

The persistent signature of tropical cyclones in ambient seismic noise.

Gualtieri Lucia^{a,e,*}, Suzana J. Camargo^a, Salvatore Pascale^{b,c}, Flavio M.E. Pons^d, Göran Ekström^a

^a*Lamont-Doherty Earth Observatory of Columbia University, 61 Route 9W, Palisades, NY 10964, USA.*

^b*Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey 08540, USA.*

^c*Geophysical Fluid Dynamics Laboratory/NOAA, Princeton, New Jersey 08540, USA.*

^d*Department of Statistics, University of Bologna, Via delle Belle Arti 41, 40126 Bologna, Italy.*

^e*Present Address: Princeton University, Department of Geosciences, Guyot Hall, Princeton NJ 08544, USA.*

Abstract

The spectrum of ambient seismic noise shows strong signals associated with tropical cyclones, yet a detailed understanding of these signals and the relationship between them and the storms is currently lacking. Through the analysis of more than a decade of seismic data recorded at several stations located in and adjacent to the northwest Pacific Ocean, here we show that there is a persistent and frequency-dependent signature of tropical cyclones in ambient seismic noise that depends on characteristics of the storm and on the detailed location of the station relative to the storm. An adaptive statistical model shows that the spectral amplitude of ambient seismic noise, and notably of the short-period secondary microseisms, has a strong relationship with tropical cyclone intensity and can be employed to extract information on the tropical cyclones.

Keywords: Ambient seismic noise; Tropical cyclones; Statistical modeling.

*Corresponding author

Email address: luciag@princeton.edu (Gualtieri Lucia)

1. Introduction

Ambient seismic noise is the ubiquitous background vibration of the solid Earth recorded worldwide by seismic stations and mainly due to ocean waves driven by winds in intense storms, such as extra-tropical storms and tropical cyclones (TCs) (Gutenberg, 1936; Bromirski, 2009). Two mechanisms are responsible for ambient seismic noise generation: (A) the primary mechanism, which is the direct coupling between ocean waves and the solid Earth in shallow water, responsible for primary microseisms (Hasselmann, 1963; Ardhuin et al., 2015, period T in the range of 10 to 20 s) and the seismic “hum” (Nishida, 2013; Rhie and Romanowicz, 2004; Ardhuin et al., 2015, $T > 50$ s), and (B) the secondary mechanism, which is the interaction amongst ocean waves, responsible for secondary microseisms (Longuet-Higgins, 1950; Hasselmann, 1963, $T < 10$ s).

Much has been done towards understanding the oceanic mechanisms that control the generation of ambient seismic noise (e.g. Longuet-Higgins, 1950; Hasselmann, 1963; Kedar et al., 2008; Ardhuin et al., 2011; Gualtieri et al., 2013; Ardhuin et al., 2015; Nishida and Takagi, 2016), allowing it to be used to infer characteristics of the sea state (e.g. Ardhuin et al., 2012; Neale et al., 2017). Recent studies have shown that ambient seismic noise sources associated with isolated TCs moving across the ocean can be located using seismic methods in the vicinity of the TCs (e.g. Gerstoft et al., 2006; Zhang et al., 2010; Gualtieri et al., 2014; Farra et al., 2016). Their signature are clearly visible on land (e.g. Ebeling and Stein, 2011; Sufri et al., 2014). Other studies have focused on TCs moving over land and on the link between seismic signals and TC energy decay at landfall (e.g. Tanimoto and Lamontagne, 2014; Tanimoto and Valovcin, 2015). Still, the relationship between seismic signals and characteristics of TCs is not yet well understood (e.g. Ebeling and Stein, 2011) due to the complexity of the non-linear and frequency-dependent energy transfer between the atmosphere and the ocean (e.g. Janssen, 2004; Ochi, 2003), as well as between the ocean and the solid Earth (e.g. Hasselmann, 1963; Ardhuin et al., 2010).

Seismic ground motion is related indirectly to the intensity of TCs through ocean gravity waves (microseisms) and infragravity waves (seismic hum) excited in turn by strong winds. Therefore, ocean wave models could be employed to study the relationship between ambient seismic noise and TCs. However, the use
35 of ocean wave models for studying ambient seismic noise generated by decades of TCs is difficult due to limitations of the wave-model data. In particular, ocean wave models, such as WAVEWATCH III (Tolman et al., 2009), use fixed grids with a resolution (0.5×0.5 degrees for WAVEWATCH III) that is too coarse for TCs, generating spatial aliasing and underestimation of the maximum wind
40 and ocean wave height (e.g. Tolman and Alves, 2005, their Figure 8). Moreover, these models use wind reanalyses as an input, which do not represent well the observed TC intensity and location (Schenkel and Hart, 2012; Murakami, 2014). In Fig. S1 in the supplementary material, we show the comparison between the TC wind speed dataset used in this study and the TC wind speed from the
45 ERA-Interim reanalysis database from the European Centre for Medium-Range Weather Forecasts, commonly used as an input for ocean wave models like WAVEWATCH III. The wind speed in the reanalysis is underestimated with respect to observations by about a factor of two (in line with the results of Murakami (2014)). We also observe that the cycle of intensification and decay
50 of TCs differs between observations and reanalysis. The wind speed is related to the spectrum of the ocean wave height (Hasselmann et al., 1973), which in turn is related to the spectral amplitude of noise sources (Hasselmann, 1963). Farra et al. (2016) modeled P-wave sources associated with typhoon Ioke and showed an error on the modeled amplitude comparable with the underestimation
55 given by the reanalysis dataset (their Fig. 6). For these reasons, we decided to rely on TC best track datasets without using information from ocean wave models.

Understanding how processes in the atmosphere and in the ocean couple into seismic waves in the solid Earth and how these can be used to monitor
60 the global environment has been listed as one of the high-priority Seismological Grand Challenges (Lay et al., 2009). Studying this coupling is becoming more

important as a new and valuable source of information on the geophysical effects of climate change at time scales not otherwise accessible and for the pre-satellite era.

65 **2. Materials and Methods**

2.1. Atmospheric and seismic datasets.

We analyze 13 years of atmospheric and seismic data recorded in and adjacent to the northwest Pacific to assess the relationship between the occurrence of TCs and the spectral characteristics of ambient seismic noise. TCs in this region having wind speed larger than 33 m/s are called typhoons and can develop throughout the year with a climatological peak between June and November. The northwest Pacific is the most active basin globally, where approximately 30% of the TCs form each year, as well as where the most intense ones tend to occur (Gray, 1968). We focus on TCs occurring in the northwest Pacific Ocean between 2000 and 2012 during the peak season activity June-November (Fig. 1A). Each TC is identified by track location, intensity and size, recorded every 6 hours. TC intensity is defined as the one-minute mean sustained surface wind speed. We use center locations and intensities of TCs in the northwest Pacific from the Joint Typhoon Warning Center best-track dataset (Chu et al., 2002) (http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/). A tropical cyclone dataset (Knaff et al., 2014, 2015), built by using storm-centered infrared imagery, is used to identify their size. The size of a tropical cyclone is defined as the squared radius of 5-kt (1 kt=0.514 m/s) winds (Knaff et al., 2014, 2015), and therefore it incorporates wind speeds larger than this threshold. We select TCs within 40° of each seismic station and, since we are interested in estimating TC intensity before landfall, we retain only that part of the track moving over the ocean. Time series of TC intensity, size, propagation speed and number of simultaneous TCs are shown in Fig. S2 and S3 in the supplementary material. A scatterplot between intensity and size of TCs is shown in Fig. S4. We keep in our dataset those storms that have been identified as typhoons – i.e.

with wind speed larger than 33 m/s – for at least two days. We do not include in our analysis tropical depressions, tropical storms, as well as short-lived (i.e. less than two days) category-1 typhoons. TCs on the Southern Hemisphere have peak season in January-March, and therefore have been excluded from
95 our analysis. Including TCs on the Southern Hemisphere did not influence our results.

We also analyze continuous broadband vertical-component seismograms recorded during the same time period (2000-2012) at seven seismic stations of the Global Seismic Network (GSN) located in the same region (Fig. 1A). We use the ver-
100 tical component long-period (LHZ) seismograms, with a sampling rate of 1 Hz. In case of stations with multiple seismometers, the primary sensor is used. The instrumental response is deconvolved from the original seismogram in order to get ground acceleration and the power spectral density (PSD, with respect to 1 $(\text{m/s}^2)^2/\text{Hz}$) is computed each 15 minutes and in 30 frequency bands, consider-
105 ing overlapping windows both in time and frequency (Berger et al., 2004). Data have been cleaned from earthquakes, glitches and spurious signals by visual inspection. A time-moving median each 6 hours is performed to obtain the same time step of the TC best-track dataset. Furthermore, to remove seasonality effects due to winter storms on the noise records and better isolate the effect
110 of TCs, long-period trends (i.e. 30 days) have been removed from the seismic data.

In Fig. 1B-C, we show spectrograms of ambient seismic noise ($T = 4 - 12$ s) recorded in 2012 at stations (B) TATO (Taipei, Taiwan) and (C) GUMO (Guam, Mariana Islands). Black lines denote the intensity of TCs – defined as
115 the one-minute mean sustained surface wind speed – moving above the ocean within 40° of each station. Long-lasting signals characterized by high power spectral density (PSD) at short period occur simultaneously with TCs. Figure S5 in the supplementary material shows spectrograms of ambient seismic noise in the microseism frequency band ($T = 4 - 20$ s) at station (A) TATO and (B)
120 GUMO between 2008 (bottom) and 2012 (top). Superposed is the TC intensity. In all cases, we observe a good agreement between the occurrence of TCs and

large-amplitude PSDs at short periods.

2.2. Statistical data processing and estimation of TC intensity.

125 We use a generalized linear model (GLM) with seismic and atmospheric data
between 2000 and 2010 to estimate TC intensity during the TC peak season
2011 and 2012. Ordinary linear regression implies a linear relationship between
a dependent variable \mathbf{Y} and a set of independent variables, or covariates, \mathbf{X} ,
assuming that the dependent variable \mathbf{Y} , conditional to the observed \mathbf{X} , is nor-
130 mally distributed. However, TC intensity is a non-negative variable, displaying
a strongly skewed marginal probability density function, which can be well ap-
proximated by a Gamma distribution (Fig. S6 in the supplementary material).
The dispersion of the distribution is not small with respect to the mean value,
so that an ordinary linear regression is not a realistic assumption, while a GLM
135 is a more appropriate choice (Agresti, 2015).

In order to estimate TC intensity from ambient seismic noise, we proceed as
follows. First, we specify a GLM of TC intensity given the ambient seismic noise
PSD using data between 2000 and 2010. Second, we use the estimated GLM
parameters to predict the TC intensity during 2011 and 2012. A limitation
140 of this method is that, in case of simultaneous TCs, we cannot estimate their
TC intensities separately. In such a case, we still estimate an equivalent TC
intensity which accounts for their cumulative effect.

Our GLM has four components: 1) a dependent variable, that is the intensity
of TCs \mathbf{v}^{TC} , 2) a matrix containing the set of independent variables \mathbf{X} , 3) a
parameter vector $\boldsymbol{\beta}$ and 4) a link function g , such that

$$g(\mu_i) = \mathbf{X}\boldsymbol{\beta} = \beta_0 + X_1\beta_1 + X_2\beta_2 + \dots \quad (1)$$

where μ_i is the expected value of the distribution of the TC intensity given the
observed values of \mathbf{X} (red in Fig. S6) and β_0 is the intercept, which accounts
for effects that are not explained by the considered covariates (Agresti, 2015).

The statistical prediction of TC intensity $\hat{\mathbf{v}}^{\text{TC}}$ can then be obtained as:

$$\hat{\mathbf{v}}^{\text{TC}} = g^{-1}(\mathbf{X}\hat{\boldsymbol{\beta}}) \quad (2)$$

where $\hat{\boldsymbol{\beta}}$ denotes the estimated parameters. Based on the observed distribution of TC intensity (Figure S6 in the supplementary materials), we assume a Gamma
 145 distribution for the dependent variable \mathbf{X} . There are then three possible choices for the function linking the variables: identity, inverse and logarithmic (Agresti, 2015). After comparing different GLMs in the context of the Akaike Information Criterion (Akaike, 1974), we choose to model the data using the identity link function.

150 3. Results and Discussion

3.1. Seismic signals in the presence and in the absence of TCs

For each station shown in Fig. 1A, we compute the median of the PSD over 13 years (2000-2012) in the presence and in the absence of TCs. Fig. 2A shows the spectra as a function of period at station TATO in the presence (blue) and
 155 the in the absence (black) of TCs as a difference with respect to the Low-Noise Model (Peterson et al., 1993), while Fig. 2B shows the difference of spectra in the presence and in the absence of TCs at all stations (dashed lines for stations on islands and solid lines for stations on the continent).

In order to assess if these differences are statistically significant at the 5%
 160 level, we perform a Wilcoxon-Mann-Whitney test (Gibbons and Chakraborti, 2011). This is a non-parametric method to test the null hypothesis that two medians are equal – or, equivalently, that their difference is zero – against the bilateral alternative hypothesis that the two medians are not equal. Fig. 2C shows the p-value as a function of period and station associated with the
 165 spectral difference in Fig. 2B. The spectral difference, statistically significant at the 5% level, is shown in light blue. The spectral difference is always statistically significant at station TATO, whereas, at the other stations, the test does not reject the null hypothesis at a few periods (gray scale).

Considering all stations together, the percentage of spectral difference that
170 is statistically significant decreases with increasing period: 93.8% at $T = 4 - 7$
s (short period secondary microseisms), 87.5% at $T = 7 - 10$ s (long period
secondary microseisms), 71.9% between $T = 10 - 20$ s (primary microseisms)
and 75.0% at $T > 20$ s (seismic hum).

The presence of TCs on the ocean results in an increase of the ambient
175 seismic noise PSD especially at periods $T < 20$ s (secondary and primary mi-
croseisms), where the difference of spectra is mostly statistically significant,
and at stations located on islands (dashed lines in Fig. 2B) – that is at TATO,
GUMO, DAV (Davao, Philippines) and QIZ (Qiongzong, China).

Notably, ambient seismic noise at stations TATO and QIZ, located on islands
180 close to the continent surrounded by relatively shallow water, is affected by
TCs at $T \leq 20$ s (secondary and primary microseisms), while ambient seismic
noise at stations GUMO and DAV, located on islands far away from the coast
surrounded by a deep-water environment, are especially affected by TCs at
 $T \leq 7$ s (short-period secondary microseism). These observations suggest that
185 the propagation of seismic waves from the sources – which, over 13 years, are
likely located in oceanic environments characterized by different depths – to the
receivers modulates the frequency content of the ambient noise records.

The ocean-continent boundary also modulates the seismic wavefield recorded
on the continent (Gualtieri et al., 2015). Seismic records at stations located on
190 the continent – INCN (Inchon, Republic of Korea), KMI (Kunming, Yunnan
Province, China), BJT (Baijiatuan, Beijing, China) – are in general weakly in-
fluenced by TCs (difference of spectra smaller than 1 dB at all periods). Notably,
the presence of the ocean-continent boundary contributes to weaken primarily
the short-period ambient seismic noise (Gualtieri et al., 2015). This is partic-
195 ularly evident comparing the spectral difference at KMI and QIZ (Fig. 2B,
red solid line and purple dashed lines, respectively), located onshore and on an
island close to the coast, respectively.

We also note that a persistent feature due to TCs on primary microseisms
($T=10 - 20$ s) is observed at stations located on islands close to the coasts

200 (TATO and QIZ), and at stations located on the continent (INCN, KMI and BJT). Stations located on islands in a deep-water environment, far away from the coasts (GUMO, DAV), show a weak signal between 10 and 20 s (Fig. 2B), which is not statistically significant (Fig 2C). This could be due to propagation effects or it could be the result of a persistent location of seismic sources of
 205 primary microseisms on the continental shelf (as theorized by Hasselmann, 1963; Ardhuin et al., 2015), far away from stations in a deep-water environment.

Since the effect of TCs on ambient seismic noise recorded at station TATO is the strongest (Fig. 2B) and statistically significant at all periods (Fig. 2), we select this station for in-depth study.

210 3.2. Ambient seismic noise as a function of TC intensity

In Fig. 3, we show the PSD of ambient seismic noise recorded at TATO as a function of TC intensity, colored by density. We only show periods up to 20 s (secondary and primary microseisms), where TCs strongly affect the ambient seismic noise PSD (Fig. 2B). We take the median of ambient seismic noise in
 215 three period bands: (A) $3.8 \leq T \leq 7$ s, short-period secondary microseisms, (B) $7 \leq T \leq 10$ s, long-period secondary microseisms and (C) $10 \leq T \leq 20$ s, primary microseisms. Despite the complexity of the generation mechanism and the energy transfer that involves the atmosphere, ocean and solid Earth, we observe that the PSD of ambient seismic noise is correlated with TC intensity.
 220 Notably, the PSD of ambient seismic noise increases for increasing TC intensity at all periods. To the first order, this relationship can be assumed as linear (black lines in Fig. 3). The slope of the linear regression and the Pearson correlation coefficient ρ decrease with increasing seismic period. The maximum correlation ($\rho = 0.42$) is found for the short-period secondary microseisms (Fig. 3A). In this
 225 period band, a similar linear trend can also be identified at other stations. Fig. 4 shows the relationship between the PSD of short-period secondary microseisms and TC intensity at the other six stations in Fig. 1, colored by density. We note that the slope of the linear fit decreases moving from islands to the continent.

Previous studies detected seismic signals associated with isolated TCs mov-

230 ing on the ocean (e.g. Gerstoft et al., 2006; Zhang et al., 2010; Ebeling and Stein,
 2011; Gualtieri et al., 2014; Farra et al., 2016). They found the strongest sig-
 nals in the same period band that shows the maximum correlation with TC
 intensity ($T = 4 - 7$, Fig. 2). At these periods, ambient seismic noise is gen-
 erated by the interaction of ocean gravity waves coming from nearly opposite
 235 directions (Longuet-Higgins, 1950; Hasselmann, 1963). In the specific case of a
 TC, winds on one side of the TC generate ocean waves travelling in the direc-
 tion of motion of the TC itself, while winds on the opposite side generate ocean
 waves travelling in the opposite direction. Thus, the tail of a TC is expected
 to be a considerable source of wave-wave interaction (Longuet-Higgins, 1952).
 240 While moving over the ocean, TCs thus carry their own sources of secondary
 microseisms, which follow the storm (e.g. Gerstoft et al., 2006; Gualtieri et al.,
 2014).

3.3. Results from an adaptive statistical model

The strong correlation between the PSD of short-period ambient seismic
 245 noise and TC intensity suggests that the PSD of ambient seismic noise can be
 employed as additional source of information about TCs. We use the GLM de-
 scribed in section 2.2 with the seismic and atmospheric data described in section
 2.1 between 2000 and 2010 to estimate TC intensity during the TC peak sea-
 sons 2011 and 2012. In our GLM, we consider two different sets of independent
 250 variables: $\mathbf{X} = \mathbf{X}^{[1]}$, which is composed only by the PSD of ambient seismic
 noise, and $\mathbf{X} = \mathbf{X}^{[2]}$ in which we also account for other TC characteristics.

In our first test, as independent variables $\mathbf{X} = \mathbf{X}^{[1]}$, we use the median of
 ambient seismic noise PSD in three period bands: $4 \leq T \leq 7$ s short-period
 secondary microseisms, $7 \leq T \leq 10$ s long-period secondary microseisms, and
 255 $10 \leq T \leq 20$ s primary microseisms. This choice implicitly allows us to account
 for the dominant period of the PSD of ambient seismic noise. We recall that
 the PSD is expressed in dB with respect to $1 \text{ (m/s}^2\text{)}^2/\text{Hz}$, which is a classic way
 to represent the spectrum of ambient seismic noise.

In order to improve our predictions, in a second test, we add additional

260 variables $\mathbf{X} = \mathbf{X}^{[2]}$: the size of TCs, defined as the squared radius of 5kt winds
(Knaff et al., 2014, 2015), the distance between TCs and the seismic station
TATO, and a variable that accounts for the number of simultaneous TCs. Only
the TC size is statistically significant at the 5% level (Table 1 , second column).
We also checked that other variables, such as the TC propagation speed, the
265 tropical/extra-tropical transition, and the parameters of TCs on the Southern
Hemisphere, are not statistically significant.

The estimated TC intensities for 2011 and 2012 peak seasons are shown
in Fig. 5. Observed data are in gray, estimated TC intensities considering
only short-period secondary-microseism PSDs are in blue, and considering both
270 short-period secondary-microseism PSDs and TC size are in red. Gray shad-
ows indicate simultaneous TCs, for which we can asses only an equivalent TC
intensity accounting for their cumulative effect on ambient seismic noise. The
Pearson correlation coefficient between observed and predicted values is $\rho = 0.60$
in 2011 and $\rho = 0.56$ in 2012 considering only ambient seismic noise (blue in Fig.
275 5). All the three seismic-noise period bands are statistically significant at the
5% level (Table 1 , first column). The estimated TC intensity improves adding
the information about the TC size (red in Fig. 5), with a Pearson correlation
coefficient of $\rho = 0.84$ in 2011 and $\rho = 0.80$ in 2012.

Including the TC size as a predictor in our GLM, the p-values associated
280 with long-period secondary microseisms ($T = 7 - 10$ s) and primary micro-
seisms ($T = 10 - 20$ s) exceed the threshold of 5%, meaning that these variables
lose significance, while the short-period secondary microseisms ($T = 4 - 7$ s)
and TC size are statistically significant (see Table 1). This effect is known
as omitted-variable bias (Agresti, 2015) and it is due to the potential correla-
285 tion amongst the independent variables $\mathbf{X} = \mathbf{X}^{[2]}$. This indicates that, while
TC size is a significant variable, the coefficients estimated in the first test for
long-period secondary microseisms and primary microseisms are spurious. Not
including TC size, the model attributes spurious significance to the variables
correlated to the TC size or to the short-period secondary microseisms. The
290 Pearson correlation coefficient between short-period secondary microseisms and

long-period secondary microseisms is $\rho \simeq 0.75$. Being generated by the same physical mechanism (Longuet-Higgins, 1950), the short-period secondary microseisms and long-period secondary microseisms bands are correlated. The Pearson correlation coefficient between short-period secondary microseisms and primary microseisms is $\rho \simeq 0.43$. One possible explanation for the positive correlation between short-period secondary microseisms and primary microseisms is that the transition between secondary and primary mechanisms – ocean wave-wave interaction and direct coupling between ocean waves and the seafloor (Ardhuin et al., 2015) – does not occur at a fixed period and, especially for very strong storms, secondary microseisms extend to periods greater than 10 s. For example, a category 1 TC having wind speed exceeding 40 m/s for at least two days is expected to generate ocean waves having a peak period of about 23 s (e.g. Hanafin et al., 2012), which corresponds to about 11.5 s secondary microseisms. Stronger TCs, characterized by stronger winds, can therefore generate longer period secondary microseisms also if these winds blow for a shorter time. Therefore, the long-period secondary microseisms and primary microseisms are redundant variables in our GLM. Finally, we also observe that the intercept is statistically significant in both models, but accounting for TC size, the estimated value of the intercept β_0 decreases of about 50%. Therefore, our final GLM relies on two independent variables: the median of the short-period secondary-microseism PSD, and the TC size.

In Fig. 6, we focus on three TCs occurring in the northwest Pacific Ocean during September–October 2012 within 40° of the seismic station TATO. Typhoons Sanba (September 10-17) and Jelawat (September 20-30) were classified as category-5 super typhoons on the Saffir-Simpson scale, while Prapiroon (October 7-19) was a category 3 typhoon. JTWC best-track data intensity (Chu et al., 2002) are shown in black, while estimates obtained using our GLM with short-period secondary microseisms and TC intensity are in red. Our model allows us to estimate well the overall TC intensification and decay for all the three TCs, meaning that short-period ambient seismic noise recorded at TATO carries information on the intensity of the TCs. The Pearson correlation

coefficients between observations and model are $\rho = 0.83$, $\rho = 0.85$, $\rho = 0.81$ for Sanba, Jelawat and Prapiroon, respectively.

We note that the estimated TC intensity sometimes is delayed by a few hours
325 to a couple of days with respect to the best-track data. As the seismic propa-
gation is nearly instantaneous with respect to the 6-hour time step, we ascribe
this delay to non-linear coupling between atmosphere and ocean, and a poten-
tially slow wind-wave growth which may take from a few hours to a few days
(Hasselmann et al., 1973). Furthermore, the linearity of our model is amongst
330 the possible causes of the small discrepancies between our model and the best-
track data, given the complexity and non-linearity of the generation mechanism
of ambient seismic noise. A limitation of employing a linear model is that the
extreme values of TC intensity are not captured as well. For example, our GLM
overestimates the wind speed of about 10-15 m/s at the beginning and at the
335 end of the event, when the storm is a tropical depression. This limitation could
be overcome by using a GLM with additional dependent variables accounting for
non-linearity in the seismic records (e.g. intermittency, Barndorff-Nielsen et al.,
2014) or by employing more sophisticated machine-learning statistical models,
such as Generalized Additive Models (Hastie and Tibshirani, 1990), which can
340 estimate and account for non-linear effects in the relationship between TC in-
tensity and seismic records in an automatic fashion. In addition to that, the
presence of simultaneous weaker tropical and extra-tropical storms could poten-
tially affect the amplitude of ambient seismic noise and be the cause of occasional
model overestimations.

345 Finally, Fig. S7 in the supplementary material shows the comparison be-
tween observed and estimated TC intensity in the 2011 and 2012 peak seasons
releasing the condition of the minimum wind speed considered in the best-track
data and including tropical storms lasting for more than two days. Notably, the
selection of storms in the best-track dataset is done excluding all tropical storms
350 (wind speed larger than 18 m/s) that last for less than two days and including
all long-lasting ones. Typhoons and tropical storms often occur simultaneously
and therefore only an equivalent TC intensity is estimated in this case (grey

shadow), but the correlation between observed (grey) and estimated intensity (blue and red) is still very good and is better when including the storm size in our statistical model. The Pearson correlation coefficient is $\rho = 0.63$ in 2011 and $\rho = 0.67$ in 2012 considering only ambient seismic noise (blue), and $\rho = 0.75$ in 2011 and $\rho = 0.80$ in 2012 considering also the storm size (red).

4. Conclusions

Despite the complexity of the energy transfer amongst atmosphere, ocean and solid Earth, our analysis of ambient seismic noise recorded in the northwest Pacific over 13 years suggests that the PSD of ambient seismic noise carries a persistent and statistically significant signature of TCs. This signature varies with frequency and station location likely due to the propagation of seismic waves across the heterogenous Earth structure and across the ocean-continent boundary, or to a persistent location of seismic sources associated with specific frequency bands.

As our statistical estimate of TC intensity demonstrates, this signature is correlated with TC intensity. Considering a single TC in the specific investigated area, we found that, compared with observed TC intensities, the error obtained estimating TC intensity from seismic ambient noise is smaller than the one obtained using state-of-the-art reanalyses (compare Figure 5b with Figure S1).

While we have shown strong predictive power of seismic noise for the estimation of TC intensity, we do not suggest that the modern-day seismic observations can augment current satellite capabilities. However, global seismic observations going back several decades can now be exploited to provide new quantitative constraints on TC activity in remote ocean basins for the pre-satellite era.

Acknowledgments

We thank the US Geological Survey (USGS) and the Incorporated Research Institutions for Seismology (IRIS) for providing openly available seismic data. L.G. acknowledges support from a Lamont-Doherty Earth Observa-

tory Postdoctoral Fellowship. S.J.C. acknowledges support from NOAA grants NA15OAR4310095 and NA16OAR4310079. S.P. was supported by the NOAA CICS grant NA14OAR4320106. We would like to thank John Knaff (NOAA) for making the tropical cyclone size data available for this study. We also wish to
385 acknowledge the editor Peter Shearer and the two anonymous reviewers, whose comments helped us to improve the quality of the manuscript.

References

- Agresti, A., 2015. Foundations of linear and generalized linear models. John Wiley & Sons.
- 390 Akaike, H., 1974. A new look at the statistical model identification. *IEEE transactions on automatic control* 19 (6), 716–723.
- Ardhuin, F., Balanche, A., Stutzmann, E., Obrebski, M., 2012. From seismic noise to ocean wave parameters: General methods and validation. *Journal of Geophysical Research: Oceans* (1978–2012) 117 (C5).
- 395 Ardhuin, F., Gualtieri, L., Stutzmann, E., 2015. How ocean waves rock the Earth: Two mechanisms explain microseisms with periods 3 to 300 s. *Geophysical Research Letters* 42 (3), 765–772.
- Ardhuin, F., Rogers, E., Babanin, A. V., Filipot, J.-F., Magne, R., Roland, A., Van Der Westhuysen, A., Queffelec, P., Lefevre, J.-M., Aouf, L., et al., 2010.
400 Semiempirical dissipation source functions for ocean waves. Part I: Definition, calibration, and validation. *Journal of Physical Oceanography* 40 (9), 1917–1941.
- Ardhuin, F., Stutzmann, E., Schimmel, M., Mangeney, A., 2011. Ocean wave sources of seismic noise. *Journal of Geophysical Research: Oceans* 116 (C9).
- 405 Barndorff-Nielsen, O. E., Pakkanen, M. S., Schmiegel, J., et al., 2014. Assessing relative volatility/intermittency/energy dissipation. *Electronic Journal of Statistics* 8 (2), 1996–2021.

- Berger, J., Davis, P., Ekström, G., 2004. Ambient Earth noise: a survey of the global seismographic network. *Journal of Geophysical Research: Solid Earth* 109 (B11).
410
- Bromirski, P. D., 2009. Earth vibrations. *Science* 324 (5930), 1026–1027.
- Chu, J.-H., Sampson, C. R., Levine, A. S., Fukada, E., 2002. The joint typhoon warning center tropical cyclone best-tracks, 1945–2000. Naval Research Laboratory Rep. NRL/MR/7540-02 16, 22.
- 415 Ebeling, C. W., Stein, S., 2011. Seismological identification and characterization of a large hurricane. *Bulletin of the Seismological Society of America* 101 (1), 399–403.
- Farra, V., Stutzmann, E., Gualtieri, L., Schimmel, M., Ardhuin, F., 2016. Ray-theoretical modeling of secondary microseism P waves. *Geophysical Journal International* 206 (3), 1730–1739.
420
- Gerstoft, P., Fehler, M. C., Sabra, K. G., 2006. When Katrina hit California. *Geophysical Research Letters* 33 (17).
- Gibbons, J. D., Chakraborti, S., 2011. *Nonparametric statistical inference*. Springer.
- 425 Gray, W. M., 1968. Global view of the origin of tropical disturbances and storms. *Mon. Wea. Rev.* 96, 669–700.
- Gualtieri, L., Stutzmann, E., Capdeville, Y., Ardhuin, F., Schimmel, M., Mangeney, A., Morelli, A., 2013. Modelling secondary microseismic noise by normal mode summation. *Geophysical Journal International* 193 (3), 1732–1745.
- 430 Gualtieri, L., Stutzmann, E., Capdeville, Y., Farra, V., Mangeney, A., Morelli, A., 2015. On the shaping factors of the secondary microseismic wavefield. *Journal of Geophysical Research: Solid Earth* 120 (9), 6241–6262.

- Gualtieri, L., Stutzmann, E., Farra, V., Capdeville, Y., Schimmel, M., Ardhuin, F., Morelli, A., 2014. Modelling the ocean site effect on seismic noise body waves. *Geophysical Journal International* 197 (2).
- 435
- Gutenberg, B., 1936. On microseisms. *Bulletin of the Seismological Society of America* 26 (2), 111–117.
- Hanafin, J. A., Quilfen, Y., Ardhuin, F., Sienkiewicz, J., Queffelec, P., Obrebski, M., Chapron, B., Reul, N., Collard, F., Corman, D., et al., 2012. Phenomenal sea states and swell from a North Atlantic storm in February 2011: a comprehensive analysis. *Bulletin of the American Meteorological Society* 440 93 (12), 1825–1832.
- Hasselmann, K., 1963. A statistical analysis of the generation of microseisms. *Reviews of Geophysics* 1 (2), 177–210.
- 445
- Hasselmann, K., Barnett, T., Bouws, E., Carlson, H., Cartwright, D., Enke, K., Ewing, J., Gienapp, H., Hasselmann, D., Kruseman, P., et al., 1973. Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). Tech. rep., Deutsches Hydrographisches Institut.
- Hastie, T., Tibshirani, R., 1990. *Generalized additive models*. Wiley Online Library.
- 450
- Janssen, P., 2004. *The interaction of ocean waves and wind*. Cambridge University Press.
- Kedar, S., Longuet-Higgins, M., Webb, F., Graham, N., Clayton, R., Jones, C., 2008. The origin of deep ocean microseisms in the North Atlantic Ocean 455 464 (2091), 777–793.
- Knaff, J. A., Longmore, S. P., Molenaar, D. A., 2014. An objective satellite-based tropical cyclone size climatology. *Journal of Climate* 27 (1), 455–476.
- Knaff, J. A., Longmore, S. P., Molenaar, D. A., 2015. CORRIGENDUM. *Journal of Climate* 28 (21), 8648–8651.

- 460 Lay, T., Aster, R., Forsyth, D., Romanowicz, B., Allen, R., Cormier, V.,
Gomberg, J., Hole, J., Masters, G., Schutt, D., et al., 2009. Seismological
grand challenges in understanding Earth’s dynamic systems. Report to the
National Science Foundation, IRIS Consortium 46, 1–18.
- Longuet-Higgins, M., 1952. Can sea waves cause microseisms? Proc. Sympos.
465 on Microseisms.
- Longuet-Higgins, M. S., 1950. A theory of the origin of microseisms. Philosoph-
ical Transactions of the Royal Society of London. Series A. Mathematical and
Physical Sciences, 1–35.
- Murakami, H., 2014. Tropical cyclones in reanalysis data sets. Geophysical Re-
470 search Letters 41 (6), 2133–2141.
- Neale, J., Harmon, N., Srokosz, M., 2017. Monitoring remote ocean waves using
P-wave microseisms. Journal of Geophysical Research: Oceans.
- Nishida, K., 2013. Earth’s background free oscillations. Annual Review of Earth
and Planetary Sciences 41, 719–740.
- 475 Nishida, K., Takagi, R., 2016. Teleseismic S wave microseisms. Science
353 (6302), 919–921.
- Ochi, M., 2003. Hurricane generated seas. Vol. 8. Elsevier.
- Peterson, J., et al., 1993. Observations and modeling of seismic background
noise. US Geological Survey Open File Report, 93–322.
- 480 Rhie, J., Romanowicz, B., 2004. Excitation of Earth’s continuous free oscilla-
tions by atmosphere–ocean–seafloor coupling. Nature 431 (7008), 552–556.
- Schenkel, B. A., Hart, R. E., 2012. An examination of tropical cyclone posi-
tion, intensity, and intensity life cycle within atmospheric reanalysis datasets.
Journal of Climate 25 (10), 3453–3475.

Independent variables \mathbf{X}	p-values ($\mathbf{X} = \mathbf{X}^1$)	p-values ($\mathbf{X} = \mathbf{X}^2$)
Intercept	$< 2 \times 10^{-16}$	$< 2 \times 10^{-16}$
Short-period secondary microseisms	$< 2 \times 10^{-16}$	$< 2 \times 10^{-16}$
Long-period secondary microseisms	$< 2 \times 10^{-13}$	0.22
Primary microseisms	$< 1.4 \times 10^{-3}$	0.26
TC Size		$< 2 \times 10^{-16}$

Table 1: P-values of the independent variables \mathbf{X} used in the GLM. The first column is referred to the case in which the independent variable $\mathbf{X} = \mathbf{X}^{[1]}$ is only ambient seismic noise (blue in Fig. S5), divided in three period bands, as explained in the text. The second column is referred to the case in which we consider as independent variables $\mathbf{X} = \mathbf{X}^{[2]}$ ambient seismic noise and some additional independent variables related to the characteristics of TCs (red in Fig. S5). Bold face indicates a variable that is statistically significant at the 5% level.

- 485 Sufri, O., Koper, K. D., Burlacu, R., de Foy, B., 2014. Microseisms from super-storm Sandy. *Earth and Planetary Science Letters* 402, 324–336.
- Tanimoto, T., Lamontagne, A., 2014. Temporal and spatial evolution of an on-land hurricane observed by seismic data. *Geophysical Research Letters* 41 (21), 7532–7538.
- 490 Tanimoto, T., Valovcin, A., 2015. Stochastic excitation of seismic waves by a hurricane. *Journal of Geophysical Research: Solid Earth* 120 (11), 7713–7728.
- Tolman, H. L., Alves, J.-H. G., 2005. Numerical modeling of wind waves generated by tropical cyclones using moving grids. *Ocean Modelling* 9 (4), 305–323.
- Tolman, H. L., et al., 2009. User manual and system documentation of WAVE-WATCH III TM version 3.14. Technical note, MMAB Contribution 276, 220.
- 495 Zhang, J., Gerstoft, P., Bromirski, P. D., Aug. 2010. Pelagic and coastal sources of P-wave microseisms: Generation under tropical cyclones. *Geophysical Research Letters* 37 (15).

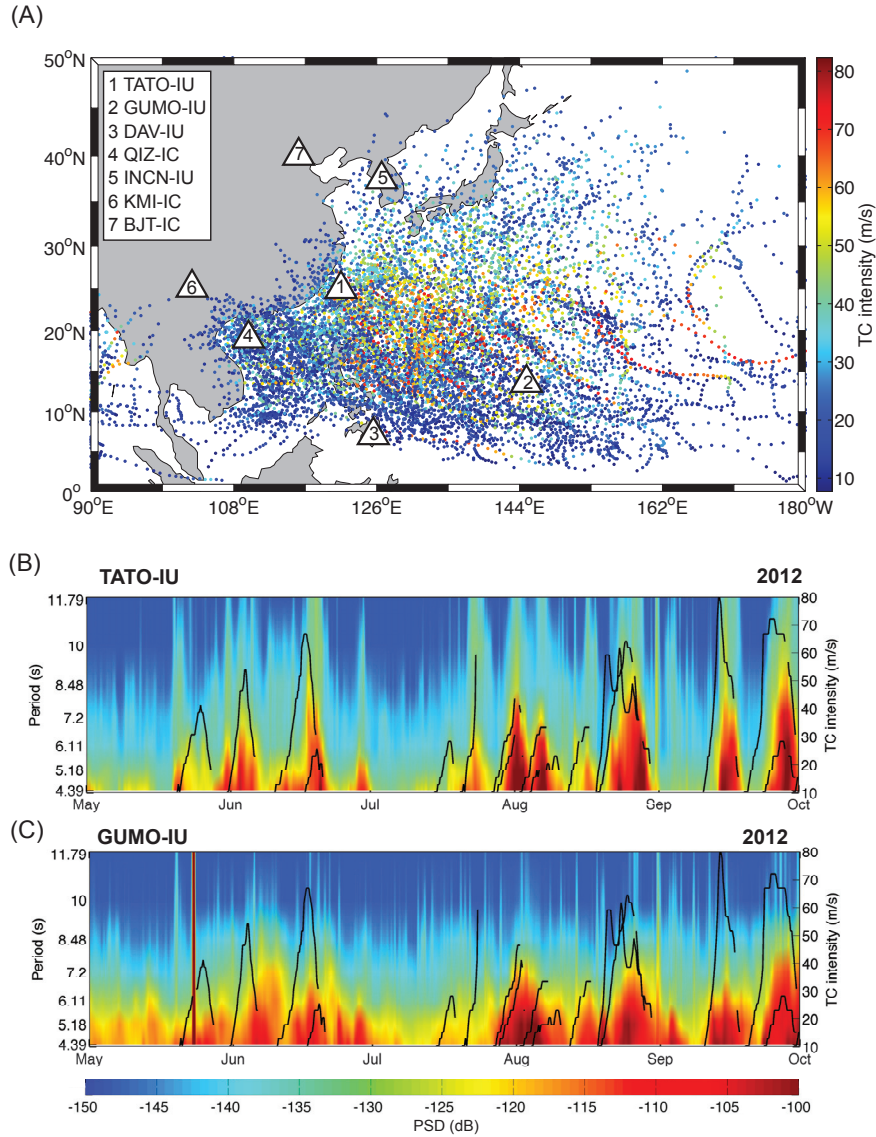


Figure 1: TCs and ambient seismic noise in the northwest Pacific Ocean. (A) Northwest Pacific Ocean typhoon tracks for the period 2000 and 2012, colored by intensity. Black triangles mark the seismic stations whose data has been used in this study. Spectrograms of ambient seismic noise recorded at station (B) TATO and (C) GUMO during 2012. Black lines denote TC intensity (scale on the right). The PSD of ambient seismic noise is defined with respect to $1 \text{ (m/s}^2\text{)}^2/\text{Hz}$. Vertical straight lines indicate times of earthquakes.

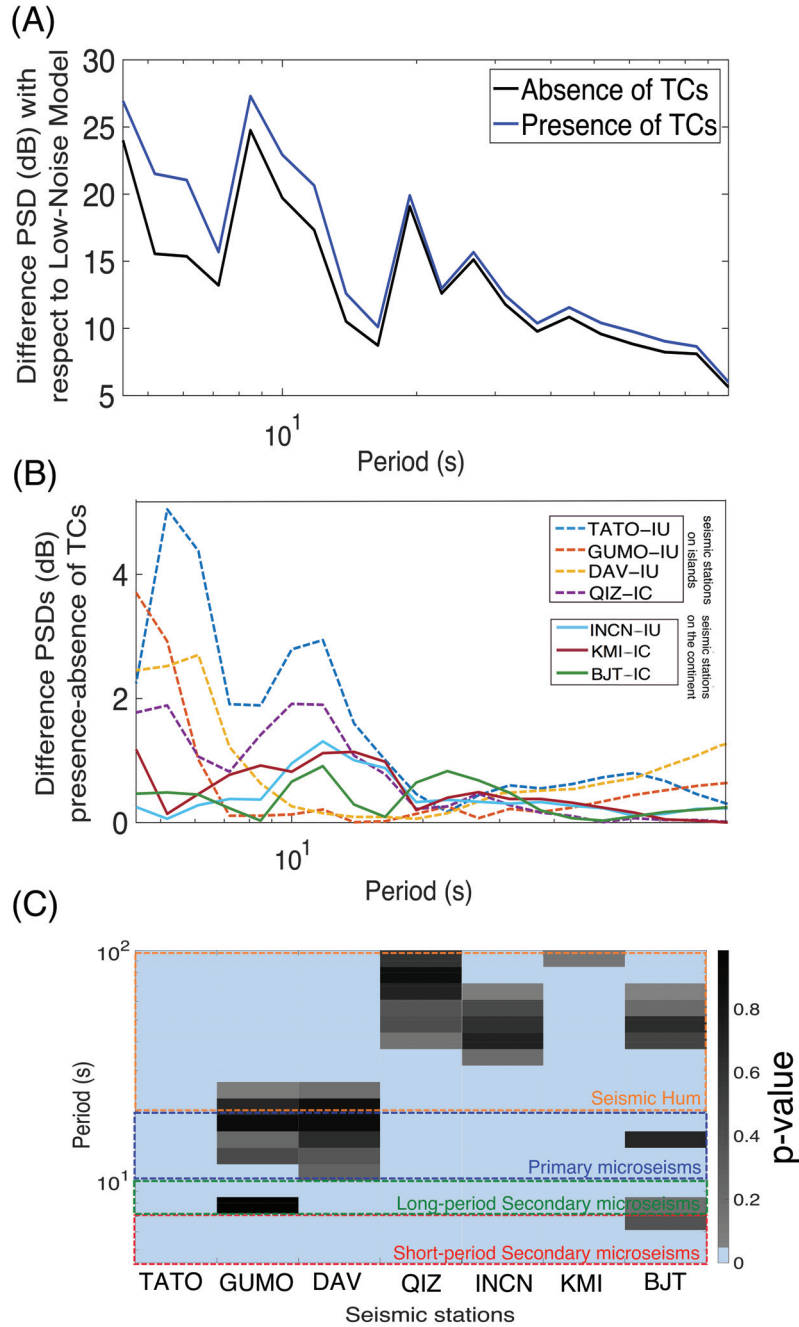


Figure 2: (A) Power spectrum computed in the presence and in the absence of TCs at station TATO during 2000 – 2012 with respect to the Low-Noise Model (Peterson et al., 1993). The blue (black) curve is computed considering only the ambient seismic noise PSD in the presence (absence) of TCs within 40° of each station. (B) Difference between the spectrum computed in the presence and in the absence of TCs at the seismic stations in Fig. 1. (C) P-value associated with the difference between the PSD measured in the presence and in the absence of TCs (panel B in this figure). Light blue indicates seismic periods at which the spectral difference is statistically significant at the level of 5%.

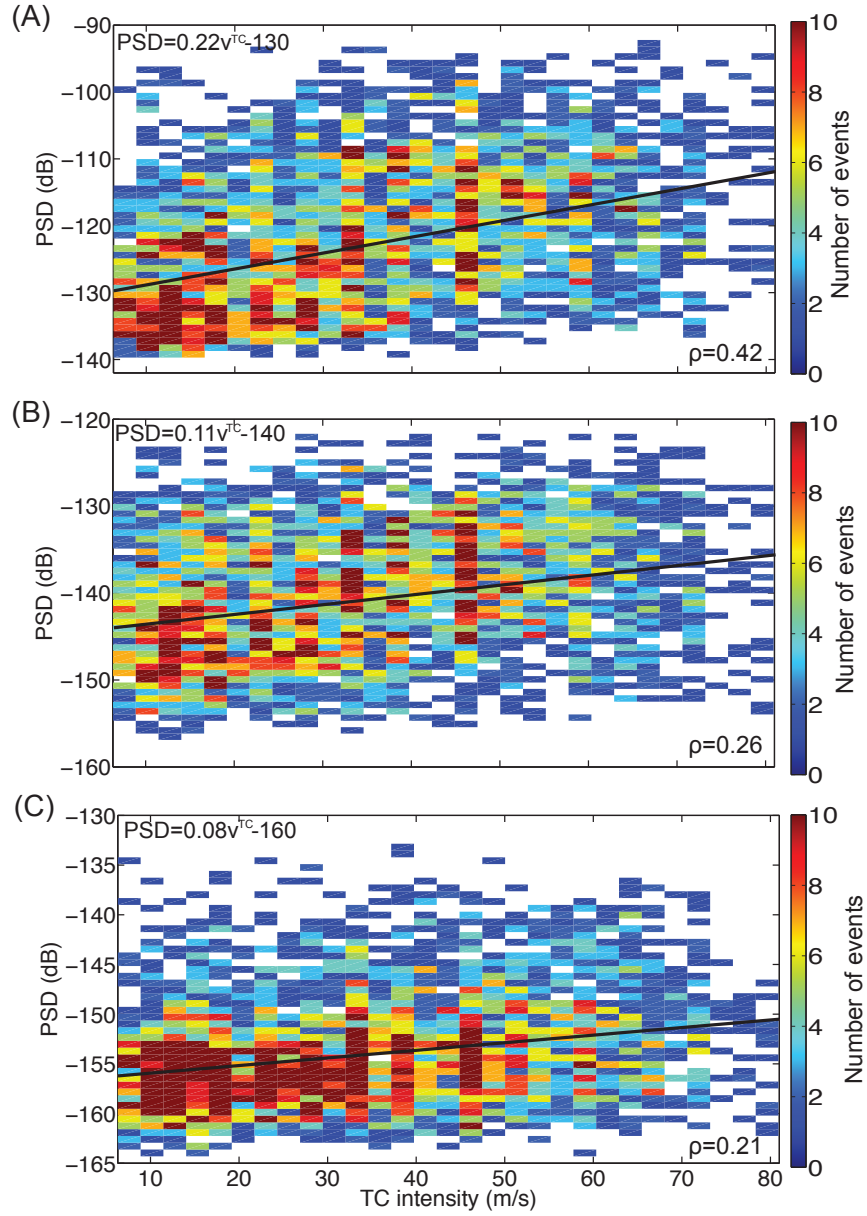


Figure 3: Ambient seismic noise at TATO vs TC intensity coloured by number of events. The PSD of ambient seismic noise has been filtered in three period bands: (A) $T=4-7$ s, short-period secondary microseisms, (B) $T=7-10$ s, long-period secondary microseisms and (C) $T=10-20$ s, primary microseisms. The linear fit between ambient seismic noise PSD and TC intensity (equation as a label) is in black. The Pearson correlation coefficient ρ between the two quantities is printed as a label on the bottom-right corner.

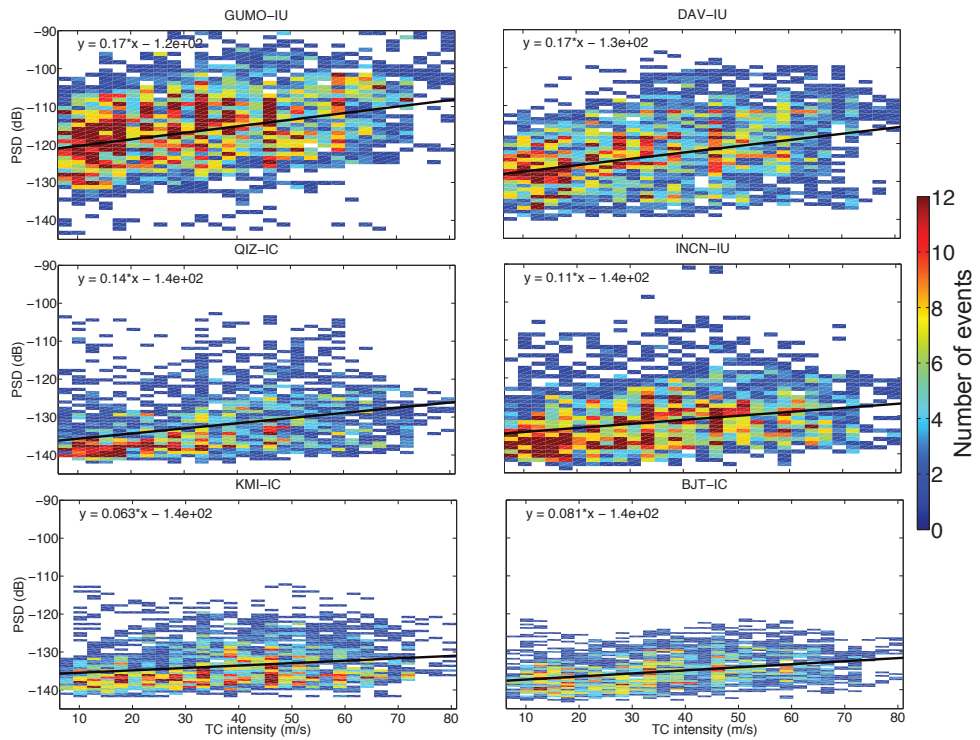


Figure 4: Scatter plots colored by density and linear fits between short-period secondary microseisms ($T=4 - 7$ s) recorded at different stations and TC intensity. The slope of the linear fit decreases moving from islands to the continent. For each station, we consider TCs moving within 40° .

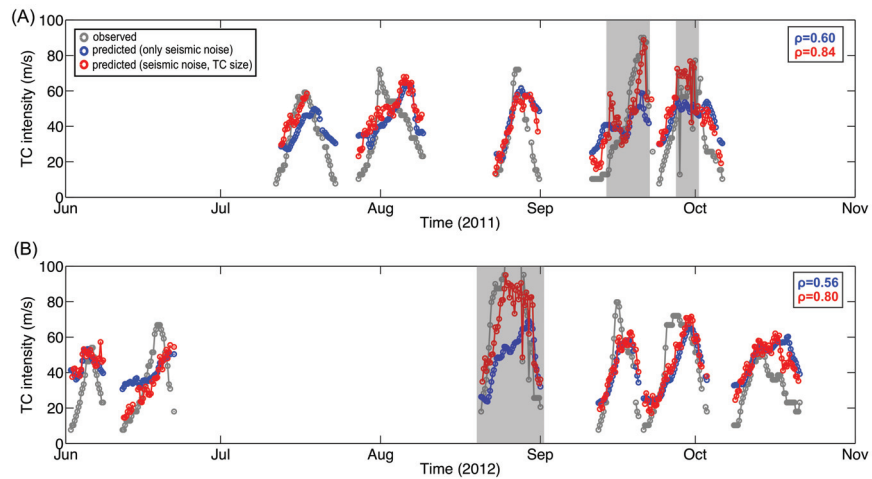


Figure 5: Observed TC intensity (gray) is compared to estimated TC intensity obtained using as two different sets of independent variables: only ambient seismic noise (blue) and ambient seismic noise with further independent characteristics of TCs (red). See Table 1 for more details on these independent variables. Gray shadows indicate the time periods when simultaneous TCs lie within 40° from the station. For each year, the Pearson correlation coefficient is shown as a label.

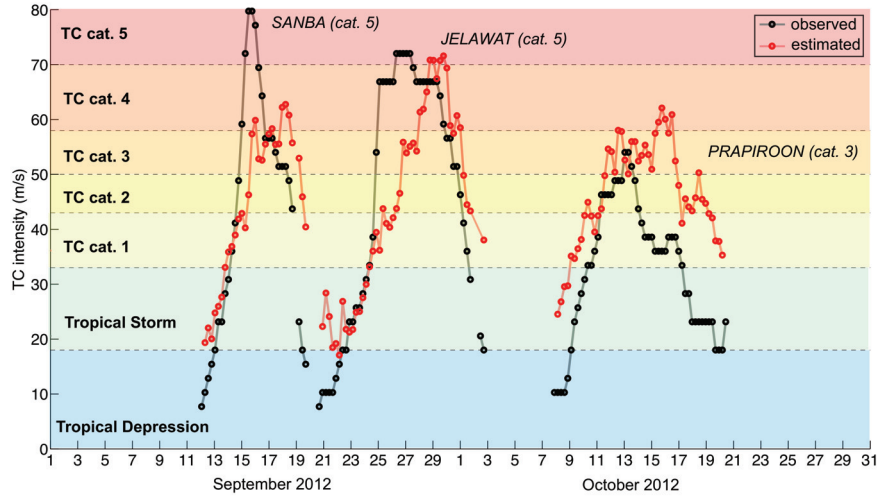


Figure 6: Observed (black) and estimated (red) TC intensity. TC intensity is estimated by a Generalized Linear Model using short-period secondary microseisms and TC size as independent variables. Background colors denote the TC category as given by the Saffir-Simpson scale.