Automatic non-volcanic tremor detection in the Mexican subduction zone

A. Husker[1*], S. Peyrat[2], N. Shapiro[2], V. Kostoglodov[1]

[1]*Instituto de Geofísica, Universidad Nacional Autónoma de México, Mexico City, Mexico*
[2]*Laboratoire de Sismologie, Institut de Physique du Globe de Paris, Paris, France*

Resumen

Desarrollo de un algoritmo de detección de tremores no-volcánicos (TNV). Un TNV ocurre en el rango de frecuencia de 1 – 10 Hz, pero el ruido ambiental alto ocurre por encima de los 2 Hz en muchas de las estaciones del MesoAmerican Seismic Experiment (MASE), entonces los sismogramas se filtran con un paso banda de 1 – 2 Hz. Se determina el efecto de sitio comparando la ventana de las codas de terremotos regionales en todas las estaciones, y luego se remueve. Se aplica un filtro mediano a la amplitud absoluta de los datos filtrados con el filtro de paso banda para remover los picos de los terremotos locales. Se quita la tendencia y el valor mediano de los archivos de datos de un día para eliminar el efecto de las tormentas y el ruido local. Se obtiene el promedio de todos los datos por día para amplificar la energía coherente de los TNV y para disminuir el ruido local. Se determina empíricamente un límite de amplitud de los TNV y se aplica a todos los datos diariamente del periodo de estudio para generar el catálogo de los TNV. Se compara el catálogo determinado automáticamente con el catálogo creado por el análisis visual de los espectrogramas diariamente. Se hace un análisis preciso durante el mes con la mayor diferencia entre los dos catálogos (mayo de 2006). Encontrandose que el algoritmo de detección automática tuvo menos falsos positivos y menos TNV indeterminados que la detección inicial de TNV visualmente, por lo tanto el algoritmo automático desarrollado de detección de TNV tiene menos errores y puede usarse efectivamente para futuros análisis de la actividad de los tremores.

Palabras clave: Tremores no volcánicos, detección de tremor, filtro mediano.

Abstract

An automatic non-volcanic tremor (NVT) detection algorithm is developed. NVT occurs in the 1 – 10 Hz frequency range, but high ambient noise dominates above 2 Hz at many of the MesoAmerican Seismic Experiment (MASE) stations, so the seismograms are bandpass filtered 1 – 2 Hz. The site amplification effect is determined by comparing the coda envelopes of regional earthquakes at all stations and then removed. A median filter is applied to the absolute amplitude of the bandpass filtered data to remove the spikes of local earthquakes. Day-long data files of each station are detrended and the median value is removed to suppress the effect of storms and local noise. An average of all data files per day is then obtained to amplify the coherent energy from NVT and diminish residual noise. An amplitude cutoff for NVT is empirically determined and applied to all daily data within the study period to generate the NVT catalog. The automatically determined catalog is compared with a catalog created by visual analysis of daily spectrograms. A detailed comparison is done of the month with the greatest difference between the two catalogs, May 2006. It is found that the automatic detection algorithm had fewer false picks and undetermined NVT than the initial visual NVT detection. Thus, the automatic NVT detection algorithm developed has fewer errors and can be effectively used for further analysis of the tremor activity.

Key words: Non-volcanic tremor, tremor detection, median filter.

Introduction

Non-volcanic tremor (NVT) has been heavily examined in the subduction zones of Japan (ex. Obara, 2002; Obara *et al*., 2004; Shelly *et al*., 2006; Ide *et al*., 2007) and Cascadia (ex. Rogers and Dragert, 2003; McClausland *et al*., 2005; Kao *et al*., 2007; Wech and Creager, 2008) and is now being studied in Mexico (Payero *et al*., 2008). In subduction zones it appears in bands perpendicular to the direction of subduction at specific distances from the trench lasting for 10's of minutes to hours (ex. Obara, 2002; Wech and Creager, 2008; Payero *et al*., 2008). It has also just been discovered on the San Andreas Fault, although it has not been as thoroughly researched there (Shelly *et al*., 2009). NVT activity increases dramatically during Slow Slip Events (SSE) in subduction zones but also occurs at times outside of the SSE (ex. Obara, 2002: Rogers and Dragert, 2003; Payero *et al*., 2008). Despite

the growing number of studies about NVT, the nature of NVT and its relationship to SSE are unclear and still being debated. Detection of NVT is the most basic step necessary to study the phenomenon. For example, one of the possible methods to analyze a space-time correlation between NVT and SSE is to examine the seismic energy released by the NVT (Kostoglodov *et al.*, 2008). Reliable determination of large quantities of NVT burst onsets and durations is a very important task in particular for this purpose.

The original method to detect tremor came from studies in Japan (Obara, 2002). The first step was to filter the NVT waveform to 1-10 Hz, the range where tremor is observed. The next step was to determine the root-mean-square (rms) amplitude of the filtered seismogram. Finally, the rms window was cross-correlated with the rms windows of seismograms from other stations to determine the NVT location from the relative arrival times of the cross-correlogram. This technique requires human input to determine what part of the seismogram is NVT and what part is a local earthquake or noise.

Since the initial tremor detection technique, two separate studies have been done to automatically detect tremor in Cascadia (Kao *et al.*, 2007 and Wech and Creager, 2008). Kao *et al.* (2007) calculate the moving average and scintillation index (SI) (Yeh and Liu, 1982) of the waveform as well as the mean and standard deviation of the SI. They use the different values in an empirically determined logic chart to classify NVT from background noise. Then the NVT classified events are compared between stations. If 3 stations within a 100 km radius detect the event, it is considered real.

Wech and Creager (2008) use a modified form of the original window and cross-correlation detection method. They cut the enveloped waveforms into 5 minute windows and correlate them. Then a grid-search localization is performed to minimize the difference between the NVT arrival time and the cross correlation maximum. If the determined location from adjacent 5 minute windows are not close enough spatially, the event is thrown out.

Both methods have drawbacks that made us decide to implement a different detection technique. The method implemented by Kao *et al.* (2007) is not intuitively straightforward. That is, it is not immediately evident why very high or low means or SI's constitute NVT. Therefore, the process of tuning the logic chart to a new area and possibly to each new station can be quite time consuming and require a learning curve. The Wech and Creager (2008) method assumes clustering of events, requires a relatively well spaced 2D network, and requires locating

NVT at the same time as detecting it. In Mexico NVT has appeared in two regions at the same time. Such a "double-event" could be thrown out with the implemented detection scheme. But perhaps the biggest problem with both methods is that they assume a point source epicenter. NVT actually may be occurring over an area or volume up to tens of kilometers of characteristic scale.

In Mexico, the National Seismological Service (SSN) seismometers are sparsely located (http://www.ssn.unam.mx/website/jsp/red_sismologica.jsp) in the zone of tremor 150 - 270 km from the trench in Central Mexico or about 100 - 230 km south of Mexico City (Kostoglodov *et al.*, 2008). In order to detect tremor, a temporarily installed array of seismometers has been used (Payero *et al.*, 2008). The array, called the Meso American Seismic Experiment (MASE), was made up of 100 seismometers spaced every 5-6 km running from Acapulco, through Mexico City to near the Gulf Coast (Pérez-Campos *et al.*, 2006). The 1D line of stations used for data analysis has not allowed for true 2D or 3D locations of NVT using the method of Wech and Creager (2008). As a result, a simple automatic tremor detection algorithm has been created which only attempts to detect the tremor occurrence. Once it is detected, it can be located using cross-correlated windows and location-inversion or grid-search. However, in the case of a 1D line, we limit the location to distance from the trench along the array profile. Also, this method recognizes that NVT occurs over an area. Subsequent modeling of the source will be used in future to locate NVT over an area or volume. This paper presents only the NVT detection method.

Automatic detection

The frequency range where NVT occurs is 1-10 Hz (Obara, 2002). NVT can be seen visually on the spectrogram of an individual station as shown in Fig. 1A. It is observed as continuous low-level (green-yellow) energy distributed evenly across 1 - 10 Hz over a period 10's of minutes to hours. Local earthquakes are easily distinguishable visually as spikes which have energy evenly spread across the 1 - 10 Hz frequency range (Fig. 1A). The spikes observed in Fig. 1A line up with the earthquakes in the gray trace of the seismogram in Fig. 1B. There are no teleseismic events in Fig. 1, but they are easily determined by the early energy across all frequencies rolling off to only low frequency (< 1 Hz) energy for 10's of minutes. Since earthquakes have much higher energy than NVT and have energy within 1 - 10 Hz, they must be removed or they will be selected as NVT incorrectly by an automatic detection routine. Therefore a large part of automatic detection is devoted to earthquake removal.
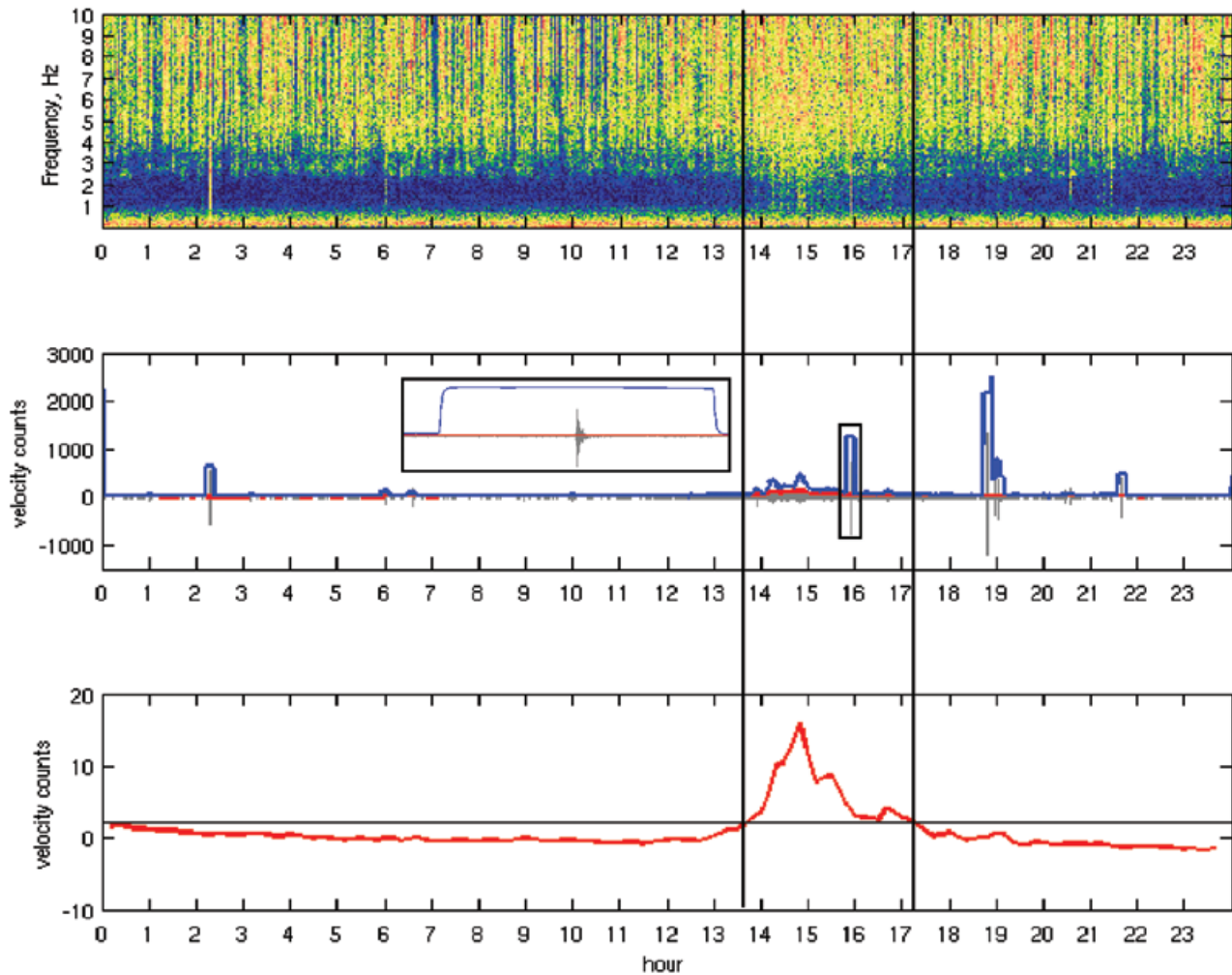
Fig. 1. (A) is an example of a day-long spectrogram. The red-blue-green color scheme represents the energy with red being the highest and blue the lowest energy of the frequency range. The 1-2 Hz frequency range has low noise (blue) overall. (B) depicts a comparison of the median filter and the standard smoothed enveloping technique (Obara, 2002) for NVT detection. The axis for the gray seismogram is on the left. The right axis is for the smoothed envelope (dark blue line) of the seismogram with no median filter, and for the output of the median filter (red line). There are many examples in the seismogram where the median filter (red) removes local earthquakes from the signal, whereas the smoothed envelope (blue) preserves the energy from local events. (C) depicts the final averaged signal from all stations (average of all red lines) with the final median removed. The black horizontal line is the empirically determined cutoff. The two vertical lines that cross the 3 sub-figures enclose the determined NVT.

Most of the MASE seismometers recorded high ambient noise above 2 Hz (Fig. 1). Therefore, we first applied a 1-2 Hz bandpass filter to remove the noise. A ±10 minute window median filter was applied to the absolute value of the filtered data to remove the spikes from local earthquakes (Fig. 1). It is important to note that this is different from a mean or averaging filter which leaves the effect of local earthquakes in the signal. The coda normalization method was used to determine and remove the site effect as presented in the following section. The trend was then removed to reduce the effect of long duration rain storms. Storms typically appear across many

stations in the same frequency range as NVT. However, storms tend to have longer duration (6+ hours for storms, compared to 30 minutes to a couple of hours for NVT) and a smaller maximum than NVT. Detrending a signal can remove the effect of half a day of storms (Fig. 2).

The individual median values were removed from each day for all stations as a further step to reduce the effect of local noise. The signal was then averaged across all stations to end with a final, day-long common signal for all stations (Fig. 1). The averaged signal amplified coherence among stations and lowered noise from lone
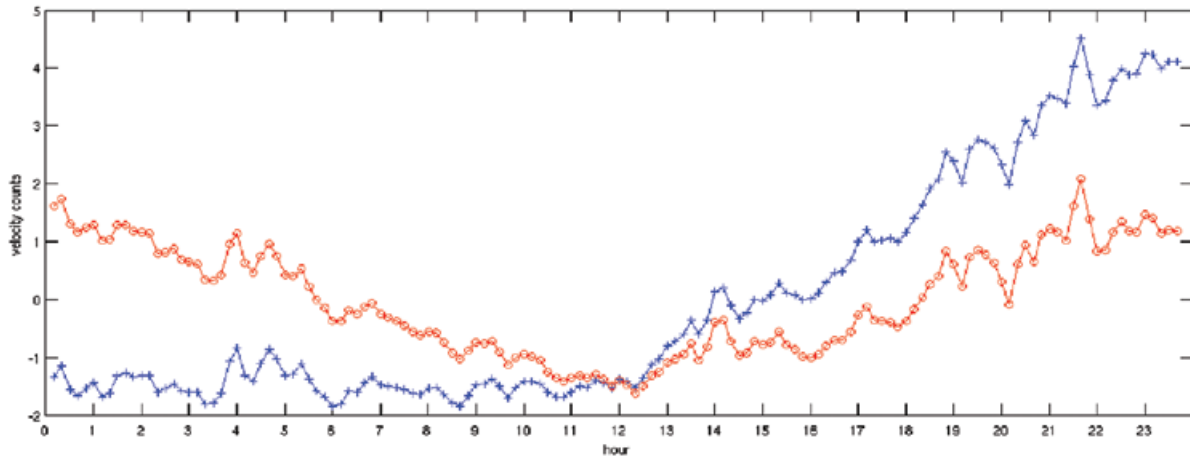
Fig. 2. A 24-hour median filtered seismic signal. The mean has been removed for clarity as detrending removes the mean automatically as well. The dark blue line with the '+' symbol is the unaltered signal. The lighter red line with 'o' symbols has had the trend removed. There is a storm in the signal for the last 7 hours. With the threshold (solid line) set at 2.25 velocity counts, the signal that still has a trend has amplitudes high enough in the last 4 hours that they would erroneously be considered NVT. The detrened signal has no amplitudes high enough to be considered NVT. There was no NVT this day.

stations. We also attempted to multiply the signals of all stations, instead of doing a summed average, in order to amplify the difference among coherent signals and noise even further. However, we were unable to find a simple consistent method to normalize the multiplied signal, which is very important as not all stations were available for the entire period of the MASE (Husker *et al.*, 2008). The method to normalize a summed average is to divide by N, the number of signals. However, when multiplying, the divisor for normalization of multiplied signals is X^N, where X must be determined as the average amplification for all stations for the entire experiment. Instead of trying to determine the correct X for the experiment, we found that the average of the daily seismograms determined NVT as well as visually analyzing spectrograms, as will be shown in the results.

The empirically determined cutoff of average amplitude of 2.25 normalized velocity counts was used to determine NVT (Fig. 1). At least two consecutive data points had to be above this cutoff in order to be distinguished from the noise. Finally, from all identified events a few were removed from the catalog when the amplification of the signal was higher at the Acapulco station on the coast than at any other station. This last step was perhaps unique to Mexico and the MASE data. The zone where NVT occurs begins ~50 km inland from the coast. Storms are always stronger at the coast. Removing high amplification events at the coast thus removed storms that passed through the previously applied storm-removal steps.

Coda normalization

The coda normalization method is based on the empirical observation that coda waves consist of S waves scattered at random heterogeneities in the earth, and at lapse times greater than the S-wave travel time the seismic energy is uniformly distributed around the source. Assuming that coda waves are composed of scattered waves (e.g., Aki and Chouet 1975; Rautian and Khalturin, 1978), the coda envelope recorded at station i during event k can be expressed as:

$$A_{ik}(f, t) = S_k(f) R_i(f) I_i(f) G(f, t),$$

where f is the frequency, t is the lapse time, $S_k(f)$ is the source term, $R_i(f)$ is the site term, $I_i(f)$ is the instrument response, $G(f, t)$ is the term describing the energy of scattered waves that is approximately equal at all stations at lapse times significantly larger than the direct S-wave travel time.

In order to compare the relative site amplifications between two stations, the coda from a single event must be measured at both stations over the same frequency band and lapse time. Then, assuming a common decay curve of coda, the envelope amplitude ratio for two different stations is independent from the event source term:

$$E_{if}(f, t) = \frac{A_{ik}(f, t) = G_i(f) I_i(f)}{A_{jk}(f, t) = G_j(f) I_j(f)}$$

Thus, the ratio of coda recorded at two stations at the same lapse time from the same event is free of source and path effects and depends only on the local site amplifications at the two stations (and instrument response).

The relative site-amplification factor of each station was estimated using up to 140 regional earthquakes with magnitudes ranging from 4 to 6 and epicentral distances less than 150 km around all the stations. The seismograms were band-pass filtered between 1 and 2 Hz and the envelope determined using the Hilbert transform. The absolute values of the envelopes were smoothed using a running average of 10 seconds. The normalization factor, E_{is}^n, for each component, n, was determined relative to a reference station ($j = s$). The reference station (SAFE) was chosen due to having the largest number of recorded events. Relative amplitude was determined by taking the average ratio of coda amplitude for each lapse-time window for time greater than about twice the direct S-wave travel time up to the end the coda. The end of the coda was assumed to be when the signal to noise ratio is equal to 2 (Fig. 3). The ratios from all events at their lapse-time windows were then averaged for each station to obtain normalization coefficients, $C_i^n = \langle E_{is}^n \rangle$.

It is interesting that normalization coefficients show a clear tendency to increase when the altitude of the station is greater than $H\sim700$ m (Fig. 4). They are almost the same for the NS and EW components (slightly higher for the NS) while the vertical component coefficients are less sensitive to the elevation change and reach only $\sim50\%$ of the horizontal ones for $H>700$ m. This observation deserves further special study.

Results

The threshold in the automatic NVT detection program was set to give nearly the same number of hours of NVT as picking by visual analysis of spectrograms: 461 hours automatically compared to 451 visually (Fig. 5). Visual inspection was done with only about 5 spectrograms for speed of the N-component for each event. Automatic detection used data from 15 - 20 stations. In general, automatic detection and visual determination of the NVT found the same events, but the exact duration of each varied between the two methods with humans determining longer durations in general (Fig. 5).
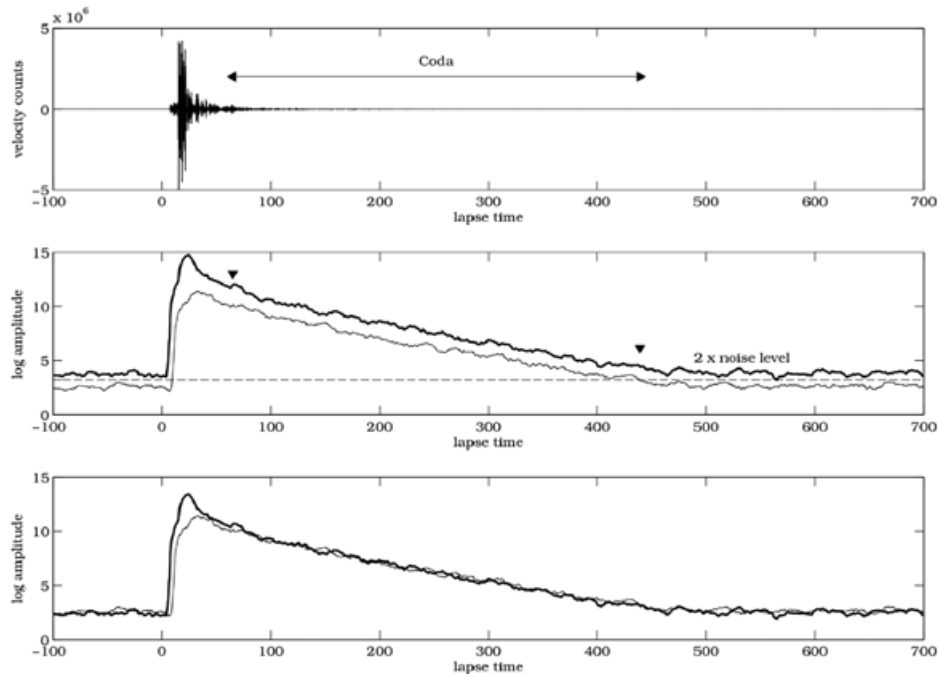

Fig. 3. A: 1-2 Hz band-passed velocity record of M_L 4.7, 2005/05/26 earthquake. The coda window used is shown by the arrow. B: Envelopes plotted as a function of lapse time after the origin time of the earthquake, for an example station (thick line), and the reference station (thin line). The dashed line shows 2 times the ambient noise level and the inverted triangles show 2 t_s (S-wave arrival) and the end of the section of the coda we used. C: The absolute amplitude of the envelope at the example station is adjusted to the reference using the coda normalization coefficient.
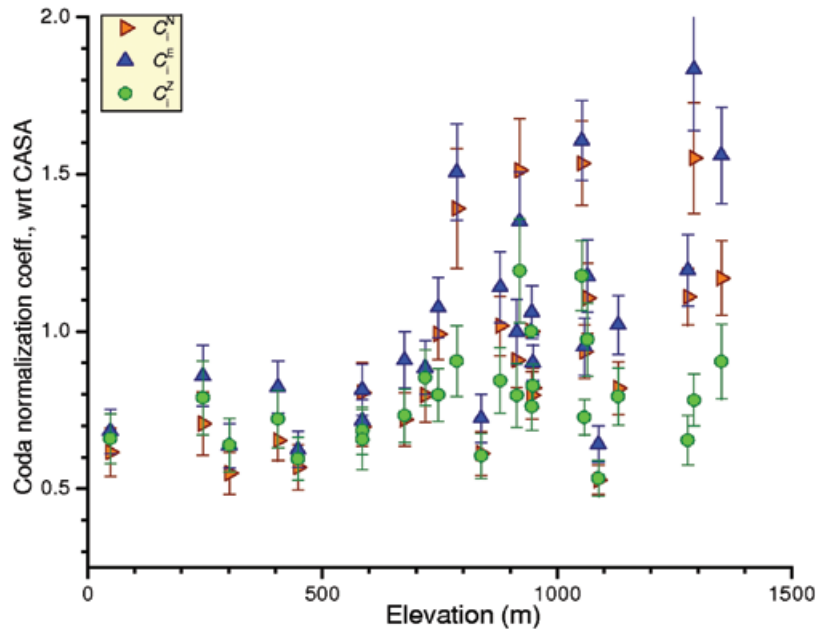
Fig. 4. Coda normalization coefficients, C_i^n, versus elevation of the seismic station. C_i^n generally increases when the elevation of the seismic station is over ~700 m.
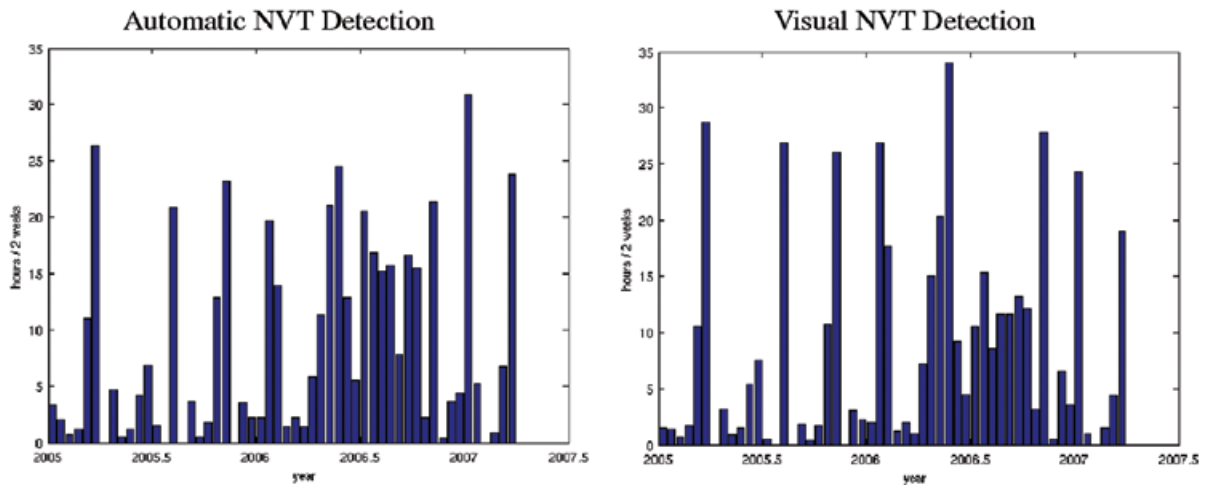

Fig. 5. A comparison of automatic (A) and visual (B) determination of NVT. The bins are the number of hours of NVT in a 2 week period. The data are from 24 MASE stations.

The greatest difference in hours of NVT measured in a single month occurred during the month of May, 2006. This was during the SSE that occurred from the end of March 2006 to the end of October 2006. The automatic detection program found almost 9 hours less NVT than visual inspection in the second half of the month. NVT for the month of May 2006 (Fig. 6) were reanalyzed using all

spectrograms with all 3 components to determine problems with the automatic detection scheme. Of the 13 additional NVT determined by the automatic method not included in the visual catalog, only 2 were found to be incorrect; one was a storm and one was a teleseism. However, there may have been NVT during the teleseism, but it was not possible to distinguish between the two. Of the 14 NVT

unidentified by the automatic method but included in the visual catalog, only 4 were determined to be real events. The rest were actually storm noise. Thus, the automatic detection actually proved better than analyzing a subset of spectrograms visually to determine NVT. Analysis of all spectrograms was necessary to be as accurate as the automatic detection method.

The most important criteria to determine real and false event determinations, was that 1 - 2 Hz energy was seen on contiguous stations for at least 20 minutes. Since the original visual detection method used only a handful of stations, at times there was ambient noise that occurred on multiple stations at the same time which lead to false positives. When all 15 - 20 stations were included it was clear that 'in-between' stations did not exhibit the same ambient energy. Undetected events also occurred during visual inspection due to only examining a handful of seismograms. False positives in the automatic detection routine were almost exclusively due to storms. Storms in general last longer than NVT. NVT rarely lasts for a few hours. Storms usually last at least a few hours. Also, on stormy days, in general there is more background noise (wind) with multiple occurrences of raised and lowered energy as the storm rises and dies so that they are identifiable visually. Unfortunately, the threshold for the automatic detection had to be set high enough to not falsely identify the majority of storms, so occasionally low energy tremor was not correctly identified as was the case with the 4 unidentified NVT in May.

Although teleseismic events have long durations and high energy within the 1-2 Hz bandwidth, only 1 of the 23 M > 7 earthquakes (Advanced National Seismic System - ANSS) recorded during the MASE installation was falsely identified as NVT (May 3, 2006 Mw = 8 Tonga earthquake). However, it appears that an episode of NVT may have occurred at the same time as mentioned above. Thus, large teleseismic earthquakes do not pose a problem in automatic detection of NVT, as large earthquakes do not retain energy above 1 Hz for the 20 minute requirement we place on a positive identification of NVT.

Conclusions

We have developed an automatic NVT detection routine which is both faster and more accurate than analysis our initial search for NVT using a subset of spectrograms. In addition, only 1 empirical value must be set for NVT selection, the cutoff for the averaged day-long amplitude of the network, making transfer of this method to other networks easy. Unlike other automatic NVT detection methods, the detection step is separated from the localization step, making it unnecessary to make assumptions of unrealistic point sources or clustering of those sources. This automatic algorithm may be easily implemented into the near-real-time analysis of daily seismological data for detecting NVT.
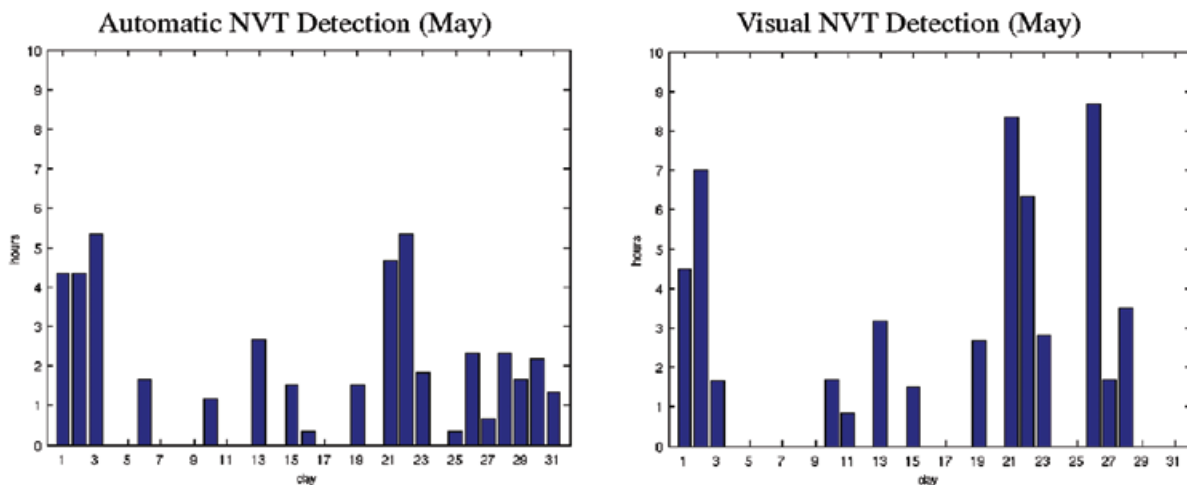

Fig. 6. A comparison of automatic (A) and visual (B) determination of NVT in May 2006. The bins are the number of hours of NVT daily. It can be seen that many of the NVT bursts are left undetected by the visual method. The NVT bursts not detected by automatic detection are due their energy being below the threshold for measurement.

Acknowledgements

Bibliography

Aki, K. and B. Chouet, 1975. Origin of coda waves; source, attenuation, and scattering effects. *J. Geophys. Res.*, *80*, 3322-3342.

Husker, A., I. Stubailo, M. Lukac, V. Naik, R. Guy, P. Davis and D. Estrin, 2008. WiLSoN: The Wirelessly Linked Seismological Network and its application in the Middle American Subduction Experiment. *Seis. Res. Lett.*, *79*, 438-443; DOI: 10.1785/gssrl.79.3.438.

Ide, S., G. C. Beroza, D. R. Shelly and T. Uchide, 2007. A scaling law for slow earthquakes. *Nature*, *447*, 76 – 79, doi:10.1038/nature05780.

Kao, H., P. J. Thompson, G. Rogers, H. Dragert and G. Spence, 2007. Automatic detection and characterization of seismic tremors in northern Cascadia. *Geophys. Res. Lett.*, *34*, L16313, doi:10.1029/2007GL030822.

Klein, F. W., 2007. User's Guide to HYPOINVERSE-2000, a Fortran program to solve for earthquake locations and magnitudes. *U.S. Geological Survey Open-File Report 02-171 revised Version 1.1*, 125 pp.

Kostoglodov, V., N. Shapiro, K. M. Larson, J. S. Payero, A. Husker, J. A. Santiago and R. W. Clayton, 2008, Nonvolcanic Tremor Activity is Highly Correlated With Slow Slip Events, Mexico, *Eos Trans. AGU, 89*(53), Fall Meet. Suppl., Abstract U31B-05.

McCausland, W., S. Malone and D. Johnson, 2005. Temporal and spatial occurrence of deep non-volcanic tremor: From Washington to northern California. *Geophys. Res. Lett.*, *32*, L24311, doi:10.1029/2005GL024349.

Obara, K., 2002. Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science*, *296*, 1679–1681.

Obara, K., H. Hirose, F. Yamamizu and K. Kasahara, 2004. Episodic slow slip events accompanied by non-volcanic tremors in southwest Japan subduction zone. *Geophys. Res. Lett.*, *31*, L23602, doi:10.1029/2004GL020848.

Payero, J. S., V. Kostoglodov, N. Shapiro, T. Mikumo, A. Iglesias, X. Pe´rez-Campos and R. W. Clayton, 2008. Nonvolcanic tremor observed in the Mexican subduction zone. *Geophys. Res. Lett.*, *35*, L07305, doi:10.1029/2007GL032877.

Pérez-Campos, X., R. Clayton, A. Iglesias, S. Singh, A. Husker, P. Davis, C. Valdes-Gonzalez and I. Stubailo, 2006. MesoAmerican Seismic Experiment; imaging the subducting slab. *Seis. Res. Lett.*, *77*, SSA Annual Meeting, no.2, pp. 292.

Rautian, T. G. and V. I. Khalturin, 1978. The use of the coda for determination of the earthquake source spectrum. *Bull. Seis. Soc. Amer.*, *68*, 923-948.

Rogers, G. and H. Dragert, 2003. Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip. *Science*, *300*, 1942– 1943.

Shelly, D. R., G. C. Beroza, S. Ide and S. Nakamula, 2006. Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. *Nature*, *442*, 188-191.

Shelly, D. R., W. L. Ellsworth, T. Ryberg, C. Haberland, G. S. Fuis, J. Murphy, R. M. Nadeau and R. Bürgmann (2009), Precise location of San Andreas Fault tremors near Cholame, California using seismometer clusters: Slip on the deep extension of the fault?, *Geophys. Res. Lett.*, *36*, L01303, doi:10.1029/2008GL036367.

Simkin, T. and L. Siebert, 2002-2009. Global Volcanism FAQs. *Smithsonian Institution*, *Global Volcanism Program Digital Information Series*, *GVP-5* (*http://www.volcano.si.edu/faq/*).

Wech, A. G. and K. C. Creager, 2008. Automated detection and location of Cascadia tremor. *Geophys. Res. Lett.*, *35*, L20302, doi: 10.1029/2008GL035458.

Yeh, K. and C.-H. Liu, 1982. Radio wave scintillations in the ionosphere. *Proc. IEEE*, *70*, 324– 360.

A. Husker[1*], S. Peyrat[2], N. Shapiro[2], V. Kostoglodov[1]

[1]*Instituto de Geofísica, Universidad Nacional Autónoma de México, Ciudad Universitaria, Del. Coyoacán 04510, Mexico City, Mexico*
[2]*Laboratoire de Sismologie, Institut de Physique du Globe de Paris, Paris, France*
Corresponding author: allen@geofisica.unam.mx