Supporting Information for "Ambient-noise tomography of the Ligurian-Provence Basin using the AlpArray onshore-offshore network: Insights for the oceanic domain structure"

A. Nouibat[1], L. Stehly[1], A. Paul[1], S. Schwartz[1], Y. Rolland[1,2], T. Dumont[1],

W. C. Crawford[3], R. Brossier[1], Cifalps Team, and AlpArray Working Group

[1]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, UGE, ISTerre, *38000 Grenoble, France*

[2]Univ. Savoie Mont Blanc, CNRS, UMR 5204, EDYTEM, *73370 Le Bourget-du-Lac, France*

[3]Institut de Physique du Globe de Paris, Sorbonne Paris Cité, *75005 Paris, France*

Content of this file

Corresponding author: A. Nouibat, (ahmed.nouibat@univ-grenoble-alpes.fr)

Text S1: Glitch removal

After detrending and demeaning the daily records, we perform a global analysis of the vertical and pressure components to check that the glitches are continuously present and that their periodicity is stable with time. At this stage, we measure the average glitch periodicity on both components. Since the periodicity fluctuates slightly from day to day, we measure the deviation from the average periodicity for each daily record and design a Dirac comb signal of the same periodicity for each record.

This Dirac comb signal is firstly used to extract an average glitch signal by cross-correlation with the corresponding daily record. A synthetic glitch time series is created by convolving the average glitch signal with the Dirac comb. The peak amplitude of the glitches is reduced by subtracting this synthetic glitch time series from the daily records. Then, we further reduce the residual artifacts by matching each glitch individually. This is achieved by searching the combination of time shift and amplitude with respect to average that best fits each residual glitch. In that aim, we compute the least-square solution of the linear system:

$$GA = r \longmapsto \left[g_-(t)\ \ g_+(t)\right] \begin{bmatrix} A_- \\ A_+ \end{bmatrix} = r(t) \tag{1}$$

where g_-, g_+ are one left-shifted and one right-shifted average glitch and A_-, A_+ their amplitudes adjustment factor. r is the residual signal that we want to erase. Finally, we subtract from r the matrix product of G and the least-square solution (A_{LSQR}).

Text S2: Seafloor noise reduction

We use transfer functions to correct the vertical-component record by removing coherent signals derived from the three other components:

$$\tilde{Z}(f) = Z(f) - T^*_{\mathbf{z},\mathbf{r}}(f)R(f) \tag{2}$$

Z and R are Fourier transforms of the vertical channel $\mathbf{z}$ and the component $\mathbf{r} = \{\mathbf{h1}, \mathbf{h2}, \mathbf{p}\}$, which is either the pressure channel $\mathbf{p}$, or one of the two horizontal channels $\mathbf{h1}$, $\mathbf{h2}$. $\tilde{Z}$ is the Fourier transform of the clean signal $\tilde{z}$. T^* is the conjugate of the transfer function between $\mathbf{z}$ and $\mathbf{r}$ and T^*R is the coherent signal between the two components. f is the frequency.

The transfer function T is computed in the frequency domain by quantifying the coherent function C between $\mathbf{z}$ and $\mathbf{r}$:

$$\begin{aligned} T_{\mathbf{z},\mathbf{r}}(f) &= C_{\mathbf{z},\mathbf{r}}(f)\sqrt{\frac{D_{\mathbf{z},\mathbf{z}}(f)}{D_{\mathbf{r},\mathbf{r}}(f)}}, \\ C_{\mathbf{z},\mathbf{r}}(f) &= \frac{D_{\mathbf{z},\mathbf{r}}(f)}{\sqrt{D_{\mathbf{z},\mathbf{z}}(f)D_{\mathbf{r},\mathbf{r}}(f)}} \end{aligned} \tag{3}$$

where $D_{\mathbf{z},\mathbf{z}}$ and $D_{\mathbf{r},\mathbf{r}}$ are the auto-spectral density functions of $\mathbf{z}$ and $\mathbf{r}$, and $D_{\mathbf{z},\mathbf{r}}$ is the cross-spectral density function.

Table S1 and Figures S1 to S8

Period band	Land-Land	Land-OBS	Land-Cleaned OBS
$5s - 10s$	3.85	3.03	3.54
$10s - 20s$	4.32	3.57	4.21
$20s - 40s$	4.03	3.18	3.56
$40s - 70s$	3.58	2.87	3.11

Table S1. Signal-to-noise ratio in different period bands for: (1) correlations between land-land stations pairs, (2) correlations between land stations and OBSs raw signals, and (3) correlations between land stations and OBSs pre-processed signals.

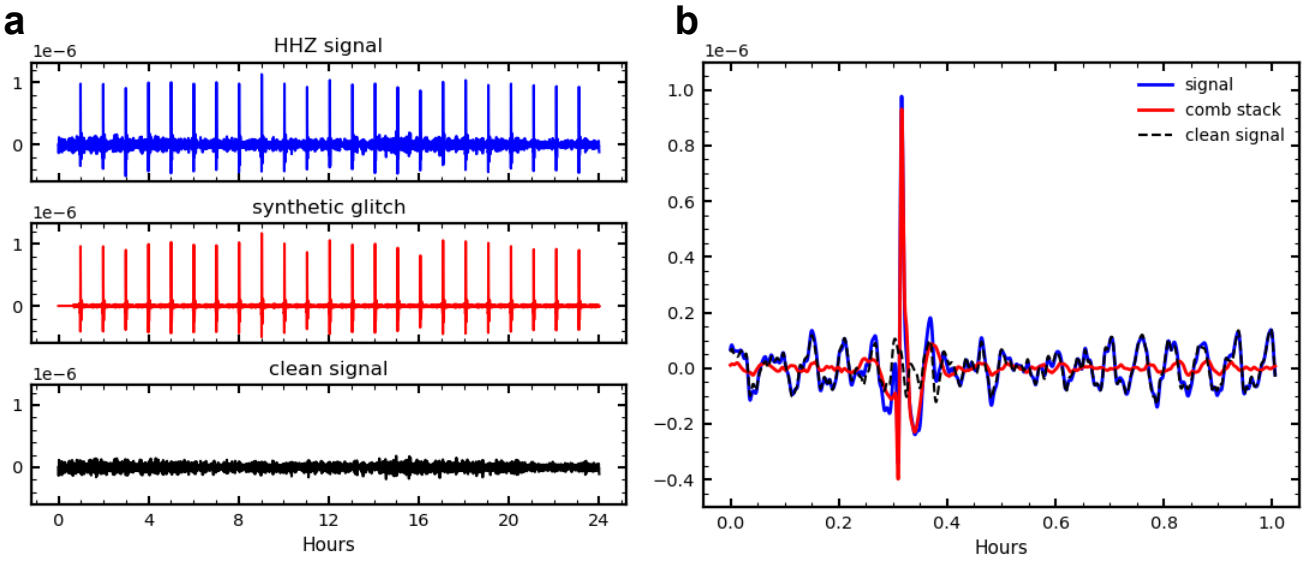

Figure S1. Example of result of the glitch removal process applied to a vertical-component 1-day record of the sea-bottom station A416A. (a) Blue: initial record band-pass filtered between 2 and 30 mHz to reduce the micro-seismic peaks and enhance the 1-hr period glitches; Red: synthetic glitch time series used to clean the signal; Black: cleaned record. (b) Blue: 1-hr length slice containing waveform of the first glitch; Red: waveform of the average glitch used to construct the synthetic glitch time series (red trace in a); Black: signal after correcting for the glitch wavelet. Note that: (1) the initial and cleaned signals are quite coherent on both sides of the glitch peak amplitude; (2) phase and amplitude are stable in the vicinity of the removed glitch. This illustrates the efficiency of the glitch removal processing.

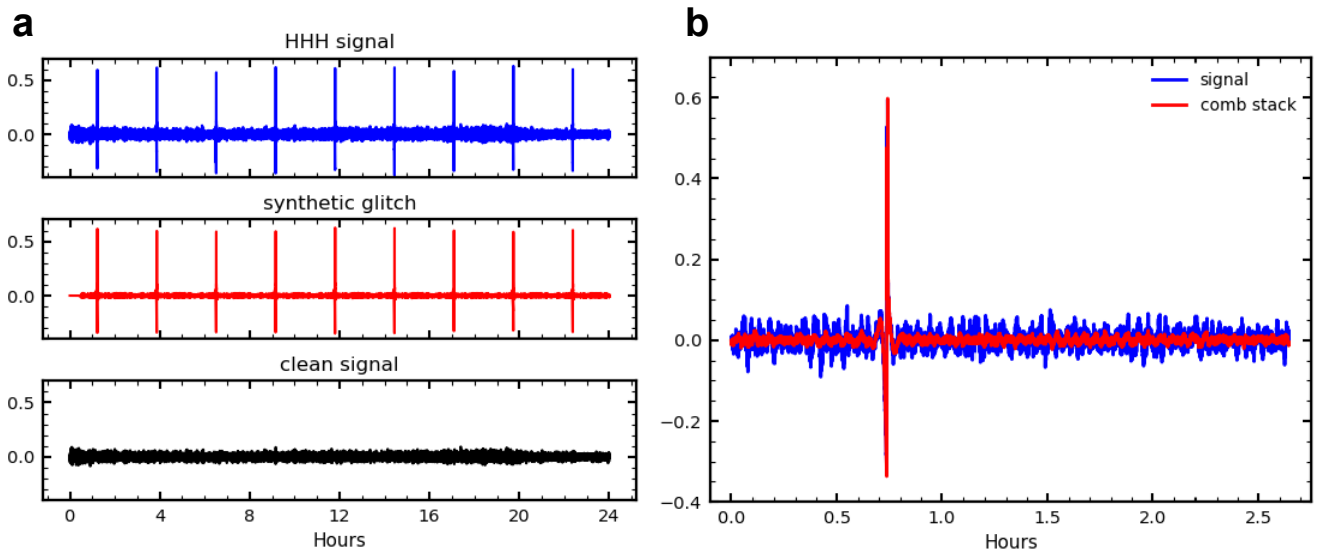

Figure S2. Example of result of the glitch removal process applied to a pressure-component 1-day record of the sea-bottom station A416A. (a) Blue: initial record band-pass filtered between 2 and 30 mHz to reduce the micro-seismic peaks and enhance the 2.65-hr period glitches; Red: synthetic glitch time series used to clean the signal; Black: cleaned record. (b) Blue: 2.65-hr length slice containing waveform of the first glitch; Red: waveform of the average glitch used to construct the synthetic glitch time series (red trace in a).

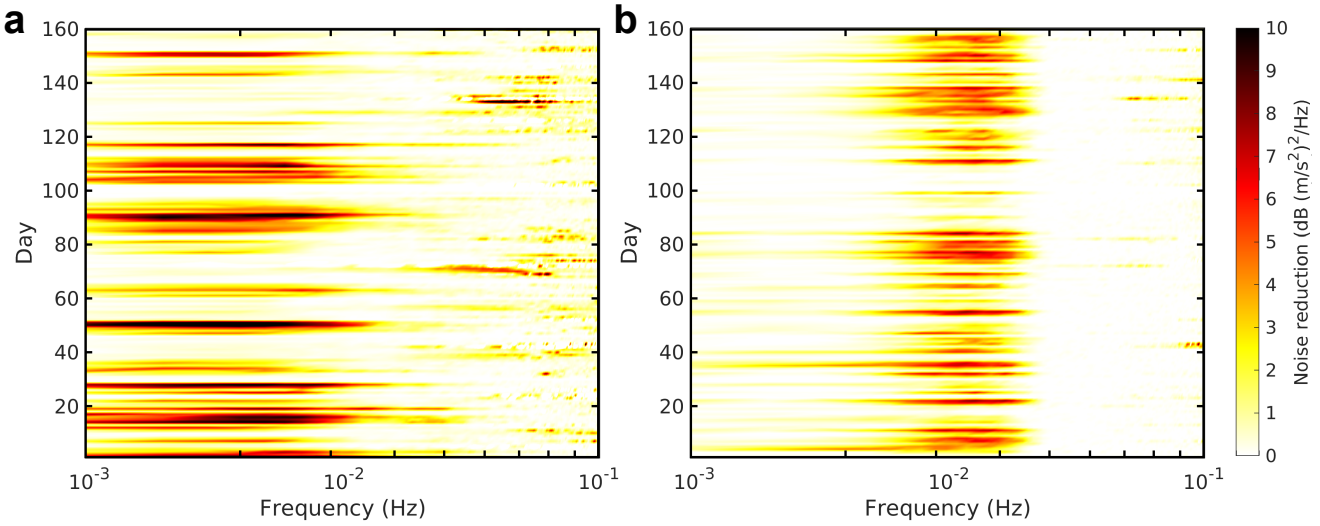

Figure S3. Reduction of the noise level after correction from the tilt- (a) and the compliance- (b) induced noises of a 5-month vertical-component record of OBS A416A. For each 1-day record, the amplitude of the reduction is computed from the difference between power spectral densities (PSDs) before and after correction. The frequency range that is most impacted by the noise reduction due to tilt correction (a) is indicated by the gray area in the example of Fig 2b in the main text. The frequency band most impacted by the correction for compliance (b) is shown as the green area in the example of Fig 2b in the main text. Note that the tilt affects signals over a wide band (below 5×10^{-2} Hz) while the compliance effect is more focused around 1.2×10^{-2} Hz. The noise reduction due to compliance correction is weaker than for tilt correction. However, we observe that tilt is maximum when compliance is minimum and vice versa. During periods of strong tilt (i.e., strong seabed currents), the tilted vertical sensor is less sensitive to signals induced by pressure variations in the vertical direction (i.e., lower compliance effect).

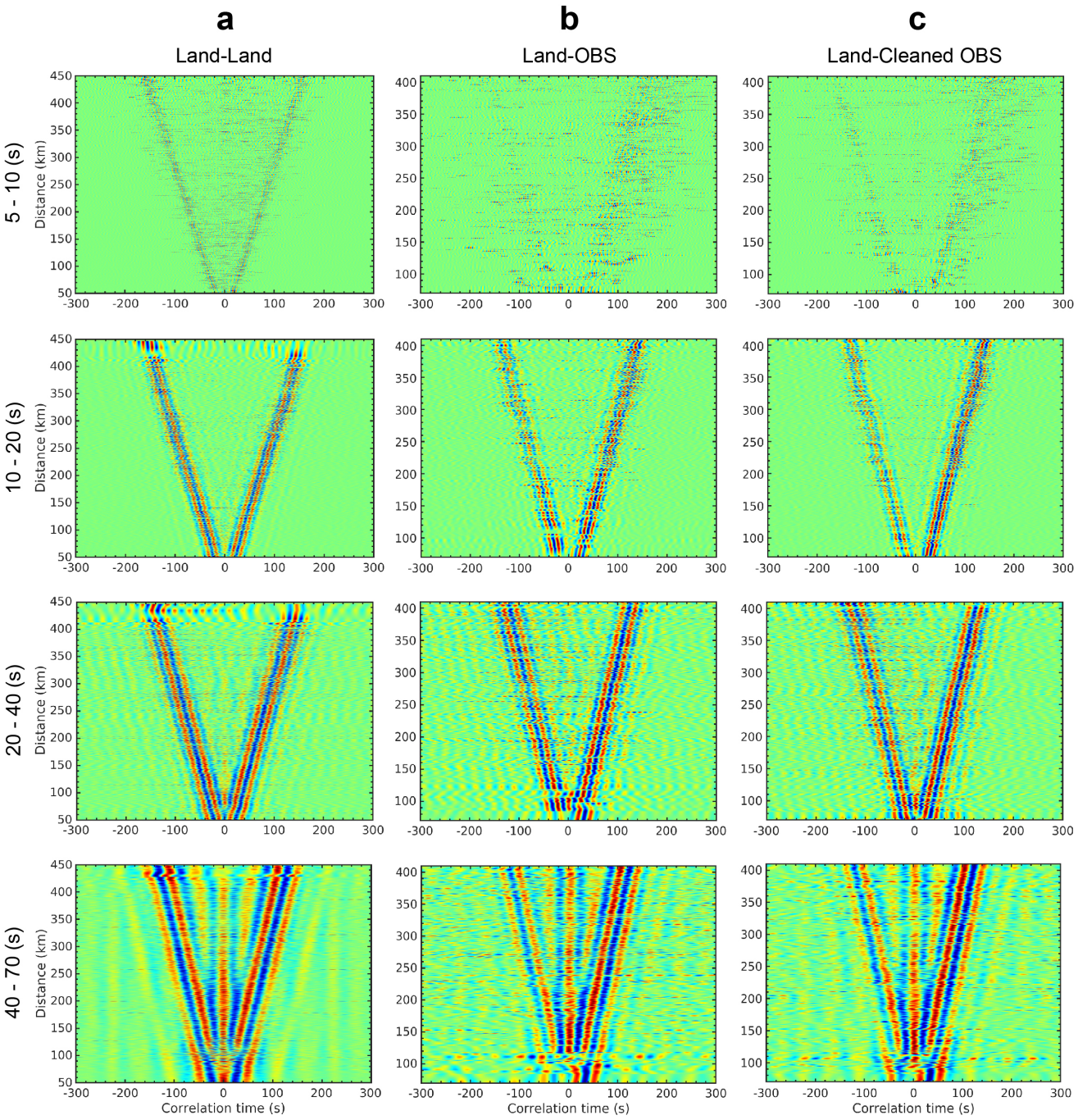

Figure S4. Time-distance plots of correlation signals obtained in different period bands. (a) Correlations between land-land stations pairs. (b) Correlations between land stations and OBSs raw signals. (c) Correlations between land stations and OBSs pre-processed signals (glitch removal & seafloor-noise reduction).

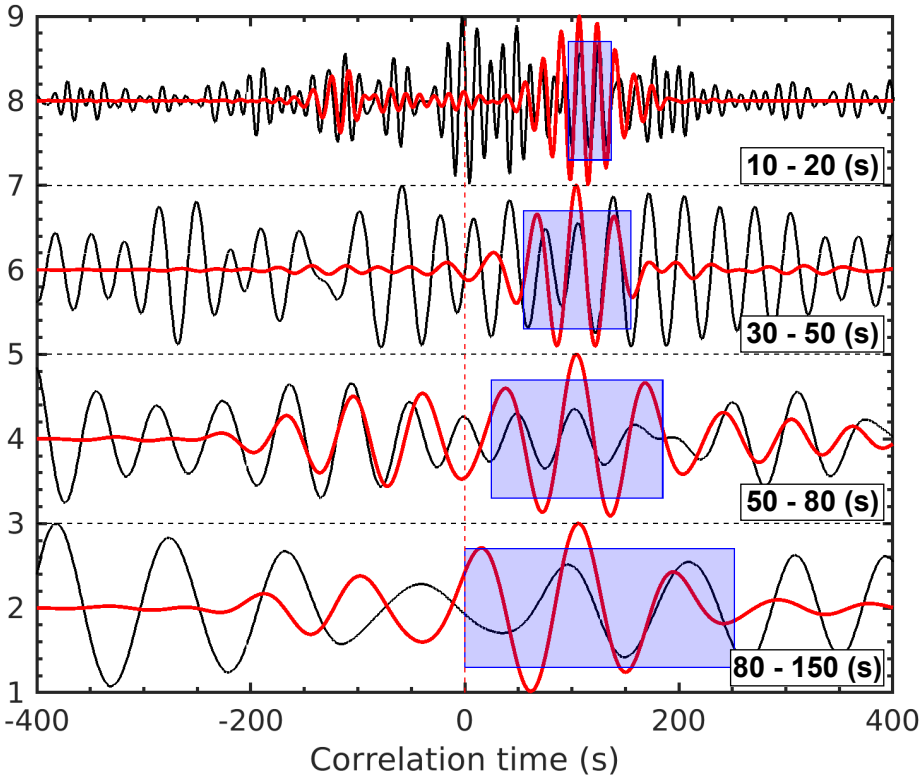

Figure S5. Example of comparison between Rayleigh waveforms reconstructed from the first order correlation or "C^1" (in black) and the iterative correlation or "C^2" (in red). Signals are normalized and filtered in four different period bands.

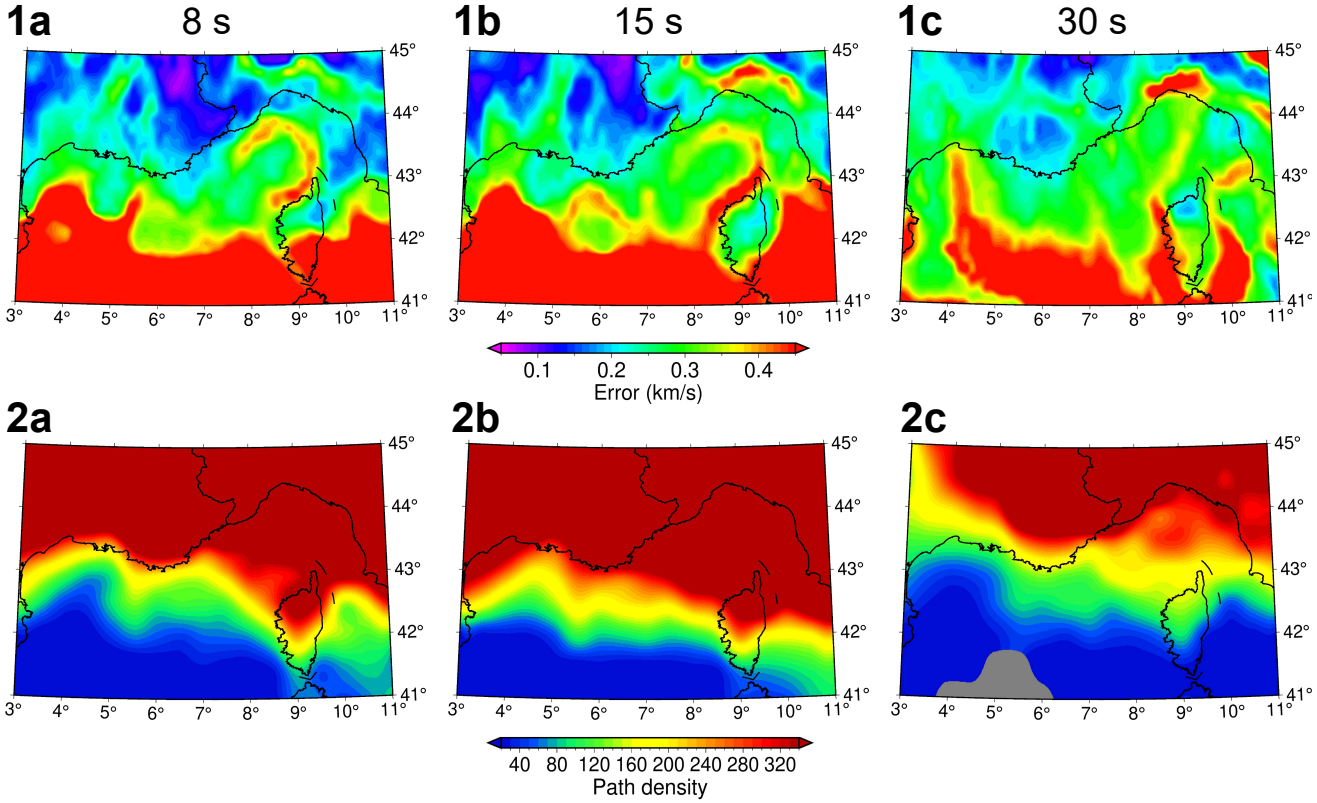

Figure S6. Maps of: (1) estimated error of the group-velocity estimates (standard deviation of the ensemble of sampled velocities) at (1a) 8 s, (1b) 15 s and (1c) 30 s; (2) path density (number of paths crossing each $0.15° \times 0.15°$ cell) at (2a) 8 s, (2b) 15 s and (2c) 30 s. Except for regions with strong velocity contrast (see Fig. 5 of the main text), uncertainty is generally low where ray coverage is high. Note that the Ligurian-Provence basin is covered by no less than 40 - 50 paths per cell, and that coverage increases in the northwestern part of the basin.

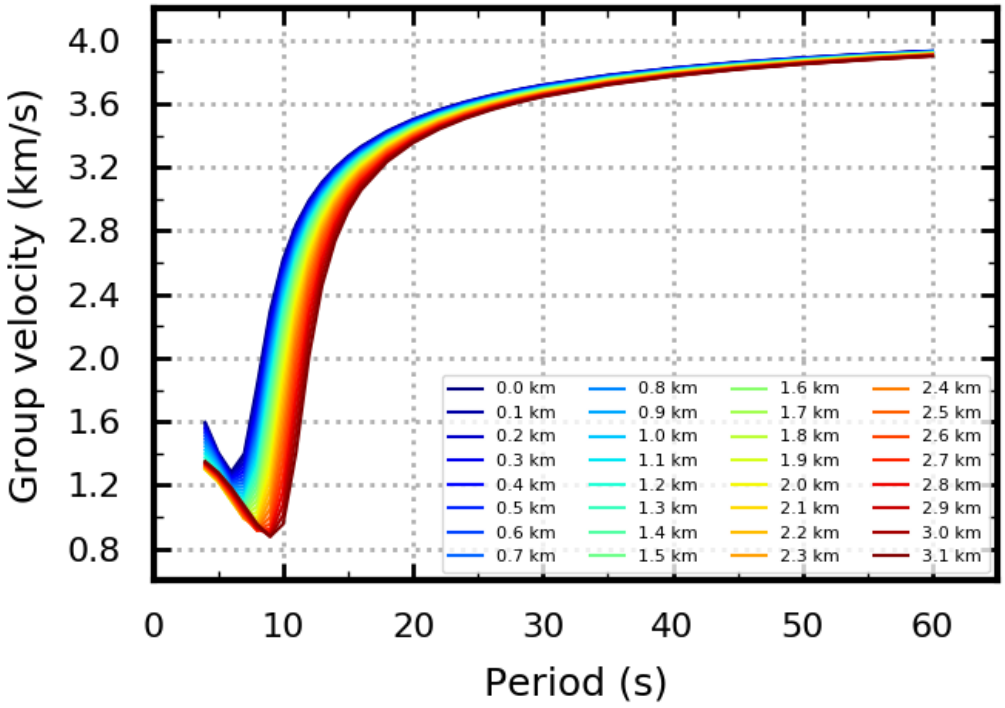

Figure S7. Influence of the water layer thickness on a synthetic Rayleigh-wave group-velocity dispersion curve at an offshore location. Dispersion curves are computed for a 3-layer crustal model consistent with a typical oceanic crust of the Ligurian-Provence basin. The underlying uppermost mantle is represented by a half space. Colors indicate dispersion curves computed with different water levels, from 0 to 3.1 km. The presence of the water layer and its thickness variations influence absolute group velocities and the shape of the dispersion curve, particularly at periods shorter than 15 - 20 s. These changes are for instance strong in the vicinity of the Airy phase (minimum of the dispersion curve).

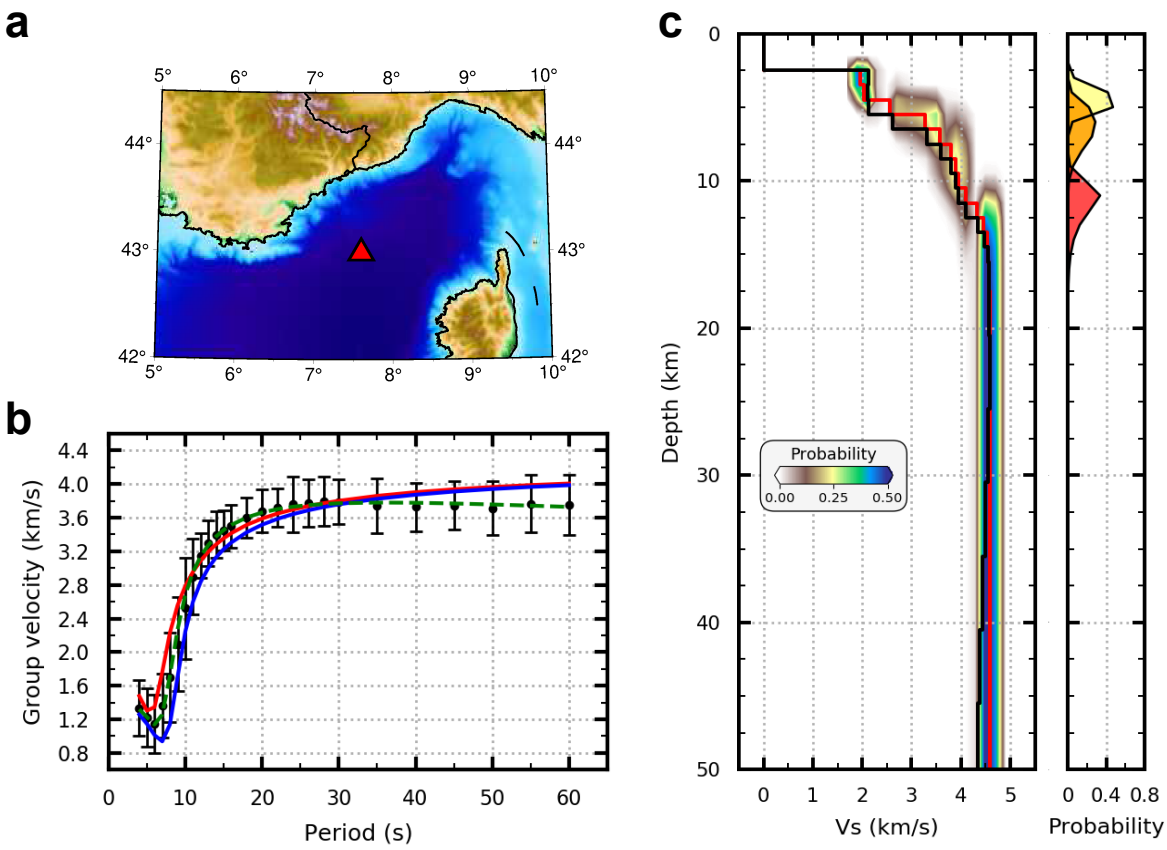

Figure S8. Example of the inversion process for V_s at an offshore location in the Ligurian sea. (a) Location map of the selected grid point (red triangle). (b) Group-velocity dispersion curve of Rayleigh-wave for the selected grid point; Black: observed dispersion curve with its uncertainties; Red: predicted dispersion curve for the average Bayesian model; Blue: predicted dispersion curve for the average Bayesian model with water layer on top; Green: predicted dispersion curve after the linear inversion. Note that the average Bayesian model does not fit well the long-period part of the dispersion curve (periods $>$ 30 s) because of the half-space mantle assumption used in the grid search. The fit is further improved after the linear inversion. (c) Resulting velocity models; Left: posterior probability distribution on shear-wave velocity from the Bayesian inversion (background colors), weighted average solution model (posterior mean) predicted from the probabilistic scheme (red line) and final solution model predicted from the linear inversion (black line). Right: posterior probability distribution on layer boundaries resulting from the Bayesian inversion. Three maxima at $\sim$ 5 km, $\sim$ 6.5 km and $\sim$ 11 km depths correspond respectively to the base of the Messinian salt unit, to the base of the sediment layer and to the Moho.