



CTBT seismic monitoring using coherent and incoherent array processing

Tormod Kværna · Steven J. Gibbons ·
Sven Peter Näsholm

Received: 20 November 2020 / Accepted: 16 June 2021
© The Author(s) 2021

Abstract The detection and location capability of the International Monitoring System for small seismic events in the continental and oceanic regions surrounding the Sea of Japan is determined mainly by three primary seismic arrays: USRK, KSRS, and MJAR. Body wave arrivals are coherent on USRK and KSRS up to frequencies of around 4 Hz and classical array processing methods can detect and extract features for most regional signals on these stations. We demonstrate how empirical matched field processing (EMFP), a generalization of frequency-wavenumber or f-k analysis, can contribute to calibrated direction estimates which mitigate bias resulting from near-station geological structure. It does this by comparing the narrowband phase shifts between the signals on different sensors, observed at a given time, with corresponding measurements on signals from historical seismic events. The EMFP detection statistic is usually evaluated as a function of source location rather than slowness space and the size of the geographical footprint valid for EMFP templates is affected by array geometry, the available signal

bandwidth, and Earth structure over the propagation path. The MJAR array has similar dimensions to KSRS but is sited in far more complex geology which results in poor parameter estimates with classical f-k analysis for all signals lacking energy at 1 Hz or below. EMFP mitigates the signal incoherence to some degree but the geographical footprint valid for a given matched field template on MJAR is very small. Spectrogram beamforming provides a robust detection algorithm for high-frequency signals at MJAR. The array aperture is large enough that f-k analysis performed on continuous AR-AIC functions, calculated from optimally bandpass-filtered signals at the different sites, can provide robust slowness estimates for regional P-waves. Given a significantly higher SNR for regional S-phases on the horizontal components of the 3-component site of MJAR, we would expect incoherent detection and estimation of S-phases to improve with 3-component sensors at all sites. Given the diversity of the IMS stations, and the diversity of the methods which provide optimal results for a given station, we advocate the development of seismic processing pipelines which can process highly heterogeneous inputs to help associate characteristics of the incoming signals with physical events.

T. Kværna (✉) · S. J. Gibbons · S. P. Näsholm
NORSAR, PO Box 53, N-2027 Kjeller, Norway
e-mail: tormod@norsar.no

S. J. Gibbons
NGI, Sognsveien 72, N-0855 Oslo, Norway
e-mail: steven.gibbons@ngi.no

S. P. Näsholm
Department of Informatics, University of Oslo,
Gaustadalléen 23 B, N-0373 Oslo, Norway
e-mail: peter@norsar.no

Keywords Seismic monitoring · Array processing · CTBT monitoring · International Monitoring System · Seismic event location · Parameter estimation · Empirical signal processing · Matched field processing

1 Introduction

We define a seismic array to be a set of seismometers, distributed in a spatial pattern, with the distance between seismometer sites typically significantly shorter than the distance from source to receiver. Array stations are central to the global monitoring of seismic disturbances which could constitute breaches of the Comprehensive Nuclear-Test-Ban Treaty (CTBT). At present, 24 stations of the primary seismic network of the International Monitoring System (IMS) are arrays and, in most cases, they demonstrably provide a superior detection capability to 3-component stations (Kværna and Ringdal 2013). The performance superiority comes both from the ability to improve the signal-to-noise ratio (SNR) through delay-and-stack beamforming and from the ability to obtain robust estimates for the backazimuth and slowness of a signal using spatial analysis of the arriving wavefront. This helps the phase association process greatly by indicating both the likely identity of a seismic arrival and the direction it came from. However, the performance of an array for a given signal will depend upon the frequency content of a signal, the array geometry, and the local geological conditions.

The IMS seismic arrays have very diverse geometries, with apertures ranging from a few hundred meters to many tens of kilometres (Gibbons 2014). A small array (with an aperture of 1 or 2 km) will likely provide excellent detection and estimation capability for high-frequency regional phases but poor noise suppression and low-resolution slowness estimation for lower frequency teleseismic arrivals. A very large array (with aperture of above 20 km) will provide high array gain and high-resolution slowness estimation for teleseismic arrivals but inadequate signal coherence for regional signals. Different procedures have attempted to mitigate the difficulties caused by the array geometry. Gibbons et al. (2008) use spectrogram beamforming to facilitate detection and estimation of high-frequency signals on large aperture arrays while Selby (2013) exploits the spatial pattern of the coherent noise-field to improve the performance of teleseismic signal detection and estimation on medium and small aperture arrays.

The region surrounding the Sea of Japan (Fig. 1) is an exceptionally interesting region in test-ban treaty monitoring. In addition to the presence of the North

Korean nuclear test site, there are many regions of high natural seismicity with the capability of generating very large and damaging earthquakes with potentially extensive aftershock sequences. Accurate and robust detection and location of seismic events in the region, covering a wide range of magnitudes, is crucial both for regional and global seismic monitoring. The detection capability for the smallest seismic events in the region is determined by three primary IMS seismic arrays: USRK (Ussuriysk, Russian Federation), KSRS (Wonju, South Korea), and MJAR (Matsushiro, Japan). These three stations have very different geometries and signal coherence attributes; the IMS network in the region provides a useful laboratory for exploring the performance of processing algorithms.

In Fig. 1 is marked a single earthquake, offshore to the South of Honshu, Japan, with magnitude 4.6. Throughout this paper, we will use recordings of signals from this earthquake on the three seismic arrays to demonstrate generic properties of the stations for signal processing relevant to CTBT monitoring. The algorithms presented in this paper were tested on many different seismic events but, for demonstration purposes, we chose a single event which fulfilled a number of requirements. First, the magnitude is of relevance for low to medium yield nuclear explosions. Second, the event was recorded well by all three arrays without data outages or faults. Third, it is part of a continuous band of seismicity such that we can observe the degree to which detection statistics apply to nearby events.

In Section 2, we pay special attention to the performance of these three stations using classical array processing: frequency-wavenumber or f - k analysis. How does direction estimation perform for signals in different frequency bands on the three different arrays, and under which circumstances do classical procedures break down? Which consequences might this have for CTBT monitoring in the region?

Harris and Kværna (2010) demonstrated how empirical matched field processing (EMFP), a generalization of f - k analysis, could be applied to classifying signals on seismic arrays. In particular, it was demonstrated how EMFP could differentiate between sources of explosions far closer to each other than could be resolved by f - k analysis: both in theory and practice. The scattering of seismic energy along the path, which degrades the performance of

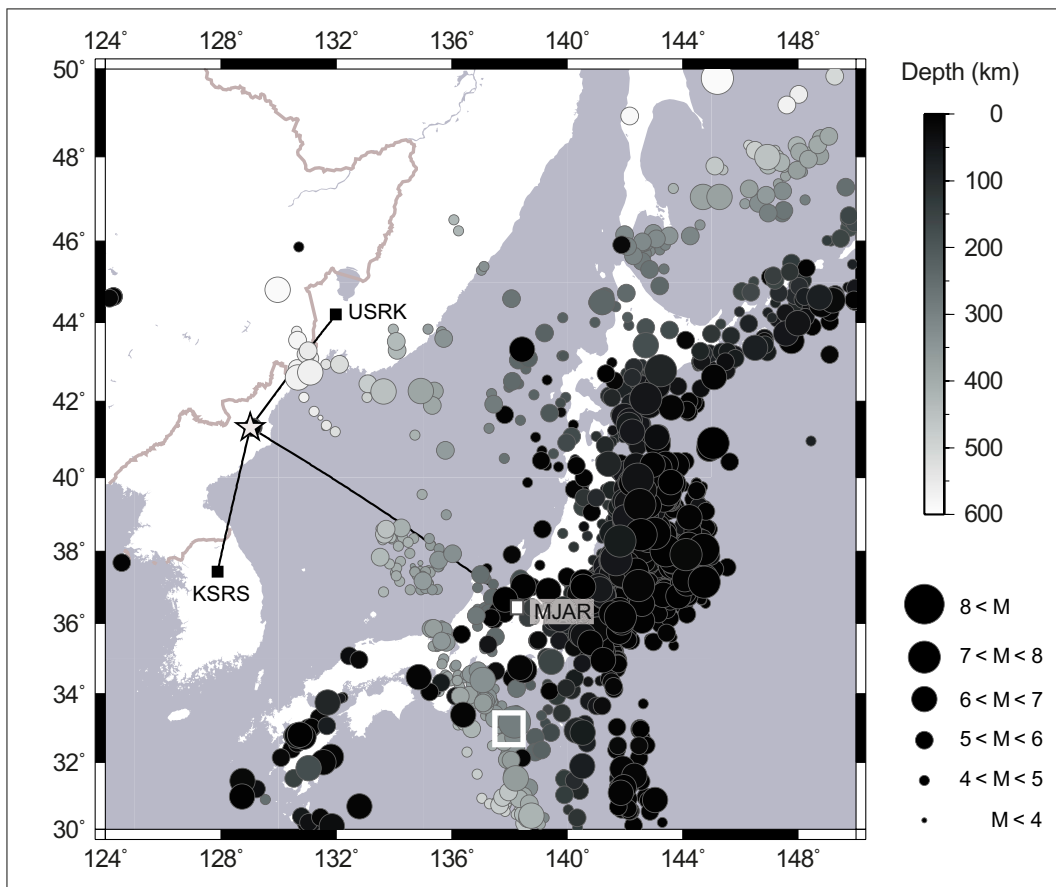


Fig. 1 Seismicity surrounding the Sea of Japan in the years 2008–2016 taken from the Reviewed Event Bulletin (REB) of the International Data Center (IDC) for the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO). Only events with REB phase readings from each of the three primary seismic arrays USRK, KSRS, and MJAR are displayed. The star

indicates the location of the North Korean Punggye-ri nuclear test site with great circle lines to the 3 arrays in the region. The white open square indicates the location of a magnitude 4.6 earthquake with REB location 32.67°N, 138.12°E and depth 323 km at time 07:42:52 UT on January 2, 2009

plane-wavefront f - k analysis, deforms the incoming wavefield and leaves a narrowband spatial seismic fingerprint on the observing array which is recognized and exploited by EMFP. Gibbons et al. (2017) demonstrated how EMFP could be applied to obtaining unbiased estimates of backazimuth and apparent velocity on a small aperture seismic array by comparing incoming wavefronts with a bank of EMFP templates constructed from previous observations. It appears that structure close to the SPITS array deflects the incoming wavefront, resulting in large biases in these parameters which need correcting prior to phase association and event location. Mapping the EMFP templates to sources in geographical space can provide indirect estimates of these parameters that

are in principle unbiased. In Section 3, we apply the procedure described in Gibbons et al. (2017) to the USRK, KSRS, and MJAR arrays and evaluate its performance on arrays of increasing aperture. We investigate how EMFP could mitigate problems associated with classical array processing.

In Section 4, we consider the special case of the MJAR array. The scattering of high-frequency seismic energy at this station has demonstrably proved problematic in the detection and correct classification of arrivals from events at regional distances. In addition to EMFP, we explore a refined incoherent signal processing procedure to make more robust seismic phase identification.

While the main focus of this paper is improving the reliability of estimates for the traditional parameters used in the event-building seismic processing pipeline, we note that the geographical parameter space which forms the basis for EMFP may point to fundamentally different ways of interpreting seismic signals. Could EMFP facilitate a more general form of seismic pipeline processing? We conclude with some perspectives for future development.

2 Classical array processing on IMS arrays in the Sea of Japan region

Figure 2 shows the geometries of the three arrays KSRS, USRK, and MJAR, drawn to the same scale. While USRK is typical of the small to medium aperture 9-element arrays developed especially for the IMS, with an aperture of around 4 km, the innermost elements are exceptionally closely spaced. KSRS and MJAR are medium aperture arrays with diameters of around 12 km. However, as Gibbons and Ringdal (2012) demonstrate, the signal coherence properties at KSRS and MJAR are very different, in particular for regional signals. Alongside each map are broadband frequency-wavenumber spectra in two different frequency bands for the first arriving P-wave at the stations from the earthquake marked in Fig. 1.

Each of the slowness estimate panels in Fig. 2 displays a projection of the incoming seismic energy into a 2-dimensional parameter space: (s_x, s_y) . It is assumed that the seismic signal propagates as a planar wavefront with the time of arrival, t_i , at seismometer site i given explicitly as

$$t_i = t_0 - \mathbf{r}_i \cdot \mathbf{s} \tag{1}$$

where $\mathbf{s} = (s_x, s_y)$ is called the slowness vector. The magnitude of the slowness vector,

$s_h = \sqrt{s_x^2 + s_y^2}$, is the inverse of the apparent velocity: $v_{app} = 1/s_h$. The scalars s_x and s_y are related to the backazimuth, Θ , with $s_x = s_h \sin(\Theta)$ and $s_y = s_h \cos(\Theta)$. t_0 is the time at which the wavefront passes over the reference site with absolute location \mathbf{R}_0 . The coordinates of instrument i at the absolute location \mathbf{R}_i are specified relative to \mathbf{R}_0 with

$$\mathbf{r}_i = \mathbf{R}_i - \mathbf{R}_0. \tag{2}$$

The slowness components s_x and s_y have units s/km, and the eastward and northward components of the relative position vector \mathbf{r}_i have units of km.

Both the relative position vectors \mathbf{r}_i and the slowness vectors \mathbf{s} have vertical components r_z and s_z although these are usually neglected on the grounds that the differences in elevation between sensors in the array is typically significantly smaller than the horizontal distances between sensors.

There are many ways of evaluating which slowness vectors, \mathbf{s} , best describe a given incoming wavefront (e.g. Rost and Thomas 2009; Ruigrok et al. 2017). If N is the number of sites in our array then we can define for a single frequency ω , and slowness vector \mathbf{s} , a so-called steering vector of N complex numbers given by

$$\boldsymbol{\varepsilon}(\omega, \mathbf{s}) = \left[e^{-i\omega \mathbf{s} \cdot \mathbf{r}_1} \dots e^{-i\omega \mathbf{s} \cdot \mathbf{r}_i} \dots e^{-i\omega \mathbf{s} \cdot \mathbf{r}_N} \right]^T. \tag{3}$$

The energy incident on an array with this particular single-frequency plane wavefront hypothesis can be expressed by

$$P(\omega, \mathbf{s}) = \boldsymbol{\varepsilon}(\omega, \mathbf{s})^H \mathbf{C}(\omega) \boldsymbol{\varepsilon}(\omega, \mathbf{s}) \tag{4}$$

where $\mathbf{C}(\omega)$ is the covariance matrix, evaluated on the waveform data, and H the Hermitian transpose operator. The covariance matrix has dimensions N by N with element $C_{lm}(\omega)$ defining the relationship between $v_l(t)$ and $v_m(t)$, the time-series belonging to channels l and m , with

$$C_{lm}(\omega) = \int d\tau \chi_{lm}(\tau) e^{-i\omega\tau} \tag{5}$$

where $\chi_{lm}(\tau)$ is the correlation between the two waveforms with

$$\chi_{lm}(\tau) = \int d\gamma w(\gamma) w(\gamma + \tau) v_l(\gamma) v_m(\gamma + \tau) \tag{6}$$

for a windowing (boxcar) function $w(t)$ which isolates the data-window of interest.

A challenge in estimating the narrowband covariance matrices from real-world seismic array data is the transience of coherent seismic signals with a given phase velocity, meaning that the phase shifts often need to be evaluated on very short data windows. The multitaper coherence routines of Prieto et al. (2009) have been found to be particularly robust for making accurate estimates of $\mathbf{C}(\omega)$. For increased

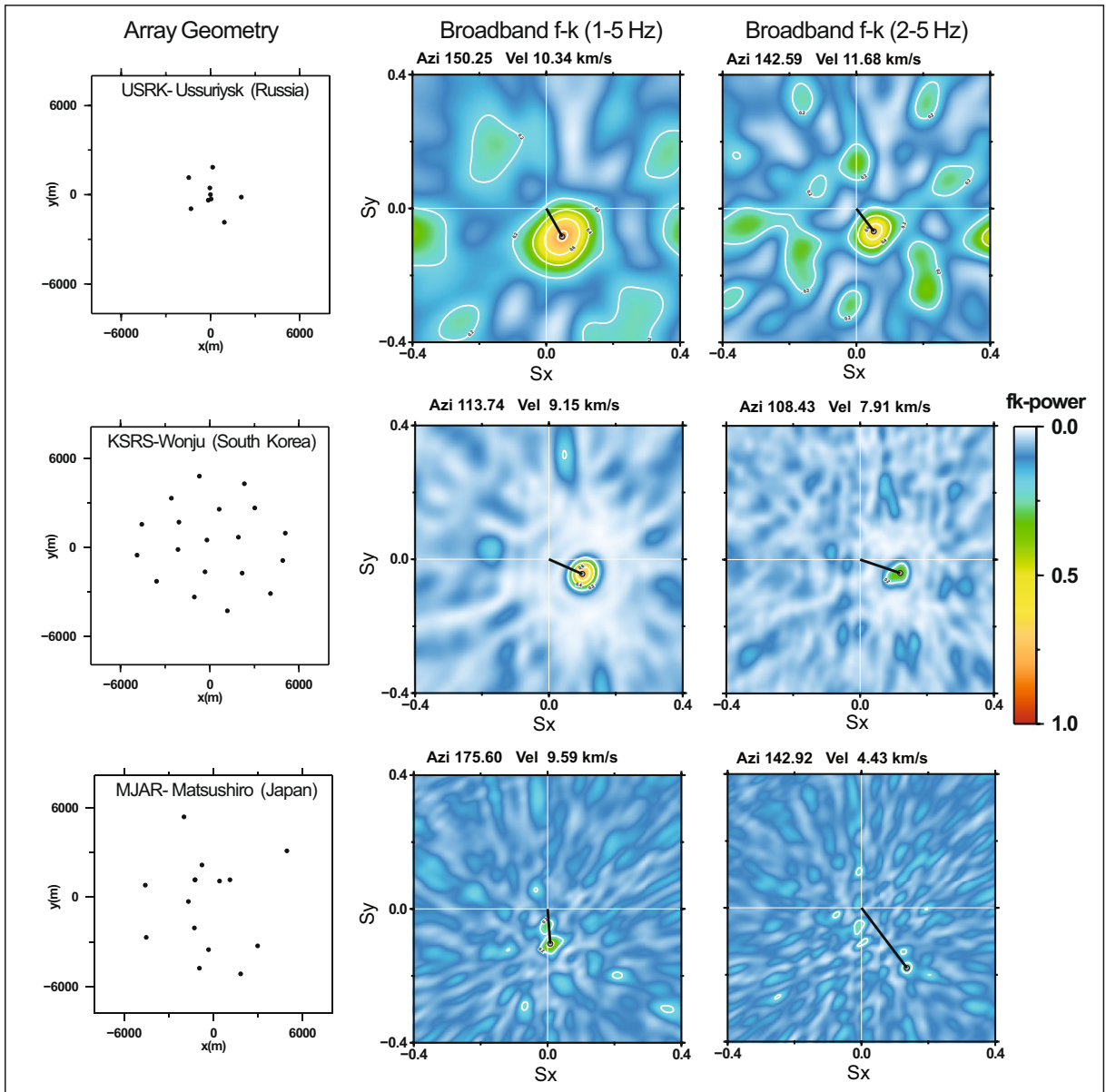


Fig. 2 For each of the primary seismic array stations USRK, KSRS, and MJAR, we display the array geometry (left), and broadband f-k estimates for the first arriving P-wave from the January 2, 2009, earthquake in the 1–5 Hz band (center) and

the 2–5 Hz band (right). Each f-k estimate was made using a 4-s-long data segment starting at the times 07:45:38.85 (USRK), 07:45:07.35 (KSRS), and 07:43:56.675 (MJAR)

stability, and reduced sensitivity to the spectral content of the signal and noise, we typically consider a relative beam-power summed incoherently over a range of distinct frequencies, k , giving a broadband estimator

$$\hat{P}(s) = \frac{\sum_k \mathbf{e}(\omega_k, s)^H \mathbf{C}(\omega_k) \mathbf{e}(\omega_k, s)}{\sum_k \text{tr} \{ \mathbf{C}(\omega_k) \}} \quad (7)$$

Evaluating this quantity over a dense grid of slowness vectors, s , provides a plot which gives an analyst an intuitive indication of the azimuthal direction a wave came from, how fast it appears to propagate over the array (identifying it as, e.g. teleseismic P, regional P, regional S), and the confidence with which we can attach to the measurement.

The frequency bands considered in Fig. 2 are 1–5 Hz and 2–5 Hz.

These two bands were chosen as they differ only by the energy below 2 Hz which is usually present for medium to large seismic events and largely absent in the signals from lower-magnitude events in the study region, as observed at the three IMS arrays considered.

Comparing the 1–5 Hz and 2–5 Hz panels for each of the stations in Fig. 2 demonstrates how significant this low-frequency energy is in generating a robust and plausible direction estimate using f-k analysis. For the arrays USRK and KSRS, the slowness estimate for this regional signal is qualitatively identical for the two frequency bands; that is to say that we obtain a similar backazimuth (direction) and apparent velocity and the same identification of phase type (a Pn-type regional arrival). For both arrays however, the relative f-k power (array gain) is lower for the 2–5 Hz band, with more significant sidelobes and greater asymmetry in the f-k pattern.

For the MJAR array, a qualitatively different slowness estimate results from the 2–5 Hz band. Whereas the 1–5 Hz band indicates a Pn arrival from the south, the 2–5 Hz band has a maximum at one of the sidelobes and returns an S-wave velocity from the south-east. The estimate without the 1–2 Hz frequency energy has generated a qualitatively misleading analysis of this signal arrival and this would affect all subsequent stages of the phase association and event generation procedure. The MJAR example displayed in Fig. 2 is typical. Gibbons et al. (2008) show a similar phenomenon for the Pn arrival from the October 8, 2006, declared DPRK nuclear test. In that case, there is little energy above the noise level between 1 and 2 Hz. For the first four declared North Korean nuclear tests (October 8, 2006, May 25, 2009, February 12, 2013, and January 6, 2016) the MJAR array did not contribute to the fully automatic event hypotheses issued by the CTBT International Data Center in their Standard Event List 3 (SEL3). All the explosions were well recorded by the array but the parameter estimation capability did not result in estimates which could be associated with the event hypotheses. (The larger explosions on September 9, 2016, and September 3, 2017, did result in signals with accurate slowness estimates for the MJAR Pn arrivals and these do contribute to the SEL3 reports for these events).

An analyst familiar with operational processing on a given network of seismic arrays will develop an intuition for how the slowness plots appear for different types of signals, in different frequency bands. Which frequency bands provide robust estimates on a given array station? Examining the slowness grids together with the array geometries (c.f. Fig. 2) provides a level of insight into the properties of the slowness grids. However, significant bias in estimates, or low-quality estimates resulting from signal incoherence, is only learned empirically.

3 Empirical matched field processing: projecting wavefield measurements into geographical space

Harris and Kværna (2010) demonstrated an empirical form of array signal processing, empirical matched field processing or EMFP, that could identify the origin of seismic signals generated by sources far closer to each other than the array should be able to resolve under classical considerations. The sources were mines, located within a few kilometres of each other, situated over 400 km away from the ARCES small aperture seismic array. A ground truth database of signals from explosions known to have taken place at each of the mines was used to build up a calibration set of waveforms from which measurements could be used to classify subsequent signals for which the source was not known.

Instead of testing a hypothesis that an incoming wavefront matches a set of phase shifts corresponding to a plane wavefront with slowness vector \mathbf{s} , we test the hypothesis that the incoming wavefront matches a set of phase shifts consistent with previous observations from a given source. We can arbitrarily denote an arrival from a source of interest α . This may, for example, refer to a Pn phase from a given site, an Sn, or Lg, phase from the same site, or first arrivals from all seismic events within a very limited geographical footprint.

Specifically, we replace the expression in Eq. 7 with

$$\hat{P}(\alpha) = \frac{\sum_k \mathbf{e}(\omega_k, \alpha)^H \mathbf{C}(\omega_k) \mathbf{e}(\omega_k, \alpha)}{\sum_k \text{tr} \{ \mathbf{C}(\omega_k) \}} \quad (8)$$

where $\mathbf{e}(\omega_k, \alpha)$ are the empirical steering vectors corresponding to the arrivals with the classification α and are the eigenvectors of the covariance matrices $\mathbf{C}(\omega_k, \alpha)$ measured from the calibration dataset. In the

source classification problem addressed by Harris and Kværna (2010), a number of alternative hypotheses, α_m , were tested where the indices m indicate different mines. The empirical steering vectors, $\boldsymbol{\varepsilon}(\omega_k, \alpha_m)$ were generated from Pn arrivals from Ground Truth explosions that took place at mine m . For each signal of unknown origin in the test dataset, the mine of origin was assumed to be the one for which the operator in Eq. 8 was largest. This classification is conceptually no different from the broadband slowness grids displayed in Fig. 2. Each grid simply displays evaluations of $N_s \times N_s$ alternative hypotheses of wave propagation using different plane-wave rather than empirical models. In almost all cases, the correct source, m , was selected. Differences in the scattering of the seismic wavefield along the marginally different pathways from each source were sufficient to give a slightly better match with the steering vectors from the true source than with the alternative hypotheses.

In principle, an empirical steering vector $\boldsymbol{\varepsilon}(\omega_k, \alpha)$ can be obtained from a covariance matrix $\mathbf{C}(\omega_k, \alpha)$ evaluated for a single segment of array data, from the signals generated by one event. However, elements of this covariance matrix may be unrepresentative of the population of events from that source: possibly due to noise or imperfections in the data. In practice, where possible, we estimate an ensemble covariance matrix $\overline{\mathbf{C}}(\omega_k, \alpha)$ averaged over a representative sample of signals from events at the source of interest:

$$\overline{\mathbf{C}}(\omega_k, \alpha) = \frac{\sum_{\epsilon \in \alpha} \mathbf{C}(\omega_k, \alpha_\epsilon)}{\sum_{\epsilon \in \alpha} 1} \quad (9)$$

with care being taken to ensure that all elements of covariance matrices are present for all signals. If data is missing from one or more channels of the array for a given event, the corresponding rows and columns of the covariance matrix will be zero. If it is unavoidable to use events, ϵ , for which there is missing data then the denominator in Eq. 9 will need to be adjusted for each element of the covariance matrix to consider the true number of contributions to that element.

In classical plane-wave f-k analysis, we typically choose a single component of the wavefield, most often the vertical component. This is because the propagation model assumes that the waveforms are identical on all channels except for the time-delays between the signals at the different sites. In EMFP, for 3-component seismic arrays, we can in principle include all components of the wavefield such that $v_l(t)$

and $v_m(t)$ do not necessarily need to measure the same component of ground motion. The elements of the covariance matrices are just estimates of the coherence and phase differences between pairs of channels without any underlying assumptions about the relation between the signals; the steering vectors are simply the eigenvectors of these covariance matrices. However, the contributions to the covariance matrices from the cross-component pairs of channels are likely to be far less significant than the contributions from channel pairs with the same component of ground motion. Further work will be required to explore the performance gain possible from including multiple components in the same system on seismic arrays where this is possible.

In this paper, we consider only vertical component sensors. In fact, we have no choice; all of the arrays USRK, KSRS, and MJAR consist of vertical-only seismometer sites, with the exception of a lone 3-component station in each of the arrays.

Before we construct the ensemble covariance matrix, we should test the internal consistency of the elements contributing to it. Ringdal et al. (2009) demonstrate an attempt to classify seismic signals from two very closely spaced mines using two competing empirical matched field steering vectors. The classification was relatively unsuccessful, with several of the events at one of the mines being incorrectly attributed to the other mine. A cluster analysis of the empirical steering vectors from individual calibration events for one of the mines showed two distinct populations with rather different sets of phase shifts between the narrowband signals on different sensors. Mines and quarries are seismically highly heterogeneous with voids and free surfaces which may result in different radiation for different sequences of shots within the same complex (e.g. McLaughlin et al. 2004). Forming two different ensemble covariance matrices from each of the populations of events at this mine allowed for a far more robust and accurate signal classification.

In the applications described above, the primary focus was source identification. EMFP was a significantly better signal classifier than waveform correlation. One reason for the poor performance of correlation is that the explosions are ripple-fired sequences, resulting in very different temporal forms of the signals. The estimates of the narrow-band coherence and phase-shifts are relatively insensitive to the temporal

form of the signal and the elements of the covariance matrices, and corresponding elements of the eigenvectors, are relatively stable characteristics of the event populations.

Gibbons et al. (2017) describe a slightly different application of EMFP. There, the aim was to provide parameter estimates for incoming wavefronts at a seismic array which provide a best possible indication of the direction the signal comes from. In other words, they wanted to provide a visual impression, analogous to the slowness grids in Fig. 2, which helps an analyst to interpret the incoming signal but which is not limited by the limitations of the plane-wave model. Gibbons et al. (2017) focused on the very small aperture (1-km diameter) SPITS array on the Arctic island of Spitsbergen. Due to the small aperture, signal coherence is relatively high for the frequencies of interest in regional seismic monitoring. However, likely due to structure very close to the array, there is often significant bias in both the backazimuth and apparent velocities measured (see Gibbons et al. 2011). It was demonstrated that a geographical labelling of empirical steering vectors, where each classification α considered is associated with the epicenter of the earthquake which generated the calibration signal, allowed an indirect estimate of the backazimuth and slowness that was not significantly biased.

The three primary IMS seismic arrays surrounding the Sea of Japan are very different from the SPITS array studied by Gibbons et al. (2017). The array apertures are far larger and we have already seen that signal incoherence can generate difficulties. How applicable is this form of EMFP to signal characterization to the Asian seismic arrays? Figure 3 displays the waveforms of the signal from the January 2, 2009, earthquake on each of the elements of the USRK array. The slowness grids in the top row of Fig. 2 were calculated using the data near the signal arrival from exactly these signals. The signal at USRK from this earthquake is impulsive and has a high SNR. The central panel of Fig. 3 displays the coherence and phase-shifts between each pair of sensors measured for two discrete frequencies: 1.25 Hz and 3.75 Hz. If \mathbf{r}_l and \mathbf{r}_m denote the relative (Cartesian) position vectors of sites l and m then the coherence and phase shift between channels m and l is displayed at the location $\mathbf{r}_m - \mathbf{r}_l$. The set of these locations is called the co-array and displaying the phase relationships superimposed onto this geometry (as opposed to a simple N by N matrix)

allows us to see if time-shifts between sensors are associated with a consistent direction. Since the size of the symbols is proportional to the coherence between the sensors, we also see how the waveform similarity diminishes with distance between the sensors.

The relative sparsity of the USRK array means that there are significant gaps in the co-array, and coherent spatial patterns of phase-shifts are difficult to discern. To the right of Fig. 3, we display how these phase shifts would appear if the wave had propagated across the array according to the plane-wave assumption. All the symbols to the right are sized to indicate perfect coherence. The colours and the superimposed dashed lines to the right provide the impression of a wave approaching from the correct direction with alternating red and blue bands cutting across the co-array. The 1.25 Hz evaluation contains essentially a single cycle while the 3.75 Hz panel indicates four or so cycles across the co-array aperture. The differences between the empirical and theoretical phase shifts indicates the degree to which the incoming seismic wavefield deviates from the plane-wave model.

Figure 4 (right) displays a magnification of a short time-window at the onset of the signal displayed in Fig. 3. Here, we see directly not only the moveout of the signals across the array but also how the temporal form of the signal varies from sensor to sensor. The human eye can perform a visual cluster analysis on these signals and discern clear groupings of similar waveforms. Differences in the shape of the signals on the different sensors explain the less than perfect coherence values (size of symbols in Fig. 3). Differences in the colour of symbols in Fig. 3) along the superimposed theoretical wavefronts indicate that the time-delays measured between the pairs of sensors at the indicated frequencies is different to that predicted by the plane-wave propagation model. To the left of Fig. 4, we display the “matched field statistics” as formulated in Eq. 8 evaluated for the covariance matrix measured at the time 07:45:38.850 GMT on January 2, 2009, and an empirical steering vector $\mathbf{e}(\omega_k, \alpha)$ for the first arriving P-wave for each of the seismic events displayed.

A single steering vector is calculated for each of the events displayed; in other words, we have not attempted to calculate an ensemble covariance matrix as discussed earlier. Each covariance matrix was evaluated over the nine vertical channels of the array for a data window with length 256 samples (6.4 s)

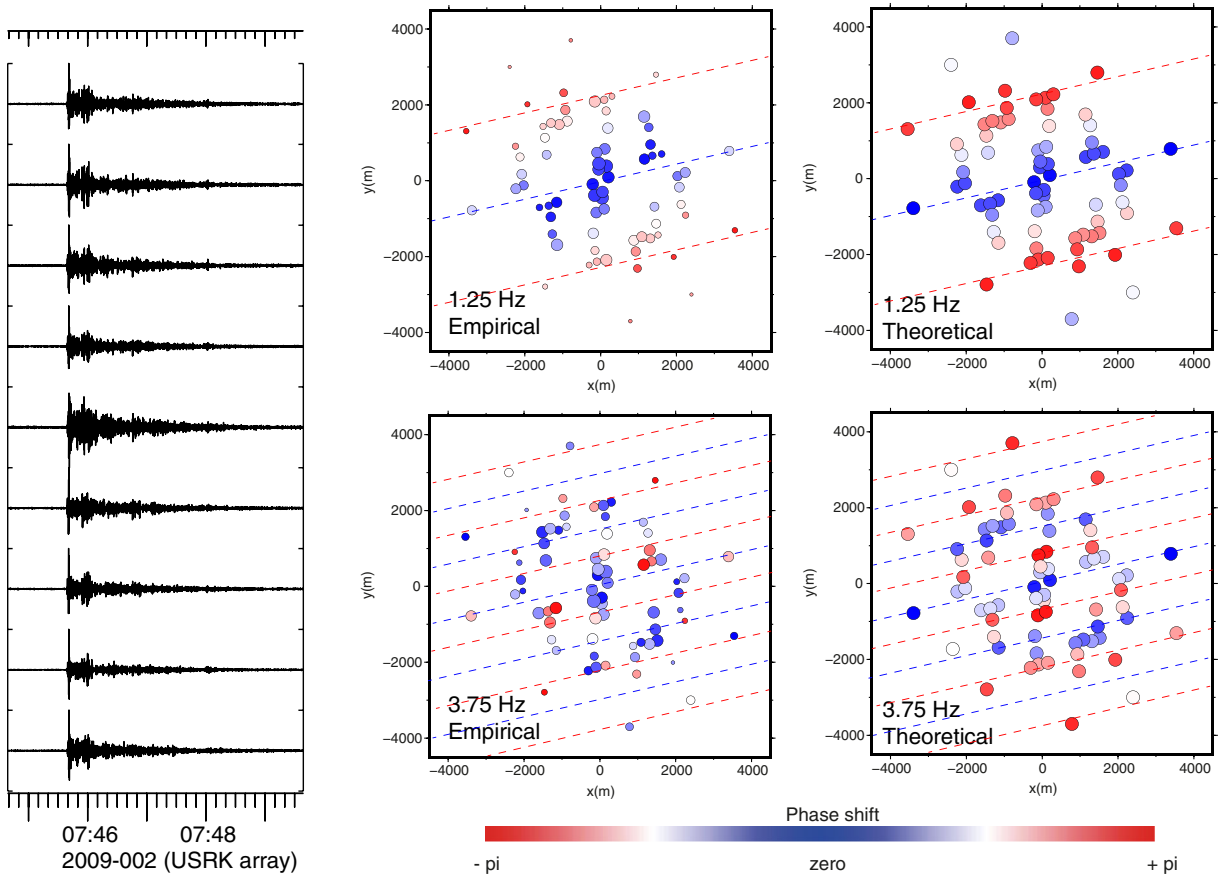


Fig. 3 Signals at USRK SHZ array sensors (trace normalized) for the January 2, 2009, earthquake (left) and bubble-plot representations of the signal coherence and phase-shifts between sensor pairs (right). The symbols are scaled linearly with coherence and the largest symbols displayed have coherence 1.0. The dashed blue and red lines represent wavefronts of the propagating plane wave with the predicted backazimuth and

apparent velocity for the frequencies indicated. Both theoretical (plane-wave model) and measured phase shifts are displayed for narrow frequency bands centered on the frequencies as indicated (details in text). Measurements are made on a 6.4-s-long data segment starting at a time 07:45:38.850. A zoom-in on the waveforms within the measurement window is shown in the right-hand panel of Fig. 4

and frequency bands between 1.09 and 5.00 Hz. The matched field statistics are uniformly weighted averages over this entire broad frequency band. We see the highest values of the matched field statistic are obtained for steering vectors calculated from seismic events close to the great circle path between the source and receiver. The matched field statistic diminishes the further the template event is from this fixed backazimuth path. It is noted that the events close to the USRK array in the same direction have relatively low matched field statistics. Most of these are deep earthquakes with the wavefield arriving from a far steeper angle than is the case for earthquakes in Japan.

There are clear parallels between the maps of geographically evaluated matched field statistics (Fig. 4) and the plane wavefront projections displayed in the uppermost panels of Fig. 2. Both types of plot indicate a direction from the array from which the wavefront has most likely arrived. The slowness plots cover densely theoretical angles of approach below the array, and the indicated direction may differ from the true direction due to heterogeneity and diffraction close to the array. The geographical estimates likely correct for bias in the direction estimate, but are sampled sparsely: only at coordinates at which previously observed seismic events have been located. This limitation almost guarantees the continuation

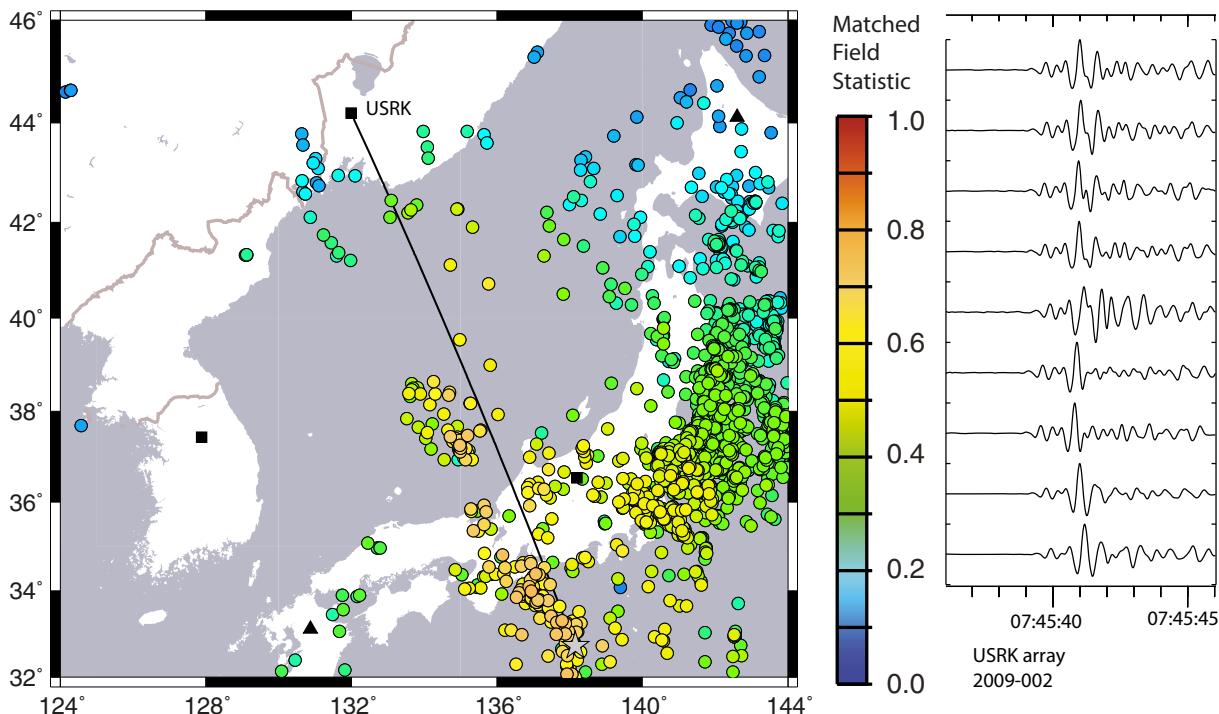


Fig. 4 Geographical matched field direction estimates on the USRK array, integrated over discrete frequencies between 1.09 and 5.0 Hz. Empirical matched field statistics calculated between the covariance matrix evaluated for the regional P arrival at the USRK array at a time 2009–002:07.45.38.850 and the steering vectors calculated from the covariance

matrices evaluated for the first P arrival at USRK for each of the events indicated. The star denotes the location of the earthquake with origin time 2009–002:07.42.51.34 and epicenter (33.72, 138.15). Station to source backazimuth is 153.4^{circ} at a distance 1279 km. To the right is a zoom-in of the signal arrival displayed in Fig. 3

of classical plane-wave f-k analysis for the study of incoming wavefields regardless of progress in empirical signal processing; a correct identification of a wavefront arriving from a location that no previous seismic events have been observed is fundamental to test-ban treaty monitoring. However, we identify a clear potential for returning enhanced parameter estimates for situations in which there exists an steering vector for an empirical source α for which

$$\widehat{P}(\alpha) > \widehat{P}(s) \tag{10}$$

for all theoretical slowness vectors s .

Figure 5 shows the signal from the 2009 earthquake recorded on the KSRS array together with the phase-relationships between the different sites displayed on the co-array. The panels of Fig. 5 are entirely analogous to the corresponding panels of Fig. 3, regarding the frequencies, the colour scheme, and the quantities

displayed. We note that the x - and y -axes of the “bubble plots” are different for the larger aperture KSRS array. The greater number of sites in the array (19 versus 9) and their spatial distribution results in a more uniform covering for elements of the co-array. The banded pattern of phase-shifts is far clearer for the KSRS array than for USRK. The empirically measured phase shifts for 1.25 Hz in Fig. 5 reproduce the pattern of bands from the theoretical steering vectors well. The diminishing signal coherence as a function of distance between sensors is clearer in Fig. 5 than in Fig. 3. For the higher frequency of 3.75 Hz, the banded pattern is difficult to discern for the empirical phase shifts and the coherence is universally lower. The fall-off of symbol size with distance from the center of the plot is less clear for the 3.75 Hz panel than for the 1.25 Hz panel, with some anomalously large symbols displayed for pairs of sensors with large intersite distances. This is a warning sign that we may be measuring rather coincidental waveform similarity

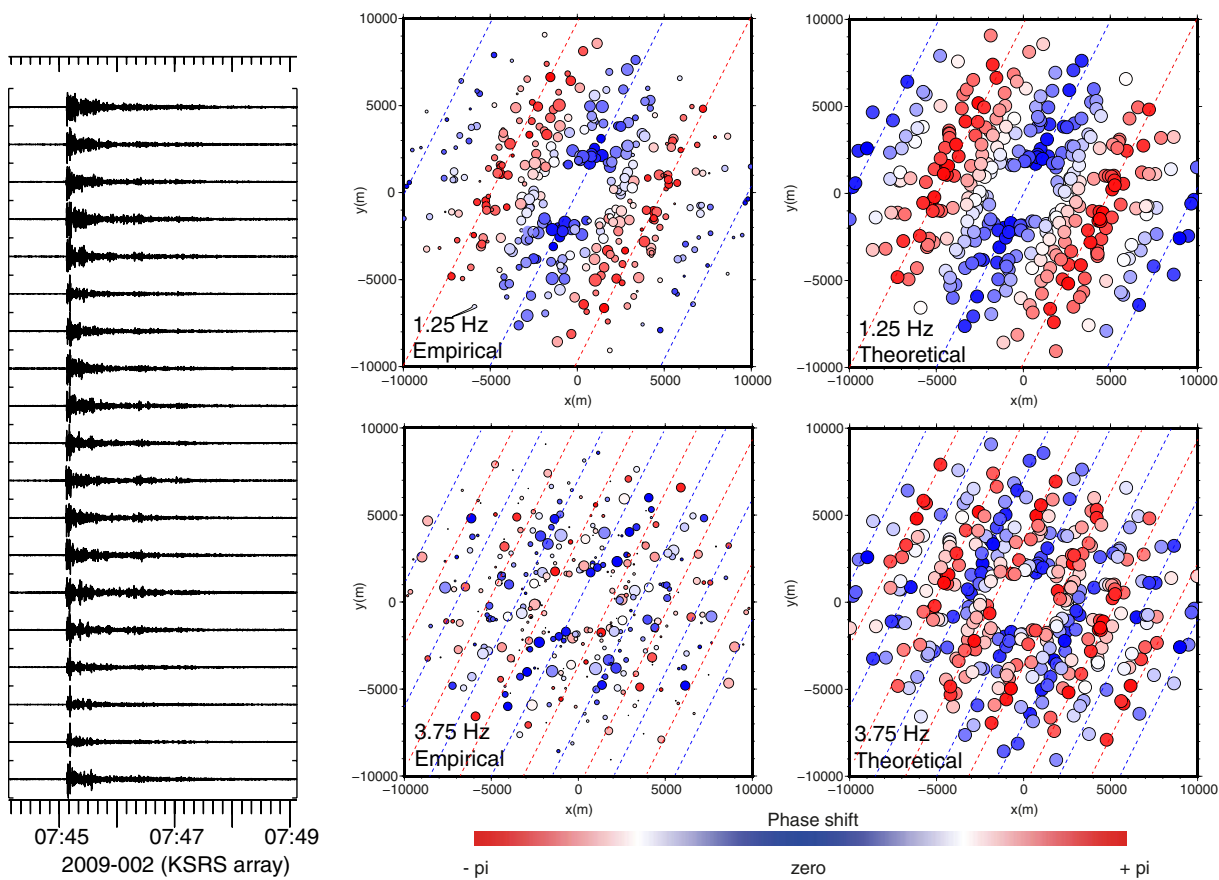


Fig. 5 Signals at the KSRS SHZ array sensors (trace normalized) for the January 2, 2009, earthquake (left) and bubble-plot representations of the signal coherence and phase-shifts between sensor pairs (right). The symbols are scaled linearly with coherence and the largest symbols displayed have coherence 1.0. The dashed blue and red lines represent wavefronts of the propagating plane wave with the predicted backazimuth and

apparent velocity for the frequencies indicated. Both theoretical (plane-wave model) and measured phase shifts are displayed for narrow frequency bands centered on the frequencies as indicated (details in text). Measurements are made on a 6.4-s-long data segment starting at a time 07:45:07.350. A zoom-in on the waveforms within the measurement window is shown in the right-hand panel of Fig. 6

that may not be characteristic of general wavefronts from that particular source region.

Figure 6 displays a close-up of the first arrival on KSRS from this earthquake. We see a larger moveout of the signal onsets (with over a second between first and last arrivals) and, again, a clear spatial clustering of the signal shapes on sensors in different parts of the array. Comparing the spatial distribution of matched field statistics in Fig. 6 with that in Fig. 4 is similar to comparing the slowness grids in the first and second rows of Fig. 2. Just as the peak in slowness space was smaller and more compact for the larger KSRS array than the smaller USRK array, the earthquake epicenters which display an elevated matched field (MF) statistic on KSRS form a much thinner band

surrounding the great circle path between the source and receiver than the corresponding epicenters in the USRK plot. The values of the MF statistics are generally lower for KSRS than for USRK. On the smaller aperture USRK array, the signal moveouts are smaller and earthquakes from a far larger region result in similar temporal shifts. The number of degrees of freedom and the size of the moveouts on KSRS are greater; it is more difficult to match the moveouts/phase-shifts if the sources are not in very similar locations.

Figure 7 shows the signals from the same earthquake on the MJAR array. At just over 300 km, this array is considerably closer to the source than the other two arrays. Since the backazimuth from station to the source is 181^{circ} (almost due South), the red and blue

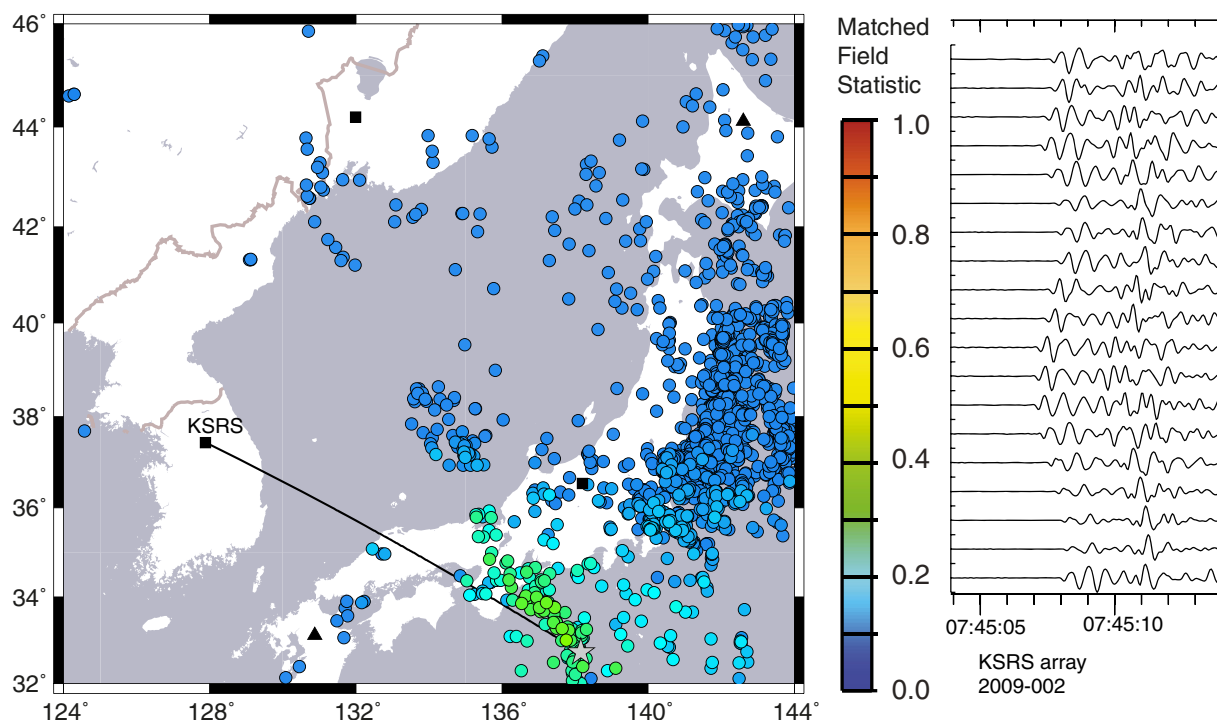


Fig. 6 Geographical matched field direction estimates on the KSRS array, integrated over discrete frequencies between 1.09 and 5.0 Hz. Empirical matched field statistics calculated between the covariance matrix evaluated for the regional P arrival at the KSRS array at a time 2009–002:07.45.07.350 and the steering vectors calculated from the covariance

matrices evaluated for the first P arrival at KSRS for each of the events indicated. The star denotes the location of the earthquake with origin time 2009–002:07.42.51.34 and epicenter (33.72, 138.15). Station to source backazimuth is 110.9^{circ} at a distance 1017 km. To the right is a zoom-in of the signal arrival displayed in Fig. 5

bands are approximately horizontal. The empirical coherence and phase-shifts measurements at 1.25 Hz indicate a clear blue band for sensors with the shortest inter-site distances in the North-South direction. However, the size of the symbols diminishes rapidly for greater inter-site distances: a clear indicator of lacking signal coherence. For the 3.75 Hz panel, there is even less of clear correspondence between symbol colour and the colour of the superimposed dashed lines.

Given these coherence properties at the array, we should not be surprised by the geographical distribution of the MF statistics in Fig. 8. A preliminary examination of the waveforms themselves emphasize the signal dissimilarity which makes coherent processing of high-frequency signals on this array so problematic. While waveforms on selected subsets of neighbouring stations show similarity, the overall lack of coherent wavelet to align would mean that a human analyst would struggle to align signals visually. The only earthquake epicenters for which the empirical

MF statistics are significantly above the background level are to the South of the station. However, unlike for the USRK and KSRS arrays, the elevated MF statistics do not form a clear band surrounding the Great Circle path between source and receiver. An almost co-located earthquake should indeed generate similar phase shifts across the array to the reference event but, given the coherence patterns displayed in Fig. 7, the significance of these similarities may not be sufficient to make a robust direction indicator. As a “spotlight detector” (i.e. a detector which only triggers on events from the immediate vicinity of the reference event) we would need a very large number of reference event steering vectors to cover existing seismicity, with a more or less unknown capability for events some distance from reference events.

In this section, we have covered only a small range of the possibilities that EMFP provides for reduced-bias parameter estimation on array stations in the

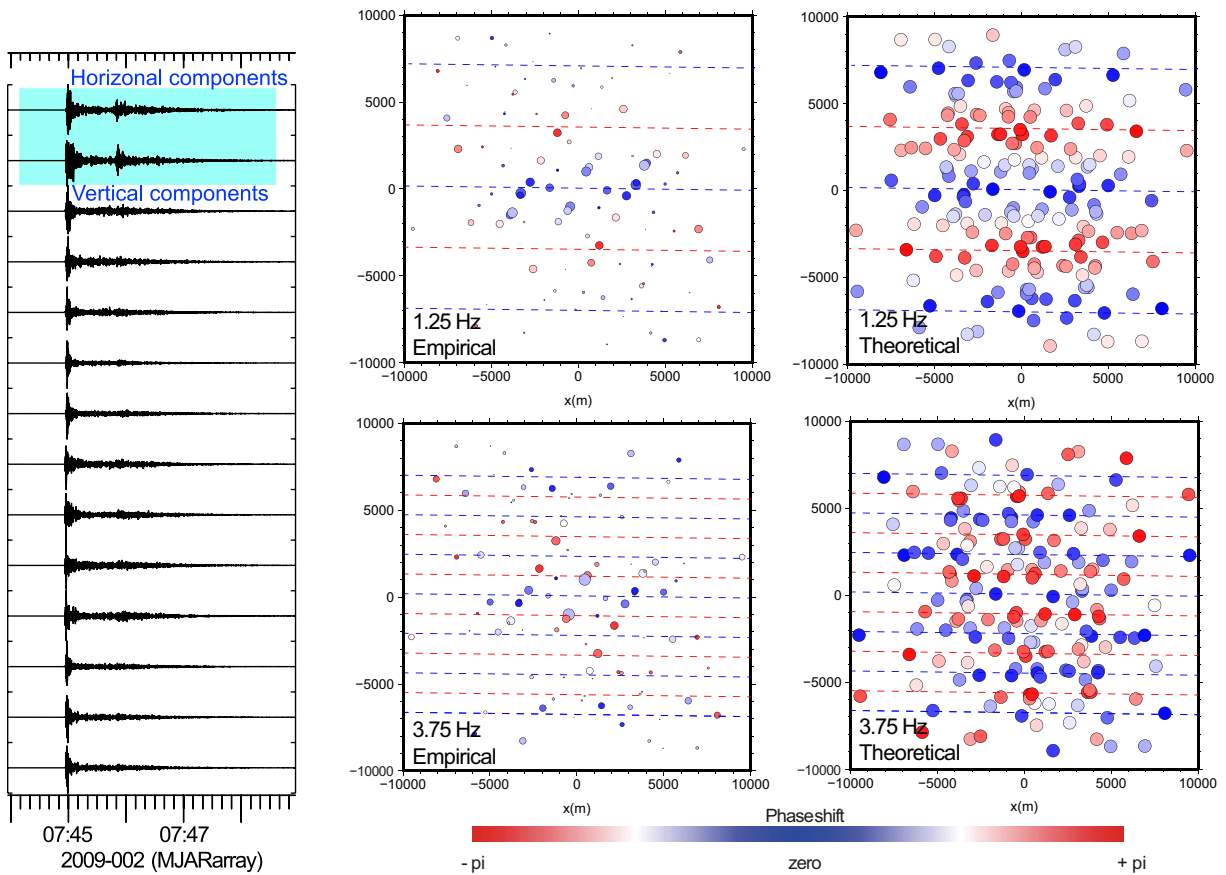


Fig. 7 Signals at the MJAR HHZ, HHN and HHE array sensors (trace normalized) for the January 2, 2009, earthquake (left) and bubble-plot representations of the signal coherence and phase-shifts between sensor pairs (right). The symbols are scaled linearly with coherence and the largest symbols displayed have coherence 1.0. The dashed blue and red lines represent wavefronts of the propagating plane wave with the predicted backazimuth and apparent velocity for the frequencies

indicated. Both theoretical (plane-wave model) and measured phase shifts are displayed for narrow frequency bands centered on the frequencies as indicated (details in text). Measurements are made on a 6.4 s long data segment starting at a time 07:43:56.675. A zoom-in on the waveforms within the measurement window is shown in the right-hand panel of Fig. 8. Note that traces from the three-component station are not used in the array processing

region around the Sea of Japan. For example, in considering only a wide-band estimator, we have not studied how the performance varies with frequency band. However, we have hopefully demonstrated the transportability of the matched field estimators from very small aperture arrays to significantly larger deployments and explained the associated limitations. The geographical parameter spaces displayed in Figs. 4, 6, and 8 are potentially powerful tools for advanced phase association algorithms when these projections are combined for different stations with complementary parameters such as arrival times.

4 Incoherent processing for high-frequency signals at MJAR

There are numerous situations where we would ideally perform coherent array processing to estimate propagation parameters of seismic wavefronts, but are prevented from doing so because of signal dissimilarity. Gibbons et al. (2008) demonstrate how robust estimates for high-frequency regional seismic signals can be obtained for a large aperture seismic array by calculating and transforming spectrograms prior to delay-and-stack beamforming. Similar procedures

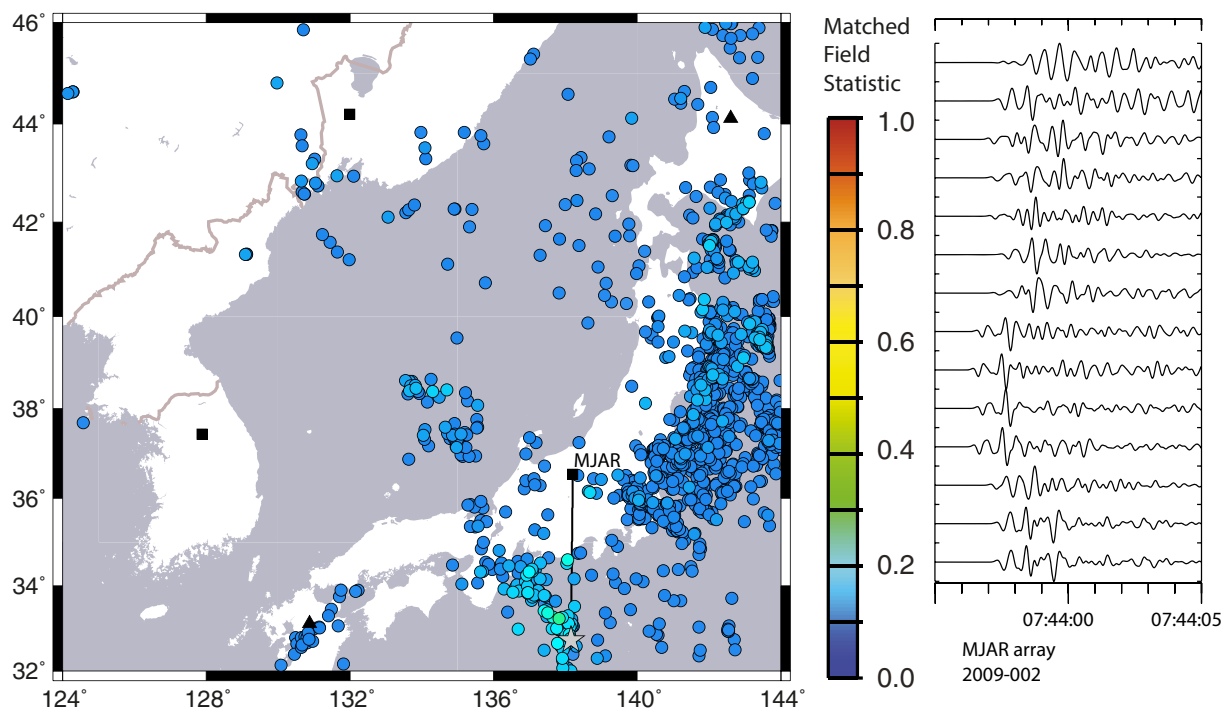


Fig. 8 Geographical matched field direction estimates on the MJAR array, integrated over discrete frequencies between 1.09 and 5.0 Hz. Empirical matched field statistics calculated between the covariance matrix evaluated for the regional P arrival at the MJAR array at a time 2009–002:07.43.56.675 and the steering vectors calculated from the covariance

matrices evaluated for the first P arrival at MJAR for each of the events indicated. The star denotes the location of the earthquake with origin time 2009–002:07.42.51.34 and epicenter (33.72, 138.15). Station to source backazimuth is 181.0^{circ} at a distance 313 km. To the right is a zoom-in of the signal arrival displayed in Fig. 7

were later applied to the detection and classification of ground-coupled airwaves from volcanoes over sparse seismic networks (Fee et al. 2016) and to parameter estimation for seismic signals on large Ocean Bottom Seismometer networks (Krüger et al. 2020). In each of these studies, the incoherent method is both possible and effective due to the relatively large move-outs of the signal arrivals over the large aperture arrays/networks.

The problem of coherent estimation of regional phases on MJAR was addressed by Gibbons et al. (2008) and it was demonstrated that spectrogram beamforming provided an efficient signal detector and a parameter estimator that provided qualitatively correct slowness estimates more reliably than coherent f-k analysis. However, due to differences in the shapes of the transformed spectrograms, the incoherent slowness estimates displayed considerable spread. Gibbons (2014) demonstrated that this problem could be mitigated by smoothing the spectrogram beamgrids with an estimate of the array response function for a

wavefront with a period appropriate for the temporal scale of spectrogram onsets.

In this section, we demonstrate a strategy for improving incoherent slowness estimates at MJAR using a slightly different approach. Instead of time-domain stacking of the incoherent detection statistic traces, Krüger et al. (2020) apply a bandpass filter in a suitably low-frequency band in order that coherent f-k analysis can be performed. This filtering captures those features of the detection statistic traces which are common to all sensors in the network and appears to provide more robust estimates than the smoothing operator proposed by Gibbons (2014). The Auto-Regressive Akaike Information Criterion (AR-AIC) method for signal onset estimation (e.g. Leonard and Kennett (1999, 2000), and references therein) has become a workhorse for automatic refinement of phase arrival time, and may constitute a higher accuracy onset-time indicator than the transformed spectrograms. Gibbons et al. (2016) demonstrated a modification to the AR-AIC procedure which

produces a continuous detection statistic with an exceptionally robust arrival time. We demonstrate here a mini-workflow which combines these techniques in a new procedure for slowness estimation on MJAR.

The spectrogram beamforming procedure of Gibbons et al. (2008) provides an approximate signal onset time for each channel, together with an estimate of the frequency range in which the greatest SNR is to be found. Since we are not processing the waveforms coherently, we can consider far higher frequencies than we would if we were to insist on some degree of waveform similarity between sensors. Figure 9a displays the raw waveforms on the vertical component sites of MJAR for the 2009 earthquake. Panel b

displays the same waveforms following bandpass filtering in a frequency band deemed to be optimal for the SNR. There are no hard and fast rules as to which frequency band to use. One alternative is to estimate from spectrograms or a cascade of filtered seismograms, another is to specify a fixed band (e.g. 2–8 Hz, 3–12 Hz). Even over the 12-km aperture of MJAR, we see a moveout of the traces on the different sensors. We also see great differences between the temporal forms of the signals on different channels.

Figure 9c displays the continuous AR-AIC detection statistic demonstrated by Gibbons et al. (2016) for each of the channels in panel b. Unlike the spectrograms, which depend upon the rate of change

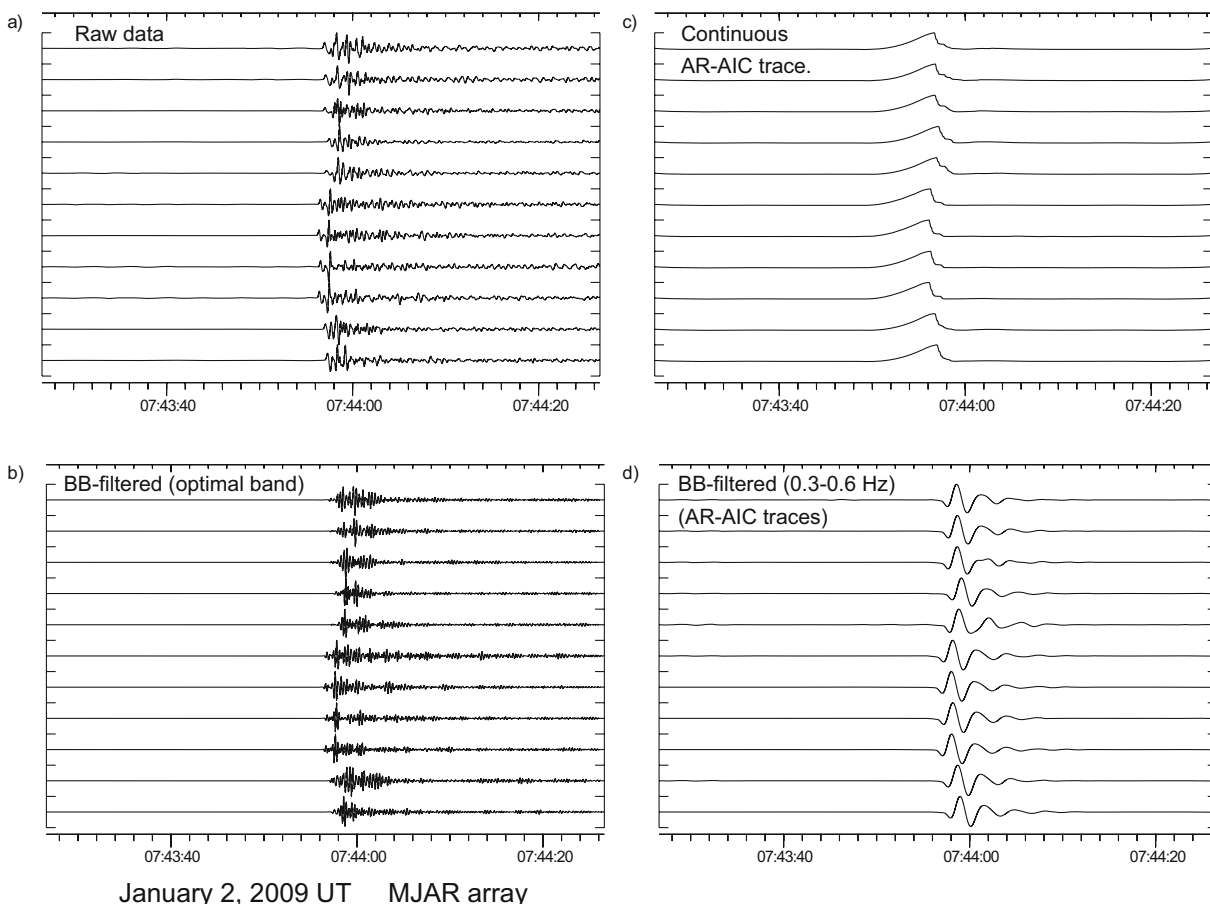


Fig. 9 Stages of preparation of MJAR array data for an incoherent slowness estimate of a high-frequency regional signal. **a** Raw waveforms from each of the vertical sensors of the MJAR array, **b** bandpass-filtered in an optimal frequency band, **c** calculation of continuous AR-AIC detection statistic traces using a uniform set of parameters, and **d** bandpass filtering of the traces

in panel c in the frequency band 0.3–0.6 Hz. The maxima of the filtered AR-AIC traces (panel d) are not to be interpreted as indications of the signal onset-times. The time-delays between the maxima in the unfiltered AR-AIC traces (panel c) determine the phase shifts between the wavelets of the traces in panel d from which the direction of arrival is estimated

of energy in the signal in the different frequency bands, the shapes of the continuous AR-AIC functions depend upon the variance of the error in the linear prediction traces for the individual traces—and are less sensitive to the signal amplitudes. While the times of the local maxima of two AR-AIC traces may be insufficiently accurate to estimate robust relative delay-times between the signal arrivals, the general shapes of the AR-AIC traces are remarkably consistent between sensors and are mostly dependent upon the parameters of the AR-AIC estimates (e.g. the length of the AR-models, the length of the data-window used for measuring the AR-coefficients, and the length of the linear prediction window). The effect of the fine detail of the signal onset, which determines the exact time of the AR-AIC local maximum, is removed by subjecting the AR-AIC function to a lower frequency bandpass (e.g. 0.3–0.6 Hz). This is essentially the procedure applied by Krüger et al. (2020) and the resulting traces are displayed in Fig. 9d). The bandpass filtered AR-AIC detection statistic traces appear as almost identical wavelets, between which the phase shifts appear to correspond very well with the delay-times between the signal arrivals.

In a final stage, we perform broadband f-k analysis on these bandpass filtered AR-AIC wavelets (Fig. 10). The f-k spectrum has a broad peak as would be anticipated for the frequency band selected. However, the

fall-off of relative power is well-behaved and we can have high confidence in the location of the peak in (S_x, S_y) -space. There is no danger of selecting a spurious sidelobe. A backazimuth of 175.6^{circ} is returned, approximately 5^{circ} off the estimated Great Circle path from the Reviewed Event Bulletin location estimate. However, other agencies' location estimates for this earthquake (ISC 2020) display some deviation from the CTBTO estimate and many of the estimates result in a backazimuth with a far smaller residual. The primary caveat to this method for an array with the dimensions of MJAR is the requirement of significant SNR on each trace. Whereas classical f-k analysis can estimate correctly the direction of a coherent incoming wavefield below the noise level, the incoherent method requires a visible signal (e.g. SNR over 2) on a significant subset of the sensors used Gibbons et al. (2008).

5 Conclusions and discussion

Monitoring compliance with the Comprehensive Nuclear-Test-Ban Treaty using the International Monitoring System requires that the signals generated by a nuclear explosion be detected, correctly interpreted, and associated with an accurate source hypothesis. We have focussed on the region surrounding the Sea of

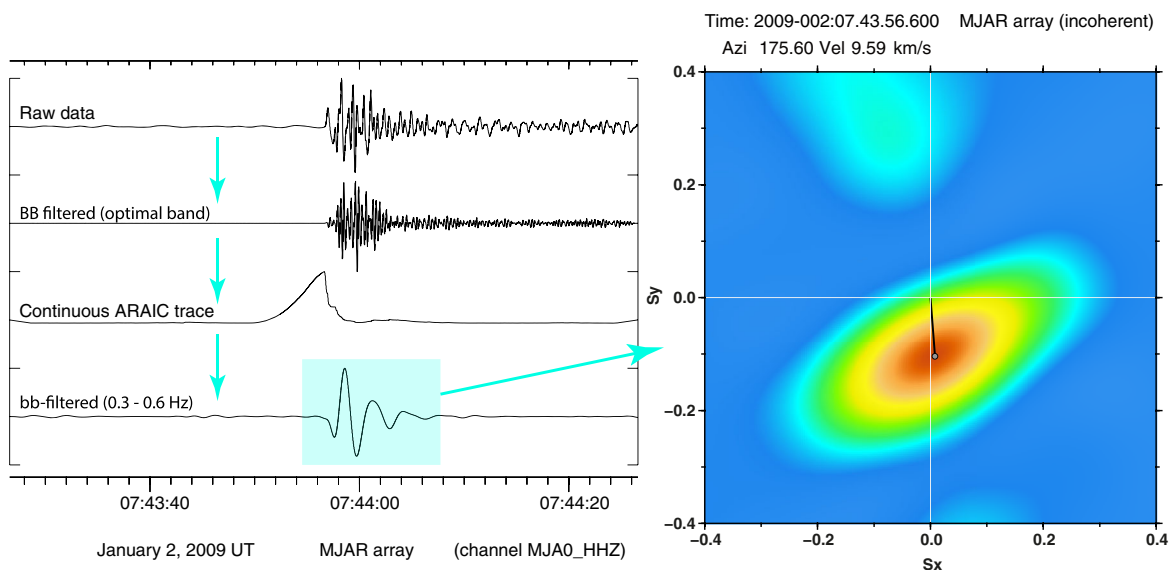


Fig. 10 A condensed step-wise summary of the incoherent processing algorithm (left) and coherent broadband f-k estimate of the bandpass filtered AR-AIC traces (right)

Japan. This region is both a prolific source of seismic signals from natural earthquakes and a region of intense monitoring scrutiny. The only nuclear test explosions to have taken place since the turn of the century have taken place at the Punggye-ri Nuclear Test Site in North Korea. The detection capability for this region is dominated by three primary IMS seismic array stations: USRK (Russian Federation), KSRS (South Korea), and MJAR (Japan). Understanding the properties of these three stations is key to evaluating and improving the monitoring capability for the region.

We have presented the geometries of the stations together with a discussion of the ability of the stations to characterize seismic signals. Many of the properties can be explained by the geometry alone. Others, such as the lack of signal coherence at MJAR, have needed to be found out empirically. We have demonstrated how empirical matched field processing (EMFP) can provide an alternative interpretation of seismic signals, mapping back into a geographical parameter space as opposed to a slowness space. On one level, this may allow for indirect estimations of the fundamental slowness parameters, apparent velocity, and backazimuth. Such estimates are likely to be subject to less bias than estimates made directly using a plane wavefront assumption since EMFP calibrates for persistent bias in measurement of wavefronts arriving from a given source. EMFP mitigates loss of signal coherence to a point, in that wavefronts that show persistent deviations from the plane-wave model can be correctly identified and classified. EMFP allows a calibration which changes with frequency.

Augmenting sets of plane-wave steering vectors with empirical vectors from sources of interest is likely to improve the ability to identify and resolve sources. Processing pipelines in which EMFP plays a major role will require dynamically updated banks of templates in order to correctly characterize all seismic signals observed so far. There needs to be an understanding of how to maintain and update the set of available templates. Our sensitivity studies have demonstrated that on small aperture arrays, a relatively small number of templates is likely to cover the geographical parameter space admirably. On a small aperture array, you can detect a phase from an earthquake at a location A using a template calibrated from an event at a location B with a considerable distance between the sources A and B. As the aperture of the

seismic array increases, the geographical footprint of a matched field template from a given source will decrease. The KSRS array would require far more empirical templates than the USRK array to provide adequate coverage of the same regions.

We have demonstrated the difficulties in slowness estimation of high-frequency regional signals on the MJAR primary seismic array in Japan. Spectrogram beamforming provides a powerful means of detecting high-frequency regional signals on this station, but the corresponding incoherent slowness estimates for such phases are often poor, mainly due to the relatively small aperture of the array. We suggest an alternative procedure for enhanced incoherent slowness estimates whereby

1. The high-frequency signals are bandpass filtered in an optimal frequency band
2. A continuous AR-AIC detection statistic is calculated for each of the traces
3. The AR-AIC traces are bandpass filtered in a lower frequency band
4. Coherent broadband f-k analysis is performed on the filtered AR-AIC traces.

The filtering of the AR-AIC traces close to the time of the local maxima results in a remarkably similar set of wavelets which are far more coherent across the array aperture than the transformed spectrograms as presented in Gibbons et al. (2008). Like the spectrogram method, the AR-AIC incoherent estimation procedure requires a somewhat higher SNR for each channel than is, in principle, required by coherent array processing. However, the incoherent procedure described here obtains stable slowness estimates for high-SNR signals for which coherent array processing fails. This is why we advocate its use as a supplementary pipeline process.

The performance of the AR-AIC-based procedure for regional S-waves at MJAR is typically poor due to the low SNR of the S_n waves on the vertical component sensors. We see in Fig. 7 that, for the case demonstrated here, the amplitudes and SNR of the S-phases are significantly higher for the horizontal components of MJAR than on the vertical. However, there is only a single site with horizontal components and so we cannot estimate the slowness vector for S-phases over the array using the workflow described above. The improvement in SNR of S-phases going from vertical only to a 3-component array can only be

determined by experimenting. Gibbons et al. (2019) showed that the increase in SNR for S-phases on the horizontal components at the ARCES array was modest; but that the spatial coherence of the horizontal-component S-phases was far higher than on the vertical components. Given the degree of signal incoherence demonstrated at MJAR in this paper, it seems unlikely that a modest improvement in the signal coherence for S-phases on the horizontal component sensors would be sufficient to provide robust direction estimates.

We have demonstrated that there are many approaches to signal detection and classification that provide very different kinds of output and very different levels of performance on the three array stations considered. Looking further afield, and including many different types of stations, will diversify still the range of methods needed to provide optimal characterization of the global wavefield. Methods for processing general seismic processing pipelines are changing. In the coming years, we are likely to see a replacement of the Global Association system for generating event hypotheses with the NET-VISA Machine Learning approach (Le Bras et al. 2020). Elsewhere in seismology, machine learning and deep learning are making steady progress into operational signal detection and characterization (Ross et al. 2018) and reliable real-time signal classification (Meier et al. 2019). Comprehensive overviews into applications of machine learning have recently been published both for seismology (Kong et al. 2019) and, more generally, acoustics (Bianco et al. 2019).

The problem of backazimuth estimation using 3-component stations using deep neural networks is addressed by Dickey et al. (2020) and it is likely that a similar procedure could be extended to sensor arrays. Advances in machine learning in seismic processing pipelines may result in us thinking differently about which inputs are needed for optimal monitoring and we may end up running many parallel processing pipelines, generating very diverse outputs. Our fundamental task is to input raw waveform data and output a comprehensive and accurate report of all seismic disturbances. We recognize that there will be many ways of reaching this goal from the available raw data and that advances in data processing and machine learning may accelerate the speed with which we need to re-evaluate our operational pipelines.

Acknowledgements The authors thank the reviewers for their time and for valuable comments to improve the quality of the manuscript. All maps in this paper are created using GMT software (Wessel and Smith 1995). Calculation of the covariance matrices was performed using the multitaper coherence software written by German Prieto and is freely available from <https://www.gaprieto.com/software> (last accessed 23 February 2021).

Funding This work was partly funded by the Air Force Research Laboratory under contract number FA9453-16-C-0006.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bianco MJ, Gerstoft P, Traer J, Ozanich E, Roch MA, Gannot S, Deledalle CA (2019) Machine learning in acoustics: theory and applications. *J Acoust Soc Amer* 146(5):3590–3628. <https://doi.org/10.1121/1.5133944>
- Dickey J, Borghetti B, Junek W (2020) BazNet: a deep neural network for confident three-component backazimuth prediction. *Pure Appl Geophys*. <https://doi.org/10.1007/s00024-020-02578-x>
- Fee D, Haney M, Matoza R, Szuberla C, Lyons J, Waythomas C (2016) Seismic envelope-based detection and location of ground-coupled airwaves from volcanoes in Alaska. *Bullet Seismol Soc Amer* 106(3):1024–1035. <https://doi.org/10.1785/0120150244>, <https://pubs.geoscienceworld.org/bssa/article/106/3/1024-1035/332161>
- Gibbons SJ, Ringdal F, Kvaerna T (2008) Detection and characterization of seismic phases using continuous spectral estimation on incoherent and partially coherent arrays. *Geophys J Int* 172(1):405–421. <https://doi.org/10.1111/j.1365-246X.2007.03650.x>
- Gibbons SJ, Schweitzer J, Ringdal F, Kvaerna T, Mykkeltveit S, Paulsen B (2011) Improvements to seismic monitoring of the European Arctic using three-component array processing at SPITS. *Bullet Seismol Soc Amer* 101(6):2737–2754. <https://doi.org/10.1785/0120110109>
- Gibbons SJ, Ringdal F (2012) Seismic monitoring of the North Korea nuclear test site using a multichannel correlation detector. *IEEE Trans Geosci Remote Sens* 50:1897–1909. <https://doi.org/10.1109/TGRS.2011.2170429>

- Gibbons SJ (2014) The Applicability of incoherent array processing to IMS seismic arrays. *Pure Appl Geophys* 171(3-5):377–394. <https://doi.org/10.1007/s00024-012-0613-2>
- Gibbons SJ, Kväerna T, Harris DB, Dodge DA (2016) Iterative strategies for aftershock classification in automatic seismic processing pipelines. *Seismol Res Lett* 87(4):919–929. <https://doi.org/10.1785/0220160047>
- Gibbons SJ, Harris DB, Dahl-Jensen T, Kväerna T, Larsen TB, Paulsen B, Voss PH (2017) Locating seismicity on the Arctic Plate boundary using multiple-event techniques and Empirical Signal Processing. *Geophys J Int* 211(3):1613–1627. <https://doi.org/10.1093/gji/ggx398>
- Gibbons, SJ, Schweitzer J, Kväerna T, Roth M (2019) Enhanced detection and estimation of regional S-phases using the 3-component ARCES array. *J Seismol* 23(2):341–355. <https://doi.org/10.1007/s10950-018-9809-y>
- Harris DB, Kväerna T (2010) Superresolution with seismic arrays using empirical matched field processing. *Geophys J Int* 182:1455–1477. <https://doi.org/10.1111/j.1365-246X.2010.04684.x>
- ISC (2020) International Seismological Centre on-line Bulletin. <https://doi.org/10.31905/D808B830>
- Kong Q, Trugman DT, Ross ZE, Bianco MJ, Meade BJ, Gerstoft P (2019) Machine learning in seismology: Turning data into insights. *Seismol Res Lett* 90(1):3–14. <https://doi.org/10.1785/0220180259>
- Krüger F, Dahm T, Hannemann K (2020) Mapping of Eastern North Atlantic Ocean seismicity from Po/So observations at a mid-aperture seismological broadband deep sea array. *Geophys J Int* 221:1055–1080. <https://doi.org/10.1093/gji/ggaa054>
- Kväerna T, Ringdal F (2013) Detection capability of the seismic network of the International Monitoring System for the Comprehensive Nuclear-Test-Ban Treaty. *Bullet Seismol Soc Amer* 103(2A):759–772. <https://doi.org/10.1785/0120120248>
- Le Bras R, Arora N, Kushida N, Mialle P, Bondár I, Tomuta E, Alamneh FK, Feitio P, Villarroel M, Vera B, Sudakov A, Laban S, Nippres S, Bowers D, Russell S, Taylor T (2020) NET-VISA from cradle to adulthood. A machine-learning tool for Seismo-Acoustic Automatic Association. *Pure Appl Geophys*. <https://doi.org/10.1007/s00024-020-02508-x>
- Leonard M, Kennett BLN (1999) Multi-component autoregressive techniques for the analysis of seismograms. *Phys Earth Planet Inter* 113(1-4):247–263. [https://doi.org/10.1016/s0031-9201\(99\)00054-0](https://doi.org/10.1016/s0031-9201(99)00054-0)
- Leonard M (2000) Comparison of manual and automatic onset time picking. *Bullet Seismol Soc Amer* 90(6):1384–1390. <https://doi.org/10.1785/0120000026>
- McLaughlin KL, Bonner JL, Barker T (2004) Seismic source mechanisms for quarry blasts: modelling observed Rayleigh and Love wave radiation patterns from a Texas quarry. *Geophys J Int* 156(1):79–93. <https://doi.org/10.1111/j.1365-246X.2004.02105.x>
- Meier MA, Ross ZE, Ramachandran A, Balakrishna A, Nair S, Kundzicz P, Li Z, Andrews J, Hauksson E, Yue Y (2019) Reliable real-time seismic signal/noise discrimination with machine learning. *J Geophys Res Solid Earth* 124:788–800. <https://doi.org/10.1029/2018JB016661>, 1901.03467
- Prieto GA, Parker RL, Vernon FL (2009) A Fortran 90 library for multitaper spectrum analysis. *Comput Geosci* 35(8):1701–1710. <https://doi.org/10.1016/j.cageo.2008.06.007>
- Ringdal F, Harris DB, Kväerna T, Gibbons S (2009) Expanding coherent array processing to larger apertures using empirical matched field processing. In: *Proceedings of the 2009 Monitoring Research Review, Ground-Based Nuclear Explosion Monitoring Technologies*. LA-UR-09-05276, Los Alamos National Laboratory, pp Tucson, Arizona, 379–388. <https://doi.org/10.6084/m9.figshare.14865342.v1>
- Rost S, Thomas C (2009) Improving seismic resolution through array processing techniques. *Surv Geophys* 30:271–299. <https://doi.org/10.1007/s10712-009-9070-6>
- Ross ZE, Meier MA, Hauksson E, Heaton TH (2018) Generalized seismic phase detection with deep learning. *Bullet Seismol Soc Amer* 108(5):2894–2901. <https://doi.org/10.1785/0120180080>, 1805.01075
- Ruigrok E, Gibbons S, Wapenaar K (2017) Cross-correlation beamforming. *J Seismol* 21(3):495–508. <https://doi.org/10.1007/s10950-016-9612-6>
- Selby ND (2013) A multiple-filter F detector method for medium-aperture seismometer arrays. *Geophys J Int* 192(3):1189–1195. <https://doi.org/10.1093/gji/ggs072>
- Wessel P, Smith WHF (1995) New version of the Generic Mapping Tools. *EOS Trans Am Geophys Union* 76:329

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.