Geophys. J. Int. (2008) **175,** 291–300 doi: 10.1111/j.1365-246X.2008.03912.x

Regional monitoring of infrasound events using multiple arrays: application to Utah and Washington State

Stephen J. Arrowsmith,[1] Rod Whitaker,[1] Steven R. Taylor,[2] Relu Burlacu,[3] Brian Stump,[4] Michael Hedlin,[5] George Randall,[1] Chris Hayward[4] and Doug ReVelle[1]

[1]*Los Alamos National Laboratory,* PO Box 1663*, Los Alamos, NM* 87545*, USA. E-mail: sarrowsmith@gmail.com*
[2]*Rocky Mountain Geophysics,* 167 *Piedra Loop, Los Alamos,* NM 87544*, USA*
[3]*University of Utah Seismograph Stations, University of Utah, Salt Lake City,* UT 84112*, USA*
[4]*Department of Geological Sciences, Dedman College, Southern Methodist University,* PO Box 750395*, Dallas,* TX 75275*, USA*
[5]*Laboratory for Atmospheric Acoustics, University of California at San Diego,* 9500 *Gilman Drive, La Jolla,* CA 92093-0223*, USA*

Accepted 2008 July 10. Received 2008 July 9; in original form 2007 December 21

SUMMARY

In this paper, we present an integrated set of algorithms for the automatic detection, association, and location of low-frequency acoustic events using regional networks of infrasound arrays. Here, low-frequency acoustic events are characterized by transient signals, which may arise from a range of natural and anthropogenic sources, examples of which include (but are not limited to) earthquakes, volcanic eruptions, explosions, rockets and bolides. First, we outline a new technique for detecting infrasound signals that works successfully in the presence of correlated noise. We use an F-statistic, sequentially adapted to ambient noise conditions, in order to obtain detections at a given statistical significance while accounting for real background noise. At each array, individual arrivals are then grouped together based on measured delay-times and backazimuths. Each signal is identified as either a first or later arrival. First arrivals at spatially separated arrays are then associated using a grid-search method to form events. Preliminary event locations are calculated from the geographic means and spreads of grid nodes associated with each event. We apply the technique to regional infrasound networks in Utah and Washington State. In Utah, over a period of approximately 1 month, we obtain a total of 276 events recorded at three arrays in a geographic region of $6 \times 4^{\circ}$. For four ground-truth explosions in Utah, the automatic algorithm detects, associates, and locates the events within an average offset of 5.4 km to the actual explosion locations. In Washington State, the algorithm locates numerous events that are associated with a large coalmine in Centralia, Washington. An example mining-explosion from Centralia is located within 8.2 km of the mine. The methodology and results presented here provide an initial framework for assessing the capability of infrasound networks for regional infrasound monitoring, in particular by quantifying detection thresholds and localization errors.

Key words: Time series analysis; Probability distributions; Seismic monitoring and test-ban treaty verifications; Wave propagation.

INTRODUCTION

Low-frequency acoustic waves with periods of 0.1–100 s, or 'infrasonic waves', can propagate in the atmosphere over large distances. Infrasound is generated by a wide range of natural and anthropogenic sources including, but not limited to, earthquakes, volcanoes, ocean waves, meteors, explosions and aircraft (Bedard & Georges 2000). Networks of infrasound arrays have the potential to detect, locate, and characterize these sources—with many potential applications ranging from monitoring and surveillance to ultimately improving our physical understanding of these phenomena. However, source effects (including coupling and directivity issues), atmospheric effects (which are dominated by effects of wind and temperature), and site-noise effects can complicate the measurement of infrasound signals at spatially separated locations. Le Pichon *et al.* (2008) developed a method for associating and locating infrasound signals at multiple arrays and applied it to data from infrasound stations in Europe. Their study demonstrated the importance of seasonal variations of atmospheric winds on network detection capability. This paper outlines an alternative approach to the infrasound association and location problem that uses a forward approach rather than an inverse method. In addition, we develop a new detection scheme that iteratively adapts to real ambient noise, removing the requirement for applying a post-detection categorization procedure to remove detections from continuous/repetitive sources. In comparison with the study by Le Pichon *et al.* (2008),

we focus on shorter datasets from much smaller regional networks with the goal of providing a preliminary assessment of the numbers and distributions of multi-array infrasound events in Washington State and Utah.

The development of the infrasonic component to the International Monitoring System (IMS), comprising a global network that will eventually include 60 infrasound arrays (Brachet & Coyne 2006), as well as the recent deployment of regional infrasonic arrays (e.g. Stump *et al.* 2004; Matoza *et al.* 2007; Stump *et al.* 2007), has led to a need for efficient algorithms to routinely process infrasonic data in order to generate preliminary event catalogues. This need has motivated this study as well as the recent study by Le Pichon *et al.* (2008). Brown *et al.* (2002) provide a broad summary of the infrasonic processing system used at the Prototype International Data Center (PIDC) for the automatic detection and location of infrasound. Similar algorithms are routinely used for the processing of seismic data (e.g. Blandford 1982; Bache *et al.* 1990). However, the infrasonic problem is inherently more complex due to the temporal variations in atmospheric conditions and higher levels of ambient noise. In this paper, we outline strategies for handling this added complexity.

An automatic system for processing infrasound network data comprises the following tasks: (1) signal detection, (2) phase identification and association, (3) event association (the association of signals at multiple sites), (4) event location and (5) event identification. This paper outlines an automatic procedure for implementing steps (1)–(4) in order to generate preliminary event catalogues for review by a human analyst.

METHODOLOGY

Signal detection

We have developed a new signal detection algorithm, which has a statistical basis, and can perform well in the presence of correlated noise. Most conventional array-based detectors are based on the assumption of uncorrelated noise, resulting in numerous spurious detections (which we term 'clutter' in this paper) from correlated noise sources. In the algorithm outlined here, we adaptively refine the null hypothesis used in the detector, using the actual ambient noise distribution at any given time.

The basis for the detection algorithm used in this study is the F-statistic (e.g. Shumway 1971; Smart 1971; Smart & Flinn 1971; Blandford 1974; Evers & Haak 2001). Following Blandford (1974), the F-statistic is defined as the ratio of the power on the beam versus the residual power, as follows:

$$F = \left(\frac{J-1}{J}\right) \times \frac{\sum_{n=n_0}^{n_0+(N-1)} \left[\sum_{j=1}^{J} x_j(n+l_j)\right]^2}{\sum_{n=n_0}^{n_0+(N-1)} \left[\sum_{j=1}^{J} \left\{x_j(n+l_j) - \left(\frac{1}{J} \cdot \sum_{m=1}^{J} x_m(n+l_m)\right)\right\}^2\right]}, \quad (1)$$

where J is the number of sensors, $x_j(n)$ is the waveform amplitude of the nth sample of the time-series from sensor j, l_j is the time-alignment lag obtained from beamforming, n_0 is the start sample index for the processing interval and N is the number of samples in the processing window.

F-statistics are calculated at each point in the waveform in a moving time-window. A standard frequency–wavenumber (F–K) analysis is implemented (e.g. Rost & Thomas 2002), whereby filtered array data are processed in each time-window to obtain the backazimuth and phase velocity of the maximum coherent energy across the array. These values of backazimuth and phase velocity are then used to compute the lag times (l_j in eq. 1), which allow for the computation of the F-statistic.

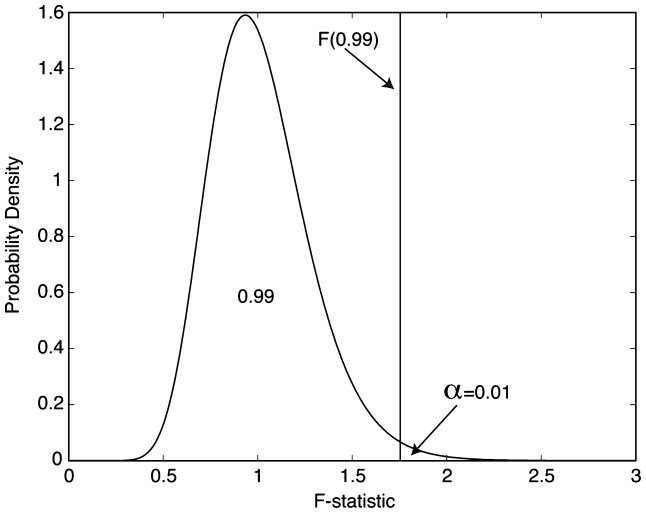

Figure 1. Probability density function of the null hypothesis (H_0)—which is 'perfectly correlated noise'—with 40 degrees of freedom on the numerator and 120 degrees of freedom on the denominator (corresponding to the values used for signal detection in Fig. 2). The 99th percentile value ($F_{0.99}$), used as the threshold in this study, is shown. The false-alarm probability (or p-value) is the area under the density function to the right of the threshold.

The calculated F-statistics follow a Central F-distribution in the presence of uncorrelated noise (Blandford 1974), allowing for the computation of p-values—providing the probabilities of obtaining F-statistics at least as extreme as the calculated values under this distribution: $p\{F(t)\}$. Under the null hypothesis of perfectly uncorrelated noise, by setting an appropriate statistical threshold (α), we can control the false alarm rate—effectively allowing us to control the numbers of spurious detections (see Fig. 1 for a graphical explanation of these terms). Unfortunately, correlated noise is often found in both seismic and infrasound data and arises from a variety of both natural and anthropogenic sources (e.g. ocean storms, surf, wind-farms, etc.). Furthermore, with the growing interest in regional seismic and infrasound data, typical array apertures have become smaller, further increasing contamination from correlated noise.

In the presence of correlated noise, the F-statistic is distributed as $cF_{2BT,2BT(N-1)}$, where B is the bandwidth, T is the time-window, N is the number of array elements and c is given by:

$$c = \left(1 + N\frac{P_s}{P_n}\right), \quad (2)$$

with $\frac{P_s}{P_n}$ denoting the signal-to-noise ratio (Shumway *et al.* 1999); where 'signal' in this instance strictly refers to correlated noise, and 'noise' strictly refers to uncorrelated noise. In practice, the ratio of uncorrelated to correlated noise is unknown, and varies as a function of time. Therefore, in this study, we divide our calculated F-statistics by a range of values of c in order to obtain the best fitting value; this is accomplished by finding the value of c that best matches the position of the peak of our distribution to the position of the peak of the Central F-distribution with $2BT$, $2BT(N-1)$ degrees of freedom. This simple distribution fitting approach

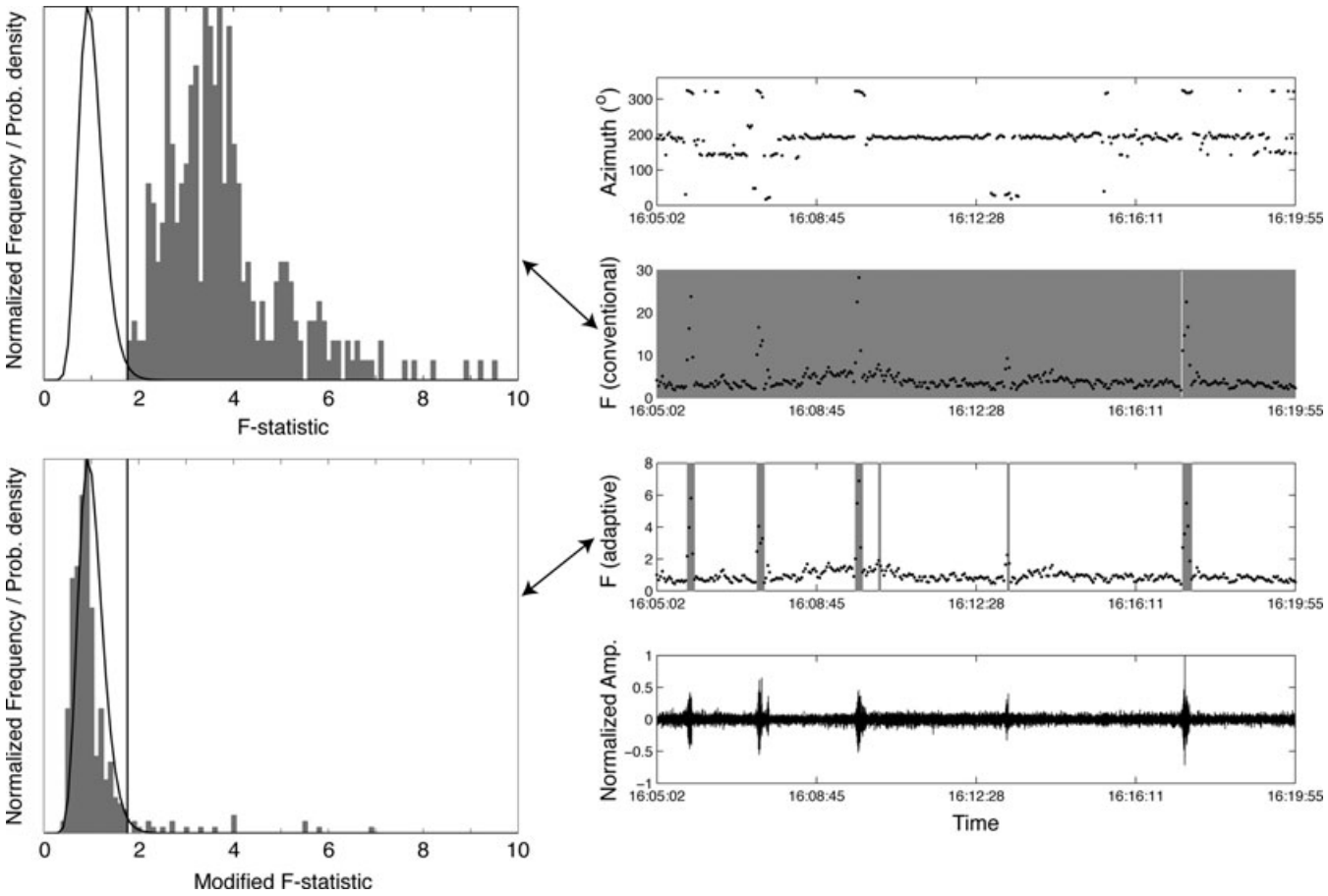

Figure 2. Example illustrating the difference between a conventional (idealized noise hypothesis, i.e. perfectly correlated noise) detector and the new (adaptive noise hypothesis) detector outlined in this study. Right-hand panels: Derived values of azimuth and F-statistic as a function of time, with the corresponding waveform at element MSH11. Gray shaded regions denote detections, with conventional detections plotted in Panel 2 and adaptive detections plotted in Panel 3. Left-hand panels: histograms of the F-statistics (top panel) and scaled F-statistics (bottom panel), in the 15-min time window shown. Black curves denote the Central F-distribution with $2BT$, $2BT(N-1)$ degrees of freedom and vertical red lines denote a p-value of 0.01.

is implemented adaptively in the detection algorithm, allowing for temporal variations in ambient noise. By fitting only the positions of the distribution peaks, the best-fitting value of c is representative of the ambient background, and not influenced by signals (which are intrinsically tails in our real distribution).

To illustrate the benefit of this method, rather than a more conventional approach (which typically assumes uncorrelated noise), Fig. 2 provides a comparison for a 15-min record at the four-element MSH1 infrasound array in Washington State (see Section 'Data set' for more details on the array design). This interval is characterized by persistent correlated noise from a wind farm located at a range of 15 km from the array, and an azimuth of $\sim 200^\circ$. In addition, there are five transient signals in the record, which clearly occur at different azimuths. The distribution of F-statistics, computed using eq. (1) for this time interval, is associated with a mean and spread that is quite different from the central F-distribution with $2BT$, $2BT(N-1)$ degrees of freedom. By applying a statistical significance threshold of 0.01, practically the whole record is flagged as 'detection' because the null hypothesis (uncorrelated noise) is violated. We have confirmed that this result is indicative of most conventional infrasound detectors that assume uncorrelated noise (e.g. Cansi 1995; Hart 2004). In contrast, by scaling the calculated F-statistics by a constant (c), and aligning the peaks of the distributions, we obtain a much better result for a statistical significance of 0.01 (and the same Central F-distribution). Although this example is quite an extreme case—correlated noise at many infrasound arrays is typically lower in amplitude and less persistent—it clearly illustrates the benefit of using an adaptive noise hypothesis.

Since the algorithm outlined here does not pick up significant clutter, in comparison with most existing detectors, we assert that it is not necessary to apply a post-detection categorization algorithm (e.g. Brachet & Coyne 2006) in order to remove clutter as a secondary step.

Phase identification and association

In this paper, we define 'phase association' as the process of associating separate phases (from a single event) at an array. Previous studies have explored the problem of phase association using seismic data, where the separate phases are well understood (e.g. P, S, etc.). Mykkeltveit & Bungum (1984) and Bache *et al.* (1990) apply phase association algorithms in order to associate S-wave arrivals with earlier P-wave arrivals. Phase velocities are used to identify the separate phases automatically. When a P-wave arrival is followed by an S-wave arrival within a specified time window, and the backazimuths are approximately the same, the two arrivals are automatically associated.

Infrasound phase identification (and subsequent association) is a more complex problem due to the time-varying nature of the propagation medium, which affects the specific phases that we observe. Furthermore, whereas phase velocity can be used for the identification of seismic arrivals, it can be ambiguous for identifying infrasound arrivals due to the complex interaction of infrasound with topography and near-surface atmospheric conditions. Refraction of infrasound can occur in the troposphere, stratosphere, and thermosphere [denoted as 'Iw', 'Is' and 'It' arrivals, respectively, by Brown *et al.* (2002)]. Modern, event-driven models of atmospheric temperatures and winds (e.g. the NRL-G2S model, described by Drob *et al.* 2003), can resolve the hourly variation of the upper atmospheric solar migrating tidal components, as well as the synoptic meteorology in the troposphere and stratosphere. However, atmospheric variations on shorter-timescales (i.e. $<$1 hr) and on shorter scale lengths (i.e. mesoscale and local scale fluctuations) are not accurately represented in the models. Furthermore, the effects of both topographic specular reflection (e.g. Arrowsmith *et al.* 2007a) and scattering of infrasound off inhomogeneities in the atmosphere (Kulichkov 2004), can result in a highly complex sequence of arrivals. Given these considerations, we follow the basic approach taken by Brown *et al.* (2002), which is to identify two types of arrivals: first arrivals ('I') and later arrivals ('Ix'). A given event will produce a single first arrival and q later arrivals, where $q \geq 0$. In this way, we can associate all arrivals from a single source without explicitly identifying the propagation paths.

We use two criteria to determine if an arrival is a first arrival: (1) arrival time—the difference in arrival time between any pair of arrivals from an individual event should be less than some threshold, $T_{\max}$ and (2) backazimuth—the difference in backazimuth between any pair of arrivals from an event should be less than some threshold, $\Delta\phi$. If an individual arrival is not associated with any preceding arrivals, it is automatically flagged as a first arrival ('I'). $T_{\max}$ can be empirically determined for each array from the predicted time difference between the fastest and slowest arrivals, taken for the most distant grid node from that array. A limitation with this approach is that signals from sources repeating with intervals less than $T_{\max}$ become associated. There is only one free parameter that needs to be explicitly defined: $\Delta\phi$.

Event association and location: grid-search method

The problem of associating first arrivals (i.e. 'I' phases) at multiple spatially separated arrays is termed 'event association' in this paper. This problem is closely related to the location problem and can be performed simultaneously. For infrasound, Le Pichon *et al.* (2008) recently outlined a method for associating and locating infrasound signals using an inverse approach. A systematic search for associations between detection bulletins provides initial locations (from cross bearings), which are then iteratively modified using an iterative least-square inversion scheme. Related work has been done on the problem of seismic event association (e.g. Blandford 1982; Ringdal & Kvaerna 1989) Ringdal & Kvaerna (1989) use a grid-based beamforming method. A network of spatially separated seismic arrays is 'steered' towards each beam location and a set of time delays for all combinations of beams, stations and phases is computed. By time-aligning the waveform data, beam traces are obtained that can be processed using standard threshold-based detection algorithms. Blandford (1982) reviews a number of seismic association algorithms that are based on grouping sets of arrivals at spatially separated seismometers. In each case, an initial estimate of the epicentre is used to identify compatible arrivals at other stations.

In this study, we use a grid-search algorithm to simultaneously identify and locate events (i.e. groups of 'I' arrivals at the different arrays that are associated). We do not have the advantage of seismic data, where sufficiently accurate earth models allow time delays to be reliably determined for a given source location, or where the differential time between the *P*- and *S*-wave arrivals can be used to estimate an initial epicentre for use in identifying compatible arrivals. In this study, a geographic grid is constructed that covers the region of interest. The grid spacing along a straight-line profile from a single array varies linearly with distance from the array. Thus, a straight-line function with two parameters must be defined: the minimum grid spacing (i.e. y-intercept) and rate of change of grid spacing (i.e. gradient). For multiple arrays we evaluate the grid spacing at any given point—from the distance to the nearest array (and corresponding straight-line function)—mapping out a 2-D grid-spacing function. The straight-line parameters (gradient and y-intercept) are chosen empirically for a given region to maximize the number of events detected while minimizing the required number of grid nodes.

Next, for each grid-node location, k, we compute a set of predicted (denoted by p-superscript) great-circle backazimuths (at each station) and maximum and minimum delay times (between each pair of arrays):

$$\begin{aligned} \Phi^k &= \left(\phi_1^p, \phi_2^p, \ldots \phi_n^p\right), \\ \mathbf{dT}^k_{\max} &= \left\{\frac{(d_1 - d_2)}{0.22}, \frac{(d_1 - d_3)}{0.22}, \ldots\right\}, \\ \mathbf{dT}^k_{\min} &= \left\{\frac{(d_1 - d_2)}{0.34}, \frac{(d_1 - d_3)}{0.34}, \ldots\right\}, \end{aligned} \tag{3a}$$

where there are n elements in $\Phi^{\boldsymbol{k}}$ (corresponding to the number of arrays) and there are $m = n!/(2(n\text{-}2)!)$ elements in $\mathbf{dT}^k_{\max}$ and $\mathbf{dT}^k_{\min}$(corresponding to the number of pairs of arrays); ϕ_i^p is the predicted backazimuth at the ith array (taken as the true great-circle backazimuth) and d_i is the distance (in kilometres) from the grid node to the ith array; finally, k denotes the kth grid node. The maximum and minimum delay times are calculated in order to ensure that we only associate signals with desired group velocities (infrasonic group velocities are typically in the range $0.22 \text{ km s}^{-1} \leq \text{v} \leq 0.34 \text{ km s}^{-1}$, Ceplecha *et al.* 1998).

The set of maximum and minimum delay-times and great-circle backazimuths for all the grid nodes are then compared with the set of observed first-arrivals at all arrays (obtained by following the methodology described in Sections 'Signal detection' and 'Phase identification and association'). At the ith array, we have a set of observed arrival times and backazimuths:

$$\begin{aligned} \mathbf{t_i} &= (t_{i1}, t_{i2}, \ldots, t_{ij_i}), \\ \Phi_{\mathbf{i}} &= (\phi_{i1}, \phi_{i2}, \ldots, \phi_{ij_i}), \end{aligned} \tag{3b}$$

with a total of j_i first-arrivals. For each pair of arrays, we therefore have a total of $j_1 \times j_2$ pairs of first-arrivals that can be considered possible two-array ('2-a') events (i.e. events with arrivals at both arrays). However, for each pair of arrays we can reduce the number of 2-a events by storing only 2-a events with delay-times that are less than the longest possible interarray delay-time ($\mathbf{dt}^{\mathbf{global}}_{\mathbf{max}}$).$\mathbf{dt}^{\mathbf{global}}_{\mathbf{max}}$ is equal to $\Delta_{ij}/0.22$ (where Δ_{ij} is the great-circle distance between the two most distant stations). This step is performed prior to comparing observations (eq. 3b) with predictions (eq. 3a) via a grid search in order to reduce computational requirements.

For n arrays, we obtain $m = n!/(2(n\text{-}2)!)$ sets of 2-a events (i.e. one set of events for each pair of stations). We then cluster 2-a events into n-a events (i.e. events with arrivals at all n arrays). For each n-a event we now have a set of observed delay-times and backazimuths,

$$\begin{aligned} \mathbf{dt} &= (dt_1, dt_2, ..., dt_m)\,, \\ \Phi &= (\phi_1, \phi_2, ..., \phi_n)\,, \end{aligned} \tag{3c}$$

where m denotes the number of pairs of arrays as defined above. These observations can be directly compared with the predicted parameters in eq. (3a) using a grid-search procedure.

In the grid-search procedure, we perform a loop over each grid-node location. For each grid node, the algorithm searches for n-a events where the set of delay-times (**dt**) lie within the maximum and minimum allowed bounds for the given grid node ($\mathbf{dT}^k_{\max}$ and $\mathbf{dT}^k_{\min}$) and where the set of backazimuths (Φ) are within $\pm\Delta\phi^\circ$ of the great-circle backazimuths (Φ^k) for the given grid node. A given event may be associated with a number of grid nodes, which map out the corresponding localization uncertainty region. A simple and robust estimate of location uncertainty may be estimated from the standard deviations of grid node extents in latitude and longitude. To account for picking errors of the automatic detector, an estimated picking error, E_p, is allowed when comparing real observations with the maximum and minimum allowed delay times (eq. 3a).

DATA SET

In this study we used data from two regional infrasound networks located in Utah (Stump *et al.* 2007) and Washington State (Matoza *et al.* 2007). The Utah infrasound network used in this study comprises three four-element arrays, with apertures of ∼100 m. Inter-array distances range from 74 to 85 km (Fig. 3). The network was deployed as part of a collaboration between the University of Utah, Southern Methodist University, Weston Geophysical and ENSCO Inc. (Stump *et al.* 2007). The arrays comprise Chaparral 2 and Chaparral 2.5 sensors, fitted with porous hoses for wind noise reduction. The data are sampled at 100 Hz. The Washington State network used in this study contains two temporary deployments, the MSH1 and MSH2 arrays, which are operated by the University of California San Diego and the Geological Survey of Canada; and one permanent IMS array, the I56US array. The MSH1 and MSH2 arrays each consist of four elements, with apertures of ∼100 m. Interarray distances for the full three-array network, range from 249 to 445 km (Fig. 3). The MSH1 and MSH2 arrays consist of MB2000 (DASE/Tekelec) broad-band aneroid microbarometres fitted with porous hoses. The data are sampled at 40 Hz, with a flat response between 100 s and 17 Hz. The I56US array also consists of MB2000 sensors, which are fitted with a pipe array for wind noise reduction. The array aperture is ∼1 km, and the data are sampled at 20 Hz.

SYNTHETIC TESTS

For the purpose of assessing network location resolution, and to get a handle on the effect of grid-search parameters (i.e. allowed group velocities and backazimuth deviations), we have performed a series of synthetic tests. For each synthetic test, an event location and origin time are defined. Given these parameters, synthetic arrival times and backazimuths are calculated at each array using pre-defined group-velocities (and taking synthetic backazimuths to be great-circle backazimuths). The set of synthetic arrival times and backazimuths thus constitute the input to the grid-search procedure described above.

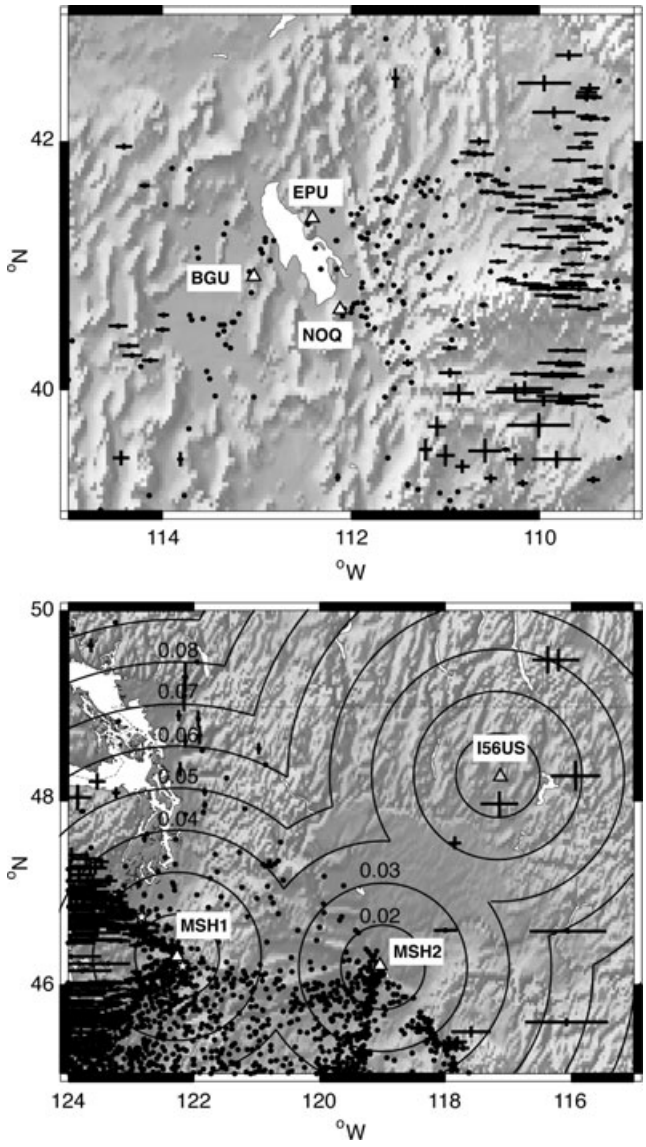

Figure 3. Maps of the Utah (top panel) and Washington (bottom panel) infrasound networks with association/location results superimposed. A fixed grid spacing of 0.01° is used in Utah whereas variable grid spacing (shown by the black lines) is used in Washington. Black points denote event locations; error bars depict corresponding standard deviations of uncertainties in latitude and longitude. White stars in the bottom panel denote the locations of the Centralia coalmine (∼46.7°N, 122.8°W) and Mount Saint Helens (∼46.2°N, 122.2°W).

Fig. 4 shows the results obtained from a set of synthetic tests for the Utah network described in the previous section. Four pre-defined event locations are set (one corresponding with the location of the Utah Test and Training Range, UTTR). For three different sets of grid-search parameters, the location uncertainties are mapped out by the locations of associated grid nodes. As the grid-search parameters are increasingly constrained, the location uncertainties collapse to the known event locations (Fig. 4). In an operational monitoring scheme it is necessary to use sufficiently relaxed parameters to ensure that no events are missed. For the Utah network, this results in corresponding location uncertainties shown in Fig. 4(c). These observations must be considered when interpreting the results described below.

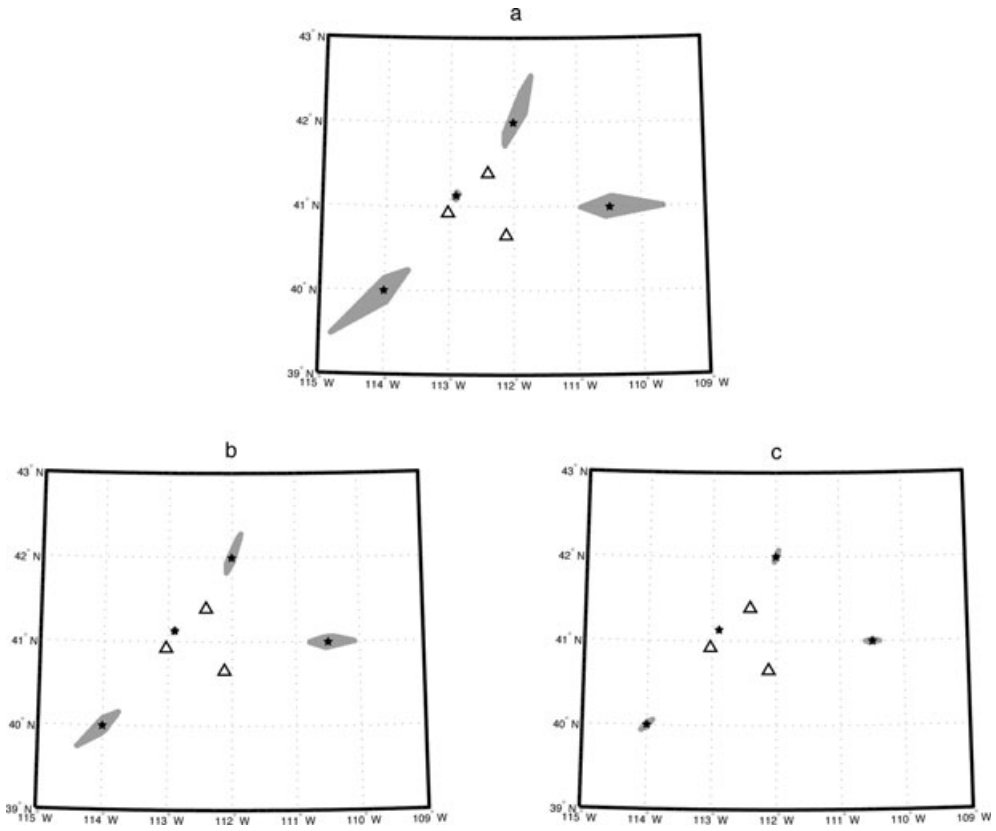

Figure 4. Results of synthetic association/location tests for four pre-defined event locations (black stars). The grid-search parameters used for each test are as follows: (a) $\Delta\phi = 6^\circ$, $v_g = 0.28{-}0.34$ km s^{-1}, (b) $\Delta\phi = 3^\circ$, $v_g = 0.32{-}0.34$ km s^{-1} and (c) $\Delta\phi = 1^\circ$, $v_g = 0.299{-}0.301$ km s^{-1}. Black stars show input event locations and grey regions highlight resultant event location uncertainties (i.e. location extents of associated grid nodes). The event located at ∼41°N, 113°W corresponds to the location of ground-truth UTTR explosions analysed in this paper.

RESULTS

Utah network

We have processed Utah network data for a period of ∼1 month (07/24/07–08/28/07) using the detection scheme described above with the following parameters: window length = 10 s, overlap = 50 per cent, frequency band = 1–5 Hz. These parameters are specifically tailored for the detection of relatively short, high-frequency arrivals. Such types of signals may arise from a range of sources including (but not limited to) some explosions, mining events, rocket launches, earthquakes, bolides and impulsive volcanic eruptions at local and regional distances (although earthquakes, bolides and volcanic eruptions can also cause lower frequency signals). Some signals will unavoidably be filtered out by our choice of parameters, including large events with purely thermospheric arrivals (i.e. high frequencies strongly attenuated), and some long duration volcanic signals.

After applying the detector to the data, we obtain sets of arrivals at each array in the Utah network (Fig. 5). For each arrival, the following parameters are obtained: start time, end time, backazimuth, phase velocity, *F*-statistic, *c*-value (eq. 2) and correlation. Clear spatial trends are observed in the arrivals at each array, as shown in Fig. 5. For example, relatively large numbers of arrivals with $C \sim 0.5$ are observed from the east at each array.

The three sets of detections are then processed using the Phase Identification/Association and Event Association algorithms previously described. We search for signals with group velocities that range from 0.28 to 0.34 km s^{-1}, filtering out thermospheric arrivals (which are typically <1 Hz and therefore already effectively removed by the bandpass filter). The following parameters are used: Backazimuth deviation ($\Delta\phi$) = 6° and picking error (E_p) = 40 s. Large deviations in backazimuth are allowed in order to compensate for measurement uncertainty (e.g. Szuberla & Olson 2004) and atmospheric effects, in particular winds (e.g. Mutschlecner & Whitaker 2005). The allowed picking error is large due to the emergent nature of many infrasonic signals. The result is a set of triplets of associated arrivals at all three arrays (or 'events'). In total, we obtain 276 events recorded at all three arrays during the period of study. Each event is associated with a set of N grid nodes, which delineates the possible source region for that event.

The final set of event locations is computed by taking the geographic means of the grid nodes associated with each event. The location uncertainties are taken to be the standard deviations of grid-node locations in latitude and longitude respectively. The locations of all events recorded at the three arrays are shown in Fig. 3. Four known rocket motor explosions, conducted at UTTR were automatically detected and located with high accuracy (see Table 1 and further discussion below), providing confidence in the automatic catalogue. We also detect a number of unknown events, including relatively large numbers of events to the east of the network. However, due to the network geometry, which was intended primarily for studying local events, there is a large uncertainty in the range of these events (refer to Fig. 4). At the time of writing, the origin of these events was unknown, but a detailed analysis of the causes of infrasound in this region is beyond the scope of this paper.

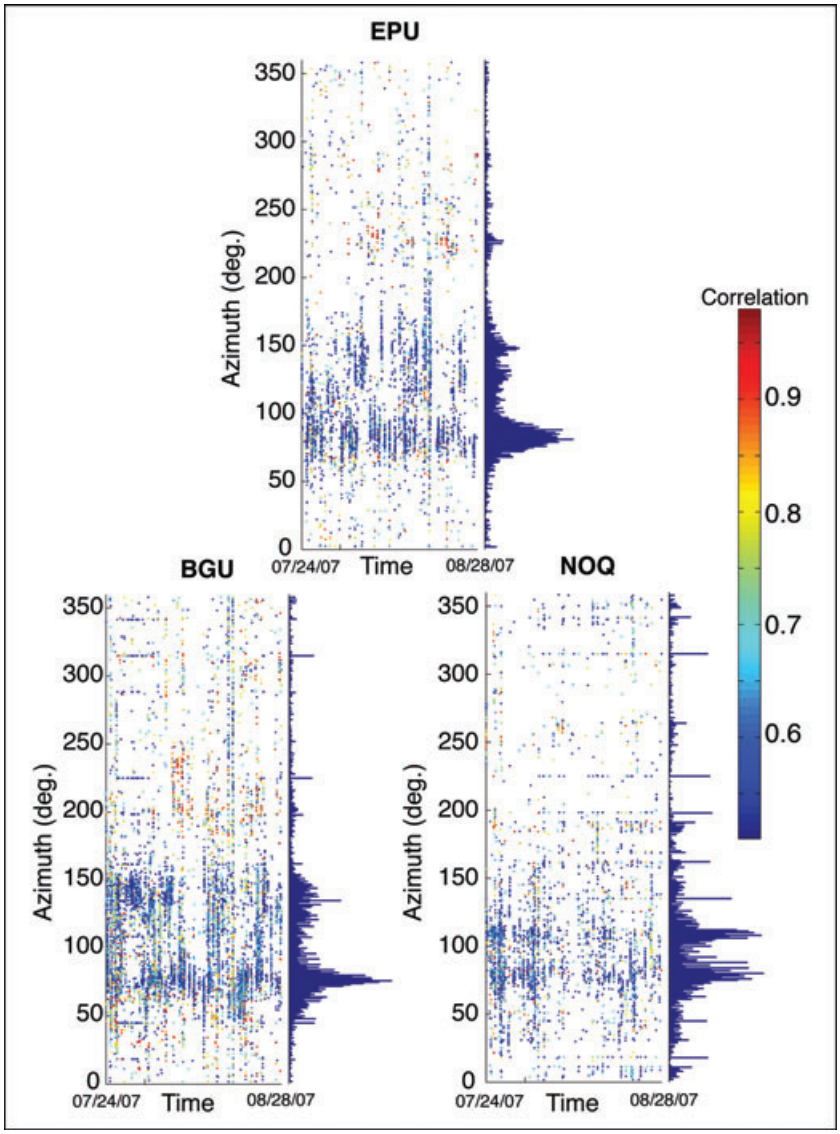

Figure 5. Automatic detections obtained at the Utah network arrays. In each panel, detections are plotted with respect to time and backazimuth and colour-coded by correlation. Some artefacts are observed, associated with lines of detections at constant azimuth, which arise from *F–K* processing of coherent signals with subsonic group velocities (i.e. winds). The histograms show numbers of detections from each 1-degree bin of backazimuth.

Table 1. Comparison of automatic results with ground-truth for the four known UTTR explosions.

	Ground-truth		Automatic Results		
Event	Location	Origin-time	Location	Dist. from GT (km)	Azimuth from GT (°)
1	41.13152, −112.89577	08/01/07 20:01:24.5	41.22, −112.86	10.3	17
2	41.13152, −112.89577	08/06/07 20:33:02.6	41.15, −112.86	3.8	55.6
3	41.13152, −112.89577	08/13/07 19:38:20.6	41.16, −112.93	4.4	317.8
4	41.13152, −112.89577	08/27/07 20:43:12.0	41.13, −112.86	3.1	93.2

It is anticipated that the linkage with seismic data, in addition to detailed infrasonic propagation modelling, will improve these event locations and allow for eventual event characterization.

Washington network

As a second test of the methodologies described above, we have applied the procedure to ∼5 months of data (11/02/04–03/20/05) from the Washington network (using the same parameters used for Utah). The same proviso applies: we are only searching for relatively short, high-frequency signals at all three arrays. After applying the Phase Identification/Association and Event Association algorithms, we obtained a total of 1824 two-array events (using MSH1 and MSH2), but no three-array events after incorporating I56US. We speculate two reasons for this: (1) we have trained the algorithm to search for short-duration/high-frequency signals, which are unlikely to be observed over long ranges and (2) we have observed high

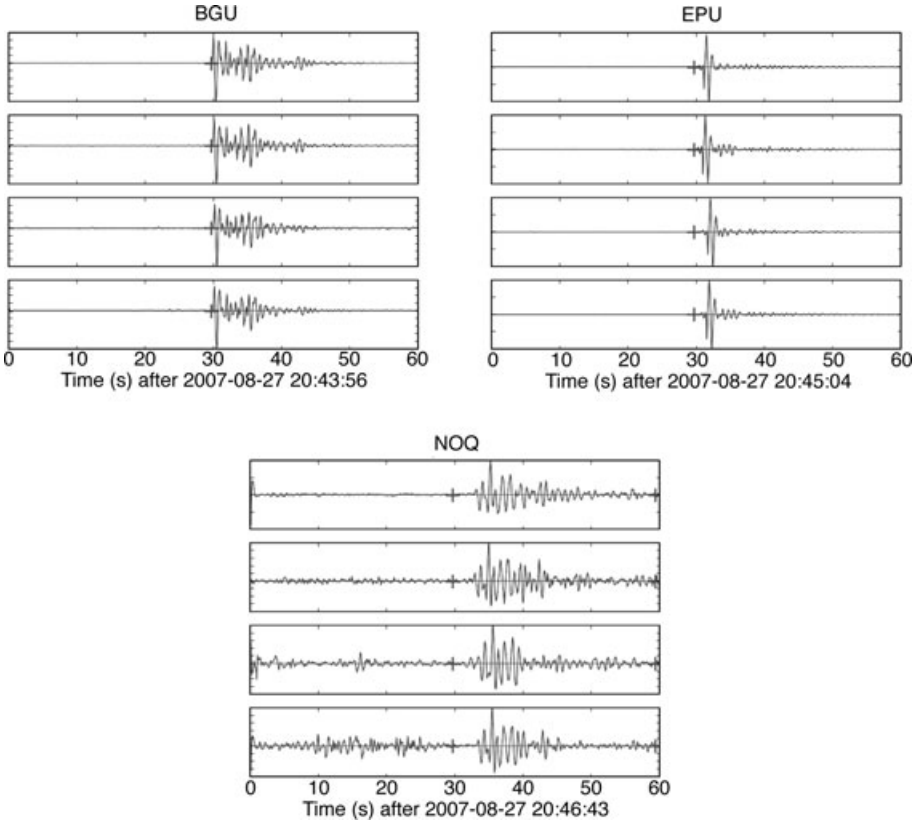

Figure 6. Automatic waveform picks for a UTTR explosion detected on 2007 August 27. Picks associated with the event are located at 30 s on each panel. The waveform amplitudes are normalized to the maximum amplitude for each trace.

levels of noise at I56US. By searching for lower-frequency signals we associated and located an eruption of Mount Saint Helens within 5 km of the caldera using all three arrays (Arrowsmith and ReVelle 2007).

The event locations and corresponding uncertainties for the two-array events (MSH1 and MSH2) are shown in Fig. 3. We highlight two notable sources, each producing numerous events during the period of study. The first source is located at ∼46.7°N, 122.8°W—close to a large surface coal mine located at Centralia, Washington. We have tentatively associated a number of signals with the mine (see the next section for an example). The second source is Mt Saint Helens, which appears to be associated with several events recorded at both arrays (Fig. 3). Unfortunately, given the network configuration, the location uncertainties of sources are large.

Example events

Although a detailed interpretation of the detected events is beyond the scope of this paper, we have selected two example events for further analysis in order to provide some confidence in the automatic results. The first event (Event 4, Table 1) was a ground-truth rocket motor explosion at the UTTR site on 2007 August 27, that was detected by all three Utah network arrays. The waveforms and automatic arrival time picks at each array are plotted in Fig. 6. Clear impulsive signals are observed at each array, with very high signal-to-noise ratios. In total, four known ground-truth explosions were detonated at the UTTR site during the time interval studied. Each explosion is associated with high-quality arrivals at each array, which are automatically associated and located within a mean distance of 5.4 km from the test site. The location offsets are larger than the estimated uncertainties due to the fact that the estimated uncertainties do not account for atmospheric winds. Further research is required to quantify this effect. Table 1 provides a summary of the automatically obtained UTTR explosion locations, with comparison to ground truth.

The second event selected for further analysis is a suspected mining explosion from the Centralia coalmine in Washington State (located at ∼46.7°N, 122.8°W) on 2004 December 21. The waveforms, shown in Fig. 7, indicate a high signal-to-noise arrival at MSH2 (located at a range of 60.6 km from the Centralia coal mine). The arrival at MSH1 (located at a range of 292.6 km) is associated with a significantly lower signal-to-noise ratio, but the signal can still be visually identified above the noise. The final location for this event is offset by 8.2 km from the mine location at an azimuth of 242°.

DISCUSSION AND CONCLUSIONS

The two case studies presented here demonstrate the effectiveness of the methodologies outlined in this paper for detecting, associating and locating events over regional scales. This assessment is further confirmed by the analysis of ground-truth events—providing confidence in the automatic results. One issue not discussed in relation to these results is the effect of seasonal wind patterns. During the summer months, stratospheric wind directions in both regions blow from east to west, inhibiting stratospheric ducting of infrasound from the west, whereas the wind direction is opposite during the winter. This seasonal effect may partially explain why the

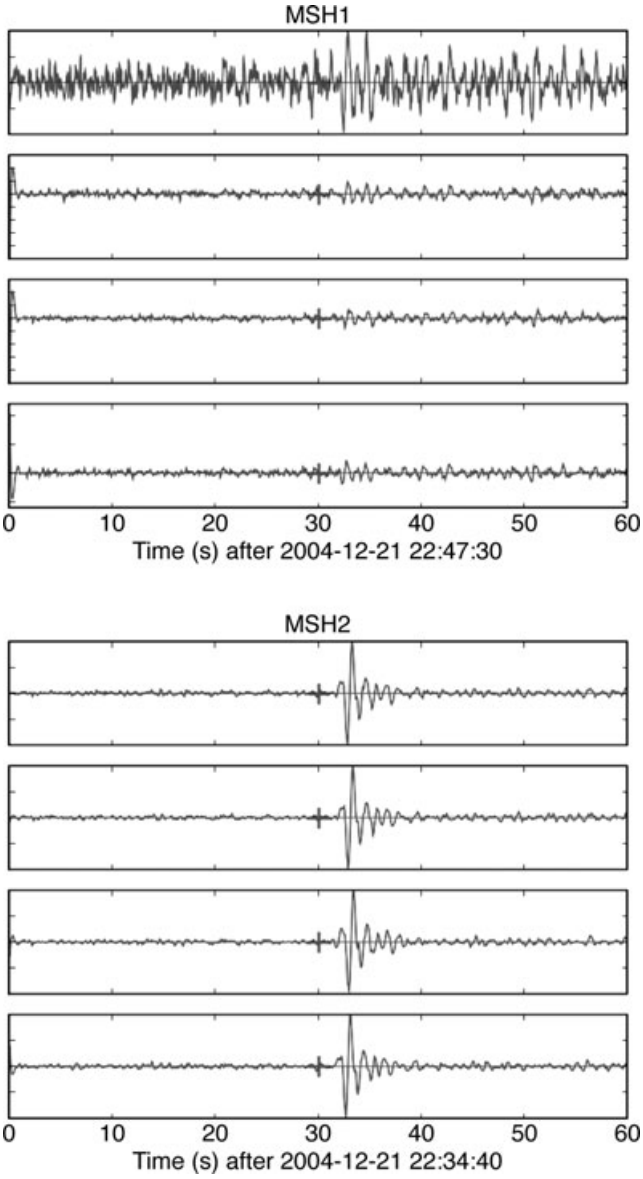

Figure 7. Automatic waveform picks for a suspected coal mine shot detected on 2004 December 21. Picks associated with the event are located at 30 s on each panel. The waveform amplitudes are normalized to the maximum amplitude for each trace.

distributions of detected events are located mainly to the east of the Utah network and to the west of the Washington network for the time periods processed. If we had sufficient data to perform a full seasonal study, we would expect to observe seasonal differences in the distributions of events, as was observed by Le Pichon *et al.* (2008).

It must be emphasized that the purpose of this suite of algorithms is to automatically generate preliminary catalogues of infrasound events at regional networks. Despite taking every effort to reduce the number of false alarms by (1) employing an adaptive noise hypothesis detector and (2) requiring each signal to be associated with signals at other arrays, any automatic method will ultimately require review by a human analyst. A manual analysis of waveforms to ensure accurate pick times, and detailed propagation modelling, should be performed for events of interest in order to refine these preliminary results. However, the manual production of a preliminary event bulletin is clearly impractical. This methodology reduces sets of acoustic time-series at spatially separated arrays into a preliminary event bulletin that can be practically reviewed by a human analyst. We stress that it is not our aim in this paper to provide a final set of event locations, nor to comprehensively assess network performances.

This paper provides an initial framework towards quantifying thresholds of detection as well as quantifying location capabilities using multiple infrasound arrays. First, we have outlined a new method for signal detection, and demonstrated its effectiveness in the presence of clutter. Next, we have presented a new method for the association and location of events recorded at multiple arrays. This suite of algorithms has been implemented into a Matlab software package at Los Alamos National Laboratory ('InfraMonitor'). We have tested this package on data from two regional infrasound networks, located in Utah and Washington State. Our results clearly demonstrate the effectiveness of this set of tools for detecting, associating and locating sources of interest. In particular, we have shown that for four ground-truth explosions in Utah, the automatic algorithm detects, associates and locates the events within an average of 5.4 km of the site. Using a separate dataset, the algorithm locates a suspected mining explosion in Washington State within 8.2 km of the mine. These results do not account for the effects of variations in atmospheric winds and temperature, and we intend to further improve these findings in the future by incorporating realistic atmospheric models.

ACKNOWLEDGMENTS

We would like to thank Dale Anderson and Bob Shumway for their statistical insights and Marv Wetovsky for editorial comments. We also thank Johan Robertsson and three anonymous reviewers for their constructive comments, which helped to improve this manuscript. We gratefully acknowledge financial support from the GNEM program of the US DOE-HQ in NA-22.

REFERENCES

Arrowsmith, S.J. & ReVelle, D., 2007. Infrasound monitoring of local, regional and global events, in *Proceedings of the 29th Monitoring Research Review*, Denver, CO, 25–27 Sept.

Arrowsmith, S.J., Drob, D., Hedlin, M.A.H. & Edwards, W., 2007. A joint seismic and acoustic study of the Washington State bolide, *J. geophys. Res.,* **112,** doi:10.1029/2006JD008001.

Bache, T.C., Bratt, S.R., Wang, J., Fung, R.M., Kobryn, C. & Given, J.W., 1990. The intelligent monitoring system, *Bull. seism. Soc. Am.,* **6,** 1833–1851.

Bedard, A.J. & Georges, T.M., 2000. Atmospheric infrasound, *Phys. Today,* **53,** 32–37.

Blandford, R.R., 1974. An automatic event detector at the Tonto Forest seismic observatory, *Geophysics,* **39,** 633–643.

Blandford, R.R., 1982. Seismic event discrimination, *Bull. seism. Soc. Am.,* **72,** S69–S87.

Brachet, N. & Coyne, J., 2006. The current status of infrasound data processing at the International Data Centre, in *Proceedings of the 28th Seismic Research Review,* Orlando, FL, 19–21 Sept.

Brown, D.J., Katz, C.N., Le Bras, R., Flanagan, M.P., Wang, J. & Gault, A.K., 2002. infrasonic signal detection and source location at the prototype International Data Centre, *Pure appl. Geophys.,* **159,** 1081–1125.

Cansi, Y., 1995. An automated seismic event processing for detection and location: the P.M.C.C. method, *Geophys. Res. Lett.,* **22,** 1021–1024.

Ceplecha, Z., Borovicka, J., Elford, W.G., ReVelle, D.O., Hawkes, R.L., Porubcan, V. & Simel, M., 1998. Meteor phenomena and bodies, *Space Sci. Rev.,* **84,** 327–471.

Drob, D.P., Picone, J.M. & Garces, M.A., 2003. The global morphology of infrasound propagation, *J. geophys. Res.,* **108,** doi:10.1029/2002JD003307.

Evers, L.G. & Haak, H.W., 2001. Listening to sounds from an exploding meteor and oceanic waves, *Geophys. Res. Lett.,* **28,** 41–44.

Hart, D., 2004. Automated infrasound signal detection algorithms implemented in MetSeis—infra tool, Sandia Report, SAND2004-1889.

Kulichkov, S.N., 2004. Long-range propagation and scattering of low-frequency sound pulses in the middle atmosphere, *Metorol. Atmos. Phys.,* doi:10.1007/s00703-003-0033-z, 47–60.

Le Pichon, A., Vergoz, J., Herry, P. & Ceranna, L., 2008. Analyzing the detection capability of infrasound arrays in Central Europe, *J. geophys. Res.,* **113,** doi:10.1029/2007JD009509.

Matoza, R.S., Hedlin, M.A.H. & Garces, M.A., 2007. An infrasound study of Mount St. Helens, *J. Volcanol. Geother. Res.,* **160,** 249–262.

Mutschlecner, J.P. & Whitaker, R.W., 2005. Infrasound from Earthquakes, *J. geophys. Res.,* **110,** doi:10.1029/2004JD005067.

Mykkeltveit, S. & Bungum, H., 1984. Processing of regional seismic events using data from small-aperture arrays, *Bull. seism. Soc. Am.,* **74,** 2313–2333.

Ringdal, F. & Kvaerna, T., 1989. A multi-channel processing approach to real time network detection, phase association, and threshold monitoring, *Bull. seism. Soc. Am.,* **79,** 1927–1940.

Rost, S. & Thomas, C., 2002. Array seismology: methods and applications, *Rev. Geophys.,* **40,** doi:10.1029/2000RG000100.

Shumway, B.H., 1971. On Detection a signal in N stationarily correlated noise series, *Technometrics,* **13,** 499–519.

Shumway, R.H., Kim, S.-E. & Blandford, R.R., 1999. Nonlinear estimation for time series observed on arrays, in *Asymptotics, Nonparametrics and Time Series,* pp. 227–258, ed. Ghosh, S., Marcel Dekker, New York.

Smart, E., 1971. Erroneous phase velocities from frequency-wavenumber spectral sections, *Geophys. J. R. astr. Soc.,* **26,** 247–253.

Smart, E. & Flinn, E.A., 1971. Fast frequency-wavenumber analysis and fisher signal detection in real-time infrasonic array data processing, *Geophys. J. R. astron. Soc.,* **26,** 279–284.

Stump, B., Jun, M.-S., Hayward, C., Jeon, J.-S., Che, I.-Y., Thomason, K., House, S.M. & McKenna, J., 2004. Small-aperture seismo-acoustic arrays: design, implementation, and utilization, *Bull. seism. Soc. Am.,* **94,** 220–236.

Stump, B., Burlacu, R., Hayward, C., Bonner, J., Pankow, K., Fisher, A. & Nava, S., 2007. Seismic and infrasound energy generation and propagation at local and regional distances: phase 1 – divine strake experiment, in *Proceedings of the 29th Monitoring Research Review,* Denver, CO, 25–27 Sept.

Szuberla, C.A. & Olson, J.V., 2004. Uncertainties associated with parameter estimation in atmospheric infrasound arrays, *J. Acous. Soc. Am.,* **115,** 253–258.