
A MATCHED FIELD PROCESSING FRAMEWORK FOR COHERENT DETECTION OVER LOCAL AND REGIONAL NETWORKS

Tormod Kværna, et al.

**NORSAR
PO Box 53
N-2027 Kjeller, Norway**

1 June 2011

Draft Final Report

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.



**AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
3550 Aberdeen Ave SE
AIR FORCE MATERIEL COMMAND
KIRTLAND AIR FORCE BASE, NM 87117-5776**

DTIC COPY

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the Air Force Research Laboratory 377 ABW Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RVPS-TR-2011-0130 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED//

Robert Raistrick, RVBYE
Program Manager

//SIGNED//

Robert Morris, PhD
Chief, AFRL/RVB

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE 01-06-2011		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 26-Jun-2010 to 02-May-2011	
4. TITLE AND SUBTITLE A Matched Field Processing Framework for Coherent Detection Over Local and Regional Networks				5a. CONTRACT NUMBER FA9453-10-C-0209	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 62601F	
6. AUTHOR(S) Tormod Kværna, D.B. Harris , S.J. Gibbons, and D. Dodge				5d. PROJECT NUMBER 1010	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER 837619	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NORSAR PO Box 53 N-2027 Kjeller, Norway				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory Space Vehicles Directorate 3550 Aberdeen Ave SE Kirtland AFB, NM 87117-5776				10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/RVBYE	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-RV-PS-TR-2011-0130	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution is unlimited. 377ABW-2011-0932; 21 June 2011.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT In this final report it is demonstrated that it is possible to generalize the single-phase matched field processing template developed under a previous contract to a template constructed from an entire seismogram. This template can be used as a basis for a detection algorithm implemented in a subspace detector framework. Furthermore, the framework can represent correlation detectors and the incoherent (narrowband) matched field detectors and gradations between the two. It is shown that detectors in the framework can be made to adapt, as new event waveforms become available, through a subspace update mechanism. A detector that begins as a matched field detector will adapt to resemble a correlation detector, if the source produces highly repeatable waveforms. This form of adaptation is demonstrated with a highly repeatable source in Fennoscandia, and somewhat different adaptive behavior that improves detection performance with a mining explosion source in the Korean peninsula. It is demonstrated that the performance of the matched field detector can improve dramatically as the number of available sensors increases, and that the matched field and subspace detectors are applicable to very large apertures over which conventional notions of waveform coherence do not apply.					
15. SUBJECT TERMS Seismic arrays, Empirical Matched Field Processing					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Robert Raistrick
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED	88	19b. TELEPHONE NUMBER (include area code)

This page is intentionally left blank.

TABLE OF CONTENTS

Summary	1
1. Introduction.....	2
2. Theory and Implementation.....	8
3. A Reference Database of Events from the Korean Peninsula	32
4. Testing of the Matched Field Detector	43
5. Detection Framework.....	59
6. Conclusions.....	75
References.....	77

LIST OF FIGURES AND TABLES

Figure	Page
1 Map Showing Study Region of Our Previous Contract on Matched Field Processing	4
2 Results of Event Classification with FK Analysis and Matched Field Processing.....	5
3 Results of Classification with Wideband Waveform Correlation.....	5
4 Representative Waveforms from Two Different Sources Types	7
5 The Real Parts of the First 10 Impulse Responses and Amplitude Frequency Responses for a Bank of Narrowband Filters.....	10
6 Complex Envelopes Obtained by Filtering the Array Waveforms Into One Particular Narrow Frequency Band	12
7 Sample Complex Envelope Covariance Matrices for Ensembles of Events from Two Sources Approximating a Source Fully Correlated Across Frequency Bands (the Finnish Munitions Demolition Site) and Uncorrelated Across Frequency Bands (Kirovsk Underground Mine)..	17
8 Distribution of Eigenvalues for the Finnish and the Kirovsk Explosion Ensembles.....	18
9 Structure of the Subspace Implementation of the Detection Framework	22
10 Sample Finnish Munitions Explosion Used to Create Templates	23
11 Segment of Data, Used to Test the Detectors, Containing Two Finnish Munitions Explosions.....	24
12 Detection Statistic for Incoherent Detector and Coherent Detector	24
13 Detail of Detection Statistics for the Incoherent Detector and the Coherent Detector Around a Design Event.....	25
14 Comparison of Detection Statistics for a Correlator and a Rank-3 Detector.....	26
15 Detail of the Correlator and Rank-3 Detector Statistics Around the First Munitions Explosion	27
16 Template Rank as a Function of Iteration Number.....	30
17 Detection Statistic Around First Finnish Munitions Explosion in Four Hour Test Sequence..	31
18 Evolution of the Detection Statistic Around the First Finnish Munitions Explosion.....	31
19 Location of the IMS Arrays KSRS (Wonju, South Korea), USRK (Ussuriysk, Russian Federation), and MJAR (Matsushiro, Japan) and the IRIS 3-Component Stations INCN (IU Network) and MDJ (IC Network).....	34
20 An Example of the Detection of a P and S Phase for an Event Within Regional Distances of the KSRS Array	35
21 Broadband F-K Estimates on the KSRS Array for a Regional P and a Regional S Phase	36
22 Approximate Locations of 19 Sites at Which Repeating Events Have Been Observed on the KSRS Teleseismic Array in the Republic of Korea.....	37

23	Waveforms from the KS01_SHZ Instrument, Aligned According to the Times of Best Correlation Coefficient for 40 Events from Four Clusters	38
24	Array-Based Waveform Correlation Detections for Six Mine-Type Clusters.....	40
25	Array-Based Waveform Correlation Detections for Six Earthquake-Type Clusters.....	41
26	Array-Based Waveform Correlation Detections for Seven Clusters of Unknown Origin.....	42
27	View from Google Earth Centered on the Location 39.413°N and 125.873°E Near the City of Sunshon in North Korea	43
28	Location of the Hukkakero Military Explosion Site and the Suurikuusikko Gold Mine in Relation to the ARCES Array and 3 Broadband 3-Component Stations of the Finnish National Network.....	45
29	Waveforms on the Channel ARA0_Sz of ARCES from Each of 12 Events from the Two Finnish Explosion Sites.....	46
30	Variation of Waveforms Over the Different Receiver Apertures.....	46
31	Matched Field Detection Statistic Traces for Two Finnish Explosion Sites Using 4 Sensors of the ARCES Array.	47
32	Matched Field Detection Statistic Traces for Two Finnish Explosion Sites Using 4 Stations of a Regional Network.....	48
33	Map Showing the Location of the KSRS Array in South Korea, the Odaesan Earthquakes, the M4 Mine and the DGY Three-Component Station.....	50
34	Sensor Distribution of the 19-Element KSRS Array in South Korea.....	51
35	Panel of 10 Events from the Odaesan Earthquake Sequence Recorded at the KS01_SHZ Sensor of the KSRS Array.	52
36	Detector Processing Results Using the Five KSRS SHZ Sensors for a One-Hour Time Interval Around the Largest Foreshock of the Odaesan Earthquake Sequence.....	53
37	Detector Processing Results Using the Five KSRS SHZ Sensors for a One-Hour Time Interval Around the Main Odaesan M_b 4.8 Earthquake on 20 January 2007.	54
38	Panel of 10 Events from the M4 Mine Recorded at the KS01_SHZ Sensor of the KSRS Array.	55
39	Detector Processing Results Using the Five KSRS SHZ Sensors for 25 January 2007.....	56
40	Detector Processing Results for a One-Hour Time Interval Around Two M4 Event Candidates That Occurred Right After 03:00 on 25 January 2007.....	57
41	Detector Processing Results for a One-Hour Time Interval Around Two M4 Event Candidates That Occurred Right After 03:00 on 25 January 2007 Using the Full 19-Element KSRS Array Geometry.	58
42	Schematic of the Detection Framework	60
43	Tables Used to Store Results from Previous Runs of the Framework.....	61
44	Tables Used to Define Stream Groups.....	62

45 Tables for STA/LTA Boot Detectors.....	63
46 Tables for Array Power Boot Detectors.....	64
47 Tables Used by the Detection List Detector	65
48 Subspace Detector Tables	66
49 Tables Used for Matched Field Detectors	67
50 Template Rank as a Function of Iteration Number.....	68
51 Detection Statistics Around the First, 20 th and 40 th Events Detected for the Detector Initialized with an Incoherent (Rank 33) Template and Allowed to Adapt as New Events Were Detected.....	69
52 Evolution of the Detection Statistic in the Vicinity of Successive Detections for the Adaptive Detector in the Detection Framework	70
53 Rank of the M4 Detector as a Function of Adaptation Step	71
54 Detection Statistics Around the Detection Times of the First, 20 th and 40 th Detections for the M4 Detector.	72
55 Evolution of the Detection Statistic for the M4 Detector.	73
56 70 M4 Detections Used in the Template Adaptation Process.....	74
57 Comparison of Correlator Detections of M4 Events and Adaptive Framework Detections	75

Table

Page

1 A List of 19 Clusters of Repeating Seismicity on and Surrounding the Korean Peninsula.....	39
--	----

Acknowledgments

We are grateful to staff at the IDC in Vienna, Austria, for preparation of the data from the KSRS array.

Data from the DGY station in South Korea are courtesy of the Korea Meteorological Administration. We are grateful to Kwanghee Kim for discussions on the 2007 Odaesan earthquake.

We are grateful to the University of Oulu for data from the RNF station (part of the Northern Finland Seismological Network, FN) and to the University of Helsinki for data from the VRF and HEF stations (part of the Finnish seismograph network).

All maps are produced using GMT software (Wessel & Smith, 1995).

This page is intentionally left blank.

Summary

The objective of this one-year feasibility study has been to develop a data-adaptive matched field procedure to detect coherently and incoherently across networks of stations at local and regional distances. The detector extends the single-phase matched field processing approach to detection using the entire waveform. We base the procedure upon a narrowband signal representation that exposes invariant spatial and temporal correlation structure of network signals from repeating sources. Wideband multichannel correlators and subspace detectors are both realizations of the resulting detector representation class, of which the matched field detector is an end member. The matched field detector is referred to as an incoherent detector (spatially coherent over a receiver aperture, but temporally incoherent due to the incoherent summation over the narrow frequency bands) and is optimal for sources displaying significant variation in the source time function from event to event. The framework is designed to adapt to the statistics of source time histories for any given target and employs an exponential age-weighting approach to effectively monitor evolving sources, such as open-cast mines.

For a given source of interest, an initial detector is formed. If only a single observation is available, the detector can be either a rank-1 coherent detector (a correlator) or an incoherent detector. If multiple observations are available, an event ensemble can be formed, allowing higher rank subspace detectors. As each subsequent similar signal is detected, a new covariance matrix is formed where the contributions from the waveforms of the existing members of the event ensemble and the newly detected signal are weighted according to processing parameters selected to control the adaptation properties of the detector. The nature of the new detector is determined by the eigenspectrum of the resulting covariance matrix. A covariance matrix with the eigenspectrum heavily concentrated in a single eigenvalue indicates that the phase structures of template components in all frequency bands are locked together and a frequency-coherent processor (correlation detector) is formed. If the energy in the eigenspectrum is dispersed over many eigenvalues, then the emerging processor is a higher-rank subspace projection operator. The rank of the detector is determined by an energy capture criterion.

We have evaluated the framework on datasets in the European Arctic, exploiting extensive existing Ground Truth information, and the Korean Peninsula, where repeating sources have been identified using diverse array processing techniques on the KSRS array. For munitions demolitions at Hukkakero in Finland, observed on the ARCES array, the incoherent detector adapts to a rank-2 subspace detector, reflecting the low variability of signals recorded from this site. A presumed open-cast quarry in South Korea, observed on KSRS, does not produce low-rank detectors and this may be the result of both source variability and low signal SNR. The performance of this detector starts off poorer than that of a single-template waveform correlation detector, missing several events detected by the correlator. As the detector captures more events and adapts to the evolving structure of the wavefield, more events are detected and the correlator

is ultimately outperformed. We also examine the 2007 Odaesan earthquake sequence in South Korea for which a higher-rank subspace detector, comprising several of the larger events, detects several events missed by waveform correlators using single events as templates. This demonstrates the subspace detector's ability to represent a greater diversity of source mechanisms and/or an increased geographical footprint. The performance of the incoherent (matched field) detector increases dramatically as the number of available sensors increases. We demonstrate the applicability of matched field and subspace detectors to far larger receiver apertures over which conventional waveform coherence does not apply.

1. Introduction

This project was initiated to extend the successes of our earlier project on event classification with matched field processing to several new objectives:

- event detection,
- use of the entire seismogram rather than a single phase,
- adaptation to the statistics of source time histories, and
- coherent processing over increasingly larger apertures.

We have succeeded in formulating and implementing a single detector representation that includes wideband multichannel correlators and matched field detectors as admissible realizations. The representation is a complex subspace detector defined on a narrowband decomposition of the seismic wavefield. The matched field detector (an end member of the representation class) directly generalizes the matched field classification algorithm (which is spatially coherent, but temporally incoherent) to incorporate the entire observed wavefield as a template. We provide two approaches for designing detectors in this representation that automatically adapt detectors to the statistics of their sources. The first designs a detector from a pre-defined ensemble of events and the second incrementally adapts the detector to source characteristics as new events are detected. We are conducting tests to determine whether this framework produces detectors that coherently combine observations over large apertures.

Before providing more detail on our accomplishments during this contract, we first summarize results obtained during our previous research contract that motivated the current work. Under this prior contract, we adapted matched field processing, a detection and localization method operating on narrowband, continuous signals in underwater sound problems to solve a seismic classification problem using wideband, transient observations. We used an empirical form of matched field processing to push the resolution limit of the ARCES array in distinguishing explosions from closely spaced mines at distance ranges from 350 to 410 kilometers (Figure 1). The sources of the explosions were correctly identified over 98% of the time from ARCES Pn observations.

As we applied it in this context, matched field processing is a generalization of FK spectral analysis. The data in a window surrounding the Pn phase are transformed into the frequency

domain and spectral covariance matrices \mathbf{R}_k for the array signal are estimated for a number of frequencies $\{f_k\}$. A frequency-wavenumber spectrum

$P(f_k, \mathbf{s}) = \boldsymbol{\varepsilon}_k^H(\mathbf{s}) \mathbf{R}_k \boldsymbol{\varepsilon}_k(\mathbf{s})$	(1)
---	-----

maps the power incident on the array in the frequency band centered on f_k as a function of vector slowness \mathbf{s} . In conventional FK analysis, $\boldsymbol{\varepsilon}_k(\mathbf{s})$, the so-called steering vector, is a vector of complex phasors defined by a plane wave model for the incident Pn phase. It serves to focus the array on a particular source, represented, in this case, by the direction and velocity of the Pn wave from that source. The frequency-wavenumber spectrum of equation (1) is a narrowband map of incident power. A wideband map may be obtained by summing the spectra over a range of frequencies:

$P(\mathbf{s}) = \sum_k \boldsymbol{\varepsilon}_k^H(\mathbf{s}) \mathbf{R}_k \boldsymbol{\varepsilon}_k(\mathbf{s})$	(2)
---	-----

The problem of locating an explosion assumed to occur at one of a collection of known mines can be treated as a discrete classification problem. Figure 1 illustrates the concept. There are two groups of mines in the Kola peninsula, each consisting of five mines. Consequently, there are ten hypotheses about the origin of an explosion occurring in this region. The rule for choosing the correct hypothesis is to choose the source that maximizes the FK spectrum, i.e. to find the value of the slowness vector \mathbf{s} that maximizes $P(\mathbf{s})$. Since there are ten possible sources, there are ten values of \mathbf{s} to test.

Empirical matched field processing performs the same analysis, but replaces the theoretical steering vector defined by the great circle back-azimuth to the source, nominal Pn velocity and plane wave model with an empirical steering vector obtained from prior observations of Pn for the source in question. The empirical steering vectors are obtained as the principal eigenvectors of covariance matrices estimated from an ensemble of events characterizing the source. There is one empirical steering vector for each frequency band.

Figure 2 indicates the performance of event classification by maximizing the FK spectra and the corresponding matched field processing statistic. We obtained 549 events with good ground truth information, which we classified by originating source. The top frame of Figure 2 shows the results of classification with the FK measurement. Most of the events are incorrectly assigned to one particular mine (K2 – Rasvumchorr). In this region, FK spectral peaks are biased to the south (i.e. exhibit a clockwise rotation in slowness space). The Rasvumchorr mine is the southernmost mine, therefore its theoretical steering vectors should tend to maximize the FK spectra. This effect probably accounts for the fact that most explosions are attributed to this mine. The middle frame displays the results of performing classification with FK analysis, but with slowness corrections, as are commonly employed to correct effects of unmodeled refraction (e.g. Schweitzer, 2001). These corrections still assume a plane wave model for the signal, but apply a correction to the slowness vector in applying that model. In this case classification

results are substantially better, with few errors between the two mining groups, but substantial errors within the Khibiny (southernmost, more distant) group. The bottom frame shows the results of classification with the matched field statistics, which exhibit nearly perfect separation of the ten mines.

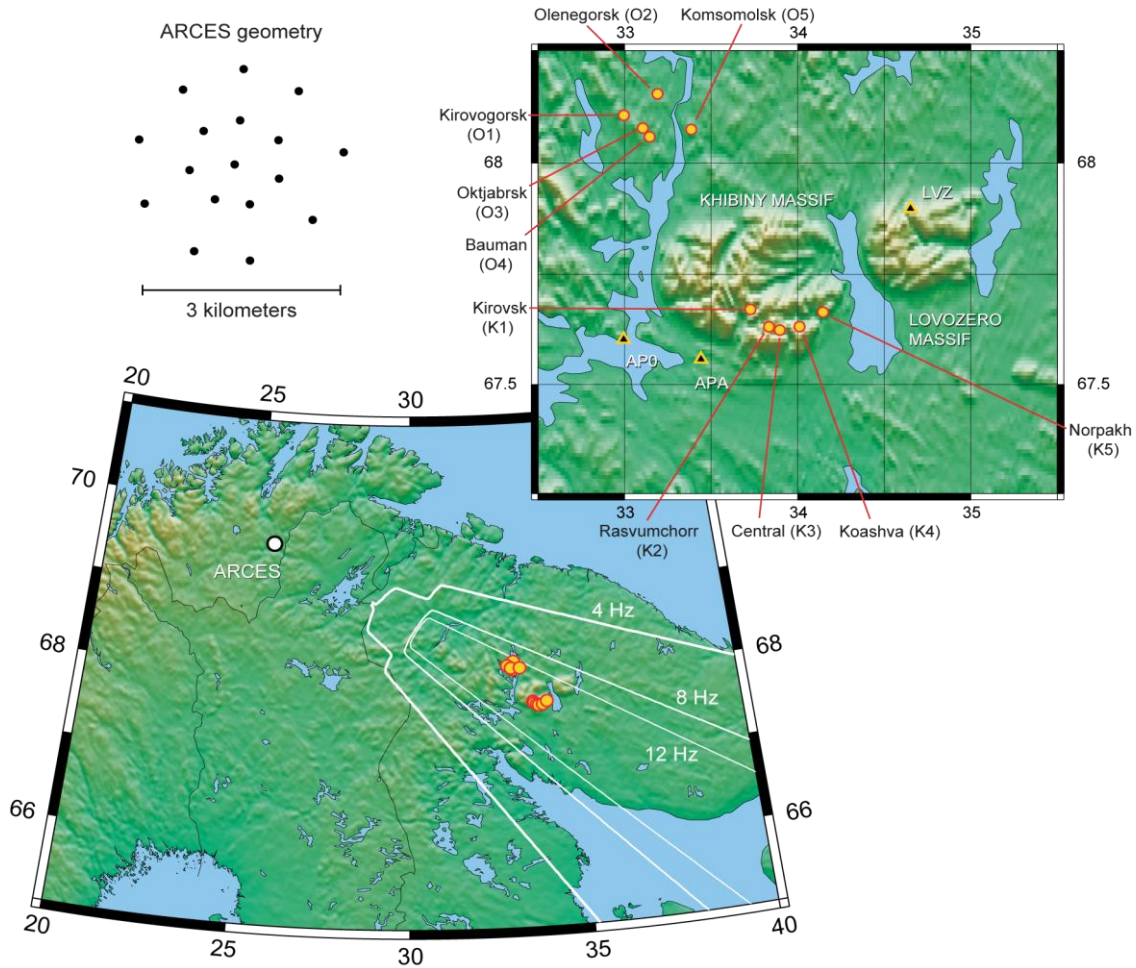


Figure 1: Study Region of Our Previous Contract

As seen in Figure 1, the ARCES array was used to distinguish events among mines in the Khibiny and Olenegorsk groups, using just Pn observations. At bottom left, the 3 dB points of the main lobe of the theoretical beampattern (at 4, 8 and 12 Hz) of the array are projected onto the geographic field of observation, assuming the array to be steered at the Rasvumchorr mine. All of the mines fall within the main lobe of the array pattern, even at 12 Hz, suggesting that conventional FK analysis will not perform well in resolving these sources.

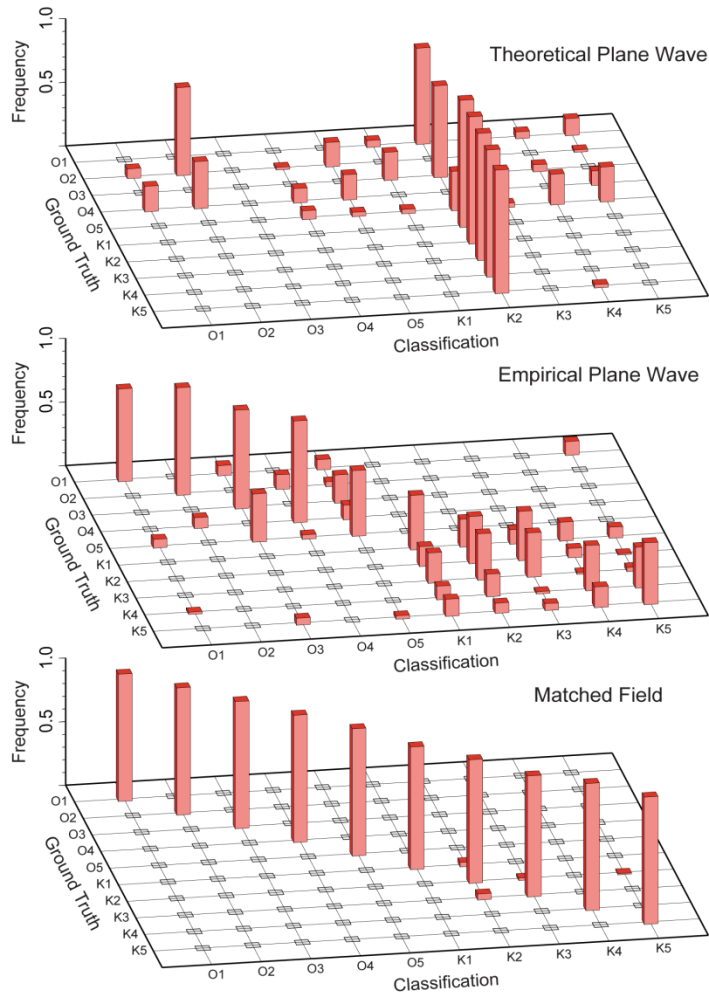


Figure 2: Results of Event Classification with FK Analysis (Top), FK Analysis with Slowness Corrections (Middle) and Matched Field Processing (Bottom)

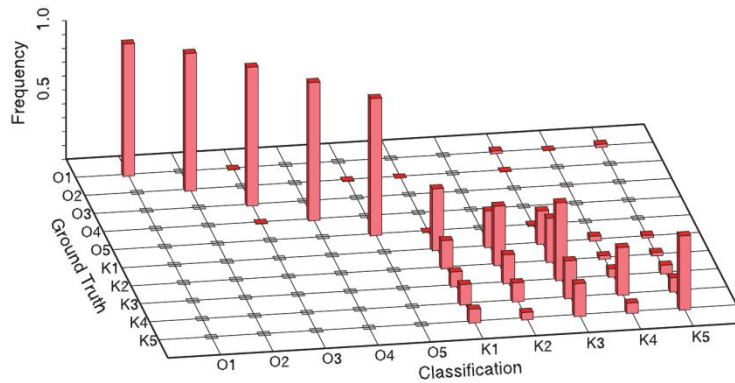


Figure 3: Results of Classification with Wideband Waveform Correlation

Waveform correlation also is an option for event classification. The approach would be to select a representative event for each mine and use the corresponding array waveforms as a templates, against which the waveforms of a new event would be compared to identify the originating mine. The rule we used in this example was to select the source that produced the maximum correlation between its template waveforms and the waveforms of a new event (a broad set of offsets was considered to eliminate the effects of waveform misalignment). Figure 3 shows the results of event classification using this approach. The results are clearly better than either of the FK methods of classification, but still not as good as the matched field approach. To reduce the likelihood of a poor choice of master templates we performed 10 randomized trials, with different events selected in each trial to define the waveform templates, and averaged the results.

We attribute the success of empirical matched field processing in this example to two factors. First, as a calibration technique, it captures the effects of refraction and scattering in the wavefield across the array aperture. These effects appear to provide distinct “fingerprints” for the ten mines that do not appear in the simple plane wave model for the signals. Second, matched field processing decomposes the signal into (approximately) disjoint frequency bands, performs the matching operation band by band, then reassembles a wideband result incoherently. There is no assumption that the signal is correlated in distinct frequency bands. From the comparison of matched field and correlation results, this approach appears to have advantages for classifying mining explosions. This observation suggests that we treat explosions at an individual mine as stochastic processes, which produce finite-duration signals that vary substantially from shot-to-shot, but nonetheless have Green’s functions describing propagation to the array that are approximately invariant. The narrowband decomposition of the signal allows us to separate the (approximately) invariant spatial structure of P_n across the array aperture from its highly variable temporal structure. To first order, matched field processing matches the spatial structure of the signal, but not the temporal structure. Waveform correlation attempts to match both, which may account for its relatively poor performance.

These results were encouraging, and led to a series of questions considered in this contract:

- 1.) Can we construct a generalization of the single-phase matched field processing method that encodes the full structure of the entire wavefield? What would this generalization look like? It would need to retain the advantage of incoherent combination of the matching statistic across frequency bands, where appropriate for temporally-variable sources.
- 2.) Since not all sources exhibit strong temporal variations, is it possible to construct a mathematical framework for detectors that allows implementation of correlators appropriate for exactly repeating sources and some form of matched field detector for temporally-variable sources? Examples of each type of source are shown in Figure 4. In principle, the framework should allow continuous gradations between two extremes: perfect coherence across frequency bands and complete incoherence. This requirement suggests that the framework should be some type of subspace detector (Harris, 2006a,b).

- 3.) Can a detector implementation be build incrementally, beginning with waveforms from a single event and evolving to the structure of an ensemble detector as more events are detected? How should this process be initialized in order to assure that events for adaptation are detected?
- 4.) Can spatially coherent processing be extended in practice to network apertures?
- 5.) Are the results obtained in Fennoscandia portable to other areas of interest?

We take up these issues in subsequent sections of this report, beginning with a mathematical description of a framework for detectors that supports the objectives of points 1 through 3 above.

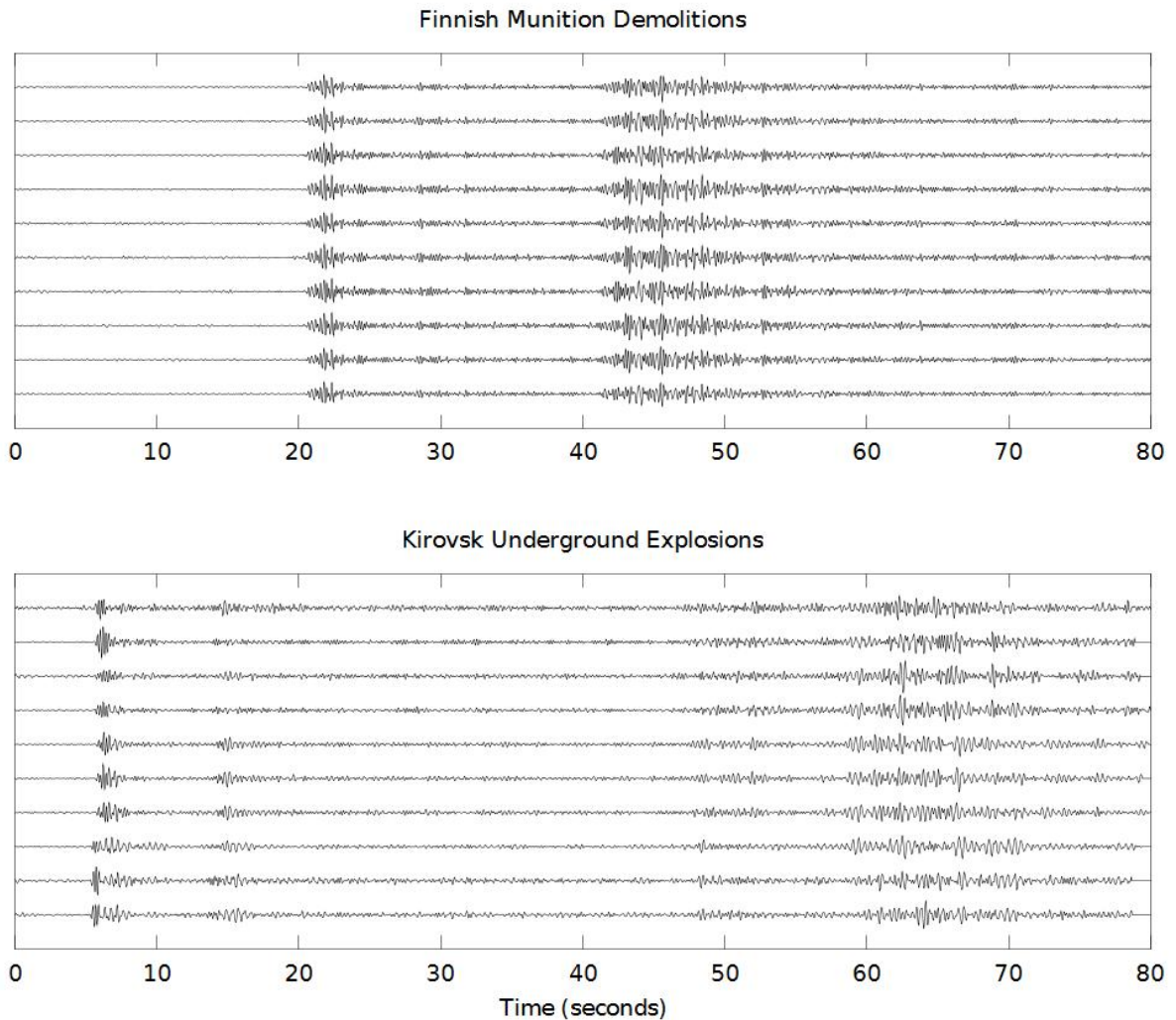


Figure 4: Representative Waveforms from Two Different Sources Types

Figure 4 shows 10 consecutive explosions from a munitions demolition site in Finland and 10 consecutive explosions from the Kirovsk underground mine in the Khibiny Massif. Note the similarity of the munitions shots and the variation among the Kirovsk shots.

2. Theory and Implementation

Narrowband representation

The requirement that the detector capture the spatial structure of the characteristic waveforms from the source, but not the fine temporal structure of the source time history can be met with a narrowband waveform representation. The signal \mathbf{r} observed from a particular source has a representation in terms of a convolution between a forcing source time history f and a collection of Green's functions \mathbf{g} that describe propagation from the source to the observing sensors:

$\mathbf{r}[n] = \sum_l f[l] \mathbf{g}[n-l]$	(3)
---	-----

Here, we are describing signals in discrete time, consistent with digital recording and using bold characters to denote vector quantities. The size of the signal vector is N_c , the number of observing channels. The choice of a scalar forcing function is a simplification for reasons of exposition; extensions to more complicated sources described by multiple forcing functions are straightforward.

It would be possible to take a Fourier transform with respect to the entire waveform to replace the time-domain convolution operation with a multiplication operation:

$\mathbf{r}(\Omega) = \sum_l \mathbf{r}[n] e^{-j\Omega n} = f(\Omega) \mathbf{g}(\Omega)$	(4)
---	-----

which performs the factorization of spatial signal structure (embodied in $\mathbf{g}(\Omega)$) from temporal structure of the source (encoded in $f(\Omega)$). Clearly the source can be normalized out in this case, leaving a vector proportional to the Green's function vector that embodies the relative phase and amplitude structure of the signal across the receiving aperture. However, as will be seen later, this would lead to a very large subspace representation for the waveforms, which is undesirable. In addition, it is overkill. Characteristically, the duration of the forcing function is short compared to that of the Green's functions. We can obtain an approximate factorization of the signal into forcing function and path terms with a narrowband representation having characteristic filter lengths that are longer than the source duration, but much shorter than the entire seismogram. The advantage of such a representation is that it can capture slowly-varying details of the spatial structure of the signal (occasioned by the coming and going of different branches of propagation). The timing of such phase arrivals can be preserved, which allows a more specific template for a particular source.

The details of the factorization involve a narrowband decomposition of the received signals with a bank of N filters represented by impulse responses $h_k[n]$:

$r_k[n] = \sum_l r[l] h_k[n-l]; \quad k = 0, \dots, N-1$	(5)
--	-----

The filters can be obtained as center-frequency translations of a prototype baseband filter with impulse responses $h_0[n]$:

$h_k[n] = h_0[n] e^{j\frac{2\pi kn}{N}} = h_0[n] W_N^{kn}$	(6)
--	-----

The discrete-time Fourier transforms of the suite of filters are related simply by:

$H_k(\Omega) = H_0\left(\Omega - \frac{2\pi k}{N}\right)$	(7)
---	-----

If the prototype is implemented as a finite approximation to the ideal lowpass filter:

$H_0(\Omega) = \begin{cases} 1 & \Omega \leq \pi/N \\ 0 & \Omega > \pi/N \end{cases}$	(8)
---	-----

then the suite of filters decomposes the frequency domain into approximately disjoint narrow bands as shown in Figure 5.

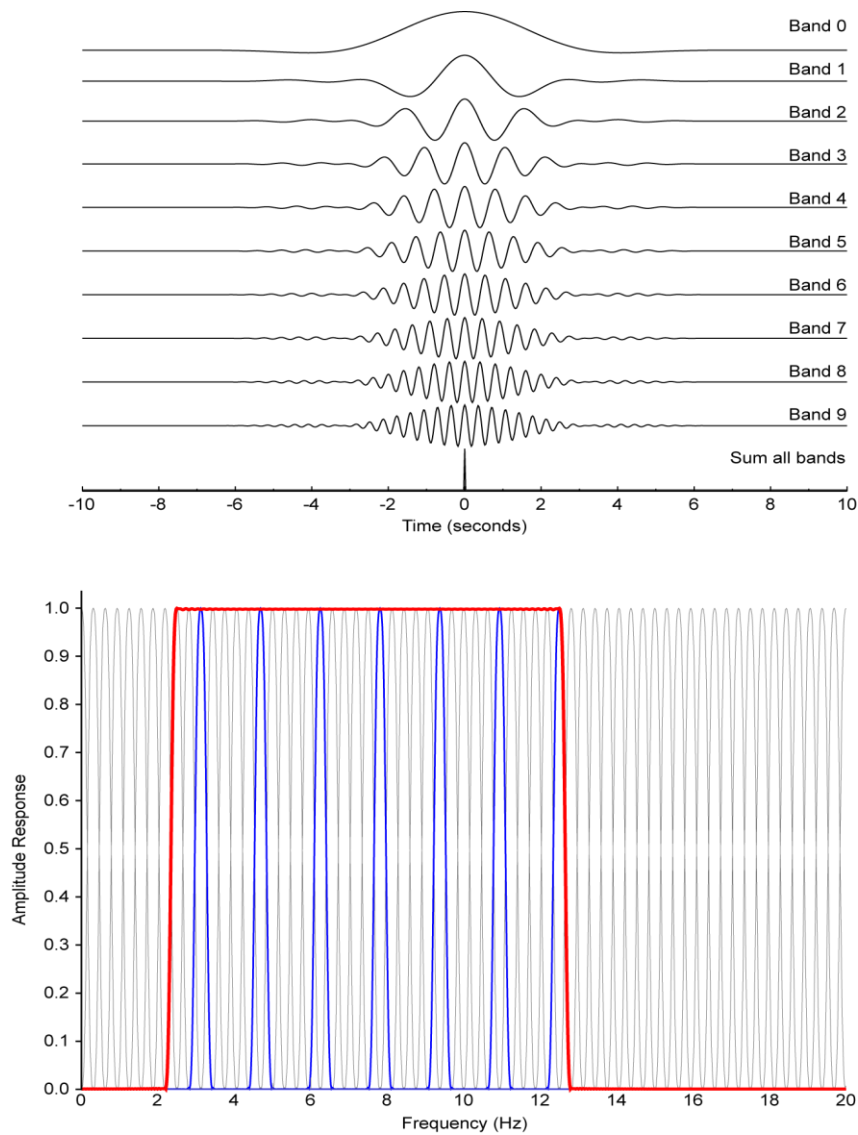


Figure 5: The Real Parts of the First 10 Impulse Responses (Top) and Amplitude Frequency Responses (Bottom) for the Bank of Narrowband Filters Defined in Equations (6-8). Seven Individual Bands Are Highlighted in Blue and the Sum of 33 Bands in the Range 2.5-12.5 Hz Is Shown in Red

Figure 5 shows seven individual bands are highlighted in blue and the sum of 33 bands in the range 2.5-12.5 Hz is shown in red.

Our particular choice of baseband impulse response is a windowed sinc function

$h_0[n] = w[n] \frac{\sin \pi n / N}{\pi n / N}; \quad n = -pN, \dots, 0, \dots, pN$	(9)
--	-----

where $w[n]$ is the zeroth-order prolate spheroidal wave sequence and p is an integer design factor that controls the duration of the impulse response and the degree to which the filter approximates the ideal lowpass response of equation (8). For this choice it is possible to show that

$\sum_{k=0}^{N-1} H_k(\Omega) = 1$	(10)
------------------------------------	------

which implies that the wideband signal can be exactly recovered from its narrowband components:

$\mathbf{r}[n] = \sum_{k=0}^{N-1} \mathbf{r}_k[n]$	(11)
--	------

Now it is apparent that the narrowband components are complex signals. In fact they can be interpreted as slowly-varying complex envelope functions modulated to a center frequency given by $\frac{2\pi k}{N}$. This fact is seen by inserting (6) into (5):

$\mathbf{r}_k[n] = \sum_l \mathbf{r}[l] h_0[n-l] W_N^{k(n-l)}$	(12)
--	------

and performing some simple manipulations:

$\mathbf{r}_k[n] = W_N^{kn} \sum_l (\mathbf{r}[l] h_0[n-l]) W_N^{-kl} = W_N^{kn} \tilde{\mathbf{r}}_k[n]$	(13)
---	------

Here the k^{th} complex envelope is defined by:

$\tilde{\mathbf{r}}_k[n] = \sum_l (\mathbf{r}[l] W_N^{-kl}) h_0[n-l]$	(14)
---	------

and can be interpreted as a demodulation operation performed on the signal (multiplication by the complex exponential, in parentheses) followed by a lowpass filter operation. This combination “snips” out the signal in a narrow band centered at frequency $\frac{2\pi k}{N}$, and removes the modulation to produce the complex envelopes shown in Figure 6. Note that the complex envelope retains relative amplitude and phase information across the array and the relative arrival times of the main phase arrivals.

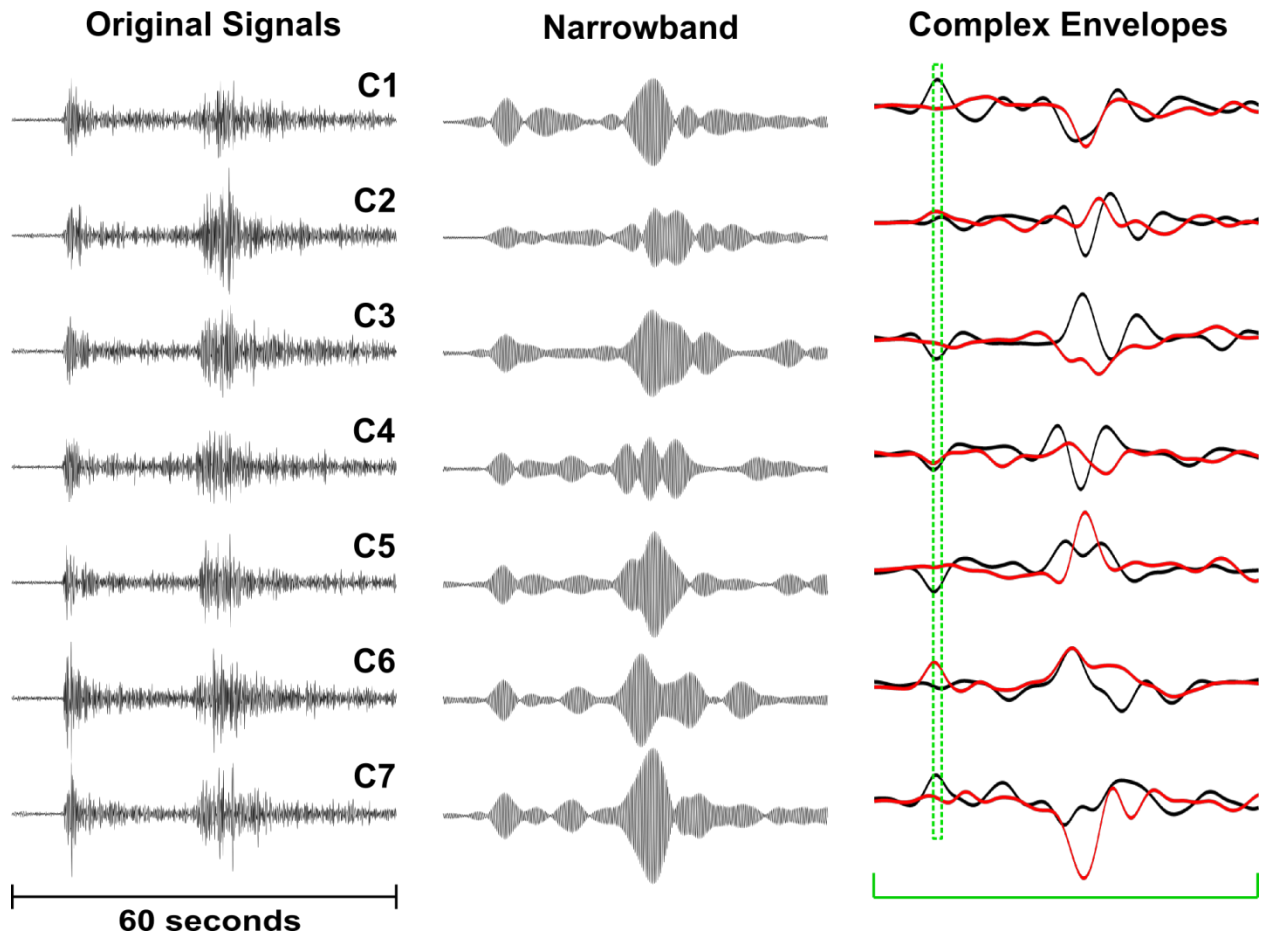


Figure 6: Complex Envelopes Obtained by Filtering the Array Waveforms Into One Particular Narrow Frequency Band – Left, the Original Signals; Center, the Waveforms Filtered Into a Band 0.3125 Hz Wide; Right, the Corresponding Complex Envelopes After Demodulation to Remove the High-Frequency Carrier

In Figure 6, the real part of the envelope is shown in black and the imaginary part is shown in red. The matched field template used for classification in our last contract used only a single sample of the complex envelope around P_n as indicated by the dashed green rectangle. The detection framework we developed as part of this contract uses the entire waveform, indicated by the green bracket.

We asserted earlier that, for short-duration sources, the complex envelopes approximately factor the spatial and temporal structure of the signal. To demonstrate this fact, we first observe that the source forcing function $f[n]$ of equation (3) can be written in terms of complex envelopes:

$f[n] = \sum_{k=0}^{N-1} W_N^{kn} \tilde{f}_k[n]$ $\tilde{f}_k[n] = \sum_l (f[l] W_N^{-kl}) h_0[n-l]$	(15)
---	------

Since we have assumed that the duration of the source is much shorter than the impulse response of the filters $h_0[n]$, we can make the approximation:

$\tilde{f}_k[n] \approx h_0[n] \sum_l (f[l] W_N^{-kl})$	(16)
---	------

Since $h_0[0] = 1$,

$\tilde{f}_k[n] \approx h_0[n] \tilde{f}_k[0]$	(17)
--	------

Inserting this approximation into (15):

$f[n] \approx h_0[n] \sum_{k=0}^{N-1} W_N^{kn} \tilde{f}_k[0]$	(18)
--	------

and the result into equation (3):

$\mathbf{r}[n] \approx \sum_{k=0}^{N-1} \tilde{f}_k[0] \sum_l \mathbf{g}[n-l] W_N^{kl} h_0[l]$	(19)
--	------

Making the substitution $m = n - l$:

$\mathbf{r}[n] \approx \sum_{k=0}^{N-1} \tilde{f}_k[0] \sum_m \mathbf{g}[m] W_N^{k(n-m)} h_0[n-m]$	(20)
--	------

And reorganizing terms:

$\mathbf{r}[n] \approx \sum_{k=0}^{N-1} W_N^{kn} \tilde{f}_k[0] \left\{ \sum_m \mathbf{g}[m] W_N^{-km} h_0[n-m] \right\}$	(21)
---	------

Observe that the term in brackets is $\tilde{\mathbf{g}}_k[n]$:

$\mathbf{r}[n] \approx \sum_{k=0}^{N-1} W_N^{kn} \tilde{f}_k[0] \tilde{\mathbf{g}}_k[n]$	(22)
--	------

Comparing this result to equations (11) and (13), we can make the identification:

$\tilde{\mathbf{r}}_k[n] \approx \tilde{f}_k[0] \tilde{\mathbf{g}}_k[n]$	(23)
--	------

which is the result we seek. Equation (23) is the narrowband equivalent of equation (4).

Stochastic Source

We think of each source of explosions as a stochastic process, periodically producing waveforms which are sample functions of the process. To be precise, we consider the forcing function (source time history) f to be the stochastic process and the Green's functions to be deterministic, but unknown. We seek to estimate the Green's functions, at least up to functions proportional to their narrowband envelopes. As described in the introduction, several types of sources have been identified characterized by the degree of correlation across frequency. We will characterize the process in terms of moments, with the mean being zero:

$E\{\tilde{f}_k[0]\} = 0$	(24)
---------------------------	------

and the second moment:

$E\{\tilde{f}_k[0] \tilde{f}_l^*[0]\} = \gamma_{kl}$	(25)
--	------

A source that is uncorrelated across frequency bands is characterized by:

$\gamma_{kl} = \delta_{kl} \sigma_k^2$	(26)
--	------

where δ_{kl} denotes the Kronecker delta function. By contrast, a perfectly correlated signal exhibits the separable correlation:

$\gamma_{kl} = \sigma_k \sigma_l.$	(27)
------------------------------------	------

Extending these observations to the complex envelopes, we assume the envelopes to be zero mean:

$E\{\tilde{\mathbf{r}}_k[n]\} = 0$	(28)
------------------------------------	------

From equation (23), they have covariance functions given by:

$E\{\tilde{\mathbf{r}}_k[m] \tilde{\mathbf{r}}_l^H[n]\} = \tilde{\mathbf{C}}_{kl}[m, n] = \gamma_{kl} \tilde{\mathbf{g}}_k[m] \tilde{\mathbf{g}}_l^H[n]$	(29)
--	------

To examine the covariance of all of the complex envelopes simultaneously, we assume that a finite number M of samples of the envelopes are computed, corresponding to the finite duration of the seismograms, and assemble these into a vector for one band:

$\tilde{\mathbf{r}}_k = \begin{bmatrix} \tilde{\mathbf{r}}_k[0] \\ \tilde{\mathbf{r}}_k[1] \\ \vdots \\ \tilde{\mathbf{r}}_k[M-1] \end{bmatrix}$	(30)
--	------

Similarly, we assemble the vectors for all bands into one very large summary vector:

$\tilde{\mathbf{r}} = \begin{bmatrix} \tilde{\mathbf{r}}_{k_{min}} \\ \tilde{\mathbf{r}}_{k_{min}+1} \\ \vdots \\ \tilde{\mathbf{r}}_{k_{max}} \end{bmatrix}$	(31)
---	------

Note that we have elected to retain a subset of bands $[k_{min}, \dots, k_{max}]$ which may be fewer in number than N , the total number of bands in the narrowband decomposition. We denote the total number of bands used in processing the data by $N_b = k_{max} - k_{min} + 1$. By construction, $\tilde{\mathbf{r}}$ is zero-mean, and its covariance is:

$\tilde{\mathbf{C}} = E\{\tilde{\mathbf{r}} \tilde{\mathbf{r}}^H\}$	(32)
---	------

The covariance matrix $\tilde{\mathbf{C}}$ will have special structure depending upon the statistical characteristics of the source. A source that is uncorrelated across frequency bands will have a block diagonal structure, with blocks:

$\tilde{\mathbf{C}}_{kl} = E\{\tilde{\mathbf{r}}_k \tilde{\mathbf{r}}_l^H\} = \delta_{kl} \sigma_k^2 \begin{bmatrix} \tilde{\mathbf{g}}_k[0] \\ \tilde{\mathbf{g}}_k[1] \\ \vdots \\ \tilde{\mathbf{g}}_k[M-1] \end{bmatrix} [\tilde{\mathbf{g}}_l^H[0] \quad \tilde{\mathbf{g}}_l^H[1] \quad \dots \quad \tilde{\mathbf{g}}_l^H[M-1]]$	(33)
---	------

Similarly, a source perfectly correlated across frequency bands has the following block structure:

$\tilde{\mathbf{C}}_{kl} = \sigma_k \begin{bmatrix} \tilde{\mathbf{g}}_k[0] \\ \tilde{\mathbf{g}}_k[1] \\ \vdots \\ \tilde{\mathbf{g}}_k[M-1] \end{bmatrix} \sigma_l [\tilde{\mathbf{g}}_l^H[0] \quad \tilde{\mathbf{g}}_l^H[1] \quad \dots \quad \tilde{\mathbf{g}}_l^H[M-1]]$	(34)
---	------

These are theoretical covariance structures, taken as an expectation over all sample functions of the random process. A finite-sample approximation to the covariance matrix can be obtained by replacing the expectation of equation 32 with a finite sum of outer products. This operation can be summarized with a data matrix of envelope vectors constructed from N_s sample events:

$\tilde{\mathbf{R}} = [\tilde{\mathbf{r}}^1 \ \dots \ \tilde{\mathbf{r}}^{N_s}]$	(35)
--	------

After removing the sample mean from the vectors, the covariance may be estimated by:

$\tilde{\mathbf{C}} \sim \tilde{\mathbf{R}} \tilde{\mathbf{R}}^H$	(36)
---	------

Sample covariance matrices are rendered as images in Figure 7 for the Finnish munition explosions and Kirovsk mine explosions shown in Figure 4. The Finnish ensemble consisted of 115 explosions and the Kirovsk ensemble consisted of 114; Figure 4 shows just the first 10 for each source. Four channels of the ARCES array were used in forming the complex envelope vectors, five frequency bands are shown (center frequencies 2.5, 5.0, 7.5, 10.0, 12.5 Hz; bandwidth 0.3125 Hz) and approximately 40 seconds of data were used for the Finnish explosions and 80 seconds for the Kirovsk explosions. Since the complex envelopes are extremely narrowband, the envelopes were heavily decimated before constructing the data matrices. We note that the Finnish source covariance approximates equation (34) and the Kirovsk source covariance approximates equation (33). Both are actually between the extremes, with the first approaching the coherent end member and the second approaching the incoherent extreme.

The eigenstructure of the covariance matrix should reveal the nature of the source. In the incoherent source case (equation 33), the number of non-zero eigenvalues should equal the number of frequency bands. In the perfectly coherent case, the number of non-zero eigenvalues should be one. Figure 8 shows the eigenspectra for the Finnish munition and Kirovsk sources. As expected, the eigenspectrum energy for the Finnish source is heavily concentrated in a single eigenvalue, due to the high repeatability of the waveforms. The eigenspectrum for the Kirovsk explosions is much more diffuse. In part, this effect is due to the variation of the source, but it also is due to the method of waveform alignment, which was based on catalog origin times. The Finnish explosions were aligned to trigger times made with a correlation detector. It is likely that the concentration of the Kirovsk eigenspectrum could be improved somewhat with correlation alignment.

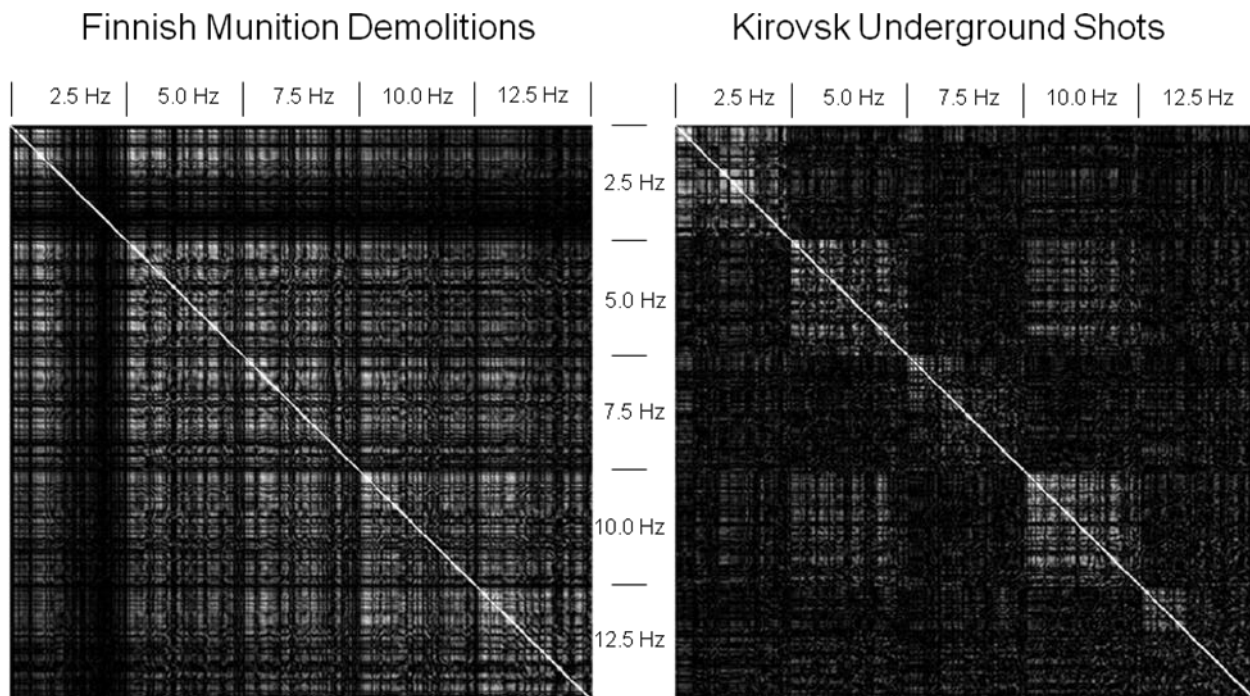


Figure 7: Sample Complex Envelope Covariance Matrices (Equation 36) for Ensembles of Events from Two Sources Approximating a Source Fully Correlated Across Frequency Bands (the Finnish Munitions Demolition Site, Left) and Uncorrelated Across Frequency Bands (the Kirovsk Underground Mine, Right)

In figure 7, all matrix elements have been normalized by the product of square roots of the diagonal elements in the same row and column. This normalization scales all values to range between 0 and 1 in absolute value. The absolute value of the results are rendered in grey scale with white being one and black zero. The Finnish explosions exhibit higher off-diagonal covariance showing that they derive from what we call an approximately coherent source. The covariance of the Kirovsk explosions approximates a block diagonal structure, suggesting that their source is approximately incoherent.

Detection Framework

The development of matched field processing templates for the entire waveform extends the much simpler approach developed for a single phase. In that approach, we estimated a collection of independent templates in the selected processing frequency bands. A template component was extracted as the principal eigenvector of the covariance matrix in each frequency band. We made no attempt to link the templates across frequency bands, which is required to extend the matching scheme to incorporate waveform correlation in addition to an incoherent matched field processor. In our current scheme, we again use the eigendecomposition of a covariance matrix (equation 36 in this case) to derive the matching template. Because the covariance matrix

incorporates all frequency bands, the eigendecomposition provides the mechanism to extract a waveform correlator or an incoherent matched field processor (or something in between) depending on the nature of the eigenspectrum.

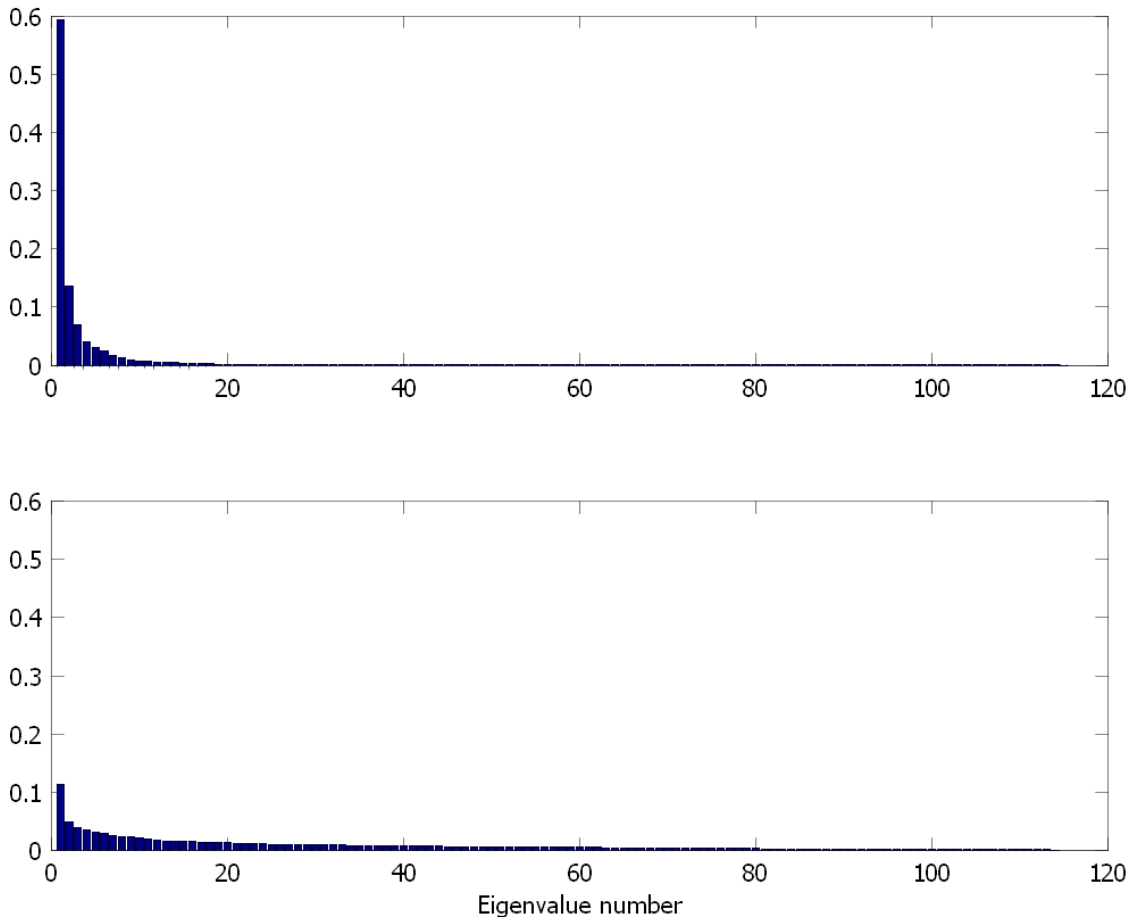


Figure 8 Distribution of Eigenvalues for the Finnish (Top) and the Kirovsk (Bottom) Explosion Ensembles

As seen in Figure 8, the total energy of the eigenspectrum has been normalized to one in each case. Better concentration of the Kirovsk explosion eigenvalues probably can be achieved with improved alignment of the waveforms

If the eigenspectrum energy is heavily concentrated in a single eigenvalue, then the corresponding eigenvector may form the template. In that case the phase structures of the template components in all frequency bands are locked together and a frequency-coherent processor (waveform correlator) emerges. If, instead, the eigenspectrum energy is dispersed over many eigenvalues, then the appropriate template is multidimensional, consisting of the eigenvectors corresponding to the dominant eigenvalues (however they are defined). In that case the processor that emerges is a higher-rank subspace projection operator.

Let the eigendecomposition of the covariance matrix be defined by:

$\tilde{\mathbf{C}} = \tilde{\mathbf{E}} \tilde{\mathbf{\Lambda}} \tilde{\mathbf{E}}^H = \begin{bmatrix} \tilde{\mathbf{\epsilon}}^1 & \tilde{\mathbf{\epsilon}}^2 & \dots & \tilde{\mathbf{\epsilon}}^{N_\epsilon} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_{N_\epsilon} \end{bmatrix} \begin{bmatrix} (\tilde{\mathbf{\epsilon}}^1)^H \\ (\tilde{\mathbf{\epsilon}}^2)^H \\ \vdots \\ (\tilde{\mathbf{\epsilon}}^{N_\epsilon})^H \end{bmatrix}$	(37)
--	------

Note that a sample covariance matrix will not be full rank, as the number of events N_ϵ comprising the ensemble is likely to be in the hundreds at most. The size of the matrix is $D \times D$, where $D = N_b \cdot N_c \cdot M$. Typical numbers for a practical implementation are: $N_b = 33$, $N_c = 17$, $M = 200$, for a total $D \sim 10^5$. It is impractical actually to form the covariance matrix and compute its eigendecomposition directly. Fortunately, the eigendecomposition can be obtained through a singular value decomposition of the data matrix (equation 35) at a reasonable computational cost.

An appropriate dimension d of the template can be determined by an energy ‘‘capture’’ kind of measurement, i.e. identifying the minimum number of eigenvalues required to represent a given fraction θ of the total eigenspectrum energy. Since the total energy is the sum of the eigenvalues, the objective is to find the integer d for which the following criterion is met:

$\frac{\sum_{i=1}^d \lambda_i}{\sum_{i=1}^{N_\epsilon} \lambda_i} > \theta$	(38)
---	------

This criterion assumes, of course, that the eigenvalues have been sorted into descending order, as shown in Figure 8. The corresponding template would be defined as the matrix of d top eigenvectors, $\tilde{\mathbf{E}}_d = [\tilde{\mathbf{\epsilon}}^1 \ \tilde{\mathbf{\epsilon}}^2 \ \dots \ \tilde{\mathbf{\epsilon}}^d]$. Note that, by construction, $\tilde{\mathbf{E}}_d^H \tilde{\mathbf{E}}_d = \mathbf{1}$.

Once the template is defined, it remains to develop a matching operation suitable for classifying or detecting new signals. The straightforward approach is to select a data window surrounding the event or detection window to be tested, compute the suite of complex envelopes, assemble them into a data vector $\tilde{\mathbf{r}}$ as in equation (31), then compute the square of the norm of the projection of $\tilde{\mathbf{r}}$ into the template subspace:

$S = \frac{ \tilde{\mathbf{E}}_d^H \tilde{\mathbf{r}} ^2}{\tilde{\mathbf{r}}^H \tilde{\mathbf{r}}}$	(39)
---	------

The statistic can be normalized as indicated so that it ranges between 0 and 1.

However, this statistic really is not desirable for purposes of detection because it is costly to compute. An alternative is to reconstruct a wideband projection operator from the narrowband eigenvectors comprising $\tilde{\mathbf{E}}_d$, then implement it in a subspace detector. This approach can be

shown to be equivalent to the algorithm implied by equation (39) in the limit that p in equation (9) grows very large. This is the approach that we adopt and describe next.

First we consider how to interpret the eigenvectors in terms of their narrowband components, essentially reversing equations (31) and (30). An individual eigenvector $\tilde{\mathbf{\epsilon}}^i$ can be partitioned into its constituent frequency bands (reversing 31):

$$\tilde{\mathbf{\epsilon}}^i = \begin{bmatrix} \tilde{\mathbf{\epsilon}}_{k_{\min}}^i \\ \tilde{\mathbf{\epsilon}}_{k_{\min}+1}^i \\ \vdots \\ \tilde{\mathbf{\epsilon}}_{k_{\max}}^i \end{bmatrix} \quad (40)$$

Each narrowband vector trace can be further partitioned into its individual samples (reversing 30):

$$\tilde{\mathbf{\epsilon}}_k^i = \begin{bmatrix} \tilde{\mathbf{\epsilon}}_k^i[0] \\ \tilde{\mathbf{\epsilon}}_k^i[1] \\ \vdots \\ \tilde{\mathbf{\epsilon}}_k^i[M-1] \end{bmatrix} \quad (41)$$

Following equations (11) and (13), the wideband template can be reconstructed as:

$$\mathbf{\epsilon}^i[n] = \sum_{k=k_{\min}}^{k_{\max}} W_N^{kn} \tilde{\mathbf{\epsilon}}_k^i[n]; \quad i = 1, \dots, d \quad (42)$$

If the data are filtered into the same collection of bands as the template:

$$\bar{\mathbf{r}}[n] = \sum_{k=k_{\min}}^{k_{\max}} W_N^{kn} \tilde{\mathbf{r}}_k[n] \quad (43)$$

then the appropriate matching statistic is:

$$S = \sum_{i=1}^d \left| \sum_{m=0}^{M-1} (\mathbf{\epsilon}^i[m])^H \bar{\mathbf{r}}[m] \right|^2 / \sum_{m=0}^{M-1} \bar{\mathbf{r}}^H[m] \bar{\mathbf{r}}[m] \quad (44)$$

This statistic could be used for purposes of event classification on a fixed event window, or for detection on a window sliding over the data.

When used as a detector the numerator and denominator in equation (44) are calculated repeatedly as the window defining $\bar{\mathbf{r}}$ slides over a continuous data stream. The numerator term

$$\sum_{m=0}^{M-1} (\boldsymbol{\epsilon}^i[m])^H \bar{\mathbf{r}}[m]$$

can be implemented efficiently as a complex multichannel correlator, and the full numerator as the incoherent combination of a bank of such correlators. Figure 9 provides a high-level block diagram of the resulting processor. One additional detail: the summation of equation 42 over a limited number of frequency bands results in template components $\boldsymbol{\epsilon}^i$ which are not orthogonal. The subspace detector implementation requires that the individual dimensions of the template be mutually orthogonal:

$\sum_{m=0}^{M-1} (\boldsymbol{\epsilon}^i[m])^H \boldsymbol{\epsilon}^l[m] = \delta_{il}$	(45)
--	------

This requirement necessitates a reorthogonalization step which can be performed with either a QR or a singular value decomposition of the matrix:

$\mathbf{E} = \begin{bmatrix} \boldsymbol{\epsilon}^1[0] & \boldsymbol{\epsilon}^2[0] & \cdots & \boldsymbol{\epsilon}^d[0] \\ \boldsymbol{\epsilon}^1[1] & \boldsymbol{\epsilon}^2[1] & \cdots & \boldsymbol{\epsilon}^d[1] \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\epsilon}^1[M-1] & \boldsymbol{\epsilon}^2[M-1] & \cdots & \boldsymbol{\epsilon}^d[M-1] \end{bmatrix}$	(46)
---	------

If data from only one event are available to characterize a source, yet there is good reason to expect that the source is incoherent over frequency bands, incoherence can be forced. The individual bands can be treated as being independent, which is what we did for the single-phase matched field processor described in the introduction. This approach is useful in developing an incremental approach to designing a matched field detector, where the detector somehow has to be initialized with waveforms from a single event. Incremental design will be described in detail in a later section.

Forcing incoherence with data from a single observation can be accomplished by replacing the data matrix of equation (35) with:

$\tilde{\mathbf{R}} = \begin{bmatrix} \tilde{\mathbf{r}}_{k_{min}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{r}}_{k_{min}+1} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \tilde{\mathbf{r}}_{k_{max}} \end{bmatrix}$	(47)
---	------

This matrix is orthogonal by construction and has rank N_b . The SVD factors can be computed trivially.

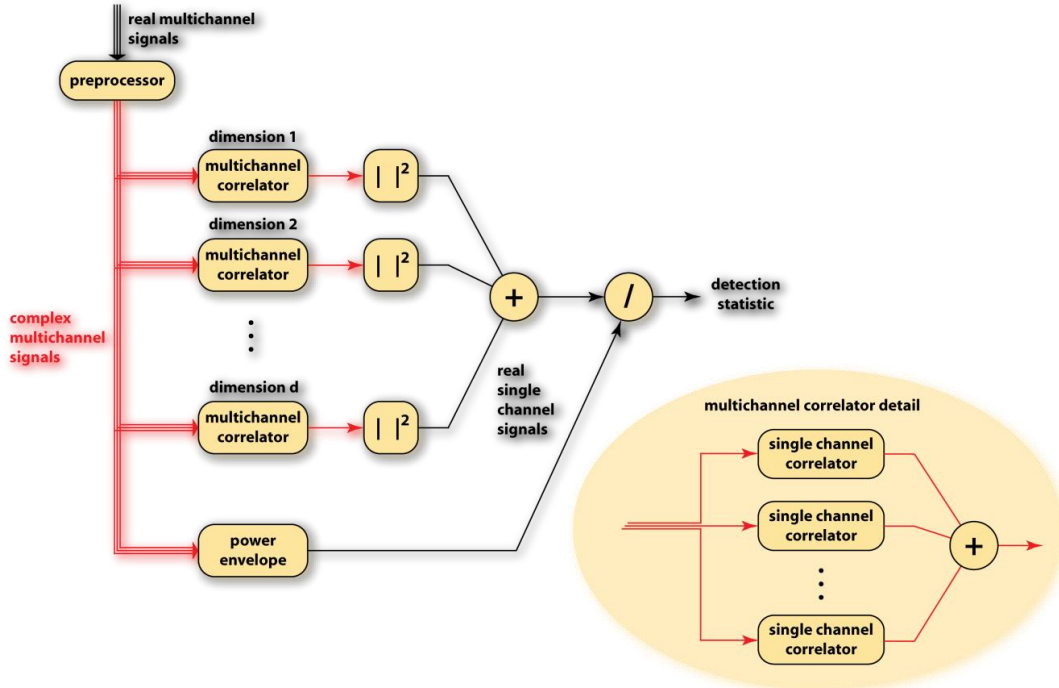


Figure 9: Structure of the Subspace Implementation of the Detection Framework

As seen in Figure 9, a matched field detector is implemented by initializing each multichannel correlator with a different narrowband template (representing d different bands) obtained from a single master event. Note that the outputs of the multichannel correlators are combined incoherently. A correlation detector is obtained by initializing a single wideband correlator with the wideband template of a single master event.

Examples

Three examples exemplify the range of implementations provided by the framework we describe. We draw data for the examples from ARCES array observations of one of the Finnish munitions explosions (Figure 10) to develop templates and apply the resulting detectors to a four hour segment of data (Figure 11) containing two Finnish munitions demolitions. The first is the event used to develop the template, so should provide a perfect match against the detectors. In these examples, we use four channels of data from the ARCES array (ARA0, ARD1, ARD4, ARD7), and processing parameters: $N = 128, k_{min} = 8, k_{max} = 40, N_b = 33$. Since the data

are sampled at 40 samples per second, the bandwidth of each of the 33 narrow bands is 0.3125 Hz. The first band is centered at 2.5 Hz and the highest frequency band is centered at 12.5 Hz.

Figures 11 and 12 contrast the detection statistics calculated by the fully incoherent detector defined by the template of equation 47, for which all 33 bands are treated independently, and the fully coherent detector defined by equation (35) with a single column (event). Both detectors have been implemented via the wideband template constructed with equation (42) with subsequent reorthogonalization. Note that in Figure 12 the noise floor is much higher for the incoherent detector, as is expected from the fact that the constituent templates in the 33 frequency bands act independently to match the data (and thus can match the background noise more effectively). The coherent processor maintains a fixed phase relationship among the bands which is only matched effectively by the two munitions explosions. Figure 13 shows detail of the two detection statistics around the design event. The incoherent processor has a very broad peak, corresponding to the sinc function of equation (9), consistent with the narrow processing bands (0.3125 Hz). The peak of the coherent processor collapses all energy to a spike, as is expected for a correlation detector.

Note that the detection statistic for the second munitions demolition event (occurring between 12000 and 13000 seconds) is lower than for the first event. This is due to some mismatch between the waveforms of the first and second events. It is known that the explosions take place at different locations within an area with an aperture of approximately 300 meters.

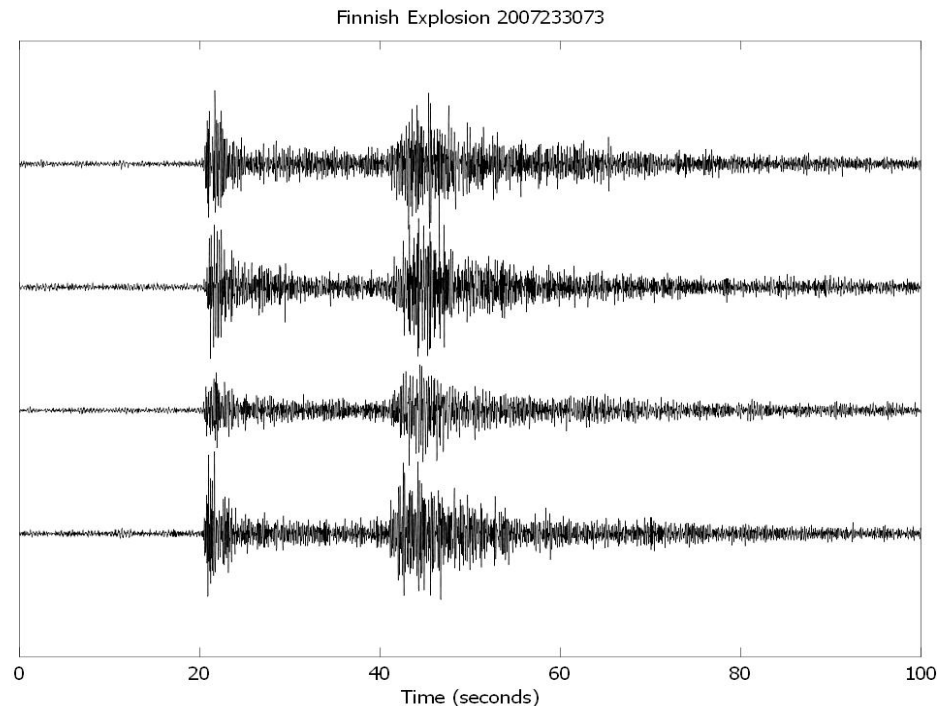


Figure 10: Sample Finnish Munitions Explosion Used to Create Templates – the Data are Filtered Into the 2.5 to 12.5 Hz Band

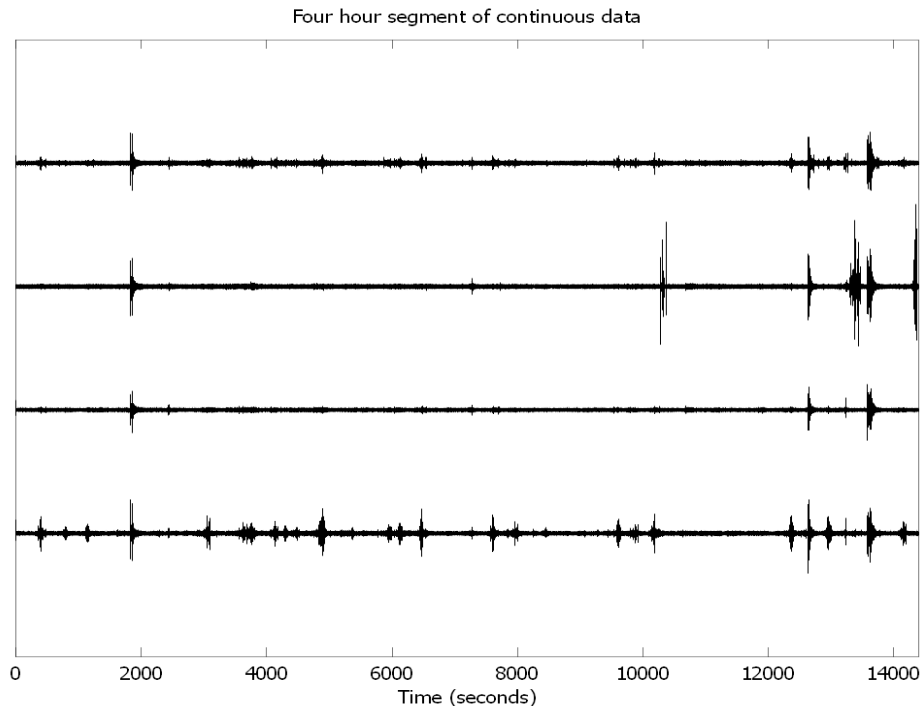


Figure 11: Segment of Data, Used to Test the Detectors, Containing Two Finnish Munitions Explosions

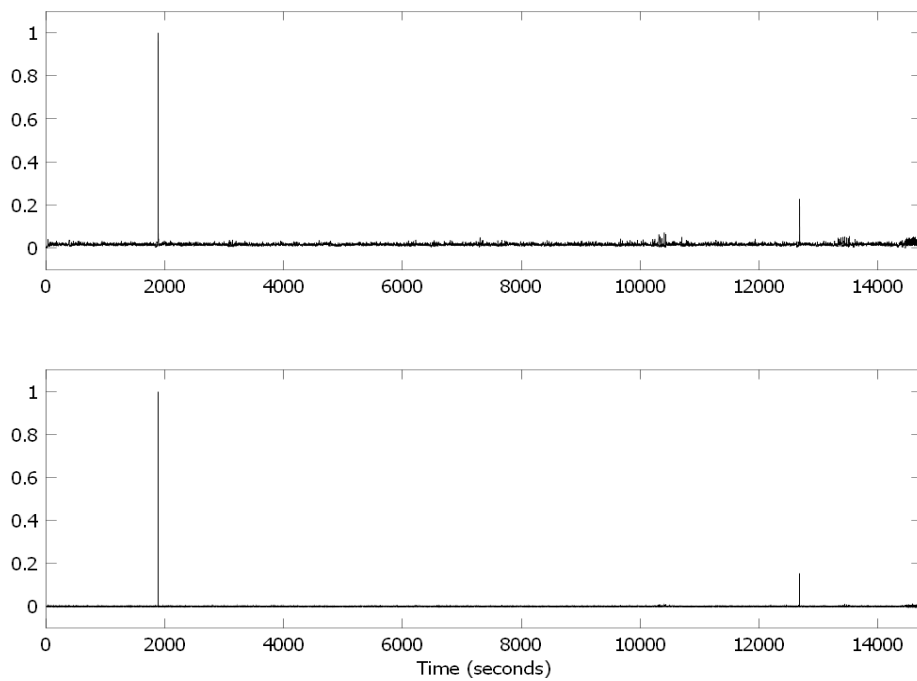


Figure 12: Detection Statistic for Incoherent Detector (Top) and Coherent Detector (Bottom)

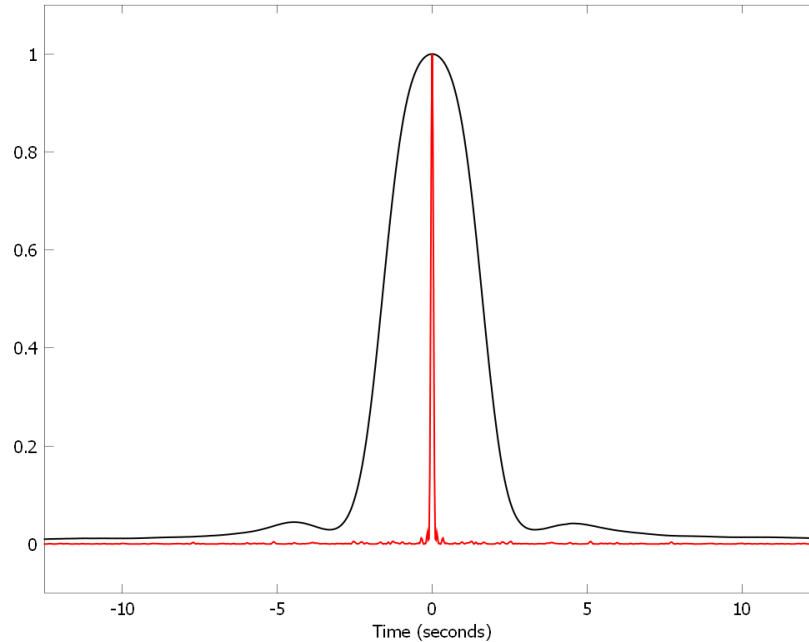


Figure 13: Detail of Detection Statistics for the Incoherent Detector (Black) and the Coherent Detector Around the Design Event, Near 1900 Seconds in Figure 12

A generalization of the correlator may be achieved by designing the template with an ensemble of events rather than an individual event. If we construct the data matrix (equation 35) with all of the events available (115) and estimate the dimension of the template using equation (38) with $\theta = 0.75$, we obtain a rank 3 processor that produces more nearly equal detection statistic values for the two events. The detection statistic of the rank-3 processor is contrasted in Figures 14 and 15 with the correlator statistic. Figure 15 provides the detail for the first event: the detection statistic for the rank-3 template drops a bit from one and spreads out slightly. The reduction in peak value is due to the fact that the template represents all events in the ensemble and does not match the one event quite as precisely as a template designed specifically for it. The slight spread is due to the higher rank of the template, which improves the match over a slightly increased range of time steps.

Adaptation with Time

It is our intention to embed the detection framework that we have described to this point into a real-time system tracking the development of aftershocks and other repeating sources. Consequently, there is a requirement to initialize a detector, perhaps with a single event and then adapt the detector to match source characteristics as more observations accrue. We have developed two options to consider so far: one is to create a correlation detector from the single event and the other is to create an incoherent matched field processor. In principle, the matched field processor is more general, being much higher rank. One might argue, also, that it is maximally non-committal about temporal phase structure, and that this is the correct stance to

take before the nature of the source is known. It could be adopted as an initiator for a source-specific detector along with a rule for updating the template when new events are detected. The objective is to obtain the ensemble result incrementally and in the limit as many events are acquired.

In principle, the source-specific detection template is obtained from the eigendecomposition of the covariance matrix (equation 37). As observed earlier, this approach is impractical due to the very large size of the covariance matrix. Instead, the data matrix (equation 35) can be factored with a singular value decomposition:

$\tilde{\mathbf{R}} = \tilde{\mathbf{U}} \tilde{\mathbf{\Sigma}} \tilde{\mathbf{V}}^H$	(48)
--	------

From equations (36) and (37) we make the identifications $\tilde{\mathbf{U}} = \tilde{\mathbf{E}}$ and $\tilde{\mathbf{\Lambda}} = \tilde{\mathbf{\Sigma}}^2$.

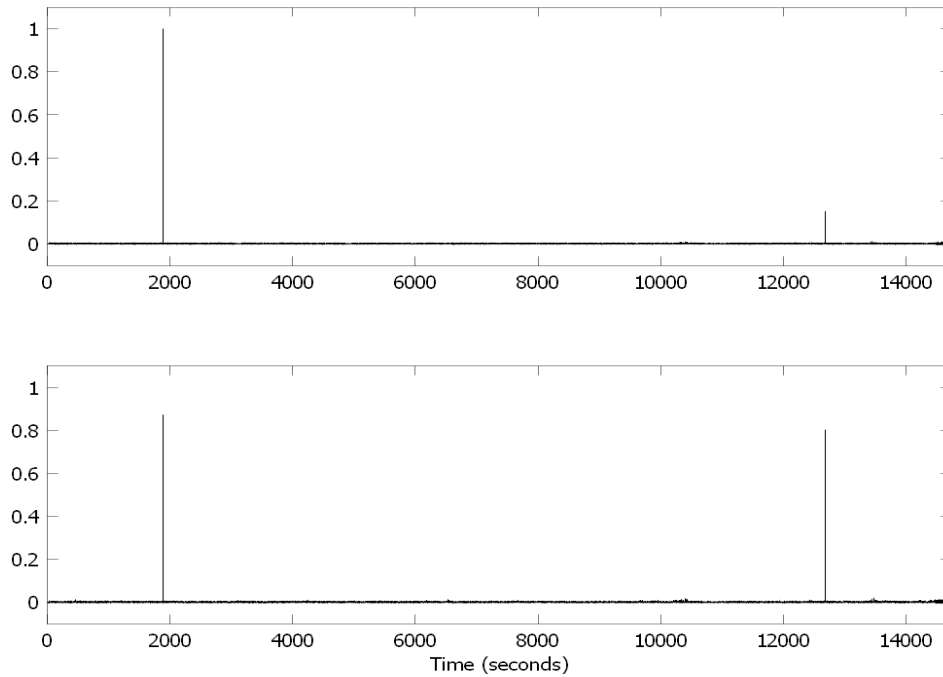


Figure 14: Comparison of Detection Statistics for the Correlator (Top Trace) and the Rank-3 Detector (Bottom Trace) – Note the Rank-3 Detector Increases the Statistic of the Second Event

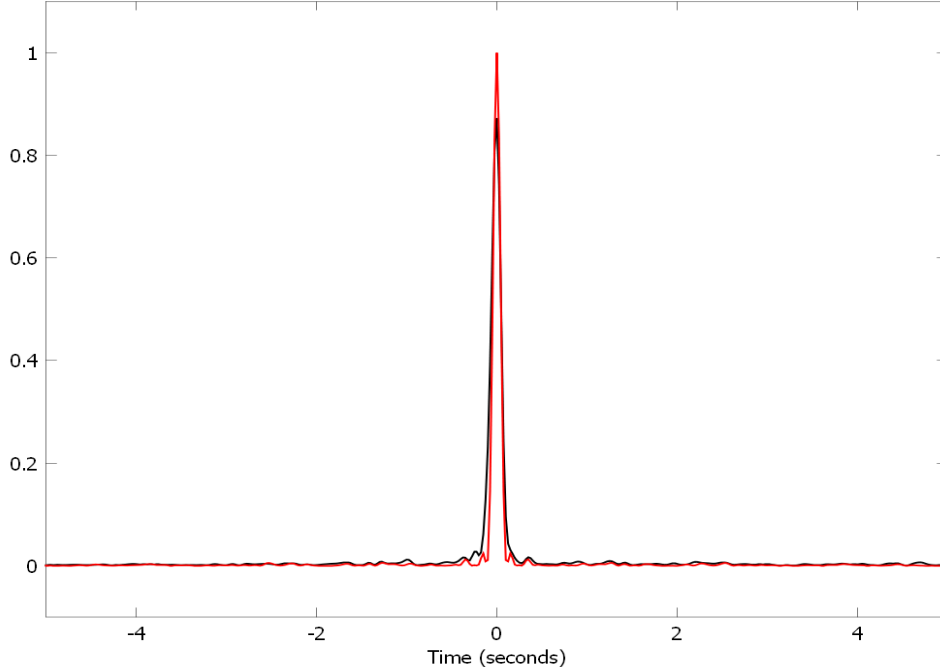


Figure 15: Detail of the Correlator and Rank-3 Detector Statistics Around the First Munitions Explosion

There are many approaches to updating the singular value decomposition (SVD) of a matrix given that the matrix is augmented with the addition of a new column (in our case, the waveform from a new event). Such SVD update algorithms are common in subspace tracking schemes. The general objective is to obtain the SVD of the augmented matrix:

$$\tilde{\mathbf{R}}^+ = [\tilde{\mathbf{r}}^1 \ \dots \ \tilde{\mathbf{r}}^{N_s} \ | \ \tilde{\mathbf{r}}^{N_s+1}] \quad (49)$$

from the SVD of $\tilde{\mathbf{R}}$ without recomputing the SVD from scratch. The specific algorithm that we will use is due to Brand (2006), with minor adaptations by Harris (2008).

We begin by computing the “innovation” in the new data vector with respect to the existing subspace basis:

$$\tilde{\mathbf{c}} = (\mathbf{I} - \tilde{\mathbf{U}}\tilde{\mathbf{U}}^H) \tilde{\mathbf{r}}^{N_s+1} \quad (50)$$

This vector is normalized:

$$\tilde{\mathbf{v}} = \tilde{\mathbf{c}}/|\tilde{\mathbf{c}}| \quad (51)$$

Note that $\tilde{\mathbf{U}}$ augmented by $\tilde{\mathbf{v}}$ is a basis for the column space of the augmented data matrix $\tilde{\mathbf{R}}^+$.

Next define:

$$\tilde{\mathbf{F}} = \begin{bmatrix} \tilde{\mathbf{U}}^H & \tilde{\mathbf{r}}^{N_g+1} \\ \tilde{\mathbf{v}}^H & \tilde{\mathbf{r}}^{N_g+1} \end{bmatrix} \quad (52)$$

and the square $(d+1) \times (d+1)$ matrix:

$$\tilde{\Phi} = \begin{bmatrix} \tilde{\Sigma}^2 & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} + \tilde{\mathbf{F}} \tilde{\mathbf{F}}^H \quad (53)$$

The eigendecomposition of this latter matrix is computed, relatively inexpensively:

$$\tilde{\Phi} = \tilde{\mathbf{G}} \tilde{\Gamma} \tilde{\mathbf{G}}^H \quad (54)$$

Then it can be shown that two of the factors $\tilde{\mathbf{U}}^+$ and $\tilde{\Sigma}^+$ of the SVD we seek (of $\tilde{\mathbf{R}}^+$) are obtained from:

$$\tilde{\mathbf{U}}^+ = [\tilde{\mathbf{U}} \quad \tilde{\mathbf{v}}] \tilde{\mathbf{G}} \quad \tilde{\Sigma}^+ = \tilde{\Gamma}^{1/2} \quad (55)$$

Note that the third factor $\tilde{\mathbf{V}}$ of the SVD of the original matrix never appears and is not used to form the template. Consequently, it does not have to be computed or carried along in the calculation. From (47), the process can be initialized with:

$$\tilde{\mathbf{U}} = \begin{bmatrix} \tilde{\mathbf{r}}_{k_{min}}^1 / |\tilde{\mathbf{r}}_{k_{min}}^1| & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{r}}_{k_{min}+1}^1 / |\tilde{\mathbf{r}}_{k_{min}+1}^1| & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \tilde{\mathbf{r}}_{k_{max}}^1 / |\tilde{\mathbf{r}}_{k_{max}}^1| \end{bmatrix} \quad (56)$$

$$\tilde{\Sigma} = \begin{bmatrix} |\tilde{\mathbf{r}}_{k_{min}}^1| & 0 & \cdots & 0 \\ 0 & |\tilde{\mathbf{r}}_{k_{min}+1}^1| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |\tilde{\mathbf{r}}_{k_{max}}^1| \end{bmatrix}$$

in the case of an incoherent processor and

$$\begin{aligned} \tilde{\mathbf{U}} &= [\tilde{\mathbf{r}}^1] \\ \tilde{\Sigma} &= [|\tilde{\mathbf{r}}^1|] \end{aligned} \quad (57)$$

in the case of a coherent processor. In the example that follows, we examine the incoherent case. We note, in addition, that the dimension of the SVD grows with each iteration. In our implementation, when we begin with an incoherent processor, we crop the SVD to maintain a dimension of N_b from iteration to iteration. By this, we mean that we trim off the lowest eigenvalue and its associated eigenvector.

In order to adjust the rate of adaptation of the template, we use an exponential age-weighting approach. The waveform from the newest event is scaled by an amount ϵ and the previous data matrix is scaled by λ ($\lambda < 1$), i.e. equation (49) is modified to:

$\tilde{\mathbf{R}}^+ = [\lambda \tilde{\mathbf{R}} \mid \epsilon \tilde{\mathbf{r}}^{N_s+1}]$	(58)
--	------

The update equations (49-55) are altered accordingly. The parameters ϵ and λ are adjusted to obtain a suitable rate of adaptation. In our tests, we used $\lambda = 0.95$ and $\epsilon = 0.1$. The use of an exponential forgetting factor of **0.95** means that older events no longer contribute after about 30 iterations.

We make use of another economy in our representation of the complex envelopes. Since they are very narrowband, they are heavily decimated (by a factor of 16 or more). Only the decimated samples are computed and stored throughout the calculations. The complex envelopes are interpolated back up to the original sampling rate when the wideband template representation is created (equation 42 with reorthogonalization).

An example of adaptation using explosions from the Finnish munitions demolitions is shown in Figures 16 through 18. In this example, we initialized the template with a single event, using the incoherent template (equation 56), with parameters $N = 128, k_{min} = 8, k_{max} = 40, N_b = 33$ and the same four stations of the ARCES array. We did not run the detector to find the events, but rather obtained the event waveforms from an archive. The point was to exercise the adaptation mechanism (equations 49-55) only to illustrate its behavior. The successive templates were used to construct detectors to process the four-hour data segment of Figure 11. The archive specifically excluded the two events that appear in the four-hour segment. We present the detection statistics calculated from this data segment as a means of observing adaptation. Although we maintained the narrowband template at rank $N_b = 33$ internally for purposes of running the adaptation mechanism, we cropped the wideband template to a lower dimension using equation (38) with $\theta = 0.75$.

The resulting rank of the wideband template as a function of iteration of the update algorithm (equations 49-55) is shown in Figure 16. Note that it takes about 9 or 10 iterations (addition of new events) with the selected parameters for the dimension of the wideband template to change. After 10 iterations, the dimension starts to drop and the decrease accelerates. After about 32 iterations, the rate of change drops significantly, and by 40 iterations the algorithm is reaching its steady state value (with occasion oscillations).

To see what type of detector emerged, it is instructive to examine the detection statistics for the four-hour segment. Figure 17 shows detail for the detection statistics for the initial incoherent detector and for the detector at iterations 20 and 40. The detection statistic for the incoherent detector is a broad peak, as in Figure 13 with the difference that its peak value is lower (a consequence of the reduced rank of the detector estimated with equation 38). After 20 iterations, the detection statistic is an interesting hybrid of the broad peak and the spike of a correlator. We had anticipated that the peak might stay approximately the same height, but become progressively narrower. However, something more complicated occurs. Apparently the template maintains two subspaces, one the subspace of the initial incoherent detector and the second the subspace of an emerging correlator. The balance between the two changes over time, with the initial subspace ultimately fading away. This behavior is more apparent in Figure 18, which shows the detection statistic as a continuum over adaptation steps.

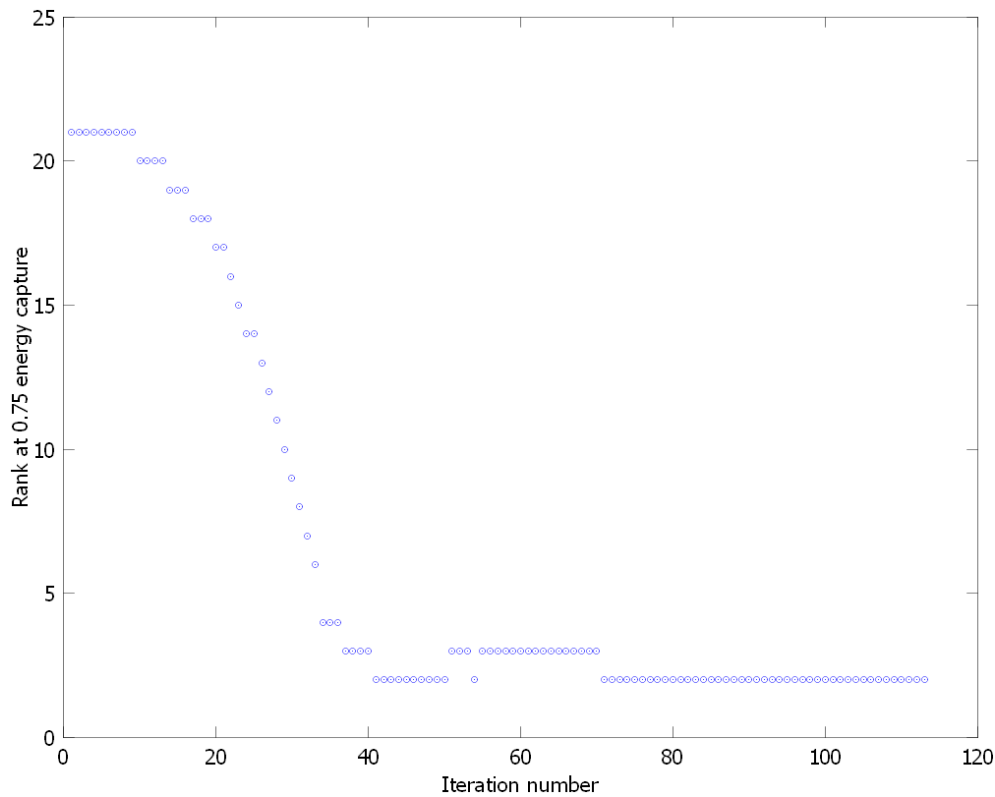


Figure 16: Template Rank as a Function of Iteration Number – the Rank Was Determined from Equation 38 with $\theta = 0.75$

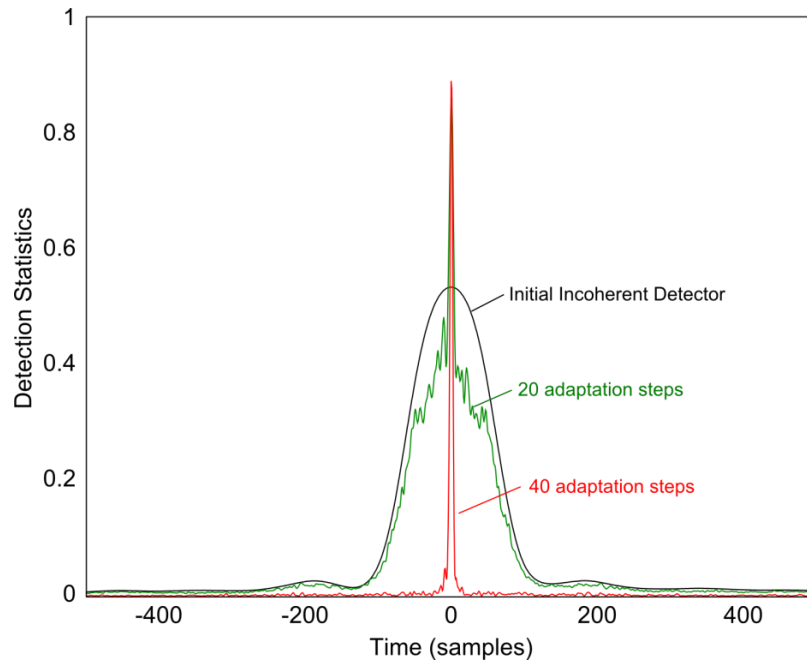


Figure 17: Detection Statistic Around First Finnish Munitions Explosion in Four Hour Test Sequence

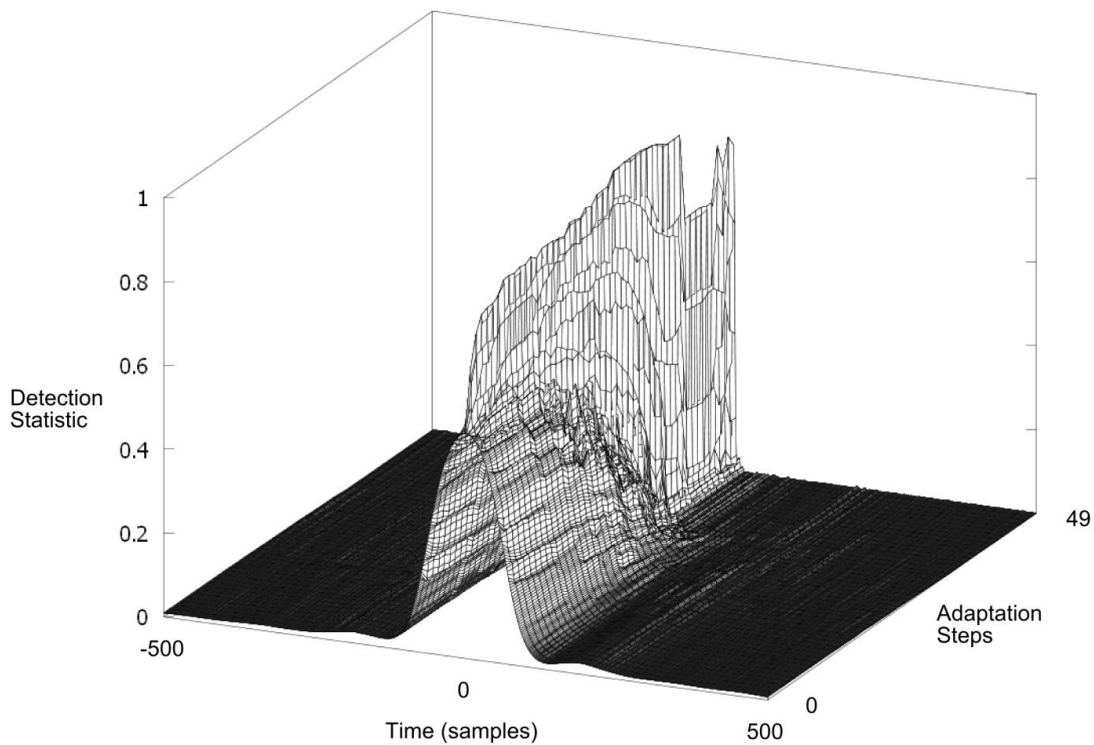


Figure 18: Evolution of the detection statistic around the first Finnish munitions explosion. The time window is 25 seconds in duration. The peak makes a relatively smooth transition from a broad peak to a narrow spike.

3. A Reference Database of Events from the Korean Peninsula

The initial testing of the new full-waveform Matched Field detector has taken place in the European Arctic, exploiting the databases of Ground Truth data collected under previous US contracts. In the current project, the new framework is to be applied to the Korean Peninsula. In order for the new framework to be tested and compared with existing procedures, it is necessary to establish sets of well-constrained events. All of the publicly available bulletins for the region have been obtained, although these in general exclude the majority of small explosive events such as mining blasts and ammunition destructions.

We therefore attempt to bootstrap lists of seismic events from repeating industrial sources using data from the KSRS array in South Korea (see Figure 19). Most of the signals from such events are characterized by high frequency regional phases. KSRS array is a teleseismic array over which signals are fairly incoherent above 4 Hz. This has the consequence that beamforming at the frequencies with the greatest SNR leads to substantial beam-loss, and that slowness estimates in these frequency bands are often qualitatively misleading. In order to detect the high frequency regional phases with a low false alarm rate, we use the spectrogram beamforming (incoherent) method described in Gibbons et al. (2008) - see Figure 20. Despite the typically low SNR in the 2-4 Hz frequency band, the slowness estimates made in this band do appear to be fairly stable. The coherent f-k estimates for the P and S phases displayed in Figure 20 are shown in Figure 21.

A procedure of taking a single event defined by the power detector, running a correlation detector, removing all of the power detections which are confidently associated with these correlation detections, and then spawning a new correlation detector from an event in the pool of remaining detections, is developed and run on KSRS data between November 2006 and February 2010. A set of 19 locations with repeating events were found (see Figure 22 and Table 1), and waveforms for four of these clusters are shown in Figure 23. We see that these four clusters alone show significant variation in both the waveform amplitudes, signal-to-noise ratio and waveform semblance. The clusters were located from KSRS data only, using manually picked P- and S-arrival times and the corresponding backazimuth estimates. However, the locations are subject to uncertainties in the velocity model, deviations of backazimuth estimates from the geographical “true” values, and possible misinterpretation of phases.

The repeating seismic sources include natural seismicity, presumed mining activity, and numerous sites of repeating events of unknown origin.

Figures 24 – 26 display the correlation detections for 19 distinct templates as a function of time for the period November 2006 to February 2010. The plots indicate significant differences in the nature of the repeatability which (a) help to identify the kind of source in question and (b) present different challenges to the Matched Field and correlation detectors.

As seen from Figure 24, some templates resulted in a large number of detections; typically one or two per day, at fairly consistent times of day. The correlation coefficients are typically low, although the detections are usually well validated using the f-k post-processing method of Gibbons and Ringdal (2006). There are frequently a number events with relatively high

correlation coefficients quite close temporally to the master event. All of these characteristics are typical of mining/quarrying activity. The events detected by the template M6 (see Figure 24) show in general a greater degree of similarity than typical ripple-fired quarry blasts, and are more sparse in time than is typical. However, the regularity of the time-of-day and the near-periodicity in their occurrence means that they are almost certainly anthropogenic events.

Based on the location estimates of each cluster, we attempted to identify the corresponding quarry using Google Earth. We were able to identify a candidate quarry for six of the anthropogenic-type event clusters, and it is on this basis that the events in labelled “probable mining” are selected (see Figures 22, 23, and Table 3.1). A Google Earth view of the quarry likely corresponding to cluster M6 (Sunshon – North Korea) is shown in Figure 27.

Other template detection patterns are very characteristic of earthquake aftershock sequences with events clustered tightly in time (see Figure 25). It is noted that there are usually very few large events in these sequences (i.e. few candidates for templates). We labelled six of the clusters “probable earthquakes” (see Figures 22, 25, and Table 1), and one of these clusters (E5) corresponds to the documented January 2007 Odaesan earthquake sequence described by Kim and Park (2010). There is an excellent correspondence between events in our cluster E5, defined from KSRS data only, and the event list provided by Kim and Park (2010), defined from local network data. This places confidence in our approach for defining the event clusters.

The remaining seven clusters are labelled “unknown” (see Figures 22, 25, and Table 1).

The mining events are likely to test the matched field detector’s ability to make the detection procedure relatively insensitive to the source-time function. The demonstrably low correlation coefficients are problematic for the classical correlation detectors.

The earthquakes present different challenges; representative waveforms may only be available from a very limited number of events generating signals with a good SNR. In addition, the majority of aftershocks may be significantly smaller resulting in a lower SNR and a different spectral content. The spatial extent of the aftershock sequence may also contribute to a lack of correlation with the master event.

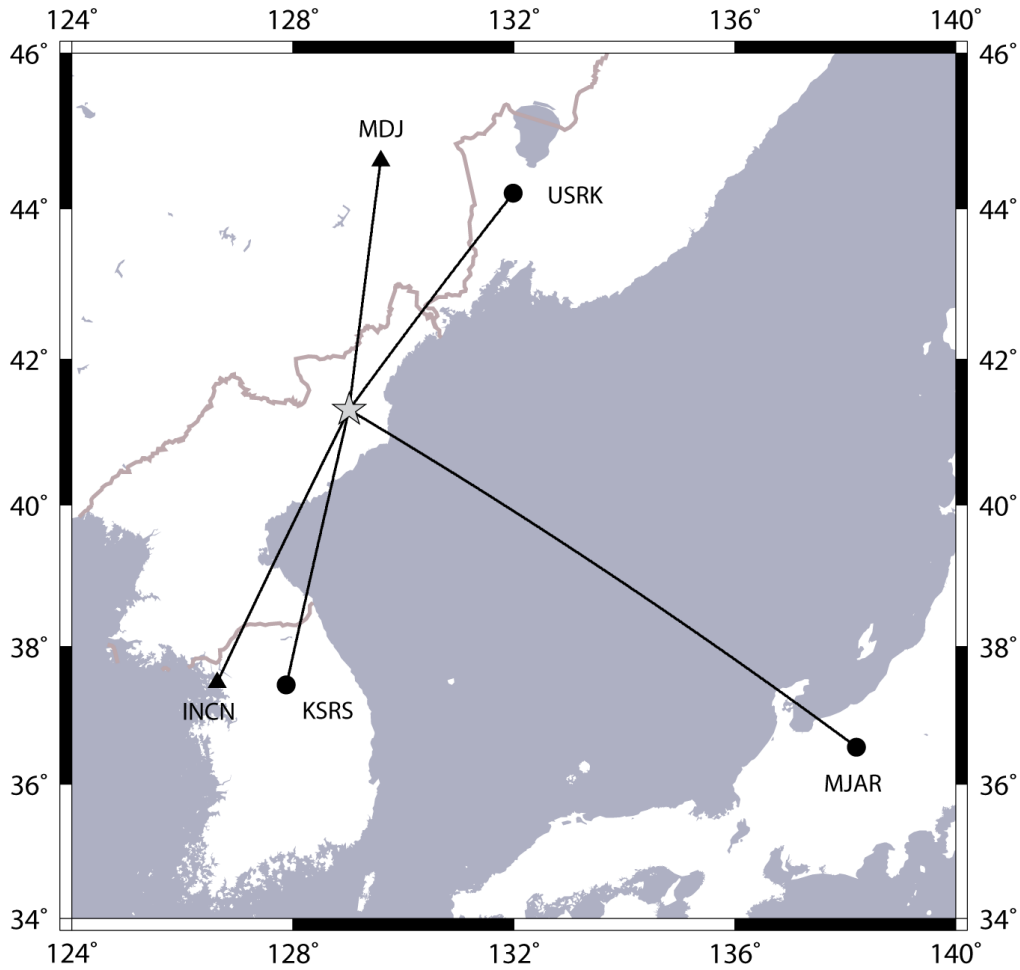


Figure 19: Location of the IMS Arrays KSRS (Wonju, South Korea), USRK (Ussuriysk, Russian Federation), and MJAR (Matsushiro, Japan) and the IRIS 3-Component Stations INCN (IU Network) and MDJ (IC Network) – the Star Denotes the Location of the DPRK Nuclear Test Site

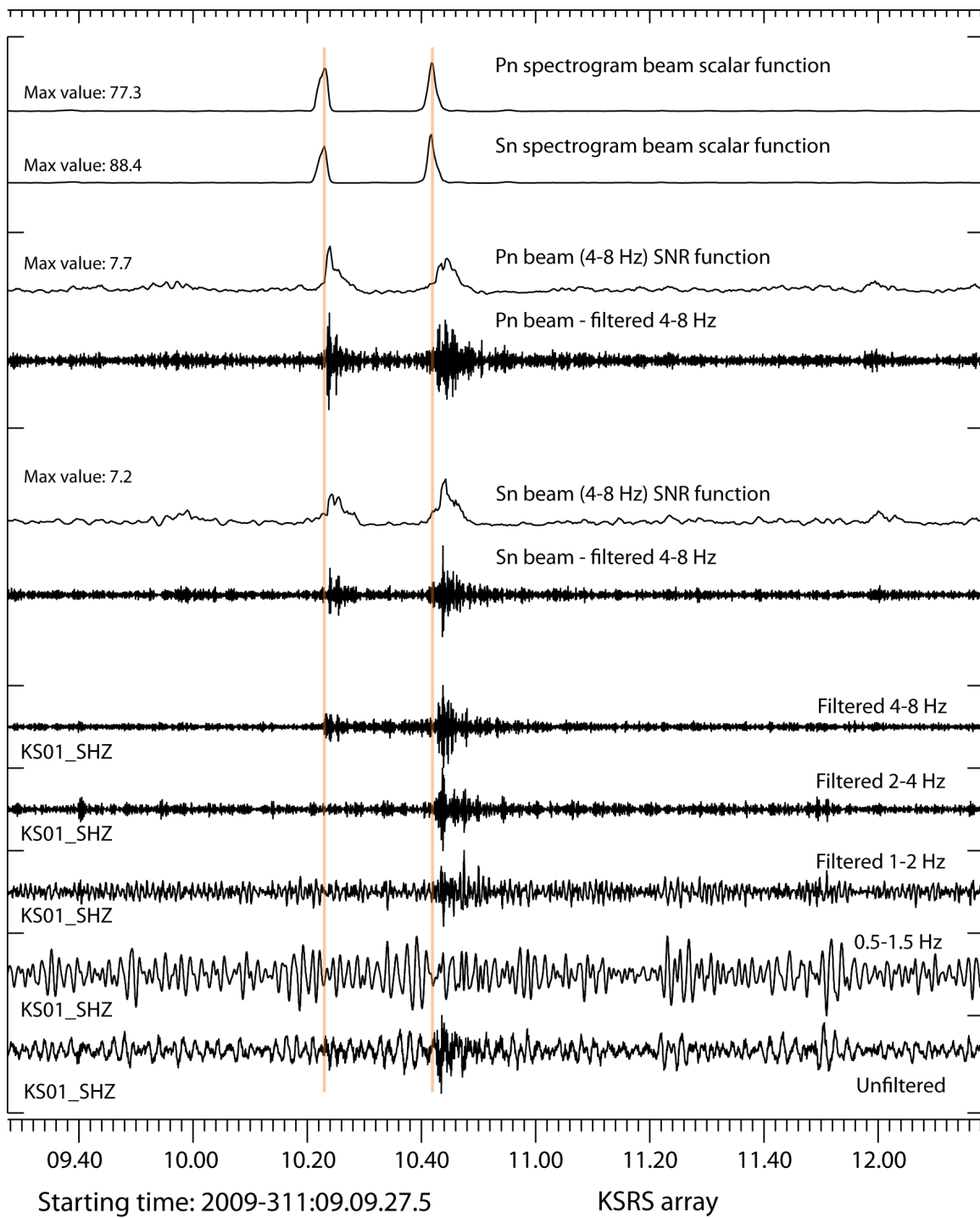


Figure 20: An Example of the Detection of a P and S Phase for an Event Within Regional Distances of the KSRS Array

As seen in Figure 20, the only frequency band at which a reasonable SNR is attained is the 4-8 Hz band. The beamforming of transformed spectrograms described by Gibbons et al. (2008) provides a promising detection statistic for these phases.

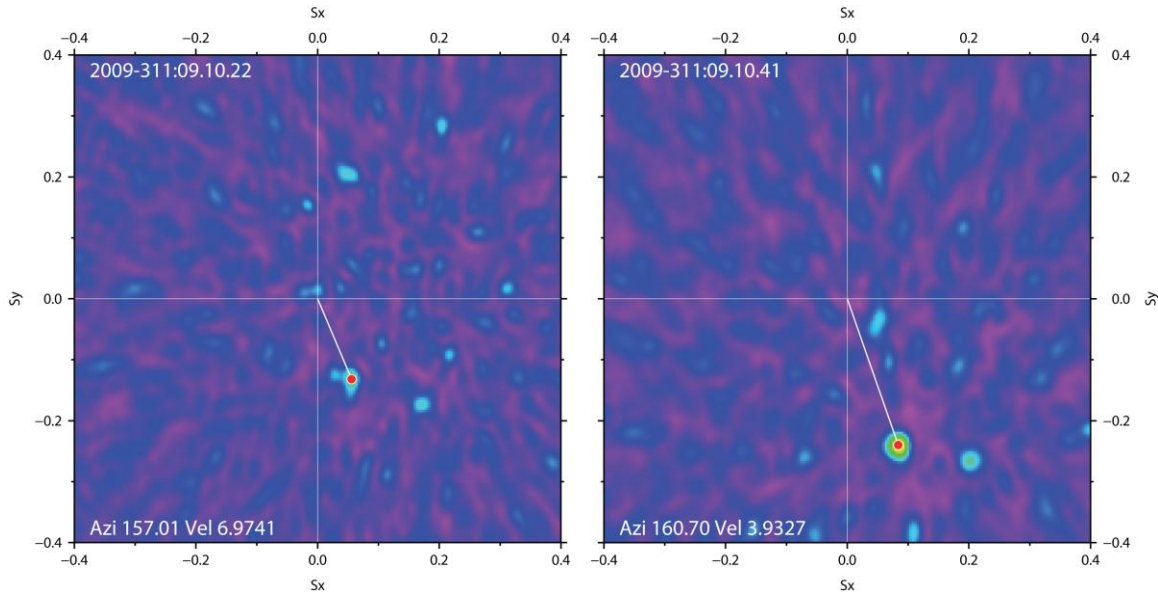


Figure 21: Broadband F-K Estimates on the KSRS Array for a Regional P and a Regional S Phase at the Times as Indicated

Estimates are made in the 2-4 Hz band which appears to produce fairly stable estimates despite the relatively poor SNR in this frequency band (c.f. Figure 20). The detections are made using the procedure of Gibbons et al. (2008) which exploits the incoherent energy at higher frequencies. However, the coherent slowness estimates in the 2-4 Hz band were found to be more stable than the incoherent slowness estimates.

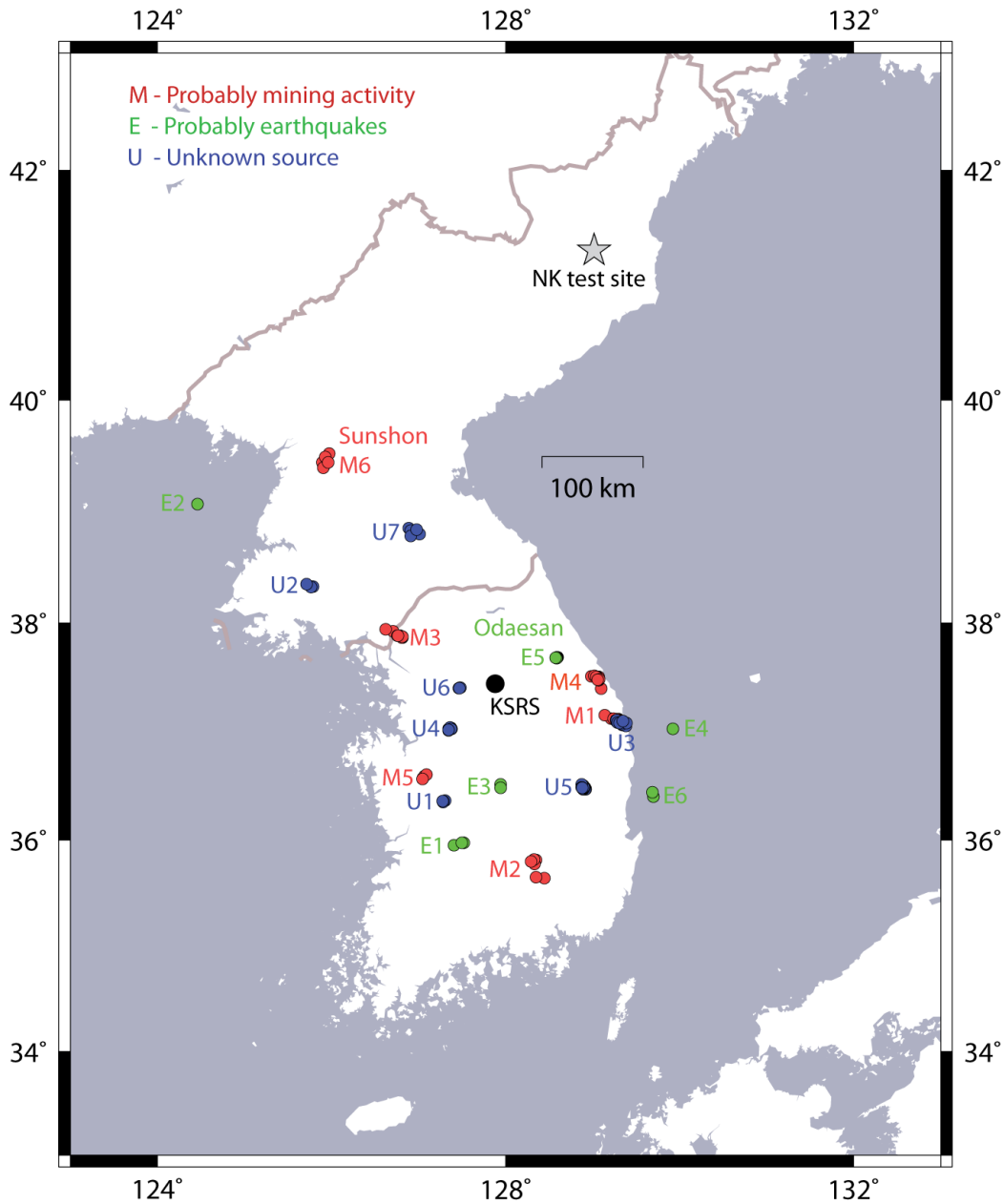


Figure 22: Approximate Locations of 19 Sites at Which Repeating Events Have Been Observed on the KSRS Teleseismic Array in the Republic of Korea

All location estimates are made using only first P and first S arrival time estimates at KSRS, together with the KSRS backazimuth estimates made in the 2-4 Hz frequency band, and a 1-dimensional velocity model.

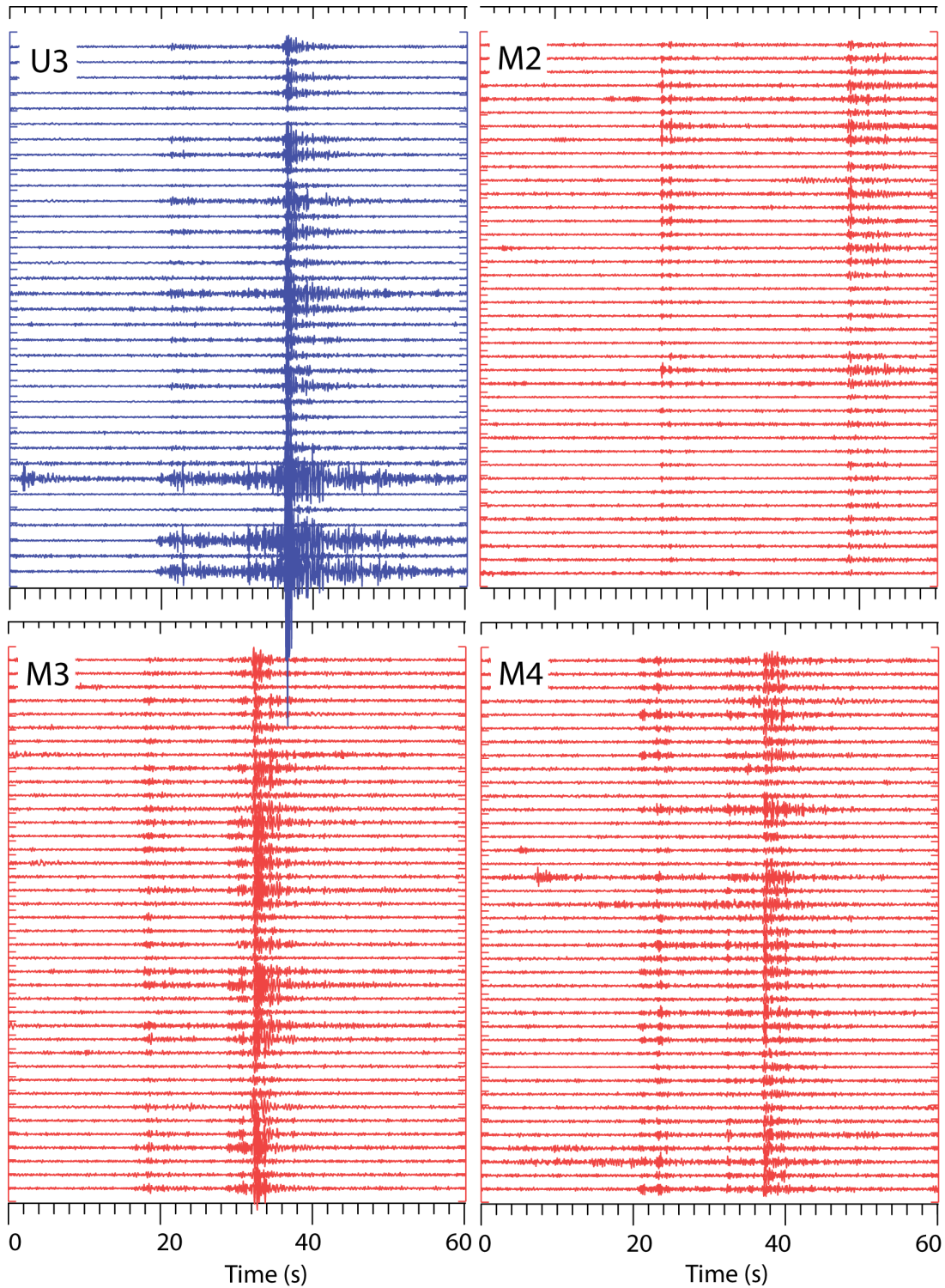


Figure 23: Waveforms from the KS01_SHZ Instrument, Aligned According to the Times of Best Correlation Coefficient for 40 Events from Four of the Clusters Displayed in Figure 22

Table 1: A List of 19 Clusters of Repeating Seismicity on and Surrounding the Korean Peninsula

Code for cluster	Example Origin Time	Lat.	Lon.	Nature
M1 KSRS_2006307A	2006-307:20.30.34	37.12	129.21	mining
M2 KSRS_2006308B	2006-334:08.55.10	35.64	128.44	mining
M3 KSRS_2006315A	2006-315:04.06.52	37.88	126.79	mining
M4 KSRS_2006319B	2006-319:02.44.47	37.60	129.31	mining
M5 KSRS_2007146A	2007-146:06.49.25	36.59	127.07	mining
M6 KSRS_2007244A	2009-039:02.58.34	39.52	125.97	mining
E1 KSRS_2006314C	2006-314:16.16.01	35.97	127.52	earthquake
E2 KSRS_2006315C	2006-315:23.19.22	39.07	124.46	earthquake
E3 KSRS_2006341A	2006-341:04.56.22	36.48	127.94	earthquake
E4 KSRS_2006353E	2006-353:16.18.19	37.03	130.02	earthquake
E5 KSRS_2007020C	2007-020:10.55.21	37.68	128.59	earthquake
E6 KSRS_2007254A	2007-254:03.33.11	36.40	129.69	earthquake
U1 KSRS_2006306A	2006-306:19.24.29	36.36	127.30	unknown
U2 KSRS_2006315B	2006-315:09.48.36	38.32	125.78	unknown
U3 KSRS_2006315D	2006-315:22.28.02	37.11	129.26	unknown
U4 KSRS_2006361A	2006-361:02.52.23	37.02	127.34	unknown
U5 KSRS_2007020A	2007-020:06.37.10	36.48	128.90	unknown
U6 KSRS_2007189A	2007-189:08.42.42	37.40	127.47	unknown
U7 KSRS_2009046A	2007-098:05.49.54	38.85	126.88	unknown

In Table 1, the “example origin time” is an estimate for the origin time of a single event in the cluster from which all other events can be readily found using a multichannel waveform correlation procedure

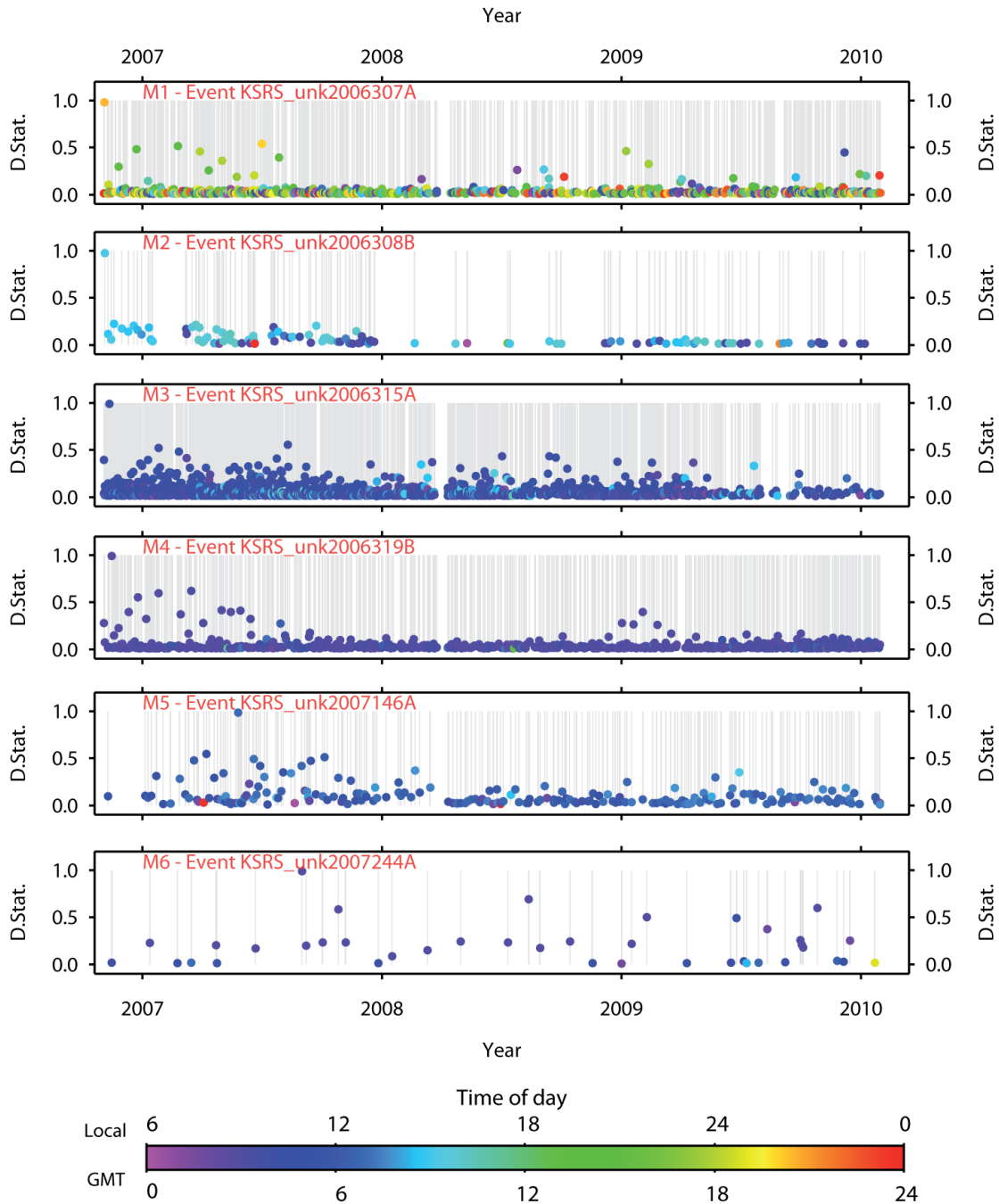


Figure 24: Array-Based Waveform Correlation Detections for Six Mine-Type Clusters (as Listed in Table 1)

In Figure 24, the regularity of the time-of-day and the near-periodicity in their occurrence indicate anthropogenic-type events, and we were able to identify a candidate quarry using Google Earth view. The color of the symbol indicates the time of day as shown on the colorbar,

and the vertical scale of each panel represents the cross-correlation coefficients between the cluster template and the detected events.

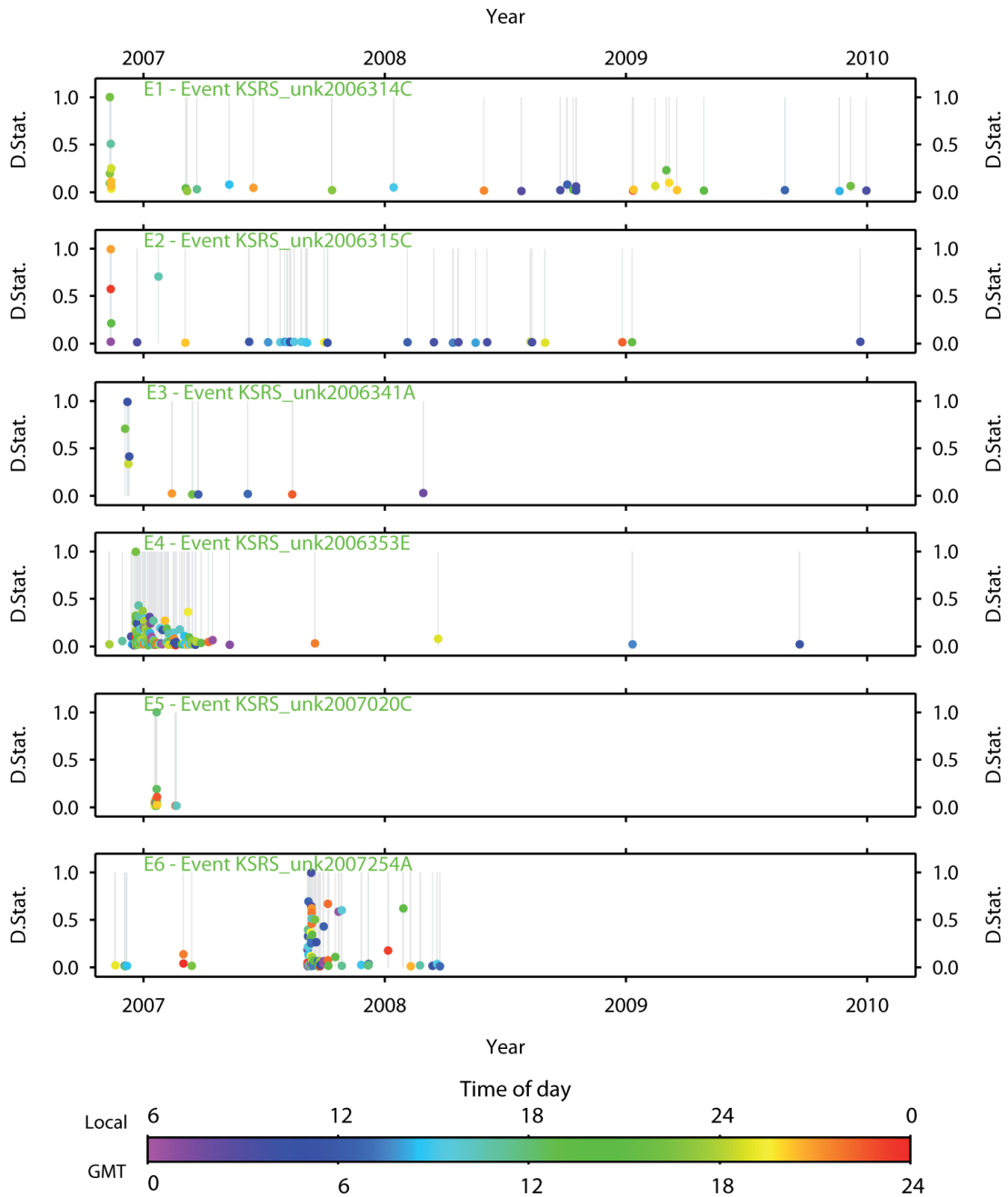


Figure 25: Array-Based Waveform Correlation Detections for Six Earthquake-Type Clusters (as Listed in Table 1)

In Figure 25, the color of the symbol indicates the time of day as shown on the colorbar, and the vertical scale of each panel represents the cross-correlation coefficients between the cluster template and the detected events.

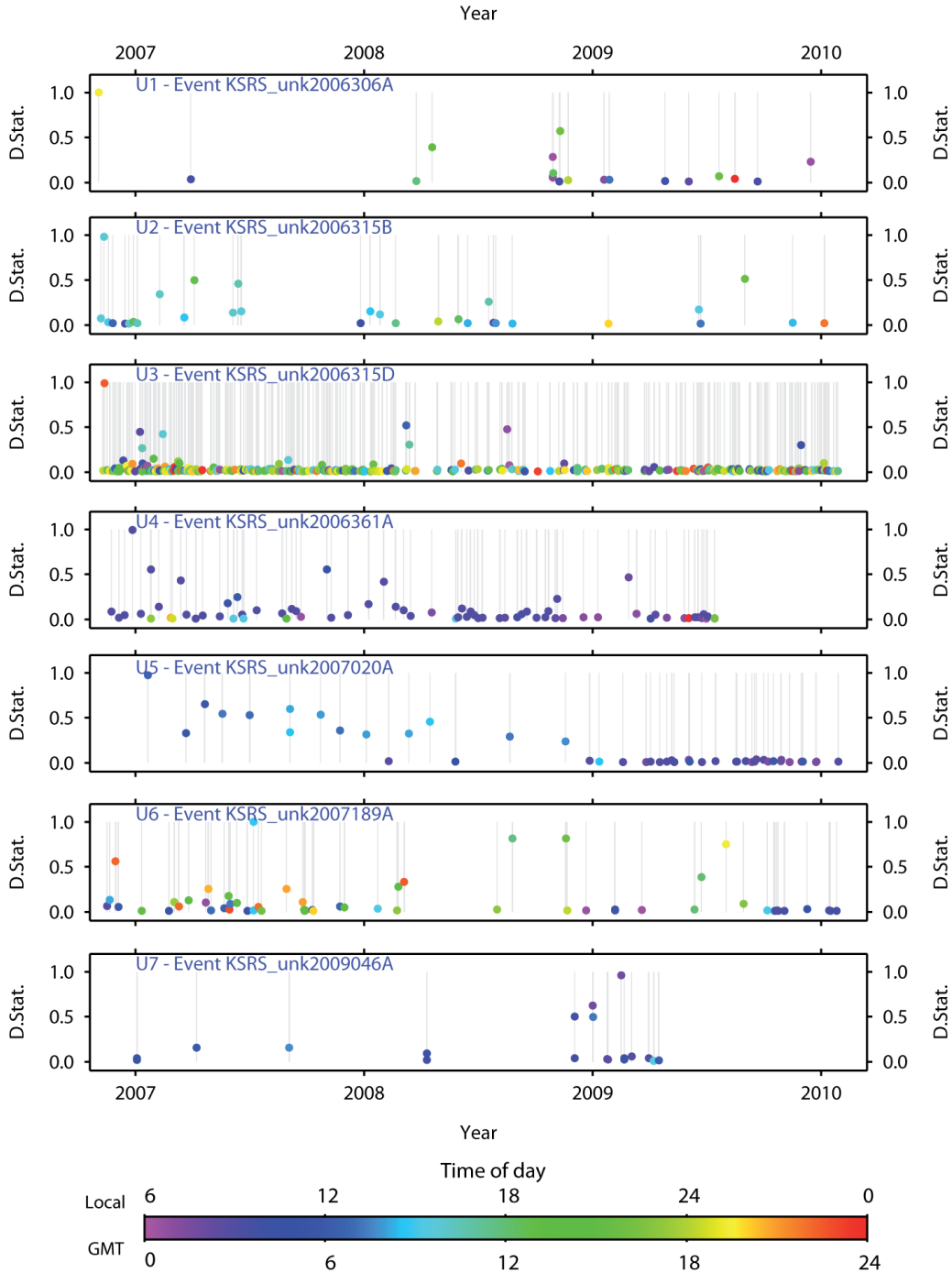


Figure 26: Array-Based Waveform Correlation Detections for Seven Clusters of Unknown Origin (as Listed in Table 1)

The regularity of the time-of-day and the near-periodicity in their occurrence indicate anthropogenic-type events for some of the clusters, but we were not able to identify a candidate quarry using Google Earth view. In Figure 26, the color of the symbol indicates the time of day as shown on the colorbar, and the vertical scale of each panel represents the cross-correlation coefficients between the cluster template and the detected events.

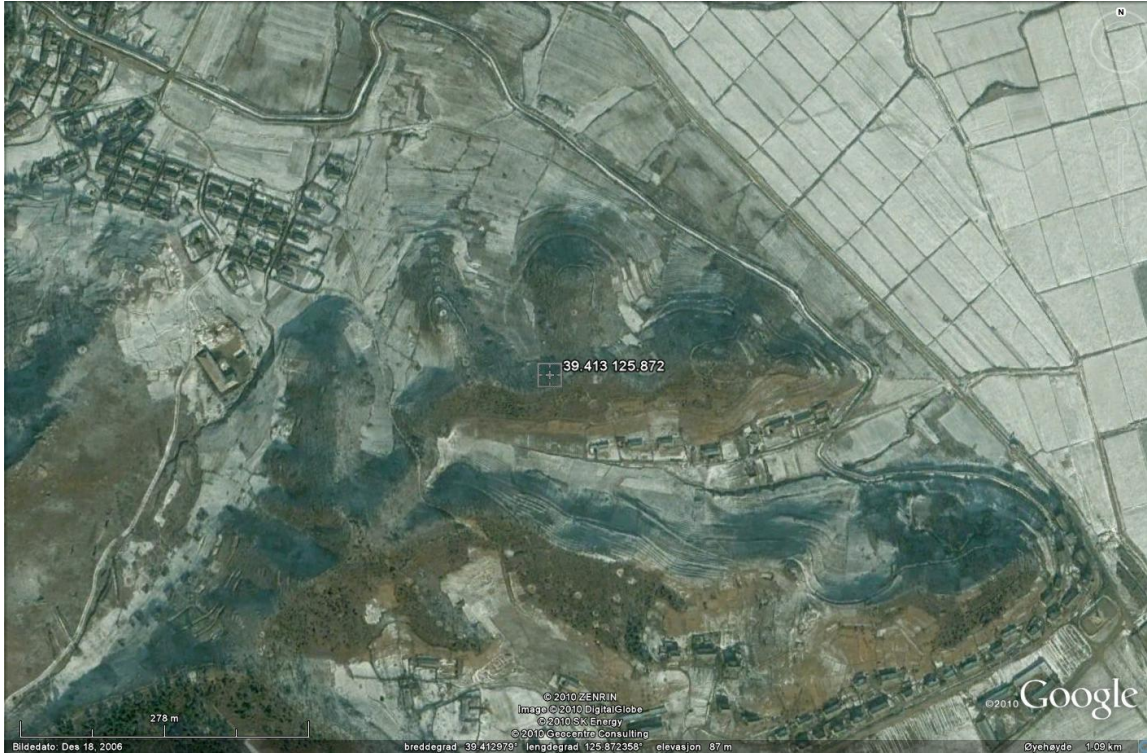


Figure 27: View from Google Earth Centered on the Location 39.413°N and 125.873°E Near the City of Sunshon in North Korea

The above view identifies a candidate quarry corresponding to the location of event cluster M6 shown in Figures 22 and 24.

4. Testing of the Matched Field Detector

Evaluation Using Sources in the European Arctic

The first stage of evaluation of the matched field detection architecture uses known sites of repeating explosions in the European Arctic. Two sites are of particular interest, since the event-to-event repeatability of explosion characteristics are rather different and because their geographical separation is small compared with the source-to-receiver distances. This is important in that it gives us greater confidence in attributing differences in detector performance to differences in source characteristics, rather than path differences or distance-related properties of the wavetrains. The region is also fortuitous in that there are high quality broadband seismometers of the Finnish National Network at comparable epicentral distances to ARCES

(see Figure 28) such that we can evaluate performance of the matched field detection process over two very different receiver apertures. The first set of receivers - four sites of the ARCES array - are within 3 km of each other and display significant waveform similarity between sensors. The second set of receivers form a wide aperture regional network.

The first site is Hukkakero (67.935°N, 25.835°E: already mentioned in the introduction) where the Finnish military destroy expired ammunition in a series of explosions (usually one or two per day) in August and September of each year. The seismic signals generated are highly repeatable, both in amplitude (the yield of each explosion is about 20000 kg TNT equivalent) and in source-time function; all events are, for example, single source explosions as opposed to multiple shot or rippled-fired. All shots are known to take place within a region with an aperture not exceeding 350 meters. The high degree of similarity between events has meant that extensive and reliable lists of events assumed to have taken place at this site have been acquired using the multi-channel waveform correlation procedure of Gibbons and Ringdal (2006).

The second site is a gold mine at Suurikuusikko operated by Agnico Eagle Mine Limited (67.9015°N, 25.39102°E). A correlation procedure has also been applied to this site, although given the considerably lower correlation coefficients frequently involved, the f-k analysis post-processing algorithm of Gibbons and Ringdal (2006) was found to be absolutely necessary to keep the false alarm rate low. High confidence can be obtained in the event-lists obtained given the times of detections. For example, no valid detections were obtained prior to June 2006 when the mine became operational. In addition, almost all detections fall within three time-of-days bands corresponding to the ends of mine shifts; there are essentially no detections at night time.

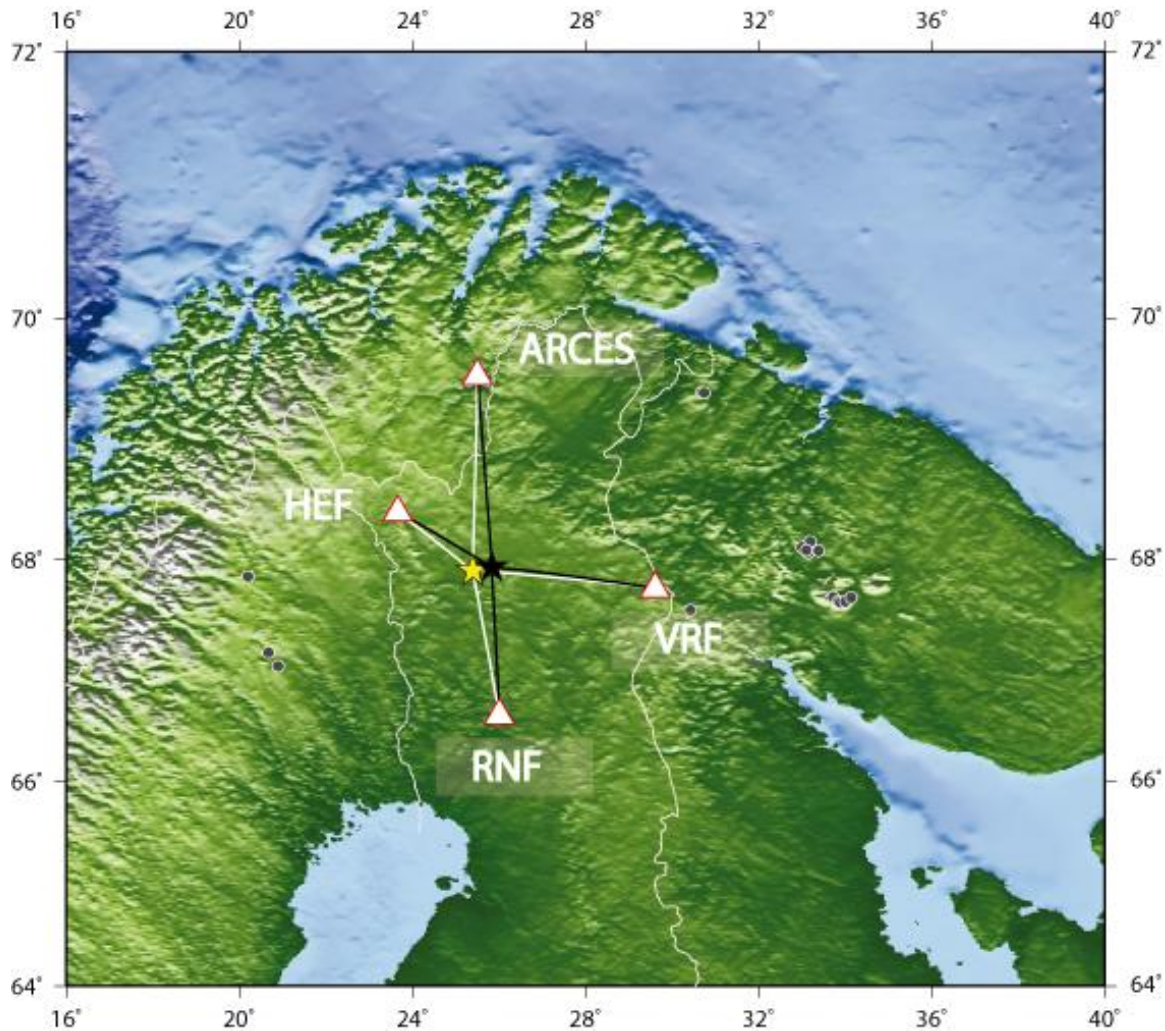


Figure 28: Location of the Hukkakero Military Explosion Site (Black Star) and the Suurikuusikko Gold Mine (Yellow Star) in Relation to the ARCES Array and 3 Broadband 3-Component Stations of the Finnish National Network

This is an ideal laboratory for studying two very nearby sources with rather different signal characteristics using two different receiver apertures. The first (consisting of 4 sites in the ARCES array) has an aperture of approximately 3 km and there is significant coherence between sensors in the frequency range examined for at least transient portions of the wavetrain. The second (consisting of the central element of ARCES and the vertical component channels of three stations of the Finnish network) has a minimum inter-site separation of over 100 km, with each station covering a different backazimuth from the sources, and no semblance of actual waveforms is to be expected.

Figure 29 displays representative waveforms from both of the sites. The Hukkakero waveforms in the left panel are almost all of approximately the same amplitude, whereas the

amplitudes of the Suurikuusikko waveforms vary by around an order of magnitude. Similarly, the variability of signals from the same event from sensor to sensor on the two different receiver apertures is displayed in Figure 30. The sensor-to-sensor variability of waveforms over ARCES for a single event (Figure 30 left) is visibly greater than the event-to-event variability observed on a single sensor of ARCES (Figure 29 left) for a series of Hukkakero explosions. This is consistent with an array aperture of 3 km and a maximum event-to-event distance of 0.3 km. The waveforms from a single Hukkakero event on the different sensors of our regional network have significant moveout due to the differences in traveltimes and there is evidently variability in the waveform structure far in excess of that observed over the small aperture array.

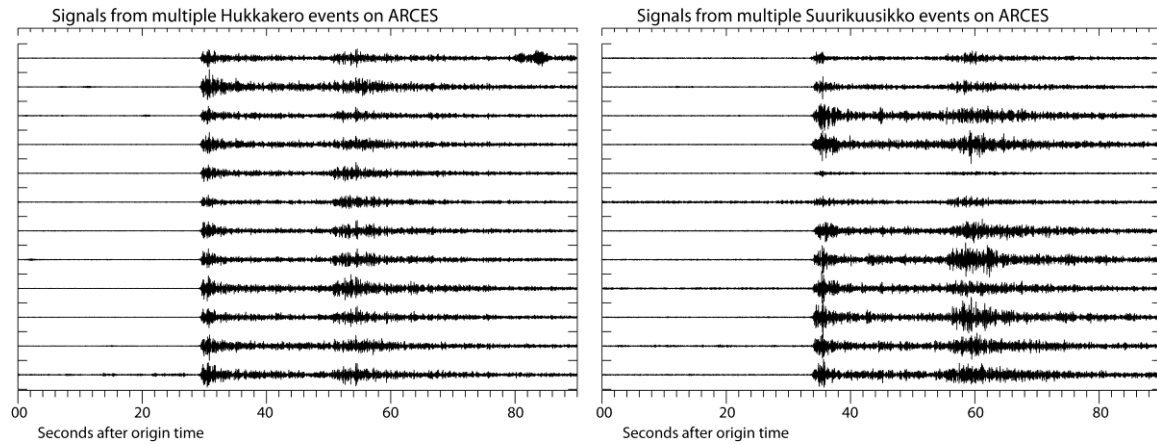


Figure 29: Waveforms on the Channel ARA0_Sz of ARCES from Each of 12 Events from the Two Sites as Indicated – For Both Left and Right Panels, the Vertical Scaling of Traces Is Identical for the Different Traces Within the Same Panel

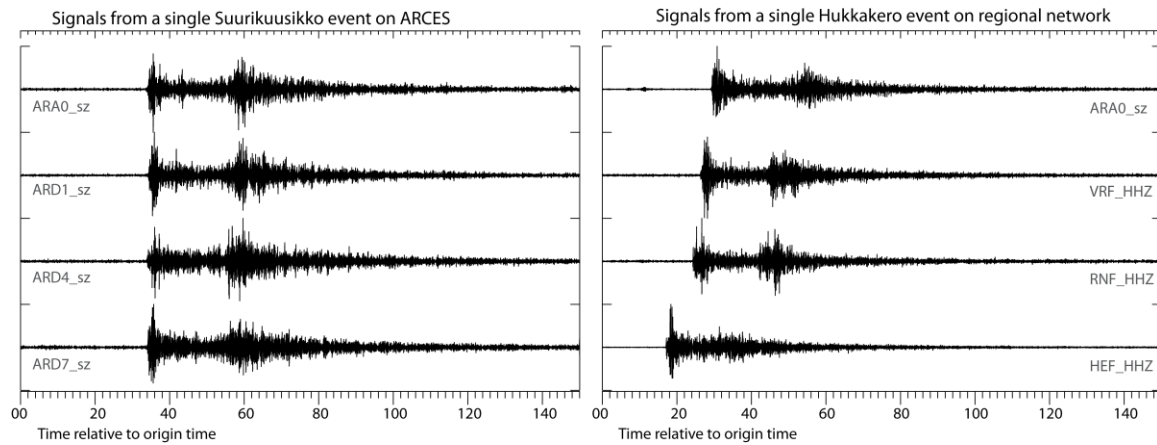


Figure 30: Variation of Waveforms from the Two Sites Over the Different Receiver Apertures

Figure 31 displays the detection statistics generated by six different matched field detectors over a single day (August 13, 2008) in which there are known to be two Hukkakero events and a

single Suurikuusikko event. The lowermost three traces comprise detection statistics from matched field detectors that use waveform templates from Hukkakero events. These traces are analogous to those displayed in Figures 12 and 14, only using different events to generate the templates. The detection statistic for both the Hukkakero explosions which took place on this day are more even for the ensemble template than for the single event template. The approximate origin time of the single event used for this template is 11.00.00 UT on August 20, 2008.

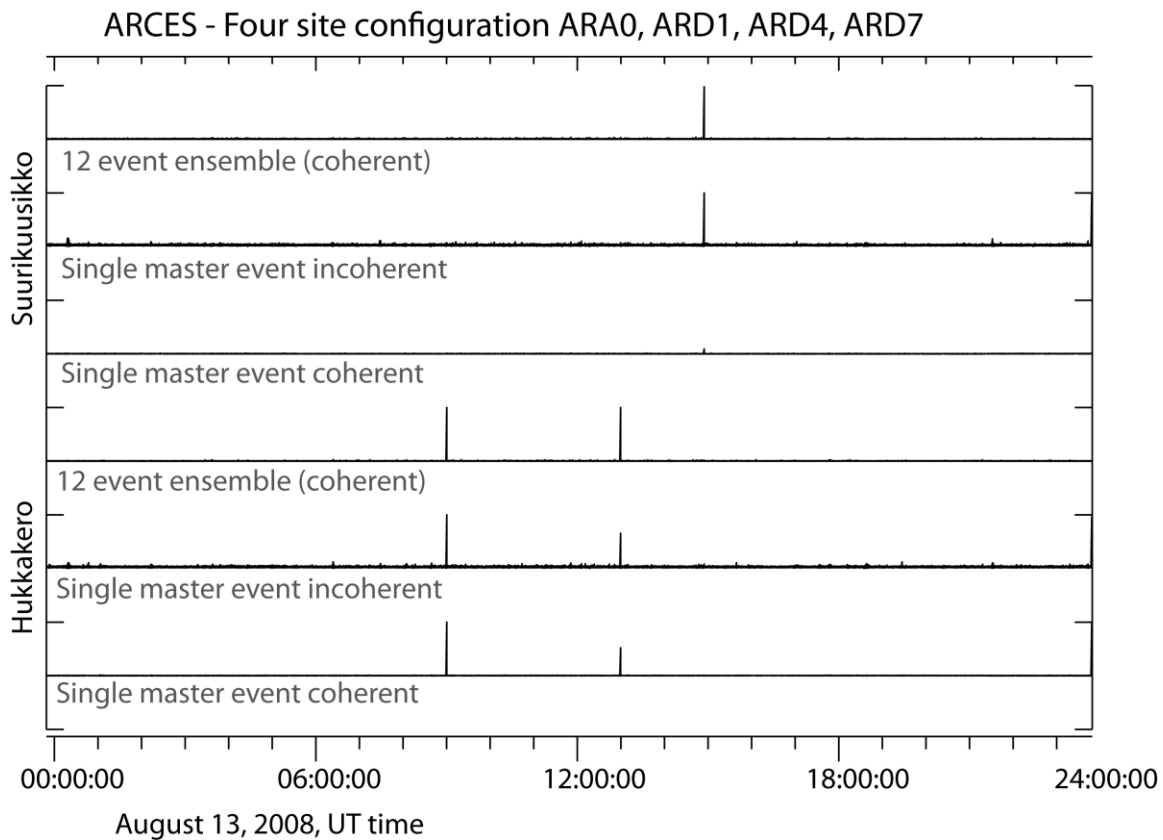


Figure 31: Matched Field Detection Statistic Traces for the Two Sites Displayed in Figure 28 Where All Calculations Are Performed Using Only the Four Short-Period Vertical Channels of the ARCES Array as Indicated

In Figure 31, for each site, three channels are displayed. The lowermost (*Single master event coherent*) is a coherent matched field detector where the template is from a single event. The next trace up (*Single master event incoherent*) uses the same event as a template but sums incoherently over the frequency bands. The top trace (*12 event ensemble: coherent*) is a matched field detector with the template constructed from the 12 events in displayed in Figure 29. All detection statistic traces have equal scale and are clipped at a value of 0.5.

The uppermost three traces in Figure 31 result from matched field detectors that employ Suurikuusikko events as templates. The single-event coherent detection statistic trace has a significant peak at the time of the August 13, 2008, Suurikuusikko blast, although the ratio of the

maximum to the standard deviation of the background values of the detection statistic is somewhat lower than for the 12 event ensemble template. The incoherent single event template detector also detects the August 13 event unambiguously. The approximate origin time of the single event used for this template is 15.54.16 UT on November 20, 2008.

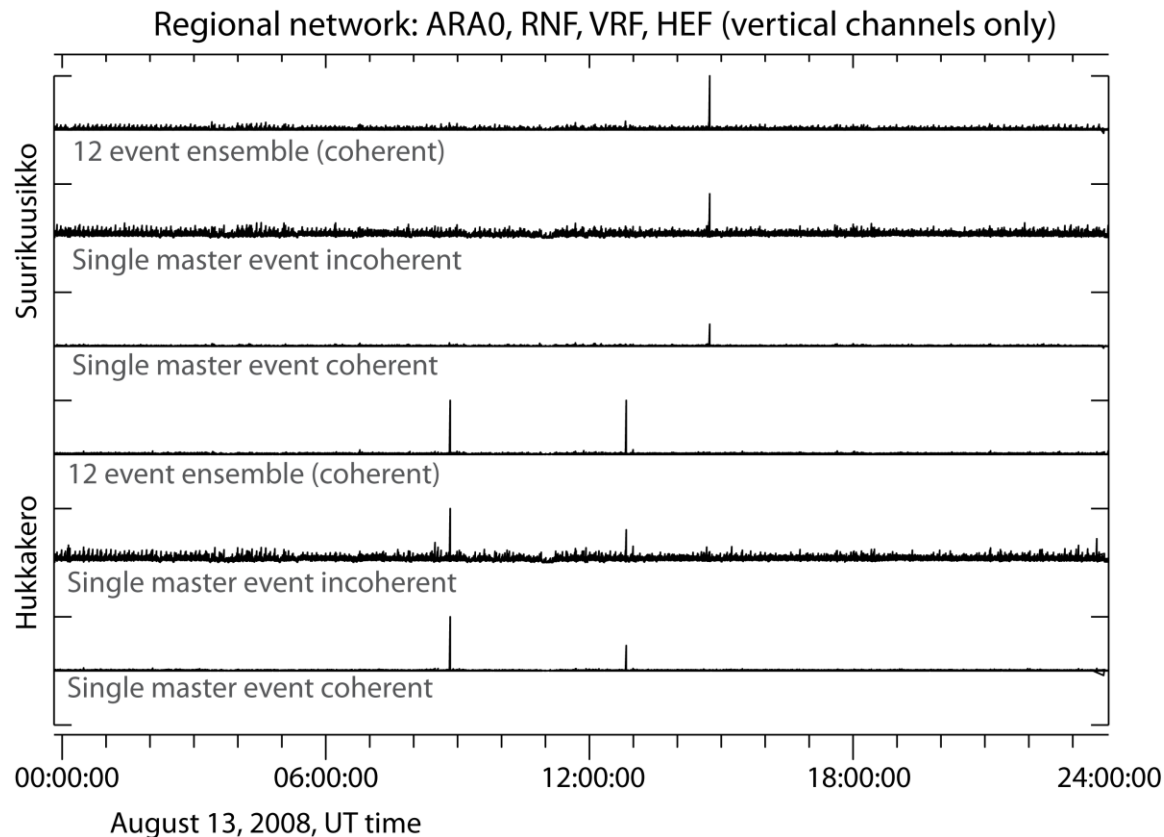


Figure 32: As for Figure 31, Except Calculations Are Performed Using the Four Site Regional Network as Displayed in Figure 28 – All Detection Statistic Traces Have Equal Scale and Are Clipped at a Value of 0.5

Figure 32 displays the corresponding results for the 4 site regional network shown in Figure 28. All of the matched field detectors manage to identify the events from the appropriate source regions with a sufficiently high detection statistic that a "zero-false-alarm-rate" can still be achieved. The background variability of the incoherent detectors appears to be somewhat higher than for the four sensor configuration of ARCES. There may be a degradation of performance resulting from the waveform dissimilarity. On the other hand, there are a number of complications which could contribute to a slight worsening of performance:

- Whereas the ARCES array has central timing, the ARA0, HEF, VRF, and RNF stations are part of three entirely independent networks and even a sub-sample time discrepancy may degrade the results.

- The HHZ channels of the 3-component seismometers are sampled at 100 Hz whereas the ARA0_sz channel is sampled at 40 Hz. All channels have been resampled on to a common grid prior to processing.
- The instruments have different responses. These calculations have been performed on raw waveform data with no correction for instrument response having been made. In particular, there is some discrepancy between the amplitude (in counts) of the different instruments.
- The 60 second long template was chosen to include all initial P-arrivals and as much of the coda at the various sites as could be included following the P-arrival at HEF (the closest station). No alignment of the waveforms was attempted and no tapering of the noise segments of the template waveforms was performed. It may be that the performance could have been improved by applying systematic shifts both to the template waveforms and the incoming waveforms to make all of the P-arrivals approach approximately simultaneously.

The main conclusion to be drawn from this part of the study is that the matched field detection framework proposed here can be extended to large receiver apertures over which signal coherence in the traditional sense of the word is non-existent. We can therefore apply the procedure to any network of sensors recording a signal from a specified source.

Evaluation using sources on the Korean peninsula

For assessment of the new full-waveform Matched Field (MF) detector, we report on a testing exercise using the KSRS array in South Korea, focusing on two of the event clusters shown in Figure 22. The first cluster consists of the so-called Odaesan earthquake with foreshock and aftershock sequences described by Kim and Park (2010), located at a distance of about 70 km north-east of KSRS. We have also been given access to 4 days of data from the station DGY operated by the Korea Meteorological Administration, located as close as a few kilometers from the earthquake sequence. The other cluster is a presumed mining site denoted M4, located about 100 km east of KSRS. The locations of KSRS, DGY and the two event clusters are shown in Figure 33.

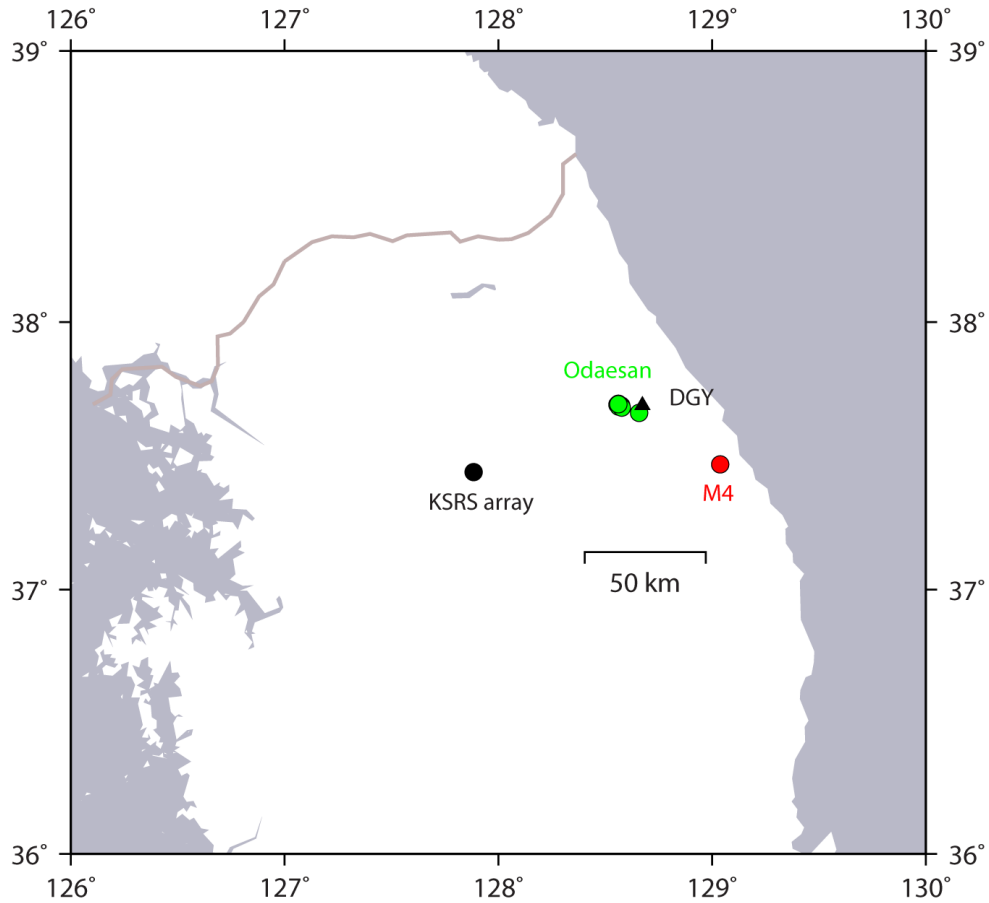


Figure 33: Map Showing the Location of the KSRS Array in South Korea, the Odaesan Earthquakes (Green Filled Circles), the M4 Mine (Red Filled Circle) and the DGY Three-Component Station Operated by the Korea Meteorological Administration (KMA)

KSRS is a 19 element array spanning an aperture of about 10 km (see Figure 34). In order to limit the processing load when testing different approaches to full-waveform MF detection, we have initially used data from only 5 of the KSRS sensors shown red in Figure 34.

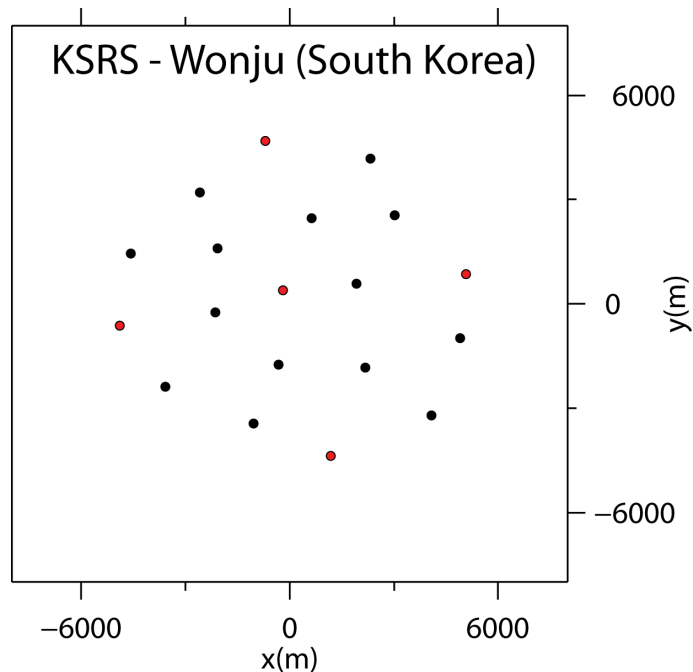


Figure 34: Sensor Distribution of the 19-Element KSRS Array in South Korea – Red Symbols Show the Location of the 5 Sensors Used for Initial Testing of the MF Detector

According to Kim and Park (2010), the local DGY data showed signals from 75 potential earthquakes occurring within the 8 days of the Odaesan earthquake sequence. The m_b 4.8 main event on 20 January 2007 was preceded by a sequence of smaller events starting 3 days earlier. As first test, we selected one of the larger foreshocks occurring on 17 January as the processing template. The corresponding waveforms at the KS01_SHZ sensor are shown in the upper trace of Figure 35. This single event was used as a template for both incoherent and coherent MF detection processing, as well as for detection using array-based waveform correlation (Gibbons and Ringdal, 2006). In addition, an ensemble of ten events distributed over the event sequence (also see Figure 35) was used for coherent MF detection processing.

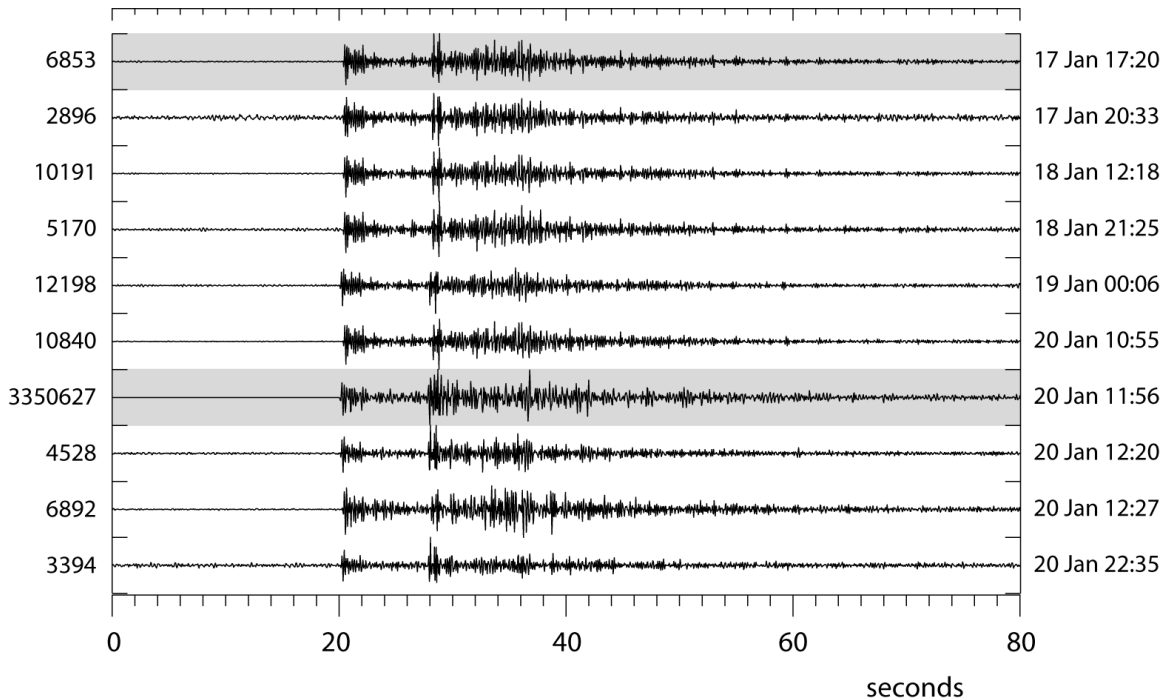


Figure 35: Panel of 10 Events from the Odaesan Earthquake Sequence Recorded at the KS01_SHZ Sensor of the KSRS Array – the Data Are Bandpass Filtered Between 2 & 8 Hz

In Figure 35, the highlighted upper trace shows the foreshock used for single-event incoherent and coherent MF detection processing, as well as for detection using array-based waveform correlation. The highlighted trace no. 7 from top shows a recording of the main m_b 4.8 event. The ensemble of all 10 events was also used for coherent MF detection processing. The event origin times are given to right of each trace, and the corresponding maximum amplitudes are given to the left.

The results from processing a one-hour time interval around the template foreshock on 17 January 2007 are shown in Figure 36 (see figure caption for details). Notice that the vertical scale is set to 0.1 for the detector outputs, such that the largest peaks are clipped. E.g. the detection trace peaks associated with the 17:20 template event have all values close to 1.0. We observe that there is a very good correspondence between the events found by full-array KSRS waveform correlation (red stars) and signals observed at the local DGY station (lower trace). However, only a couple of these events can be observed on the top KS01 trace. Both the single template coherent MF detector (trace no. 3) and ensemble coherent MF detector (trace no. 5) show clear detection peaks at the time of the reported events (red stars). The single template incoherent detector (trace no. 2) show, as expected, larger noise variability, but clear peaks show up at the times of the larger events.

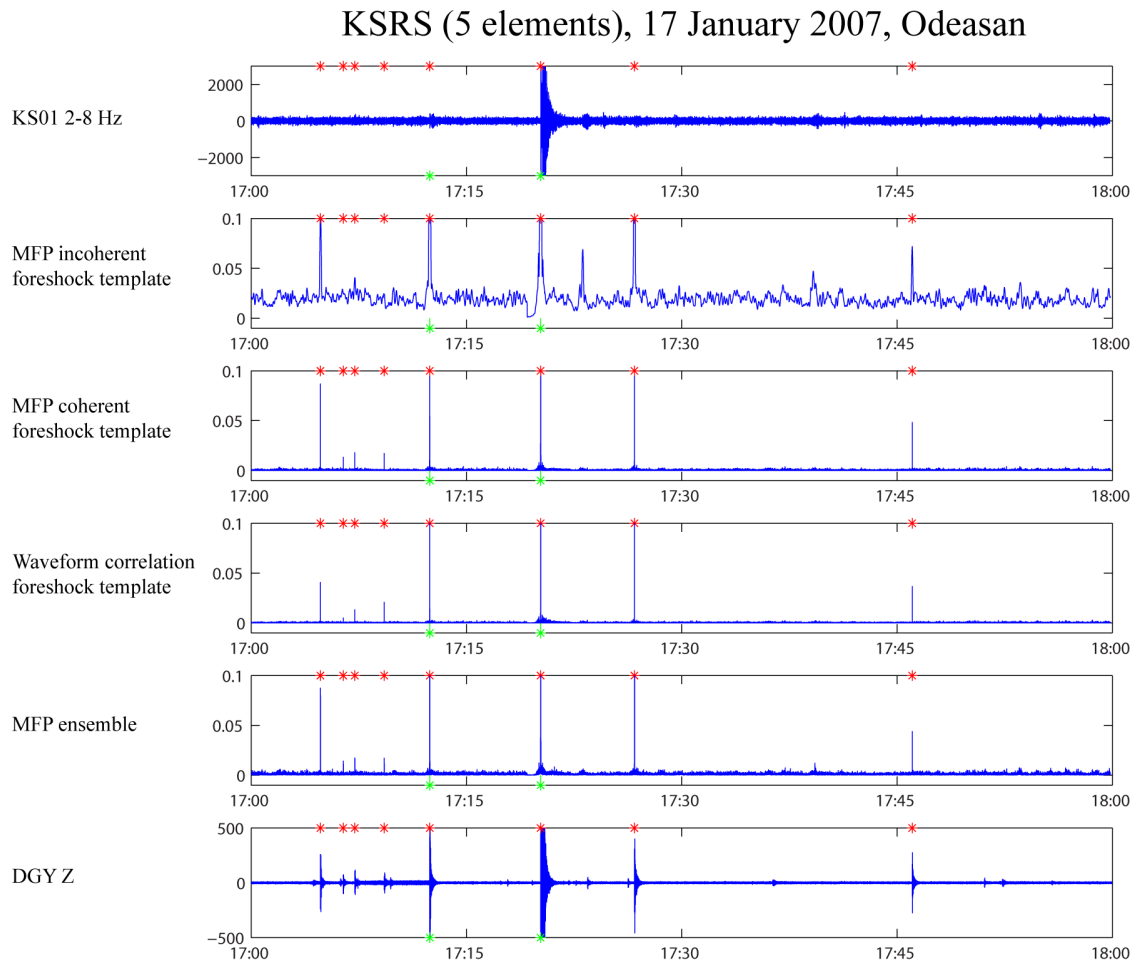


Figure 36: Detector Processing Results Using the Five KSRS SHZ Sensors KS01, KS08, KS11, KS14 and KS17 for a One-Hour Time Interval Around the Largest Foreshock of the Odaesan Earthquake Sequence

Figure 36 shows the following:

- The upper trace shows recordings at one sensor of the KSRS array (KS01) filtered in the passband 2-8 Hz. Traces 2-4 from the top show different detector outputs where a 60 s data segment of the event seen around 17:20 is used as the template.
- Trace no.2: Single event incoherent matched field detector.
- Trace no. 3: Single event coherent matched field detector.
- Trace no. 4: Waveform correlation (single event).
- Trace no. 5 shows the detector output using the ensemble of ten events of the Odaesan earthquake sequence.
- Trace no. 6 shows the vertical-component recordings at the station DGY.

The red stars indicate event candidates found by waveform correlation utilizing all 19 sensors of the KSRS array, and green symbols indicate larger events reported by Kim and Park (2010).

To get an impression of the background noise variability, the maximum vertical scale of the detector outputs is set to 0.1. As a result, the largest detection “peaks” are clipped.

The results from processing a one-hour time interval around the main event (m_b 4.8) on 20 January 2007 are shown in Figure 37. We see here that waveform correlation (red stars) using the full KSRS array with the 17 January foreshock template was only able to find the 4 larger events. As seen from the local DGY data in the lower trace, there are several smaller Odaesan event candidates during this time interval. However, the ensemble coherent MF detector (trace no. 5) shows clear detection peaks corresponding to several of these smaller events. This indicates that the characteristics of the events occurring during this time interval differ significantly from the template foreshock that occurred three days earlier, but that the subspace spanned by the 10 event ensemble enabled the coherent MF detector to trigger on these events.

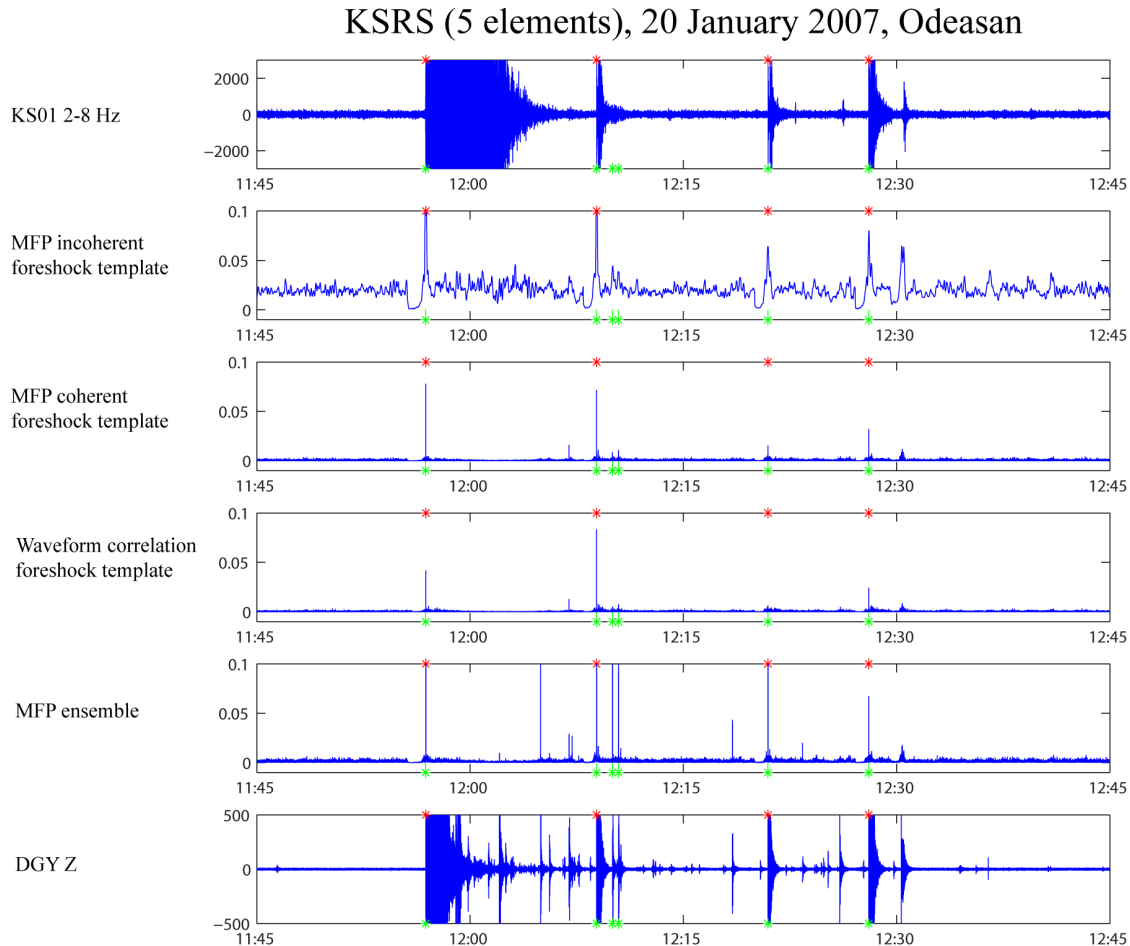


Figure 37: Detector Processing Results Similar to Figure 36, but for a One-Hour Time Interval Around the Main Odaesan M_b 4.8 Earthquake on 20 January 2007

For the M4 mine cluster (see Figure 33), the KSRS array-based waveform correlation procedure triggered on numerous events during the time period November 2006 to February 2010. This is clearly observed in panel no. 4 of Figure 24. As seen from the color scale, almost all of the triggers occur around 03 GMT (09 local time), which indicate anthropogenic origin. A candidate open-pit mine was found by Google Earth view at coordinates 37.47°N, 129.04°E.

An ensemble of ten M4 events is shown in Figure 38. We initiated full-waveform MF detection processing using both the event ensemble and a single master event with origin time 02:44 GMT on November 15 2006 (see trace no.2 of Figure 38).

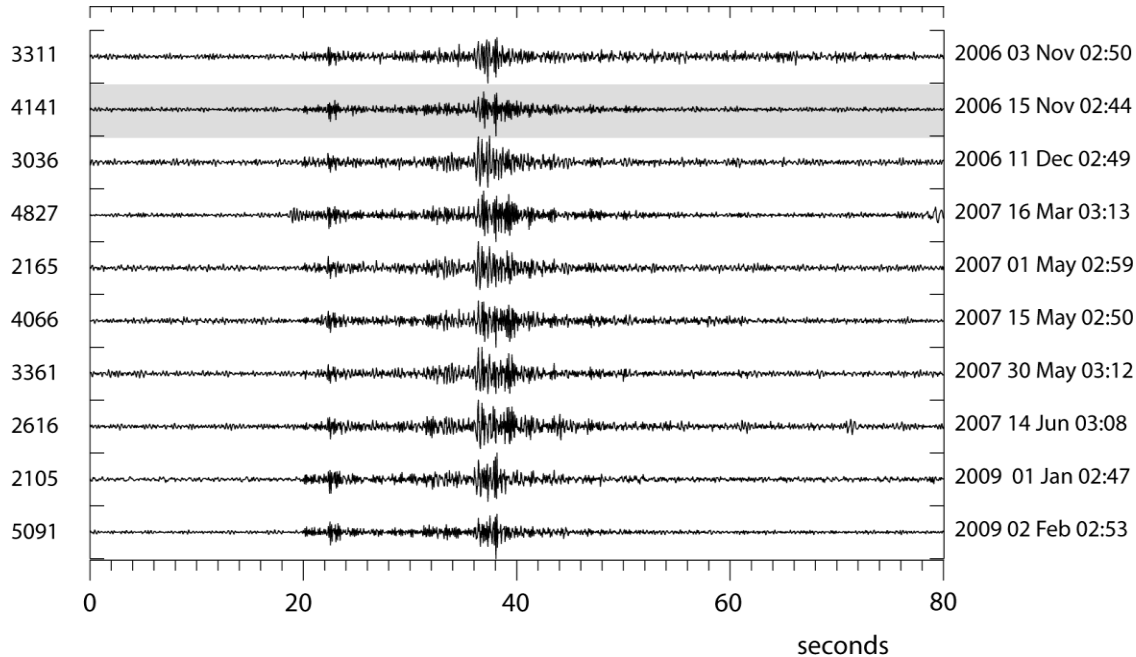


Figure 38: Panel of 10 Events from the M4 Mine Recorded at the KS01_SHZ Sensor of the KSRS Array. The Data Are Bandpass Filtered Between 2 and 8 Hz

In Figure 38, the highlighted trace no. 2 shows the event used for single-event incoherent and coherent MF detection processing, as well as for detection using array-based waveform correlation. The ensemble of all 10 events was also used for coherent MF detection processing. The event origin times are given to right of each trace, and the corresponding maximum amplitudes are given to the left.

The results from processing 24 hours on 25 January 2007 are shown in Figure 39. The full-array KSRS waveform correlation (red stars) triggers on two instances around 03:00 GMT, and we see that a clear peak is seen on all detector outputs of traces 2-5. Outside this small time interval, the single event coherent MF detector (trace no. 3), the ensemble coherent MF detector (trace no. 5) and the array-based waveform correlator (trace no. 4) all show very low detection statistics. On the other hand, the single event coherent MF detector (trace no. 2) show significantly higher noise variability.

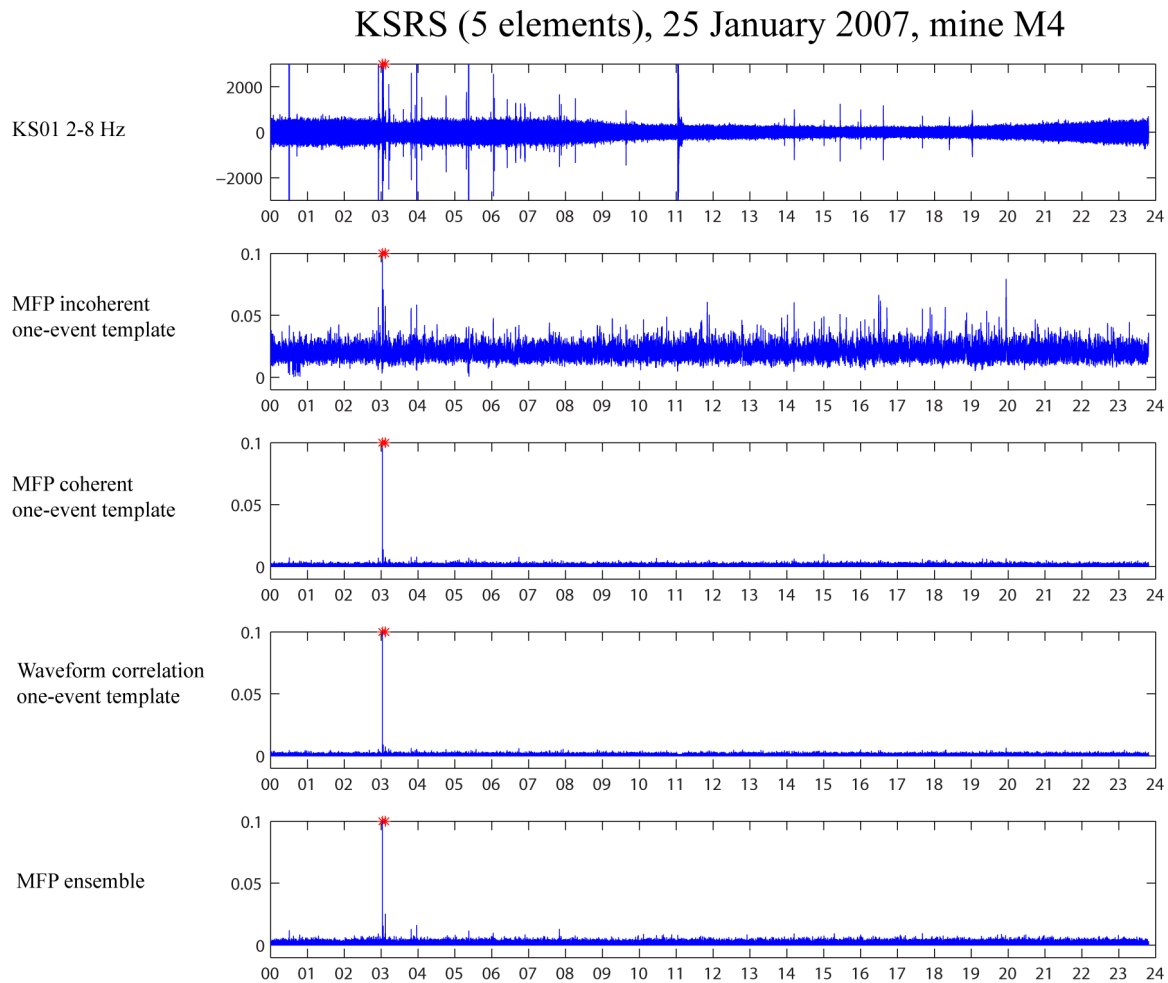


Figure 39: Detector Processing Results Using the Five KSRS SHZ Sensors KS01, KS08, KS11, KS14 and KS17 for 25 January 2007

Figure 39, shows the following:

- The upper trace shows recordings at one sensor of the KSRS array (KS01) filtered in the passband 2-8 Hz.
- Traces 2-4 from top show different detector outputs where an M4 mine event on 15 November 2006 is used as a template.
- Trace no.2: Single event incoherent matched field detector.
- Trace no. 3: Single event coherent matched field detector
- Trace no. 4: Waveform correlation (single event)
- Trace no. 5 shows the detector output using an ensemble of ten M4 mine events.

The red stars indicate event candidates found by waveform correlation utilizing all 19 sensors of the KSRS array, and green symbols indicate events reported by Kim and Park (2010). To get

an impression of the background noise variability, the maximum vertical scale of the detector outputs is set to 0.1. As a result, the largest detection “peaks” are clipped.

When zooming in the a one-hour time interval between 02:30 and 03:30, we can from Figure 40 separate the two M4 event detections found by waveform correlation (which utilized all 19 sensors of the KSRS array - red stars). The detection peak associated with the second event is very small for both the single event coherent MF detector and the array-based waveform correlator (panels 3 and 4). The ensemble coherent MF detector (panel no. 5) show significant increases at the time of the second event, which again indicate that the subspace spanned by the 10 event ensemble enabled the coherent MF detector to trigger on this small event. The single event MF incoherent detector statistics (panel no. 2), calculated from the 5-element KSRS sub-configuration has rather high variability during noisy conditions, but several peaks exceeding the background noise level can be observed within the time interval of the declared events.

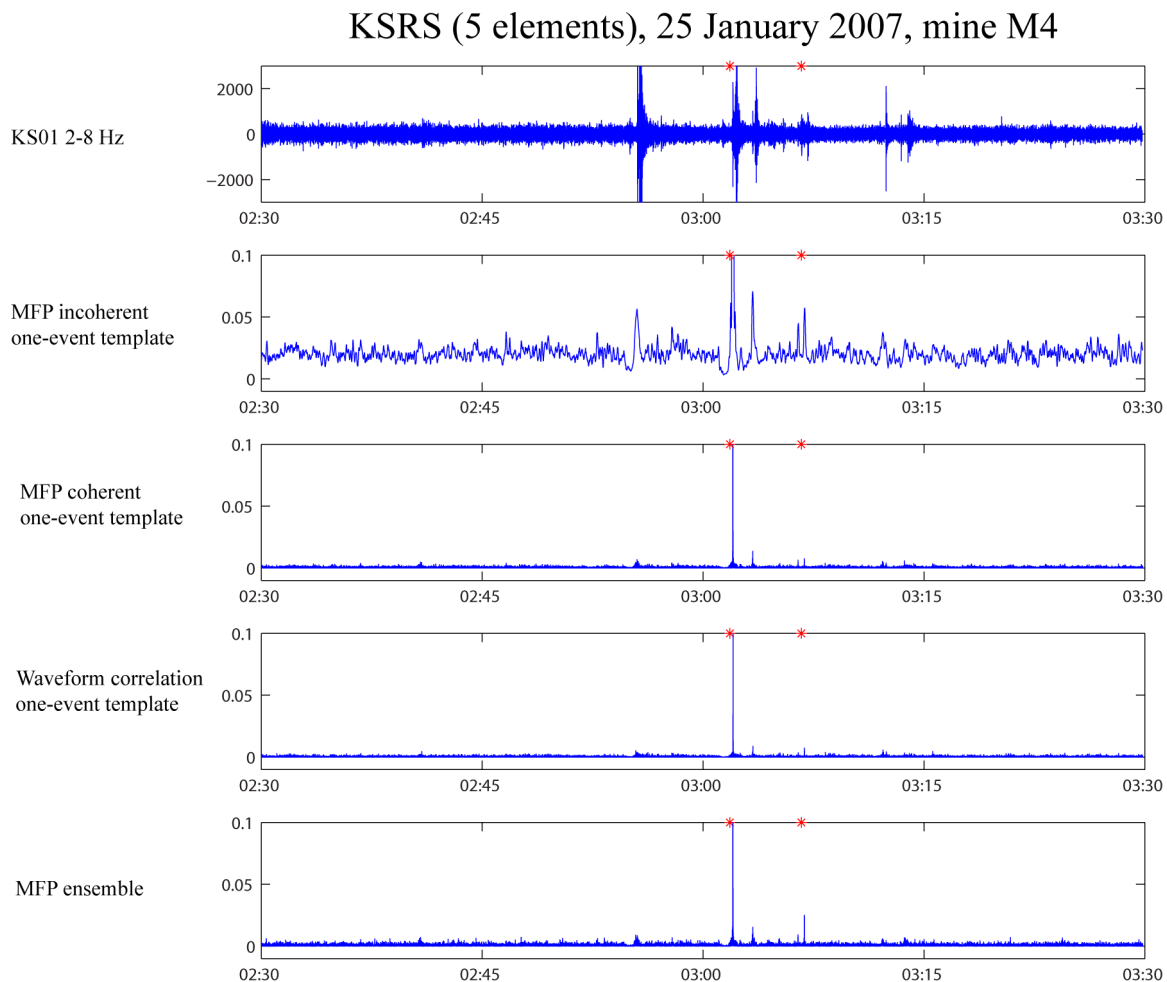


Figure 40: Detector Processing Results Similar to Figure 39, but for a One-Hour Time Interval Around the Two M4 Event Candidates That Occurred Right After 03:00 on 25 January 2007

In a final run we included the full KSRS array with 19 sensors in the processing for the M4 mine. The results are shown in Figure 41, and are directly comparable to the results for the 5-element KSRS sub-geometry shown in Figure 40. Noise reduction and enhancements of the detection peaks are observed for all outputs, but the largest improvement is observed for the single event MF incoherent detector shown in panel no. 2. In fact, the single event MF incoherent detector and the ensemble MF coherent detector (panel no. 5) indicate that an additional event from the M4 mine occurred 1-2 minutes after the first event.

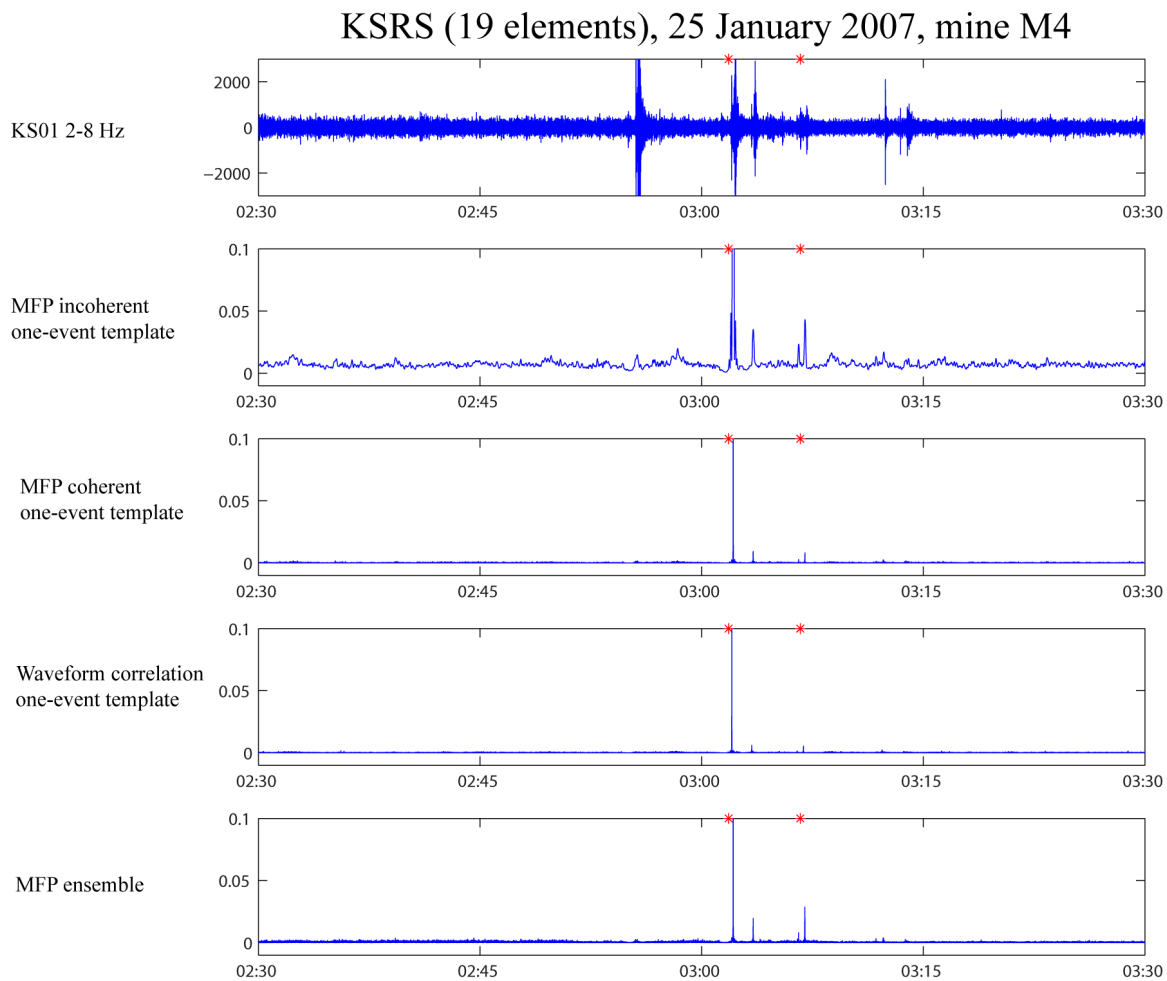


Figure 41: Detector Processing Results Similar to Figure 40, but Now Using All 19 Sensors of the KSRS Array, as Opposed to the 5 Sensors Used in Figure 40

This final example provides evidence that matched field detectors are less sensitive than correlators to source time history variations and give some indication that further research is merited. Matched field detectors have higher background noise than correlators because of the

large subspaces they employ. The large subspaces enable MF detectors to match noise better in addition to intended signals; there is a tradeoff between increasing the dimension of the detector to capture more signal energy, which is desirable, and increasing noise capture. This example shows that this disadvantage of MF detectors is partially offset by increases in the number of data streams (stations) being processed. Background noise reduction increases faster than signal loss in the detection statistic as the number of stations increases. This observation suggests that MF detectors may be suitable for situations where a very large number of stations is available for detection. These could be stations in a network that normally would not be used for coherent processing, but may be used that way with the calibration that matched field processing provides. The same observation can be made for the wideband MFP generated from an event ensemble (a classical subspace detector in this implementation), which also assures that diverse signals can be matched (panel 5). The ensemble design has potential to calibrate a larger geographic footprint and mechanism variations in addition to source time history variations. In this case, the fact that the ensemble processor and the single-event incoherent (matched field) processor performed well, provides evidence that the later events differed from the first event (with a large detection statistic) primarily in source time history variations.

5. Detection Framework

In this section, we first describe the detection framework we built to test matched field detectors on continuous data streams, then give several examples.

The detection processing framework is a research database and associated java programs intended to facilitate the development and testing of algorithms for operating suites of subspace and matched field detectors. The framework is a generalization of the system described in Harris and Dodge (2011). It allows retrospective processing of sequences of data using various system configurations. Results are saved in a database, making it easy to compare the results obtained using different configurations of the system. The new features introduced in this system are:

1. more flexible specification of and input of waveform data,
2. archival of state for all framework runs so that effects of parameter changes can be tracked and documented,
3. better management of non-database artifacts (detection statistics and detection segments),
4. speed enhancements, where possible, with multithreading techniques (parallelization),
5. support for multiple streams,
6. support for adaptive matched field detectors.

Management of Input Waveform Data

The new system is an improvement on the previous system, which read data only from flat files. Waveform data now are read either from a WFDISC_STAGE table that is part of the DETECTOR schema or else from an external continuous WFDISC table. Data may still be input from non-database sources (e.g. SAC file collections, WFDISC flat files, SEED volumes), but

these data are preloaded into the WFDISC_STAGE table. As external data are staged, segments are merged where possible. By ensuring a known data source for any detection in the system, it is no longer necessary to write out (and then manage) detection segments. Instead, the system includes a program that allows extraction of detection segments as required. In the previous system detection segments were written to disk as detections were made. The problem with that approach is that if the detector is later retired as a result of being subsumed into a higher-rank detector, all the files must be moved into a new directory and the old directory deleted. Also, the decrease in IO means that the framework is able to run faster. Having the results stored in the database also facilitates collaboration because users can see the same data even if they are not using the same computer.

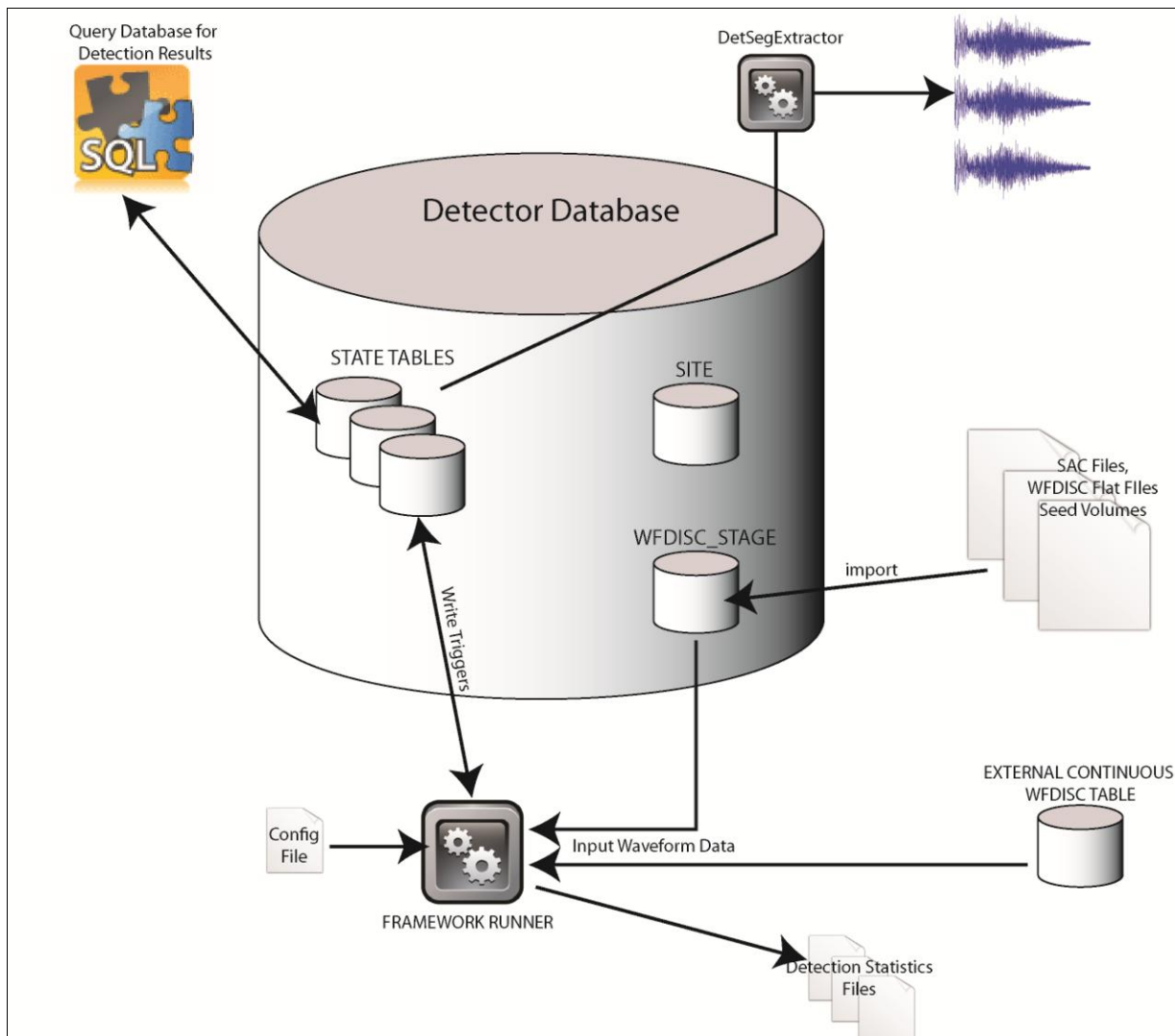


Figure 42: Schematic of the Detection Framework

There are still two unresolved issues related to input waveform management. The first is dealing with gaps in the data. The Stream Server is able to recognize gaps in the data and to fill those gaps with zeros. This allows the system to operate without crashing when gaps occur. However, a mechanism must be created to notify the stream processors of the location of gaps so that they may take appropriate action. For example, in some cases it might be necessary to skip to the end of the last gap in a data buffer and re-start the preprocessor. Alternatively, if gaps are restricted to one or a few channels, it may be possible to temporarily modify the detector template(s) so as not to use the affected channels.

The second unresolved issue is interpolation of the data streams. Multi-channel data are not guaranteed to have identical sample rates, and may not have samples at the same instant. We plan in the near future to include an interpolation mechanism to solve this problem.

Preservation of State

In our work with the previous version of the framework, we often found it necessary to reproduce the results of a previous run in order to answer some question. Without a way to track everything that went into a run, it could be difficult in some cases. In this implementation of the framework we have introduced versioning of the results to solve that problem.

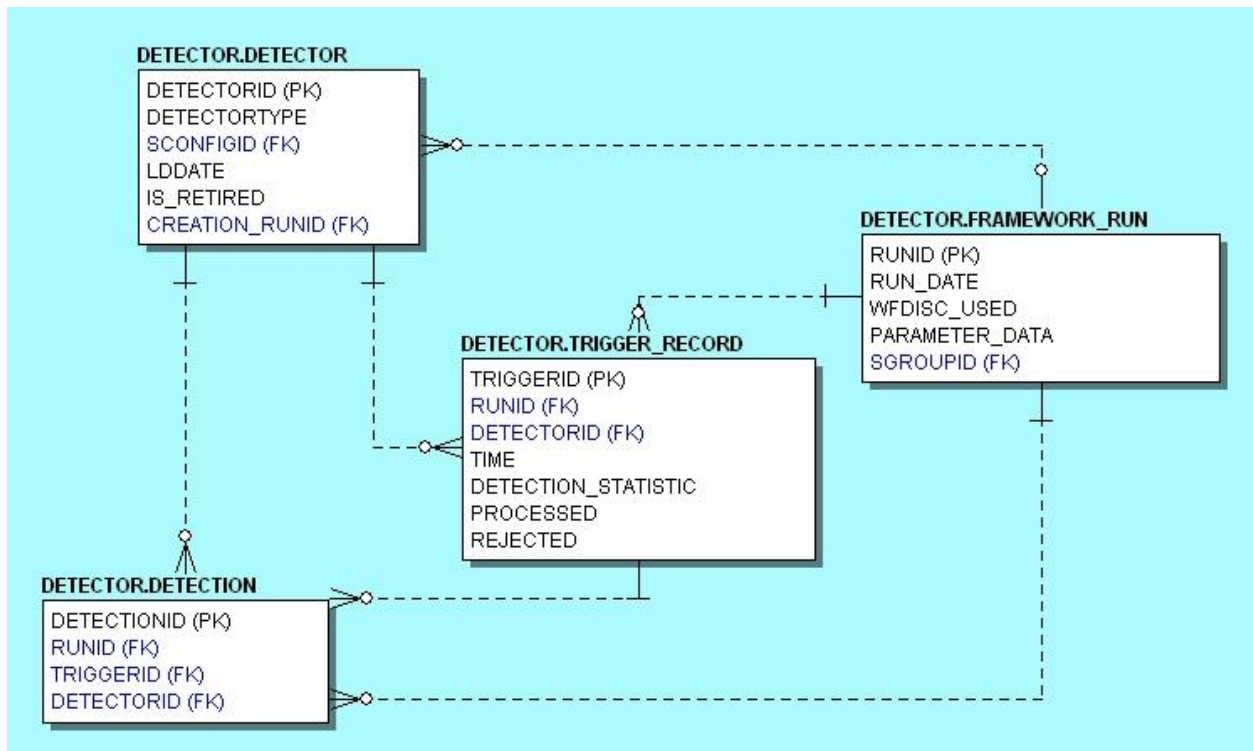


Figure 43: Tables Used to Store Results from Previous Runs of the Framework

The FRAMEWORK_RUN table shown above in Figure 43 contains summary information about each run. Each run of the framework software automatically creates a new row in this table with a unique RUNID. The contents of the parameter file used for the run are stored in the PARAMETER_DATA column as a binary object. This makes it easy to retrieve the exact set of parameters that went into a given run. The DETECTION and TRIGGER_RECORD tables now include RUNID as a foreign key making it easy to form queries for data specific to a particular run or to form queries that compare results from different runs.

Detection statistics are automatically placed in subdirectories named with the corresponding RUNID and DETECTORID. This prevents overwriting of detection statistic files and makes it easier to compare statistics from different runs. As mentioned previously, detection segments are no longer written by the framework so there is no need to manage them.

Operation with Multiple Streams

A stream is a collection of channels assigned to a specific pre-processor and with a collection of dedicated detectors. With this framework it is possible simultaneously to process multiple data streams with each stream processed in its own thread. This is required to support simultaneous operation of detector collections that use different data pre-processing strategies. In addition, it is likely to prove useful in designing algorithms for network detection.

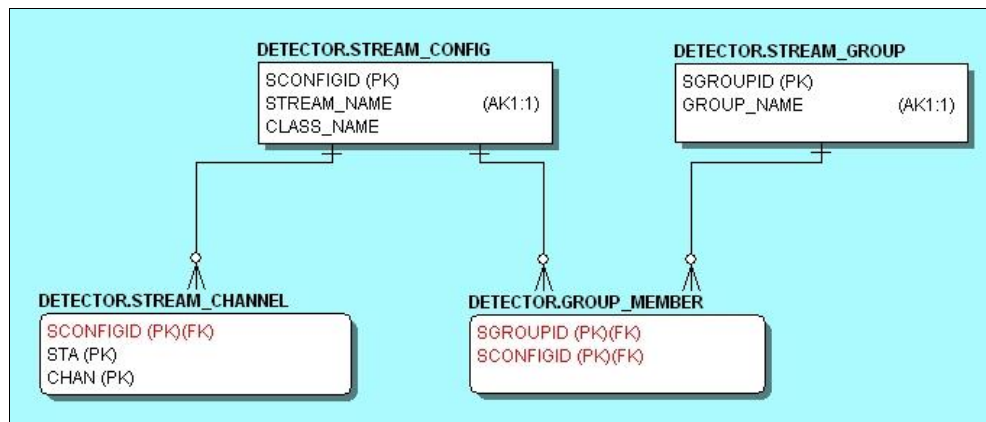


Figure 44: Tables Used to Define Stream Groups

Figure 44 shows the tables involved in managing multiple streams. Streams are defined in the STREAM_CONFIG table and have unique names. Each stream has a single pre-processor identified by CLASS_NAME. Each stream also has a collection of channels which are listed in the STREAM_CHANNEL table. Finally, each stream also has a collection of detectors (not shown) identified by SCONFIGID.

A STREAM_GROUP is a collection of streams that are meant to be processed simultaneously. The framework software requires that a single STREAM_GROUP be specified in the configuration file. In principal, a stream configuration may be used with more than one

group, so a many-to-many relation between STREAM_GROUP and STREAM_CONFIG is defined using GROUP_MEMBER as the association table.

Detector Management

The wideband stream operates with one or more “boot” detectors responsible for spawning new correlation or matched field detectors. There are currently three types of boot detectors defined:

1. STA/LTA
2. ARRAY POWER
3. DETECTION LIST

In addition, the framework supports correlation and matched field detectors. Each of these detector types requires different configuration parameters, so although all detectors have an entry in the DETECTOR table, each has its own configuration table.

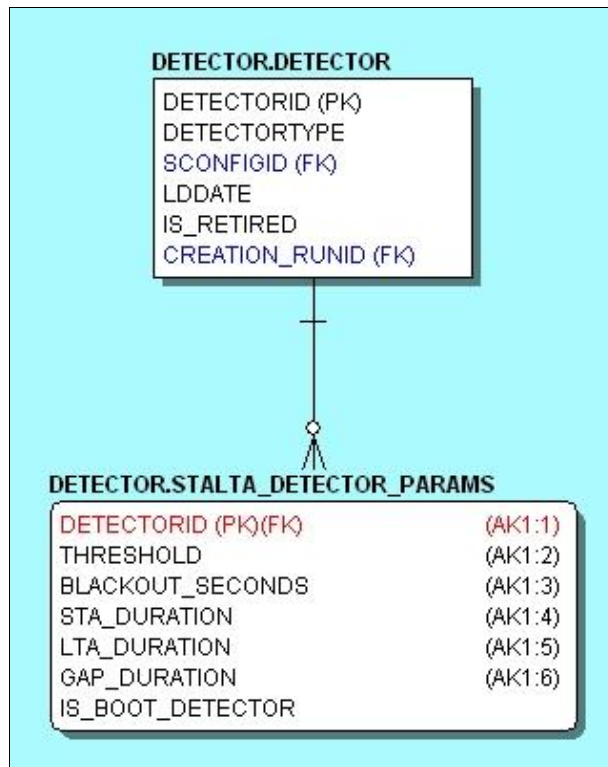


Figure 45: Tables for STA/LTA Boot Detectors

Figure 45 shows the tables used to store information about STA/LTA boot detectors. These may be created manually, but there is a configuration tool that can be used for this task. Currently, these detectors can only be created for a wideband stream. However their detections may be used to create matched field detectors on a compatible narrowband stream. The IS_BOOT_DETECTOR column is somewhat misnamed. This column can contain one of two

values ('y','n'). When set to 'y', the detector will be loaded into the framework when the configuration is run, and it will be used in creation of other detectors.

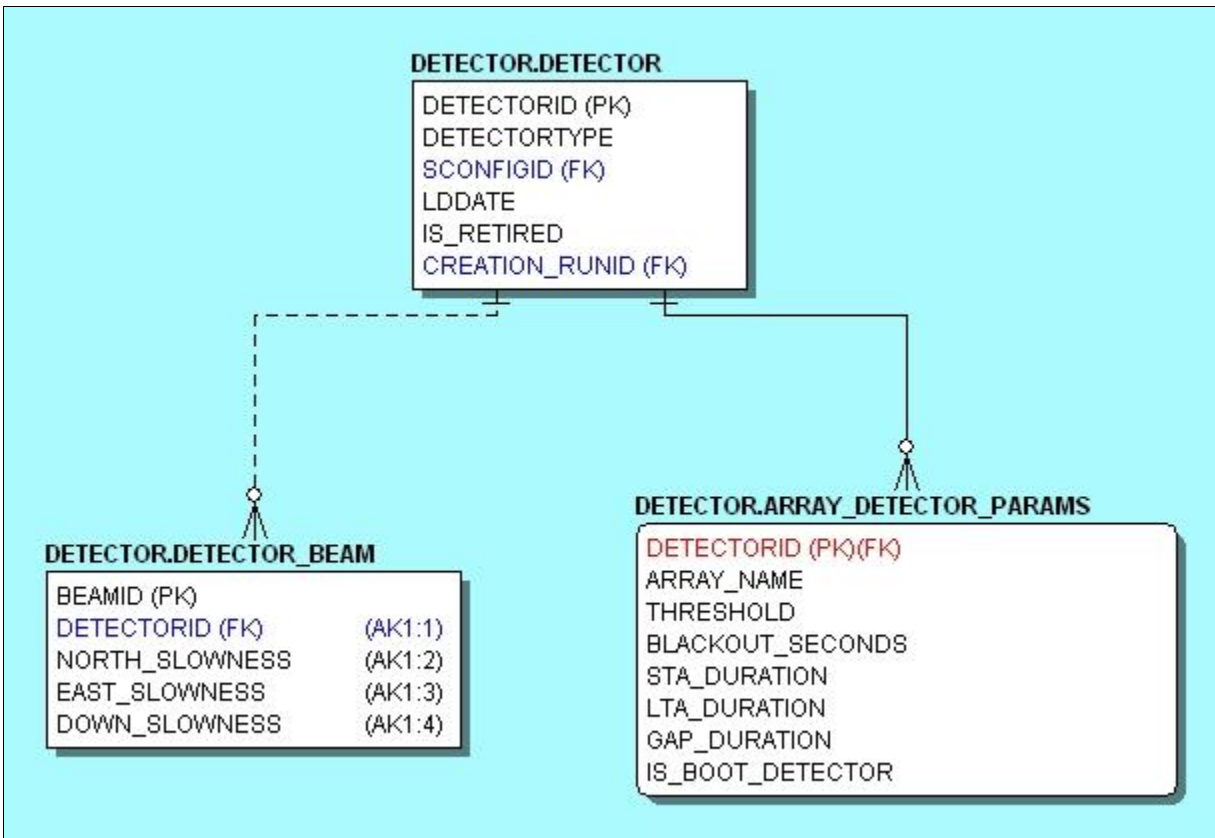


Figure 46: Tables for Array Power Boot Detectors

Figure 46 shows the tables used in configuration of array power boot detectors. These detectors are similar to the STA/LTA detector, but the detection algorithm is applied to one or more beams defined in the DETECTOR_BEAM table.

The detection list detector (Figure 47) is able to generate detections from lists imported from independent processes, for example a parallel pipeline detection algorithm or an event catalog. The catalog is contained in a file with a format like:

Time	SNR	Vel	Phas	Azi	fkp	Qual	freq	Ampl	FKID
2007-001:03.01.15.200	4.6	4.5	S	102.4	0.16	3	2.87	115.8	5072
2007-001:06.55.17.450	4.8	7.5	Pgn	94.6	0.23	3	5.45	83.6	5073
2007-001:06.55.33.812	4.8	4.3	S	92.8	0.31	3	4.29	118.3	5074
2007-001:06.55.42.312	4.2	7.4	Pgn	93.8	0.32	2	3.53	163.2	5075
2007-001:06.55.58.312	7.6	4.2	S	91.5	0.29	2	4.87	316.9	5076
2007-001:06.56.37.562	7.3	4.3	S	92.5	0.39	2	3.18	174.8	5077

A detection is declared at points in the data stream matching the detection time from column 1. The detector can be configured to be more selective by imposing conditions on the values in the remaining columns. For example, by setting the MIN_SNR and MAX_SNR columns in the DETECTION_LIST_DETECTOR_PARAMS table, detections may be formed from the catalog only if their SNR values fall within the defined range. The sgConfig program can be used to create a detection_list detector. However, it currently cannot populate the DETECTION_LIST_PHASE table. This table is used to specify which phases are allowable as detections. Currently, the detection_list detector runs only on the wideband stream. However, its detections may be used to spawn matched field detectors on a compatible narrowband stream.

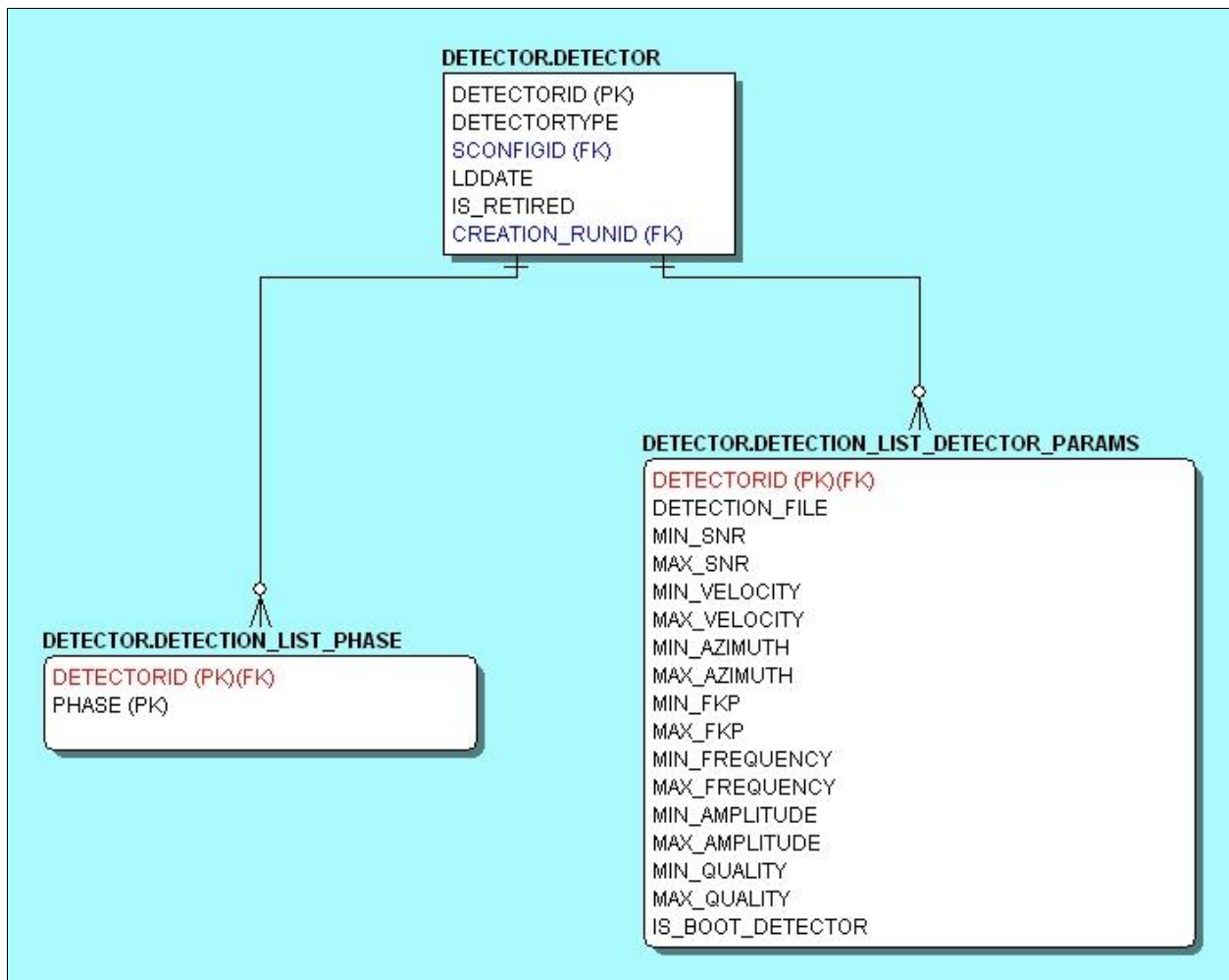


Figure 47: Tables Used by the Detection List Detector

The framework supports automatic creation of subspace detectors on the wideband stream. This behavior is enabled in the configuration file where parameters controlling detector operation also are specified. When a new subspace detector is created by the framework, the values from the configuration file will be copied into a new row in the SUBSPACE_DETECTOR_PARAMS

table (Figure 48) corresponding to the new detector. The template for the new detector will also be written into DETECTOR_TEMPLATE_DIMENSION. The framework can use these data to re-create the detector in subsequent runs of the framework.

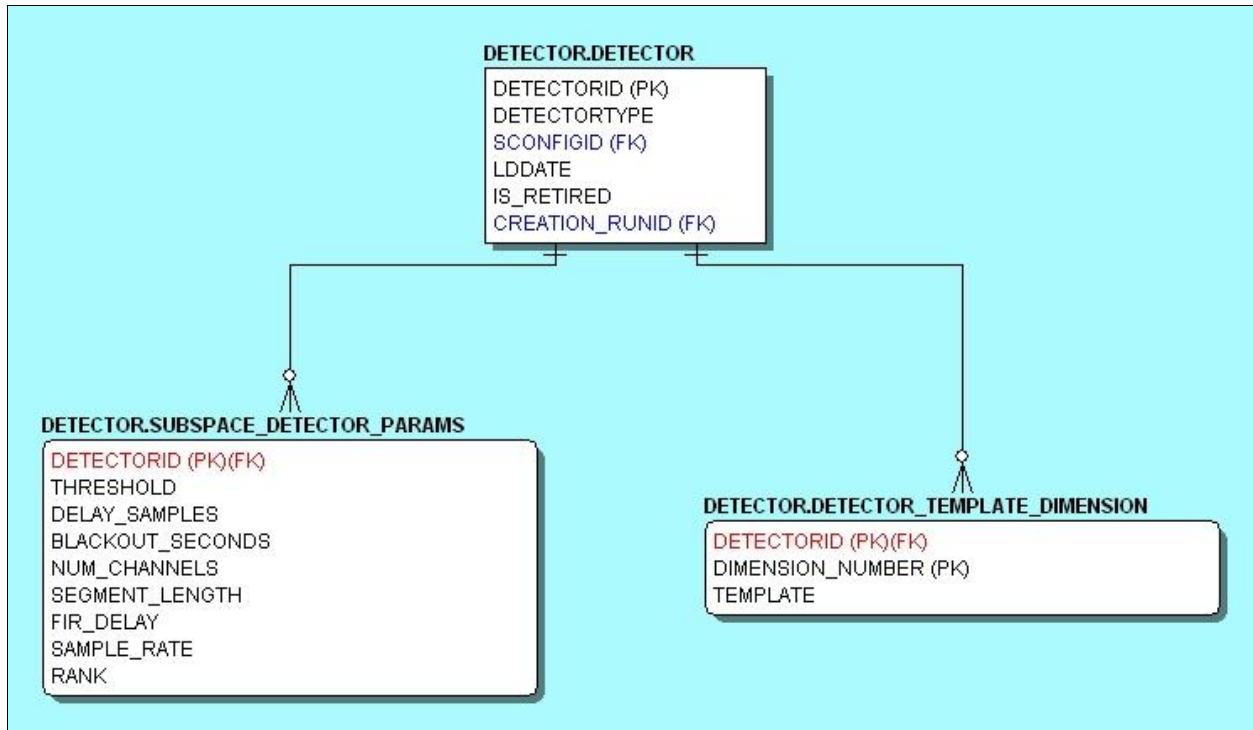


Figure 48: Subspace Detector Tables

It is also possible for the framework to create matched field detectors in response to detections on the wideband stream. The principal limitation of this approach is that incoherent matched field detectors require a lot of memory. Creation of two or three such detectors may cause a memory fault on a 32-bit JVM. Also since we are still experimenting with these detectors we require more control over their generation. Instead, we run the framework in a mode where a single matched field detector is created from a collection of SAC files at the beginning of a run. That detector is used for the entire run although it may evolve as events are detected.

Figure 49 shows the tables used to store matched field detectors. The MF_DETECTOR_PARAMS table holds some parameters used in creating new templates and in managing triggering. The MF_DETECTOR_VERSION table holds different versions of the MFTemplate on which the detector is based. If updating is turned on, each time a matched field detector makes a detection, the detector is updated with the template data for the newly-detected event. At that point, the MF_DETECTOR_VERSION table gets a new row containing the newly-created MFTemplate and a row is added to the MF_DETECTOR_HISTORY table containing the event template data. By storing all this data, it is possible to study the evolution of the detector.

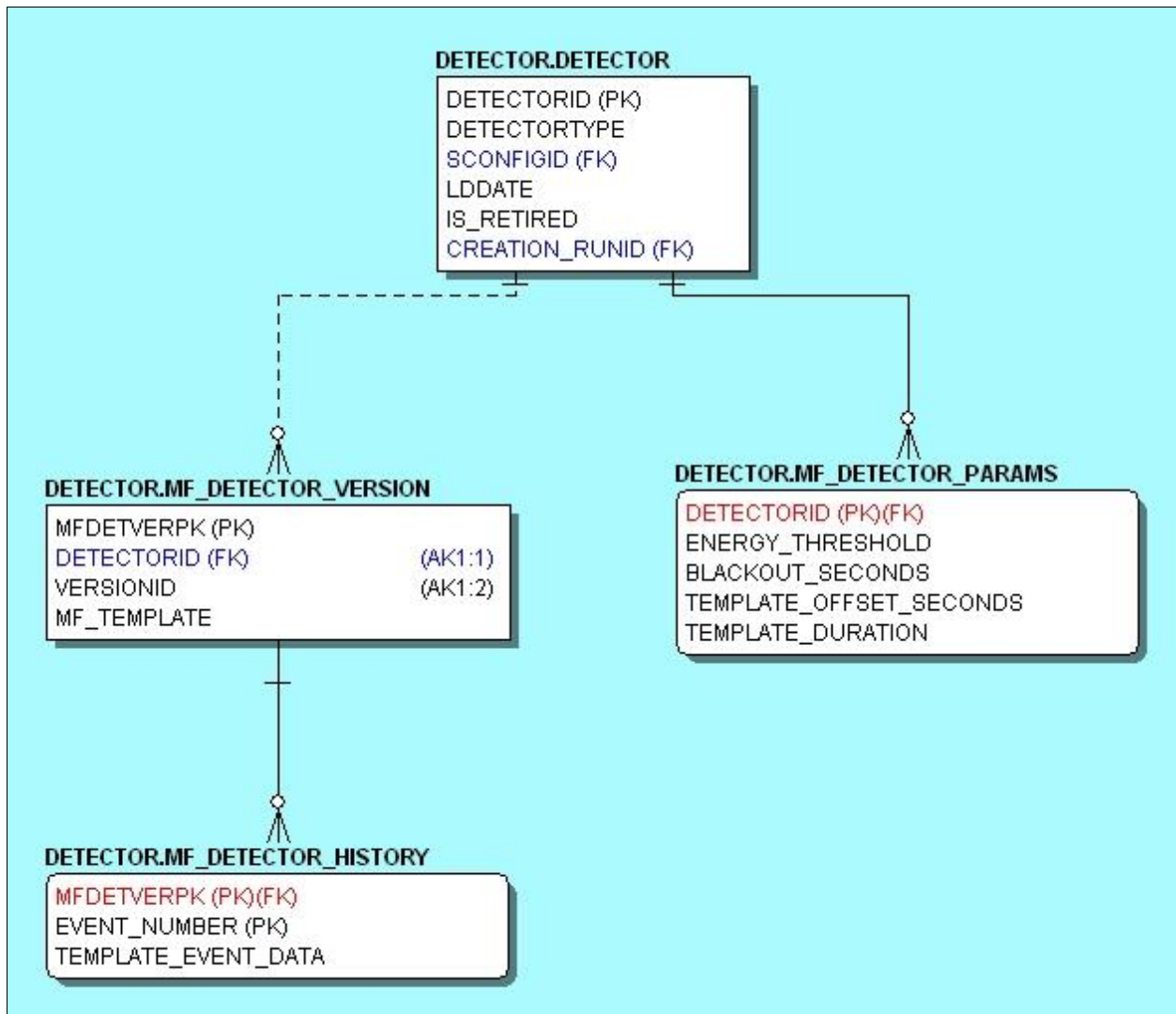


Figure 49: Tables Used for Matched Field Detectors

Matched Field Detector Examples

We now show two examples of matched field detector operation within the detection framework. These were run with the system configuration that permits construction of a matched field detector from flat files followed by application to a long data stream held in the database.

The first example used ARCES observations of a Finnish munitions demolition to define a matched field detector, which was then applied to 25 days of continuous data to detect additional explosions at that source. The detector was permitted to adapt to the source using each newly-detected waveform. This process tested whether the update scheme would perform as demonstrated in the offline tests (Figures 2.16, 2.17, 2.18) when embedded in a live processing

framework actually performing detections. The tests of chapter 2 were conducted with adaptation using pre-cut waveform windows defined from catalog events, and the test we show now used sequential detections made with the adapting template.

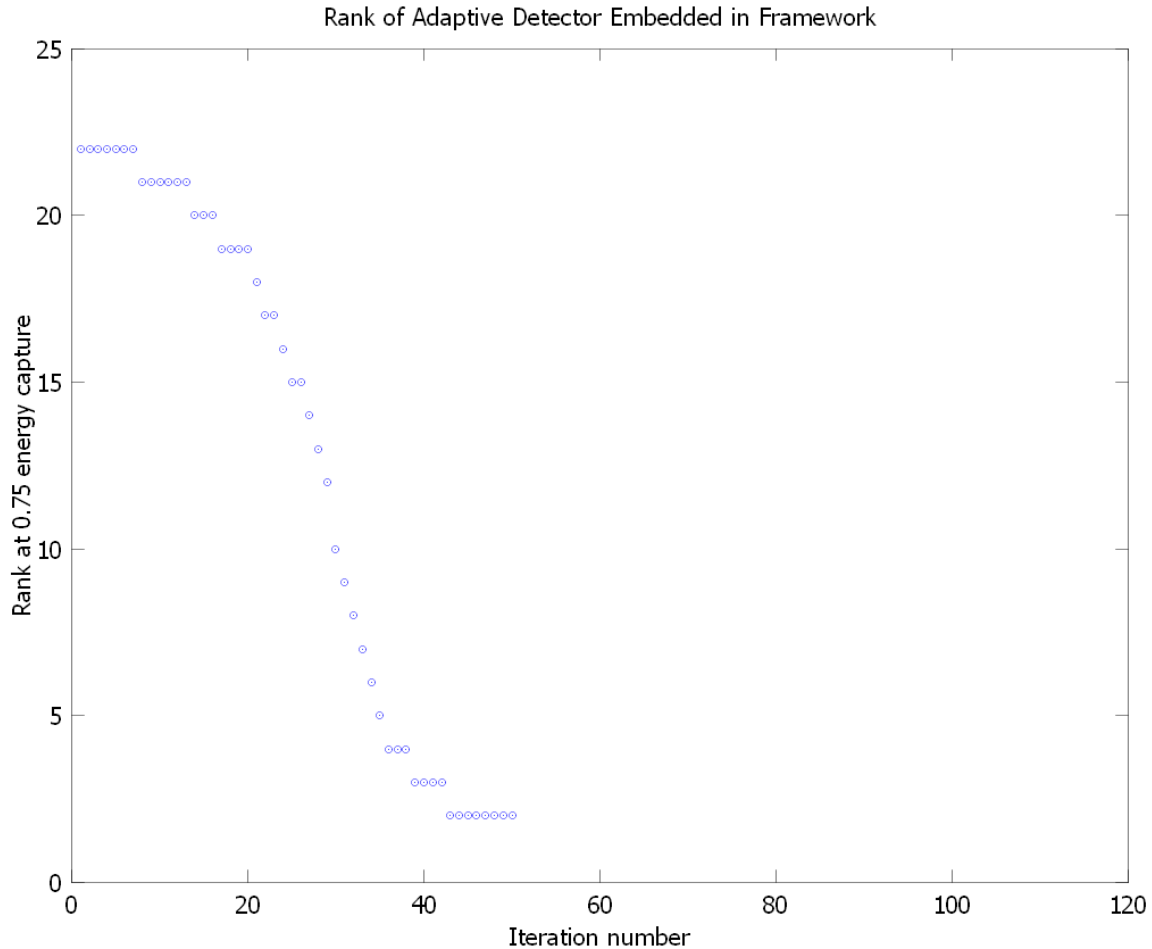


Figure 50: Template Rank as a Function of Iteration Number from Equation 38 with $\theta = 0.75$ for the Embedded Update Algorithm

Figure 50 shows the detector rank as a function of adaptation step. It has much the same behavior as figure 2.16, since the parameters controlling the update process were the same ($\theta = 0.75, \epsilon = 0.1, \lambda = 0.95$). The detector was designed with 33 independent bands, as before beginning at 2.5 Hz and ranging up to 12.5 Hz. With $\theta = 0.75$, the rank begins at 22 and reaches 2 after about 40 events have been detected. Figure 51 shows the change in the detection statistic around the event detections. It begins with the broad peak of the matched field detector and narrows to a spike after 40 steps. The interim statistic at 20 iterations shows the curious hybrid behavior of both a broad peak and a spike. Eventually the broad peak disappears and the detector adapts to resemble a correlator. Figure 52 shows the full evolution of the detector. The transition to a correlation-like detector is not as smooth as in figure 2.18, mostly because the events being detected are changing from detection to detection. Early on it seems that some

events respond better to the narrowband subspace of the template and others to the developing wideband subspace. Eventually the narrowband subspace disappears. This example confirms that the update algorithm works as expected when embedded in the detection framework.

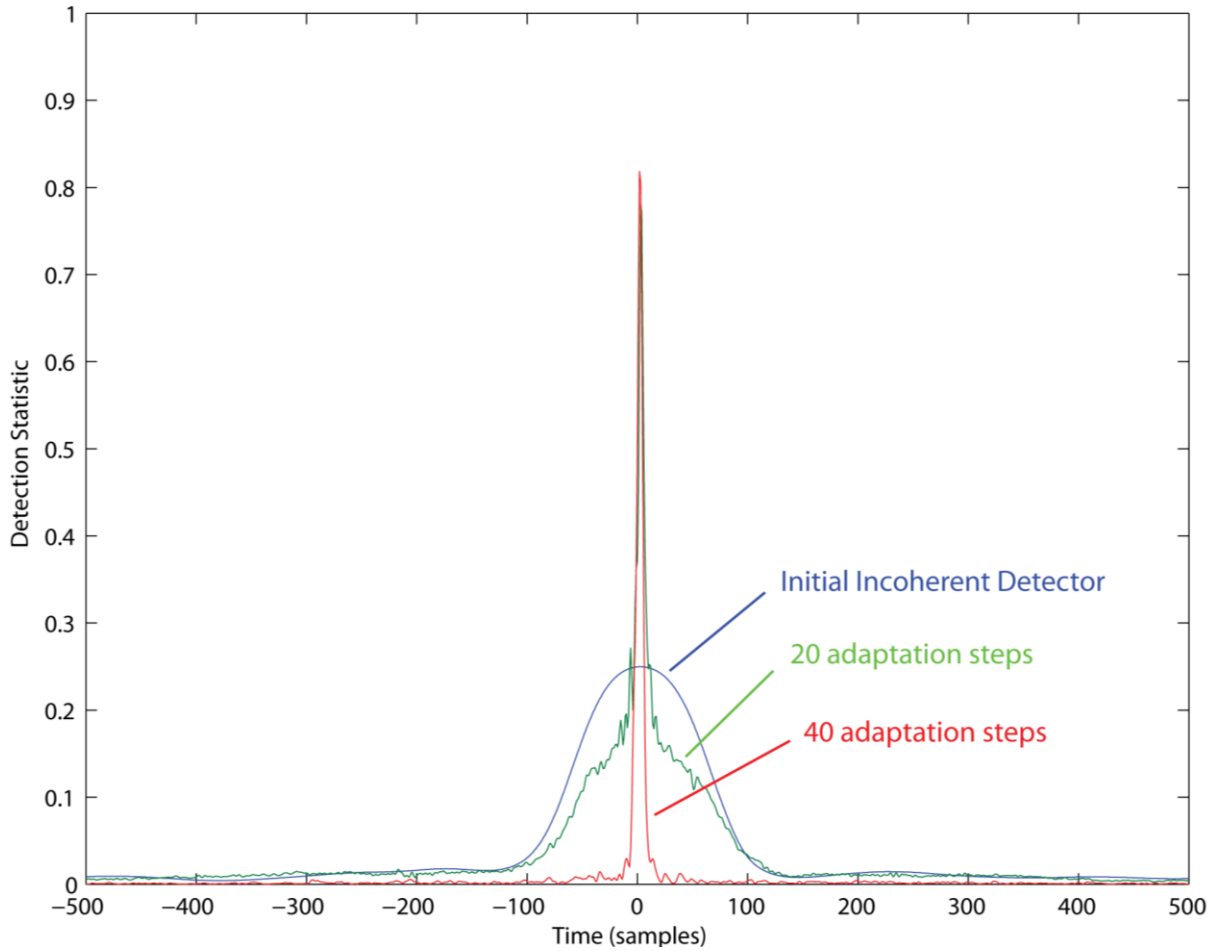


Figure 51: Detection Statistics Around the First, 20th and 40th Events Detected for the Detector Initialized with an Incoherent (Rank 33) Template and Allowed to Adapt as New Events Were Detected

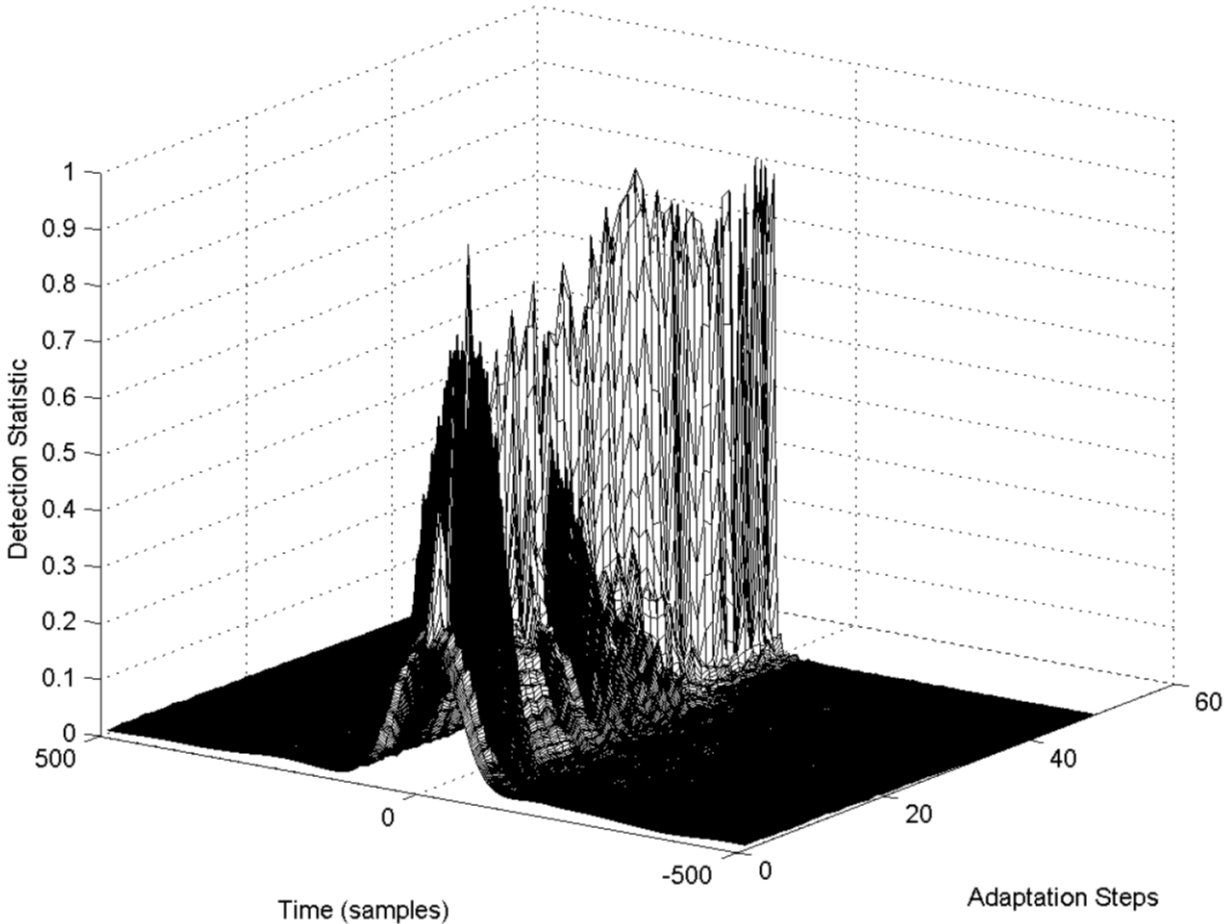


Figure 52: Evolution of the Detection Statistic in the Vicinity of Successive Detections for the Adaptive Detector in the Detection Framework

In figure 52, the time window is 25 seconds in duration. When compared to Figure 2.18, the detection statistic does not make a smooth transition from a broad peak to a spike, because the events being detected are different in each adaptation step. The dual subspace characteristic of the template representation is apparent here as some events respond better to the wideband “spike” subspace and others to the smoother narrowband subspace. Eventually the narrowband subspace disappears.

Our second example is drawn from the Korean peninsula dataset, in particular, detection of the M4 cluster events. We chose one event (2009:274:03:09:13.2) to develop a matched field processor and processed 89 days of data (2009:210 – 2009:299) from the KSAR array, using 11 of the 19 vertical elements (8 elements were found to have significant narrowband background signals that corrupt the templates and artificially inflate the detection statistic). The template length was short (20 seconds), but encompassed the main phases and coda of the observations. The detector was allowed to update with the parameters that we had used with the Finnish munitions explosions ($\theta = 0.75, \epsilon = 0.1, \lambda = 0.95$).

This source behaves differently from the Finnish source, as can be seen in Figure 53. The detector was designed to operate in 19 bands (ranging from 1.875 to 7.5 Hz) and, with $\theta = 0.75$ began at rank 10. The rank actually grew by one before starting to drop and reached equilibrium at rank 8 after 40-50 iterations.

For this run of the detector, we modified the framework to use an adaptive detection threshold with the objective of allowing the threshold to drop with the rank of the detector. The algorithm for determining the threshold examined the detection statistic in a window approximately 820 seconds long, estimating its mean and standard deviation. The mean was subtracted from the detection statistic and this residual was compared against a threshold equal to 10 times the standard deviation to declare detections. This approach produced 116 detections in the 89 day time interval examined. We set a secondary threshold of 0.1 to define detections which were used to update the detector template. With this modification, only 70 detections were used to update the detection template.

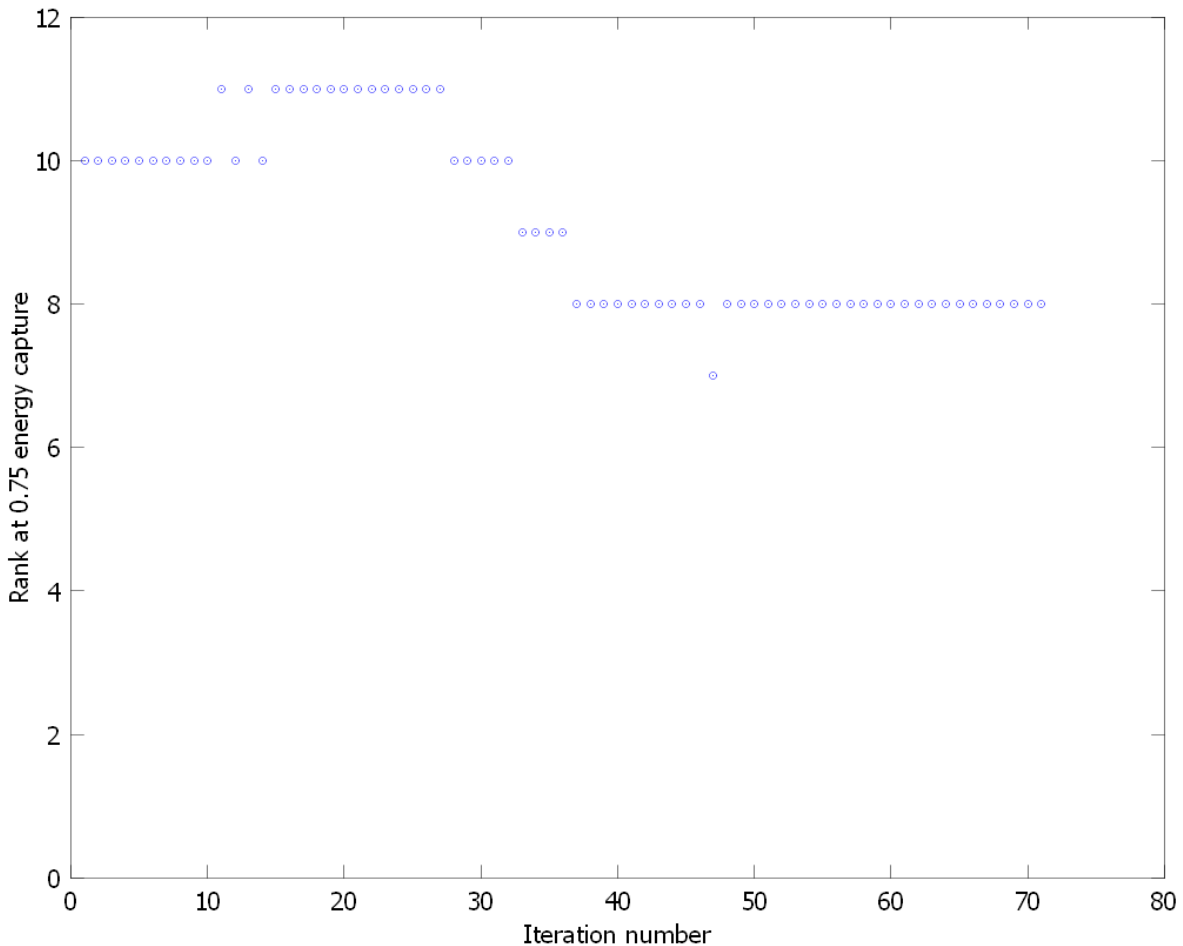


Figure 53: Rank of the M4 Detector as a Function of Adaptation Step – This Source Does Not Tend to Produce Low-Rank Detectors, Perhaps Because of the Low SNR of the Signals

Despite the fact that the rank did not change very much, the detection statistics around the detections do show a change in the character of the detector. Figure 54 shows the detection statistic of the initial detection with the matched field detector, and two subsequent detections after 20 and 40 iterations, respectively. The detector does become more focused, but the detection statistics are lower than in the case of the Finnish munitions source because the M4 events have much lower SNR. Figure 55 shows the complete evolution of the detection statistic, which is much more ragged than Figure 52, again, we think, because of the low SNR of the events. The collection of 70 event waveforms used to update the detector template is shown in Figure 56.

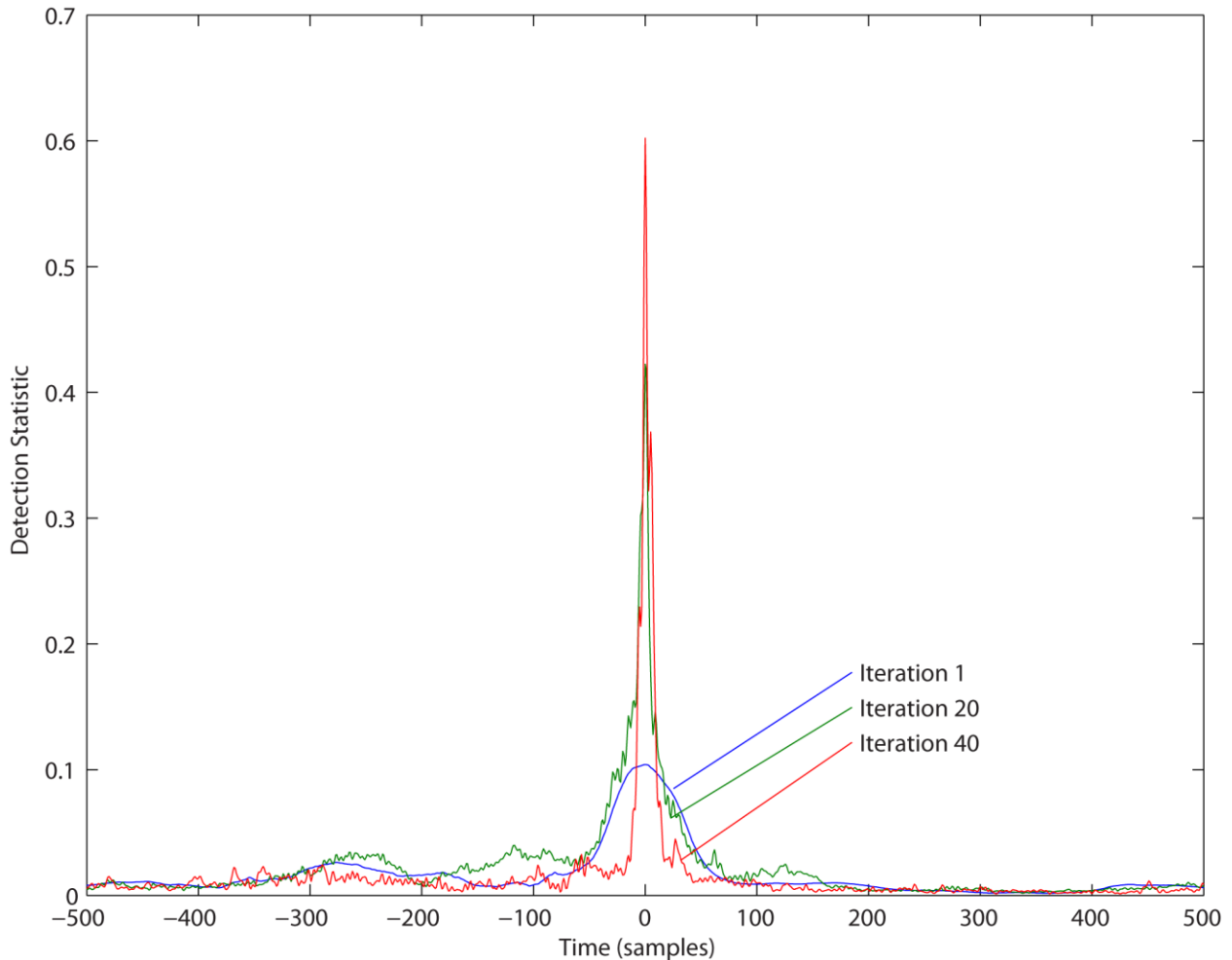


Figure 54: Detection Statistics Around the Detection Times of the First, 20th and 40th Detections for the M4 Detector

As seen in Figure 54, the detection statistics do change from a broad peak to a narrower peak, but not so narrow as was the case for the Finnish munitions source. The detection statistics are low because the events have relatively poor signal to noise ratio.

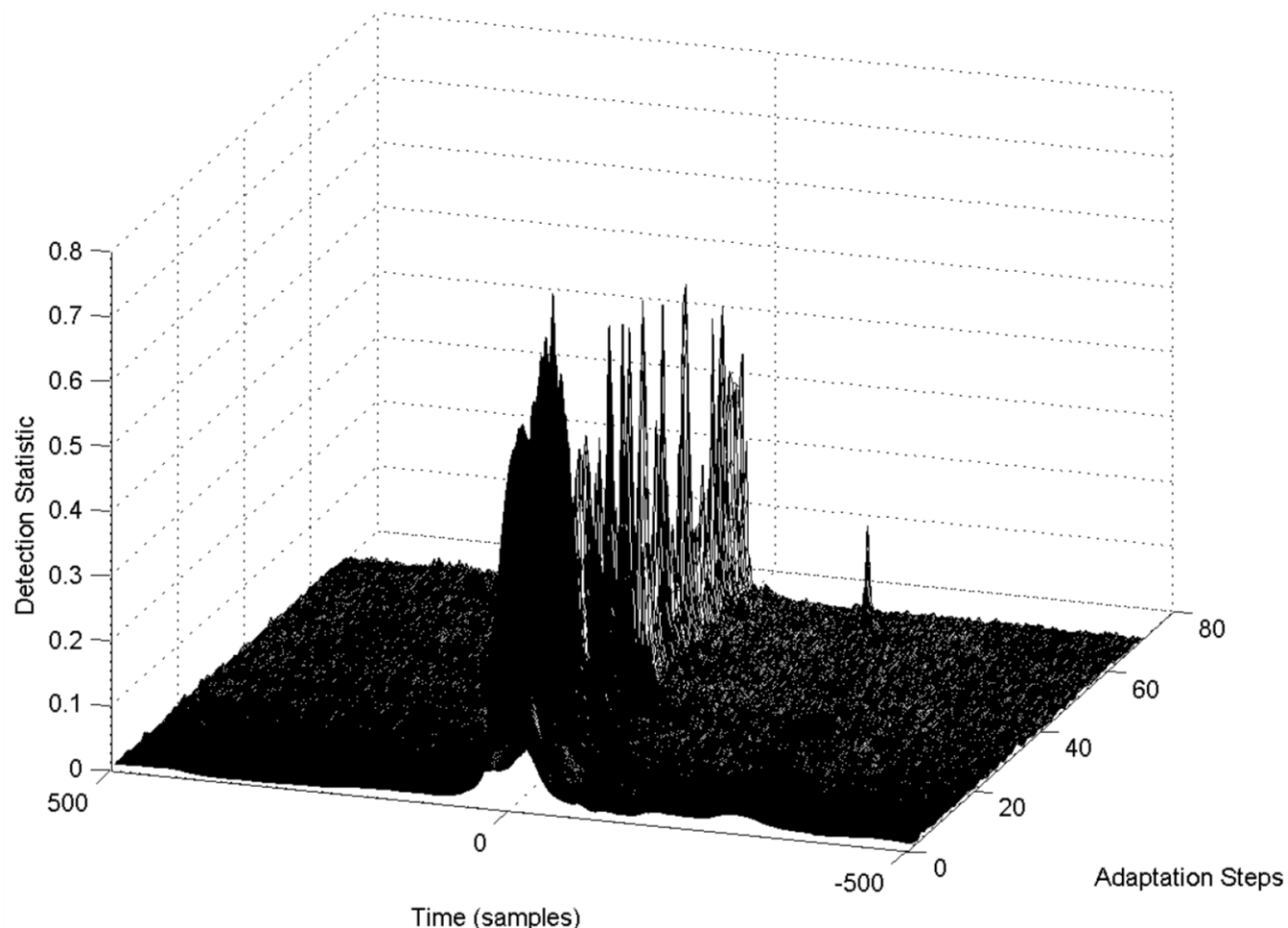
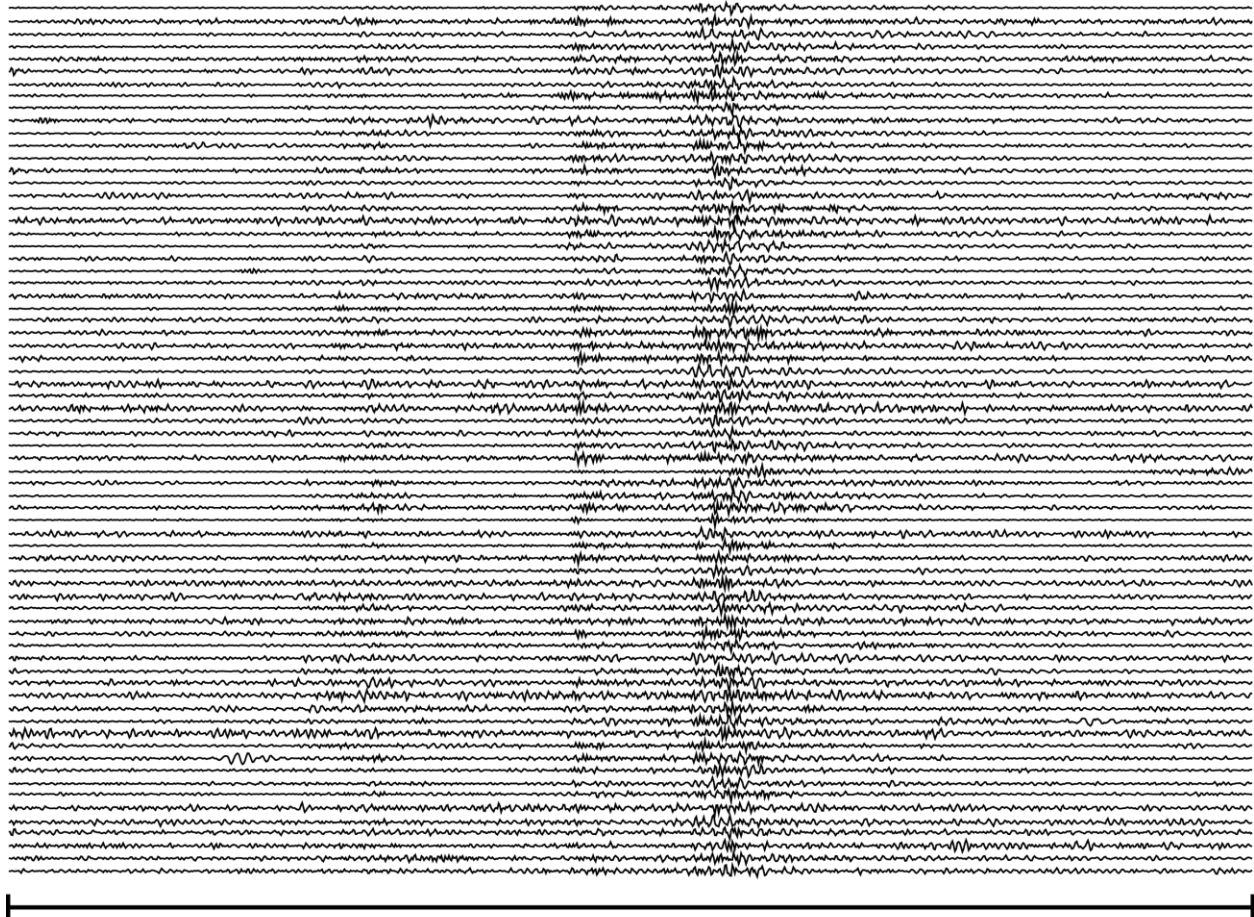


Figure 55: Evolution of the Detection Statistic for the M4 Detector – the Detection Statistic Is Ragged Because of the Poor SNR of the Events – the Narrowband Subspace Disappears After 30 Iterations

While we have not finished research on tuning this algorithm, we think it shows promise for adapting to the characteristics of sources with complex time histories. Figure 57 compares detections made by the adapting framework detector with those made by the correlation detector used to discover and characterize the M4 source (as described in Chapter 3). The framework detector begins as a matched field processor, which necessarily has a detection statistic with a higher noise floor than a correlator. Our adaptive algorithm for setting the detection threshold raises the threshold initially, which results in many of the smaller M4 events being missed by the framework detector. However, as the detector “burns in” with adaptation using the “cleaner” events with larger detection statistics, it captures more of the events found by the correlator (there are still misses). In addition, it finds some events missed by the correlator. We are not yet convinced by this example that this adaptive framework detector is a good substitute for a well-

designed correlation detection algorithm, but the adaptation algorithm does seem to perform as anticipated.



50 seconds

Figure 56: The 70 M4 Detections That Were Used in the Template Adaptation Process – Note That Very Weak P Phase(S), Sn and Lg Are Present – One Channel of the KSRS Array Is Shown Filtered Into the Detection Band (2-8 Hz)

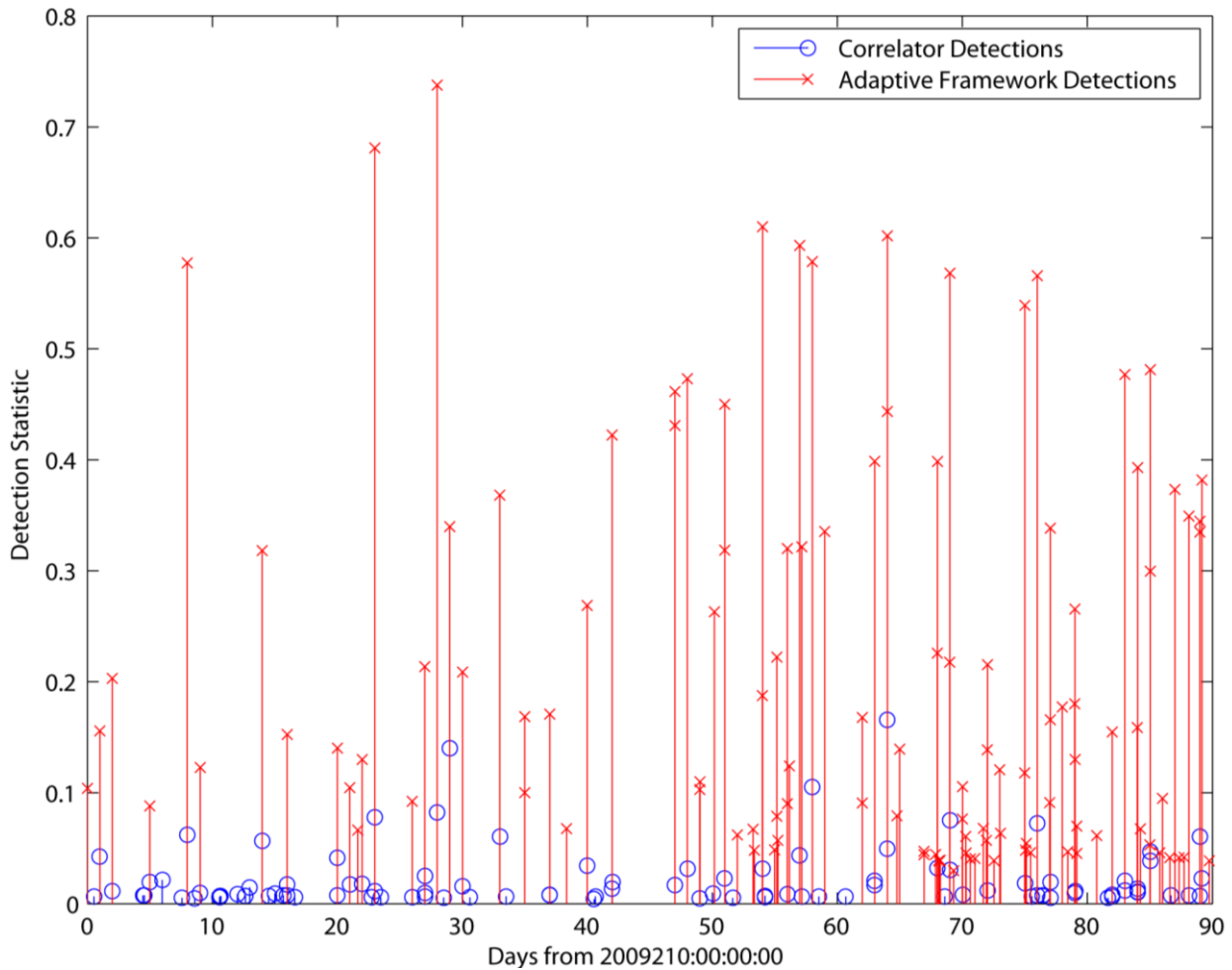


Figure 57: Comparison of Correlator Detections (Blue) of M4 Events and Adaptive Framework Detections (Red)

As seen in Figure 57, at first the adaptive framework detector misses many of the lower SNR events found by the fixed correlator (up through about 17 days) but as the detector burns in it captures larger numbers of the events defined by the correlator (and even some missed by the correlator).

6. Conclusions

In this project we have demonstrated that it is possible to generalize the single-phase matched field processing template developed in our earlier contract to a template constructed from an entire seismogram. The general template is multichannel and multidimensional, with the number of dimensions being controlled by the number of narrow bands used to decompose the wideband master event waveform. This template can be used as a basis for a detection algorithm implemented in a subspace detector framework. Furthermore, the framework can represent

correlation detectors and the incoherent (narrowband) matched field detectors and gradations between the two. We have shown that detectors in the framework can be made to adapt, as new event waveforms become available, through a subspace update mechanism. A detector that begins as a matched field detector will adapt to resemble a correlation detector, if the source produces highly repeatable waveforms. We have demonstrated this form of adaptation with a highly repeatable source in Fennoscandia, and we have also demonstrated somewhat different adaptive behavior that improves detection performance with a mining explosion source in the Korean peninsula. We have demonstrated that the performance of the matched field detector can improve dramatically as the number of available sensors increases, and have demonstrated the applicability of matched field and subspace detectors to far larger apertures over which conventional notions of waveform coherence do not apply.

We have not yet proven conclusively that matched field detectors will perform substantially better than correlation detectors for sources with highly variable source time histories. We think that tests with strongly variable sources such as the large open-pit explosions at the Black Thunder mine in Wyoming or the Kostamuksha mine in Russia will be required to determine whether matched field processors perform better than correlators for these types of sources. We have built a framework that simplifies research on the automated generation and adaptation of detectors that should support such definitive tests.

REFERENCES

- Brand, M., “Fast Low-Rank Modifications of the Thin Singular Value Decomposition,” *Linear Algebra and its Applications*, **514**(1), 2006, pp. 20-30.
- Gibbons, S.J. & Ringdal, F., “The Detection of Low Magnitude Seismic Events Using Array-Based Waveform Correlation,” *Geophys. J. Int.*, **165**, 2006, pp. 149–166.
- Gibbons, S. J., Ringdal, F., and Kväerna, T., “Detection and Characterization of Seismic Phases Using Continuous Spectral Estimation on Incoherent and Partially Coherent Arrays,” *Geophys. J. Int.*, **172**, 2008, pp. 405-421.
- Harris, D. B., *Subspace Detectors: Theory*, UCRL-TR-222758, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94550-9234, 2006a.
- Harris, D. B., *Subspace Detectors: Efficient Implementation*, UCRL-TR-223177, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94550-9234, 2006b.
- Harris, D. B., *Covariance Modifications to Subspace Bases*, LLNL-TR-409155 (internal report), Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94550-9234, 2008.
- Harris, D.B. & Dodge, D.A., “An Autonomous System for Grouping Events in Developing Aftershock Sequence,” *Bull. Seism. Soc. Am.*, **101**, 2011, pp. 764-774.
- Kim, K.-H. and Park, Y., “The 20 January 2007 M_L 4.8 Odaesan Earthquake and its Implications for Regional Tectonics in Korea,” *Bull. Seism. Soc. Am.*, **100**, 2010, pp. 1395-1405.
- Schweitzer, J., “Slowness Corrections – One Way to Improve IDC Products,” *Pure Appl. Geophys.* **158**, 2001, pp. 375-396.
- Wessel, P. and W. H. F. Smith (1995), “New version of the Generic Mapping Tools released,” *Eos Trans. AGU*, **76**, 1995, p. 329.

DISTRIBUTION LIST

DTIC/OCP 8725 John J. Kingman Rd, Suite 0944 Ft Belvoir, VA 22060-6218	1 cy
AFRL/RVIL Kirtland AFB, NM 87117-5776	2 cys
Official Record Copy AFRL/RVBYE/Robert Raistrick	1 cy