Global Trends in Extremal Microseism Intensity

1. Supplementary Material

1.1. Data Selection

We analyze microseism power spectral densities using continuous 1 sample/s digital vertical-component time series (LHZ) from the Global Seismographic Network (GSN) and predecessor networks at sites selected for long continuous recording history and wide geographic distribution (Table S1). Seismic data from long-running GSN and predecessor stations allows for good resolution of SFM spectral power through the entire digital history of the station, early 1970s, and back to the early 1990s to late 1980s for the DFM. All seismic data channels were retrieved from the archives of the Incorporated Research

Richard C. Aster, Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, NM 87801, USA (aster@ees.nmt.edu).

Daniel E. McNamara, United States Geological Survey, Box 25048, Stop 966, Denver Federal Center, Denver, CO 80225, USA.

Peter D. Bromirski, Integrative Oceanography Division 0209, Scripps Institution of Oceanography, 9500 Gilman Drive, La Jolla, CA 92093, USA.

Institutions in Seismology (IRIS) Data Management System (DMS) and from the US Geological Survey (USGS) Albuquerque Seismological Laboratory.

1.2. Power Spectral Density Probability Density Function Processing

The long-term and regional variation of microseism power is observed by computing instrument corrected power spectral density (PSD) probability density functions (PDFs) using the spectral methods presented in *McNamara and Buland* [2004]. Spectral methods follow the original (Welch's) algorithm used to develop the GSN New Low and High Noise Models (NLNM, NHNM) [*Peterson*, 1993]. PSDs are computed from continuous, overlapping (50%) time series segments (LH channels: 3-hour segments sampled at 1 sample/s (10,800 seconds)). All available data are included; there is no removal of earthquakes, system transients or data glitches. The instrument transfer function is deconvolved from each time segment, yielding ground acceleration. Each 3-hour time series segment is divided into 13 subsegments (2,700 seconds), overlapping by 75%. Each sub-segment is processed by: 1) removing the mean, 2) removing the long-period trend, 3) tapering using a 10% cosine function, 4) transforming via fast Fourier transform to obtain the amplitude spectrum, and 5) squaring the amplitude spectrum to obtain the power spectrum [*McNamara and Buland*, 2004]. The final 3-hour PSD estimate is calculated as the average of the 13 sub-segment PSDs. Due to sub-segment averaging, the final PSD estimate has a 95% level of confidence that the spectral point lies within -2.14 dB to +2.87 dB of the estimate. The averaged PSD is then smoothed by computing full octave averages in 1/8-octave intervals resulting in 96 spectral estimates for each 3-hour time segment. 3-hour time windows are chosen for good resolution at long-periods (30 - 1,000 s). Long-periods are critical for frequency dependent culling of earthquakes and sensor/instrument transients that contaminant the small variations that we hope to detect in the SFM and DFM bands. PSDs are then combined into PDFs where distribution of spectral power as a function of frequency/period is then readily visualized (Figure S1).

1.3. Passband Selection

When working with older digital seismic data it is necessary to properly accommodate for changes in recording parameters and instrumentation. To ensure that PSD estimates truly reflect microseism power, significant effort was required to comprehensively identify, understand, and obviate influences of any instrumental artifacts and transients in these bands. Each unique data channel (i.e. IU.ANMO.00.LHZ; Figures S1a, b and HG.ALQ.–.LHZ; Figures S1c, d) will have several instrument epochs throughout its history that may include variations in absolute amplitude due to erroneous estimation of the gain constant from inaccurate calibration methods [*Ringler et al.*, 2010].

With modern GSN and similar seismographic instrumentation, both the short-period DFM (5 - 9 s) and longer-period SFM (14 - 20 s) are well resolved (IU.ANMO.00.LHZ; Figures S1a and b). In a typical example, seismic data from the current instrumentation at IU.ANMO.00.LHZ is collected with a KS54000 borehole seismometer at a depth of 100 m with a Quanterra 680 digitizer. During the 1980s, the SFM and longer-period portion of the DFM became gradually resolvable (SR.ANMO.–.LHZ; Table S1). However, from 1972 through the mid-1980s the DFM is not resolvable due to filtering (HG.ALQ.–.LHZ; Figure S1c, d; Table S1), and we are restricted to longer period SFM band (14 - 20 s) data for this time period. Secondary microseisms are generally observable beginning in the late 1980s to early 1990s for all stations, depending on specific station history.

1.4. Transient Culling

Large shallow earthquakes are efficient radiators of seismic surface waves in the SFM and DFM bands (Figure S2a). Furthermore, microseism signal may be contaminated by sporadic instrumentation problems [*McNamara et al.*, 2009]. Three-hour PSDs are sufficiently short so that power levels are sensitive to surface wave contamination from nearby and distant earthquakes (Figure S1a, c) and recording system transients. We eliminate false microseism events arising from high power transients due to earthquakes and instrumentation problems in the two microseism period bands (SFM, DFM) by computing probability density function (PDF) percentage point statistics over the length of individual instrument epochs. An instrument epoch is typically several years in length and is defined as a time-period in which the instrumentation, station metadata, and sensors were consistent (a single instrument transfer function). PDF percentiles are used as station baselines for culling out high- and low-power transients in the short-term 3-hour PSDs. Reliable rejection of earthquake and nonphysical outlier PSD estimates was achieved when we applied station specific baselines, defined by the (1^{st} through the 80^{th} percentiles of the PDF distribution for each modern instrument epoch [*McNamara et al.*, 2009] (Figure S1a, b). The 5^{th} percentile was often used as the lower bound of the station baseline for older stations with numerous occurrences of instrumentation problems (Figures S1c, d).

Figure S1a, b shows total and post-culling PSD PDFs for IU.ANMO.00.LHZ for an instrument epoch. Also shown, to demonstrate the frequency-dependent culling procedure, are 3-hour PSDs from four earthquakes and two Pacific storms. In general, significant earthquakes at local, regional, and teleseismic distances generate long-period energy that is easily detectable above the (80th percentile of the distribution at periods > 30 s. For example, the earthquakes shown in Figure S1a are easily culled from further microseism power measurements by this discriminant.

Microseism events are generally low power at periods > 30 s and high power in the DFM and SFM bands. By culling PSDs using periods > 30 s, we remove earthquakes and other transients that would bias microseism power levels. This allows for large power level microseism signals in the DFM and SFM bands to remain even if they exceed station baseline upper bound (80th percentile). This is observed from two individual 3-hour PSDs with known southern California wave events from Pacific Storms (12/14/00-green curve and 01/19/01-black curve in Figure S1b).

1.5. Microseism Index Method

Microseism index computation is similar to earthquake detection algorithms that make use of the ratio of short-term versus long-term amplitude ratios. In the case of detecting microseism events, rather than earthquakes, we use short-term SFM and DFM band averages in continuous 3-hour intervals, rather than seconds, and long-term threshold over many years, rather than minutes. The microseism index is defined as the number of event trigger hours per month exceeding a long-term threshold. An event must be at least six continuous hours. Microseism index hours are the total number of hours of triggering, typically evaluated here over the course of a winter storm season. Because the microseism index is measured relative to a long-term threshold microseism level, and corrected to a common datum, at each site, it has the advantage of being relatively insensitive to instrument response variations or uncertainties, such as the transfer function gain constant, within the SFM and DFM bands considered in our analysis. Details of the method are discussed in the following sections.

1.6. Correction of PSD Power to a Common Datum

To compare power levels across the entire span of the dataset, it is necessary to correct older data for amplitude biases due to instrumentation and recording system limitations. First, using only the smoothed and culled 3-hour PSDs, we compute PSD average power levels for the DFM and SFM bands, for each instrument epoch. The DFM band (5 - 9 s) power level is the average of 9 spectral estimates and the longer-period SFM band (14 - 20 s) power level is the average of 7 spectral estimates. An instrument epoch is defined as a time-period in which the station's instrument response did not change and is demonstrated by a single instrument transfer function. Then individual 3-hour PSD SFM and DFM band averages are corrected to a common datum by subtracting the difference between the instrument epoch median and the current instrument epoch median. All 3-hour PSD medians are corrected to a common datum, which is assumed to be the most recent instrument epoch. In general, the correction has the effect of increasing power levels for epochs with clear instrumentation issues and decreasing power levels for time periods with particularly high microseism power levels. In effect, the datum correction produces a more conservative estimate of absolute microseism power levels resulting in flatter trends as a function of time.

1.7. Microseism Event Detection: Monthly Index

Once all culled 3-hour PSD SFM and DFM band averages are corrected to a common datum the microseism index is taken as the number of events exceeding the station's long-term threshold power level. So that every year has a significant number of events we selected the 95th percentile of the corrected dataset as the long-term threshold. A microseism event is declared when 6 continuous hours (i.e. 3 overlapping 3-hour PSDs) exceed the 95th percentile of the corrected dataset. A 6-hour time span is short enough to be impacted by large oceanic microseism storms. A minimum of two 3-hour PSDs must be available to declare a microseism event. Monthly event detections are then used to compute winter storm season total annual microseism event duration for the combined and corrected datasets.

1.8. Data Deposition

All data are freely available through the IRIS Data Management System (www.iris.edu/data). Detailed information on instrumentation at each station and transfer functions, used for the instrument correction procedure for each station instrument epoch, was retrieved from the IRIS DMC meta-data aggregator database at http://www.iris.edu/mda.

References

Aster, R., D.E. McNamara, P. Bromirski (2008), Multi-decadal climate-induced variability in microseisms, *Seism. Res. Lett.*, 79, doi: 10.1785/gssrl.79.2.194.

McNamara, D., and R. Buland (2004), Ambient noise levels in the continental United States, *Bull. Seism. Soc. Am.*, 94, 1517-1527.

McNamara, D.E., C. R. Hutt, L. S. Gee, H. M. Benz, and R. P. Buland (2009), A Method to Establish Seismic Noise Baselines for Automated Station Assessment, *Seism. Res. Lett.*, 80, 628-637.

Peterson, J. (1993). Observation and Modeling of Seismic Background Noise, *USGS Technical Report* 93-322.

Ringler, A.T., L.S. Gee, C.R. Hutt, and D.E. McNamara (2010), Temporal variations in global seismic station ambient noise power levels, *Seism. Res Lett.*, in review.

2. Figure and Table Captions

Figure S1. (a) Vertical acceleration power spectral density (PSD) probability density function (PDF) [*McNamara and Buland*, 2004] constructed from continuous 1 sample/s data from station ANMO near Albuquerque, New Mexico (Figure 1) for a single instrument epoch between 1 November 2000 and 19 November 2002, displaying 22,989 3-hour 50% overlapping power spectral density estimates. NLNM and NHNM indicate low and high-noise PSD bounds, respectively, for GSN stations [*Peterson*, 1993]. Earthquake signals, such as the representative Peru (M 8.4; yellow curve) and mid-Atlantic (M 5.7; red curve) teleseisms, and two regional earthquakes (M 2.5 at 94 km; black line, and M 4.0 at 293 km; blue line, from ANMO), generate strong long-period (> 30 s) surface wave energy relative to Earth's background spectrum. PSDs dominated by earthquake signals are thus readily recognized by their exceedance, in 95% or more of the PSD period bins, above the PDF 80th percentile between 30 and 200 s, and are automatically excised prior to further analysis. 3-hour PSDs affected by station downtime and instrumentation problems are similarly excised by detecting and removing those that fall below the 1st percentile in 95% or more period bins between 30 and 200 s. (b) Transient removal: PSD PDF for the same time interval as (a) with transients culled, showing the DFM and SFM power peaks characteristic of the global microseism signal. Two prominent Pacific storm-associated 3-hour PSDs during this time demonstrate the spectral characteristics of individual extreme storm events. (c) HG.ALQ.–.LHZ PSD PDF computed with 10,226 3-hour PSDs for an instrument epoch from 1976, day 201 through 1978, day 199. Instrument epoch baseline PDF percentiles (5th - 80th) are shown as dashed black lines. (d) After culling, 4133 PSDs remain, and restricting to winter months only (November - March), 2079 PSDs remain for HG.ALQ.–.LHZ. 3-hour PSDs, computed during a time when the seismometer was malfunctioning, fall below the lower bound (5th percentile; black line). Such transients are easily detected and rejected.

Figure S2. Departures from station mean index hours for the SFM plotted at station locations (Figures 1, 2). Red and blue indicate positive and deviations, respectively (scale at bottom).

Figure S3. Departures from station mean index hours for the DFM plotted at station locations (Figures 1, 2). Red and blue indicate positive and deviations, respectively (scale at bottom).

Figure S4. Microseism index slopes and 1σ uncertainties (Figure 2; Table S3) plotted as a function of latitude.

Figure S5. Individual year regressions of DFM and SFM index hour results by station (Figure 1; Table S1), where DFM and SFM values have been normalized by respective station means in each case. Number of years of data with both DFM and SFM determinations, and correlation coefficients c are indicated in each panel, as is the approximate distance from the nearest ocean coast. Stations showing anomalously high DFM years include CHTO, ESK, GUMO, KIP, MAJO. Years during which the Multivariate El Niño Index (MEI; Figure 3) show El Niño (high MEI) or La Niña (low MEI) during the appropriate hemisphere wave year (November - March for the Northern Hemisphere and May through September for the Southern Hemisphere) are indicated in red and green, respectively. Note that many, although not all, outlier years are associated with ENSO excursions, particularly for Pacific affinity stations [*Aster et al.*, 2008] (e.g., GUMO, KIP, MAJO, SNZO) but also for some Atlantic (e.g., ESK, KONO) stations. Correlations between the normalized index hours range from ≈0.9 (KONO) to nearly zero for some Pacific island stations (KIP, GUMO). Lower right panel shows DFM versus SFM trends for all years (annual index hours) by station. A generally greater number of DFM outliers globally produces a general skew of the points above the equity line in the lower right panel.

Table S1. Station locations and histories (Figure 1).

Table S2. Annual Microseism Index Determinations, 95% threshold.

Table S3 (a-d). Parameters from L_2 regression (Figure 2) of winter season microseism index hours. R_2 refers to squared fit of regression line to data. Slope, norm in a and c indicate regression slope normalized by mean index hours. M-K confidence is calculated from the Mann-Kendall statistic for trend. Slope significance is slope divided by the standard deviation. Station averages are calculated from station data normalized by data mean.