ADAPTIVE WAVEFORM CORRELATION DETECTORS FOR ARRAYS: ALGORITHMS FOR AUTONOMOUS CALIBRATION

Frode Ringdal[1], Steven J. Gibbons[1], and David B. Harris[2]

NORSAR[1] and Lawrence Livermore National Laboratory[2]

Sponsored by National Nuclear Security Administration
Office of Nonproliferation Research and Engineering
Office of Defense Nuclear Nonproliferation

Contract Nos. DE-FC52-05NA26604[1] and W-7405-ENG-48[2]

ABSTRACT

Correlation detection is a relatively new approach in seismology that offers significant advantages in increased sensitivity and event screening over standard energy detection algorithms. The basic concept is that a representative event waveform is used as a template (i.e., matched filter) that is correlated against a continuous, possibly multichannel, data stream to detect new occurrences of that same signal. These algorithms, therefore, are effective at detecting repeating events, such as explosions and aftershocks at a specific location.

We present evidence that correlators can be used to lower detection thresholds by 0.5 to 1 magnitude units over standard energy detection algorithms. Thus, they offer roughly the same enhancement in detection performance that is achieved by replacing single sensors with arrays. This enhancement can be achieved *in addition to* the sensitivity increase afforded by arrays.

The fact that correlators are specific to individual sources makes them attractive as event classifiers. This feature can be exploited to construct efficient screens for mining explosions or earthquake aftershocks. This major research topic is addressed in this project.

In principle, the deployment of correlation detectors against seismically active regions could involve very large numbers of very specific detectors. To meet this challenge, we are examining two strategies:

- use of *subspace detectors*, an extension of correlators, which allows representation and detection of signals exhibiting some range of variation, and
- autonomous calibration of many (subspace and correlation) detectors in a flexible, adaptive detection framework, subject to analyst review.

Because array-based correlation detectors are new to seismology, a significant amount of research on how to tune these detectors is needed to inform later calibration efforts that will arise if they are adopted for operational use.

We have begun to explore these issues through a series of representative case studies drawn from important monitoring problems involving detection of weak sources and effective screening of mining explosions and earthquake swarms and aftershocks. We focus upon two geographical regions for these case studies: (a) the European Arctic and (b) Central Asia, using available arrays and seismic networks in and near these regions. Examples of such case studies are presented.

OBJECTIVE

The overall objective of this proposed three-year study is to develop and test a new advanced, automatic approach to seismic detection using waveform correlation, with special application to seismic arrays. The principal goal is to develop an adaptive processing algorithm. By this we mean that the detector is initiated using a basic set of reference ("master") events to be used in the correlation process, and then an automatic algorithm is applied successively to provide improved performance by extending the set of master events selectively and strategically. These additional master events are generated by an independent, conventional detection system. A periodic analyst review will then be applied to verify the performance and, if necessary, adjust and consolidate the master event set.

RESEARCH ACCOMPLISHED

Introduction

Correlation detection is a relatively new approach in seismology that offers significant advantages in increased sensitivity and event screening over standard energy detection algorithms. The basic concept is that a representative event waveform is used as a template (i.e., matched filter) that is correlated against a continuous, possibly multichannel, data stream to detect new occurrences of that same signal. These algorithms, therefore, are effective at detecting repeating events, such as explosions and aftershocks at a specific location. Matched filters have been used for decades in radar, sonar, and communications processing to detect occurrences of weak signals in background noise (van Trees, 1968). The principal unique features in the seismological application are the following:

- that the signal to be detected is obtained empirically, and is not under the control of the system designer,
- for arrays, the template is multichannel (i.e., the correlator cannot be factored into separate spatial and temporal components) as a consequence of the strong scattering and multipathing experienced by seismic waves, and
- the template is strongly a function of the specific source location, mechanism, and excitation time history.

Correlators use the fine spatial and temporal structure of the signal to produce exquisitely sensitive detection algorithms. We present evidence that correlators can be used to lower detection thresholds by 0.5 to 1 magnitude units over standard energy detection algorithms. Thus, they offer roughly the same enhancement in detection performance that is achieved by replacing single sensors with arrays (Figure 1). This enhancement can be achieved *in addition to* the sensitivity increase afforded by arrays. This remarkable additional capability is a consequence of the fact that array correlators add the exploitation of temporal signal structure to the exploitation of spatial structure currently afforded by array beamforming. Array correlators also provide a means of compensating signal decorrelation across an array (or indeed network) aperture, allowing a type of high-frequency, spatially coherent beamforming across a very large aperture. The sensitivity of correlators is a consequence of their very specific waveform templates, a fact which also imposes certain limitations:

- confinement to very small geographical footprints (typically 1-2 wavelengths at the dominant signal frequency)
- confinement to events that exhibit very little variation in source characteristics, such as the source mechanism and time history.

Subspace detectors generalize correlators by continuously matching the (multichannel) data stream against an optimal linear combination of a collection of waveform templates. The signal subspace basis (i.e. waveform templates) can be chosen to represent the variability observed in a particular source as a consequence of spatial, mechanism, or temporal variation of the source.

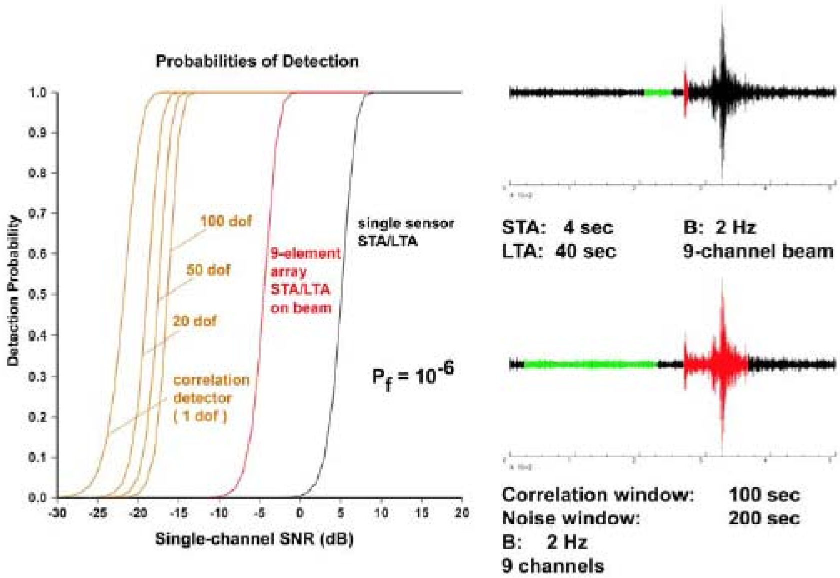

Figure 1. Comparison of theoretical detection performances of simple energy detectors and the class of correlation detectors (including subspace detectors). The suite of curves to the left show probabilities of detection at a fixed (low) false alarm probability of 10^{-6} and at varying signal-to-noise ratios for a variety of detectors. To the far right of this suite of curves is the detection probability curve for a single sensor using a short-term average/long-term average (STA/LTA) algorithm and a characteristically small window duration (4 sec) usually chosen to detect individual phases; note that it has a threshold above SNR=1 (0 dB). The next curve to the left is the detection probability for an STA/LTA detector operating on an 9-sensor array beam, assuming perfect signal coherence. To the far left is the detection probability for a correlator operating on a 9-sensor array and a 100-second detection window (correlation detection allows the coherent collection of signal energy over large windows). Intermediate are a suite of curves for subspace detectors of varying dimension degrees of freedom (dof) that could be chosen to represent a degree of signal variability. Note that all correlators exhibit an increase in performance comparable to or exceeding that afforded by replacing a single sensor with an array.

Since correlation detectors are relatively new to seismology, significant work is required as to how to tune these detectors in order to inform subsequent calibration efforts that will arise as and when they are adopted for operational use. Issues that will arise include the following:

- How many detectors are required to cover a particular source region?
- What distribution of events (number of events, magnitude range, range of source locations) is required to calibrate a detector?
- How are tuning parameters (such as signal duration, bandwidth, number of channels) optimized in order to maximize the probability of detection at a fixed false alarm rate?
- What is the optimal dimension of a subspace detector given the trade-off between false-alarm rates, detection probability, enhanced geographical coverage, and insensitivity to variation in source mechanism and time history?
- Is it possible to devise an autonomously operating framework which can continuously spawn and update correlation detectors (under analyst review) as new sources become active?
- Is it more efficient to deploy a large number of correlators or a small number of subspace detectors to cover a given geographical region?

All of these issues are to be addressed through a series of representative case studies drawn from important monitoring problems involving the detection of weak sources and the effective screening of mining explosions and earthquake swarms and aftershocks. Examples of such case studies are presented in this paper.

Detecting small chemical explosions

NORSAR has recently applied array-based waveform cross-correlation techniques to detect small chemical explosions in Sweden, with considerable success (Gibbons and Ringdal, 2004, Stevens et al., 2004). This work was initiated under a Defense Threat Reduction Agency (DTRA)-funded joint project with Science Applications International Corporation (SAIC). Between 1986 and 1989, a total of 11 chemical explosions were carried out in two underground chambers at a site in Älvdalen, central Sweden. Explosions with yields 10, 1000, and 5000 kg were performed in each of the two chambers, one with size 300 m^3 and one with size 200 m^3. The signals from these events were sought in order to provide a comparison with a later series of explosions in a different chamber at the same site; it was, however, not known at which times these events had taken place.

To detect the signals from the Älvdalen explosions for which the origin time was unknown, we developed a procedure using array correlation which proved highly successful. The basic procedure was as follows:

1. select a reference event among the known explosions and a seismic array for which data exists both for the reference event and the time period in which the sought-after event was known to have occurred
2. select a frequency band with an optimum SNR
3. for each individual channel, correlate the filtered signal from the reference event against the data stream containing the sought-after signal
4. sum (i.e., beamform) the single channel correlation traces
5. apply a standard power detector to the beam

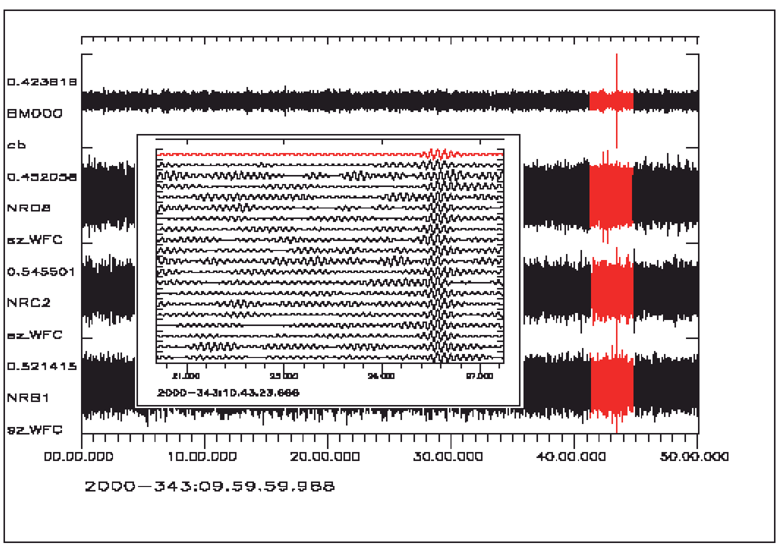

Figure 2. Example of detecting a small chemical explosion by cross-correlation. The source was a detonation of 500 kg TNT within an underground chamber with volume 1000 m^3 at a distance of 142 km from the NORES array. The master event was a 10000 kg detonation in the same chamber 5 days later. The waveforms were filtered within a narrow frequency band (14.0 - 18.0 Hz), which was necessitated by the high-frequency nature of these signals. Note the high SNR gain on the top trace, which is the beam of the individual correlation traces.

All of the explosions of 1000 kg and more were detected by this procedure using NORES and NORSAR array data. In addition, we succeeded in detecting an explosion with a yield of only 500 kg TNT which we knew had taken place but had been entirely unable to detect the signals using conventional processing. A significant aspect of the correlation detector is the low false alarm rate. To test the reliability of the procedure, the detector was run on NORSAR array data over selected time-segments totalling over one year; the process triggered 5 times only, each time corresponding to a confirmed event at Älvdalen. No event at this site of which the authors are aware has failed to be detected by this procedure, and on no occasion did the test produce a detection without there having been a confirmed explosion at the site.

Detecting rockbursts in the Barentsburg mine, Spitsbergen

Detecting sequences of earthquakes and rockbursts is an obvious application of the waveform correlation technique. We present an illustrative example, which clearly demonstrates the power of the approach. This example is from the Barentsburg mine in Spitsbergen, using data from the Spitsbergen array at approximately 50 km from the mine (see Figure 3). Since Kola Regional Seismological Center operates two in-mine seismic stations in Barentsburg, we have been able to verify that for selected time periods during which we had access to the in-mine data, the events in this swarm (see Figure 4) are in fact real and located in or near the mine. This reconfirms the ability of the array correlation detector to operate at a high sensitivity combined with a low false alarm rate.

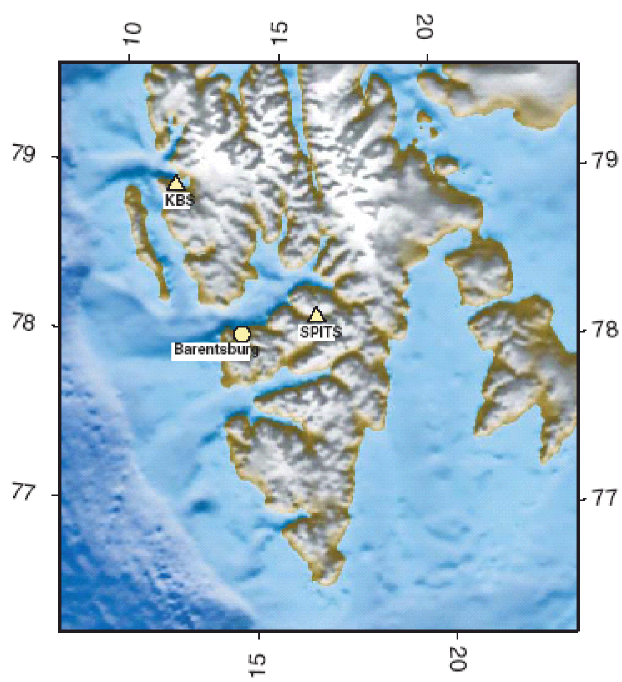

Figure 3. Locations of the Spitsbergen seismic array and the Barentsburg coal mine on Svalbard. The distance between the array and the mine is approximately 50 km.

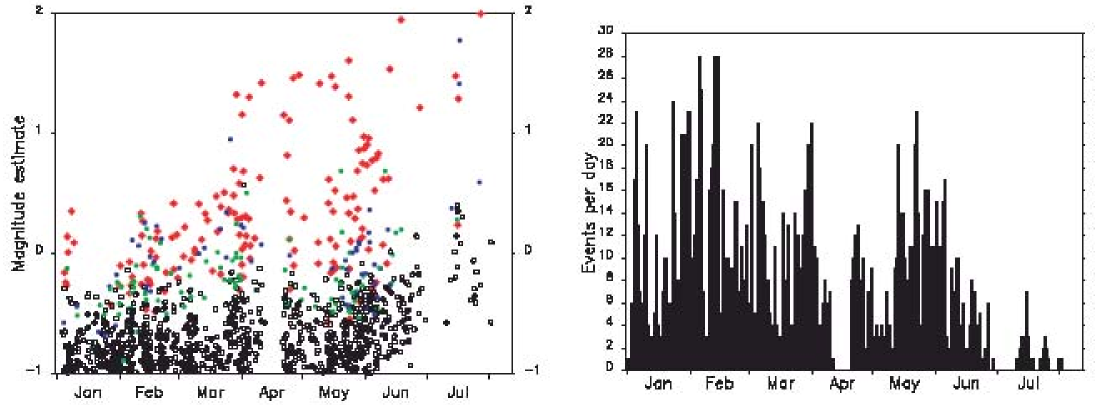

Figure 4. Sequence of waveform correlation detections of events assumed to come from the Barentsburg coal m ine on Svalbard. A fatal rockburst occured on July 27, 2004, and data from the Spitsbergen array (at a distance of approximately 50 km) was searched for similar waveforms using the signal from this m_b = *2.0* event (filtered between 3.0 and 6.0 Hz) as the master waveform. A total of 7378 detections were made on the correlation beam between January 1, 2004, and August 12, 2004, the majority of which were discarded automatically based on the results of f-k analysis (Gibbons and Ringdal, 2005). The plots illustrate the 1578 events which, following manual examination, were deemed likely to come from the mine. The histogram (right) illustrates the frequency with which these events occured, which is known to correspond well with activity at the mine. The other panel (left) displays the magnitude estimated from the waveform correlation for these events, scaled to the master event which has been fixed at 2.0. Black circles indicate correlation detections which were not detected using traditional processing on the SPI array. Red symbols indicate correlation detections for which both a P and S detection with a suitable slowness estimate and arrival time were measured. Blue symbols indicate that only a suitable P detection was made, and green symbols indicate that only an S detection was made.

__Array processing__

In the Älvdalen study, we found that beamforming on correlation traces is particularly effective in suppressing noise while preserving the "signal" (i.e., the correlation peak). While standard array beamforming requires coherent signals and incoherent noise, the "correlation beamforming" requires only that the noise be incoherent; the requirement for spatially coherent signals is replaced by a requirement of similarity between the detected and master waveform. Thus, even across the NORSAR array, the correlation traces line up perfectly with no loss in the beamforming. This applies regardless of frequency. For example, in detecting the Älvdalen explosions, we applied a filter of 14-18 Hz, as illustrated in the preceding section (Figure 2).

When applying waveform correlation over a small aperture array, correlation detections can occur when two unrelated incoming wavefronts exhibit a coincidental degree of similarity over a short time-window. It can be demonstrated that the correlation traces emulate a plane wavefront traversing the array with a slowness approximately equal to the difference between the slownesses of the two correlating data segments. It is therefore possible to perform f-k analysis on the correlation traces; a clearly non-zero slowness at the time of the maximum correlation almost guarantees the occurrence of a false alarm since the correlating wavefronts come from different directions. In our processing of the Barentsburg sequence, we have used the f-k results to effectively screen out such false alarms.

__Subspace detectors__

Waveform correlators, including the multichannel correlator just described frequently provide exquisitely sensitive detectors of repetitive signals from geographically compact sources. However, it is often (even usually) the case that repeating sources exhibit a degree of variability in the signals that they generate. Variability can be caused by the events from the source being distributed over a region of small, but nonetheless nonvanishing extent (such as the aperture of an aftershock sequence). Signal diversity also can be caused by variations in the source time history (as is frequently the case with ripple-fired mining explosions) or the source mechanism. It is desirable to develop detectors that are tolerant of signal diversity while still achieving much of the sensitivity of correlation detectors.

Subspace detectors (Scharf and Friedlander, 1994) are a generalization of waveform correlators that provide a flexible mechanism for trading diversity in signal representation for detector sensitivity. Subspace detectors add an uncertain signal model to the usual formulation of the detection problem. In detection problems, a window is scanned continuously along a data stream, and, for each window position, a test of two alternative hypotheses is carried out: H_0, that the window contains noise only, and H_1, that the window contains signal plus noise. In the subspace detector formulation, the signal being sought is assumed to be a linear combination of orthonormal basis functions. The basis functions are known, but the coefficients are not and must be estimated continuously as the window is scanned along the data stream. The coefficients are chosen to maximize the correlation between the linear combination of basis functions and the data in the detection window. This step-wise optimization feature provides subspace detectors with the flexibility required to deal with source variability.

The problem of subspace detector design is to find a low-dimension collection of waveform basis functions that spans the range of signals anticipated for a source. An empirical procedure for accomplishing this objective is outlined in Figure 5. In this approach to subspace design, a collection of waveforms from events defining the expected variation of the source is assembled, the waveforms are aligned as columns in a data matrix, and the waveform basis is constructed from a singular value decomposition (SVD) of the data matrix. The dimension of the waveform basis can be determined by counting the significant singular values in the SVD. Alternatively, it is possible to calculate the probabilities of detection and false alarm under an assumption of uncorrelated, normally-distributed background noise. Then the optimum dimension can be found to maximize the probability of detection for a fixed false alarm rate (Neyman-Pearson criterion). In Figure 5, the events defining the source are assembled by clustering a large collection of events detected with an STA/LTA algorithm on the basis of waveform correlation. The events from a single cluster are assumed to arise from a single source

This procedure has been tested numerous times on mining explosions, earthquake swarms, and aftershocks. The following example, drawn from detection of the San Ramon, California, swarm of 2002, indicates that subspace detectors can provide an increase in performance over waveform correlators when properly designed. Figure 6 shows the location of the swarm and two stations used in the example. Event waveform data from the broadband station KCC, 240 kilometers from the swarm, was used to construct a correlator and several subspace detectors to compare detection performance. A cluster of 19 events detected with an STA/LTA algorithm was used to design subspace detectors; the waveform of the largest event (an ML 3.9 event) was used in a correlator. The subspace detectors and the correlator were multichannel algorithms, using all three channels of the three-component broadband station.

The distributions by magnitude of the detections made by the correlator and a 9-dimension subspace detector are shown in Figure 7. The Berkeley Seismological Laboratory (BSL) catalog was used for ground truth on the swarm events. The histograms show that the correlator missed some of the high magnitude events, but the subspace detector captured all events down to duration magnitude (Md) 1.5.

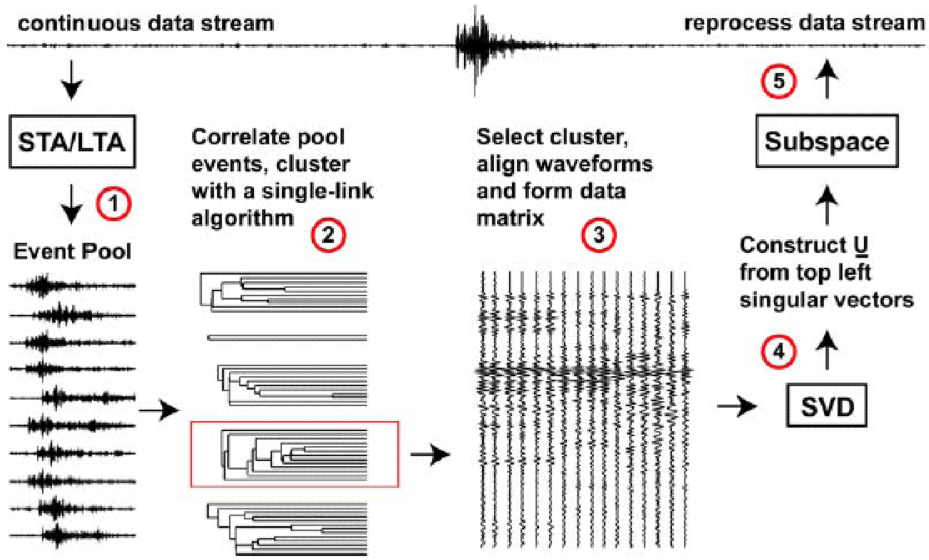

Figure 5. A “bootstrap” procedure for designing and applying subspace detectors empirically to reduce the detection threshold for a specific source with five steps.

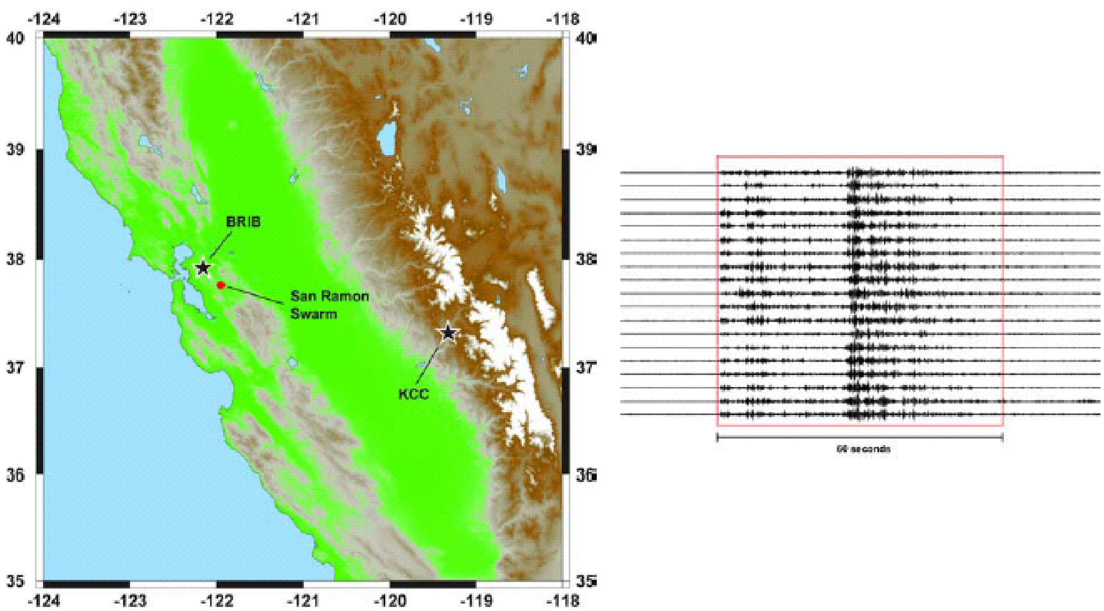

Figure 6. The 2002 San Ramon, California, swarm provided a convenient test of subspace detector design and performance. Data from two Berkeley Digital Seismic Network broadband stations (KCC, BRIB) were used in the test example to implement the detector and help assess its performance. At right are the waveforms of the 19 events used to develop subspace detectors and a correlator. A degree of signal variation is evident in the waveforms.

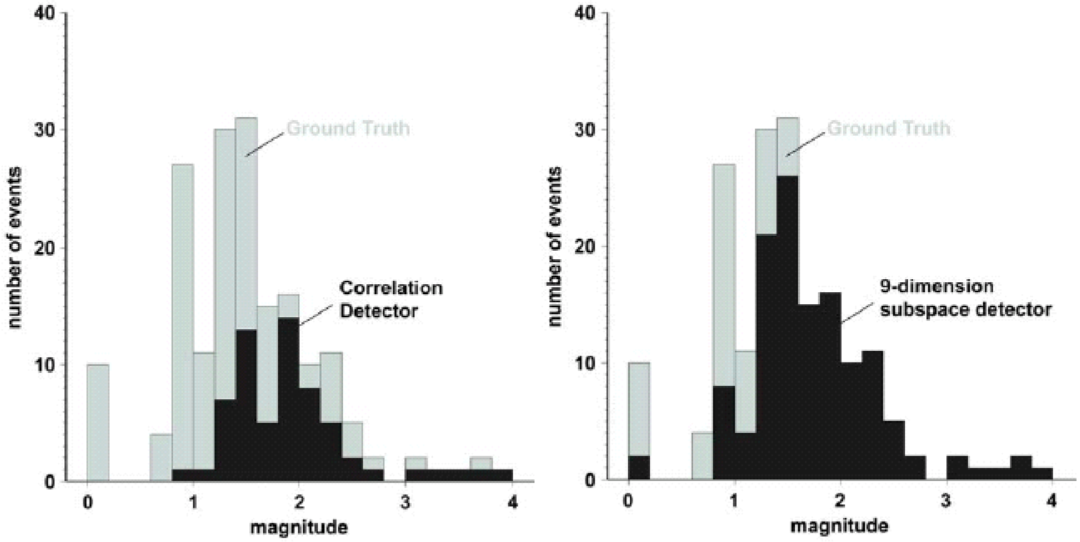

Figure 7. A 9-dimension subspace detector detected twice as many San Ramon swarm events as the correlation detector when both were operated at the same theoretical false alarm rate. The histograms above com pare the distribution of detections by magnitude to the ground truth event distribution obtained from the BSL catalog. The subspace detector has a threshold around Md 1.5 at 240 kilometers range.

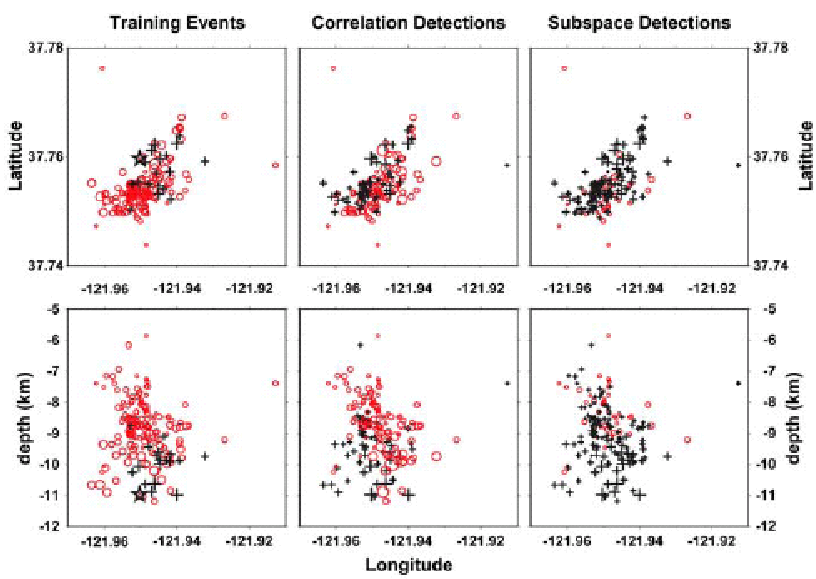

Figure 8. Subspace algorithm detections have broader spatial distribution than the correlator detections, but both miss a large fraction of the shallower events in the sequence. The maps above show the distribution of San Ramon sequence events in latitude and longitude (top maps) and in longitude and depth (bottom maps). The maps at left show the distribution of master events used in the design of the correlator (star) and subspace detectors (star and crosses).

Plots of the spatial distribution of the detected events are shown in Figure 8. These show that the subspace detectors are sensitive to events over a wider geographic footprint than the correlator based upon a single event waveform. The reason for this effect is clear from the spatial distribution of the events used to design the subspace detector (left column of Figure 8): these form a more complete sampling of the geographic distribution of events in the swarm.

One of the interesting questions to be addressed in this proposed project is whether subspace detectors provide more efficient detection and lower detection thresholds than multibeam correlators (or vice versa). These two approaches essentially provide different representations for the source: the first by constructing a minimal waveform basis for the signals to be detected and the second by using a bundle of individual waveform correlators to represent a diverse source.

CONCLUSIONS AND RECOMMENDATIONS

The project is still in an initial stage. We expect that the work will result in the development of a new advanced, automatic approach to detection/location using array processing in combination with waveform correlation. Based on the promising preliminary results described in this paper, we expect that the application of the method will result in significantly improved detection of low-magnitude events, combined with a low false alarm rate at sites for which adequate calibration information is available.

While it is not our purpose to develop an operational system at this point, the project is intended to provide automatic algorithms that will facilitate future implementation in an operational environment.The overall purpose is to develop practical methods that can be implemented into operational monitoring, and that will be useful for improving the quality of automatic bulletins and for reducing the routine workload of the analysts. Among the products of our research will be examples of clusters of events that are guaranteed to produce effective correlation detectors, including "correlation templates" in selected cases.

The development of correlation detectors is challenging because of the large number of sources that surround many stations, requiring distinct dedicated detectors. The number of detectors required is compounded by the fact that the detailed structure of signals may change over time for many of these sources, requiring detector updates. We plan to address these problems by automating the correlation detector development to the extent possible, under analyst review. The framework that we propose to develop is intended to allow exploration of a number of different strategies for autonomous detector development.

REFERENCES

Gibbons, S. J., and F. Ringdal (2004), A waveform correlation procedure for detecting decoupled chemical explosions, *Semiann. Tech. Summary,* NORSAR Sci. Rep. 2-2004, Kjeller, Norway, pp 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, *Semiann. Tech. Summary,* NORSAR Sci. Rep. 1–2005, Kjeller, Norway, pp 35-48.

Scharf, L. L. and B. Friedlander (1994), Matched subspace detection, *IEEE Transactions on Signal Processing*, 42 (8): 2146–2157.

Stevens, J., Rimer, N., Xu, H., Murphy, J., Barker, B., Gibbons, S., Lindholm, C., Ringdal, F., Kværna, T., and I. Kitov (2004), Analysis and simulation of cavity-decoupled chemical explosions, in *Proceedings of the 26th Seismic Research Review, Trends in Nuclear Explosion Monitoring*, LA-UR-04-5801, Vol. 1: pp. 495–502.

van Trees, H. L. (1968), Detection, Estimation, and Modulation Theory. New York: John Wiley & Sons, Inc.