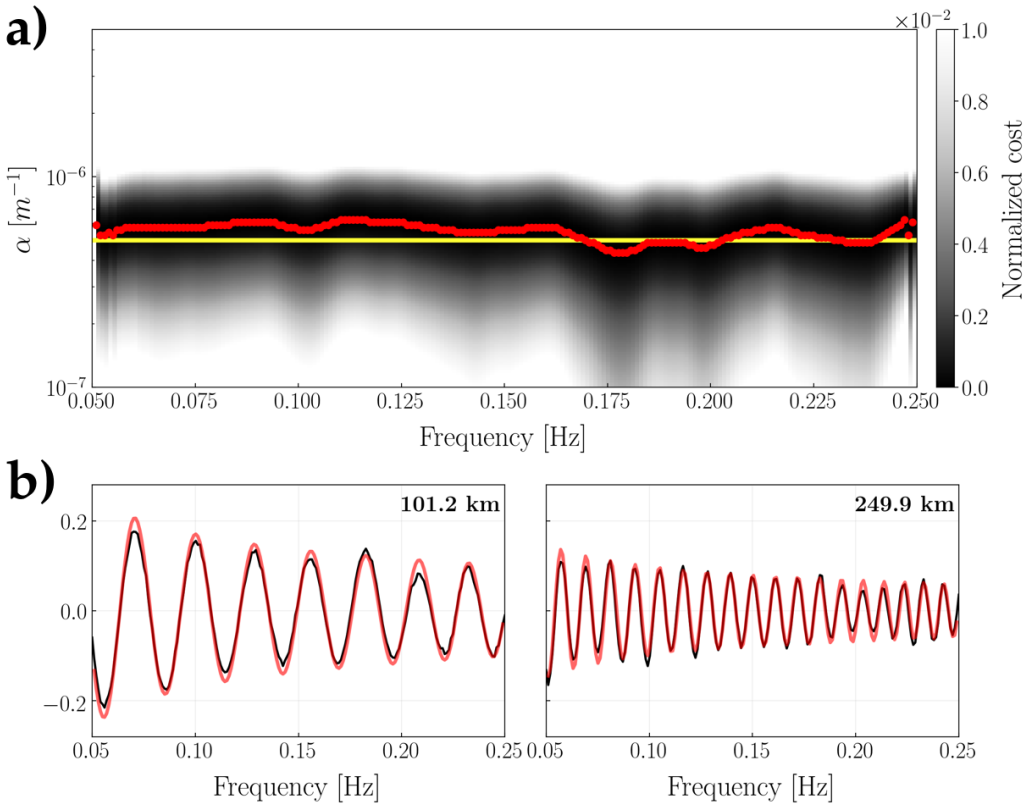

Figure 4. (a) Cost function $C(\alpha, \omega)$ associated with the numerical experiment of Section 4.1 (uniform source distribution) shown (after normalization) as a function of attenuation coefficient and frequency. Red dots mark the values of α for which $C(\alpha, \omega)$ is minimized at each frequency; the yellow line indicates the assumed attenuation model $\alpha = 5 \times 10^{-7}$ m^{-1}, used for generating synthetic recordings. (b) Normalized cross correlations (black) fitted by the model (red) obtained by substituting into eq. (8) the values of $\alpha(\omega)$ which minimize $C(\alpha, \omega)$. Within each subplot, the inter-station distance is indicated on the upper right.

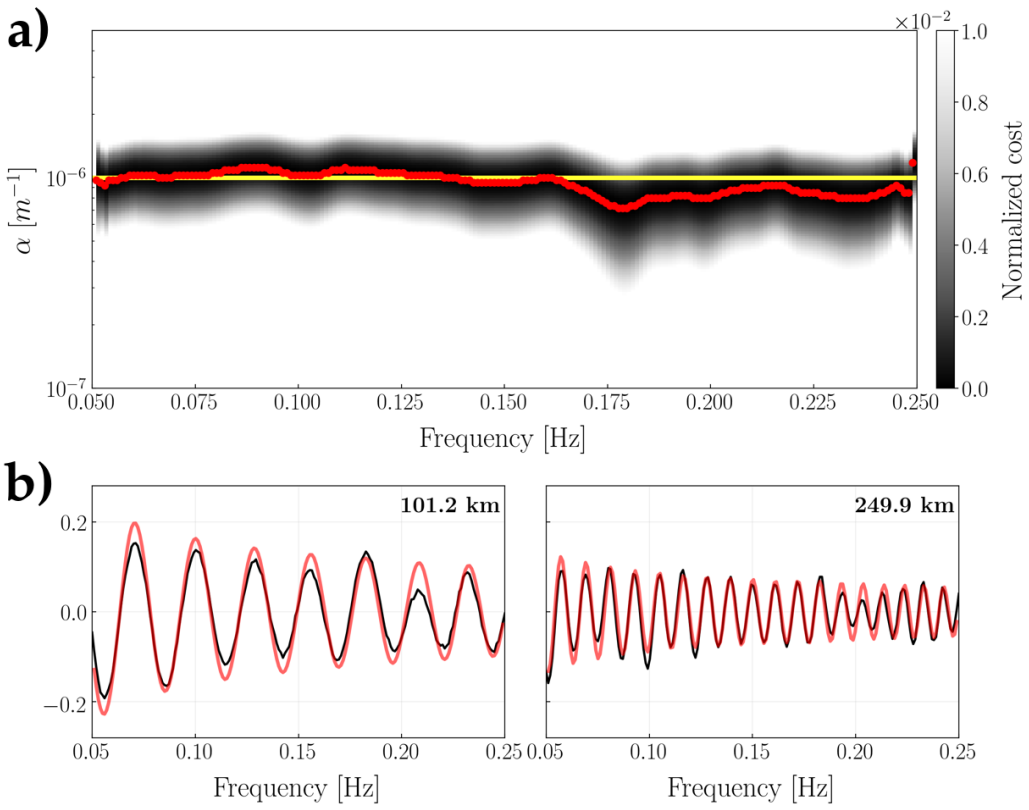

Figure 5. Same as Fig. 4 but but synthetic data were obtained assuming constant attenuation $\alpha = 1 \times 10^{-6}$ m^{-1}.

The cost function $C(\alpha, \omega)$ is evaluated by means of a 1-D grid search over 275 values of α evenly spaced on logarithmic scale between 5×10^{-8} and 1×10^{-4} m^{-1}. Figs. 4a and 5a show that, on average, the minima of $C(\alpha, \omega)$ correspond to the values of α used for generating synthetic recordings. The datafit obtained by substituting into eq. (8) the values of $\alpha(\omega)$ retrieved by minimizing the cost function $C(\alpha, \omega)$ is shown for both numerical tests in subpanel (b), and can be considered good at all the investigated interstation distances.

4.2 Azimuth-dependent source density

In a second numerical simulation, the spatial distribution of sources is modified while all other parameters are left unchanged; the nonuniformity in the source distribution is implemented by generating random values k between 0 and 2π, and obtaining source azimuth from k via the formula $\theta = k + \frac{1}{2}\cos(k - \frac{4}{5}\pi)$; $r = R\sqrt{n}$, with $0 \leq n < 1$, as above. The spatial distribution of sources thus obtained is characterized by a higher density to the South-West of the array (Fig. 1b). Synthetic data are generated using the phase velocity $c = c(\omega)$ of Section 4.1, and a constant attenuation coefficient $\alpha = 1 \times 10^{-6}$ m^{-1}. In analogy with the first numerical test, seismic ambient noise has been simulated for 25,000 realizations, with $h(\omega) = 1$ and random phase ϕ between 0 and 2π.

The cancellation of cross-terms, as inferred from the real and imaginary parts of normalized cross correlations, is illustrated in Fig. 6 for the same pair of receivers employed in Fig. 2. As expected for a nonuniform distribution of noise sources (see Section 2), the imaginary part of Fig. 6c is larger than that shown in Fig. 2c. The source spectrum $h(\omega)$ is shown in Fig. 7 to be less accurate than that obtained in the ideal case of a uniform source distribution (Fig. 3). This result should not surprise, as the ambient noise used within the experimental setup in question has been simulated in a fashion that violated the assumptions made in Section 2. On the other hand, the average of $|h(\omega)|$ over frequency is equal to 0.987 and only 1.3% smaller than the true $h(\omega)$ employed for generating synthetic recordings. This indicates that eq. (6) allows to estimate the average source spectrum to a relatively high degree of accuracy, even if the assumption of diffuse ambient field is not exactly met.

Following the same procedure as in Section 4.1.3, we obtained minima of $C(\alpha, \omega)$ which correspond, on average, to the true attenuation $\alpha = 1 \times 10^{-6}$ m^{-1} (Fig. 8a).

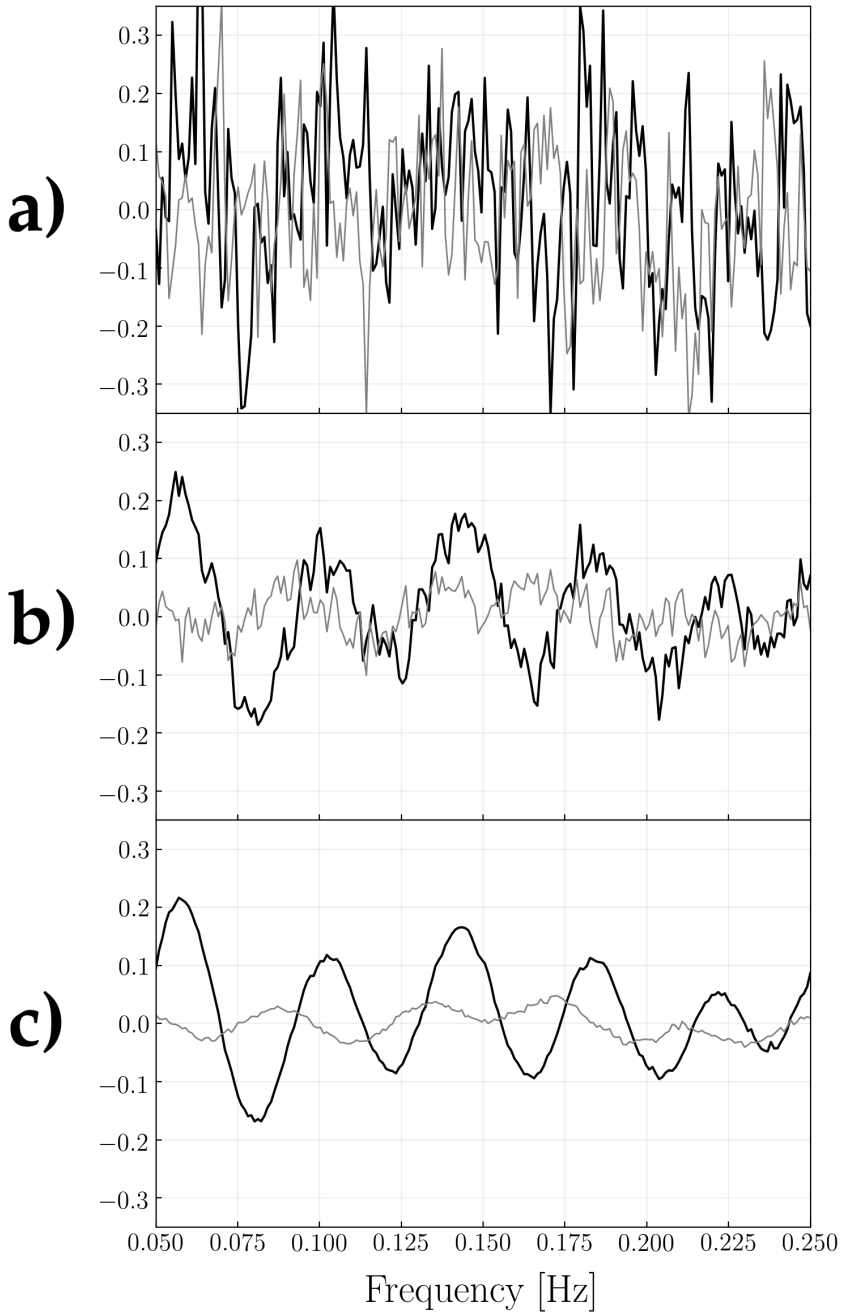

Figure 6. Same as Fig. 2, but obtained through the experimental setup of Section 4.2 (azimuth-dependent source density). Inter-station distance is 67,600 m. Real (black) and imaginary (gray) parts of the cross correlation obtained from the ensemble-average over 25, 500, and 25,000 realizations are shown in subpanel (a), (b), and (c), respectively.

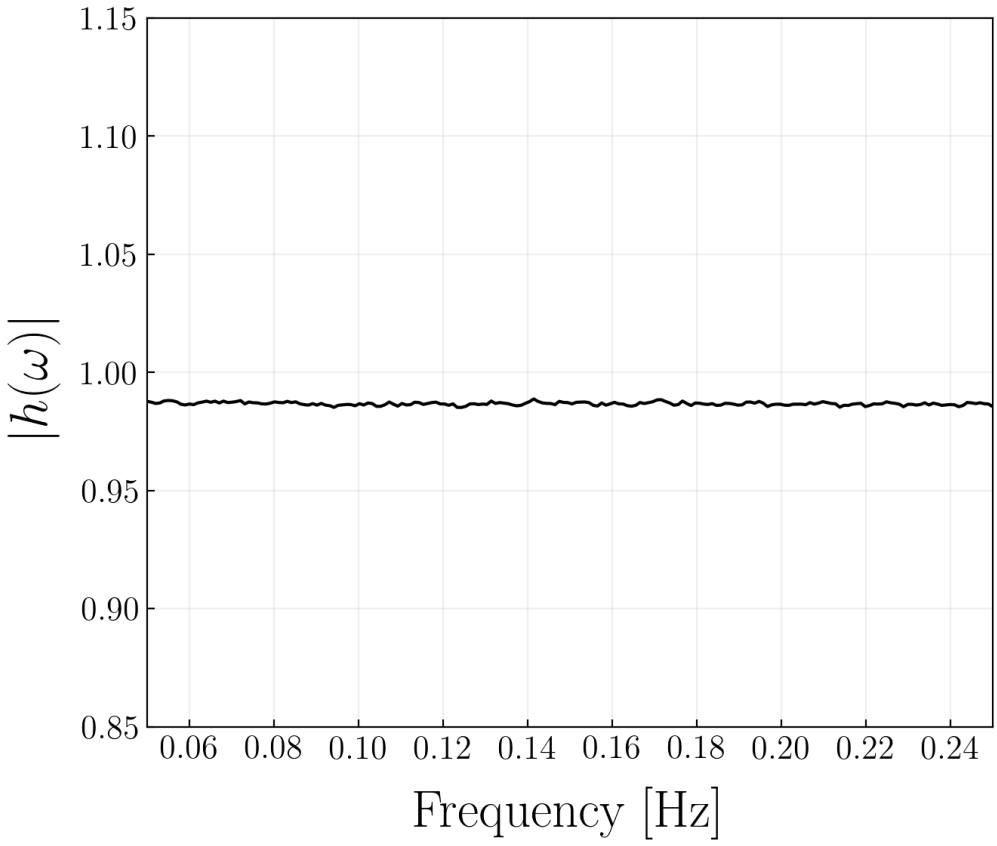

Figure 7. Same as Fig. 3, but obtained through the experimental setup of Section 4.2 (azimuth-dependent source density). The absolute value of source amplitude $|h(\omega)|$ has been retrieved from synthetic data by taking the square root of the RHS of eq. (6). Note that noise has been simulated using a constant $h(\omega) = 1$.

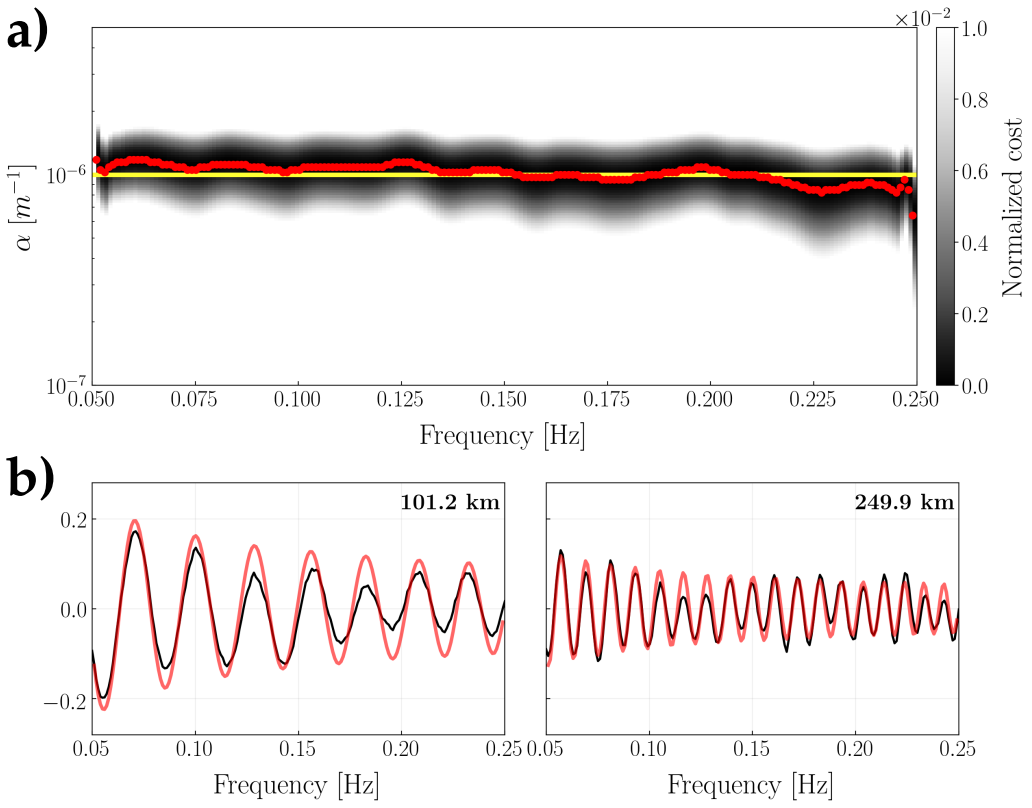

Figure 8. (a) Cost function $C(\alpha, \omega)$ associated with the numerical experiment of Section 4.2 (azimuth-dependent source density) shown (after normalization) as a function of attenuation coefficient and frequency. Red dots mark the values of α for which $C(\alpha, \omega)$ is minimized at each frequency; the yellow line indicates the assumed attenuation model $\alpha = 1 \times 10^{-6}$ m^{-1}, used for generating synthetic recordings. (b) Normalized cross correlations (black) fitted by the model (red) obtained by substituting into eq. (8) the values of $\alpha(\omega)$ which minimize $C(\alpha, \omega)$. Within each subplot, the inter-station distance is indicated on the upper right.

The datafit obtained by substituting into eq. (8) the best values of $\alpha(\omega)$ is shown in Fig. 8b for the same station pairs employed in Figs. 4b and 5b.

The above results show that, even if the spatial distribution of noise sources is slightly nonuniform, the value of $\alpha(\omega)$ can be reconstructed correctly from the cross correlation of ambient noise: we have achieved this, as anticipated, by neglecting possible lateral heterogeneities in $\alpha(\omega)$, and minimizing a cost function where as many azimuths of propagation as possible are simultaneously included. In practice, this means that surface-wave attenuation can be estimated based on ambient noise, even when the noise field is not exactly diffuse. This is indeed the case in most practical applications.

4.3 Absence of near-field sources

Sources are uniformly distributed in space, as in Section 4.1, but starting at a minimum distance of 900×10^3 m from the station that defines the center of the array (Fig. 1c). We implement 25,000 realizations with the same phase velocity $c = c(\omega)$ as before, attenuation $\alpha = 1 \times 10^{-6}$ m^{-1}, and $h(\omega) = 1$. Again, a random phase ϕ between 0 and 2π, newly generated at each realization, is assigned to each source.

In analogy with the experiments above, we verified the emergence of coherent signal in the cross correlations due to the cancellation of cross-terms. The amplitude of the imaginary part of the cross-spectra, not shown here for brevity, is similar to that obtained for an azimuthally heterogeneous source distribution (see Fig. 6). On the other hand, the real part is systematically larger than in the uniform-source-distribution case. As in Sections 4.1 and 4.2, we then used the synthetic data to quantify source spectrum $h(\omega)$ and attenuation $\alpha(\omega)$, as illustrated in Figs. 9 and 10a. We infer from the results thus obtained that the absence of near-field sources leads to a significant underestimate of both $h(\omega)$ and $\alpha(\omega)$ (the latter by a factor of about 5), in agreement with the theoretical findings of Tsai (2011).

5 Preliminary application to a small subset of USArray

5.1 Data set

We downloaded continuous vertical-component recordings from 33 broad-band receivers belonging to the transportable component of the USArray network (Fig. 11) and operating between February 2007 and August 2008. Each seismogram has been demeaned,

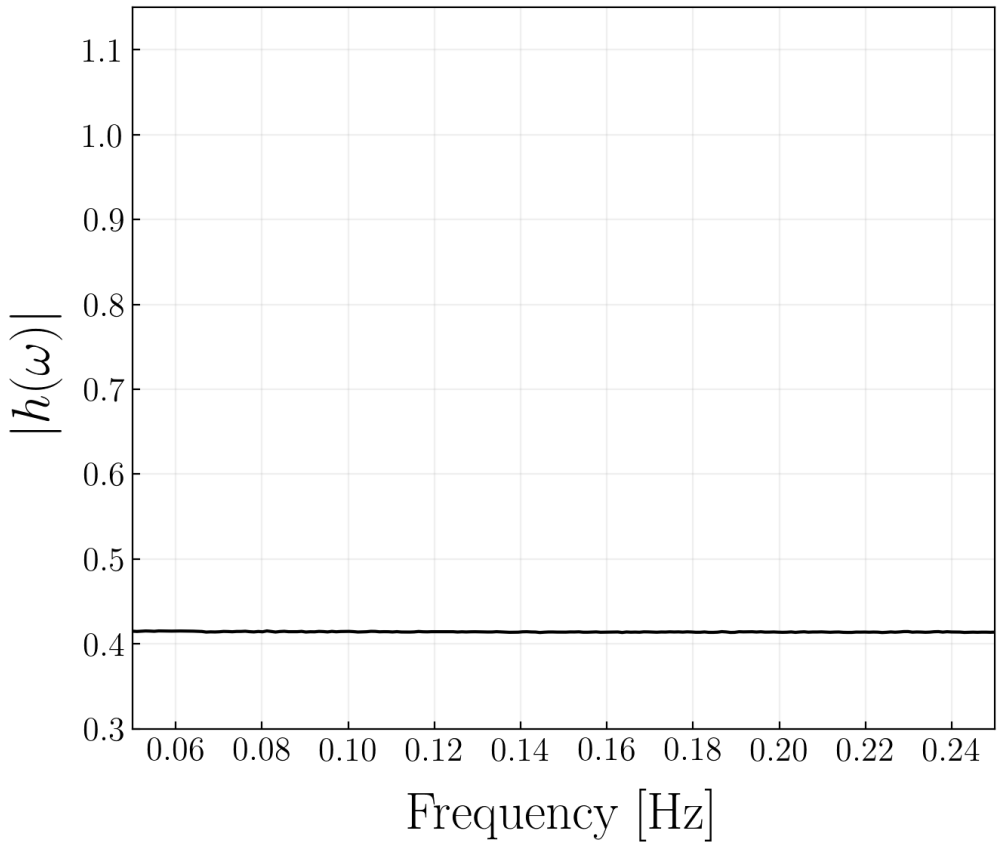

Figure 9. Same as Figs. 3 and 7, but obtained through the experimental setup of Section 4.3 (no sources in the near field). The absolute value of source amplitude $|h(\omega)|$ has been retrieved from synthetic data by taking the square root of the RHS of eq. (6). Note that noise has been simulated using a constant $h(\omega) = 1$.

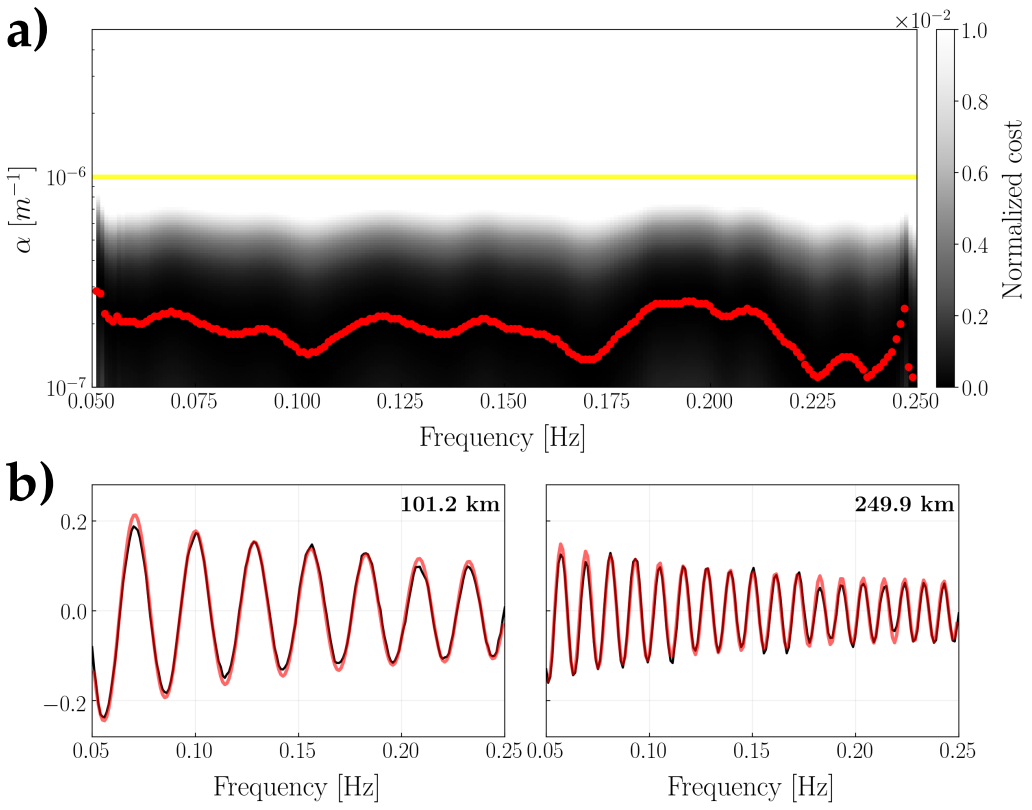

Figure 10. (a) Cost function $C(\alpha, \omega)$ associated with the numerical experiment of Section 4.3 (no sources in the near field) shown (after normalization) as a function of attenuation coefficient and frequency. Red dots mark the values of α for which $C(\alpha, \omega)$ is minimized at each frequency; the yellow line indicates the assumed attenuation model $\alpha = 1 \times 10^{-6}$ m^{-1}, used for generating synthetic recordings. (b) Normalized cross correlations (black) fitted by the model (red) obtained by substituting into eq. (8) the values of $\alpha(\omega)$ which minimize $C(\alpha, \omega)$. Within each subplot, the inter-station distance is indicated on the upper right.

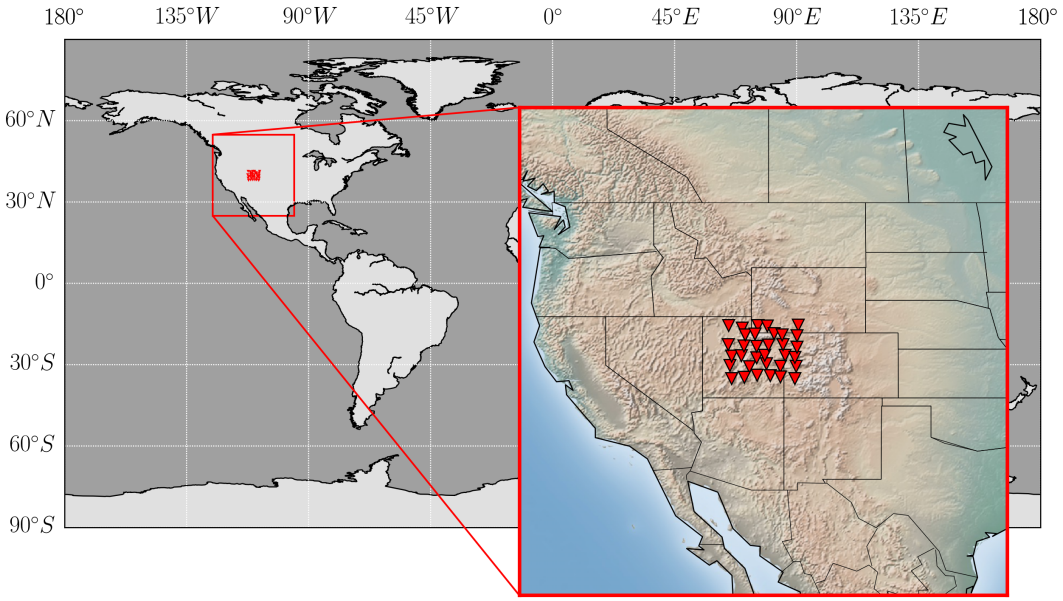

Figure 11. Seismic stations (red triangles) from the USArray project transportable network, forming the data set described in Section 5.1

detrended, tapered (5%), and bandpass-filtered between 0.01 and 0.5 Hz before deconvolving the instrumental response to displacement; eventual gaps present in the waveforms have been zero-padded, in order to obtain continuous time-series.

The data thus collected allowed us to determine 509 empirical Green's functions (i.e. LHS of eq. 8), by ensemble averaging cross-spectra calculated in 6-hour long windows. To reduce the effects of temporal variability and/or seasonality of noise sources, we only used pairs of receivers that recorded simultaneously for more than 9 months. The normalized cross-correlations served us to retrieve Rayleigh-wave dispersion curves in the frequency range between 0.3 and 0.04 Hz, by means of Kästle et al. (2016)'s automated algorithm.

5.2 Signal-to-noise ratio

We show in Fig. 12 four normalized cross-correlations associated with receiver pairs that are characterized by significantly different inter-station distances. The fact that the imaginary part of the empirical Green's function is nonzero indicates that the assumptions described in Section 2 are not exactly met by our observations, because the ambient wavefield is not perfectly diffuse (Boschi & Weemstra, 2015). To estimate possi-

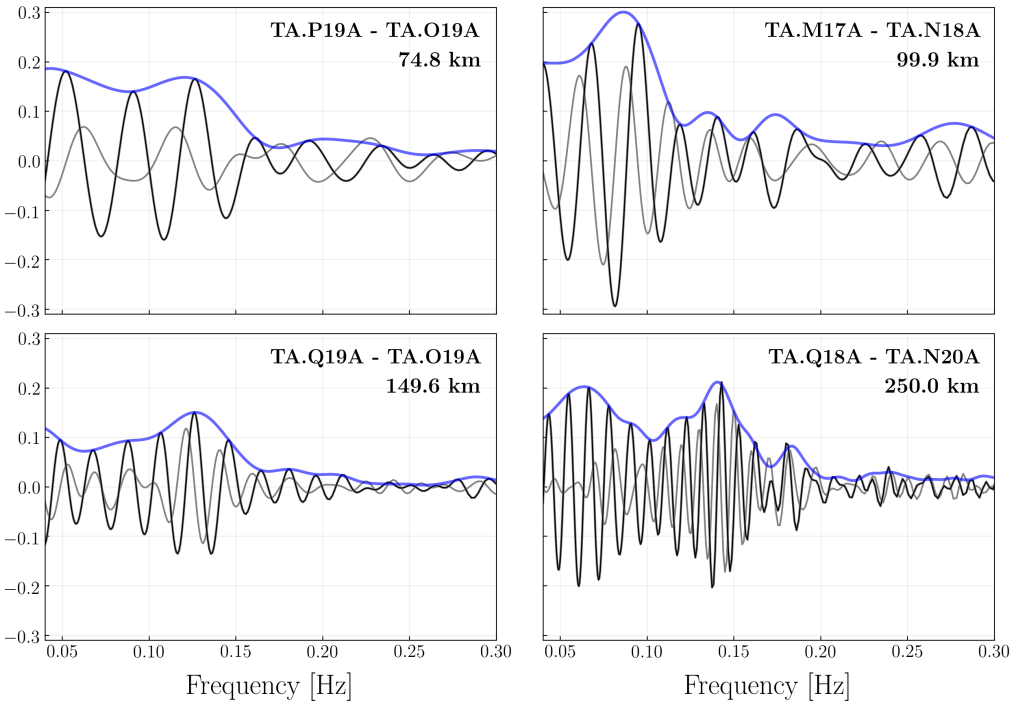

Figure 12. Envelopes (blue), real (black), and imaginary part (gray) of the normalized cross correlations calculated for 4 different pairs of receivers. Within each subplot, station codes and inter-station distance are indicated on the upper right.

ble azimuthal biases introduced in the recordings, we therefore performed a signal-to-noise ratio (SNR) analysis; this allows to assess the presence of preferential directionality of the noise sources, thus giving indication of the diffusivity of the ambient wavefield. The analysis has been carried out by narrow-bandpass filtering and inverse-Fourier transforming all the available cross-spectra; in the time domain, the SNR is then calculated by taking the ratio of the maximum signal amplitude to the maximum of the trailing noise (e.g. Yang & Ritzwoller, 2008; Kästle et al., 2016). In this analysis, "signal" refers to the segment of ambient-noise cross correlation that contains the Rayleigh-wave fundamental mode propagating between the two relevant receivers. In practice, this corresponds to the temporal window identified by a velocity range between 2 and 4.2 km s^{-1}.

We infer from visual inspection of the results thus obtained (Fig. 13) that the ambient field is relatively isotropic within the study area, at least in the frequency band associated with the *primary* microseisms, i.e. from ∼10 s to ∼20 s period, peaking at ∼14 s (e.g. Friedrich et al., 1998). The SNR at the central periods of the *secondary* microseisms band is characterized by a relative maximum along the SW-NE direction (see the periods of 6 and 8 s in Fig. 13). This was also noted by, e.g., Landès et al. (2010) and Tian and Ritzwoller (2015), who identified in the central Pacific Ocean a probable source region of secondary microseisms (see Fig. 8 of Tian & Ritzwoller, 2015). However, the preferential directionality of noise emerging from our SNR analysis is less prominent. This result confirms the known seasonality of ambient noise sources (e.g. Tanimoto et al., 2006; Hillers et al., 2012). Ensemble-averaging over several months of recordings reduces this effect, and the resulting empirical Green's functions better approximate those that would be obtained from a truly diffuse ambient field.

5.3 Results and discussion

To retrieve the attenuation coefficient within the study area, we performed a 1-D grid search over 275 values of α evenly spaced on a logarithmic scale between 5×10^{-8} and 1×10^{-4} m^{-1}; in analogy with Section 4, minimization of the cost function $C(\alpha,\omega)$ allowed us to identify the best fitting value of α at each frequency, as shown in Fig. 14a. The datafit obtained by substituting into eq. (8) the values of $\alpha(\omega)$ which minimize $C(\alpha,\omega)$ is shown in Fig. 14b for four different station pairs.

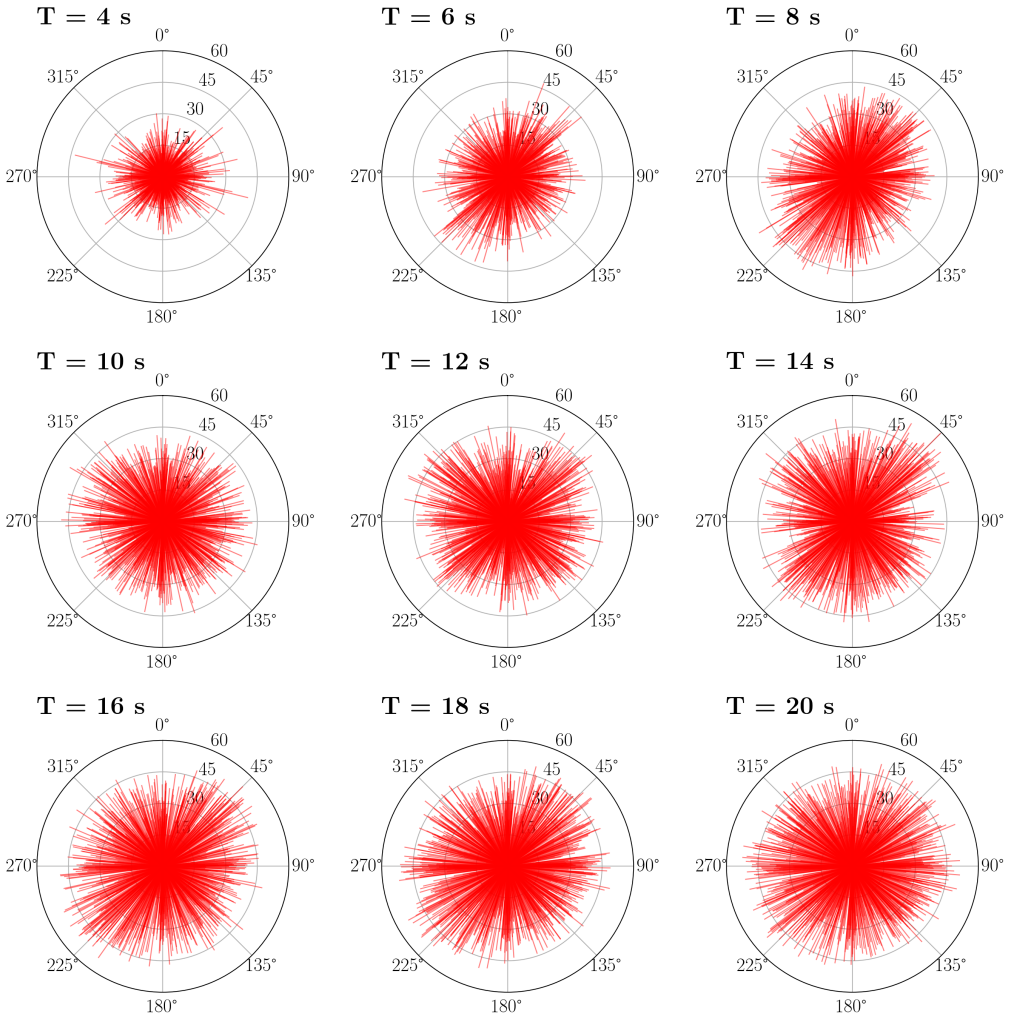

Figure 13. Signal-to-noise ratio at different periods as a function of azimuth, as inferred from the normalized cross-correlations. The length of the red segments is determined by the value of SNR, while their orientation coincides with the azimuth/back-azimuth of the respective station pair. 0° corresponds to the north, 90° to the east, etc.

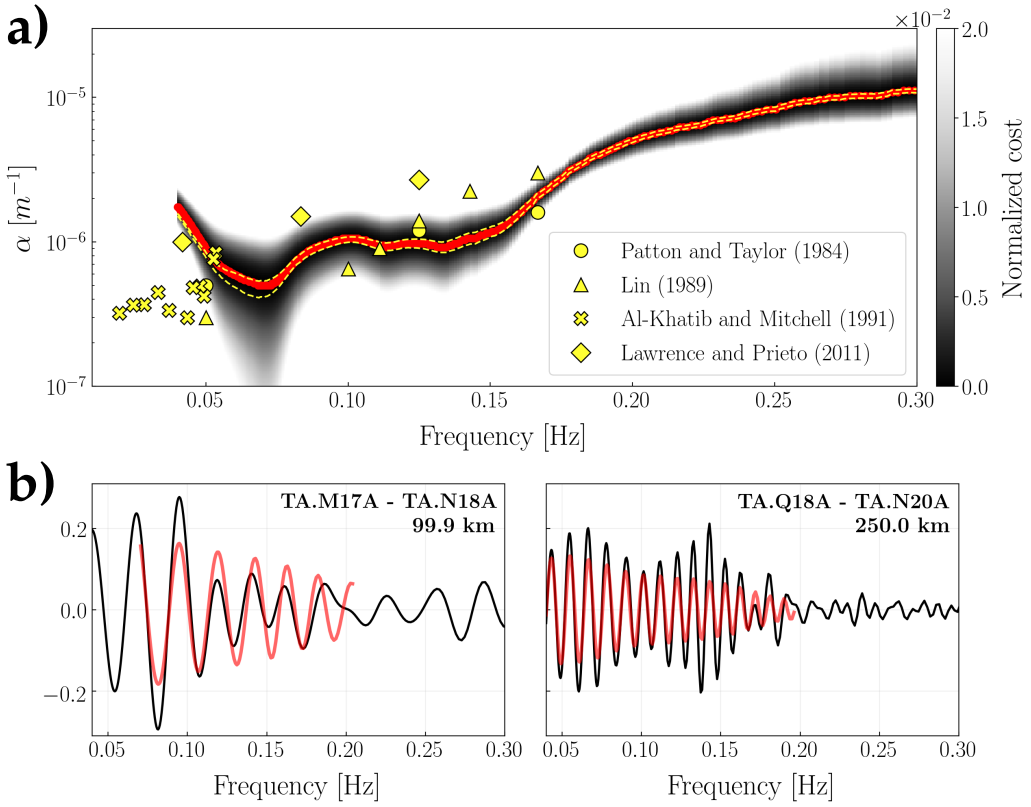

Figure 14. (a) Cost function $C(\alpha, \omega)$ shown (after normalization) as a function of attenuation coefficient and frequency. The red dots mark the values of α for which $C(\alpha, \omega)$ is minimized at each frequency. The dashed yellow line is calculated, at each frequency, as $\mu \pm \sigma$, where μ and σ indicate mean and standard deviation of the values of α retrieved from the bootstrap analysis. Yellow marks indicate average measurements of alpha as collected in the vicinity of the study area in previous studies (i.e. Patton & Taylor, 1984; Lin, 1989; Al-Khatib & Mitchell, 1991; Lawrence & Prieto, 2011, as specified in the legend). (b) Normalized cross correlations (black) fitted by the model (red) obtained by substituting into eq. (8) the values of $\alpha(\omega)$ which minimize the cost function $C(\alpha, \omega)$. The datafit is shown for the same station pairs of Fig. 12. Within each subplot, station codes and inter-station distance are indicated on the upper right. The frequency band spanned by the models is determined by the availability of phase-velocity measurements.

To assess the uncertainty of this result, we performed a bootstrap analysis: we minimized $C(\alpha, \omega)$ 100 times, randomly removing 20 per cent of the cross correlations at each iteration. The resulting set of $\alpha(\omega)$ allowed us to estimate the statistical robustness of the values of attenuation retrieved from the inversion; in this regard, its average approximately coincides with the red curve showed in Fig. 14a, with the largest differences being $\sim 2{\times}10^{-7}$ m^{-1} at 0.04 Hz, whereas its standard deviation is at least one order of magnitude smaller than the mean values at all frequencies, varying from 3.18×10^{-7} m^{-1} at 0.3 Hz to 3.42×10^{-8} m^{-1} at 0.04 Hz.

Fig. 14a also shows that our estimates of α, and their dependence on ω, are similar to those found by Patton and Taylor (1984) and Lin (1989) from earthquake-based Rayleigh waves; at the same frequencies, the values proposed by Lawrence and Prieto (2011) based on seismic ambient noise are slightly larger. At higher frequencies (> 0.2 Hz), our measurements fit well those that would be obtained by linearly extrapolating the values of α reported by Lin (1989). At frequencies lower than $\sim$0.065 Hz (periods $\gtrsim 16$ s), on the contrary, we observe an increase of α, in disagreement with what reported in previous studies. We ascribe this to the lack of ambient-noise signal near the upper boundary of the secondary microseism energy band ($\sim$20 s). This would result, for most station pairs, in a decrease of the envelopes of the empirical Green's functions (Fig. 12) and the subsequent overestimate of α in our inversion. We infer that our estimates of α at such frequencies are not reliable; in the rest of the frequency range under study our observations appear to be in good continuity with those measured from earthquake-based Rayleigh waves by Al-Khatib and Mitchell (1991).

As shown in Section 4.3, attenuation is significantly underestimated if the distribution of noise sources is limited to the far field of the receivers. If this was the case in the real world, we should observe a significant discrepancy between ambient-noise- and earthquake-based attenuation estimates, the latter being systematically larger than the former. Our estimates, however, are compatible with those obtained from earthquakes by previous authors in the area of interest. This suggests that ambient noise in the frequency range relevant to this study might be generated in the relative vicinity of our receiver array, i.e. within the continent; alternatively, other complex non-homogeneities in the distribution of noise sources might compensate for the lack of sources in the near field. This issue merits further attention, but is beyond the scope of our current study.

6 Conclusions

We have validated numerically the method proposed by Boschi et al. (2019) to quantify the attenuation of Rayleigh waves from the cross correlation of seismic ambient noise. We achieved this by simulating the displacement associated with 200,000 impulsive sources and recorded by 29 receivers. In all our simulations, we imposed realistic values of attenuation ($\alpha = 5 \times 10^{-7}$ m^{-1} and $\alpha = 1 \times 10^{-6}$ m^{-1}) and phase velocity. We conducted three different experiments. Firstly we presented the "ideal" case of a uniform distribution of noise sources; then we implemented two different spatially heterogeneous source distributions: one characterized by an azimuth-dependent source density, the other by the absence of noise sources in the near field of the receivers. For each experimental setup, we first verified the cancellation of the "cross-terms", predicted by the theory (eq. (4)) in case of a diffuse ambient wavefield and a laterally homogeneous source spectrum; we then verified that the source spectrum is reconstructed accurately, as predicted by the theory, if density of sources ρ and attenuation coefficient α are known. Finally, we performed an inversion to measure α from normalized cross correlations of synthetic recordings, through the cost function $C(\alpha, \omega)$. The definition of $C(\alpha, \omega)$ involves a sum over all available station pairs and therefore all available propagation azimuths; importantly, this reduces the unwanted effects of nonuniformites in source distribution. We successfully retrieved the correct values of α in the experiments involving noise sources in both near and far field of the receivers, with good accuracy over a broad frequency range. This result confirms that it is possible to estimate attenuation reliably, even if the assumption of a diffuse wavefield is not exactly met by the data. On the other hand, we inferred from the third experiment that when noise sources are absent in the near field of the receivers both source spectrum and attenuation are significantly underestimated.

We finally compiled a data set of noise recordings using 33 broadband receivers distributed within part of the Colorado plateau and of the Great Basin. We first used this data set to quantify the diffusivity of the ambient wavefield, calculating the signal-to-noise ratio (SNR) as a function of azimuth within the area of interest. The SNR proved to be rather homogeneous in the energy band characteristic of the primary microseisms (centered at the period of 14 s), but revealed a SW-NE preferential directionality of the noise sources within the secondary microseism band (6-8 s); this observation is compatible with what reported in previous studies. When inverting the data to constrain α, the effects of SNR inhomogeneity with respect to azimuth are reduced both by ensemble av-

eraging over time, and implicit averaging over azimuth in the definition of $C(\alpha, \omega)$. The resulting estimates of α, confirmed by a bootstrap analysis, range from $\sim 1{\times}10^{-5}$ m^{-1} at 0.3 Hz to $\sim 4.5 \times 10^{-7}$ m^{-1} at 0.065 Hz; in this frequency range, those values are compatible with previous observations made on the basis of both earthquake-generated and ambient Rayleigh waves.

Acknowledgments

Our many exchanges with Emanuel Kästle, Kees Weemstra, and Sebastian Lauro were very beneficial to this study. We thank the makers of Obspy (Beyreuther et al., 2010). Graphics were created with Python Matplotlib (Hunter, 2007). The facilities of IRIS Data Services, and specifically the IRIS Data Management Center (http://ds.iris.edu/ds/nodes/dmc/), were used for access to waveforms, related metadata, and/or derived products used in this study. We used publicly-available seismic data from the Transportable Array (TA) seismic network (https://doi.org/10.7914/SN/TA). The Grant to Department of Science, Roma Tre University (MIUR-Italy Dipartimenti di Eccellenza, ARTICOLO 1, COMMI 314 - 337 LEGGE 232/2016) is gratefully acknowledged.

References

Abramowitz, M., & Stegun, I. A. (1964). *Handbook of Mathematical Functions*. National Bureau of Standards–Applied Mathematics Series.

Al-Khatib, H. H., & Mitchell, B. J. (1991). Upper mantle anelasticity and tectonic evolution of the western United States from surface wave attenuation. *Journal of Geophysical Research: Solid Earth*, *96*(B11), 18129–18146.

Asten, M. W. (2006). On bias and noise in passive seismic data from finite circular array data processed using SPAC methods. *Geophysics*, *71*(6), V153–V162.

Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A. L., Lin, F., Moschetti, M., ... Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, *169*(3), 1239–1260.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, *81*(3), 530–533.

Boschi, L., Magrini, F., Cammarano, F., & van der Meijde, M. (2019). On seis-

mic ambient noise cross-correlation and surface-wave attenuation. *Geophysical Journal International*, *219*(3), 1568–1589.

Boschi, L., & Weemstra, C. (2015). Stationary-phase integrals in the cross-correlation of ambient noise. *Reviews of Geophysics*, *53*, doi:10.1002/2014RG000455.

Campillo, M., & Roux, P. (2014). Seismic imaging and monitoring with ambient noise correlations. In B. Romanowicz & A. M. Dziewonski (Eds.), *Treatise of geophysics. vol. 1.* Elsevier.

Cupillard, P., & Capdeville, Y. (2010). On the amplitude of surface waves obtained by noise correlation and the capability to recover the attenuation: a numerical approach. *Geophysical Journal International*, *181*(3), 1687–1700.

De Boor, C., De Boor, C., Mathématicien, E.-U., De Boor, C., & De Boor, C. (1978). *A practical guide to splines* (Vol. 27). springer-verlag New York.

Ekström, G. (2014). Love and Rayleigh phase-velocity maps, 5–40 s, of the western and central USA from USArray data. *Earth and Planetary Science Letters*, *402*, 42–49.

Friedrich, A., Krüger, F., & Klinge, K. (1998). Ocean-generated microseismic noise located with the Gräfenberg array. *Journal of Seismology*, *2*(1), 47–64.

Hillers, G., Graham, N., Campillo, M., Kedar, S., Landès, M., & Shapiro, N. (2012). Global oceanic microseism sources as seen by seismic arrays and predicted by wave action models. *Geochemistry, Geophysics, Geosystems*, *13*(1).

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in science & engineering*, *9*(3), 90–95. doi: 10.1109/MCSE.2007.55

Jones, E., Oliphant, T., Peterson, P., et al. (2001). *SciPy: Open source scientific tools for Python.* Retrieved from `http://www.scipy.org/` ([Online; accessed])

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. *Geophysical Journal International*, *207*, 1493–1512.

Landès, M., Hubans, F., Shapiro, N. M., Paul, A., & Campillo, M. (2010). Origin of deep ocean microseisms by using teleseismic body waves. *Journal of Geophysical Research: Solid Earth*, *115*(B5).

Lawrence, J. F., & Prieto, G. A. (2011). Attenuation tomography of the western United States from ambient seismic noise. *Journal of Geophysical Research: Solid Earth*, *116*(B6).

Lin, W. (1989). *Rayleigh wave attenuation in the Basin and Range province.* St Louis University MO, USA.

Menon, R., Gerstoft, P., & Hodgkiss, W. S. (2014). On the apparent attenuation in the spatial coherence estimated from seismic arrays. *Journal of Geophysical Research: Solid Earth*, *119*(4), 3115–3132.

Mitchell, B. J. (1995). Anelastic structure and evolution of the continental crust and upper mantle from seismic surface wave attenuation. *Reviews of Geophysics*, *33*(4), 441–462.

Patton, H. J., & Taylor, S. R. (1984). Q structure of the Basin and Range from surface waves. *Journal of Geophysical Research: Solid Earth*, *89*(B8), 6929–6940.

Prieto, G., Lawrence, J., & Beroza, G. (2009). Anelastic earth structure from the coherency of the ambient seismic field. *Journal of Geophysical Research: Solid Earth*, *114*(B7).

Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, *36*(8), 1627–1639.

Tanimoto, T., Ishimaru, S., & Alvizuri, C. (2006). Seasonality in particle motion of microseisms. *Geophysical Journal International*, *166*(1), 253–266.

Tian, Y., & Ritzwoller, M. H. (2015). Directionality of ambient noise on the Juan de Fuca plate: Implications for source locations of the primary and secondary microseisms. *Geophysical Journal International*, *201*(1), 429–443.

Tsai, V. C. (2011). Understanding the amplitudes of noise correlation measurements. *Journal of Geophysical Research: Solid Earth*, *116*(B9).

Weemstra, C., Boschi, L., Goertz, A., & Artman, B. (2013). Seismic attenuation from recordings of ambient noise. *Geophysics*, *78*(1), Q1–Q14.

Weemstra, C., Snieder, R., & Boschi, L. (2015). On the estimation of attenuation from the ambient seismic field: inferences from distributions of isotropic point scatterers. *Geophysical Journal International*, *203*(2), 1054–1071.

Weemstra, C., Westra, W., Snieder, R., & Boschi, L. (2014). On estimating attenuation from the amplitude of the spectrally whitened ambient seismic field. *Geophysical Journal International*, *197*(3), 1770–1788.

Yang, Y., & Ritzwoller, M. H. (2008). Characteristics of ambient seismic noise as a source for surface wave tomography. *Geochemistry, Geophysics, Geosystems*, *9*(2).

Yokoi, T., & Margaryan, S. (2008). Consistency of the spatial autocorrelation method with seismic interferometry and its consequence. *Geophysical Prospecting*, *56*(3), 435–451.