

**ADAPTIVE WAVEFORM CORRELATION DETECTORS FOR ARRAYS: ALGORITHMS FOR  
AUTONOMOUS CALIBRATION**

Frode Ringdal<sup>1</sup>, Steven J. Gibbons<sup>1</sup>, and David B. Harris<sup>2</sup>

NORSAR<sup>1</sup> and Lawrence Livermore National Laboratory<sup>2</sup>

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

Contract Nos. DE-FC52-05NA26604<sup>1</sup> and W-7405-ENG-48<sup>2</sup>

**ABSTRACT**

It can be demonstrated that waveform correlation detectors can detect seismic events up to an order of magnitude smaller than traditional energy detectors. Such detectors work by measuring the degree of similarity between the incoming data stream and a template of representative waveforms from an event at a site of interest. Whereas STA/LTA detectors require a significant arrival of energy over a relatively short time interval, correlation detectors utilize the full available waveform and can exploit the characteristic coda which is typically unique to a given source region and often contains the largest amplitudes in the wavetrain. Correlation detectors over arrays are particularly effective since correlation coefficient traces are coherent over arrays or networks of arbitrary dimensions, even when the waveforms at the different sites are incoherent. The greatest disadvantage of such detectors is the requirement of a template waveform. There are many unresolved issues regarding the use of such detectors in operational seismic monitoring. For instance, for a given master event, how far away can a subsequent seismic event be, and how different can the source mechanism and magnitude be, such that it is still detected by the same waveform template.

We have examined an interesting sequence of earthquakes in the north of Norway which has allowed many of these issues to be addressed simultaneously. The events are approximately 600 km from the (large-aperture) NORSAR and (small-aperture) ARCES array stations. Local stations within 20 km of the epicenters have provided confirmation and stable magnitude estimates for all events. Using a waveform template extracted from an  $m_b=3.5$  event, the NORSAR and ARCES arrays were both able to detect events down to  $m_b=0.5$ , including many aftershocks concealed within the coda of larger events. This is a clear order of magnitude detection threshold improvement for ARCES and an even greater effective improvement for the NORSAR array over which these high frequency regional signals are incoherent. The false alarm rate at the small-aperture arrays is far higher than for the NORSAR array since the high similarity of the waveforms across small arrays, necessary for traditional array processing, appears to degrade the performance of the array-based correlation detectors. False alarms are usually identified by measuring misalignment of the correlation coefficient traces, although this becomes more challenging for weaker events. We present an example whereby the alignment of correlation maxima allows us to detect and measure a case of erroneous instrumental timing.

Subspace detectors generalize the matched filter concept to detect signals from a particular source that exhibit significant variation, expressed as a signal subspace of dimension greater than one. We have developed interactive software that assists the development of subspace detectors from a collection of master events in the source region of interest by allowing the user to select the dimension of the detector to maximize the probability of detection at a fixed false alarm rate. This new software has been tested on especially well-constrained event clusters in eastern California observed by the NVAR array. Double-difference locations from the Hauksson SCSN catalog provide ground truth with a number of stations in close proximity to the sequences of interest. The results from this study are mixed. Signals from one sequence, near the Long Valley Caldera, are well-characterized by a subspace of dimension one; an array correlation detector performs well in this case, as in the northern Norway example. Two other sequences show large signal diversity for events within fairly compact source regions ( $\sim 3-4$  km). In these cases subspace detectors perform better than correlation detectors, but still do not capture the majority of events. The issue appears to be smaller events with source mechanisms substantially different from those of larger master events. We are developing a number of strategies for dealing with this problem, including an iterative approach for expanding the subspace as event magnitude decreases.

### **OBJECTIVE**

The overall objective of this 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.

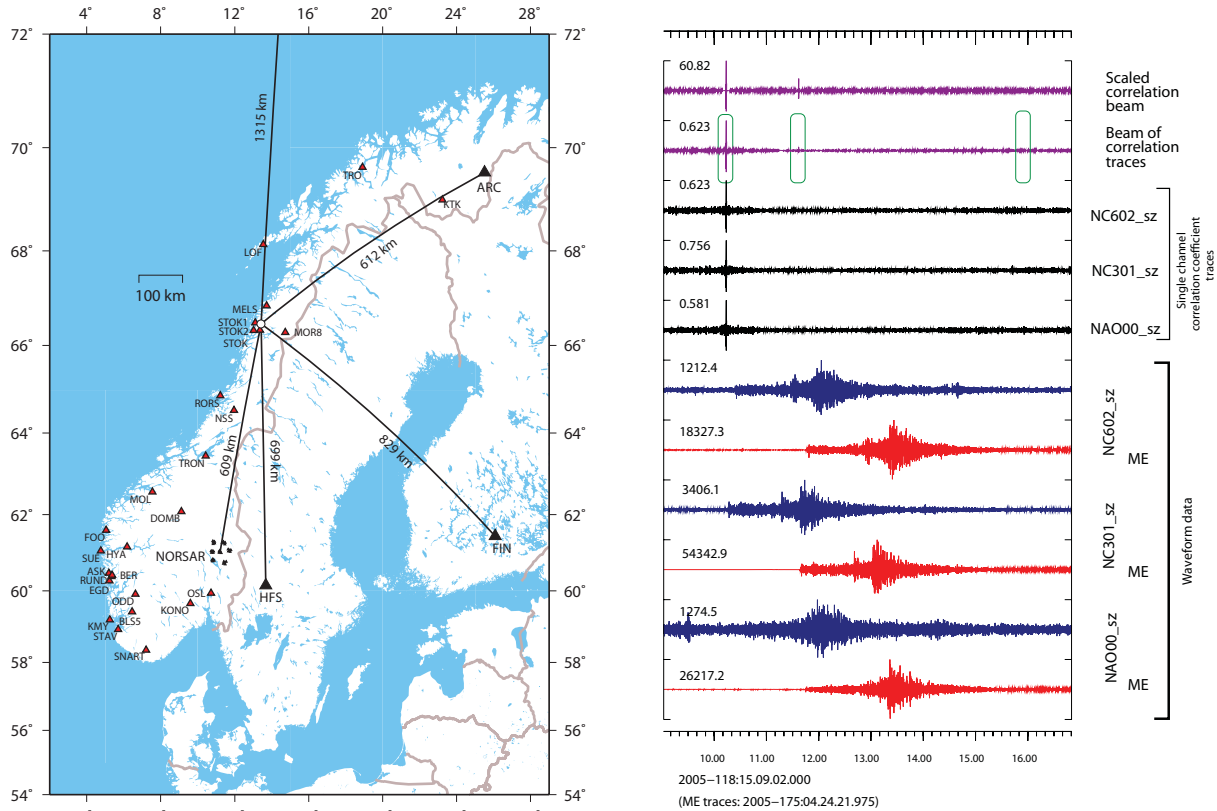
### **RESEARCH ACCOMPLISHED**

In the first year of this project, we have focused on three different aspects of waveform correlation detectors. Firstly, we have addressed how great an improvement in the detection threshold for low-magnitude events is possible under various scenarios for repeating events in a small geographical region. Secondly, we have examined how the use of correlation detectors over arrays and networks can be exploited to expose and measure problems with instrumental timing. Thirdly, we have investigated how the use of higher dimension subspace detectors allows an improvement in event detectability even when significant waveform variation is observed between different events from a highly confined region. These issues are discussed in the following sections.

#### **The Detection of Low-Magnitude Seismic Events Using Array-Based Waveform Correlation: A Case Study in Northern Norway**

The Rana region of Norway is the site of constant intraplate seismicity (Hicks et al. 2000) and, on June 24 2005, was the site of an  $m_b = 3.5$  event which was well recorded at all the Fennoscandian International Monitoring System (IMS) array stations in addition to the National Seismic Networks of Norway and Finland (Figure 1). An  $m_b = 2.4$  event on April 28, 2005, had been located to almost the same location and subsequent analysis of the waveforms recorded at regional distances indicated very high correlation coefficients between the two events at frequencies up to 10 Hz. The high waveform similarity indicates that the spatial separation between the events is very small (see Geller and Mueller, 1980) and waveform templates were extracted from the June 24 event to attempt to detect occurrences of smaller events from this location which were not detected using traditional array processing. We focused upon the NORSAR array at a distance of approximately 600 km from the earthquake epicenters. The large inter-site distances on this teleseismic array preclude the effective processing of high-frequency regional phases using traditional array methods for small-aperture arrays due to the lack of waveform similarity between sites. However, as was demonstrated by Gibbons and Ringdal (2006), the correlation coefficient channels are coherent over arbitrarily spaced networks and arrays even when the waveforms are not (the condition of waveform similarity between sites is replaced by a condition of waveform similarity between events). The waveform templates for the initial investigation were bandpass filtered between 2.0 and 8.0 Hz, and a 120.0 second long segment was cut for each of the 42 short-period vertical channels. The array correlation beam rarely exceeded a value of 0.005, and a detection threshold of 0.03 was exceeded a total number of 32 times throughout the calendar year 2005. On only three occasions did these detections correspond to events from the Rana region which were detected by the IMS arrays: the April 28 event, the June 24 event (a trivial self-correlation detection), and a third event on December 15 with an estimated magnitude 3.0.

Due to the high seismicity in this region, several 3-component stations of the Norwegian National Seismic Network (NNSN) are placed close to the site of the Rana earthquakes (Figure 1); the permanent stations at Stokkvågen (STOK), Meløy (MELS) and Mo i Rana (MOR8) were augmented in the summer of 2005 with temporary stations STOK1 and STOK2, both within 15 km of the assumed epicenter. These stations have provided us with the necessary confirmation of the occurrence of events in this region corresponding to 31 of the 32 correlation detections on the NORSAR array (the remaining marginal detection appears to be a false alarm). Several of the NORSAR detections occurred with a short time separation and one such sequence of detections (on April 28) is displayed in Figure 1. The first and largest correlation peak corresponds to the  $m_b=2.4$  event which was detected at regional distances. The second peak occurs approximately 90 seconds later, during the coda of the main event. Even without the observations at local distances, the evidence that this correlation peak corresponds to a true aftershock is quite strong in that similar correlation detections are observed at essentially the same time at all the Fennoscandian IMS arrays.

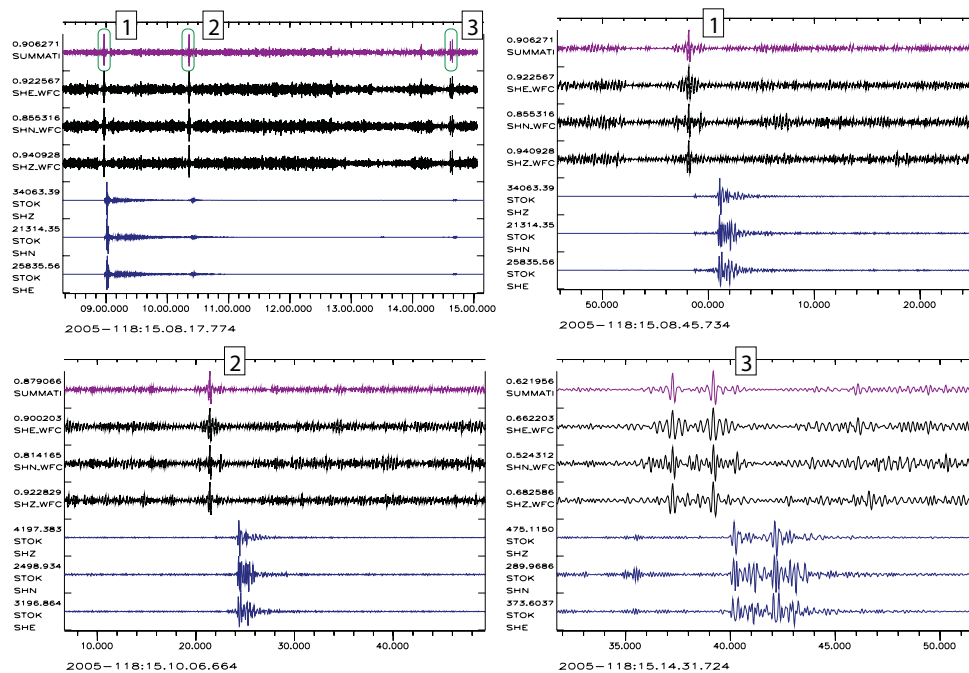


**Figure 1. (left)** Approximate location of the Rana earthquake sequence relative to the four Fennoscandian IMS array stations (black triangles) and the seismic stations of the Norwegian National Seismic Network (NNSN: red triangles). The line leading Northwards leads towards the SPITS IMS seismic array on the island of Spitsbergen. **(right)** Detection of an  $m_b=1.4$  aftershock to an  $m_b=2.4$  earthquake on April 28, 2005, in the Rana region of Northern Norway using multichannel waveform correlation on the large aperture NORSAR array at a distance of over 600 km. The master event signals (red traces) result from a magnitude 3.5 earthquake on June 24, 2005, in the same region, and are aligned with the April 28 data (blue traces) according to the first aftershock detection. Only the main shock can be detected on the single channel correlation coefficient trace; the numerous aftershocks are only detected using array processing of the correlation coefficient channels. The green rings mark detections on the array correlation beam. The third detection comprises two distinct correlation maxima, both close to the detection threshold.

Due to the short epicentral distances, the NNSN recordings give the clearest picture of this sequence and the STOK recording is shown for all 3 components in Figure 2. Unlike on the regional distance seismograms, the aftershock signals are observed clearly over the rapidly diminishing coda. The third and most marginal detection is seen to correspond to two distinct events with very similar waveforms separated by no more than 2 seconds. The radiation pattern from these earthquakes is such that the STOK station observes very little P-energy (always over an order of magnitude smaller than S) and, even at the short distances involved, the P-arrivals from these small events are not observed above the noise. Careful inspection of the correlation beam on the NORSAR array reveals two peaks resolved with the correct inter-event spacing. The aftershocks were not detected when templates of only 60 seconds of the P-coda were used; detections were however made when 60 second long templates of the Lg phase and coda were used.

The similarity of the STOK waveforms from event to event, together with the excellent signal-to-noise ratio (SNR) for the S-phases, provides excellent control of the event magnitudes. Applying a logarithmic scaling argument to the

STOK waveforms indicates that the smallest of the detected events were approximately  $m_b=0.5$ . This is at least an order of magnitude detectability improvement for the ARCES array and an even greater effective improvement for NORSAR given the lack of array gain for high frequency signals. All events detected by the 42 element NORSAR array could also be detected by the 25 element ARCES array. However, with the threshold set to detect the smallest of these events, the ARCES array generated many more false alarms. The tendency of small-aperture arrays to produce spurious correlation detections was documented by Gibbons and Ringdal (2006), who found that the vast majority of such false alarms could be screened out automatically by measuring the alignment of the correlation coefficient traces. This method is less effective for marginal detections when far longer waveform templates are used, and it remains one of the greatest challenges for the remainder of this project to find robust methods of maintaining high sensitivity together with a low false alarm rate for small-aperture regional arrays. The northern Norway study is described in detail in Gibbons et al. (2006).



**Figure 2.** Part of the April 28, 2005, Rana earthquake sequence as recorded by the Stokkvågen 3-component station of the NNSN at a distance of approximately 15 km from the earthquake epicenter. The top left panel shows a seven minute long time window which (correcting for the difference in traveltimes) corresponds to the interval shown in Figure 1 for the NORSAR array. The remaining panels provide a close-up view at the times indicated. In each panel, the blue traces are STOK data from April 28, bandpass filtered between 2.5 and 8.0 Hz, the black traces show the componentwise correlation coefficients with filtered STOK data from the June 24, 2005, event (waveforms not shown), and the magenta trace indicates the mean of the 3 single component correlation traces.

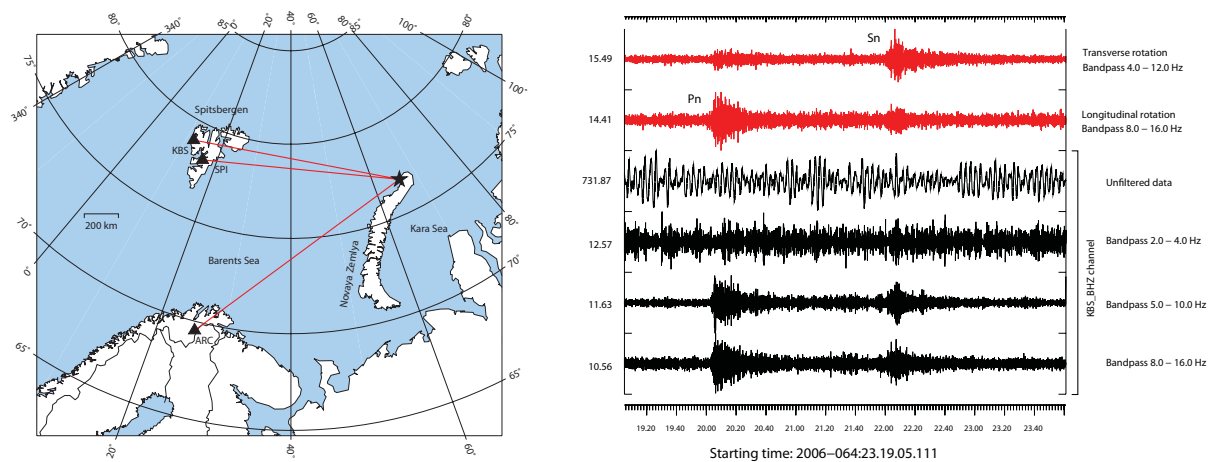
### The Use of Repeating Seismic Events to Control and Measure Erroneous Instrumental Timing

Gibbons and Ringdal (2006) pointed out that if two seismic events are co-located, since the travel time to a given station is identical for both events, the time separating the start of the waveform template for the first event and the maximum of the correlation coefficient channel (whereby the second event is detected) should be identical for all stations. This is the basis by which we can perform coherent beamforming of correlation coefficient traces over sparse networks even when the actual waveforms show no similarity whatsoever. If this is not the case (and the difference cannot be ascribed to waveform dissimilarity - whether due to differences in the seismic sources or to a low SNR) then we have to conclude that there is an inconsistency in instrumental timing at one (or both) of the stations at the time of one (or both) of the events. An event in the vicinity of Novaya Zemlya on March 5, 2006, was

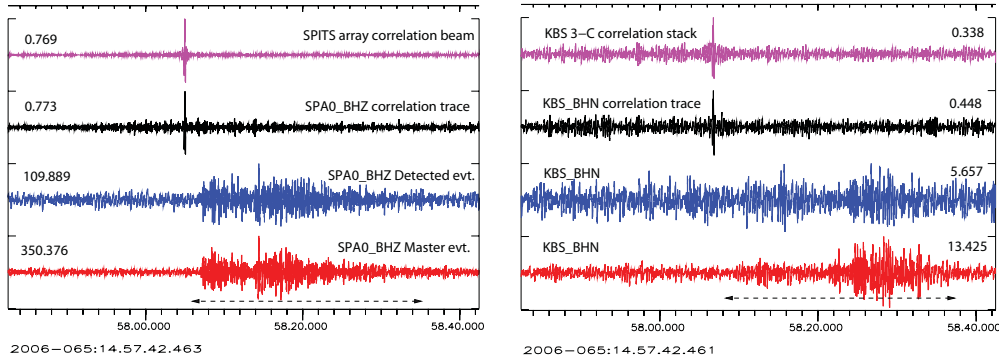
## 28th Seismic Research Review: Ground-Based Nuclear Explosion Monitoring Technologies

well-recorded by the SPITS and ARCYES arrays and also by the broadband station KBS on Spitsbergen (Figure 3). The arrival times for the Pn and Sn phases for the KBS station could not be reconciled with those at the array stations and the residuals obtained by various attempts to locate the event indicated that a consistent offset in the timing at KBS was to blame.

The operators of the station confirmed that a technical fault had occurred with the station on February 17, 2006, which was repaired on March 22, 2006. Continuous real-time correlation detectors have been run on SPITS data at NORSAR for some time to detect microseismic activity at the Barentsburg coal mine, the source of many almost-repeating seismic signals. Many of the Barentsburg events were recorded by both the SPITS array and the KBS station. Measurements of correlation maxima at both stations for many events before, during, and following the period effected by the technical fault allowed us to measure that at the time of the Novaya Zemlya event, the time stamp at the KBS station was approximately 8.07 seconds earlier than the actual UTC time (Figure 4). With the corresponding correction applied to the arrival time estimates, a well-constrained location is obtained for the Novaya Zemlya event including the KBS phase determinations. This study is described in full in Gibbons (2006).



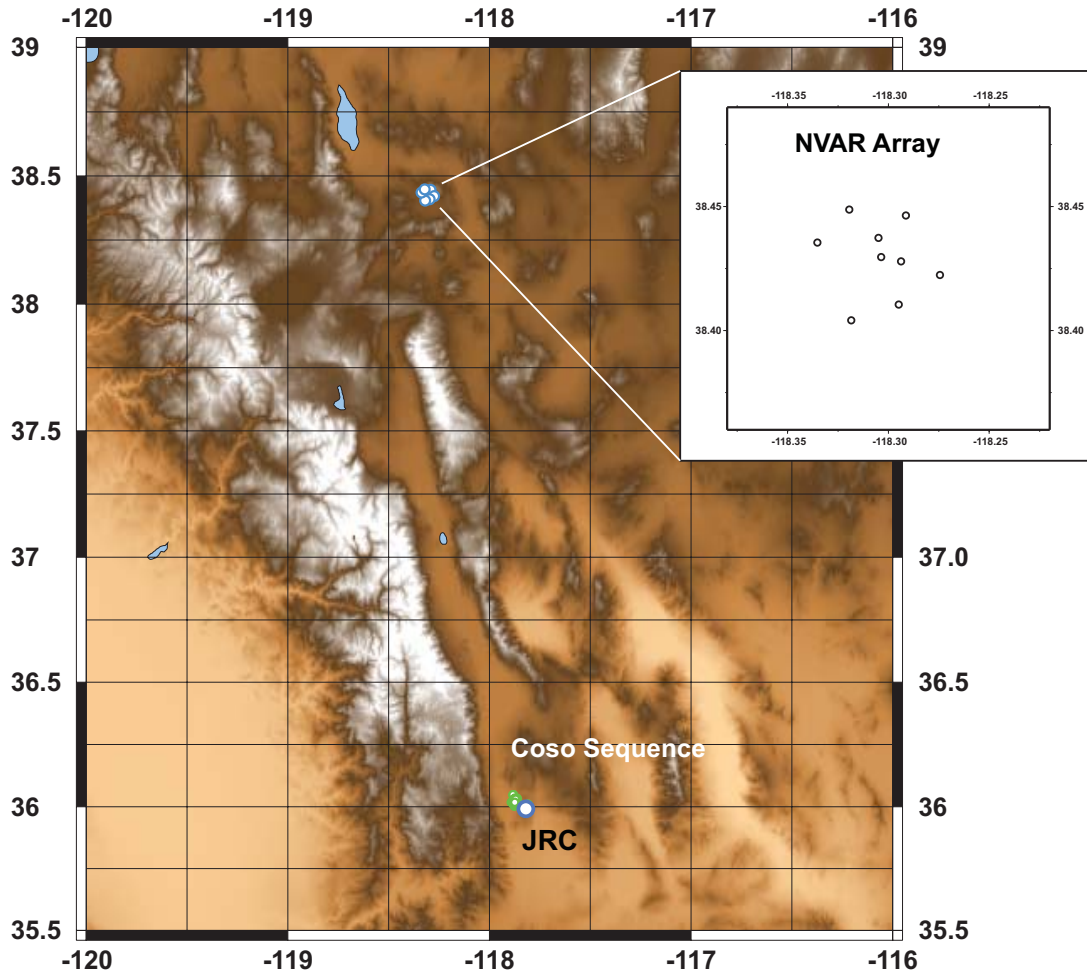
**Figure 3. Left panel: location of the IMS arrays ARCYES and SPITS and the IRIS/GEOFON/AWI 3-component station KBS at Ny Ålesund, Kings Bay, on Spitsbergen together with the fully automatic GBF location estimate for the March 5, 2006, event on or close to Novaya Zemlya. Right panel: waveforms from the Novaya Zemlya event recorded at the KBS station. Due to an instrumental technical fault, the time-stamp for the KBS station could not be relied upon and attempts to locate the Novaya Zemlya event using phase determinations from ARCYES, SPITS, and KBS demonstrate that the time-stamp at KBS must be too early by several seconds.**



**Figure 4.** Detection by waveform correlation of the first identified rockburst at the Barentsburg coal mine, Spitsbergen, following the March 5, 2006, Novaya Zemlya event using the SPITS array (distance 50 km) and the KBS 3-component station (distance 120 km). The master event used is the first identified Barentsburg event following the repair of the KBS station on March 22, 2006. All data are bandpass filtered between 3.0 and 6.0 Hz and a 30.0 second waveform template is extracted for all available channels at both stations beginning at the estimated onset time of the first P-arrival. The SPITS waveform template begins at a time 2006-081:23.25.07.15000 and the interpolated correlation coefficient maximum at SPITS occurs at a time 2006-065:14.58.05.00229. The KBS waveform template begins at a time 2006-081:23.25.17.00760 and the interpolated correlation coefficient maximum at KBS occurs at a time 2006-065:14.58.06.78268.

#### Detection Studies in the Western US

To gain experience with correlation detector performance in a quite different tectonic setting, we are testing and applying correlation and subspace detectors to event clusters in eastern California using the NVAR array. The general setting and array geometry are shown in Figure 5. Studies in this region benefit from the double-difference catalogue produced by Hauksson et al. (2003), which provides good relative event locations important to understanding how event waveform correlations are influenced by waveform proximity.



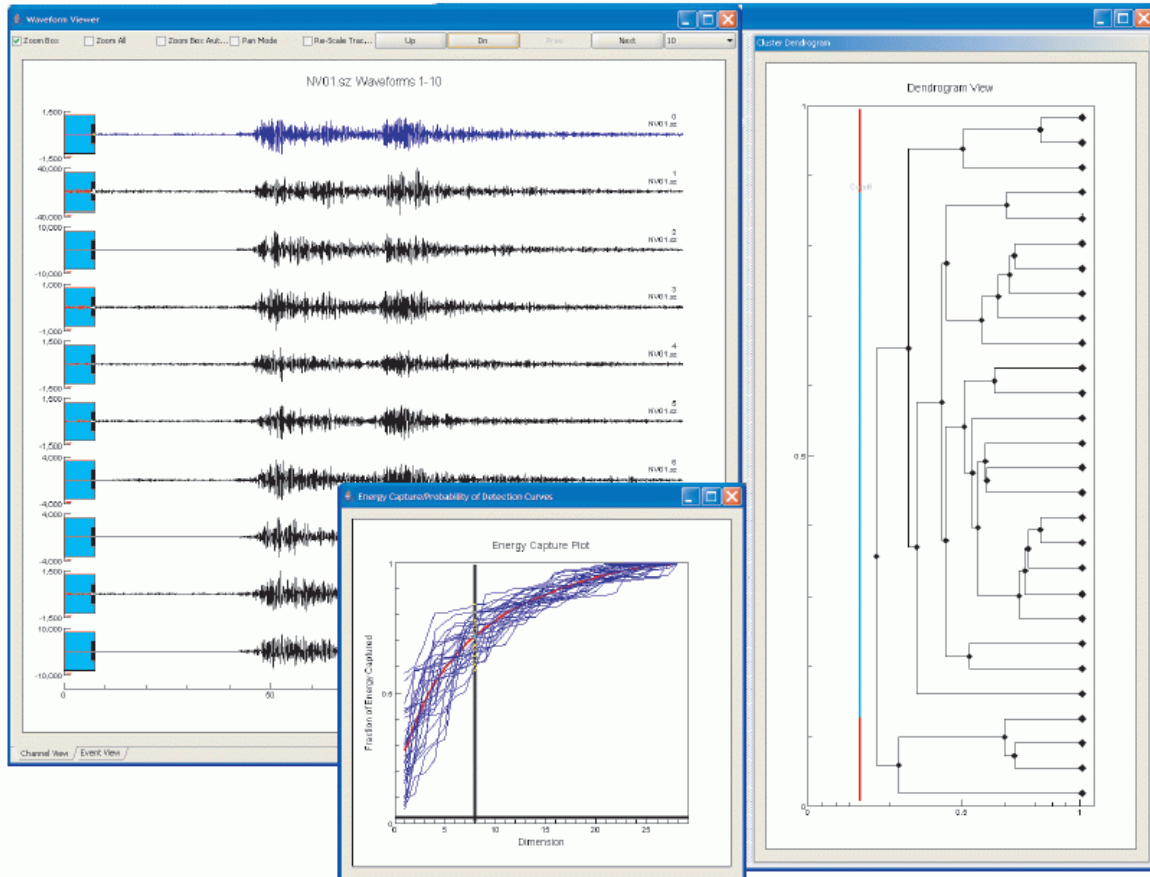
**Figure 5.** Location of the NVAR array (blue and white symbols) in western Nevada and the 2001 Coso sequence (green and white symbols) in southeastern California. Nine elements of the array were used to construct and apply detectors. The other elements of the array had significant data gaps or poor signal to noise ratio. The Coso sequence and two others in eastern California were chosen for study because of the availability of double-difference locations (Hauksson et al., 2003). While the overall sequence consisted of more than 3500 events (as reported in the catalog), 406 occurring in a 3-day period and in a fairly compact region were selected for study. Reported catalog magnitudes ranged from 0.4 to 3.9 for these events.

We describe one particular cluster here approximately 270 kilometers south of the NVAR array. This cluster of 406 events is part of a larger sequence just north of the geothermal region in Coso, southeastern California (Figure 5). The smaller cluster occurred in 2001 (Julian days 195-197) and had reported magnitudes ranging from 0.4 to 3.9. A local broadband station, JRC, within 5 kilometers of the cluster provides further ground truth information on the sequence.

A set of 29 events with magnitudes above 2.8 were used to characterize the events in the cluster - these are shown in Figure as green crosses. They were drawn from a longer time interval (195-211), but included many of the larger events in the detection interval. Waveforms from the 29 events were extracted from the NVAR data stream; a sample of these recorded at station NV01 are displayed in Figure 6. The figure shows screenshots from a new tool written to select and align waveform data, and compute subspace representations with user-selected parameters (filter band,

## 28th Seismic Research Review: Ground-Based Nuclear Explosion Monitoring Technologies

representation dimension, etc.). Besides a waveform display, used to screen bad events and channels, the tool has a correlator that aligns the waveforms, and a dendrogram display allowing the user to set clustering thresholds and select clusters. It also has a series of subspace panels that display the fidelity of representation (energy capture, shown in the figure) of waveforms and the probability of detection (at a fixed, selectable false alarm probability) both as a function of the subspace dimension. These analytical tools allow the user to select a subspace (correlator operator) dimension. The tool finally allows the user to write a detector definition file which can be used by a separate code to run the detector against a continuous (multichannel) data stream. In this instance, an 8-dimension detector was constructed and applied against the 9 channels of continuous NVAR data for the 3-day interval 2001:195-197.



**Figure 6.** Screen shots from a new tool showing several of the steps in constructing a subspace detector for the Coso sequence with data from the NVAR array. A waveform display tool (at left) allows the user to screen bad events and bad channels. A correlator aligns the waveforms and provides measurements for building a dendrogram display (at right). The dendrogram allows the user to define and select event clusters. A series of subspace panels (one shown at center: energy capture as a function of subspace dimension) allow the user to select a window on the aligned waveforms and choose a subspace dimension.

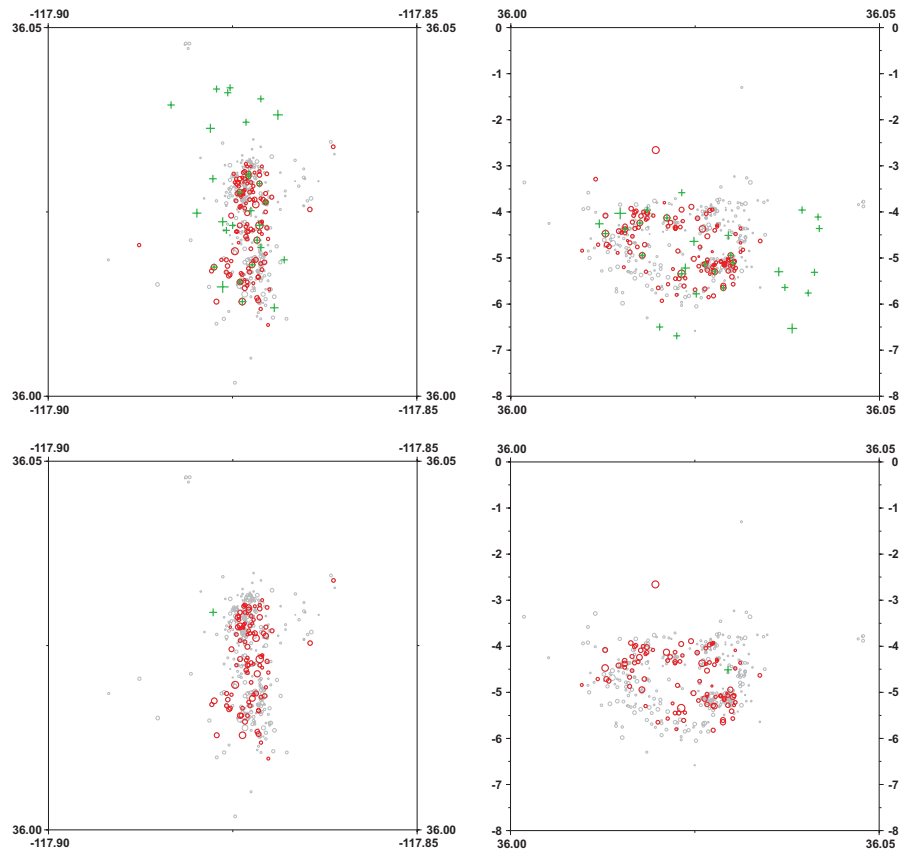


Figure 7. Detail maps showing the spatial distribution of training events (green), detected events (red) and undetected events (grey). Results for the 8-dimension subspace detector are on top and for a correlation detector are on the bottom. The maps show the distribution of events in latitude and longitude (left) and latitude and depth (right). Symbol size is scaled to event magnitude.

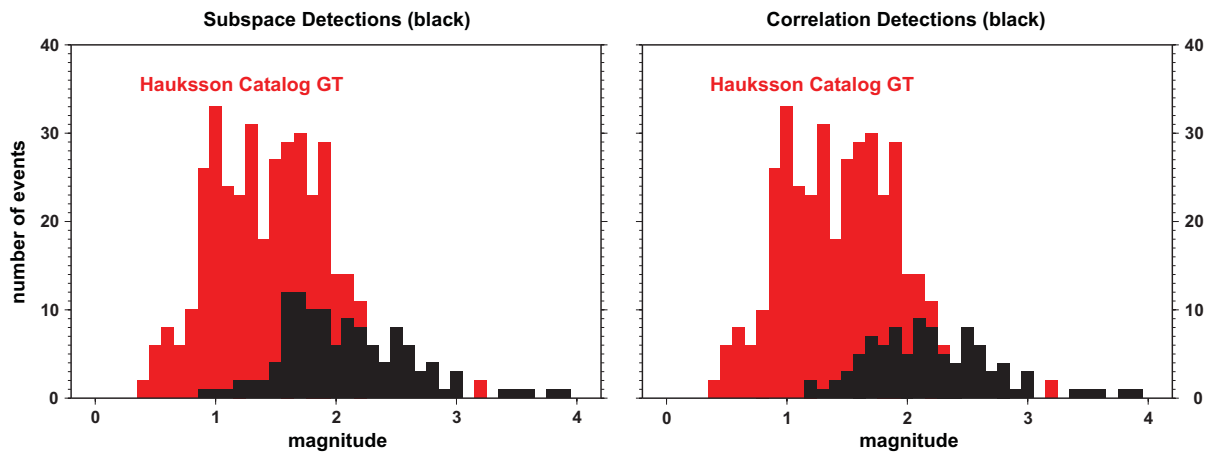


Figure 8. Histograms of detected events in this sequence in black for the subspace detector (left) and the correlator (right). The histogram for the Hauksson catalog is shown in red behind the detector histograms. Waveform representation is a significant challenge for this sequence.

## 28th Seismic Research Review: Ground-Based Nuclear Explosion Monitoring Technologies

The results of the detection operation are shown in Figures 7 and 8. Figure 7 contains maps of the 406 events comprising the cluster in question. The green crosses indicate the reference events, the red circles indicate detected events and the grey circles indicate undetected events. Results for the 8-dimension multichannel subspace detector are shown at the top: 120 events were detected after adjusting the detection threshold (to 0.03) to allow only 2 presumed false alarms (defined as events not reconcilable against the Hauksson catalog). At left in Figure 8 is a histogram (in black) of the 120 detections reconciled against the catalog as a function of catalog magnitude. Behind the histogram of detections is another histogram (in red) of all the catalog events. Comparison of the two histograms makes it clear that the subspace detector missed at least half the events between magnitudes 1.5 and 2.2, and two magnitude 3.2 events (these were in coda of still larger events).

For comparison, a multichannel correlation detector was developed using one event (shown as the green cross in the bottom of Figure 7) and a similar detection operation was performed. Adjusting the detection threshold (to 0.0125) to allow the same number (2) of unreconciled triggers as in the subspace detector case, this detector found 95 events distributed about as widely as for the subspace detector. The correlation detector threshold was substantially lower than the subspace detector because of the fact that the 8 basis functions of the subspace detector significantly raise the noise floor of the detection statistic (while simultaneously raising the values of the statistic for actual events).

We can only speculate about the reasons for the relatively poor performance of both detector types in this case (compared to almost complete detections in other instances). From cursory examinations of the three-component signals observed at the local station JRC, it does appear that the mechanisms of the events not detected are significantly different from the larger events used for design. One hypothesis consistent with the observations is that the larger events have principal axis orientations aligned with the dominant stress field in the region, and thus have relatively uniform mechanisms. This hypothesis would explain neatly why the correlation detector performed nearly as well as the subspace detector designed from 29 events: the additional 28 events would add relatively little information to the representation relevant to the smaller events. In this view, the smaller events might have relatively random orientations.

### **CONCLUSIONS AND RECOMMENDATIONS**

In the first year of this project, we have demonstrated a significant lowering of the detection threshold using the multichannel waveform correlation method in a number of different scenarios; continuous real-time online correlation detectors have been implemented for an increasing number of test cases on all of the arrays operated by NORSAR. In the example presented here, the signal from an  $m_b=3.5$  event in northern Norway was successfully used as a template to detect events in the immediate vicinity down to  $m_b=0.5$  at distances of over 600 km. The improvement in detection capability appears to be greater for large aperture seismic arrays. False alarms (i.e. spurious correlations) are generally identified by measuring the alignment of peaks in the correlation coefficient traces at different sites in the array. This becomes a greater challenge for very weak signals on small-aperture arrays and will be a main focus of future work. The alignment of correlation peaks has also proved very useful in the identification, measurement, and correction of erroneous timing at seismic stations. This property alone should motivate a large scale initiative to identify sources of repeating seismic events.

The greatest drawback of matched filter detectors is the requirement that we have a representative waveform to use as a detection template. Whilst we are encountering surprisingly many situations in which correlation detectors are highly effective, we have also identified a number of situations where very closely spaced events display a very high degree of waveform dissimilarity and where correlation detectors subsequently miss a large number of events. In a case study in the western United States, using NVAR array data from the Coso earthquake sequence from 2001 (at a distance of 270 km), multidimensional subspace detectors were found to perform only modestly better than a correlation detector. Large numbers of events went undetected by either correlation or subspace detectors, suggesting that the smaller magnitude events in this sequence may display greater diversity in mechanism than the larger events used to design the detection templates.

**ACKNOWLEDGEMENTS**

We are grateful to Mathilde Böttger Sørensen and Jens Havskov from the Department of Earth Sciences at the University of Bergen (UiB), Norway for providing data from the STOK station. The Norwegian National Seismic Network is operated by the UiB and supported by Oljeindustriens Lands Forening (OLF) and the Faculty of Mathematics and Natural Sciences, UiB.

**REFERENCES**

- Geller, R. J. and Mueller, C. S. (1980). Four similar earthquakes in Central California, in *Geophys. Res. Lett.*, 7: 821-824.
- Gibbons, S. J. (2006). On the identification and documentation of timing errors: an example at the KBS station, Spitsbergen, *Seism. Res. Lett.*, in press.
- Gibbons, S. J., Böttger Sørensen, M., Harris, D. B. and Ringdal, F. (2006). The detection and location of low magnitude earthquakes in northern Norway using multichannel waveform correlation at regional distances, manuscript submitted to *Phys. Earth Planet. Int.*, under review.
- Gibbons, S. J. and Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation, in *Geophys. J. Int.*, 165: 149-166.
- Hauksson, E., W-C. Chi, P. Shearer, and A. Michael (2003). Comprehensive waveform cross-correlation of southern California seismograms: Part 1. Refined hypocenters obtained using the double-difference method and tectonic implications (abstract), Fall Ann. Meeting, American Geophys. Union. Dec. 8-12, San Francisco, CA.
- Hicks, E. C., Bungum, H. and Lindholm, C. D. (2000). Seismic activity, inferred crustal stresses and seismotectonics in the Rana region, Northern Norway, *Quaternary Science Reviews*, 19: 1423-1436.