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:

- Synthetic data are generated, to validate inverse theory relating seismic ambient noise and surface-wave attenuation
- Inversion of synthetic data suggests that ambient noise, even when not perfectly "diffuse", is a useful constraint to attenuation
- Preliminary application to North American data yields attenuation estimates compatible with previous studies

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

–1–

Abstract

We evaluate, by numerical tests, whether surface-wave attenuation can be determined from ambient-noise data. 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 the source spectrum can be reconstructed from ambient recordings (provided that the density of sources and the attenuation coefficient are known) and that 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), consistent with previous studies. 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. The result is compatible with previous observations based on both earthquake-generated and ambient Rayleigh waves. In this regard, the 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 *am-*

plitude, on the other hand, is less accurately reconstructed by cross correlation (e.g., Lehujeur & Chevrot, 2020). 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) derived a mathematical expression for the multiplicative factor relating normalized cross correlations to the Rayleigh-wave Green's function; numerical evaluation showed it to be equal to 1, under the assumption of a spatially uniform source distribution, confirming the speculation of Prieto et al. (2009).

It can be inferred from the mentioned theoretical contributions, as well as from that of Nakahara (2012), that – in an idealized situation (spatial and spectral uniformity of source distribution) – surface-wave attenuation can be measured from seismic ambient-noise data. The purpose of this study is to quantify the error that is caused, in such attenuation estimates, by applying the algorithm of Boschi et al. (2019, 2020) to recordings of non-idealized, more realistic seismic noise. We model those by numerical simulation of the signal resulting from a suite of increasingly non-uniform source distributions. We then invert the synthetic data and evaluate the similarity of the thus obtained attenuation estimates to the attenuation model used to compute the data.

Our inversion technique differs from those implemented by other authors in several aspects. First of all, we determine phase velocity through a preliminary, independent inversion: this is preferable to constraining velocity and attenuation at the same time, since phase velocity is directly related to the "zero crossings" of the Green's function in the frequency domain and independent of its amplitude (e.g., Ekström et al., 2009; Boschi et al., 2013). The so obtained velocity is then a fixed parameter of the attenuation inversion. Secondly, we do not average ambient-noise cross-correlation data over azimuth and/or distance, and we minimize the sum of misfits associated with each inter-station pair. Earlier work by Menon et al. (2014), as well as the synthetic tests presented here, indicate that this contributes to "regularizing" the inversion, resulting in an ap-

parently more accurate solution. Third, cross correlations are systematically processed in the time-domain, prior to inversion, so as to remove any signal associated with surface-wave overtones.

The results of our numerical tests, discussed in Section 4, indicate that our approach can potentially contribute to constraining surface-wave attenuation at regional and continental scale, where Rayleigh-wave fundamental mode dominates the seismic ambient noise. A preliminary application to USArray is given in Section 5. The relative success of our tests both confirms the validity of the theory (summarized in Section 2) and shows that our inversion strategy (described in detail in Section 3) is sound.

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 generated at $\mathbf{x}_S$ and recorded at $\mathbf{x}$ (e.g. Boschi et al., 2019, App. A),

$$G_{2D}^d(\mathbf{x}, \mathbf{x}_S, \omega) = -\frac{i}{4\sqrt{2\pi}c^2} H_0^{(2)}\left(|\mathbf{x} - \mathbf{x}_S|\sqrt{\frac{\omega^2}{c^2} - \frac{2i\alpha\omega}{c}}\right), \tag{1}$$

where i, ω, c, and α denote imaginary unit, angular frequency, phase velocity and attenuation coefficient, respectively; $|\mathbf{x} - \mathbf{x}_S|$ is the distance between the impulsive source and the receiver, 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. $\alpha \ll \omega/c$, eq. (1) can be reduced to the more convenient, approximate form

$$G_{2D}^d(\mathbf{x}, \mathbf{x}_S, \omega) \approx -\frac{i}{4\sqrt{2\pi}c^2} H_0^{(2)}\left(\frac{\omega|\mathbf{x} - \mathbf{x}_S|}{c}\right) e^{-\alpha|\mathbf{x} - \mathbf{x}_S|}, \tag{2}$$

employed throughout this 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 $h(\omega)G_{2D}^d(\mathbf{x}_A, \mathbf{x}, \omega)e^{i\phi}$. The vertical-component, Rayleigh-wave dis-

–4–

108 placement associated with a set of N_S sources with the same amplitude $h(\omega)$, but different phase,

109 is then

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

111 where the index j identifies the source. Let a second receiver, located at $\mathbf{x}_B$, also record the sig-

112 nal emitted by the same sources. In the frequency domain, the cross-correlation of recordings made

113 at $\mathbf{x}_A$ and $\mathbf{x}_B$ is then given by

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

115 (e.g., Press et al., 1992, sec. 13.2), where * denotes complex conjugation.

116 Seismic surface-wave ambient noise is usually described by eq. (3), with the additional re-

117 quirement that the phases ϕ_1, ϕ_2, $\phi_3, \ldots$ be random (uniformly distributed between 0 and 2π).

118 Under this assumption, it has been shown (Weemstra et al., 2014; Boschi & Weemstra, 2015) that,

119 if the statistical expectation (e.g., Bendat & Piersol, 2011, eq. (3.8)) $E[\ldots]$ of eq. (4) is taken,

120 the $j \neq k$ terms ("cross terms") cancel out and

$$E[s(\mathbf{x}_A, \omega) s^*(\mathbf{x}_B, \omega)] = h^2(\omega) \sum_{j=1}^{N_S} G_{2D}^d(\mathbf{x}_A, \mathbf{x}_j, \omega) G_{2D}^{d*}(\mathbf{x}_B, \mathbf{x}_j, \omega). \tag{5}$$

122 In the following, we shall drop the symbol $E[\ldots]$ for the sake of simplicity. In a passive seismol-

123 ogy experiment as the ones we are concerned with, statistical expectation cannot be evaluated

124 directly from the data, which are seismic traces of finite duration. It is assumed that the finite-

125 time cross correlation of seismic ambient noise coincides with its expectation. Our work hypoth-

126 esis, similar e.g. to Boschi and Weemstra (2015), is that this condition is approximately verified,

127 in ambient-noise seismology, if sufficiently long time series (typically one year) are considered:

128 in that case, a large number of sources distributed over the entire real plane, with random un-

129 correlated phases, should eventually be sampled. In practice, we call "diffuse" (e.g., Kinsler et

130 al. (1999), section 12.1; Boschi and Weemstra (2015), section 3) a wave field that meets this re-

131 quirement. (The purpose of this study, in a sense, is to evaluate how far real data are from such

132 an idealized situation.)

133　　　As shown by Boschi et al. (2019), under this assumption the sum at the right-hand side (RHS)

134　of eq. (4) can be replaced by an integral, and combined with the reciprocity theorem for a lossy

135　membrane (Section 2.2 of Boschi et al., 2019) to yield

$$
s(\mathbf{x}_A, \omega) s^*(\mathbf{x}_B, \omega) \approx -\frac{h^2(\omega)\rho}{2\sqrt{2\pi}\alpha\omega c} \Im[G_{2D}^d(\mathbf{x}_A, \mathbf{x}_B, \omega)], \tag{6}
$$

137　where the operator $\Im[\ldots]$ maps a complex number into its imaginary part, ρ is the surface den-

138　sity of noise sources, and the factor $1/\sqrt{2\pi}$ arises from the correction of the algebraic error found

139　in Boschi et al. (2019) (Boschi et al., 2020). Equation (6) stipulates that the amplitude of the

140　cross-correlation of ambient-noise recordings carries the information on the surface-wave atten-

141　uation coefficient α. This means that α can be retrieved from the data, if inter-station phase ve-

142　locity and spatial density and power spectral density of noise sources are known.

143　**2.3 Power spectral density as normalization term**

144　　　As shown by Boschi et al. (2019), the RHS of eq. (6) can be manipulated algebraically, to

145　find an expression for the cross correlation of ambient noise where the source parameters $h^2(\omega)$

146　and ρ conveniently cancel out. In the following, we rederive the result of Boschi et al. (2019) in

147　a slightly different fashion, showing explicitly that our formulation is in agreement with that of

148　Nakahara (2012). (The result of Nakahara (2012) had been overlooked by Boschi et al. (2019).)

149　　　Let us rewrite eq. (2) as

$$
G_{2D}^d(x_1, x_2, \omega) \approx -\frac{i}{4\sqrt{2\pi}c^2}\left[J_0\left(\frac{\omega|\mathbf{x}_A - \mathbf{x}_B|}{c}\right) - i Y_0\left(\frac{\omega|\mathbf{x}_A - \mathbf{x}_B|}{c}\right) \right] e^{-\alpha|\mathbf{x}_A - \mathbf{x}_B|} \tag{7}
$$

151　(e.g. Boschi & Weemstra, 2015), where J_0 and Y_0 denote zeroth-order Bessel functions of the first

152　and second kind (e.g. Abramowitz & Stegun, 1964), respectively. Substituting eq. (7) into (6)

153　yields

$$
s(\mathbf{x}_A, \omega) s^*(\mathbf{x}_B, \omega) \approx -\frac{h^2(\omega)\rho}{16\pi\alpha\omega c^3} J_0\left(\frac{\omega|\mathbf{x}_A - \mathbf{x}_B|}{c}\right) e^{-\alpha|\mathbf{x}_A - \mathbf{x}_B|}. \tag{8}
$$

155　It follows from eq. (8) that cross-correlating the signal recorded at an arbitrary location $\mathbf{x}$ with

156　itself, one obtains

$$
|s(\mathbf{x}, \omega)|^2 \approx \frac{h^2(\omega)\rho}{16\pi\alpha\omega c^3}. \tag{9}
$$

–6–

Similar to Boschi et al. (2019), this formulation provides an analytical expression for $h^2(\omega)$ via eq. (9), which can be rewritten

$$\begin{aligned}
h^2(\omega) &\approx \frac{16\pi\alpha\omega c^3}{\rho} \, |s(\mathbf{x},\omega)|^2 \\
&\approx \frac{16\pi\alpha\omega c^3}{\rho} < |s(\mathbf{x},\omega)|^2 >_{\mathbf{x}} \, ,
\end{aligned} \tag{10}$$

where, following the assumption that h is the same for all noise sources, we replaced $|s(\mathbf{x},\omega)|^2$ with its average $< |s(\mathbf{x},\omega)|^2 >_{\mathbf{x}}$ over all available receiver locations $\mathbf{x}$. Upon dividing eq. (8) by (9), we find

$$\frac{s(\mathbf{x}_A,\omega)s^*(\mathbf{x}_B,\omega)}{< |s(\mathbf{x},\omega)|^2 >_{\mathbf{x}}} \approx J_0\left(\frac{\omega|\mathbf{x}_A - \mathbf{x}_B|}{c}\right) \mathrm{e}^{-\alpha|\mathbf{x}_A - \mathbf{x}_B|}, \tag{11}$$

where, again, $|s(\mathbf{x},\omega)|^2$ is replaced by the average $< |s(\mathbf{x},\omega)|^2 >_{\mathbf{x}}$. Importantly, equation (11), adopted in this study, coincides with eq. (52) of Aki (1957), except for the damping term $\mathrm{e}^{-\alpha|\mathbf{x}_A - \mathbf{x}_B|}$, and is consistent with eqs. (9) of Prieto et al. (2009) and (31) of Nakahara (2012).

Equation (11) should be compared with eq. (30) of Boschi et al. (2019), i.e.

$$\frac{s(\mathbf{x}_A,\omega)s^*(\mathbf{x}_B,\omega)}{< |s(\mathbf{x},\omega)|^2 >_{\mathbf{x}}} \approx \frac{c}{\omega\,\pi\,I(\alpha,\omega,c)} J_0\left(\frac{\omega|\mathbf{x}_A - \mathbf{x}_B|}{c}\right) \frac{\mathrm{e}^{-\alpha|\mathbf{x}_A - \mathbf{x}_B|}}{\alpha}, \tag{12}$$

where the mentioned algebraic error (a factor $1/\sqrt{2\pi}$) has been corrected and

$$I(\alpha,\omega,c) = \int_0^\infty \mathrm{d}r \; r \; \left|H_0^{(2)}\left(\frac{\omega r}{c}\right)\right|^2 \mathrm{e}^{-2\alpha r}. \tag{13}$$

Comparing eqs. (12) and (11) provides an analytical expression for the integral at the right-hand side of eq. (13),

$$I(\alpha,\omega,c) \approx \frac{c}{\pi\alpha\omega}, \tag{14}$$

which Boschi et al. (2019) were unable to solve analytically. We show in Fig. 1 that our numerical evaluation of $I(\alpha,\omega,c)$ (which exploits Gaussian quadrature, as provided by the SciPy Python library, Jones et al., 2001) is consistent with the expression at the right-hand side of eq. (14), if the conditions for the theoretical validity of our formulation are met. This further validates the formulation of Boschi et al. (2019).

The left-hand side (LHS) of both (11) and (12) 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 to eqs. (11) and/or (12); it follows that these equations 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 constant in space). In addition, if the LHS of eqs. (11) and (12) is calculated as an

186 ensemble-average of relatively small temporal windows with respect to the entire recording time,

187 the normalization term $< |s(\mathbf{x}, \omega)|^2 >_\mathbf{x}$ mitigates the effect of possible anomalous, ballistic sig-

188 nals like, e.g., large or nearby earthquakes (Boschi et al., 2019). This is often necessary when work-

189 ing with real data and commonly accomplished by one-bit normalization or spectral whitening

190 (e.g. Bensen et al., 2007); these empirical normalization terms, however, albeit useful for retriev-

191 ing phase or group velocities since they leave the phase of the cross correlations unchanged, are

192 doomed to introduce a bias in their amplitude and therefore in the resulting estimates of α (Weemstra

193 et al., 2014).

194   We emphasize that eq. (11) (or (12)) only holds if all our theoretical assumptions on the

195 nature of ambient noise and propagation medium are valid, i.e., (i) the seismic ambient field is

196 diffuse, (ii) source density and frequency content are both constant in space, and (iii) $\alpha \ll \omega/c$;

197 in such scenario, both its RHS and LHS are purely real. When working with observational data

198 these assumptions are not strictly verified, and the numerical value of the LHS, as obtained from

199 the data, is only approximately equal to the theoretical model at the RHS; this is why in ambient-

200 noise literature empirical Green's functions commonly show a non-zero imaginary part, and are

201 referred to as "complex coherency" (e.g. Weemstra et al., 2014) to distinguish them from the true

202 Green's function.

3 Inverse problem

204   Equation (11) allows to formulate an inverse problem to determine α from cross correla-

205 tions of recorded ambient signal. Because equation (11) holds for all station pairs, it is desirable

206 that the cost function be related to the weighted sum of the squared differences between LHS

207 and RHS of (11), calculated for each station pair; since the RHS is an oscillatory function of ω

208 (through the Bessel function J_0), and α only affects its envelope but not its oscillations (e.g. Pri-

209 eto et al., 2009; Boschi et al., 2019), we introduce the envelope function env to define the cost

210 function

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

(15)

212 where $\mathbf{x}_i$ and $\mathbf{x}_j$ denote the positions of the two receivers, and the sum spans all possible sta-

213 tion pairs. The weight $|\mathbf{x}_i - \mathbf{x}_j|^2$ is chosen based on the fact that larger inter-station distances

214 are associated with smaller amplitudes of the cross correlations, due to geometrical spreading,
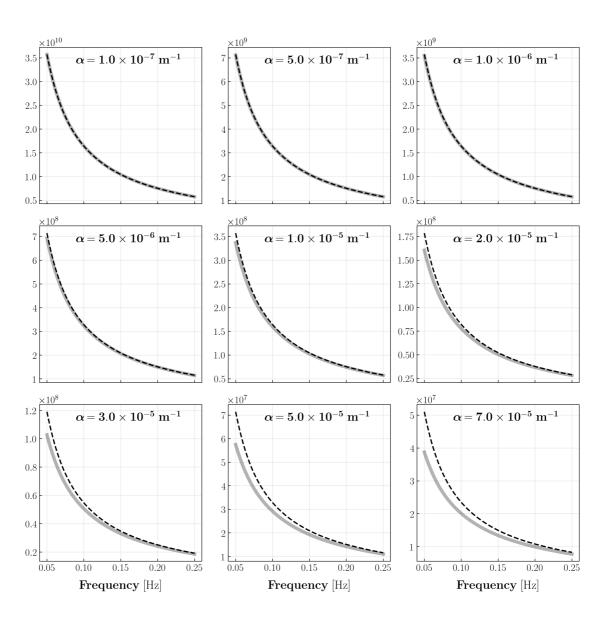
Figure 1. $I(\alpha, \omega, c)$ (gray, solid) and its analytical counterpart $\frac{c}{\pi\alpha\omega}$ (black, dashed) as functions of frequency, shown for different values of attenuation coefficient α. The phase velocity $c = c(\omega)$ is chosen to vary smoothly between 0.05 Hz (where $c = 3526$ ms^{-1}) and 0.25 Hz (2851 ms^{-1}), as illustrated in the supplementary materials (Fig. S1). The same values of $c(\omega)$ are employed in the numerical simulations of Section 4.

215 which would result in smaller absolute values of misfit if not weighted accordingly. $c_{ij}(\omega)$ denotes

216 the average phase velocity, at frequency ω, between receivers at $\mathbf{x}_i$ and $\mathbf{x}_j$; for each station pair,

217 we determine its value before minimizing $C(\alpha, \omega)$ by application of the method described by Kästle

218 et al. (2016). The minimum of $C(\alpha, \omega)$ can then be found through some form of "grid-search"

219 over α, for a discrete set of values of ω. The formula (15) for $C(\alpha, \omega)$ was selected after exper-

220 imenting several other options, as partly documented in Boschi et al. (2019). After a suite of pre-

221 liminary tests, we chose to implement the envelope function by fitting a combination of cubic splines

222 (De Boor, 1978) to the maxima of the absolute value of the real part of their arguments; the en-

223 velopes are then smoothed by means of a running average performed with a Savitzky-Golay fil-

224 ter (Savitzky & Golay, 1964). Smoothing is motivated by the fact that, if the anelastic proper-

225 ties of the Earth are assumed to be smoothly varying with depth, the same behavior is expected

226 for the amplitude of adjacent peaks of the real coherency; abrupt amplitude variations are as-

227 cribed to a non-perfectly diffuse wavefield or, in the case of real recordings, simply to noisiness

228 of the empirical Green's function.

229   The summation over receiver pairs i, j at the RHS of (15) involves all the available receivers

230 and, if the array has good azimuthal coverage, most azimuths of wave propagation. Minimizing

231 $C(\alpha, \omega)$ therefore involves finding one function $\alpha(\omega)$ such that a good fit is simultaneously achieved

232 at all azimuths; this has a regularizing effect on the inversion, and should reduce the effects of

233 non-homogeneity in azimuthal source distribution.

234   Previous studies (e.g. Prieto et al., 2009; Weemstra et al., 2013) formulated inverse prob-

235 lems whose data consisted of azimuthally averaged cross correlations calculated over several sta-

236 tion pairs; this was based on the idea that azimuthal averaging is necessary to retrieve a reliable,

237 purely real empirical Green's function (e.g. Asten, 2006; Yokoi & Margaryan, 2008). Menon et

238 al. (2014), however, show that slightly different inter-station distances or a laterally inhomoge-

239 neous phase velocity would introduce a phase offset of the cross correlations involved in the av-

240 erage. This would result in an "attenuation-like" effect (Fig. 2 of Menon et al., 2014), i.e., in a

241 fictitious decrease of the amplitude of the averaged coherency and thus in a bias of the estimates

242 of α.

4 Numerical Validation

244   We simulate ambient signal via a very large number of randomly distributed, uncorrelated

245 point sources. We next solve an inverse problem, as described above, to retrieve the theoretical

²⁴⁶ value of α; we also verify numerically the emergence of coherent signal in the cross correlations

²⁴⁷ due to cancellation of cross-terms in eq. (4), and the validity of eq. (10), which relates recorded

²⁴⁸ ambient noise and the frequency spectrum of ambient-noise sources. The simulation is carried

²⁴⁹ out in four different experimental setups. First, we present the "ideal" case of a spatially uni-

²⁵⁰ form distribution of sources (Fig. 2a). Since real-world ambient sources are not distributed uni-

²⁵¹ formly (e.g. Hillers et al., 2012), we next discuss the cases of an azimuthally heterogeneous source

²⁵² distribution (Fig. 2b) and of a distribution characterized by absence of noise sources in the vicin-

²⁵³ ity of the receivers (Fig. 2c). Finally, we show the results obtained through a "patchy" distri-

²⁵⁴ bution, characterized by a variable density of sources both in space and in azimuth (Fig. 2d).

4.1 Uniform source distribution

²⁵⁵

²⁵⁶ Our first experimental setup consists of 200,000 point sources randomly distributed both

²⁵⁷ in the near and far field of 29 receivers, within a circle of radius $R = 1 \times 10^7$ m centered at the

²⁵⁸ receiver array (Fig. 2a). Source locations are defined by their polar coordinates θ, r with respect

²⁵⁹ to one station located at the center of the array; random values of θ between 0 and 2π, and of

²⁶⁰ n between 0 and 1 are generated, and $r = R\sqrt{n}$ (the square root results in a linear growth of

²⁶¹ the number of sources with increasing distance from the center of the circle, hence constant source

²⁶² density in space). The receivers are randomly deployed in the central part of such distribution

²⁶³ on 4 concentric circles, with radii of 45, 90, 135, and 180×10^3 m.

4.1.1 Simulation of seismic ambient noise

²⁶⁴

²⁶⁵ In each of two experiments, synthetic data are generated using different models of atten-

²⁶⁶ uation, i.e. (i) constant attenuation with $\alpha = 1 \times 10^{-6}$ m^{-1}, and (ii) frequency-dependent $\alpha =$

²⁶⁷ $\alpha(\omega)$, chosen to vary linearly from 3×10^{-7} at 0.05 Hz to 1×10^{-6} m^{-1} at 0.25 Hz. (A third

²⁶⁸ experiment, with constant $\alpha = 5 \times 10^{-7}$ m^{-1} is illustrated in the supplementary materials.)

²⁶⁹ In all numerical simulations we employed a fixed, frequency-dependent phase velocity $c = c(\omega)$,

²⁷⁰ which decreases monotonously (and almost linearly) between 0.05 Hz (where $c = 3526$ ms^{-1})

²⁷¹ and 0.25 Hz (2851 ms^{-1}), with a slight kink around 0.07 Hz where its derivative with respect to

²⁷² time decreases with increasing frequency (Fig. S1). We consider these values to be realistic, based,

²⁷³ e.g., on Mitchell (1995) and Ekström (2014).

²⁷⁴ Each numerical test consisted of 25,000 *realizations* (Cupillard & Capdeville, 2010; Weem-

²⁷⁵ stra et al., 2015). At each realization every source emits an independent signal of constant am-
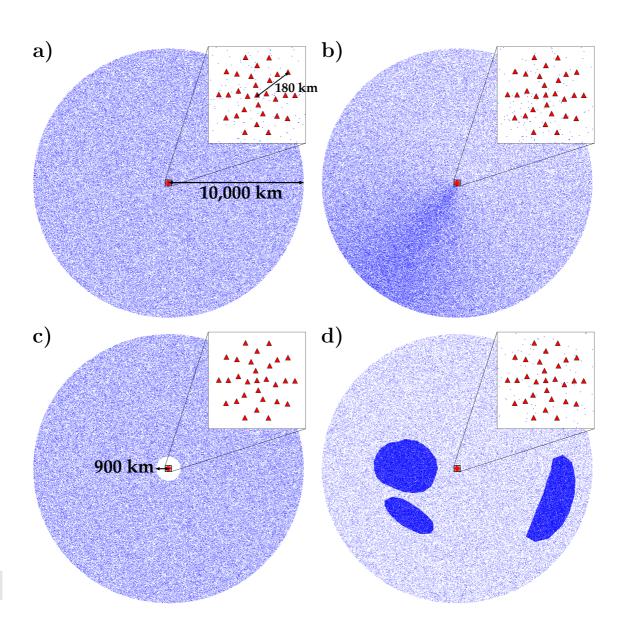
Figure 2. Sources (blue dots) and stations (red triangles) used for simulating seismic noise. (a), (b), (c), and (d) indicate the source distributions used in Sections 4.1 (uniform source distribution), 4.2 (azimuth-dependent source density), 4.3 (no sources in the near field), and 4.4 ("patchy" source distribution) respectively.
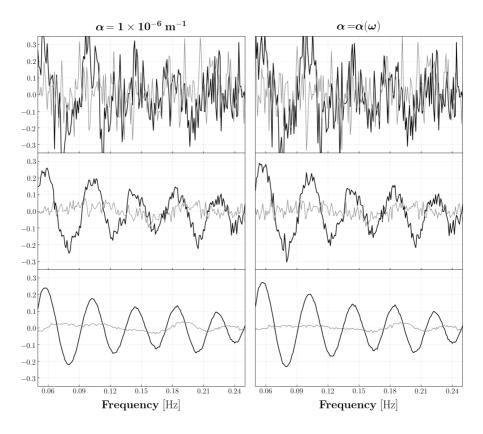
$$\alpha = 1 \times 10^{-6} \text{ m}^{-1} \qquad\qquad \alpha = \alpha(\omega)$$

Figure 3. Real (black) and imaginary (gray) parts of LHS of eq. (16), obtained for a pair of receivers with inter-station distance of 67,600 m by ensemble-averaging over (top) 25, (middle) 500, and (bottom) 25,000 realizations. Results are shown for both $\alpha = 1 \times 10^{-6}$ and $\alpha = \alpha(\omega)$, used in the experimental setup of Section 4.1 (uniform source distribution). Note the slightly different amplitudes of the real coherencies, as expected for media characterized by different attenuation coefficients.

plitude $h(\omega) = 1$ 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. (11) 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,

$$\frac{s(\mathbf{x}_A, \omega) s^*(\mathbf{x}_B, \omega)}{< |s(\mathbf{x}, \omega)|^2 >_{\mathbf{x}}} = \frac{1}{N_R} \sum_{k=1}^{N_R} \frac{s_k(\mathbf{x}_A, \omega) \, s_k^*(\mathbf{x}_B, \omega)}{< |s_k(\mathbf{x}, \omega)|^2 >_{\mathbf{x}}}. \tag{16}$$

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. (16), are shown in Fig. 3 for a pair of receivers with inter-station distance of 67,600 m. For both chosen values of α, the increase in
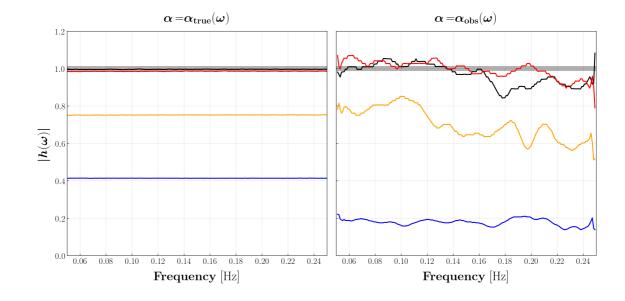
–13–

Figure 4. Absolute value of source amplitude $|h(\omega)|$, retrieved in the synthetic tests described in sections 4.1 (uniform source distribution, black), 4.2 (azimuth-dependent source density, red), 4.3 (absence of near-field sources, blue), and 4.4 ("patchy" source distribution, orange). In all cases, synthetic data are computed with $h(\omega) = 1$ (gray) and $\alpha = 1 \times 10^{-6}$ m^{-1}. $|h(\omega)|$ is calculated by substituting into eq. (9) both the true value of α (left) and the ones inferred from the cross correlations (right), and then taking the square root; the estimates of α have been obtained by minimizing the cost function $C(\alpha, \omega)$ (Section 3). In the case of a uniform source distribution, analogous results are obtained through the synthetics generated with $\alpha = 5 \times 10^{-7}$ m^{-1} and $\alpha = \alpha(\omega)$, as shown in Fig. S5.

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). An analogous result is obtained through the synthetic recordings generated with $\alpha = 5 \times 10^{-7}$ m^{-1}, as illustrated in Fig. S2.

Equation (10) 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 the same for all sources (Section 2.2); we show in Fig. 4 that, implementing eq. (10), $h(\omega) = 1$ is retrieved correctly, at least to the second decimal digit, in case of a uniform source distribution. This result validates numerically the derivation of eq. (10). Fig. 4 also shows that, when the true α is unknown, the observed values of α (obtained by minimization of the cost function $C(\alpha, \omega)$, as explained in the following paragraph) can be used to implement eq. (10), allowing one to achieve only slightly worse accuracy.

–14–

4.1.3 Retrieval of the attenuation coefficient

The cost function $C(\alpha, \omega)$ in eq. (15) 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}. Fig. 5 shows that, on average, the minima of $C(\alpha, \omega)$ correspond to the values of α used for generating synthetic recordings, i.e. the synthetic test is successful. In the same Fig., we show for both numerical tests the datafit obtained by substituting into eq. (11) the values of $\alpha(\omega)$ retrieved by minimizing $C(\alpha, \omega)$, and can be considered good at all the investigated inter-station distances. An analogous result is obtained through the synthetic recordings generated with $\alpha = 5 \times 10^{-7}$ m^{-1}, as illustrated in the supplementary materials (Fig. S7).

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. 2b). 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.

We verified the emergence of coherent signal in the normalized cross correlations at increasing number of realizations, as illustrated in Fig. S3. As expected, the cross terms cancel out, but, in comparison with the previous section, we found slightly larger imaginary parts of the cross-spectra; this can be ascribed to the nonuniform distribution of the noise sources (see Section 2). The source spectrum $h(\omega)$, retrieved from the synthetics, is shown in Fig. 4 to be closed to the true value of 1; this indicates that eq. (10) allows estimating the frequency content of the noise sources accurately, 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}. This is illustrated in Fig. 6, together with the datafit obtained by substituting into eq. (11) the best values of $\alpha(\omega)$ for the same station pairs employed in Fig. 5.
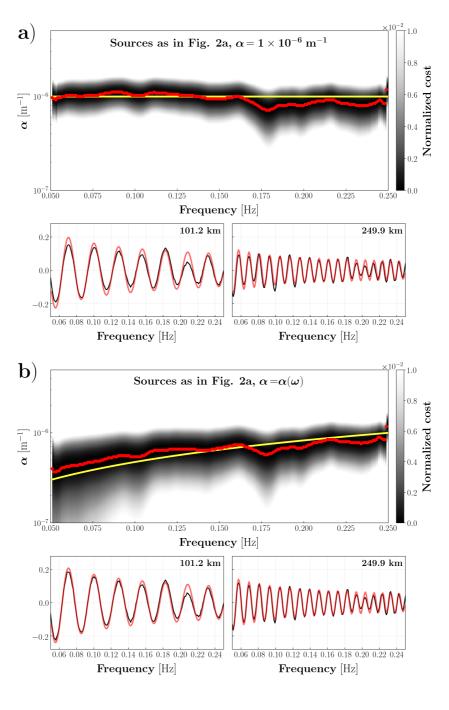
Figure 5. (a) Top panel: cost function $C(\alpha, \omega)$ associated with the numerical experiment of Section 4.1 (uniform source distribution) shown (after normalization) as a function of attenuation coefficient and frequency; red dots mark the values of α for which $C(\alpha, \omega)$ is minimized at each frequency; the yellow line indicates the assumed attenuation model $\alpha = 1 \times 10^{-6}$ m^{-1}, used for generating synthetic recordings. Bottom: normalized cross correlations (black) fitted by the model (red) obtained by substituting into eq. (11) the values of $\alpha(\omega)$ which minimize $C(\alpha, \omega)$. Within each subplot, the inter-station distance is indicated on the upper right. (b) Same as (a), but obtained through the synthetic recordings generated with a frequency-dependent $\alpha = \alpha(\omega)$, varying between 3×10^{-7} m^{-1} at 0.05 Hz and 1×10^{-6} m^{-1} at 0.25 Hz.
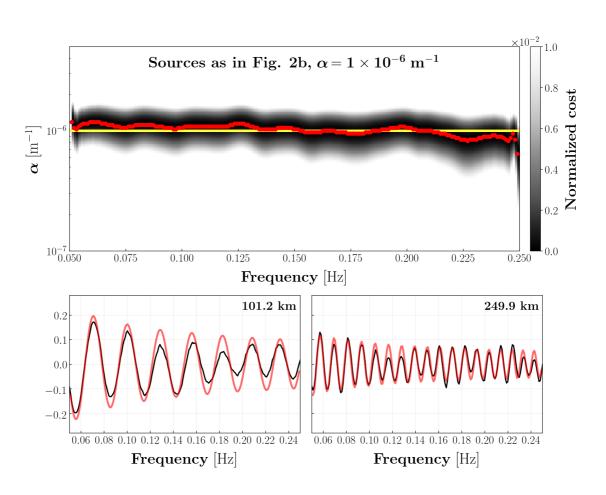
–16–

Figure 6. Top panel: 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. Bottom: normalized cross correlations (black) fitted by the model (red) obtained by substituting into eq. (11) the values of $\alpha(\omega)$ which minimize $C(\alpha, \omega)$. Within each subplot, the inter-station distance is indicated on the upper right.

327 The above results show that, even if the spatial distribution of noise sources is slightly nonuni-

328 form, the value of $\alpha(\omega)$ can be reconstructed correctly from the cross correlation of ambient noise:

329 we have achieved this, as anticipated, by neglecting possible lateral heterogeneities in $\alpha(\omega)$, and

330 minimizing a cost function where as many azimuths of propagation as possible are simultane-

331 ously included. In practice, this means that surface-wave attenuation can be estimated based on

332 ambient noise, even when the noise field is not exactly diffuse. This is indeed the case in most

333 practical applications.

4.3 No near-field sources

335 Sources are uniformly distributed in space, as in Section 4.1, but starting at a minimum

336 distance of 900×10^3 m from the station that defines the center of the array (Fig. 2c). We im-

337 plement 25,000 realizations with the same phase velocity $c = c(\omega)$ as before, attenuation $\alpha =$

338 1×10^{-6} m^{-1}, and $h(\omega) = 1$. Again, a random phase ϕ between 0 and 2π, newly generated at

339 each realization, is assigned to each source.

340 In analogy with the experiments above, we verified the emergence of coherent signal in the

341 cross correlations due to the cancellation of cross-terms. The amplitude of the imaginary part

342 of the cross-spectra turned out to be similar to that obtained for an azimuthally heterogeneous

343 source distribution (see Fig. S4). On the other hand, the real part is systematically larger than

344 those observed in the previous experiments. As in Sections 4.1 and 4.2, we then used the syn-

345 thetic data to quantify source spectrum $h(\omega)$ and attenuation $\alpha(\omega)$ (Figs. 4 and 7). We infer from

346 the results thus obtained that the absence of near-field sources leads to a significant underesti-

347 mate of both $h(\omega)$ and $\alpha(\omega)$ (the latter by a factor of about 5), in agreement with the theoret-

348 ical findings of Tsai (2011).

4.4 "Patchy" source distribution

350 100,000 sources are uniformly distributed in space, as in Section 4.1, and 100,000 sources

351 are concentrated in discrete regions (Fig. 2d). As in the previous section, we implemented 25,000

352 realizations using the phase velocity $c = c(\omega)$, attenuation $\alpha = 1 \times 10^{-6}$ m^{-1}, and $h(\omega) = 1$,

353 attributing to each source a random phase ϕ between 0 and 2π newly generated at each realiza-

354 tion.

355 We first used the so generated synthetic recordings to compute the normalized cross-correlations,

356 whose smoothness indicate that the cross-terms canceled out, as predicted by the theory (Fig.
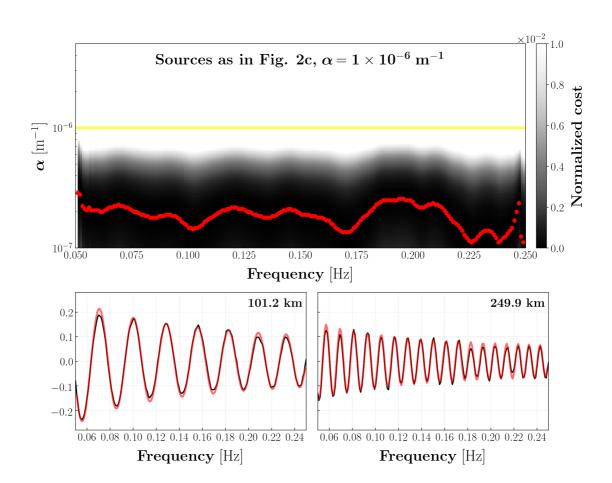
–18–

Figure 7. Same as Fig. 6, but obtained through the numerical experiment of Section 4.3 (no sources in the near field).
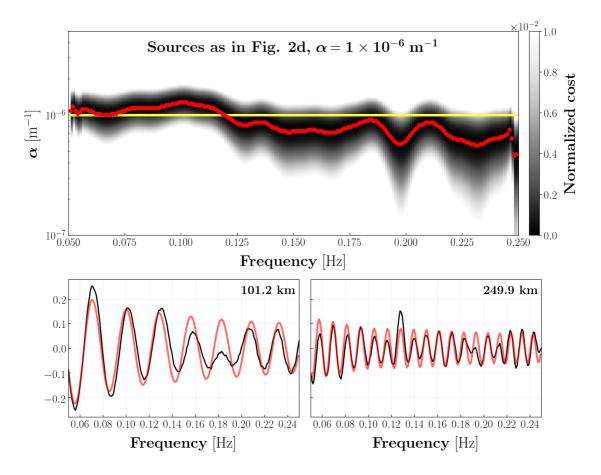
Figure 8. Same as Figs. 6 and 7, but obtained through the numerical experiment of Section 4.4 ("patchy" source distribution).

S5). Due to the strong lateral heterogeneity characterizing this source distribution, we found, on average, larger imaginary parts of the cross-spectra than what we observed in the experiments above. We then used the cross-spectra to reconstruct $h(\omega)$ (Fig. 4), finding values $\sim 25\%$ less than the true frequency spectrum used to generate the synthetics: a much better result than what we achieved in the case of a distribution characterized by absence of near-field sources. Finally, we inverted the synthetic data using eq. (15) to obtain estimates of the attenuation coefficient that fluctuate around the true value of α. Overall, we observed a lower accuracy in this experimental setup than those of Sections 4.1 and 4.2 (compare Fig. 8 with Figs. 5 and 6), but still acceptable considering the extreme spatial and azimuthal heterogeneity of the source distribution.
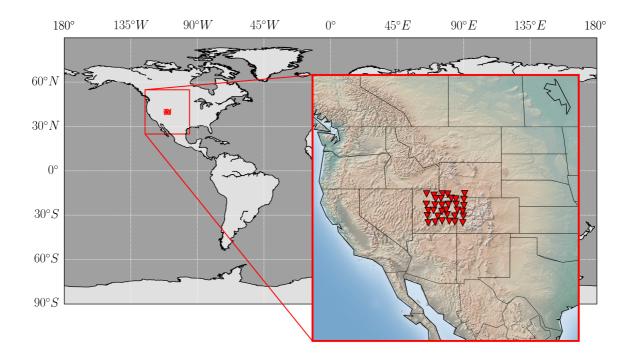
Figure 9. Seismic stations (red triangles) from the USArray project transportable network, forming the data set described in Section 5.1

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. 9) and operating between February 2007 and August 2008. Each seismogram has been demeaned, 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. From all those continous data we computed 509 empirical Green's functions (i.e. LHS of eq. (11)), 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 cross-correlated pairs of receivers that recorded simultaneously for more than 9 months.

We next "pre-process" the normalized cross-correlation, to better isolate the fundamental-mode amplitude signal that is relevant to the subsequent attenuation inversion, i.e., the "stationary points" of the frequency-domain cross correlation curves (Fig. 10). This processing was not carried out in the synthetic tests discussed above, but is likely to be necessary when dealing with real-world ambient signal. The procedure we have designed consists of three steps. (i) We inverse-
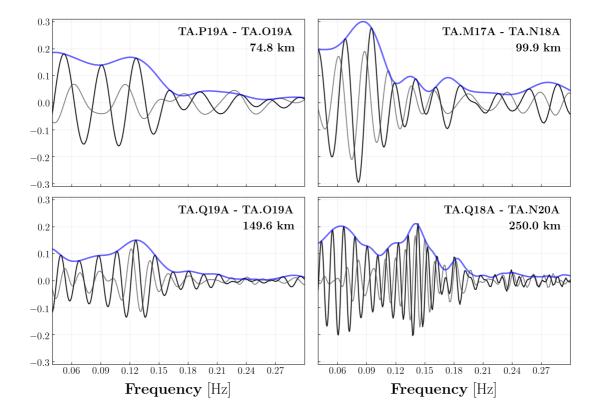
Figure 10. Real (black) and imaginary (gray) parts of smoothed (Section 5.1), normalized cross correlations calculated for four different station pairs. The envelope of the real parts, used in the inversion for α, is shown in blue. At the upper right corner of each subplot, station codes and inter-station distance are specified.

Fourier transform our frequency-domain cross correlations to the time domain; (ii) we zero-pad the resulting time-domain traces at times corresponding to velocities lower than 2 km s^{-1} and higher than 5 km s^{-1} (the same cosine taper is applied at the two ends of this interval), i.e., all signal that is much faster or slower than the typical fundamental-mode surface wave; (iii) we forward-Fourier transform the padded cross correlations back to the frequency domain. Through this procedure, most of the signal that is not associated with the Rayleigh-wave fundamental mode (i.e., the Rayleigh-wave overtones; the body waves) is eliminated. By the properties of the Fourier transform, this has also the effect of "smoothing" the frequency-domain cross correlation (Fig. 10).

The thus achieved (normalized) cross-correlations served us to retrieve Rayleigh-wave dispersion curves in the frequency range between 0.3 and 0.04 Hz (we used the Kästle et al. (2016)'s automated algorithm) and the envelopes to be used in the inversion.

–22–

5.2 Signal-to-noise ratio

We show in Fig. 10 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 possible 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. 11) 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 $\sim$10 s to $\sim$20 s period, peaking at $\sim$14 s (e.g. Friedrich et al., 1998): compare, e.g. with Fig. 12 of Kästle et al. (2016), where a strong nonuniformity in the source distribution can be appreciated from exactly the same analysis. Our measurements of SNR at the central periods of the *secondary* microseisms band are characterized by a relative maximum along the SW-NE direction (see the periods of 6 and 8 s in Fig. 11). 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.
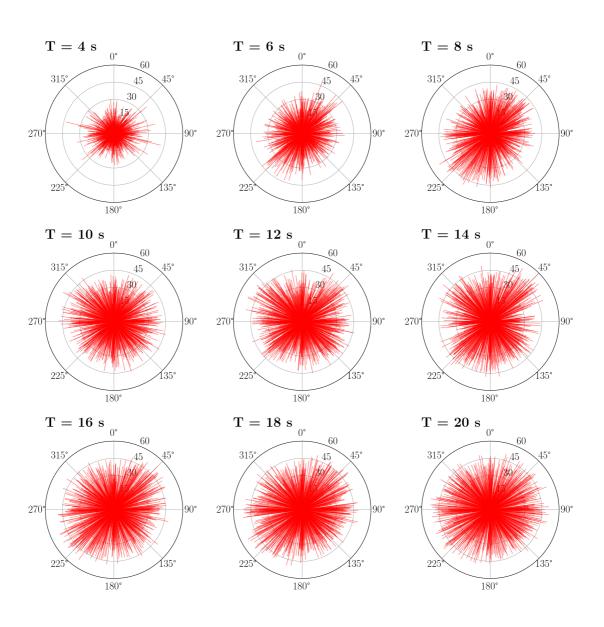
Figure 11. Signal-to-noise ratio at different periods as a function of azimuth, as inferred from the normalized cross-correlations. The length of the red segments is determined by the value of SNR, while their orientation coincides with the azimuth/back-azimuth of the respective station pair. 0° corresponds to the north, 90° to the east, etc.

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. 12. In the same Fig. we also show, for four different station pairs, the datafit obtained by substituting into eq. (11) the values of $\alpha(\omega)$ which minimize $C(\alpha, \omega)$. The source spectrum of the study area, as inferred from our measurements of $\alpha(\omega)$, $c(\omega)$, average power spectral density, and assuming a constant source density ρ, is shown in the supplementary materials (Fig. S8).

We assessed the uncertainty associated with α via a bootstrap analysis: in practice, 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. 12a, 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.

We infer from Fig. 12 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 ~ 0.065 Hz (periods $\gtrsim 16$ s), on the contrary, we observe an increase of α, in disagreement with some of the early observations shown in Fig. 12.

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
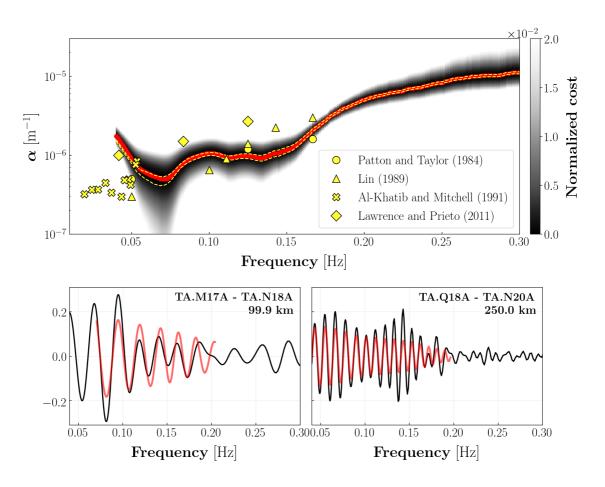
–25–

Figure 12. Top panel: 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). Bottom: normalized cross correlations (black) fitted by the model (red) obtained by substituting into eq. (11) the values of $\alpha(\omega)$ which minimize the cost function $C(\alpha, \omega)$. The datafit is shown for the same station pairs of Fig. 10. 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.

454 of sources in the near field. This issue merits further attention, but is beyond the scope of our

455 current study.

6 Conclusions

457 We have validated numerically the method proposed by Boschi et al. (2019, 2020) to quan-

458 tify the attenuation of Rayleigh waves from the cross correlation of seismic ambient noise. We

459 achieved this by simulating the displacement associated with 200,000 impulsive sources and recorded

460 by 29 receivers. In all our simulations, we imposed realistic values of phase velocity and atten-

461 uation ($\alpha = 5 \times 10^{-7}$ m^{-1}, $\alpha = 1 \times 10^{-6}$ m^{-1}, and a frequency-dependent $\alpha = \alpha(\omega)$ ranging

462 from 3×10^{-7} to 1×10^{-6} m^{-1}). Synthetic data were generated and inverted based on a suite

463 of different setups. We first considered the "ideal" case of a uniform distribution of noise sources;

464 then we implemented three different spatially heterogeneous source distributions: the first char-

465 acterized by significant azimuthal variations in source density, another by the absence of noise

466 sources in the near field of the receivers, and the third involving rapid source density variations

467 with both distance and azimuth (i.e., a "patchy" source distribution). For each experimental setup,

468 we first verified the cancellation of the "cross-terms", predicted by the theory (eq. (4)) when the

469 ambient wave field is diffuse and the spectrum of emitted noise does not change as a function of

470 source location; we then verified that the source spectrum is reconstructed accurately, as predicted

471 by the theory, if the source density and attenuation coefficient α are known. Finally, we performed

472 an inversion to measure α from normalized cross correlations of synthetic recordings, through the

473 cost function $C(\alpha, \omega)$. The definition of $C(\alpha, \omega)$ involves a sum over all available station pairs

474 and therefore all available propagation azimuths; importantly, this reduces the unwanted effects

475 of nonuniformites in source distribution. We successfully retrieved the correct values of α in syn-

476 thetic experiments where noise sources had been deployed in both the near and far field, with

477 good accuracy over a broad frequency range. This result confirms that it is possible to estimate

478 attenuation reliably, even if the assumption of a diffuse wavefield is not exactly met by the data.

479 On the other hand, we inferred from the third experiment that when noise sources are absent from

480 the near field of the receivers both source spectrum and attenuation are significantly underes-

481 timated.

482 Having validated our method by synthetic tests, we compiled a data set of noise recordings

483 using 33 broadband receivers distributed within part of the Colorado plateau and of the Great

484 Basin. We first used this data set to quantify the diffusivity of the ambient wavefield, calculat-

485 ing the signal-to-noise ratio (SNR) as a function of azimuth within the area of interest. The SNR

–27–

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 averaging 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

References

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

Aki, K. (1957). Space and time spectra of stationary waves with special reference to microtremors. *Bulletin of the Earthquake Research Institute, University of Tokyo*, *35*, 415–456.

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.

Bendat, J. S., & Piersol, A. G. (2011). *Random data: analysis and measurement procedures*

517   (Vol. 729). John Wiley & Sons.

518 Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A. L., Lin, F., Moschetti, M., . . . Yang, Y.

519   (2007).   Processing seismic ambient noise data to obtain reliable broad-band surface

520   wave dispersion measurements. *Geophysical Journal International*, *169*(3), 1239–1260.

521 Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J.   (2010).

522   ObsPy: A Python toolbox for seismology.   *Seismological Research Letters*, *81*(3), 530–

523   533.

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

525   ent noise cross-correlation and surface-wave attenuation.   *Geophysical Journal Interna-

526   tional*, *219*(3), 1568–1589.

527 Boschi, L., Magrini, F., Cammarano, F., & van der Meijde, M. (2020). Erratum: On seismic

528   ambient noise cross-correlation and surface-wave attenuation.   *Geophysical Journal In-

529   ternational*, *222*(2), 1090–1092.

530 Boschi, L., & Weemstra, C. (2015). Stationary-phase integrals in the cross-correlation of am-

531   bient noise. *Reviews of Geophysics*, *53*, doi:10.1002/2014RG000455.

532 Boschi, L., Weemstra, C., Verbeke, J., Ekström, G., Zunino, A., & Giardini, D.   (2013).   On

533   measuring surface wave phase velocity from station–station cross-correlation of ambient

534   signal. *Geophysical Journal International*, *192*(1), 346–358.

535 Campillo, M., & Roux, P.   (2014).   Seismic imaging and monitoring with ambient noise cor-

536   relations. In B. Romanowicz & A. M. Dziewonski (Eds.), *Treatise of geophysics. vol. 1.*

537   Elsevier.

538 Cupillard, P., & Capdeville, Y. (2010). On the amplitude of surface waves obtained by noise

539   correlation and the capability to recover the attenuation: a numerical approach.   *Geo-

540   physical Journal International*, *181*(3), 1687–1700.

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

542 Ekström, G. (2014). Love and Rayleigh phase-velocity maps, 5–40 s, of the western and cen-

543   tral USA from USArray data. *Earth and Planetary Science Letters*, *402*, 42–49.

544 Ekström, G., Abers, G. A., & Webb, S. C.   (2009).   Determination of surface-wave phase

545   velocities across USArray from noise and Aki's spectral formulation.   *Geophysical Re-

546   search Letters*, *36*, doi:10.1029/2009GL039131.

547 Friedrich, A., Krüger, F., & Klinge, K.   (1998).   Ocean-generated microseismic noise located

548   with the Gräfenberg array. *Journal of Seismology*, *2*(1), 47–64.

549 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.

Kinsler, L. E., Frey, A. R., Coppens, A. B., & Sanders, J. V. (1999). *Fundamentals of acoustics.*

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).

Lehujeur, M., & Chevrot, S. (2020). Eikonal Tomography Using Coherent Surface Waves Extracted From Ambient Noise by Iterative Matched FilteringApplication to the Large-N Maupasacq Array. *Journal of Geophysical Research: Solid Earth*, *125*(6), e2020JB019363.

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.

Nakahara, H. (2012). Formulation of the spatial autocorrelation (SPAC) method in dissipative media. *Geophysical Journal International*, *190*(3), 1777–1783.

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.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical*

Recipes in Fortran, Second Edition. Cambridge University Press.

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.

a) 10,000 km, 180 km

b)

c) 900 km

d)

a) Sources as in Fig. 2a, $\alpha = 1 \times 10^{-6}\ \mathrm{m}^{-1}$

101.2 km

249.9 km

b) Sources as in Fig. 2a, $\alpha = \alpha(\omega)$

101.2 km

249.9 km

Sources as in Fig. 2b, $\alpha = 1 \times 10^{-6}$ m^{-1}

Sources as in Fig. 2d, $\alpha = 1 \times 10^{-6}$ m^{-1}