Detecting events in the urban seismic wavefield using a novel nodal array in Singapore: earthquakes, blasts and thunder quakes

Karen H Lythgoe[1], Aidan Loasby[1], and Shengji Wei[1]

[1]Earth Observatory of Singapore

November 23, 2022

Abstract

Receiver density is key to being able to detect and characterise seismic events at the noise level. This is particularly important in urban environments where high cultural noise levels can obscure seismic event signals at a single station. Here we catalogue the seismicity and describe the basic data features of a dense nodal array that was deployed in the city state of Singapore for a 1 month period in 2019. We utilise array methods to detect and characterise seismic events, the first based on waveform similarity (Li et al 2018) and the second (presented here) on spectral energy. Distant earthquakes are easily detected using the waveform similarity method, but local events are more difficult to detect in this way. We therefore develop a spectrogram stacking approach that highlights the location of anomalous coherent spectral energy. Overall, we identify 76 distant earthquakes and 35 local events. Out of the local events, 22 are determined to be from blasting works, while 13 remain from an origin that we cannot yet determine. We also find that lightning produces a plentiful supply of natural seismic sources through the conversion of acoustic waves propagating through the atmosphere (thunder), to seismic waves. We record hundreds of thunder quakes with a high signal to noise ratio and over a wide frequency range. We suggest that a tropical region such as Singapore has high potential to further advance thunder-quake studies.

Detecting events in the urban seismic wavefield using a novel nodal array in Singapore: earthquakes, blasts and thunder quakes

Karen H. Lythgoe[1], **Aidan Loasby**[1], **Shengji Wei**[1]

[1]Earth Observatory of Singapore, Nanyang Technological University, Singapore

Key Points:

- We develop a new method for event detection and characterization over frequency space using a nodal array in the city state of Singapore.
- High frequency energy is elevated during lightning storms which is due to ground movement from hundreds of thunder quakes.
- During 1 month we detected 76 distant earthquakes and 35 local events, some originating from blasting works and some of unknown origin.

Corresponding author: Karen Lythgoe, karen.lythgoe@ntu.edu.sg

Abstract

Receiver density is key to being able to detect and characterise seismic events at the noise level. This is particularly important in urban environments where high cultural noise levels can obscure seismic event signals at a single station. Here we catalogue the seismicity and describe the basic data features of a dense nodal array that was deployed in the city state of Singapore for a 1 month period in 2019. We utilise array methods to detect and characterise seismic events, the first based on waveform similarity (Li et al., 2018) and the second (presented here) on spectral energy. Distant earthquakes are easily detected using the waveform similarity method, but local events are more difficult to detect in this way. We therefore develop a spectrogram stacking approach that highlights the location of anomalous coherent spectral energy. Overall, we identify 76 distant earthquakes and 35 local events. Out of the local events, 22 are determined to be from blasting works, while 13 remain from an origin that we cannot yet determine. We also find that lightning produces a plentiful supply of natural seismic sources through the conversion of acoustic waves propagating through the atmosphere (thunder), to seismic waves. We record hundreds of thunder quakes with a high signal to noise ratio and over a wide frequency range. We suggest that a tropical region such as Singapore has high potential to further advance thunder-quake studies.

1 Introduction

Analyzing the urban seismic wavefield is important not only for unraveling tectonic and geological features but also for building a smart city. However, within an urban environment, the challenges of seismology are inherently increased by the strength and complexity of the seismic noise. Recent advances in instrumentation now make dense passive seismic surveys in urban areas feasible. The highly centralized and portable seismic nodes allows many instruments to be deployed rapidly, directly into the ground without bulky equipment. The relatively low cost of nodes also allows dense instrumentation. Dense arrays have proved to be efficient in solving many challenges in seismology, including improving seismic event detection by stacking or tracing coherent signals (e.g. Gibbons and Ringdal (2006); Hansen and Schmandt (2015); Meng and Ben-Zion (2018b); Gradon et al. (2019)). There have been few dense passive seismic surveys in urban environments since the first nodal array deployed in Long Beach, California in 2011 (F.-C. L. Lin et al., 2013; Riahi & Gerstoft, 2015).

In this study we focus on detecting and characterising discrete seismic events in Singapore, a densely populated city state. Knowledge of background seismicity is critical for the successful seismic monitoring of future underground developments. This is particularly prescient for land and resource scarce Singapore, where expanding the capacity of the city underground is an imperative reality. Vast underground storage caverns have been built and future subsurface construction plans include a potential geothermal energy plant (Zhao et al., 2002; Zhou & Zhao, 2016). Singapore also has geological faults, including a significant tectonic fault located between granite and the metasedimentary Jurong Group in Bukit Timah (Figure 1, Leslie et al. (2019); Lythgoe et al. (2020)). The seismic activity level of the faults is unknown, however neighbouring geological faults have been reactivated due to post-seismic stresses from large earthquakes at the nearby Sumatra subduction zone (Shuib et al., 2017; Yong et al., 2017). Discrete seismic events also offer the possibility to use such sources for seismic imaging, in locations where it is difficult to use an active seismic source.

To investigate the ambient seismic wavefield of Singapore and to detect seismic events, we deployed 88 seismic nodes across Singapore from Feb 27th to April 7th, 2019. Seismic event detection is a fundamental and routine process in the seismological community and various methods have been developed to maximise the number of events that are detected (e.g. Withers et al. (1998); Gibbons and Ringdal (2006); Yoon et al. (2015); Chamarczuk et al. (2020)). In this noisy urban environment, the traditional single station detection method based on waveform amplitude (Withers et al., 1998) proved to be ineffective. Instead we utilise array detection methods that take advantage of the small inter-station spacing. Here we use two array detection methods. The first method is to measure the waveform similarity between a station and its nearest neighbours to create an array coherence function (Li et al., 2018), which is then used to guide the detection. The second method is based on spectral energy, which detects coherent anomalous energy in spectrograms across the array. We develop the second method to, 1) overcome the need to choose a specific frequency band for waveform analysis, 2) save computational cost and 3) aid event classification.

We find that array detection methods allow us to detect events with amplitudes near to or even below noise levels. We detect seismic events from regional and teleseismic earthquakes, as well as local seismic events, some from surprising sources. Detected anthropogenic events include blasting events from underground construction. One interesting seismic source are thunder and lightning storms, which create impulsive seismic signals across a wide frequency range.

Here, we first describe the array and the basic features of the dataset. We then introduce the array detection methods and catalogue the seismic events, both distant and local, observed during the 1 month observation period using an existing and our new array detection technique. We finish with a discussion of the results, including possible future seismic source analysis and uses of the detected events.

2 Nodal array and basic data features

A nodal array, comprised of 88 5-Hz Fairfield Z-land nodes, was deployed across Singapore for a continuous period from 27th February to 7th April 2019. The aim of the survey was to i) investigate the urban seismic wavefield and detect seismic events and ii) image the subsurface structure, particularly across fault zones (e.g. Lythgoe et al. (2020) and following efforts). Instruments were therefore located across the island, with denser deployments around fault zones (Figure 1). The station spacing ranged from 100 m for deployment across the fault zones, to 8 km for a node deployed on a nearby island. Sites were located in public and private areas, including schools, nature reserves, weather stations, parks and roadsides, and so the sites had a wide range of (normally high) ambient noise levels.

The data recovery rate was over 98% - with data unusable from one station that had no GPS signal for the entire deployment, therefore the clock drift could not be corrected. Basic pre-processing procedures were applied to the data, including automatic correction for clock drift, removal of instrument response and de-trending. Data was resampled from 250 Hz to 125 Hz (62.5 Hz Nyquist Frequency) in order to make data size more manageable for this study. The instruments recorded three-component data, although we use only the vertical component for event detection.

The seismic wavefield in urban areas is dominated by vibrations from anthropogenic sources, for example trains (Green et al., 2017), traffic (Riahi & Gerstoft, 2015), airplanes (Meng & Ben-Zion, 2018a), foot-traffic (Díaz et al., 2017) and construction sites (Albert & Decato, 2017). Anthropogenic sources also dominate in Singapore, as shown by the correlation between ground velocity amplitude and the timing of man-made activity - for example Figure 1c shows the quietest times at a school are at night and during the lunch

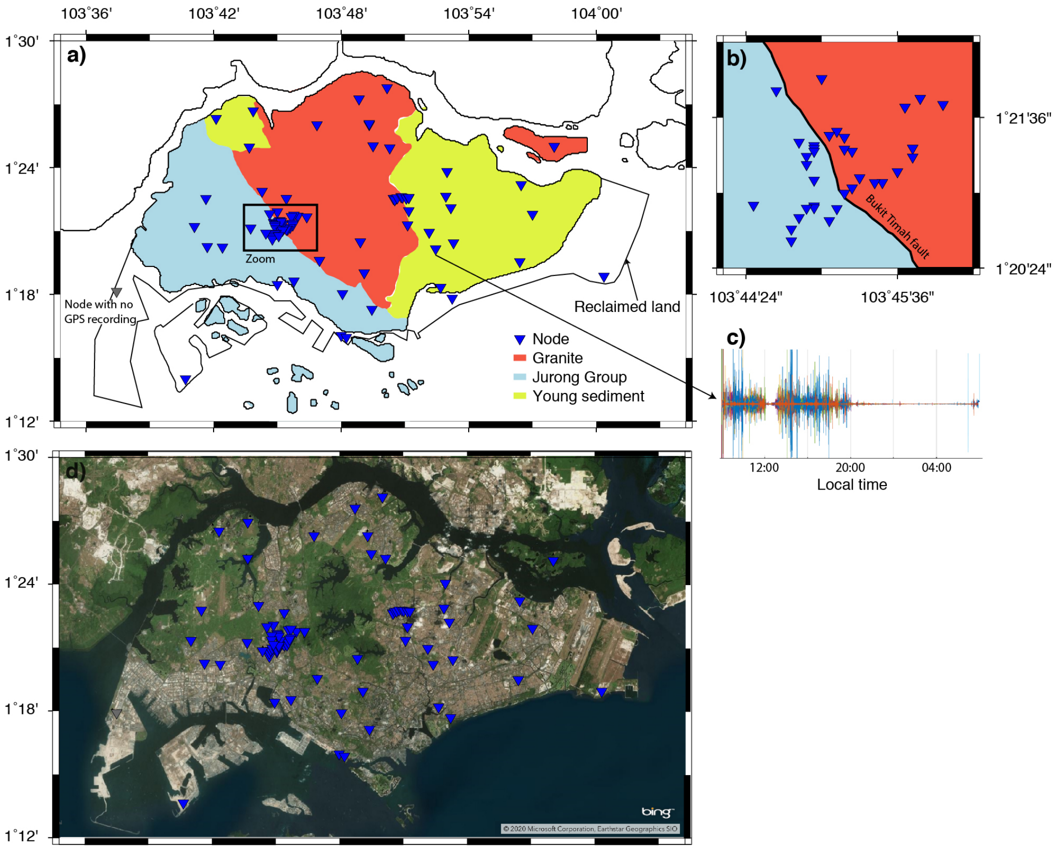

Figure 1. a) A simplified geological map of Singapore showing locations of nodes as blue triangles. White land areas within the Singapore coastline are reclaimed lands. b) Zoom to area of dense deployment across the Bukit Timah fault. c) Seismic data recorded by a node deployed in a school for a period of one day. Quiet times are during the lunch break and between 8pm and 8am. d) A satellite photo of Singapore with location of nodes as blue triangles.

break. At another site, shown in calendar view in Figure 2, spectrograms show clear diurnal variation and a reduction of man-made signal on Sundays (Saturdays are often a part working day in Singapore). It appears that the man-made signals are the highest at frequencies between 2 and 10 Hz. The spectrogram also shows a near constant low frequency energy around 0.1 Hz, which is the background microseismic energy generated by the coupling from ocean waves to the solid earth (Hasselmann, 1963). The rest of the paper focuses on detecting and characterising discrete events within this urban wavefield.

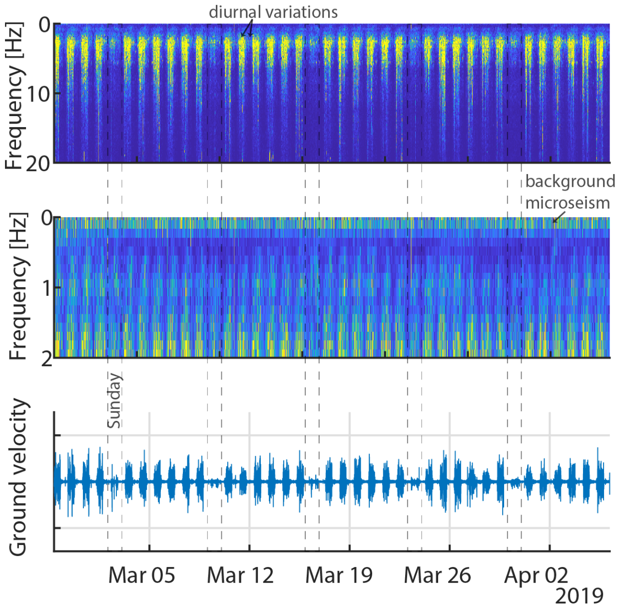

Figure 2. Spectrogram for frequencies less than 20 Hz (top row) and frequencies less than 2 Hz (middle row), plus the waveform (bottom row) for a node deployed in a park, after removing the instrument response.

3 Event detection methodology

To detect discrete seismic events, we first use a single station method and then use two array methods. The traditional single station detection method is based on changes in the short term average (STA) over the long term average (LTA) amplitude at a single station (Withers et al., 1998). In the noisy urban environment, we found this single station detection method to be ineffective, detecting only 8 events over the deployment time (Table 1). Seven of these events were detected during the night time, suggesting the single station method is hampered by day time cultural noise.

Seismic arrays offer the advantage of using the coherency of signal between nearby stations to identify seismic events. We can therefore overcome the limitations imposed by high amplitude cultural noise by exploiting the seismic array. We first apply a detection method based on waveform similarity, as previously suggested by (Li et al., 2018). This method works effectively to detect relatively low frequency signals from distant earthquakes. We then develop a detection method based on stacking anomalous spectral energy and apply it to the data. We find that this method is more effective than the waveform similarity method at frequencies where man-made noise dominates and the output is useful for event classification.

3.1 Array detection method using waveform similarity

Following Li et al. (2018), we generate an array coherence function which detects events when the coherence is greater than a threshold value. The method is based on the principle that waveforms at nearby stations are expected to be very similar for a common source, while noise is sufficiently random. The array coherence function is generated by stacking local 'similarity' functions at each station in the array. The 'similarity' function is the sum of cross-correlation coefficients measured between a station and its neighbours in a moving time-window. A time shift is allowed to obtain the maximum correlation coefficient, in order to account for small travel time differences between stations. Thus, the total array coherence is a measure of the waveform similarity between neighbouring stations. Figure 2 in Li et al. (2018) shows a representative workflow.

The detection threshold is set as the median amplitude in a sliding time window plus ten times the median absolute deviation (Li et al., 2018). We use a maximum distance of 4 km between each station and its neighbours. We examine two frequency ranges, 0.5 - 3 Hz and 5 - 10 Hz, with the aim to detect distant and local events respectively. A sliding window of 3 seconds and 1 second is used for the low and high frequency ranges respectively, with windows having 50% overlap with the previous time window. We find that the 5-10 Hz coherence function is too noisy to enable clear detection. We note that closer station spacing may be required to detect events in the higher frequency range using this method.

One benefit of this method is the ability to approximately locate local seismic sources using the time lags from cross-correlations. We do this in a grid search method by i) calculating travel times for each point on the grid using a 1D velocity model; ii) converting the travel times to lag times between master and neighbouring stations; iii) extracting cross-correlation coefficients at the corresponding lag times; iv) stacking all cross-correlation coefficients, such that each grid point corresponds to a stacked correlation value. In this way we define the best location as the location that has the highest correlation value. We use a 1D velocity model for Singapore calculated from the joint inversion of receiver functions and surface waves (Macpherson et al., 2013). In order to minimise the dependence on the velocity model, we take the highest cross-correlation coefficient from the neighbouring 5 samples.

A drawback of the waveform similarity method is that it is restricted to the frequency range selected. A poorly chosen filter will miss important signals. This problem is exacerbated in urban areas, where the use of a wide band-pass filter is likely to capture significant cultural noise. It remains possible to run the detection algorithm multiple times across a variety of frequency bands, however the computational expense makes this impractical. The high frequency range is also limited by the minimum station spacing - if stations are not sufficiently close then high frequency local events will not have similar waveforms. Finally, detected signal must travel at an apparent velocity greater than the minimum moveout speed, which is set by the maximum lag allowed in cross-correlation. Signal that is travelling slower than this minumum moveout speed will be missed.

3.2 Array detection method using anomalous and coherent spectral energy

To overcome the limitations above, we develop a method that identifies array coherent anomalous spectral energy. By searching the full frequency space, we require no a priori assumptions of the expected form of seismic source. The result of the algorithm contains an added dimension (frequency) compared to the waveform similarity method, providing information that greatly aids event classification. Based upon the Short-Time Fourier Transform (implemented in MATLAB), the method is fast; analysing one day of seismic data (88 stations) in approximately 5 minutes using one CPU.

A schematic of our workflow is shown in Figure 3. We begin by calculating the spectrogram for each station. Spectrograms are calculated with windows of 1.6 s with 50% overlap and plotted in 0.25 Hz frequency bins. We then perform outlier decomposition, where each individual spectrogram is decomposed into a binary image where a '1' represents a pixel of anomalously high energy and '0' is a pixel below a certain threshold. The threshold is defined as 1 median absolute distributions (MAD) above the median. We use MAD and median statistics as they are less sensitive to extreme outliers than mean and standard deviation. Thresholding is performed on each frequency row individually, using a moving time window comparing each pixel to the median amplitude in

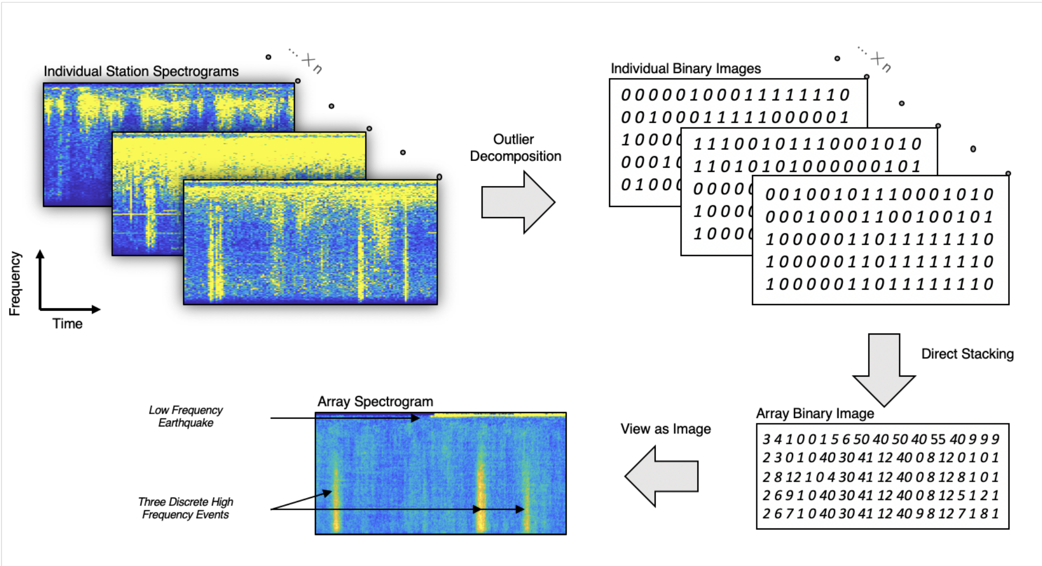

Figure 3. Workflow showing our methodology to detect anomalous, coherent spectral energy.

the surrounding 1 hour. Events can then be automatically detected based on finding anomalous 'islands' within each binary image (for example by setting a minimum number of connected '1' pixels at each station, and then requiring a minimum number of stations). However we find it more instructive to create images, that we term 'array spectrograms', which show the time and frequency of anomalous and coherent energy across the array. The array spectrogram is formed by stacking the binary images for each station, such that the final amplitude is a measure of the number of stations that have anomalous energy at that pixel. Thus high amplitudes indicate that many stations within the array detect anomalously high energy at that time and frequency. Viewing the array spectrogram as an image gives a useful overall view of the frequency and amplitude of coherent signal within the array (Figure 3). Using this approach we detect both distant and local events. In the following discussion we initially used waveform coherence to detect events and followed this by spectrogram stacking. For distant events, we rely waveform coherence and verify the events using the spectrogram approach. For local events, the waveform coherence did not produce clear detections and so we rely on spectrogram stacking.

Table 1. Comparison of earthquake detection between traditional single station method and array waveform similarity method

Earthquake type	Single station detection (STA/LTA)	Array detection by waveform similarity
Regional	6	42
Teleseismic	2	23
Previously Unreported	1	11

4 Distant earthquake detection

We first discuss the detection of distant earthquakes since they are clearly detected using both array methods. Singapore lies on the relatively stable Sunda continental shelf, however it is surrounded by active subduction zones and regularly experiences shaking from earthquakes at the closest subduction zone in Sumatra (Pan & Sun, 1996). During our 40-day deployment time, we detected a total of 76 regional and teleseismic earthquakes, some of which were unreported in global catalogues (Figure 4 and Table 1).

Seismic signals from regional and teleseismic earthquakes are dominated by relatively low frequency energy. For instance, earthquake signals from Sumatra have energy up to 10 Hz, while more distant earthquakes have dominant energy at even lower frequencies (Figure 4). At these low frequencies, the amplitude of the man-made noise is low, hence earthquake waveforms exhibit high coherence between nearby stations. As a result, distant earthquakes are relatively easy to detect using our array similarity function, even in the middle of the day when the cultural noise level is high (Figure 4). We detect 11 distant earthquakes that are unreported in global and regional catalogues (Table 1), however we do not attempt to locate all of these events here due to the small aperture of our array. Figure 4c shows an example event that is unreported. Here we use the azimuth, determined from array moveout, and P-S differential time to approximately locate this event to Sumatra. Supplementary Video 1 shows the long-period seismic waves of an event in Sumatra travelling through Singapore from west to east.

5 Local event detection

Local seismic events have a different character to regional and teleseismic earthquakes. These events produced lower amplitude seismic signals, yet retain high frequency

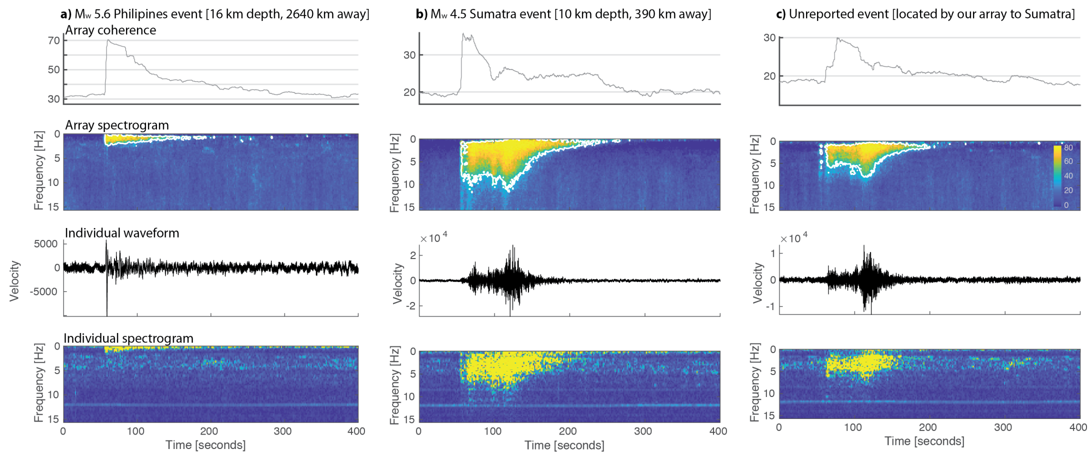

Figure 4. Examples of distant earthquakes detected by our array - a) an earthquake in the Philippines, b) an earthquake in Sumatra, c) an earthquake not reported in global earthquake catalogues, which we locate using the Singapore array to Sumatra. Panels from top to bottom: array coherence function; array spectrogram; waveform at one station (raw); corresponding single station spectrogram for the waveform.

energy (Figure 5), indicating that they are small magnitude local events. Figure 5 shows an example of two events detected within 15 minutes of each other. The low amplitude of the events compared to background noise makes them difficult to detect using the waveform similarity function. However their anomalous spectral content across the array makes them identifiable on the array spectrogram.

To locate the events, we use the cross-correlation lag times already calculated for the similarity function (Supplementary Figure 1). Both events are located in the northeast of Singapore at a known construction site. Figure 6 shows the location and apparent moveout for Event 2 in Figure 5. The characteristic of two events occurring close in time is typical of blasting patterns at this site. The timing in the early evening, is also a typical characteristic of blasting, since permits are normally given for approximately 5-6 pm once the site is clear and workers have left. Therefore we conclude that these events are from construction site blasting works.

In total we detect 22 local events that we determine to be from blasting works. These are characterised as blasting events based on their location and character, such as in the events above, or whether they are in a blasting catalogue that we have available to us

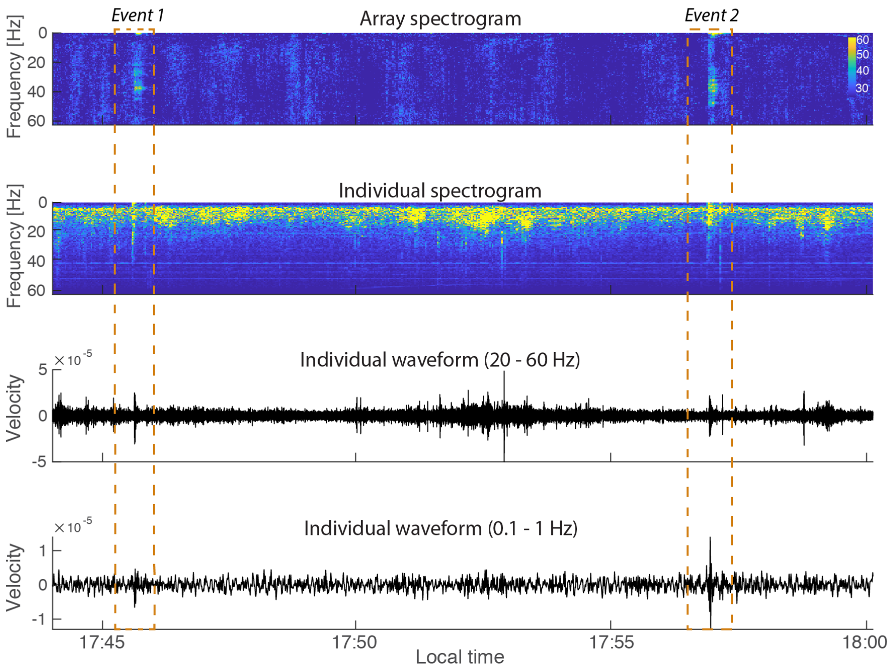

Figure 5. Two local events on 3rd March 2019 marked by orange boxes. Panels from top to bottom: array spectrogram; spectrogram at one station; waveform at same station filtered 20 - 60 Hz; same waveform filtered 0.1 -1 Hz.

from several sites in Singapore. The origin of the remaining events are still to be determined. Several of these unknown events occur during the night and so we do not expect them to be man made.

6 Thunder quakes

During event detection we identified hundreds of short-duration, impulsive signals that have a characteristic high spectral energy across all frequencies. These signals usually occur in bursts, most frequently between 2-6 pm and commonly repeat every few minutes. Figure 7 shows an example of these signals over a 30 minute period, during a thunder and lightning storm. We term these signals thunder quakes for reasons that follow. Lightning is a discharge of electricity, which induces a shock wave that converts electrical energy to acoustic energy (which humans hear as thunder). Holmes et al. (1971)

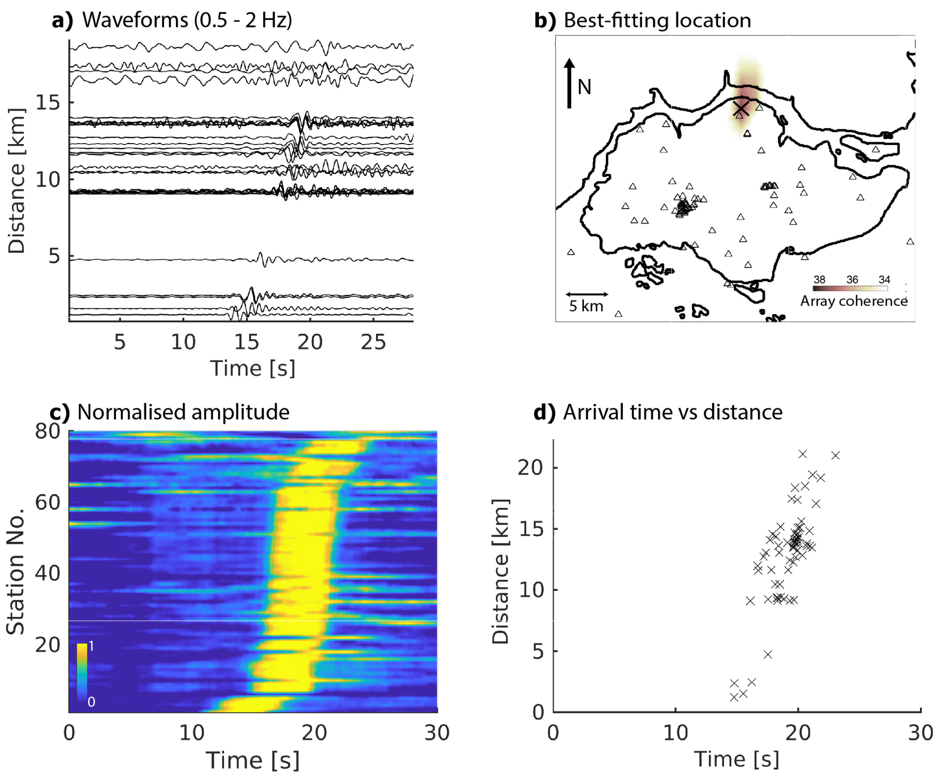

Figure 6. Local event on 3rd March 2019 (Event 2 in Figure 5). a) Waveforms recorded at stations in array, filtered 0.5 - 2 Hz. b) Map of Singapore with best-fitting location of source (cross) obtained by grid search of lag times. c) Moving mean of the absolute amplitude in 0.5 - 2 Hz frequency band vs stations ordered in distance from the best-fitting location. d) Move-out across the array, where distance is from the best-fitting location. Crosses mark the time of maximum amplitude from c).

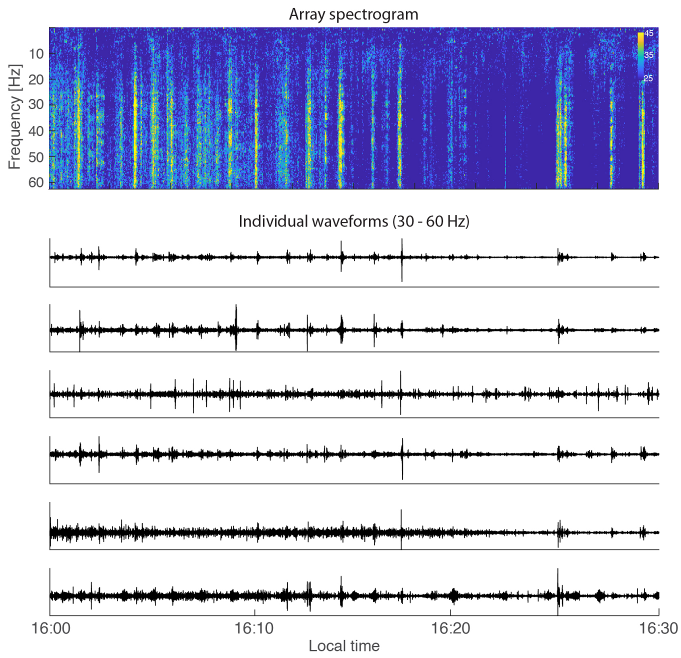

Figure 7. Seismic observations during a lightning storm on 11th March 2019. Top panel: array spectrogram. Bottom panels: representative waveforms at randomly selected stations around the island, filtered 30 - 60 Hz.

showed that microphone recordings of acoustic waves induced by lightning have frequencies ranging from 4 to 125 Hz, consistent with our observations.

Singapore has one of the highest incidents of lightning strikes in the world, with an average of 184 lightning days a year (Meteorological Service of Singapore, 2020). In March 2019 alone, there was over 6000 cloud to ground lightning strikes in Singapore (Figure 8, Meteorological Service of Singapore (2020)). We compare the number of lightning strikes in Singapore over time, with the amplitude of high frequency seismic energy in Figure 8. High frequency seismic energy is calculated as the ratio of the average energy in the 40-60 Hz band to the total seismic energy at less than 60 Hz. Lightning is reported by the Meteorological Survey of Singapore's Lightning Detection System, which is a network of 4 lightning sensors located island wide. We find a clear positive correlation between the number of lightning strikes and the percentage of high frequency energy. In general, more lightning strikes creates more relative high frequency energy. The scaling is also dependent on the time of day at which the storms occurred due to changes in relative ambient noise levels. Lightning in Singapore occurs most frequently between

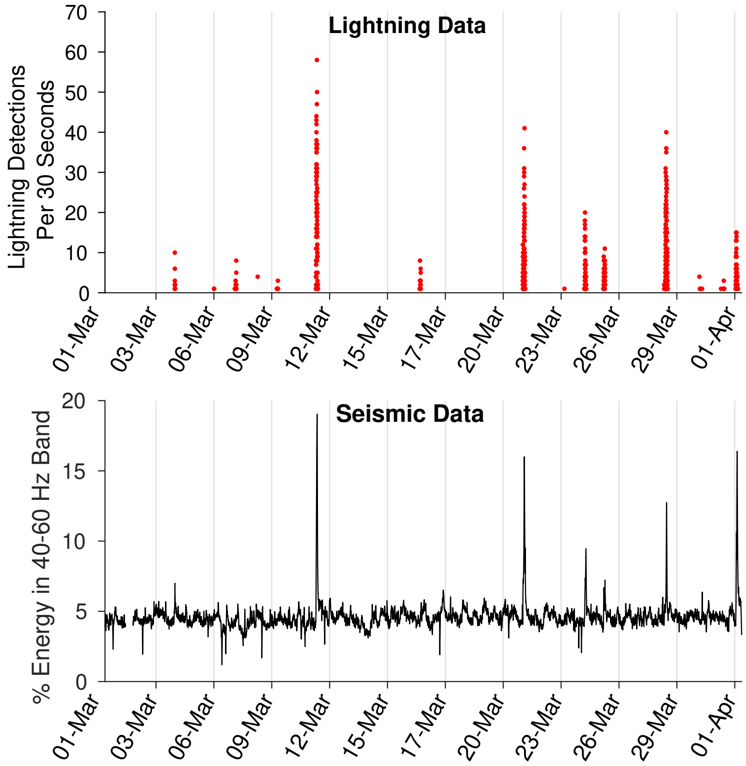

Figure 8. Correlation between high frequency seismic energy and frequency of lightning detected in Singapore. Seismic energy is calculated as the ratio of the average energy in the 40-60 Hz to the total seismic energy less than 60 Hz averaged for the whole array. Lightning strikes are cloud to ground lightning detected by Singapore's lightning detection system.

2 and 6 pm due to the generation of storms by diurnal heating. However a night-time lightning storm on 1st April is evident from the larger percentage of high frequency seismic energy for the number of lightning strikes detected, due to quieter ambient seismic energy levels at night. The rare occurrences of night-time lighting storms are particularly valuable since they occur when cultural noise is a minimum and we use this for more detailed analysis of an individual event.

Figure 9 shows an example of a teleseismic earthquake arriving in Singapore during a lightning storm. The earthquake's seismic waves are dominant at low frequencies (0.1 - 1 Hz). The signal from the earthquake is difficult to see in the spectrogram of an individual station, however they can be clearly identified on the array spectrogram. We note that for the thunder signals, the stacked signal from the entire array shows less low frequency energy (< 10 Hz) compared to higher frequencies, while the single station shows

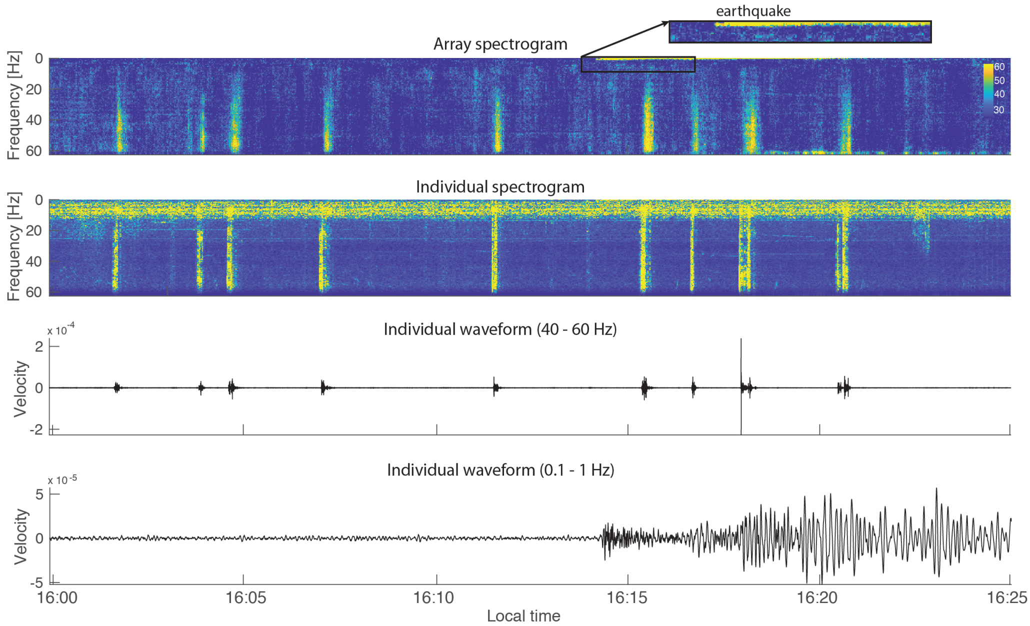

Figure 9. The concurrent arrival of seismic waves from an earthquake in Sumatra and a series of discrete lightning quakes. Seismic waves from the earthquake arrive at low frequencies, while the lightning quakes are across all frequencies.

fairly uniform energy at almost all frequencies. This is because the cultural noise peaks at $\sim$ 10 Hz, therefore, the MAD we used for one-bit transform suppress the real signal along with the noise.

One thunder event occurring at night is shown in Figure 10. The high frequency waveforms for the event are arranged by distance from the estimated source location in Figure 10a. We estimate the source location using a 3D grid search of hand picked first arrival times (Figure 10b). We also attempt to locate the source based on a grid search of differential lag times, as used to locate a local blasting event above (we re-calculate the coherence functions with a longer lag time to account for the slower moveout velocity). However the best-fitting location from this method is incompatible with the first arrival times since the effectiveness of the waveform similarity method is limited due to low waveform coherence (Supplementary Figure 2). The energy of the event can be traced across the array (Figure 10c) although it is difficult to pick first arrivals at all stations - stations with a robust pick are shown as triangles in Figure 10b. The apparent moveout velocity of the event is 350 m/s (Figure 10d). We therefore use a constant velocity of 350 m/s in the grid search location and find the best-fitting location to be at an elevation of 4 km. Given that the moveout velocity is similar to the speed of sound, and that the source is elevated by 4 km, we conclude that the seismic signal is due to the conversion of energy from an acoustic wave in the atmosphere to an elastic wave in the ground.

Figure 11 shows the waveforms for the thunder event at all three components on several nearby stations that have relatively clear first arrivals. There is a long coda after the first arrival, which could be a combination of surface waves and trailing acoustic waves from subsequent thunder claps. Although the waveforms are aligned on their first arrivals, they are no coherent envelopes or spikes in the data that can be traced as sub-events. This is true even for stations spaced 100 m apart as shown in Supplementary Figure 3. The highly different waveforms shows that local site effects are playing an important role in modulating the signal. For instance, the responses of nearby buildings to the acoustic waves can generate strong seismic signals (Kanamori et al., 1991). Near surface conditions can also play an important role in shaping the signal given the very high frequency nature of the source. There is an interesting azimuthal variation of amplitude, with stations to the east having lower amplitudes than stations to the west of the source. This may be due to atmospheric conditions such as prevailing wind, or be caused by a change in air-ground coupling due to different geological units.

7 Discussion

The traditional single station detection method (Withers et al., 1998) proved to be ineffective in this urban environment. However employing array detection techniques allowed us to detect events with amplitudes near to or below noise levels. Overall, we identify 76 distant earthquakes in the recording time (Figure 12). Distant earthquakes are easily detected using the waveform similarity method and they are characterised as having coherent low frequency content. We match our detected events to global event catalogues, however several are not present in global catalogues and are likely from regions close to Singapore, such as Sumatra.

Local events are more difficult to detect using the waveform similarity method, which is due to their higher dominant frequencies. Higher frequency signal has lower waveform coherence since stations are more than 1 wavelength from each other (for example for a dominant frequency of 10 Hz and velocity of 2500 m/s, one wavelength is 250 m, which is lower than the average station spacing) and also cultural noise has greater amplitude.

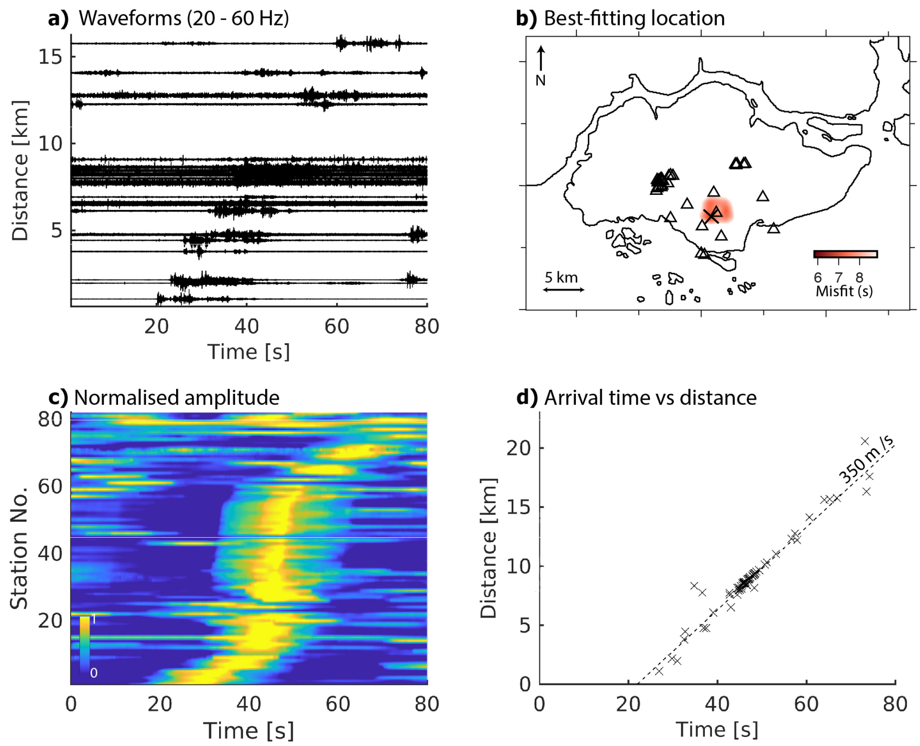

Figure 10. Seismic recording of a thunder quake. a) Waveforms recorded at stations in array, filtered 20 - 60 Hz. b) Map of Singapore with best-fitting location of source (cross) obtained by grid search of picked first arrival times. c) Moving mean of the absolute amplitude in 20 - 60 Hz frequency band vs stations ordered in distance from the best-fitting location. d) Moveout across the array, where distance is from the best-fitting location. Crosses mark the time of maximum amplitude from c). Dashed line corresponds to a moveout velocity of 350 m/s.

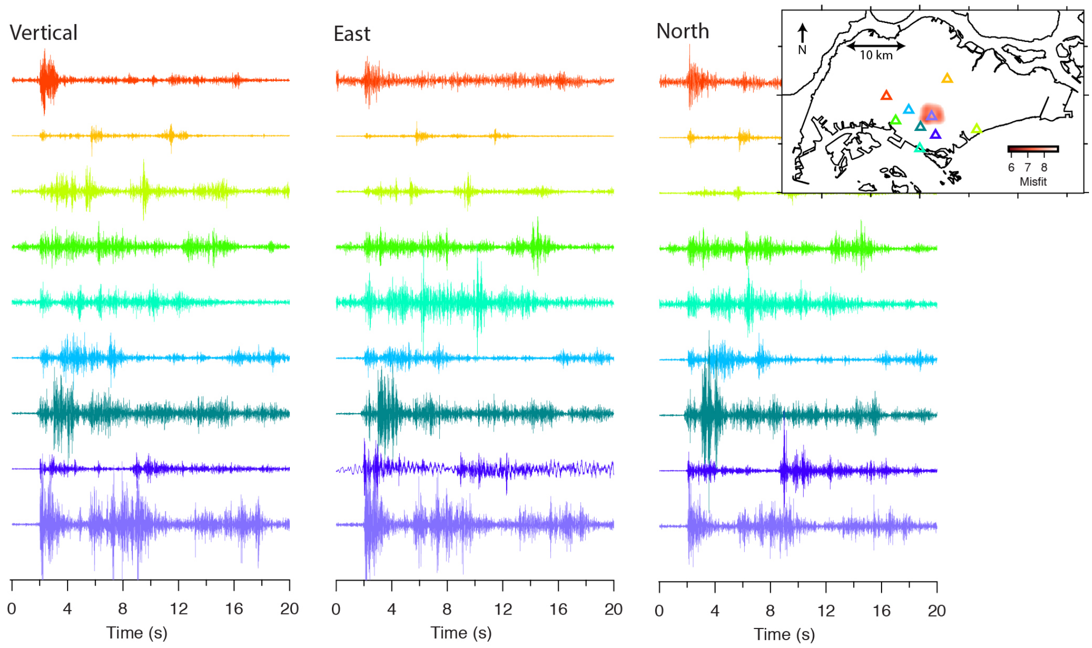

Figure 11. Waveforms induced by thunder at 9 nearby stations. Waveforms are aligned at their first arrival, bandpass filtered 20 - 60 Hz and are not normalised.

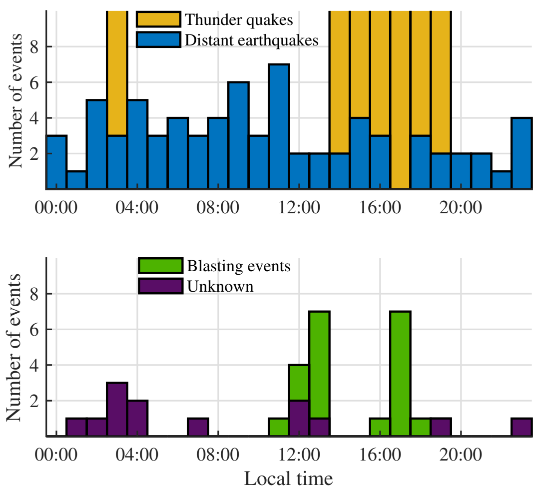

Figure 12. Total number of events per hour detected in each event classification. For comparison purposes the number of events is capped at 10, although hundreds of thunder quakes are detected (we do not count them all).

The optimal solution to detect local seismic events with this method may be to have several arrays of dense stations (100 m station spacing) across the city. We therefore use a spectrogram stacking approach and find 35 local events with this method (Figure 12, detected local event catalog in Table S1). We determine whether these events are from blasting, by first cross-referencing with a blasting catalogue from several sites that we have available to us. We also characterise events to be from blasting if they occur at times when blasting is permitted (generally 1pm and 5pm) at known construction sites.

Within this 1 month period we did not detect local earthquakes, although there are 13 local events whose origin we cannot determine. The timing and location of several of these events indicates that they are likely not from a blasting source (Figure 12). More work is needed to characterise these unknown events, for example moment tensors would help to indicate their source characteristics. Refined locations are also necessary and methods such as back projection should produce more reliable locations. Machine learning also offers promise as a way to distinguish between earthquakes and blasting sources (Miao et al., 2020).

We have shown that lightning produces a plentiful supply of natural seismic sources through the conversion of acoustic waves propagating through the atmosphere (thunder), to elastic waves in the ground. We record thunder with a high signal to noise ratio and over a wide frequency range, which makes it a high quality seismic source. Some of the stations show clear first arrivals (Figure 11), which provide a chance to locate the origin of the thunder, given the condition that the acoustic wave speed in the atmosphere does not change in space. Azimuthal variations in amplitude (Figure 11) indicate that there may be propagation effects due to atmospheric conditions such as wind and temperature. Particle motion analysis for a station located ~ 8 km horizontally from the source, shows retrograde elliptical motion, indicating Rayleigh waves at this station (Supplementary Figure 4). T.-L. Lin and Langston (2007) previously used a combination of a surface and borehole seismometer at a site in the USA to show that an atmospheric wave from thunder can produce locked Rayleigh waves, with the energy trapped in a thin low velocity near surface soil layer because the base layer wave velocity is larger than the speed of sound in air. T.-L. Lin and Langston (2009b) subsequently extracted phase velocities from the induced Rayleigh waves to constrain the near surface velocity structure. Thunder may therefore be a ubiquitous source for near surface structure and site response studies in Singapore.

It is unclear whether the first arriving seismic energy from thunder is produced at the site or at some distance away. Incident slowness differences between acoustic pressure and vertical ground velocity at one site, prompted T.-L. Lin and Langston (2009a) to suggest that the seismic waves initiated away from the station. Additionally Kanamori et al. (1991) showed that P-waves can be generated by motion of high-rise buildings due to an atmospheric shock wave generated by a space shuttle, and that these P-waves can arrive before the shock wave at some seismic stations. The moveout of the thunder shown in Figure 10 is similar to the speed of sound in air, indicating that the first arriving energy is from air-coupled waves at most stations. Zhu and Stensrud (2019) find different moveout velocities for 18 thunder events recorded along a DAS cable, which they suggest is due to a mixture of thunder generated from cloud-to-cloud lightning and cloud-to-ground lightning. The event we analyse in Figure 10 is likely from a cloud-to-cloud source at a high elevation. Together with spectral analysis and detailed location methods such as back-projection, our dataset could be used to refine the source properties of thunder in the future. We may also be able to differentiate between the categories of lightning (for example cloud to cloud and cloud to ground) and elucidate how atmospheric weather couples with the solid Earth. Singapore is likely one of the best places for such studies, given it is near the equator and has very frequent thunder and lightning storms.

8 Conclusions

We show that a dense nodal array can record a rich dataset with only 1 month of observation in an urban environment. Such receiver density allows us to detect and characterise events at the noise level. We utilise new methods to detect and characterise events using arrays, the first based on waveform similarity (Li et al., 2018) and the second (presented here) on spectral energy. Events detected originate from distant earthquakes, man-made blasts and thunder from lightning strikes. Further work on subsurface imaging beneath Singapore will employ a variety of seismic sources recorded in this period, including distant earthquakes (for example receiver functions, Lythgoe et al. (2020)), ambient noise retrieved surface waves and possibly local seismic sources as identified here.

Acknowledgments

We thank the following Singapore agencies for assistance in data acquisition: Meteorological Service of Singapore, NParks, Ministry of Education and the Public Utilities Board.

We thank the Centre for Geohazard Observation at NTU, along with all other field assistants for their help with the field survey, particularly Jeffrey Encillo, Choong Yew Leong, Weiwen Chen and Mark Lim. Due to the sensitive nature of the data, not all of it can be shared. We make available datasets of the examples and events presented here which will be accessible from the NTU data repository (DOI to be provided). Data for review purposes is accessible via this link: `https://entuedu-my.sharepoint.com/:f:/g/personal/karen_lythgoe_staff_main_ntu_edu_sg/EtIpEEWWS8VGvBYp7CguTjkBhpNYbFfDjNhjnQ9va3_gXQ?e=TuZdIu`. This work comprises Earth Observatory of Singapore contribution no. XXX. This research is partly supported by the National Research Foundation Singapore and the Singapore Ministry of Education under the Research Centres of Excellence initiative (Project Code Number: XXXX). Karen Lythgoe is funded by an NTU Presidential Postdoctoral Fellowship.

References

Albert, D. G., & Decato, S. N. (2017). Acoustic and seismic ambient noise measurements in urban and rural areas. *Applied Acoustics*, *119*, 135–143.

Chamarczuk, M., Nishitsuji, Y., Malinowski, M., & Draganov, D. (2020). Unsupervised learning used in automatic detection and classification of ambient-noise recordings from a large-n array. *Seismological Research Letters*, *91*(1), 370–389.

Díaz, J., Ruiz, M., Sánchez-Pastor, P. S., & Romero, P. (2017). Urban seismology: On the origin of earth vibrations within a city. *Scientific reports*, *7*(1), 1–11.

Gibbons, S. J., & Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation. *Geophysical Journal International*, *165*(1), 149–166.

Gradon, C., Moreau, L., Roux, P., & Ben-Zion, Y. (2019). Analysis of surface and seismic sources in dense array data with match field processing and Markov chain Monte Carlo sampling. *Geophysical Journal International*, *218*(2), 1044-1056.

Green, D. N., Bastow, I. D., Dashwood, B., & Nippress, S. E. (2017). Characterizing broadband seismic noise in central london. *Seismological Research Letters*, *88*(1), 113–124.

Hansen, S. M., & Schmandt, B. (2015). Automated detection and location of microseismicity at mount st. helens with a large-n geophone array. *Geophysical Research Letters*, *42*(18), 7390–7397.

Hasselmann, K. (1963). A statistical analysis of the generation of microseisms. *Reviews of Geophysics*, *1*(2), 177–210.

Holmes, C., Brook, M., Krehbiel, P., & McCrory, R. (1971). On the power spectrum and mechanism of thunder. *Journal of Geophysical Research*, *76*(9), 2106–2115.

Kanamori, H., Mori, J., Anderson, D. L., & Heaton, T. H. (1991). Seismic excitation by the space shuttle columbia. *Nature*, *349*(6312), 781–782.

Leslie, G. A., Dodd, T. J., Gillespie, M. R., Kendall, R. S., Bide, T. P., Kearsey, M. R., Timothy I.and Dobbs, ... Chiam, K. (2019). Ductile and brittle deformation in singapore: A record of mesozoic orogeny and amalgamation in sundaland, and of post-orogenic faulting. *J. of Asian Earth Sciences.*, *180*.

Li, Z., Peng, Z., Hollis, D., Zhu, L., & McClellan, J. (2018). High-resolution seismic event detection using local similarity for large-n arrays. *Scientific reports*, *8*(1), 1646.

Lin, F.-C. L., Li, D., Clayton, R. W., & Hollis, D. (2013). High-resolution 3d shallow crustal structure in long beach, california:application of ambient noise tomography on a dense seismic array. *Geophysics*, *78*(4), 45–56.

Lin, T.-L., & Langston, C. A. (2007). Infrasound from thunder: A natural seismic source. *Geophysical research letters*, *34*(14).

Lin, T.-L., & Langston, C. A. (2009a). Thunder-induced ground motions: 1. observations. *Journal of Geophysical Research: Solid Earth*, *114*(B4).

Lin, T.-L., & Langston, C. A. (2009b). Thunder-induced ground motions: 2. site characterization. *Journal of Geophysical Research: Solid Earth*, *114*(B4).

Lythgoe, K. H., Qing, M. O. S., & Wei, S. (2020). Large-scale crustal structure beneath singapore using receiver functions from a dense urban nodal array. *Geophysical Research Letters*. doi: https://doi.org/10.1029/2020GL087233

Macpherson, K. A., Hidayat, D., Feng, L., & Goh, S.-H. (2013). Crustal thickness and velocity structure beneath singapore's seismic network. *J. of Asian Earth Sciences*, *64*, 245–255.

Meng, H., & Ben-Zion, Y. (2018a). Characteristics of airplanes and helicopters recorded by a dense seismic array near anza california. *Journal of Geophysical*

Research: Solid Earth, *123*(6), 4783–4797.

Meng, H., & Ben-Zion, Y. (2018b). Detection of small earthquakes with dense array data: Example from the san jacinto fault zone, southern california. *Geophysical Journal International*, *212*(1), 442–457.

Meteorological Service of Singapore, M. (2020). *Climate of Singapore.* Retrieved 2020-02-17, from `http://www.weather.gov.sg/climate-climate-of-singapore`

Miao, F., Carpenter, N. S., Wang, Z., Holcomb, A. S., & Woolery, E. W. (2020, 03). High-Accuracy Discrimination of Blasts and Earthquakes Using Neural Networks With Multiwindow Spectral Data. *Seismological Research Letters*. doi: 10.1785/0220190084

Pan, T. C., & Sun, J. (1996). Historical earthquakes felt in singapore. *Bull. Seismol. Soc. Am.*, *86*(4), 1173–1178.

Riahi, N., & Gerstoft, P. (2015). The seismic traffic footprint: Tracking trains, aircraft, and cars seismically. *Geophysical Research Letters*, *42*(8), 2674–2681.

Shuib, M. K., Manap, M. A., Tongkul, F., Bin Abd Rahim, I., Jamaludin, T. A., Surip, N., ... Ahmad, Z. (2017). Active faults in peninsular malaysia with emphasis on active geomorphic features of bukit tinggi region. *Mal. J. Geosci.*, *1*(1), 13–26.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., & Trujillo, J. (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bulletin of the Seismological Society of America*, *88*(1), 95–106.

Yong, C. Z., Denys, P. H., & Pearson, C. F. (2017). Present-day kinematics of the sundaland plate. *Journal of Applied Geodesy*, *11*(3), 169–177.

Yoon, C. E., O'Reilly, O., Bergen, K. J., & Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search. *Science advances*, *1*(11), e1501057.

Zhao, J., Chen, C., & Cai, J. (2002). A hydrogeological study of the sembawang hot spring in singapore. *Bulletin of Engineering Geology and the Environment*, *61*(1), 59–71.

Zhou, Y., & Zhao, J. (2016). Assessment and planning of underground space use in singapore. *Tunnelling and Underground Space Technology*, *55*, 249–256.

Zhu, T., & Stensrud, D. J. (2019). Characterizing thunder-induced ground motions using fiber-optic distributed acoustic sensing array. *Journal of Geophysical Research: Atmospheres*.

Detecting events in the urban seismic wavefield using a novel nodal array in Singapore: earthquakes, blasts and thunder quakes

Karen H. Lythgoe[1], Aidan Loasby[1], Shengji Wei[1]

[1]Earth Observatory of Singapore, Nanyang Technological University, Singapore

Contents of this file

1. Figure S1
2. Figure S2
3. Figure S3
4. Figure S4
5. Table S1

Additional Supporting Information (Files uploaded separately)

1. Movie S1

Caption for Movie S1. Long period seismic waves passing through Singapore. Colours on map correspond to ground velocity. Waveforms are from three stations located across Singapore (triangles 1, 2 and 3).

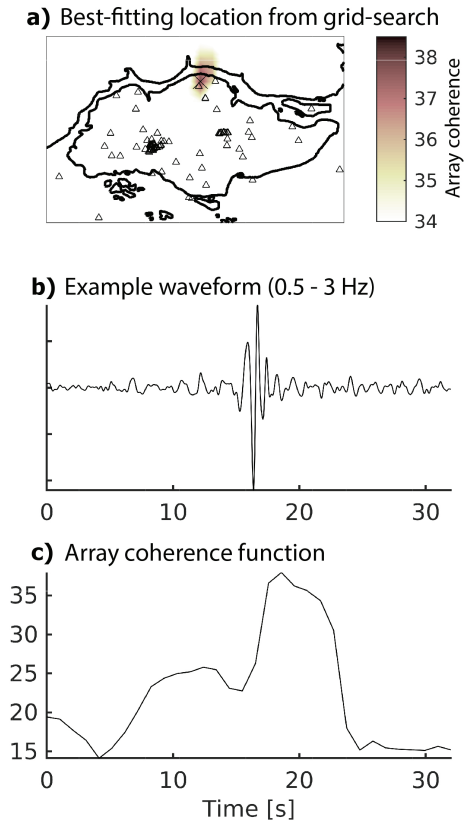

Figure S1. Local event at 17:57 on 3rd March 2019. a) Location obtained by grid search, with the best-fitting location defined as the point that maximises the array coherence. b) Example waveform filtered 0.5 - 3 Hz. c) Array coherence function (in 0.5 - 3 Hz frequency band).

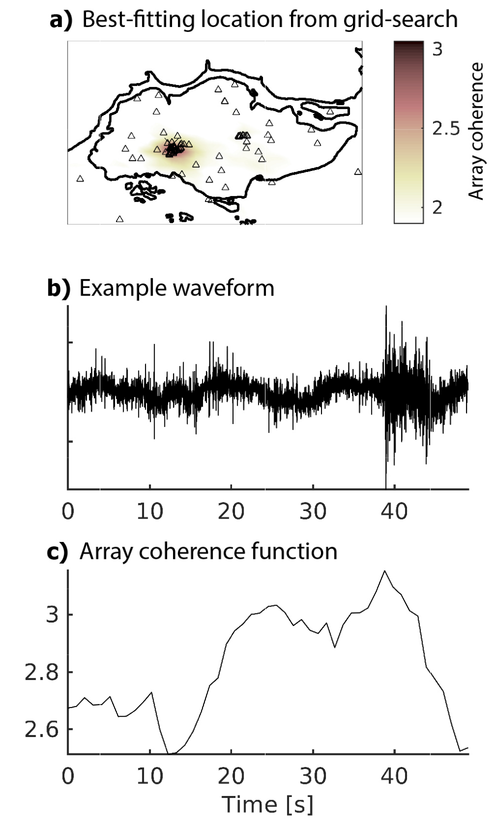

Figure S2. Thunder quake at around 2 am on 1st April 2019. a) Location obtained by grid search, with the best-fitting location defined as the point that maximises the array coherence. b) Example waveform (raw) c) Array coherence function.

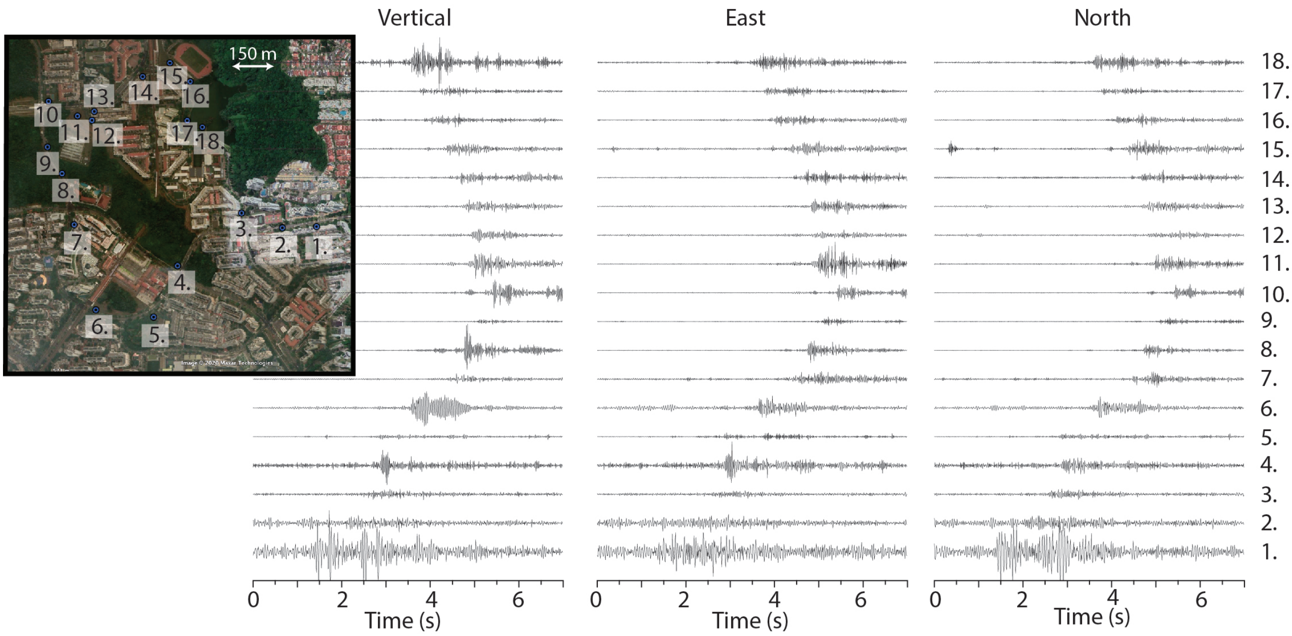

Figure S3. Waveforms induced by thunder at 18 stations within a 1 km x 1 km area. Waveforms are highly variable across the area likely affected by nearby buildings and near-surface structures.

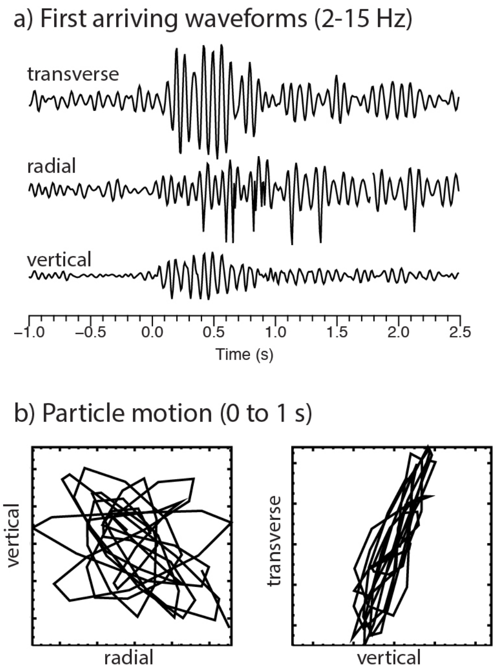

Figure S4. a) First arriving waveforms induced by a thunder source at red coloured station in Figure 10 (same as station number 6 in Figure S3). Horizontal components have been rotated to radial and transverse directions for the best-fitting event location. b) Particle motion for the first 1 second is consistent with a Rayleigh wave.

Table S1. Approximate origin time of detected local events

Probable blasting events

UTC Day	UTC	Local Day	Local Time
01/03/2019	09:10	01/03/2019	17:10
03/03/2019	09:56	03/03/2019	17:56
03/03/2019	09:56	03/03/2019	17:56
05/03/2019	03:56	05/03/2019	11:56
05/03/2019	08:38	05/03/2019	16:38
07/03/2019	09:54	07/03/2019	17:54
08/03/2019	04:05	08/03/2019	12:05
11/03/2019	05:00	11/03/2019	13:00
13/03/2019	05:39	13/03/2019	13:39
13/03/2019	11:48	13/03/2019	19:48
15/03/2019	05:23	15/03/2019	13:23
17/03/2019	09:34	17/03/2019	17:34
18/03/2019	05:00	18/03/2019	13:00
18/03/2019	09:15	18/03/2019	17:15
19/03/2019	05:12	19/03/2019	13:12
20/03/2019	10:00	20/03/2019	18:00
20/03/2019	09:32	20/03/2019	17:32
21/03/2019	04:32	21/03/2019	12:32
22/03/2019	04:15	22/03/2019	12:15
25/03/2019	05:20	25/03/2019	13:20
27/03/2019	04:20	27/03/2019	12:20
28/03/2019	05:15	28/03/2019	13:15

Unknown origin events

UTC Day	UTC	Local Day	Local Time
01/03/2019	20:02	02/03/2019	04:02
08/03/2019	11:20	08/03/2019	19:20
11/03/2019	17:07	12/03/2019	01:07
11/03/2019	19:13	12/03/2019	03:13
12/03/2019	04:02	12/03/2019	12:02
15/03/2019	15:46	15/03/2019	23:46
17/03/2019	23:31	18/03/2019	07:31
18/03/2019	18:36	19/03/2019	02:36
18/03/2019	19:00	19/03/2019	03:00
21/03/2019	20:17	22/03/2019	04:17
23/03/2019	19:43	24/03/2019	03:43
27/03/2019	05:30	27/03/2019	13:30
28/03/2019	04:06	28/03/2019	12:06