

RESEARCH ARTICLE

10.1002/2016JB013356

Key Points:

- Method to detect and locate infrasound from sustained explosive volcanic eruptions
- Uses global International Monitoring System infrasound data
- Global, multiyear infrasonic catalog of sustained explosive volcanic eruptions

Supporting Information:

- Supporting Information S1
- Movie S1

Correspondence to:

R. S. Matoza,
matoza@geol.ucsb.edu

Citation:

Matoza, R. S., D. N. Green, A. Le Pichon, P. M. Shearer, D. Fee, P. Mialle, and L. Ceranna (2017), Automated detection and cataloging of global explosive volcanism using the International Monitoring System infrasound network, *J. Geophys. Res. Solid Earth*, 122, doi:10.1002/2016JB013356.

Received 13 JUL 2016

Accepted 17 MAR 2017

Accepted article online 24 MAR 2017

Automated detection and cataloging of global explosive volcanism using the International Monitoring System infrasound network

Robin S. Matoza¹ , David N. Green² , Alexis Le Pichon³ , Peter M. Shearer⁴ , David Fee⁵ , Pierrick Mialle⁶, and Lars Ceranna⁷ 

¹Department of Earth Science and Earth Research Institute, University of California, Santa Barbara, California, USA, ²AWE Blacknest, Reading, UK, ³CEA, DAM, DIF, Arpajon, France, ⁴Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA, ⁵Wilson Alaska Technical Center and Alaska Volcano Observatory, University of Alaska Fairbanks, Fairbanks, Alaska, USA, ⁶CTBTO, Vienna, Austria, ⁷BGR, Hannover, Germany

Abstract We experiment with a new method to search systematically through multiyear data from the International Monitoring System (IMS) infrasound network to identify explosive volcanic eruption signals originating anywhere on Earth. Detecting, quantifying, and cataloging the global occurrence of explosive volcanism helps toward several goals in Earth sciences and has direct applications in volcanic hazard mitigation. We combine infrasound signal association across multiple stations with source location using a brute-force, grid-search, cross-bearings approach. The algorithm corrects for a background prior rate of coherent unwanted infrasound signals (clutter) in a global grid, without needing to screen array processing detection lists from individual stations prior to association. We develop the algorithm using case studies of explosive eruptions: 2008 Kasatochi, Alaska; 2009 Sarychev Peak, Kurile Islands; and 2010 Eyjafjallajökull, Iceland. We apply the method to global IMS infrasound data from 2005–2010 to construct a preliminary acoustic catalog that emphasizes sustained explosive volcanic activity (long-duration signals or sequences of impulsive transients lasting hours to days). This work represents a step toward the goal of integrating IMS infrasound data products into global volcanic eruption early warning and notification systems. Additionally, a better understanding of volcanic signal detection and location with the IMS helps improve operational event detection, discrimination, and association capabilities.

1. Introduction

How often do the world's volcanoes erupt and with what magnitude? Attempts to keep track of Earth's presently active volcanism compile geological and geophysical data, reports, and publications from multiple sources [e.g., *Simkin and Siebert*, 2000; *Siebert and Simkin*, 2002; *Siebert et al.*, 2010; *Mason et al.*, 2004; *Croweller et al.*, 2012; *Global Volcanism Program*, 2013]. However, many potentially active volcanoes are not equipped with dedicated local monitoring instruments and are only observed using satellite-based methods, which may be limited, e.g., by meteorological cloud cover [*Webley and Mastin*, 2009]. Volcanic activity may be underreported or poorly quantified in remote regions, which lack ground-based instrumentation and observers.

We explore the potential use of the global International Monitoring System (IMS) infrasound network, currently under construction, to build a quantitative acoustic catalog of current global explosive volcanic activity to complement other data sources. The IMS infrasound network is designed to detect atmospheric explosions with an equivalent yield of 1 kt of TNT anywhere on the planet with at least two stations as part of the verification measures for the Comprehensive Nuclear-Test-Ban Treaty (CTBT) [*Christie and Campus*, 2010]. However, simulations show that long-range stratospheric ducting of infrasound significantly enhances the detection capability, with explosions of a couple hundred tons TNT equivalent being detectable globally at two or more stations on the complete network [*Le Pichon et al.*, 2009; *Green and Bowers*, 2010]. Data holdings from the IMS infrasound network go back to about 2002 and at present the IMS has 49 certified stations out of a planned total 60; data from 41 stations from 2005 to 2010 are used in this study (Figure 1). Each infrasound station consists of an array of at least four infrasonic sensors with a flat response from about 0.01 to 8 Hz

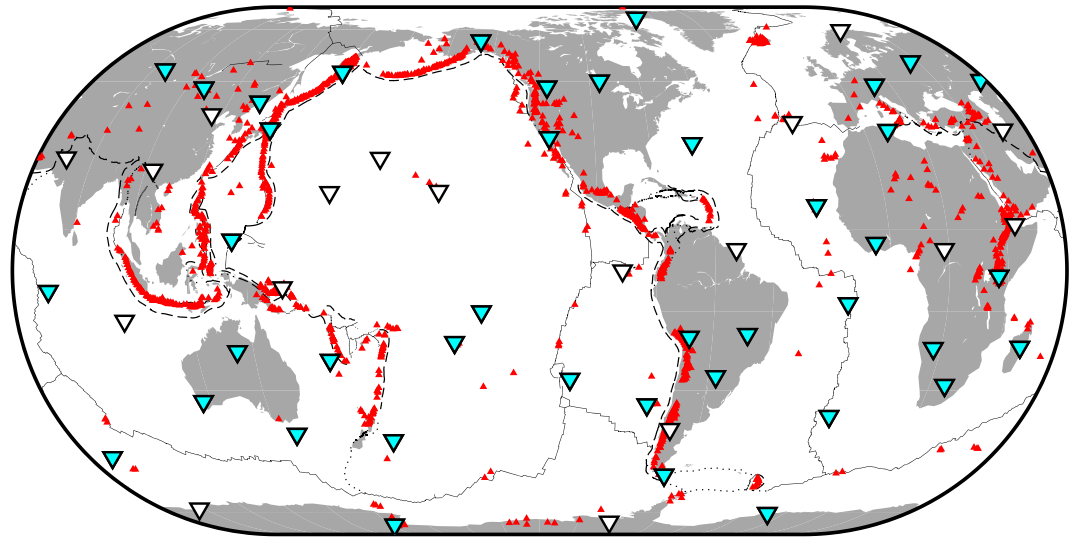


Figure 1. Global potentially active holocene (past 10,000 years) volcanoes (red triangles) [Siebert and Simkin, 2002; Global Volcanism Program, 2013], tectonic plate boundaries (black lines) [Coffin et al., 1998], and the IMS infrasound network (inverted triangles) [Christie and Campus, 2010]. Planned stations (59) of the full 60-station IMS infrasound network are shown in white, while the 41 stations considered in this study are shown in cyan (see section 3.1).

(sampled at 20 Hz; IMS minimum requirement states that the response should be flat between 0.02 and 4 Hz) and a sensitivity of about 0.1 mPa per count. The average station spacing for the complete network will be 1920 km in the Northern Hemisphere and 2027 km in the Southern Hemisphere [Christie and Campus, 2010]. Explosive volcanic eruptions are among the most powerful sources of infrasound observed on Earth, and infrasound is efficiently ducted in atmospheric waveguides with low attenuation. Consequently, infrasound recordings of explosive eruptions are routinely made at ranges of hundreds to thousands of kilometers on the IMS [Fee and Matoza, 2013, and references therein].

Recent studies have used IMS infrasound data to detect, locate, and provide detailed chronologies of remote explosive volcanism for individual eruption case studies [Fee et al., 2010b; Matoza et al., 2011a, 2011b; Green et al., 2013; Fee et al., 2013; Caudron et al., 2015]. Dabrowa et al. [2011] investigated the use of the IMS infrasound network for detecting explosive volcanism globally. They worked forward from the Smithsonian Institution Global Volcanism Program (GVP) database [Siebert and Simkin, 2002; Global Volcanism Program, 2013], systematically but manually examining the relevant infrasound data recorded at the time of known eruptions in the GVP database. Here we build on these efforts and attempt to go backward from the infrasound data to identify possible volcanic eruption signals. We perform automated searches through multiyear IMS infrasound station detection lists to identify signals from sustained explosive volcanic eruptions.

Explosive eruptions can also inject large volumes of ash into heavily travelled aviation routes, posing a significant societal and economic hazard [Casadevall et al., 1994; Casadevall and Thompson, 1995; Mazzocchi et al., 2010]. In the context of previous work developing regional volcano-acoustic early warning systems [Kamo et al., 1994; Garces et al., 2008; Matoza et al., 2007; Fee et al., 2010a; De Angelis et al., 2012; Olivieri et al., 2013], the present work represents a step toward the goal of integrating IMS data products into global volcanic eruption early warning and notification systems. A global infrasonic volcanic notification system is currently being developed in support of the Volcanic Ash Advisory Centers (VAACs) designated by the International Civil Aviation Organization [Mialle et al., 2015].

2. Motivation for the Approach

Methodologies to detect and locate infrasound sources using multiple infrasound arrays contain the following basic elements: (1) signal detection and estimation of wavefront parameters on the individual infrasound arrays (e.g., IMS stations) using array processing to produce a detection list for each station, (2) grouping and selection of signals of interest and removal of unwanted repetitive coherent infrasound signals from the list, (3) phase identification, (4) association of candidate phases across multiple stations for an assumed

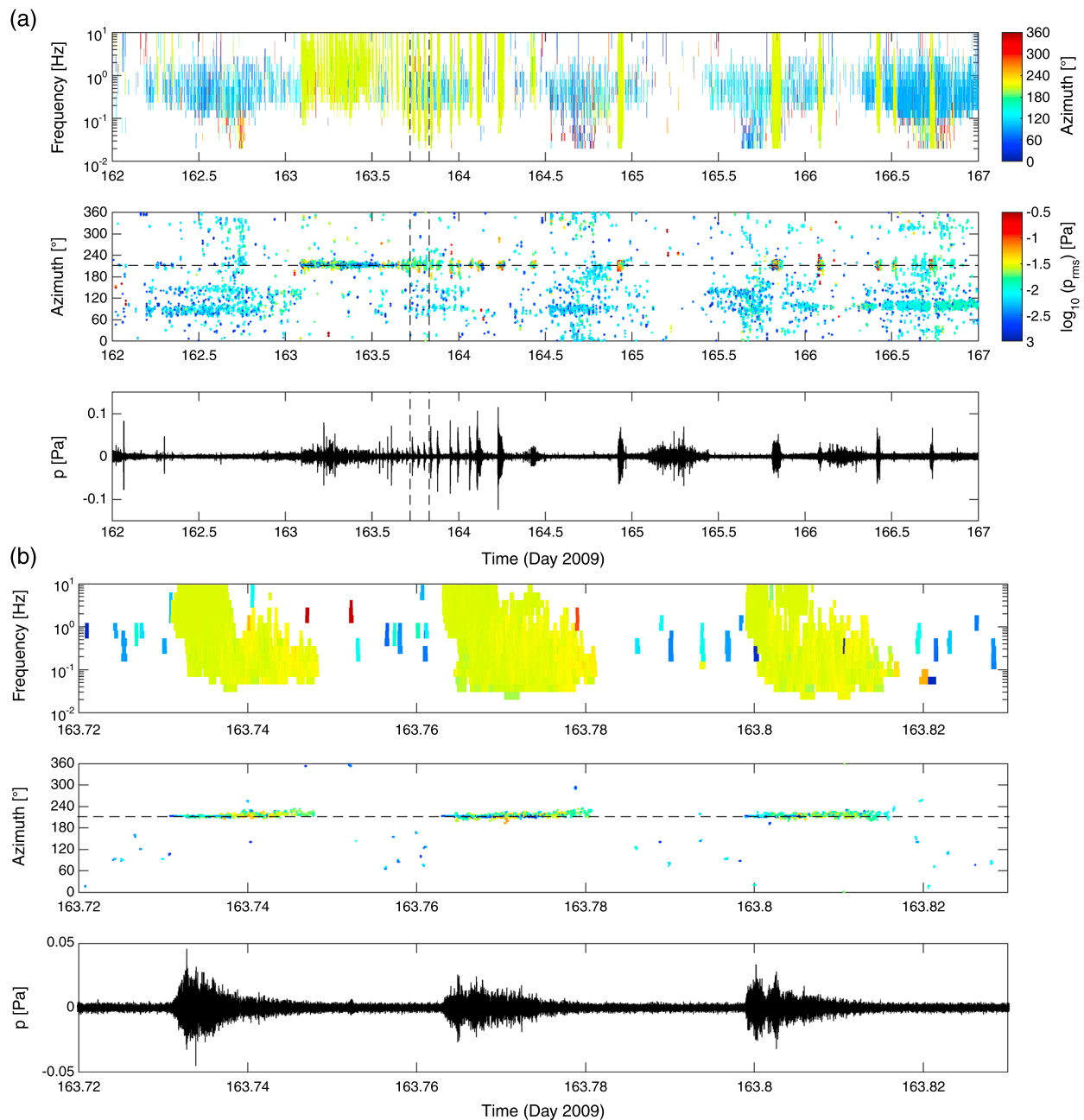


Figure 2. PMCC processing of IS44 (Kamchatka, range 643 km) array data from the June 2009 eruption of Sarychev Peak, Kuriles [Matoza et al., 2011a]. (a) Five days: 11–16 June 2009, or Days 162–167 2009. (bottom) Infrasonic waveform time-domain beam (0.5–5 Hz). (middle) PMCC pixel backazimuth variation with time, with symbol colors representing pixel root-mean-square amplitude (log scale, see color bar); the dashed line at 211.6° is the true backazimuth to Sarychev Peak. (top) Rectangles show the time and frequency space spanned by each pixel, and colors represent the pixel backazimuth (lime green is 211.6° backazimuth to Sarychev Peak). The pixels overlap in time with time-steps 10% of the window length; only the start and end times of pixel sequences are visible in this representation. The vertical dashed lines show the time interval expanded in Figure 2b. (b) Expanded view showing three individual explosive eruption events. The azimuth deviation from true observed at a station depends on the range from the source, propagation path taken, and magnitude and variability of the crosswinds, among other factors. The azimuth deviations in this example are on the order of $\pm 15^\circ$ [Matoza et al., 2011a].

propagation model, (5) source location, and (6) further source parameter estimation. We note that the processing performed in (1) aims to isolate correlated infrasonic signals from uncorrelated wind noise. Previously proposed approaches have combined all or subsets of these elements in different ways [e.g., Brown et al., 2002; Evers and Haak, 2005; Le Pichon et al., 2008; Arrowsmith et al., 2008; Arrowsmith and Whitaker, 2008; Brachet et al., 2010; Modrak et al., 2010; Park et al., 2015; Arrowsmith et al., 2015].

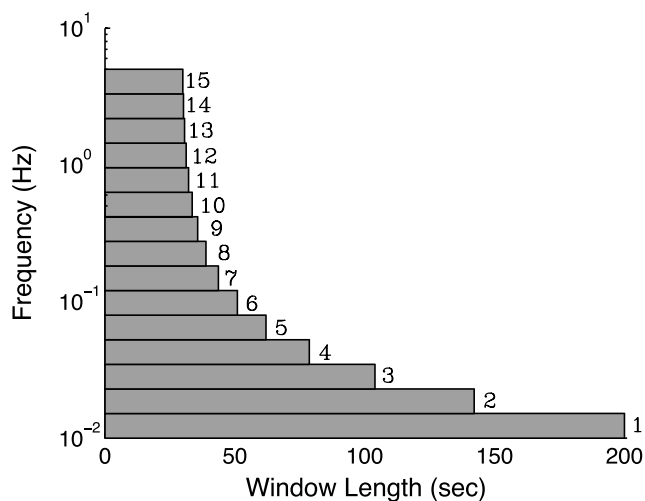


Figure 3. The 15 log-spaced frequency bands and time-window lengths used for PMCC processing in this study (see Matoza *et al.* [2013b, Table 1] for more details). The windows advance through the waveform data with a time-step of 10% of the window length.

Atmospheric variability makes phase identification (3) and signal association (4) challenging [Brown *et al.*, 2002]. Usually, the assumed source is an impulsive transient; thus, the phase identification, timing accuracy, and travel time corrections are all critical for providing confidence in multistation signal associations. Celerity is defined as the total range travelled divided by the propagation time. Travel times may be estimated from an assumed constant celerity [Le Pichon *et al.*, 2008], group velocity range [Arrowsmith and Whitaker, 2008; Modrak *et al.*, 2010; Nippres *et al.*, 2014], or computed using ray tracing with atmospheric specifications [e.g., Garcés *et al.*, 1998; Virieux *et al.*, 2004; Drob *et al.*, 2010; Matoza *et al.*, 2011a; Blom *et al.*, 2015]. We propose a different approach to signal detection, association, and location that is targeted specifically toward signals from volcanic eruptions and where the celerity assumption is less critical.

We exploit the following a priori assumed characteristics of volcanic infrasound, based on knowledge from previous studies [e.g., Matoza *et al.*, 2009; Fee *et al.*, 2010b; Matoza *et al.*, 2011a, 2011b; Dabrowa *et al.*, 2011; Fee and Matoza, 2013]:

1. Explosive volcanic eruptions represent fixed point source locations. This is different to the case of microbaroms, which have a large spatially distributed and temporally evolving source region [Landès *et al.*, 2012] and which represent a dominant source of coherent unwanted infrasound signals [Matoza *et al.*, 2013b].
2. Large sustained explosive eruptions typically produce long-duration volcanic jet noise signals (continuous infrasound signals lasting from hours to days [e.g., Matoza *et al.*, 2013a]) or sequences of multiple shorter duration signals (e.g., a series of explosions over several days lasting a few minutes to hours in duration each [e.g., Matoza *et al.*, 2011a]). At a given array, we therefore expect a large number of signal detections coming from the same backazimuth (varying slightly due to atmospheric propagation effects) throughout an eruption (Figure 2).
3. Large sustained explosive eruptions tend to produce broadband signals, even at long ranges (<5000 km) with efficient tropospheric and/or stratospheric ducting. This generates a large number of detections when array processing is performed as a time-frequency grid search (Figures 2 and 3); thus, eruption infrasound spans a larger area in time-frequency space compared to more band-limited signals such as microbaroms. We note that tropospheric ducts are more unstable than stratospheric ducts; long-range propagation (>1000 km) usually involves stratospheric ducting [e.g., Le Pichon *et al.*, 2009].

We work with array processing detection lists available for each IMS station (section 3.1). By working with station detection lists rather than raw time series data [e.g., Walker *et al.*, 2011], we greatly reduce the computational burden, reduce wind noise sensitivity, and enable fast processing of multiyear data for all available stations of the IMS infrasound network. Alternative network-based approaches have used raw waveform data and performed reverse time migration (RTM) using the waveform or waveform envelope from denser

networks of single sensors [e.g., Walker *et al.*, 2011]. In this case, the waveform may consist of wind noise or real infrasound signal. By working with array processing detection lists, wind noise is excluded.

3. Method

3.1. Array Processing Detection Lists

We start with Progressive Multichannel Correlation (PMCC) array processing results [Cansi, 1995; Cansi and Klinger, 1997]. PMCC is the array processing algorithm in use by the International Data Center (IDC) of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO). The PMCC detection lists used in this study [Matoza *et al.*, 2013b] result from a systematic global reprocessing of waveform data from all IMS infrasound stations available at the IDC from 1 April 2005 to 31 December 2010 using a log frequency band configuration (0.01–5 Hz) that is more suitable for broadband volcanic signals [Matoza *et al.*, 2011a]. The data used for this study are from the 39 IMS infrasound stations considered by Matoza *et al.* [2013b] (see their Figure 1) plus two additional stations: IS56 (Newport, WA) and IS57 (Piñon Flat, CA) that have subsequently become available for reprocessing. Since the IMS network is under construction, station availability varies during the time period considered from 2005 to 2010, with the number of operating stations gradually increasing with time (see supporting information Figure S1).

PMCC estimates wavefront parameters (e.g., backazimuth, apparent velocity, and root-mean-square (RMS) amplitude) of coherent plane waves using correlation time delays between successive array element triplets or subnetworks [Cansi, 1995]. PMCC performs a grid search for coherent signals in advancing time-windows over a set of frequency bands defined with band-pass filters, forming a pixelated time-frequency representation of coherent infrasound (Figures 2 and 3). PMCC records the wavefront properties of the dominant coherent arrival in a given time-window and frequency-band pair as a *pixel*. PMCC then groups pixels with similar wavefront properties into *families*, suppressing spuriously correlated and ungrouped pixels [Cansi and Klinger, 1997]. To speed the search, we work with summary output files containing only family-level information (summary properties of all contributing pixels), which consist of a list of detections for each station with a start and end time; backazimuth; minimum, mean, and maximum frequency (statistics are based on all pixels contributing to a family); apparent velocity; RMS amplitude; and number of pixels contributing to the family. The processing of Matoza *et al.* [2013b] uses a minimum and maximum of 10 and 50 pixels, respectively, per family. We note that our subsequent algorithm is not limited to the PMCC method and could easily be adapted to use output from a different array processor (e.g., the F detector [Blandford, 1974]). We also note that the IDC PMCC processing is designed for an operational system with a coarser time-frequency parameterization than used for the reprocessing of Matoza *et al.* [2013b] (Figure 3). In particular, the IDC uses a low-frequency limit of 0.07 Hz and only 11 bands with a maximum window length of 60 s (compare with Figure 3).

3.2. Brute-Force, Grid-Search, Cross-Bearings Method

Our method combines infrasound signal association and source location using a brute-force, grid-search, cross-bearings approach (Figure 4). Our method is based loosely on the methods of Landès *et al.* [2012] and Matoza *et al.* [2011b]. We take advantage of the assumed signal characteristics of infrasound from sustained explosive volcanic eruptions (section 2), employing long-duration (e.g., 2-day) detection averaging to give more weight to sustained signals and achieve robustness with respect to assumed celerity. We employ several measures to reduce sensitivity to unwanted infrasound signals such as from microbaroms and persistent local infrasound sources [Matoza *et al.*, 2013b]. Our method includes the removal of a prior model of unwanted persistent coherent infrasound signals (clutter) from all stations globally (section 3.3). The full workflow is illustrated in Figure 5. Values of adjustable parameters used for the results presented in this study are stated here in parentheses. If values are varied for different processing runs, they are listed in Table 1. Our algorithm works by building grid functions, as follows:

1. We define a grid of trial source nodes on the Earth's surface with a regular spacing in latitude and longitude (0.2°) (Figure 4). We define a constant association celerity v (0.33 km/s or 0.3 km/s; Table 1) and a backazimuth deviation tolerance ψ_{dev} (2° , 5° , or 10° used in different cases, see section 3.3 and Table 1). We define a desired time interval τ , in which to perform the search (2 or 10 days, see section 3.3).
2. We read the station detection lists and station coordinates available for the time interval $\tau + \beta$, where β is a buffer of additional time to allow for propagation delay ($\beta > v/d_{\text{max}}$, where d_{max} is the maximum distance considered). We assume that a station is functioning normally and is available if the station detection list is not empty for this time (this is a simple assumption that should be refined in future work to account for, e.g.,

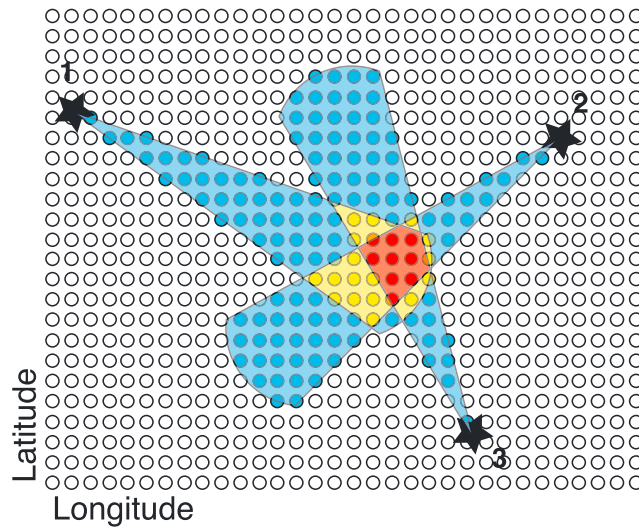


Figure 4. Cartoon illustrating the brute-force, grid-search, cross-bearings algorithm. Three stations are indicated by stars, and trial source locations on the Earth’s surface are indicated by circles. The figure shows one array detection (e.g., a PMCC pixel) at each station. The backazimuth of the detection is used to define a sector on the Earth within a specified azimuth deviation tolerance and for a maximum great-circle distance. In this example, blue, yellow, and red trial source nodes (grid points) have values of 1, 2, and 3, respectively. The numerical values of each grid point accumulate as further detections are considered from all stations.

variability in station noise levels and station data quality, e.g., timing errors that may degrade array processing results). We parse the station detection lists using a specified minimum, maximum, and mean frequency (f_{\min} , f_{\max} , and f_{mean} , Table 1). Pixels contributing to PMCC families define f_{\min}^{fam} , f_{\max}^{fam} , and $f_{\text{mean}}^{\text{fam}}$ for the family in our input detection lists, and we require that $f_{\min}^{\text{fam}} \geq f_{\min}$, $f_{\max}^{\text{fam}} \leq f_{\max}$; and $f_{\text{mean}}^{\text{fam}} \leq f_{\text{mean}}$ (all three conditions must be satisfied). Values of f_{\min} and f_{\max} (Table 1) are chosen to capture broadband detections (e.g., based on eruption signals from Sarychev Peak [Matoza et al., 2011a]); signals with f_{mean} above about 2 Hz are less likely to have propagated long range and are more likely to be local sources [Brachet et al., 2010]. We experiment (Table 1) with setting $f_{\min} = 0.1$ Hz to reduce the effect of known persistent sources, primarily microbaroms (microbarom frequency range extends below 0.1 Hz). We note that these constraints do not necessarily permit broadband signals only; provided the conditions are satisfied, a narrowband signal will pass but will have lower weight in the algorithm because of the small number of contributing pixels.

3. For each time interval τ , we perform a new grid search over the latitude-longitude grid of trial source nodes. For every trial source node, we sort all available stations with respect to distance from the trial source node. We then progress through the stations, starting with the closest.
4. For each station, we proceed through all detections. We time adjust the detections with the assumed celerity given the trial source-receiver distance. In a grid function $\mathbf{G}(\tau)$, we count the number of detections (number of pixels) that can be associated with the trial source node from the station during the time interval τ , within the allowed azimuth deviation and a maximum distance d_{\max} (5000 km) from each station (Figure 4). We sum the results from all stations. The values in $\mathbf{G}(\tau)$ gradually accumulate as the corresponding trial source nodes are associated with more detections from more stations.
5. As we progress through the stations, we also keep track of the number of pixels linking to a trial source node from the current station. We define a minimum number of pixels m to consider that a station is *linked* with a trial source node (we use $m = 500$ for most runs; Table 1). If the nearest n^{near} functioning stations (n^{near} varied between 0, 1, and 2; Table 1) do not link with the current trial source node, the trial source node is no longer considered (we exit out of the loop over stations so that more distant stations are not considered for this grid node). Thus, we require that the closest n^{near} functioning stations to a trial source must detect. This also reduces processing time by not searching on additional stations unless a near-station detection warrants further consideration for a trial source node.
6. We store additional grid functions to keep track of the number and distribution (azimuthal gap) of stations linked to each source node. Azimuthal gap is the largest open azimuth between azimuthally adjacent

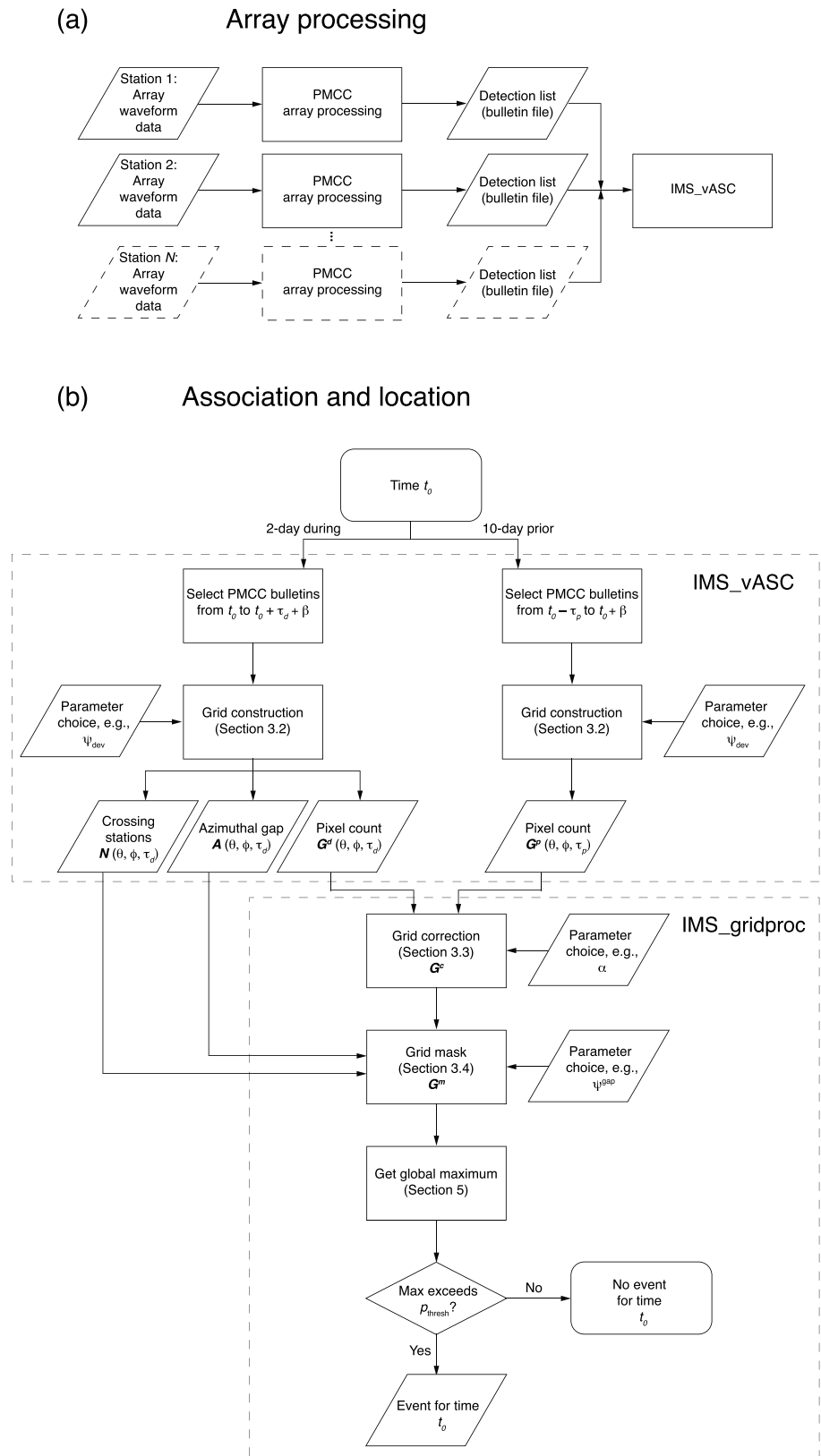


Figure 5. Flowchart illustrating our method. (a) Array processing is performed for all individual stations [Matoza et al., 2013b]. (b) The array processing results are used in the combined association and location procedure (section 3). IMS_vASC and IMS_gridproc are two Fortran 90 programs that perform the procedures indicated.

Table 1. Parameters Used in Figures and Processing Runs^a

| | ψ_{dev} (deg) | v (km/s) | d_{max} (km) | f_{min} (Hz) | f_{max} (Hz) | f_{mean} (Hz) | m | n^{near} | n^{cross} | ψ^{gap} (deg) | α |
|-----------|---------------------------|------------|-----------------------|-----------------------|-----------------------|------------------------|-----|-------------------|--------------------|---------------------------|----------|
| Figure 6 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 10.0 | - | 0 | - | - | 1.5 |
| Figure 7 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 500 | 0 | 3 | 220 | 1.5 |
| Figure 8 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 500 | varied | 3 | 220 | 1.5 |
| Figure 9 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 500 | 2 | 3 | 220 | 1.5 |
| Figure 10 | 5, 10 | 0.33 | 5000 | 0.01 | 5.0 | 5.0 | 500 | 0 | 3 | 220 | 2.0 |
| Figure S2 | 5, 10 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 500 | 0 | 3 | 220 | 2.0 |
| Figure S3 | 5, 10 | 0.33 | 5000 | 0.01 | 5.0 | 5.0 | 500 | 0 | 3 | 220 | 2.0 |
| Run 1 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 500 | 2 | 3 | 220 | 1.5 |
| Run 2 | 2, 5 | 0.30 | 5000 | 0.1 | 5.0 | 2.0 | 500 | 2 | 3 | 220 | 1.5 |
| Run 3 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 500 | 2 | 2 | 220 | 1.5 |
| Run 4 | 5, 10 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 500 | 1 | 3 | 220 | 2.0 |
| Run 5 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 500 | 0 | 2 | 220 | 1.5 |
| Run 6 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 500 | 1 | 2 | 220 | 1.5 |
| Run 7 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 50 | 1 | 2 | 220 | 1.5 |
| Run 8 | 2, 5 | 0.33 | 5000 | 0.1 | 5.0 | 2.0 | 50 | 1 | 2 | 315 | 1.5 |
| Run 9 | 5, 10 | 0.33 | 5000 | 0.01 | 5.0 | 5.0 | 500 | 1 | 3 | 220 | 1.5 |
| Run 10 | 5, 10 | 0.33 | 5000 | 0.01 | 5.0 | 5.0 | 500 | 1 | 3 | 220 | 2.0 |
| Run 11 | 5, 10 | 0.33 | 5000 | 0.01 | 5.0 | 5.0 | 500 | 1 | 3 | 220 | 3.0 |
| Run 12 | 5, 10 | 0.33 | 5000 | 0.01 | 5.0 | 5.0 | 500 | 1 | 3 | 220 | 5.0 |

^aWhere two values are listed and separated by a comma, they refer to “during” \mathbf{G}^{d} and “prior” \mathbf{G}^{p} , respectively. Results from Runs 1, 4, 8, and 10 are shown in Figure 11, while results from other runs are shown in supporting information Figures S4–S6.

recording stations [e.g., Bondár *et al.*, 2004]. In a later stage (section 3.4), we use this information to impose a two- or three-station detection and maximum azimuthal gap (section 3.4).

The end result of the above procedure is an integer grid function $\mathbf{G}(\tau)$, where G_{ij} is the number of all associated pixels from all stations at a trial source latitude θ_i and longitude ϕ_j during time τ . We also have grid function $\mathbf{N}(\tau)$, where N_{ij} counts the number of stations linked with source i, j with at least m pixels per station, and grid function $\mathbf{A}(\tau)$, where A_{ij} is the azimuthal gap of all stations linked to source i, j with at least m pixels per station.

We note that for a distance of 5000 km, propagation time with a celerity of 0.33 km/s is 4.2 h. In this case, the adjustment for an assumed celerity will consider detections for up to 4.2 h after τ ; this correction is more critical when using a shorter time stacking interval (e.g., 1 h) (see supporting information Movie S1). We also note that Landès *et al.* [2012] and Matoza *et al.* [2011b] normalized the contributions from each station such that each station contributes equally to the grid function. We avoid such a normalization here as our method relies on emphasizing the large number of detections on stations closest to the source.

3.3. Grid Correction for a Prior Model

For simplicity in this section we assume that $n^{\text{near}} = 0$ (i.e., we temporarily ignore our requirement that the nearest one or two stations must detect a source). In the basic algorithm described in section 3.2 (with $n^{\text{near}} = 0$), a persistent source of infrasound from the same backazimuth (e.g., local surf or anthropogenic infrasound) will smear out into a sector of large grid values extending from the station to the maximum distance. An example of this is shown in Figure 6, in which the global grid \mathbf{G} during the eruption of Sarychev Peak is dominated by persistent infrasound from a dam near IS10 in Canada (D. McCormack, NRCAN, personal communication, 2016). In Figure 6a, even though the eruption of Sarychev Peak creates a local maximum in \mathbf{G} (see expanded region in Figure 6e), clutter at IS10 complicates identification of this feature. The rates of PMCC detections persistently registered by stations of the IMS vary by several orders of magnitude globally [Matoza *et al.*, 2013b]; thus, we need a way to reduce the influence of large numbers of irrelevant detections from some stations.

We assume that the persistent coherent infrasound signals at each station are stationary in rate and wavefront properties during a time prior to the current time interval. By performing separate runs of the algorithm

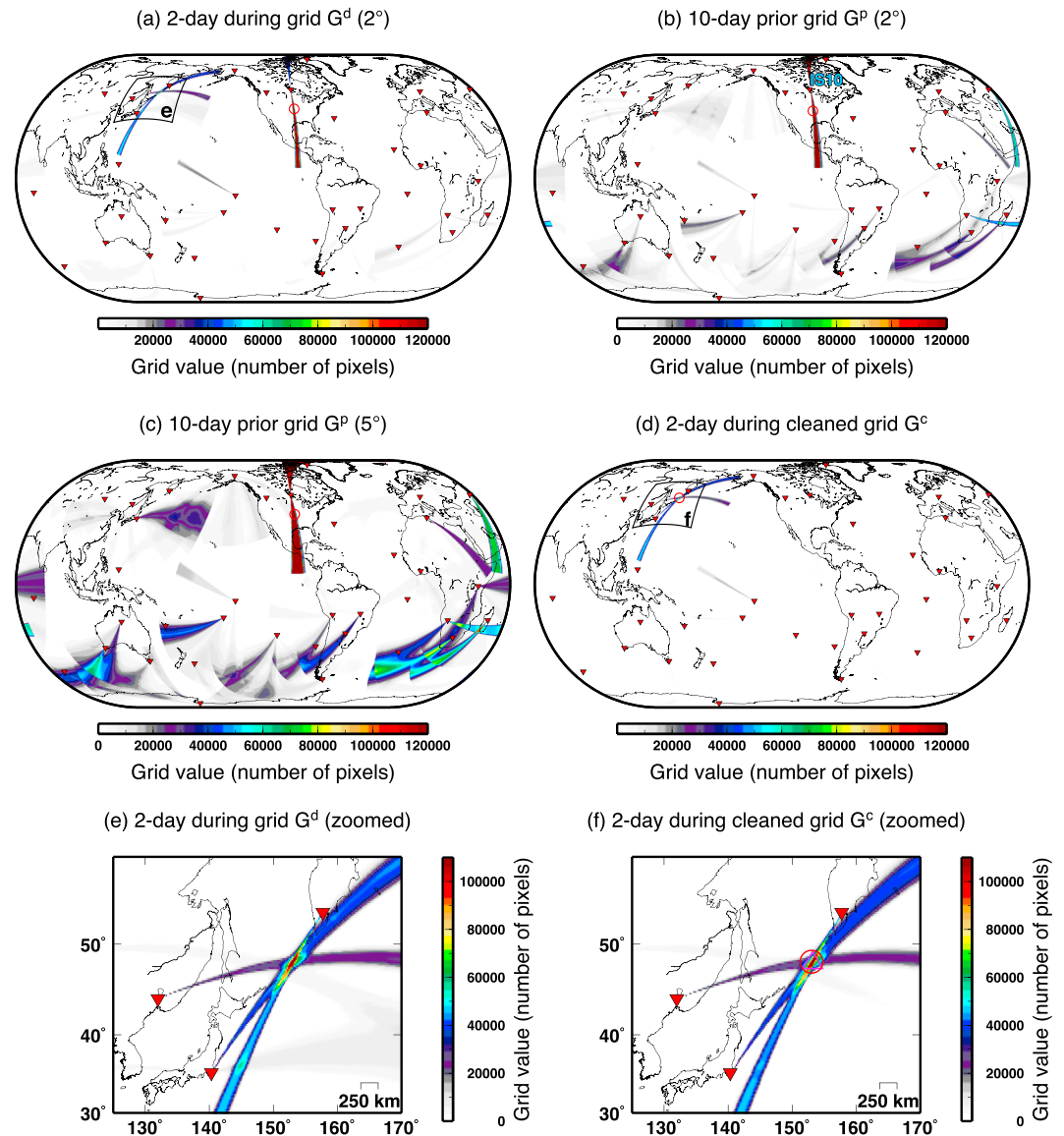


Figure 6. Example application of the grid correction procedure (equation (1)) to the June 2009 eruption of Sarychev Peak, Kuriles. For illustration in this figure, we set $n^{\text{near}} = 0$ (see section 3.2) and do not impose requirements on the minimum number of detecting stations or azimuthal gap (see section 3.4). We use $f_{\text{min}}^{\text{fam}} \geq 0.1$ Hz, $f_{\text{max}}^{\text{fam}} \leq 5$ Hz, $f_{\text{mean}}^{\text{fam}} \leq 10$ Hz (to essentially switch off the filter on $f_{\text{mean}}^{\text{fam}}$ and include strong local infrasound for illustration purposes), $d_{\text{max}} = 5000$ km, and $v = 0.33$ km/s (see Table 1). (a) The 2-day “during” grid G^d is constructed for the time from 00:00 UTC 11 June (Day 162) to 00:00 UT 13 June (Day 164) 2009 using the 37 stations available during this time (inverted red triangles) and using an azimuth deviation tolerance of 2° . The box labeled “e” indicates the area expanded in Figure 6e, which shows a clear local maximum in the grid function at the true location of Sarychev Peak (open magenta triangle in Figure 6f). However, the global maximum value of G^d (red circle in Figure 6a) occurs near station IS10, Canada due to the presence of persistent infrasound from a nearby dam (clutter). (b) The 10-day “prior” grid G^p , constructed for the time from 00:00 UTC 1 June (Day 152) to 00:00 UT 11 June (Day 162) 2009 with the same azimuth deviation tolerance of 2° , shows the same grid feature near IS10. (c) A 10-day “prior” grid G^p is constructed using the expanded azimuth deviation tolerance of 5° , to account for clutter azimuth variability. (d) The “cleaned” grid G^c (equation (1)) is constructed using $\alpha = 1.5$. The feature from IS10 is consequently removed and the grid global maximum (red circle) now coincides with the true location of Sarychev Peak; box shows area expanded in Figure 6f.

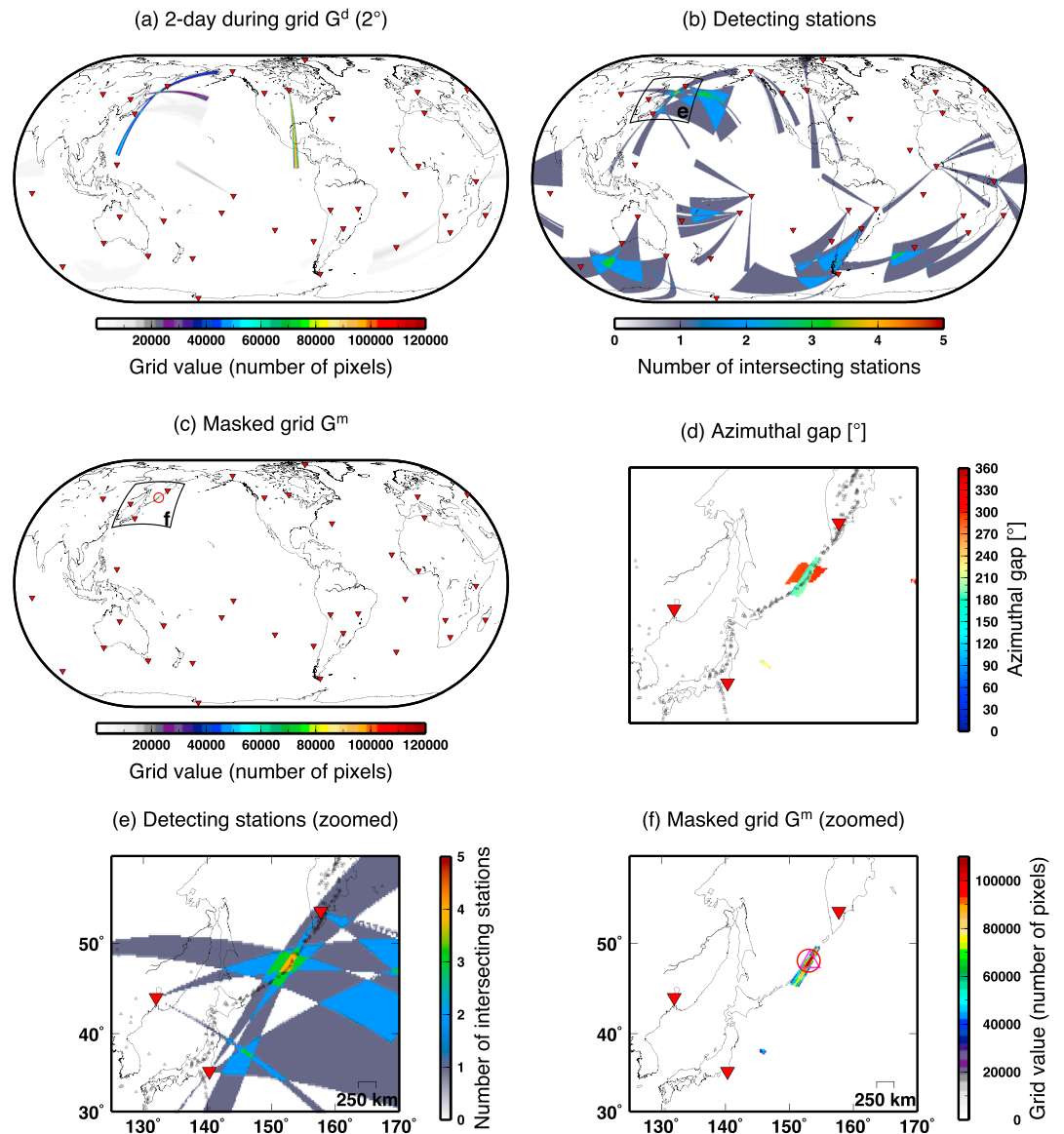


Figure 7. Example application of the three-station detection and azimuthal gap constraints (section 3.4) to the June 2009 eruption of Sarychev Peak, Kuriles. For illustration in this figure, we set $n^{\text{near}} = 0$. Here we use $f_{\text{min}}^{\text{fam}} \geq 0.1$ Hz, $f_{\text{max}}^{\text{fam}} \leq 5$ Hz, $f_{\text{mean}}^{\text{fam}} \leq 2$ Hz (different to Figure 6 and consistent with, e.g., the catalog Run 1 shown in Figure 11a), $d_{\text{max}} = 5000$ km, and $v = 0.33$ km/s (Table 1). (a) The 2-day “during” grid \mathbf{G}^d is constructed for the time from 00:00 UTC 11 June (Day 162) to 00:00 UT 13 June (Day 164) 2009 using the 37 stations available during this time (inverted red triangles) and using an azimuth deviation tolerance of 2° . (b) Grid function \mathbf{N} , which records the number of stations linked to the trial source node with at least $m = 500$ pixels, and corresponding to \mathbf{G}^d (section 3.2). (c) The masked grid \mathbf{G}^m (equation (2)) corresponding to three-station detection with a maximum azimuthal gap of 220° with the cleaned grid \mathbf{G}^c (not shown, cleaned using 10-day prior grid with azimuth tolerance of 5° and $\alpha = 1.5$). The global maximum location is shown by a red circle. (d) The azimuthal gap grid \mathbf{A} for the zoomed region (indicated by box in Figure 7c); azimuthal gap A_{ij} is only calculated at trial source (θ_i, ϕ_j) if the number of intersecting stations N_{ij} is greater than 3. (e) Zoom of Figure 7b, showing the grid function \mathbf{N} in the box indicated in Figure 7b. (f) Zoom of Figure 7c, showing the masked grid function \mathbf{G}^m in the box indicated in Figure 7c. The grid global maximum is shown by a red circle and the true location of Sarychev Peak by a magenta triangle. Small black triangles in Figures 7d and 7e are holocene potentially active volcanoes [Siebert and Simkin, 2002; Global Volcanism Program, 2013].

described in section 3.2, we construct \mathbf{G}^P for a long-term data sample (here $\tau_p = 10$ days) prior to a given time t_0 (Figures 6b and 6c). We then use this to predict the number of unwanted detections contributing to the grid \mathbf{G}^d (Figure 6a) for a shorter time interval of interest beginning at t_0 ; here $\tau_d = 2$ days, where p and d denote “prior” and “during”, respectively. Assuming that the rate of unwanted detections sampled in τ_p remains unchanged, we require that the number of detections must exceed the background rate by more than α times, using the “cleaned” grid function \mathbf{G}^c for time τ_d :

$$\mathbf{G}^c = \mathbf{G}^d - \alpha \frac{\tau_d}{\tau_p} \mathbf{G}^P. \quad (1)$$

A plot of \mathbf{G}^c for $\alpha = 1.5$ is shown in Figure 6d for the June 2009 eruption of Sarychev Peak. Equation (1) may overcorrect for the rate of clutter, especially if the rate of clutter is higher during τ_p than during τ_d . \mathbf{G}^c will have negative values at overcorrected trial source nodes. This is acceptable for our application, and we can simply ignore any negative values. We are only interested in strongly positive values $G_{ij}^c \gg 0$, where a transient feature during τ_d greatly exceeds the background rate of associations in τ_p at a trial source node (θ_i, ϕ_j) . We also have the desirable property that for a source at (θ_u, ϕ_v) , inactive during τ_p and suddenly active in τ_d with no background detections, $G_{uv}^c = G_{uv}^d$; thus, we preserve the number of contributing pixels and can set a threshold value for detection.

Choosing $\alpha > 1$ (Table 1) makes the algorithm more robust to small changes in the rate of clutter over time. After some experimentation (see Table 1), we also found it helpful to use a wider azimuth tolerance ψ_{dev} (e.g., 5°) for \mathbf{G}^P (Figure 6c) and a smaller azimuth tolerance (e.g., 2°) for \mathbf{G}^d (Figure 6a). In this case, the cleaned grid \mathbf{G}^c is more robust with respect to statistical fluctuations in the azimuth values of clutter at individual stations.

In addition to the removal of the artifact from dam clutter near IS10, Figures 6a–6c capture the influence of the Southern Ocean storms on the grid functions, which is removed in the cleaned grid (Figure 6d). The effect of oceanic clutter on station IS30, Japan is also captured in Figure 6e (sector extending to the east), which is effectively removed in Figure 6f.

As a caveat, we note that this method may fail during seasonal equinoxes particularly for midlatitude stations, where the background coherent infrasound sources registered at the arrays are not stationary [Le Pichon *et al.*, 2009]. Finally, we emphasize that additionally using $n^{\text{near}} = 1$ or 2 (requiring detection on the nearest one or two stations) also helps dramatically reduce artifacts from clutter (see section 4.1).

3.4. Additional Constraints

In a final processing stage, we use the grid functions \mathbf{N} and \mathbf{A} as a mask on the cleaned grid \mathbf{G}^c . The final “masked” grid \mathbf{G}^m , which we use for source identification, is defined

$$G_{ij}^m = \begin{cases} G_{ij}^c, & \text{if } N_{ij} \geq n^{\text{cross}} \text{ and } A_{ij} \leq \psi^{\text{gap}} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

This corresponds to an n^{cross} -station detection requirement with a maximum azimuthal gap of ψ^{gap} . An example of this procedure using $n^{\text{cross}} = 3$ and $\psi^{\text{gap}} = 220^\circ$ for the June 2009 eruption of Sarychev Peak is shown in Figure 7. We note that grid functions \mathbf{N} and \mathbf{A} are created during the construction of \mathbf{G}^d . It would be an improvement to instead have \mathbf{N} and \mathbf{A} correspond to the cleaned grid \mathbf{G}^c . However, this modification would require significantly restructuring our algorithm and is beyond the scope of the present work. It is flagged here as a possible improvement in future work.

4. Eruption Case Studies

4.1. Sarychev Peak, June 2009

Figures 6–8 demonstrate the application of this method to the 11–16 June 2009 sustained explosive eruption of Sarychev Peak, Kuriles. This VEI 4 eruption [Siebert and Simkin, 2002; Global Volcanism Program, 2013] was well recorded by six stations of the IMS infrasound network at ranges from 640 (IS44, Kamchatka) to 6400 km (IS31, Kazakhstan) [Matoza *et al.*, 2011a]. The infrasound data from this event are described in detail by Matoza *et al.* [2011a], who used the data to estimate source locations, provide a high-temporal resolution chronology of the explosive eruption, and investigate atmospheric propagation of the signals. The data from IS44 for this eruption are shown in Figure 2. Owing to the extensive IMS infrasound recordings from this eruption, we use it as the primary test data set in developing this algorithm.

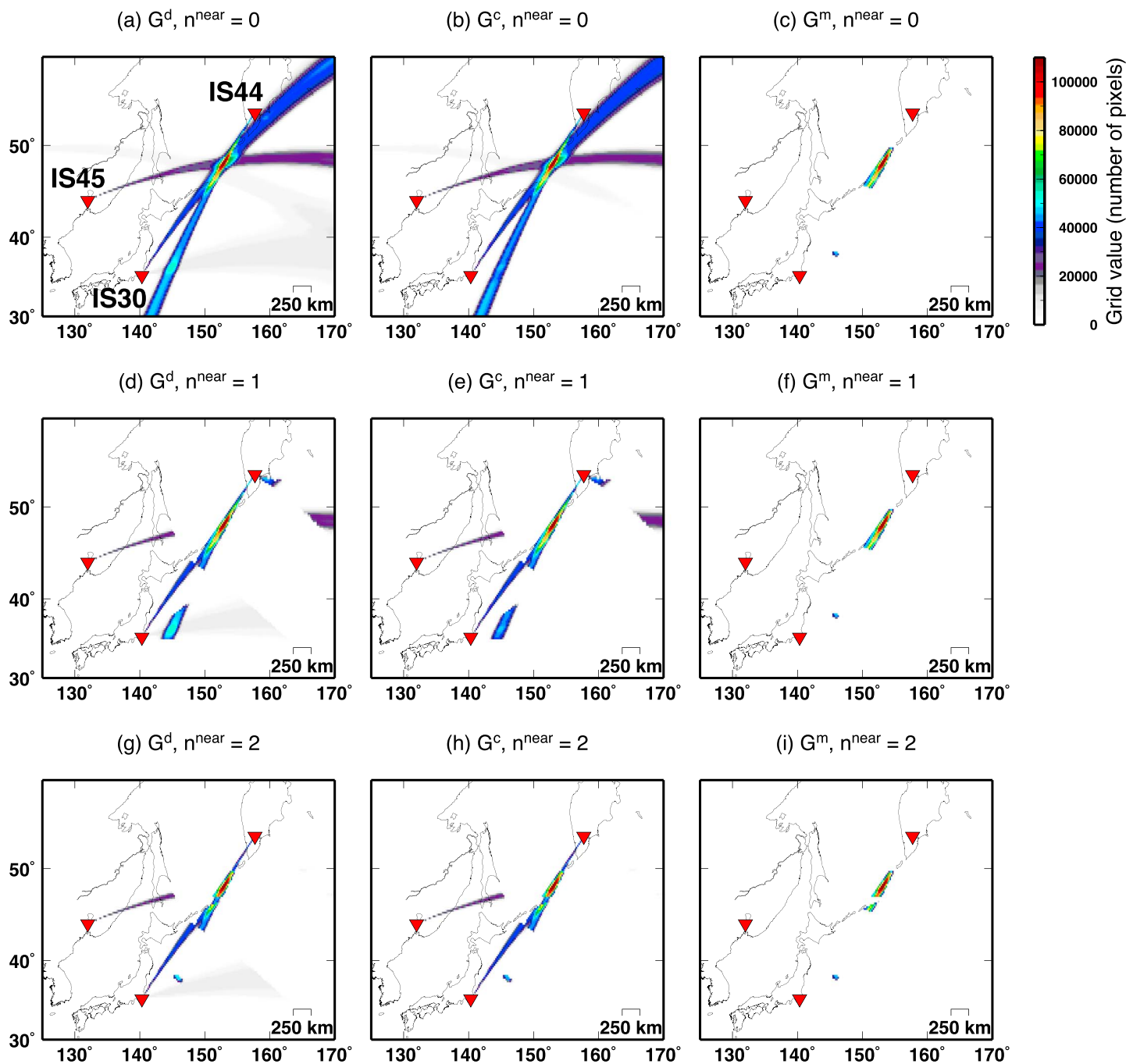


Figure 8. Demonstration of the near-station detection constraint, with n^{near} nonzero (see section 3.2). We show the “during” grid G^d , the “cleaned” grid G^c , and the “masked” grid G^m for three cases, in which $n^{\text{near}} = 0, 1, 2$ as indicated in the title labeling. The grids are constructed using the same data times as Figures 6 and 7. All parameters used are identical with those used in Figure 7 (see Table 1), varying only n^{near} . All panels are displayed with the same color scale, shown in the top right.

Supporting information Movie S1 is an animation of a regional grid G^d at short time snapshots τ_d of 1 h, which are progressively summed over the 5 day duration of the event. One hour is much shorter than the time stacking duration τ_d of 2 days used throughout this study, but the movie illustrates how the values in the grid function G^d accumulate as more detections are added for multiple stations from a long-duration, broadband signal originating from a fixed source location.

Figure 6 is described in section 3.3 and illustrates the grid cleaning procedure using a 10-day prior grid (equation (1)). Whereas the grid constructed for a region around Sarychev Peak (supporting information

Movie S1) has a clear maximum corresponding to the eruption, the global grid is contaminated by clutter from IS10 and the grid correction procedure (equation (1)) is necessary to isolate the eruption event within global data (Figure 6). The minimum number of intersecting stations and azimuthal gap constraints (section 3.4, equation (2)) are demonstrated in Figure 7. The mask procedure (equation (2)) dramatically reduces the long sector features extending away from each individual station, with the masked grid \mathbf{G}^m reducing the possible source location solution to a small elliptical region around the true location of Sarychev Peak (Figure 7f). The final source location solution (defined as the global maximum of \mathbf{G}^m for a given time τ_d) is only 19.1 km from the true location of Sarychev Peak (Figure 7f). An automated scan through the Smithsonian GVP database reveals that Sarychev Peak is the closest potentially active volcano to the automatically obtained source solution; thus, in this example, the algorithm automatically identifies the correct erupting volcano from within the global IMS data.

Figure 8 illustrates the effects of systematically varying the parameter n^{near} (with values of 0, 1, and 2) (section 3.2). We added the near-station detection constraint (n^{near} , section 3.2) after preliminary experimentation with this method without using a near-station detection constraint (equivalent to $n^{\text{near}} = 0$). These experiments produced examples in which long-range intersecting backazimuths apparently identify and locate a source that is not recorded by a station closer to the apparent source, as would usually be expected. Although nonzero n^{near} likely eliminates many spurious associations, Figure 8 shows a possible problem with an $n^{\text{near}} = 2$ restriction. In Figures 8a and 8b, it is clear that if IS44 had not detected the eruption (for example, if the northeast waveguide had not been present or if noise levels had been high at IS44 at the time), the combination of IS30 and IS45 would still have detected the eruption (although with lower overall grid values in \mathbf{G}^d). However, if IS44 had not detected the event and n^{near} was chosen as ≥ 1 , the eruption would not have been identified in our automatic procedure, despite the signals recorded at IS30 and IS45. Although this problem does not arise for the Sarychev Peak eruption (Figure 8), there may be geometries (especially in sparser regions of the IMS network), where the n^{near} constraint is too restrictive. For this reason, in section 5, we produce different versions of the catalog with different choices of $n^{\text{near}} = 0, 1, 2$ (Table 1).

4.2. Eyjafjallajökull, April–May 2010

We further test and validate the method using the 2010 eruption of Eyjafjallajökull, Iceland (Figure 9). An initial burst of infrasound from this event was recorded from 14 to 15 April 2010, with the major long-duration infrasound recorded from 18 April to 20 May 2010 [Matoza *et al.*, 2011b]. We only use IMS stations in this study, but additional data from regional infrasound array networks (not used here) greatly enhance the detection capability of this event [Matoza *et al.*, 2011b]. Figure 9 shows the application of this method using the time 18–20 April 2010 for \mathbf{G}^d and 8–18 April 2010 for \mathbf{G}^p . The parameters used for the run shown in Figure 9 are identical with those used in Figures 8g–8i. The algorithm automatically detects the eruption and locates the source to within 68.6 km of Eyjafjallajökull. The nearest volcano in the GVP database to the obtained source solution is Vestmannaeyjar in this case (40.4 km from the obtained source location). Thus, the automated algorithm does not correctly identify the true volcano but picks out a volcano close to the true one.

4.3. Kasatochi

The 7–8 August 2008 eruption of Kasatochi, AK, USA, provides an illustrative example for which our algorithm struggles to identify the known eruptive event given the parameter choices used so far. This eruption was recorded by six IMS stations out to a maximum range of 5247 km [Fee *et al.*, 2010b]. Station IS59, Hawaii (range 3996 km), is an important station for the detection of Kasatochi; however, IS59 data are not included in the database used for this study since the raw waveform data are not available in the IDC data archive [Matoza *et al.*, 2013b]. Nevertheless, four other IMS stations (IS53, IS30, IS56, and IS18) within a distance of 5000 km detected the Kasatochi eruption [Fee *et al.*, 2010b]. After some experimentation, we found that the Kasatochi event is well identified using our method if we use (1) an expanded azimuth deviation tolerance ψ_{dev} of 5° for \mathbf{G}^d and 10° for \mathbf{G}^p and (2) a more broadband frequency range of $f_{\text{min}} = 0.01$ Hz, $f_{\text{max}} = 5.0$ Hz, and $f_{\text{mean}} = 5.0$ Hz (Figure 10). The expanded azimuth deviation tolerance (1, above) allows for greater azimuthal deflection from crosswinds during long-range infrasonic propagation. The more broadband frequency parameters (2, above) permit lower frequency signals than previously considered, corresponding to six additional PMCC processing bands (see Figure 2), greatly increasing the overall number of detections from each station associated with this eruption. During PMCC processing, individual pixels (Figure 2) are grouped into families of pixels with similar wavefront parameters. Thus, pixels at low frequencies (≤ 0.1 Hz) may join with pixels from higher frequencies (≥ 0.1 Hz) to form a family. Therefore, our frequency parsing, which is based on family-level information

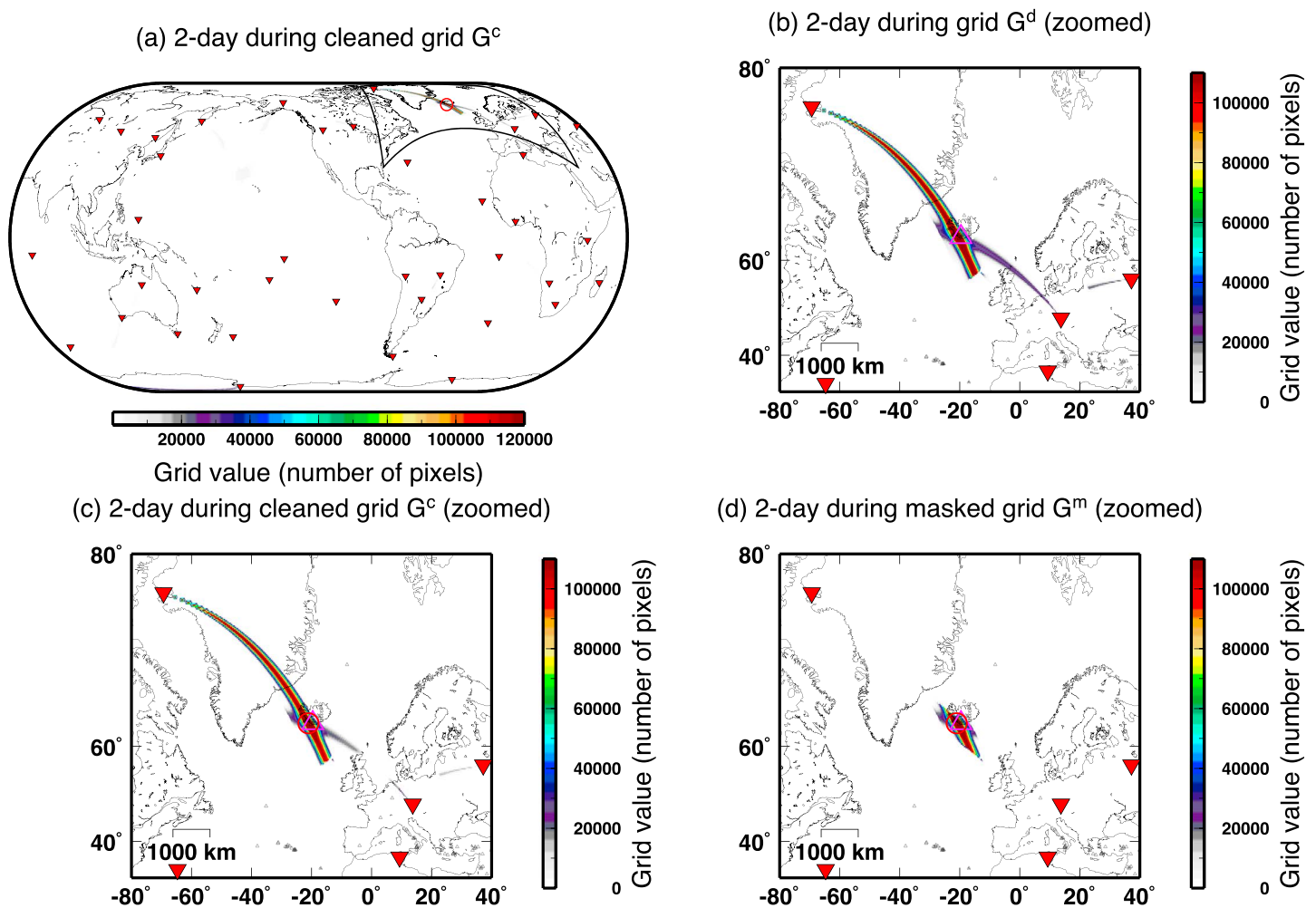


Figure 9. Example application to the 2010 eruption of Eyjafjallajökull, Iceland. The 2-day “during” grid G^d is constructed for the time from 00:00 UTC 18 April (Day 108) to 00:00 UT 20 April (Day 110) 2010 with an azimuth deviation tolerance of 2° (b). A 10-day “prior” grid with an azimuth deviation tolerance of 5° is used to form the cleaned grid G^c using $\alpha = 1.5$ (a and c). Here we use $n^{\text{near}} = 2$, $f_{\text{min}}^{\text{fam}} \geq 0.1$ Hz, $f_{\text{max}}^{\text{fam}} \leq 5$ Hz, $f_{\text{mean}}^{\text{fam}} \leq 2$ Hz, $d_{\text{max}} = 5000$ km, and $v = 0.33$ km/s (Table 1). (d) The masked grid G^m corresponds to three-station detection with a maximum azimuthal gap of 220° and $m = 500$ pixels. The grid global maximum is shown by a red circle and the true location of Eyjafjallajökull by a magenta triangle. The missing part of the sector extending from IS26 (Germany) in Figure 9c results from “overcorrection” from nonstationary prior North Atlantic microbarom signals at IS43 (Dubna, Russian Federation).

(section 3.2), may have the unwanted effect of suppressing contributions from truly broadband families that extend across the frequency range from 0.01 to 5 Hz (Figure 2). Consequently, we have produced a separate version of the catalog (section 5) using this more broadband processing ($f_{\text{min}} = 0.01$ Hz, $f_{\text{max}} = 5.0$ Hz, and $f_{\text{mean}} = 5.0$ Hz). We note that our original choice of $f_{\text{min}} = 0.1$ Hz was designed to reduce the influence of microbaroms; however, the processing described in sections 3.3 and 3.4 is effective against microbaroms. Supporting information Figures S2 and S3 show the same processing with expanded azimuth deviation and broadband frequencies applied to the Sarychev Peak eruption.

For the results shown in Figure 10 (see Table 1 for parameters used), the algorithm locates the source to within 176 km of Kasatochi. The nearest volcano in the GVP database to the obtained source solution is Takawangha in this case (8.1 km from the obtained source location). The effect of the parameter choices on the detection or nondetection of the 2008 Kasatochi eruption is further discussed in section 6.2.

5. Global Acoustic Catalogs 2005–2010

We applied the algorithm described in section 3 systematically to the multiyear 2005–2010 IMS reprocessed PMCC detection lists of Matoza *et al.* [2013b] (section 3.1), varying the parameters as indicated in Table 1 to

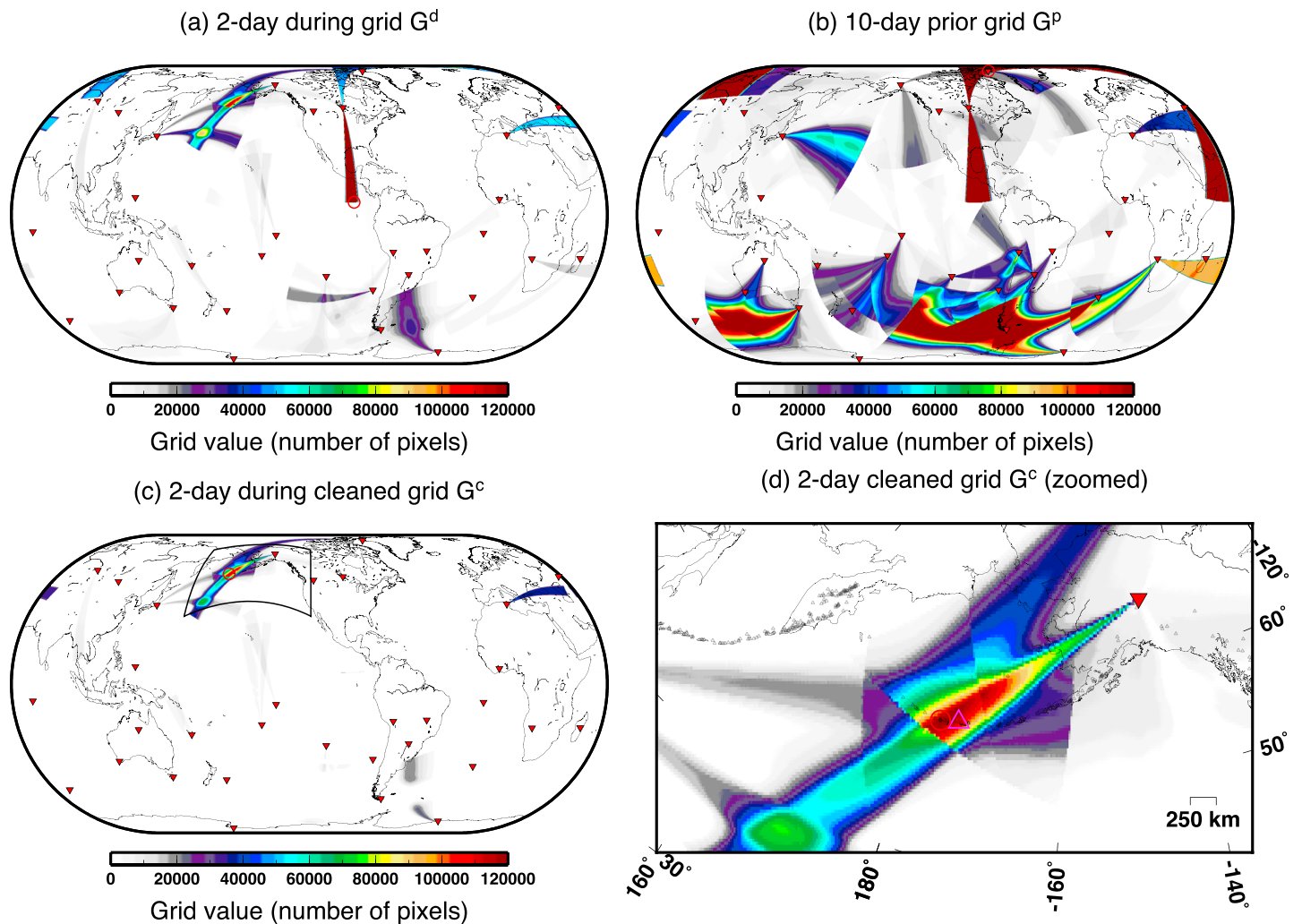


Figure 10. Example application to the 2008 eruption of Kasatochi, AK. (a) The 2-day “during” grid G^d is constructed for the time from 00:00 UTC 7 August (Day 220) to 00:00 UT 8 August (Day 221) 2008 with an azimuth deviation tolerance of 5° . (b) A 10-day “prior” grid with an azimuth deviation tolerance of 10° is used to form (c and d) the cleaned grid G^c using $\alpha = 2$. Here we use $n^{\text{near}} = 0$, $f_{\text{min}}^{\text{fam}} \geq 0.01$ Hz, $f_{\text{max}}^{\text{fam}} \leq 5$ Hz, $f_{\text{mean}}^{\text{fam}} \leq 5$ Hz, $d_{\text{max}} = 5000$ km, and $v = 0.33$ km/s (Table 1). The grid global maximum in each case is shown by a red circle and the true location of Kasatochi by a magenta triangle.

produce different catalog versions for comparison. The PMCC detection lists primarily cover the time from 1 April 2005 to 5 December 2010, with temporal evolution in station availability occurring as the IMS network is still under construction (see supporting information Figure S1). We proceed through each year (from 2005 to 2010) in daily increments t_0 starting on Day 11, using 2-day windows τ_d for the “during” grid and the 10 days prior for τ_p ; thus, each day of data is considered twice (as the first then the second of the 2 days in τ_d). We define an event at time t_0 simply as the global maximum of the masked grid G^m , provided that the value of the global maximum exceeds a defined threshold p_{thresh} . Consequently, only one event is permitted globally in a given 2 day time interval τ_d starting at t_0 (t_0 increments in daily intervals). We set $p_{\text{thresh}} = 500$ pixels for the results shown here (a low, but nonzero, value to allow a sensitive catalog). There are 2075 days in the time period considered between 1 April 2005 to 5 December 2010; thus, the maximum number of possible events in each catalog is 2075. Our results (Figure 11 and supporting information Figures S4–S6) include many days for which no event is defined given the parameters choices used to construct G^m (e.g., Run 1 has 336 events globally, i.e., 336 days for which G^m satisfies the criteria to produce an event). This provides confidence that our algorithm is effective at suppressing spurious associations (see section 6.1).

The value of $p_{\text{thresh}} = 500$ pixels relates to a minimum area in time-frequency space, with this relationship dependent on the PMCC band configuration (number of bands, frequency range of bands, window-length, and time-step of each band) [Cansi, 1995; Cansi and Klinger, 1997; Le Pichon et al., 2010]. The PMCC detection

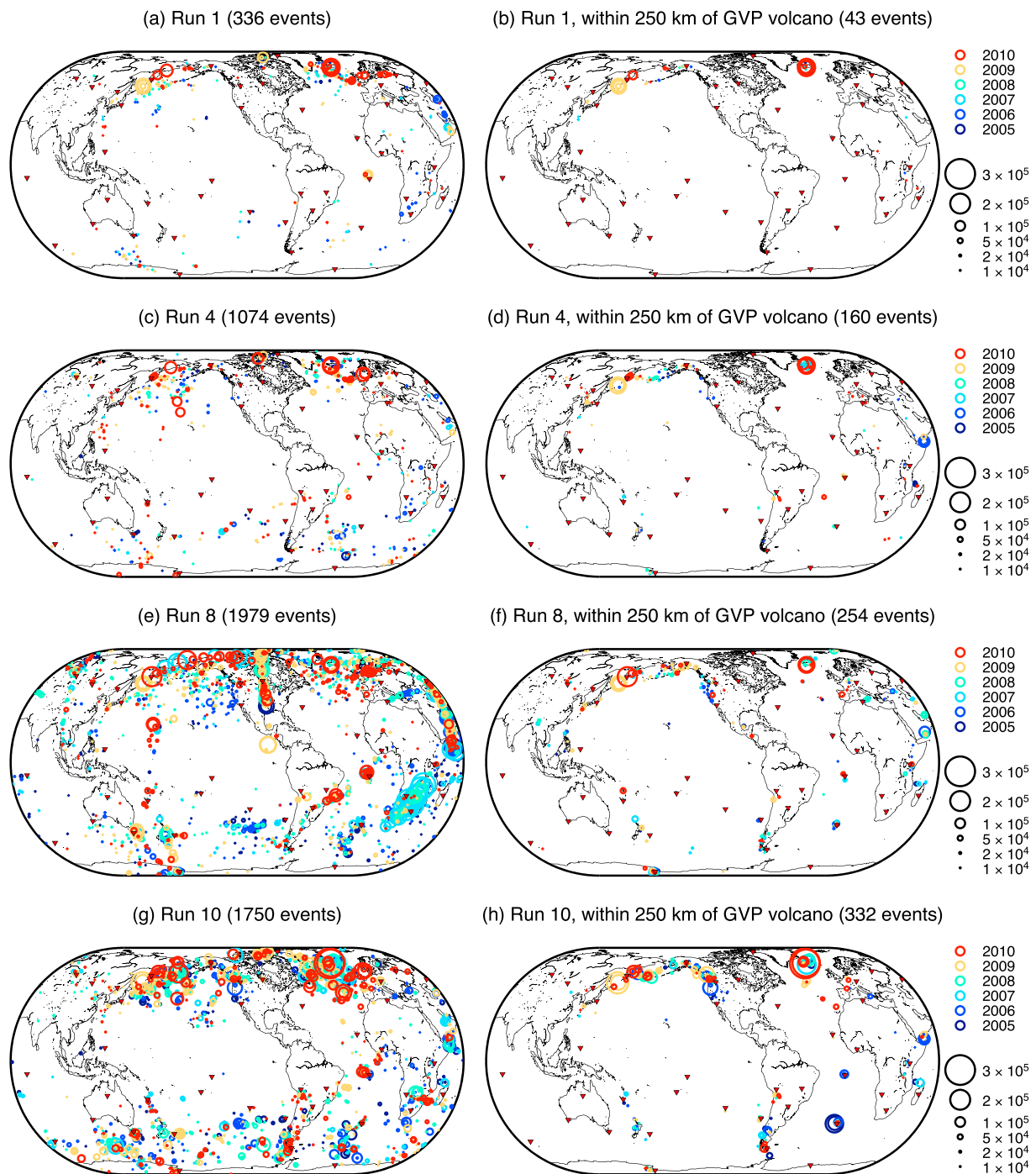


Figure 11. (a–h) Global infrasound catalogs for data from 1 April 2005 to 31 December 2010. Infrasound events are shown as circles colored by the year of occurrence (see legend on right) with symbol size proportional to the number of pixels in the grid G^m at the detection location (black circles in legend indicate number of pixels). Events represent the global maximum of G^m in a 2 day time period (only one event is allowed globally during each 2 day time period) (see section 5). Different results are shown for separate applications of the algorithm with different parameters choices (Runs 1, 4, 8, and 10; see Table 1). Figures 11b, 11d, 11f, and 11h) represent the same catalogs as Figures 11a, 11c, 11e, and 11g), respectively, but where only events within 250 km of a known GVP volcano [Siebert and Simkin, 2002; Global Volcanism Program, 2013] are shown. Red inverted triangles show the 41 stations used (supporting information Figure S1 shows the temporal availability of data from each station). The number of events in each case is displayed in the title. See supporting information Figures S4–S6 for similar plots of all runs listed in Table 1.

lists used in this study [Matoza *et al.*, 2013b] were produced using 15 frequency bands with window-lengths varying from 30 s (for the highest frequency band 3.304–5.000 Hz) to 200 s (for the lowest frequency band 0.0100–0.0151 Hz), each advancing in time-steps of 10% of the window-length. Thus, for a continuous broadband signal spanning all 15 bands, the higher-frequency bands contribute proportionally more pixels for a given signal duration than the lower frequency bands because the higher-frequency bands use a shorter window length and time step (at each time-step, a new pixel is created). The threshold value of p_{thresh} (500 pixels) could be achieved with a single station (array) using the PMCC configuration of Matoza *et al.* [2013b] for a broadband signal occupying all 15 bands with only 146 s worth of time-stepping, corresponding to a relatively short minimum signal duration. This only considers a single station. For the results shown here we use two- or three-station detection with at least, e.g., $m = 500$ pixels per station (Table 1). We note that although the events can be relatively short in duration, the celerity assumption is still not critical because all detections within a 2 day time period are combined together in the constructed grid (section 3.2), with information on their relative arrival times within the 2 day time period discarded.

Figures 11 and S4–S6 show the resulting global infrasonic catalogs for the 12 different “Runs” explained in Table 1. Run 1 (Figures 11a and 11b) is the most conservative catalog using the most restrictive constraints ($\psi_{\text{dev}} = 2^\circ$ is used for \mathbf{G}^{d} and $\psi_{\text{dev}} = 5^\circ$ for \mathbf{G}^{p} ; the detections are restricted using $f_{\text{min}} = 0.1$ Hz, $f_{\text{max}} = 5$ Hz, $f_{\text{mean}} = 2$ Hz; three-station detection is required with an azimuthal gap $\leq 220^\circ$; and the nearest two stations must detect the trial source node). We use a tropospheric association celerity (0.33 km/s) for most of the runs based on the observations by Matoza *et al.* [2011a], who found that a reduction celerity of 0.33 km/s was needed to temporally align signals on remote stations for the June 2009 Sarychev Peak eruption. We test the sensitivity to celerity in Run 2 (supporting information Figures S4c and S4d), which is identical to Run 1 except that a stratospheric association celerity of 0.3 km/s is used. There is only a small difference (<3% difference in total number of events) between the results from Run 1 (Figures 11a and 11b) and Run 2 (supporting information Figures S4c and S4d), indicating that our results are relatively robust and insensitive to the assumed celerity because a long time stacking interval is used ($\tau_d = 2$ days). Relaxing the three-station detection constraint (Run 1) to two-station detection ($n^{\text{cross}} = 2$) (Run 3; supporting information Figures S4e and S4f) only increases the total number of events by 4%. Run 4 is identical to Run 1 except for using a wider azimuth deviation tolerance ($\psi_{\text{dev}} = 5^\circ$ for \mathbf{G}^{d} and $\psi_{\text{dev}} = 10^\circ$ for \mathbf{G}^{p}) and requiring only that the closest station must detect ($n^{\text{near}} = 1$). Runs 5 and 6 (supporting information Figure S5) are identical with Run 3 (supporting information Figure S4) except that n^{near} is set to 0 and 1, respectively, in order to test the sensitivity to this near-station detection constraint. Of the parameters varied in Runs 1 to 6 (supporting information Figures S4 and S5), the allowed azimuth deviation has the largest effect (Run 4, Figures 11c and 11d). Reducing n^{near} from 2 to 1 to 0 has the expected effect of increasing the number of events (compare Runs 3, 6, and 5 in supporting information Figures S4e, S5c, and S5a, respectively).

Runs 7 and 8 are performed to illustrate the importance of the parameter m , the number of pixels required to count a station as recording a trial source node and therefore to associate a station (section 3.2). For illustration, in Runs 7 and 8, we use a small value of $m = 50$ pixels, which we consider to be too low for a robust detection on a given station considering background clutter. Run 8 (Figures 11e and 11f) differs from Run 7 (supporting information Figures S5e and S5f) in having a wider allowed azimuthal gap (315°). It is noteworthy that Runs 7 and 8 produce spurious event location artifacts, which have similar spatial distributions to events found in the SEL3 infrasonic catalog produced by the Preparatory Commission of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) International Data Center (IDC) (see section 6).

Runs 9–12 (supporting information Figure S6) are broadband runs with parameters chosen based on the 2008 Kasatochi eruption (section 4.3). Here we use $f_{\text{min}} = 0.01$ Hz, $f_{\text{max}} = 5$ Hz, and $f_{\text{mean}} = 5$ Hz, which permits a much greater number of detections considering the PMCC frequency band configuration shown in Figure 3. We also allow the greater azimuth deviation tolerance found necessary to associate the Kasatochi signals ($\psi_{\text{dev}} = 5^\circ$ for \mathbf{G}^{d} and $\psi_{\text{dev}} = 10^\circ$ for \mathbf{G}^{p}) and set $n^{\text{near}} = 1$. We performed an additional four runs, identical to Runs 9–12 except with $n^{\text{cross}} = 2$; however, the results were identical with Runs 9–12, so we do not include them here. Microbaroms, one of the most dominant (in terms of number of pixels) and the most globally pervasive source of infrasonic clutter [Matoza *et al.*, 2013b], contribute much more significantly to the grid functions when using these frequency parameters. In supporting information Figure S6, we vary the parameter α (equation (1)) systematically with values of 1.5, 2.0, 3.0, and 5.0 in Runs 9, 10, 11, and 12, respectively, in order to investigate if α may be tuned to reduce spurious associations primarily from increased microbarom

contamination. However, we find that the parameter α has little effect in this case with less than 20% reduction in total events (supporting information Figure S6).

A major result shown in Figure 11 and supporting information Figures S4–S6 (e.g., Figure 11a) is that the June 2009 eruption of Sarychev Peak and the April–May 2010 eruption of Eyjafallajökull represent the largest events (detection features) identified by a blind search through the global multiyear data set. This demonstrates that this method is well adapted to emphasizing signals from sustained explosive volcanic eruptions.

6. Discussion

Our catalogs span the dates from 1 April 2005 to 5 December 2010, a time when the IMS infrasound network was in relatively early stages of development (the number of certified stations increased from 24 to 43 during this period). We anticipate significant improvements in results when applying this algorithm to more recent data from 2010 to present (beyond the scope of the present work), and in future as the IMS infrasound network achieves full 60-station global coverage (Figure 12) (there are 49 certified stations at the time of writing). Figure 12 shows the computed great circle distances from each potentially active GVP volcano to the nearest two or three IMS stations and serves as a first-order proxy for infrasonic network detection capability of global volcanism. Figure 12 gives a preliminary impression for how volcanic eruption detectability will improve by completing the IMS network and also reveals the volcanic regions (e.g., the Caribbean, Southeast Asia, and Aleutians) that would most benefit from additional regional infrasound arrays. Additional infrasound arrays deployed in regions with dense volcanism [e.g., *Guilbert et al.*, 2005; *Matoza et al.*, 2007; *Garces et al.*, 2008; *Matoza et al.*, 2011b; *De Angelis et al.*, 2012; *Tailpied et al.*, 2013; *Caudron et al.*, 2016] have not been considered in this study. Including such data in future work would likely lead to significant improvements by augmenting the network density and decreasing the acoustic source power of explosive eruption detectable with a globally integrated network [*Le Pichon et al.*, 2009; *Green and Bowers*, 2010]. The utility of adding dedicated regional infrasound arrays is highlighted by the case of the 2006–2008 eruptions of Tungurahua, Ecuador, which are not identified in our global search using IMS data but were well recorded on dedicated regional infrasound arrays designed for volcano-acoustic monitoring [*Garces et al.*, 2008; *Fee et al.*, 2010a]. IMS coverage of the region around Tungurahua was sparse during this time (see Figure 12b), and atmospheric conditions are less favorable for long-range propagation at equatorial latitudes [*Le Pichon et al.*, 2009; *Green and Bowers*, 2010]. In addition to infrasound propagation effects, noise levels experienced by key stations in volcanic regions play an important role in controlling volcanic signal detection. This is especially the case for island stations, e.g., IS06, Cocos Islands is critical for observing Indonesian volcanoes.

In sections 6.1 and 6.2 we compare our results achieved using the available IMS network from 2005 to 2010 with the CTBTO IDC SEL3 infrasonic catalog [*Mialle et al.*, 2010] and with the Smithsonian GVP catalog of known volcanic eruptions [*Siebert and Simkin*, 2002; *Global Volcanism Program*, 2013] to validate our method and evaluate its performance.

6.1. Comparison With Existing Standard Event List (SEL3) Infrasonic Catalog

The CTBTO IDC produces an automated infrasound event listing (Standard Event List, SEL3) and an analyst-reviewed Reviewed Event Bulletin (REB) [*Brachet et al.*, 2010]. As with our method, the SEL3 catalog also builds infrasound events using backazimuth cross-bearings but uses a different procedure than used in this study, as described by *Brachet et al.* [2010] and *Mialle et al.* [2010]. The SEL3 catalog incorporates both seismic and infrasonic data, but we limit our comparison to infrasound-only events in the SEL3. This is important as it excludes events that may be primarily built from seismic data, with just a minor contribution from infrasound stations. The automated infrasound processing (SEL3) is only available from February 2010 (Figure 13); therefore, our catalogs only overlap for 11 months. The SEL3 algorithm is designed to identify potential short-duration transient events and therefore does not have the same constraints available in our algorithm, which is designed specifically for sustained explosive volcanic eruption signals (section 2).

According to the description by *Brachet et al.* [2010] with updates by *Mialle et al.* [2010], the algorithm behind SEL3 requires that one of the associated detections has ≥ 20 pixels, each station must have more than 7 pixels, and two-station events have a minimum of 12 pixels. Since a different PMCC time-frequency parameterization is used, the numbers of pixels are not directly comparable between the IDC processing and our results; however, our threshold of 500 pixels is likely much higher than that used for IDC association. The SEL3 has a

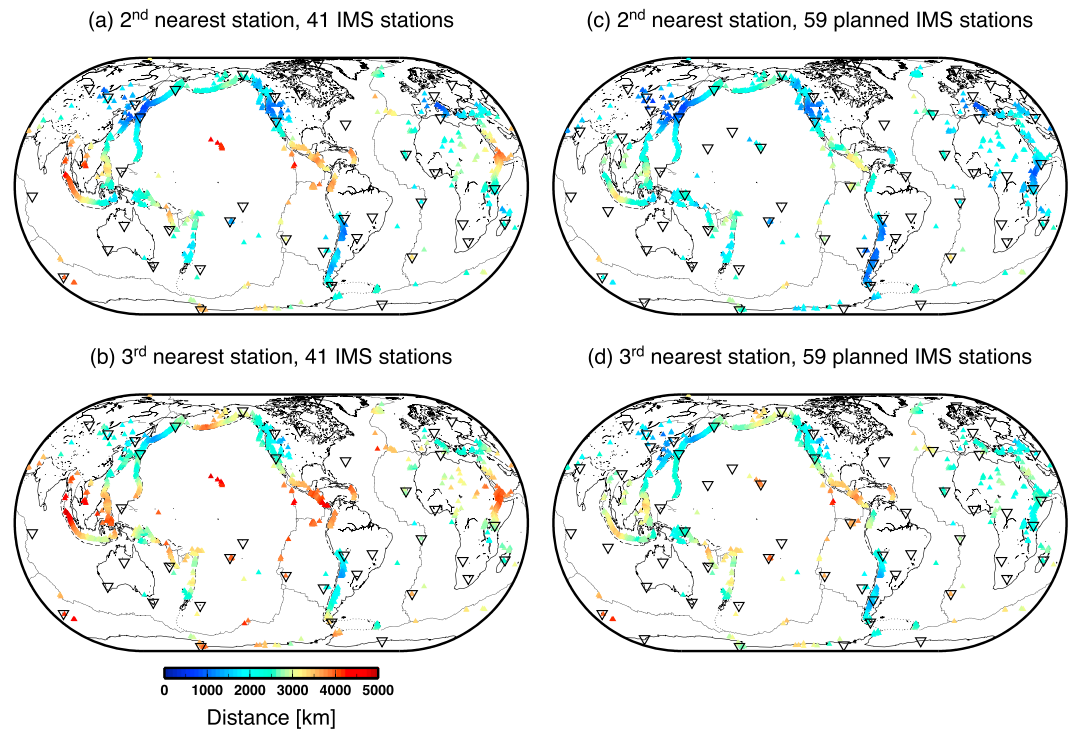


Figure 12. Distance from each potentially active volcano in the Smithsonian GVP database [Siebert and Simkin, 2002; Global Volcanism Program, 2013] (inverted triangles with color scale corresponding to distance) to the nearest two or three IMS stations (open inverted triangles) given (a and b) the 41-station network configuration used in this study and (c and d) the planned 59-station IMS network (location of sixtieth station IS28 is to be determined). The color scale is the same in each case.

distance limit of $\Delta = 60^\circ$ (slightly larger than our 5000 km) and a requirement that the range of the first detecting station should not exceed 40° (4440 km). This is different to the near-station detection constraint used in this study. SEL3 also incorporates a distance-dependent central-frequency function that removes possible “high-frequency, long-distance” arrivals from any potential association [Brachet et al., 2010].

Figure 13 shows a qualitative comparison of events from our Run 8 for April 2005 to December 2010 with the SEL3 events from February 2010 to May 2016. Run 8 was created for illustrative purposes, having a small threshold value of $m = 50$ pixels for linking and associating stations (section 5), which we do not consider to be a robust criterion. Run 8 also has a wide maximum azimuthal gap of 315° . The SEL3 does not have an azimuthal gap constraint. The SEL3 has many more events during the year 2010 (3493 events) than Run 8 (329 events) (Figure 13). We identify similar spurious events in the results from Run 8 (Figure 13a) and SEL3 (Figure 13c); linear features extending northeastward from IS41 (Paraguay) and southward from IS05 (Tasmania) are observed in both Run 8 (Figure 13a) and SEL3 (Figure 13c and 13d). The presence of these features in Run 8 and their absence in, for example, Run 1 (Figure 11a), indicates the utility of using the near-station detection and azimuthal gap constraints and more critically indicates the importance of choosing a threshold value of pixels appropriate for the target signals in detection, association, and location algorithms.

We also observe linear features in Run 8 (Figure 13a) that are not observed in SEL3 (Figure 13c). For example, the linear features extending southward from IS31 (Kazakhstan) and IS10 (Canada) are well-known sources of clutter (gas fields near IS31 [Smirnov et al., 2015] and the dam near IS10). The linear feature pointing southwest from IS32 (Kenya) is not observed in the 2010–2016 SEL3 results. The problem is most acute in 2007 (Figure 11e) and may relate to a known problem with correlated noise at IS32 during this time. We suggest that these linear features are caused by persistent clutter from the same direction at a given station (e.g., IS41). This persistent clutter becomes spuriously linked with occasional signals at additional stations, whose backazimuth trajectories happen to cross the well-defined clutter trajectory from the first station. The absence of these linear features in our other processing runs (see Figure 11 and supporting information Figures S4–S6) indicates that such features can be eliminated using additional constraints and most effectively by using a

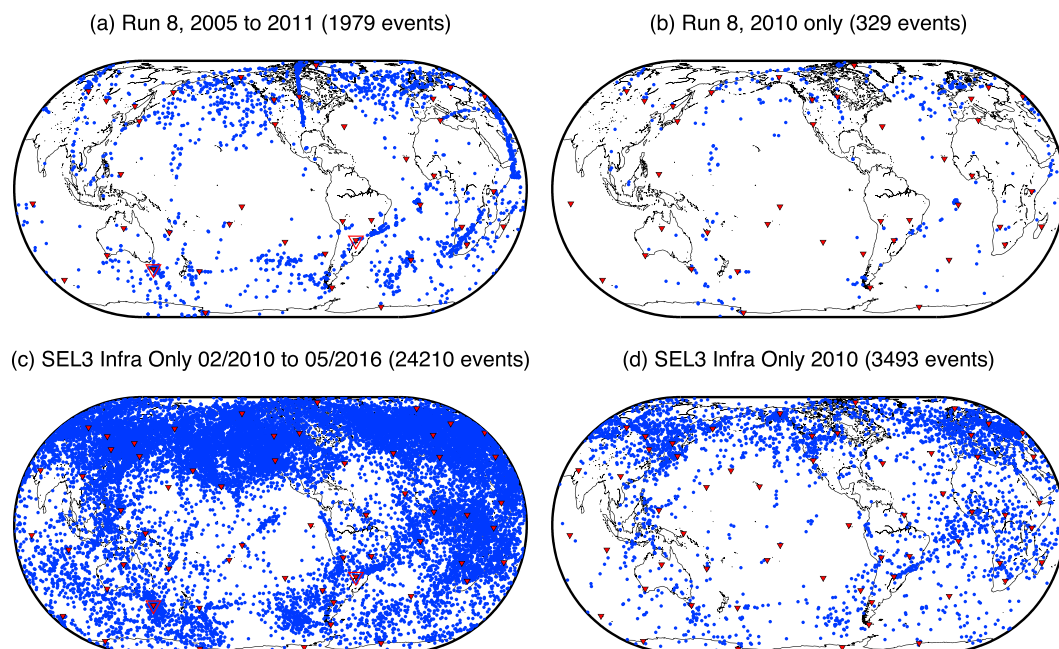


Figure 13. Comparison of (a) the “sensitive” Run 8 (Table 1) catalog for 2005–2010 and (b) the Run 8 catalog for 2010 only, with (c) the February 2010 to May 2016 IDC automated SEL3 infrasound-only catalog and (d) the 2010 (February to December) infrasound-only SEL3. The events in each case are shown by blue dots and the IMS infrasound stations by red inverted triangles. Similar spatial distributions in apparent events (artifacts) are seen in Figures 13a, 13c, and 13d, e.g., linear features extending from stations IS41 (Villa Florida, Paraguay) and IS05 (Hobart, Tasmania) (stations indicated by large red triangles in Figures 13a and 13c).

larger pixel threshold for defining a station association ($m = 500$ pixels). A linear feature also extends from IS48 (Tunisia) toward Mount Etna (Figure 13a) [Assink *et al.*, 2014]. The linear feature from IS48 to Mount Etna (Figure 13a) apparently results from real volcanic signals of interest at a given station (IS48) becoming spuriously linked with clutter (nonvolcanic signals) at other stations.

The SEL3 processing flow includes an algorithm that is run after PMCC to clean the detection lists, which removes approximately 90% of repetitive clutter [Brachet *et al.*, 2010]. However, $\sim 10\%$ of clutter passes through this algorithm and can form spurious associations, resulting in the linear features.

Events over the North Sea area in 2010 (Figure 11) are likely to be real events from sonic boom activity [Le Pichon *et al.*, 2008; Green and Nippres, 2015]. Although the sonic booms are transient events, significant numbers of repetitive events occur within a ~ 30 min time window, which is a signal feature that our algorithm is designed to detect (section 2).

In comparing the numbers of events in Figure 13 between SEL3 and our catalog Run 8, we reiterate that our catalog currently only permits one event per day globally. This is a major restriction, and a next step in future work (beyond the scope of the present study) is to modify these simple detection criteria to permit multiple simultaneous events occurring in different locations on Earth. This could be done, for example, by (1) identifying multiple local maxima in the grid function or (2) dividing the Earth’s surface into multiple separate regions (e.g., VAAC jurisdictions or specific volcanic regions) and performing the processing separately for each region.

6.2. Comparison With GVP Catalog of Known Volcanic Eruptions

In Figures 14 and 15, we compare our Run 1 and Run 12 (Table 1) infrasonic catalogs with the Smithsonian Institution GVP database of known volcanic eruptions during this time (2005–2010) [Siebert and Simkin, 2002; Global Volcanism Program, 2013]. Run 1 has strict criteria, while Run 12 is a more broadband run based on knowledge about the Kasatochi eruption signals (section 4.3 and Table 1). Figure 14a only shows events from the GVP database with Volcanic Explosivity Index (VEI) [Newhall and Self, 1982] of 3 or greater. The GVP eruption dates shown in Figure 14a are eruption “start dates”; since “eruptions” can be prolonged sequences with varied activity, these dates do not necessarily coincide with the climactic explosive eruption event that will

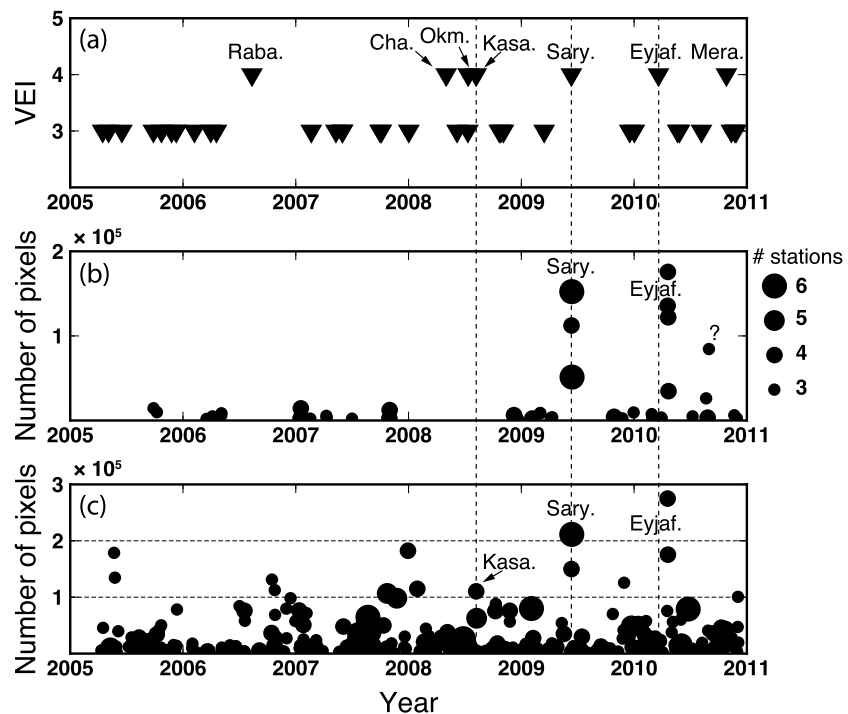


Figure 14. Comparison of (a) Smithsonian GVP catalog [Siebert and Simkin, 2002; *Global Volcanism Program*, 2013], (b) our infrasonic catalog Run 1 (Table 1), and (c) our infrasonic catalog Run 12 (Table 1) in the time from 2005 to 2010. Figures 14b and 14c only show events within 250 km of a known GVP volcano (see Figure 15). Symbol sizes in Figures 14b and 14c are proportional to the number of detecting stations (see legend on the right); note that Sarychev Peak is registered by six stations out to the defined maximum distance of 5000 km. Figure 14a only shows events with Volcanic Explosivity Index (VEI) greater than or equal to 3. VEI 4 eruptions are labeled: Raba. (Rabaul), Cha. (Chaiten), Okm. (Okmok), Kasa. (Kasatochi), Sary. (Sarychev Peak), Eyjaf. (Eyjafjallajökull), and Mera. (Merapi). The same eruptive events captured in the infrasonic catalogs are labeled in Figures 14b and 14c and further highlighted by vertical dashed lines. A blindly detected event in Kamchatka (see Figure 11b) is indicated by a “?” (we attribute this event to Kliuchevskoi, see Figure 15). Arbitrary thresholds of 1×10^5 and 2×10^5 pixels are indicated in (c) with dashed horizontal lines for illustration.

produce the largest acoustic signals. For example, the 2006 VEI 4 eruption of Rabaul, Papua New Guinea, has a start date of 11 August 2006 in the GVP database, but based on the detailed eruption reports, the largest explosive event and plume height occurred on 7 October 2006. The durations of the eruptions shown in Figure 14a range from days to years in the GVP database [Siebert and Simkin, 2002; *Global Volcanism Program*, 2013]. Matching individual infrasound-generating events to individual eruption phase times within a prolonged eruption sequence therefore requires manual examination of the GVP eruption reports; performing this systematically for all identified events is beyond the scope of the present study. Assigned integer VEI values are also approximate and cover a large range (e.g., a “VEI 4” can have a column height from 10 to 25 km and an ejecta volume from 10^8 to 10^9 m³, and VEI does not directly correspond to the gas exit velocity) [Newhall and Self, 1982]. Thus, there may be significant variability in acoustic source power between eruptions with the same assigned integer VEI value. The relationship between acoustic power and gas exit velocity was reviewed by Matoza *et al.* [2013a].

During the time period spanned by our infrasonic catalog, seven VEI 4 eruptions are reported in the GVP database (2006 Rabaul, Papua New Guinea; 2008 Chaiten, Chile; 2008 Okmok, AK, USA; 2008 Kasatochi, AK, USA; 2009 Sarychev Peak, Kuriles, Russian Federation; 2010 Eyjafjallajökull, Iceland; and 2010 Merapi, Indonesia), of which three are clearly identified in our catalogs (2008 Kasatochi; 2009 Sarychev Peak; 2010 Eyjafjallajökull) (Figure 14). The three VEI 4 events identified in our catalogs all lasted multiple days, consistent with our algorithm being targeted toward sustained eruption signals. The 2008 Kasatochi eruption is identified in Run 12 (Figure 14c), but not Run 1 (Figure 14b). Thus, for the time period considered and with the IMS network geometry available during 2005–2010, the global detection threshold with our algorithm is around VEI 4. However, as illustrated in Figure 12, we anticipate considerable improvement for the complete

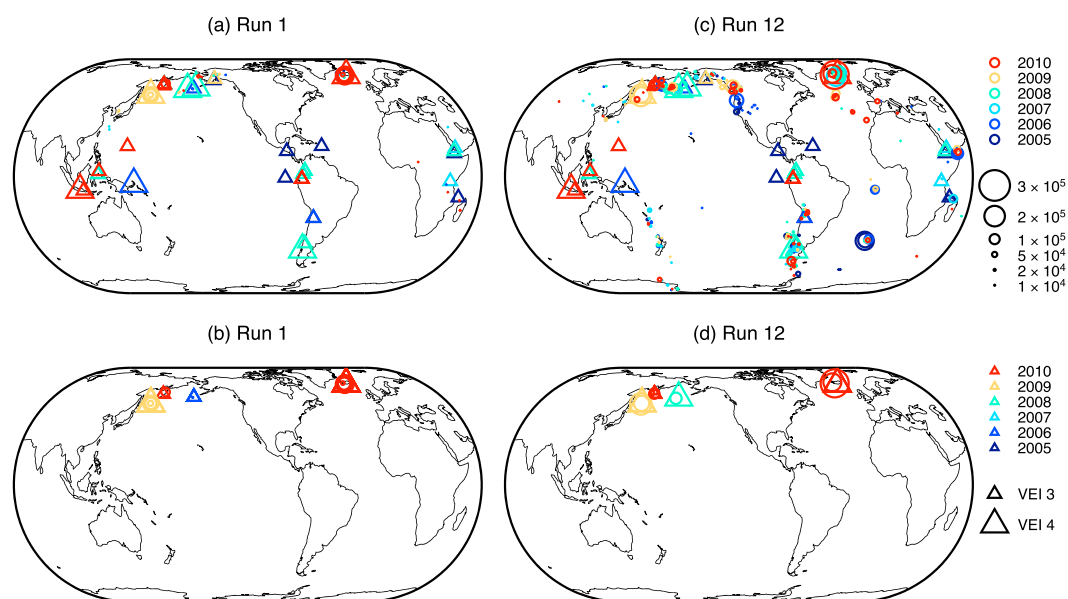


Figure 15. Map view comparison of Smithsonian GVP catalog [Siebert and Simkin, 2002; Global Volcanism Program, 2013] and our infrasonic catalogs Run 1 and Run 12 (Table 1) in the time from 2005 to 2010. The same symbols and legend (right) are used in all figure parts. As with Figure 11, infrasonic events are shown as circles colored by the year of occurrence (legend at the top right) with symbol size proportional to the number of pixels (black circles in top right legend indicate number of pixels). Smithsonian GVP cataloged eruptions ($VEI \geq 3$) are shown as triangles colored by the year of occurrence (legend at the bottom right) with symbol size corresponding to the VEI (black triangles in the bottom right legend indicate VEI symbol size). (a) All Run 1 infrasonic events within 250 km of any potentially active GVP volcano compared with the locations of all cataloged GVP eruptions during this time period. (b) The subset of Run 1 infrasonic events and GVP cataloged events that match both in location (within 250 km distance) and time within ± 90 days. (c and d) The same as Figures 15a and 15b, respectively, but for Run 12 infrasonic events. In Figure 15b, the volcanoes identified are (from left to right) Sarychev Peak (2009), Kizimen or Kliuchevskoi (2010), Cleveland (2006), and Eyjafjallajökull (2010). In (d), the volcanoes identified are (from left to right) Sarychev Peak (2009), Kizimen or Kliuchevskoi (2010), Kasatochi (2008), and Eyjafjallajökull (2010).

60-station IMS infrasound network and with additional regional infrasound arrays. Furthermore, our algorithm is relatively strict in requiring two- or three-station detection (both Runs 1 and Run 12 require three-station detection), but there may be additional information on smaller VEI eruptions recorded by one or two stations that could be exploited with a different algorithm. For example, *Dabrowa et al.* [2011] found that 62% of manually inspected events in the GVP database (many with $VEI < 4$) were detected by at least one IMS infrasound station.

Upon manual inspection, the 7 October 2006 eruption of Rabaul is well recorded on station IS22 (New Caledonia) and marginally detected with a low number of pixels on station IS07 (Warramunga, Australia). Thus, although this eruption is recorded by the IMS infrasound network, it does not meet our criteria for association. This event was also manually analyzed and discussed by *Dabrowa et al.* [2011], who noted signals on IS22 and IS07, with possible weak signals at IS36 (Chatham Island, New Zealand; distance of 5354 km exceeds our 5000 km cutoff). As noted by *Dabrowa et al.* [2011], the 2006 eruption of Rabaul was an outlier in terms of detectable frequency content; the eruption was only seen clearly at IS07 above the microbarom peak (> 0.2 Hz), with a slightly lower minimum frequency of detection at IS22. Additionally, during 2006, IS07 consisted of a mix of both low- and high-frequency (L and H) array elements, which have different dimensions of spatial wind filters [e.g., *Hedlin et al.*, 2003]. Array processing with such a design is challenging, resulting in low numbers of PMCC pixels. This station was recently reconfigured as an eight-element array of high-frequency spatial filters and should be more effective at detecting volcanic eruption signals in the future. Therefore, the lack of infrasonic event in our catalog from the 7 October 2006 Rabaul eruption can be partially attributed to a technical station issue that has now been resolved. Similarly, the May 2008 eruption of Chaiten, Chile was not well recorded by IMS infrasound stations probably due to technical issues with the closest critical stations (e.g., IS02, Argentina had excessive flooding of the sensors [Pautet and Cavahlo, 2009] and IS41, Paraguay had a reduced number of operating elements).

The 2008 eruption of Okmok was recorded by 5 IMS infrasound stations (IS53, IS56, IS59, IS30, and IS18) as documented in detail by *Fee et al.* [2010b]. However, IS59 (Hawaii) is a critical station for this event [*Fee et al.*, 2010b] and is absent in the data considered for our study because IS59 data for 2005–2010 are not available from the IDC (supporting information Figure S1). The absence of IS59 largely explains the lack of infrasound event for Okmok in our catalog. Manual inspection of the grid functions (G^d and G^c) for this eruption reveals that intersecting sectors from IS53 (Fairbanks, AK) and IS30 (Japan) form a local maximum in the grid corresponding to Okmok; IS56 does not contribute significantly to the grids; and the resulting grid G^c at this time has a global maximum elsewhere on the globe. Therefore, modifying our processing to allow events to be defined by regional maxima, rather than simply the global maximum, would permit smaller grid features (events) such as the 2008 Okmok eruption to be included in our infrasonic catalogs.

Detection of the 2010 eruption of Merapi is likely limited by being near the equator and having unfavorable long-range propagation [*Le Pichon et al.*, 2009; *Green and Bowers*, 2010]. As shown in Figure 12, the complete 60-station IMS infrasound network will have improved coverage of volcanic eruptions in Indonesia, most critically by station IS06 (Cocos Islands, Australia), which was certified in December 2011 (after the time period covered by our study).

In addition to the seven VEI 4 events, there are 31 VEI 3 events from 2005–2010 shown in Figures 14 and 15. Manually cross referencing the GVP eruption report descriptions of individual explosive eruptive phases to the infrasonic catalog for each event is beyond the scope of the present study. However, here we perform a first-order comparison to estimate the number of missed VEI 3 events. Figure 15 compares Run 1 and Run 12 event locations with the positions of GVP VEI ≥ 3 eruptions during 2005–2010. In Figures 5b and d, we show only events common to both infrasound and GVP catalogs, where infrasound events and GVP eruptions were determined to match if the locations are within 250 km and the dates are within ± 90 days (to account for a possible discrepancy between reported eruption start time and the time of the most explosive phases). For Run 1 (Figure 15b), the volcanoes identified are Sarychev Peak (VEI 4, June 2009), Kizimen or Kliuchevskoi (VEI 3, August 2010), Cleveland (VEI 3, May 2006), and Eyjafjallajökull (VEI 4, April 2010). For Run 12 (Figure 15d), the volcanoes identified are Sarychev Peak (VEI 4, June 2009), Kizimen or Kliuchevskoi (VEI 3, December 2010), Kasatochi (VEI 4, August 2008), and Eyjafjallajökull (VEI 4, April 2010). Figure 15 indicates that 28 of the 31 VEI 3 eruption events reported by the GVP (90%) were not identified in our infrasonic catalog (missed events).

Of the identified VEI 3 events, the apparent Run 1 infrasonic detection of Cleveland, AK, on 3–4 May 2006 has a low number of pixels and is difficult to interpret with confidence. Eruption reports indicate a thermal anomaly and a continuous plume from Cleveland seen on satellite imagery on 2 May, with no further activity detected after 2 May. However, it is possible that an explosion went unreported since many of Cleveland's explosions were often "hidden" and difficult to detect [*De Angelis et al.*, 2012]. We manually investigated the Run 1 infrasound event in Kamchatka, 30 August 2010 (labeled by ? in Figure 14b and shown in map view in Figure 15b). We examined the raw waveform data and array processing detection lists (PMCC, *F*-stat). This event is formed primarily from a large number of PMCC detections at IS44 (Kamchatka Krai, Russian Federation) combined with detections at IS53 (Fairbanks, USA) and marginal detections at IS45 (Ussuriysk, Russian Federation) and IS30 (Isumi, Japan). Although it is difficult to associate signals at IS44 with a particular volcano in Kamchatka, due to multiple volcanoes falling along the same backazimuth, we attribute the infrasound event to Kliuchevskoi. The GVP bulletins report strombolian activity at Kliuchevskoi at this time, consistent with the extended sequence of short-duration transient events observed in the IS44 waveform. Our interpretation of the event is that a prolonged signal as viewed from IS44 satisfied the three-station detection requirement on 30 August 2010 when station IS53 began to detect the signal (e.g., source signal amplitude increased due to increased activity or favorable atmospheric propagation conditions were established). A consistent infrasound-only event is also reported in the REB. This event represents a volcanic eruption signal identified through the automated blind search process; we did not have prior knowledge of the event. We similarly attribute the Run 12 1–3 December 2010 events to Kliuchevskoi (Figure 15d).

The results shown in Figure 14b further highlight that the June 2009 eruption of Sarychev Peak and the April–May 2010 eruption of Eyjafjallajökull are dominant events in our catalog in terms of numbers of pixels and numbers of detecting stations. Run 1 (Figure 14b) does not capture the 2008 Kasatochi eruption but qualitatively has a lower false-alarm rate (no nonvolcanic events approach the pixel values of Sarychev Peak and Eyjafjallajökull). Run 12 (Figure 14c) includes the 2008 Kasatochi eruption but has a greater number of false alarms (see arbitrary thresholds of 1×10^5 and 2×10^5 pixels for illustration). A quantitative analysis of the

false-alarm rate is beyond the scope of this study; indeed, the false-alarm rate is difficult to evaluate because the GVP catalog may not be complete. However, we performed a preliminary analysis by manually inspecting the waveforms and array processing detection lists for the seven largest (number of pixels) events from Run 12 (Figure 14c) with pixel values between 1×10^5 and 2×10^5 that were not already attributed to known eruptions (Kasatochi, Sarychev, and Eyjafjallajökull). We systematically attributed each of these events to other nonvolcanic sources. Five of these events are connected to the sudden onset of microbarom activity, e.g., related to sudden changes in propagation conditions (violating the assumption of stationary clutter used in section 3.3); the other two are associated with anthropogenic sources (wind farms, gas flares). It is clear from comparing the results from Run 1 and Run 12 that relaxing the constraints on azimuth and frequency content increases the chance of including unwanted microbarom sources. Thus, as expected, the catalog can be made more sensitive (e.g., to detect Kasatochi), but at the expense of an increased rate of false alarms.

7. Conclusions

We have experimented with a brute-force, grid-search, cross-bearings method to search systematically through global multiyear data from the IMS infrasound network and identify sustained explosive volcanic eruption signals originating anywhere on Earth. The algorithm corrects for a background prior clutter rate in a global grid. We applied the algorithm to IMS infrasound data available at the CTBTO IDC from 1 April 2005 to 5 December 2010, a time when the IMS infrasound network was in relatively early stages of development. Our algorithm automatically detects, associates, and locates three of the seven VEI 4 eruptions reported globally during this time period by the Smithsonian GVP [Siebert and Simkin, 2002; *Global Volcanism Program*, 2013]. We anticipate significant improvements in results when applying this algorithm to more recent data from 2011 to present, in future as the IMS infrasound network achieves full 60-station global coverage, and by incorporating data from additional infrasound arrays in regions of dense volcanism. The June 2009 eruption of Sarychev Peak and the April–May 2010 eruption of Eyjafjallajökull represent the largest events identified by a blind search through the global multiyear 2005–2010 IMS data set, demonstrating that this method is well adapted to emphasizing signals from sustained explosive volcanic eruptions. This algorithm could also be adapted to target nonvolcanic signals of interest within clutter for operational monitoring purposes. Extensions to this work include the incorporation of more realistic infrasound propagation and attenuation constraints (e.g., criteria for the vertical structure of the along-path wind), incorporation of more realistic station noise levels and data quality information, modification of the simple detection criteria to permit multiple simultaneous events occurring in different locations on Earth, more detailed automated event characterization (e.g., origin time, acoustic intensity, signal duration, and frequency content metrics) enabled by querying the raw waveform data and array processing results at times of cataloged events, and adaptation of the method for use in a real-time global volcanic notification system.

Acknowledgments

Data from the CTBT IMS infrasound network are available upon request through the CTBTO vDEC platform. The manuscript was improved after reviews by Emanuele Marchetti and an anonymous reviewer. This work was partially supported by a CTBTO Young Scientist Research Award and NSF awards EAR–1113338, EAR–1546139, and EAR–1614855.

References

- Arrowsmith, S., G. Euler, O. Marcillo, P. Blom, R. Whitaker, and G. Randall (2015), Development of a robust and automated infrasound event catalogue using the International Monitoring System, *Geophys. J. Int.*, *200*, 1411–1422, doi:10.1093/gji/ggu486.
- Arrowsmith, S. J., and R. Whitaker (2008), Inframonitor: A tool for regional infrasound monitoring, in *Proceedings of the 2008 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, pp. 837–843, LA-UR-08-05261, Portsmouth, Va.
- Arrowsmith, S. J., R. Whitaker, S. R. Taylor, R. Burlacu, B. Stump, M. Hedlin, G. Randall, C. Hayward, and D. Re Velle (2008), Regional monitoring of infrasound events using multiple arrays: Application to Utah and Washington State, *Geophys. J. Int.*, *175*, 291–300, doi:10.1111/j.1365-246X.2008.03912.x.
- Assink, J. D., A. Le Pichon, E. Blanc, M. Kallel, and L. Khemiri (2014), Evaluation of wind and temperature profiles from ECMWF analysis on two hemispheres using volcanic infrasound, *J. Geophys. Res. Atmos.*, *119*, 8659–8683, doi:10.1002/2014JD021632.
- Blandford, R. R. (1974), An automatic event detector at the Tonto Forest seismic observatory, *Geophysics*, *39*(5), 633–643, doi:10.1190/1.1440453.
- Blom, P. S., O. Marcillo, and S. J. Arrowsmith (2015), Improved Bayesian infrasonic source localization for regional infrasound, *Geophys. J. Int.*, *203*, 1682–1693, doi:10.1093/gji/ggv387.
- Bondár, I., S. C. Myers, E. R. Engdahl, and E. A. Bergman (2004), Epicentre accuracy based on seismic network criteria, *Geophys. J. Int.*, *156*(3), 483–496, doi:10.1111/j.1365-246X.2004.02070.x.
- Brachet, N., D. Brown, R. Le Bras, Y. Cansi, P. Mialle, and J. Coyne (2010), Monitoring the Earth's atmosphere with the global IMS infrasound network, in *Infrasound Monitoring for Atmospheric Studies*, edited by A. L. Pichon, E. Blanc, and A. Hauchecorne, chap. 3, pp. 77–118, Springer, Netherlands.
- Brown, D. J., C. N. Katz, R. Le Bras, M. P. Flanagan, J. Wang, and A. K. Gault (2002), Infrasonic signal detection and source location at the Prototype International Data Centre, *Pure Appl. Geophys.*, *159*(5), 1081–1125, doi:10.1007/s00024-002-8674-2.
- Cansi, Y. (1995), An automatic seismic event processing for detection and location: The P.M.C.C. method, *Geophys. Res. Lett.*, *22*(9), 1021–1024.
- Cansi, Y., and Y. Klinger (1997), An automated data processing method for mini-arrays, *Newslett. European-Mediterr. Seismol. Centre*, *11*, 2–4.

- Casadevall, T. J., et al. (1994), Volcanic ash and aviation safety, in *Proceedings of the First International Symposium on Volcanic Ash and Aviation Safety*, U.S. Geol. Surv. Bull., 2047, 450 pp., DIANE Publ., Washington, D. C.
- Casadevall, T. J., and T. B. Thompson (1995), *World Map of Volcanoes and Principal Aeronautical Features*, U.S. Geol. Surv. Geophys. Invest. Ser. Map, GP-1011, Reston, Va.
- Caudron, C., B. Taisne, M. Garces, A. Le Pichon, and P. Mialle (2015), On the use of remote infrasound and seismic stations to constrain the eruptive sequence and intensity for the 2014 Kelud eruption, *Geophys. Res. Lett.*, 42, 6614–6621, doi:10.1002/2015GL064885.
- Caudron, C., B. Taisne, A. Perttu, M. Garcés, E. A. Silber, and P. Mialle (2016), Infrasound and seismic detections associated with the 7 September 2015 Bangkok fireball, *Geosci. Lett.*, 3(1), 26, doi:10.1186/s40562-016-0058-z.
- Christie, D., and P. Campus (2010), The IMS infrasound network: Design and establishment of infrasound stations, in *Infrasound Monitoring for Atmospheric Studies*, edited by A. L. Pichon, E. Blanc, and A. Hauchecorne, chap. 2, pp. 29–75, Springer, Netherlands.
- Coffin, M. F., L. M. Gahagan, and L. A. Lawver (1998), Present-day plate boundary digital data compilation, *Inst. Geophys. Tech. Rep.*, 174, Univ. of Texas.
- Crowther, H. S., et al. (2012), Global database on large magnitude explosive volcanic eruptions (LaMEVE), *J. Appl. Volcanol.*, 1, 1–13, doi:10.1186/2191-5040-1-4.
- Dabrowa, A. L., D. N. Green, A. C. Rust, and J. C. Phillips (2011), A global study of volcanic infrasound characteristics and the potential for long-range monitoring, *Earth Planet. Sci. Lett.*, 310, 369–379, doi:10.1016/j.epsl.2011.08.027.
- De Angelis, S., D. Fee, M. Haney, and D. Schneider (2012), Detecting hidden volcanic explosions from Mt. Cleveland Volcano, Alaska with infrasound and ground-coupled airwaves, *Geophys. Res. Lett.*, 39, L21312, doi:10.1029/2012GL053635.
- Drob, D. P., M. Garces, M. Hedlin, and N. Brachet (2010), The temporal morphology of infrasound propagation, *Pure Appl. Geophys.*, 167(4–5), 437–453, doi:10.1007/s00024-010-0080-6.
- Evers, L. G., and H. W. Haak (2005), The detectability of infrasound in the Netherlands from the Italian volcano Mt. Etna, *J. Atmos. Sol. Terr. Phys.*, 67(3), 259–268, doi:10.1016/j.jastp.2004.09.002.
- Fee, D., and R. S. Matoza (2013), An overview of volcano infrasound: From hawaiian to plinian, local to global, *J. Volcanol. Geotherm. Res.*, 249, 123–139, doi:10.1016/j.jvolgeores.2012.09.002.
- Fee, D., M. Garces, and A. Steffke (2010a), Infrasound from Tungurahua Volcano 2006–2008: Strombolian to Plinian eruptive activity, *J. Volcanol. Geotherm. Res.*, 193, 67–81, doi:10.1016/j.jvolgeores.2010.03.006.
- Fee, D., A. Steffke, and M. Garces (2010b), Characterization of the 2008 Kasatochi and Okmok eruptions using remote infrasound arrays, *J. Geophys. Res.*, 115, D00L10, doi:10.1029/2009JD013621.
- Fee, D., R. S. Matoza, K. L. Gee, T. B. Neilsen, and D. E. Ogden (2013), Infrasonic crackle and supersonic jet noise from the eruption of Nabro Volcano, Eritrea, *Geophys. Res. Lett.*, 40, 4199–4203, doi:10.1002/grl.50827.
- Garces, M., D. Fee, D. McCormack, R. Servranckx, H. Bass, C. Hetzer, M. Hedlin, R. Matoza, and H. Yepes (2008), Prototype ASHE volcano monitoring system captures the acoustic fingerprint of stratospheric ash injection, *Eos Trans. AGU*, 89(40), 377–379.
- Garces, M. A., R. A. Hansen, and K. G. Lindquist (1998), Traveltimes for infrasonic waves propagating in a stratified atmosphere, *Geophys. J. Int.*, 135(1), 255–263, doi:10.1046/j.1365-246X.1998.00618.x.
- Global Volcanism Program (2013), *Volcanoes of the World v. 4.5.3*, edited by E. Venzke, Smithsonian Institution. [Available at <http://dx.doi.org/10.5479/si.GVP.VOTW4-2013> accessed on June 2016.]
- Green, D., and A. Nippres (2015), *Five Years of Infrasound Arrivals in the International Data Center Bulletins: A review*, Infrasound Technology Workshop (ITW), Vienna, Austria.
- Green, D. N., and D. Bowers (2010), Estimating the detection capability of the International Monitoring System infrasound network, *J. Geophys. Res.*, 115, D18116, doi:10.1029/2010JD014017.
- Green, D. N., L. Evers, D. Fee, R. Matoza, M. Snellen, P. Smets, and D. Simons (2013), Hydroacoustic, infrasonic and seismic monitoring of the submarine eruptive activity and sub-aerial plume generation at South Sarigan, May 2010, *J. Volcanol. Geotherm. Res.*, 257, 31–43, doi:10.1016/j.jvolgeores.2013.03.006.
- Guilbert, J., P. Harjadi, M. Purbawinata, S. Jammes, A. Le Pichon, and B. Feignier (2005), Monitoring of Indonesian volcanoes with infrasound: Preliminary results, Infrasound Technology Workshop, Tahiti.
- Hedlin, M. A. H., B. Alcoverro, and G. D’Spain (2003), Evaluation of rosette infrasonic noise-reducing spatial filters, *J. Acoust. Soc. Am.*, 114, 1807–1820.
- Kamo, K., K. Ishihara, and M. Tahira (1994), Infrasonic and seismic detection of explosive eruptions at Sakurajima Volcano, Japan, and the PEGASAS-VE early-warning system, in *Proceedings of the 1st International Symposium on Volcanic Ash and Aviation Safety*, edited by T. C. Casadevall, U.S. Geol. Surv. Bull., 2047, pp. 357–365, Denver, Colo.
- Landès, M., L. Ceranna, A. Le Pichon, and R. S. Matoza (2012), Localization of microbarom sources using the IMS infrasound network, *J. Geophys. Res.*, 117, D06102, doi:10.1029/2011JD016684.
- Le Pichon, A., J. Vergoz, P. Herry, and L. Ceranna (2008), Analyzing the detection capability of infrasound arrays in Central Europe, *J. Geophys. Res.*, 113, D12115, doi:10.1029/2007JD009509.
- Le Pichon, A., J. Vergoz, E. Blanc, J. Guilbert, L. Ceranna, L. Evers, and N. Brachet (2009), Assessing the performance of the International Monitoring System infrasound network: Geographical coverage and temporal variabilities, *J. Geophys. Res.*, 114, D08112, doi:10.1029/2008JD010907.
- Le Pichon, A., R. Matoza, N. Brachet, and Y. Cansi (2010), Recent enhancements of the PMCC infrasound signal detector, *Inframatics*, 26, 5–8.
- Mason, B. G., D. M. Pyle, and C. Oppenheimer (2004), The size and frequency of the largest explosive eruptions on Earth, *Bull. Volcanol.*, 66(8), 735–748, doi:10.1007/s00445-004-0355-9.
- Matoza, R. S., M. A. H. Hedlin, and M. A. Garces (2007), An infrasound array study of Mount St. Helens, *J. Volcanol. Geotherm. Res.*, 160, 249–262.
- Matoza, R. S., D. Fee, M. Garces, J. M. Seinar, P. A. Ramon, and M. A. H. Hedlin (2009), Infrasonic jet noise from volcanic eruptions, *Geophys. Res. Lett.*, 36, L08303, doi:10.1029/2008GL036486.
- Matoza, R. S., A. Le Pichon, J. Vergoz, P. Herry, J. M. Lalonde, H. Lee, I. Che, and A. Rybin (2011a), Infrasonic observations of the June 2009 Sarychev Peak eruption, Kuril Islands: Implications for infrasonic monitoring of remote explosive volcanism, *J. Volcanol. Geotherm. Res.*, 200, 35–48, doi:10.1016/j.jvolgeores.2010.11.022.
- Matoza, R. S., et al. (2011b), Long-range acoustic observations of the Eyjafjallajökull eruption, Iceland, *Geophys. Res. Lett.*, 38, L06308, doi:10.1029/2011GL047019.
- Matoza, R. S., D. Fee, T. B. Neilsen, K. L. Gee, and D. Ogden (2013a), Aeroacoustics of volcanic jets: Acoustic power estimation and jet velocity dependence, *J. Geophys. Res. Solid Earth*, 118, 6269–6284, doi:10.1002/2013JB01030.
- Matoza, R. S., M. Landès, A. Le Pichon, L. Ceranna, and D. Brown (2013b), Coherent ambient infrasound recorded by the International Monitoring System, *Geophys. Res. Lett.*, 40, 429–433, doi:10.1029/2012GL054329.

- Mazzocchi, M., F. Hansstein, and M. Ragona (2010), The 2010 volcanic ash cloud and its financial impact on the European airline industry, *CEifo Forum, Ifo Institute for Economic Research at the University of Munich*, 11(2), 92–100.
- Mialle, P., P. Bittner, N. Brachet, D. Brown, J. Given, R. Le Bras, and J. Coyne (2010), Infrasound event analysis into the IDC operations.
- Mialle, P., et al. (2015), Towards a volcanic notification system with infrasound data: Use of infrasound data in support of the VAACs in the framework of ARISE project, in *World Meteorological Organization 7th International Workshop on Volcanic Ash (IWVA/7)*, Anchorage, Alaska.
- Modrak, R. T., S. J. Arrowsmith, and D. N. Anderson (2010), A Bayesian framework for infrasound location, *Geophys. J. Int.*, 181, 399–405, doi:10.1111/j.1365-246X.2010.04499.x.
- Newhall, C. G., and S. Self (1982), The Volcanic Explosivity Index: An estimate of explosive magnitude for historical volcanism, *J. Geophys. Res.*, 87(C2), 1231–1238.
- Nippress, A., D. N. Green, O. E. Marcillo, and S. J. Arrowsmith (2014), Generating regional infrasound celerity-range models using ground-truth information and the implications for event location, *Geophys. J. Int.*, 197, 1154–1165, doi:10.1093/gji/ggu049.
- Park, J., S. J. Arrowsmith, C. Hayward, B. W. Stump, and P. Blom (2015), Automatic infrasound detection and location of sources in the western United States, *J. Geophys. Res. Atmos.*, 119, 7773–7798, doi:10.1002/2013JD021084.
- Pautet, L., and J. Cavahlo (2009), Through-the-sensor maintenance monitoring tool, Infrasound Technology Workshop, Brasilia.
- Siebert, L., and T. Simkin (2002), *Volcanoes of the world: An illustrated catalog of holocene volcanoes and their eruptions*. Smithsonian Institution, Global Volcanism Program Digital Information Series GVP-3. [Available at <http://www.volcano.si.edu/world/>.]
- Siebert, L., T. Simkin, and P. Kimberly (2010), *Volcanoes of the World*, 3rd ed., Univ. of California Press, Berkeley.
- Simkin, T., and L. Siebert (2000), Earth's volcanoes and eruptions: An overview, in *Encyclopedia of Volcanoes*, edited by T. Simkin and L. Siebert, pp. 249–261, Academic Press, San Diego, Calif.
- Smirnov, A., I. Sokolova, and V. Dubrovin (2015), *The Variety of Infrasound Sources Recorded by Kazakhstani Stations*, CTBT Science and Technology (SnT) Workshop, Vienna, Austria.
- Tailpied, D., A. Le Pichon, E. Marchetti, M. Ripepe, M. Kallel, L. Ceranna, and N. Brachet (2013), Remote infrasound monitoring of Mount Etna: Observed and predicted network detection capability, *InfraMatics*, 2(1), 1–11, doi:10.4236/inframatics.2013.21001.
- Ulvieri, G., M. Ripepe, and E. Marchetti (2013), Infrasound reveals transition to oscillatory discharge regime during lava fountaining: Implication for early warning, *Geophys. Res. Lett.*, 40, 3008–3013, doi:10.1002/grl.50592.
- Virieux, J., N. Garnier, E. Blanc, and J. X. Dessa (2004), Paraxial ray tracing for atmospheric wave propagation, *Geophys. Res. Lett.*, 31, L20106, doi:10.1029/2004GL020514.
- Walker, K. T., R. Shelby, M. Hedlin, C. de Groot-Hedlin, and F. Vernon (2011), Western US infrasonic catalog: Illuminating infrasonic hot spots with the USArray, *J. Geophys. Res.*, 116, B12305, doi:10.1029/2011JB008579.
- Webley, P., and L. Mastin (2009), Improved prediction and tracking of volcanic ash clouds, *J. Volcanol. Geotherm. Res.*, 186(1–2), 1–9, doi:10.1016/j.jvolgeores.2008.10.022.