Geophysical Journal International

Geophys. J. Int. (2015) **200,** 1411–1422
GJI Seismology

doi: 10.1093/gji/ggu486

Development of a robust and automated infrasound event catalogue using the International Monitoring System

Stephen Arrowsmith, Garrett Euler, Omar Marcillo, Philip Blom, Rod Whitaker and George Randall

EES-17, Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545, USA. E-mail: sarrowsmith@gmail.com

Accepted 2014 December 18. Received 2014 December 17; in original form 2014 May 2

SUMMARY

Methods for detecting, associating and locating infrasound events recorded on the global International Monitoring System (IMS) infrasound network are presented. By using likelihood arguments, and reducing the use of empirically determined parameters, our techniques enable us to formally quantify the false alarm rate at both station and network levels, and to calculate confidence areas for event localization. We outline a new association technique that uses graph theory for associating arrivals at multiple spatially separated stations, and perform Monte Carlo simulations to quantify the performance of the scheme under different scenarios. The detection, association and location techniques are applied to 10 large events in the Reviewed Event Bulletin of the Comprehensive Nuclear Test Ban Treaty Organization. Out of 10 events, a total of seven were automatically detected and associated. By analysing the three missed events, we identify improvements that might be made to improve the algorithms.

Key words: Time-series analysis; Probability distributions; Acoustic-gravity waves; Seismic monitoring and test-ban treaty verification; Statistical seismology.

1 INTRODUCTION

The International Monitoring System (IMS) infrasound network is a global network of infrasound stations designed to monitor atmospheric nuclear tests (Christie & Campus 2010). At the time of writing, the network comprised 47 stations with the eventual goal of comprising 60 stations. Despite being designed for monitoring atmospheric tests, the network detects a variety of geophysical events such as bolides, volcanic eruptions and earthquakes, providing an excellent data set for researchers. The task of processing the data from all IMS stations to automatically detect, associate and locate events for further study presents a significant challenge due to the amount of data (the IMS will generate around 8 GB d^{-1} of data when completed, assuming 8 arrays per station, sampling rate at 40 samples per second and 5 bytes per sample), the sparseness of the network and the effect of the dynamic nature of the atmosphere on infrasound propagation. For this task, one is interested in optimizing the detection of large transient events that generate detections at multiple arrays while minimizing the potential detection of false detections. The Comprehensive Nuclear Test Ban Treaty Organization (CTBTO) has made significant progress in developing algorithms for processing infrasound data automatically (Brachet *et al.* 2010) but significant challenges remain, particularly in terms of estimating and reducing the false alarm rate (FAR). Here, we outline an alternative framework for processing IMS data where our goals are to minimize the number of tuning parameters that are not tied to likelihood arguments, to provide a technique that does not rely on parametrizing the Earth with nodes and is therefore more efficient than existing methods, and to quantify the probability of detecting real events and the FAR at both station and network levels. Through developing this process we hope to shed light on the capability of the IMS infrasound network for detecting low-yield atmospheric tests, which may not be detected by any other means.

A series of recent studies have developed and applied methods for processing data from regional infrasound networks to generate event catalogues (Arrowsmith *et al.* 2008; Le Pichon *et al.* 2008; Briels 2010; Park *et al.* 2014). This paper extends the methods developed an earlier publication for processing data from regional infrasound networks (Arrowsmith *et al.* 2008). First, infrasound detection methodology (identification of waveform signals at an individual site) as relevant to IMS data processing is discussed. Next, we discuss a new approach for associating detections at multiple stations that employs graph theory (West 2000). The implementation of Monte Carlo simulations enables us to estimate the probability of detection and false alarm at the network level. Next, we summarize how the resultant associations are processed to obtain locations at global distances, leveraging methods developed for infrasound location which had been developed previously (Modrak *et al.* 2010; Arrowsmith *et al.* 2014; Marcillo *et al.* 2014). The resulting set of techniques are illustrated by analysing infrasound generated by a large event that were detected on multiple stations across the IMS infrasound network: the eruption of Kelud volcano in Indonesia on

© The Authors 2015. Published by Oxford University Press on behalf of The Royal Astronomical Society.

2014 February 13. Finally, by processing 10 large events in the Reviewed Event Bulletin (REB), we quantify the performance of our techniques with real data.

2 DETECTION

In the context of infrasonic signal analysis, detection is the process of identifying signals recorded on spatially separated microbarometers. The method can be separated into parameter estimation (identifying the slowness vector and uncertainties) and hypothesis testing (determining if the estimated parameters have statistical merit). These are reviewed briefly below to justify our choice of methodology and to provide some parameters that are relevant to processing IMS infrasound data.

In selecting and developing a detector for processing IMS infrasound data there are several considerations. First, the detector should reliably detect transient signals with a minimum number of false alarms. In the case of very low amplitude signals (SNR < 1), detection can become extremely challenging. Here, we focus on detecting signals with SNR ~ 1 and above. Future R&D should examine methods to detect low amplitude signals in a noisy record; however, it is important to note that such signals can dramatically increase the probability of false association. Additionally, one should formally account for temporally varying noise. We have addressed this using an adaptive F detector as outlined in Arrowsmith *et al.* (2009). Finally, as the IMS produces large volumes of data, it is necessary that the method be efficient so that it can be employed in real time.

There is a wide-body of literature on parameter estimation using stations, largely from the engineering community (Krim & Viberg 1996; van Veen & Buckley 1998; Shumway *et al.* 2008). Such methods typically assume a plane wave model, of the form:

$$y_j(t) = \sum_{k=1}^{K} s_k \left(t - \mathbf{r}'_j \mathbf{q}_k\right) + n_j(t), \tag{1}$$

where s_k is a plane wave crossing an array with element coordinates $\mathbf{r}_j$, $\mathbf{q}_k$ is the slowness vector and n_j is the incoherent background noise. The estimated slowness vector, $\mathbf{q}_k = (\theta_1, \theta_2)'$, is related to the trace velocity and backazimuth of the incident planewave by $c_k = \|\mathbf{q}_k\|^{-1}$ and $\alpha_k = \tan^{-1}(\theta_{2k}/\theta_{1k})$, respectively. For the application discussed here, the slowness vector is computed in the frequency domain. We estimate $\mathbf{q}$ in a moving time window by selecting the value that maximizes the F-statistic,

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

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

For the work discussed here, a single coherent signal is considered within each window; though is should be noted that methods have been developed which search for multiple signals in any given time window (Shumway *et al.* 2008). From a practical perspective, estimating parameters for multiple incident signals may reduce bias when one of the signals' backazimuth is similar to the ambient microbaroms ($\mathbf{q}_s \cong \mathbf{q}_{mb}$). For more typical scenarios, where a signal and microbarom would come from different directions, the moving time window estimate of $\mathbf{q}_k = (\theta_1, \theta_2)'$ will alternate between $\mathbf{q}_s$ and $\mathbf{q}_{mb}$ depending on the relative power of these signals.

Hypothesis testing can be based on some statistical model of the noise, some statistical model of the signal, or both. The detection hypothesis is formulated in terms of a null hypothesis (only noise is present in the record), H_0, and an alternate hypothesis (the record contains both signal and noise), H_1. To maximize the probability of detection for a given probability of false alarm, one defines a threshold, γ, and makes a decision that H_1 is valid if,

$$\frac{p(\mathbf{x}; H_1)}{p(\mathbf{x}; H_0)} > \gamma, \tag{3a}$$

where $\mathbf{x}$ represents some parameter that is used to distinguish between signal and noise (e.g. power), $p(\mathbf{x}; H_0)$ and $p(\mathbf{x}; H_1)$ are probability distributions for $\mathbf{x}$ under the two hypotheses, and the threshold, γ, is found by specifying a probability of false alarm $P_{\mathrm{FA}} = \alpha$,

$$P_{\mathrm{FA}} = \int_{\left\{\mathbf{x}: \frac{p(\mathbf{x};H_1)}{p(\mathbf{x};H_0)} > \gamma\right\}} p(\mathbf{x};\ H_0)\, \mathrm{d}\mathbf{x} = \alpha, \tag{3b}$$

where the combination of eqs (3a) and (3b) is referred to as the Neyman–Pearson theorem (Kay 2008). In practice, $p(\mathbf{x}; H_1)$ is dependent on the signal-to-noise ratio, and therefore unknown. Because of this, we focus on controlling the false alarm rate under H_0. The advantage of using this p-value scheme is that Fisher's combined probability test provides a natural means to combine complementary detectors (Arrowsmith & Taylor 2013).

For infrasound monitoring, the probability distribution $p(\mathbf{x}; H_0)$ should vary as a function of time to account for ambient noise conditions. Following the classic paper by Blandford (1974), we employ the F-statistic detector defined in eq. (2), which is a ratio of the beam power to the residual power. In the presence of uncorrelated noise F can be shown to be distributed as a central F distribution with $2BT, 2BT(N–1)$ degrees of freedom (where BT is the time-bandwidth product and N is the number of sensors in a station array).

Infrasonic noise is often correlated, typically due to microbaroms (Arrowsmith *et al.* 2009). In the presence of a single continuous correlated signal,

$$F_{|H_0} \sim F_{2BT,2BT(N-1)\lambda} \approx C\, F_{2BT,2BT(n-1)}, \tag{4}$$

where λ is the non-centrality parameter (Shumway *et al.* 1999; Shumway 2002). In this relation we use the statistical convention where $\sim$ denotes 'distributed as' and $\approx$ denotes 'approximately distributed as'. Based on eq. (4), one can remap the observed F statistics in a given time window by simply dividing F by a constant C so that F/C follows the central F-distribution for uncorrelated noise with $2BT$, $2BT(N–1)$ degrees of freedom. The purpose of this step is to reduce false detections associated with coherent continuous-wave sources such as microbaroms or wind farms. For this modified F-detector, Arrowsmith *et al.* (2009) have estimated receiver operating characteristic (ROC) curves using real infrasound data. Because F in a given processing time window is selected by maximizing over all values of $\mathbf{q}$, it can be shown that this approach is not valid in the presence of pure uncorrelated noise (M. Charbit, personal communication, 2013). However, due to the ubiquity of microbaroms and other sources of correlated infrasonic noise, we find that white noise is not representative of real infrasound noise. As found in a study by Park (2013), the value of C varies

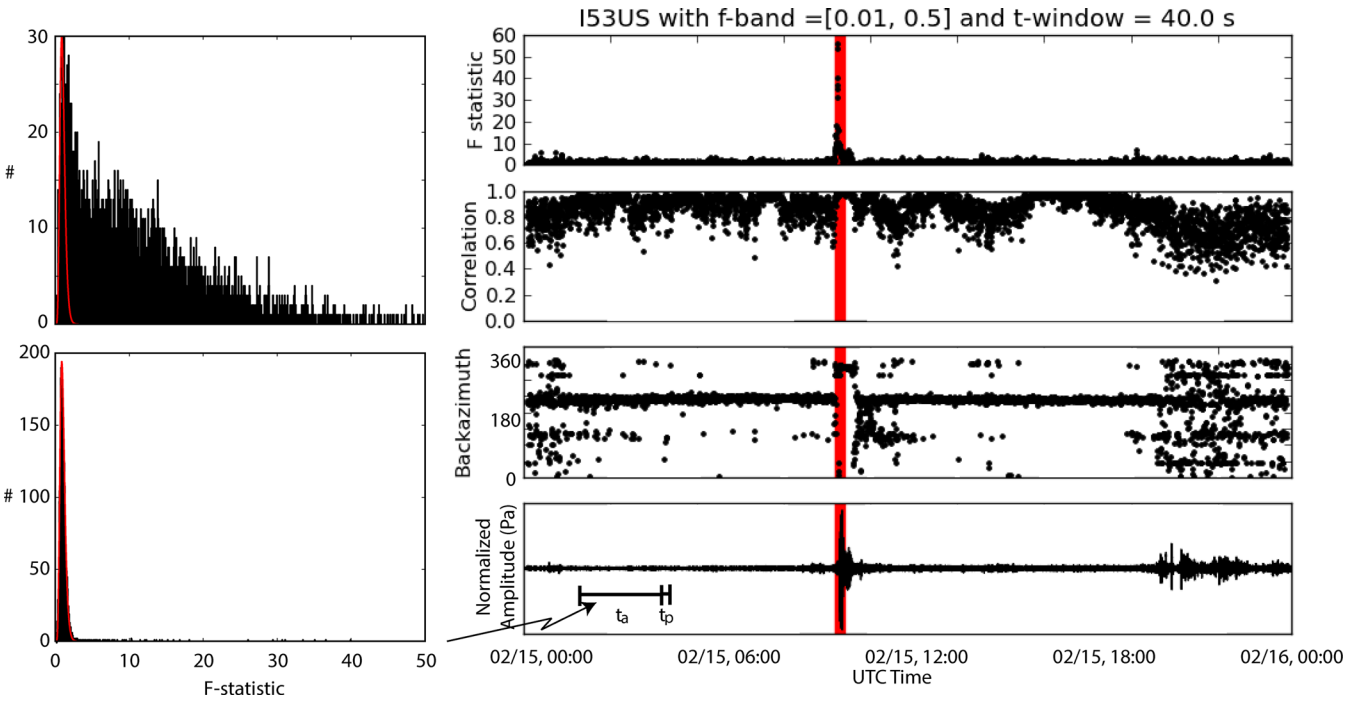

Figure 1. Illustration of the adaptive *F*-detector as applied to data recorded at I53US in Alaska on 2013 February 15. On the right-hand side the four panels represent (from top to bottom), the *F*-statistic for each processing time window, the cross-correlation coefficient averaged between each pair of elements, the backazimuth and the filtered waveform on a single channel. The red bar denotes the automatic detection based on a *p*-value threshold of 0.01. On the left-hand-side, the top plot shows the distribution of *F*-statistics computed using eq. (3) (black histogram), with theoretical *F* distribution for $2BT$, $2BT(N–1)$ degrees of freedom (red curve). The bottom plot shows the same theoretical curve with the scaled *F*-statistics (F/C). The detected signal is associated with the Chelyabinsk bolide. Note the persistent backazimuth of $\sim$200° that is associated with microbaroms.

as a function of time-of-day and is typically high during periods of microbarom-dominated noise at nighttime and decreases as uncorrelated wind noise increases during the day. Consequently, the adaptive F detector can formally account for these diurnal variations.

In our IMS implementation of the adaptive F detector, we adaptively update H_0 on the basis of the distribution of estimated F statistics in a time window, t_a, referred to as the adaptive time window. For a given processing time window, t_p, the detection threshold is estimated based on the scaling parameter, C, required to properly scale the calculated F statistics in the adaptive time window preceding the processing time window. Figs 1 and 2 illustrate the implementation of the adaptive F detector for SNR $\gg 1$ and SNR ~ 1. The example in Fig. 1 shows an automatic detection of a signal at I53US from the Chelyabinsk bolide on 2013 February 15. A more detailed analysis of the infrasound signals from this event can be found in Le Pichon *et al.* (2013). The example in Fig. 2 shows an automatic detection at I56US of a suspected bolide that fell over the North Pacific on 2010 December 25, discussed in Arrowsmith *et al.* (2013). Both examples illustrate the effects of persistent correlated noise from microbaroms that result in non-uniformly distributed backazimuth estimates.

We process IMS data in two frequency bands that span the ranges 0.01–0.5 and 0.5–3 Hz; the sets of detection parameters we have chosen for these bands are shown in Table 1. We use *P*-values of 0.001 and 0.01 for detecting signals in the low-frequency and high-frequency bands, respectively, based on tuning using events from the REB compiled by the CTBTO. A full summary of the parameters for detection can be found in Table 1. To reduce the false alarm rate we require that the *p*-value threshold be exceeded for a minimum duration of 360 and 80 s for the low- and high-frequency bands, respectively. Based on calibration using real events, we have found that these constraints significantly reduce the number of false alarms while ensuring that the detector does not miss regional and global signals with SNR > 1.

Once a number of detections at a single station has been obtained, the next step is to group separate phases from a single event together. Detections from long duration signals can occasionally be incorrectly broken down into separate detections based due to temporal drops in SNR while vertical multipathing can result in multiple arrival phases from a single event. Before running association at the network level it is necessary to merge such detections together. The IDC refer to this step as detection categorization (Brachet *et al.* 2010). Detections are merged together on the basis of proximity in detection time, backazimuth and frequency into groups referred to as metafamilies. Metafamilies are subsequently analysed and those with certain characteristics, such as a duration exceeding a defined threshold, are identified as noise and removed from further analysis (Brachet *et al.* 2010). Since the adaptive F detector is designed to remove detections from correlated noise, we do not throw detections away prior to association. Instead, our 'detection categorization' step simply clusters detections that are close in detection time and backazimuth, using thresholds of 3600 s in time, chosen to equal the value used by the IDC to form metafamilies (Brachet *et al.* 2010), and 10° in backazimuth to account for large additive deviations in measurement (Szuberla & Olson 2004) and wind bias (Mutschlecner & Whitaker 2005). In the following discussion, metafamilies are referred to as detections and their individual components as subdetections.

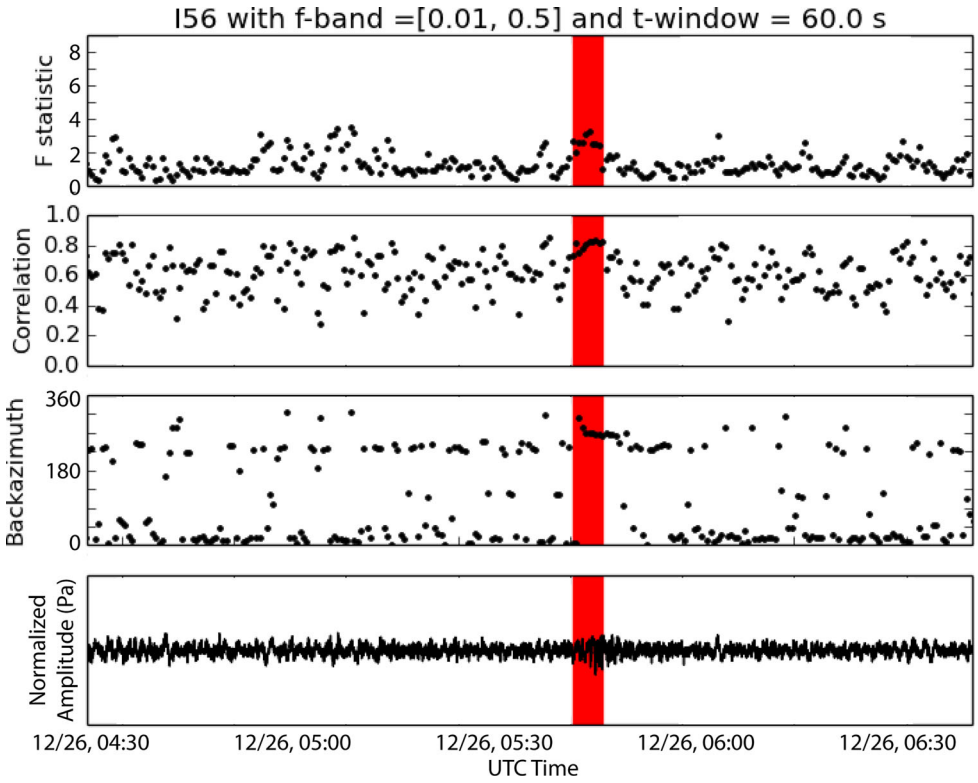

Figure 2. Illustration of the adaptive *F*-detector as applied to data recorded at I56US in Washington State on 2010 December 26. The signal at around 05:45 is associated with a suspected bolide that fell over the North Pacific. The red bar denotes the automatic detection based on a *p*-value threshold of 0.01. Note the persistent backazimuths of ∼200° and ∼10° that are associated with microbaroms.

Table 1. Detection parameters used for IMS processing. Refer to Fig. 1 for an explanation of the time windows.

Frequency band (Hz)	Adaptive time window (s)	Processing time window (s)	Overlap (per cent)	*p* Value threshold (P_{FA})	Minimum duration (s)
0.01–0.5	900	60	50	0.01	360
0.5–3	1800	40	50	0.001	80

3 ASSOCIATION

Association is the process by which signals recorded at multiple separate stations are grouped together based on consistent features to form an event hypothesis. Infrasound association is often based purely on the arrival time and back azimuth of signals recorded at multiple spatially separated stations (Arrowsmith *et al.* 2008). Efforts to explore the inclusion of additional waveform parameters such as frequency (Brachet *et al.* 2010), slowness, or amplitude in the association scheme require further research to formally define models of the relevant observable characteristics. In the case of using amplitude and frequency content, one might utilize the expression developed by Le Pichon *et al.* (2012) from large-scale numerical modelling, where the transmission loss is parametrized as a function of propagation range and the effective sound speed ratio between the ground and stratosphere. Our requirements for an association algorithm are that it (i) does not miss events detected at more than six stations, (ii) does not associate too many false events, (iii) is efficient and scalable to large numbers of detections and (iv) minimizes the number of ad-hoc parameters.

The association algorithm presented here is broken down into a four-step process, which is demonstrated using the eruption of Kelud volcano in Indonesia on 2014 February 13. The crux of the algorithm is the formation of a robust initial (seed) event hypothesis that forms the basis of Stages 1–3. Our method down-weights the influence of large backazimuth deviations using an iterative procedure.

3.1 Stage 1: Forming two-station associations

The detection information is used to construct inputs consisting of detection times and backazimuths at multiple spatially separated stations,

$$\mathbf{t}=\begin{cases}\mathbf{t}_1=(t_{11},t_{12},\ldots,t_{1n_1})\\ \vdots\\ \mathbf{t}_m=(t_{m1},t_{m2},\ldots,t_{mn_m})\end{cases}\quad\Theta=\begin{cases}\Theta_1=(\theta_{11},\theta_{12},\ldots,\theta_{1n_1})\\ \vdots\\ \Theta_m=(\theta_{m1},\theta_{m2},\ldots,\theta_{mn_m})\end{cases}, \tag{5}$$

where m is the number of stations and n_j is the number of detections at station j. The associator begins by considering each pair of stations within great-circle distances of 90°. For a given pair of stations, we consider each detection at Station a, searching for possibly associated detections at Station b. For each detection at Station a, d_{ai}, the algorithm first finds the set of detections at Station b, $\mathbf{d}_b$, such that the arrival time of each detection at Station b, t_{bi}, is within a range $t_{ai} - T_{\max} \leq t_{bi} \leq t_{ai} + T_{\max}$ defined on the basis of the arrival time at Station a (t_{ai}) and the maximum delay time

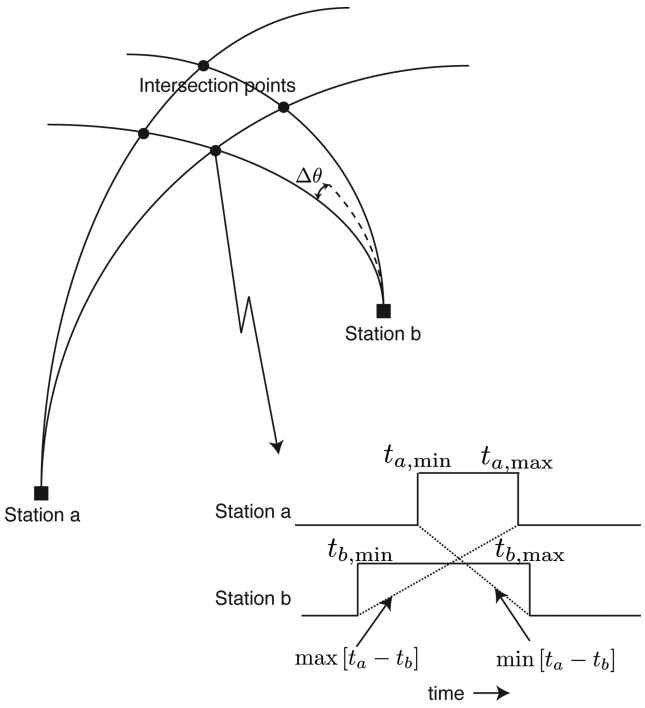

Figure 3. Illustration of the calculation of the four intersection points given two detections and subsequent testing for compatible time delays. Four intersection points are readily calculated given any two detections and associated uncertainties. For a given intersection point, the arrival times at the *j*th station, $t_{j,\min}$ and $t_{j,\max}$ are calculated based on the full range of infrasound celerities (0.22–0.34 km s^{-1}) for a source at $t = 0$. Next, the maximum and minimum possible arrival time difference between the pair of stations is calculated as shown schematically in the bottom right inset.

for the given station pair ($T_{\max} = \Delta_{ab}/v_{\min}$). The great circle distance between Stations a and b, Δ_{ab}, is calculated using spherical geometry. This step eliminates associations that are not physically plausible on the basis of incompatible arrival times.

Next, for each plausible detection pair, four bounding great circle intersection points given by $(\theta_{ai} - \Delta\theta, \theta_{ai} + \Delta\theta, \theta_{bi} - \Delta\theta, \theta_{bi} + \Delta\theta)$ and the locations of Stations a and b are calculated as shown in Fig. 3. The intersection points define a leading order bound of the possible source region on the sphere. The backazimuth deviation, $\Delta\theta$, is assumed constant for all detections and is set to a sufficiently large value that it accounts for variations due to both measurement and model components. It should be noted that an estimate of $\Delta\theta$ from the shape of the FK spectrum at a given station would provide only an estimate of the measurement uncertainty and does not account for wind bias along the propagation path. If the possible source location exceeds a threshold distance of 10 000 km from either Station a or Station b, the two-station association is not considered because the possible origin location is poorly resolved. This threshold is an empirically tuned parameter and an equivalent threshold could be based on the area of the polygon defined by the four intersection points defined in Fig. 3.

Using the set of intersection points obtained from the backazimuth variation, one can remove two-station associations with incompatible time delays. For each of the four intersection points, time delays between possible arrivals at Stations a and b are calculated using a uniform celerity PDF that bounds possible infrasound celerities. The uniform PDF in celerity translates to a uniform PDF in time at each station, as shown in Fig. 3. By calculating the minimum and maximum delay times for each intersection point, as illustrated for one point in Fig. 3, the global minimum and maximum define the full range of possible delay times. If the difference in arrival times observed at Stations a and b, $\Delta t = t_a - t_b$, falls within the global maximum and minimum times calculated from all four intersection points, the pair of detections are stored as a possible two-station association.

3.2 Stage 2: Constructing robust multistation associations

The next step links the resulting two-station associations together using common detections to identify multistation associations. In the following discussion, we refer to multi-station associations as graphs following graph theory (West 2000). Fig. 4(a) shows a 48-hr period of IMS data corresponding to the eruption of Kelud volcano where 16 graphs are formed using common detections of two-station associations. In this representation, the coloured circles denote detections and the lines connecting them indicate two-station associations formed in Stage 1. The majority of the detections in these graphs are associated with only one other detection and are likely not due to real events. We have found that real events generate graphs containing highly linked detections (detections that are associated with multiple other detections). To eliminate low-connectivity graphs, we remove detections that are associated with fewer than three other detections. As can be seen in Fig. 4(a), the connectivity in a graph is greater for the true associations. Fig. 4(b) illustrates that by removing detections with fewer than three associations then constructing graphs on the basis of single associations, we obtain three robust graphs.

3.3 Stage 3: Examining physical consistency of multistation associations

The next step considers the physical consistency of each graph by analysing the locations of the event hypotheses (Fig. 5). The event hypotheses for the largest two graphs in Fig. 4 are shown in Fig. 5. Because we cannot check the physical consistency of graphs with only two robust detections they are not carried beyond Stage 2. A robust geographic mean is calculated by iteratively removing the event hypothesis with the largest distance from the mean location and recalculating the mean. This repeats until the mean is determined using only 50 per cent of the event hypotheses for the graph. The purpose of this step is to calculate a more robust geographic mean that is less influenced by outliers. A mean origin time is also computed for these event hypotheses. Using the full graph, we then remove event hypotheses with distances greater than 2000 km from the robust geographic mean. The value of 2000 km is an empirically tuned parameter that provides additional robustness of the event hypothesis. For the graph with four detections in Fig. 4, the geographic distribution of event hypotheses is so poor that the graph collapses when event hypotheses larger than 2000 km from the robust mean are removed (Fig. 5).

3.4 Stage 4: Searching for signals at remaining stations using the seed association

The final step is to use the geographic and time estimates obtained in Stage 3 to search for associations at more distant stations. Given the origin time and location, it is straightforward to search for signals that are within a time window and backazimuth tolerance at each station. For this step we allow a large tolerance on arrival

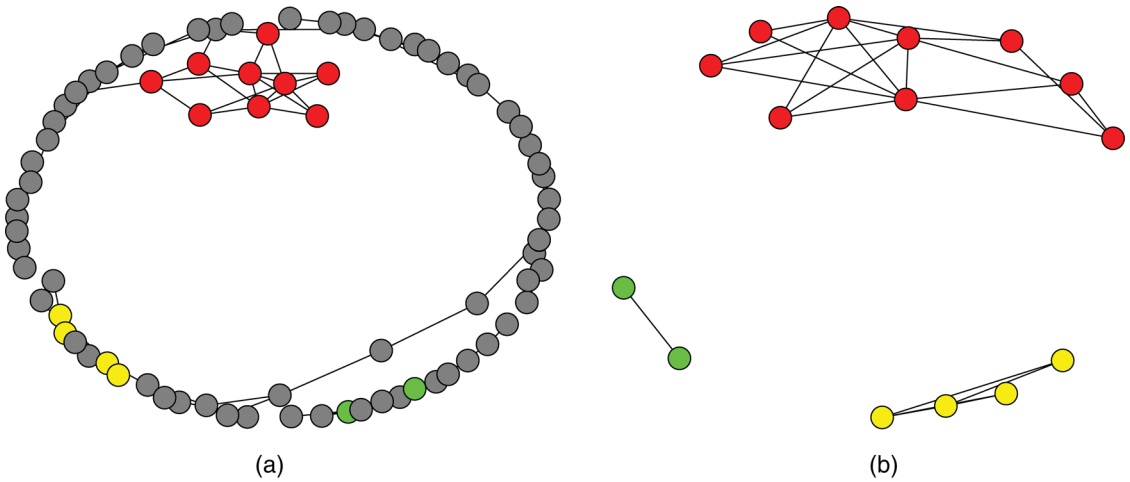

Figure 4. Detection linkages for a 2-d period of IMS data covering the eruption of Kelud volcano. Panel (a) represents all detection linkages; there are 16 separate graphs where each circle (grey, yellow, or red) represents an individual detection. Panel (b) represents detection linkages for detections with three or more two-station associations during the 2-d time period. Note that the graph comprising only two detections is constructed because each detection (i.e. green point in a and b) is connected to at least two other detections that are themselves connected with fewer detections.

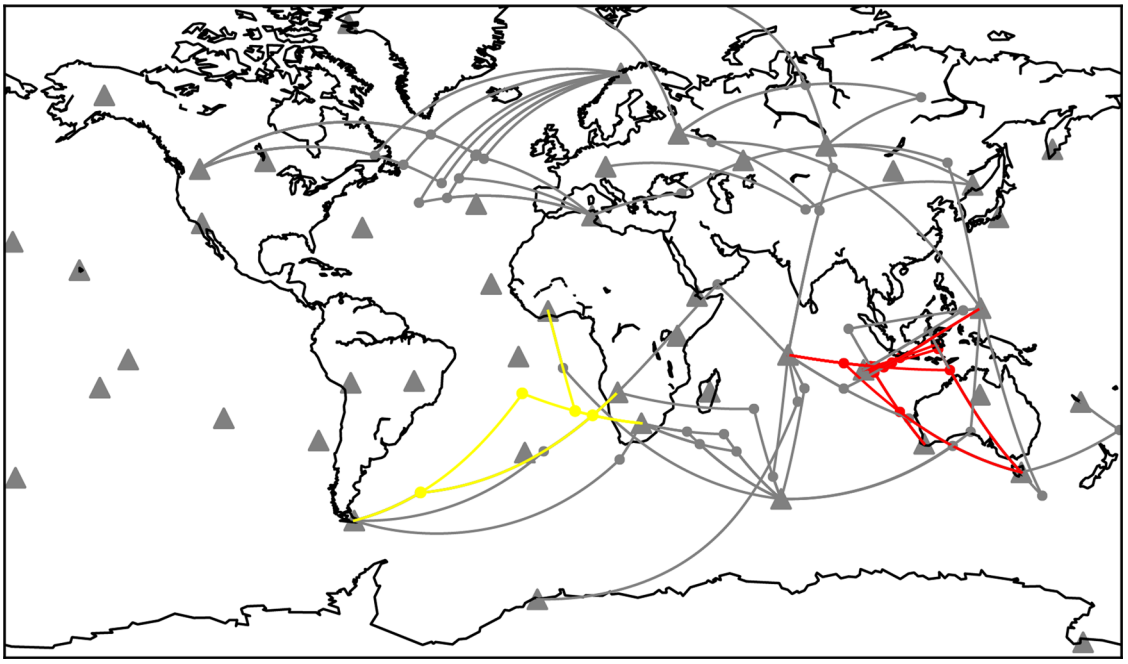

Figure 5. The geographic consistency of the graphs with nine detections and four detections in Fig. 4. In this plot, circles represent event hypotheses and lines represent detections. All associated detections for days 2014 February 13–2014 February 14 are shown in grey with detections and corresponding event hypotheses for the 9-detection graph and 4-detection graph denoted in red and yellow respectively. The 9-detection graph corresponds to the eruption of Kelud volcano, Indonesia on 2014 February 13.

time (corresponding to a uniform PDF on celerity from 0.22 to 0.34 km s^{-1} and an allowed tolerance on backazimuth of $\pm 15^{\circ}$ from the great circle backazimuth, calculated using standard spherical trigonometry equations.

4 ASSOCIATION SYNTHETIC TESTS

To assess the association method we perform tests to characterize (i) the probability of correct association (the probability of correctly associating signals from a real event, referred to as p_d) and (ii) the probability of false association (the probability of associating unrelated detections together, referred to as p_f). To characterize p_d we simulate synthetic events recorded by IMS stations. A given simulated event has a random location on the globe and a random origin time. The event generates signals at a random subset of IMS stations. The randomness is a crude simulation of source, path and site effects, which can affect the detection of a given source at any particular site. For the purposes of the association method outlined above, each signal is characterized by an arrival time and a backazimuth estimate. We can simulate these estimates by sampling a celerity and backazimuth at each station from probability density functions (PDF's). The PDF's are chosen to provide rough estimates of the expected distributions inside 2000 km, where any phase is possible, and beyond 2000 km where stratospheric returns are most likely. We define the backazimuth distributions at each station using

Table 2. Set of probability of detection parameters used in the Monte Carlo simulations.

Parameter set	Probability of detection <2000 km	Probability of detection 2000–10 000 km	Average number of stations detecting the event
1	1.0	0.5	11
2	0.5	0.5	10
3	0.5	0.4	8
4	0.5	0.3	6
5	0.5	0.2	4
6	0.5	0.1	2

a Gaussian PDF centred on the backazimuth of the great circle path for the randomly chosen location. The standard deviation in the backazimuth distribution is varied as outlined below. Inside 2000 km, the PDF for celerity is a uniform distribution between 0.22 and 0.34 km s^{-1}, while beyond 2000 km the PDF is a Gaussian distribution centred on 0.28 km s^{-1} with a standard deviation of 0.03 km s^{-1}.

We have performed a series of Monte Carlo simulations to simulate events covering a range of different sizes (event size is parametrized by the probability of detection as detailed in Table 2). For a given event size, five scenarios are considered where the PDF on backazimuth is varied such that the standard deviation can take five different values (0.5°, 3°, 5°. 7° and 15°). For each value of the standard deviation, 500×6 simulations are conducted and p_d is simply calculated from

$$p_d = \frac{\#\ \text{Synthetic Events Found}}{\#\ \text{Synthetic Events}}.$$

Thus, the total number of simulations performed is 6 (number of probability-of-detection parameters in Table 2) × 5 (number of backazimuth standard deviation parameters) × 500 (number of trials with the same set of parameters but different realizations) = 15 000.

A summary of the probability of association simulations are summarized in Fig. 6. For illustration, Fig. 7 provides an example event for a specific set of simulation parameters (a synthetic event at 53N, 27.5E, standard deviation in backazimuth of 6° and event size corresponding to Set 2 in Table 2). For this event, an initial event hypothesis is formed on the basis of detections at five stations (IS42, IS26, IS43, IS34 and IS11). These detections form an event hypothesis shown by the black star in the top right of Fig. 7. After searching for arrivals at the remaining stations, an additional four detections are added at IS10, IS57, IS56 and IS53. The results in Figs 6 and 7 indicate that the association methodology described in Section 3 is robust for high backazimuth deviations. The main factor affecting the probability of association is the number of stations detecting the event (a proxy for the size of the events). Large events have a high probability of association, consistent with our detection formulation that is designed to detect good SNR events.

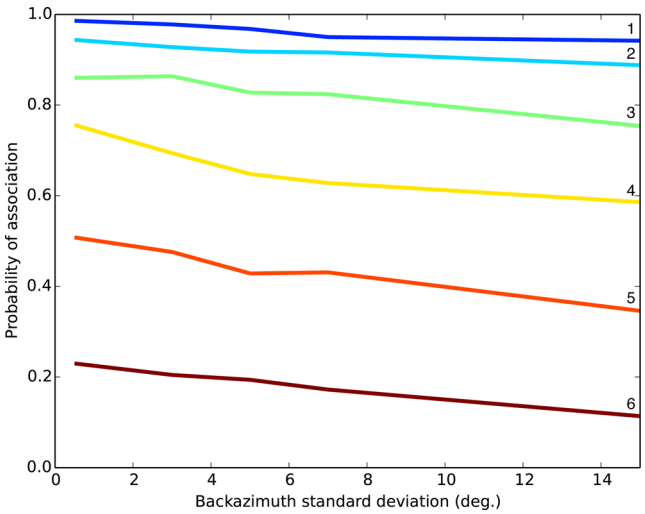

Figure 6. Analysis of the effects of backazimuth standard deviation and event size (numbers corresponding to coloured lines refer to rows in Table 2) on the probability of detection. Each point on one of the lines is estimated from 500 simulations.

To characterize the probability of false association, p_f, we performed a second series of Monte Carlo simulations that includes only unrelated detections. Here, we refer to unrelated detections to denote detections that are not generated from a single causative event. In this second series of simulations we generate sets of unrelated detections at each station in the IMS network in a 24-hr period. Each unrelated detection is random in arrival time and backazimuth. The number of unrelated detections at a given station is pulled from a uniform PDF with limits at 0 and N, where N is the maximum possible number of detections. This approach accounts for the fact that the number of unrelated detections at stations will vary in practice due to ambient noise conditions, local events and other effects. For this test, we define:

$$p_f = \frac{\#\ \text{Synthetic Events Found}}{\#\ \text{Monte Carlo Simulations}}.$$

The value of N is non-intuitive and so we consider the relationship between the probability of false association and the average number of unrelated detections across the IMS network for a given N. For each value of N we perform 500 Monte Carlo simulations. Fig. 8 shows the relationship between the average number of unrelated detections in a 24-hr time period and p_f. The likelihood of a given event being real must therefore be considered in the context of the number of detections in a given 24-hr time period, where 24 hr is the approximate time it takes for the slowest propagating infrasound waves to reach the antipode. For reference, the average number of detections per day across the network, estimated by processing all IMS data in 2010 using the Adaptive F-detector outlined in Section 2 is also plotted in Fig. 8.

5 LOCATION

Once one has obtained a set of associated arrivals, one can locate the event using the Bayesian Infrasound Source Location (BISL) algorithm described in detail by Modrak *et al.* (2010) and Marcillo *et al.* (2014). Because the method is described elsewhere, we provide only a brief summary here as relevant to IMS data processing. The BISL method is based on the construction of a posterior PDF that evaluates the likelihood of a set of models (represented by the vector **m**) for the source configuration given a set of measurements

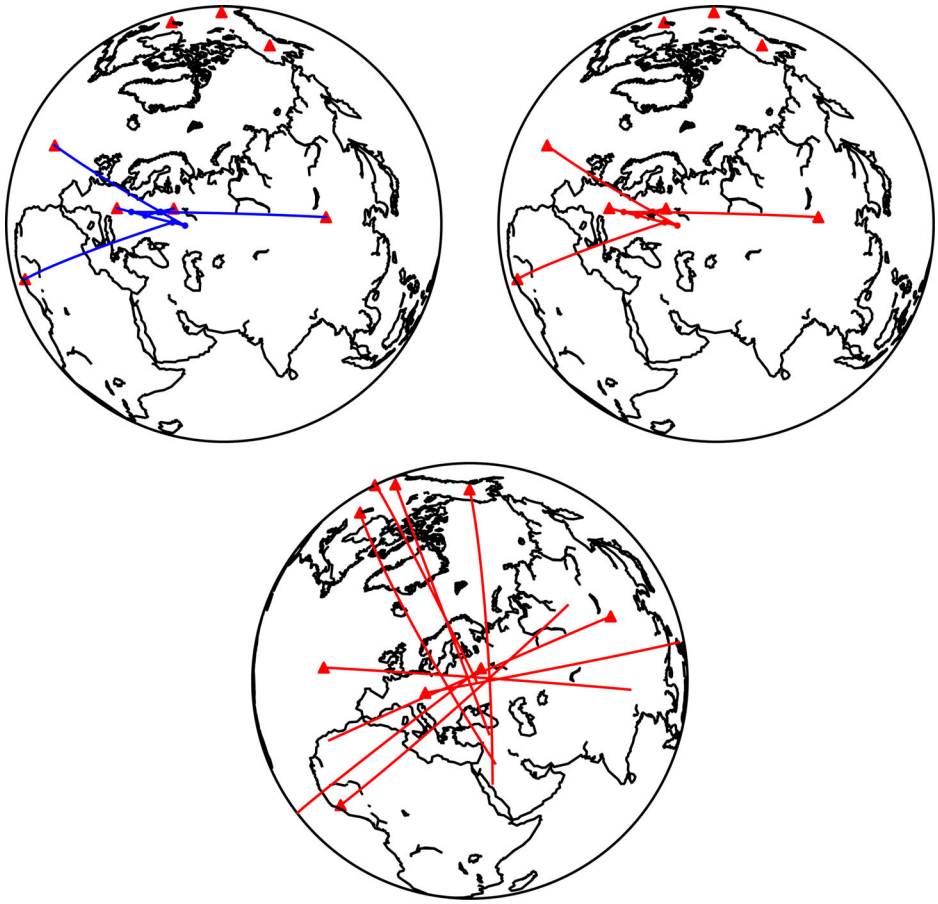

Figure 7. An example synthetic test for an event at 53°N, 27.5°E and backazimuth standard deviation of 6°. Stage 1 is denoted by the top left panel where individual two-array associations are denoted by blue dots and corresponding detections by blue lines, Stages 2 and 3 are denoted by the top right panel, where the red lines denote a single clustered event and the black star denotes the geographic mean of the intersection points. Stage 4 is denoted by the bottom panel where each line now denotes a detection that is associated with the seed event. For this specific simulation, a single nine-station association was identified from nine detections. The initial seed was formed by five stations and four stations were added in Stage 4.

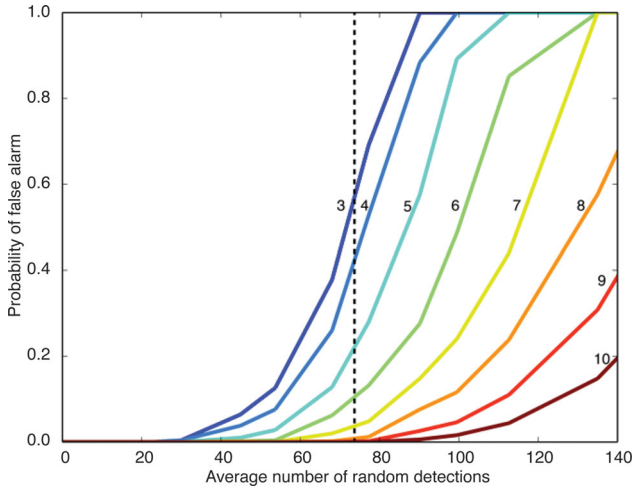

Figure 8. Analysis of the probability of false association, p_f, as a function of the minimum number of stations in a false association (each line represents the p_f of obtaining at an association of at least x stations). The average number of detections across the IMS network is indicated by the dashed line.

(represented by the vector **d**) compiled by the associator. The models are represented by a multidimensional parameter space that can include the source parameters, that is, spatial coordinates (x_0, y_0), origin time (t_0) and propagation/medium parameters, that is, celerity (v) and backazimuth deviation ($\Delta\theta$). The posterior PDF, $P(\mathbf{m}|\mathbf{d})$, can be written as:

$$P(\mathbf{m}|\mathbf{d}) = c(\mathbf{d})\, P(\mathbf{m})\, P(\mathbf{d}|\mathbf{m}), \tag{6}$$

where, $P(\mathbf{m})$ and $P(\mathbf{m}|\mathbf{d})$ are the prior and likelihood PDF's and $c(\mathbf{d})$ is a normalization function. For the implementation described in this work $\mathbf{m} = \{x_0, y_0, v, t_0\}$, and $\mathbf{d} = \{t_i, \theta_i\}$. $P(\mathbf{m}|\mathbf{d})$ is a product of backazimuth $\Theta(\theta_i|\mathbf{m})$ and arrival time $\Phi(t_i|\mathbf{m})$ constraints (Modrak *et al.* 2010) defined as follows:

$$P(\mathbf{d}|\mathbf{m}) = \prod_i^n \Theta(\theta_i|\mathbf{m})\Phi(\mathbf{t}_i|\mathbf{m}), \tag{7}$$

where n is the number of stations in the network. The separate components are defined as:

$$\Theta(\theta_i|\mathbf{m}) = \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left[-\frac{1}{2}\left(\frac{\gamma_i}{\sigma_\theta^2}\right)^2\right], \tag{8a}$$

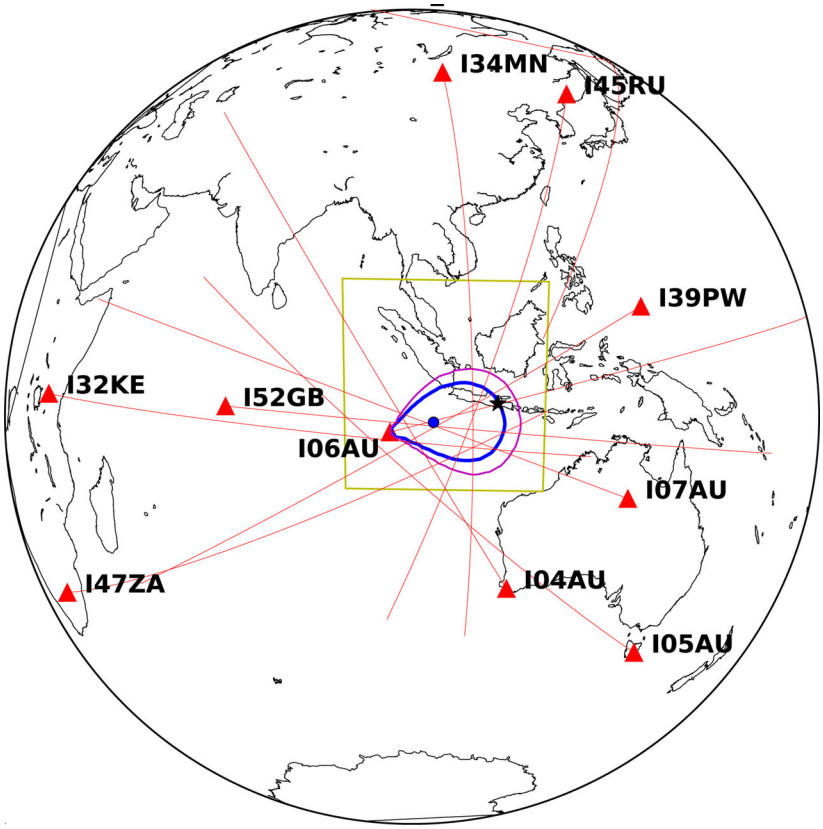

Figure 9. Automatic location solution for a 2-d period corresponding to 2014 February 13–2014 February 14 (shown in Fig. 5). The maximum *a posteriori* location is shown by the blue dot, 95 and 99 per cent credibility contours are shown by blue and magenta polygons, and detecting stations shown by red triangles with red lines the projections of estimated backazimuths along their great circle paths. The black star denotes the actual location of Kelud volcano and the yellow box denotes the BISL search region defined based on the association solution.

and

$$\Phi\left(\mathbf{t}_i|\mathbf{m}\right) = \frac{1}{\sqrt{2\pi\sigma_\phi^2}} \exp\left[-\frac{1}{2}\left(\frac{\varepsilon_i}{\sigma_\phi}\right)^2\right], \tag{8b}$$

where the differences between observed and predicted backazimuths and arrival times given by:

$$\gamma_i = \theta_i - \tilde{\theta}_i, \tag{9a}$$

and,

$$\varepsilon_i = t_i - \left(t_0 + \frac{d_i}{v}\right), \tag{9b}$$

where $\tilde{\theta}_i$ and d_i are the great-circle backazimuth and radial distance at station i for a given node, respectively, and σ_θ and σ_ϕ are the corresponding standard deviations. P (**m**) accounts for any potential prior information that can be used to constrain the parameters. See Marcillo *et al.* (2014) and Nippress *et al.* (2014) for a discussion of the implementation of model and data priors for enhancing the performance of BISL. BISL allows assigning probabilities to each point of the posterior PDF. Using the posterior PDF we can determine a point estimate or a credibility region.

The application of BISL to IMS data is illustrated in Fig. 9 for the automatic association corresponding to eruption of Kelud volcano that was discussed in Section 3. At the time of writing, propagation priors for this region of the globe have not been constructed

Table 3. Large infrasound events in the Reviewed Event Bulletin and results from this study. Note that the REB reported three events on 2014 December 13 that were associated with the eruption of Kelud.

Event #	Date	Latitude	Longitude	Description	# Associated detections in REB	# Associated detections in this study
1	02/07/2009	56.0779	76.7068	Russian mining explosion?	6	Not associated
2	10/08/2009	−4.2261	121.1155	Large bolide near Sulawesi	14	6
3	12/26/2010	40.0244	156.9966	Large bolide over North Pacific	11	8
4	04/18/2011	0.831	−126.5921	Large bolide over Central Pacific?	8	6
5	05/21/2011	64.688	−17.368	Eruption of Grimsvotn, Iceland	6	Not associated
6	05/05/2012	76.7401	−10.5816	Bolide east of Greenland?	8	4
7	10/22/2012	51.7055	117.1126	Russian mining explosion?	6	Not associated
8	02/15/2013	54.0563	61.8062	Large bolide over Chelyabinsk, Russia	17	5
9	10/18/2013	56.1074	160.9198	Eruption of Klyuchevsky, Kamchatka	6	7
10*	02/13/2014	−8.5353	109.7683	Eruption of Kelud, Indonesia	13	13

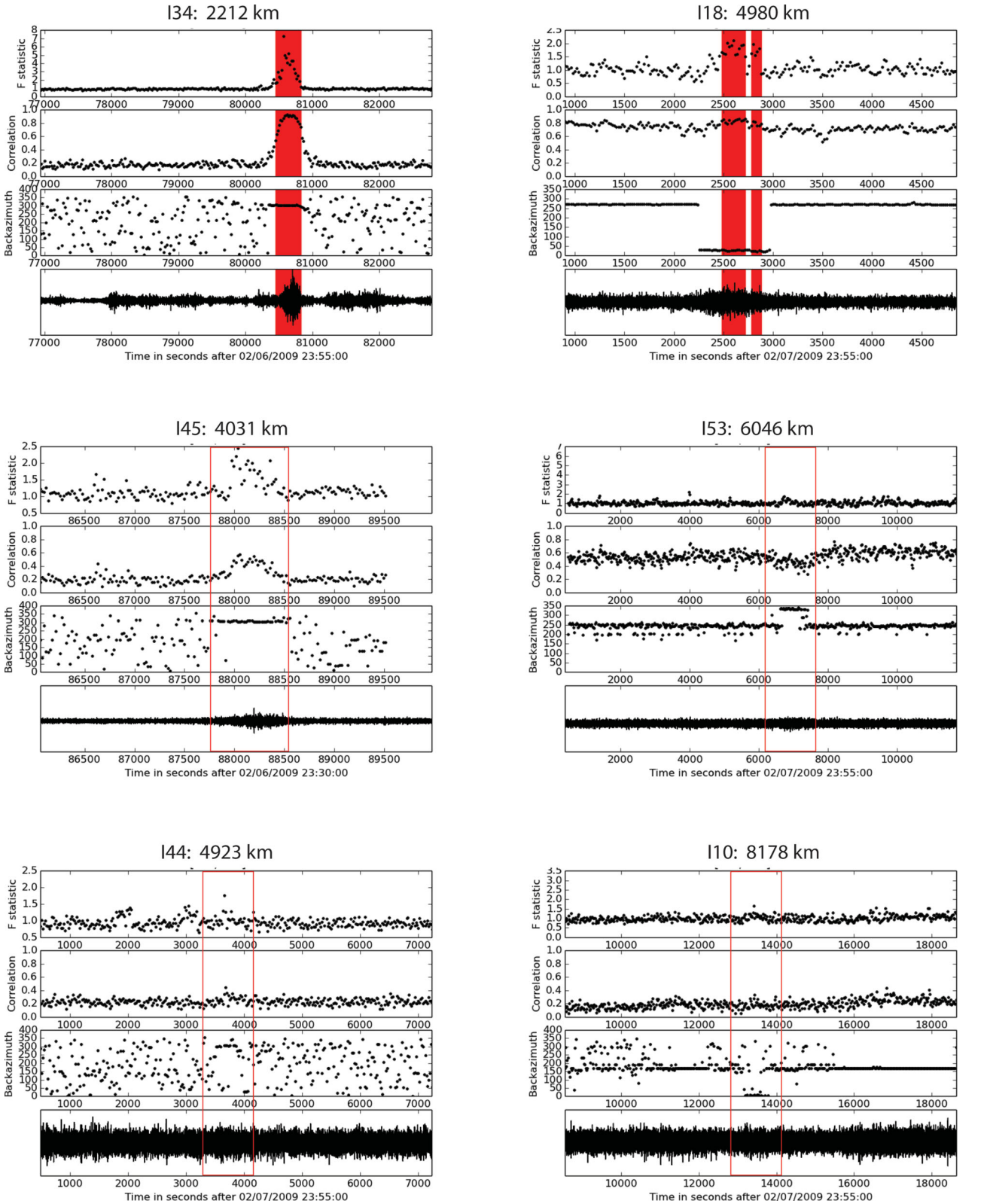

Figure 10. Results of our detection processing for each of the REB detections reported for Event 1 in Table 3. Each panel was processed using the 0.5–3 Hz band, results using the 0.01–0.5 Hz band did not declare any detections. Although we only obtained detections at two stations, I34MN and I18DK (red-shaded boxes), we did obtain direction-of-arrival estimates consistent with REB detections at the remaining four stations (see consistent backazimuths within the open red boxes at I45RU, I44RU, I53US and I10CA).

although the methods of their use have been presented by Marcillo *et al.* (2014), and a uniform celerity distribution has been used in estimating the source location. The resultant 95 and 99 per cent credibility contours are biased towards the west of the event (Fig. 9) due to crosswinds. Research is currently underway on the development of propagation priors for global scales enabling reduction in bias and improvement in the precision of the location estimates. The solution in Fig. 9 illustrates the final product of the full set of algorithms presented in this paper—a location and origin time estimate from the maximum *a posteriori* estimate and corresponding uncertainties at a specified credibility level.

6 RESULTS FROM PROCESSING LARGE EVENTS IN THE REB

The final outcome of International Data Centre processing is the REB, a catalogue of events that have been analyst-reviewed. The REB provides a useful set of reviewed events for testing the automatic detection and association methods outlined in this paper. Based on the estimated false-alarm rate at the association level (Fig. 8), we focus on events in the REB since 2009 with >5 infrasound arrivals (pertinent details about the events are summarized in Table 3) in order to provide a preliminary assessment of the true probability of missed detection and association with real data. By focusing on large events, the REB associations are expected to be well constrained.

For each of the 10 large REB events in Table 3, we processed the data at all available IMS stations in either a 24- or 48-hr period, depending on whether detections in the REB were contained within a single day or spanned two consecutive days. The data were processed automatically using the detection and association techniques described in this paper without modifying the input parameters described in this paper and with no analyst review. Of the 10 events, seven were automatically detected and associated (Table 3). In most cases the number of associated detections for the associated events was less than the number of associated detections in the REB. The undetected events constituted two suspected mining explosions in Russia and the eruption of Grimsvotn, Iceland that are associated with only six detections in the REB, at the low end for the current setup presented here.

Focusing on the three missed events, we performed a detailed analyst review of each REB detection in order to determine what caused our algorithms to miss the events. In each case, fewer than three of the detections listed in the REB were automatically formed by our processing. A more detailed summary of the detection analysis for one of these events, Event 1 in Table 3, is shown in Fig. 10. For this event, two automatic detections were obtained with very similar parameter estimates to the REB detections. However, at the remaining four stations with REB detections, our detector did obtain stable parameter estimates that matched the REB estimates but our F detector did not declare detections in either of the two frequency bands in Table 1. However, the stable estimates of backazimuth for many of these signals, which differs from the ambient background at I18DK, I53US and I10CA, suggests that using a complimentary detector based on this property, such as the Hough transform (Brown *et al.* 2008), may provide enhanced detection of low SNR signals. Such complimentary detectors could be used in the case where there are at least two strong detections (e.g. the detections at I34MN and I18DK for Event 1) or it could be implemented as a multivariate p-value detector such as has been done at local distances (Arrowsmith & Taylor 2013). In any case, we have not explored the development of automated multivariate detection for small events in this study.

7 CONCLUSIONS AND FUTURE RESEARCH NEEDS

Moving towards routine analysis of data streams requires robust automated techniques to detect, associate and locate transient events with a high probability of detection and low false alarm rate. Such techniques can contribute to our understanding of the number and distributions of global infrasound events, enabling us to develop more comprehensive data sets that will provide new insights on both source and propagation physics. For the nuclear monitoring mission faced by the CTBTO, such techniques are critical to reliably detect large atmospheric explosions.

This paper provides a probabilistic framework for processing large quantities of IMS infrasound data. The results of applying this framework to both synthetic and real data suggest that this framework performs well for detecting very large events recorded at five or more stations, but that further research is needed before we can robustly detect and associate small events that may only be recorded by a couple of stations. Indeed, the results shown in this paper highlights how increased station density would result in a dramatic reduction of the false alarm rate. The current IMS infrasound network is so sparse that the reliable detection and association of a 1 kT explosion, which may only be detected at one or two arrays, is beyond the reach of current methods without an unacceptably high false alarm rate.

At the detection level, the use of a multivariate detector that uses multiple properties in the waveform such as stability in backazimuth and power may enable the detection of lower SNR events. At the association level, the use of additional physical constraints is needed to increase the probability of association and decrease the probability of false alarm for small events. The current implementation uses only arrival time and backazimuth constraints. Including additional constraints from atmospheric propagation modelling is left for future research. At present, our group at LANL is exploring methodologies to encapsulate propagation catalogues in probabilistic models, building on work by Marcillo *et al.* (2014). In principle, by using both full wave and geometric models it should be possible to incorporate constraints on frequency, amplitude and possibly other physical parameters. We note that the CTBTO currently does include some constraints on frequency for event association using observations of ground-truth events (Brachet *et al.* 2010). We believe, however, that the current ground-truth data set is inadequate for capturing the true stochastic properties of the atmosphere. We advocate for a model-based approach, benchmarked using data where available. We are beginning to explore such an approach for the development of celerity and backazimuth deviation priors using ray theory, validated with observations in the Western United States (Morton & Arrowsmith 2014).

ACKNOWLEDGEMENTS

Thanks to David Green at AWE/Blacknest, Kyle Jones at Sandia, Dale Anderson and Josh Carmichael at LANL and Dean Clauter at AFTAC for helpful discussions.

REFERENCES

Arrowsmith, S.J. & Taylor, S.R., 2013. Multivariate acoustic detection of small explosions using Fisher's combined probability test, *J. acoust. Soc. Am.,* **133**(3), doi:10.1121/1.4789871.

Arrowsmith, S.J. *et al.*, 2008. Regional monitoring of infrasound events using multiple arrays: application to Utah and Washington State, *Geophys. J. Int.,* **175,** 291–300.

Arrowsmith, S.J., Whitaker, R.W., Katz, C. & Hayward, C., 2009. The *F*-detector revisited: an improved strategy for signal detection at seismic and infrasound arrays, *Bull. seism. Soc. Am.,* **99,** 449–453.

Arrowsmith, S.J., Marcillo, O. & Drob, D.P., 2013. A framework for estimating stratospheric wind speeds from unknown sources, *Geophys. J. Int.,* **195,** 491–503.

Arrowsmith, S.J., Norris, D., Whitaker, R. & Anderson, D., 2014. Sources of error model and progress metrics for acoustic/infrasonic analysis: location estimation, *Pure appl. Geophys.,* **171,** 587–597.

Blandford, R.R., 1974. An automatic event detector at the Tonto forest seismic observatory, *Geophysics,* **39**(5), 633–643.

Brachet, N. & Brown, D., LeBras, R., Cansi, Y., Mialle, P. & Coyne, J., 2010. Monitoring the Earth's atmosphere with the global IMS infrasound network, in *Infrasound Monitoring for Atmospheric Studies,* pp. 77–118, eds Le Pichon, A., Blanc, E. & Hauchecorne, A., Springer.

Briels, S.A., 2010. *Infrasound Source Location,* p. 90, Delft University of Technology.

Brown, D., Whitaker, R., Kennett, B.L.N. & Tarlowsk, C., 2008. Automatic infrasonic signal detection using the Hough transform, *J. geophys. Res.,* **113,** doi:10.1029/2008JD009822.

Christie, D.R. & Campus, P., 2010. The IMS infrasound network: design and establishment of infrasound stations, in *Infrasound Monitoring for Atmospheric Studies,* pp. 29–75, eds Le Pichon, A., Blanc, E. & Hauchecorne, A., Springer.

Kay, S.M., 2008. *Fundamentals of Statistical Signal Processing Volume II: Detection Theory,* Prentice Hall, 560 pp.

Krim, H. & Viberg, M., 1996. Two decades of array signal processing research: the parametric approach, *IEEE Signal Process. Mag.,* **13**(4), 67–94.

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.

Le Pichon, A., Ceranna, L. & Vergoz, J., 2012. Incorporating numerical modeling into estimates of the detection capability of the IMS infrasound network, *J. geophys. Res.,* **117**(D5), doi:10.1029/2011JD016670.

Le Pichon, A., Ceranna, L., Pilger, C., Mialle, P., Brown, D., Herry, P. & Brachet, N., 2013. The 2013 Russian fireball largest ever detected by the CTBTO, *Geophys. Res. Lett.,* **40,** doi:10.1002/grl.50619.

Marcillo, O., Arrowsmith, S.J., Whitaker, R., Anderson, D., Nippress, A., Green, D. & Drob, D.P., 2014. Using physics-based priors in a Bayesian algorithm to enhance infrasound source location, *Geophys. J. Int.,* **196,** 375–385.

Modrak, R.T., Arrowsmith, S.J. & Anderson, D.N., 2010. A Bayesian framework for infrasound location, *Geophys. J. Int.,* **181**(1), 399–405.

Morton, E.A. & Arrowsmith, S.J., 2014. The Development of Global Probabilistic Propagation Look-Up Tables for Infrasound Celerity and Back-Azimuth Deviation, *Seism. Res. Lett.*, **85**(6), doi:10.1785/0220140124.

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

Nippress, A., Green, D., Marcillo, O. & Arrowsmith, S.J., 2014. Generating regional infrasound celerity-range models using ground-truth information and the implications for event location, *Geophys. J. Int.,* **197**(2), 1154–1165.

Park, J., 2013. Infrasound signal processing from regional arrays and seismic characteristics of North Korean nuclear explosions, *PhD thesis,* Southern Methodist University, Dallas, TX.

Park, J., Stump, B., Hayward, C., Arrowsmith, S.J. & Blom, P., 2014. Automatic infrasound detection and location of sources in the Western United States, *J. geophys. Res.,* doi:10.1002/2013JD021084.

Shumway, R.H., 2002. Detection and location capabilities of multiple infrasound arrays, Final Scientific Report. DTRA01-00-C0082, Defense Threat Reduction Agency.

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.

Shumway, R.H., Smart, E. & Clauter, D.A., 2008. Mixed signal processing for regional and teleseismic arrays, *Bull. seism. Soc. Am.,* **98**(1), 36–51.

Szuberla, C.A.L. & Olson, J.V., 2004. Uncertainties associated with parameter estimation in atmospheric infrasound arrays, *J. acoust. Soc. Am.,* **115**(1), 253–258.

van Veen, B. & Buckley, K.M., 1998. Beamforming: a versatile approach to spatial filtering, *IEEE ASSP Mag.,* **5**(2), 4–24.

West, D.B., 2000. *Introduction to Graph Theory,* Pearson.