

One Magnitude Unit Reduction in Detection Threshold by Cross Correlation Applied to Parkfield (California) and China Seismicity

by David P. Schaff and Felix Waldhauser

Abstract From a local high-resolution base catalog at Parkfield, California, 5076 earthquakes (M 0.2 to 6) are used to study the comparative performance of a correlation detector and standard energy detector on the sparse regional network of continuously operating stations. Eighty-six percent of the events detected by a standard energy detector can also be detected by cross correlation. Correlation detection is able to find additional events by lowering the detection threshold by about 1 unit beyond what standard processing detects for Parkfield, a factor of 10 increase in number of events such as those predicted by Gutenberg–Richer. Most event separation distances for events that correlate at Parkfield are less than 1 km. The distribution of magnitude differences for events that correlate at Parkfield is not distinguishable from the input magnitude distribution. More robust measures to quantify reduction in detection threshold are introduced. Detection magnitude threshold reduction of about 1 unit holds for large-scale application to the 18,886 events in China and 5,076 events in Parkfield with false-alarm rates of a few percent. Large and small events are seen to correlate well enough for detection. Two examples are shown with magnitude differences as large as 2.3 and 3.3 units. The correlation detector also finds two cases of buried aftershocks in the coda of mainshocks that were previously unreported in the Annual Bulletin of Chinese Earthquakes (ABCE).

Introduction

Waveform cross correlation is a basic tool that has found applications in many fields and across many disciplines. But when it comes to observational seismology, waveform cross correlation has been used comparatively little for event detection as compared to earthquake location. Most work and practical applications in seismic event detection have concentrated on the standard power detector technique, where the energy in a short-term average window (STA) is divided by a long-term average window (LTA) and a detection is triggered when this ratio exceeds some signal-to-noise ratio (SNR) threshold (Freiberger, 1963).

A correlation detector, also known as a matched filter, is the optimal means of detecting a known signal in the presence of Gaussian, white noise (Van Trees, 1968). Some earlier work on correlation detectors on a local scale (Withers *et al.*, 1999) and on a global scale (Young *et al.*, 1996) used raw data preprocessed with an STA/LTA filter to reduce the influence of mechanism and source differences and matched the waveforms with synthetics computed in a 1D velocity model. Other earlier work also used matched filters for global detection of earthquakes (Shearer, 1994; Ekstrom, 2006). Subspace detectors are related to correlation detectors in that they match the waveforms to a linear combination of basis waveforms (Harris, 2006; Harris and Paik, 2006). This has the advantage

of reducing the number of master templates needed and accounting for less than perfect waveform matches. Cross correlation as a detector has been applied to the NORSAR array in Norway with great success (Gibbons and Ringdal, 2004, 2005, 2006; Gibbons *et al.*, 2007). Shelly *et al.* (2006) introduced a novel, direct, scientific application of a correlation detector to identify low-frequency earthquakes in nonvolcanic tremor. More recent work has used matched filters to detect early (repeating) aftershocks of the 2004 Parkfield earthquake (Lengline and Marsan, 2009; Peng and Zhao, 2009).

We investigate the potential use of correlation detectors on a broad regional scale to improve seismic monitoring and to quantify the reduction in magnitude detection thresholds. Smaller case studies at NORSAR demonstrated a full magnitude unit reduction in detection threshold (Gibbons and Ringdal, 2006; Gibbons *et al.*, 2007). Our previous work has similarly confirmed from semi-empirical analysis (Schaff, 2008) and a case study in Xiuyan, China, (Schaff, 2010) that an order of magnitude improvement is possible comparing a correlation detector for similar events with a standard STA/LTA detector. This unit reduction in detection threshold is achieved with acceptably low false-alarm rates of about one per day or a probability of false alarm of 6.1E–7 (Schaff, 2008).

Schaff (2009) was the first study to apply correlation detectors on a large scale to about 19,000 events in China. It was seen that a large percentage (85%) of the events that a standard STA/LTA detector finds are also detected by cross correlation. This was consistent with our findings that 95% of the events in northern California correlate well enough for location purposes (Schaff and Waldhauser, 2005). There also was a 70% increase in additional events that were detected by cross correlation, but not by the standard detector in China. To quantify the reduction in detection threshold, Schaff (2009) used a simple 95% confidence lower limit that amounted to 0.2 magnitude units in reduction. This was less than expected based on work from the previous case studies (Gibbons and Ringdal, 2006; Gibbons *et al.*, 2007; Schaff, 2010), but it seemed that completeness of the catalog to lower magnitudes may have been the limiting factor. When detection threshold reduction was considered as a function of station distance, it was found that at large station distances the reduction reached 0.9 units and presumably was no longer affected by the catalog not being complete to smaller magnitudes. For this paper we follow similar procedures that were performed for the large-scale application in China (Schaff, 2009), but we now apply it on a large scale to a new region at Parkfield, California. We also introduce more representative and robust measures to quantify the reduction in detection threshold, providing a comparison between results for Parkfield and China.

Large-Scale Application to Parkfield, California

At Parkfield the magnitudes range from 0.2 to 6 for our data set. The tectonic setting comprises a creeping section that has lots of repeating events and streaks that correlate well and a locked portion with events that are less correlated (Waldhauser, Ellsworth, *et al.*, 2004; Thurber *et al.*, 2006). The strategy we employ is to analyze events obtained from a more dense local catalog and perform the correlation detections on a sparse network of regional stations. This allows us to answer two key questions: the first is the percentage of seismicity that we are able to detect by cross correlation assuming that the base catalog is complete to some lower magnitude; the second is to quantify the reduction in detection thresholds since we have independent magnitude information. Waveforms from 5076 events in the base catalog from Thurber *et al.* (2006) that were accurately located using the local dense short-period network at Parkfield, California, were correlated against each other at seven continuously operating regional stations available from IRIS DMC (Fig. 1). There are actually more events in the catalog of Thurber *et al.* (2006), close to 9000, but 5076 events were the ones for which waveforms were available at the regional stations. *P*, *S*, and *Lg* waves were all analyzed with 15 s, 20 s, and 50 s windows, respectively; 10 s, 15 s, and 30 s lags were searched over, respectively, forward and backward for the time-domain cross correlation. *Lg* waves are high frequency, multiply reflected, large amplitude arrivals on regional seismograms with long durations coming

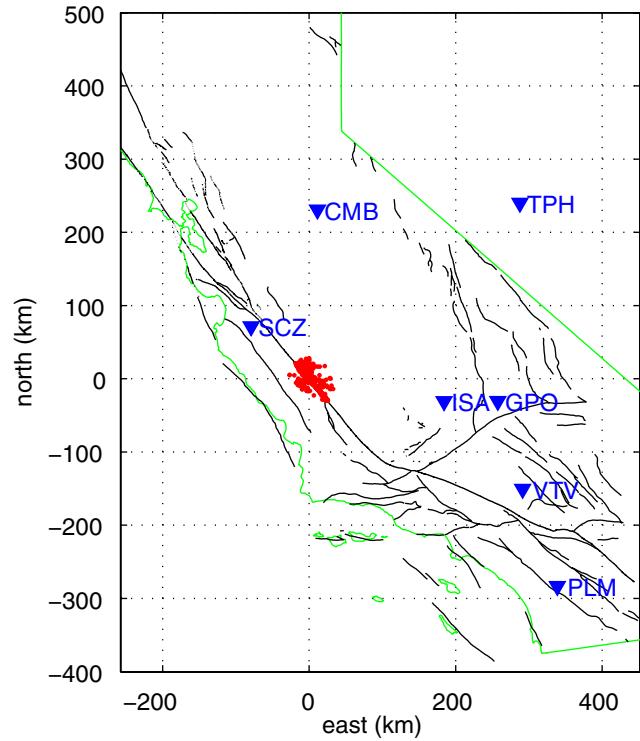


Figure 1. Five thousand seventy-six events at Parkfield (dots) processed at seven regional broadband stations (triangles) for a correlation and standard detector. Faults and California state border are shown. The color version of this figure is available only in the electronic edition.

in after the *S* waves. The windows were centered based on the predicted travel times for the three waves. The windows are chosen to be long to increase the significance of the cross-correlation values with a higher time-bandwidth product. The filter bands were 0.75 to 2 Hz, 0.5 to 3 Hz, and 0.5 to 5 Hz, respectively. All pairs of events that occurred within 6 km of each other were considered amounting to 53 million correlations and about 2 weeks of continuous processing time on a four processor computer. We compute all possible correlations using every event as a master event against all the other events within reasonable distance thresholds to quantify the best results obtainable from a complete library of master templates.

We employ the same method of using a scaled cross-correlation coefficient (SCC) on cross-correlation traces that are averaged across all three components for the particular phase to enhance the detection signal by constructive interference that was utilized in Schaff (2008, 2009, 2010) to determine a trigger whether it be a true detection or a false alarm. Each point in the cross-correlation trace, CC_i , is scaled by the mean absolute value of the moving window (length N) n points before the point

$$SCC_i = \frac{CC_i}{\frac{1}{N} \sum_{j=i-N-n}^{i-1} |CC_j|}. \quad (1)$$

We use a moving window of 20 s and choose n to be 4 samples to avoid side-lobes of the cross-correlation function. Intuitively, SCC is a measure of the statistical significance of a detection spike because it quantifies the deviation of the cross-correlation coefficient from an empirical distribution of background values based on a moving window throughout the correlation trace.

In our China study (Schaff, 2009) we estimated false alarms by computing the false-alarm rate per day at a single station by running a single master template over 36 days of continuous data. Then we distributed this false-alarm rate over the individual pairwise observation matrices for all stations to compute the total number of false alarms (Schaff, 2009). For Parkfield, we improve our estimation of false-alarm rates in a more empirical approach by applying the algorithms in an identical manner with the same parameters except for shifting the windows 120 s before the expected P -wave arrival. The idea behind this is that the windows should contain only noise and so any trigger for a given threshold would then be considered as a false alarm. This is the most robust method we know of for estimating false alarms for our comparisons since all the stations and event pairs are the same and the windows are centered on the noise characteristic right before the signal comes in for each individual station. It is possible since events (especially aftershocks) are seen to cluster in time and space that actual events may be detected in this 120 s window before another event; therefore, we could overestimate the number of false alarms using this procedure. However, for this data set only 1 event out of the 5076 in the catalog is within 120 s of the origin time of another event. These two events are 15 km apart and do not correlate well. Therefore, we conclude that our estimate of false alarms is accurate and that most likely none of the triggers are real events in our study area.

Table 1 shows how we determine the number of true detections for each station and phase using this procedure. For example, the P waves at station SCZ had 4096 observations with $\text{SCC} \geq 6$ for time windows centered on the expected signal arrival. This threshold was initially chosen because 6 was empirically determined to have low false-alarm rates (Schaff, 2008). Comparing this to 623 observations with an $\text{SCC} \geq 6$ for the same processing except on noise allows us to estimate the number of true detections as the difference between the two or 3473 (Table 1).

Figure 2 shows the detection triggers as a function of SCC for P waves at station SCZ for the windows centered on the signal and the windows centered on the noise. The estimate of the true number of detections is the difference between these two curves. As can be seen the choice of SCC threshold is rather arbitrary since there is a trade-off as expected that lower thresholds produce more detections but also more false alarms. The choice depends on the desired application. For example, if the operator wants to be sure that none of the events are false alarms for this station and phase, a threshold of $\text{SCC} = 10.2$ has zero false alarms and 357 true detections. If, on the other hand, it is important

Table 1
Number of True Detections with $\text{SCC} \geq 6$

Station	Phase	Detections
SCZ	P	3473
SCZ	S	7658
SCZ	Lg	10612
ISA	P	1225
ISA	S	6228
ISA	Lg	12986
CMB	P	720
CMB	S	3002
CMB	Lg	3787
VTV	P	-37
VTV	S	23
VTV	Lg	404
TPH	P	27
TPH	S	268
TPH	Lg	515
PLM	P	13
PLM	S	281
PLM	Lg	479

not to miss any events, a lower threshold can be chosen, such as in the case of nuclear monitoring where it is essential that any events that might be nuclear explosions not be missed.

Figure 3 shows the distributions of the interevent separations for the event pairs that are detected with $\text{SCC} \geq 6.3$ occurring for two or more stations or phases (Fig. 3a) and compare it with the input interevent separations searched over (in this case, it was an earlier run considering separations out to 20 km instead of 6 km as indicated in Fig. 3b). To maximize the number of true detections based on the combined criteria, 6.3 was chosen empirically. From Figure 3a it is clearly seen that most detections occur within a kilometer of each other. This is independent confirmation that these are mostly true detections because the false alarms would tend to reproduce the distribution of the input observation event pair matrix seen in Figure 3b, which is not what is observed. Based on the distribution in Figure 3, we use 1 km as a cutoff for interevent separation distance.

The stations in Table 1 are listed in order of increasing station distance. It is observed that the two closest stations give the most detections. It is also seen that the Lg phase gives the most detections, even though it does not propagate as efficiently in California as compared with China. Stations farther away produce relatively few true detections compared with the number of false alarms. Therefore, we decided based on this initial screening of all seven stations and phases to use only Lg waves at stations SCZ and ISA. We further required a selection criteria of event separations less than 1 km, lags searched over to be 0.3 s, and $\text{SCC} \geq 5$. We did this to maximize the amount of detections while trying to minimize the number of false alarms. If an event pair meets the detection threshold, both events are considered detected because all events are treated as masters. Only one event pair has to meet these criteria at both stations to count as a detection. If we use selection criteria similar to what we used for China, however,

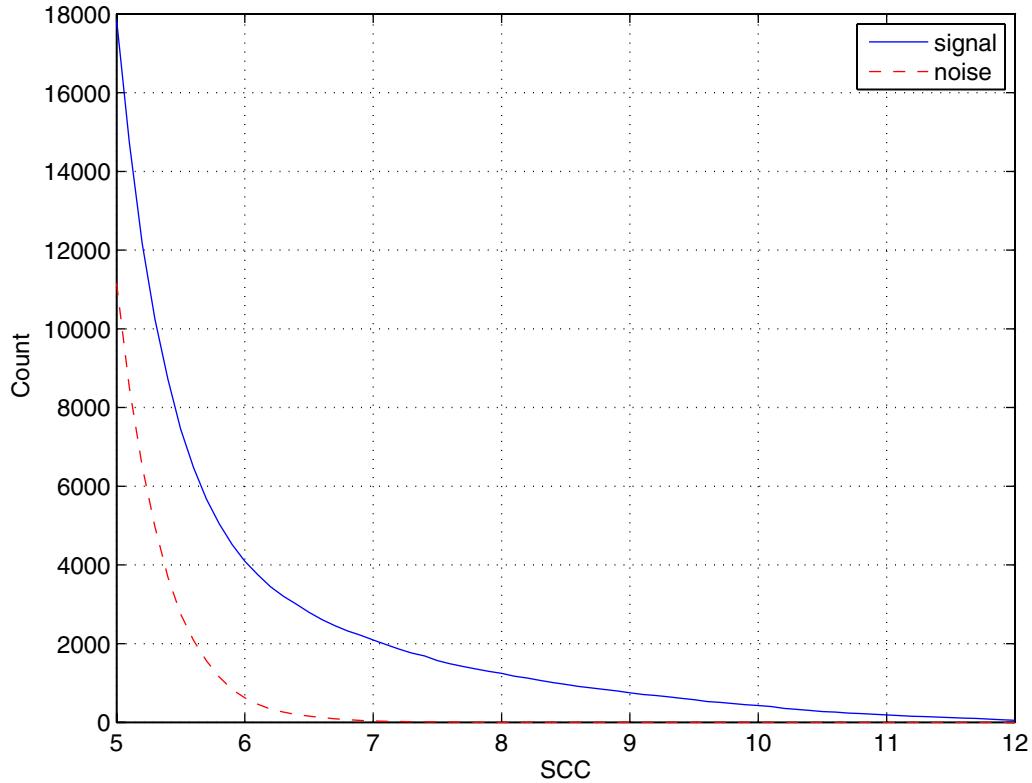


Figure 2. Number of detection triggers as a function of SCC for P waves at station SCZ for the windows centered on the signal and the windows centered on the noise. The color version of this figure is available only in the electronic edition.

with $\text{SCC} \geq 6$, all event separations and 30 s lags show that the overall statistics are not grossly different. Because we have good control on the locations at Parkfield with our high resolution catalog, we wanted to see how well the correlation detector could perform. Lags of 0.3 s correspond to as much

as 1 km relative location error for a group velocity of 3.5 km/s for the Lg waves, which is still a rather conservative estimate. Typical average relative location errors for the Parkfield catalog range from a few tens to a few hundred meters (Waldhauser, Ellsworth, *et al.*, 2004). Using 30 s

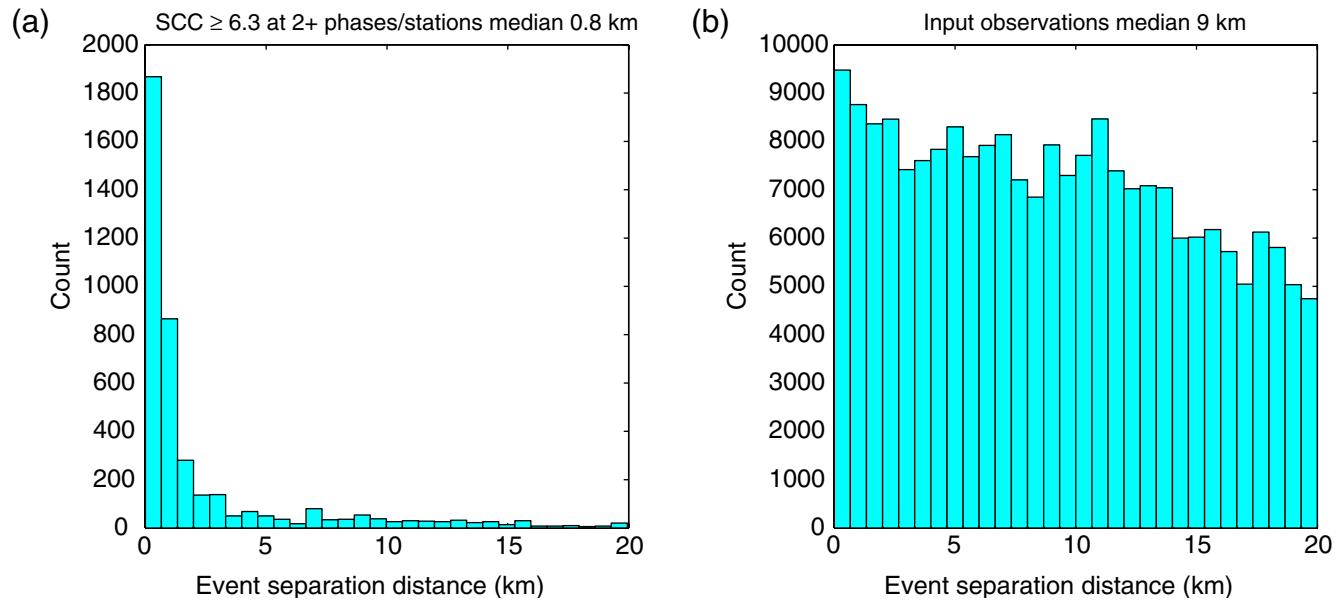


Figure 3. Interevent separation for (a) detected pairs and (b) input observation matrix for Parkfield events. The color version of this figure is available only in the electronic edition.

(two orders of magnitude larger) is then overly conservative for the high-resolution catalog we have for Parkfield, although it was helpful for China due to greater mislocation errors in the Annual Bulletin of Chinese Earthquakes (ABCE).

Using these selection criteria we find that 1357 events out of the 5076 or 27% are detected. The false alarms using these same criteria are 1.4%, so we estimate that the true detections are approximately 25.6%. We compare with the same simulated prototype International Data Center (“pIDC”) procedures for a standard STA/LTA detector as used for China (Schaff, 2009), except counting a trigger if it occurs within 5 s of the first arriving P wave this time since the locations are more accurate. We use an STA = 1 s, LTA = 60 s. Triggers at the pIDC range from 3.0 to 4.5, so we use 3.2 as a common value. Overlapping filter bands, 0.5–1, 0.75–1.5, 1–2, 1.5–3, 2.5–5, 4–8 Hz are used to increase the signal-to-noise ratio (SNR). And finally 3 or more stations triggers are required to call the event a detection. The “pIDC” procedures find 140 events out of the 5076. Figure 4 shows the magnitude distributions. It can readily be seen that the correlation detector finds approximately 10 times the number of events that the “pIDC” detects, which is what would be predicted from a magnitude unit reduction in detection threshold based on a Gutenberg–Richter magnitude–

frequency distribution. The correlation detector also finds 120 out of the 140 events that the “pIDC” detector finds or 86%.

The 95% confidence lower limits for the distributions in Figure 4 are: all (0.8), correlation (1.1), and “pIDC” (1.4). This represents a 0.3 magnitude unit reduction between the two techniques, which is not as great as expected from prior case studies where a full unit was measured (Gibbons and Ringdal, 2006; Gibbons *et al.*, 2007; Schaff, 2010). This time, however, the completeness of the catalog at Parkfield is not the reason we supposed it was for China (Schaff, 2009) since the base catalog has magnitudes well below that for both the correlation and “pIDC” detected distributions. Further examination brings up some issues and explanations. If we assume that a 1.4 lower limit is really representative of a 95% confidence level, we see that this gives 132 events for the “pIDC” out of a possible 1344 in the catalog or 9.8%, which is far below the 95% level and immediately presents a problem. Looking at the “pIDC” distribution another way, we see that 22 events out of 2581 possible in the catalog have magnitudes less than 2 or 0.85%. A false-alarm rate as low as 1% therefore could be the cause for these lower magnitude events that are detected in the “pIDC”. The reason is that the Gutenberg–Richter magnitude–frequency relationship means there are orders of magnitude for more smaller events than

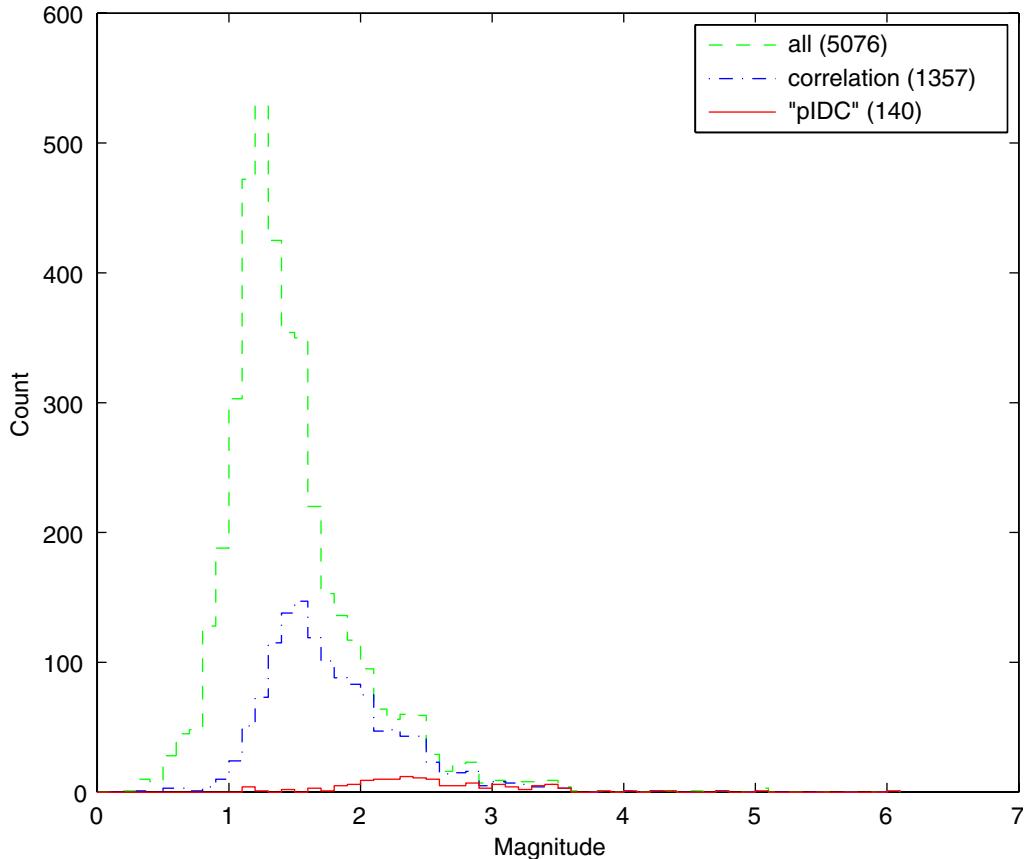


Figure 4. Magnitude distribution for the Parkfield catalog and correlation and “pIDC” detectors. Numbers in parentheses show total events. The color version of this figure is available only in the electronic edition.

larger events. Therefore, for a given false-alarm rate it is more likely that the false alarm will be a lower magnitude event than a larger magnitude. This causes a problem for determining lower limit magnitude detection thresholds, especially for the case when the total number of events is small (140 for the “pIDC”) compared with the background catalog (5076). So if 22 events out of the 140 for the “pIDC” are actually false alarms, that is 16% of 140. Therefore, a 95% threshold would include false alarms. One idea for a solution is to use a different threshold (e.g., 100% – 16% = 84%). If we consider a 90% lower limit, then correlation is at 1.2 whereas the “pIDC” is 1.8, therefore the reduction in threshold is 0.6 units. An 85% confidence lower limit has correlation (1.3) and “pIDC” (1.9) with a threshold reduction of 0.6 unit. This is closer to what is expected and presumably reflects the improvement more accurately since they are not as influenced by the presence of false alarms.

More Robust and Representative Measures of Reduction in Detection Threshold

We explore a more representative measure of detection thresholds that takes into account the predominance of smaller magnitude events according to the Gutenberg–Richter magnitude-frequency distribution. The final detected magnitude distribution is the product of the input distribution with the distribution associated with each detection method that acts as a filter on the input distribution. What we are really interested in quantifying are the statistics of the distribution associated with each filter and not the product of the two distributions, which is severely dominated by smaller events. To get the distribution of each detection filter, we therefore must take the final detected distribution and divide it by the input magnitude distribution of the base catalog.

Looking at Figure 4 again a more intuitive limit may be the first value where the detector finds 50% or more of the available events within a given magnitude range. This would avoid the problem seen before that a 95% threshold actually only captured 9.8% of the events. We create a graph of what we term a normalized probability density function (PDF) associated with the detection filter in Figure 5 by taking the “pIDC” detector curve of Figure 4 and dividing it by the curve for the entire catalog. This changes the number of detections into percentages that are not skewed by the large numbers of small events. If all the events were detected in the catalog, then the PDF would be flat at unity (100% detection) across all magnitudes.

The left vertical line on Figure 5a indicates the lower magnitude threshold of 3.0 corresponding to the first point at which 50% of the background catalog is detected in that magnitude bin. The right red line on Figure 5a marks where at least 70% of the events in the catalog are detected at 3.5. An energy detector should be able to detect more large magnitude events. Drops in the curve for higher magnitudes may be due to gaps in data coverage for certain events and the statistics of small numbers. Figure 5b shows the normalized PDF for the

correlation detector by taking the correlation detector distribution from Figure 4 and dividing it by the distribution for all the events. The 50% red line on Figure 5b corresponds to a magnitude bin of 1.6 and the 70% red line falls on the 1.9 bin. The difference between the two detectors for this detection threshold measure is a reduction of 1.4 units for the 50% level and 1.6 units for the 70% level. This improvement is consistent with what we observed before with a full magnitude unit reduction in the previous studies (Gibbons and Ringdal, 2006; Gibbons *et al.*, 2007; Schaff, 2010). It also is a more representative and intuitive measure since a 50% value corresponds to the point where half of the earthquakes in the catalog at that magnitude range are actually detected instead of a 95% confidence limit giving the misleading information that 95% of the events are detected whereas only 9.8% actually were. Additionally by normalizing according to the PDF of the background catalog we account for the Gutenberg–Richter magnitude-frequency relationship and reduce the impact of false alarms at lower magnitudes since a percentage is considered instead of a number.

An alternative way to measure the detection threshold is to convert the normalized PDF to a normalized cumulative density function (CDF) in Figure 5c,d. The CDF is computed as the cumulative sum of the PDF in Figure 5a,b and then normalized to one. A 95% confidence lower limit on the normalized CDF for the “pIDC” detector corresponds with magnitude 2.2 on Figure 5c. For the correlation detector the 95% confidence lower limit corresponds with a magnitude 1.3 in Figure 5d, so the reduction in threshold is 0.9 units using the normalized CDF as a measure. At the 90% confidence lower limit the values are 1.5 for the correlation detector and 2.4 for the “pIDC” detector with a reduction in threshold again of 0.9 units. It seems that the normalized CDF is more robust of a measure of threshold since it was a 0.9 reduction for both 95% and 90% confidence limits, whereas the normalized PDF ranged from 1.2 to 1.4. The CDF curves are also smoother than the PDF curves. Finally, a magnitude threshold that captures a normalized percentage of all earthquakes above a certain value instead of within a specific magnitude range seems more robust and less sensitive to the details and shape of the particular bins of the magnitude distribution of the background catalog.

From Figure 5d it appears that M 1.3 is a good measure of the detection limit for the correlation detector for these stations at Parkfield. Figure 6 shows all the events in the high resolution catalog at Parkfield with $M \geq 1.3$. Events in black are detected by cross correlation and events in gray went undetected. It can be seen that the detected events are spatially well distributed across the fault and at all depths. They tend to occur in tightly clustered areas. The undetected events occur in more diffuse areas. The density of events within 1 km around detected events above M 1.3 is greater than that around the undetected events, which may be the main reason why they went undetected. The reason why the larger events were not detected is because of the interevent distance thresholds being restricted to 1 km. A run with separation out to 20 km

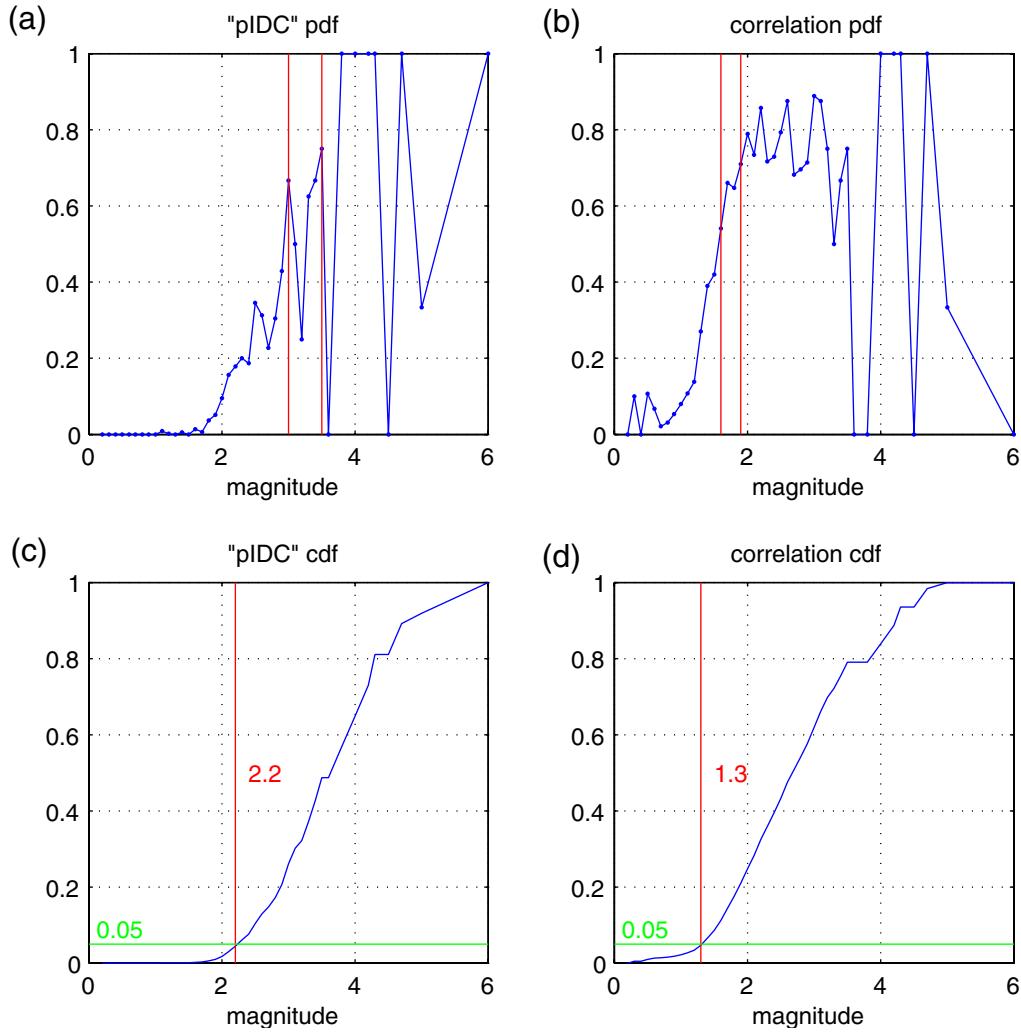


Figure 5. Normalized probability density function (pdf) for (a) the “pIDC” computed as that distribution in Figure 4 divided by the distribution for all the events at Parkfield and for (b) correlation as that distribution in Figure 4 divided by the distribution for all the events. Vertical lines denote points where 50% and 70% of the events in that magnitude bin are first detected. Normalized cumulative density function (cdf) computed from pdf in Figure 5a and 5b as the cumulative sum normalized to one for (c) “pIDC” and (d) correlation. Intersection of vertical and horizontal lines corresponds to the 95% confidence lower limit for the cdf. The color version of this figure is available only in the electronic edition.

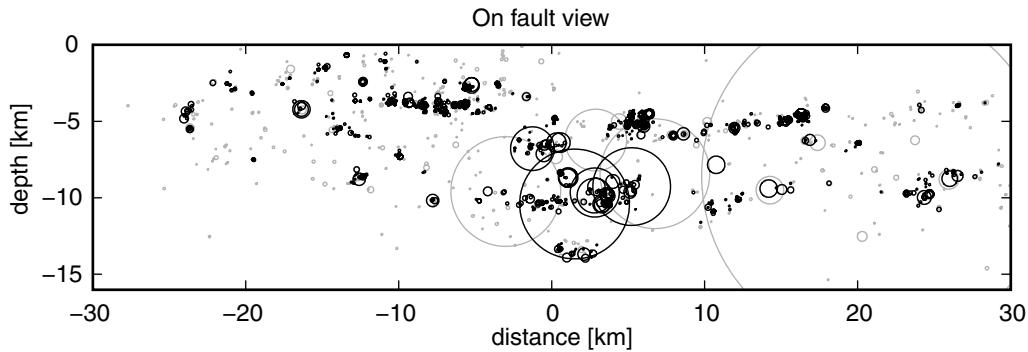


Figure 6. On fault view of events with $M \geq 1.3$ in the Parkfield catalog. Black circles are detected by cross correlation and gray circles are undetected. Circle size is estimated source area based on a 3 MPa stress drop.

captured all of these larger magnitude events. The reason for this is that larger magnitudes tend to correlate better with other large magnitudes even with greater separation distances as opposed to closer smaller events (Schaff *et al.*, 2004).

Large-Scale Application to China Revisited

For comparison, we reanalyze the results obtained on a large scale for China (Schaff, 2009) with these more representative and robust measures of magnitude detection threshold reduction. A summary of that paper is as follows. For China similar processing was carried out for 18,886 events in the ABCE. Correlations for all pairs with interevent separation distances of 150 km were performed on 50 s windows centered on the Lg waves searching over 30 s lags forward and backward. For China, only Lg waves were considered, which performed well since they propagate well over long distances for that tectonic region. The seismograms were filtered from 0.5 to 5 Hz. The results showed 12,902 events were detected by correlation out of 18,886 or 68%. Correlation finds 7063 out of 8358 events found by the “pIDC” detector or 85%. Correlation finds an additional 5839 events over the 8358 from the “pIDC” or a 70% increase. The magnitude distributions are in Figure 7. The 95% confidence lower limit for the detection thresholds was 2.8 for the cor-

relations and 3.0 for the “pIDC” or a 0.2 unit reduction in threshold. In that paper we noted, however, that 2.8 is also the 95% confidence level for all the events, so the correlations may be able to detect smaller events if the catalog was complete to lower magnitudes. By analyzing the detection threshold reduction as a function of station distances, we observed that the reduction increased with increasing station distance approaching 0.9 magnitude units for large distances. Presumably these were not as affected by the completeness of the catalog as for shorter station distances (Schaff, 2009).

Figure 8a,b displays the normalized PDFs for the “pIDC” and correlation detectors for the curves in Figure 7 for China. The 50% detection levels are 3.7 for the “pIDC” and 2.1 for correlation amounting to a 1.6 unit reduction in threshold. For the 70% level (chosen as the local maximum for the “pIDC”) the values are 4.2 for “pIDC” and 2.4 for correlation corresponding to a 1.8 reduction in threshold. Similarly the normalized CDFs for China are shown in Figure 8c,d. The 95% confidence lower limit is 3.0 for “pIDC” and 2.2 for correlation, corresponding to a reduction in threshold of 0.8 units. For the 90% confidence lower limit it is 3.5 for the “pIDC” and 2.5 for correlation corresponding to a 1.0 unit reduction in threshold.

Table 2 summarizes and compares the results for reduction in detection thresholds using three different measures for

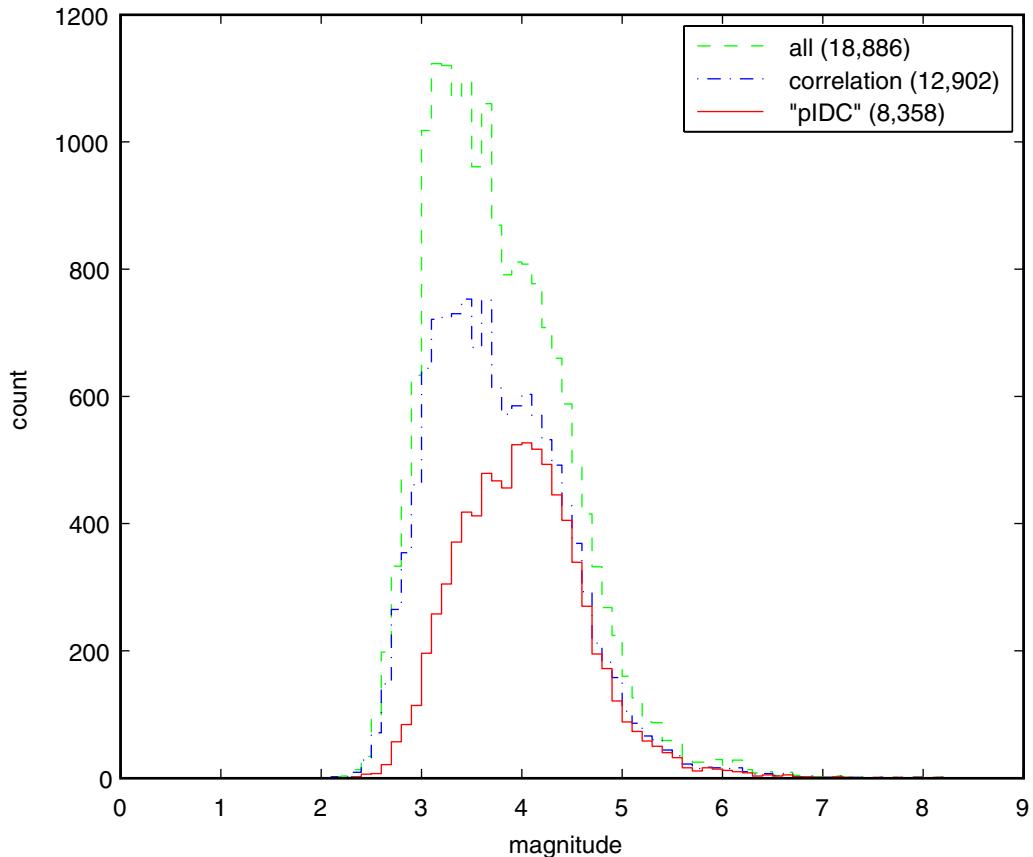


Figure 7. Magnitude distribution for ABCE catalog for China and correlation and “pIDC” detectors. Numbers in parentheses show total events. Figure is adapted from figure 2 of Schaff (2009). The color version of this figure is available only in the electronic edition.

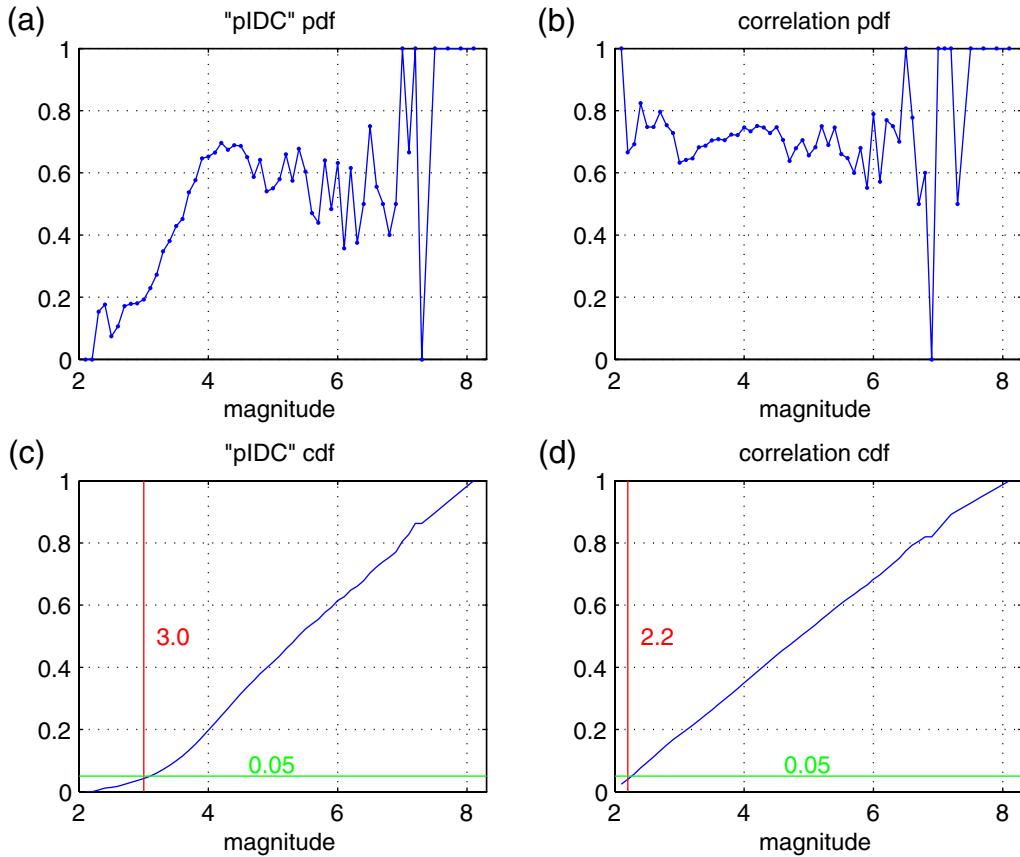


Figure 8. Normalized probability density function (pdf) for (a) the “pIDC” computed as that distribution in Figure 7 divided by the distribution for all the events in China and for (b) correlation as that distribution in Figure 7 divided by the distribution for all the events. Normalized cumulative density function (cdf) computed from pdf in a,b as the cumulative sum normalized to one for (c) “pIDC” and (d) correlation. Intersection of vertical and horizontal lines corresponds to the 95% confidence lower limit for the cdf. The color version of this figure is available only in the electronic edition.

large-scale application to China and Parkfield, California. The normalized PDF and CDF measures give more representative and intuitive results and are consistent with the findings of an order of magnitude improvement from the semi-empirical synthetic runs (Schaff, 2008) and the smaller case studies in Xiuyan, China, (Schaff, 2010) and at NORSAR (Gibbons and Ringdal, 2006, Gibbons *et al.*, 2007). The values for the normalized PDF and CDF measures in Table 2 are averages of those given in the text for different confidence limits.

For the case of Parkfield the normalized measures help to reduce the impact of false alarms. For China the presence of false alarms is not so much an issue because the false-alarm rate is low compared with the total number of detected events and relative to the background catalog. But for China the completeness of the catalog is an issue for lower magnitudes; the normalized measures of detection threshold appear less sensitive to that as well and more robust.

Large and Small Event Correlations

Since we are presumably going to take a larger event as a master template to try to detect a smaller event in the noise, it

is helpful to know what range of magnitude differences will still correlate to provide satisfactory detections. Where the corner frequency is higher than the filter band used theoretically this is not as much of an issue. When the corner frequency is within the passband, source finiteness is a factor that will change the shape of the waveforms and therefore the degree to which they correlate. Prior work has shown that events of similar magnitude tend to correlate well with each

Table 2
Detection Threshold Reduction of Correlation
versus “pIDC”

	Magnitude Reduction
China	90% confidence limit
	Normalized PDF
	Normalized CDF
Parkfield, California	90% confidence limit
	Normalized PDF
	Normalized CDF

other (small with other nearby small events, and large with other large events even if the source areas are not overlapping; see Schaff *et al.*, 2004).

Figure 9 plots the distribution of the magnitude differences for our analysis at Parkfield for the detected event pairs and the input observation matrix. The two distributions are virtually indistinguishable, and many pairs with separations larger than 1 magnitude unit are detected. This is a desirable feature since we hope to detect smaller events with larger events and also include detections for less than perfect waveform matches due to source complexities. For detection purposes it appears that larger magnitude separations are possible than for locations involving correlation measurements.

Figure 10 shows an example of more extreme magnitude differences from the 1999 Xiuyan earthquake sequence (Schaff and Richards, 2004; Schaff, 2010): a magnitude 5.5 event that correlates with a 3.2 event at the 0.5 level (a 2.3 unit difference). The travel time is reduced to the Lg -wave arrival. Note that the Lg wave is barely discernible to the eye for the magnitude 3.2 event. The first arriving P wave (at -80 s) is well below the noise. Figure 11 shows the waveforms zoomed into the Lg -wave window that was correlated and with normalized amplitudes. The averaged three-component cross-correlation trace is a well-defined spike (not shown) because the value is quite high (0.5) for these window lengths and frequency content. On the superimposed traces it can be seen that indeed several of the wiggles, peaks, and troughs do line up. Figure 12 shows another example from the same sequence of events at Xiuyuan where a magnitude 5.8 correlates with a magnitude 2.5 at the 0.26 level (an astonishing 3.3 unit difference or a factor of about 2000 in amplitude). Schaff (2008) reported that for similar data using the three-component averaging that a CC threshold of 0.27

had a high degree of statistical significance occurring with a false-alarm rate of zero in the 62 million samples or 36 days of continuous records considered. Still at these levels it is harder to see the similarities in the waveforms by eye (Fig. 12b), but the cross-correlation traces show a clear detection spike on all three components that align to the nearest sample, which is independent confirming evidence that the detection is real (Fig. 12c). In this case we know it is real because the events are known along with the correct times for their corresponding Lg -wave windows where the detection spike is observed. Gibbons *et al.* (2007) also report correlation detections are possible with as great as a three magnitude unit difference.

Buried Aftershocks

The ultimate goal of this work is to be able to detect new events that are not in existing catalogs. We found two examples of a small aftershock detected seconds after a mainshock in our 1999 Xiuyan, China, case study. Detection of early aftershocks was also found at Parkfield using matched filters (Lengline and Marsan, 2009; Peng and Zhao, 2009). These are new events discovered that are not in the ABCE because they are in the coda of the mainshocks. Figure 13 shows the first one. The waveforms are shown in Figure 13a,b,c. Figure 13d shows the CC traces for the three components overlaid. The main shock detection spike clearly comes in around 75 s. The aftershock detection spike around 120 s is also clearly seen on all three components aligning to the nearest sample (three independent tests confirming a true detection). The location of the beginning of the aftershock, denoted by the arrow, is shown to be buried in the coda of the waveforms of the mainshock in Figure 13a,b,c.

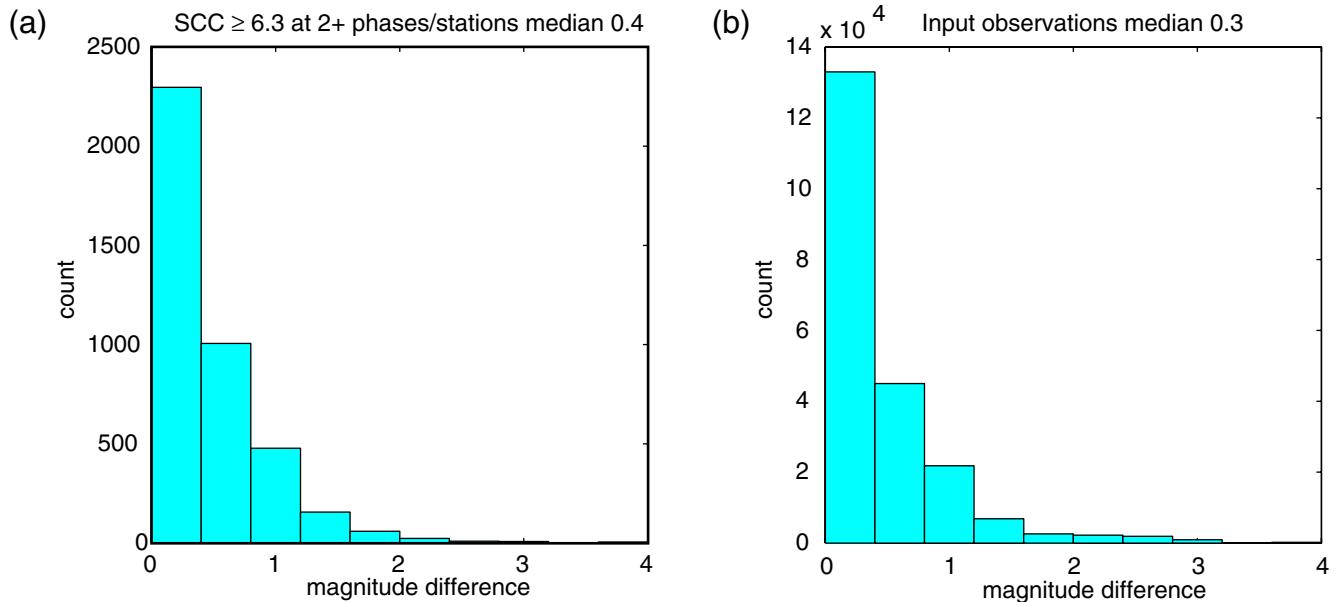


Figure 9. Distribution of magnitude differences for (a) correlation detected event pairs and (b) the input observations at Parkfield. The color version of this figure is available only in the electronic edition.

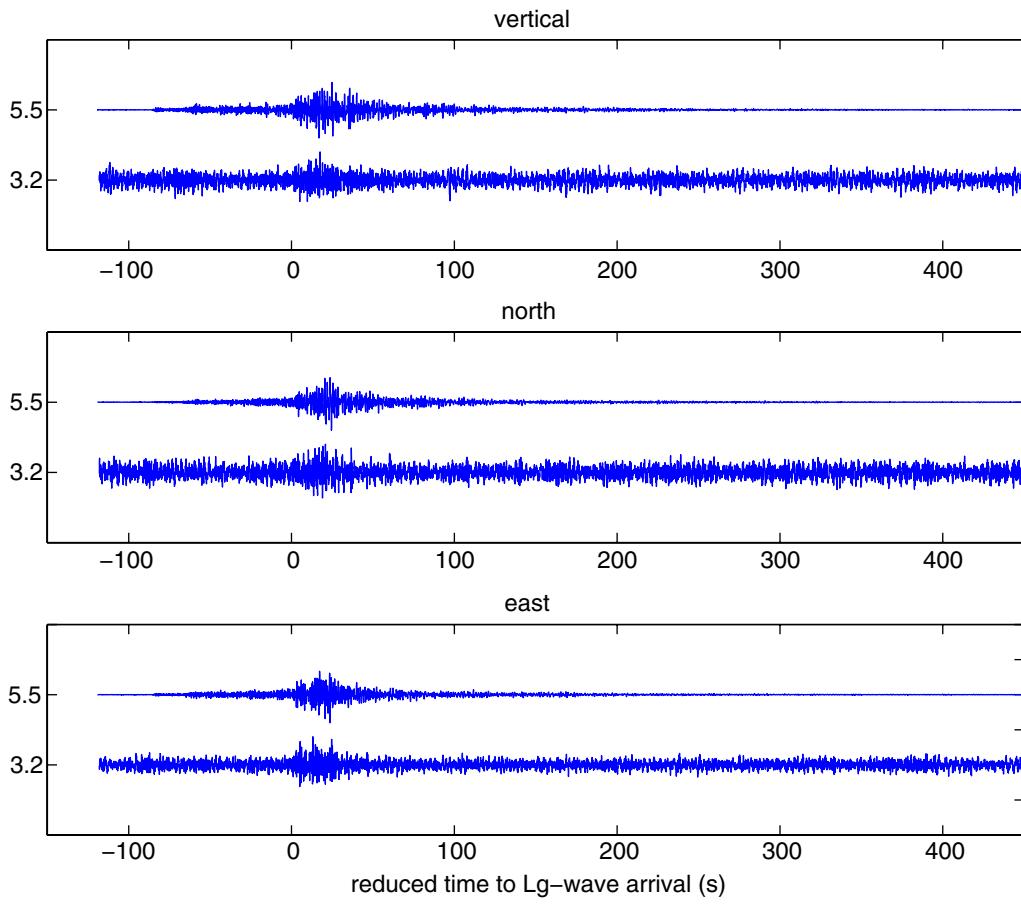


Figure 10. Waveforms for a magnitude 5.5 event and 3.2 in the Xiuyan, China, earthquake sequence that correlate with $CC = 0.5$. Lg wave at 0 s reduced travel time. The color version of this figure is available only in the electronic edition.

The master template is 50 s long, so that is why the correlation trace is offset by 25 s. The CC values are 0.3, 0.32, and 0.44 for the three components. Overlaying the waveforms at the correct time, it is seen that several of the peaks and troughs do line up by eye on the superimposed traces. The amplitude of the aftershock is about one-third that of the main shock, corresponding to an event about 0.44 magnitude units lower.

This highlights an additional application of a correlation detector pertaining to the issue of discrimination. On 16 August 1997 a small seismic event occurred near the former Soviet nuclear test site on the island of Novaya Zemlya. Because of its proximity to the test site, serious concerns were raised as to whether this event was a clandestine nuclear test in violation of the Comprehensive Nuclear Test Ban Treaty signed 11 months previously by Russia. By taking a data window around the mainshock and cross correlating forward in time, convincing evidence was found of an aftershock occurring four hours later (Richards and Kim, 1997; Gibbons and Ringdal, 2006). Some large nuclear tests have had aftershocks, but not ones with this small a magnitude. Also a second seismic event can follow closely in space and time to a nuclear test due to cavity collapse, but the waveform for such an event is very different. The fact that this small seismic

event had an aftershock with a similar waveform, taken alone, was very strong evidence that it was an earthquake.

Conclusions

The main finding of this research is that a correlation detector can lower thresholds of detection by an order of a magnitude over standard detectors on a large scale for the crustal earthquakes at regional distances in China and Parkfield, California. Importantly, this capability is achieved with acceptably low false-alarm rates as demonstrated by running the codes in an identical manner on the noise right before the expected arrival signals. Similar results of this order of magnitude improvement in detection thresholds have been demonstrated in smaller case studies using real seismic data (Gibbons and Ringdal, 2006; Gibbons *et al.*, 2007; Schaff, 2010) and semi-empirical synthetic tests (Schaff, 2008).

It is important to bear in mind that these results are for similar events. Detection capability will decrease as the underlying waveform similarity decreases due to increasing interevent separation distances, mechanisms differences, and source time function complexities. Previous work has shown, however, that even semi-similar events with less than perfect waveform matches still provide useful detections (Schaff,

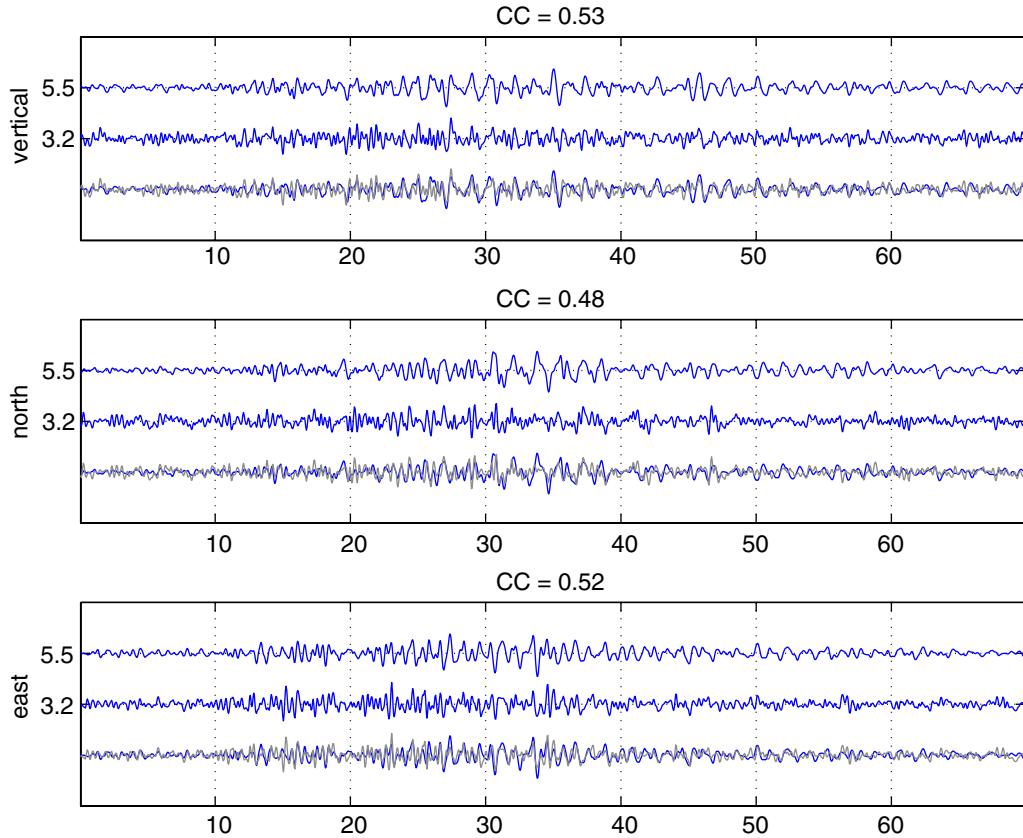


Figure 11. Zoom in on Lg waveforms for magnitude 5.5 and 3.2 events from Figure 10. Bottom trace in each panel shows the events superposed. x axis is in seconds for each panel. The color version of this figure is available only in the electronic edition.

2010). Also for the regions we have studied, correlation detects a high percentage of the seismicity that a standard detector finds: for China it is 85% (Schaff, 2009) and for Parkfield it is 86%.

For Parkfield, the increase in the number of events detected is approximately a factor of 10 just as Gutenberg–Richter predicts for a magnitude unit reduction. Normalized PDFs and CDFs are more robust measures of detection limits, less sensitive to false alarms, and more representative of the magnitude–frequency distribution than unnormalized confidence limits. High-resolution Parkfield locations show that the majority of the detections are for events with 1 km or less interevent separation distances, but this is magnitude and window-length dependent. Lg waves give most detections. For some stations this may be due to larger amplitudes and durations of energy. For the two closest stations the S - and Lg -windows start at nearly the same spot, so the parameters used for Lg may be the reason for the better detections: longer windows and wider frequency bands (higher time-bandwidth product).

Since we will often want to detect smaller events with a larger master waveform template, it is helpful to know what magnitude differences can be used before source finiteness degrades waveform similarity too much. Two examples were shown: one with a 2.3 magnitude unit difference and one with a 3.3 magnitude unit difference that correlate well enough for detection (correlation maximums are significantly above

background levels). This is consistent with the purely synthetic tests of Schaff (2010) that showed source time function durations of 15 s produced statistically significant detections with delta source time functions with zero duration. There is a need, however, to comprehensively study on a large scale the effect of source finiteness for less than perfect matches that still may be useable for correlation detection.

We have noted how a correlation detector can aid in discriminating explosions from earthquakes for the 1997 event near Novaya Zemlya by looking for aftershocks. We report on two examples of new aftershocks buried in the coda of mainshocks that can be detected and were not listed in the ABCE. Waveform cross correlation can also be used in discrimination when a signal correlates with a known explosion in the case of mining activity (Israelsson, 1990; Harris, 1991; Riviere-Barbier and Grant, 1993) and nuclear explosions (Shearer and Astiz, 1997; Thurber *et al.*, 2001; Fisk, 2002; Waldhauser, Schaff, *et al.*, 2004). The detection of buried aftershocks in the coda of mainshocks also has application to studies of anomalous early aftershock decay rate (Peng *et al.*, 2006).

Detection of seismic events in the presence of other seismic signals is a problem because the background noise is higher. This was a difficulty during the days following the great M 9.0 Sumatra–Andaman Islands earthquake on 26 December 2004. There were so many intersecting signals traversing the globe it was hard to make detections and

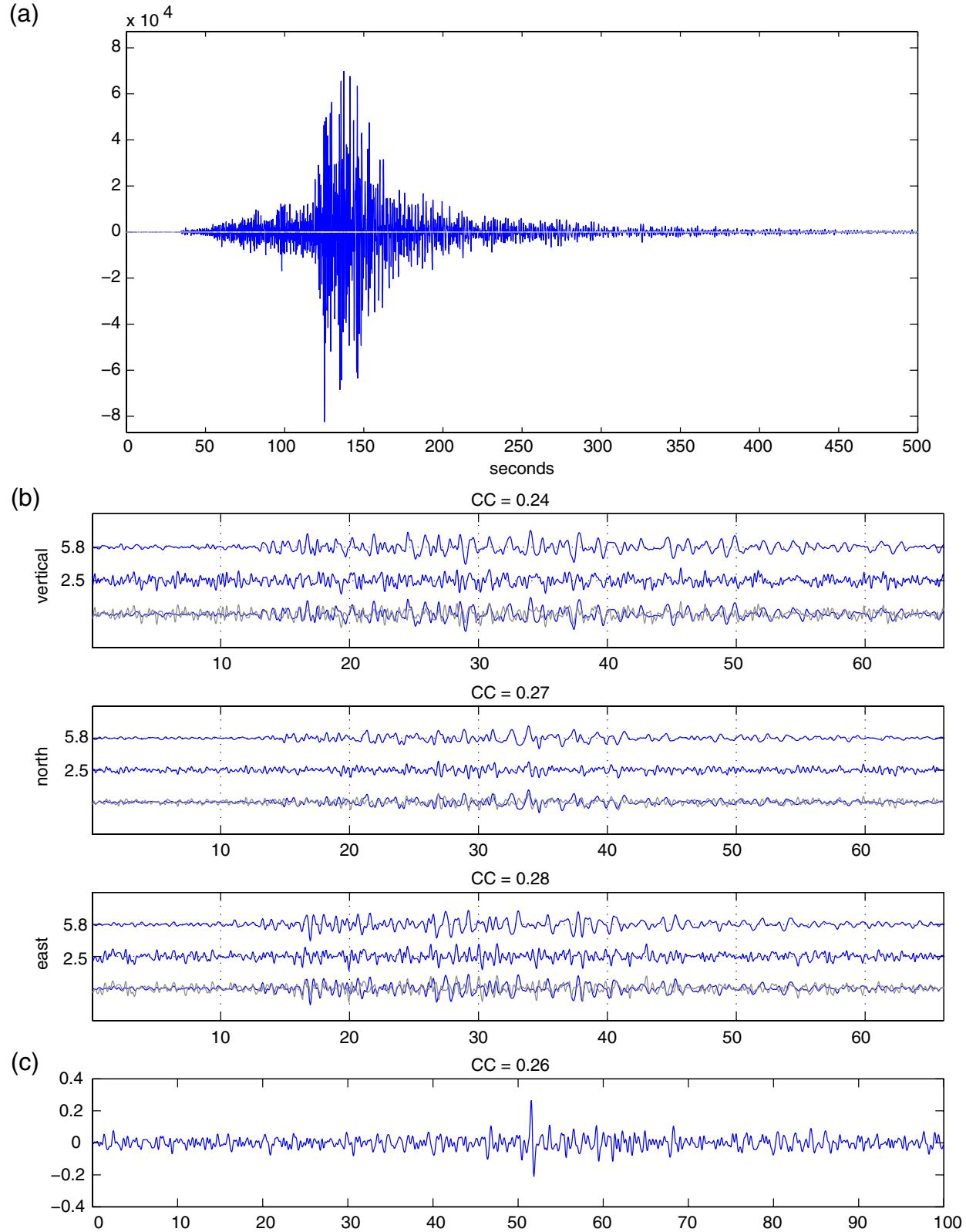


Figure 12. (a) Waveforms overlayed and to scale for a magnitude 5.8 (dark) and 2.5 (barely visible horizontal gray line at $y = 0$) that correlate with each other. (b) Lg waveforms on three components for $M 5.8$ and 2.5 events. Bottom trace in each panel shows the events superposed. x axis is in seconds. Correlation coefficient is in title of each panel. (c) Averaged correlation trace across the three components with a significant detection spike coming in after 50 s. x axis is in seconds. Correlation coefficient is in title of panel. The color version of this figure is available only in the electronic edition.

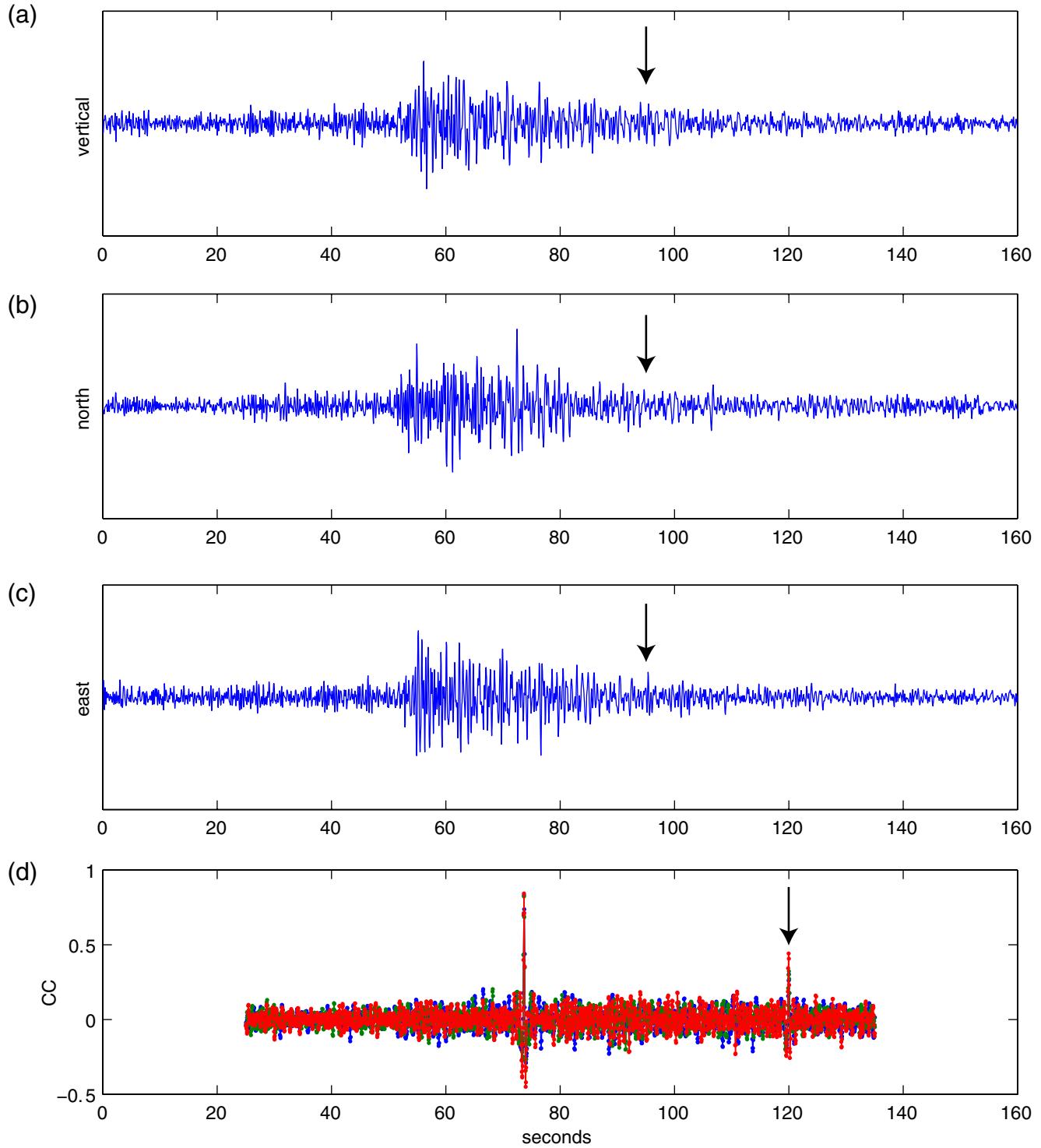


Figure 13. Waveforms for a detected mainshock for all three components and a new buried aftershock that was detected, onset shown by arrow. *x* axis is in seconds for each panel. Bottom panel shows that the correlation traces have detection spikes aligning to the nearest sample for both the mainshock and the buried aftershock. The color version of this figure is available only in the electronic edition.

associations. Data centers were swamped, such as the National Earthquake Information Center and the International Data Center in Vienna; in the latter case, the International Data Center even ceased bulletin production for several days. Correlation detectors may be able to extract more of the aftershocks in the days immediately following.

Data and Resources

Data are from the Berkeley Digital Seismic Network, Caltech Regional Seismic Network, GEOSCOPE, Leo Brady Network archived at IRIS and freely available at <http://www.iris.edu> (last accessed April 2008). The relocated earthquake

catalog can be obtained at <http://www.ldeo.columbia.edu/~felixw/DDcatalogs/index.html> (last accessed April 2008) and is described in Thurber *et al.* (2006). Code was written in MATLAB for the computations and graphics.

Acknowledgments

We are grateful to two anonymous reviewers and to Paul Richards for helpful comments. We thank colleagues in China who have given us basic information over the last decade on the effort to produce the Annual Bulletin of Chinese Earthquakes. This work is supported by Air Force Research Laboratory contract no. FA8718-05-C-0022 and U.S. Geological Survey award no. G09AP00067. This is Lamont-Doherty Earth Observatory Contribution no. 7381.

References

- Ekström, G. (2006). Global detection and location of seismic sources by using surface waves, *Bull. Seismol. Soc. Am.* **96**, 1201–1212, doi [10.1785/0120050175](https://doi.org/10.1785/0120050175).
- Fisk, M. D. (2002). Accurate locations of nuclear explosions at the Lop Nor test site using alignment of seismograms and IKONOS satellite imagery, *Bull. Seismol. Soc. Am.* **92**, 2911–2925.
- Freiberger, W. F. (1963). An approximate method in signal detection, *Q. Appl. Math.* **20**, 373–378.
- Gibbons, S. J., and F. Ringdal (2004). A waveform correlation procedure for detecting decoupled chemical explosions, *NORSAR Scientific Report: Semiannual Technical Summary No. 2–2004*, NORSAR, Kjeller, Norway, 41–50.
- Gibbons, S. J., and F. Ringdal (2005). The detection of rockbursts at the Barentsburg coal mine, Spitsbergen, using waveform correlation on SPITS array data, *NORSAR Scientific Report: Semiannual Technical Summary No. 1–2005*, NORSAR, Kjeller, Norway, 35–48.
- 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.
- Gibbons, S. J., M. B. Sørensen, D. B. Harris, and F. Ringdal (2007). The detection and location of low magnitude earthquakes in northern Norway using multi-channel waveform correlation at regional distances, *Phys. Earth Planet. In.* **160**, 285–309.
- Harris, D. (2006). Subspace Detectors: Theory, *Lawrence Livermore National Laboratory Internal Report UCRL-TR-222758*, 46 pp. (<http://www.llnl.gov/tid/lof/documents/pdf/335299.pdf>).
- Harris, D., and T. Paik (2006). Subspace detectors: Efficient implementation, *Lawrence Livermore National Laboratory Internal Report UCRL-TR-223177*, 36 pp. (<http://www.llnl.gov/tid/lof/documents/pdf/336400.pdf>).
- Harris, D. B. (1991). A waveform correlation method for identifying quarry explosions, *Bull. Seismol. Soc. Am.* **81**, 2395–2418.
- Israelsson, H. (1990). Correlation of waveforms from closely spaced regional events, *Bull. Seismol. Soc. Am.* **80**, no. 6, 2177–2193.
- Lengliné, O., and D. Marsan (2009). Inferring the coseismic and postseismic stress changes caused by the 2004 $M_w = 6$ Parkfield earthquake from variations of recurrence times of microearthquakes, *J. Geophys. Res.* **114**, B10303, doi [10.1029/2008JB006118](https://doi.org/10.1029/2008JB006118).
- Peng, Z., and P. Zhao (2009). Migration of early aftershocks following the 2004 Parkfield earthquake, *Nat. Geosci.* **2**, 877–881, doi [10.10138/Ngeo697](https://doi.org/10.10138/Ngeo697).
- Peng, Z., J. E. Vidale, and H. Houston (2006). Anomalous early aftershock decay rate of the 2004 $M_w = 6$ Parkfield, California, earthquake, *Geophys. Res. Lett.* **33**, L17307, doi [10.1029/2006GL206744](https://doi.org/10.1029/2006GL206744).
- Richards, P. G., and W.-Y. Kim (1997). Testing the nuclear test-ban treaty, *Nature* **389**, 781–782.
- Riviere-Barbier, F., and L. Grant (1993). Identification and location of closely spaced mining events, *Bull. Seismol. Soc. Am.* **83**, 1527–1546.
- Schaff, D. P. (2008). Semi-empirical statistics of correlation detector performance, *Bull. Seismol. Soc. Am.* **98**, 1495–1507, doi [10.1785/0120060263](https://doi.org/10.1785/0120060263).
- Schaff, D. P. (2009). Broad-scale applicability of correlation detectors to China seismicity, *Geophys. Res. Lett.* **36**, L11301, doi [10.1029/2009GL038179](https://doi.org/10.1029/2009GL038179).
- Schaff, D. P. (2010). Improvements to detection capability by cross correlating for similar events: A case study of the 1999 Xiuyan, China, sequence and synthetic sensitivity tests, *Geophys. J. Int.* **180**, 829–846, doi [10.1111/j.1365-246X.2009.04446.x](https://doi.org/10.1111/j.1365-246X.2009.04446.x).
- Schaff, D. P., and P. G. Richards (2004). Lg -wave cross correlation and double-difference location: Application to the 1999 Xiuyan, China, sequence, *Bull. Seismol. Soc. Am.* **94**, 867–879.
- Schaff, D. P., and F. Waldhauser (2005). Waveform cross-correlation-based differential travel-time measurements at the Northern California Seismic Network, *Bull. Seismol. Soc. Am.* **95**, 2446–2461, doi [10.1785/0120040221](https://doi.org/10.1785/0120040221).
- Schaff, D. P., G. H. R. Bokelmann, W. L. Ellsworth, E. Zanzerka, F. Waldhauser, and G. C. Beroza (2004). Optimizing correlation techniques for improved earthquake location, *Bull. Seismol. Soc. Am.* **94**, 705–721.
- Shearer, P. M. (1994). Global seismic event detection using a matched filter on long-period seismograms, *J. Geophys. Res.* **99**, 13,713–13,725.
- Shearer, P. M., and L. Astiz (1997). Locating nuclear explosions using waveform cross-correlation, in *Proc. of the 19th Annual Seismic Research Symposium on Monitoring CTBT*, 301–309.
- Shelly, D. R., G. C. Beroza, S. Ide, and S. Nakamura (2006). Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip, *Nature* **442**, 188–191.
- Thurber, C. H., C. Trabant, F. Haslinger, and R. Hartog (2001). Nuclear explosion locations at the Balapan, Kazakhstan, nuclear test site: The effects of high-precision arrival times and three-dimensional structure, *Phys. Earth Planet. Inter.* **123**, 283–301.
- Thurber, C., H. Zhang, F. Waldhauser, J. Hardebeck, A. Michael, and D. Eberhart-Phillips (2006). Three-dimensional compressional wave-speed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region, *Bull. Seismol. Soc. Am.* **96**, S38–S49.
- Van Trees, H. L. (1968). *Detection, Estimation and Modulation Theory*, John Wiley and Sons, New York.
- Waldhauser, F., D. Schaff, P. G. Richards, and W.-Y. Kim (2004). Lop Nor revisited: Underground nuclear explosion locations, 1976–1996, from double-difference analysis of regional and teleseismic data, *Bull. Seismol. Soc. Am.* **94**, no. 5, 1879–1889.
- Waldhauser, F., W. L. Ellsworth, D. P. Schaff, and A. Cole (2004). streaks, multiplets, and holes: High-resolution spatio-temporal behavior of Parkfield seismicity, *Geophys. Res. Lett.* **31**, L18608, doi [10.1029/2004GL020649](https://doi.org/10.1029/2004GL020649).
- Withers, M., R. Aster, and C. Young (1999). An automated local and regional seismic event detection and location system using waveform correlation, *Bull. Seismol. Soc. Am.* **89**, 657–669.
- Young, C., H. Harris, J. Beiriger, S. Moore, J. Trujillo, M. Withers, and R. Aster (1996). The waveform correlation event detection system project phase 1: Issues in prototype development and testing, Sandia Report, SAND96-1916, 36 pp.

Lamont-Doherty Earth Observatory of Columbia University
Palisades, New York 10964
dschaff@LDEO.columbia.edu