A study of the seismic noise from its long-range correlation properties

L. Stehly,1,2 M. Campillo,1 and N. M. Shapiro3,4

Received 21 December 2005; revised 5 May 2006; accepted 21 June 2006; published 14 October 2006.

We study the origin of the background seismic noise averaged over long time by cross correlating of the vertical component of motion, which were first normalized by 1-bit coding. We use 1 year of recording at several stations of networks located in North America, western Europe, and Tanzania. We measure normalized amplitudes of Rayleigh waves reconstructed from correlation for all available station to station paths within the networks for positive and negative correlation times to determine the seasonally averaged azimuthal distribution of normalized background energy flow (NBEF) through the networks. We perform the analysis for the two spectral bands corresponding to the primary (10–20 s) and secondary (5–10 s) microseism and also for the 20–40 s band. The direction of the NBEF for the strongest spectral peak between 5 and 10 s is found to be very stable in time with signal mostly coming from the coastline, confirming that the secondary microseism is generated by the nonlinear interaction of the ocean swell with the coast. At the same time, the NBEF in the band of the primary microseism (10–20 s) has a very clear seasonal variability very similar to the behavior of the long-period (20–40 s) noise. This suggests that contrary to the secondary microseism, the primary microseism is not produced by a direct effect of the swell incident on coastlines but rather by the same process that generates the longer-period noise. By simultaneously analyzing networks in California, eastern United States, Europe, and Tanzania we are able to identify main source regions of the 10–20 s noise. They are located in the northern Atlantic and in the northern Pacific during the winter and in the Indian Ocean and in southern Pacific during the summer. These distributions of sources share a great similarity with the map of average ocean wave height map obtained by TOPEX-Poseidon. This suggests that the seismic noise for periods larger than 10 s is clearly related to ocean wave activity in deep water. The mechanism of its generation is likely to be similar to the one proposed for larger periods, namely, infragravity ocean waves.


1. Introduction

[2] It has been recently demonstrated that the time cross-correlation function of random seismic wavefields such as seismic coda [Campillo and Paul, 2003] or seismic noise [Shapiro and Campillo, 2004] computed between a pair of distant stations contains, at least partially, the actual Green function between the two stations [Campillo, 2006]. This provides us with a possibility to retrieve the propagation properties of deterministic seismic waves along long paths by analyzing microseisms only. The emergence of the Green function is effective only after a sufficient averaging. In the case of diffuse coda waves, the averaging is performed over a set of earthquakes [Campillo and Paul, 2003; Paul et al., 2005]. With the seismic noise (in the following, we use the term noise for the microseism which actually have no relation with instrumental noise), it is assumed that the averaging is provided by randomization of the noise sources when considering long time series [Shapiro and Campillo, 2004; Sabra et al., 2005a]. Another important process contributing to the randomization is the scattering of seismic waves on heterogeneities within the Earth that is significantly strong at periods less than 40 s. Reconstruction of Rayleigh waves from the seismic noise is sufficiently efficient and accurate to lead to high-resolution imaging at the regional scale [Shapiro et al., 2005; Sabra et al., 2005b]. Further optimization of seismic imaging based on noise correlation requires better understanding of the origin of the seismic noise and of the spatial and temporal distribution of its sources [Pederson et al., 2006; Schulte-Pelkum et al.,...
2004]. In particular, it is important to establish conditions under which the noise can be considered as well randomized. To be more precise, a perfect randomization is not necessary but at least a distribution of sources covering a sufficiently large surface is required when integrating over time.

[5] Ambient seismic noise is mostly made of surface waves [e.g., Friedrich et al., 1998; Ekström, 2001]. Therefore its sources are likely close to the Earth’s surface. Observed noise amplitudes cannot be explained by the background seismicity [Tanimoto and Um, 1999] and main noise sources are believed to be loads caused by pressure perturbations in the atmosphere and the ocean. Moreover, the mechanisms of generation of seismic noise are not the same in different period bands. At relatively short periods (<20 s), the two strongest peaks of the seismic noise, i.e., the primary and the secondary microseisms, are believed to be related to the interaction of the sea waves with the coast [Gutenberg, 1951]. The primary microseism has periods similar to the main swell (10–20 s), while the secondary microseism that is the strongest peak in the noise spectrum originates from the nonlinear interaction between direct and reflected swell waves that results in half period (5–10 s) pressure variations [Longuet-Higgins, 1950]. This interaction results in variations of pressure at the sea bottom that do not exhibit the rapid exponential decay with depth expected for primary gravity waves.

[4] The long-period noise or “the hum” has been shown to exhibit a spectra corresponding to the normal modes of the Earth [Nawa et al., 1998; Suda et al., 1998; Tanimoto et al., 1998; Roult and Crawford, 2000; Kobayashi and Nishida, 1998; Nishida et al., 2000]. The origin of the long periods has been attributed to the so-called infragravity waves, a ocean wave mode that exists at long period and which has been studied for its role in sediment transport in coastal zone. According to Webb et al. [1991], infragravity waves propagate in free waters, and result in long-period pressure fluctuations at the ocean bottom. For long-period noise, Tanimoto [2005] ruled out the effect of atmospheric pressure variations since it was observed by Watada et al. [2001] to be much smaller than pressure at the ocean bottom for periods larger than 70 s. Rhie and Romanowicz [2004] and Tanimoto [2005] proposed the infragravity waves as the source of the long-period noise. Observations at an ocean bottom broadband station [Dolenc et al., 2005] show a strong link between infragravity waves and the local level of excitation of shorter-period ocean waves. This suggests a local generation of the long-period infragravity waves from the primary ocean swell. Dolenc et al. [2005] also observed a correlation of the amplitude of the infragravity waves with tides, a point that could be related to the interaction of waves and currents [see Longuet-Higgins and Stewart, 1964; Kobayashi and Nishida, 1998; Nishida et al., 2000]. Noise excitation also exhibits strong seasonal variations. Using array analysis, Rhie and Romanowicz [2004] have shown that sources of the long-period (150–500 s) seismic noise are dominantly located in the Northern Hemisphere oceans during the northern winter and migrate to the Southern Ocean during the southern winter. This behavior is well correlated with the seasonal variation of the amplitudes of ocean waves suggesting that the “hum” is produced by some sort of atmosphere-ocean-seafloor coupling.

[5] In the present paper, we study the origin and the seasonal variability of the relatively short-period noise (between 5 and 40 s) with a particular emphasis on the primary microseism (10–20 s band). One of our main motivations is that a better understanding of the distribution of noise source in space in time is needed for optimization of the noise-based seismic tomography. Using several networks in North America, Africa, and Europe, we determine the direction of the average azimuthal distribution of normalized background energy flow (NBEF) across each array by measuring the degree of symmetry of time cross correlations computed between pairs of stations and locate the apparent noise sources.

[6] Results of our analysis show that while the sources of the secondary microseism remain stable in time, the sources of the primary microseism exhibit strong variability very similar to long period noise (hum) and well correlated with sea wave conditions.

2. Asymmetry of Cross Correlation

[7] The idea of using random noise to reconstruct the Green function has already been applied successfully in various fields of physics such as helioseismology [e.g., Duvall et al., 1993; Gilles et al., 1997], acoustics [Weaver and Lobkis, 2001], or oceanography [Roux and Kuperman, 2003]. In seismology, Aki [1957] already proposed to use the noise to retrieve the dispersion properties of the subsoil. Shapiro and Campillo [2004] reconstructed the surface wave part of the Green function by correlating seismic noise at stations separated by distance of hundreds to thousands of kilometers, and measured their dispersion curves at periods ranging from 5 to about 150 s. Later this method has been used for seismic imaging in California [Shapiro et al., 2005]. In the case of a spatially homogeneous distribution of noise sources, the cross correlation is expected to be nearly symmetric in amplitude and in arrival time with its positive and negative parts corresponding to the Green function of the medium and its anticausal counterpart, respectively [e.g., Lobkis and Weaver, 2001; Van Tiggelen, 2003; Snieder, 2004; Sánchez-Sesma and Campillo, 2006]. In practice, as we will see below, the causal and anticausal parts of the cross correlation may strongly differ in amplitude. This amplitude factor depends directly on the energy flux of the waves traveling from one station to the other [Van Tiggelen, 2003; Paul et al., 2005]. In others words, in the case of a perfectly isotropic distribution of sources, the energy flux between two stations is the same in both directions and the resulting cross correlation between these stations is symmetric (Figure 1a). On the other hand, if the density of sources is larger on one side than on the other, the amounts of energy propagating in both directions are different. In this case, the resulting cross correlation is not symmetric anymore in amplitude (although the arrival time remains the same) (Figure 1b). An important consequence is that the asymmetry of the cross correlation computed between several pairs of stations of a network can be used to measure the main direction of the energy flux across the array. Making such
measurements at different arrays will allow us to determine the location of main sources of the seismic noise.

3. Origin of Seismic Noise Observed in California

We first consider one pair of stations in California (MLAC and PHL, Figure 2a). We analyze 1 year (2003) of continuous vertical records. Before computing cross correlations, records are corrected from the instrumental response and band passed within different bandwidths. To reduce the contribution of the most energetic arrivals, we disregard completely the amplitude and consider 1-bit signals only [Derode et al., 1999; Campillo and Paul, 2003; Shapiro and Campillo, 2004]. Figure 2 shows the cross correlations of different months of records band-passed between 5 and 10 s, i.e., around the secondary microseism. Positive time delay indicates waves propagating from MLAC to PHL, whereas negative time indicates waves propagating from PHL to MLAC. In this period range, the form of cross correlations is very stable in time and very asymmetric. The amplitude of the anticausal part of the cross correlations is much larger than the one of the causal part. The main arrival is the fundamental Rayleigh wave. The Green function is poorly reconstructed in the causal part. This indicates that most of the noise is propagating from PHL to MLAC, i.e., from the coastline to the continent. This confirms that at periods between 5 and 10s most of the noise is dominated by waves generated by nonlinear interaction between the swell and the coastline.

The behavior is very different when considering periods between 10 and 20 s (Figure 3). The cross correlations in this band exhibit a clear seasonal variation. During the northern winter (October to March), the amplitude of the causal part of the correlation is larger than the amplitude of the anticausal part. This indicates that most of the Rayleigh wave energy propagates from MLAC to PHL (NE to SW). During the northern summer (May-September), the opposite is observed: the noise is dominated by waves propagating from PHL to MLAC (SW to NE). Moreover, during the whole year, the Rayleigh waves are visible both at positive and negative times. All this shows that in the period band of the primary microseism, an important contribution of the noise observed in California is coming from the east, having likely its source in the Atlantic Ocean. The waveforms

Figure 1. Schematic illustration of the effect of inhomogeneous noise sources distribution on the degree of symmetry of cross correlation. (a) Symmetric cross correlation between 1 and 2 obtained when the sources of noise are evenly distributed. (b) Asymmetric cross correlation (but symmetric travel times) associated with a nonisotropic distribution of sources.
observed at positive and negative times are not exactly identical because of the differences in the spectrum of the noise coming from east or west.

To determine the main direction of the normalized background energy flow (NBEF), we used 23 stations located in southern California separated by distances of a few hundreds kilometers (Figure 4). We kept only paths longer than 120 km (two wavelengths at 20 s) and shorter than 450 km. This resulted in 136 paths or 272 azimuths when using both the causal and the anticausal parts.

Using the procedure described above, we computed cross correlations of continuous vertical records for each of the 136 paths during the year 2003. We considered three period bands: 5–10, 10–20, and 20–40 s. For the first two period bands, cross correlations were stacked in a moving window of 15 days. For each stack, amplitudes of the causal and anticausal parts were determined by taking the maximum of their envelopes in a time window corresponding to the Rayleigh wave group velocity. We correct our measurement from the geometrical attenuation of the Rayleigh wave with distance, by multiplying the amplitude by the square root of the interstation distance. This way, we measured normalized amplitudes of seismic noise for two azimuths from each cross correlation. Combining measurements from all stations pairs we obtain the distribution of the normalized amplitude with respect to azimuth. Maxima of this distribution indicate main directions of NBEF across the array. Seasonal variation of the normalized amplitude and the direction of the NBEF during the year 2003 are illustrated in Figures 5a and 5b. Amplitudes were normalized between 0 and 1 for the entire set of azimuths. One must remark that these azimuthal distributions were computed after 1-bit normalization. This means that high-amplitude events, likely associated with the strongest storms are down weighted by our processing. Figures 5a 5b 5c–5d are not directly characteristic of the actual absolute noise.
energy but of the time-averaged normalized energy. Note that it is the relevant measure for analyzing the noise in the context of Green function reconstruction using the procedure initiated by Shapiro and Campillo [2004]. This procedure can be applied to networks with spatial distribution of station which are not suitable for F-K analysis.

[12] Between 5 and 10s, the main direction of the NBEF remains constant over the whole year similar to the observation made for the path MLAC-PHL. Most of the noise is coming from a range of azimuth going from 200° to 225° (SW) confirming that in this period band, most of the seismic noise is caused by the secondary microseism locally generated by ocean waves along the coastline.

[13] The noise behavior is very different in the period range of the primary microseism (10–20 s) where the NBEF exhibits a very strong and sharp seasonal variation (Figures 5c and 5d). Its main direction remains around 220°–225° (S-SW) during the summer. However, during the winter, most of the noise is coming from the north with two clear preferential directions: 315° and 45°. The change between these two main regimes is very rapid and takes place in March and October.

Figure 4. Location of seismic stations used in the network analysis in California.

Figure 5. (a) Normalized amplitude of the cross correlation of seismic noise at periods between 5 and 10 s as a function of time and azimuth determined from the network analysis in California. (b) Location map of the used stations. Main directions of the NBEF during winter and summer are shown with blue and red arrows, respectively. (c) and (d) Same as Figures 5a and 5b at periods between 10 and 20 s.
Figure 6. Normalized amplitude of the cross correlation averaged during (left) the winter and (right) the summer versus azimuth for various frequency bands (a) 5–10 s, (b) 10–20 s, and (c) 20–40 s. The color scale is the same as on Figure 5.
This difference between the directions of NBEF of the primary and secondary microseisms is surprising because it indicates that the two main spectral noise peaks observed locally do not have the same region of origin. This could be the result of the attenuation of the seismic waves which is stronger for shorter periods and cancels the contributions of distant sources in the 5–10 s period band. Another explanation could be that the actual regions of generation of the seismic noise is different in the different period bands. This hypothesis is supported by the observation that the azimuthal distribution is dominated in winter by a flux from azimuths between 180 and 225° for the period band 5–10 s while there is no significant contribution from this directions for the period band 10 to 20 s. Before exploring further the origin of noise in the primary microseism band (10–20 s), let us consider longer period, i.e., the 20–40 s period band.

While at periods between 5 and 10 s most of the noise is coming from the coast during the whole year, the noise provenance has a clear seasonal variability at longer periods. Moreover, the normalized amplitude versus azimuth for the 10–20 and 20–40 s period bands exhibit main features which are similar, although not completely identical. Most of the background noise energy is coming from the northeast (possibly North Atlantic) during the winter and the main direction switches to the southwest during the summer. This similarity suggests that the average primary microseism may originate from the same regions as the longer-period noise, which has been considered to be excited by the infragravity ocean waves.
propagating away from the coast [Webb et al., 1991; Rhie and Romanowicz, 2004; Tanimoto, 2005]. We will investigate further the origin of the noise in the band of the primary microseism by considering simultaneously several networks.

4. Origin of the 10–20 s Noise

[17] Using the method described in section 3, we analyzed the noise in several other regions of the world: east coast of United States, western Europe, and Tanzania (Figure 7). Similar to the Californian network, our analysis is based on the amplitudes of the reconstructed causal and anticausal Rayleigh waves. The data for the Northern Hemisphere are from 2003 and the Tanzanian data are from 1994 to 1995, the period of the 97-005 PASSCAL experiment [Owens et al., 1997]. For each of these networks we estimated the normalized noise amplitudes as functions of azimuth during the winter (October to March) and the summer (April to September) for the period bands of the primary and the secondary microseisms. Similar to the observations in California, the NBEF of the 5–10 s noise remains stable over the whole year at all networks. This observation is compatible with the idea that the generation of the secondary microseism is mostly controlled by the bathymetry and the geometry of the coastlines. The noise in this period is dominated by the activity in a region close to the station since the absorption limits the contribution of distant source zones. It explains the stability of the distribution during the year.

[18] The behavior of the 10–20 s noise is very different. Indeed the effect of attenuation is weaker for longer periods and sources at the global scale are contributing. In this period band, the regions where the noise is originating are different during the summer and the winter. A first-order observation from Figure 8 is that the major part of the noise seems to come from the north during winter, and from the south during the summer.

[19] In spite of their limited number, the considered networks illuminate a large part of the Earth surface. In the following, we use them for a rough identification of the source regions that generate the 10–20 s noise consistently recorded by the four networks. We use a simple back projection procedure. For every considered pair of stations of a given network, we trace the great circle connecting the stations and divide it in two equal parts corresponding to two opposite directions of wave propagation between the stations. Normalized amplitudes measured from positive and negative parts of the cross correlation computed between these stations are attributed to the corresponding half

Figure 8. Average normalized amplitude of cross correlations of noise versus azimuth, for various networks (a) during winter between 5 and 10 s, (b) summer 5–10 s, (c) winter 10–20 s, and (d) summer 10–20 s.
great circle. This is an approximation since in the 10–20 s period band the path followed by Rayleigh waves is not exactly a great circle because of the lateral heterogeneity of the crust and the mantle. We then define a $2^\circ \times 2^\circ$ grid on the surface of the Earth and take the mean value of the normalized correlation amplitude associated with all great circles crossing every cell. To be able to use all four networks simultaneously, we assume that the seasonal noise variability is stable over different years and combine the 2003 data from the Northern Hemisphere with the 1994 data from Tanzania (results for individual networks are given in Appendix A).

The results for the four networks in the 10–20 s band are presented in Figure 9. During the Northern Hemisphere winter, the observed average noise is apparently dominated by a source zone in the northern Atlantic Ocean. Two weaker sources are visible in the northern Pacific, one close to Alaska and the other close to Japan. During the Northern Hemisphere summer, most of the noise originates from the southern Indian Ocean and the southern Pacific. It has to be noted that the resolution of these patterns is limited by the small number of networks. An example is the band in the Pacific Ocean which is due to the fact that the source zone, the California, eastern United States, and western Europe networks are on the same great circle. In spite of the limited resolution, these results clearly show that contrary to the secondary microseism, the 10–20 s noise is not generated locally but is excited by sources acting at the global scale and having a clear seasonal variability. An attempt to locate these sources suggests they have to be searched in the central parts of the oceans, although with our limited resolution we cannot exclude that a part is generated on the coast or in shallow water areas. Specifically, with the form of normalization and averaging used here, we do not contradict the observation that the arrival of strong storms along the coasts is associated with high-amplitude microseisms. The hypothesis that the average background noise energy originates in the zones of storm activity is confirmed by strong similarities between the maps of the apparent sources of the 10–20 s seismic noise and the global wave height maps obtained from the radar altimeter data collected by the satellite TOPEX-Poseidon (Figure 9). Both maps show clear seasonal variations with maxima located in deep oceans in the Southern and the Northern hemispheres during the summer and the winter, respectively. A similar seasonal variability was recently reported for the very long period (>150 s) Earth hum [Rhie and Romanowicz, 2004].

5. Conclusions

Our results clearly demonstrate that the origins of the primary and the secondary microseisms are different. Sources of the secondary microseism are stable in time and likely associated with coastlines, confirming that this part of the seismic noise is generated by the nonlinear interaction of the ocean swell with the coast. In the same time, the primary microseism exhibits a very clear seasonal variability very

Figure 9. Comparison between seasonal variations of the location of seismic noise sources and significant wave height. (a) and (c) Geographical distribution of the mean back-projected normalized amplitude of the 10–20 s noise cross correlations (see text for detailed explanation of the back projection) during the winter and the summer. (b) and (d) Global distribution of the square of wave height measured by TOPEX/Poseidon during the winter and the summer.
Figure A1. Origin of the background noise at 10–20 s sensed during the winter by the different networks for (a) California, (b) east coast of the United State, (c) western Europe, and (d) Tanzania.

Figure A2. Origin of the background noise at 10–20 s sensed during the summer by the different networks for (a) California, (b) east coast of the United States, (c) western Europe, and (d) Tanzania.
kinds of mechanisms of generation are compatible with the location of background noise source regions in the zones of highest sea waves.

Appendix A

[24] Figures A1 and A2 show the apparent origin of the background noise seen by the different networks in winter and summer respectively. The California network is the denser and is therefore the most favorable for the reconstruction of the Rayleigh waves. In order to test the importance of the California network in the maps of Figure 9, we present the same map computed without including it. The resulting maps (Figure A3) indicate that the maps built with or without the California network share the same features.

[25] Acknowledgments. All the seismic data used in this study have been obtained at the IRIS Data Management Center (http://www.iris.edu/) and include GSN, GEOSCOPE, GEOFON, Southern California Seismic Network, Berkeley Digital Seismic Network, and Tanzania PASSCAL seismic experiment. We are grateful to all scientists and technical staff who provided the data. The ocean wave height map was obtained from the Physical Oceanography Distributed Active Archive Center (PO.DAAC, http://podaac.jpl.nasa.gov/). This research has been supported by contract DE-FC52-05NA26607 from the National Nuclear Security Administration, by the National Science Foundation grant EAR0450082, by CNRS/Institut National des Sciences de l’Univers (program DyETI), and by the Commissariat a l’Energie Atomique.

References

Campillo, M., and A. Paul (2003), Long-range correlations in the diffuse seismic coda, Science, 299, 547–549.


