

COMPREHENSIVE SEISMIC DETECTION AND ESTIMATION USING MATCHED FIELD PROCESSING

Steven J. Gibbons, et al.

**NORSAR
PO Box 53
Gunnar Randers vei 15
N-2027 Kjeller, Norway**

18 November 2017

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 AFMC/PA 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-RV-PS-TR-2018-0003 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED//

Dr. Frederick Schult
Program Manager, AFRL/RVBYE

//SIGNED//

Dr. Thomas R. Caudill, Acting Chief
AFRL Battlespace Environment Division

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 (DD-MM-YYYY) 18-11-2017		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 08 Oct 2015 – 18 Oct 2017	
Comprehensive Seismic Detection and Estimation Using Matched Field Processing				5a. CONTRACT NUMBER FA9453-16-C-0006	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 62601F	
6. AUTHOR(S) Steven J. Gibbons, David B. Harris ¹ , Tormod Kværna, and Douglas A. Dodge ²				5d. PROJECT NUMBER 1010	
				5e. TASK NUMBER PPM00027855	
				5f. WORK UNIT NUMBER EF128707	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NORSAR, PO Box 53 Gunnar Randers vei 15 N-2027 Kjeller, Norway ¹ Deschutes Signal Processing, LLC, 81211 E. Wapinitia Road, Maupin, OR 97037 ² Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 9455				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory Space Vehicles Directorate 3550 Aberdeen Avenue 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-2018-0003	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited (AFMC-2018-0112 dtd 20 Jun 2018)					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Empirical Matched Field Processing (EMFP) is a narrowband method for signal detection and estimation which recognizes a characteristic phase and amplitude structure over a sensor array. EMFP can provide enhanced detection capability for signals of special interest from a target source region and we here explore its application as a comprehensive front-end detector for general seismic signals. On small aperture seismic arrays, we demonstrate that a relatively small number of matched field templates are required to cover the parameter space of anticipated signals. In addition, estimates of apparent velocity and back azimuth made directly from the empirical steering vectors are likely to be more accurate than classical f-k estimates since bias from path and near-receiver characteristics is corrected for empirically. On larger aperture arrays, more templates are required to cover parameter space. On seismic arrays for which signal coherence in the frequency band of interest is poor, a given matched field detector will provide a very signal-specific detector and vast numbers of templates are needed for a comprehensive detector. We demonstrate how classical slowness estimates for high frequency signals can be enhanced by considering only selected sensor pairs. Similarly, by limiting the influence of correlations between the most distant sensors in the generation of the matched field templates, the footprint of matched field detectors may be broadened to be less source-specific. EMFP detectors can confidently identify characteristic phase shifts for low SNR signals and we demonstrate detection of the small event at or close to the DPRK nuclear test-site using matched field templates from previously observed events. The path and tectonic setting of the source can also determine the variability of matched field templates as a function of source location, particularly if longer time-windows and 3-component process are used. This has been demonstrated using both empirical observations and computation of synthetic waveforms for long-period surface waves.					
15. SUBJECT TERMS nuclear explosion monitoring, signal detection, seismic monitoring, array processing					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			Dr. Frederick Schult
			Unlimited	120	19b. TELEPHONE NUMBER (include area code)

This page is intentionally left blank.

Table of Contents

List of Figures.....	ii
Acknowledgements.....	vi
Summary.....	1
1 Introduction	2
2 Methods, Assumptions, and Procedures.....	13
2.1 An Anatomy of a Matched Field Detector.....	14
2.2 Parameter Estimation Using Empirical Matched Field Processing.....	29
2.3 Parameter Estimation on Seismic Arrays Using Limited Sensor Pair Correlation	39
2.4 Matched Field Detection: A Case Study. The North Korea Nuclear Test-Site	43
2.5 The Physical Basis for Matched Field Processing	50
2.5.1 Introduction.....	50
2.5.2 Data and Methods.....	51
3 Results and Discussions	57
3.1 Anatomy of a Matched Field Detector	57
3.2 Parameter Estimation on Seismic Arrays and Consequences for Matched Field Detectors.	59
3.3 Limited Sensor Pair Correlation.....	79
3.4 Case Study: The North Korea Nuclear Test-Site	82
3.5 The Physical Basis for Matched Field Processing	86
4 Conclusions	102
References.....	105
List of Abbreviations	109

List of Figures

Figure 1: Seismic signal detectors, for known and unknown signals	2
Figure 2: The relationship between the geometry of a seismic array and its co-array	3
Figure 3: Phase differences between the ground motions at different sites of an array represented by colored symbols at the different sites of the co-array	4
Figure 4: A graphical display of the comparison between the covariance matrix of an unknown wavefront with two possible phase and coherence patterns – one theoretical and one empirical	5
Figure 5: An illustration of the procedure to generate a detection statistic trace for a single matched field “beam” under an ongoing aftershock sequence, based on a single P-phase.....	7
Figure 6: Matched field detection of the signal from a DPRK nuclear test (I).....	8
Figure 7: Matched field detection of the signal from a DPRK nuclear test (II).....	9
Figure 8: Detection of the signal from the 2013 DPRK nuclear test on the KSRS array using matched field templates as indicated	10
Figure 9: Schematic representation of a matched field beam deployment on two arrays both with general beams and empirical beams aimed at specific wavefronts	11
Figure 10: Demonstration of how empirically calibrated f-k analysis results in an increased array gain (in particular at higher frequencies) and parameter estimates that are more consistent between the different frequency bands.....	12
Figure 11: “Mock helicorder plot” of a single channel short period vertical seismometer on the KKAR array.....	14
Figure 12: Events in the 8 October 2005 Kashmir earthquake aftershock sequence (red symbols) in relation to the KKAR (Karatau) array in Kazakhstan at a distance of approximately 9 degrees	15
Figure 13: Three events in the 8 October 2005 Kashmir earthquake sequence which demonstrate why waveform correlation may be a poor signal classifier for such events.....	16
Figure 14: Narrowband f-k spectra (2 Hz and 4 Hz) for the Pn arrival at KKAR for the three events displayed in Figure 13	17
Figure 15: Calculation of narrowband steering vectors using theoretical (plane-wave) covariance matrices (a) and empirical covariance matrices (b)	19
Figure 16: Narrowband covariance matrices for the frequencies as shown displayed with the phase (color) and coherence (symbol size) plotted as a function of the vector displacement between the two sensors (i.e. at the coordinates of the co-array).....	20

Figure 17: Calculation of matched field “pseudospectrograms” from an array data stream.....	22
Figure 18: The full matched field pseudospectrogram for the empirical steering vector calculated for the Kashmir mainshock Pn arrival for a 20 minute long time interval on October 8, 2005 (top panel)	24
Figure 19: Matched field pseudospectrograms (both direct and transformed) for the four array stations KKAR, BVAR, WRA, and ILAR	26
Figure 20: Steps to transform the matched field pseudospectrogram to a scalar detection trace	28
Figure 21: The northern mid-Atlantic ridge	29
Figure 22: Location estimates for 971 events using the Bayesloc program and both regional and teleseismic phases	31
Figure 23: Comparison of event location estimates	32
Figure 24: Consistency of time-residuals for Bayesloc origin estimates.....	33
Figure 25: Robustness of Bayesloc event locations to choice of regional velocity model	34
Figure 26: Deviation in backazimuth estimates using classical plane-wave models at SPITS	35
Figure 27: Display of potential utility of empirical matched field processing for providing an empirically corrected direction estimate for a detected phase arrival at SPITS	36
Figure 28: Study region in the Far East for the USRK, KSRS and MJAR seismic arrays.....	37
Figure 29: Geometries of the USRK, KSRS, and MJAR seismic arrays	38
Figure 30: Observation of a magnitude 4.5 earthquake near to the city of Donetsk recorded on the large aperture AKASG array (Malin) in Ukraine	40
Figure 31: f-k analysis for the AKASG arrival on correlation trace virtual co-array	41
Figure 32: Slowness estimation for the teleseismic P-arrival at NOA from the declared October 9, 2006, underground nuclear test in North Korea, using 2-4 Hz f-k analysis on the correlation co-arrays for different subsets of sensors.	42
Figure 33: The five seismic events at or very close to the North Korea nuclear test between September 3, 2017, and October 12, 2017, recorded on station IC.MDJ.....	43
Figure 34: Matched field detection statistics for a ten minute long data window on the KSRS array on October 12, 2017, for Pn templates as indicated	44
Figure 35: Matched field detection statistics for a ten minute long data window on the KSRS array on October 12, 2017, for Pn templates as indicated	45
Figure 36: Matched field detection statistics for a ten minute long data window on the KSRS array on October 12, 2017, for Pn templates as indicated	46
Figure 37: The five seismic events at or very close to the North Korea nuclear test between September 3, 2017, and October 12, 2017, recorded on the IMS station USRK (Russian Federation).....	47

Figure 38: Matched field detection statistics for a ten minute long data window on the KSRS array on October 12, 2017, for Pn templates as indicated	49
Figure 39: Study regions for this part of the project	51
Figure 40: Global CMT project mechanisms for the two study regions.....	52
Figure 41: Preprocessing steps for estimating the matched field template at 35 seconds period from a Rayleigh wave window.....	55
Figure 42: Scalar detection statistics for a segment of data from the KKAR array attempting to detect P-wave arrivals from earthquakes in the region surrounding the 8 October 2005 Kashmir main shock.	58
Figure 43: Geographical matched field direction estimates on the SPITS array.....	61
Figure 44: Geographical matched field direction estimates on the ARCES array	62
Figure 45: Backazimuth estimates for first P-arrivals at SPITS from the Bayesloc-located events displayed in Figure 22 b) using (a) classical f-k analysis and (b) a matched field comparison with the templates evaluated from each of the other events.....	64
Figure 46: Provisional zone map for events along the ridge between 71 and 84 degrees north	66
Figure 47: Beam deployment in slowness space.....	67
Figure 48: The construction of ensemble covariance matrices for first P-arrivals at SPITS for events from two of the clusters, obtained by calculating median values of each real and each imaginary value over all events within one of the given zones.	68
Figure 49: Time-frequency displays of matched field statistics for “beams” based upon the empirical steering vectors calibrated for first P-arrivals for ensembles for the zones of seismicity indicated.....	70
Figure 50: Plane-wave f-k analysis for Pn and Sn phases for an event on the ridge to the North West of the SPITS array	71
Figure 51: Evaluation of slowness vectors on the vertical, radial and transverse components at the SPITS array for the October 11, 2010, Novaya Zemlya earthquake	73
Figure 52: Geographical matched field direction estimates on the USRK array.....	75
Figure 53: Backazimuth and apparent velocity measured on the USRK array for events displayed in Figure 52, both using analyst-measured estimates (as reported in the Reviewed Event Bulletin of the IDC) and using the predicted values from the best-fit matched field statistics.	76
Figure 54: Geographical matched field direction estimates on the KSRS array.....	77
Figure 55: Geographical matched field direction estimates on the MJAR array.....	78
Figure 56: Direction estimates at ARCES for the Pn, Sn, and Lg phases from the largest June 24, 2005, earthquake discussed by using 2-5 Hz filtered short-period vertical waveforms.....	80

Figure 57: Vectors of empirical steering vectors for the Pn arrival on the KSRS array for four events at or close to the North Korea nuclear test site.....	83
Figure 58: Matched field statistics for the North Korea events as listed for the KSRS array.....	84
Figure 59: Empirical matched field processing response patterns for sources in regions of relatively homogeneous mechanism show well-behaved drop-off from a maximum at the source location	86
Figure 60: In regions with events of more diverse mechanism, the EMFP response is not monotonic with distance from the source location.....	87
Figure 61: Detail of the mid-Atlantic ridge region with transform faults and spreading centers	88
Figure 62: Vertical (black) and north (red) component waveforms for the two events of Figure 61	89
Figure 63: EMFP responses for an event at the northwestern end of the Zagros (left) and for another at the southeastern end.....	90
Figure 64: Waveforms from events along the mid-Atlantic ridge to the GRSN typically are much simpler than those along a path from the Zagros.....	91
Figure 65: Detail of the matched field response to event 2013099115249, a magnitude 6.4 event in the central Zagros.....	92
Figure 66: Domain for spectral element synthetics calculations for the north mid-Atlantic ridge events.....	93
Figure 67: Comparison of the S2.9ea synthetics and observations in the 25-50 second band for a magnitude 5.5 event on October 6, 2011.....	94
Figure 68: Comparison of the Z components of the moment tensor (black) superimposed by their finite-difference approximations (red)	96
Figure 69: Comparison of the N components of the moment tensor (black) superimposed by their finite-difference approximations (red)	97
Figure 70: Comparison of the E components of the moment tensor (black) superimposed by their finite-difference approximations (red)	98
Figure 71: Comparison between actual empirical matched field responses to data from event 2013326193058 (location indicated by the white star) and the responses predicted by the synthetics using mechanisms obtained from the Global CMT Project [Dziewonski et al., 1981; Ekström et al., 2012].	100
Figure 72: Comparison of synthetic MFP responses to event 2013326193058 under two different mechanism scenarios.....	101

Acknowledgements

Seismic data from many stations of the International Monitoring System (IMS) was obtained from the International Data Center (IDC) of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) in Vienna.

The data from the KKAR-array is now available from the IRIS Data Management Center (DMC) at http://ds.iris.edu/SeismiQuery/by_station.html (network KZ).

Summary

Empirical Matched Field Processing (EMFP) is a narrowband method for signal detection and estimation which recognizes a characteristic phase and amplitude structure over a sensor array. EMFP can provide enhanced detection capability for signals of special interest from a target source region and we here explore its application as a comprehensive front-end detector for general seismic signals. On small aperture seismic arrays, we demonstrate that a relatively small number of matched field templates are required to cover the parameter space of anticipated signals. In addition, estimates of apparent velocity and backazimuth made directly from the empirical steering vectors are likely to be more accurate than classical f-k estimates since bias from path and near-receiver characteristics is corrected for empirically. On larger aperture arrays, more templates are required to cover parameter space. On seismic arrays for which signal coherence in the frequency band of interest is poor, a given matched field detector will provide a very signal-specific detector and vast numbers of templates are needed for a comprehensive detector. We demonstrate how classical slowness estimates for high frequency signals can be enhanced by considering only selected sensor pairs. Similarly, by limiting the influence of correlations between the most distant sensors in the generation of the matched field templates, the footprint of matched field detectors may be broadened to be less source-specific. EMFP detectors can confidently identify characteristic phase shifts for low SNR signals and we demonstrate detection of the small event at or close to the DPRK nuclear test-site using matched field templates from previously observed events. The path and tectonic setting of the source can also determine the variability of matched field templates as a function of source location, particularly if longer time-windows and 3-component process are used. This has been demonstrated using both empirical observations and computation of synthetic waveforms for long-period surface waves.

1 Introduction

Methods for detecting seismic signals can be categorized in many ways. One classification of methods is whether or not we know what the signal looks like a priori (Figure 1). Given only a single channel, a signal which we do not know the shape of is detected using a power detector - resulting in a trigger when a sudden increase in amplitudes on a time-series is encountered (see e.g. Withers et al., 1998, for an overview). If we do know the shape of an anticipated signal beforehand, we can use correlation or matched filter techniques (e.g. Gibbons and Ringdal, 2006), or more general subspace detectors (e.g. Harris, 2006; Barrett and Beroza, 2014), which recognize the ripple-for-ripple similarity between the incoming waveform and a template signal. Methods such as Fingerprint and Similarity Thresholding (FAST; Yoon et al., 2015) attempt to make pattern detectors somewhat more general at the price of reduced sensitivity.

When we have an array of sensors (see Figure 2) the distinction is more subtle. The spatial distribution of sensors means that, although we don't know the shape of the actual ripples beforehand, we are looking for patterns of anticipated time-delays or phase shifts between the different sites. Even the delay-and-stack methods of detecting general signals on a seismic array are therefore pattern detectors to a certain degree.

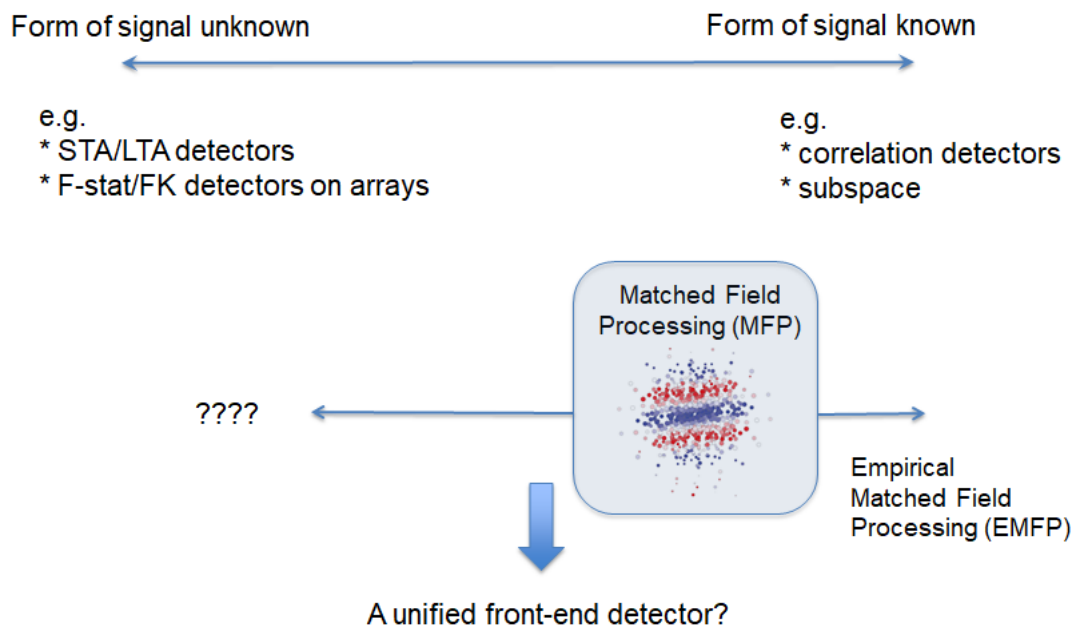


Figure 1: Seismic signal detectors, for known and unknown signals. Can Empirical Matched Field Processing (EMFP) bridge the gap between the detectors for known signals and the detectors for unknown signals in a global monitoring context?

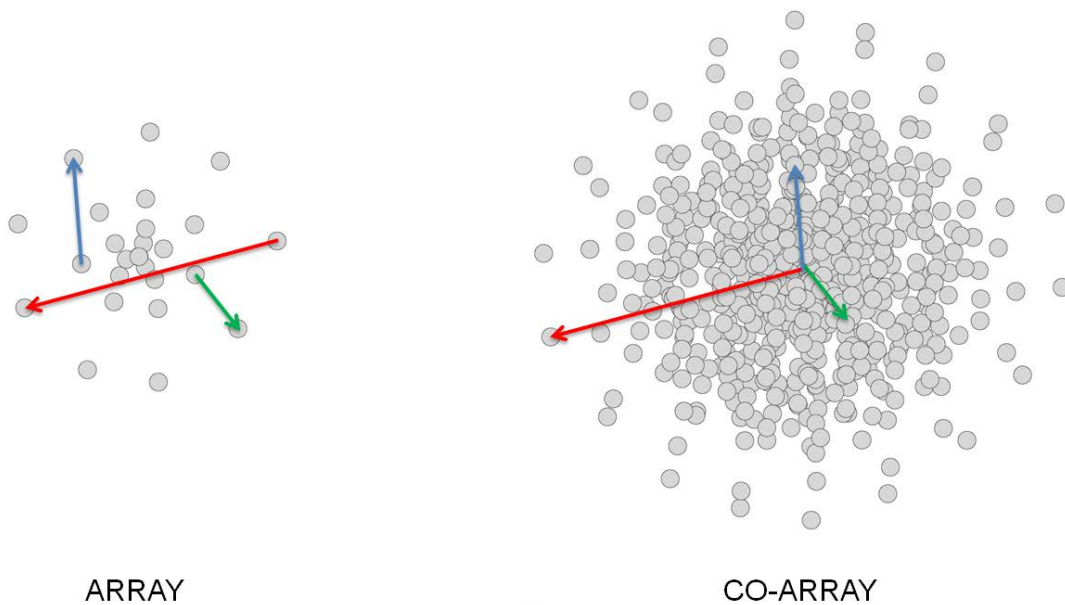


Figure 2: The relationship between the geometry of a seismic array and its co-array. The locations of the elements of the co-array are given by the set of vector differences between each location in the array.

For a given array of sensors, the co-array (Figure 2) is a useful construct since it allows us to visualize the covariance matrix evaluated on a short time-window of data on a seismic array. A typical example is shown in Figure 3, both for an ideal or theoretical signal and for a real-world signal coming from the same direction. For a theoretical plane wavefront (left), the signals across the array aperture are perfectly coherent and the phase-shifts conform exactly to a set of time-delays defined by a steady progression of a plane wavefront over the sensors. On real world seismic arrays (right), the situation is more complicated.

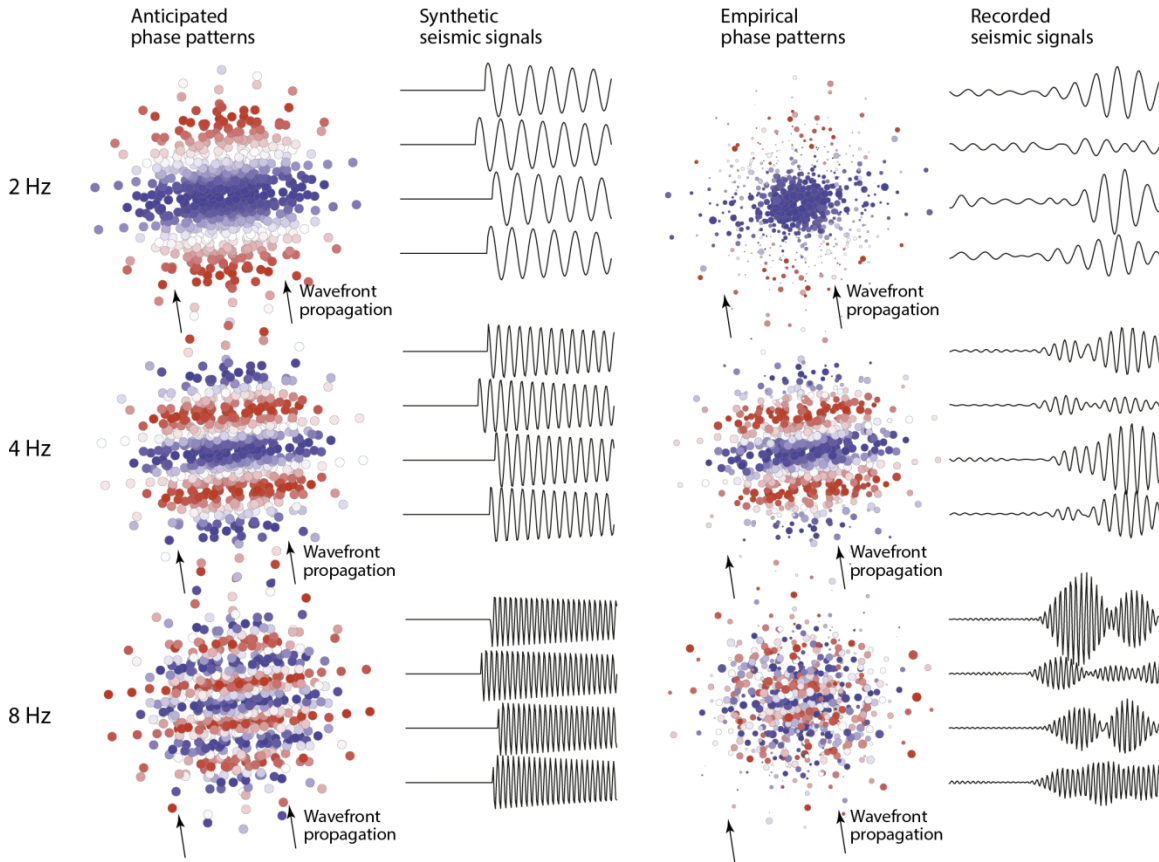


Figure 3: Phase differences between the ground motions at different sites of an array represented by colored symbols at the different sites of the co-array. To the left are phase shifts at three distinct frequencies corresponding to a theoretical plane-wave propagating across an array from a backazimuth of 175 degrees and with an apparent velocity of 7.5 km/s. To the right are the corresponding phase shifts measured from actual observations of a seismic wavefront which is predicted to arrive with the same parameters.

At 4 Hz, the empirical and theoretical covariance matrices look very similar, with a banded structure of phase shifts. However, we note that the symbols in the center of the co-array are larger than those at the periphery – meaning that the signals on closely-spaced sensors are more similar than those on more distant sensors. At 8 Hz, for which the seismic wavelengths are shorter and the wavefield more sensitive to small-scale heterogeneities in the real earth, the pattern becomes more fragmented and the signal coherence is diminished for all sensor separations. The difference between the theoretical and empirical covariance matrices is also large at 2 Hz. This is most likely the result of low signal-to-noise ratio (SNR). A signal at 2 Hz would likely be highly coherent over this array aperture, but at 2 Hz the waveforms are dominated by a diffuse noise field and hence the coherence is low.

Matched field processing involves comparing the phase, coherence and amplitude relations between sensors of an array or network with those anticipated from a given signal. The idea is displayed in Figure 4. In practice, a statistic is calculated from the covariance matrix of the incoming wavefront and a prescribed steering vector (this is a quadratic form and will be described later). A theoretical steering vector is constructed from a set of modeled time-delays and an empirical steering vector is the principal eigenvector of the covariance matrix evaluated for a previously observed signal. Matched field processing (MFP) using empirical steering vectors is referred to as empirical matched field processing (EMFP).

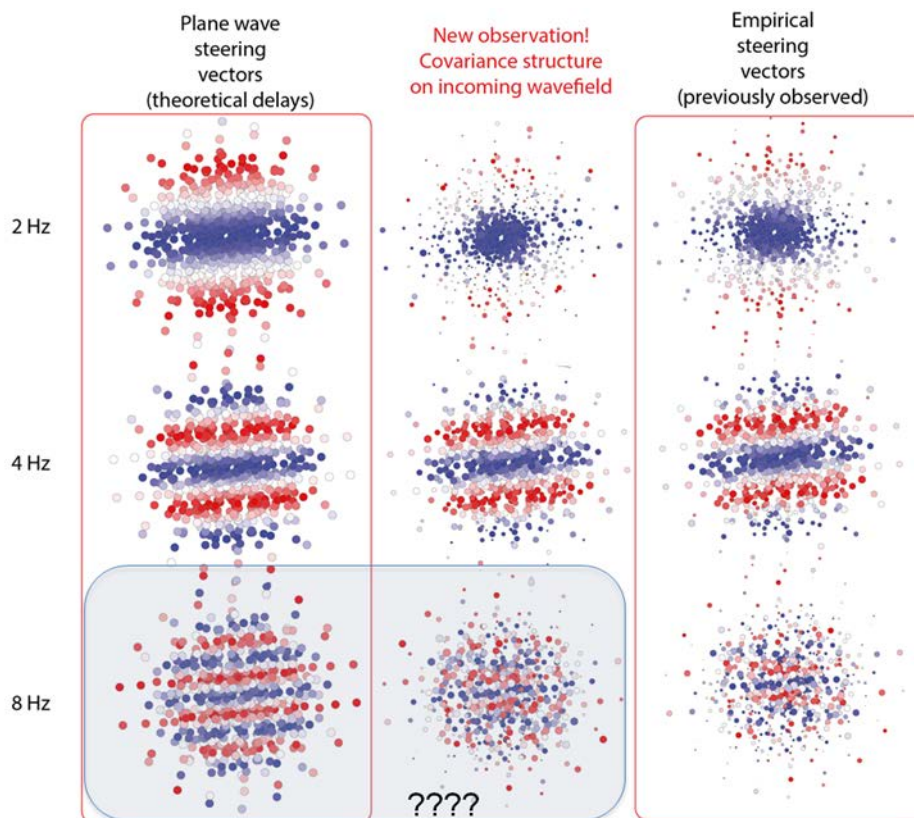


Figure 4: A graphical display of the comparison between the covariance matrix of an unknown wavefront with two possible phase and coherence patterns – one theoretical and one empirical.

EMFP was demonstrated by Harris and Kværna (2010) to provide an excellent classifier of seismic signals from very closely spaced active mines. The small differences in the paths taken from the different sites (on the Kola Peninsula in northwestern Russia) to the ARCES seismic array (at a distance of approximately 440 km) were sufficient to provide differences in the signal covariance matrices that were sufficiently characteristic that the small differences in the phase shifts observed could actually distinguish between blasts from the different mines. With only a few km between each mine, the sources are far closer to each other than a seismic array at this distance should be able to resolve using theoretical considerations. EMFP is actually exploiting

the diffraction of the seismic wavefront which degrades the performance of theoretical frequency wavenumber (or f-k) analysis.

The EMFP classification was also demonstrated to provide far superior identification of the sources than waveform correlation (Harris and Kværna, 2010). The reason for this is that the seismic signals from the mining explosions differ strongly temporally from one event to the next. The sources are ripple-fired shots with many temporally spaced charges detonating within fractions of a second of each other. This results in a very different sequence of ripples, depending upon the orientation and source time-function of each event. EMFP however breaks the signal down into narrow frequency bands such that temporal differences are mitigated; the phase shifts between each sensor of the array are still characteristic for the appropriate path.

EMFP has been shown to be an effective signal classifier, but how do we perform continuous detection? Figure 5 shows a sequence of signals from aftershocks of the 8 October 2005 Kashmir earthquake. For an effective characterization of the sequence we would like a detector which provided a trigger for every event coming from the aftershock region but not in the absence of such an event. Waveform correlation has been demonstrated by Slinkard et al. (2013) to be a poor classifier of such sequences with very few pairs of events with sufficiently similar signals. In particular, the difference in source dimensions between the main event (M 7.6) and the aftershocks (typically below magnitude 5.5) mean that the signal from the main shock is a poor representation of the Greens function for any given aftershock. However, for a seismic array at some distance from the source region, the phase shifts between sensors for incoming wavefronts from different events in the sequence are likely to show significant similarity.

Figure 5 demonstrates how EMFP for a single phase arrival can be used to characterize the aftershock sequence from a large earthquake observed at a single seismic array. A phase amplitude template characterized by the steering vectors, is obtained for the initial P-wave arrival from the main event crossing the array. A continuous comparison is made between these template phase shifts and the phase shifts inherent in the incoming wavefield measured over the array at any given time. For each time window, a statistic indicating the degree of likeness is obtained for each narrow frequency band resulting in a time-frequency display that is reminiscent of a spectrogram but where, instead of energy in the signal, we measure the goodness of fit between the instantaneous wavefield over the array and our template wavefront.

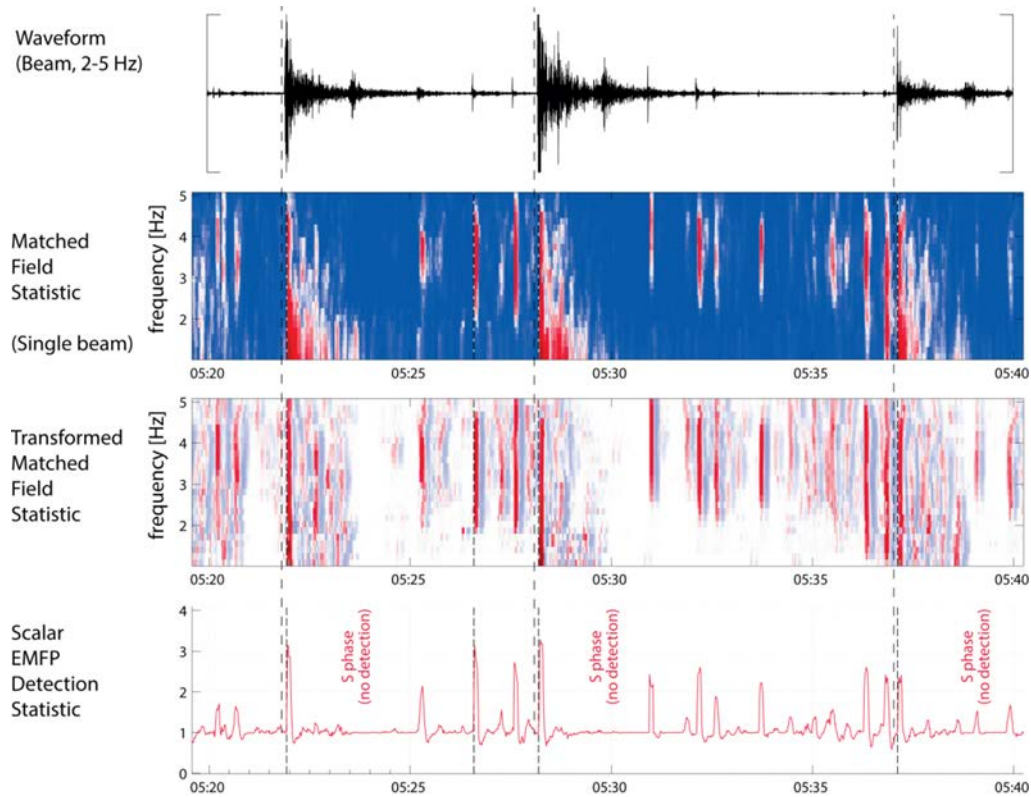


Figure 5: An illustration of the procedure to generate a detection statistic trace for a single matched field “beam” under an ongoing aftershock sequence, based on a single P-phase. The covariance matrix is calculated across all sensors of the array, in many narrow frequency bands, for the initial P-phase from a large earthquake, and empirical steering vectors are obtained as the principal eigenvectors of these matrices. A continuous matched field statistic trace is generated for each narrow frequency band using these steering vectors, and transformations of these single-frequency traces are converted to a scalar detection statistic trace by taking averages across selected frequency ranges. New occurrences of the template phase structure over the array appear as vertical bars in the “pseudo-spectrograms” and then as local peaks in the scalar trace. The waveforms shown are from the KKAR array in Kazakhstan on October 8, 2005. The main event is the magnitude 7.6 Kashmir earthquake at a distance of approximately 9 degrees from the station.

The properties of this “pseudospectrogram” vary from arrival to arrival. For the largest aftershocks in the sequence, the likeness is strong for all frequencies and the wavefield structure persists for a long duration into the coda. For smaller events, the likeness is only significant for higher frequencies for which the SNR is high. A differencing transformation (Gibbons et al., 2008) is applied for each frequency band that replaces the matched field statistic with a derivative function which is maximized at the phase arrival time, giving the impression of vertical bars in the time-frequency plot. A transformation across frequency bands (most simply a mean value) then generates a scalar trace on which significant local maxima occur at the times of P-wave arrivals from the different aftershocks. It is important to note that

unrelated energy arrivals (e.g. the S waves) do not result in detections, whereas all P-arrivals from a source region with an aperture of the order ~ 70 km do register a trigger. The procedure will be discussed in greater detail in Section 2.

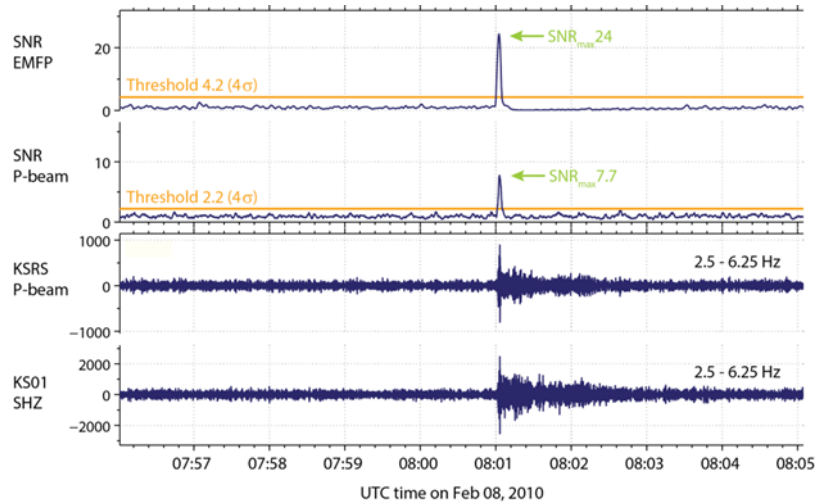


Figure 6: Matched field detection of the signal from a DPRK nuclear test (I). The signal from the February 12, 2013, DPRK nuclear test, recorded on the KSRS array, is submerged, scaled by a factor of 0.0125, into background noise on the same channels at the time shown on February 8, 2010. A beam is formed with apparent velocity 8.5 km/s and backazimuth of 15 degrees and an SNR (STA/LTA) trace is formed. At the top is the SNR trace of the empirical MFP scalar detection statistic, generated in the manner displayed in Figure 5, using the Pn phase of the signal from the May 25, 2009, DPRK nuclear test as the matched field template. Thresholds for the different SNR traces are determined from statistics of the (log) background variability of these traces over long time-windows.

To demonstrate that the MFP detector really does perform robustly below the detection threshold for classical (beam/STA:LTA) detectors, we submerge copies of the 2013 DPRK nuclear test signal, recorded on the KSRS array, with different scaling factors into background noise on the array and attempt to detect the initial P-phase using a matched field template obtained from the 2009 nuclear test signal. Figure 6 shows a case where the 2013 signal is submerged into background noise with a 1/80 scaling factor. We form a classical beam SNR and also an SNR trace on the output from the matched field detector as displayed in Figure 5. Thresholds for the two traces are obtained by examining the statistical variability of the traces over long time-durations. A clear detection is made for the signal on both SNR traces.

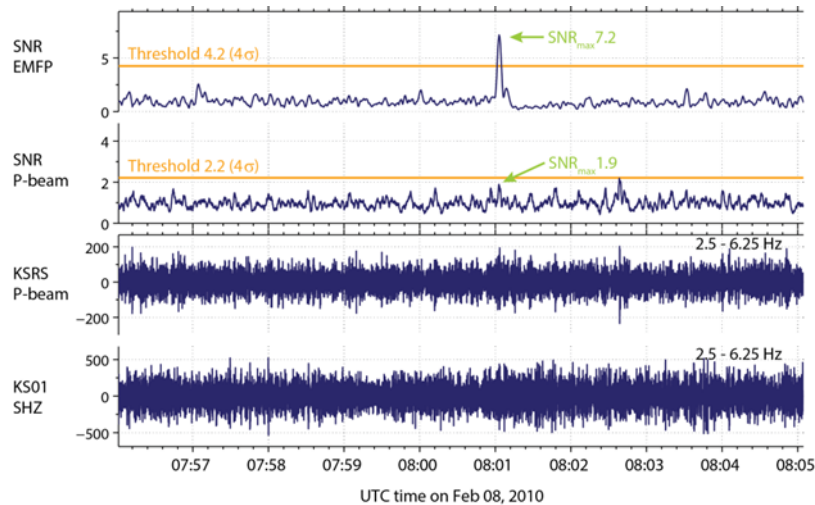


Figure 7: Matched field detection of the signal from a DPRK nuclear test (II). Exactly the same as for Figure 5 except that the scaling factor now applied to the signal from the 2013 nuclear test is now 0.0025 (i.e. a factor of 5 lower than that displayed in Figure 6). The submerged signal is no longer observed over the background noise, even on the beam, and the SNR calculated from the beam fails to exceed the empirically determined threshold. The SNR of the empirical MFP trace does exceed the statistically determined detection threshold indicating that, despite the fact that no significant increase in signal amplitudes is observed, the spatial structure of the wavefield immediately after the signal arrival fits the matched field template significantly better than the spatial structure of the wavefield immediately before the signal arrival.

The same is done in Figure 7, the only difference being that the scaling factor now applied is 1/400, taking the signal amplitude right down to the noise level. No detection is made on the classical beam SNR trace. The SNR trace derived from the MFP statistics still exceeds the threshold comfortably at the time of the signal onset. A detection statistic that measures the rapid increase in the similarity between the phase structure of the wavetrain and the target phase structure is significant at the time of the signal arrival, whereas a detection statistic that measures an increase in waveform amplitude is not significant at this time.

The system displayed in figures 6 and 7 has been optimized using an empirical template. In the absence of previous observations, a narrow band theoretical matched field template can be used that may still exceed the detection capabilities of beam/SNR-based algorithms if the increase in spatial coherence of the wavefield with the specified phase shifts is more significant than the increase in waveform amplitude. Such templates will usually consist of plane-wave phase shifts, although we may have additional models for wave propagation close to the array which may predict deviations from the simplest plane-wave models. Figure 8 illustrates the detection of the 2013 DPRK nuclear test signal at KSRS using plane-wave matched field detectors and empirical matched field detectors with the templates taken from the events in

2006 (distance ~2km) and 2009 (distance < 500 m). In the lower frequency band, there is a little difference in the performance between the 3 detectors. At higher frequencies, both empirical templates outperform the plane-wave MFP detector with the template from the closer 2009 event outperforming the template from the 2006 event.

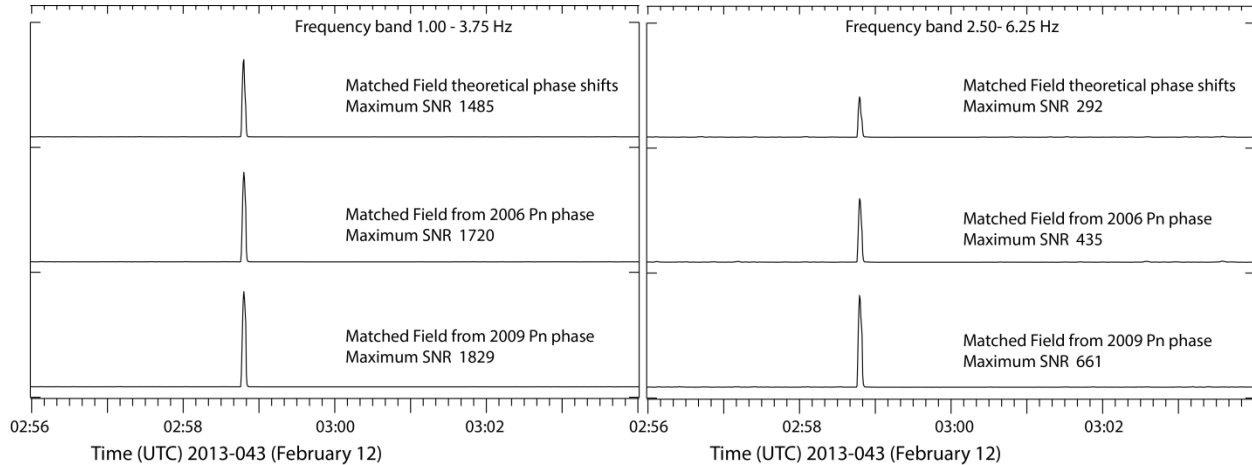


Figure 8: Detection of the signal from the 2013 DPRK nuclear test on the KSRS array using matched field templates as indicated. For both frequency bands, the empirical matched field templates perform better than the plane-wave template (top) with the template from the closer 2009 event signal performing better than the template from the 2006 event signal. However, the difference is more pronounced at the higher frequencies indicating that the empirical MFP detector exploits significantly more of the critical high frequency energy than the plane wave matched field detector.

We have demonstrated that a specially targeted matched field beam will detect events in that given direction effectively. However, the purpose of the current contract is to explore the extent to which EMFP can be extended to cover the spectrum of detectors from the highly source-specific spot detectors to the general detectors for unknown signals. Which considerations are most important when designing generalized detectors using the principles of MFP? Figure 9 depicts deployments of matched field “beams” for seismic arrays covering the theoretical slowness space and augmented by source-specific, empirical, matched field detectors. A general beam deployment has been augmented by a few specially chosen site-specific matched field templates, looking for e.g. Pn, Pg, Sn, and Lg phases from regional events at a particular site. In addition, a few other templates aimed at targets at teleseismic distances are deployed. Other templates may be added for spotlight detectors on target sites, or even for repeating sites that generate signals that require safe classification for screening.

In principle, every well-classified detected phase can result in a new matched field detector. In practice, this could escalate rapidly to a deployment containing an intractable number of empirical MFP detectors. Repeating sources such as mines may be adequately represented with one or a few matched field templates; other very limited regions of slowness space may correspond to quite broad geographical regions and may need very dense coverage of matched field detectors to provide optimal detection and parameter estimation.

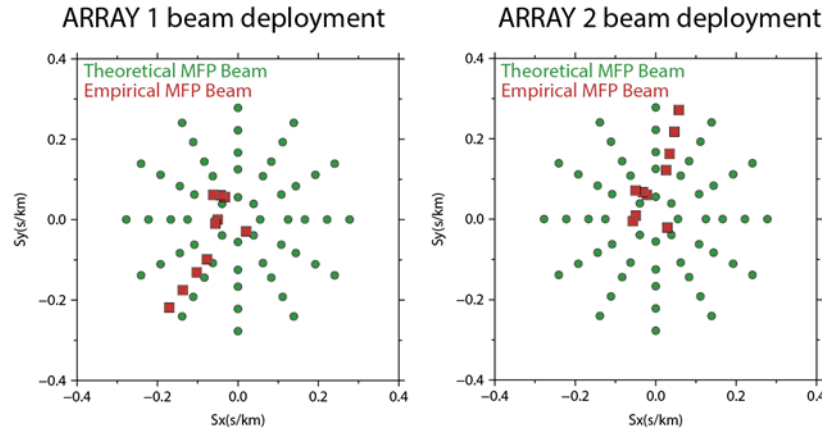


Figure 9: Schematic representation of a matched field beam deployment on two arrays both with general beams and empirical beams aimed at specific wavefronts. After the deployment of a new seismic array, an initial beam deployment would consist only of the (green) matched field “beams” with the theoretical phase shifts. As events are recorded and located, the (red) empirical matched field beams can be deployed in addition to detect new occurrences of previously observed phases.

Following signal detection, an estimate must be made of parameters which describe the propagation of a wavefront, in order that the signal’s origin can best be estimated.

Under existing procedures, slowness space is scanned using plane-wave phase shifts which frequently results in bias and waveform misalignment (hence the need for SASCs – Slowness and Azimuth Station Corrections). For example, the top panels of Figure 10 show f-k analysis for the P-phase of an aftershock from the October 8, 2005, Kashmir earthquake recorded at KKAR, performed at 2 Hz and 4 Hz, together with the theoretical slowness vector indicated by a red arrow. The estimate at 2 Hz is close to the expectation whereas the estimate at 4 Hz gives a qualitatively different estimate (a teleseismic apparent velocity). The frequency dependency of plane-wave slowness estimates is one reason for the variability of broadband f-k estimates, which in turn requires large tolerance bounds in phase association algorithms.

A significant part of the project investigates how EMFP can result in calibrated parameter estimates for placing the direction of signal in the classical slowness space. However, we also investigate the potential for mapping signals directly into the geographical space of the source regions: a procedure which may change significantly the nature of standard seismic processing pipelines.

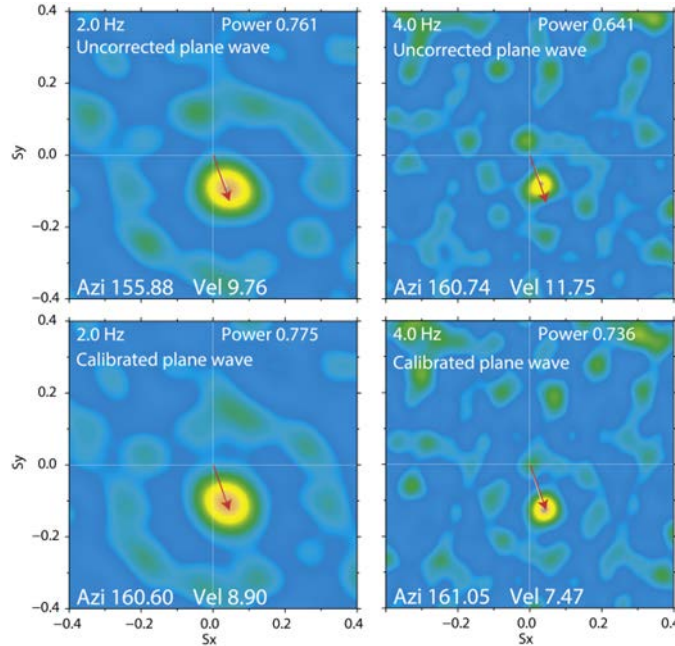


Figure 10: Demonstration of how empirically calibrated f - k analysis results in an increased array gain (in particular at higher frequencies) and parameter estimates that are more consistent between the different frequency bands. All of the above slowness grids are evaluated at the time of a P-arrival at the KKAR array from an aftershock of the 8 October 2005 Kashmir earthquake. The “calibrated” slowness grids are evaluated using imposed phase shifts calculated from the P-arrival from the main shock which is attributed an apparent velocity of 8 km/s and a backazimuth of 160 degrees. The aftershock is approximately 50 km to the North West of the main shock from which the template is calculated. The array is at a distance of approximately 9 degrees from the source region.

Most of the calculations performed in this project are at frequencies above 1 Hz since this is the band in which energy is primarily detected for body waves from low magnitude seismic events of monitoring interest. However, we wish also to examine the physical basis for the observed variability in matched field templates. Which factors of the propagation medium are most significant in determining the form of the received signal? We therefore wish to study variability of matched field templates in a frequency band that allows comparison with results obtained through model calculations. With anything but the largest supercomputers, the calculation of synthetics requires that the comparison be carried out at long periods. Consequently, we intend to investigate the variation of matched field templates for long-period surface waves.

2 Methods, Assumptions, and Procedures

Each of the following subsections describes a study of a different aspect of the implementation and understanding of empirical matched field processing in online processing pipelines.

Section 2.1 describes what we call the “anatomy of a matched field signal detector” with special reference to earthquake aftershock sequences. This section lays down the foundations of how general online systems are envisaged.

Section 2.2 examines the performance of empirical matched field processing applied to estimating the direction of arrival of a signal on a seismic array, guided by previous observations from well constrained events. We examine in particular how the slowness resolution is related to the array geometry and the near-array geologic conditions, and discuss the implications for deployments of matched field beams in a processing pipeline.

Section 2.3 describes how we can mitigate the problems in estimating the direction of arrival of signals on an array using so-called Limited Sensor Pair Correlation. While the work in this section is classical array processing (i.e. not empirical matched field processing) we describe why this has consequences for how matched field processing can be applied in general pipelines.

Section 2.4 presents a case study of EMFP detection: small events at, or close to, the North Korea nuclear test site.

Section 2.5 examines the physical basis of empirical matched field processing: what we can learn from comparing synthetics and observations for long period surface waves recorded on a network from sources in different tectonic settings.

2.1 An Anatomy of a Matched Field Detector

How do we exploit the covariance structure of a wavefield over an array of sensors to obtain a detection statistic on which triggers can be declared? We have alluded to a possible procedure in the introduction and will describe the procedure in greater detail here.

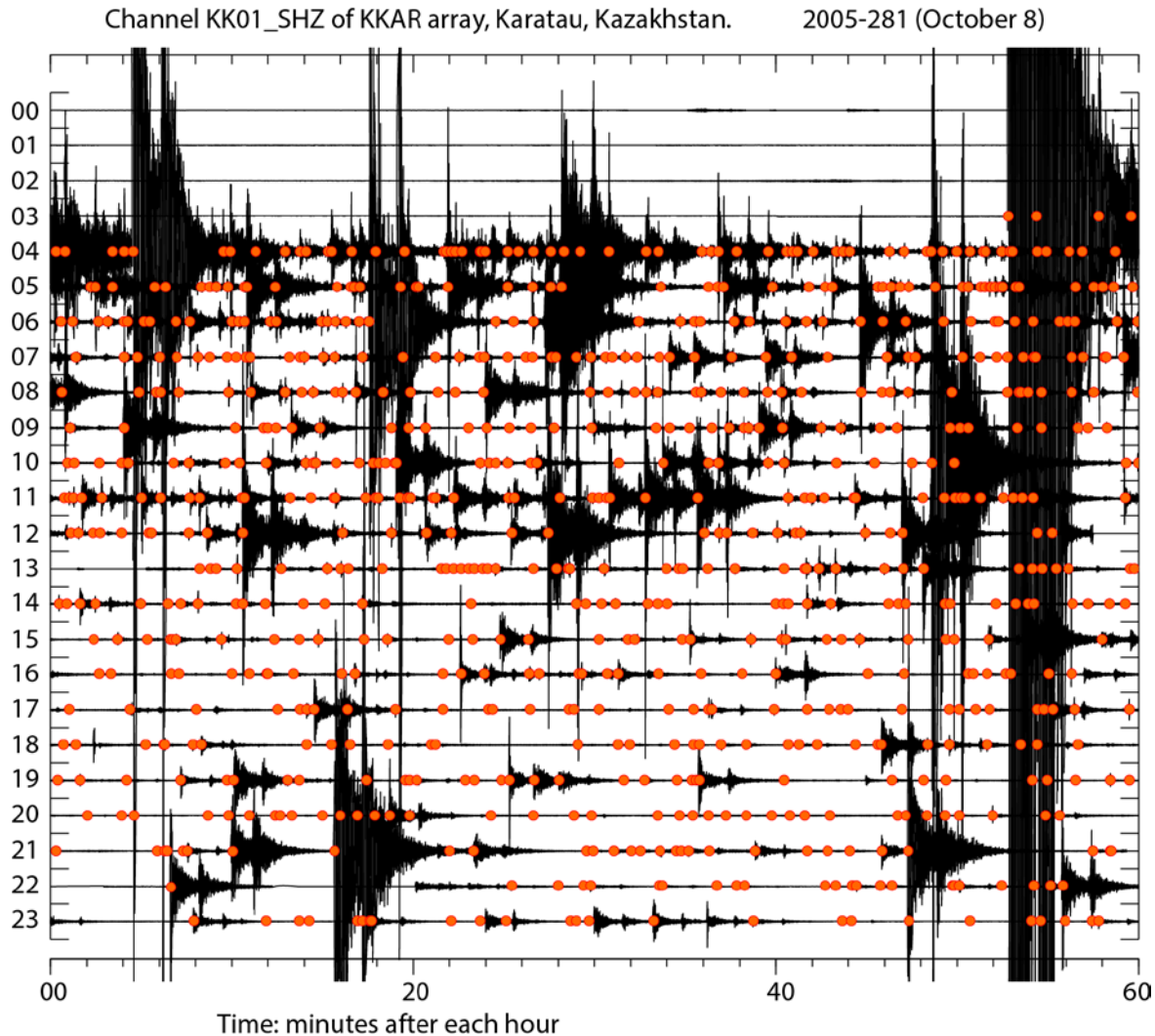


Figure 11. “Mock helicorder plot” of a single channel short period vertical seismometer on the KKAR array.

We illustrate the procedure in the context of an extensive aftershock sequence, for which we take the 8 October 2005 Kashmir earthquake as an example (Figure 11). An aftershock sequence is a good case study for illustrating the advantages of matched field processing. Firstly there is the need to detect a vast number of signals and classify them with a high degree of confidence. Secondly, given that seismic arrays form the backbone of the global seismic monitoring

infrastructure (Figure 12), there are properties of the signals from aftershocks which lend themselves to empirical matched field processing.

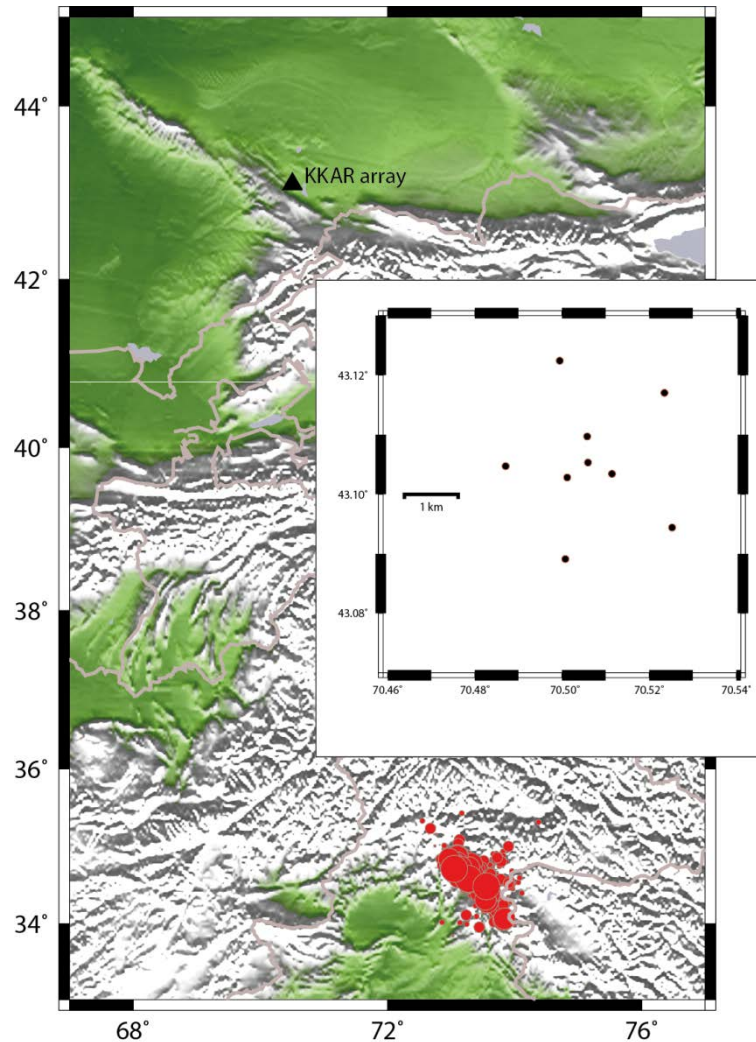


Figure 12: Events in the 8 October 2005 Kashmir earthquake aftershock sequence (red symbols) in relation to the KKAR (Karatau) array in Kazakhstan at a distance of approximately 9 degrees. The geometry of the 9 element array is displayed in the inset diagram.

Aftershock sequences from large earthquakes can create huge problems for data centers attempting to create comprehensive seismic bulletins. A vast increase in the number of events which need processing is compounded by a significant deterioration in the quality of automatic event bulletins which form the basis for the analyst-reviewed bulletins. This deterioration results from the challenges faced by phase association algorithms; arrivals from key stations are frequently missing as they are drowned out by the coda from earlier events in the sequence, and spurious associations are likely within the vast number of detections on each station.

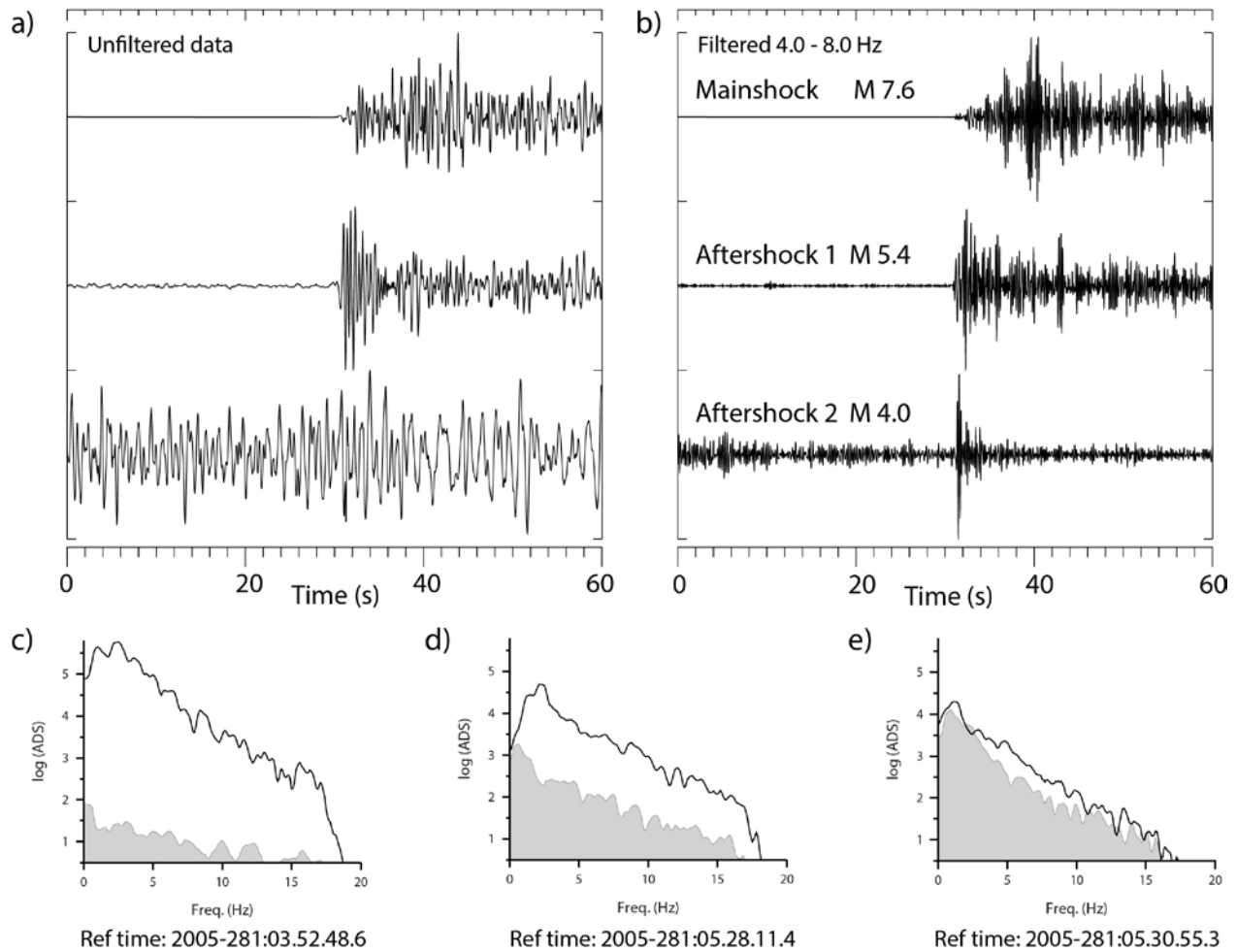


Figure 13: Three events in the 8 October 2005 Kashmir earthquake sequence which demonstrate why waveform correlation may be a poor signal classifier for such events. Panels (a) and (b) display the aligned waveforms (raw and filtered as indicated) and (c), (d), and (e) display the spectra of the signal (black line) and preceding noise (solid shape) for the events with arrival times as specified. Note that the preceding noise is likely to include a significant contribution from the coda wave energy for the previous event.

Figure 13 displays the waveforms on the KKAR array in Kazakhstan for three events in the Kashmir sequence. One event (Figure 13 c) is the main shock, one event (Figure 13 d) is an aftershock of magnitude 5.3, and one event (Figure 13 e) is an aftershock with magnitude below 4.0. Visual inspection suggests that a waveform correlation procedure may be unsuccessful in characterizing the sequence, as was confirmed in an extensive study by Slinkard et al. (2013). Firstly is the difference in the signal spectra, with vast amounts of low frequency energy dominating the signal from the mainshock. Secondly is the difference in the time-duration, with the signal from the smallest of the events barely extending past the initial Pn arrival.

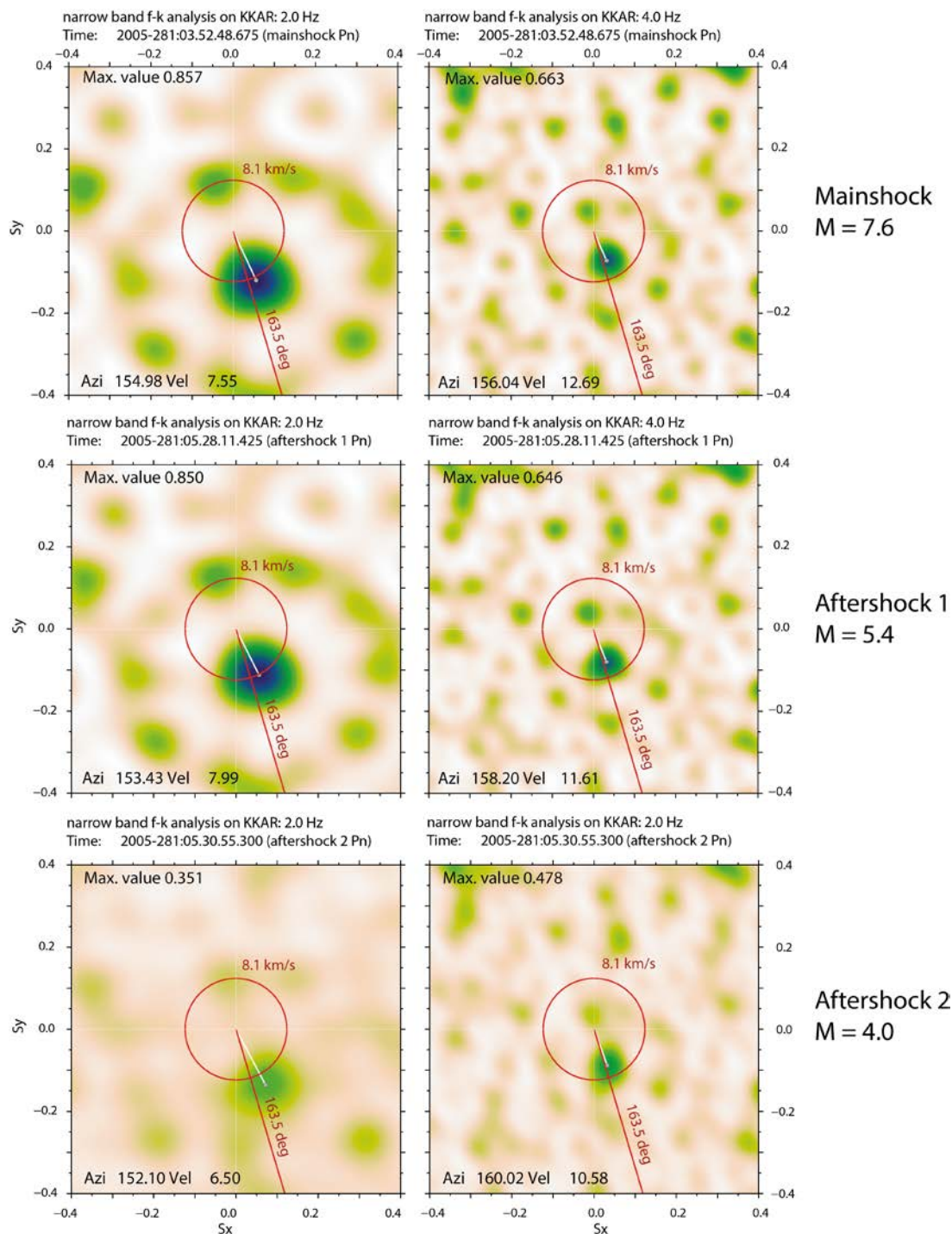


Figure 14: Narrowband f-k spectra (2 Hz and 4 Hz) for the Pn arrival at KKAR for the three events displayed in Figure 13. The FORTRAN 90 multitaper routines of Prieto et al. (2009) were used to calculate the narrow frequency band phase shifts.

Figure 14 displays the narrowband classical f-k spectra at 2 Hz and at 4 Hz for the first arrivals at KKAR from the events displayed in Figure 13. In spite of the large differences in the spectral content between the events, the projection of energy over slowness space is remarkably similar for each arrival. As frequently (if not usually) observed for arriving wavefronts at seismic arrays, there is a deviation or bias both in the estimated backazimuth and in the apparent velocity. This bias is different for the different frequencies examined and this is the reason for the frequency-dependent bias and variability of broadband f-k estimates (e.g. Gibbons et al., 2010, and references therein). If the wavefront, when filtered in different narrow frequency bands, appears to come from different directions then the broadband slowness vector estimate will be a function of how the different frequencies contribute to the total broadband estimate. This will depend, for example, on the signal-to-noise ratio (SNR) at the different frequencies.

At 2 Hz, there is an approximate bias of 10 degrees (with backazimuth estimates around 153 degrees, when the great circle backazimuth is approximately 163 degrees). The apparent velocity is slightly slower than that predicted by the 1-dimensional ak135 velocity model. The relative power for the 2 Hz estimate is low for the smallest of the three events, entirely consistently with the low SNR at this frequency (c.f. Figure 13 c).

At 4 Hz, the azimuthal bias is smaller, but the apparent velocity is significantly higher: more consistent with an arrival from near-teleseismic distances rather than regional. The consistency of the narrow band f-k spectra between the three events, despite the different large spectral differences between the three signals, strongly suggests that a classification of the signals based upon characteristic phase shifts is possible.

Figure 15 displays the covariance matrices at different frequency bands together with the so-called steering vectors which define the phase shifts for a given slowness vector. In classical f-k analysis, we are simply finding which theoretical steering vector out of many best fits the observations. In Figure 15 a), the covariance matrices and steering vectors are specified directly from the theory – given an array-geometry and a speed and direction of a plane wavefront. The size and colors of the symbols in the covariance matrices, $\mathbf{R}$, indicate the coherence and phase (c.f. Prieto et al., 2009) between the waveforms on the different sensors for a given frequency. That all the symbols are full-sized indicates that the signals are perfectly coherent as is the case with the plane-wavefront model. The empirical steering vectors, $\mathbf{\Xi}$, are the principal eigenvectors of the covariance matrices.

In Figure 15 b) we see the actual covariance matrices calculated from the waveform data at the time of the first arrival from the Kashmir main shock. The empirical covariance matrix looks almost identical to the theoretical covariance matrix for the lowest frequency (1 Hz) but

significantly different at the highest frequency (5 Hz). In particular, the diminished coherence – indicated by the smaller symbols, is a result of the scattering and refraction of the higher frequency and smaller length-scale wavefield. It is the understanding that EMFP captures the characteristic deformation of the wavefront and exploits this for signal identification (c.f. Harris and Kværna, 2010).

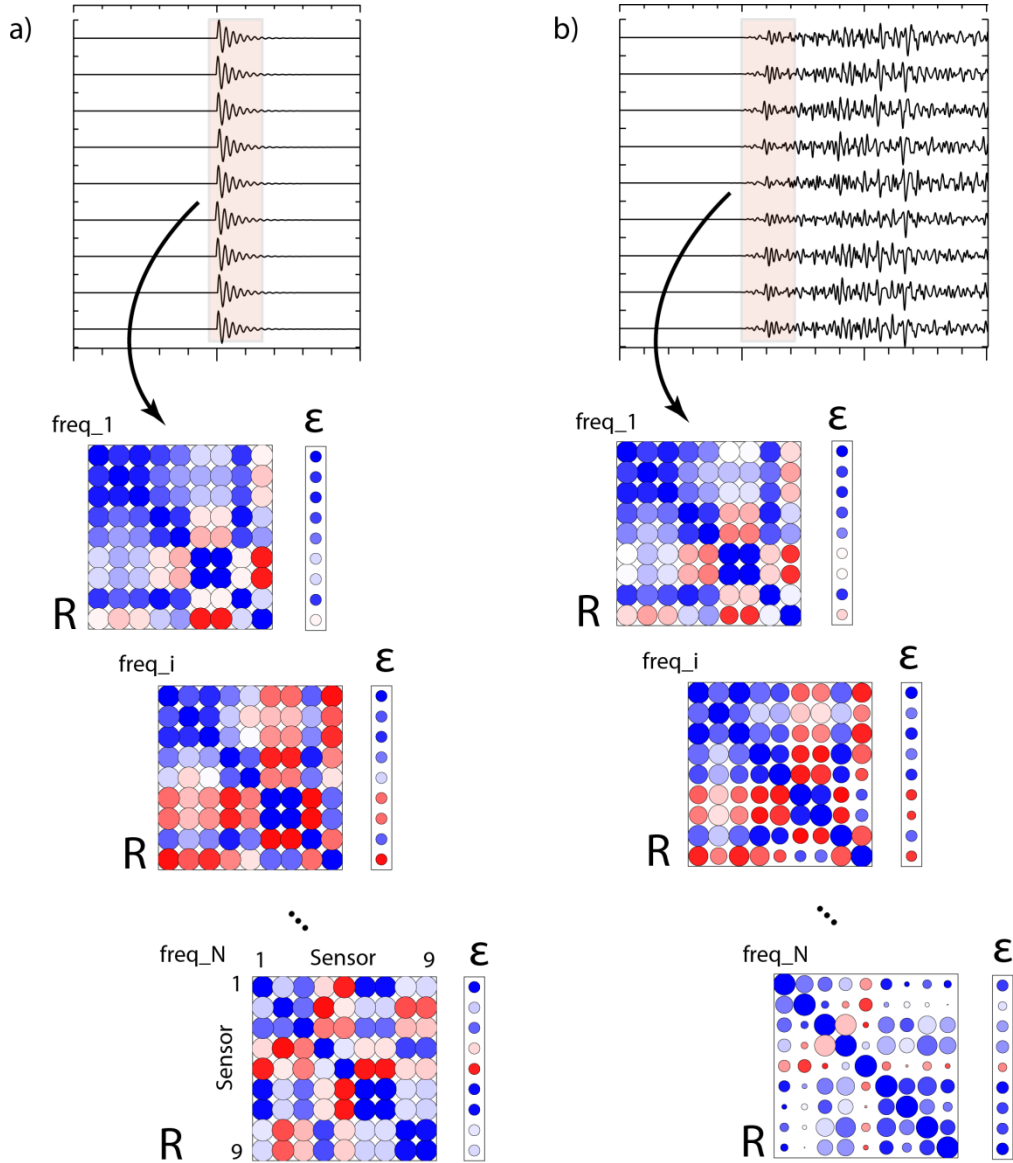


Figure 15: Calculation of narrowband steering vectors using theoretical (plane-wave) covariance matrices (a) and empirical covariance matrices (b). The steering vector for each band is the principal eigenvector for the corresponding covariance matrix.

In Figure 16 we display the covariance matrices at 2 and 4 Hz for four different wavefronts but with the locations of the elements being defined by the position in the co-array rather than in

the block-matrix form displayed in Figure 15. At left are the phase shifts at the two frequencies predicted by the plane-wavefront model. The phase shifts at 4 Hz have, intuitively, more “bands” across the array aperture than the phase shifts at 2 Hz. To the right, we see the elements of the empirical covariance matrix calculated from the Pn arrival from the mainshock. Surprisingly, the phase shifts calculated at 4 Hz indicate fewer bands than anticipated from the theoretical time-delays. The pattern displayed however is entirely consistent with the f-k spectra displayed in Figure 14; increasing the apparent velocity will make the colored bands appear broader.

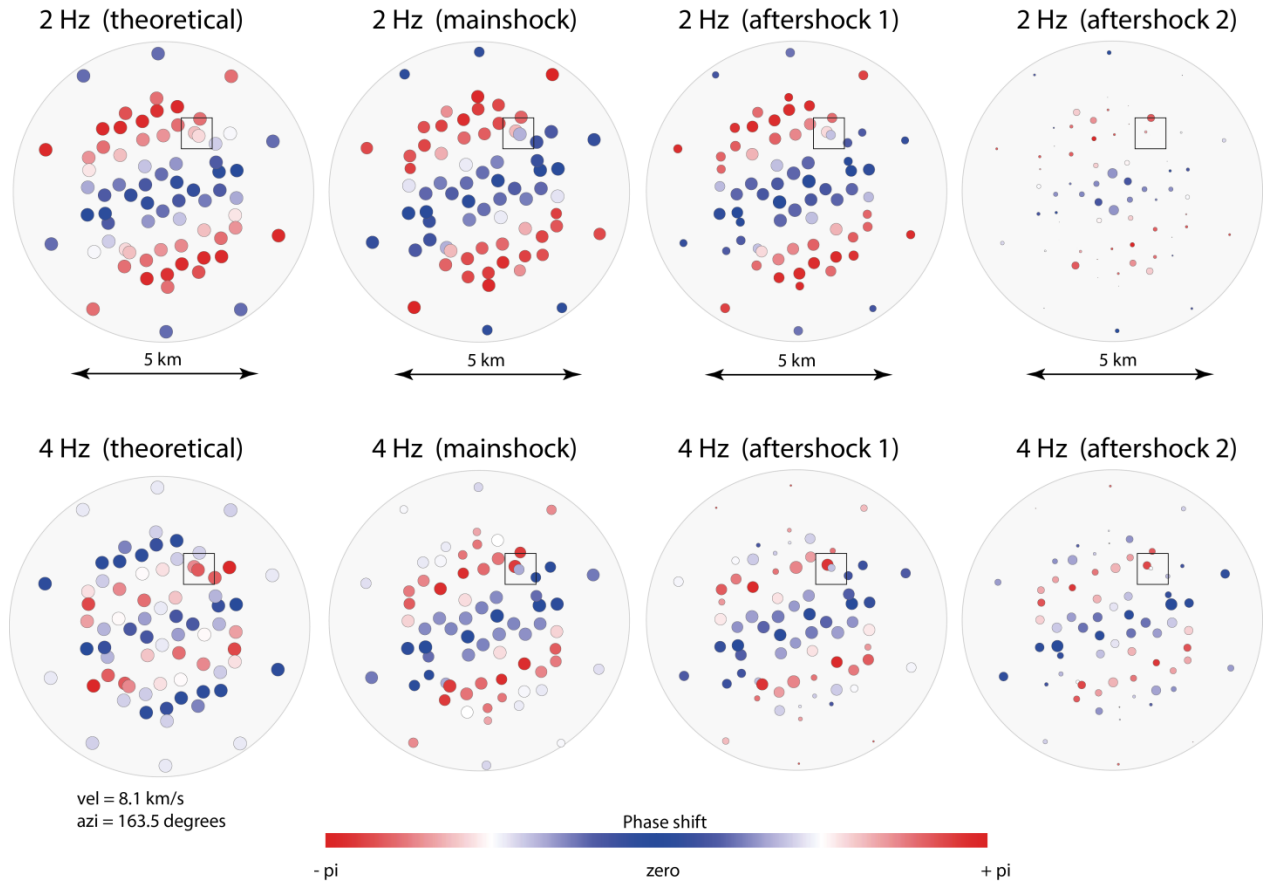


Figure 16: Narrowband covariance matrices for the frequencies as shown displayed with the phase (color) and coherence (symbol size) plotted as a function of the vector displacement between the two sensors (i.e. at the coordinates of the co-array). The square displayed is simply to highlight two points in the co-array with almost identical locations (i.e. vector displacements in the original array).

The small square displayed in each of the so-called bubble plots highlights two elements of the covariance matrix which have almost identical coordinates in the co-array. This is to say that,

given the identical vector displacements between the sensors, these points should show almost identical colors. This is the case for the theoretical phase shifts but not for the empirical phase shifts; this is a clear indication that the wavefield is somehow being deformed as it crosses the array aperture, and we seek an empirical correction to this. The elements of the covariance matrices for “aftershock 1” and “aftershock 2” are very similar to those for the mainshock, except for the diminished coherence at all frequencies for the low-SNR aftershock 2.

What we conclude from Figure 16 is that the sets of phase shifts appear to form a robust characteristic of the incoming wavefronts which may allow a confident identification of such signals in a way that the correlation detectors cannot provide.

Figure 17 displays the first step of how we might exploit this characteristic to form a continuous detection statistic on which triggers may be declared. We step along the multi-channel datastreams calculating covariance matrices on each of many short and overlapping time-windows. We seek to explore how general we can make matched field detectors and, given the typical transience of a given phase pattern of a seismic arrival, we select relatively short time-windows with length of the order 3 to 5 seconds. The optimal length of the window length selected will depend upon the aperture of the array and the frequency-content of the signals of interest, and the length selected will have consequences for the time-resolution of a resulting scalar detection statistic. Windows of length 3-5 seconds are typical of the length used to perform f-k calculations and are also typical of the duration for which an arriving wavefront retains coherence before the less coherent coda-waves often dominate.

For the time-window highlighted, a total of N covariance matrices are formed: one for each of the available frequencies selected. For a given set of steering vectors – either derived from theory (Figure 15 a) or empirically (Figure 15 b) – we can calculate a scalar of the quadratic form $\mathbf{e}^H \mathbf{R} \mathbf{e}$ and set these values in a two-dimensional array with time along one axis and frequency along the other. Due to the similarity to the spectrogram, in which we record energy as a function of both time and frequency, we refer to this two-dimensional function as a matched field *pseudo-spectrogram*. This provides, as a function of time and frequency, how well the wavefront crossing the sensors in an array fits with the set of steering vectors applied. In all the examples provided in this report, the covariance matrices are fully normalized such that the value of the pseudospectrogram for any one time and frequency is limited between zero and one. This is to say that our detector is to respond to recognize patterns of coherent energy only, and not the signal amplitude. The original waveform data exhibits a phenomenal dynamic range (c.f. Figure 11) and there will be both advantages and challenges resulting from the far lower dynamic range from which our detection statistic can be constructed.

There is a deep blue “background level” in the pseudospectrogram interspersed with red pixels corresponding to Pn arrivals from the approximate direction of the main shock Pn signal. For the three largest events, the red pixels cover the full range of frequencies considered (1-5 Hz) while there are many other “vertical red bars” which are only visible at higher frequencies (e.g. above 2.5 Hz or 3 Hz). It is noted that the red pixels also begin to disappear towards the highest frequency considered (5 Hz). This is to do with loss of coherency between the sensors of the array, rather than a falloff in the seismic energy. Harris and Kväerna (2010) demonstrate how empirical matched field processing extends coherent array processing to wider apertures for the purpose of recognizing signals generated by seismic events at a specific source location but, for aftershocks covering a wide region (> 50 km), differences in the propagation path are likely to alter the phase shifts applying to the higher frequency parts of the wavefield. This means that for small events, we may have a relatively limited frequency band to work with between the lower limit, at which the SNR becomes problematic, and the higher limit, at which the characteristic phase-shift assumption fails.

For the larger events, coherent energy at low frequencies with phase shifts characteristic of the Pn arrival persist well into the P-coda and, in some cases, right up until the Sn-arrival. The characteristic phase shifts at higher frequencies are very transient and the high frequency part of the coda wave energy is highly incoherent. The smaller events, generally only visible at the higher frequencies, also have highly transient matches with the template steering vectors.

The black line in the lowermost panel of Figure 18 shows a simple arithmetic mean of the values in the pseudospectrogram and makes it clear that while the colored bars in the time-frequency plot are readily identified features, more advanced signal processing is needed to extract these features from this image. The declaration of triggers from time-frequency plots has led previous studies (e.g. Taylor et al., 2010) to apply techniques from image processing to transform relevant features into robust detection statistics. Whereas Taylor et al. (2010) were trying to identify transients visible over multiple frequencies, Gibbons et al. (2008) wanted to identify the starting points of signals on spectrograms in which the energy was likely to be sustained at a high level over a long duration. It was necessary to generate a transformation which emphasized times at which a significant increase in energy was observed - whether or not this increased energy persisted.

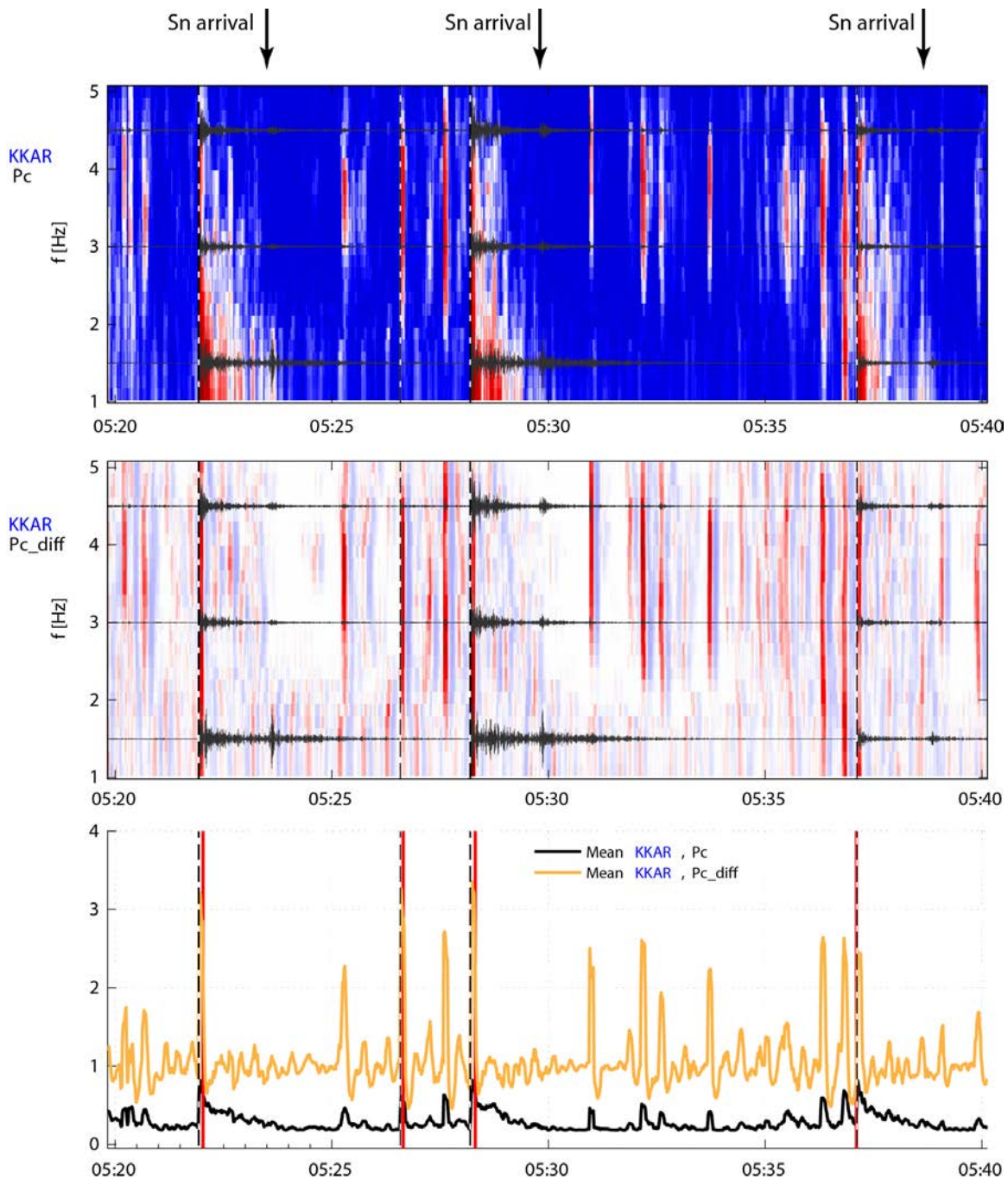


Figure 18: The full matched field pseudospectrogram for the empirical steering vector calculated for the Kashmir mainshock Pn arrival for a 20 minute long time interval on October 8, 2005 (top panel). The middle panel shows a non-linear transformation of the pseudospectrogram designed to emphasize signal arrivals (see text). The lowermost panel shows scalar functions that, for a given time-sample, are simple means of the values in the original and transformed pseudospectrograms. All pseudospectrograms calculated for the 1.09-5.00 Hz band, 256 samples at 40 Hz sampling rate and 25 narrow bands.

Gibbons et al. (2008) use the notation $A(f)_{t+}$ to denote the value in a spectrogram (as a function of frequency, f) evaluated in a time-window immediately following a time t , and the notation $A(f)_{t-}$ to denote the corresponding function evaluated in a time-window immediately prior to time t . A non-linear transformation

$$S(f, t) = [A(f)_{t+} - A(f)_{t-}]A(f)_{t+}$$

was identified as a means of obtaining a peak that was proportional both to the value of the spectrogram and the increment. (Note that the logarithm that needed to be applied to the amplitude density spectra calculated by Gibbons et al., 2008, is not necessary here given the limited dynamic range of the matched field pseudospectrogram.) The middle panel of Figure 18 displays the $S(f, t)$ transform of the $A(f)$ pixels displayed in the top panel of Figure 18 and we see that the arrivals of energy consistent with our imposed steering vector are now more visible as vertical bars. The mean over frequencies (the orange line in the lowermost panel of Figure 18) shows a local maximum for each of these vertical bars, although the time-series is noisy and probably not suitable for a detector; further manipulation will be necessary.

In Figure 19 we display the matched field pseudospectrogram and its transform for four different arrays: KKAR, BVAR, ILAR, and WRA. It is clear that KKAR is the most sensitive station with detections of many arrivals in the time-duration displayed. BVAR in northern Kazakhstan is further away and records fewer of the smallest events, although is still one of the most sensitive IMS stations to events in this sequence. At greater distances, WRA (Australia) and ILAR (Alaska) detect significantly fewer aftershocks although the performance of both arrays is excellent given both low background noise in the frequency band of interest and high coherence of the waveforms between sensors. At WRA and ILAR the seismic energy is best observed at lower frequencies than for the closer arrays and the 1-5 Hz frequency band captures most of the energy in the signal – not just most of the coherent energy.

The approximate spectrogram alignment resulting from time-delaying the images points to the possibility of “geographical beamforming” of the matched field statistic traces. This is to say that one would exploit the coherence between sensors locally (over the aperture of an array) and then search the source region using incoherent stacking of the single-array outputs. The emphasis here though will be to support existing phase reporting and association algorithms in the best possible way.

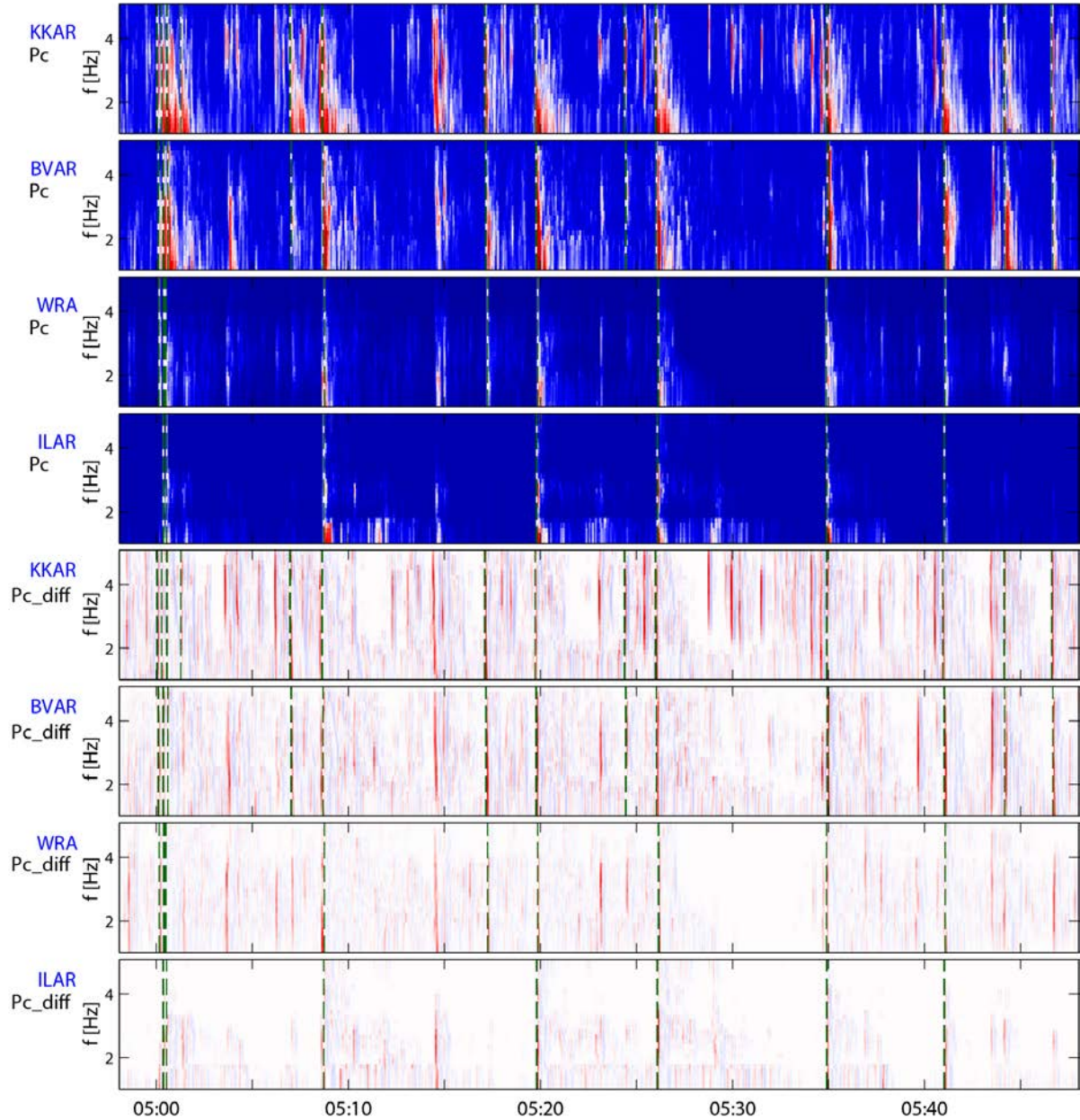


Figure 19: Matched field pseudospectrograms (both direct and transformed) for the four array stations KKAR, BVAR, WRA, and ILAR. The more distant arrays (WRA and ILAR) detect far fewer events than the closer arrays (KKAR and BVAR) and also have energy concentrated at lower frequencies. The images have been projected back in time with the traveltimes from the main-shock in order to align features approximately. (Some misalignment will occur due to the differences in traveltimes to different locations in the aftershock region.)

Figure 20 a) displays for three frequencies the numerical values of the matched field pseudospectrogram displayed in the uppermost panel of Figure 18. With all frequencies displayed, and with a suitable color scale, arrivals are visible in the two dimensional image. Considerable effort is still required to form a sufficiently robust scalar detection statistic. In Figure 20 b) we see the $S(f, t) = [A(f)_{t+} - A(f)_{t-}]A(f)_{t+}$ transformation displayed for the same three frequencies. In order to reduce the influence of fluctuations which occur over very narrow frequency bands we wish to reduce the dynamic range even further prior to a zero offset stack over the frequency bands. Inspired by the N-th root methods applied to seismic array traces (Muirhead and Datt, 1976) we note that the function $S(f, t)$ consists of small values either slightly above or slightly below zero. Performing the $\exp_{10}$ operation on these traces makes all values positive, either above or slightly less than unity, such that a subsequent subtraction of unity (Figure 20 c) results in traces covering a similar range to those displayed in the panel above but with reduced significance for those fluctuations close to zero.

An empirical comparison between the stack displayed in Figure 20 d) and optimized array beams appears to indicate a robust detection trace with few false-alarms (i.e. triggers at times at which there is no evidence of a wavefront arriving at the array from a direction consistent with the applied steering vector). The procedure described in Figures 17 and 20 is necessary for every set of steering vectors required in a beam deployment with a computational cost proportional to the number of such “beams” necessary to cover the parameter space adequately. For the purpose of monitoring an aftershock sequence (as displayed here) a single matched field beam may suffice, provided that we are able to scan the parameter space for each detection to confirm that the signal has indeed come from the assumed direction. The classical slowness scan (e.g. Figure 14) is likely to provide us with a good estimate of the number of beams necessary to monitor the full parameter space.

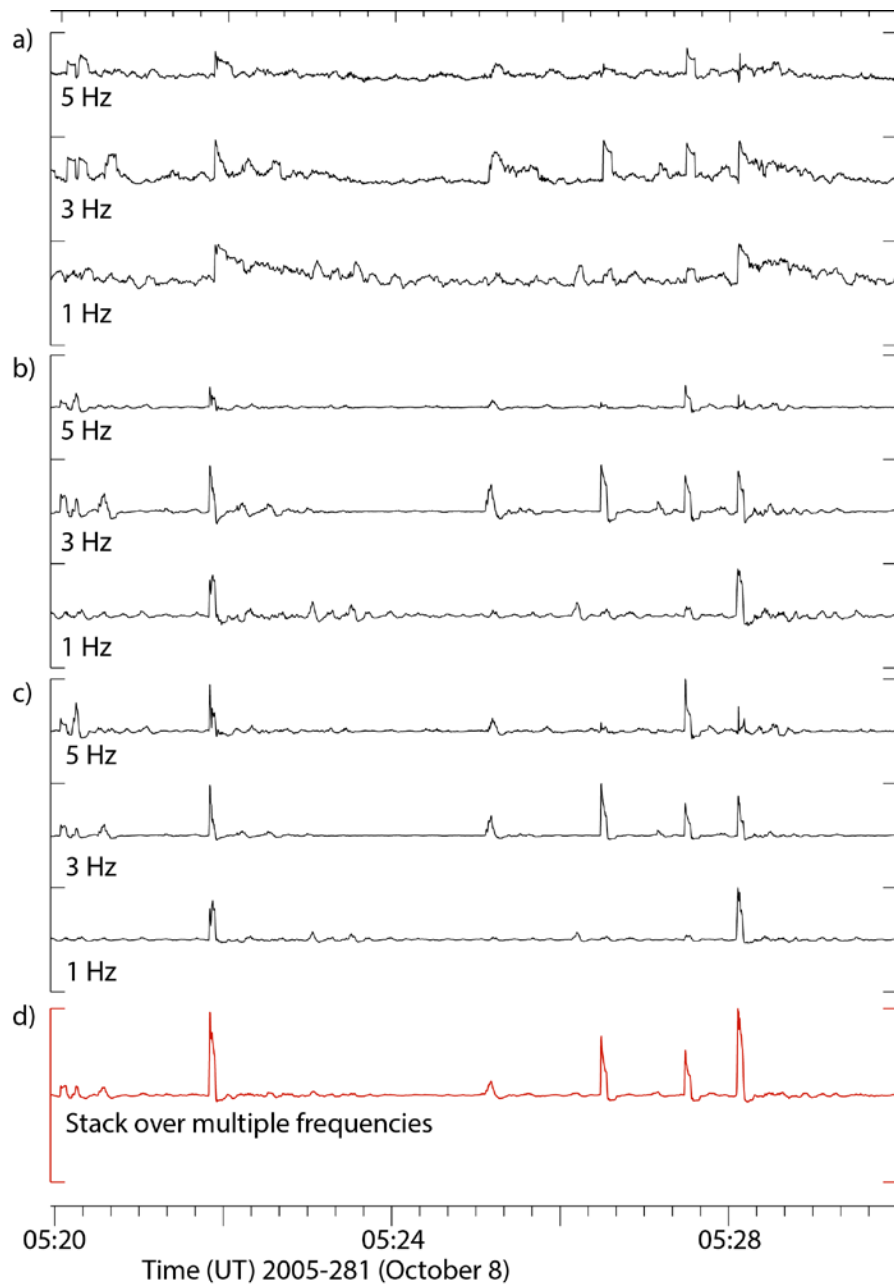


Figure 20: Steps to transform the matched field pseudospectrogram to a scalar detection trace. Panel (a) shows the numerical values of the pseudospectrogram displayed at the top of Figure 18 for 3 distinct narrow frequency bands. Panel (b) shows these scalar time-series transformed as displayed in the middle panel of Figure 18. Panel (c) shows $\exp_{10}(\cdot) - 1$ of each of the traces in panel (b), and panel (d) shows the stack over all frequencies (25 narrow bands) of the functions displayed in panel (c).

2.2 Parameter Estimation Using Empirical Matched Field Processing

We seek to understand how combinations of theoretical and empirical templates will cover the parameter space that describes the set of incoming wavefronts, anticipated at a given station, that are of interest for monitoring purposes. This is to say that we want to be able to detect all signals of interest and classify them correctly, using a Matched Field method of the type described by Harris and Kväerna (2010). How do we implement this in operational systems? An interesting test-case here is seismicity in the European Arctic detected and classified using the SPITS array. The case is interesting because this small aperture array (diameter approximately 1 km) is critical for monitoring needs (due in particular to its proximity to Novaya Zemlya) and produces slowness and backazimuth estimates that require calibration (see Gibbons et al., 2011).

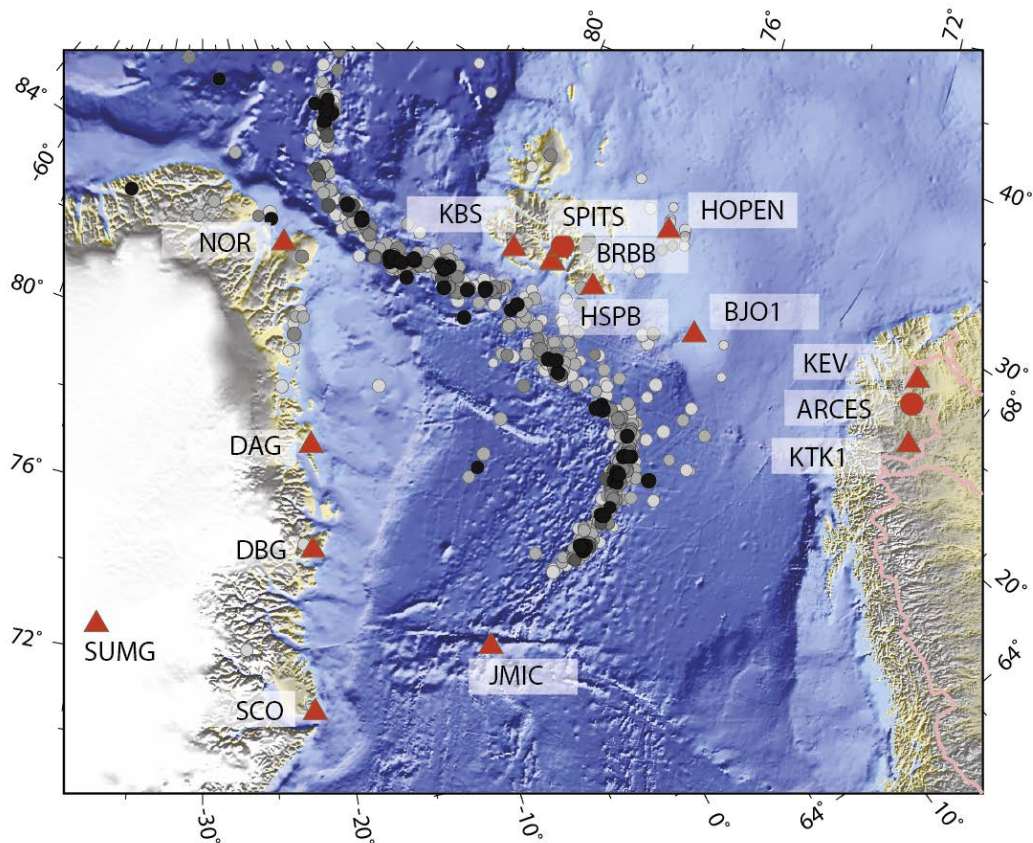


Figure 21: The northern mid-Atlantic ridge. Events north of 72 degrees from Engen et al. (2003) together with the stations in the current network which are most sensitive to low-magnitude events on the spreading ridge. The events are marked with circles with size proportional to magnitude and with the shade a function of the number of associated phase readings (essentially all teleseismic). The palest symbols have as few as 10 defining phases and the darkest have over 100.

Figure 21 displays the location estimates for seismicity between 1955 and 1999 north of 72 degrees as compiled by Engen et al. (2003). Since this time, the highly sensitive SPITS array and additional 3-component stations on Greenland and in the European Arctic have resulted in a significant improvement in the detection and location capability for events in the region. However, most of the available event location estimates (e.g. the routine solutions calculated by NORSAR, and those in the Reviewed Event Bulletin – REB - of the International Data Center in Vienna) use backazimuth measured at SPITS as a location constraint. This disqualifies such solutions from being used to calibrate parameter estimation using the SPITS array.

It was necessary to obtain a set of new and reliable event location estimates that were independent of the parameter estimates obtained from SPITS. The Bayesian hierarchical multiple event location algorithm Bayesloc (<https://www-gs.llnl.gov/nuclear-threat-reduction/nuclear-explosion-monitoring/bayesloc>; see Myers et al., 2007, Myers et al., 2009) was chosen for this purpose. The seismicity covers a range of magnitudes with many significant events with magnitudes greater than 5 that are recorded well teleseismically, in addition to many events of magnitude 3 and below which are recorded only by a subset of the stations displayed in Figure 21. The larger events which are very well constrained using teleseismic phases have far smaller location uncertainties. These uncertainties are reflected in the event priors and this is likely to provide additional constraints on the “regional only” events when they are relocated in the multiple event procedure.

Over 15 years of continuous data was obtained for the stations on the Greenland side of the ridge. Very few of the arrivals on stations such as NOR and DBG have been used in location estimates found in, for example, the bulletin of the International Seismological Center, and it is only very recently that these stations have been used in the routine NORSAR event location estimates.

The preliminary set of relocations using Bayesloc is displayed in Figure 22. These solutions use regional P and S arrivals (largely P_n and S_n) from the stations displayed in Figure 21, in addition to teleseismic P readings taken from the Reviewed Event Bulletin. Note that the initial event location estimates were provided by the HYPOSAT program (Schweitzer, 2001a). The HYPOSAT program is useful in many respects as a single event algorithm, providing very robust preliminary location estimates. Phases with anomalously high time-residuals, for example, are excluded from the inversion such that the arrival input to Bayesloc is already screened for obvious outliers. HYPOSAT in addition provides an initial phase label for each arrival (Pg, Pb, P_n, etc.) even if we only specify “P1” (for first arriving P-phase) – and, once in Bayesloc, the probabilistic phase labelling approach of Myers et al. (2009) is applied.

To give an impression of the improvement that the Bayesloc solutions provide over previously available event bulletins, we compare the solutions displayed in Figure 21 with three available bulletins in Figure 23 a), b), and c) – and with the input solutions from the HYPOSAT algorithm in Figure 23 d). Only panel (d) of Figure 23 shows exactly the same events as displayed in Figure 22. The events displayed in panels a), b), and c) have significant overlap with the events displayed in Figure 22. The ISC solutions for example are only available two or so years before the present. The REB solutions are limited to events recorded by more than one International Monitoring System (IMS) station; there are many events displayed in Figure 23 for which SPITS is the only IMS station that recorded the event well enough for inclusion.

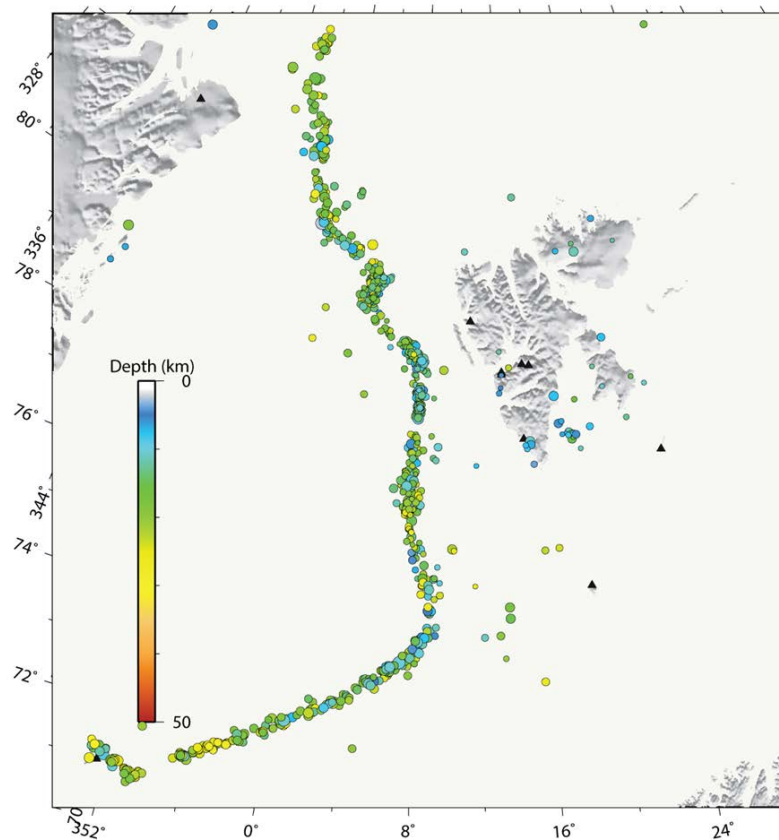


Figure 22: Location estimates for 971 events using the Bayesloc program and both regional and teleseismic phases.

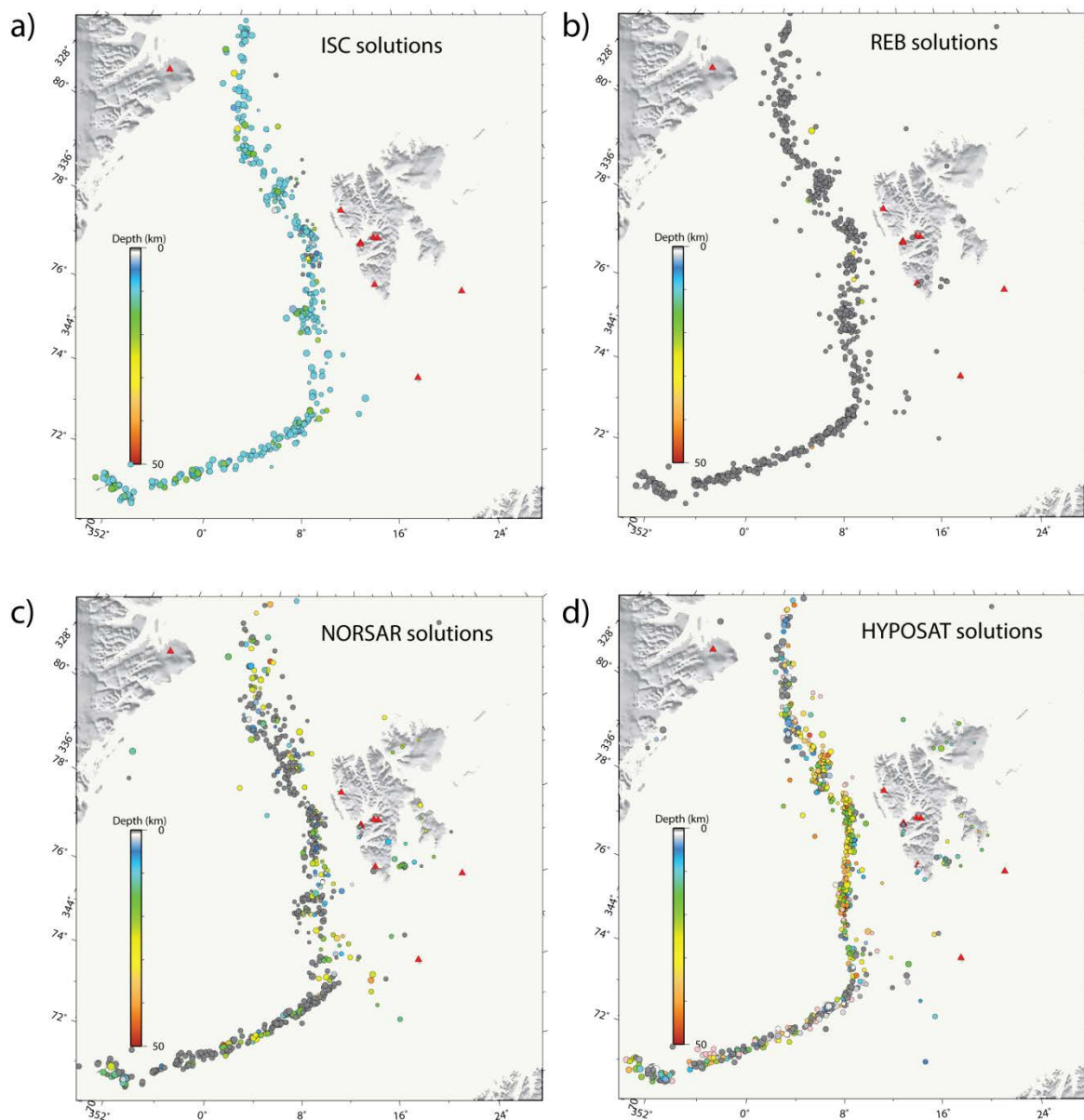


Figure 23: Comparison of event location estimates. (a) Events in the period 1998-2011 from the bulletin of the International Seismological Center (www.isc.ac.uk), (b) Events in the period 2000-2015 from the Reviewed Event Bulletin of the International Data Center for the Comprehensive Nuclear-Test-Ban Treaty, (c) Events in the period 1998-2015 from the routine event location carried out at NORSAR (<http://www.norsardata.no/NDC/bulletins/regional/>), and (d) the events displayed in Figure 22, located one by one using the HYPOSAT program.

A useful insight into the properties of the Bayesloc solutions is provided by examining the residuals associated with one given phase for each event (Figure 24). For a single 1-d velocity model, we display the time residual for the first regional P-arrival at SPITS for the multiple event location estimates (Figure 24 a) and for the single event locations (Figure 24 b). The residual for the multiple event locations varies exceptionally smoothly as a function of source location, whereas it varies significantly with location for the single event location estimates. Like most single event location algorithms, HYPOSAT seeks a solution which minimizes a measure of the total time residual vector. Bayesloc, given the additional constraints provided by the better observed events, favors those solutions with residuals which are consistent with the residuals of neighboring events. The procedure was repeated for several different velocity models and the Bayesloc location estimates were found to vary significantly less than the single event algorithm location estimates; only the residual distributions differ. The significance of this is that, given a continuous band of seismicity containing many events with exceptionally good teleseismic location constraints, Bayesloc can provide quite robust event location estimates that are relatively insensitive to the exact structure of the applied regional velocity model. The spatial statistics of the regional phase residuals provides implicit correction terms for the applied velocity model.

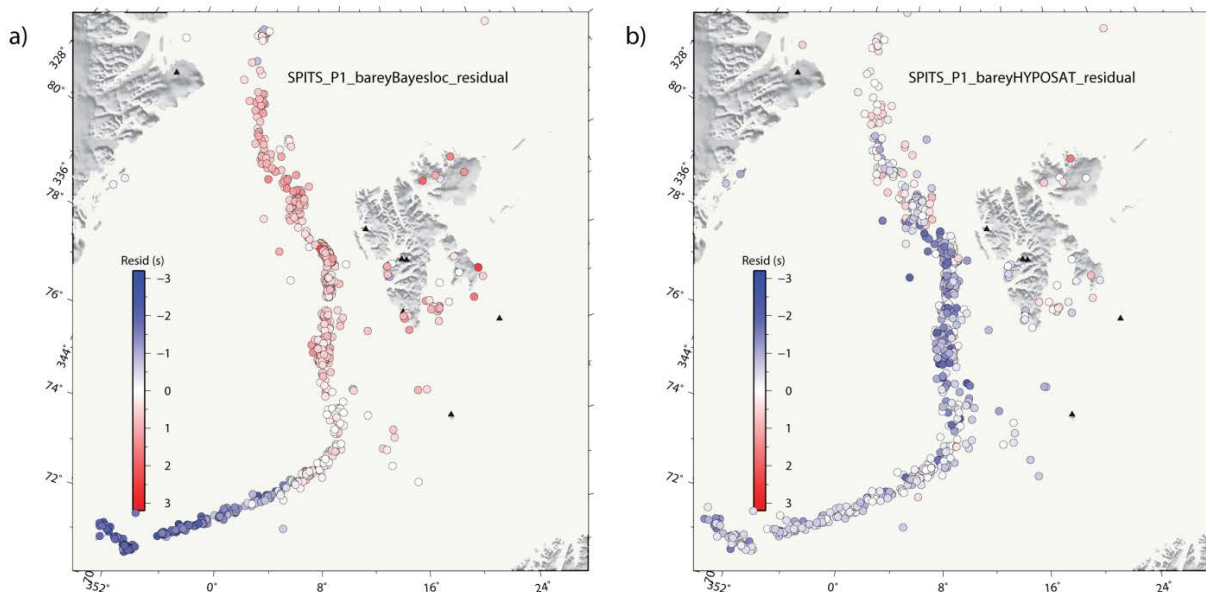


Figure 24: Consistency of time-residuals for Bayesloc origin estimates. Time residual in seconds for the initial P-arrival at SPITS relative to the BAREY velocity model (see Hicks et al., 2004) for (a) the mean solutions provided by Bayesloc and (b) the preliminary single event locations provided by HYPOSAT.

Finally, in Figure 25, we demonstrate how little sensitive to the applied regional velocity model the Bayesloc solutions are. The different velocity models for the crust and uppermost mantle result in different regionalized correction terms but the hypocenters are affected little.

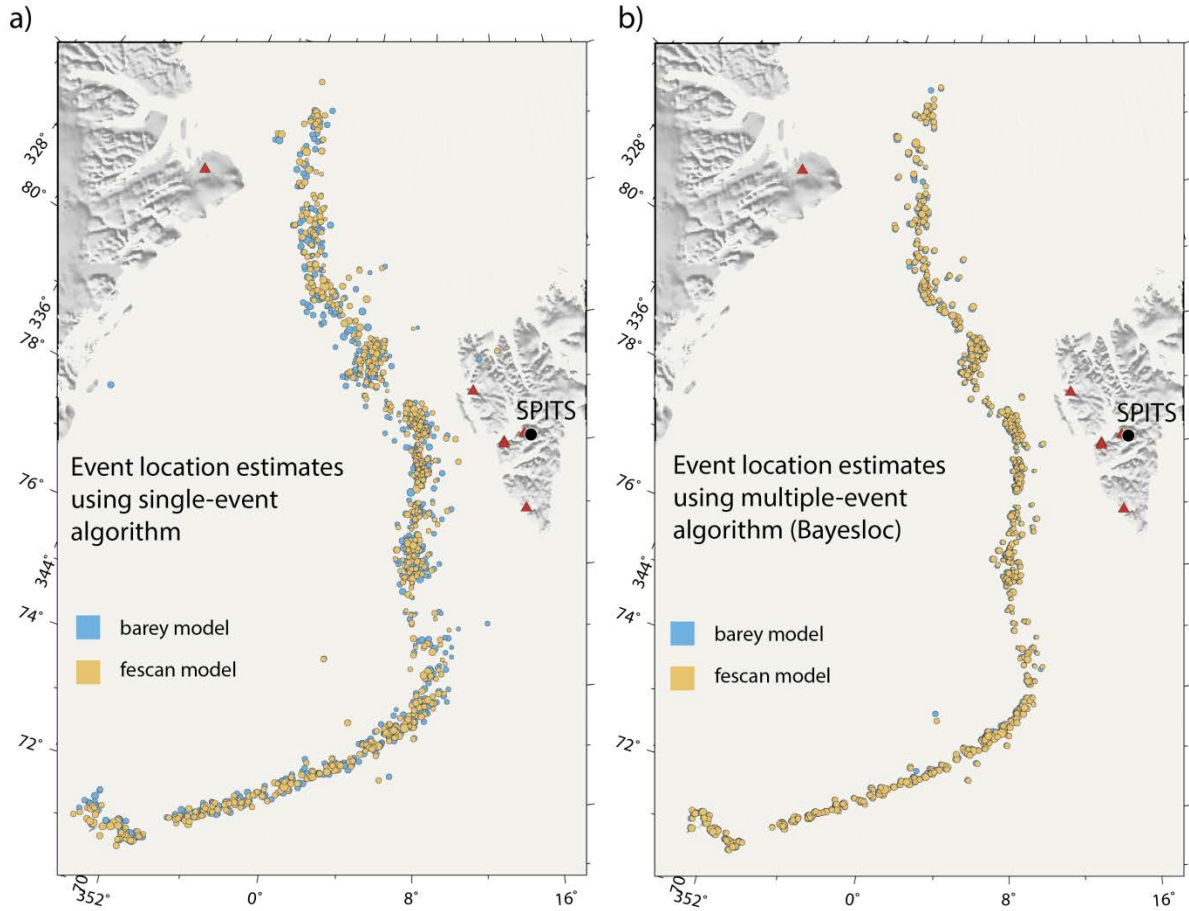


Figure 25: Robustness of Bayesloc event locations to choice of regional velocity model. Event location estimates for earthquakes on the mid-Atlantic ridge between 71 and 84 degrees North using (a) individually located events and (b) the Bayesloc routine. Two different velocity models, *fescan* (Mykkeltveit and Ringdal, 1981) and *barey* (Hicks et al., 2004) have been used. The sets of events displayed are identical for both panels as are the sets of arrival times used.

Given our set of Bayesloc solutions – not explicitly constrained by slowness parameter estimates measured on the SPITS array – we can examine how the array backazimuth estimate varies with the geographical backazimuth (Figure 26). As indicated by Gibbons et al. (2011), there is an enormous variation of the apparent velocity indicated for different source regions. For events to the South of the SPITS array, regional P-phases can generate apparent velocity estimates indicative of regional S-phases. For events to the north of SPITS, regional P-phases generate apparent velocity estimates indicative of teleseismic P-phases. In addition, and of possibly

greater consequence for event location using slowness estimates as a constraint, the apparent backazimuth deviates from the true backazimuth in a significant and almost sinusoidal manner. Existing correction procedures involve mapping a slowness vector using a look-up table from a set of observations to the likely real-world velocity and azimuth (e.g. Bondar et al., 1999, Schweitzer, 2001b). In Figure 16 we display how we can perform this correction empirically using matched field processing at an earlier stage.

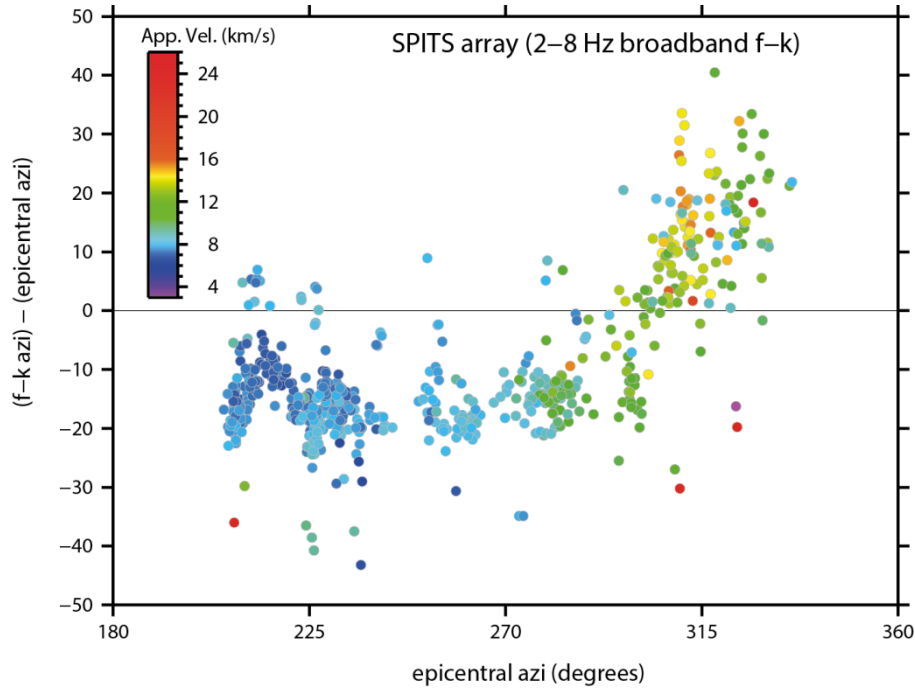


Figure 26: Deviation in backazimuth estimates using classical plane-wave models at SPITS. Variation of the parameter estimates obtained using broadband f - k analysis for the first P -arrival at SPITS as a function of geographical backazimuth obtained using a subset of the Bayesloc locations displayed in Figure 22.

Figure 27 displays how – using the Bayesloc event location estimates – we are able to obtain an empirical parameter estimate using matched field processing in the context of previously observed events. We know from Figure 26 (and previous studies) that a backazimuth correction is likely to be needed and we demonstrate here how matched field calculations can provide an improved parameter estimate which is implicitly calibrated from previous observations. We demonstrate also in Figure 27 how the EMFP estimate maps the direction of the seismic signal back into geographical parameter space and not just slowness space.

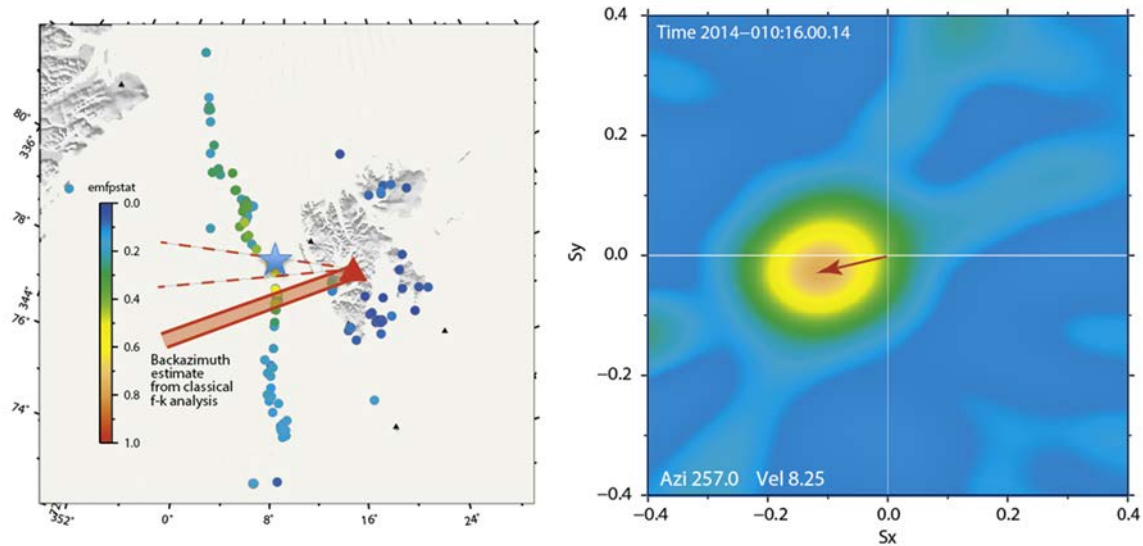


Figure 27: Display of potential utility of empirical matched field processing for providing an empirically corrected direction estimate for a detected phase arrival at SPITS. In the right hand panel we display the slowness plot obtained using conventional and uncalibrated f - k analysis for an arrival on January 10, 2014. In the left hand panel, we see the geographic implication of this slowness estimate for the event location, but also the matched field statistic (see Harris and Kværna, 2010) calculated between the detected phase and the empirical steering vectors calibrated from the P -arrivals from the events displayed. The blue star indicates the independent location estimate for the detected event – not used for calibration purposes.

The European Arctic study indicates the performance of EMFP on small aperture seismic arrays optimized for processing high-frequency regional seismic arrivals. To extend the study to arguably more challenging arrays, we consider also a part of the world which is both of programmatic relevance and monitored by several seismic arrays with very different properties. The region in question is that within regional distances of the MJAR, KSRS, and USRK arrays of the International Monitoring System (IMS) and this includes the Korean peninsula, Japan, and regions of western Russia and China (Figure 28).

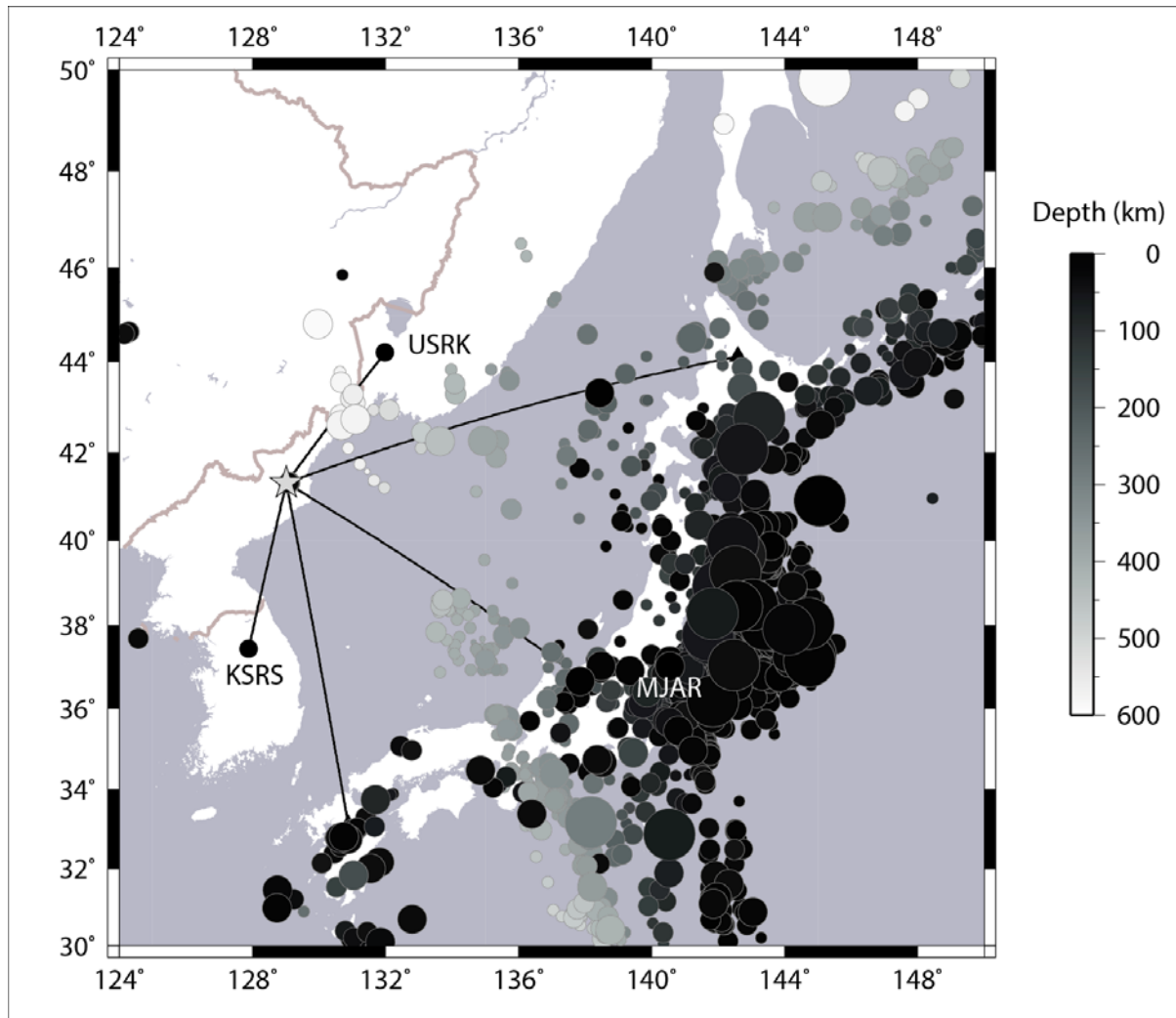


Figure 28: Study region in the Far East for the USRK, KSRS and MJAR seismic arrays. Event locations were obtained using Bayesloc (Myers et al., 2007, 2009) based upon events in the Reviewed Event Bulletin (REB) of the International Data Center (IDC) of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO). The criterion for events (between 2008 and 2016) was that an SNR > 10 was recorded for the P phases on all of the three arrays displayed. The star indicates the location of the North Korean nuclear test site.

A vast amount of natural seismicity is recorded in this region, in addition to significant industrial seismicity and the North Korean nuclear test site. The geometries of these three stations are displayed in Figure 29. From the geometry alone, and from coherence analysis of real world signals (Gibbons and Ringdal, 2012), we anticipate significantly different performance from the matched field detectors. Results are presented in Section 3.2.

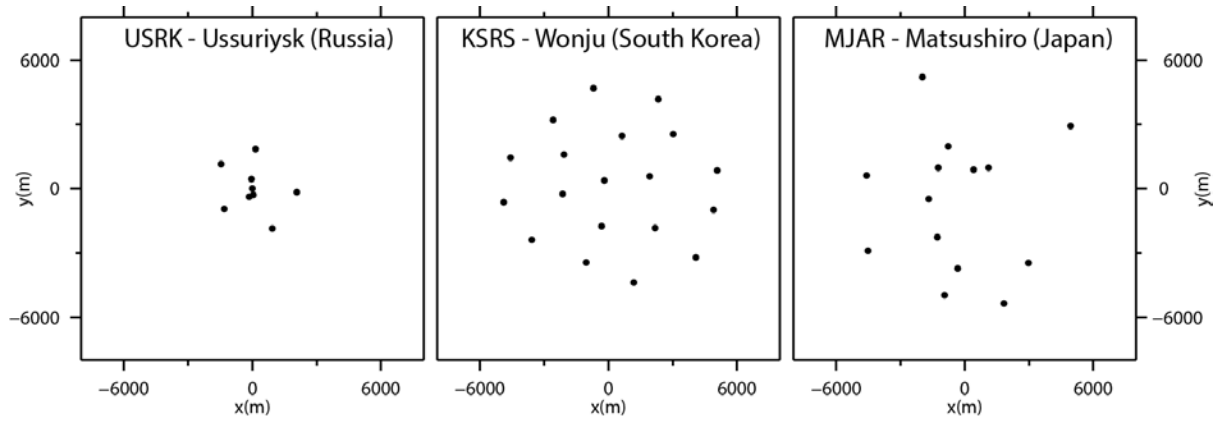


Figure 29: Geometries of the USRK, KSRS, and MJAR seismic arrays. *In addition to the differences resulting from the array geometry, there is a significant difference between the degree of waveform similarity observed between sensors of comparable spacing. The KSRS and MJAR arrays have a similar aperture but, due to geologic and topographic factors, wavefronts crossing the MJAR array are far less coherent than those crossing KSRS. Modified from Gibbons and Ringdal (2012).*

2.3 Parameter Estimation on Seismic Arrays Using Limited Sensor Pair Correlation

In classical f-k analysis, the processing of high frequency signals on large aperture arrays has long proved problematic (see Gibbons et al., 2008, and references therein). With empirical matched field processing, we attempt to mitigate the loss of coherence. However, it may be informative to understand ways of mitigating the problems from a classical array processing point of view in order to give insight into how EMFP may be adapted to cope better with large aperture arrays. Figure 30 highlights the problem of classical f-k analysis of a high frequency signal on a large aperture array. The regional Pn phase has a high SNR and is well-recorded by all sensors of the array. However, it is clear that the shape of the waveform is different on sensors in different regions of the array (Figure 30 c). The f-k spectrum in panel (d) does provide an approximately correct backazimuth and slowness estimate, but the image is speckled and the favored slowness vector is only selected in marginal preference to many other candidate slowness vectors.

There are in fact two distinct factors at play here. *Incoherence* refers to the dissimilarity of signals resulting from geological heterogeneity. We use the term *aberration* to describe the fact that coherent wavelets may be delayed by time-shifts that are not consistent with the applied wavefront model. The non-symmetric shape of the peak in the f-k spectrum indicates that both of these factors contribute to the poor direction estimate.

Ruigrok et al. (2017) draw together a large volume of literature to summarize the potential advantages in estimating the slowness using cross-correlation functions rather than original waveforms. One crucial observation is that we can typically remove a given channel-to-channel correlation function from the calculation without fundamentally losing data. Specifically, if we were to remove channels on which the correlation was poor, we could likely expect an improvement in the robustness of the slowness estimate.

Fig. 31 b) displays the so-called co-array of the Malin array (displayed in Fig. 30 a). Note that there is no point in the center of the co-array; this is because we do not calculate the autocorrelation functions. To the left (Fig. 31 a) is the f-k spectrum of the set of all cross-correlation functions for the entire co-array (the pale symbols in Fig. 31 b). The f-k power is low and, like the f-k spectrum of the original waveforms (Fig. 30 d), indicates a slowness estimate within the appropriate range but without convincing contours of beam gain. To the right (Fig. 31 c) is the f-k spectrum of the cross-correlation functions restricted to the pairs of channels, i and j for which the inter-site spacing is limited to under 6 km (the dark symbols in Fig. 31 b). In this example, the f-k maximum is far clearer and the contours of f-k power diminish far more smoothly and with greater symmetry. While the theoretical resolution of the calculation displayed in Fig. 31 c) is lower than that in Fig. 31 a), it is clear that the correlation functions

from the sensor pairs with greater separation are likely to distort and diminish the peak of the f-k spectrum due to both signal incoherence and aberration.

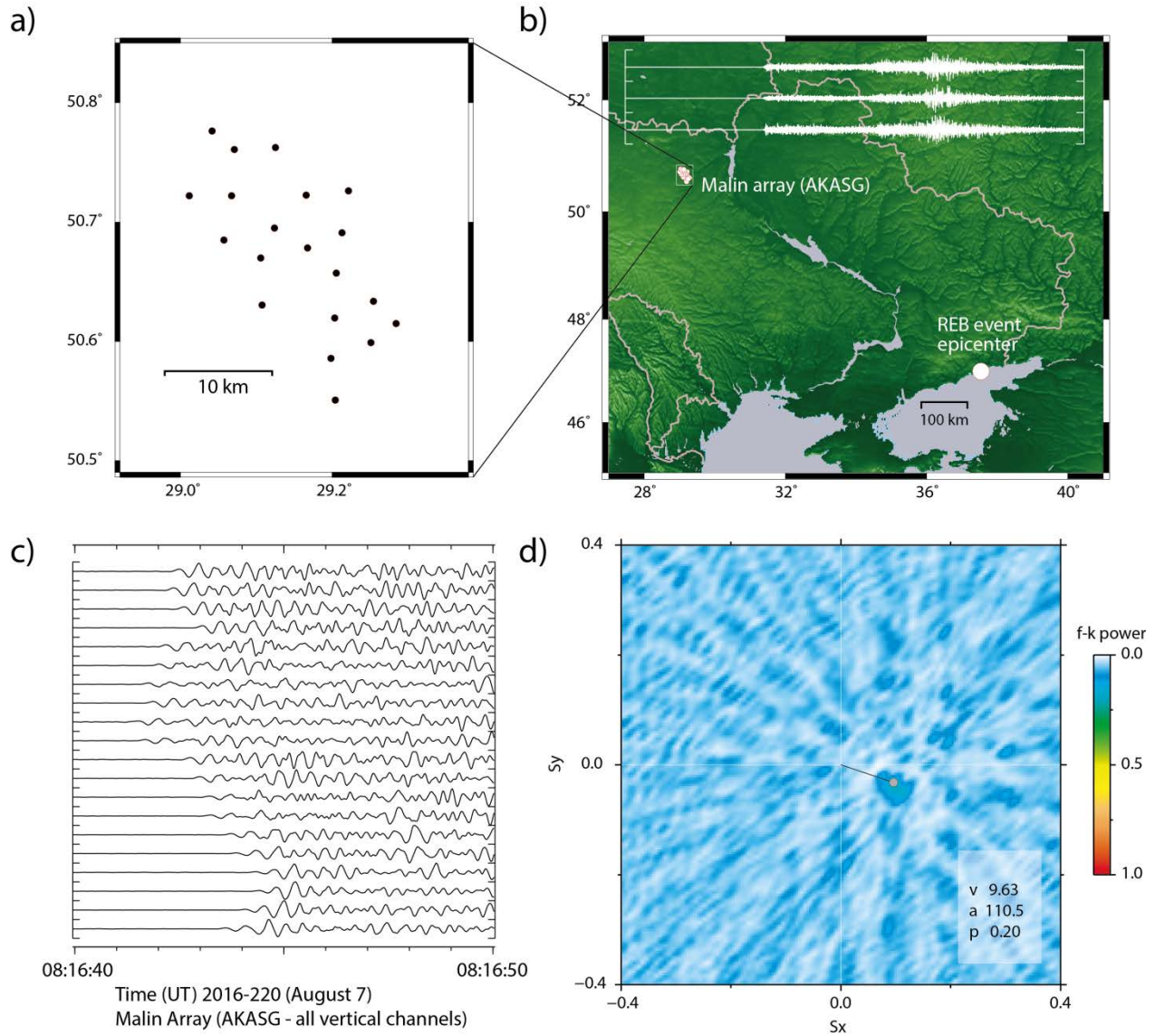


Figure 30: Observation of a magnitude 4.5 earthquake near to the city of Donetsk recorded on the large aperture AKASG array (Malin) in Ukraine. (a) shows the array geometry and (b) its location in the geographical context with waveforms from three sites bandpass filtered 1-4 Hz. (c) shows a zoom-in of the Pn-arrival in this frequency band and (d) the frequency-wavenumber spectrum for this arrival. For (d), v , a , and p denote v_{app} (km/s), backazimuth (degrees), and relative power respectively.

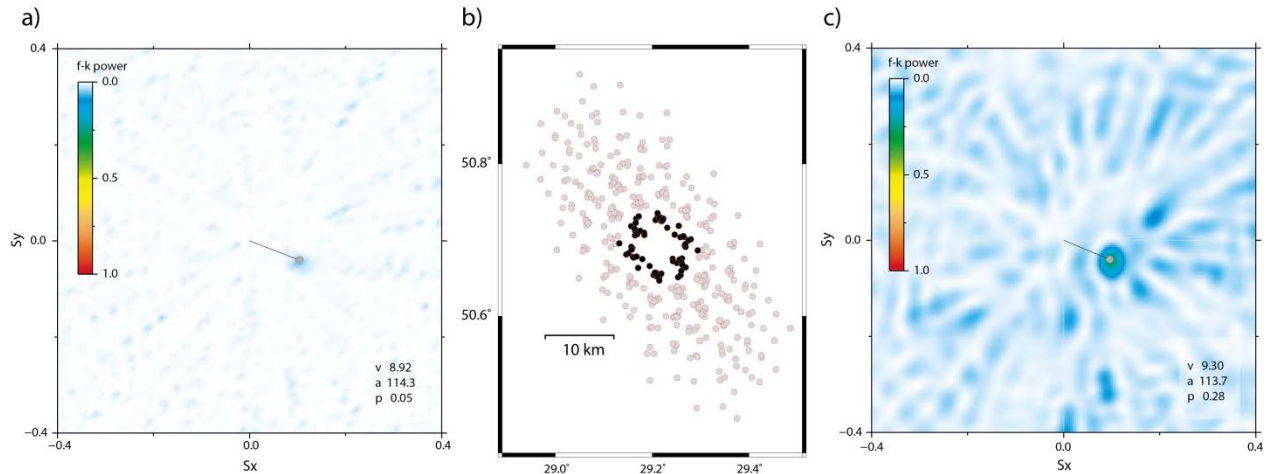


Figure 31: f-k analysis for the AKASG arrival on correlation trace virtual co-array. (a) using the full array (i.e. the pale symbols of the sensor-separation co-array in panel b) and (c) using the co-array limited to sensor separations less than 6 km (i.e. the dark symbols of the sensor-separation co-array in panel b). For each panel, v , a , and p denote v_{app} (km/s), backazimuth (degrees), and relative power respectively. The locations of the elements of the co-array are vector distances between the true geographical sensor locations such that the absolute locations of the co-array sites are undetermined. We arbitrarily fix the center of the co-array to 50.6911 N, 29.2131 E.

In Fig. 32 we display a teleseismic signal, with satisfactory SNR in a relatively narrow frequency band (2-4 Hz), on the large aperture NORSAR array. The signal is considered in detail by Gibbons et al. (2008), demonstrating that a slowness estimate using classical f-k analysis over the full array does not provide a robust estimate. In addition, it is demonstrated that a completely incoherent slowness estimate over the full-aperture NORSAR array had a similar resolution to a coherent f-k estimate using a only one single small-aperture sub-array. Fig. 32 b) displays the NOA co-array with two limited subsets of sensor-pairs displayed. One, consisting of all sensor pairs with site separations less than 26 km, contains only a fraction of the total number of sensor-pairs but does include subarray-to-subarray correlations. The second, restricted to sensor-to-sensor distances less than 6 km does not even include all sensor-pairs within a given subarray. The f-k spectrum of the first of these correlation co-array subsets is displayed in Fig. 32 a) A speckled image is obtained with local maxima covering very limited regions of slowness space. The global maximum indicated does not fall in the center of the region with the greatest values of f-k power. Restricting the distance between sensors further results in a far more convincing slowness grid (Fig. 32 c) with the global maximum falling in the center of a symmetric form.

The difficulty of estimating the slowness of a high frequency seismic arrival over a large array aperture is that a small change in slowness space corresponds to a difference in the time-delay which is large in comparison with the period of the signal. A 4 Hz signal has a period of 0.25 s

and a small difference in the slowness vector can easily result in the 0.125 s required to make the signal entirely out of phase. For a perfectly coherent signal which is described perfectly by the plane wavefront model, this property results in a very high accuracy and precision in parameter estimation. Should the coherence or fidelity of the delay-time model diminish across the array aperture then such significant changes in the delay-time will result in a rapidly changing pattern of array-gain or coherence with slowness and the estimate is (a) not likely to be stable and (b) not easy for an analyst to visualize. LSPC in effect makes a large aperture seismic array appear to have a far smaller aperture, while still exploiting phase shifts recorded between many different pairs of sensors.

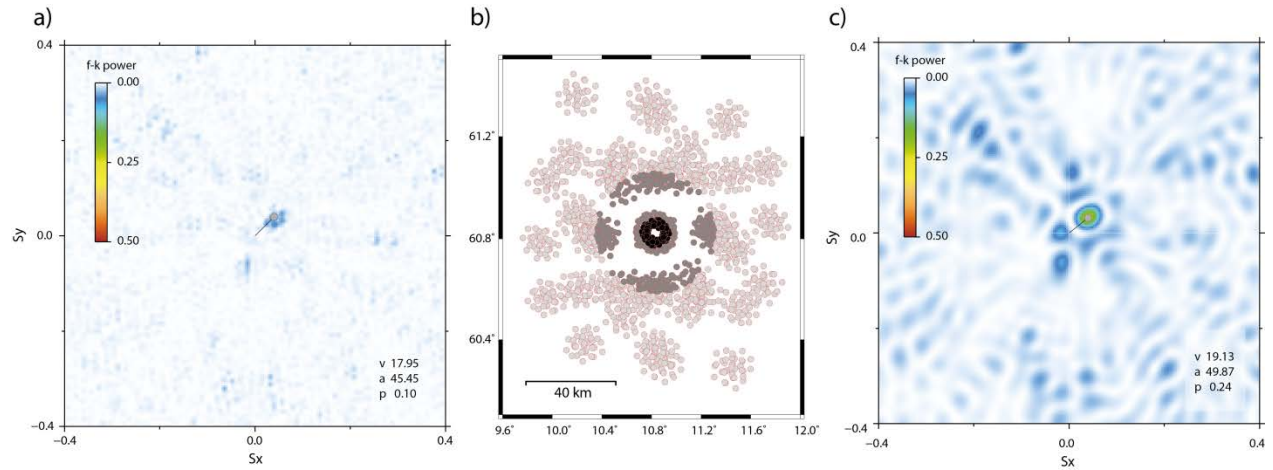


Figure 32: Slowness estimation for the teleseismic P-arrival at NOA from the declared October 9, 2006, underground nuclear test in North Korea, using 2-4 Hz f - k analysis on the correlation co-arrays for different subsets of sensors. (b) displays the complete co-array geometry with two shaded subsets: less than 6 km (black symbols) and less than 26 km (grey symbols). We arbitrarily fix the center of the co-array to 61.0397 N, 11.2148 E. (a) shows the f - k spectrum for the 26 km sensor-separation subset and (c) shows the f - k spectrum for the 6 km sensor-separation subset. For each panel, v , a , and p denote v_{app} (km/s), backazimuth (degrees), and relative power respectively.

2.4 Matched Field Detection: A Case Study. The North Korea Nuclear Test-Site

We wish to use continuous empirical matched field detectors to identify the signals from underground nuclear explosions carried out by the Democratic Peoples' Republic of Korea (the DPRK or North Korea). "Front end" EMFP detectors, formed using the kind of procedure described in Section 2.1, were set up on the KSRS and USRK arrays using empirical templates for Pn arrivals from previous declared explosions. These detectors are now operational and running in near real-time. On September 3, 2017, the sixth declared DPRK underground nuclear test was carried out. The confirmed nuclear test corresponded to a seismic event of body wave magnitude (mb) 6.1 and was followed some 8 minutes later by a seismic event of mb 4.1, generally referred to as an aftershock. Two further seismic events from on, or very close to, the Punggye-ri Nuclear Test Site were detected and identified on September 23, 2017, and then a further event on October 12, 2017. Figure 33 displays the waveforms from these 5 events recorded on the IC.MDJ station (Mudanjiang, Heilongjiang Province, China).

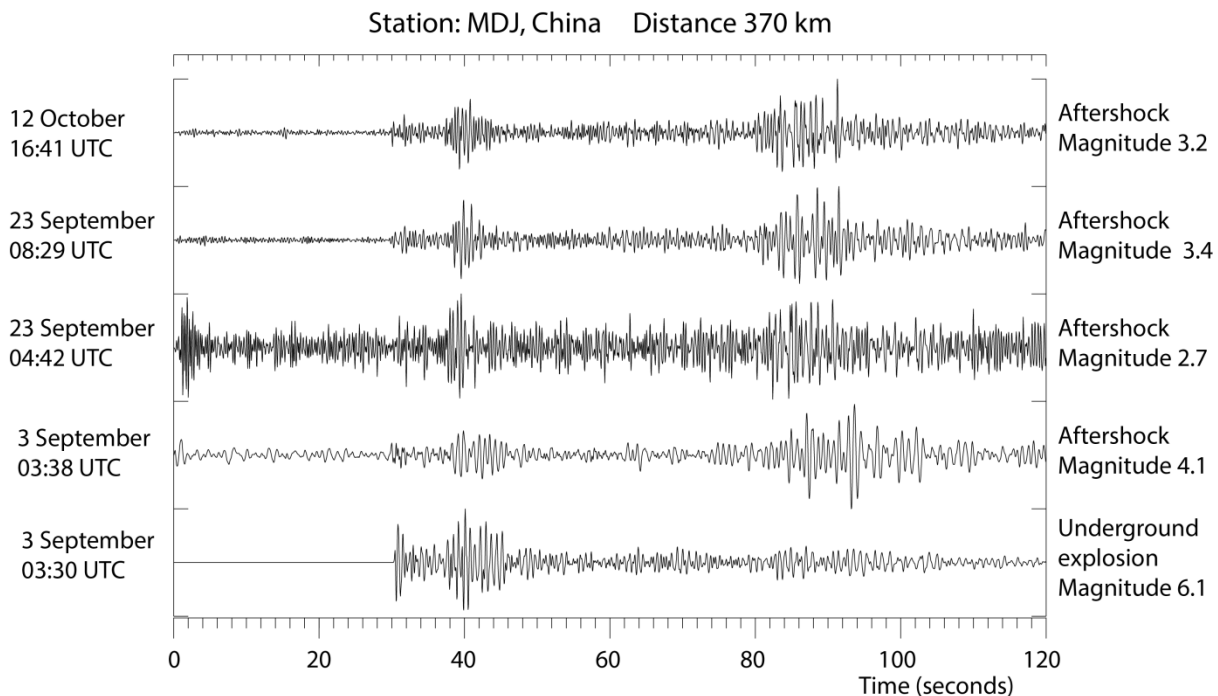


Figure 33: The five seismic events at or very close to the North Korea nuclear test between September 3, 2017, and October 12, 2017, recorded on station IC.MDJ. This station is the closest open station to the test site at a distance of 370 km. All waveforms bandpass filtered between 1 and 4 Hz.

We have demonstrated how a matching of narrowband phase-shifts over a seismic array can generate a scalar detection statistic which indicates a sudden correspondence of phase-shifts in

the incoming data with those observed in a theoretical or empirical template. At the time of the DPRK event on October 12, 2017, we had 10 available Pn templates applicable to the DPRK test-site. In addition to the recordings of the six declared nuclear tests, we have recordings of the September 3, 2017, aftershock, the two events on September 23, 2017, and the theoretical plane wavefront template (corresponding to 8.2 km/s apparent velocity and 12 degrees backazimuth). All 19 sensors of the KSRS array were used for template generation.

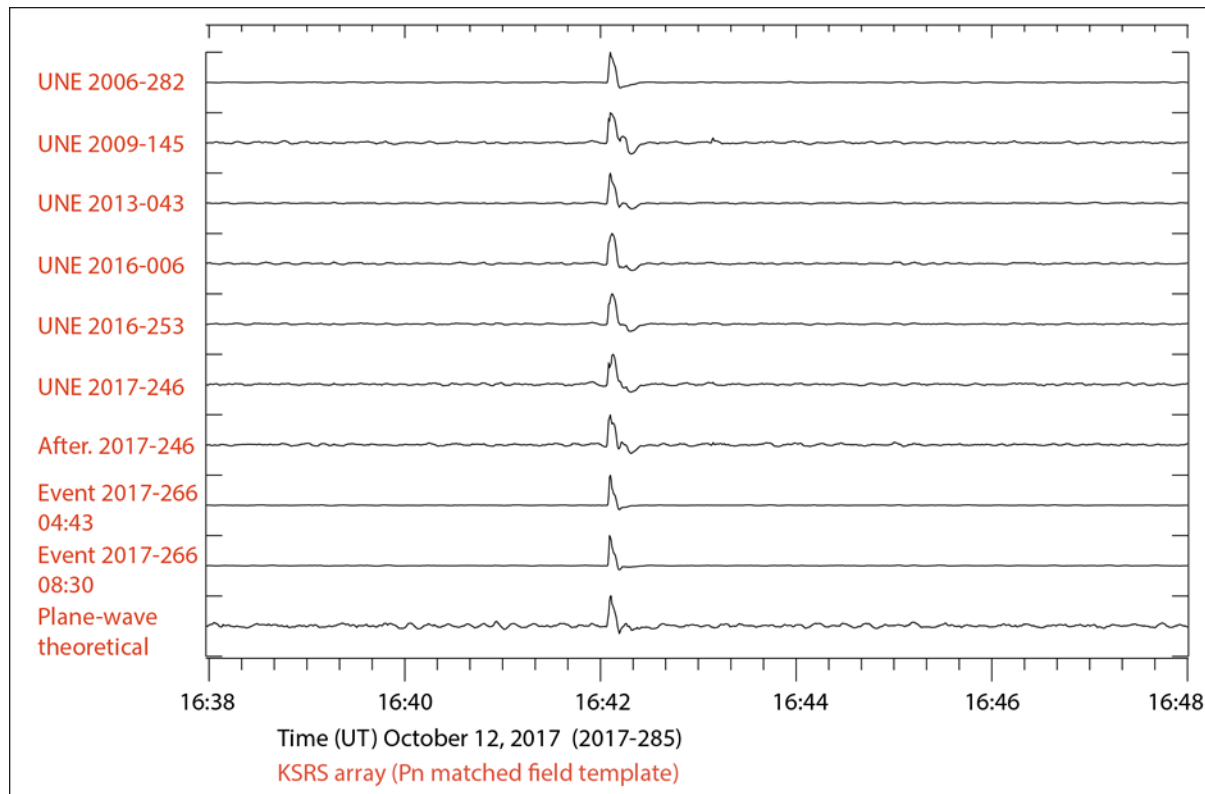


Figure 34: Matched field detection statistics for a ten minute long data window on the KSRS array on October 12, 2017, for Pn templates as indicated.

Figure 34 displays the scalar matched field detection statistics for each of these 10 Pn templates for the time-interval surrounding the October 12, 2017, event. Each of the templates generates a significant rise at the time of the Pn arrival and a low background level at all other times (this is to say no false detections in this time-interval). The plane-wave theoretical template performs less well than any of the empirical templates

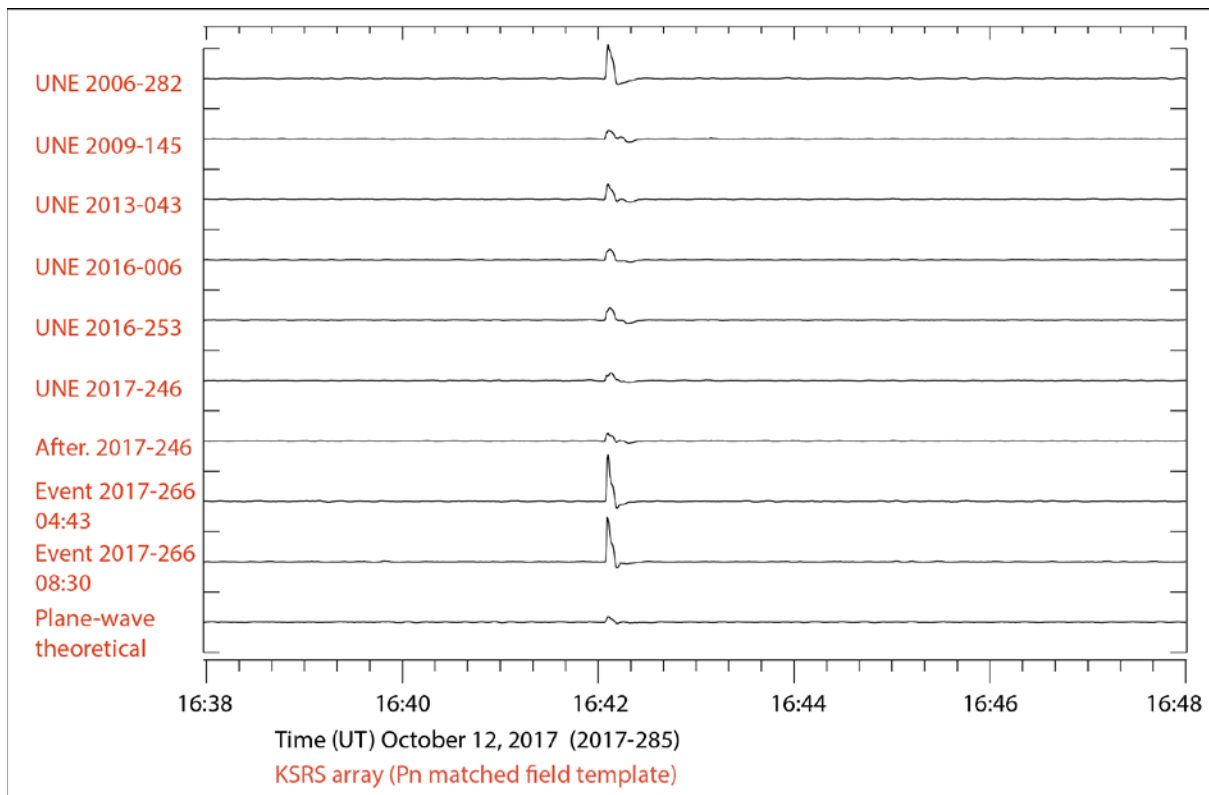


Figure 35: Matched field detection statistics for a ten minute long data window on the KSRS array on October 12, 2017, for Pn templates as indicated. Displayed to a common scale.

Figure 35 displays the same traces as Figure 34, only shown to an identical vertical scale. We see that 3 of the matched field templates perform significantly better than the other seven: the three well-performing matched templates are the two events from September 23, 2017, and the declared DPRK nuclear test on October 9, 2006.

A similar procedure was followed for the USRK array Pn arrival. There is unfortunately no recording of the October 9, 2006, declared DPRK nuclear test on this array and so only 9 matched field templates are available for this arrival. (We can of course generate matched field templates with potentially different performance by considering different subsets of sensors, window length duration, and frequency band.) Figure 36 displays the equivalent detection statistics to those displayed for the KSRS array for these nine templates.

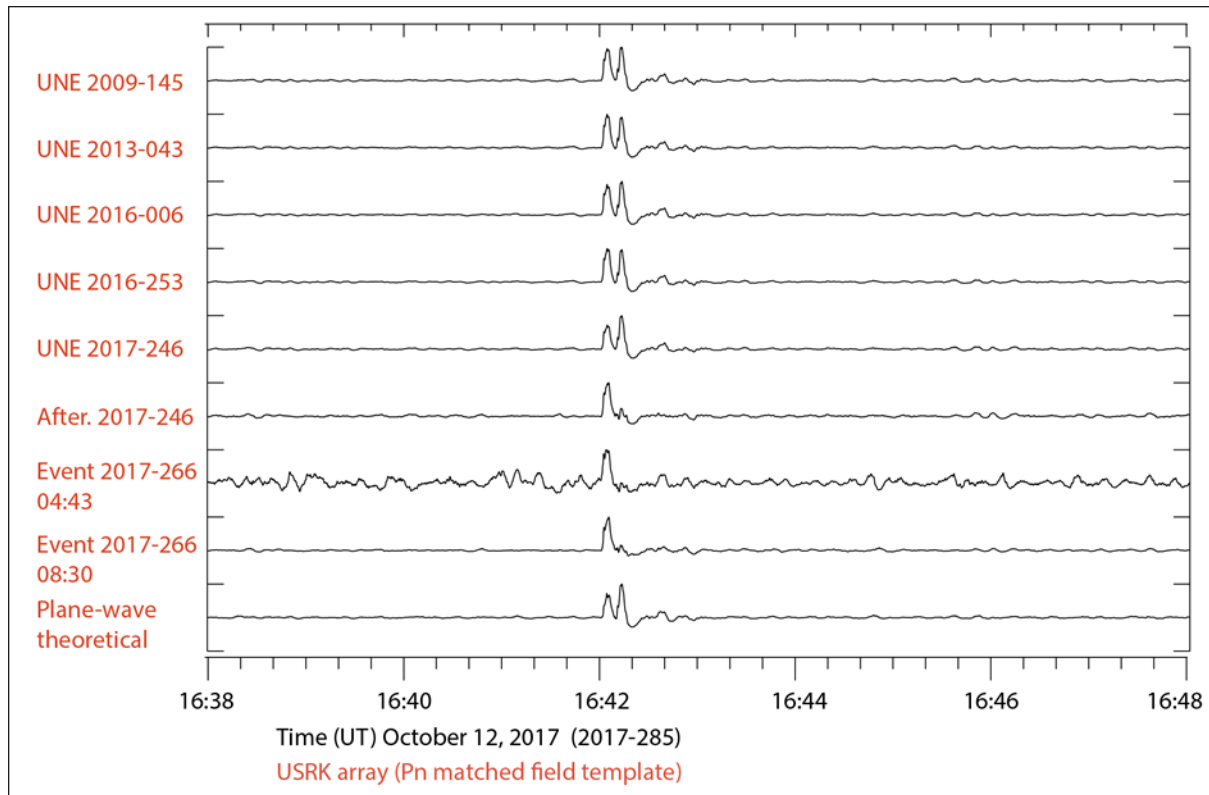


Figure 36: Matched field detection statistics for a ten minute long data window on the USRK array on October 12, 2017, for Pn templates as indicated.

There are significant differences between the USRK detection statistics displayed in Figure 36 and the corresponding traces for the KSRS array in Figures 34 and 35. The background noise level is somewhat higher for the USRK array. While signals for Punggye-ri events are usually recorded on USRK with a higher signal-to-noise ratio (SNR) than on KSRS, there are only 9 sites at USRK compared with 19 at KSRS. The smaller array aperture, in addition, means greater noise coherence and less variety in the signal waveforms across the array aperture. This means that there is less unique information resulting from the scattering which can improve the contrast in the matched field statistics from the different sources (Harris and Kværna, 2010).

For most of the templates, we see two distinct peaks. These peaks correspond to the first Pn arrival and the later Pg arrival some 10 seconds later. (The Pg phase is far more powerful for this site recorded at USRK than it is for KSRS.) Inspection of the matched field pseudospectrograms from which these scalar traces were generated, it is clear that the Pn detection is clearer at higher frequencies and the Pg detection is clearer at lower frequencies. The Pg phase detection is far less clear for the September 3 aftershock and the September 23 events than for the other templates. It is unfortunate that no signal for the 2006 event exists as a second observation in a different direction of a good correspondence between the October 2006 event and the October

12, 2007, event narrowband phase patterns would be significant. The differences between the September 3, 2017, events and the three subsequent events are seen more clearly in the higher frequency band 4-8 Hz on the USRK array (Figure 37).

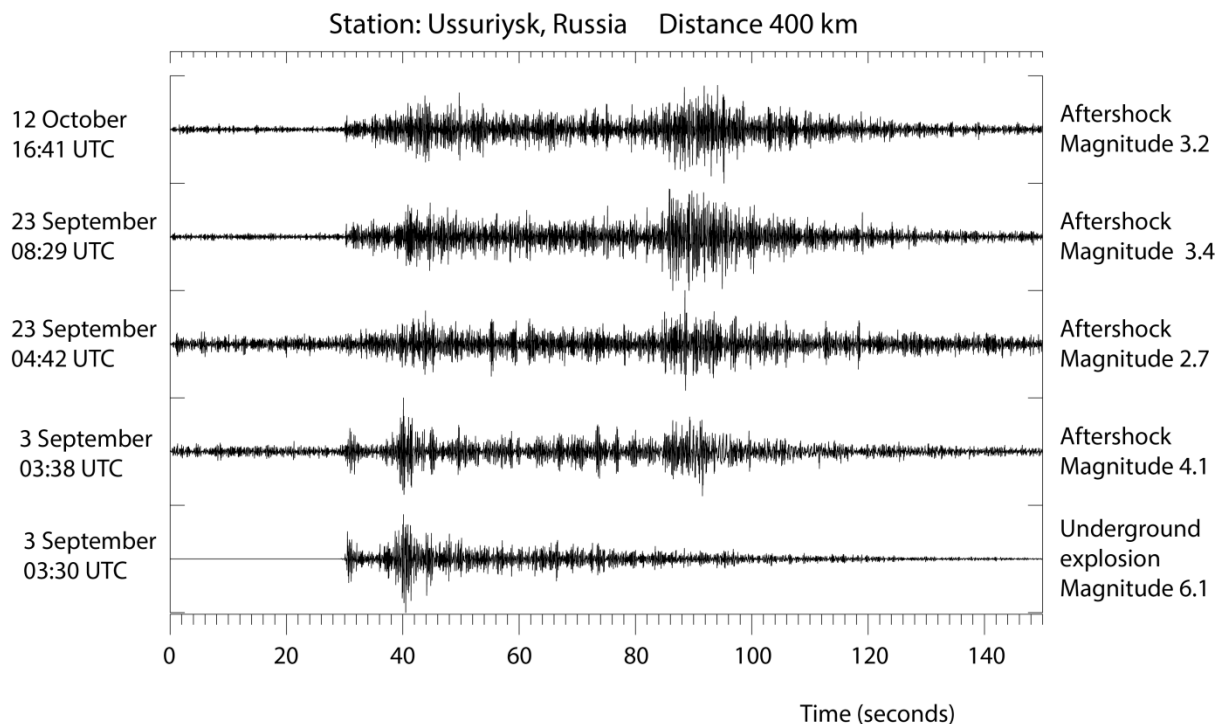


Figure 37: The five seismic events at or very close to the North Korea nuclear test between September 3, 2017, and October 12, 2017, recorded on the IMS station USRK (Russian Federation). All waveforms bandpass filtered between 4 and 8 Hz.

In earlier work carried out in this contract, we have demonstrated how the stability of classical f-k calculations can be improved by only considering sensor pairs with limited spatial separation. Instead of performing f-k analysis on the actual waveforms, we calculate correlation traces between sensors and form a “virtual” array from these traces; the coordinates of the virtual sensors – at which the correlation traces are stored – are the components of the translation vector from the first to the second sensor. Classical f-k analysis can then be performed on this “correlation co-array”. In this co-array, we are free to remove given sensor pairs; by removing sensor pairs for which the signal correlation is poor, we can generate more stable direction estimates.

All matched field processing is performed by calculating a covariance matrix between the different channels of an array. The matched field steering vector for a given covariance matrix is the principal eigenvector of that matrix. The element of the covariance matrix corresponding to two sensors at which the signals do not correlate well may result in a steering vector which is

highly specific to the exact signal from which the covariance matrix was generated and may well not represent signals from nearby sources very well. Imagine three sensors: i , j , and k . The signals on sensors i and j are very similar as are the signals on sensors j and k . The sensors i and k , being further apart, have less similarity and the estimated phase shifts between channels i and k may be at odds with the phase shifts measured between channels i and j and the phase shifts made between channels j and k . Setting the element i,k of the covariance matrix to zero is saying that there is no correlation between elements i and k : the phase shifts in the empirical steering vector are determined entirely by the elements i,j and j,k of the covariance matrix. We repeat the calculation displayed in Figure 34 with the P_n templates for the DPRK events on the KSRS array but by setting the elements of the covariance matrix for sensor separations greater than 4 km to zero.

The results are displayed in Figure 38. The responses to the different templates have become more uniform than before. The plane-wave theoretical steering vector performs considerably worse than it did for all elements of the co-array. The best-performing empirical steering vectors from Figure 34 now perform rather less well whereas the worse-performing vectors from Figure 34 are now better. In this sense we have broadened the footprint of the matched field detectors – made them less sensitive to the exact source location from which the template was calibrated.

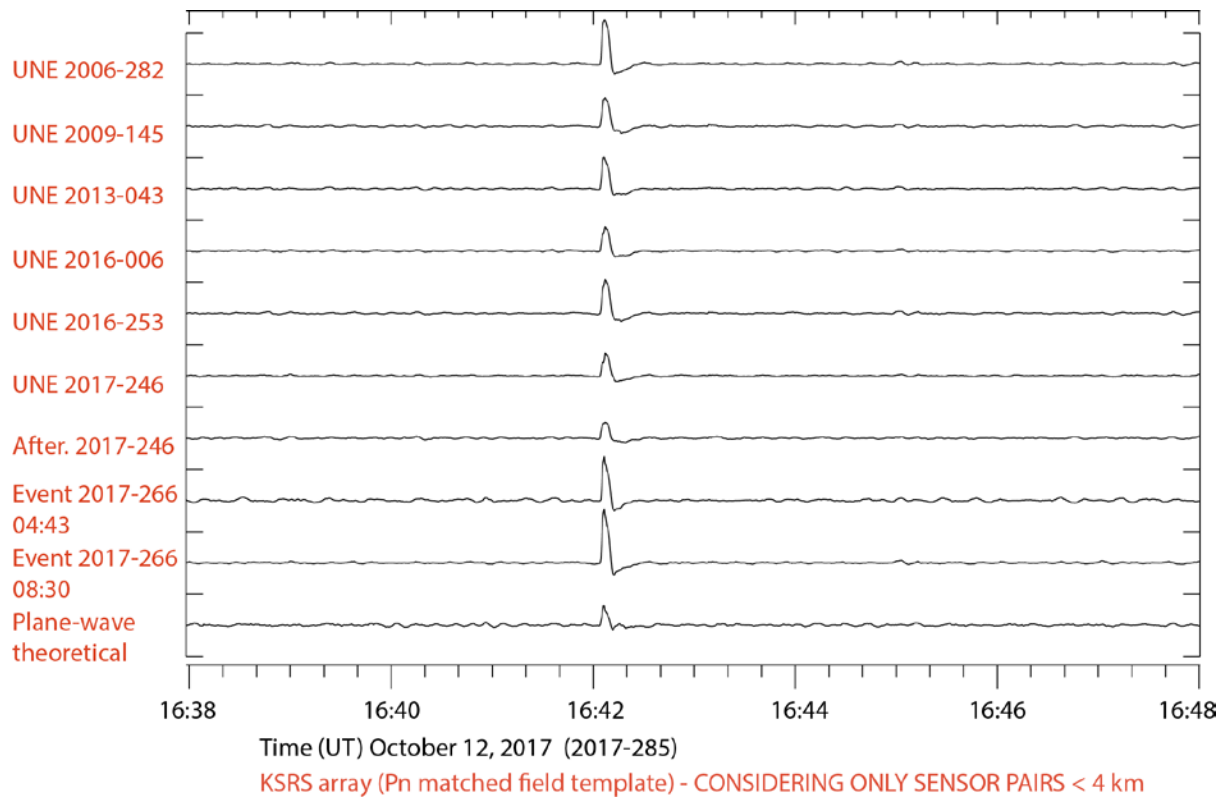


Figure 38: Matched field detection statistics for a ten minute long data window on the KSRS array on October 12, 2017, for Pn templates as indicated. The traces are calculated in the same way as for Figure 34 except that all elements in the covariance matrices corresponding to sensor pairs with inter-site distances greater than 4 km are set to zero.

2.5 The Physical Basis for Matched Field Processing

2.5.1 Introduction

Empirical matched field processing is reliant on the availability of historic seismicity as a source of calibration templates for an observing aperture. The number of templates required to calibrate propagation from a source region to a receiving aperture is a function of the size of the source region, the size (resolution) of the aperture and, we imagine, the degree of heterogeneity of the region.

One of the principal objectives of our research is to link characteristics of the physical medium to the degree of variation observed in matched field templates. We anticipate that regions with greater heterogeneity will show greater diversity in templates (due to more pronounced scattering) than more homogeneous regions. An ability to predict the degree of template diversity from models of heterogeneity in the crust and upper mantle could be useful operationally to estimate the number and distribution of master events required to provide an adequate calibration for matched field processing detectors in any given region.

In order to establish the link between structural heterogeneity and template diversity, we need also to model template variation with synthetics. This requirement colors the event groups and regions we examine; we will have more computational options for modeling if the length of propagation paths from source region to observing array is not too great. With available computational resources, we confine our efforts at simulation to long periods and, principally, surface waves. Thus, we seek the perfect storm of observational conditions: (1) an array of sufficient aperture to resolve variations in wavefield structure, but with sufficient density to avoid severe aliasing problems, (2) source regions of two types: one homogeneous and one heterogeneous, with adequate numbers of events to characterize template behavior across the regions, and (3) paths between source regions and arrays of lengths that allow modeling computationally. Simultaneous satisfaction of these conditions is not trivial to fulfill and has led us to examine a number of possibilities.

In this project, we collected and examined data from four regions and from four observing arrays or networks: the Sumatran arc to Pilbara, the Sumatran arc to Warramunga, the Pacific Antarctic Ridge (PAR) to the southern California network (SCEC), the Central Atlantic Ridge to the German Regional Seismic Network (GRSN) and the Zagros (and adjoining areas of Iran) to the GRSN. The Pilbara data were too spotty and noisy, the Warramunga array too small for adequate resolution of wavefield heterogeneity, the distance from the Pacific Antarctic Ridge (PAR) to southern California limited our computational options, was somewhat sparsely sampled by events and was recorded with too poor SNR by SCEC (though the PAR was a nicely

homogeneous region). Finally, we chose the northern mid-Atlantic Ridge (MAR) and the Zagros as observed by the GRSN. The MAR is our model for an area with relatively simple structure and the Zagros was chosen as an area with greater heterogeneity.

2.5.2 Data and Methods

Figure 39 shows the locations of study events from the National Earthquake Information Center (NEIC) catalog (magnitude 5 and greater) in the two study regions: the mid-Atlantic ridge and Iran. Data were downloaded from the BGR and GFZ web sites via their FDSN interfaces. We used ObsPy to remove instrument responses to produce displacement seismograms for all three components of 13 stations of the GRSN (GRA1, GRB1, GRC1, FUR, BFO, BRG, BUG, STU, OX, TNS, CLL, CLZ, and WET). For the time period 1997-2017, observations for 233 events along the mid-Atlantic ridge and 169 events in the Zagros/Iran region were obtained. The events were visually screened for acceptable SNR (> 5 generally) and superimposed events, leaving 162 usable mid-Atlantic ridge events and 116 usable Zagros/Iranian events.

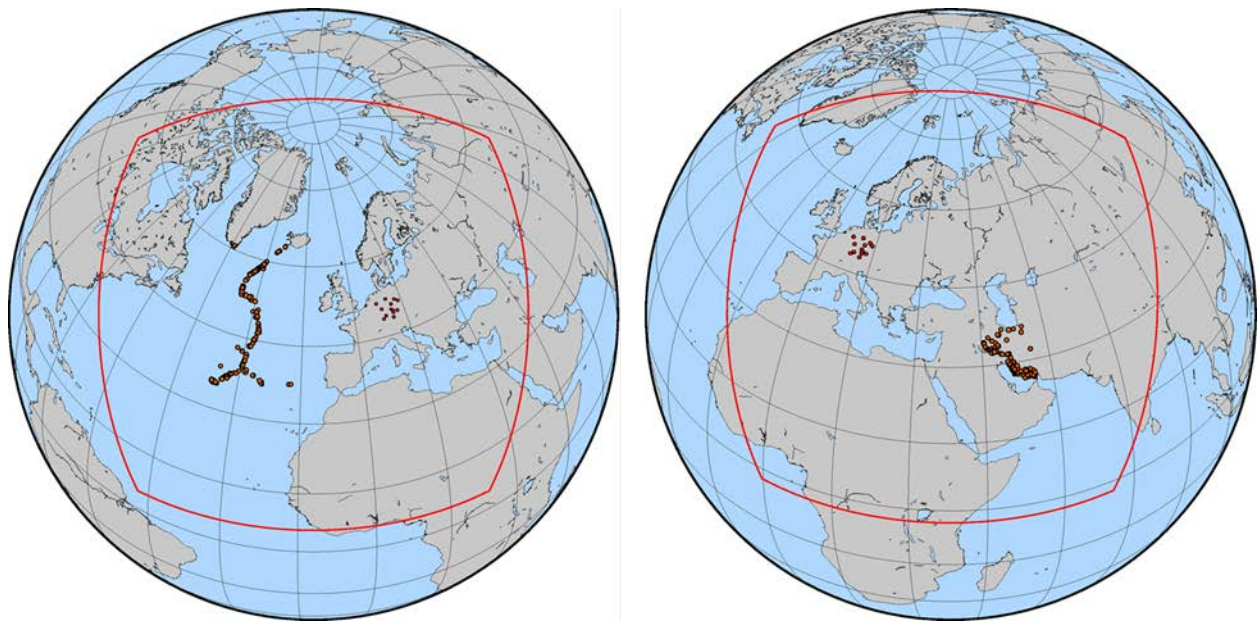


Figure 39: Study regions for this part of the project. Thirteen stations of the GRSN were used as the receiving array because of the stability of that network over 20 years of observation and its proximity to two highly active regions, the northern mid-Atlantic ridge (MAR) and the Zagros mountains (and adjacent areas) of Iran. The MAR was chosen as a relatively simple source region and Iran as more complex region, both in propagation medium heterogeneity and variability among source mechanisms.

Figure 39 also shows (with red boundary outlines) the computational domains that were used to obtain simulation waveforms for all of these events. We will have more to say about the simulations later.

In addition to obtaining waveform data, we also obtained mechanism estimates for as many of the events as possible from the Global CMT project (Dziewonski et al., 1981; Ekström et al., 2012). Not all had GCMT mechanisms, but those that did are mapped in Figure 40. Our overall impression is that the events along the mid-Atlantic ridge had somewhat more homogeneous mechanisms than the events in Iran and the Zagros. We note, in particular, that the paths from the MAR events to the GRSN are more broadside to the nodal axes of the mechanisms of those events, while the paths from the Zagros/Iranian events to the GRSN are more nodal to the majority of event mechanisms. The “endfire” or nodal paths (with respect to source mechanism) result in less consistent matched field processing responses for the Zagros/Iran events than the MAR events.

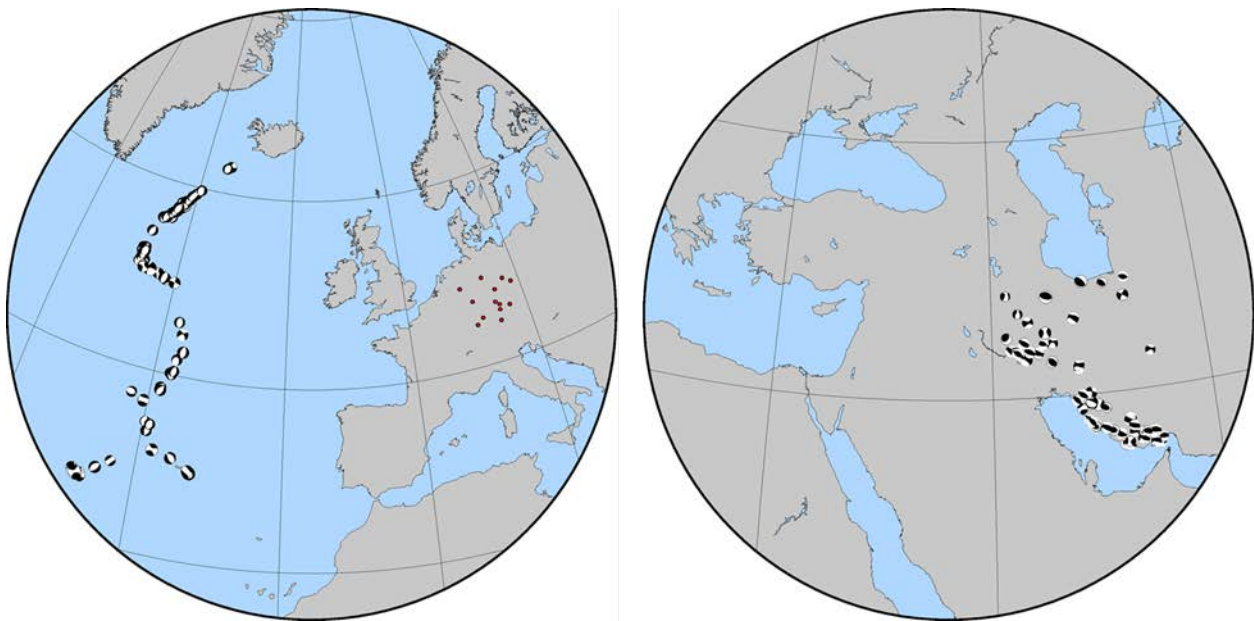


Figure 40: Global CMT project mechanisms for the two study regions. Events along the northern mid-Atlantic ridge have more consistent mechanisms, and in Iran somewhat more divergent mechanisms.

Matched field processing (MFP) as considered here produces a map of an energy incident on the array (GRSN) as a function of geographic source location. It is conceptually similar to the frequency-wavenumber (FK) spectrum, which produces a map of incident energy as a function of vector wavenumber, or, alternatively, vector slowness. Both MFP and FK analysis perform the mapping through evaluation of the quadratic form of equation (1):

$\mathbf{e}^H \mathbf{R} \mathbf{e}$	(1)
--------------------------------------	-----

The spectral covariance matrix $\mathbf{R}$ represents the power of waves incident upon the array, encoding the details of relative phase and amplitude at a single frequency across array elements as measured in a given analysis window, usually selected to encompass a single seismic phase (as in Figure 41). If $r_m(t)$ is the complex analytic form of the waveform observed on the m th channel of the array and filtered into a narrow band about the single analysis frequency, then a single element of $\mathbf{R}$ is obtained by:

$R_{mn} = \int_t^{t+T} r_m(\tau) r_n^*(\tau) d\tau$	(2)
---	-----

In this definition, t denotes the start time of the window, T its duration, and $*$ complex conjugate. In practice, several successive overlapping windows may be employed to generate “snapshots” of the covariance element, which then are averaged to enhance numerical stability (and to mitigate errors in automatic window placement).

The steering vector $\mathbf{e}$ encodes the characteristics of propagation - speed, direction, etc. – of a model for the waves incident upon the array. In the case of classical FK analysis, the m th element of this vector is given by:

$e_m = c_m e^{i2\pi f \mathbf{s} \cdot \mathbf{x}_m}$	(3)
---	-----

The model for propagation is a plane wave; $\mathbf{s}$ is the slowness vector describing the direction and speed of the plane wave, f is the analysis frequency, $\mathbf{x}_m$ is a vector describing the location of the m th channel sensor in space, and, for three-component arrays, c_m is a direction cosine describing the orientation of the sensor component. The map of power incident on the array is obtained from (1), scanning the steering vector over a two-dimensional domain of the slowness vector $\mathbf{s}$ surrounding the origin.

This is the point at which MFP and FK analysis differ. In MFP, the steering vector is a propagation template which may be other than a model for a plane wave. In model-based MFP, $\mathbf{e}$ is obtained as the solution of the Helmholtz equation, i.e. the time-independent (single frequency) wave equation governing propagation through a model of the medium containing the sources and receivers of interest. Alternatively, in empirical MFP, $\mathbf{e}$ is obtained as a

measurement from observations of a master event or events. The template is obtained as the principal eigenvector of the covariance matrix for the event observations. The concept for calibration is to obtain a large number N of template events ideally covering the source region of interest uniformly. For each event, the target phase is windowed (Figure 41) after filtering the observed waveforms into a narrow band surrounding the frequency of interest. In our analysis, the target frequency was chosen to be 1/35 Hz, as surface waves were well observed for both target regions at this frequency. Then, the covariance matrices $\mathbf{R}_\nu$; $\nu = 1, \dots, N$ are computed for the collection of events. The eigenstructure is computed for each matrix:

$\mathbf{R}_\nu = \mathbf{E}_\nu \mathbf{\Lambda}_\nu \mathbf{E}_\nu^H$	(4)
---	-------

and the eigenvector $\mathbf{e}_\nu$ corresponding to the largest eigenvalue is extracted as the steering vector representing propagation from the source location of event ν to the array. If the wavefield is a plane wave with direction and speed given by their nominal values, then this process will reproduce the plane wave steering vector used by FK analysis. Otherwise, it will capture effects of scattering and unmodeled refraction not present in the nominal plane wave model and lead to better detection and estimation performance.

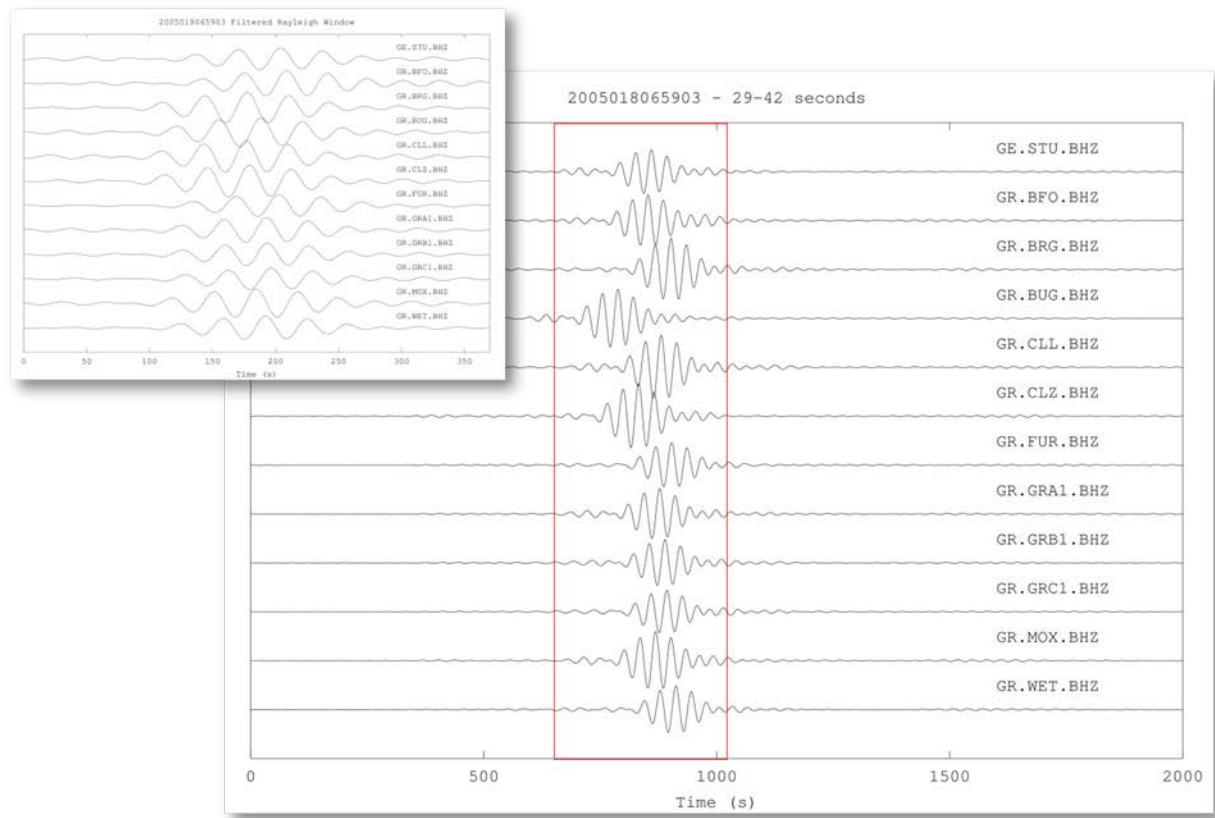


Figure 41: Preprocessing steps for estimating the matched field template at 35 seconds period from a Rayleigh wave window. The data were filtered to a narrow band about the target period, then windowed (~350 s window) after first removing a set of propagation delays for a reference source region location.

When using the GRSN as an array, the large aperture results in substantial moveout across the stations and a larger window than that required by a typical teleseismic array. This requirement can make it difficult to isolate individual phases for analysis or templating. In this circumstance, it is desirable to pick a reference location in the source region, calculate the propagation delay for the target phase to each station in the network using a reducing velocity, then consistently remove those station delays for every event in the source region. This shifting operation produces the “corrected” Rayleigh phase observations seen in the inset to Figure 41, using a source region reference location at 50.4N -28.9E and a reducing velocity of 3.75 km/sec. It can be shown readily that this correction results in no change to the value of equation (1), for a single phase in an infinite window, provided it is made consistently to the target covariance matrix and all the covariance matrices from which the steering template vectors are derived.

Now when applying MFP as an estimator of source location for a particular target event (as FK analysis is an estimator of wavefield slowness), the procedure is to scan equation (1) over all steering vectors $\mathbf{e}_\nu$; $\nu = 1, \dots, N$ and select the source location that maximizes the resulting estimated power. The power of (1) can be normalized by the total power incident on the array simply by dividing the quadratic form by the trace of $\mathbf{R}$:

$\frac{\mathbf{e}_\nu^H \mathbf{R} \mathbf{e}_\nu}{\text{tr}\{\mathbf{R}\}}$	(5)
--	-----

It is easily shown that this normalization produces a metric that ranges between 0 and 1, not unlike a correlation coefficient.

3 Results and Discussions

In this section we provide a brief overview of the results from each of the studies described in the previous section.

3.1 Anatomy of a Matched Field Detector

We have described in detail a procedure for detecting signals on seismic arrays based only upon characteristic sets of phase shifts between sensors. We used as an example the M=7.6 8 October 2005 Kashmir earthquake and attempted to detect aftershocks on the KKAR array in Kazakhstan reliably using a matched field template obtained from the first arrival from the main shock. Figure 42 displays 3 different detection statistics: the continuous AR-AIC trace (Gibbons et al., 2016), a theoretical matched field detector, and the empirical matched field detector.

The AR-AIC detection trace has the great advantage that the arrival times indicated are exceptionally accurate. The local maxima of this trace can be demonstrated to be very close to analyst onset time estimations. The disadvantage of this detection trace is that it will respond to any arrival with a significant increase in energy and/or frequency content, regardless of which direction it comes from. (It will also generate detections in the presence of noise sources.) In Figure 42 we see that we detect clearly the Sn arrivals from the larger Kashmir aftershocks in addition to the Pn phases. Selection of a detection threshold for this trace is also a complicated issue given a very non-normal distribution of the points in the time-series.

The detection traces from the matched field templates – both theoretical and empirical – detect only Pn arrivals from the required direction. The “footprint” of these detectors is related to the array response function and we would need to run an empirical study to determine how close to the associated slowness a foreign signal would need to be to be detected confidently, but falsely, by this procedure. A number of the Pn phases detected are right down at the noise level in the optimal beam, but the events are confirmed by recordings on closer stations.

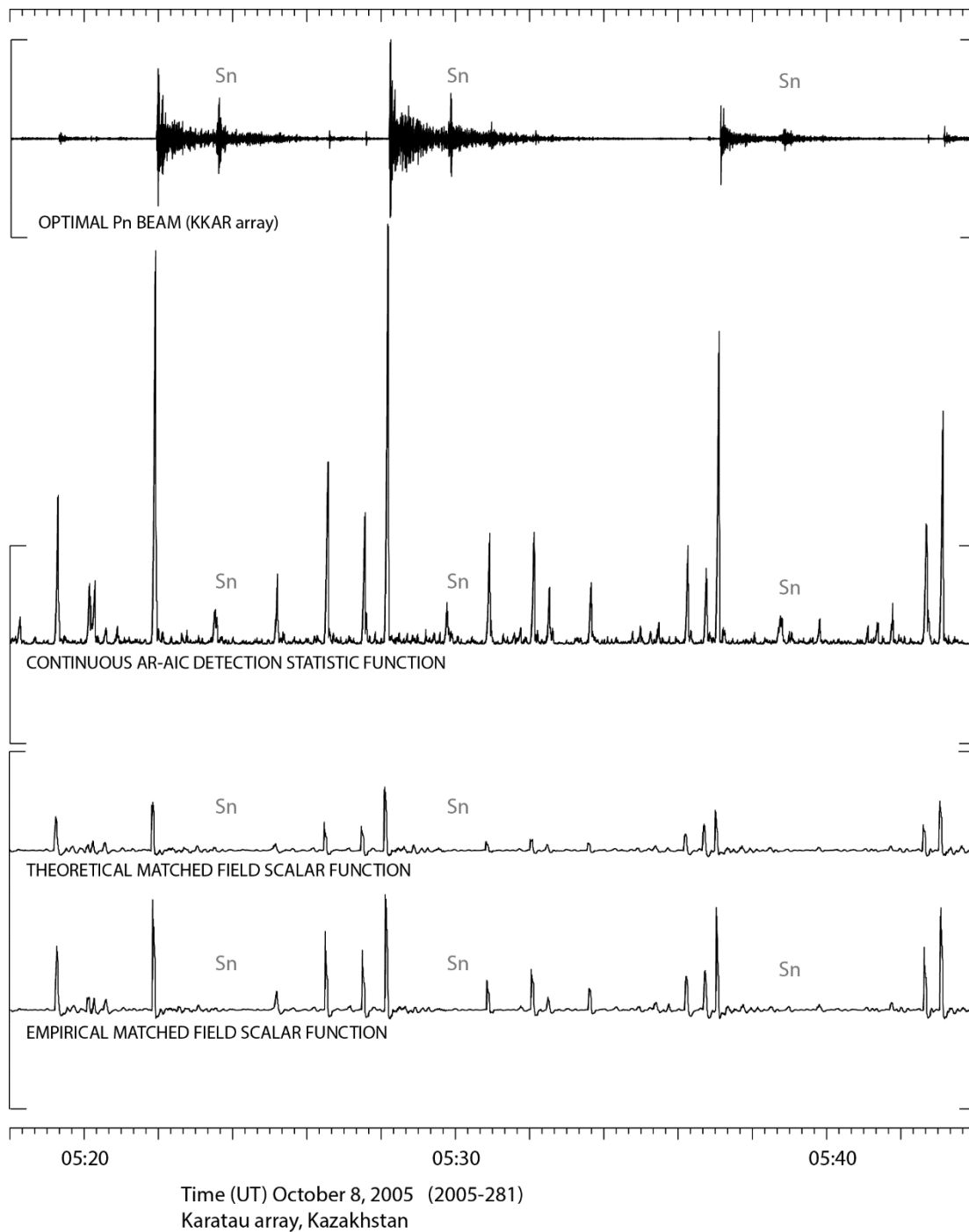


Figure 42: Scalar detection statistics for a segment of data from the KKAR array attempting to detect P-wave arrivals from earthquakes in the region surrounding the 8 October 2005 Kashmir main shock.

We note that the empirical matched field detector is significantly better than the theoretical matched field detector. This is easily understood in terms of the observations of the phase shifts of the wavefront at different frequencies. The theoretical detector assumes that all frequencies of the wavefront appear to cross the array with the same backazimuth and apparent velocity; we demonstrated that this is clearly not the case with a significant difference between the apparent wavefront at 2 Hz and the apparent wavefront at 4 Hz. These wavefield characteristics were common to several aftershocks of very different size and it is these differences that are being exploited by the more sensitive empirical matched field detector.

We observe several of the weaker Pn signals which are only just visible in the theoretical matched field detector trace but which are clear detections in the empirical matched field detector trace.

Even with relatively short time-windows over which the covariance matrices and matched field statistics are evaluated, the time-resolution of the matched field detection traces is limited. Waveform coherence is a property with a finite length (i.e. not instantaneous) and this is especially the case when coherent coda energy is observed with a similar apparent slowness to the initial arrival. We recommend a calculation of both the matched field detection trace and the continuous AR-AIC trace on an optimal beam for accurate onset time estimation.

3.2 Parameter Estimation on Seismic Arrays and Consequences for Matched Field Detectors

The question of the number of matched field beams required to cover seismicity in a given geographical region can be addressed by calculating EMFP statistics for the covariance matrix evaluated for a given arrival against empirical steering vectors calculated for vast numbers of seismic events covering the region in question.

Figure 43 displays matched field statistics evaluated for a given arrival at SPITS as a function of the location of the event used to generate the empirical steering vectors. The black line in Figure 43 indicates the backazimuth measured for the arrival using classical array processing for the SPITS Pn arrival and the red line indicates the direction towards the geographical locations of those events whose empirical steering vectors generated the highest matched field statistics. Given the Bayesloc location estimate of the event in question, we see that the empirical backazimuth estimate provides a far more accurate estimate than the plane-wave estimate generated from the theoretical steering vectors.

We note however that, in the direction indicated, there are two “crossings” of the mid-Atlantic ridge – one at a distance of about 300 km and one at a distance of about 900 km. The matched field statistics obtained from the single Pn phase, in this direction, are unable to differentiate between the two regions. This is to say that, without additional information, we are unable to map this arrival directly in geographical space.

Figure 44 shows the equivalent plot for the Pn phase at the ARCES seismic array from the same event evaluated against empirical steering vectors calculated for the initial ARCES Pn arrivals for the same set of events. Again, we land far closer to the correct direction than the backazimuth estimated from classical f-k analysis at the array. What is demonstrated both of these figures is that we avoid the bias in backazimuth obtained when using classical f-k analysis with a plane wavefront assumption (the black line) and that the empirical matched field statistic varies quite smoothly as a function of geographical source location. However, in this plot, we have evaluated over 1000 matched field statistics, many of which are almost identical. We wish to use EMFP as a detection algorithm and so we need to devise a manageable “beam deployment” which covers the parameter space adequately but without excessive duplication.

SPITS matched field statistic

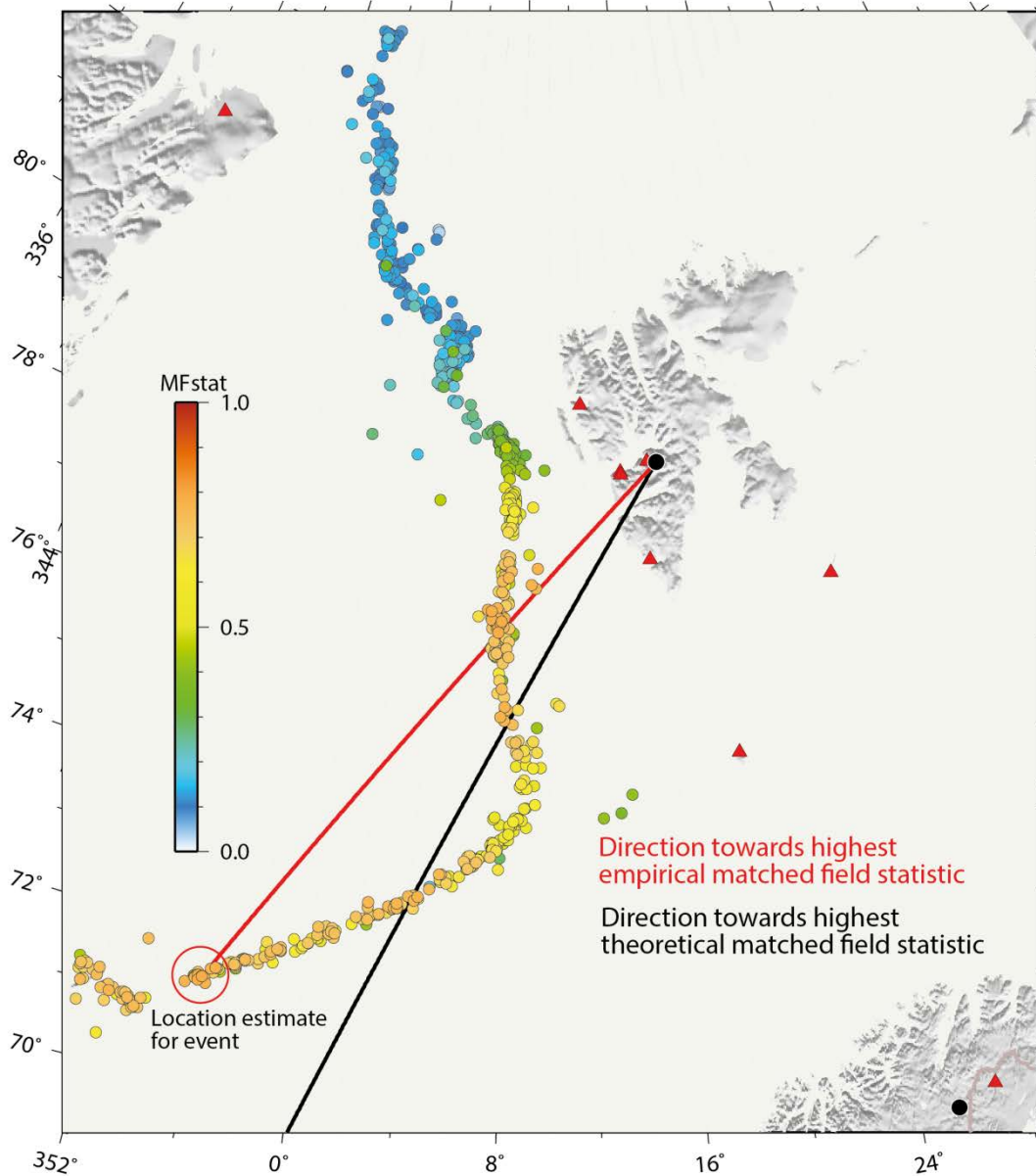


Figure 43: Geographical matched field direction estimates on the SPITS array. Matched field statistic evaluated on the SPITS array at a time 2015-276:01.48.27.425 as a function of the location of the template events together with the best estimate of the true location of the event at 2015-276:01.46.24. The backazimuth as indicated by the best theoretical (plane wave) matched field statistic is displayed.

ARCES matched field statistic

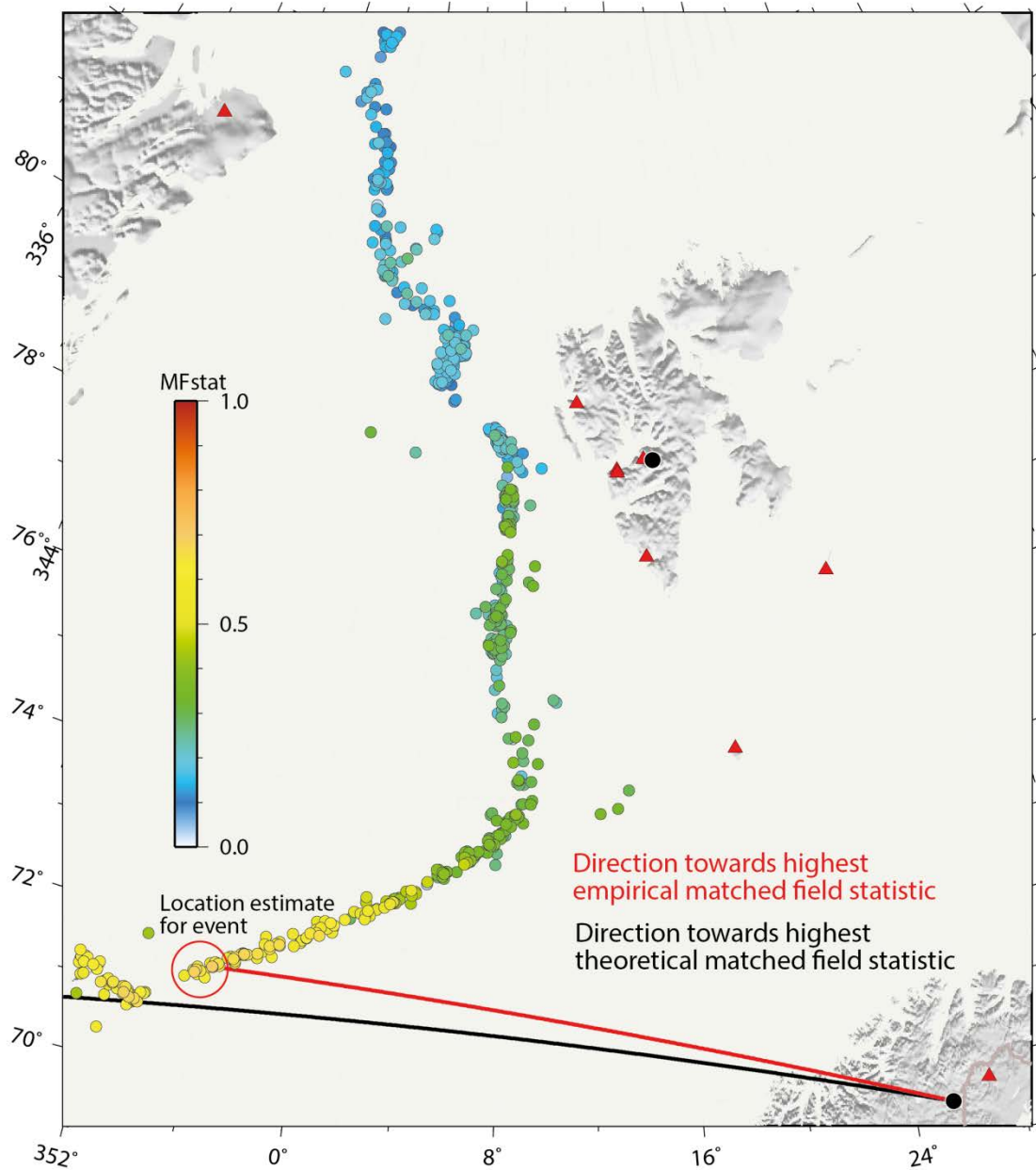


Figure 44: Geographical matched field direction estimates on the ARCES array. Matched field statistic evaluated on the ARCES array at a time 2015-276:01.48.44 as a function of the location of the template events together with the best estimate of the true location of the event at 2015-276:01.46.24. The backazimuth as indicated by the best theoretical (plane wave) matched field statistic is displayed.

Figure 45 (panel a) displays the azimuth residual calculated using classical f-k analysis as a function of the epicentral backazimuth for each of the events (this is to say how much in error the black line of Figure 43 is from the great circle backazimuth). Figure 45 (panel b) displays the residual calculated using the empirical matched field statistics (this is to say how much in error the red line of Figure 43 is from the great circle backazimuth). A backazimuth estimated for a given arrival at SPITS using the empirical steering vectors is almost certain to be within 10 degrees of the true backazimuth and, in the majority of cases, within 5 degrees. The spread in the residuals for a given source region is quite similar for the plane-wave (panel a) and empirical (panel b) calculations, again as we might expect given that the same window length and frequency band are used for both calculations. However, the empirical calculations eliminate the bias.

In figures 43 and Figure 45 we have simply taken a single backazimuth estimate, corresponding to the location estimate for the event with the best match. As we see in Figure 43, there is often little to differentiate between the values of the matched field statistics obtained using different steering vectors and a regionally averaged estimate may give a better direction indication. Considering a longer time-window or a different frequency band, focusing on higher frequencies, may differentiate more between the different source regions. Alternatively, empirical steering vectors calculated from ensemble covariance matrices calculated for many events in a concentrated region Harris and Kværna (2010) may give a more robust estimate. The wide spread in the estimates is very much a consequence of the very small aperture of the SPITS array, with the resulting limitations in slowness resolution; a larger aperture array is likely to provide a higher resolution.

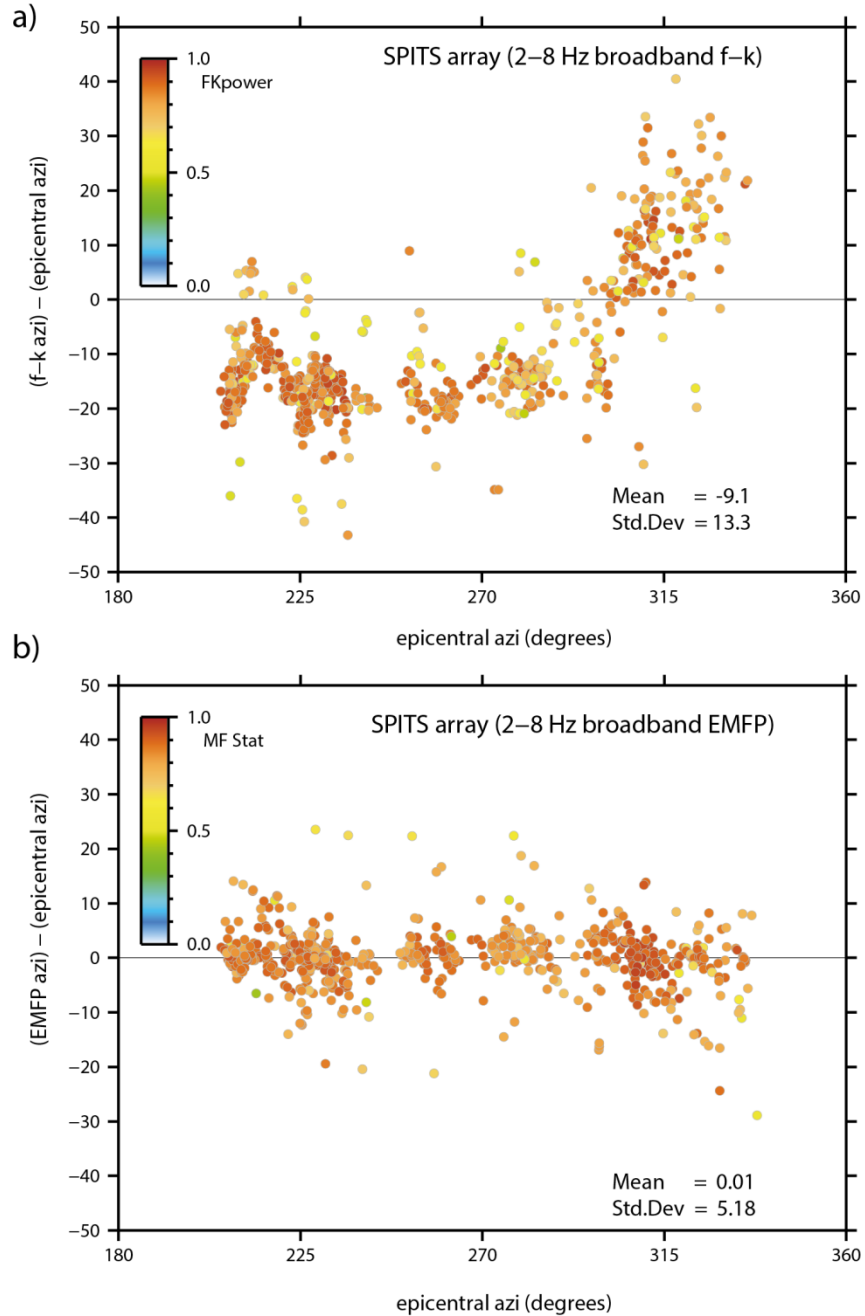


Figure 45: Backazimuth estimates for first P-arrivals at SPITS from the Bayesloc-located events displayed in Figure 22 b) using (a) classical f-k analysis and (b) a matched field comparison with the templates evaluated from each of the other events. Panel (a) is almost identical to Figure 26 except that the colour of each symbol indicates the relative power rather than the apparent velocity. In panel (b), for a given event, the matched field statistic for the first P-arrival has been calculated against the P-arrival for every other event (as displayed for a single event in Figure 43 the epicentral azimuth of the event with the best match is used to calculate the residual).

The matched field slowness estimate is exactly that; we do not map uniquely into geographical space with this single phase. The great circle path towards the true location of the event crosses the plate boundary twice and, from the EMFP statistic alone, we cannot differentiate between the two candidate source regions. It is worth noting that all contributions to the observed phase shifts between signals on the different sensors are corrected for implicitly by the EMFP estimates. This includes elevation differences, such that no assumptions about the local wavespeed at the array need to be made.

We wish to divide the seismicity into zones over which we can calculate “ensemble covariance matrices” (see Harris and Kværna, 2010). This is to say that we want to find the most representative and stable estimates possible of the phase and amplitude relationships between the narrowband signals on the different sensors of the SPITS array for events in a given source region. A covariance estimate taken from a single event may be biased due to low signal-to-noise ratio or, for example, a timing anomaly. If we can take the covariance matrices for many very closely spaced events and perform a careful averaging procedure for the phase-differences between each pair of sensors, we will hopefully be able to avoid any corruption of the ensemble covariance matrix for that region resulting from a data anomaly.

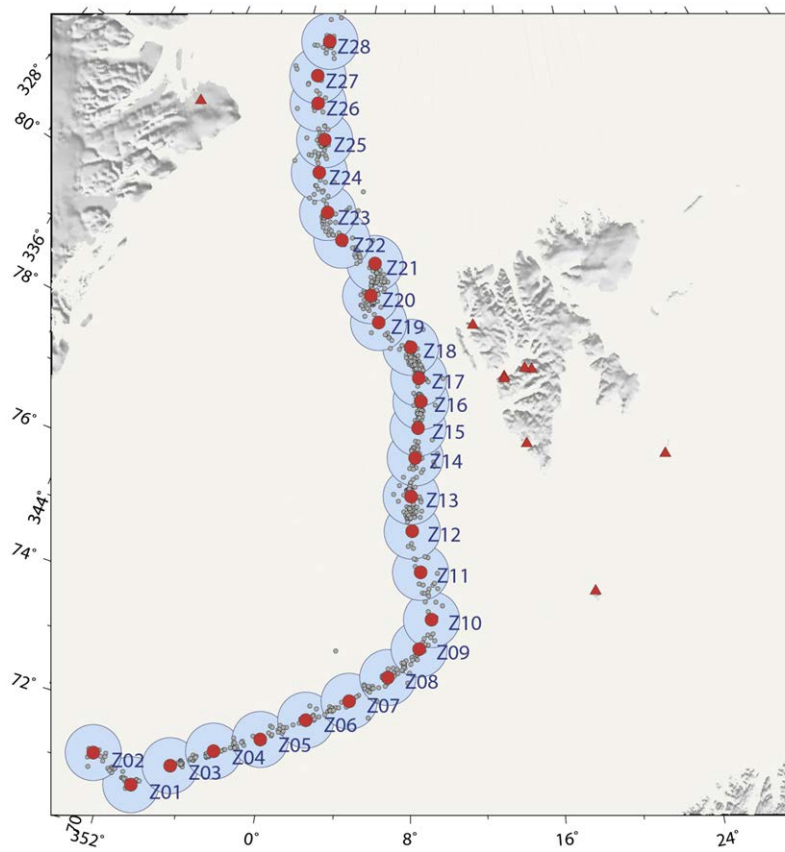


Figure 46: Provisional zone map for events along the ridge between 71 and 84 degrees north. The zones have radius of 70 km and overlap.

Figure 46 displays a trial zonation of the seismicity along the mid-Atlantic ridge. The spacing was chosen based upon considerations of geographical distance between events.

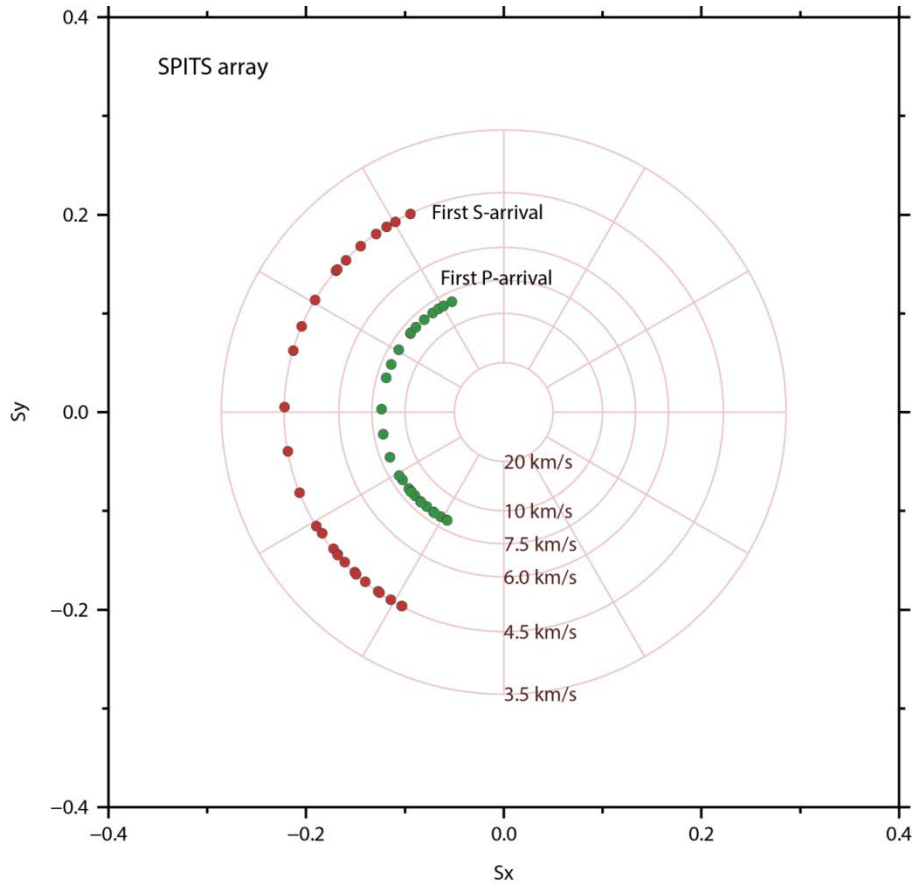


Figure 47: Beam deployment in slowness space. For each of the zones displayed in Figure 46, the theoretical slowness of the first P and the first S arrival is calculated for the central point.

One of the questions that we wish to address is the extent to which we should organize our signal classification by the source region, as opposed to the classical slowness and backazimuth parameter space which is usually used to classify incoming arrivals on seismic arrays. The study designed here should be able to test whether or not there is an advantage in processing primarily using empirical templates classified in geographical space, and then returning the corresponding theoretical slowness and backazimuth parameters (if needed) to the phase association algorithm. For example, Zones Z03 and Z13 in Figure 46 have essentially the same backazimuth and slowness, despite being in two very distinct geographical regions. Scattering of the wavefield along the path may result in quite distinct patterns in the phase relationships which allow for a resolution greater than the theoretical capability of the array (c.f. Harris and Kværna, 2010).

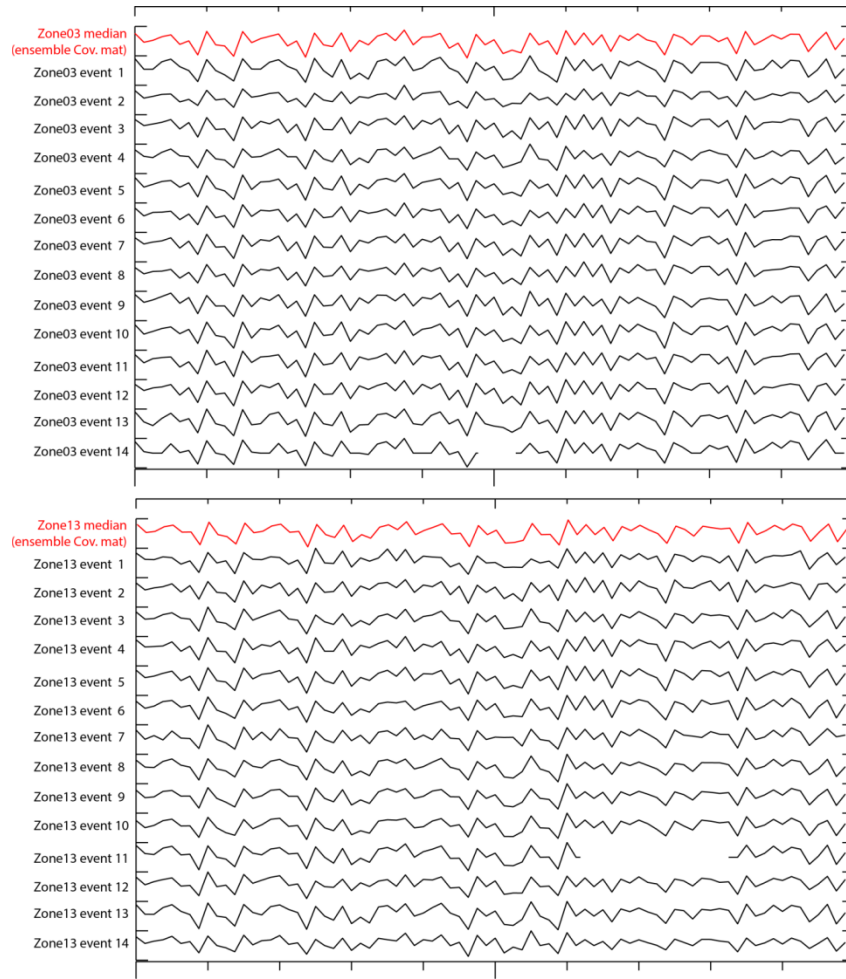


Figure 48: The construction of ensemble covariance matrices for first P-arrivals at SPITS for events from two of the clusters, obtained by calculating median values of each real and each imaginary value over all events within one of the given zones. Zones Z03 and Z13 are displayed which have a very similar backazimuth from the SPITS array. The gaps are zeros resulting from an absence of data from specific channels at a given time and have little effect on the median ensemble averages.

The covariance matrices themselves are calculated by taking a short time-window of the array data surrounding the initial P-arrival and calculating the coherence and phase shift between each pair of channels using the multitaper coherence routines provided by Prieto et al. (2009). This provides, for an N-channel datastream, for every one of the NB narrow frequency bands, $N*(N-1)/2$ complex phase shifts. So while each covariance matrix consists of $NB*N*(N-1)/2$ complex numbers, we display in Figure 48 a short segment of the corresponding $NB*N*(N-1)$ real numbers simply arranged sequentially. There is no physical interpretation of the lines displayed in Figure 25 other than that the relative heights of the various real and imaginary parts should share characteristics from event to event. By constructing the ensemble covariance

matrices by taking medians of each element for each event in the cluster, we obtain an apparently stable estimate which is not greatly influenced by outliers in the population.

The SPITS array was upgraded in 2004 from being a vertical-only array (with a single 3-component instrument) to having 6 out of 9 sites equipped with 3-component instruments (Gibbons et al., 2011). For the current study, we compiled a set of template events which included events prior to 2004 and so, for simplicity, only the 9 vertical sensors have initially been used. Gibbons et al. (2011) demonstrated that the horizontal components not only had far greater amplitudes than the vertical components for regional S-phases, but also that the coherence between sensors was often far greater for the horizontal ground motion. A final matched field beam deployment for the SPITS array will inevitably have to exploit both vertical and horizontal ground motion for optimal detection of all phases of interest, and a detection recipe will need to take into account that the different sets of channels may be required for the optimal detection of different types of signals.

For each of the zones of seismicity displayed in Figure 46, we calculate empirical steering vectors as the principal eigenvectors of the ensemble covariance matrices calculated. Figure 49 displays the matched field statistics calculated as a function of time and frequency for six of these first P-arrival empirical templates, evaluated over a short time-window (slightly in excess of 5 minutes) on August 9, 2015. This data segment includes the signal from a magnitude 3.4 event from the North West along the spreading ridge. These six steering vectors constitute part of the empirical “beam deployment” displayed in Figures 46 and 47 with the color of each pixel indicating the degree to which the pattern of phase-shifts at the time and frequency indicated match the phase-shifts specified in the empirical steering vector. As we would hope, at the time of the P-arrival from our signal, a strong matching is observed across all of the narrow frequency bands for a number of our matched field detector traces. The values obtained are highest for the empirical vectors from zones 21 to 24 which all correspond to a similar backazimuth from the SPITS array (~325 degrees). As the direction of the events from which the empirical steering vectors are calculated deviates from this backazimuth, the match between the steering vector phase shifts and the phase shifts in the incoming data diminishes.

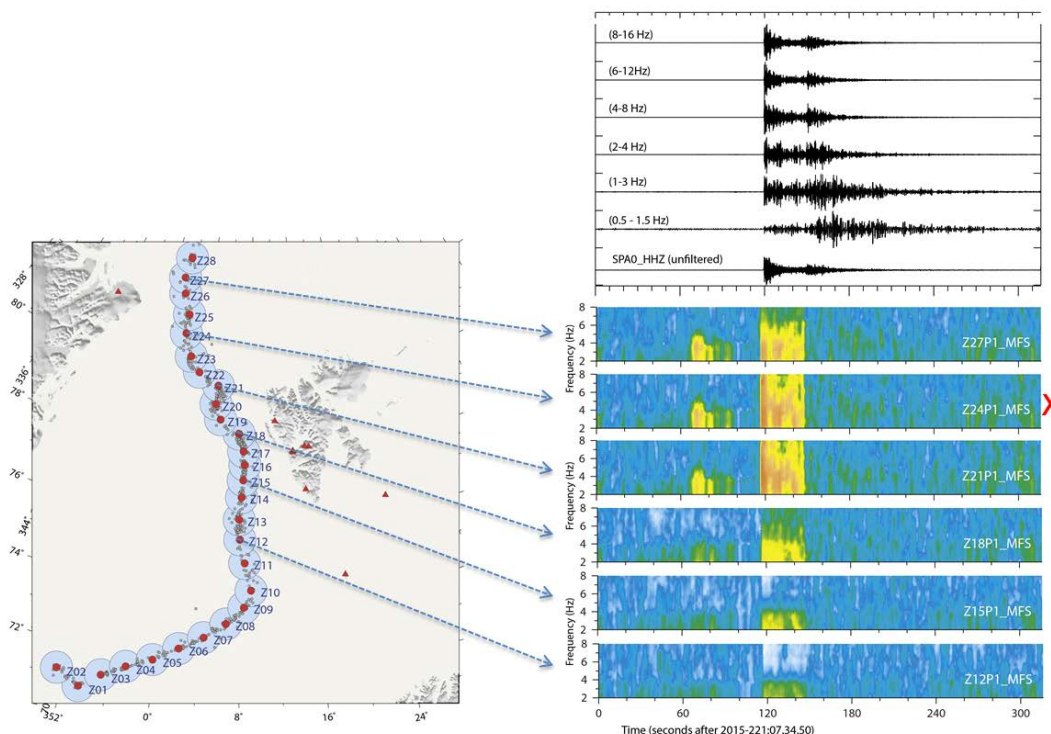


Figure 49: Time-frequency displays of matched field statistics for “beams” based upon the empirical steering vectors calibrated for first P-arrivals for ensembles for the zones of seismicity indicated. The waveform for the channel SPA0_HHZ is displayed for reference along the same time-scale bandpass filtered in the bands as indicated. Note that approximately 50 seconds before the main signal, there is significant peak in the matched field statistics, especially at lower frequencies. This is a low amplitude signal with a similar backazimuth and apparent velocity.

The challenge of controlling operational detectors based upon these principles is to be able to declare detections based upon the information contained in such time-frequency plots, together with absolute measures of the waveform amplitudes. As discussed by Gibbons et al. (2011), the detection of events from this source region at the SPITS array has always been challenging since the S-phases have typically been quite incoherent on the vertical components and f-k analysis always results in typical P-wave velocities, making it very difficult for phase association algorithms to interpret the detections correctly. This is illustrated in Figure 50.

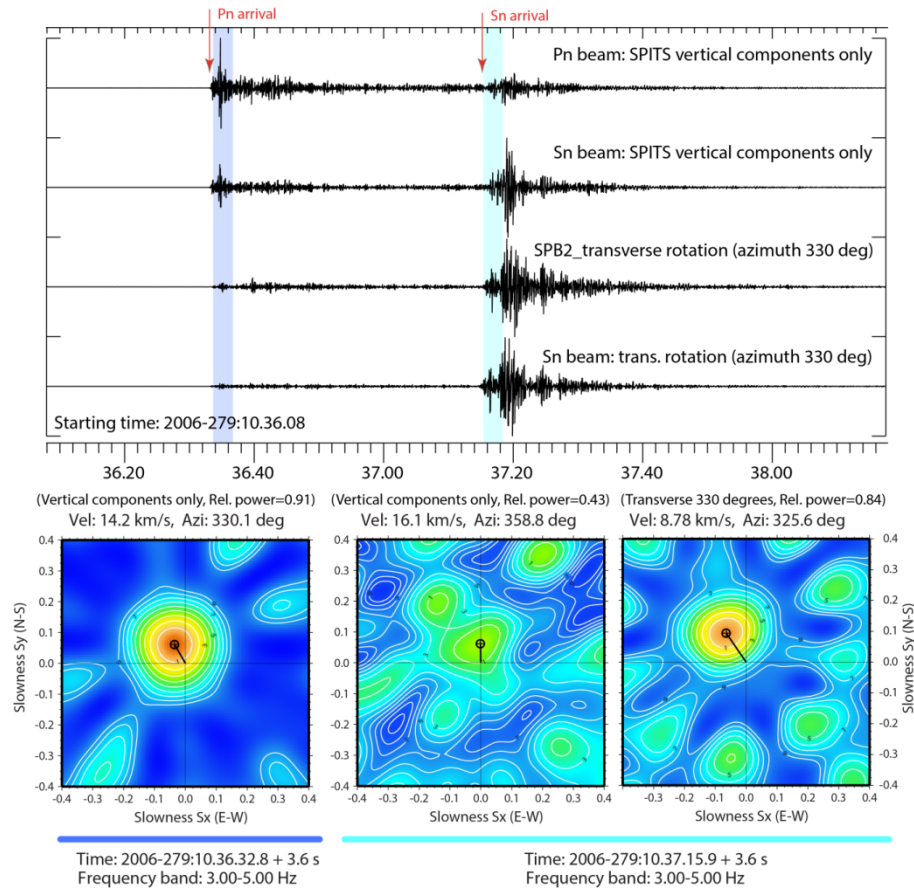


Figure 50: Plane-wave f - k analysis for Pn and Sn phases for an event on the ridge to the North West of the SPITS array. The slowness estimate for the Pn phase is evaluated only on the vertical traces whereas the slowness estimate for the Sn phase is evaluated both on the vertical and horizontal components. The coherence of the Sn phase on the horizontal traces is far greater than that for the Sn phase on the vertical traces. (Reproduced from Gibbons et al., 2011).

The time-frequency plots in Figure 49 show that throughout the P-coda of the regional event displayed, there is a good match between the phase structure of the coda wavetrain and the first P-arrival template. On the arrival of the Sn phase, this phase structure completely disappears and we return to the “background level” of these “P-beam” detection statistics. The corresponding plots for the Sn matched field beams do not show a convincing correspondence with the template at the time of the Sn arrival. This result is a disappointment given an expectation that representative matched field templates could be calculated for all phases from ensemble covariance matrices evaluated over a broad region. It works well for the P-arrival but poorly for the S-arrival on the vertical sensors for events in this source region. It is to be expected from Figure 50 that an S-phase detector would work far better on the horizontal sensors given the improved signal coherence.

The phase structure of coda waves on the SPITS array for regional events was studied by Schissole and Schweitzer (2004). They demonstrated that for some paths (for example from Novaya Zemlya to Spitsbergen) the coda waves were highly coherent and indicated the correct backazimuth and characteristic P- and S-wave velocities, whereas for other paths (for example from Mohns Ridge to Spitsbergen) there was little stability in the coda backazimuth estimates and that little evidence of S-wave apparent velocities could be detected. The work of Schissole and Schweitzer (2004) has significant implications for the likely effectiveness of a time-frequency matched field detector beam recipe. It is quite likely that a full wavetrain analysis, rather than a first-arrival-only detector is likely to provide the information necessary to be able to make a signal classification. The sudden disappearance of the P-wave-consistent coda wave energy, combined with an increase in energy and decrease in coherence on the vertical traces, could in fact be exploited as a wavefield characteristic used to identify the most likely source of a signal.

Schissole and Schweitzer (2004) were limited to the vertical component array, but we can now extend our analysis to the 3-component subarray at SPITS. Figure 51 shows f-k analysis of the coda waves for the October 11, 2010, M = 4.4 Novaya Zemlya earthquake on three different components at SPITS. Unlike the ridge event example displayed in Figure 50, the S-wave energy is highly coherent on the vertical sensors. The coherence between the horizontal sensors is even greater following the Sn arrival, but it is also worthy of note that the coherent energy in the P-wave coda for this event is present on all components - but propagating with different apparent velocities.

In summary, we have made significant progress in evaluating a matched field frontline pipeline detector for the SPITS array. We have been able to calibrate empirical steering vectors for the detection of initial P-wave arrivals from events on the ridge. It is likely that for robust detection and identification of signals – in particular S-phases - we ought to consider all components and examine the full wavetrain characteristics and not just the initial signal onset.

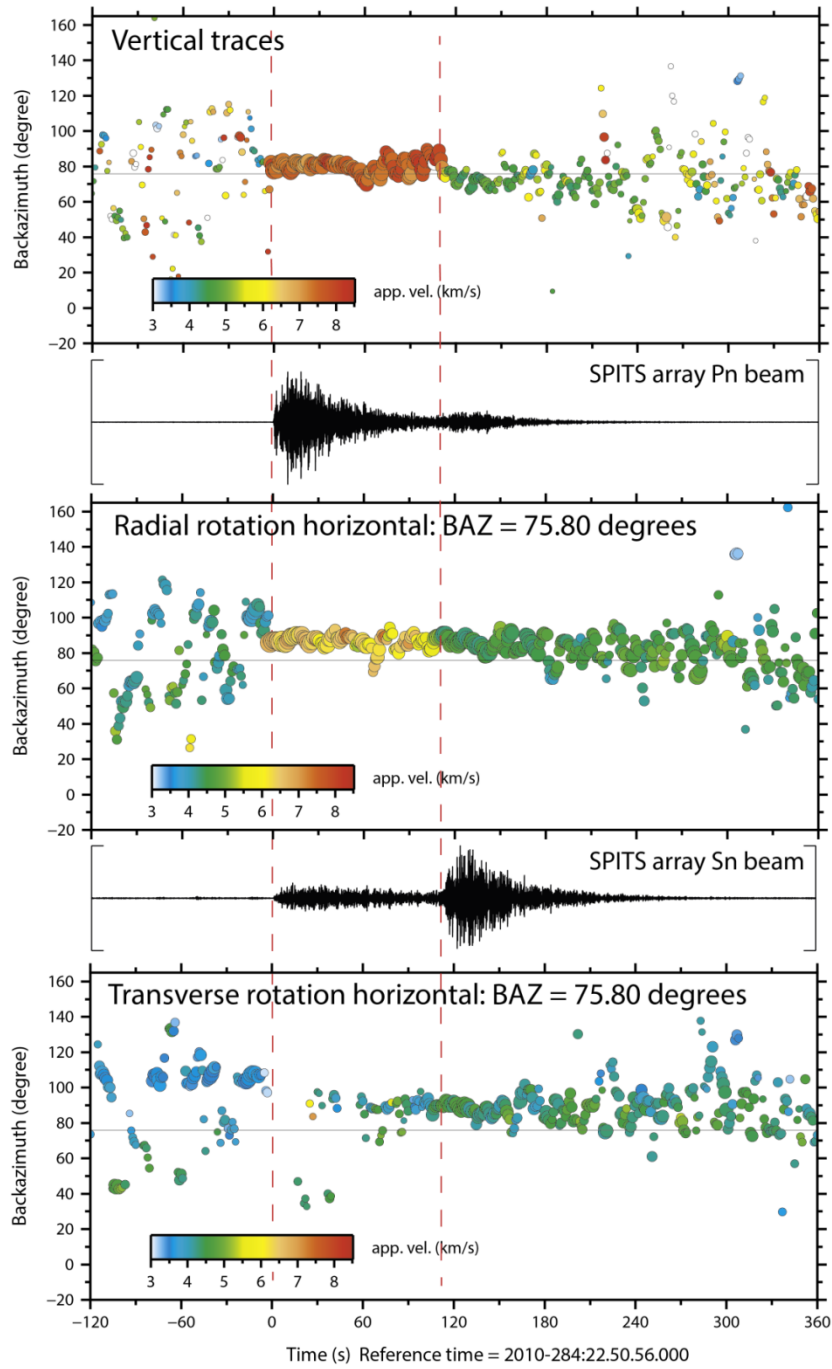


Figure 51: Evaluation of slowness vectors on the vertical, radial and transverse components at the SPITS array for the October 11, 2010, Novaya Zemlya earthquake. An evaluation is made every second using a 4-second long time-window with a 2-8 Hz broad frequency band used. The size of the symbols is proportional to the relative power in the f-k beam (essentially a measure of coherence).

Empirical steering vectors were calculated for each arrival present in the NORSAR waveform archive for all of the events displayed in Figure 17 and, for a number of phase arrivals, we evaluated the correspondence with the complete set of empirical steering vectors. An example, for an earthquake south of Japan, is displayed for the three arrays in figures 29, 30, and 31.

The results for the USRK array, for an earthquake recorded on January 2, 2009, are displayed in Figure 52. The covariance matrix was evaluated over the nine vertical channels of the array for a window with length 256 samples (6.4 seconds) and frequency bands between 1.09 Hz and 5.00 Hz were considered. As for the SPITS examples described in the previous reporting period, the highest values of the matched field statistics are close to the great circle path linking the event hypocenter and the station. As we move away from this line, the statistic diminishes indicating a decrease in the similarity in the observed phase shifts. There is clearly also a distance effect with little similarity observed between the phase shifts observed for the closest events to the station and the phase shifts observed in the signal from the January 2 event. One explanation for this is likely to be that closest events are very deep, such that the arriving wavefronts are likely to reach the array at a very steep angle of inclination.

In Figure 53 we see how the analyst-measured backazimuth and slowness estimates for the USRK observations vary with those inferred from a nearest-match empirical matched field estimate. In the vast majority of cases, the preferred backazimuth estimate falls within 5 degrees of the great-circle backazimuth to the final solution location. The comparison is flawed in that it is clear that the analyst-measured solutions have been artificially clipped within an error of ± 5 degrees. We see that there is significant variability in the measured apparent velocity which is not an issue when we base our parameter estimates on the projection into geographical space.

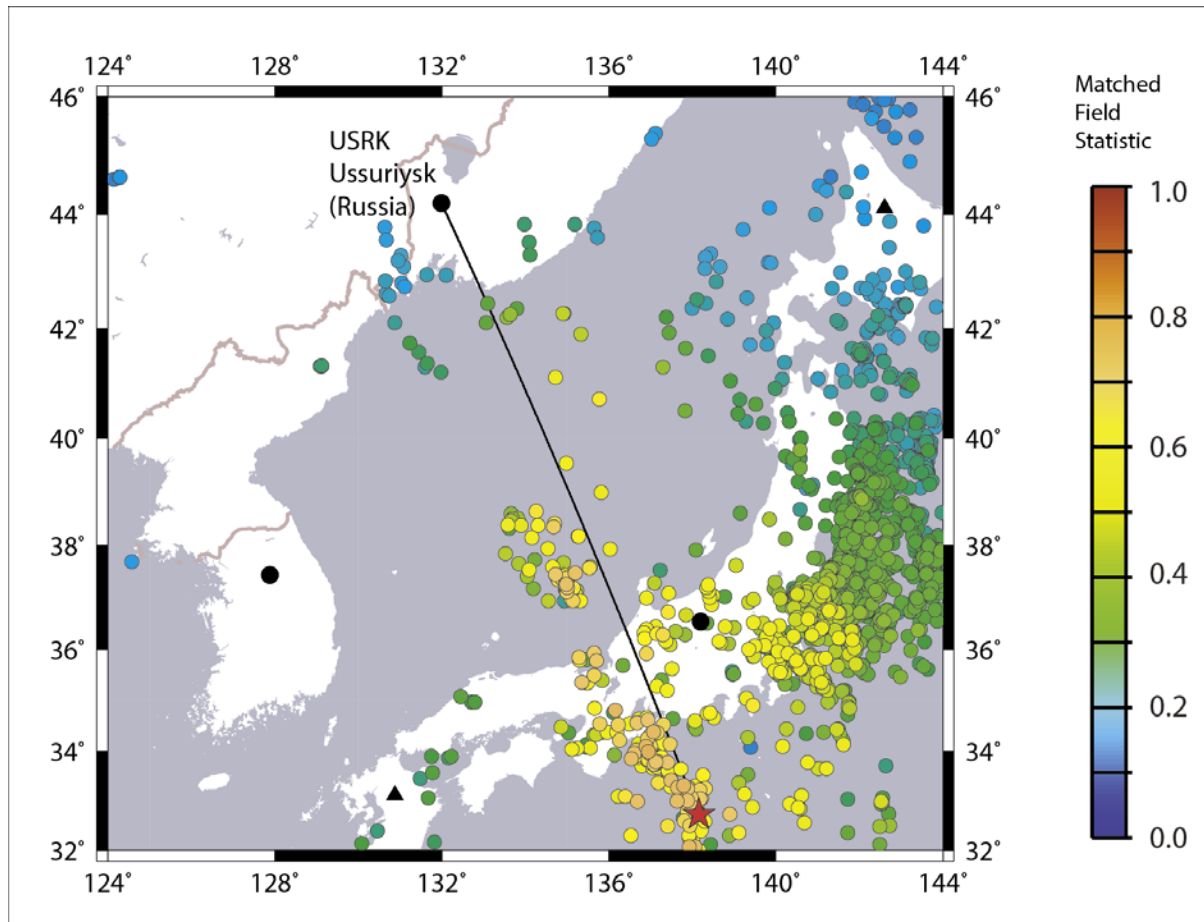


Figure 52: Geographical matched field direction estimates on the USRK array. Empirical matched field statistics calculated between the covariance matrix evaluated for the regional P arrival at the USRK array at a time 2009-002:07.45.38.850 and the steering vectors calculated from the covariance matrices evaluated for the first P arrival at USRK for each of the events indicated. The red star denotes the location of the earthquake with origin time 2009-002:07.42.51.34 and epicenter (33.72, 138.15).

The corresponding image for the KSRS array is displayed in Figure 54. The matched field statistics evaluated for this larger aperture array are significantly lower in absolute terms and only exceed the background level in a rather narrow range of backazimuth from the array. It is noted however that the “background level” for this 10km plus aperture array is significantly lower than for the far smaller aperture USRK array. So, while the colors on figures 52 and 54 indicate a weaker correspondence for KSRS than for USRK, the significance of a given value is greater for the larger aperture array.

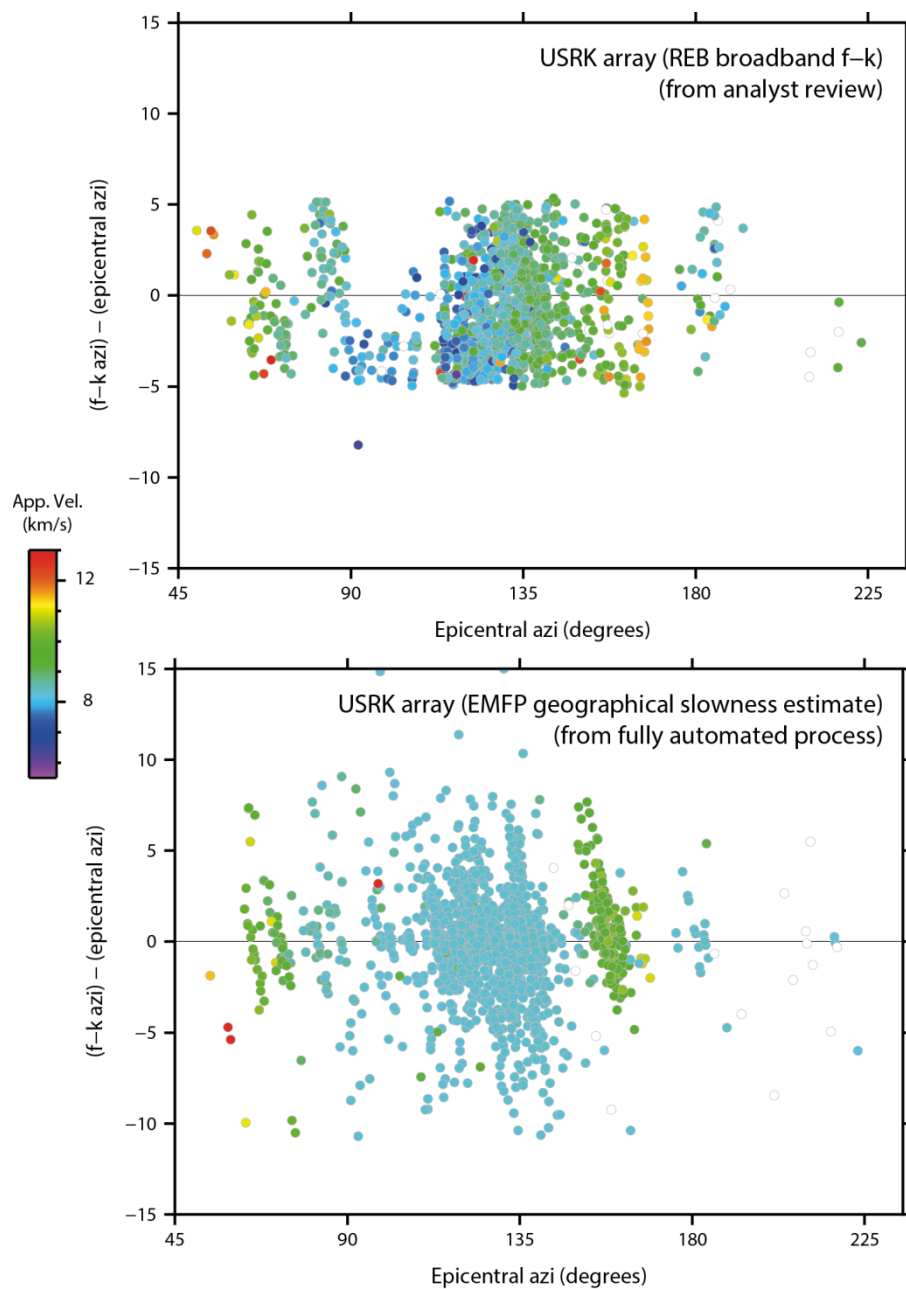


Figure 53: Backazimuth and apparent velocity measured on the USRK array for events displayed in Figure 52, both using analyst-measured estimates (as reported in the Reviewed Event Bulletin of the IDC) and using the predicted values from the best-fit matched field statistics.

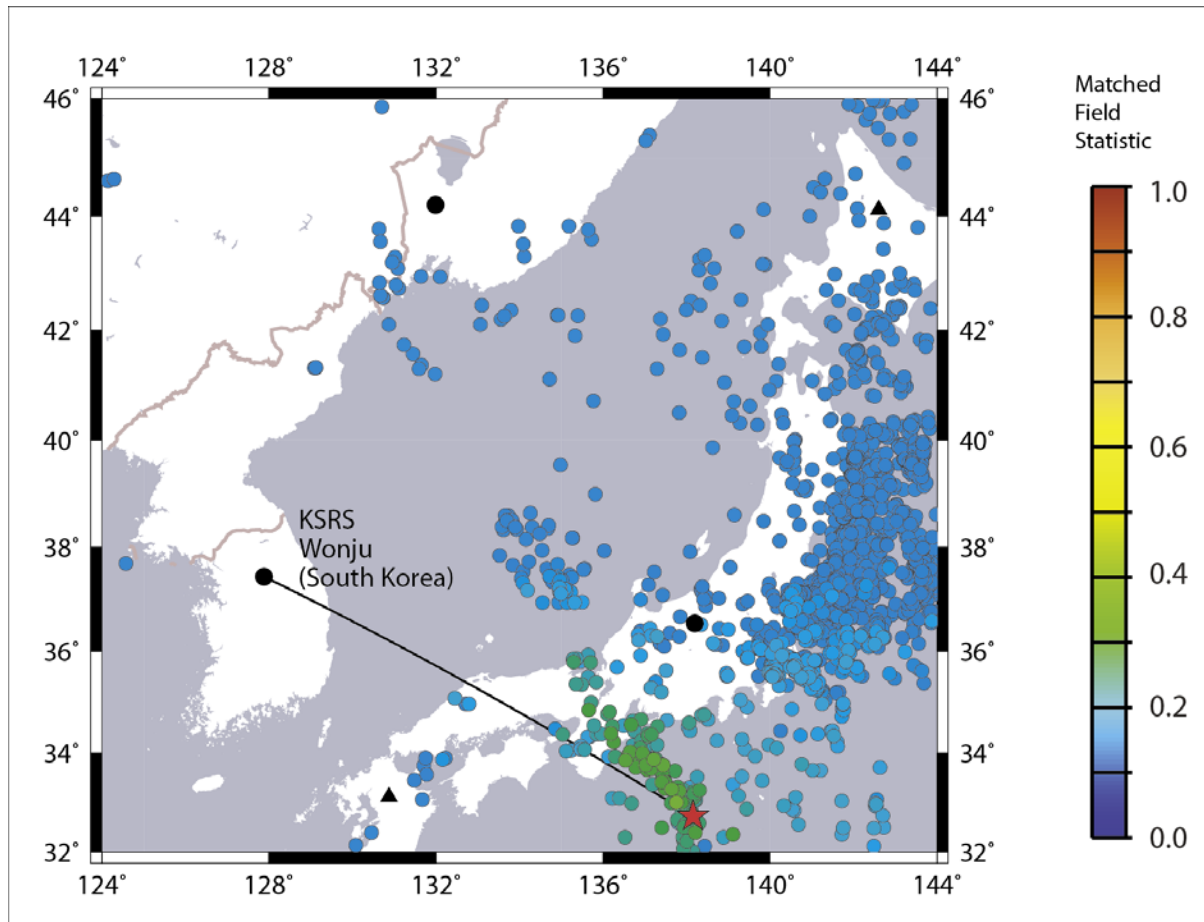


Figure 54: Geographical matched field direction estimates on the KSRS array. Empirical matched field statistics calculated between the covariance matrix evaluated for the regional P arrival at the KSRS array at a time 2009-002:07.45.07.350 and the steering vectors calculated from the covariance matrices evaluated for the first P arrival at KSRS for each of the events indicated. The red star denotes the location of the earthquake with origin time 2009-002:07.42.51.34 and epicenter (33.72, 138.15).

It comes as little surprise that the geographical response of the MJAR array is even narrower, with values of the matched field statistic only exceeding the background level in a very limited region surrounding the estimated event epicenter. Observing many plots similar to those displayed in figures 52, 54, and 55 we conclude that the empirical matched field statistics do indeed have a footprint which maps back the geographical source region with a pattern which is consistent with expectations. With a number of observing phases which are associated correctly, it is likely that a good preliminary estimate of the source region can be obtained from a numerical evaluation of matched field statistics even without explicit reference to arrival times. This alludes to a possible algorithm for evaluating source hypotheses.

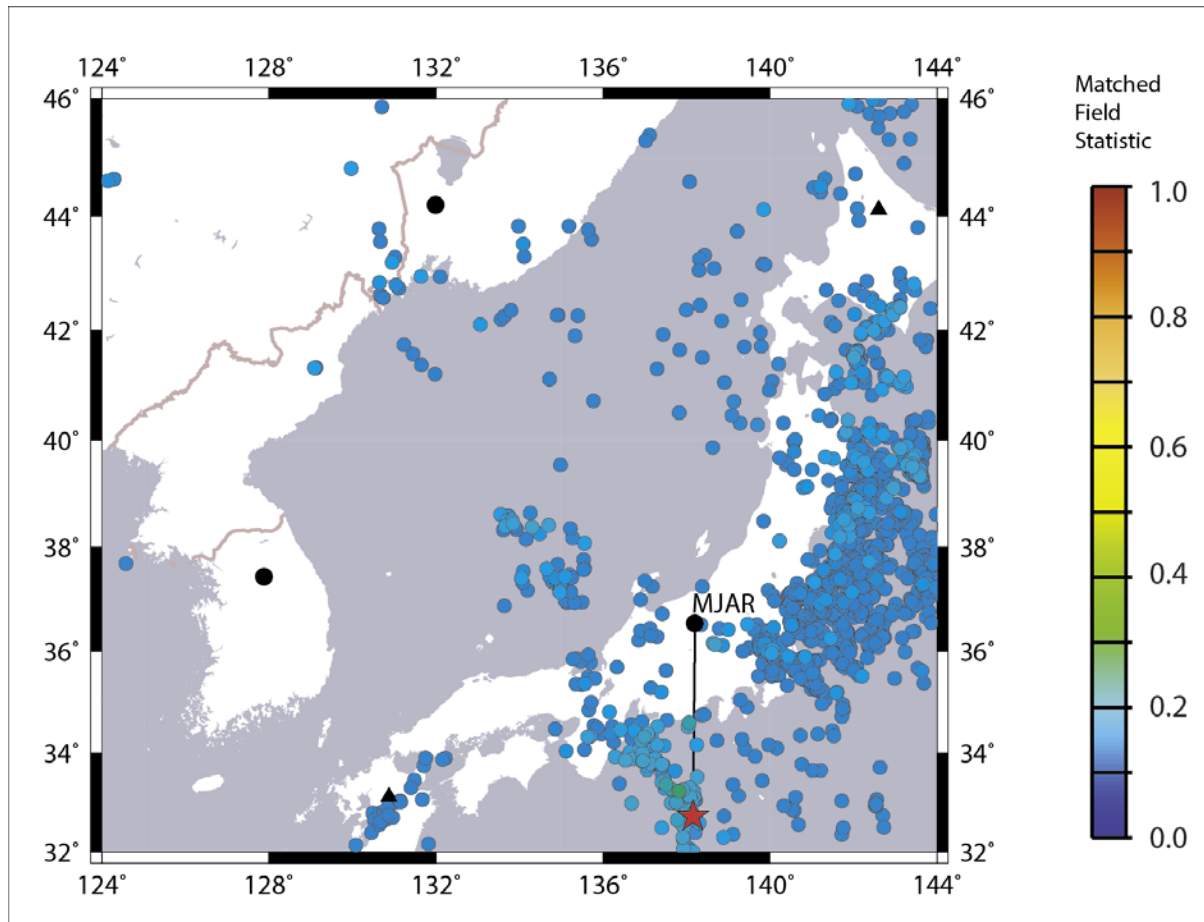


Figure 55: Geographical matched field direction estimates on the MJAR array. Empirical matched field statistics calculated between the covariance matrix evaluated for the regional P arrival at the MJAR array at a time 2009-002:07.43.56.675 and the steering vectors calculated from the covariance matrices evaluated for the first P arrival at MJAR for each of the events indicated. The red star denotes the location of the earthquake with origin time 2009-002:07.42.51.34 and epicenter (33.72, 138.15).

We have highlighted the specific dependence of the geographical resolution of the matched field statistic on the array aperture and coherence characteristics of the relevant seismic waves on the sensors. This is to say that, like in classical beamforming, a small aperture array (such as USRK or even more so SPITS) will be a relatively blunt instrument in location estimation with a high response over a wide parameter space. A large or sparse array (such as MJAR), for signals in a similar frequency band, will act more as a spotlight or pattern detector and have a highly source-specific response. This will require a far denser deployment of matched field beams to be able to cover the appropriate parameter space.

3.3 Limited Sensor Pair Correlation

There are complications associated with the estimation of apparent velocity and backazimuth for seismic arrivals on array stations. The problems fall into two categories; those resulting from incoherence (waveform dissimilarity) and those resulting from aberration (deviation from predicted relative delay times). In practice, these issues combine such that diverse custom-designed techniques are often required to estimate signals for which the receiver arrays are not optimal. For a detected phase, we advocate Limited Sensor Pair Correlation (LSPC) for improving the robustness of estimates and the ease with which an analyst can interpret maps of array-gain over slowness space. Considering only sensor-pairs with small geographical separations can mitigate both incoherence and aberration; incoherence since the waveform similarity is less diminished over short distances, and aberration because the deviation from the predicted time-delay will be far smaller for closely spaced sensors, such that the timing anomaly may be small relative to the dominant period of the signal.

Slowness estimates obtained by stacking cross-correlation functions between different channels have been performed for many years and the consideration of correlations only between closely-spaced sensors is familiar to practitioners of Progressive Multichannel Cross Correlation or PMCC (Cansi 1995). However, there are distinct advantages when analyzing seismic arrivals in examining the array gain over the full region of slowness space. Correlation beamforming (Ruigrok et al. 2017) provides us with a technique of investigating the direction of incoming wavefronts based on waveform similarity between sensors. In this work we have focused on calculating fully-normalized cross-correlation traces between sites and constructing a virtual "correlation co-array" over which we can apply standard array processing methods. It is instructive to do so since we can often see in these correlation "waveforms" how the coherence breaks down and apply these observations to decide the choice of sensor-pairs to be used. The other advantage of this strategy is that these correlation-coefficient traces are trivial to calculate and so the procedure can readily be implemented in existing processing pipelines without needing to code new implementations of the actual estimation routines.

We emphasize that a large number of different algorithms are available for direction estimation of incoming wavefronts on arrays and point out that performing each algorithm both on the original waveforms and on the correlation co-array may provide insight and inform future decisions about how the estimation of such signals is attempted. It is the estimation of short period signals over medium and large aperture arrays for which the improvement in parameter estimation is likely to be most significant. However, we have also demonstrated that the estimation of regional secondary phases on small aperture seismic arrays can benefit from the consideration of only the more closely-spaced sensors (Figure 56).

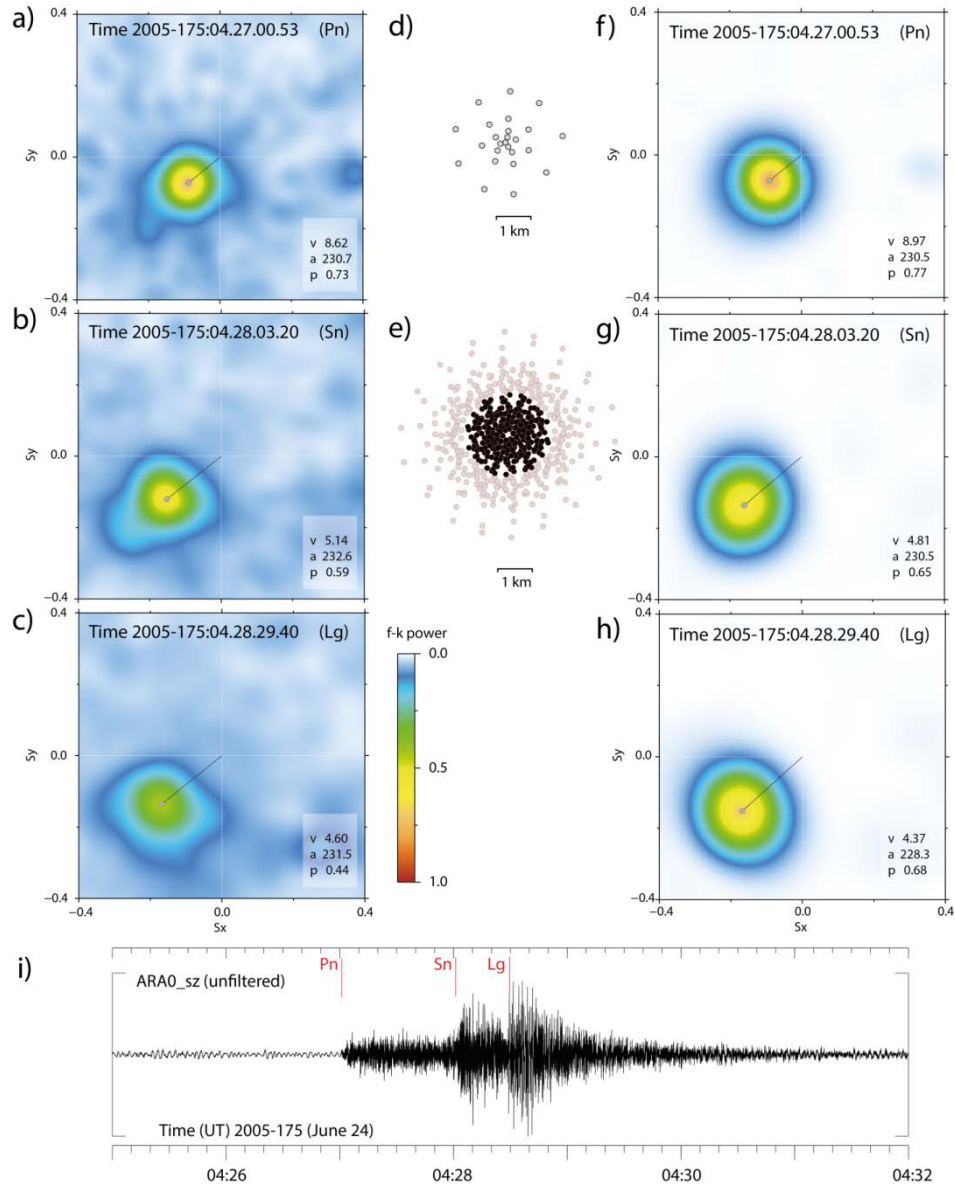


Figure 56: Direction estimates at ARCES for the Pn, Sn, and Lg phases from the largest June 24, 2005, earthquake discussed by using 2-5 Hz filtered short-period vertical waveforms. Panels (a), (b) and (c) show the classical f-k spectra for Pn, Sn, and Lg respectively for the full array (displayed in panel d). Panels (f), (g) and (h) show the corresponding estimates using the f-k analysis on the correlation co-array with sensor-separation limited to 1.2 km (the shaded symbols in panel e). For each panel, v , a , and p denote v_{app} (km/s), backazimuth (degrees), and relative power respectively. Panel (i) displays the unfiltered vertical waveform from the central site of the array.

LSPC gives added weight to waveform similarity where present and diminishes the influence of higher amplitude scattered energy which is not coherent between sensors. Although we focus on inter-sensor distance as the main criterion for omitting sensor-pairs, different choices may be applicable as determined by empirical studies.

We clarify also that information is discarded by the limitation of the set of sensor-pairs chosen. There may be valuable phase and amplitude information in the correlations between the more distant sensors which may provide high resolution information on the source and/or path (c.f. Harris & Kvaerna 2010). We have illustrated how static corrections for individual sites on an array can improve the stability and accuracy of slowness estimates; the disadvantage of this approach is the lack of continuity of observations as a function of slowness space. The effect of LSPC is to make a seismic array appear to have a smaller aperture. The loss of theoretical slowness resolution may be offset by an improvement in the stability of slowness estimates applicable to the scale (e.g. global) of monitoring being performed. The IMS network of seismic arrays is highly heterogeneous with great differences in array aperture and corresponding slowness resolution and estimation capability. LSPC may allow the network to be treated as more uniform.

3.4 Case Study: The North Korea Nuclear Test-Site

Matched field beams were set up for the USRK and KSRS seismic arrays to detect signals from possible nuclear events at the Punggye-ri test-site in North Korea. These are in continuous operation and employ templates constructed from numerous events which have been observed at or close to the test-site. A small event was detected on both arrays on October 12, 2017, and this event provided an opportunity to evaluate the performance of the different matched field detectors.

10 Pn templates were available on the KSRS array, including signals from all 6 declared underground nuclear tests, 3 additional lower magnitude events, and the theoretical plane wavefront template. For the USRK array, only 9 were available since the first test on October 8, 2006, was not recorded by this station.

The 6 declared explosions took place in the mountains at Punggye-ri and it can be demonstrated that the correlation properties of the waveforms from the different events depend not only on the relative location of the sources but also on the locations of the stations (Gibbons et al., 2017). This is almost certainly due to near-source physics, with topographical effects and the associated reflected energy being the most likely candidate for the differences. This is one of the motivations behind using empirical matched field processing as a tool for detection; changes in the differences in phase shifts across the aperture of a remote array may be less severe than the changes in the temporal form of the signal. (Note that there is no guarantee that this will be the case. Harris and Kværna, 2010, demonstrated that relatively small changes in source location were able to create differences in the phase structure of the Pn arrival over a distant small aperture array that allowed events from the different locations to be identified.)

All of the matched field templates were able to detect the signal from the October 12, 2017, event confidently. However, it is evident that some of the templates provided far more sensitive detectors than others (Figure 35) – the signals of unknown origin and the signal from the 2006 declared underground nuclear test provided far more sensitive templates than the plane-wave theoretical template or the templates from the more recent five nuclear tests. Inspired by the LSPC-work, elements of the covariance matrix corresponding to sensors separated by more than 4 km were set to zero prior to the extraction of empirical steering vectors (the principal eigenvectors). This appeared to have some effect of “levelling” of the performance of the different detectors.

In Figure 57 we display the empirical steering vectors, both from the full covariance matrix and from the covariance matrix with terms from more distant sensor pairs set to zero. We

demonstrate here that steering vectors are fundamentally the same; the small differences observed are predominantly at high frequencies.

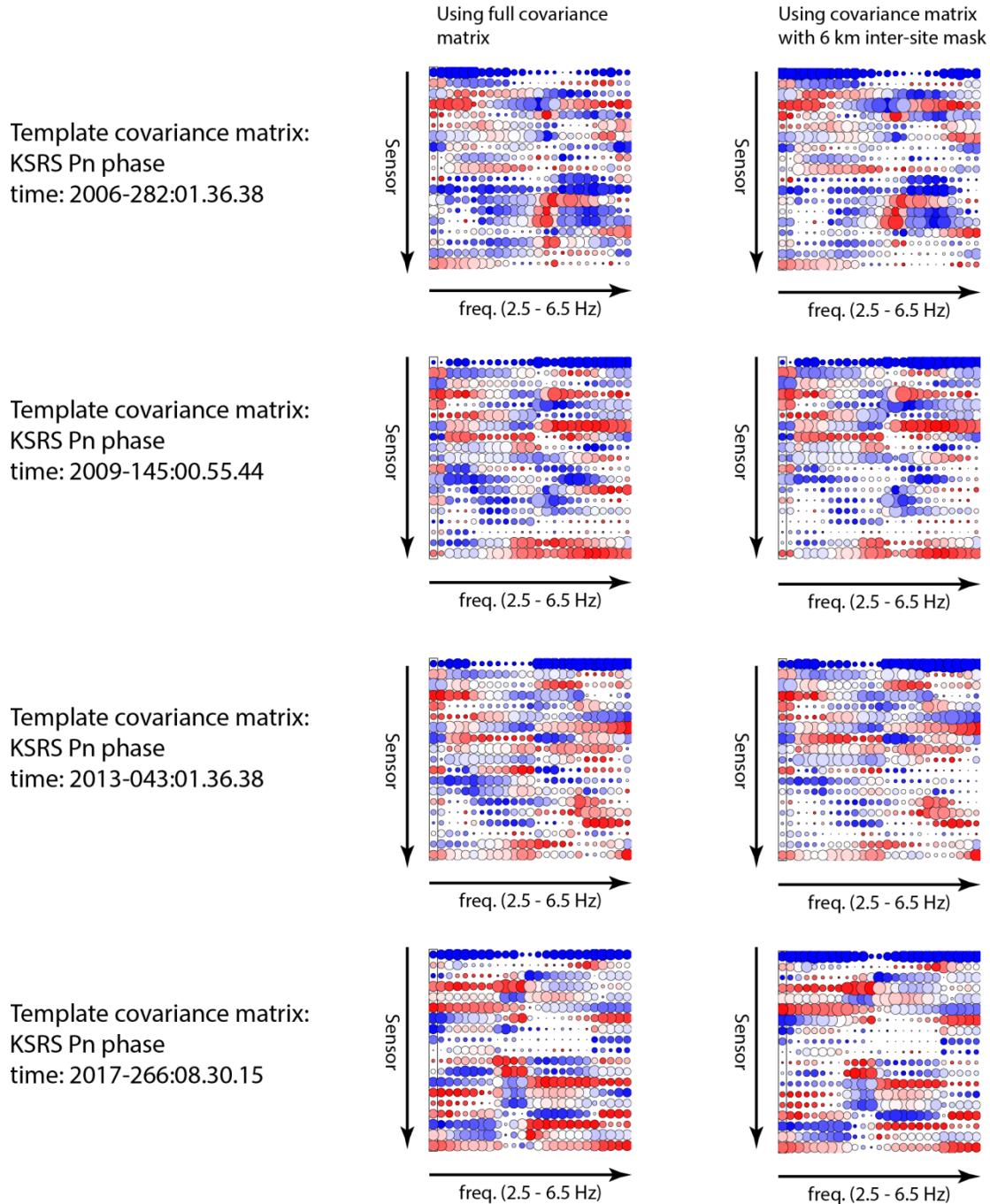


Figure 57: Vectors of empirical steering vectors for the Pn arrival on the KSRS array for four events at or close to the North Korea nuclear test site. For the raw data with 20 Hz sampling and 19 sensors, a window of 128 samples is used to calculate the covariance matrices and 25 narrow bands between 2.50 and 6.25 Hz are considered.

A more aggressive masking of the covariance matrix (i.e. setting a greater number of elements to zero) is likely to make a more marked difference to the steering vectors.

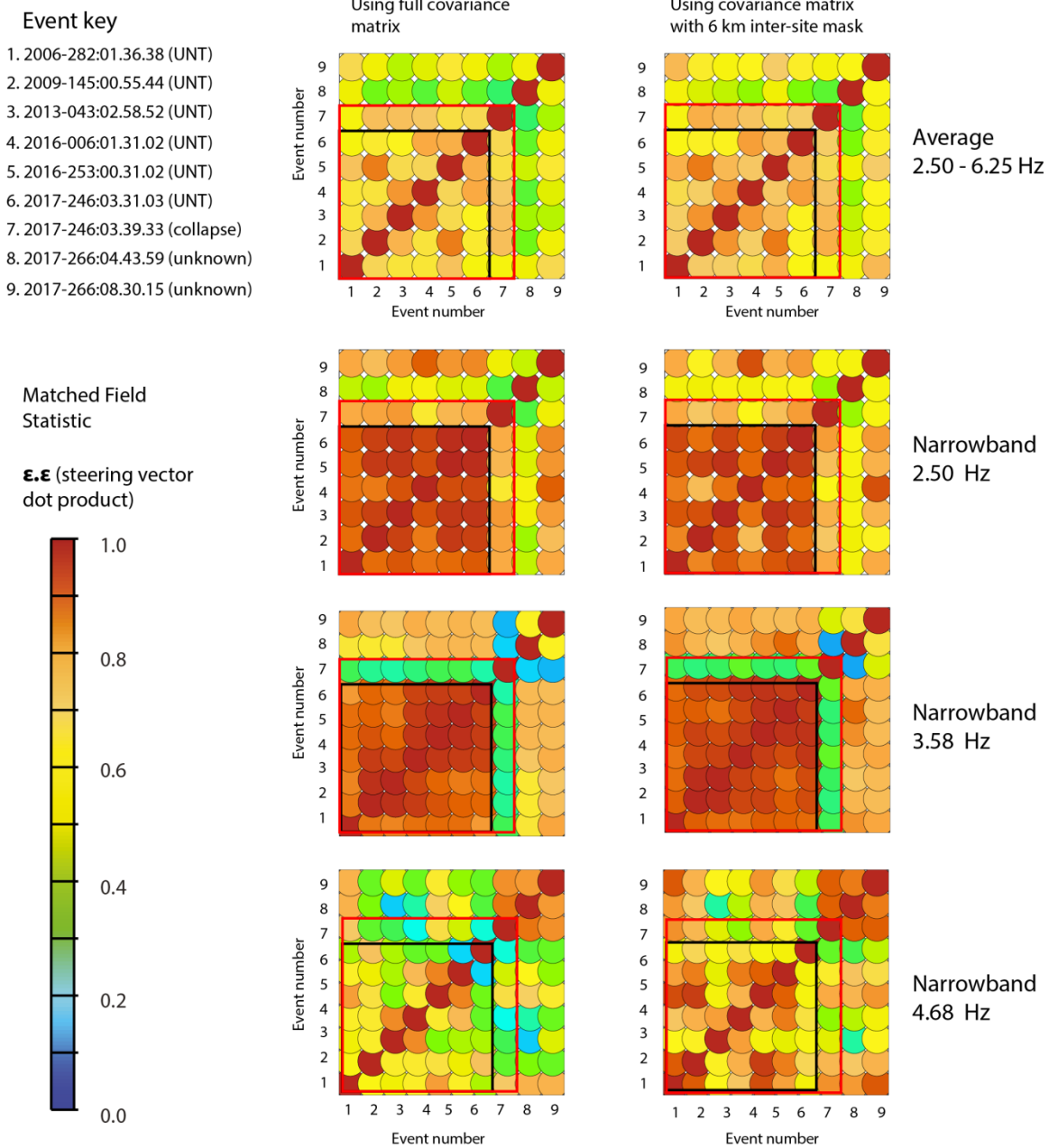


Figure 58. Matched field statistics for the North Korea events as listed for the KSRS array.

Figure 58 shows by means of a scalar product how similar the empirical steering vectors derived from the first Pn arrivals at KSRS are for the 9 events. This similarity metric is similar to those often provided by comparing correlation coefficients obtained between waveform segments from different events. We note that the behavior is very different for the different frequencies

and that this will be a significant function of the SNR at this frequency in addition to the intrinsic signal properties. For this reason, the frequency-averaged similarity statistics (in the top panels) is probably not very significant for the performance of the matched field detectors presented in Section 2.4. The pattern at 4.68 Hz (bottom row) is quite typical of the higher frequencies with significantly higher similarity between the empirical steering vectors for the unknown events (8 and 9) with the 2006 underground nuclear test (1) than for the other underground nuclear tests. There are many observations in Figure 58 which are not easy to explain, for example the great dissimilarity between the empirical steering vectors for the presumed collapse event (7) and all the other events at the middle frequencies. Many experiments were carried out with different length windows and the result is quite stable. We note that the signal from this event is in high coda energy from the preceding nuclear test and this may have greater effect on these frequencies than others.

The panels to the right are comparisons between the empirical steering vectors when the covariance matrices have zero terms for elements corresponding to channel pairs with sensors greater than 6 km apart. There is an overall slight increase in the similarity statistics between the events generally, particularly for the higher frequencies. With this treatment, we hope that the detector's footprint is increased somewhat such that it will detect confidently an event a slightly greater distance from the template event source location. Just as the LSPC processing described in the previous section increased the robustness and stability of the estimate at the expense of high resolution, we hope that this form for matched field processing will allow us to scan wider regions (albeit with slightly lower sensitivity than the full covariance matrix detectors would provide). Understanding the properties of such detectors will require careful study for individual arrays and sources, for example by comparing the predicted and empirical steering vectors (c.f. Figure 57) in order to find out at which points signal coherence breaks down, or changes, and whether special signal processing considerations would be able to optimize detection and parameter estimation.

3.5 The Physical Basis for Matched Field Processing

Mid-Atlantic ridge

We applied the procedure just described to observations from both source regions. Thirteen broadband stations were used in the array and all three components for each station, resulting in a theoretical 39 channels to calculate the metric of (5). Some modification had to be made for dead channels – in this circumstance the covariance matrix and templates were constructed using just the channels that were operating during both the target event and a particular template event. Consequently, the stations employed sometimes changed from one template event to the next, depending on station availability.

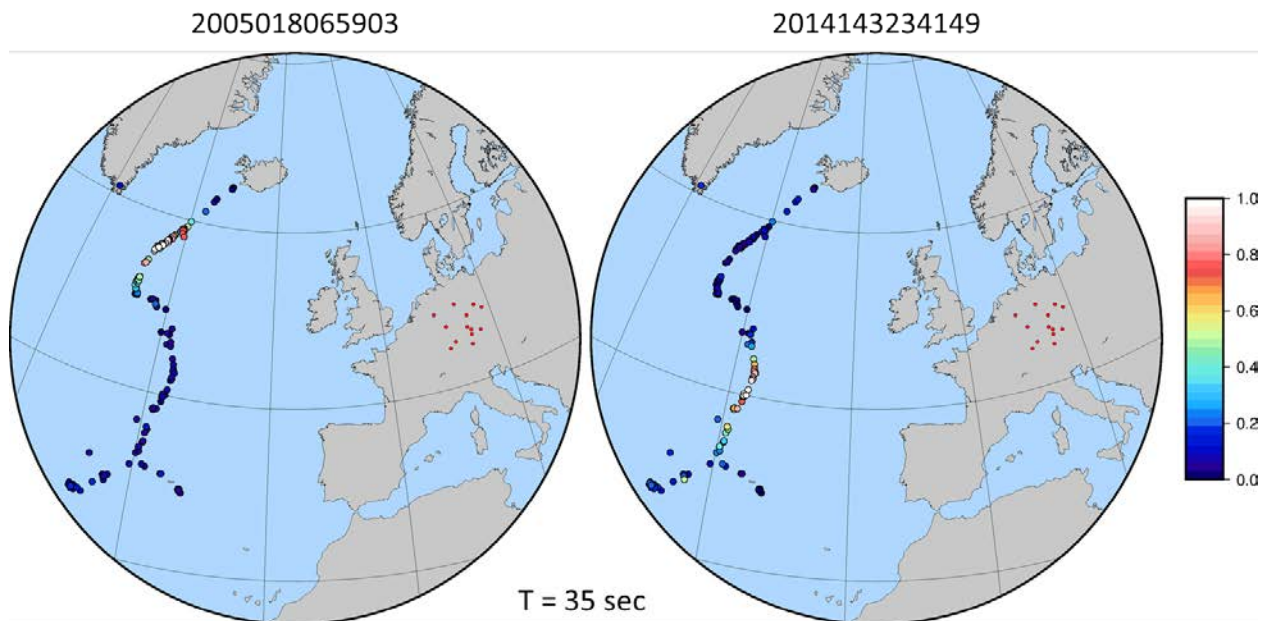


Figure 59: Empirical matched field processing response patterns for sources in regions of relatively homogeneous mechanism show well-behaved drop-off from a maximum at the source location. The grid spacing is 15 degrees.

Figure 59 shows the matched field response pattern for two events selected from stretches of the mid-Atlantic ridge dominated by events with fairly uniform mechanisms. The first was at 57.049N -33.811E, and the second, further to the south was at 45.099N -27.8369. As expected, the MF metric peaks at the correct source location and then declines more or less monotonically as distance increases along the ridge from the correct location. Since the trend of the ridge is approximately normal to the path to the array center, the response pattern is

approximately a cross section of the main lobe of the array response pattern. Resolution of source location is best in this direction and would be predicted approximately by simple plane wave considerations. Such considerations also predict that events distributed along radial lines from the array would have completely ambiguous response patterns. We'll see that when we get to the Zagros events.

In those regions of the ridge with more variation in source mechanism, the response patterns are not so well behaved. Figure 60 shows the response patterns for two events near transitions from spreading centers to transform faults. The first event is located at 52.666N -35.028E and the second at 53.1535N -35.1019E. These events are barely 50 kilometers apart, but have significantly different responses.

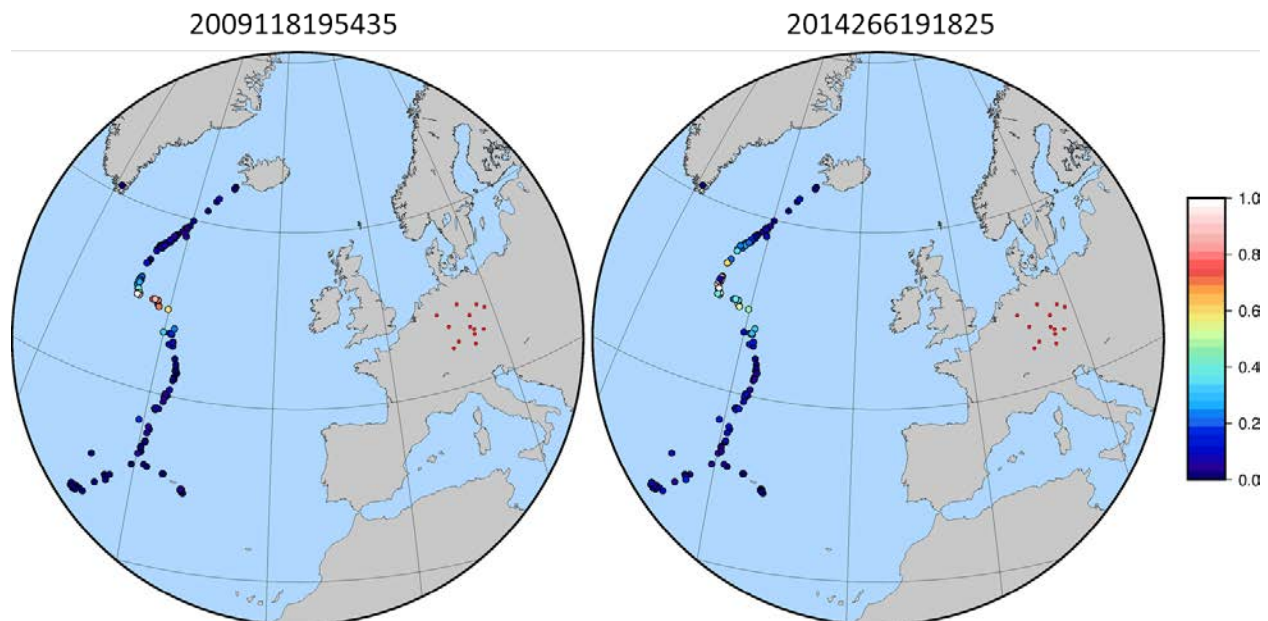


Figure 60 *In regions with events of more diverse mechanism, the EMFP response is not monotonic with distance from the source location. The grid spacing is 15 degrees.*

The differences are easier to appreciate in the more detailed Figure 61. At left is the response pattern for event 2009118195435 (location indicated by the white star), from its NEIC location and GCMT project mechanism, a transform event. Note that it has a response above 0.7 for other events along the same transform fault with similar GCMT mechanisms, and also values above about 0.6 for 3 events along a shorter transform fault farther to the south and east. All other responses are below about 0.4. We infer that the other events are spreading center events (confirmed in some locations with GCMT mechanisms) and that the pattern of high and low responses is a consequence of the differences in mechanisms.

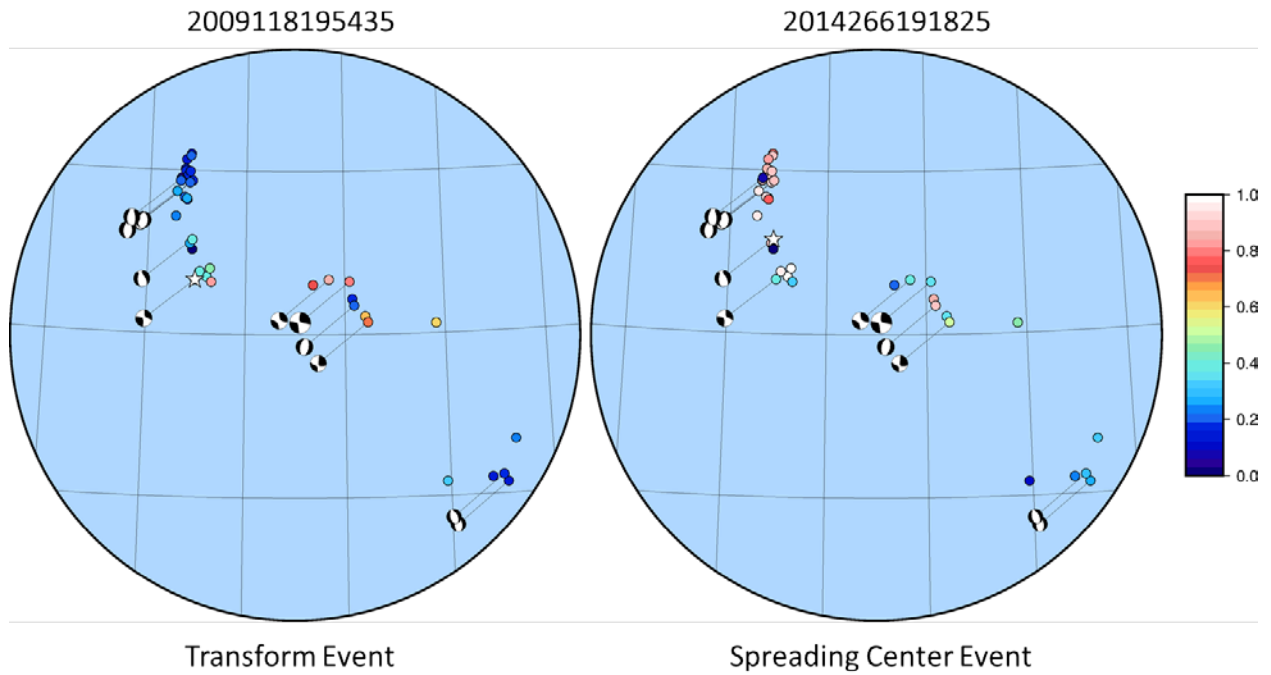


Figure 61: Detail of the mid-Atlantic ridge region with transform faults and spreading centers. A transform event responds to templates from events along transform faults, but not spreading center events, and a spreading center event does the reverse. Grid spacings are 2 degrees in these maps. GCMT mechanisms are indicated where available.

The map at the right side of Figure 61 shows the reverse pattern. In this case, the target event 2014266191825 (location indicated by the white star) is, from its position and GCMT mechanism, a spreading center event. This event “lights up” the nearby events on spreading center segments of the ridge, but has lower values for the transform events. Most strikingly, it has values above 0.8 for two templates along a short spreading center segment between the two transform faults. In the left part of the figure, those same two templates had values below 0.3.

So we have established that empirical MFP with the GRSN as a receiving aperture and using all three components is quite sensitive to mechanism in this region. How MFP distinguishes the mechanisms is apparent in Figure 62, which shows waveforms from the vertical channels of the network (black) superimposed on the waveforms from the north channels of the network (red). The north channels are approximately transverse to the direction of travel. The transform event (2009118195435) generated significant Love waves that lead and are generally larger than the Rayleigh waves. They dominate the north channels. The spreading center event (2014266191825) appears to be dominated by Rayleigh waves, which generally are diminished

on the north channels due to their approximate transverse orientation. Thus waves from these two events have distinctly different polarizations, and, in the case of the transverse event a mixture of phases with different wave speeds.

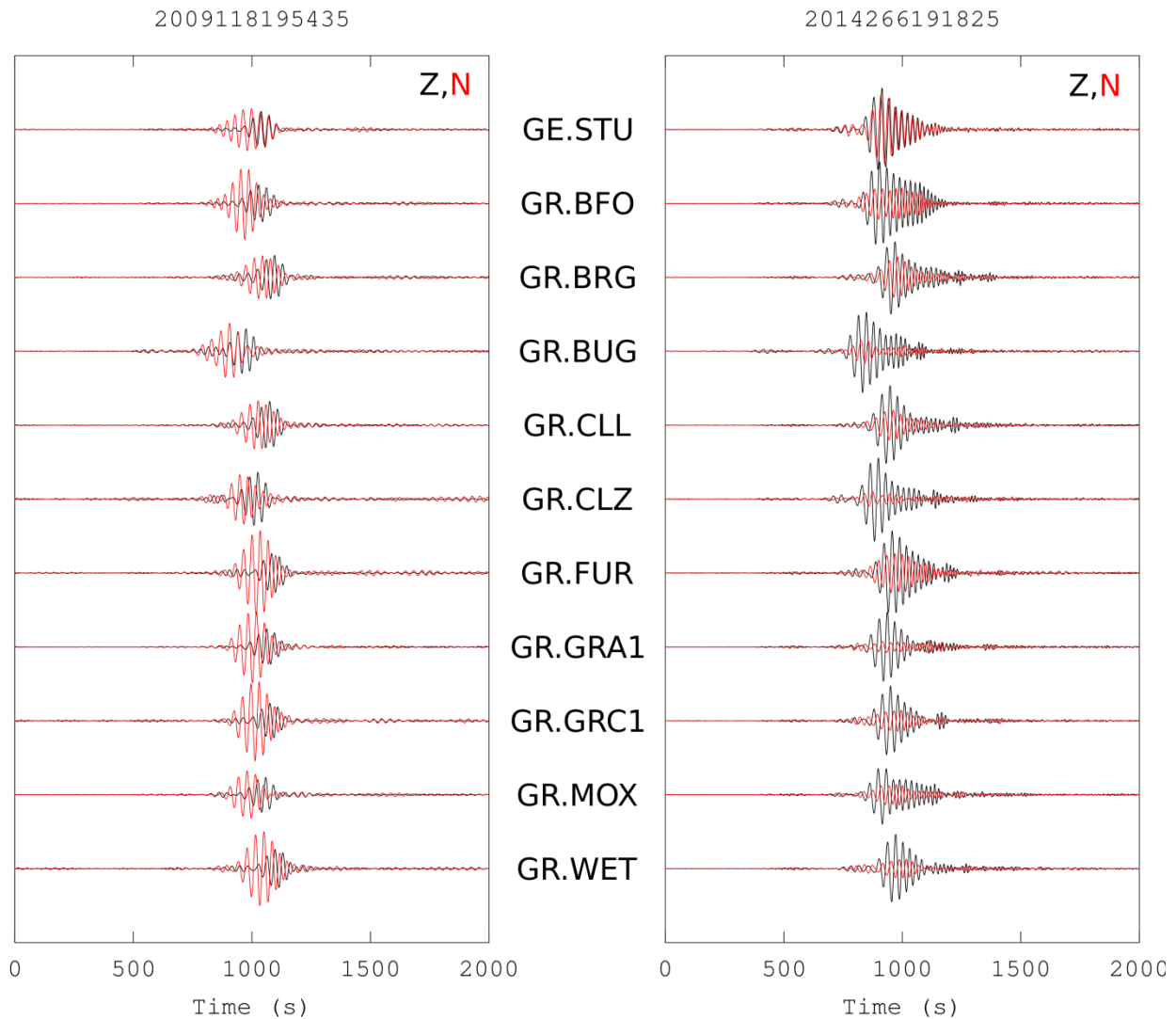


Figure 62: Vertical (black) and north (red) component waveforms for the two events of Figure 61. The transform event waveforms are at left and the spreading center waveforms are on the right. Pronounced Love waves are evident in red on the north channels at left, but are largely absent on the same channels on the right. Different polarizations and combinations of wave speeds in these arrivals allow MFP to distinguish the two mechanisms as seen in Figure 61.

Zagros/Iran

Matched field response patterns for events in Iran are more chaotic than for events along the mid-Atlantic ridge. Figure 63 shows the MFP responses at 35 second period in the Zagros and Iran to two events, one at the northwest end of the Zagros mountains and the other at the opposite southeast end of the mountain chain. Because the events are more distant from the GRSN than the MAR events, a larger surface wave window was used (duration from 700 to 1150 seconds). The surface waves were substantially more complex than the typical mid-Atlantic ridge events.

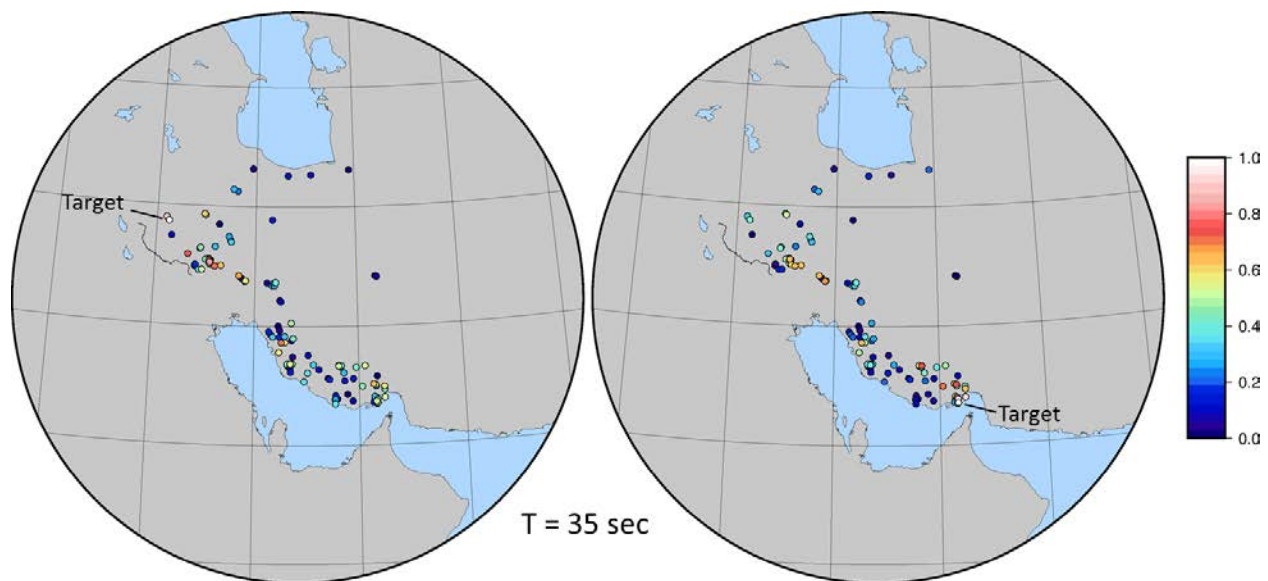


Figure 63: EMFP responses for an event at the northwestern end of the Zagros (left) and for another at the southeastern end. The patterns of responses appear almost random, but do have some higher values scattered along a line from the receiving aperture through each event. Since the Zagros events are distributed approximately along a radial line from the observing stations, high values are expected along the length of the Zagros for these two events. Grid spacing is 5 degrees in these maps.

The empirical matched field responses appear almost random, but do have higher values along a radial line from the GRSN through the target events. We think that several effects contribute to the scattered nature of the responses, chief among them the fact that the GRSN is oriented in a nodal direction from the majority of events in this region. Other contributing factors include the greater diversity of mechanisms in the Zagros/Iran region and the greater heterogeneity of the medium. The surface waves are more complicated for the Zagros events as can be seen in Figure 64, which compares waveforms in the 17.5 second to 70 second band for an Mw 7.1 mid-

Atlantic ridge event (2015044185912) with those from a Zagros event (2013099115249). The nodal direction of the path from the Zagros event and greater structural complexity produce greater waveform complexity.

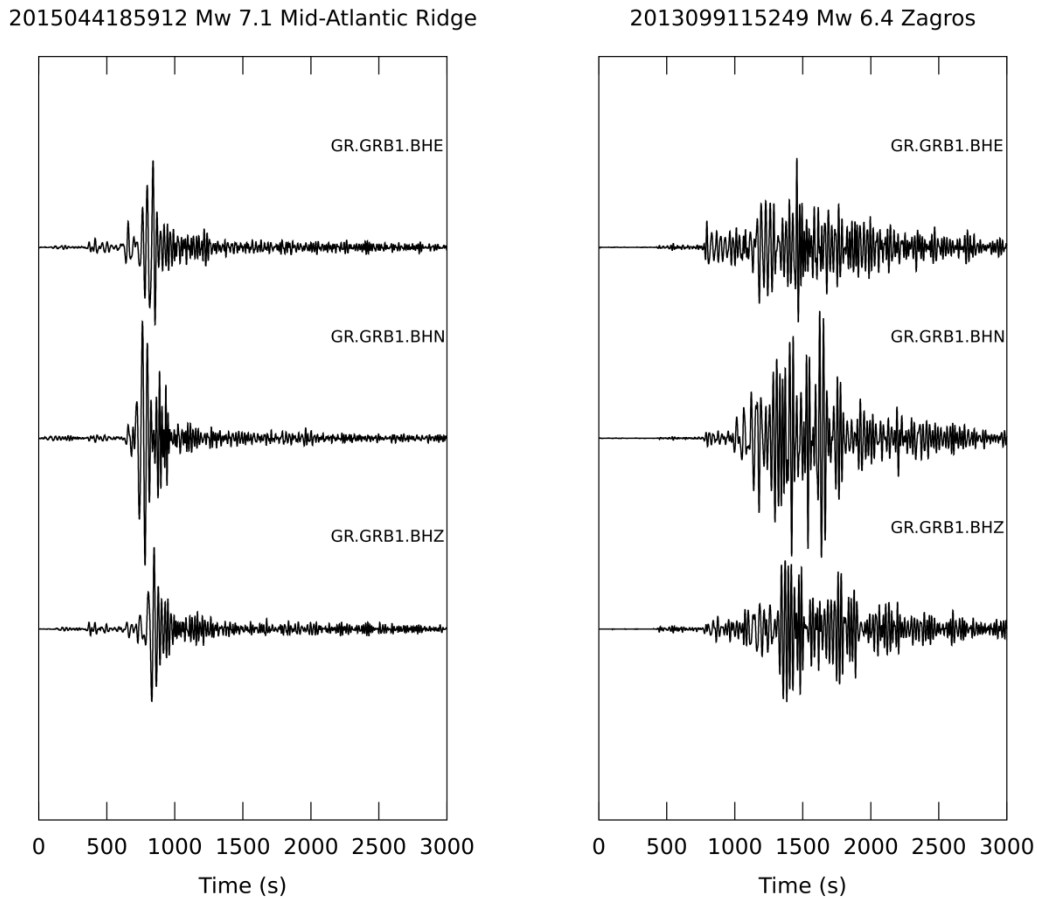


Figure 64 Waveforms from events along the mid-Atlantic ridge to the GRSN typically are much simpler than those along a path from the Zagros.

Detail of the matched field processing response to the Mw 6.4 Zagros event is shown in Figure 65. Global CMT Project mechanisms, where available, are shown for the events. To an extent, template events with a mechanism similar to the target event mechanism have high responses. Lower responses may be due to slight differences in orientation (which are amplified by the nodal path effect) and strong medium heterogeneity, perhaps causing different wavefield scattering effects at different source locations.

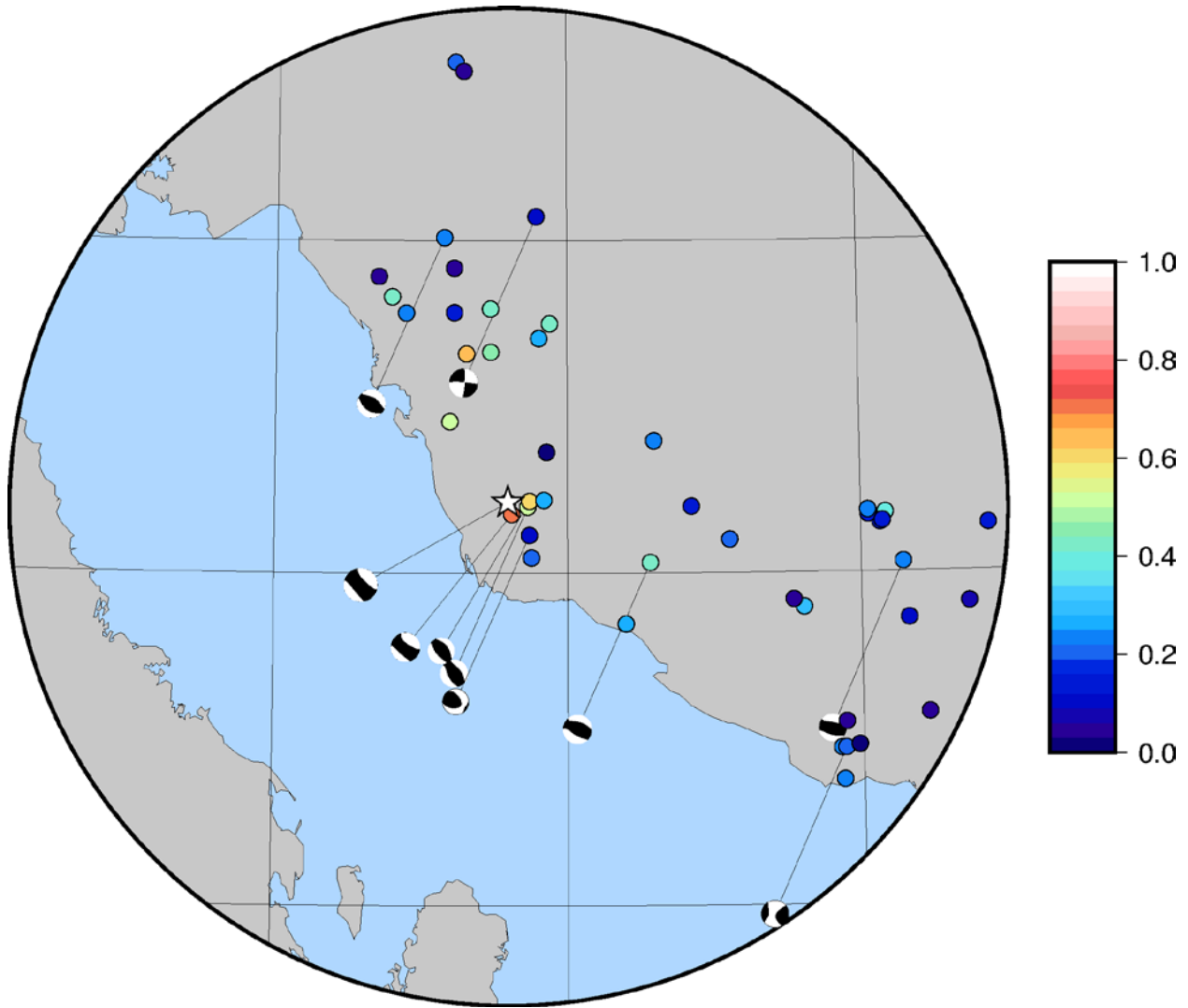


Figure 65. Detail of the matched field response to event 2013099115249, a magnitude 6.4 event in the central Zagros. (Location indicated by white star). The grid spacing is 2 degrees in this map.

Computations

Simulations of the observed seismograms were conducted with SPEC-FEM3D-GLOBE (Tromp et al., 2008) an MPI spectral-element elastic wavefield solver optimized for implementation on highly concurrent architectures (e.g. supercomputers, multicore servers). We ran the code on dual 8-core Xeon servers with 128 GB of memory at NORSAR and a 64-core Knights Landing server with 192 GB of memory at Deschutes Signal Processing LLC. As shown in Figure 39, there were two computational domains, one each for the two source regions studied in this part of

the project. Figure 66 shows the computational domain for the MAR source region. The computational domain is a single chunk of SPECFEM3D-GLOBE [Specfem3D globe manual (2015)] outlined in red in the figure. It extends from the surface region depicted to the core. At the resolution chosen for our computations (NEX=256), the computational mesh requires about 50GB of memory. The resolution is suitable to generate synthetics with negligible numerical dispersion at periods down to about 17 seconds.



Figure 66: Domain for spectral element synthetics calculations for the north mid-Atlantic ridge events.

The event indicated by the white star is 2011279003932. The domain is a single “chunk” that may be used to generate regional synthetics with SPECFEM3D-GLOBE. The events and observing stations are well away from the boundaries, which minimizes effects of absorbing boundary conditions on the calculations.

The SPECFEM3D-GLOBE distribution has a large number of earth models to choose from, including S20RTS, S40RTS, S2.9ea and S362ANI. The mantle models can be combined with

several crustal models, among them CRUST2.0. For our simulations, we chose S2.9ea with CRUST2.0. This model has higher mantle resolution under Eurasia than most of the alternatives.

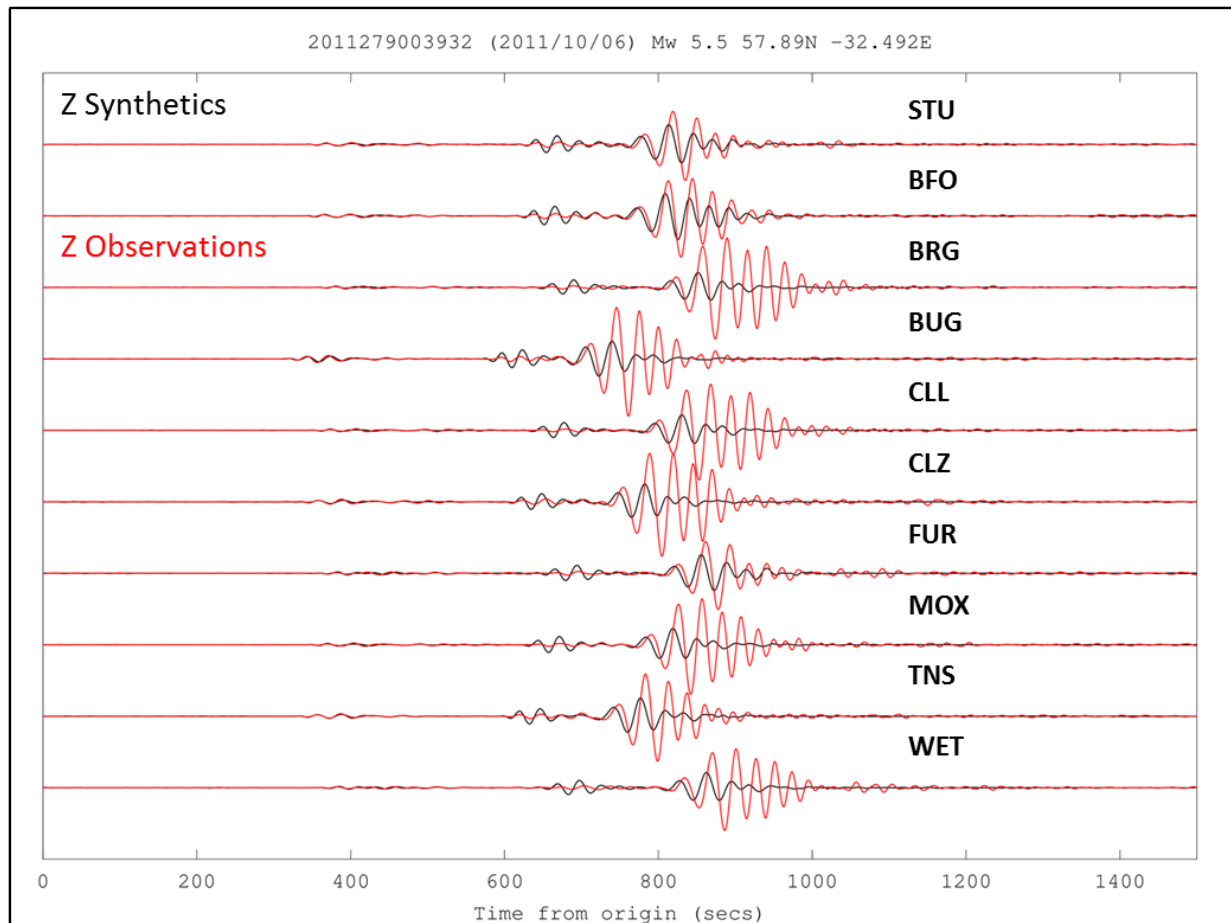


Figure 67: Comparison of the S2.9ea synthetics and observations in the 25-50 second band for a magnitude 5.5 event on October 6, 2011. The location of this event was shown as the white star in Figure 66. The P wave synthetic amplitudes and arrival times are good. The S wave synthetic amplitudes are too large, but the timing generally is good. The synthetic Rayleigh wave amplitudes are too small and the arrival time is a bit early; however, their relative phase (but not amplitude) across the network is good.

Figure 66 compares the vertical component recordings across the GRSN with corresponding S2.9ea synthetics for a magnitude 5.5 event along the spreading center just south of Iceland. We used the GCMT solution as the source mechanism for the simulation. The timing of the phases generally is quite good, but the amplitudes are significantly in error. The S phases are significantly larger in the synthetics than in the observations and the Rayleigh waves are smaller. The Rayleigh waves in the synthetics are early compared to their counterparts in the observations, though the relative phase across the network appears to be accurate. The

relative phase probably is more important for matched field processing. However, the relative amplitudes also are important in matched field processing when the covariance matrix is not diagonally normalized.

Despite the evident mismatch between the synthetics and observations, we proceeded to conduct larger scale reciprocal calculations to see whether synthetics could tell us anything about the behavior of empirical matched field processing, and also to examine the workload and work flow required to generate the library of synthetics necessary for model-based matched field processing on a scale suitable to monitoring applications. We anticipate that models eventually will produce accurate synthetics in the long-period band, so this exercise could be viewed as preparation to use the accurate adjoint tomography models (AT) likely to be available in the next 5 years.

Models developed with adjoint tomography are much more likely to predict MFP performance accurately since they are developed by matching waveforms directly rather than measurements (like group delay dispersion curves) that result in significant information loss (such as phase amplitudes). At least two adjoint tomography models have been developed that are desirable to test in our study areas. Neither is available as part of the standard SPECFEM3D-GLOBE distribution. The first, EU60 (Zhu et al., 2015), is a high-resolution, anisotropic model developed for the Europe and the north Atlantic. We have inquired about the availability of this model, but not received a reply. The second, GLAD-M15 is a more recent global AT model developed by a large consortium (Bozdag et al., 2016). GLAD-M15 was developed on the Oak Ridge Titan supercomputer using 3.8 million measurements on three-component observations of 253 earthquakes. We have inquired about the availability of this model and received a positive response. Plans are underway to supersede GLAD-M15 with a model developed with thousands of events using GLAD-M15 as a starting point (Bozdag et al., 2017).

MFP requires path calibrations from a large number of source locations to a modest number of receiver sites. Consequently, the model-based calibrations are most efficiently calculated using reciprocal methods (Eisner and Clayton, 2001). Reciprocal calculations place the source at the receiver location and record the field at the multiple source locations. A full construction of the collection of Greens functions describing an arbitrary moment tensor source requires three separate simulations, one for each component (z , n , e) of the receiver and recording of the strain tensor at each source location. The strain tensor has 6 unique components, one each corresponding to the M_{rr} , M_{rp} , M_{rt} , M_{pp} , M_{pt} and M_{tt} components of the moment tensor. When this project began, SPECFEM3D-GLOBE did not have an option for point force sources, but only moment tensor sources. Dr. Hom Nath Gharti of Princeton University volunteered to help us by modifying the code to support point force sources, and we served as beta testers for the modification. SPECFEM3D-GLOBE still lacks an option to record the strain tensor, which is the other requirement for

performing reciprocal calculations. However, the strain tensor at a particular source location can be approximated by recording the displacement field at locations perturbed around the source, then approximating the gradients with finite differences (Eisner and Clayton, 2001). We conducted a validation test of this approximation with a source at the location of event 2001096093309 (53.672 - 35.28) along the MAR to station GRFO in Germany.

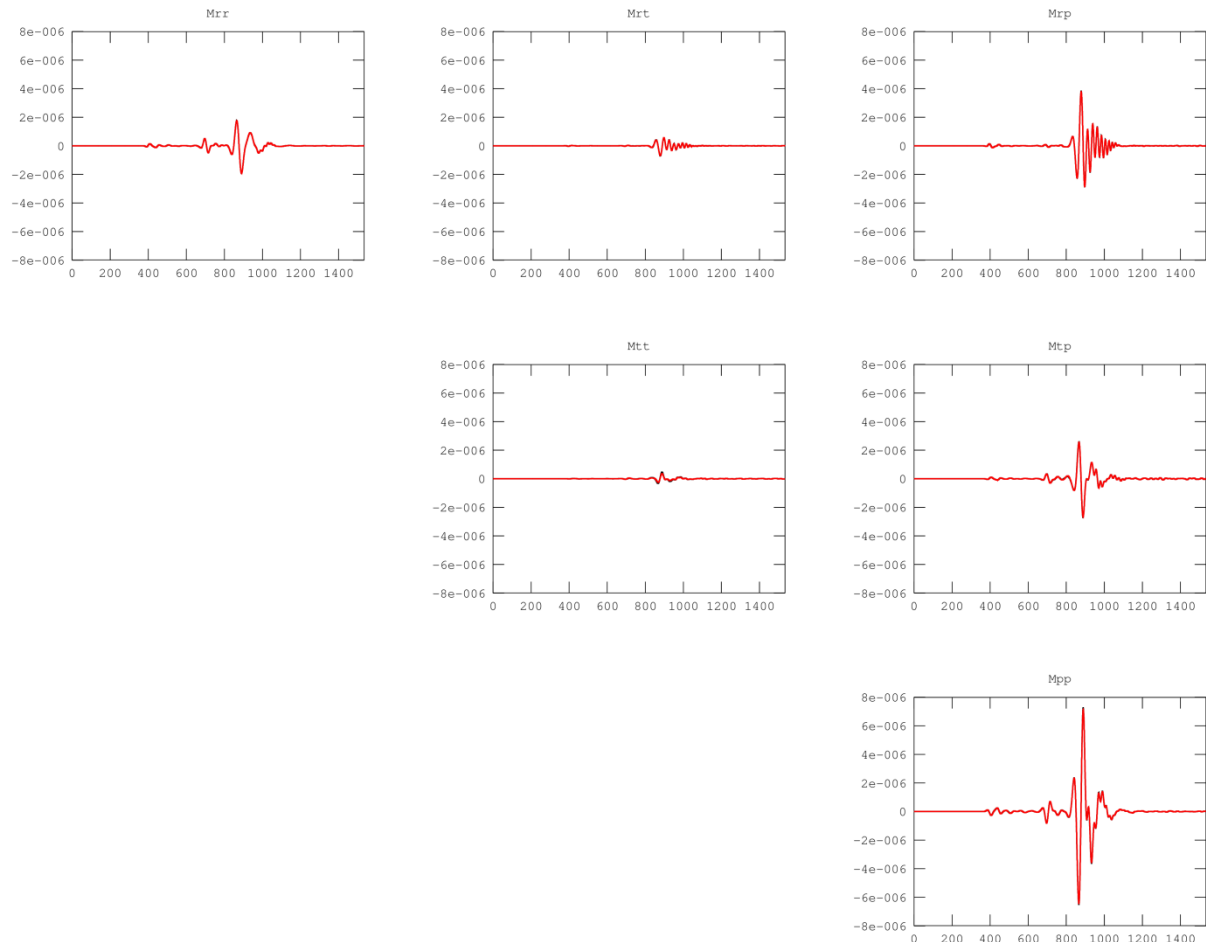


Figure 68 Comparison of the Z components of the moment tensor (black) superimposed by their finite-difference approximations (red). The traces are shown in true relative amplitude.

The test consisted of six forward calculations with the source placed at the event location, one for each of the moment tensor components, and three reciprocal calculations with the source placed at the station location, one for each of the three components of motion. For each of the forward calculations, the displacement is recorded along the three components of motion at the station location. For the reciprocal calculation, the strain tensor is estimated from the displacement field around the source location. To estimate the elements of the strain tensor, we calculate derivatives of the displacement field, which we estimate with central finite

differences at the source location. This calculation requires saving the three-component displacement field at six perturbed locations around the source location, which we took to be one kilometer offsets in each of the principal directions (up/down, east/west, north/south). Along with the field at the source location, the full archive for one reciprocal calculation at one source location consists of 21 recorded waveforms (21 output SAC files). Since the calculation is repeated for each orientation (z, n, e) at the receiver, there are 63 files involved in computing the wavefield and the strain tensor at the source.

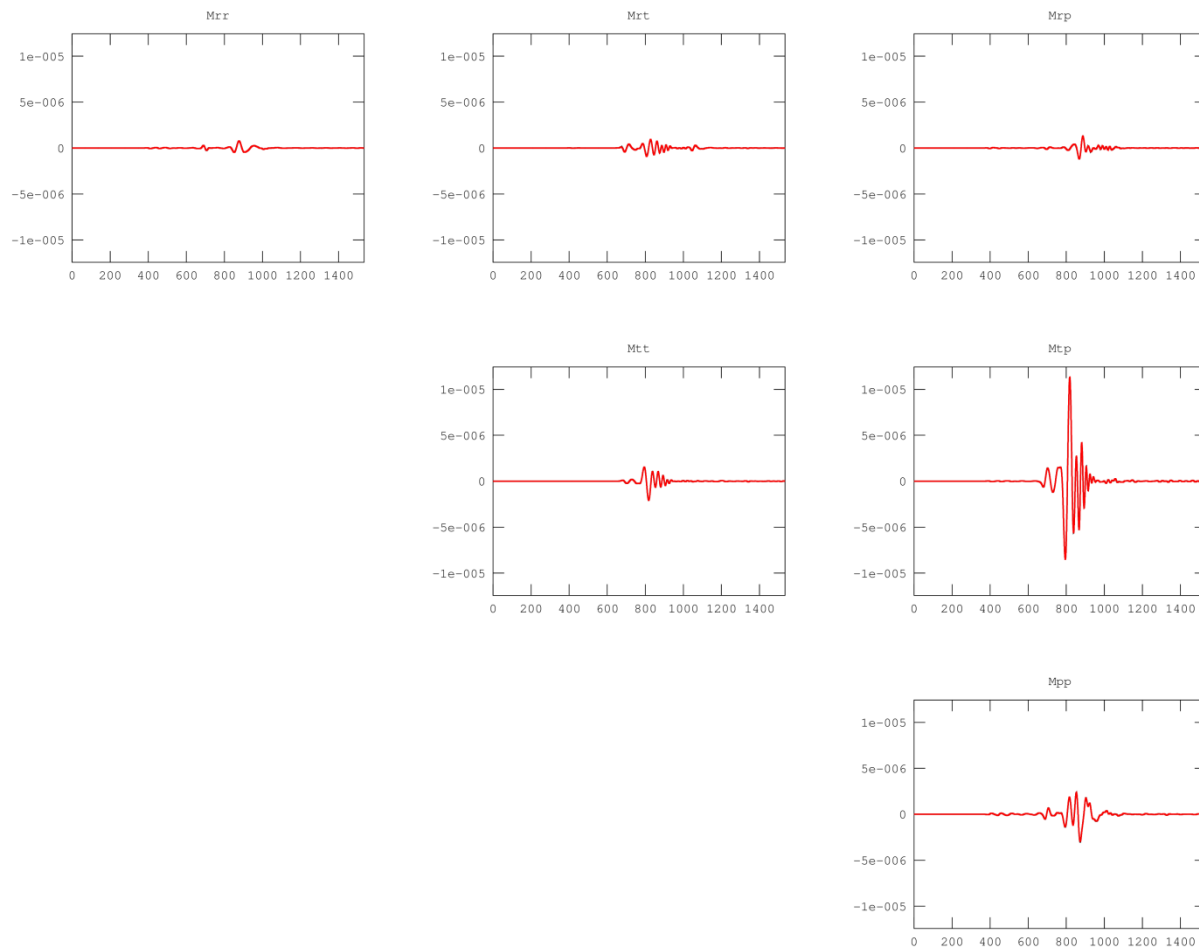


Figure 69 Comparison of the N components of the moment tensor (black) superimposed by their finite-difference approximations (red). The traces are shown in true relative amplitude.

Figure 68 shows the Z components of motion (at the receiver) for the six moment tensor components (at the source) computed with a forward calculation (black) superimposed by the same components computed with the reciprocal calculation (red). The traces are shown in true relative amplitude among all components. The agreement is excellent. Figures 68 and 69 respectively show the corresponding N and E components of motion at the receiver.

These figures establish that the reciprocal calculation of moment tensor components is accurate with SPECFEM3D-GLOBE. We describe this fact in such detail because it was a significant effort, opening up a new use for this important code. When the code is modified to save strain tensors, the bookkeeping associated with the files produced by reciprocal calculations will be substantially reduced.

Larger scale calculation of a library of Greens functions for the 227 event locations to the 13 stations of the GRSN shown in Figure 63, involves 185,913 files (63 files per source x 13 stations = 819 files per source x 227 sources). This is a significant bookkeeping effort as data in these files have to be organized and reduced from multiple independent runs of SPECFEM3D-GLOBE to generate gradients for the moment tensor components. We developed specialized software in Java to construct libraries from these files for the mid-Atlantic ridge source region and the Zagros region.

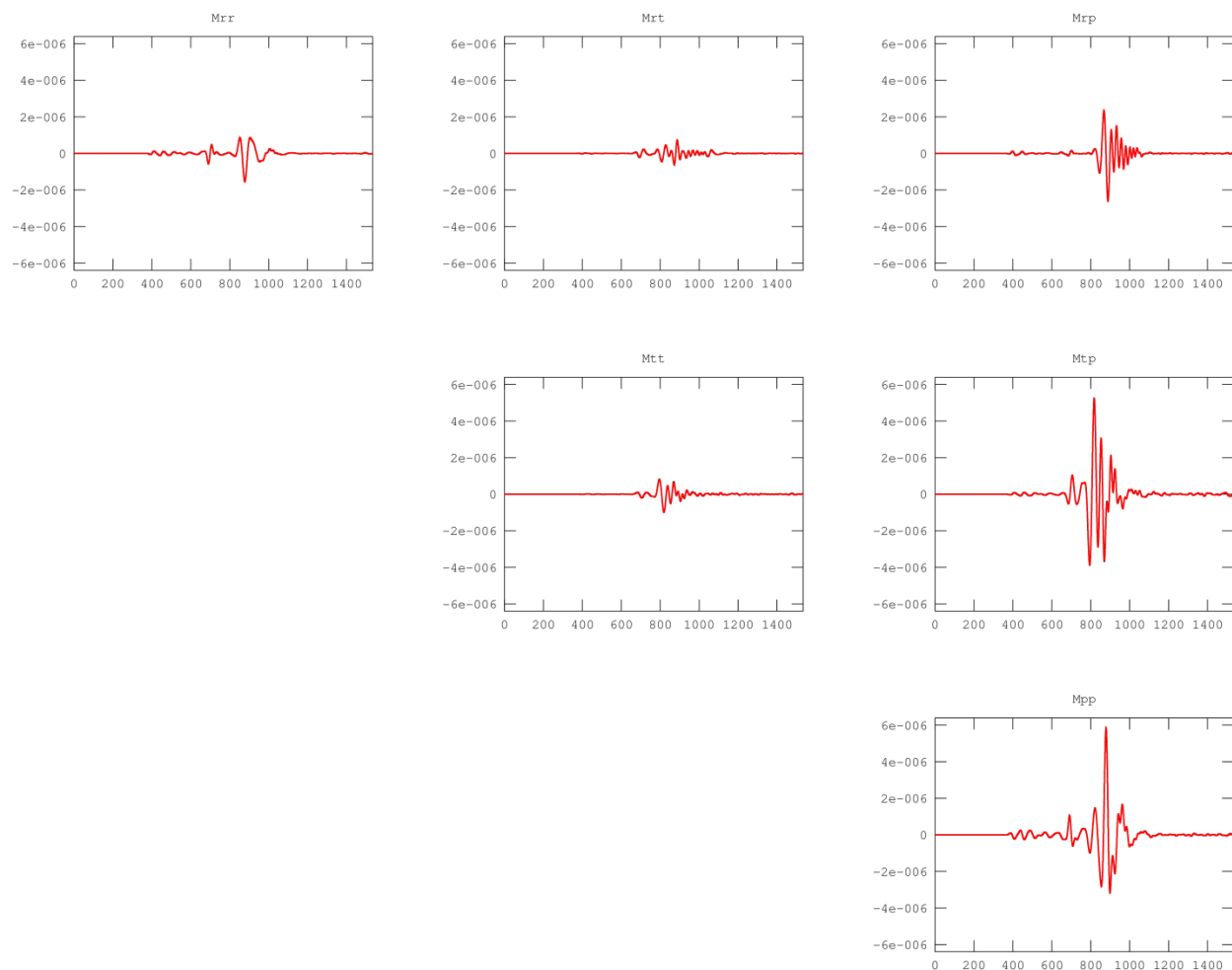


Figure 70: Comparison of the E components of the moment tensor (black) superimposed by their finite-difference approximations (red). The traces are shown in true relative amplitude.

The number of runs required was 39 (13 stations x 3 components of motion) for each of the two regions. To capture the surface waves for all events (including the most distant from the GRSN) required calculations to 40 minutes propagation time for the MAR region and 50 minutes for the Zagros region. On the Knights Landing (Xeon Phi) architecture, a typical Zagros calculation required about 3.5 hours wall time for the viscoelastic wavefield solver. The full suite of calculations for one computational domain took about 140 hours.

What this analysis tells us is that the generation of synthetics for attempting model-based assessment of MFP performance or even model-based MFP is computationally feasible to about 20 seconds period with even modest machines. The accuracy of synthetics will need to improve before model-based MFP will begin to rival empirical MFP, and this may occur in the next 5 years with the availability of higher-resolution models developed with adjoint tomography.

We used the synthetics library created for the Iran events in an attempt to examine whether the high variability of matched field templates observed there is due to heterogeneous media or variations in source mechanism. Figure 71 shows the MFP response to the event 2013326193058 at the northwest end of the Zagros (reproduced from Figure 63). At right in Figure 71 is the response predicted by the Greens functions from the synthetics library using mechanisms from the Global CMT project (Dziewonski et al., 1981; Ekström et al., 2012). The synthetics predict better (higher) responses to the target event along the Zagros chain than obtained empirically. However, the synthetic results still exhibit a significant amount of variation. The synthetics do not correctly predict observed MFP responses at many or most event locations. In part, this failure might be attributed to the fact that the observing array (GRSN) is approximately nodal to many of the mechanisms in the Zagros source region. Small uncertainties in orientation of the mechanism principal axes may produce large relative effects in the observed waveforms.

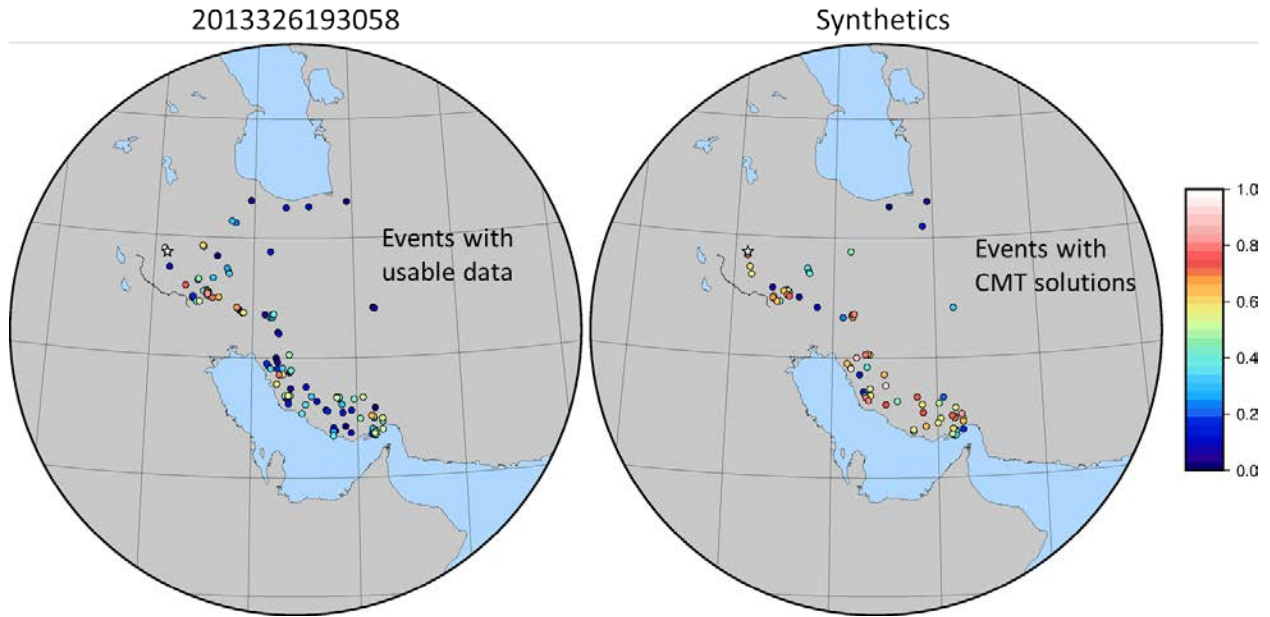


Figure 71: Comparison between actual empirical matched field responses to data from event 2013326193058 (location indicated by the white star) and the responses predicted by the synthetics using mechanisms obtained from the Global CMT Project [Dziewonski et al., 1981; Ekström et al., 2012]. The synthetics show substantial variability, but over-predict the responses along the Zagros.

In an effort to separate the effects of medium heterogeneity and mechanism variability, we changed the simulation to give all events a single (explosion) mechanism. This change ensures the mechanisms are properly matched between the templates and the target event, with only structural propagation effects controlling the MFP response to the data. Figure 72 contrasts the MFP response to 2013326193058 given the GCMT mechanisms (left) to the response if events had explosion mechanisms (right). In the latter case, the response is uniformly high, with the exception of events south of the Caspian that are far off-azimuth (as seen from the GRSN). This simulation suggests that the variability of mechanisms dominates the variation seen in the Iranian MFP responses.

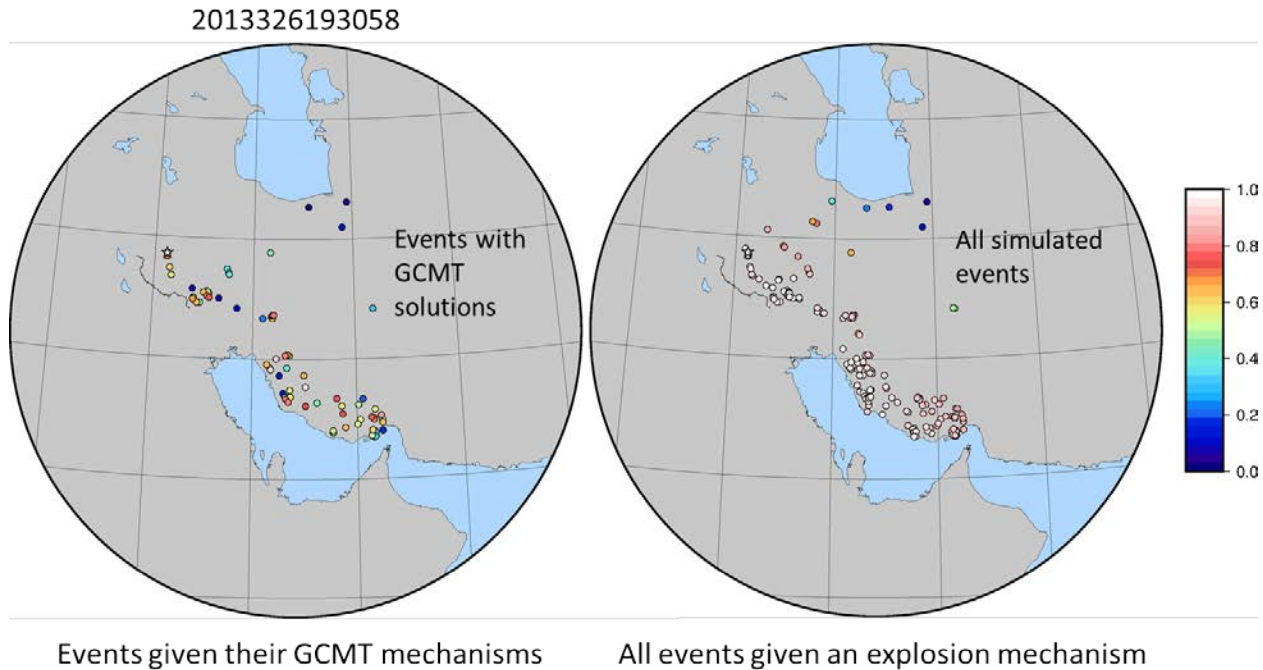


Figure 72: Comparison of synthetic MFP responses to event 2013326193058 under two different mechanism scenarios. At left, the events that have Global CMT project solutions are given those mechanisms, while at right, all simulated events (including the target event) are given the same (explosion) mechanism. There is relatively little variation in MFP response along the Zagros in the latter case, suggesting that mechanism has the dominant effect on MFP response rather than propagation medium heterogeneity.

Since the actual empirical MFP responses (left panel, Figure 71) are so much lower than their synthetic counterparts (right panel, Figure 71), we suspect that the effect of medium heterogeneity probably is much greater than indicated by the right panel of Figure 72. This effect is an indication that the model S2.9ea is not capturing the full heterogeneity of the earth in this region at the resolution of surface waves at 35 seconds period. We imagine testing this hypothesis by repeating the calculations with an adjoint tomography model. For the north Atlantic to Europe paths operative in the mid-Atlantic ridge part of this study, the ideal model would be EU60 [Zhu et al., 2015], which in addition to having higher resolution also captures anisotropy. Anisotropy can have a very large effect on surface waves originating along the MAR, given the significant structural fabric of the crust emanating from the spreading center.

4 Conclusions

Matched Field Processing (MFP) is a narrow frequency band approach to signal processing which exploits the spatial structure of a wavefield propagating over an array or network of sensors. At a given time, the narrowband phase-shifts (and, if required, amplitude relations) between the signals recorded on the different sensors can be measured and the pattern of observed phase-shifts can be compared with a template of anticipated phase-shifts for a wavefront of interest. The phase-shift template can be derived from theoretical considerations (as is the case with plane wavefront models) or can be measured from previous observations. The latter approach is referred to as empirical matched field processing or EMFP. In seismology, EMFP has previously been used both as a signal detector and a signal classifier and falls into the category of so-called pattern detectors. It has demonstrated advantages over waveform correlation detectors for sources of seismic waves such as ripple-fired quarry blasts. This is because highly differing source-time histories for seismic events change the temporal structure of the wavetrain significantly while the narrowband phase structure between distant sensors remains relatively constant.

In this project we have explored issues in the application of EMFP as a front-end detector in an autonomous seismic processing pipeline. Any seismic array for which a beam deployment provides a comprehensive detector of incoming signals can exploit matched field processing since classical f-k analysis is but a special case of MFP.

Using an example of the aftershock sequence from a large earthquake recorded on a seismic array, we have sketched out in detail a procedure for implementing a matched field detector for identifying signals for which the phase-shift pattern closely matches that of the template signal. Its significance for the aftershock sequence is that we are able to reliably identify aftershock arrivals using only the first P-arrival from the main shock. We describe why EMFP mitigates the problems associated with classifying the sequence using correlation-type detectors. We demonstrate that the performance of the empirical matched field detector is significantly better than the performance of the plane-wave matched field detector. This is explained by examining the phase-shifts at different frequencies for observed signals in the sequence. The matched field detector captures the true properties of arriving wavefront which differ significantly to those predicted from the presumed backazimuth and apparent velocity.

We describe a matched field beam-deployment for the SPITS array based upon ensemble covariance matrices constructed from regional clusters of seismicity. For such a small aperture array, it is relatively easy to cover the entire slowness space with a deployment of matched field beams which can easily be run in real-time. The beams can be constructed either in slowness or geographical space. By basing direction-of-arrival estimates on matches with empirical

templates rather than predicted phase-shifts we are able to mitigate the significant bias observed on many arrays.

Considering seismicity surrounding Japan recorded on the three arrays USRK (Russia), KSRS (South Korea), and MJAR (Japan) we demonstrate how the performance of empirical matched field detectors is likely to vary with the array geometry. On the fairly small aperture USRK array, a full coverage of slowness space can be provided with a relatively limited number of matched field beams. On the somewhat larger KSRS array, many more beams are necessary to be confident of detecting all seismic signals. The MJAR array is of a similar aperture to KSRS but the coherence between signals on neighboring sensors is far lower for MJAR than for KSRS due to the significant geological heterogeneity close to MJAR. The phase-shifts between sensors at MJAR at high frequencies are likely to be far more source-specific than for the USRK or KSRS arrays.

The problem of how to classify high frequency signals recorded on large aperture arrays has persisted for several decades. There are two fundamental problems: *incoherence*, the fundamental difference in the shape of the waveforms at neighboring sensors, and *aberration*, the delaying of coherent waveforms by time-offsets which are not consistent with the wavefront propagation model. We demonstrate how both of these effects can be mitigated by first cross-correlating the waveforms between the different sensors and then performing subsequent array-processing on limited sets of these correlation traces. We show examples of greatly enhanced slowness estimates for high frequency signal arrivals on large aperture arrays using the LSPC (Limited Sensor Pair Correlation) method and demonstrate that it can also apply to small aperture arrays, for which regional S-phases often lack coherence.

The properties of LSPC can be emulated in EMFP by setting the elements in the covariance matrices which correspond to pairs of sensors with large inter-site distances to zero. This widens the footprint of the detector at the expense of some degree of sensitivity to targets identical to that from which the template was taken. We demonstrate that the principal eigenvectors obtained from these “windowed covariance matrices” generally resemble those obtained from the full covariance matrices. Where and when they differ significantly can often provide insight into the performance of matched field detectors which employ these steering vectors.

Having built a prototype front-end detector, we put it to test on the USRK and KSRS arrays at regional distances from the North Korea nuclear test site. Matched field templates have been formed for Pn phases from the test site using both theoretical phase-shifts (plane wave) and empirical phase shifts (from previously observed seismic events at the test site). With these templates, multiple detectors have been run in parallel in near real-time – looking for evidence

of seismic disturbances at or close to the test-site. A small event on October 12, (magnitude ~ 3) was detected clearly by all of the matched field templates – on both USRK and KSRS arrays. There were significant differences between the qualities of the detections for this event from the various templates. The templates obtained from the two events on 23 September 2017 and the declared underground nuclear test on 8 October 2006 performed far better than the templates from the later declared nuclear tests. Using the windowed covariance matrices, the detectors become somewhat less source specific and the performance becomes rather more uniform for the different templates.

We examined the idea that the number of matched field templates required to calibrate a source region would be driven by the heterogeneity of the propagation medium. Our hypothesis was that relatively simple regions would require fewer templates and more complex regions more templates, i.e. that matched field processing would increase resolution in complex regions, but require a larger number of calibration events to do so. This notion was borne out, at least by the long-period test we performed, but for a different reason: the variation among matched field templates for the three-component array we used was driven by the diversity of event mechanisms. Mechanism variation effectively masked heterogeneity. We attempted to separate these effects with synthetics, but the 3D velocity model we chose (s2.9ea, one used frequently with SPECFEM3D-GLOBE, our choice of anelastic solver) probably did not have sufficient resolution to model the full wavefield effects of heterogeneity. Tests with synthetics did show the effects of template diversity with mechanism variation, but to a lesser degree than actually experienced empirically. It is likely that medium heterogeneity plays a significant role, especially (with surface waves) when anisotropy is significant as it is around spreading centers, effectively amplifying the radiation patterns of different mechanisms.

The sensitivity of templates to mechanism suggests the possibility that matched field processing, performed as it was here at long periods, may allow simultaneous detection and discrimination. It could be used as a screen for explosions among earthquakes, essentially pushing the full-waveform moment tensor discriminant from its place as a post-processing function forward into the detection front-end of pipelines. This is a topic for future research.

Finally, in attempting to model the observations with synthetics, we examined the techniques and costs required to enable model-based matched field processing. We anticipate that, at long periods, adjoint tomography models such as EU60 and the successor to GLAD-M15 will provide much more realistic predictions of matched field processing performance and enable matched field techniques to be applied in areas lacking historic observations. At periods greater than 20 seconds, the computational requirements are well within reach of even modestly-equipped institutions, requiring hundreds or a few thousand hours of simulation on next generation servers for networks with tens or a few hundred observing sensors.

References

- Barrett, S. A. and G. C. Beroza (2014), An Empirical Approach to Subspace Detection, *Seismological Research Letters*, **85**, pp. 594-600, [doi:10.1785/0220130152](https://doi.org/10.1785/0220130152).
- Bondar, I., R. G. North, and G. Beall (1999), Teleseismic slowness-azimuth station corrections for the International Monitoring System Seismic Network, *Bulletin of the Seismological Society of America*, **89**, pp. 989-1003.
- Bozdag, E., D. Peter, M. Lefebvre, D. Komatitsch, J. Tromp, J. Hill, N. Podhorszki, and D. Pugmire (2016), Global adjoint tomography: first-generation model, *Geophysical Journal International*, **207**, pp. 1739-1766, <https://doi.org/10.1093/gji/ggw356>.
- Bozdag, E., M. Lefebvre, W. Lei, R. Orsvuran, D. Peter, Y. Ruan, J. Smith, D. Komatitsch, and J. Tromp (2017), *Geophysical Research Abstracts*, **19**, EGU2017-17212.
- Cansi, Y. (1995), An automatic seismic event processing for detection and location: The P.M.C.C. Method, *Geophysical Research Letters*, **22**, pp. 1021-1024, [doi:10.1029/95gl00468](https://doi.org/10.1029/95gl00468).
- Dziewonski, A. M., T.-A. Chou, and J. H. Woodhouse (1981), Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *Journal of Geophysical Research*, **86**, pp. 825-852, [doi:10.1029/JB086iB04p02825](https://doi.org/10.1029/JB086iB04p02825).
- Eisner, L. and R. W. Clayton (2001), A reciprocity method for multiple-source simulations, *Bulletin of the Seismological Society of America*, **91**, pp. 553-560, <https://doi.org/10.1785/0120000222>.
- Ekström, G., M. Nettles, and A. M. Dziewonski (2012), The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, *Physics of the Earth and Planetary Interiors*, **200-201**, pp. 1-9, [doi:10.1016/j.pepi.2012.04.002](https://doi.org/10.1016/j.pepi.2012.04.002).
- Engen, Ø., O. Eldholm, and H. Bungum (2003), The Arctic plate boundary, *Journal of Geophysical Research - Solid Earth*, **108**, B2. 2075, [doi:10.1029/2002jb001809](https://doi.org/10.1029/2002jb001809).
- Gibbons, S. J. and F. Ringdal (2006), The detection of low magnitude seismic events using array-based waveform correlation, *Geophysical Journal International*, **165**, pp. 149-166, [doi:10.1111/j.1365-246x.2006.02865.x](https://doi.org/10.1111/j.1365-246x.2006.02865.x).
- Gibbons, S. J., F. Ringdal, and T. Kväerna (2008), Detection and characterization of seismic phases using continuous spectral estimation on incoherent and partially coherent arrays, *Geophysical Journal International*, **172**, pp. 405-421, [doi:10.1111/j.1365-246x.2007.03650.x](https://doi.org/10.1111/j.1365-246x.2007.03650.x).
- Gibbons, S. J., T. Kväerna, and F. Ringdal (2010), Considerations in Phase Estimation and Event Location Using Small-aperture Regional Seismic Arrays, *Pure and Applied Geophysics*, **167**, pp. 381-399, <http://dx.doi.org/10.1007/s00024-009-0024-1>.

- Gibbons, S. J., F. Pabian, S. P. Näsholm, T. Kværna, and S. Mykkeltveit (2017), Accurate relative location estimates for the North Korean nuclear tests using empirical slowness corrections, *Geophysical Journal International*, **208**, pp. 101-117, <http://dx.doi.org/10.1093/gji/ggw379>.
- Gibbons, S. J. and F. Ringdal (2012), Seismic Monitoring of the North Korea Nuclear Test Site Using a Multichannel Correlation Detector, *IEEE Transactions of Geoscience and Remote Sensing* **50**, pp. 1897-1909, [doi:10.1109/tgrs.2011.2170429](https://doi.org/10.1109/tgrs.2011.2170429).
- Gibbons, S. J., J. Schweitzer, F. Ringdal, T. Kværna, S. Mykkeltveit, and B. Paulsen (2011), Improvements to Seismic Monitoring of the European Arctic Using Three-Component Array Processing at SPITS, *Bulletin of the Seismological Society of America*, **101**, pp. 2737-2754, [doi:10.1785/0120110109](https://doi.org/10.1785/0120110109).
- Gibbons, S. J., T. Kværna, D. B. Harris, and D. A. Dodge (2016), Iterative Strategies for Aftershock Classification in Automatic Seismic Processing Pipelines, *Seismological Research Letters*, **87**, pp. 919-929, [doi:10.1785/0220160047](https://doi.org/10.1785/0220160047).
- Harris, D. B. (2006), *Subspace Detectors: Theory*, UCRL-TR-222758, Lawrence Livermore National Laboratory, 700 East Avenue, Livermore, CA.
- Harris, D. B. and T. Kværna (2010), Superresolution with seismic arrays using empirical matched field processing, *Geophysical Journal International*, **182**, pp. 1455-1477, [doi:10.1111/j.1365-246x.2010.04684.x](https://doi.org/10.1111/j.1365-246x.2010.04684.x).
- Hicks, E. C., T. Kværna, S. Mykkeltveit, J. Schweitzer, and F. Ringdal (2004), Travel-times and Attenuation Relations for Regional Phases in the Barents Sea Region, *Pure and Applied Geophysics*, **161**, pp. 1-19, [doi:10.1007/s00024-003-2437-6](https://doi.org/10.1007/s00024-003-2437-6).
- Muirhead, K. T. and R. Datt (1976), The N-th root process applied to seismic array data, *Geophysical Journal of the Royal Astronomical Society*, **47**, pp. 197-210, [doi:10.1111/j.1365-246x.1976.tb01269.x](https://doi.org/10.1111/j.1365-246x.1976.tb01269.x).
- Myers, S. C., G. Johannesson, and W. Hanley (2007), A Bayesian hierarchical method for multiple-event seismic location, *Geophysical Journal International*, **171**, pp. 1049-1063, [doi:10.1111/j.1365-246x.2007.03555.x](https://doi.org/10.1111/j.1365-246x.2007.03555.x).
- Myers, S. C., G. Johannesson, and W. Hanley (2009), Incorporation of probabilistic seismic phase labels into a Bayesian multiple-event seismic locator, *Geophysical Journal International*, **177**, pp. 277-289, [doi:10.1111/j.1365-246x.2008.04070.x](https://doi.org/10.1111/j.1365-246x.2008.04070.x).
- Mykkeltveit, S. and F. Ringdal (1981), Phase identification and event location at regional distances using small-aperture array data, in "Identification of seismic sources - Earthquake or underground explosions" (eds. E. S. Husebye and S. Mykkeltveit), Reidel Publishing Company, pp. 467-481.

- Prieto, G. A., R. L. Parker, and F. L. Vernon (2009), A Fortran 90 library for multitaper spectrum analysis, *Computers & Geosciences*, **35**, pp. 1701-1710, [doi:10.1016/j.cageo.2008.06.007](https://doi.org/10.1016/j.cageo.2008.06.007).
- Ruigrok, E., S. Gibbons and K. Wapenaar (2017), Cross-correlation beamforming, *Journal of Seismology*, **21**, pp. 495-508, [doi:10.1007/s10950-016-9612-6](https://doi.org/10.1007/s10950-016-9612-6).
- Schissole, E. and J. Schweitzer (2004), Study of regional variations of the coda characteristics in the Barents Sea using small-aperture arrays, Semiannual Technical Summary Report 1-2004, NORSAR, PO Box 53, Gunnar Randers vei 15, N-2027 Kjeller, Norway, February 2004, pp. 70-80.
- Schweitzer, J. (2001a), HYPOSAT – An Enhanced Routine to Locate Seismic Events, *Pure and Applied Geophysics*, **158**, pp. 277-289, [doi:10.1007/pl00001160](https://doi.org/10.1007/pl00001160).
- Schweitzer, J. (2001b), Slowness Corrections — One Way to Improve IDC Products, *Pure and Applied Geophysics*, **158**, pp. 375-396, [doi:10.1007/pl00001165](https://doi.org/10.1007/pl00001165).
- Slinkard, M. E., D. B. Carr, and C. J. Young (2013), Applying Waveform Correlation to Three Aftershock Sequences, *Bulletin of the Seismological Society of America*, **103**, pp. 675-693, [doi:10.1785/0120120058](https://doi.org/10.1785/0120120058).
- Specfem3D globe manual (2015)
https://geodynamics.org/cig/software/specfem3d_globe/specfem3d_globe-manual.pdf.
- Taylor, S. R., S. J. Arrowsmith, and D. N. Anderson (2010), Detection of Short Time Transients from Spectrograms Using Scan Statistics, *Bulletin of the Seismological Society of America*, **100**, pp. 1940-1951, <http://dx.doi.org/10.1785/0120100017>.
- Tromp, J., D. Komatitsch, and Q. Liu (2008), Spectral-element and adjoint methods in seismology, *Communications in Computational Physics*, **3**, pp. 1–32, <https://authors.library.caltech.edu/9615/>.
- Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo (1998), A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bulletin of the Seismological Society of America*, **88**, pp. 95-106.
- Yoon, C. E., O. O' Reilly, K. J. Bergen, and G. C. Beroza (2015), Earthquake detection through computationally efficient similarity search, *Science Advances*, **1** (11), e1501057, [doi:10.1126/sciadv.1501057](https://doi.org/10.1126/sciadv.1501057).
- Zhu, H., E. Bozdog, and J. Tromp (2015), Seismic structure of the European upper mantle based on adjoint tomography, *Geophysical Journal International*, **201**, pp. 18-52. [10.1093/gji/ggu492](https://doi.org/10.1093/gji/ggu492).

List of Abbreviations

CTBT	– Comprehensive Nuclear-Test-Ban Treaty
EMFP	– Empirical Matched Field Processing
MFP	– Matched Field Processing
IDC	– International Data Center
IMS	– International Monitoring System
REB	– Reviewed Event Bulletin
SNR	– Signal-to-Noise Ratio

DISTRIBUTION LIST

DTIC/OCF	
8725 John J. Kingman Rd, Suite 0944	
Ft Belvoir, VA 22060-6218	1 cy
AFRL/RVIL	
Kirtland AFB, NM 87117-5776	1 cy
Official Record Copy	
AFRL/RVBYE/Dr. Frederick Schult	1 cy

This page is intentionally left blank.