https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

A decade-scale catalog of coherent atmospheric infrasound transients from a three-element array at Jang Bogo Station, Terra Nova Bay, East Antarctica (2017–2026)

Yongcheol Park[1], Jinhoon Jung[1], Won Sang Lee[1]
[1] Korea Polar Research Institute (KOPRI), Incheon, Republic of Korea
Correspondence: Yongcheol Park (ypark@kopri.re.kr)

Abstract. Long, continuous infrasound records from the Antarctic coast are rare, but they provide passive, all-weather observations of cryospheric, oceanic, and atmospheric processes. We present a decade-scale catalog of coherent atmospheric infrasound transients recorded by the three-element KIS (KOPRI InfraSound) array at Jang Bogo Station, Terra Nova Bay, East Antarctica. Across 2,507 array-days spanning 2017–2026, we identified 23,768 coherent acoustic transients using an STA/LTA coincidence detector followed by inter-element coherence and slowness analysis. Of these, 22,711 were recorded in the 100 Hz HDF stream and 1,057 in the 20 Hz BDF stream, which bridges a 2021–2022 configuration gap. Each catalog entry includes detection time, duration, inter-element coherence, wavefront back-azimuth, apparent velocity, signal-to-noise ratio, nominal-equivalent pressure, directional sector, and recording stream. The apparent-velocity distribution is sharply concentrated around 333 m s^{-1}, consistent with acoustic propagation. The median event duration is 6.5 s, and the catalog contains both impulsive and emergent waveform classes. The catalog and its per-event waveform windows are openly archived under a CC BY 4.0 license at Zenodo (https://doi.org/10.5281/zenodo.21882394) (Park et al., 2026c), and the catalog is mirrored at the Korea Polar Data Center (https://doi.org/10.22663/KOPRI-KPDC-00003159) (Park et al., 2026e). Worked examples show that the catalog captures the coherent infrasound of the 2022 Hunga Tonga eruption and a transient from the direction of David Glacier that follows a local glacial earthquake. The dataset provides a reusable resource for cryo-acoustic and infrasound-monitoring studies on the Antarctic margin. Physical interpretation of the directional source families is presented in a companion analysis paper.

1 Introduction

The polar cryosphere radiates elastic and acoustic energy prolifically, and a substantial fraction of it enters the atmosphere. Fracturing sea ice, calving and disintegrating icebergs, and the crevassing and stick–slip motion of glaciers all radiate infrasound as well as seismic and hydroacoustic energy (Podolskiy and Walter, 2016; Aster and Winberry, 2017; Richardson et al., 2010). Atmospheric infrasound propagates efficiently over open water and ice, is sensitive to sources at and above the surface, and operates through polar night and cloud. A small-aperture array that resolves back-azimuth and cross-element coherence is therefore a natural monitor of Antarctic coastal dynamics (Cansi, 1995; Rost and Thomas, 2002).

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Infrasound-array monitoring in Antarctica has been developed chiefly at Syowa Station and in the Lützow-Holm Bay region (Yamamoto et al., 2013; Murayama et al., 2015; Murayama et al., 2017; Kanao et al., 2024). Global monitoring is provided by the CTBTO International Monitoring System (Hupe et al., 2022). Long, continuous, consistently processed transient catalogs from the Antarctic coast nonetheless remain scarce. Here we describe such a catalog, comprising a decade of coherent infrasound transients from the KIS array at Jang Bogo Station, Terra Nova Bay, in the western Ross Sea. This paper documents the array configuration, data acquisition, detection and parameter-estimation procedures, catalog contents, and quality-control framework, so that the dataset can be reused independently. The physical attribution of the transients to source families (marine, glacial, sea-ice, and wind sources) is addressed in a separate analysis by the authors, whose seismic–infrasound association table and analysis code are openly archived (Park et al., 2026a).

2 The KIS infrasound array

The KIS (KOPRI InfraSound) array consists of three Chaparral Physics Model 25 differential-pressure infrasound sensors installed at Jang Bogo Station on the northern shore of Terra Nova Bay. The sensors are digitized by the station's Nanometrics Centaur data-acquisition system. The three elements (KIS1–KIS3) form a triangular array with an aperture of approximately 150 m (Fig. 1b, c). Element coordinates are given in Table 1. Jang Bogo Station faces the open water and seasonal sea ice of Terra Nova Bay and the western Ross Sea to the east (Fig. 1a). Mt Melbourne lies 33 km to the north-northeast; the Nansen Ice Shelf and, at 230° and ~125 km, the grounding zone of David Glacier lie to the southwest; and the Drygalski Ice Tongue, fed by David Glacier, lies to the south. Each element is a Chaparral Physics Model 25 sensor in its high-gain configuration (2.0 V Pa^{-1}), fitted with a porous-hose rosette that reduces wind noise, and timing is provided by the GPS-disciplined Centaur digitizer. The rosettes were mounted directly on the sensors on 31 October 2016; in the earlier configuration the porous hose was connected to the sensor enclosure and suppressed wind noise ineffectively, which is why the catalog begins after that upgrade (Sect. 3).

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

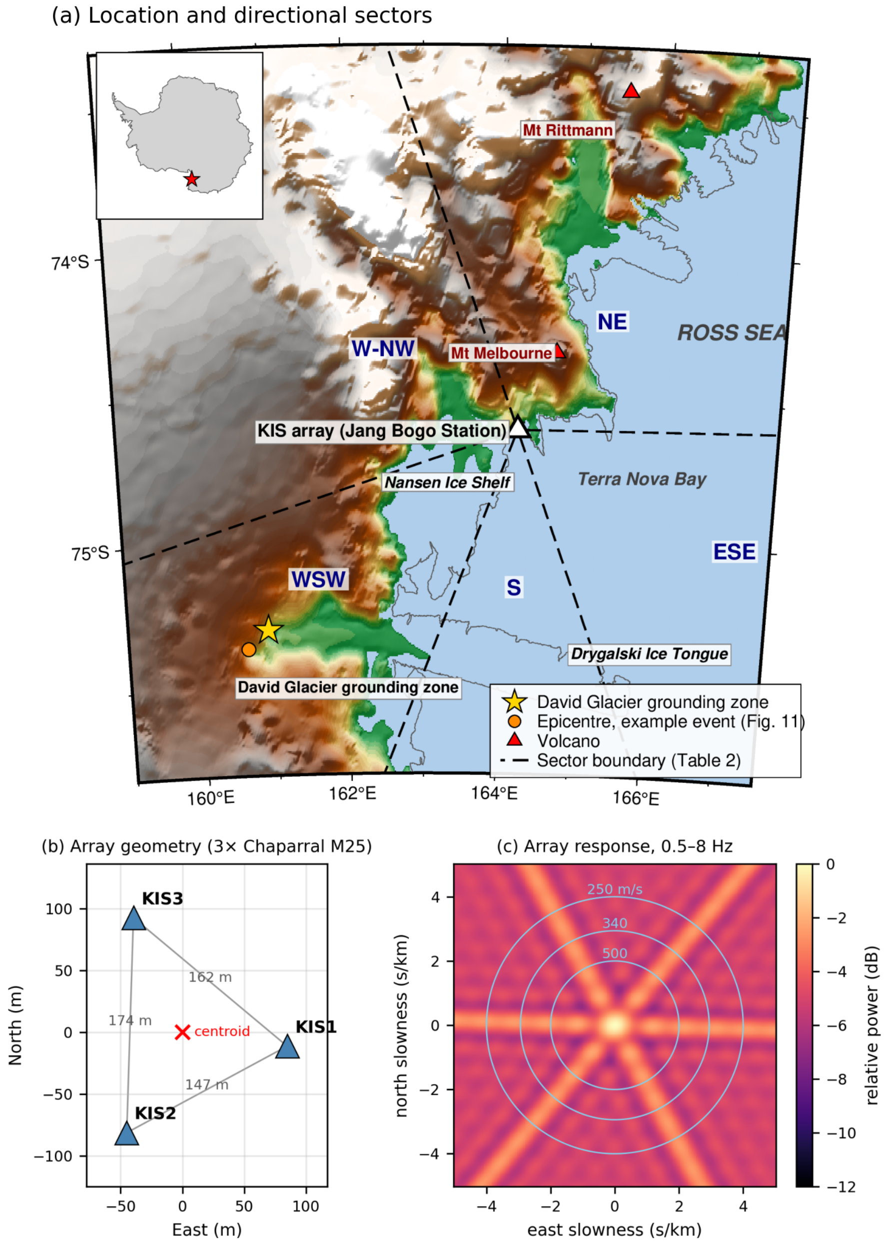

Figure 1. (a) Location of the KIS array at Jang Bogo Station, Terra Nova Bay, over Bamber DEM topography (Bamber et al., 2009) with the MODIS Mosaic of Antarctica (MOA) coastline (Scambos et al., 2007), showing the five directional-sector boundaries of Table 2 (dashed lines), the David Glacier grounding zone (star; 75.3° S, 160.9° E), and the epicentre of the local earthquake associated with the example event of Fig. 11 (circle). Inset: location in Antarctica. (b) KIS three-element array geometry (KIS1/2/3; side lengths 147–174 m). (c) Theoretical array response (transfer function) in slowness space, band-averaged over the 0.5–8 Hz detection band. The band-averaged response is dominated by a central lobe across the apparent-velocity range used in this study (250–500 m s^{-1}), supporting stable slowness estimation for the broadband transients analyzed here.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Table 1. Array element coordinates

Element	Latitude (°)	Longitude (°)	Elevation (m)
KIS1	−74.618203	164.221786	106
KIS2	−74.618836	164.217392	106
KIS3	−74.617264	164.217575	106
KIS array center	−74.61795	164.21925	—

The instrument response is flat across the 0.5–8 Hz analysis band, with a sensitivity of 8.0×10^5 counts Pa^{-1}. Pressure in pascals is therefore recovered by dividing the recorded counts by this factor. This conversion uses the manufacturer's nominal sensitivity. An independent cross-check against a reference barometer has not yet been performed, so absolute-pressure values should be treated as calibrated to the nominal sensitivity (Sect. 6). Below the analysis band the response is not flat: the sensor has a triple pole at 0.051 Hz, so the amplitude response relative to 1 Hz is 0.71 at 0.1 Hz, 0.34 at 0.05 Hz, and 0.007 at 0.01 Hz; quantitative amplitudes below about 0.01 Hz are therefore not reliable. The full instrument response (poles, zeros, and stage gains) is provided as FDSN StationXML with the archived Hunga Tonga excerpt (Sect. 7.2; Code and data availability).

3 Data acquisition and coverage

The array has operated since 2016. It recorded at 100 Hz (channel HDF) until 14 November 2021 and again from 14 December 2022 (no data are available for 19 November–13 December 2022); between 16 November 2021 and 18 November 2022 it recorded at 20 Hz (channel BDF). Both streams fully cover the 0.5–8 Hz detection band. Data before 31 October 2016 belong to an earlier, lower-quality configuration and are excluded. The catalog spans the consistently processed 2017–2026 record, comprising 2,507 array-days. The record used here ends on 8 January 2026: 2026 contributes eight operational array-days, and the last cataloged transient is on 7 January 2026. The HDF and BDF streams show broadly overlapping parameter distributions, with modest stream-dependent differences (Appendix A; Fig. A2). Of these, 75 array-days (mostly 2022) come from the BDF stream and fill what would otherwise be a gap. Fig. 2 summarizes the coverage, including data gaps, the recording-stream change, and the year-to-year variation in operational days. Annual event counts (Sect. 5) reflect this changing availability as well as source activity and should not be read directly as source-activity trends.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

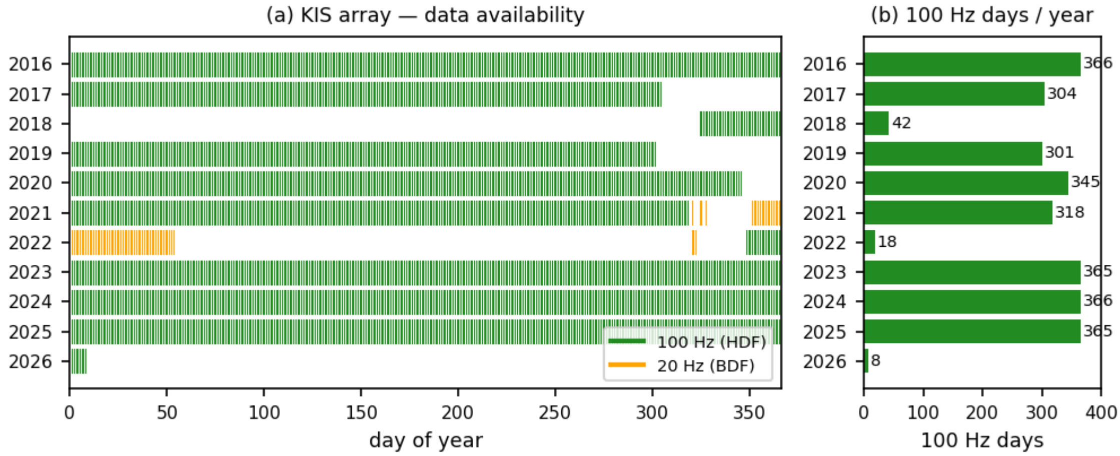

Figure 2. Data coverage: number of array-days per year, 2016–2026 (100 Hz HDF and 20 Hz BDF streams), showing the growth, gaps, and recording-stream change of the KIS record that underlie the 2,507-array-day catalog.

4 Catalog construction

Continuous KIS records were band-pass filtered to 0.5–8 Hz and scanned for coherent transients with a recursive short-term/long-term-average (STA/LTA) coincidence trigger (2 s / 80 s windows). The trigger requires simultaneous triggering on at least two of the three elements (Allen, 1978). For every trigger we computed the inter-element coherence as the mean of the normalized cross-correlation peaks of the KIS1–KIS2 and KIS1–KIS3 element pairs, and estimated the wavefront back-azimuth and apparent velocity from the corresponding inter-element travel-time delays (the lag at each cross-correlation peak). With element position vectors r_1, r_2, r_3, the slowness vector s solves G s = Δt for G = $[r_2 - r_1; r_3 - r_1]$. The back-azimuth is $\mathrm{atan2}(s_x, sy) + 180°$ and the apparent velocity is $1/|s|$ (Rost and Thomas, 2002). Only triggers with coherence ≥ 0.6 and apparent velocity between 250 and 500 m s^{-1} were retained as coherent acoustic transients (Fig. 3). This gate removes incoherent wind noise and strongly super-acoustic trace-velocity artifacts. Per-event back-azimuth and apparent-velocity uncertainties are provided in the catalog. They are computed from the array geometry and the coherence-dependent time-delay estimation error following the array-processing method of Szuberla and Olson (2004), and represent the formal estimation uncertainty; un-modeled propagation effects such as cross-wind advection and wavefront curvature are not included. The signal-to-noise ratio is the peak Hilbert-envelope amplitude in the analysis window divided by the day-median envelope amplitude; the duration is the STA/LTA trigger duration; and the peak acoustic pressure is the maximum absolute count amplitude in the window divided by the nominal sensitivity. The complete detection and parameter-estimation settings are listed in Table A1. All waveform processing used ObsPy (Beyreuther et al., 2010; Krischer et al., 2015). The 20 Hz BDF records were resampled to 100 Hz before detection. Because their native timing resolution is coarser (50 ms versus 10 ms), the back-azimuth and apparent velocity of BDF-derived events carry larger uncertainty. Every event is therefore flagged by its recording stream (Sect. 6).

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

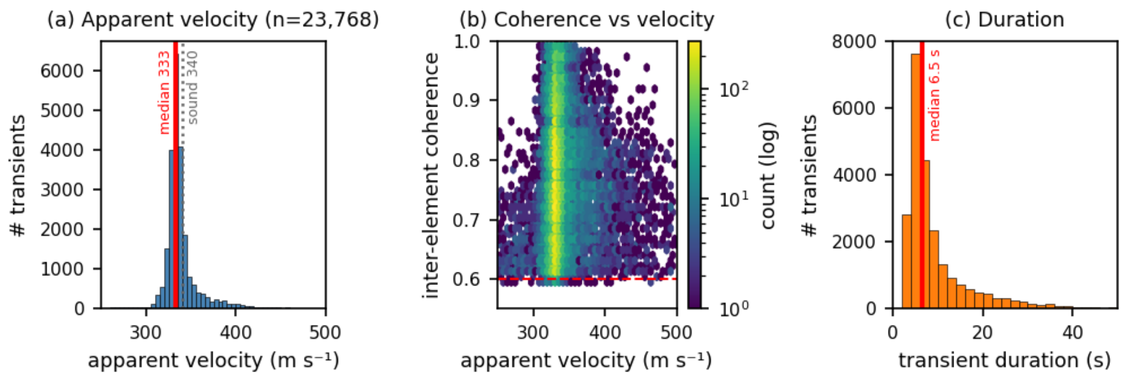

Figure 3. Detection and acoustic character of the catalog (2017–2026, n = 23,768). (a) Apparent-velocity distribution. A sharp peak at the median apparent velocity of 333 m s^{-1}, well within the acoustic-velocity range, is consistent with propagating infrasound. (b) Inter-element coherence versus apparent velocity (two-dimensional density, logarithmic color scale). The coherence ≥ 0.6 and 250–500 m s^{-1} gates isolate a tight, high-coherence population centered on the acoustic velocity. (c) Transient duration, a right-skewed distribution with a median of 6.5 s.

5 Catalog content and classification

The catalog contains 23,768 coherent acoustic transients. Their median apparent velocity of 333 m s^{-1} is well within the acoustic-velocity range (Fig. 3a). The coherence ≥ 0.6 and 250–500 m s^{-1} gates isolate a tight, high-coherence population centered on the acoustic velocity (Fig. 3b). The median event duration is 6.5 s (Fig. 3c), and waveform-shape clustering distinguishes impulsive and emergent waveform classes. Nominal-equivalent peak acoustic pressures, calculated using the manufacturer's sensitivity, span roughly 0.001–0.5 Pa, with a median near 0.017 Pa. Representative waveforms and spectrograms of each directional sector are shown in Fig. 4. The monthly occurrence and per-year detection rate of the cataloged transients are shown in Fig. 5.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

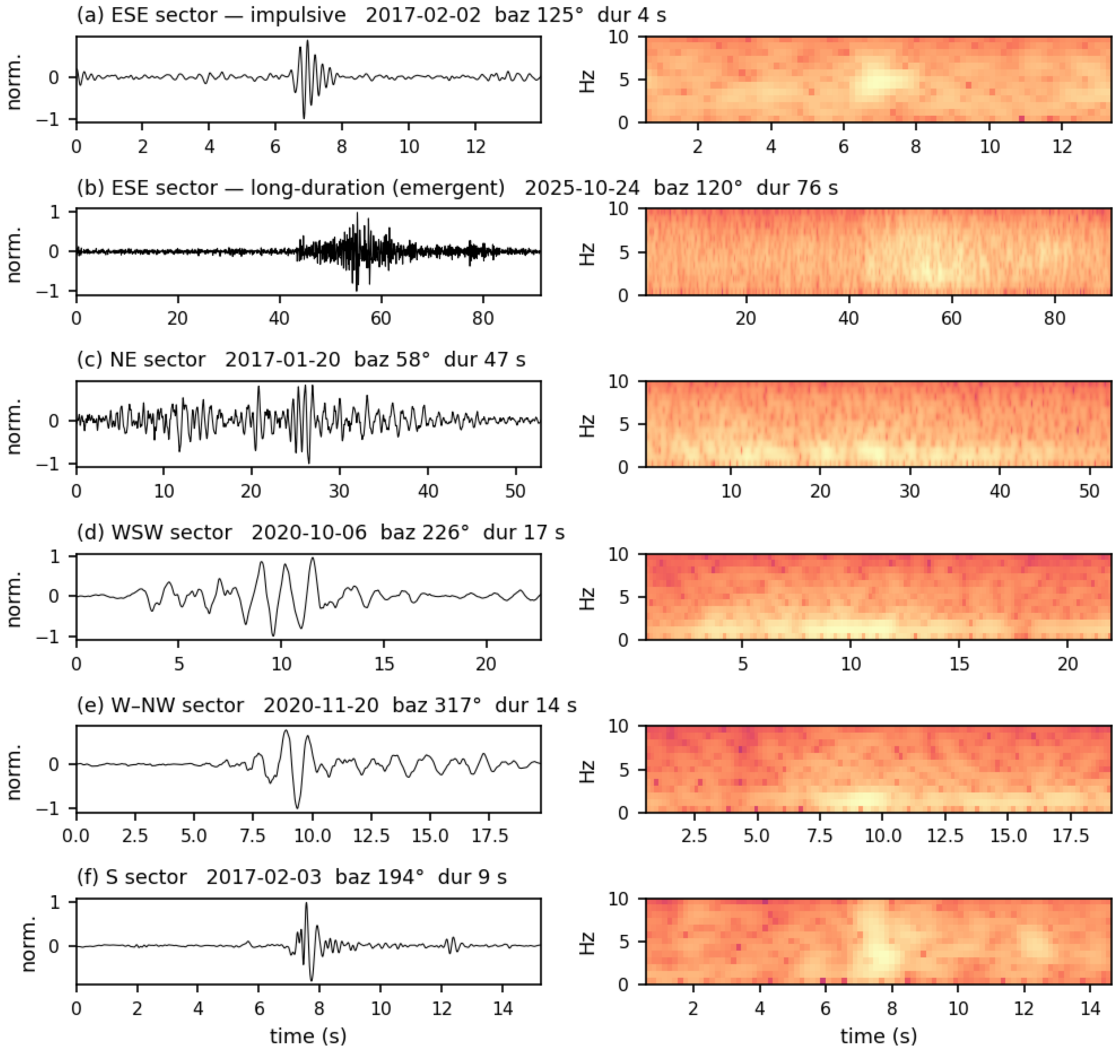

Figure 4. Representative transient waveforms by directional sector (KIS1). Waveforms are band-pass filtered 0.5–8 Hz (left) and spectrograms span 0–10 Hz with color showing relative power (right). Panel titles give the cataloged STA/LTA trigger duration; the plotted windows include additional pre- and post-event signal for context. The ESE sector shows both short impulsive signals and a smaller population of long-duration emergent signals (up to ~76 s). The remaining rows are examples from the NE, WSW, W–NW and S sectors. Source attribution for the sectors is outside the scope of this data description.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

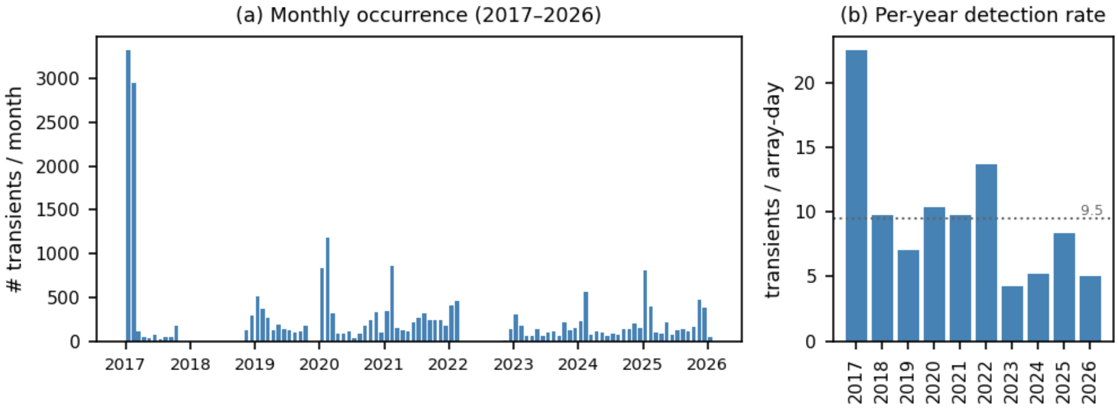

Figure 5. Full-period occurrence of the 23,768 coherent acoustic transients (2017–2026). (a) Monthly occurrence. (b) Per-year detection rate (transients per array-day; dotted line = overall 9.5). Annual rates partly reflect changing data availability and station configuration (Sect. 3) and should not be read as a direct source-activity trend; the pronounced 2017 maximum is dominated by an exceptional January–February east-southeast episode, whose physical attribution is not addressed here. Per-year rates are normalized by total operational array-days (HDF + BDF); 2026 (8 array-days) and 2018 (42) rest on far fewer days than the other years and are correspondingly uncertain.

Each event is assigned to one of five directional sectors, defined by objective back-azimuth bins that tile the full compass. The sector composition is summarized in Table 2. These are directional bins, not source interpretations. To indicate what users may expect in each sector, Table 2 also lists the geographic setting along each bearing (Fig. 1a) and, where independently documented and directly relevant to data use, brief source context; the entries do not assign individual transients to sources. The physical attribution of the sectors is outside the scope of this data description; the association table and analysis code of a separate study by the authors are openly archived (Park et al., 2026a).

Table 2. Directional-sector composition of the catalog

Sector	Back-azimuth range	Count	Share	Geographic setting and relevant source context (Fig. 1a)
ESE	90–160°	11,586	48.7 %	Terra Nova Bay and western Ross Sea (open water, polynya, seasonal sea ice); marine seismic surveys (Sect. 6)
NE	340–90° (through N)	6,014	25.3 %	Mt Melbourne (26°, 33 km); western Ross Sea
S	160–200°	2,541	10.7 %	Terra Nova Bay fast ice and Drygalski Ice Tongue; Ross Island (Mt Erebus, 168°, 335 km) in the far field
W-NW	250–340°	1,824	7.7 %	Transantarctic Mountains and the margin of the East Antarctic plateau; the prevailing direction of katabatic wind
WSW	200–250°	1,803	7.6 %	Nansen Ice Shelf; David Glacier grounding zone (230°, ~125 km)

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

All	23,768	100 %

To provide a classification independent of geographic direction, the HDF-stream waveforms were additionally described by 13 spectral and envelope-shape features (Table A2). Shape features were successfully computed for 22,709 of the 22,711 HDF-stream events; these were standardized and grouped by Ward hierarchical clustering. The optimal number of clusters, chosen by silhouette score, is two (Fig. 6). The classes are an impulsive high-frequency class and an emergent low-frequency class (Fig. 7). The resulting shape class is assigned to each classifiable HDF-stream event as a supporting, direction-independent characterization, and is reproducible from the archived waveform windows following the feature definitions in Table A2 (Park et al., 2026c).

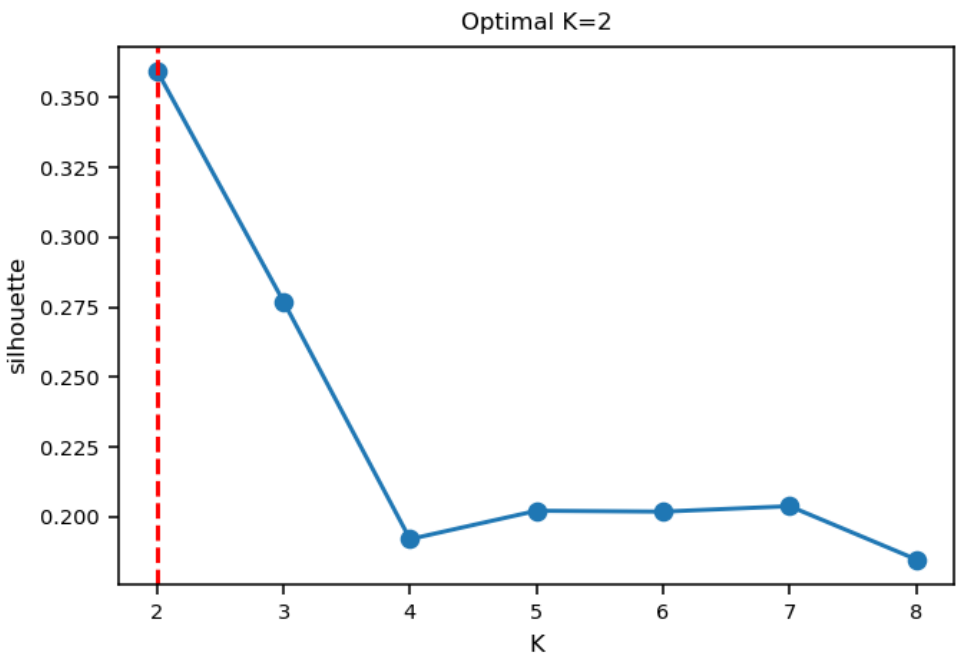

Figure 6. Silhouette score of the Ward waveform-shape clustering as a function of the number of clusters K. The score is maximized at K = 2, the value adopted for Fig. 7.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

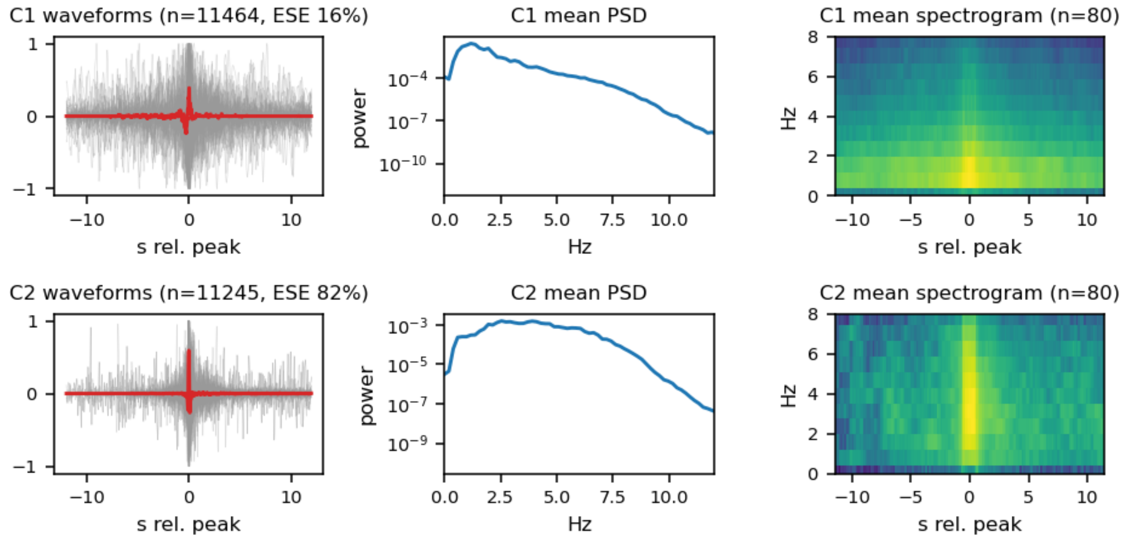

Figure 7. Waveform-shape hierarchical clustering (13 spectral and envelope features, Ward). The optimal K = 2 separates an impulsive high-frequency class from an emergent low-frequency class. The clustering covers the 22,709 HDF-stream (100 Hz) events with computable shape features (two of the 22,711 HDF events are unclassified); the 1,057 BDF-stream (20 Hz) events are not shape-classified owing to their coarser sampling. In each left-hand waveform panel the grey traces are the individual aligned events and the red curve is the cluster median.

The catalog includes the event identifier, UTC detection time (with year and day of year), duration, inter-element coherence, back-azimuth, apparent velocity, signal-to-noise ratio, peak acoustic pressure, directional sector, recording stream (HDF or BDF), and per-event back-azimuth and apparent-velocity uncertainties; complete column definitions and units are given in the archive README (Park et al., 2026c). The archived waveform windows are unfiltered raw digitizer counts at the native sampling rate: the 0.5–8 Hz band defines the detection procedure, not the distributed data, so users may apply any frequency band up to the Nyquist frequency (50 Hz for HDF, 10 Hz for BDF). Every cataloged transient is shorter than the 150 s window (maximum duration 97 s).

6 Quality control and uncertainties

• **Wind detection bias.** 96.8 % of the detections occur at wind speeds below 5 m s^{-1}, compared with 68.7 % of the array observation time with valid weather-station measurements; the per-bin detection rate falls monotonically from 0.24 h^{-1} below 2 m s^{-1} to below 0.01 h^{-1} above 8 m s^{-1}. Wind speed is from the co-located Jang Bogo automatic weather station (10-min sampling; operated by the Korea Polar Research Institute), matched to each detection by nearest time (within 15 min); the event-matched wind speeds for all cataloged events are archived with the catalog (see Code and data availability). Strong katabatic wind raises the incoherent noise floor and suppresses the coherence-based detector, so occurrence statistics carry a wind- and season-dependent detectability imprint. Users comparing occurrence rates or sector composition should account for wind conditions and, where possible, use calm-restricted or wind-matched subsets. Strong wind at Jang Bogo is

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

overwhelmingly katabatic outflow from the west-northwest: of the 10-min mean wind speeds above 15 m s⁻¹ in the station record, 73 % blow from the WNW sector and 96 % from the WNW–WSW sectors. The strongest local wind-noise conditions are therefore associated predominantly with WNW–WSW katabatic flow; we do not attempt to quantify a source-azimuth-specific detection efficiency here.

- **Back-azimuths** are not corrected for cross-wind advection.
- **Absolute-pressure calibration** relies on the nominal manufacturer sensitivity and has not been cross-checked against an independent barometer; relative amplitude comparisons between events are robust, but absolute values carry a calibration uncertainty.
- **Velocity gate.** The catalog contains coherent acoustic transients only (250–500 m s^{-1}); strongly super-acoustic wind energy (> 500 m s^{-1}) is deliberately excluded and the population is correspondingly incomplete for wind-dominated arrivals.
- **Recording stream.** Events are flagged as HDF (100 Hz) or BDF (20 Hz). The 1,057 BDF events (mostly 2022) fill the 2021–2022 configuration gap but carry larger back-azimuth and apparent-velocity uncertainty owing to the 50 ms native timing resolution. Users requiring the highest directional precision should treat BDF-derived events separately or explicitly account for their larger uncertainties.
- **Non-stationary completeness.** Annual counts reflect operational coverage and configuration as well as source activity (Sect. 3).
- **Marine seismic surveys.** The ESE sector contains a large anthropogenic component: at least 53 % of its events occur on days that show a regular air-gun shot-interval signature attributed to marine seismic surveys in the western Ross Sea by a separate analysis of the authors, whose classifier code is archived together with the association table (Park et al., 2026a). ESE occurrence statistics should therefore not be read as natural source activity without screening for survey days.

7 Catalog validation

7.1 Manual review and false-positive rate

We assessed catalog reliability by expert visual review of a stratified random sample of 650 catalog entries, 550 drawn from the 100 Hz HDF stream and 100 from the 20 Hz BDF stream, spanning all years and all five directional sectors. For each event, the three-element, 0.5–8 Hz waveforms were visually inspected and classified as genuine, false, or uncertain. No event was classified as definitely false. For zero confirmed false detections among 650 inspected events, the two-sided 95 % Clopper–Pearson upper bound is 0.57 %. Four additional events (0.6 %) were classified as uncertain (examined below). The 0.57 % bound therefore applies specifically to detections classified as false under the visual-review protocol and does not treat the uncertain category as false; if all four uncertain events were conservatively counted as false, the questionable-event fraction would still be only 0.62 %. This is a sample-based visual check, not an exhaustive screening of the full catalog. Because the review sample was stratified, the confidence bound describes the pooled reviewed sample and is not a design-weighted confidence interval for the full catalog.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Four events (0.6 %) were classed uncertain, and we re-examined each of these individually. All four lie just above the coherence acceptance threshold (coherence 0.65–0.69, signal-to-noise ratio ≤ 8). These cases reflect low detection confidence; each nevertheless shows a coherent arrival near the detection time. About one third of the catalog has coherence below 0.70, and the uncertain classifications concentrate in this near-threshold population. Users who require a more conservative event set can down-weight or remove this near-threshold population using the coherence and signal-to-noise columns. Per-event back-azimuth and apparent-velocity uncertainties are provided in the catalog; they scale inversely with coherence (median 0.6° in back-azimuth for HDF events and about 2° for the coarser-sampled BDF events). BDF-stream events additionally carry larger timing uncertainty (Sect. 6) and are flagged by recording stream.

An objective property of the catalog corroborates the manual result and shows that the acoustic-velocity concentration is not an artifact of the acceptance gate. Before the velocity gate is applied, the coherent triggers (inter-element coherence ≥ 0.6) already concentrate sharply about the acoustic speed. In a representative re-scan (every tenth array-day, 2017–2026), 94 % of coherent triggers fall within the 250–500 m s^{-1} band with a median of 334 m s^{-1}, and the gate removes only a small super- and sub-acoustic tail (Fig. A1a). In the full coherence–velocity plane the high-coherence transients form a distinct population centered on the acoustic speed, clearly separated from the low-coherence, low-velocity incoherent-noise background (Fig. A1b). The retained catalog is therefore a genuine acoustic population selected primarily by coherence, not an artifact imposed by the velocity gate.

7.2 End-to-end check against a known global source: the 2022 Hunga Tonga eruption

We examined the KIS records for the atmospheric pressure wave from the climactic eruption of Hunga Tonga–Hunga Haʻapai on 15 January 2022 (Matoza et al., 2022). This serves as an independent, end-to-end check of the array timing and coherent response. The great-circle distance from the volcano to the array is 6134 km, and the bearing to the source is 23°. At the acoustic-gravity (Lamb) wave celerity of about 0.31 km s^{-1}, the direct arrival is predicted at 09:44 UTC, about 5.5 h after the main explosion. Band-pass filtering to 0.001–0.05 Hz reveals a large, coherent wave packet beginning at about 09:50 UTC (Fig. 8). The arrival is nearly identical across all three elements, with a nominal-equivalent (uncalibrated) amplitude of about 2 Pa, some 500 times the pre-arrival background. Because this millihertz signal lies far below the ~0.1 Hz corner of the sensor response (Sect. 2), the nominal-equivalent amplitude greatly underestimates the true pressure amplitude and is used here only as a relative measure. It appears at the distant CTBTO/IMS array I55US (Marty, 2019) about 10 min later, consistent with its greater distance. The wave is also observed on its subsequent passages around the globe. An antipodal arrival on 16 January and weaker return passages on 16–17 January occur at the times predicted for a Lamb wave circling the Earth (Fig. 9).

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

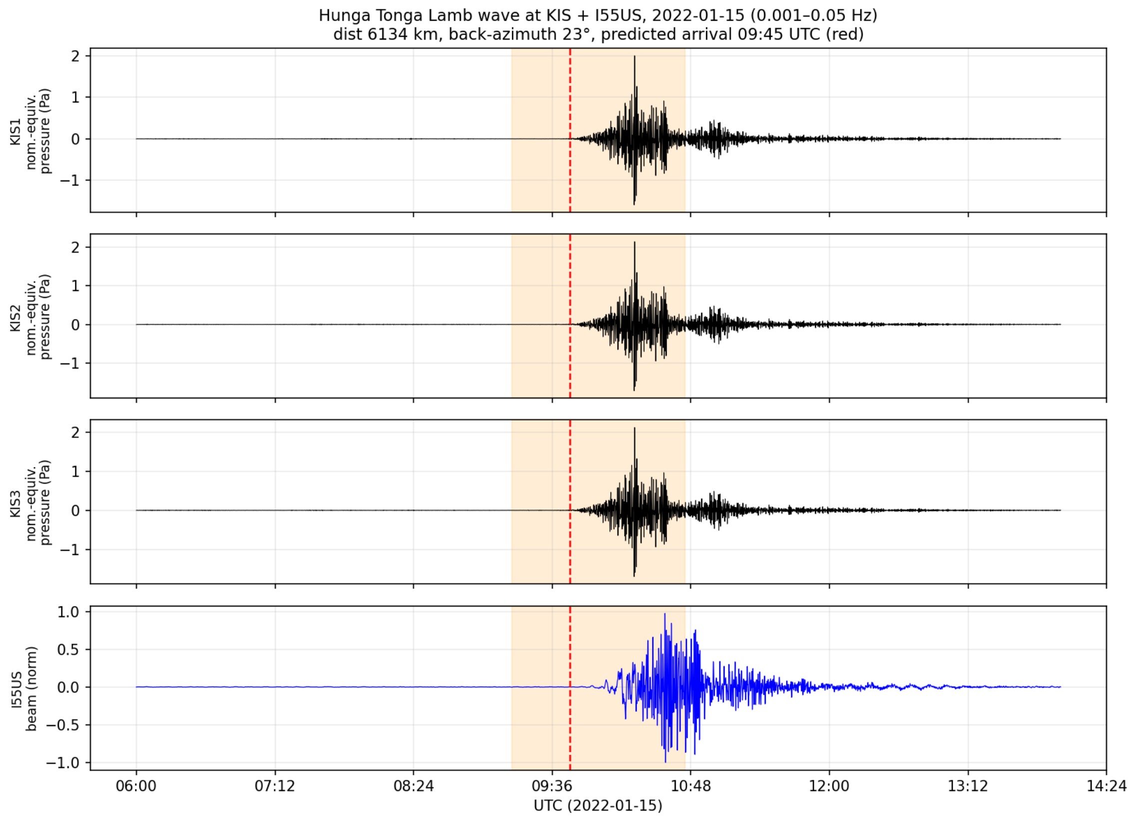

Figure 8. The 2022 Hunga Tonga Lamb wave at KIS and I55US, band-pass filtered to 0.001–0.05 Hz for 15 January 2022. The three KIS elements (top) show a coherent ~2 Pa (nominal-equivalent) arrival beginning near the predicted 09:45 UTC (red dashed line). The I55US beam (bottom) shows the same arrival about 10 min later. The vertical axis is nominal-equivalent pressure (counts divided by the nominal sensitivity), not a barometer-calibrated value; because the sensor response is strongly attenuated in this band (Sect. 2), it underestimates the true amplitude.

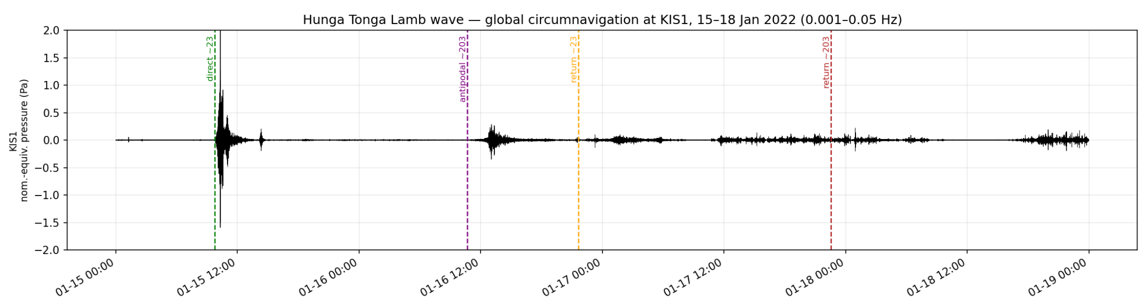

Figure 9. Global circumnavigation of the 2022 Hunga Tonga Lamb wave at KIS (element KIS1, 0.001–0.05 Hz, 15–18 January 2022). Dashed lines mark the predicted times of the direct arrival, the antipodal (long-arc) passage, and two subsequent return passages. Pressure is shown as nominal-equivalent pressure using the manufacturer's sensitivity.

The Lamb wave itself is a millihertz signal below the 0.5–8 Hz detection band and therefore has no catalog entry. The eruption is nevertheless clearly represented in the catalog (Fig. 10). Beginning at 09:49:49 UTC, at the onset of the Lamb-

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

wave arrival, the catalog contains 36 coherent transients within 2.5 h. They are the 36 entries, among the 37 consecutive entries KIS16425–KIS16461, that fall within 5–40° of back-azimuth between 09:45 and 12:30 UTC (the selection window of Fig. 10); the single excluded entry (KIS16454) arrives from 191°. Their back-azimuths span 12–26° (mean 19°), close to the 23° great-circle bearing to the volcano, and they have coherence ≥ 0.77 and apparent velocities of 328–430 m s^{-1}. The background rate from this bearing in January 2022 was about 0.5 events per day, and the sequence includes the largest nominal-equivalent peak pressure in the entire catalog (0.54 Pa; KIS16439). Mt Melbourne lies at a similar bearing (26°), but the timing ties these transients to the eruption. The ~4° offset of the mean back-azimuth from the great-circle bearing is consistent with uncorrected propagation effects such as cross-wind advection (Sect. 6). No comparable sequence accompanies the later Lamb-wave passages (Fig. 10). Users can reproduce both the Lamb-wave observation and these catalog entries, because the continuous three-element records for 14–18 January 2022 are openly archived together with the instrument response (Code and data availability). The example thus supports the consistency of the array timing and shows that the catalog captures a well-defined distant source. Because the Lamb wave lies well below the flat-response band, we do not use it to validate absolute-pressure calibration.

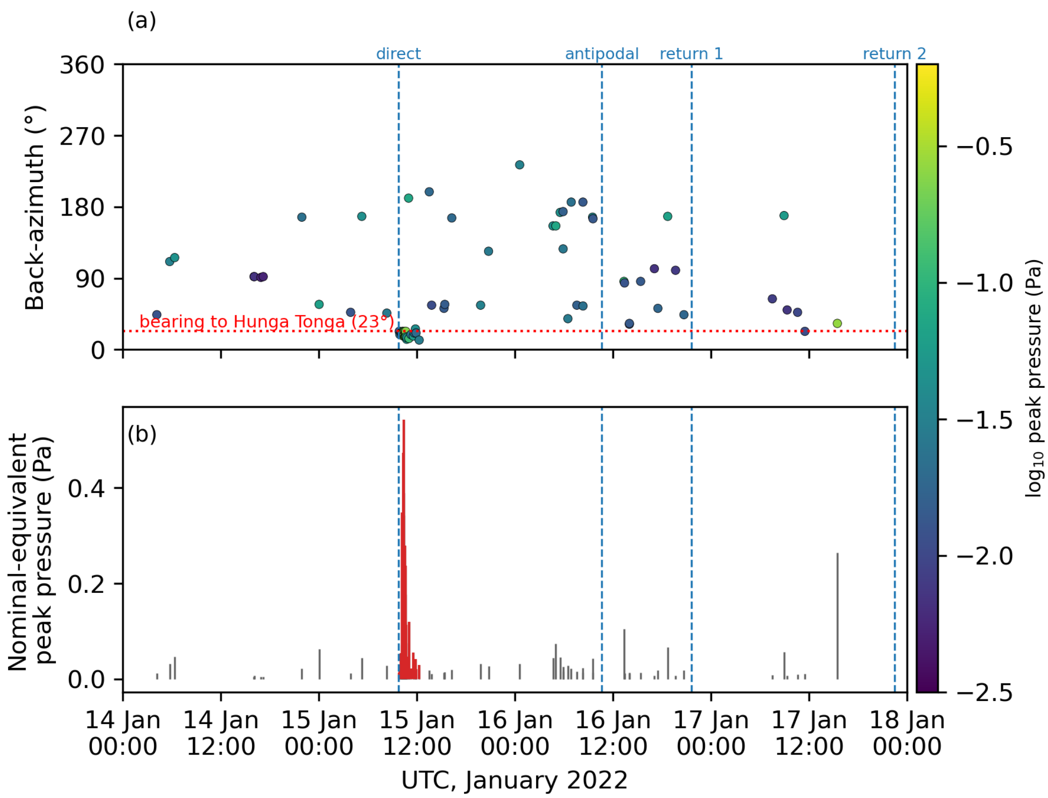

Figure 10. The 2022 Hunga Tonga eruption in the transient catalog, 14–17 January 2022. (a) Back-azimuth of all cataloged events, colored by nominal-equivalent peak pressure; the red dotted line marks the 23° great-circle bearing to the volcano. (b) Nominal-equivalent peak pressure; red bars are the 36 transients selected within 5–40° of back-azimuth between 09:45 and 12:30 UTC on

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

15 January, whose back-azimuths span 12–26°. Blue dashed lines mark the predicted Lamb-wave passages for a celerity of 0.31 km s^{-1} (direct, antipodal, and two return passages).

7.3 A cryospheric example: transients from the direction of David Glacier

Glacial sources are a primary target for cryospheric users of the catalog. Figure 11 shows event KIS06943 (30 November 2018, 18:14:04 UTC) from the west-southwest sector. Its back-azimuth of 229° points to the grounding zone of David Glacier, where the glacier discharges into the Drygalski Ice Tongue (230°, ~125 km; Fig. 1a). The transient is a short (9.8 s), impulsive, highly coherent (coherence 0.98) arrival with a signal-to-noise ratio of 96, an apparent velocity of 336 m s^{-1}, and most of its energy between 0.5 and 2 Hz. It arrives 417 s after a local earthquake located near the David Glacier grounding zone (75.364° S, 160.623° E; 133 km from the array), corresponding to a celerity of 0.318 km s^{-1}, within the range expected for tropospheric acoustic propagation. The earthquake–infrasound association is taken from the openly archived association table of the companion analysis paper (Park et al., 2026a), which also contains the code used to test such pairs against a clustering-preserving null and to locate the David Glacier source independently with the IMS array I55US. The example shows how catalog back-azimuths, detection times, and archived waveform windows can be combined with independent cryoseismic observations. Users should note that roughly nine in ten west-southwest transients have no cataloged seismic counterpart (Park et al., 2026a), so the sector as a whole should not be equated with seismogenic glacier events.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

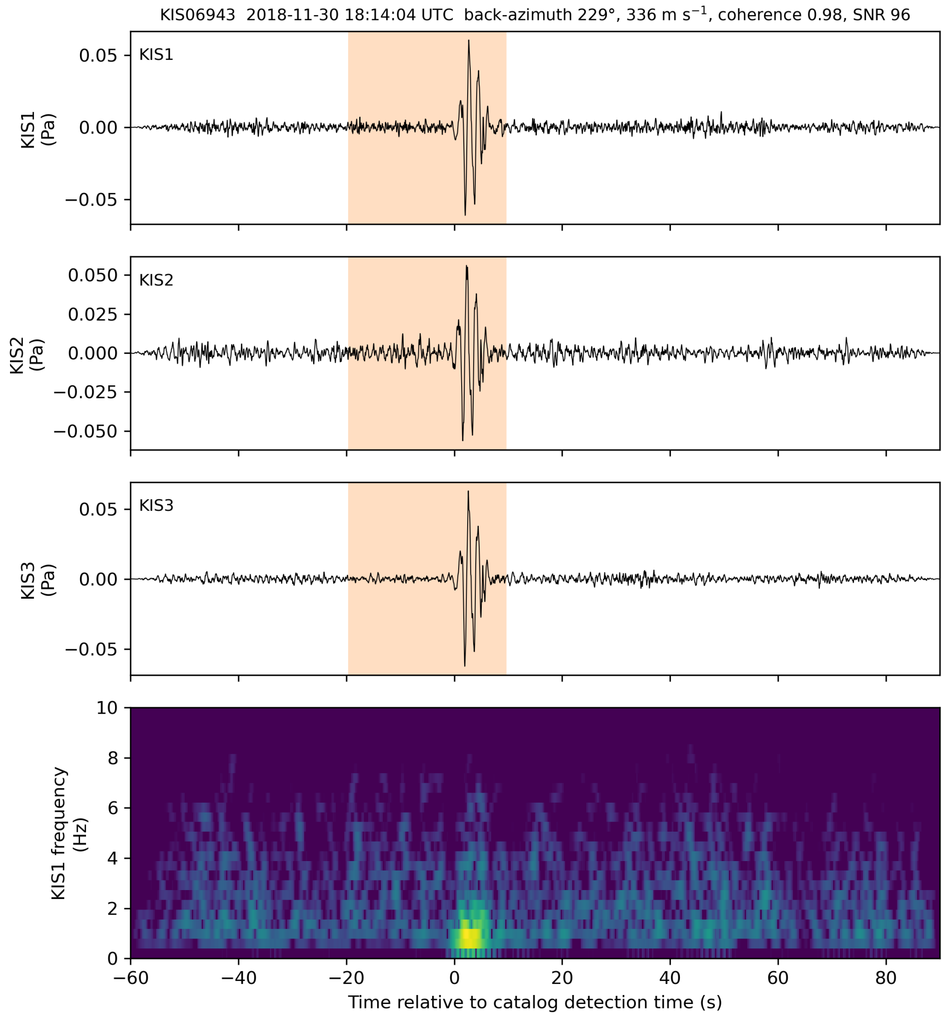

Figure 11. A cataloged transient from the direction of David Glacier (event KIS06943, 30 November 2018, 18:14:04 UTC; back-azimuth 229°, apparent velocity 336 $m\ s^{-1}$). Top three panels: KIS1–KIS3 waveforms from the archived event window, band-pass filtered to 0.5–8 Hz and shown as nominal-equivalent pressure relative to the catalog detection time. The shaded interval is the arrival window predicted from the origin time and epicentral distance (133 km) of the associated local earthquake for celerities of 0.311–0.334 $km\ s^{-1}$, the interquartile range of the seismic–infrasound pairs in the archived association table (Park et al., 2026a). Bottom panel: KIS1 spectrogram.

8 Conclusions

We present a consistently processed catalog of 23,768 coherent atmospheric infrasound transients recorded at Jang Bogo Station over 2,507 array-days during 2017–2026. The catalog includes event timing, waveform and array parameters, directional sectors, recording-stream information, and per-event directional uncertainties, together with archived waveform windows and processing code. Manual review, pre-gate velocity analysis, stream comparison, and the Hunga Tonga Lamb-

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

wave observation provide complementary checks on catalog reliability and instrument timing. Worked examples show that the catalog captures the 2022 Hunga Tonga eruption and an example transient from the direction of David Glacier associated with a local glacial earthquake. The dataset is intended to support future cryo-acoustic and infrasound studies on the Antarctic margin, while users should account for wind-dependent detectability, recording-stream differences, and the nominal nature of the absolute-pressure calibration.

Appendix A: Processing parameters and stream comparability

Table A1 lists the detection and parameter-estimation settings used to build the catalog. Table A2 defines the 13 waveform-shape features used for the clustering of Sect. 5; the features were median-imputed for missing values, clipped to the 1st–99th percentile, standardized (z-score), and clustered by Ward linkage, with the number of clusters chosen by silhouette score.

Table A1. Detection and parameter-estimation settings

Parameter	Value
Analysis band	0.5–8 Hz (4th-order Butterworth, zero-phase)
STA / LTA windows	2 s / 80 s (recursive STA/LTA)
Trigger on / off ratios	3.5 / 1.5
Coincidence	at least 2 of 3 elements
Analysis window	trigger duration (clipped to 8–60 s for parameter estimation), with 3 s pre- and 6 s post-padding; catalogued duration retains the full STA/LTA trigger duration
Inter-element coherence	mean of the normalized cross-correlation peaks of pairs KIS1–KIS2 and KIS1–KIS3
Time delays	lag at the peak of the normalized cross-correlation
Slowness	s solves $G \cdot s = \Delta t$ with $G = [r_2 - r_1; r_3 - r_1]$; back-azimuth = atan2($s_x$, s_y) + 180°; apparent velocity = reciprocal of the slowness magnitude
Acceptance gate	coherence ≥ 0.6 and apparent velocity 250–500 m s^{-1}
Duration (cataloged)	STA/LTA trigger duration
Signal-to-noise ratio	peak Hilbert-envelope amplitude / day-median envelope amplitude (KIS1)
Peak pressure	maximum absolute count amplitude in the window / nominal sensitivity (8.0×10^5 counts Pa^{-1})

Back-azimuth / velocity uncertainty	provided per event (columns baz_uncertainty_deg, app_velocity_uncertainty_ms); array geometry + coherence-based time-delay error (Szuberla and Olson, 2004)
BDF stream	20 Hz, resampled to 100 Hz before detection

Table A2. Waveform-shape features used for the clustering of Sect. 5

Feature	**Definition**
peakf	frequency of the peak of the 0.5–8 Hz power spectrum
centroid	spectral centroid (power-weighted mean frequency)
bandwidth	spectral bandwidth (power-weighted standard deviation of frequency)
flatness	spectral flatness (geometric-to-arithmetic mean ratio of the spectrum)
e_lo	fractional spectral energy, 0.5–2 Hz
e_mid	fractional spectral energy, 2–4 Hz
e_hi	fractional spectral energy, 4–8 Hz
dur	envelope duration above half the peak amplitude
rise	time from 10 % to the peak of the Hilbert envelope
decay	time from the peak to 37 % of the Hilbert envelope
kurt	kurtosis of the waveform
crest	crest factor (peak-to-RMS amplitude ratio)
zcr	zero-crossing rate

Fig. A1 shows that the acoustic apparent-velocity concentration is present before the velocity gate is applied, corroborating the manual validation of Sect. 7.1.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

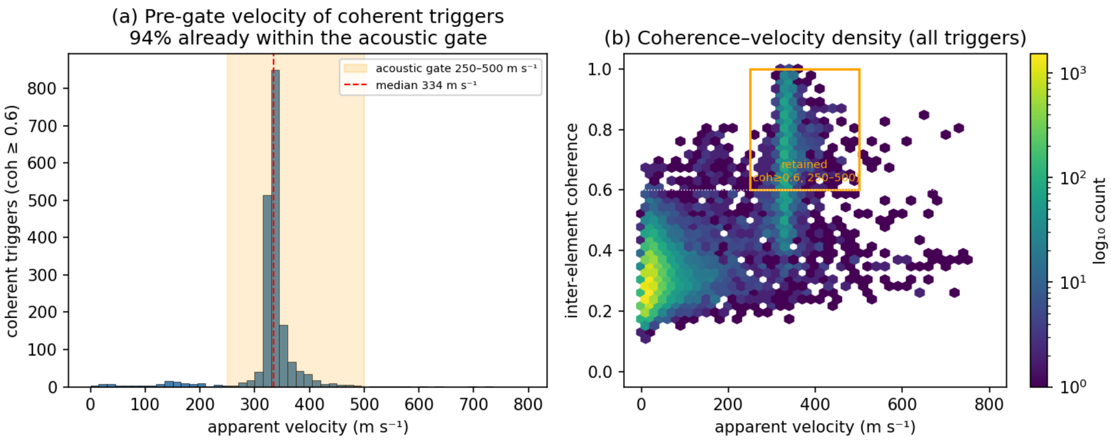

Figure A1. Pre-gate coherence and apparent velocity of all STA/LTA triggers from a systematic 1-in-10 subsample of available array-days during 2017–2026. (a) Apparent-velocity distribution of coherent triggers (inter-element coherence ≥ 0.6); 94 % already fall within the 250–500 m s^{-1} acoustic gate (shaded), with a median of 334 m s^{-1}. (b) Two-dimensional coherence–velocity density of all triggers (logarithmic color scale); the high-coherence transients form a distinct population at the acoustic speed, separated from the low-coherence incoherent-noise background. The acceptance gate (coherence ≥ 0.6, 250–500 m s^{-1}; orange box) isolates this population, showing that the acoustic-velocity concentration is present before application of the velocity gate.

The catalog combines a 100 Hz (HDF) and a 20 Hz (BDF) recording stream (Sect. 3). To characterize the two streams, we compared their apparent-velocity, coherence, and duration distributions (Fig. A2). The distributions overlap substantially but are not identical: the two-sample Kolmogorov–Smirnov statistics are D = 0.21, 0.14, and 0.12 for apparent velocity, coherence, and duration, respectively, and BDF events show modest systematic shifts toward higher coherence and longer duration. The recording-stream flag therefore allows users to account explicitly for these differences.

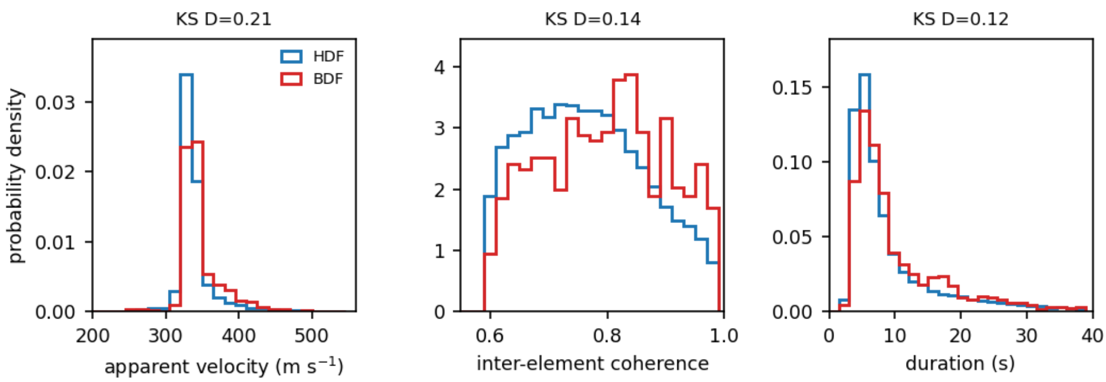

Figure A2. Comparison of the HDF (100 Hz) and BDF (20 Hz) recording streams. Probability densities of (left) apparent velocity, (center) inter-element coherence, and (right) duration for the 22,711 HDF and 1,057 BDF catalog events; the two-sample Kolmogorov–Smirnov statistic D is shown for each panel. The distributions overlap substantially but are not identical; BDF events show modest shifts toward higher coherence and longer duration (Appendix A).

Code and data availability

The event catalog and the per-event waveform windows associated with each catalog entry are openly available at Zenodo, https://doi.org/10.5281/zenodo.21882394 (Park et al., 2026c), under a CC BY 4.0 license, with anonymous access that requires no registration or login. The record comprises the event catalog (`KIS_infrasound_transient_catalog_2017-2026_v2.2.csv`, 23,768 rows; version 2.2 adds per-event back-azimuth and apparent-velocity uncertainties, and version 2.1 removed 39 spurious detections from two corrupt-data days (17–18 February 2019) present in v2), the array element coordinates (`KIS_array_stations.csv`), the per-event three-element waveform windows (detection time −60 s to +90 s, KIS1/KIS2/KIS3, one miniSEED per event, unfiltered raw-count miniSEED waveforms at the native sampling rate, with the nominal sensor sensitivity for pressure conversion documented in the README), an index manifest, an event-matched wind-speed table for all cataloged events (`KIS_event_wind.csv`), and a README documenting all column definitions, units, methods, and quality control. The event catalog is additionally mirrored at the Korea Polar Data Center (https://doi.org/10.22663/KOPRI-KPDC-00003159) (Park et al., 2026e). The continuous full-record KIS waveforms underpin ongoing KOPRI studies and are available from the corresponding author under a KOPRI data-sharing agreement; the archived event windows allow independent inspection and recomputation of waveform-derived parameters such as inter-element coherence, back-azimuth, apparent velocity, and nominal-equivalent peak pressure for each retained event. The original full-record detection scan cannot be reproduced from the event windows alone. To make the Hunga Tonga example (Sect. 7.2) fully reproducible, the continuous three-element KIS records for 14–18 January 2022 (20 Hz, raw counts, miniSEED) and an FDSN StationXML file with the full instrument response are openly archived at Zenodo, https://doi.org/10.5281/zenodo.22786159 (Park et al., 2026d), under a CC BY 4.0 license. The seismic–infrasound association table used in Sect. 7.3 is archived at https://doi.org/10.5281/zenodo.21910213 (Park et al., 2026a). Infrasound data from array I55US used in the Hunga Tonga cross-check are from the CTBTO International Monitoring System and are openly available through the EarthScope (IRIS) FDSN data center.

The detection and array-processing code is archived at Zenodo, https://doi.org/10.5281/zenodo.21867891 (Park et al., 2026b), under the MIT license.

All archived data and code are accessible to reviewers and readers without registration, login, or any personalized information.

Author contributions

The contributions follow the CRediT taxonomy. YP: conceptualization, methodology, software, formal analysis, investigation, data curation, visualization, and writing (original draft). JJ and WSL: resources and investigation (annual operation and maintenance of the KIS infrasound array and data collection at Jang Bogo Station), and writing (review and editing). All authors discussed the results and approved the manuscript.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We thank the staff at Jang Bogo Station for operating and maintaining the KIS infrasound array.

Financial support

This research was supported by the Korea Institute of Marine Science & Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries (RS-2023-00256677; PM26020).

References

Allen, R. V.: Automatic earthquake recognition and timing from single traces, Bulletin of the Seismological Society of America, 68, 1521–1532, 10.1785/BSSA0680051521, 1978.

Aster, R. C. and Winberry, J. P.: Glacial seismology, Reports on Progress in Physics, 80, 126801, 10.1088/1361-6633/aa8473, 2017.

Bamber, J. L., Gomez-Dans, J. L., and Griggs, J. A.: A new 1 km digital elevation model of the Antarctic derived from combined satellite radar and laser data – Part 1: Data and methods, The Cryosphere, 3, 101–111, 10.5194/tc-3-101-2009, 2009.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J.: ObsPy: A Python toolbox for seismology, Seismological Research Letters, 81, 530–533, 10.1785/gssrl.81.3.530, 2010.

Cansi, Y.: An automatic seismic event processing for detection and location: The P.M.C.C. Method, Geophysical Research Letters, 22, 1021–1024, 10.1029/95GL00468, 1995.

Hupe, P., Ceranna, L., Le Pichon, A., Matoza, R. S., and Mialle, P.: International Monitoring System infrasound data products for atmospheric studies and civilian applications, Earth System Science Data, 14, 4201–4230, 10.5194/essd-14-4201-2022, 2022.

Kanao, M., Murayama, T., Yamamoto, M.-Y., Ishihara, Y., Muramatsu, D., Iwakuni, M., and Nogami, M.: Detection of infrasound sources using three-years array data in 2019-2021 deployed at the Lützow-Holm Bay region, Antarctica, Annals of Geophysics, 67, 10.4401/ag-9073, 2024.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J.: ObsPy: a bridge for seismology into the scientific Python ecosystem, Computational Science & Discovery, 8, 014003, 10.1088/1749-4699/8/1/014003, 2015.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Marty, J.: The IMS Infrasound Network: Current Status and Technological Developments, in: Infrasound Monitoring for Atmospheric Studies: Challenges in Middle Atmosphere Dynamics and Societal Benefits, edited by: Le Pichon, A., Blanc, E., and Hauchecorne, A., Springer, Cham, 3–62, 10.1007/978-3-319-75140-5_1, 2019.

Matoza, R. S., Fee, D., Assink, J. D., Iezzi, A. M., Green, D. N., Kim, K., Toney, L., Lecocq, T., Krishnamoorthy, S., Lalande, J. M., Nishida, K., Gee, K. L., Haney, M. M., Ortiz, H. D., Brissaud, Q., Martire, L., Rolland, L., Vergados, P., Nippress, A., and Park, J.: Atmospheric waves and global seismoacoustic observations of the January 2022 Hunga eruption, Tonga, Science, 377, 95–100, 10.1126/science.abo7063, 2022.

Murayama, T., Kanao, M., Yamamoto, M.-Y., and Ishihara, Y.: Infrasound Signals and Their Source Location Inferred from Array Deployment in the Lützow–Holm Bay Region, East Antarctica: January-June 2015, International Journal of Geosciences, 08, 181–188, 10.4236/ijg.2017.82007, 2017.

Murayama, T., Kanao, M., Yamamoto, M.-Y., Ishihara, Y., Matsushima, T., and Kakinami, Y.: Infrasound array observations in the Lützow-Holm Bay region, East Antarctica, Polar Science, 9, 35–50, 10.1016/j.polar.2014.07.005, 2015.

Park, Y., Jung, J., and Lee, W. S.: Analysis data and code for "Decade-scale discrimination of coherent infrasound sources at Terra Nova Bay, East Antarctica, and dual-array location of a David Glacier source", Zenodo [dataset], 10.5281/zenodo.21910213, 2026a.

Park, Y., Jung, J., and Lee, W. S.: KIS infrasound transient detector: code for the Terra Nova Bay (Antarctica) 2017–2026 coherent-infrasound-transient catalogue (v2), Zenodo [code], 10.5281/zenodo.21867891, 2026b.

Park, Y., Lee, W. S., and Jung, J.: A decade-scale catalogue of coherent atmospheric infrasound transients at Terra Nova Bay, East Antarctica (2017-2026), with per-event waveform windows, Zenodo [dataset], 10.5281/zenodo.21882394, 2026c.

Park, Y., Lee, W. S., and Jung, J.: KIS continuous infrasound records during the 2022 Hunga Tonga eruption (14–18 January 2022), Terra Nova Bay, East Antarctica, Zenodo [dataset], 10.5281/zenodo.22786159, 2026d.

Park, Y., Lee, W. S., and Jung, J.: 2017-2026 infrasound transient catalog (v2.2), Korea Polar Data Center (KPDC) [dataset], 10.22663/KOPRI-KPDC-00003159, 2026e.

Podolskiy, E. A. and Walter, F.: Cryoseismology, Reviews of Geophysics, 54, 708–758, 10.1002/2016rg000526, 2016.

Richardson, J. P., Waite, G. P., FitzGerald, K. A., and Pennington, W. D.: Characteristics of seismic and acoustic signals produced by calving, Bering Glacier, Alaska, Geophysical Research Letters, 37, L03503, 10.1029/2009GL041113, 2010.

Rost, S. and Thomas, C.: Array Seismology: Methods and Applications, Reviews of Geophysics, 40, 10.1029/2000rg000100, 2002.

Scambos, T. A., Haran, T. M., Fahnestock, M. A., Painter, T. H., and Bohlander, J.: MODIS-based Mosaic of Antarctica (MOA) data sets: Continent-wide surface morphology and snow grain size, Remote Sensing of Environment, 111, 242–257, 10.1016/j.rse.2006.12.020, 2007.

https://doi.org/10.5194/essd-2026-647
Preprint. Discussion started: 22 September 2026

Szuberla, C. A. L. and Olson, J. V.: Uncertainties associated with parameter estimation in atmospheric infrasound arrays, Journal of the Acoustical Society of America, 115, 253–258, 10.1121/1.1635407, 2004.

Yamamoto, M.-Y., Ishihara, Y., and Kanao, M.: Infrasonic Waves in Antarctica: A New Proxy for Monitoring Polar Environment, International Journal of Geosciences, 04, 797–802, 10.4236/ijg.2013.44074, 2013.