Extracting low signal-to-noise ratio events with the Hough transform from sparse array data

Averbuch, Gil; Assink, Jelle D.; Smets, Pieter S.M.; Evers, Läslo G.

DOI
10.1190/GEO2017-0490.1

Publication date
2018

Document Version
Final published version

Published in
Geophysics

Citation (APA)
Averbuch, G., Assink, J. D., Smets, P. S. M., & Evers, L. G. (2018). Extracting low signal-to-noise ratio events with the Hough transform from sparse array data. *Geophysics*, *83*(3), WC43-WC51. https://doi.org/10.1190/GEO2017-0490.1

GEOPHYSICS, VOL. 83, NO. 3 (MAY-JUNE 2018); P. WC43–WC51, 12 FIGS., 1 TABLE.
10.1190/GEO2017-0490.1

Extracting low signal-to-noise ratio events with the Hough transform from sparse array data

Gil Averbuch[1], Jelle D. Assink[2], Pieter S. M. Smets[1], and Läslo G. Evers[3]

ABSTRACT

Low-frequency acoustic, i.e., infrasound, waves are measured by sparse arrays of microbarometers. Recorded data are processed by automatic detection algorithms based on array-processing techniques such as time-domain beam forming and f-k analysis. These algorithms use a signal-to-noise ratio (S/N) value as a detection criterion. In the case of high background noise or in the presence of multiple coinciding signals, the event's S/N decreases and can be missed by automatic processing. In seismology, detecting low-S/N events with geophone arrays is a well-known problem. Whether it is in global earthquake monitoring or reservoir microseismic activity characterization, detecting low-S/N events is needed to better understand the sources or the medium of propagation. We use an image-processing technique as a post-processing step in the automatic detection of low S/N events. In particular, we consider the use of the Hough transform (HT) technique to detect straight lines in beam-forming results, i.e., a back azimuth (BA) time series. The presence of such lines, due to similar BA values, can be indicative of a low-S/N event. A statistical framework is developed for the HT parameterization, which includes defining a threshold value for detection as well as evaluating the false alarm rate. The method is tested on synthetic data and five years of recorded infrasound from glaciers. It is shown that the automatic detection capability is increased by detecting low-S/N events while keeping a low false-alarm rate.

INTRODUCTION

Sensor arrays are used in many branches of geophysics for the detection of coherent energy. The use of arrays aids in the improvement of signal-to-noise ratio (S/N) and permits estimation of the wavefront parameters by the use of beam forming (Rost, 2002). The beam-forming process combined with a correlation detector, e.g., Fisher statistics, provides three important values: (1) S/N; (2) back azimuth (BA), the wavefront's horizontal angle of arrival; and (3) apparent velocity (AV), the horizontal projection of the velocity vector. This can be related to the vertical incidence angle of the wave (Melton and Bailey, 1957; Shumway, 1971; Olson, 2004).

Infrasound refers to low-frequency sound waves, ranging between 0.01 and 20 Hz that propagate in the atmosphere. It is measured by microbarometer arrays that are sensitive to air-pressure fluctuations ranging from mPa up to tens of pascals. Pressure variations due to incoherent wind noise are reduced by the use of microbarometer arrays. Infrasound is primarily a monitoring technique for the verification of the comprehensive nuclear-test-ban treaty (CTBT). Other common sources include, for example, explosions, earthquakes, volcano eruptions, calving glaciers, and sonic booms (Evers, 2008; Le Pichon et al., 2010). Long-range infrasound propagation is affected by wind and temperature variations between the earth's surface and the lower thermosphere (Drob et al., 2003). Consequently, infrasound recordings can be used as a passive remote-sensing technique to estimate wind and temperature conditions in a range of altitudes at which such measurements are rare (Assink et al., 2013; Smets et al., 2016).

In seismology, sensor arrays of different sizes, in terms of aperture and number of elements, are used. Globally, seismic sensors are used to detect earthquakes and explosions (Gibbons and Ringdal, 2006, 2012). On a regional scale, sensor arrays can assist in the analysis of body and surfaces waves (Harmon et al., 2008; Vidal et al., 2011; Panea et al., 2014). On a local scale, arrays are used

Manuscript received by the Editor 27 July 2017; revised manuscript received 13 February 2018; published ahead of production 23 February 2018; published online 19 April 2018.
[1]Delft University of Technology, Department of Geoscience and Engineering, Faculty of Civil Engineering and Geosciences, Delft, The Netherlands. E-mail: g.averbuch@tudelft.nl.
[2]Royal Netherlands Meteorological Institute, R&D Department of Seismology and Acoustics, De Bilt, The Netherlands. E-mail: jelle.assink@knmi.nl.
[3]Delft University of Technology, Department of Geoscience and Engineering, Faculty of Civil Engineering and Geosciences, Delft, The Netherlands and Royal Netherlands Meteorological Institute, R&D Department of Seismology and Acoustics, De Bilt, The Netherlands. E-mail: pieter.smets@knmi.nl.

in exploration seismology to detect microseismic activity in a reservoir or a mine (Potvin and Hudyma, 2001; Chambers et al., 2010; Boué et al., 2013).

One of the shared challenges in seismological and infrasound studies is the detection of low-S/N events due to weak or interfering sources or high background noise levels. For example, body waves can be masked by surface waves or high background noise. To detect low-magnitude seismic events, stacking of correlation traces within the co-array is used to increase the signal's S/N before correlating it with the recordings of other arrays (Gibbons and Ringdal, 2006). In ambient noise interferometry, beam-forming methods are used to distinguish between time segments with either surface or body waves. The detection of low-amplitude body waves proves to be difficult because body-wave amplitudes decay faster (Forghani and Snieder, 2010; Draganov et al., 2013). Furthermore, microseismic source localization can be accomplished by using surface arrays (Williams-Stroud et al., 2010; Verdon et al., 2017). The monitoring of microseismic events in the reservoirs and mines is used to provide information on hydraulic fracturing, changes in the local stress field, imaging, and for seismic hazard analysis (Chambers et al., 2010; Maxwell et al., 2010). The detection of such events can be difficult in cases in which background noise levels are high due to anthropogenic activity, leading to misleading results (Ge, 2005; Warpinski, 2009). Therefore, low-S/N event detection is desirable.

In the detection of low-S/N signals, it is essential to estimate the probability of missed events and false alarms as can be significant in such conditions. The statistical framework of Fisher statistics has been used to estimate these probabilities from F-distributions (Evers, 2008), under the assumption that noise is uncorrelated. A detection is defined by setting a threshold to the Fisher ratio. The Fisher ratio F and the single channel S/N are related by $F = N \cdot \mathrm{S/N}^2 + 1$, where N is the number of recording instruments. Although false alarms can be attributed to the Fisher detector statistics, missed events can be attributed to low S/N values. Analysis of noise over infrasound arrays shows that noise can be correlated (Arrowsmith et al., 2008). Consequently, noise may obtain a higher F. In addition, concurrent signals can also appear as incoherent noise to the signal of interest and thereby reduce the S/N values. For operational purposes, the threshold for automatic detectors is a S/N equal or greater than one (Le Pichon et al., 2009). This threshold will be used in the remainder of this study.

To reduce the missed events rate, by detecting events with S/N below one, a new approach based on the work of Brown et al. (2008) is proposed. In contrast to incoherent noise, for which the BA and the AV values appear to be random, coherent infrasound is characterized by coherent BA and AV values as a function of time. These characteristics can be recognized even when noise levels are high. The Hough transform (HT) is an image-processing technique for detecting shapes such as straight lines, circles, and ellipses in a pixilated data set (Nixon, 2008; Lezama et al., 2015). This technique has applications in graphical element recognition, statistical shapes analysis, detection of grids, geophysics, geology, archaeology, and remote sensing (Duda and Hart, 1972; Song and Lyu, 2005; Hall et al., 2006; Lezama et al., 2015).

Thus, low-S/N infrasound events can be detected by identifying straight lines in the resolved BA (Brown et al., 2008). The key is to define a threshold for detection of a line within a given window length. In the work of Brown et al. (2008), two threshold levels are used. The first is based on the relation between the potentially detected signal duration and its S/N, whereas the second iteratively requires the signal length, F, and S/N to be larger than a varying threshold values. Brown et al. (2008) show that signals with a minimum duration of 84 s and S/N larger than 0.5 are detectable with false-alarm rate of one per day.

Hitherto, an arbitrary threshold number has been used to determine whether an infrasound signal is present or not. In this paper, a new method, based on the binomial distribution and noise investigation, is introduced to define suitable settings for the HT. These settings can be optimized for specific signal lengths. The method is tested on the synthetic and real data sets to investigate its performance at low S/Ns. Similar methods have been used to reduce missed events in time-range radar observations. Elazar (1995) shows that the binomial distribution can be used to determine the expected HT false-alarm rate applied to such data sets. For an event to be detected, it must pass the S/N threshold and the HT threshold, which is based on known optimal parameters (Carlson et al., 1994; Elazar, 1995).

The remainder of this paper is organized as follows: It starts with describing the Fisher detector, which is used as the primary data processing tool. In the following two sections, the principles of detecting straight lines using the HT and the method of setting the HT settings are discussed. Thereafter, the HT is applied on both synthetic and real data. Finally, results are discussed and summarized.

BEAM FORMING USING FISHER STATISTICS

Infrasonic events are recorded by microbarometer arrays. The recordings are processed using beam-forming techniques that provide us with the signal S/N, BA, and AV. The Fisher detector is used to evaluate each beam using a delay-and-sum approach. Beam forming is carried out over a grid of slowness (reciprocal of velocity) vectors, assuming a plane-wave signal model. For every vector, F is calculated. F is defined as the ratio of the variation between recordings to the variation within a recording. The null hypothesis that of no coherent signal is present is accepted in the case that F is unity. A detection is made for higher F values, which follows from Fisher statistics (Melton and Bailey, 1957; Olson, 2004). As mentioned, infrasonic noise distributions do not strictly follow the theoretical distribution and incoherent noise may be associated with higher F than unity.

The Fisher ratio can be written as

$$F = \frac{T(N-1)}{N(T-1)} \frac{\sum_{t=1}^{T} \left(\sum_{n=1}^{N} x_{nt} \right)^2 - \frac{1}{T} \left(\sum_{t=1}^{T} \sum_{n=1}^{N} x_{nt} \right)^2}{\sum_{t=1}^{T} \sum_{n=1}^{N} x_{nt}^2 - \frac{1}{N} \sum_{t=1}^{T} \left(\sum_{n=1}^{N} x_{nt} \right)^2}, \tag{1}$$

where T is the number of samples, N is the number of the sensors, and x_{nt} represents the sample value for a specific sensor n and time sample t. The F-ratio is calculated for all vectors in the slowness grid. The maximum F-value corresponds to the slowness vector, which describes the wavefront of the (most powerful) signal in the time-segment of analysis. The BA and AV can be determined from this slowness vector.

The Fisher detector is applied to time segments from continuous recordings and changes of the wavefield with regard to S/N, BA, and AV, are evaluated over time. To detect a possible event, consecutive, self-similar values in BA are identified. Because AV values are more variable (see Figure 2 in Evers and Haak, 2005) and can vary due to different propagation paths, only the BA results are used here as an input to the HT.

DETECTING STRAIGHT LINES

The human brain is "wired" in a way that special neuron groups are responsible for recognizing geometrical patterns. Some of these neurons groups are especially sensitive to straight lines with horizontal and vertical orientations (Furmanski and Engel, 2000). However, straight lines can be masked by other geometric dotted shapes or obscured by randomly distributed points, making them difficult for the human eye to detect. In the case of complex images, a digital computer detector, such as the HT is useful.

Considering a data set with N_p points, the HT maps each pair of points $(x_i, y_i), (x_j, y_j) i \neq j \; i, j = 1 \ldots N_p$ onto a new domain defined by the slope between the points τ and the intercept with the axis origin μ (in the original domain). The new domain is defined as

$$\mu = y_i + \tau x_i, \tag{2}$$

$$\tau = -\frac{y_j - y_i}{x_j - x_i}. \tag{3}$$

The number of collinear points in the original data set will be given by the number of the identical parameter pairs in the new domain; that is, if pairs of points have the same slope and intercept, they lie on the same line.

Due to the resemblance of the HT with the well-known τ-p transform (also known as slant stack), let us distinguish between the two. The goal of the τ-p transform is to apply a plane-wave decomposition to the recorded data. It is realized by first applying a linear moveout correction to the traces and then summing their amplitudes. Applying it to a range of slowness parameters will provide a slant-stack gather. It contains the relevant slowness and dip parameters of the original data (Yilmaz, 2001). On the contrary, the HT is applied to beam-forming results (pixilated data instead of traces) and counts how many pairs of BA results have the same slope-intercept parameters.

A unique characteristic of an infrasound event is that the BA values are self similar in time. Using the HT as a postprocessing tool for detecting horizontal lines will enhance the detection of events. This is useful for events with low S/N that do not pass the S/N detection threshold, and therefore are missed.

A significant difference between detecting straight lines in a "regular" image compared with beam-forming results is that in the latter, time plays an important role. Because infrasound events occur in a finite time interval, this characteristic should be taken into account. Therefore, the Hough window (HW) is introduced, which is a sliding window that runs over the beam-forming results. Within an HW, a possible HT detection may occur (Figure 1).

SETTING THE THRESHOLD USING THE BINOMIAL DISTRIBUTION AND NOISE ANALYSIS

The HT is applied to results obtained using the beam-forming method discussed in the "Beam forming using Fisher statistics" section, without setting a threshold on the S/N. Therefore, all BA values are used as input data points for the HT, irrespective of their S/N. This process is divided into three steps. The first step is to define the minimum signal length to be detected. The number k will be set to the number of data points (Fisher detector results) that correspond to this length. Because long-duration signals can be constructed from short ones, this choice does not prevent us from detecting such long-duration signals.

For low-S/N signals, it is possible that some of the estimated BA values deviate from their true value. Therefore, the detection performance of the HT may be limited by only considering k consecutive aligned points. By allowing for k aligned points in a n points data set ($n > k$), the signal data points do not have to be in consecutive order to be detected by the HT.

The second step is to specify the number n, which is the size of the HW. The HW size is necessarily smaller than the size of the data set. When only incoherent noise is present, the results will be randomly distributed over the BA range. The probability of noise to have a specific azimuth p_{azimuth} is

$$p_{\text{azimuth}} = \frac{1}{(\Phi_{\max} - \Phi_{\min})/\Delta\Phi}, \tag{4}$$

where $\Phi_{\max}$ and $\Phi_{\min}$ bound the BA range with a step size of $\Delta\Phi$. A binomial distribution is used to estimate the probability P_r to get k aligned points out of n HW points in the BA results. In the case that no coherent signal is present, the probability is

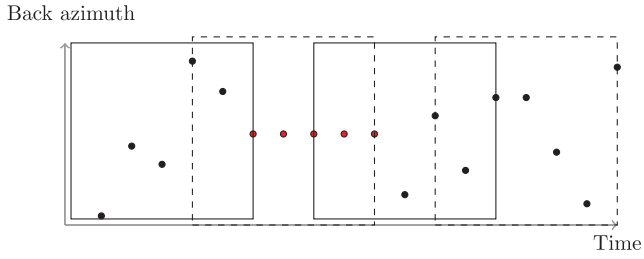

Figure 1. The HW as a sliding window over the beam-forming results. The black dots are due to incoherent noise, and they are therefore randomly distributed. The red dots result from a coherent signal with a constant BA.

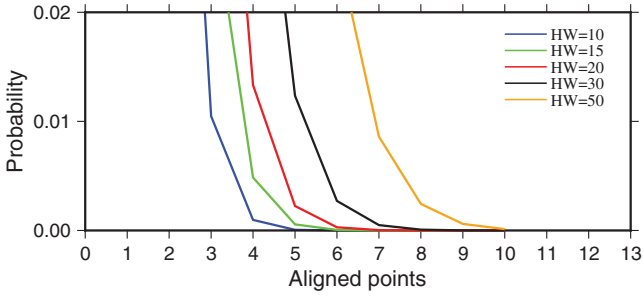

Figure 2. The probability of detecting k aligned points out of an n point HW in the presence of incoherent noise. The blue, green, red, black, and orange curves correspond to 10, 15, 20, 30, and 50 points HW, respectively.

$$P_r(k, n, p_{\text{azimuth}}) = \binom{n}{k} p_{\text{azimuth}}^{k} (1 - p_{\text{azimuth}})^{n-k}. \quad (5)$$

Figure 2 shows the probability of getting k aligned points from a certain BA in a HW containing n points. Choosing settings (k, n) with a low probability P_r reduces the false-alarm rate.

The third step is the estimation of the false-alarm rate. To do so, a data set consisting of incoherent noise is constructed. Applying the HT with the desired settings on this noise data set will provide us with an estimation of the false alarm rate.

SYNTHETIC DATA

The method is tested on synthetic data sets to conduct a controlled experiment. The results are used to evaluate the method and to fine-tune the HT parameters. The first data set consists of incoherent noise and is used to evaluate the false alarm rate. The second and third data sets contain a long signal and a repetitive impulsive source, with different signal amplitudes. Random noise is added to each of the recordings and its amplitudes determine the S/N of the events. This way, the detection capability of low-S/N events is evaluated.

The synthetic data sets are constructed with a sampling rate of 20 Hz for an eight-element microbarometer array with the station coordinates of array I18DK, which is part of the International Monitoring System (IMS) for the CTBT (Figure 3). The first data set contains 24 h of random noise with a uniform distribution. The synthetic data are band-pass filtered between 1 and 5 Hz. The Fisher detector settings are time bin 25.6 s, 50% overlap, AV range of 250–450 m/s with 5 m/s interval, and BA range of 0°–360° with a 2° interval. These settings are used throughout this paper.

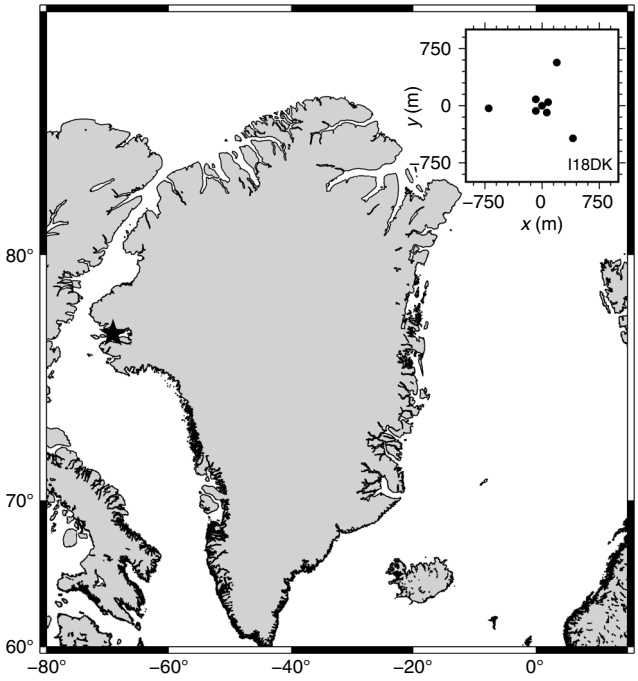

Figure 3. Map of Greenland showing the location of I18DK (black star). The array layout is shown at the top right.

Figure 4 shows the Fisher detector results for the noise data set. As expected, S/N values are very low and the BA values are randomly distributed within the analysis domain. The HT detection rate is high for low threshold values and large HW sizes and decreases as detection conditions become more strict, i.e., high threshold values and small HW sizes (Figure 5).

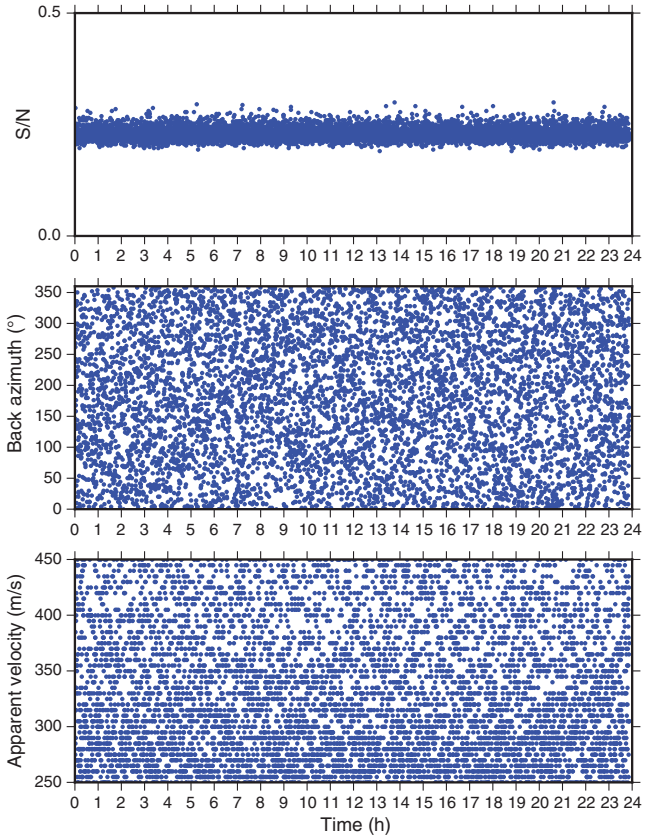

Figure 4. Fisher detector results from processing 24 h of synthetic noise. The frames from top to bottom show S/N, BA, and AV as a function of time.

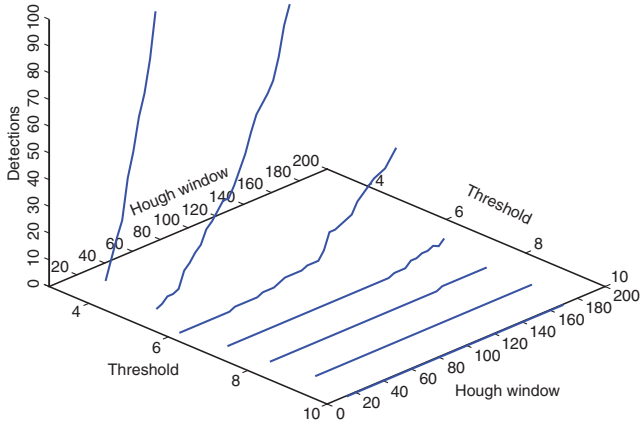

Figure 5. False-alarm rate as a function of HW and threshold. The HT detector is applied to the Fisher results of the synthetic noise. The false-alarm rate is higher if the threshold is lower and the HW is larger.

The calculated probabilities reveal a different behavior than the expected probabilities from the binomial distribution. With the given threshold value and HW size, the probability does not correspond to the theoretical value. We interpret this deviation to be related to the processing of a finite sample of points within a HW, in contrast to a larger population, for which the binomial statistic would hold better. Furthermore, the synthetics undergo filtering and array processing, which is sensitive to the array configuration affecting its response (Evers, 2008). This means that the input for the HT is a subset of the entire data distribution. Therefore, the number of possible BA values is lower and the probability higher. The distribution of the cumulative BA values for one day of synthetic data is presented in Figure 6b. A chi-square fitting is used to find the best associated probability for p_{azimuth}. Figure 6a shows the HT calculated detection probability (blue dots) and theoretical probabilities (black lines) of the binomial distribution for 10 p_{azimuth} values. The best fit provides a mean value of $p_{\text{azimuth}} = 0.0105$ (with significance of 0.024). This means that the BA values are distributed over approximately 100 possible values instead of 180 values that were estimated theoretically.

The HT's threshold and HW values were set to 10 and 14, respectively. This threshold is equivalent to 128 s, which corresponds to a typical time scale of regional signals. Testing these settings on the noise data set provides zero false alarms (Figure 4). The use of these values may lead to missed events for signals shorter than the characteristic time scale. Those can still be retrieved by setting a lower threshold value and accepting a higher false alarm rate.

Geophysical signal durations can vary from a fraction of a second to minutes and hours depending on the type of source and the propagation path (Evers and Haak, 2005; Evers et al., 2007). Therefore, two synthetic data sets are tested. The first data set is a continuous wave with a finite duration. The second data set consists of repetitive, impulsive signals. Both synthetic signals have a spectral content of 1–5 Hz, with a center frequency of 2.5 Hz and a BA of 90° relative to the center of the array.

The first synthetic signal consists of 1000 s of samples with 700 s of coherent signal. It simulates a quasi-continuous, nonmoving source (calving glaciers or volcanic eruption). Column A in Figure 7 shows the beam-forming and HT results for the high-S/N signal. Most of the signal has S/N values higher than one. As stated earlier, the threshold for automatic detectors is an S/N equal or greater than one, which means that the signal will be detected. At 750 s (the tail of the signal), S/N values decrease. This part of the signal would have been missed by an automatic detector. Nevertheless, the HT also picks up the low-S/N part of the signal. In total, the HT detected 54 data points that are related to the signal.

After increasing the noise levels (column B, Figure 7), the signal S/N values are lower than 0.5. Thus, the entire signal would have been missed. Due to the high noise levels, BA values skew and therefore less points are aligned. Consequently, a lower HT detection rate

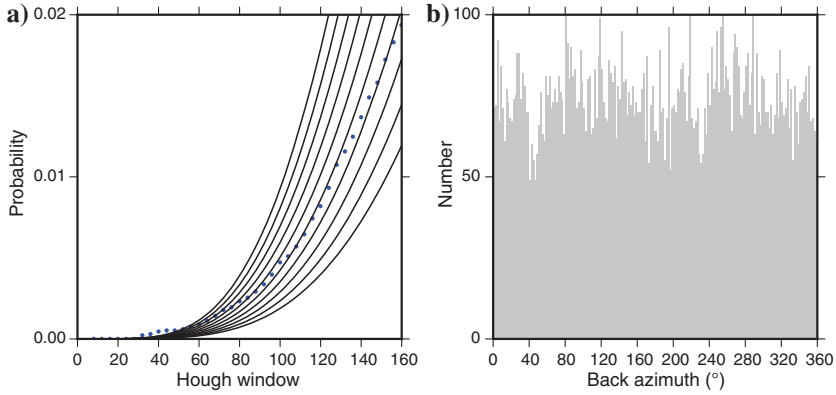

Figure 6. (a) The calculated detection probability from a synthetic noise data set. The blue dots represent the HT detection probability for 24 h of analyzed noise, as calculated from data. The HT threshold value is 5, and HW values range from 8 to 160. The black curves represent the theoretical probabilities of the binomial distribution for 10 p_{azimuth} values between 0.009 and 0.0135. The best fit is found for $p_{\text{azimuth}} = 0.0105$. (b) Distribution of cumulative BA values for one day of synthetic noise. The nonuniform distribution is consistent with our interpretation that the processed data is a subset of the original data.

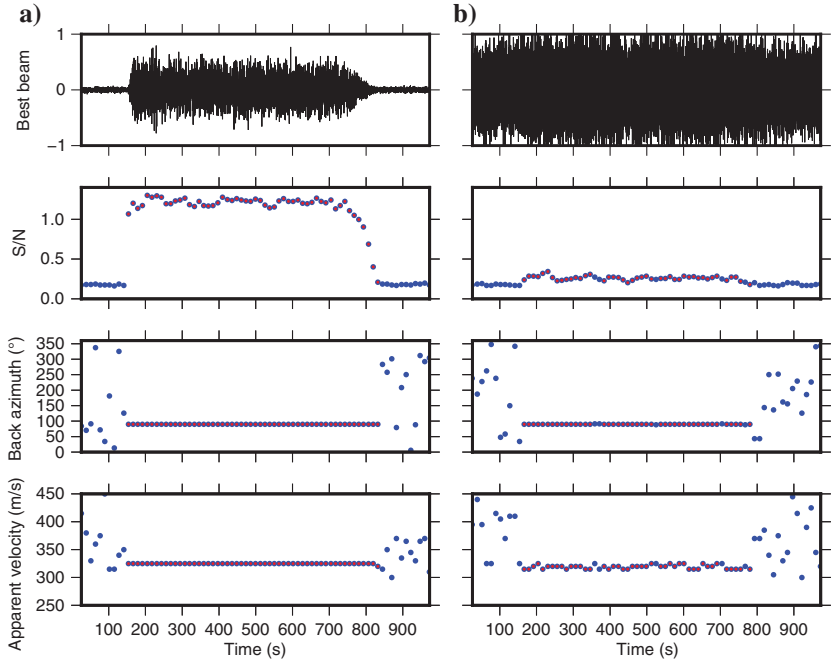

Figure 7. A synthetic long-duration signal with different S/N values. Columns A and B show a high- and a low-S/N signal, respectively. The frames from top to bottom show S/N, BA, and AV as a function of time. The blue and red dots are the Fisher and HT detector results, respectively. For the high-S/N signal, the HT detected 54 data points. The HT detected 44 data points for the low-S/N signal. Due to its low S/N, this signal would have been missed by an automatic detection algorithm that only uses S/N as a threshold.

is obtained. The 44 data points were detected by the HT (81.5% of the high-S/N signal).

A second synthetic signal simulates a repetitive impulsive source. From a seismic prospecting point of view, active seismic areas in reservoirs and mines can act as a repetitive microseismic sources. In mining, when microseismic activity exceed a certain threshold, the mine (or part of it) will be closed for safety reasons (Potvin and Hudyma, 2001). Figure 8 shows 10 repetitive impulsive signals with high and low S/Ns. The time interval between the signals is 20 s. Column A shows the beam-forming and HT detection of the high-S/N signals. The HT detects 16 data points. Column B shows the results for the low-S/N signal. Because all the signal data points (200–400 s) have S/N values approximately 0.5, the signal would have been missed. The HT detects 12 related data points, which is 75% of the high-S/N signal.

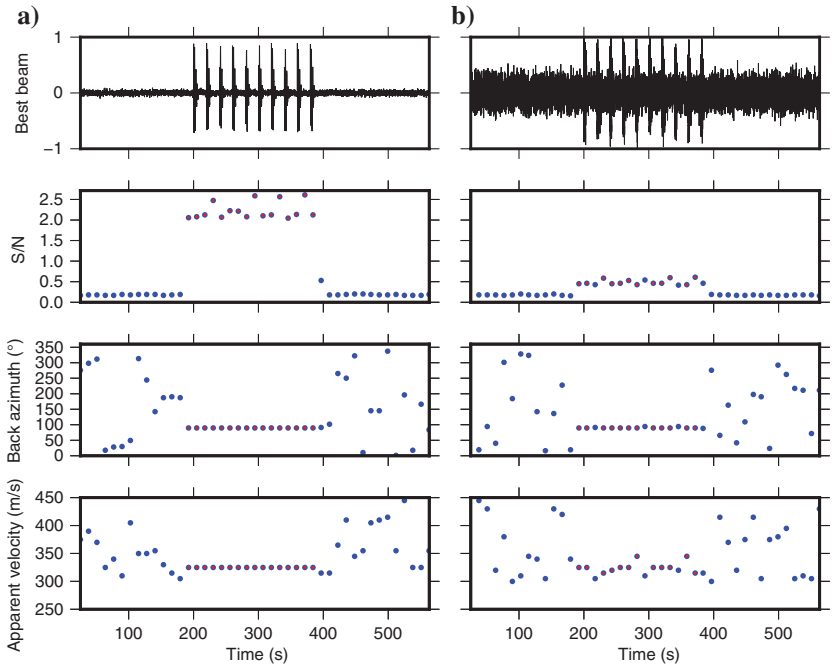

Figure 8. As Figure 7, for repetitive, impulsive signals with different S/N values. For the high-S/N signal, the HT detected 16 data points. The HT detected 12 data points for the low-S/N signal.

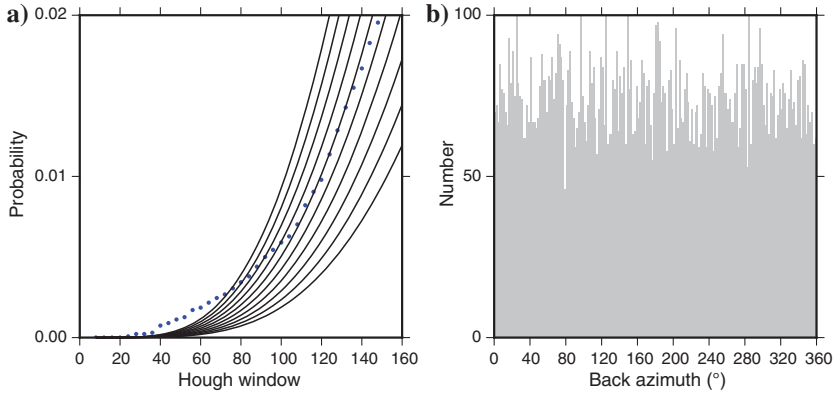

Figure 9. As Figure 6, for the noise data set constructed using measurements at I18DK.

INFRASOUND RECORDINGS FROM GREENLAND

The new method was tested on real data from five years of continuous recordings at IMS infrasound array I18DK, Greenland (Figure 3). The array is surrounded by land- and sea-terminating glaciers that act as active infrasound sources due to calving and glacial run-off. Simultaneous arriving signals from different glaciers cause a decrease in the S/N values because one source acts as noise to another. Therefore, analyzing I18DK's infrasound data set elucidates the merit of using the HT.

To evaluate the false-alarm rate, an incoherent noise data set was constructed from one year (2010) of recordings. This data set is created by randomly attributing recordings from one array element to another array element. Doing so, coherent events are no longer extracted by the Fisher detector because phase differences over the array elements are randomly mixed (Brown et al., 2008).

The distribution of the BA for one day of constructed noise (Figure 9b) reveals a similar behavior to that found for the synthetic noise. The calculated and theoretical detection probabilities are presented in Figure 9a. The best fit is found for $p_{\text{azimuth}} = 0.011$ (with a significance of 0.061). Figure 10 shows the Fisher detector and HT results.

Figure 11 shows the HT detection rate as a function of threshold and HW for 10 days of noise data. The detection rate has similar behavior to the one calculated from the synthetic data (Figure 5). A threshold value of 10 and a 14 data points HW were used as the HT settings. The expected false alarm rate is 576 data points per year. There are approximately 6750 data points per day.

Table 1 summarizes results of analyzing five years of infrasound data. The second column presents the number of automatic detections by considering an S/N threshold value greater than one. The column $H_{S/N<1}$ shows the HT detections of data points with S/N values lower than one. Most of the HT detections have S/N values lower than one (column 3). This column represents the number of missed events in case of not using the HT. A low number of HT detections with S/N values larger than one (see column 4 of Table 1) can be the result of interfering signals from the glaciers surrounding the array. Their presence is causing the Fisher detector to switch between sources because the Fisher detector returns the event parameters of the dominant signal in each time bin. Instead of providing a continuous detection from the same BA, it provides pseudo short-duration signals from different sources. These are too short to be detected by the HT with its current settings.

Figure 12 shows that the BA of the HT detections coincides with the directions of nine glaciers surrounding the array. The glaciers have been identified using Google Earth. During the summer months from August to October, the

infrasound activity is at its highest due to calving and glacial runoff. AV values lower than 300 m/s can be considered to be nonphysical. Such values can be explained by the larger uncertainties in the slowness vector estimation for low-S/N signals.

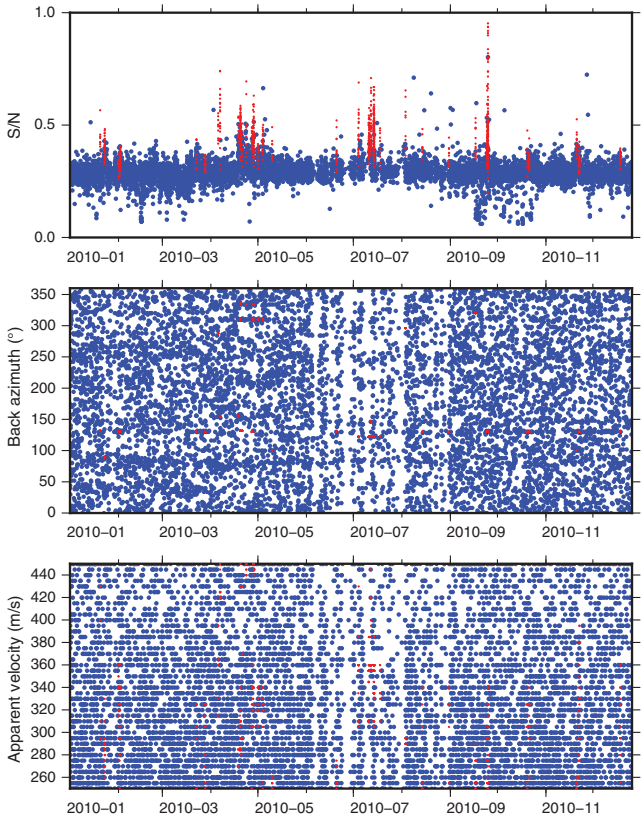

Figure 10. Noise data Fisher analysis and HT results. The frames from top to bottom show S/N, BA, and AV as a function of time. The blue and red dots correspond to the Fisher detector results and HT detections, respectively. The HT detects 576 data points.

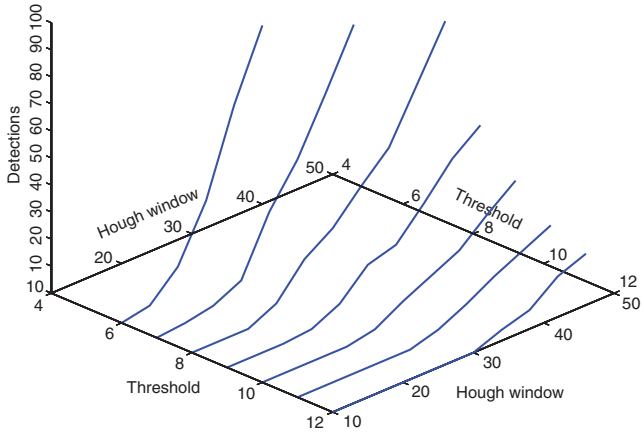

Figure 11. False-detection rates as a function of HW and threshold. The HT detector was applied to 10 days of Fisher detector results of the noise data set. Because the threshold is lower and the HW is larger, the false-alarm rate is higher, as can also be seen in Figure 5.

Table 1. $F_{S/N\geq1}$: number of Fisher detections with S/N ≥ 1. $H_{S/N<1}$: HT detections with S/N < 1. $H_{S/N\geq1}$: HT detections with S/N ≥ 1.

Year	$F_{S/N>1}$	$H_{S/N<1}$	$H_{S/N>1}$
2010	27,139	126,362	13,859
2011	34,667	143,709	25,751
2012	39,554	209,760	29,808
2013	10,096	56,224	3383
2014	41,049	149,715	33,168

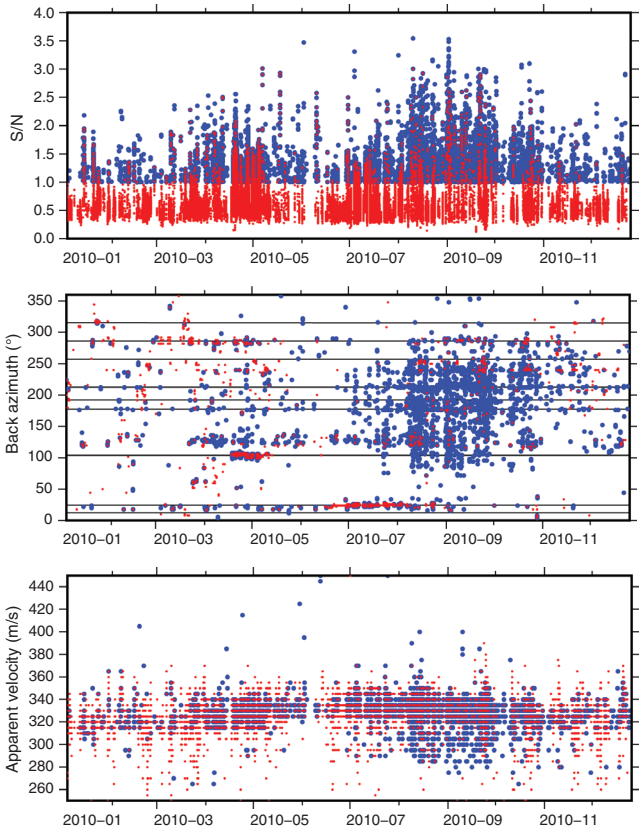

Figure 12. The 2010 Fisher analysis and HT results. The frames from top to bottom show S/N, BA, and AV as a function of time. Fisher detections with S/N ≥ 1 and HT detections are represented by the blue and red points, respectively.

CONCLUSION

Automatic detection algorithms for infrasonic signals are based on an S/N threshold value, and are prone to miss events with S/N values lower than the set threshold value. In this paper, the problem of detecting coherent but low-S/N signals is addressed, to reduce the number of missed events. A statistical framework to tune the parameters in an automatic detection system based on the HT is introduced. This methodology allows for defining suitable settings for the HT by analyzing incoherent noise and assessing the expected false-alarm rate. Even though the method was implemented for

infrasonic signals, it can be adapted to seismic array analysis, using any beam-forming procedure.

Results from the synthetic and real data show the benefit of using the HT in a postprocessing step. Assessing the expected false-alarm rate for synthetic noise allows us to define HT settings with less than one false alarm per day. In addition, these same settings provide correct detections of very low S/N events. The detection capabilities of repetitive impulsive signals are highly dependent on the time intervals between each event, the sampling rate, and the HT settings. In specific cases, the Fisher detector and the HT settings can be optimized for better detection performance, i.e., for monitoring microseismic activity.

Testing the same HT settings on Greenland's noise data provided 576 false alarms per year. For five years of real data, the same settings provided, on average, four times more detections than the S/N-based detector. Obviously, the glaciers are not the only infrasound sources. There are many more natural and anthropogenic signals that can register as sources. Thus, there are also detections that do not associate with glaciers. Some of the detections south of the array could be associated with the movement and breaking of sea ice.

The relatively low number of high-S/N detections could be attributed to the existence of simultaneous sources. These cause the BA values of the high-F detections to switch back and forth instead of "locking" to a stable direction. The threshold value determines the minimum signal length that can be detected. As the desired signal gets shorter in time, the threshold value gets lower and the expected false alarm rate gets higher even when using a more strict detection criterion (small HW). The expected low false alarm rate combined with the alignment of the HT detections with the known directions of the glaciers illustrates the benefit of the proposed technique.

Recalling the need to detect low-S/N events in various fields of seismology, the proposed method can increase their detection. For body-wave interferometry, detecting more time segments that contain low-S/N body waves arrivals can provide a better illumination of the array. Although containing low-S/N events, it will improve the interferometry results. Furthermore, the HT can help detect active microseismic sources in mines and reservoirs. Assuming that one can determine a characteristic time length for microseismic events, it is possible to calculate a corresponding number of data points from the beam-forming stage (the number of data points can vary according to the array-processing settings) and therefore, using the HT, can evaluate the number of microseismic events for a time period.

Although using the HT decreased the number of missed events, there are still other active sources that are missed. As a default, the Fisher detector picks only the highest S/N event per time bin. In the described case study, we are aware that there are multiple active sources, but only one is being detected. Future work will focus on the application of this method to a version of the Fisher detection algorithm that can process multiple sources.

ACKNOWLEDGMENTS

The authors thank the CTBTO and station operators for the high-quality IMS data and products. IMS data can be accessed through the vDec (see https://www.ctbto.org/specials/vdec/). All figures have been created using Generic Mapping Tools (Wessel et al., 2013). G. Averbuch is funded through the Marie Curie Action WAVES from the European Union within H2020, grant 641943. P.S.M. Smets is funded through the ARISE2 project from the European Union within H2020, project 653980. L.G. Evers' contribution is partly funded through a VIDI project from the Dutch Science Foundation (NWO), project 864.14.005. The authors are grateful for helpful reviews by D. Draganov, S. Shani Kadmiel, the editor, and the anonymous reviewers.

REFERENCES

Arrowsmith, S. J., R. Whitaker, S. R. Taylor, R. Burlacu, B. Stump, M. Hedlin, G. Randall, C. Hayward, and D. ReVelle, 2008, Regional monitoring of infrasound events using multiple arrays: Application to Utah and Washington State: Geophysical Journal International, **175**, 291–300, doi: 10.1111/j.1365-246X.2008.03912.x.

Assink, J. D., R. Waxler, W. G. Frazier, and J. Lonzaga, 2013, The estimation of upper atmospheric wind model updates from infrasound data: Journal of Geophysical Research Atmospheres, **118**, 707–710.

Boué, P., P. Roux, M. Campillo, and B. de Cacqueray, 2013, Double beamforming processing in a seismic prospecting context: Geophysics, **78**, no. 3, V101–V108, doi: 10.1190/geo2012-0364.1.

Brown, D. J., R. Whitaker, B. L. N. Kennett, and C. Tarlowski, 2008, Automatic infrasonic signal detection using the Hough transform: Journal of Geophysical Research, **113**, D17.

Carlson, B. D., E. D. Evans, S. L. Wilson, and M. I. T. Lincoln, 1994, Search radar detection and track with the Hough transform. Part III: Detection performance with binary integration: IEEE Transactions on Aerospace and Electronic Systems, **30**, 116–125, doi: 10.1109/7.250412.

Chambers, K., J. Kendall, S. Brandsberg-Dahl, and J. Rueda, 2010, Testing the ability of surface arrays to monitor microseismic activity: Geophysical Prospecting, **58**, 821–830, doi: 10.1111/j.1365-2478.2010.00893.x.

Draganov, D., X. Campman, J. Thorbecke, A. Verdel, and K. Wapenaar, 2013, Seismic exploration-scale velocities and structure from ambient seismic noise (>1 Hz): Journal of Geophysical Research: Solid Earth, **118**, 4345–4360.

Drob, D. P., J. M. Picone, and M. Garcés, 2003, Global morphology of infrasound propagation: Journal of Geophysical Research: Atmospheres, **108**, doi: 10.1029/2002JD003307.

Duda, R. O., and P. E. Hart, 1972, Use of the Hough transformation to detect lines and curves in pictures: Communications of the ACM, **15**, 11–15, doi: 10.1145/361237.361242.

Elazar, M., 1995, Search radar track-before-detect using the Hough transform: Ph.D. thesis, Calhoun.

Evers, L. G., 2008, The inaudible symphony: On the detection and source identification of atmospheric infrasound: Ph.D. thesis, TU Delft, Delft University of Technology.

Evers, L. G., L. Ceranna, H. W. Haak, A. Le Pichon, and R. W. Whitaker, 2007, A seismoacoustic analysis of the gas-pipeline explosion near Ghislenghien in Belgium: Bulletin of the Seismological Society of America, **97**, 417–425, doi: 10.1785/0120060061.

Evers, L. G., and H. W. Haak, 2005, The detectability of infrasound in The Netherlands from the Italian volcano Mt. Etna: Journal of Atmospheric and Solar-Terrestrial Physics, **67**, 259–268, doi: 10.1016/j.jastp.2004.09.002.

Forghani, F., and R. Snieder, 2010, Underestimation of body waves and feasibility of surface-wave reconstruction by seismic interferometry: The Leading Edge, **29**, 790–794, doi: 10.1190/1.3462779.

Furmanski, C. S., and S. A. Engel, 2000, An oblique effect in human primary visual cortex: Nature Neuroscience, **3**, 535–536, doi: 10.1038/75702.

Ge, M., 2005, Efficient mine microseismic monitoring: International Journal of Coal Geology, **64**, 44–56, doi: 10.1016/j.coal.2005.03.004.

Gibbons, S. J., and F. Ringdal, 2006, The detection of low magnitude seismic events using array-based waveform correlation: Geophysical Journal International, **165**, 149–166, doi: 10.1111/j.1365-246X.2006.02865.x.

Gibbons, S. J., and F. Ringdal, 2012, Seismic monitoring of the North Korea nuclear test site using a multichannel correlation detector: IEEE Transactions on Geoscience and Remote Sensing, **50**, 1897–1909, doi: 10.1109/TGRS.2011.2170429.

Hall, P., N. Tajvidi, and P. E. Malin, 2006, Locating lines among scattered points: Bernoulli, **12**, 821–839.

Harmon, N., P. Gerstoft, C. A. Rychert, G. A. Abers, M. S. de la Cruz, and K. M. Fischer, 2008, Phase velocities from seismic noise using beamforming and cross correlation in Costa Rica and Nicaragua: Geophysical Research Letters, **35**, 1–6, doi: 10.1029/2008GL035387.

Le Pichon, A., E. Blanc, and A. Hauchecorne, 2010, Infrasound monitoring for atmospheric studies: Springer Science & Business Media.

Le Pichon, A., J. Vergoz, E. Blanc, J. Guilbert, L. Ceranna, L. Evers, and N. Brachet, 2009, Assessing the performance of the international monitoring system's infrasound network: Geographical coverage and temporal variabilities: Journal of Geophysical Research: Atmospheres, **114**, D8.

Lezama, J., J.-M. Morel, G. Randall, and R. G. von Gioi, 2015, A contrario 2D point alignment detection: IEEE Transactions on Pattern Analysis and Machine Intelligence, **37**, 499–512, doi: 10.1109/TPAMI.2014.2345389.

Maxwell, S. C., J. Rutledge, R. Jones, and M. Fehler, 2010, Petroleum reservoir characterization using downhole microseismic monitoring: Geophysics, **75**, no. 5, 75A129–75A137, doi: 10.1190/1.3477966.
Melton, B. S., and L. F. Bailey, 1957, Multiple signal correlators: Geophysics, **22**, 565–588, doi: 10.1190/1.1438390.
Nixon, M., 2008, Feature extraction and image processing: Academic Press.
Olson, J. V., 2004, Infrasound signal detection using the Fisher F-statistics: InfraMatics, **06**, 1–7.
Panea, I., D. Draganov, C. A. Vidal, and V. Mocanu, 2014, Retrieval of reflections from ambient noise recorded in the Mizil area, Romania: Geophysics, **79**, no. 3, Q31–Q42, doi: 10.1190/geo2013-0292.1.
Potvin, Y., and M. R. Hudyma, 2001, Seismic monitoring in highly mechanized hardrock mines in Canada and Australia: Keynote Address in the Proceedings of the Fifth International Symposium on Rockburst and Seismicity in Mines (RaSiM 5), 267–280.
Rost, S., 2002, Array seismology: Methods and applications: Reviews of Geophysics, **40**, 1008, doi: 10.1029/2000RG000100.
Shumway, R. H., 1971, On detecting a signal in N stationarily correlated noise series: Technometrics, **13**, 499–519, doi: 10.1080/00401706.1971.10488814.
Smets, P. S. M., J. D. Assink, A. Le Pichon, and L. G. Evers, 2016, ECMWF SSW forecast evaluation using infrasound: Journal of Geophysical Research: Atmospheres, **121**, 4637–4650.
Song, J., and M. R. Lyu, 2005, A Hough transform based line recognition method utilizing both parameter space and image space: Pattern Recognition, **38**, 539–552, doi: 10.1016/j.patcog.2004.09.003.
Verdon, J. P., J. M. Kendall, S. P. Hicks, and P. Hill, 2017, Using beamforming to maximize the detection capability of small, sparse seismometer arrays deployed to monitor oil field activities: Geophysical Prospecting, **65**, 1582–1596, doi: 10.1111/1365-2478.12498.
Vidal, C. A., J. van der Neut, D. Draganov, G. Drijkoningen, and K. Wapenaar, 2011, Retrieval of reflections from ambient-noise field data using illumination diagnostics: 81st Annual International Meeting, SEG, Expanded Abstracts, 1613–1617.
Warpinski, N. R., 2009, Integrating microseismic monitoring with well completions, reservoir behavior, and rock mechanics: SPE Tight Gas Completions Conference, 15–17.
Wessel, P., W. H. F. Smith, R. Scharroo, J. Luis, and F. Wobbe, 2013, Generic mapping tools: Improved version released: Eos, Transactions American Geophysical Union, **94**, 409–410, doi: 10.1002/eost.v94.45.
Williams-Stroud, S., J. Kilpatrick, and B. Cornette, 2010, Moving outside of the borehole: Characterizing natural fractures through microseismic monitoring Sherilyn: First Break, **28**, 89–94.
Yilmaz, Ö., 2001, Seismic data analysis: Processing, inversion, and interpretation of seismic data: SEG.