IMPROVING MAGNITUDE DETECTION THRESHOLDS USING MULTISTATION, MULTIEVENT, AND MULTIPHASE METHODS

David P. Schaff and Felix Waldhauser

Lamont-Doherty Earth Observatory, Columbia University

Sponsored by Air Force Research Laboratory

Contract No. FA8718-05-C-0022

ABSTRACT

Research was conducted on a correlation detector (multievent and multiphase) for semiempirical synthetic tests and a regional case study in Xiuyan, China. The semiempirical runs took a 50 s window on an Lg-wave recorded at 750 km distance filtered from 1 to 3 Hz and embedded it 300,000 times in real, continuous, background seismic noise. The noise was selected for 36 days spread throughout the year to capture diurnal and seasonal variations. No screening for random, unknown signals in the noise was performed. A correlation detector has a 50% probability of detection with 1.5 false alarms per day for a signal-to-noise ratio (SNR) of 0.32, which corresponds to a full magnitude unit reduction in detection threshold over a standard STA/LTA technique. A scaled cross correlation coefficient performs slightly better with 1 false alarm per day and has fewer false triggers on unknown, random signals. Summing the cross correlation traces together for all three components enhances the detection signal similar to beamforming. A correlation detector summing the three components together has a 96% probability of detection with zero false alarms in 36 days for a SNR of 0.32. The case study looked at 90 events in Xiuyan, China, recorded at stations 500 to 1500 km away. Clear detection spikes are observed on all three components agreeing to the nearest sample allowing for constructive interference on the summation of the correlation traces. In contrast, detection maximums for unknown, random signals do not align to the nearest sample on the three components and destructively interfere with the summation. This is probably the single best indication of a true detection, the fact that three independent components all show clear spikes to the nearest sample. It is also observed that semisimilar events can provide useful detections (events that are not exactly co-located or don't have identical mechanisms). Two examples of an aftershock buried in the coda of a larger event demonstrate new detections that were not previously reported in the available catalogs. A scaled cross correlation detector is able to detect 90 out of 90 or 100% of the events in the catalog using Pg, Pn, and Lg phases, whereas a standard STA/LTA detector like that employed by the Prototype International Data Center (pIDC) finds 10 out of 90 or 11%. Comparison with a known local/regional catalog shows this represents a 1.3 unit reduction in magnitude threshold.

OBJECTIVES

Research will be conducted on techniques for generating a multiphase detection bulletin derived by two means: station array processing (multistation) and source array processing (multievent). Comparison and quantification of improvement over standard P-wave, single-event, single-station procession will be evaluated.

RESEARCH ACCOMPLISHED

Introduction

The first year of this project has focused on a correlation detector or multievent technique. Waveform cross correlation has a long history of improving locations and identifying events (e.g., Poupinet et al., 1984; Harris, 1991). Recently it appears that large percentages of events may be similar enough to enable correlation detectors to be applied on a broad scale (Schaff and Richards, 2004; Schaff and Waldhauser, 2005). Figure 1 shows an Lg-wave recorded 750 km away embedded in various levels of real background seismic noise. The waveforms are filtered from 1 to 3 Hz. Signal-to-noise ratio (SNR) values are shown to the left and are computed as a mean absolute value like an STA/LTA filter uses. A standard trigger used by the pIDC of 3.2 shows a corresponding signal that an energy detector (STA/LTA) would find in this case corresponding to a magnitude 3.5. The relationship between SNR and magnitude is

$$\text{Mag} = \log(\text{SNR}) + 4.3 - \log(\text{SNR}_{\text{Mag}=4.3}) \; .$$

Reducing the detection threshold by 0.5 magnitude units it can be seen that the SNR drops to one. It is impossible for an STA/LTA filter to detect a signal at the noise level. To reduce the detection threshold by a full magnitude unit to 2.5 corresponds to an SNR of 0.32 or where the signal is one-third of the noise level. Figure 2 displays the cross correlation traces for each of these signals correlated with the master template. Clear detection spikes can be seen all the way down to magnitude 2.5 giving us the first indication that a correlation detector can reduce the detection threshold by a full magnitude unit. In the next section we will explore if such detections occur with acceptably low false alarm rates.

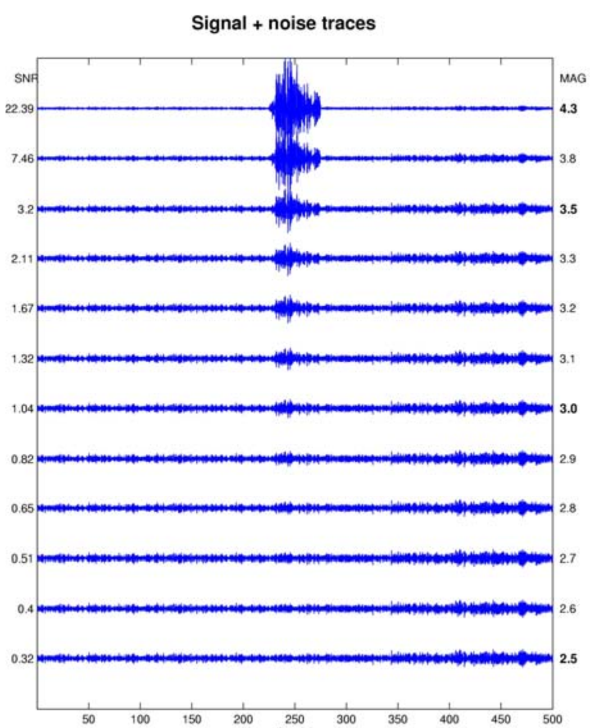

Figure 1. Lg-wave signals embedded in increasing noise.

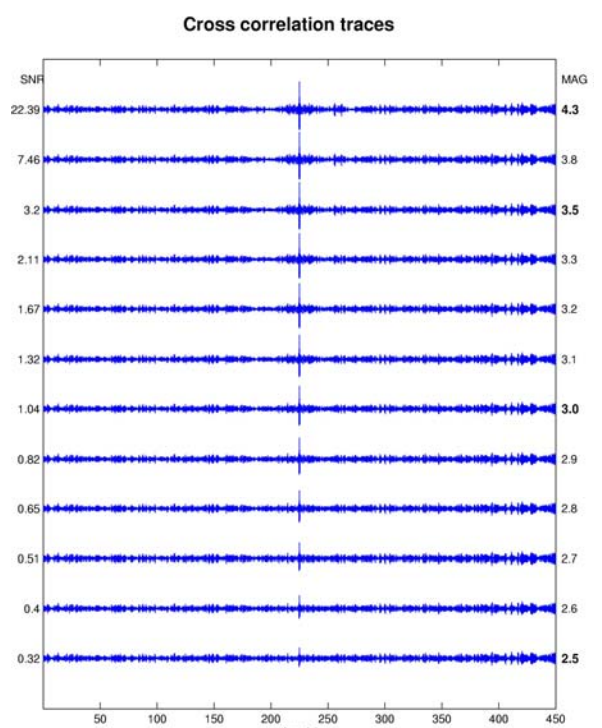

Figure 2. Cross correlation traces corresponding to Figure 1.

Figure 3 shows the maximum cross correlation coefficient (CC) as a function of SNR highlighting the three magnitudes of interest. It can be seen that detecting half a magnitude unit lower at 3.0 would easily be picked up with a cross correlation coefficient above 0.8. Remember that this corresponds to a signal right at the noise level. To detect a full magnitude unit lower, however, at 2.5 has a low CC of 0.367. Normally this type of measurement would be discarded for location purposes even though there was a clear detection spike relative to background levels. The reason can be seen in Figure 4 where you have two dissimilar traces in the top panel. The cross correlation function of these two waveforms corresponds to the first trace in the middle panel with a coefficient of 0.364. Therefore, even though these two waveforms are unrelated they have a similar coefficient to the case of an identical signal buried in noise (0.367). The second trace in the middle panel is the last correlation trace from Figure 2 corresponding to a SNR of 0.32. A detection threshold of 0.35 would trigger on both of these examples. The first case would be a false alarm, whereas only the second is the true detection that we want to capture. One solution to this problem is to note that in the second case the maximum is high relative to background values. We can apply an STA/LTA filter to the cross correlation traces, which is shown in the bottom panel. We choose for the window length of the STA one sample and the window of the LTA 20 s. Here the maximums differ by a substantial amount (5.8 for the dissimilar case and 8.3 for the identical signal buried in noise). Gibbons and Ringdal (2006) employ a similar procedure that they call a "scaled CC," where they scale the CC by the root mean square (RMS) within some window of the background levels. Because our STA/LTA filter uses mean absolute value we divide by that instead of RMS but the effect is basically the same. It can readily be seen that a scaled CC threshold of 6 would weed out the dissimilar case and would leave us with the true detection. We will use this value later as our threshold in our case study in Xiuyan, China.

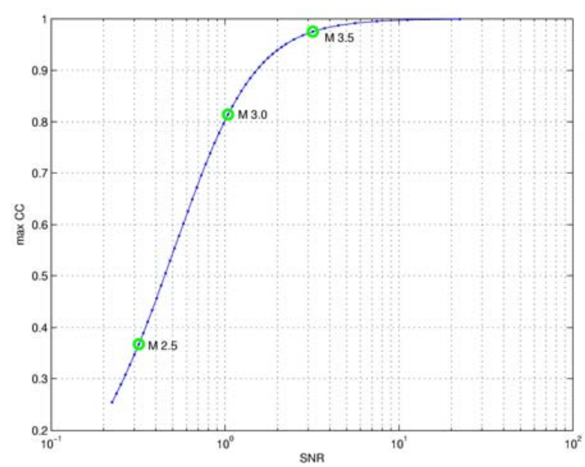

Figure 3. Maximum CC as a function of SNR from Figure 2.

Next we examine how the correlations perform for all three components. In Figure 5 we cross correlate signals for the BHN, BHE, and BHZ components in the top panel with the same signals plus noise in the second panel (SNR = 0.32) to obtain the cross correlation traces in the third panel with CCs around 0.3 as before and clear detection spikes at 225 s. If we stack the cross correlation traces we see that the spikes constructively interfere and give a new maximum of 0.97. There is an overall enhancement in the spike because the background levels deconstructively interfere. We can estimate the level of this enhancement. Define $R = \frac{\max cc}{\sigma}$. Assume σ^2 is the variance for all three components. For the stack the variances add, $\sigma_{stack}^2 = 3\sigma^2$. Then $R_{stack} = \frac{3\max cc}{\sqrt{3\sigma^2}} = \sqrt{3}R = 1.732R$. We see that the variances of the three components do add to 0.0137, which is close to 0.0139 of the stack, indicating that they are approximately normally distributed. Also R_{stack} is enhanced by 1.7 times the average R as expected. This is the same improvement in signal enhancement that is achieved by beam forming.

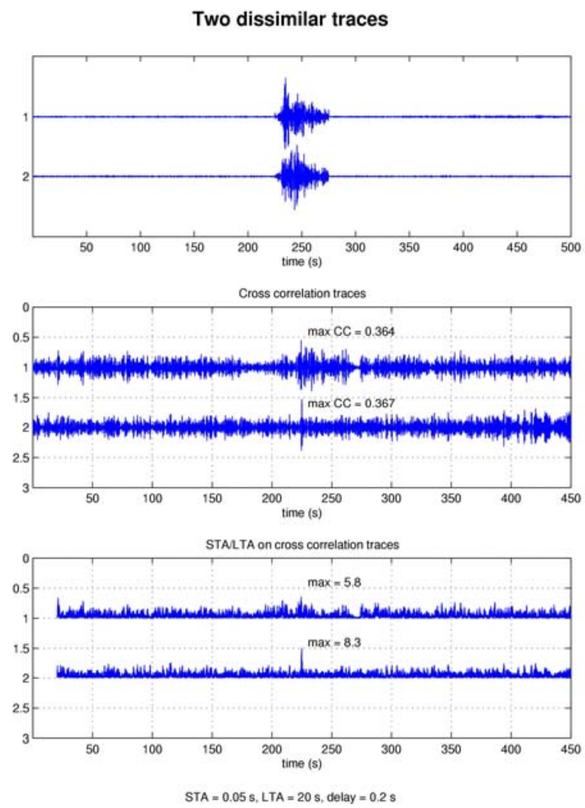

Figure 4. Two dissimilar traces would produce a false alarm with CC but not with a scaled CC.

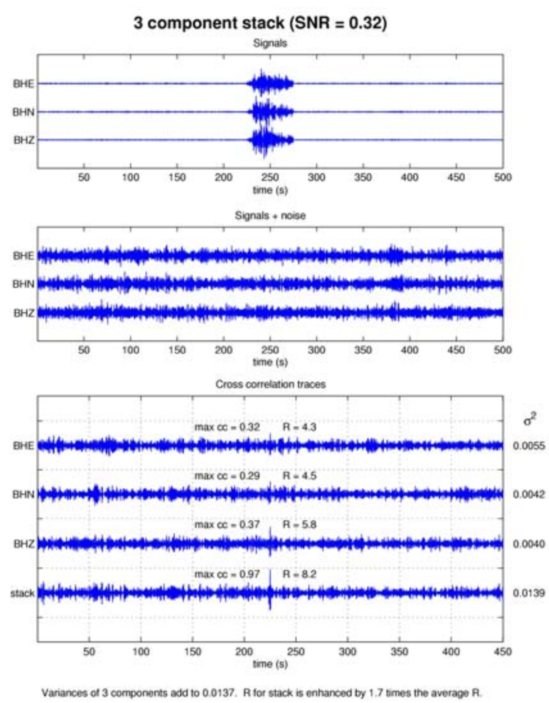

Figure 5. Three components enhance the detection spike.

Semiempirical Synthetic Runs

Now we look more in detail at statistics of detection and false alarm rates. We take 36 days of real seismic noise spread throughout the year to capture diurnal and seasonal variations. The noise also contains random seismic signals of unknown origin that have not been removed. For the master signal we choose a 50 s window on the Lg-wave of Figure 1 recorded at 750 km distance. The waveforms are all filtered from 1 to 3 Hz. The sample rate is 20 Hz. We embed the signal in the noise ~300,000 times. The noise comprises 62 million samples. Figure 6 shows the histograms for the signal and noise distributions for CC for a SNR of 0.32. There is a clear separation between the two allowing for detections to be made with a certain threshold. The mean CC is about 0.35 for the signal buried in noise as before. A Receiver Operating Characteristic plot in the lower panel shows probability of detection as a function of probability of false alarm, which can be computed from the top panel. For a probability of detection of 0.5, there is less than one in a million chance of a false alarm. Given the number of samples per day at 20 Hz this corresponds to 1.5 false alarms per day. This is a reasonable false alarm rate, and so we therefore conclude that a correlation detector is able to detect a signal buried in the noise one full magnitude unit lower than a standard energy detector. Figure 7 shows the same signal and noise distributions for a SNR of 0.32 but this time for the scaled CC or signed STA/LTA on the CC trace. Again a clear separation of the distributions is obvious. The mean of the signal buried in noise is around 6. This time, however, the probability of detection at 0.5 corresponds to a slightly lower false alarm rate of one per day.

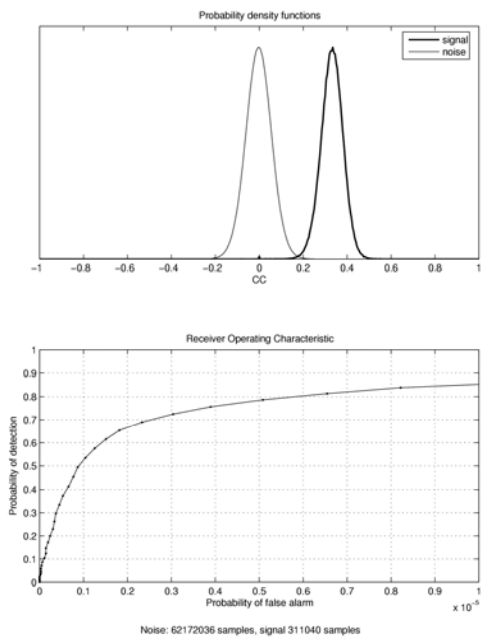

Figure 6. Statistics for a signal buried in noise (SNR = 0.32) for CC.

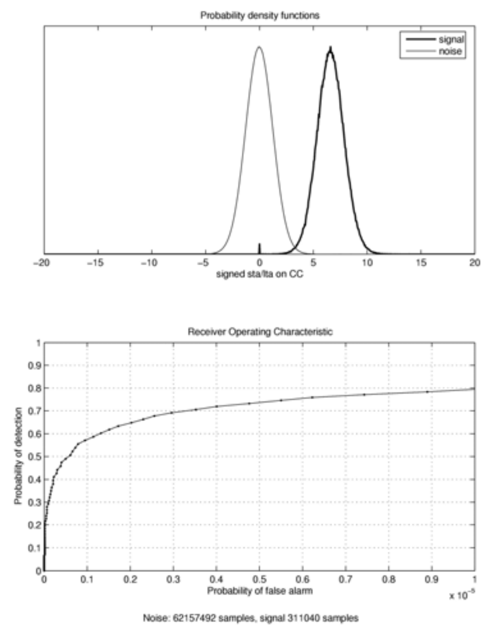

Figure 7. Statistics for a signal buried in noise (SNR = 0.32) for scaled CC.

If we increase the SNR to 1.012, corresponding to half a magnitude unit reduction in detection threshold over an STA/LTA filter then we see in Figure 8 that the signal and noise distributions are extremely well separated. The mean CC for the signal is above 0.7 indicating a high degree of similarity. In this case there are zero false alarms for the entire 36 days considered at a probability of detection of 99.996%. This is rather remarkable that even though the signal is at the noise level there is nearly a 100% probability of detection with a zero false alarm rate. Figure 9 is the same as Figure 6 for a signal buried in noise at the SNR = 0.32 level except all three components are used to enhance the detection as seen in Figure 5. Note that the mean CC is about the same but that the width of both the signal and noise distributions has been reduced by summing the variances. The narrower distributions cause less overlap of the probability density functions and produce a better Receiver Operating Characteristic curve. This time there are zero false alarms in 36 days with a probability of detection of 96.5% compared to the one false alarm per day rate at the 50% level before.

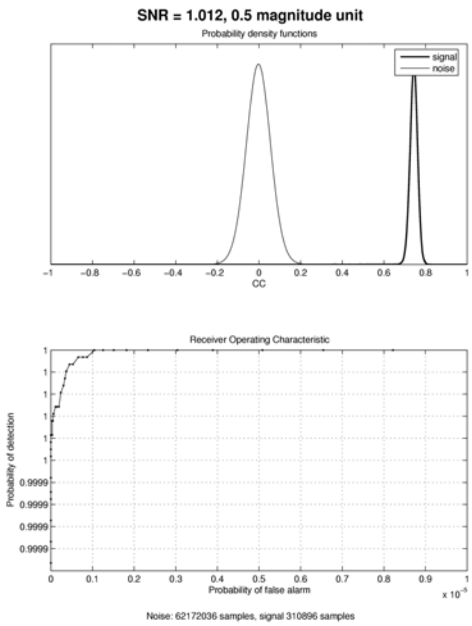

Figure 8. Statistics for a signal buried in noise (SNR = 1.012) for CC.

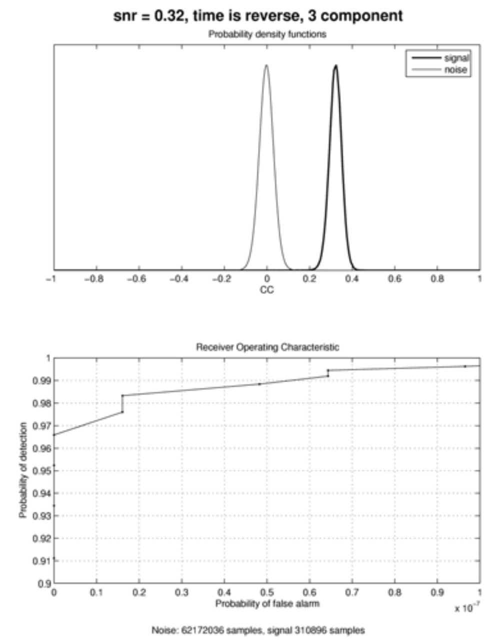

Figure 9. Statistics for a signal buried in noise (SNR = 0.32) for CC with three component enhancement.

1999 Xiuyan, China, Case Study

Ninety events are examined in the 1999 Xiuyan, China, earthquake sequence recorded at stations 500 to 1,500 km away. The cross correlation matrix for station IC.BJT is shown in Figure 10. The windows chosen are centered on the Lg-waves filtered from 0.5 to 5 Hz. Clusters of similar events appear with warm colors as blocks on the diagonal. Several values of the CC are quite high above 0.8. Other colors in cyan are at the 0.35 range that we were looking at before for the detection of a small signal buried in noise. The question is whether these values provide reliable detections. Figure 11 shows the cross correlation traces for the events in the cluster from indices 56 to 79. It is seen that there are clear detection spikes on the vertical, north, and east components. In addition, it can be seen that the spikes constructively interfere and are enhanced on the average of the three components compared to the background levels. This is the clearest indication that we have a true detection for these events—the fact that the spikes all align to the nearest sample for basically three independent tests. Visual inspection of the cross correlation traces for the other clusters in Figure 10 shows similar behavior.

If the waveforms do not correspond to a similar event this phenomenon is not observed on the three components. Figure 12 shows the case where correlations are made with an unknown dissimilar signal. The bottom panel shows the cross correlation traces in different colors for the waveforms on the three components of the top panel. The annotated text occurs where the cross correlation is a maximum for each component (CC). Although the values are in a similar range to the case of a buried signal in noise (0.27, 0.34, 0.23) it is seen that they occur over 100 s apart from each other. The average does not constructively interfere but is much less at 0.16 and is 17 s away from the nearest peak on any component. This is strikingly different than the correlation spikes that align to the nearest sample for the case in Figure 11.

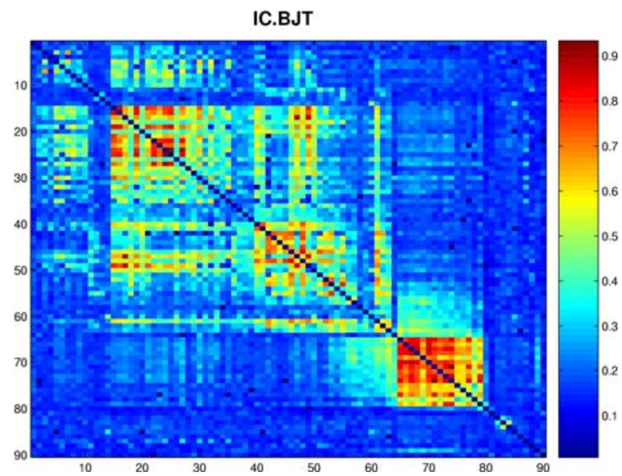

Figure 10. Cross correlation matrix for 90 events in Xiuyan, China.

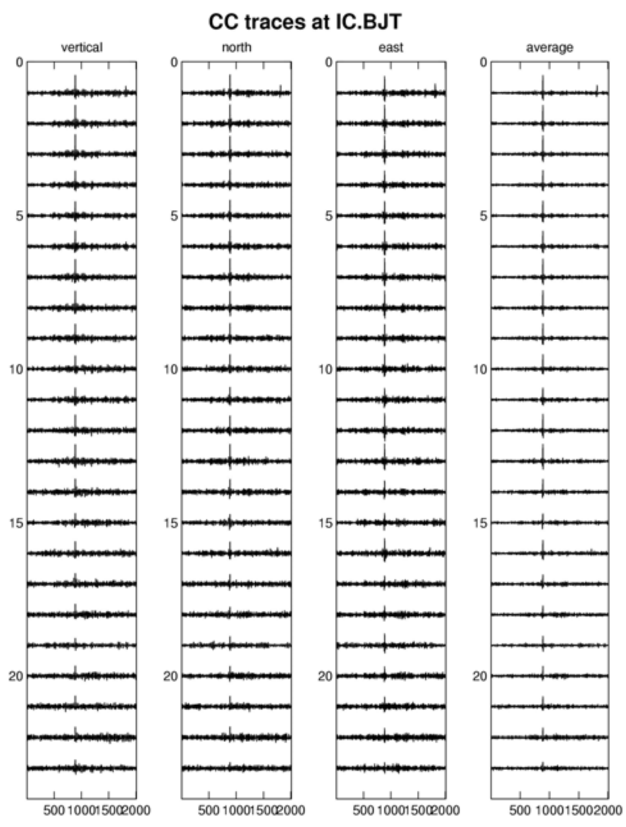

Figure 11. Cross correlation traces for 24 events.

To automatically detect the spikes above the background levels, we use a scaled CC, using a threshold of 6. Running this for all five regional stations and the phase Pg, Pn, and Lg gives the results shown in Figure 13. A blue dot means that event pair in the matrix satisfied the detection threshold of 6. Again similar events are arranged to be blocks on the diagonal. The number beside each phase is the number of events at that station matching the criterion out of 90. The largest amplitude Lg-wave that also has the longest duration of energy in the window produces the best detections and at station MDJ detects 90 out of 90, or 100% of the events. Stations BJT and HIA also have a high number of detected events for Lg. Pg is the next most detected phase and then Pn. Multiphase detections and detections at multiple stations further ensure that these are true detections.

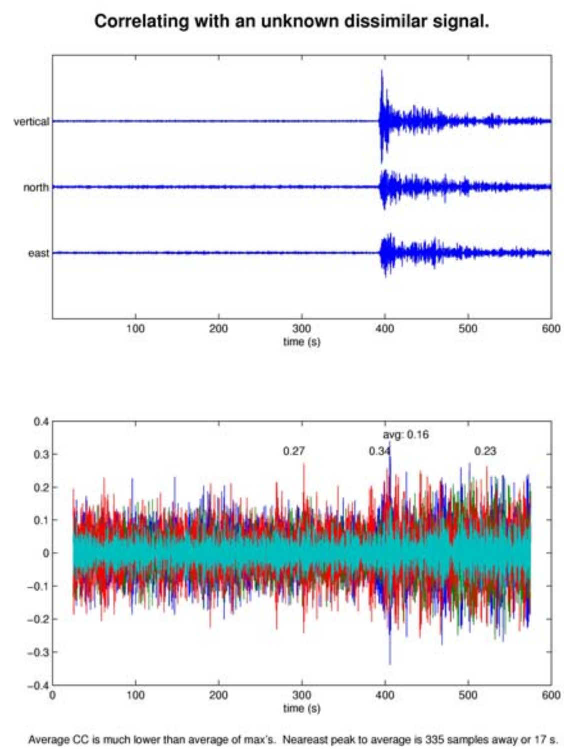

Figure 12. Cross correlation traces for three components of a dissimilar event.

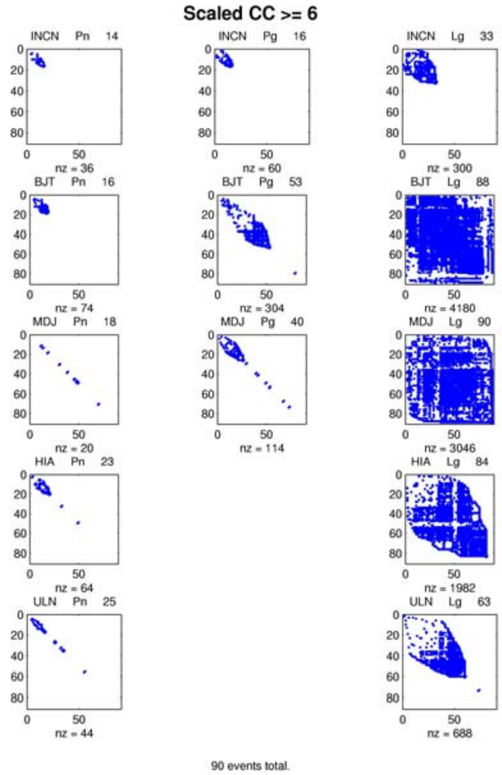

Figure 13. Detection matrices based on scaled CC for 90 events.

CONCLUSIONS AND RECOMMENDATIONS

- CC can detect one magnitude unit lower than STA/LTA.
- Three component averaging enhances the same as beam forming.
- CC has 50% probability of detection at 1.5 false alarm/day for SNR 0.32.
- Scaled CC has 50% probability of detection at 1 false alarm/day for SNR 0.32 because it has fewer triggers on unknown random signals.
- CC on 3 comp. has 96% probability of detection with zero false alarms in 36 days for SNR 0.32.
- 1999 Xiuyan, China, case study shows a CC detector finds 90 out of 90, or 100% of the events, whereas a STA/LTA detector like the pIDC finds 10 out 90, or 11%.
- Comparison with a known local/regional catalog shows this represents a 1.3 unit reduction in magnitude threshold.

ACKNOWLEDGEMENTS

We are grateful to the operators of the Chinese Digital Seismic Network (CDSN) and to the IRIS data management center for archiving the waveforms used in this project.

REFERENCES

Gibbons, S. J. and F. Ringdal (2006). The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Int.* 165: 149–166.

Harris, D. B. (1991). A waveform correlation method for identifying quarry explosions, *Bull. Seism. Soc. Am.* 81: 2395–2418.

Poupinet, G., W. L. Ellsworth, and J. Frechet (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras Fault, California, *J. Geophys. Res.* 89: 5719–5731.

Schaff, D. P. and P. G. Richards (2004). Repeating seismic events in China, *Science* 303: 1176–1178.

Schaff, D. P. and F. Waldhauser (2005). Waveform cross-correlation-based differential travel-time measurements at the Northern California Seismic Network, *Bull. Seism. Soc. Am.* 95: 2446–2461.