ARTICLE

OPEN

Check for updates

Remotely imaging seismic ground shaking via large-N infrasound beamforming

Jacob F. Anderson [1✉], Jeffrey B. Johnson[1], T. Dylan Mikesell [2] & Lee M. Liberty[1]

Seismic ground motion creates low-frequency atmospheric sound (infrasound) that is detectable at remote sensor arrays. However, earthquake infrasound signal analysis is complicated by interference between multiple waves arriving at sensors simultaneously, reducing the accuracy and detail of ground motion detection. Here we show that individual waves in complicated wavefields can be resolved by recording infrasound on large-N arrays and processing with CLEAN beamforming. Examining both a local (ML3.5, purely tropospheric infrasound propagation) and regional earthquake (ML6.5, upper-atmospheric returns), we detect infrasound from tens of km away and up to several hundred km away respectively. Source regions span arcs of approximately 90°, indicating that although detection bias does occur (most likely from atmospheric winds) the recorded infrasound sources are widely dispersed and not simply epicentral. Infrasound-based remote detection of ground motion over wide areas can complement point measurements by seismometers and spur innovations in earthquake research and real-time hazard monitoring.

[1] Department of Geosciences, Boise State University, Boise, ID 83725, USA. [2] Norwegian Geotechnical Institute, Oslo, Norway. ✉email: jacobanderson152@boisestate.edu

Ground shaking exerts normal tractions on the atmosphere, exciting infrasound waves; due to the strong wave speed contrast between the ground and air, this seismic-coupled infrasound propagates at steep incidence angles to the ground surface (Fig. 1). In this paper, we use slowness (time to travel a horizontal distance across a sensor array) as an observable indicator of a wave's direction and origin: earthquake infrasound detections can be due to either (1) "primary" or "local" infrasound that converts from seismic to infrasound at the recording site and propagates upward (often escaping atmospheric waveguides) with low slowness (high horizontal trace speed), or (2) diffracted "secondary" infrasound that converts from seismic to infrasound at topographic features and propagates to recording sites subhorizontally (i.e., with high slowness around 3 s km^{-1}, the inverse of the speed of sound in air)[1–3].

A single infrasound array can record secondary infrasound from nearby or distant ground shaking, making secondary infrasound a promising low-latency ground motion monitoring and research tool. Infrasound waves that travel distances as far as tens of km ("proximal" secondary infrasound[4]) follow simple, nearly linear paths; however, waves traveling longer distances are influenced by refractions in the troposphere[5] or upper atmosphere[6], which depend on time-variable atmospheric wind and temperature structures. Although atmospheric models can be used to predict infrasound arrival patterns, unresolvable short-time-scale atmospheric variability affects their accuracy[5]. Proximal infrasound, confined to the troposphere, experiences milder atmospheric path effects, but effects from topographic obstruction and diffraction can be noticeable[7].

The potential value of secondary infrasound in ground motion monitoring motivates efforts to backproject it to its sources[4,8]. Key steps in the backprojection process include (1) calculating the backazimuth of secondary infrasound, (2) estimating seismic/infrasound travel times from the earthquake hypocenter to candidate conversion points to the infrasound array, and (3) identifying seismic-to-infrasound conversion points considering the total travel time and infrasound backazimuth. The backprojection results can be complicated by ambiguity in the type (and speed) of seismic waves and potential multipathing.

The first step of the backprojection process–finding the direction toward infrasound sources–requires beamforming analyses of data from infrasound arrays (multiple sensors

separated by approximately tens of meters). Beamforming with infrasound arrays is a common method for locating diverse phenomena, e.g. volcanic explosions, nuclear tests, ocean swells, lahars, and avalanches[9–13]. Usual beamforming methods identify a single best-fit propagation direction per time interval analyzed, which is acceptable when the wavefields being analyzed have one dominant signal and minor background noise. However, secondary earthquake infrasound arrives from many directions simultaneously and it cannot be fully described using single-solution beamforming. We present two methodological improvements that enable beamforming of simultaneous arrivals from multiple directions, enabling scientists to more fully exploit information on infrasound sources and atmospheric structure contained in the wavefield.

First, we use the CLEAN beamforming algorithm, which identifies distinct wavefield components by iteratively deconvolving the array response from the slowness spectrum (Fig. 2). CLEAN has been applied to various physical problems, e.g., microseism seismic signals[14] and microbarom infrasound[15], but has seen little use for studying rapidly changing wavefields like secondary earthquake infrasound.

Second, we improve the resolution of CLEAN beamforming by recording data with an unprecedented "large-N" infrasound array. We define large-N functionally as being many more sensors than the minimum needed to identify a source direction or location (3 and 4, respectively), enabling source analyses to identify multiple sources and perform well even with poor signal-to-noise ratio. The number of sensors described as large-N varies (as much as hundreds[16] or thousands[17] of sensors in seismology, where the method is more established), and sensor counts will probably increase as technological improvements make installing large networks more feasible. Large-N arrays benefit from having a compact array response enabling distinction of signals arriving from different directions, (Fig. 3) while avoiding issues with spatial aliasing and loss of coherence between sensors[18], with ancillary benefits of noise reduction and redundancy in case of sensor failure.

The utility of large-N infrasound arrays, likely to include enhanced detection of weak sources and changes in atmospheric structure, remains largely untested. However, the related field of seismology has recently been transformed by the ability to distinguish weak signals and multiple sources through the use of large-N seismometer arrays, leading to revolutionary improvements in seismic imaging at local[16,17] and continental[19] scales, with benefits scalable to array size up to at least thousands of sensors.

Despite its benefits, the large-N strategy (in this paper, $N = 17–22$ sensors) is difficult because of the lifetime costs and installation/maintenance effort for instrumentation. Passive-source seismology overcame similar challenges through the use of low-cost, low-effort "nodal" seismometers[16]. In infrasound, the Gem infrasound logger[20] fills a similar niche with its low cost, low power needs, and rapid cable-free deployment; its use in our large-N array made installation and maintenance of this site feasible.

Fig. 1 Generation of earthquake infrasound. The earthquake infrasound begins with upward-propagating primary infrasound converted at the recording site's ground surface, followed by secondary infrasound arrivals that propagate sub-horizontally from the two hills and arrive at the array almost simultaneously. Slowness component labels describe horizontally-diffracted secondary infrasound (3 s km^{-1} horizontal slowness) and vertical primary infrasound (near-zero horizontal slowness).

Results

We examine two earthquakes recorded during low-noise night-time conditions at the PARK infrasound array (Fig. 3), which we installed to monitor local aftershocks following the M_L6.5 2020-03-31 Stanley, Idaho earthquake[21,22]. The first earthquake we study is a M_L3.5 aftershock[23,24] that occurred beneath PARK and produced more than 1 min of correlated infrasound, beginning with primary infrasound and followed by proximal secondary infrasound from many backazimuths. The second earthquake we
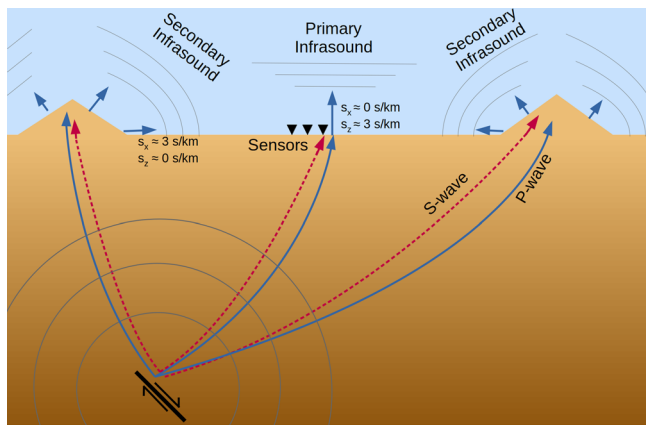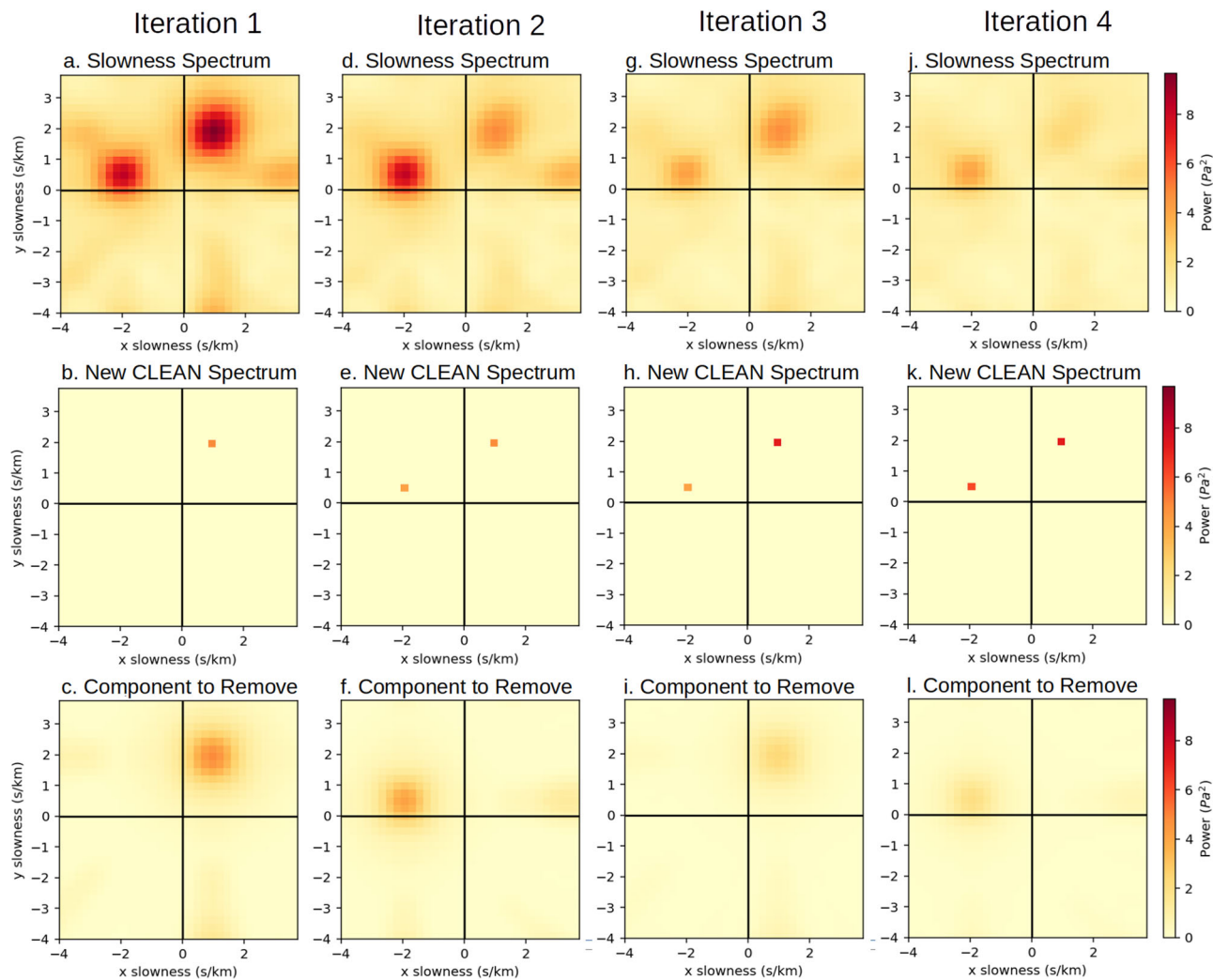
Fig. 2 Demonstration of the iterative CLEAN method. a In each iteration, identify the maximum power of the slowness spectrum and its slowness coordinates (sx,sy). **b** Add a delta function with the corresponding slowness coordinates and power to the clean spectrum. **c** Define the component to remove as the array response centered at that slowness coordinate with power equal to the spectrum's max power times a factor phi (we used phi = 0.1 when analyzing earthquakes in this paper, but use a larger value in this figure for demonstration clarity). **d** The new slowness spectrum is calculated by subtracting panel c from panel a, which may result in a new slowness vector having the maximum power. **e–l** Repeating this process many times removes coherent signals from the slowness spectrum and accounts for them as discrete waves in the clean spectrum instead. Although the component to subtract and the results are shown as slowness maps in this figure, the actual subtraction is performed in the cross-spectral domain.

study is a $M_L 6.5$ regional earthquake[25,26] ("Monte Cristo") that occurred 718 km away with a backazimuth of −160°. The regional earthquake includes tens of minutes of infrasound, including primary infrasound, proximal secondary infrasound, and distant secondary infrasound. Of the 22 sensors originally deployed in PARK, 20 recorded the local aftershock and 17 recorded the regional earthquake.

Local $M_L 3.5$ event. With the local earthquake example (0.6 km S, 7.5 km deep, ~1 km uncertainty via relocation analysis[23]), primary infrasound from incident seismic waves is detected in the seconds after the earthquake onset, and is followed by tens of seconds of secondary infrasound (Fig. 4). Results from the CLEAN analysis are plotted as image colors and show initial low-slowness primary infrasound followed by a longer phase dominated by high-slowness secondary infrasound. Secondary infrasound waves come from multiple directions that converge over time toward the northwest. Standard beamforming analyses only identify the slowness with maximum semblance or power for each time window; such results are plotted over the CLEAN

results as dots. Considerably more details can be observed in the CLEAN results than with standard beamforming.

Regional $M_L 6.5$ Monte Cristo event. With the regional earthquake example[25,26] (720 km, -160° backazimuth), we observe a lower signal-to-noise ratio than the local event, with pervasive background noise before and during the event (e.g., the 90°–180° quadrant, and around −135° to −115°, Fig. 5c). The noise around −135° to −115° is attributed to a waterfall 8 km from the array and is frequently present at this array[27]; we lack specific candidate sources for the 90°-180° noise. We suppressed the noise by calculating the pre-event power distribution at each azimuth and plotting only the signals that significantly exceed background noise (Fig. 5d). Because all arrivals travel at minimum the full distance from the hypocenter to the array, converting from seismic to infrasound waves en route, and because seismic waves travel much faster than infrasound, we expect waves that converted near the array to arrive earliest, waves that converted more distantly to arrive at intermediate times, and waves converted near the epicenter to arrive later still.
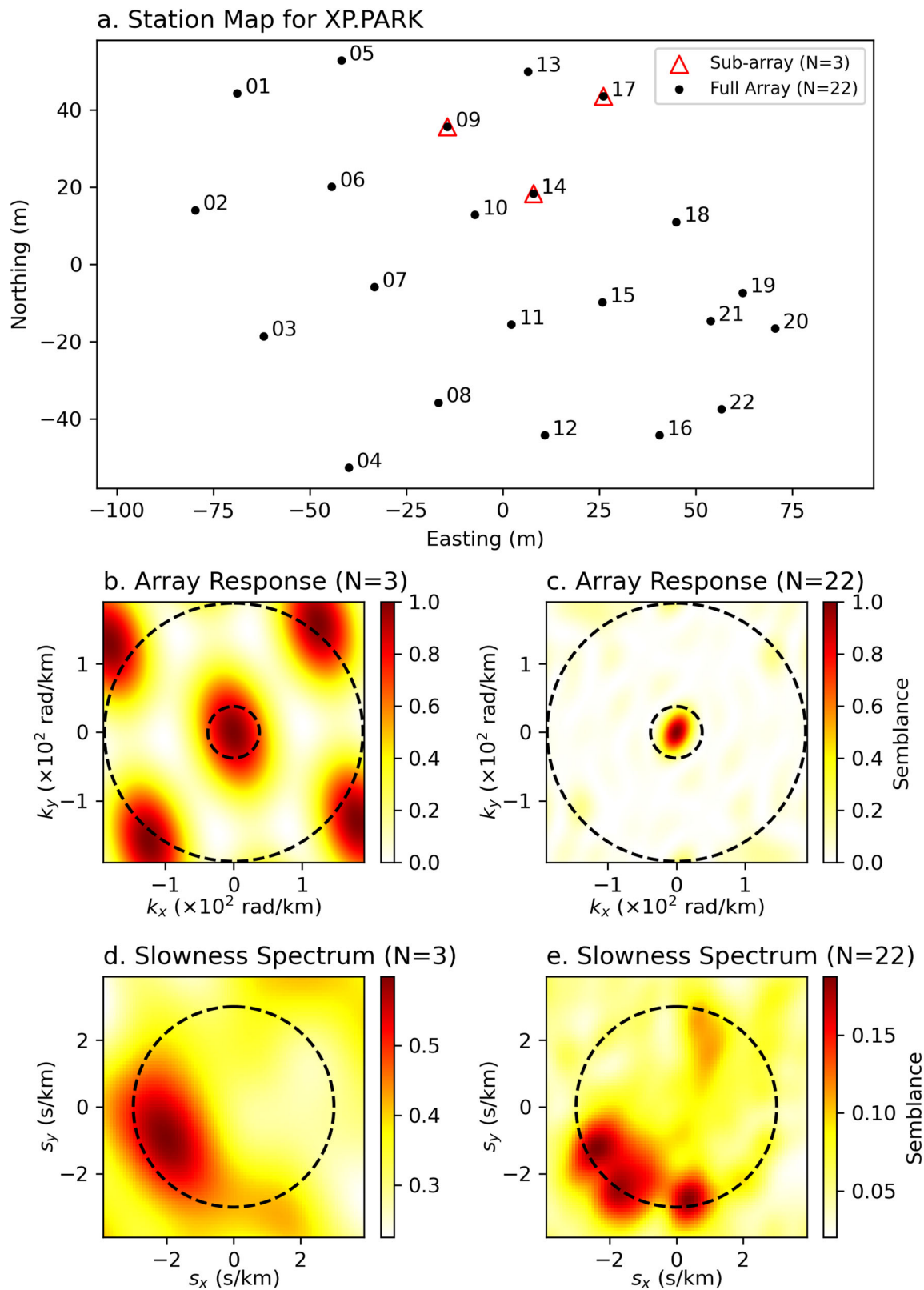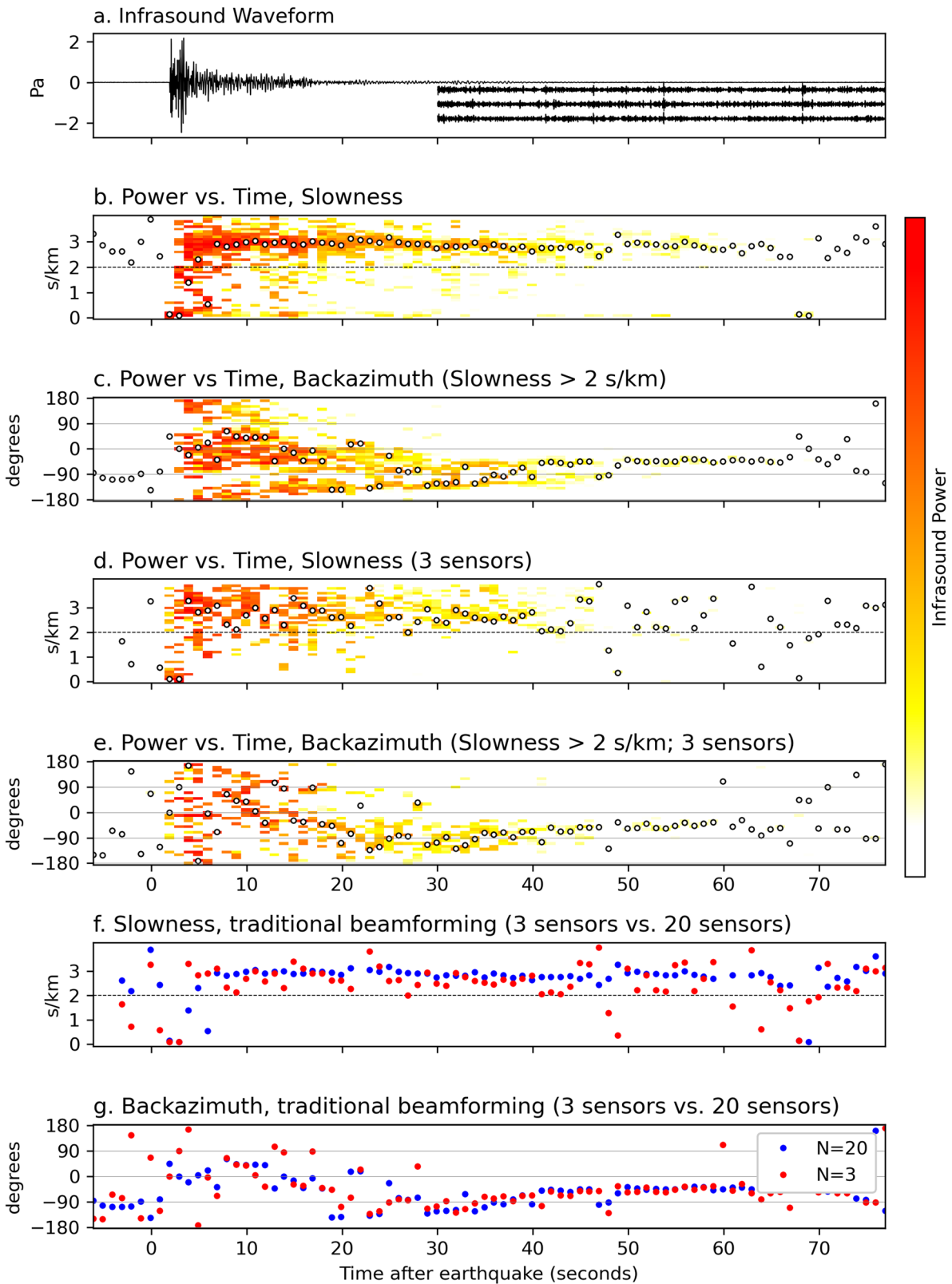
Fig. 3 Effect of large-N arrays in conventional beamforming (not CLEAN). a Array geometries used in this and other figures: the PARK 22-element large-N array and a three-element subset. **b** The three-element array's response has a broad main lobe and close sidelobes, compared to **c** the compact response of the large-N array. Dashed circles in (**b**, **c**) are wavenumbers corresponding to infrasound with frequency of 2 Hz (inner) and 10 Hz (outer). **d** The slowness spectrum of secondary earthquake infrasound calculated using the three-element array is blurry and less detailed than (**e**), the slowness spectrum of the same data calculated using the large-N array. Dashed circles in (**d**, **e**) indicate 3 s km^{-1} (the slowness of sound).

a. Infrasound Waveform

b. Power vs. Time, Slowness

c. Power vs Time, Backazimuth (Slowness > 2 s/km)

d. Power vs. Time, Slowness (3 sensors)

e. Power vs. Time, Backazimuth (Slowness > 2 s/km; 3 sensors)

f. Slowness, traditional beamforming (3 sensors vs. 20 sensors)

g. Backazimuth, traditional beamforming (3 sensors vs. 20 sensors)

Time after earthquake (seconds)

In contrast to the local event's minute of detectable signals, waves from the regional event arrive in a complicated pattern spanning over 40 min (Fig. 5). Like the local event, we interpret the first signals as primary infrasound. Body wave arrivals modeled using the 1-D IASP91 Earth model range from 95–214 s, in approximate agreement with the primary infrasound arrival times of 120–280 s. Secondary infrasound follows the primary infrasound, and due to

its early arrival time (150–500 s after earthquake), we infer it originated locally (within tens of km of the array). During an ensuing gap in secondary infrasound from the regional earthquake, two local signals are detected: primary and secondary infrasound from a small local earthquake between 670–710 s, and an inferred vehicle at 910–935 s (whose rapidly changing backazimuth agrees with highway speeds on nearby ID-21). Secondary infrasound from

Fig. 4 Analysis of the M_L3.5 local event (2020-04-14 03:27 UTC, 0.6 km S of the station, 7.5 km below the surface). a Infrasound waveform (from PARK.14). Subsequent panels show CLEAN results as an image and traditional beamforming results as dots; traditional beamforming results necessarily agree with CLEAN results but only give a single solution and fail to show the wavefield's complexity. Inset shows three traces (subarray: PARK.09, PARK.14, PARK.17) with amplitude magnified by 30x, showing that coherent signal continues to arrive after amplitude drops. **b** CLEAN beamforming of 20 sensors shows that the wavefield is dominated by low-slowness primary infrasound for the first ~5 s and by secondary infrasound subsequently, although low-power primary infrasound is still detected for tens of seconds afterward as different seismic waves continue to arrive. The low-slowness feature around time 67–70 s is weak but coherent primary infrasound, perhaps from a different small aftershock. The dashed line at 2 s km^{-1}, corresponding to an infrasound incidence angle of 42°, indicates the boundary between inferred seismic-coupled primary infrasound and other infrasound. **c** Coherent ambient noise before the primary infrasound is visible with traditional beamforming (dots with consistent backazimuth of -90 to -100); it is also present in the CLEAN results, but the image colors in this plot do not show it due to its low power. Secondary earthquake infrasound arrives from many directions initially, whereas the latest-arriving infrasound comes strictly from the northwest. Repeating the analysis using only three stations yields consistent but less-precise results for both (**d**) horizontal slowness and (**e**) backazimuth. A comparison of single-source beamforming between the three-element sub-array and the full array in horizontal slowness (**f**) and backazimuth (**g**) shows that the three-element sub-array tends to have slownesses that differ more from the expected 3 s km^{-1}, as well as greater scatter in both slowness and backazimuth.

the regional earthquake arrives from the south and east between 900–2200 s; its arrival times indicate that it originated far from both the array and the earthquake epicenter. Finally, infrasound arrives from the epicenter's backazimuth at times similar to the 2470 s arrival time predicted using the infraGA atmospheric ray-tracing model. A subsequent signal from 2600–2800 s with variable horizontal slowness and rapidly changing backazimuth is attributed to an aircraft. The full array is required to see most of these features (Fig. 5d); pervasive noise obscures them with the 3-element array (Fig. 5g).

Sources of secondary infrasound. Backprojected possible secondary infrasound sources correspond partly to topographic features (Fig. 6). For the local event (beneath PARK), secondary infrasound generation strictly occurs on the near side of mountain crests, suggesting that topographic obstruction prevents infrasound radiated on the far side of crests from reaching the infrasound array (Fig. 6a). In the regional event, non-local secondary infrasound is first observed from ranges south and southeast of PARK and not from the nearer lowlands (Fig. 6b). In both events, mountains and basins both appear within source regions, possibly due to uncertainty from seismic waves propagating at multiple speeds. Source regions do not strictly coincide with mountain ranges in either event. We credit the detection of a wide range of secondary infrasound sources to the superior resolving capabilities of CLEAN beamforming with a large-N array.

In both earthquakes, secondary infrasound overwhelmingly arrives from ~90°-wide sectors that coincide with the dominant backazimuths of diffuse background noise (excluding noise from a nearby waterfall), suggesting that both noise and secondary infrasound may be ducted by favorable atmospheric conditions (Figs. 4c, 5c, d). In the regional earthquake, the region of candidate infrasound sources includes many seismic stations (Fig. 6b, c), whose amplitudes during the earthquake do not explain whether infrasound was detected from nearby, supporting the inference that the region from which infrasound could be detected (Fig. 6b) is determined by infrasound propagation effects rather than by seismic amplitude. Importantly, the direction toward the epicenter is not strongly related to the backazimuths of secondary infrasound: the local earthquake epicenter approximately coincides with the array, and the regional earthquake epicenter is on the edge of the detected source region.

Effectiveness of advanced array analysis. For both the local (Fig. 4) and regional earthquakes (Fig. 5), the large-N array provides much more detailed information about earthquake and ambient noise infrasound wavefields than the three-element sub-

array, regardless of whether single-solution or CLEAN beamforming is used (Fig. 4). In particular, the large-N array provides excellent horizontal slowness resolution, indicated by peak slowness matching sonic slowness (3 s km^{-1}) for most time windows during the aftershock's secondary infrasound period starting at 5 s (Fig. 4b). By contrast, estimates of horizontal slowness using the three-element array have considerable scatter during the secondary infrasound period (Fig. 4d), resulting in many signals with lower slowness estimates being misclassified as primary infrasound (seismic arrivals) and omitted from the secondary infrasound image (Fig. 4e). The superior performance of the large-N array is clear when their single-source beamforming results are overlain (Fig. 4f, g). The poor horizontal slowness resolution of the three-element array is even more pronounced for the regional event (Fig. 5b), probably due to its low signal-to-noise ratio (Fig. 5a), which complicates the separation of primary infrasound. Even a compact source like the aircraft, obvious with the large-N array, is nebulous with the 3-element sub-array. Additionally, the process of identifying and removing likely background noise is less effective with the sub-array than with the large-N array (Fig. 5d, g). These performance trends hold for intermediate numbers of sensors (e.g., $N = 6$, shown in Supplementary Figs. 1–2). Finally, large-N arrays have the considerable benefit of redundancy; the loss of 2–5 sensors in a 22-element array had little effect on its resolving capabilities, whereas 3-element arrays lose their ability to calculate a wave's backazimuth with the loss of a single sensor.

The benefits of CLEAN vary between our two case studies. In the high signal-to-noise ratio local earthquake, the traditional beams (dots in Fig. 4) yield an incomplete representation of the wavefield, but do capture major features revealed by CLEAN (Fig. 4b, c). However, in the low signal-to-noise ratio regional earthquake, CLEAN is required for a step that we found essential—calculating the background noise distribution at each azimuth and plotting only the signals that significantly exceed background noise (Fig. 5c, d).

By design, CLEAN approximates a wavefield as the convolution of the array response with a finite number of delta functions in slowness space. Consequently, it is inherently prone to misrepresenting diffuse sounds as a combination of discrete waves[14]. When studying diffuse wavefields like secondary infrasound, this weakness can be mitigated by using large-N arrays because their compact array response leads to diffuse wavefields being approximated as many closely-spaced points instead of a few sparse points.

Large-N arrays incur costs including funds for instrument purchasing, field effort and site footprint, and computational expense. These can be mitigated by choice of instrumentation (using low-cost, rapid-deploy, low-maintenance devices) and
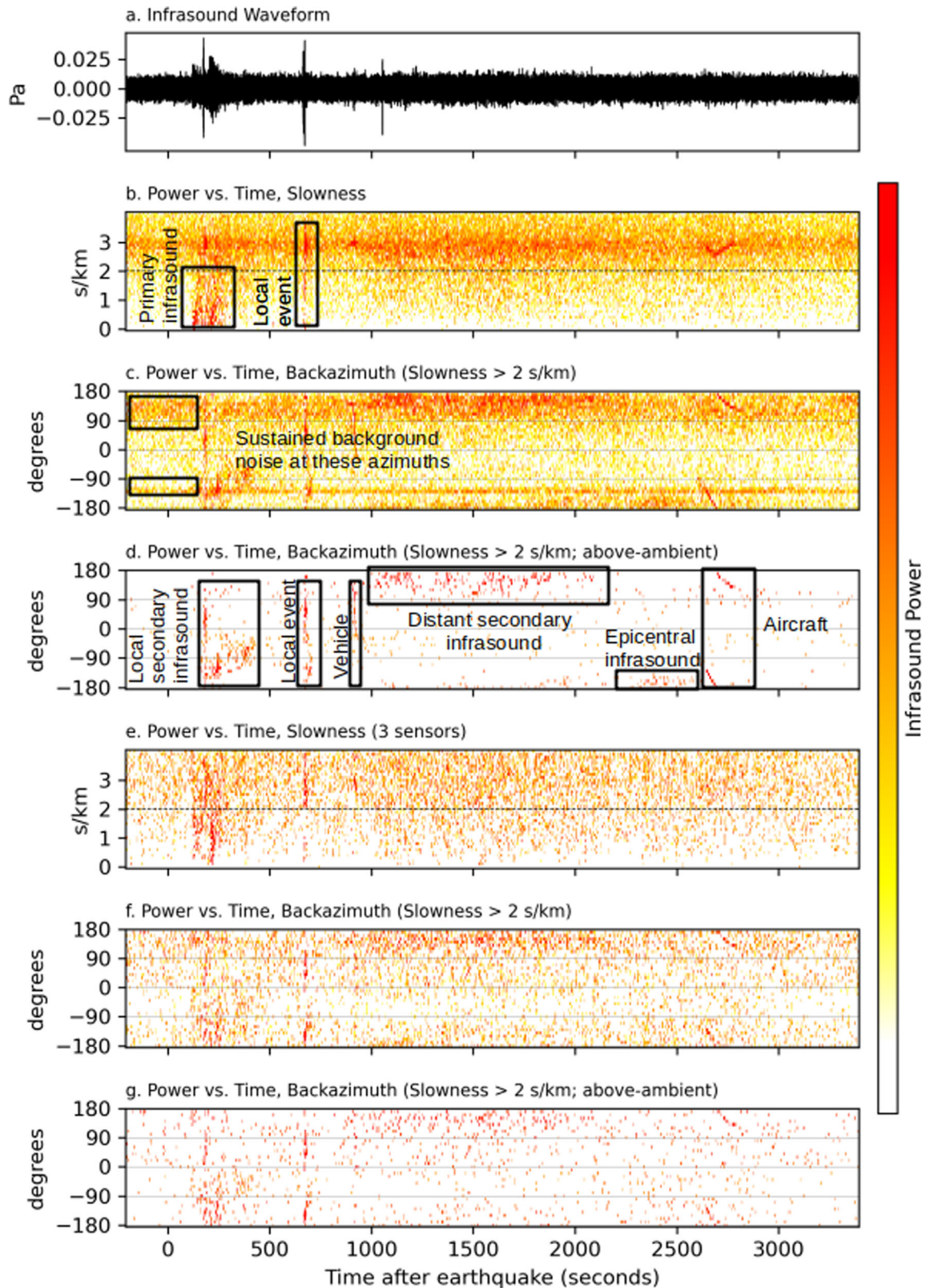
Fig. 5 The M_L6.5 Monte Cristo, NV regional earthquake (2020-05-15 11:03 UTC, 720 km distance, −160° backazimuth). a Infrasound waveform.
b Analysis of 17 sensors shows that the wavefield consists mainly of high-slowness secondary infrasound, except for arrivals of low-slowness primary infrasound coupled with seismic arrivals from the regional event and a subsequent local event. **c** Infrasound is dominated by continuous noise from two backazimuth ranges (90°–180°, and -135° to -115°), though other signals also appear. **d** Clusters of arrivals are visible after removing background noise and are attributed to ground shaking and other sources. **e** With the reduction from 17 to 3 sensors, primary infrasound from the regional earthquake is still clear but the local earthquake is not. Horizontal slowness is poorly resolved in general; unlike in (**b**), 3 sensors fail to detect the sustained horizontally-propagating infrasound. **f** The 90°−180° backazimuth range is still dominant here, but the −130° to −110° range is not visible. The features seen in (**b**) are less visible. **g** The features that were obvious in (**d**) are unresolvable.
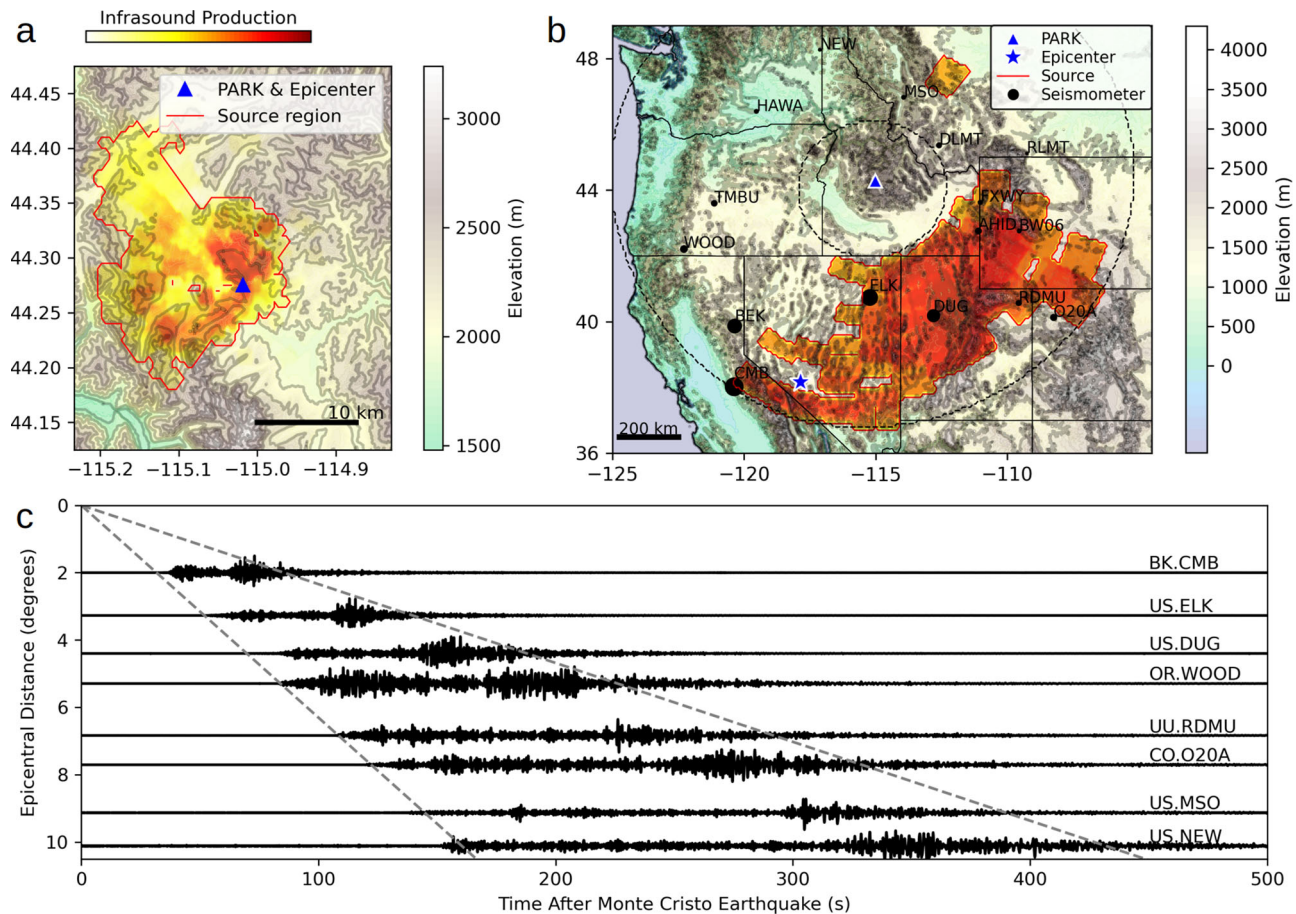
Fig. 6 Maps of inferred secondary infrasound sources. Because seismic waves propagate at a range of speeds, backprojections are less certain in range than in backazimuth; these maps err on the side of inclusion and therefore contain some points that do not radiate infrasound. a Central Idaho topographic map showing secondary infrasound identified in Fig. 4c for the local Stanley earthquake. The earthquake epicenter coincides with the PARK station so they are plotted as the same point. Secondary infrasound observed at PARK originates only out to nearby mountain crests, mainly to the southwest, west, and northwest. **b** Map of the western US showing secondary infrasound in the regional earthquake identified in Fig. 5d between 1000–2500 s after the earthquake (limits for these arrival times shown in dashed lines); local secondary infrasound (150–500 s) is excluded because of the very high uncertainty in seismic travel times. Secondary infrasound observed at PARK originates mostly from the southwest, south, and southeast, and the source region follows some mountain ranges. Amplitudes of select seismic stations are indicated by radii of black dots. **c** Select vertical seismograms from stations in (**b**), with dashed lines indicating average maximum/minimum seismic wave speeds.

analysis settings (e.g., number and size of analysis time windows). Researchers must assess their constraints and analytical needs when determining array size.

The cost of CLEAN is mainly configuration effort (running tests to optimize tuning parameters) and runtime. For example, analyzing the 85-s local earthquake on a modern laptop, traditional beamforming (the *array_processing()* function in obspy[28]), took 0.6 s (3 elements) and 16 s (20 elements), whereas the CLEAN analysis took 58 s (3 elements) and 3368 s (20 elements). Computational expense of advanced methods may restrict their use in real-time applications, or when large amounts of data must be processed.

Advancing acoustic earthquake monitoring. We distinguish primary infrasound from secondary infrasound solely on the basis of horizontal slowness. Co-deploying a seismic array alongside the infrasound array would reduce the potential for ambiguity between steeply incident secondary infrasound (with unusually low horizontal slowness) and primary infrasound coupled to slow seismic waves (with unusually high horizontal slowness). Additionally, it would elucidate air-to-ground and ground-to-air coupling (common in primary earthquake infrasound and for

high-amplitude infrasound waves[29]); this will reduce ambiguity in secondary infrasound backprojection by determining conditions that are favorable to seismic-infrasound conversions. We also note the value of monitoring ground shaking remotely on other planets, where installing large networks of seismic sensors is costly (e.g., Mars) or impossible (e.g., Venus). Ground-based or balloon-borne[30] large-N acoustic arrays may accelerate the study of tectonics and planetary structure elsewhere in the solar system.

Methods

Large-N arrays. The main advantage of large-N arrays in signal processing is its superior array response. An array's response is the wavenumber spectrum that would be estimated for a wave with vertical incidence (i.e., $k_x = k_y = 0$ rad/km) without any noise or other waves present. At a given frequency, the wavenumber spectrum an array would observe can be modeled as the 2-dimensional convolution of the array response with the wavefield's true wavenumber spectrum (typically a sparse set of impulses representing a small number of discrete waves). Therefore, compact array responses are desirable because they make it possible to distinguish distinct waves that have slightly different wavenumber vectors; compact array responses are

achieved by including sensors spaced very far apart in the array. Simultaneously, all array responses repeat periodically, and maximizing the distance between repeated main lobes of the array response can avoid problems of wavenumber aliasing in narrow-band signals. Avoiding aliasing is best accomplished by including sensors spaced close together in the array. These two objectives can be accomplished together by using a large number of sensors in an array, with large overall dimensions and small spacing between adjacent sensors. We used this approach to optimize the response of our infrasound array used in this paper, which consisted of 22 sensors (Fig. 3). Although sensor loss/dropout and restoration means this number varied throughout the campaign, the large number of sensors in the array meant that its array response was not strongly changed by the loss of a few sensors.

Because infrasound is non-dispersive at the frequencies studied in this paper, any real wave with broadband spectral content will have a range of wavenumber vectors but a fixed slowness vector ($\mathbf{s} = \frac{\mathbf{k}}{2\pi f}$); therefore, the frequency-to-wavenumber correspondence $\frac{2\pi f}{c} = |\mathbf{k}|$ can be shown on array response plots (Fig. 3b, c), and array analyses of real data are generally shown as slowness spectra rather than wavenumber spectra (Fig. 3d, e).

CLEAN beamforming. CLEAN is a method to iteratively account for and remove energy in the cross-spectrum until most of its coherent energy can be accounted for as plane waves with estimated power spectra and slowness vectors. Although CLEAN is not an acronym, its spelling in all-capitals is established in the literature[14],[15]. Figure 7 shows the steps in this process as equations and flow charts.

Before processing any recorded infrasound, the first step is to calculate array response information for the sensor array over a grid of slowness vectors (Fig. 7 steps A–C). This only needs to be calculated once, and can then be used for all data recorded from the sensor array as long as its geometry remains the same. In this paper, we assume that our data consists purely of plane waves and represent the array response using Bartlett weights, which is appropriate for our dataset because infrasound sources are distant enough to treat as plane waves and the signal changes rapidly enough that Bartlett beamforming is warranted. Datasets where signals can be treated as stationary for long periods of time can benefit from using Capon instead of Bartlett beamforming[15].

Having calculated Bartlett weights, the next steps are pre-processing the data (in this paper, band-pass filtering to maximize signal-to-noise ratio) and calculating the cross-spectrum of a recording. These steps must be done once per time window that is beamformed, before the CLEAN iterations begin (Fig. 7, steps 1–2). The cross-spectrum is a third-order tensor that, for each frequency, contains information on the power spectral density at all sensors as well as how power and phase differ among sensors. CLEAN beamforming treats the cross-spectrum as a linear combination of the array response (in the form of the weights in step C) and the distribution of infrasound power over wavenumber vectors (the CLEAN spectrum, which is to be calculated) and estimates wavefield composition over a series of iterations. In each iteration, the best-fit slowness vector is identified via grid search, scaled by phi, and added to the CLEAN spectrum (Fig. 7, steps 3-5). The result of these steps in the first iteration is identical to the result that would be found by single-source beamforming (e.g., the array_processing() function in obspy). Then, the corresponding component of the cross-spectrum is modeled using information in the array response (Fig. 7, step 6, using the beamforming weights calculated in step C). Finally, that component is scaled by the coefficient phi (in this paper, phi = 0.1) and subtracted from the cross-spectrum (Fig. 7, step 7). This process is repeated, with power subtracted from the cross spectrum and added to the CLEAN spectrum in each iteration, until some stopping criterion is satisfied, indicating that most of the coherent energy has been accounted for in previous iterations (Fig. 7, step 8). In this paper, we use a stopping criterion using the F-statistic as an indicator of the presence of coherent signals[15]. The GitHub repository https://github.com/ajakef/Earthquake_Infrasound_Paper[31] contains our Python package cleanbf and scripts to download data and to use cleanbf to produce Figs. 3–6 and supplementary Figs. 1–2.

Details of our beamforming application. Before beginning array processing, we applied a band-pass filter to improve the earthquake's signal-to-noise ratio (Fig. 7 step 1). Frequency limits of the filters were selected after examining spectra of earthquake signals and ambient noise. When processing data from the local event, we used frequency limits of 2–25 Hz and beamformed cross-spectra calculated for 2 s windows with 50% overlap. For the regional event, we used frequency limits of 1–20 Hz; due to the much longer duration, slower rates of change, greater computational expense, and diminished plot detail compared to the local event, we beamformed cross-spectra calculated for 4 s windows with 0% overlap. Additionally, because infrasound at the frequencies studied is non-dispersive, we imposed the additional physics-based constraint that each wave component has one slowness vector (s_x, s_y) that applies to all frequencies; this would not be appropriate in different settings when waves are strongly dispersive. In each iteration we searched a slowness grid where s_x and s_y ranged from $+/-4 \, \text{s} \, \text{km}^{-1}$ with spacing $0.01 \, \text{s} \, \text{km}^{-1}$. Because infrasound propagates with a slowness up to $3 \, \text{s} \, \text{km}^{-1}$ in typical earth-surface conditions, this grid is large enough to capture all realistic infrasound.

Backprojection. In this paper, "backprojection" refers to the process of identifying possible source regions for an individual beamforming result (i.e., the infrasound power associated with a specific backazimuth-time combination). As in reference[4], back-projections are accomplished by identifying all points on the map with the correct backazimuth and whose total estimated traveltime (from hypocenter to map point to infrasound array) matches the observed arrival time. Seismic waves travel at multiple speeds, so for each backprojection we estimate a minimum and maximum seismic wave speed. We do the same for infra-sound waves, although they vary in speed less than seismic waves do. Because wave speeds are treated as a range instead of a single value, any individual beamforming result will be backprojected to a line segment on the map oriented on the observed backazimuth from the infrasound station, where the nearest point on the segment corresponds to slow wave speeds and the farthest point corresponds to fast wave speeds.

For the local event, we followed reference[4] in using a seismic wave speed range of 1400–5600 m/s. We set the infrasound speed range to 328–330 m/s, corresponding to the surface-level temperature of -3 °C measured by our infrasound loggers (reasonable for the low-altitude direct-wave propagation for this local secondary infrasound in low-wind early-morning conditions).

For the regional event, we empirically determined seismic wave speeds by first calculating cumulative energy at several seismic stations over the western US during this event, then defined the maximum and minimum wave speeds corresponding to the times when cumulative energy reached 1% and 90% of total signal energy, resulting in a range of 2600–7000 m/s. We set the infrasound wave speed range to 320–340 m/s (corresponding to -20 to 14 °C ignoring wind), the wide range reflecting the varying atmospheric conditions encountered between the surface and the upper stratosphere, over six degrees of latitude between epicenter and infrasound station.
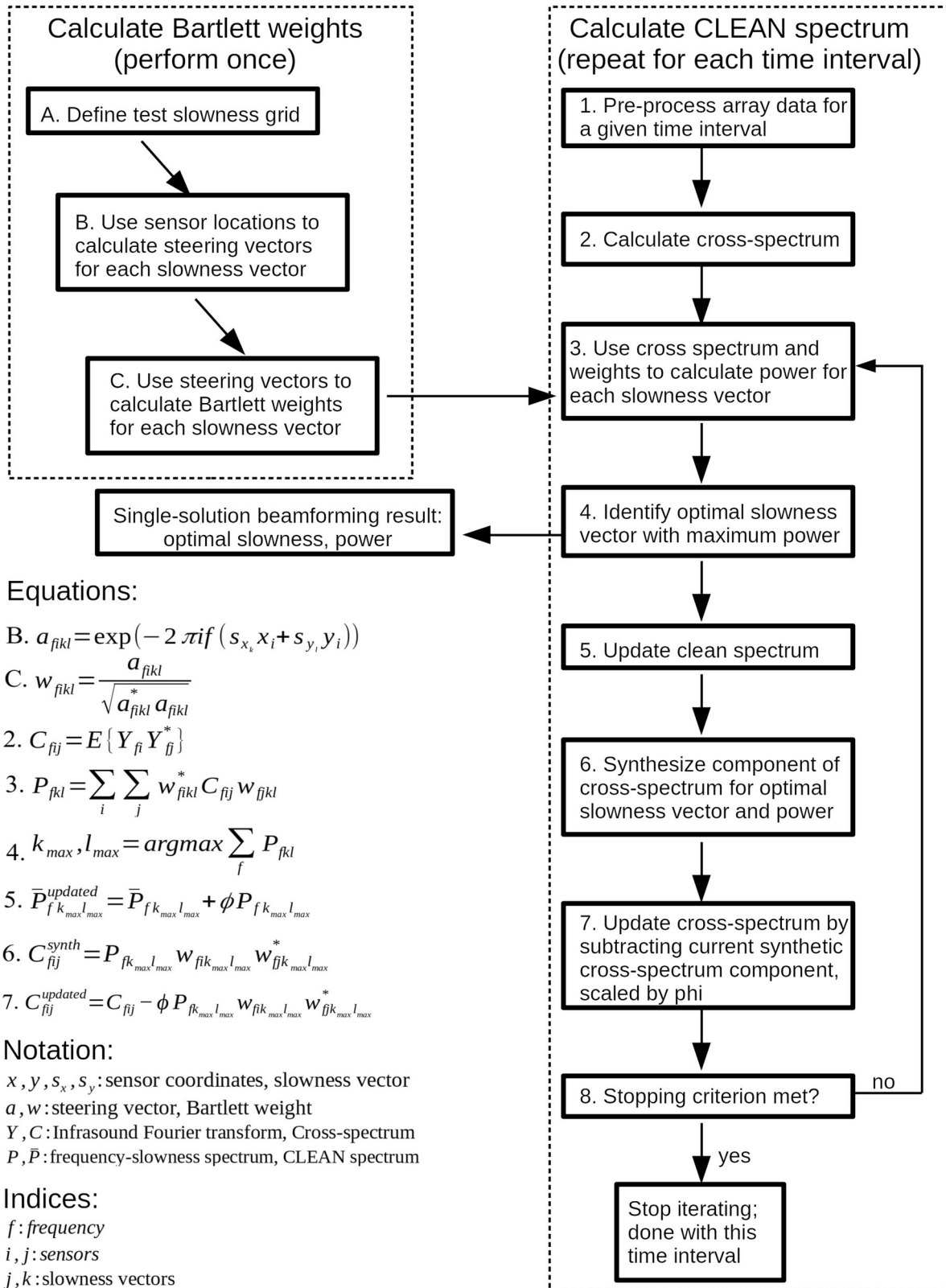
Calculate Bartlett weights (perform once)

A. Define test slowness grid

B. Use sensor locations to calculate steering vectors for each slowness vector

C. Use steering vectors to calculate Bartlett weights for each slowness vector

Single-solution beamforming result: optimal slowness, power

Calculate CLEAN spectrum (repeat for each time interval)

1. Pre-process array data for a given time interval

2. Calculate cross-spectrum

3. Use cross spectrum and weights to calculate power for each slowness vector

4. Identify optimal slowness vector with maximum power

5. Update clean spectrum

6. Synthesize component of cross-spectrum for optimal slowness vector and power

7. Update cross-spectrum by subtracting current synthetic cross-spectrum component, scaled by phi

8. Stopping criterion met? no

yes

Stop iterating; done with this time interval

Equations:

B. $a_{fikl} = \exp\left(-2\pi i f\left(s_{x_k} x_i + s_{y_l} y_i\right)\right)$

C. $w_{fikl} = \dfrac{a_{fikl}}{\sqrt{a_{fikl}^* a_{fikl}}}$

2. $C_{fij} = E\left\{Y_{fi} Y_{fj}^*\right\}$

3. $P_{fkl} = \sum_i \sum_j w_{fikl}^* C_{fij} w_{fjkl}$

4. $k_{max}, l_{max} = argmax \sum_f P_{fkl}$

5. $\overline{P}_{f k_{max} l_{max}}^{updated} = \overline{P}_{f k_{max} l_{max}} + \phi P_{f k_{max} l_{max}}$

6. $C_{fij}^{synth} = P_{f k_{max} l_{max}} w_{fi k_{max} l_{max}} w_{fj k_{max} l_{max}}^*$

7. $C_{fij}^{updated} = C_{fij} - \phi P_{f k_{max} l_{max}} w_{fi k_{max} l_{max}} w_{fj k_{max} l_{max}}^*$

Notation:

x, y, s_x, s_y : sensor coordinates, slowness vector

a, w : steering vector, Bartlett weight

Y, C : Infrasound Fourier transform, Cross-spectrum

$P, \overline{P}$: frequency-slowness spectrum, CLEAN spectrum

Indices:

f : frequency

i, j : sensors

j, k : slowness vectors

Fig. 7 Flow chart demonstrating the CLEAN method. Steps A–C depend only on array geometry, not on recorded data, and are run only once for each array. Steps 1–8 show the analysis of a specific data interval and include the CLEAN iterative loop; all of steps 1–8 must be performed for every new data interval studied.

The backprojection results (Fig. 6a, b) divide the map into a grid of cells, each having a total energy including the energy from all beamforming results that can be assigned to the cell, scaled by the distance. Because of the uncertainty in distance due to the range of possible seismic wave speeds, any individual beamforming result can be assigned to several map cells; consequently, the maps show more seismic energy than was actually detected at PARK.

ARTICLE

Data availability

Code availability

the repository https://github.com/ajakef/Earthquake_Infrasound_Paper[31] contains Python code used in this paper, including the version of "cleanbf" (a library that implements CLEAN beamforming), as well as code to reproduce Figs. 4–6 and Supplementary Figs. 1–2.

References

1. Le Pichon, A., Guilbert, J., Vega, A., Garcés, M. & Brachet, N. Ground-coupled air waves and diffracted infrasound from the Arequipa earthquake of June 23, 2001. *Geoph. Res. Lett.* **29**, 33-1–33-4 (2002).
2. Arrowsmith, S. J., Johnson, J. B., Drob, D. P. & Hedlin, M. A. The seismoacoustic wavefield: a new paradigm in studying geophysical phenomena. *Rev. Geophys.* **48**, 4003 (2010).
3. Averbuch, G., Assink, J. D. & Evers, L. G. Long-range atmospheric infrasound propagation from subsurface sources. *J. Acoust. Soc. Am.* **147**, 1264–1274 (2020).
4. Johnson, J. B., Mikesell, T. D., Anderson, J. F. & Liberty, L. M. Mapping the sources of proximal earthquake infrasound. *Geoph. Res. Lett.* **47**, e2020GL091421 (2020).
5. Averbuch, G., Ronac-Giannone, M., Arrowsmith, S. & Anderson, J. Evidence for short temporal atmospheric variations observed by infrasonic signals: 1. The Troposphere. *Earth Space Sci.* **9**, e2021EA002036 (2022).
6. Pilger, C. et al. Infrasound and seismoacoustic signatures of the 28 September 2018 Sulawesi super-shear earthquake. *Nat. Hazards Earth Syst. Sci.* **19**, 2811–2825 (2019).
7. Khodr, C., Green, D. N. & Azarpeyvand, M. Three-dimensional topographic effects on infrasound propagation across Ascension Island. *Geoph. J. Int.* **231**, 1558–1572 (2022).
8. Shani-Kadmiel, S., Assink, J. D., Smets, P. S. & Evers, L. G. Seismoacoustic coupled signals from earthquakes in central Italy: epicentral and secondary sources of infrasound. *Geoph. Res. Lett.* **45**, 427–435 (2018).
9. Yokoo, A., Suzuki, Y. J. & Iguchi, M. Dual infrasound sources from a Vulcanian eruption of Sakurajima volcano inferred from cross-array observation. *Seismol. Res. Lett.* **85**, 1212–1222 (2014).
10. Assink, J. D., Averbuch, G., Smets, P. S. M. & Evers, L. G. On the infrasound detected from the 2013 and 2016 DPRK's underground nuclear tests. *Geoph. Res. Lett.* **43**, 3526-3533 (2016).
11. Landès, M., Ceranna, L., Le Pichon, A. & Matoza, R. S. Localization of microbarom sources using the IMS infrasound network. *J. Geophys. Res. Atmos.* **117**, D6 (2012).
12. Bosa, A. et al. Tracking secondary lahar flow paths and characterizing pulses and surges using infrasound array networks at Volcán de Fuego, Guatemala. *Volcanica* **4**, 239–256 (2021).
13. Johnson, J. B., Anderson, J. F., Marshall, H. P., Havens, S. & Watson, L. M. Snow avalanche detection and source constraints made using a networked array of infrasound sensors. *J. Geophys. Res. Earth Surf.* **126**, e2020JF005741 (2021).
14. Gal, M. et al. Deconvolution enhanced direction of arrival estimation using one-and three-component seismic arrays applied to ocean induced microseisms. *Geoph. J. Int.* **206**, 345–359 (2016).
15. den Ouden, O. F. et al. CLEAN beamforming for the enhanced detection of multiple infrasonic sources. *Geoph. J. Int.* **221**, 305–317 (2020).
16. Hansen, S. M. & Schmandt, B. Automated detection and location of microseismicity at Mount St. Helens with a large-N geophone array. *Geoph. Res. Lett.* **42**, 7390–7397 (2015).
17. Inbal, A., Ampuero, J. P. & Clayton, R. W. Localized seismic deformation in the upper mantle revealed by dense seismic arrays. *Science* **354**, 88–92 (2016).
18. Green, D. N. The spatial coherence structure of infrasonic waves: analysis of data from International Monitoring System arrays. *Geoph. J. Int.* **201**, 377–389 (2015).
19. Schaeffer, A. J. & Lebedev, S. Imaging the North American continent using waveform inversion of global and USArray data. *Earth Plan. Sci. Lett.* **402**, 26–41 (2014).
20. Anderson, J. F., Johnson, J. B., Bowman, D. C. & Ronan, T. J. The Gem infrasound logger and custom-built instrumentation. *Seismol. Res. Lett* **89**, 153–164 (2018). 1.
21. Liberty, L. M., Lifton, Z. M. & Mikesell, T. D. The 31 March 2020 Mw 6.5 Stanley, Idaho, earthquake: seismotectonics and preliminary aftershock analysis. *Seismol. Res. Lett.* **92** **2A**, 663–678 (2021).
22. USGS. *M 6.5—Stanley, Idaho.* https://earthquake.usgs.gov/earthquakes/eventpage/us70008jr5/executive (2022).
23. Wilbur, S. F. *Machine-Learning Reveals Aftershock Locations for Three IDAHO Earthquake Sequences* (Thesis, Boise State University, 2022).
24. USGS. *USGS Magnitude 2.5+ Earthquakes, Past Day.* https://earthquake.usgs.gov/earthquakes/eventpage/us70008vuc/executive (2023).
25. Ruhl, C. J. et al. Complex fault geometry of the 2020 Mww 6.5 Monte Cristo Range, Nevada, earthquake sequence. *Seismol. Res. Lett* **92**, 1876–1890 (2021).
26. USGS, https://earthquake.usgs.gov/earthquakes/eventpage/nn00725272/executive (2023).
27. Scamfer, L. T. & Anderson, J. F. Exploring background noise with a large-N infrasound array: waterfalls, thunderstorms, and earthquakes. *ESS Open Archive.* https://doi.org/10.22541/essoar.169111786.61306257/v1 (2023).
28. Krischer, L. et al. ObsPy: a bridge for seismology into the scientific Python ecosystem. *Comp. Sci. Disc.* **8**, 014003 (2015).
29. McKee, K., Fee, D., Haney, M., Matoza, R. S. & Lyons, J. Infrasound signal detection and back azimuth estimation using ground-coupled airwaves on a seismo-acoustic sensor pair. *J. Geoph. Res. Solid Earth* **123**, 6826–6844 (2018).
30. Brissaud, Q. et al. The first detection of an earthquake from a balloon using its acoustic signature. *Geoph. Res. Lett.* **48**, e2021GL093013 (2021).
31. GitHub. GitHub repository earthquake infrasound paper. *Zenodo* https://doi.org/10.5281/zenodo.8388128 (2005).
32. Berti, C., Liberty, L., Stachnik, J. & Mikesell, D. *Aftershock Deployment for Stanley, ID Earthquake 2020* [Data set]. *Int. FDSN* https://doi.org/10.7914/SN/XP_2020 (2020).

Acknowledgements

Author contributions

T.D.M., J.B.J., L.M.L., and J.F.A. acquired funding. J.F.A. and J.B.J. conceptualized the study. J.F.A. collected data, researched beamforming methods, wrote the "cleanbf" software and data analysis code, and wrote the manuscript. J.F.A., J.B.J., T.D.M., and L.M.L. contributed to revising the draft. J.F.A. and T.D.M. curated the data.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-023-01058-z.

Correspondence and requests for materials should be addressed to Jacob F. Anderson.

Peer review information *Communications Earth & Environment* thanks Siddharth Krishnamoorthy, Jordan Wilson Bishop and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Luca Dal Zilio and Joe Aslin. A peer review file is available.

Reprints and permission information is available at http://www.nature.com/reprints

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