

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

4 **Karen H. Lythgoe¹, Aidan Loasby¹, Shengji Wei¹**

5 ¹Earth Observatory of Singapore, Nanyang Technological University, Singapore

6 **Key Points:**

- 7 • We develop a new method for event detection and characterization over frequency
8 space using a nodal array in the city state of Singapore.
- 9 • High frequency energy is elevated during lightning storms which is due to ground
10 movement from hundreds of thunder quakes.
- 11 • During 1 month we detected 76 distant earthquakes and 35 local events, some orig-
12 inating 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).

45 In this study we focus on detecting and characterising discrete seismic events in
46 Singapore, a densely populated city state. Knowledge of background seismicity is crit-
47 ical for the successful seismic monitoring of future underground developments. This is
48 particularly prescient for land and resource scarce Singapore, where expanding the ca-
49 pacity of the city underground is an imperative reality. Vast underground storage cav-
50 erns have been built and future subsurface construction plans include a potential geother-
51 mal energy plant (Zhao et al., 2002; Zhou & Zhao, 2016). Singapore also has geologi-
52 cal faults, including a significant tectonic fault located between granite and the metased-
53 imentary Jurong Group in Bukit Timah (Figure 1, Leslie et al. (2019); Lythgoe et al.
54 (2020)). The seismic activity level of the faults is unknown, however neighbouring ge-
55 ological faults have been reactivated due to post-seismic stresses from large earthquakes
56 at the nearby Sumatra subduction zone (Shuib et al., 2017; Yong et al., 2017). Discrete
57 seismic events also offer the possibility to use such sources for seismic imaging, in loca-
58 tions where it is difficult to use an active seismic source.

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

74 We find that array detection methods allow us to detect events with amplitudes
75 near to or even below noise levels. We detect seismic events from regional and teleseis-
76 mic earthquakes, as well as local seismic events, some from surprising sources. Detected
77 anthropogenic events include blasting events from underground construction. One in-

78 interesting seismic source are thunder and lightning storms, which create impulsive seis-
79 mic signals across a wide frequency range.

80 Here, we first describe the array and the basic features of the dataset. We then in-
81 troduce the array detection methods and catalogue the seismic events, both distant and
82 local, observed during the 1 month observation period using an existing and our new ar-
83 ray detection technique. We finish with a discussion of the results, including possible fu-
84 ture seismic source analysis and uses of the detected events.

85 **2 Nodal array and basic data features**

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

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

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

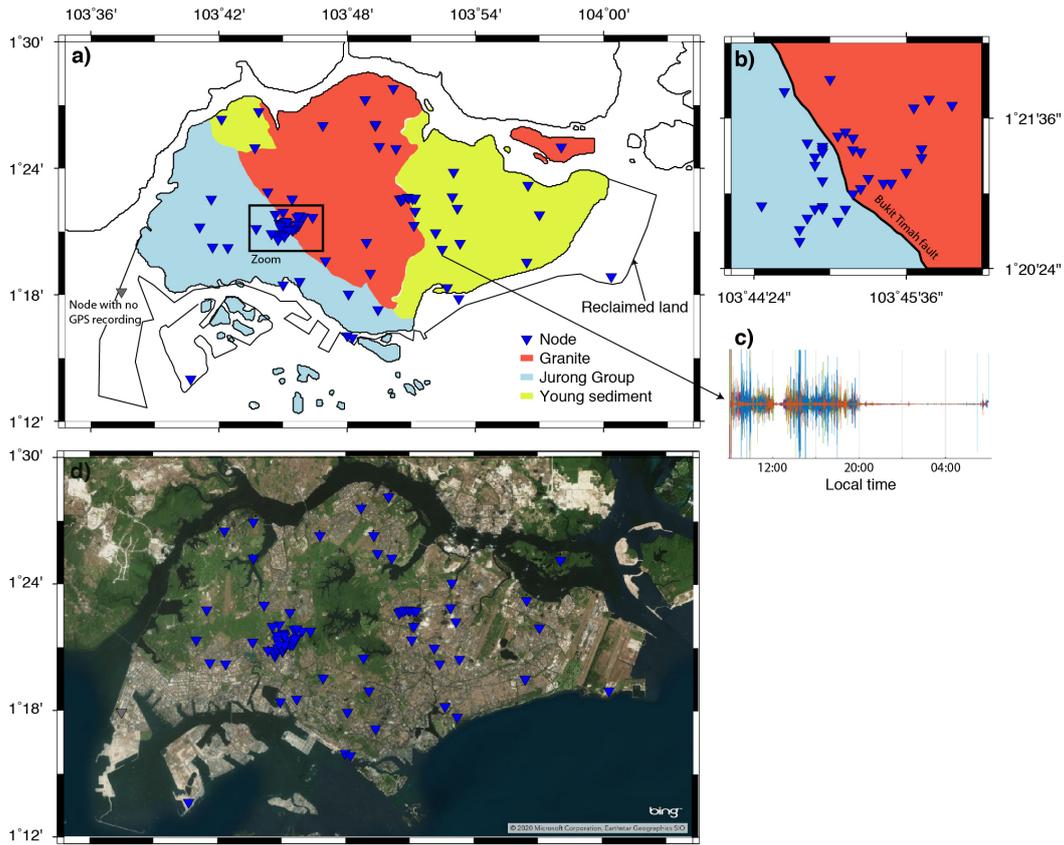


Figure 1. a) A simplified geological map of Singapore showing locations of nodes as blue triangles. White land areas within the Singapore coastline are reclaimed lands. b) Zoom to area of dense deployment across the Bukit Timah fault. c) Seismic data recorded by a node deployed in a school for a period 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.

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

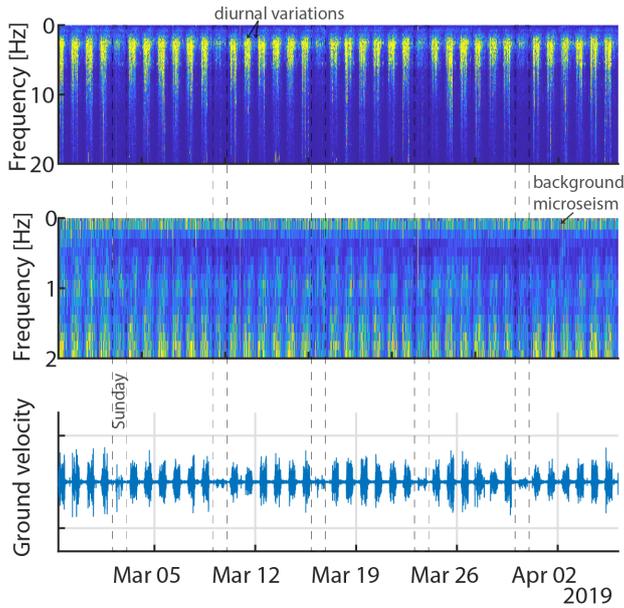


Figure 2. Spectrogram for frequencies less than 20 Hz (top row) and frequencies less than 2 Hz (middle row), plus the waveform (bottom row) for a node deployed in a park, after removing the instrument response.

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 wave-

132 form similarity method at frequencies where man-made noise dominates and the output
133 is useful for event classification.

134 **3.1 Array detection method using waveform similarity**

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

145 The detection threshold is set as the median amplitude in a sliding time window
146 plus ten times the median absolute deviation (Li et al., 2018). We use a maximum dis-
147 tance of 4 km between each station and its neighbours. We examine two frequency ranges,
148 0.5 - 3 Hz and 5 - 10 Hz, with the aim to detect distant and local events respectively.
149 A sliding window of 3 seconds and 1 second is used for the low and high frequency ranges
150 respectively, with windows having 50% overlap with the previous time window. We find
151 that the 5-10 Hz coherence function is too noisy to enable clear detection. We note that
152 closer station spacing may be required to detect events in the higher frequency range us-
153 ing this method.

154 One benefit of this method is the ability to approximately locate local seismic sources
155 using the time lags from cross-correlations. We do this in a grid search method by i) cal-
156 culating travel times for each point on the grid using a 1D velocity model; ii) convert-
157 ing the travel times to lag times between master and neighbouring stations; iii) extract-
158 ing cross-correlation coefficients at the corresponding lag times; iv) stacking all cross-
159 correlation coefficients, such that each grid point corresponds to a stacked correlation
160 value. In this way we define the best location as the location that has the highest cor-
161 relation value. We use a 1D velocity model for Singapore calculated from the joint in-
162 version of receiver functions and surface waves (Macpherson et al., 2013). In order to

163 minimise the dependence on the velocity model, we take the highest cross-correlation co-
 164 efficient from the neighbouring 5 samples.

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

176 **3.2 Array detection method using anomalous and coherent spectral en-** 177 **ergy**

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

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

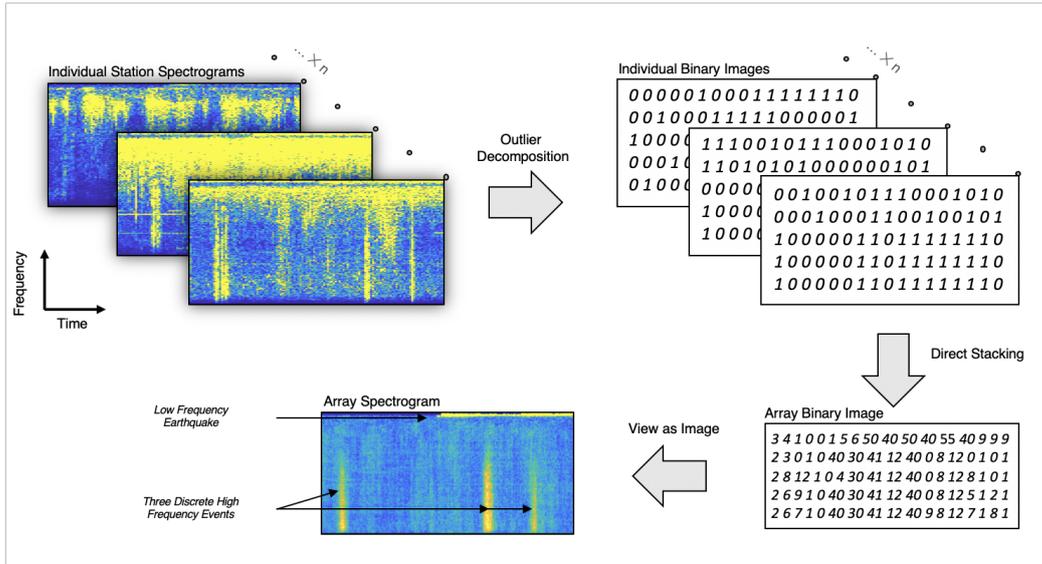


Figure 3. Workflow showing our methodology to detect anomalous, coherent spectral energy.

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

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 |

210 4 Distant earthquake detection

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

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

230 5 Local event detection

231 Local seismic events have a different character to regional and teleseismic earth-
 232 quakes. These events produced lower amplitude seismic signals, yet retain high frequency

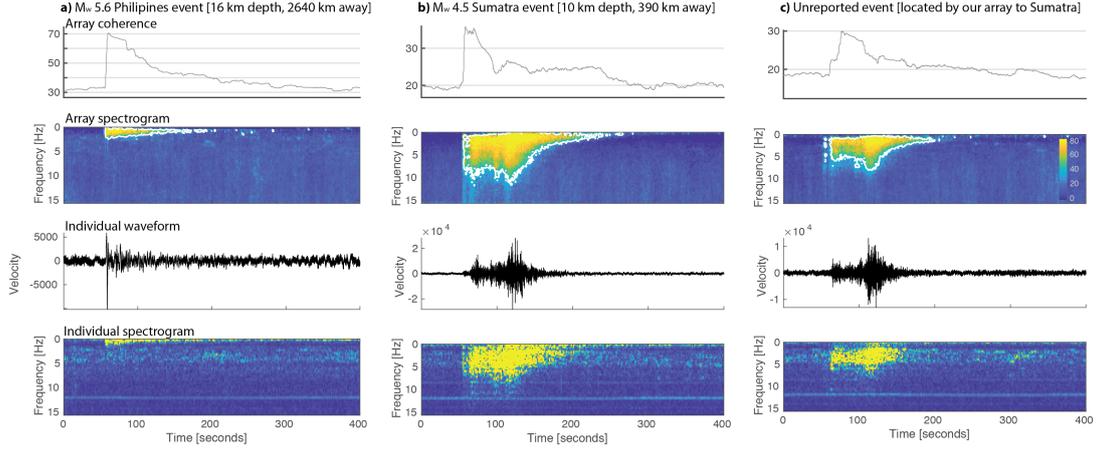


Figure 4. Examples of distant earthquakes detected by our array - a) an earthquake in the Philippines, b) an earthquake in Sumatra, c) an earthquake not reported in global earthquake catalogues, which we locate using the Singapore array to Sumatra. Panels from top to bottom: array coherence function;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

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

238 To locate the events, we use the cross-correlation lag times already calculated for
 239 the similarity function (Supplementary Figure 1). Both events are located in the north-
 240 east of Singapore at a known construction site. Figure 6 shows the location and appar-
 241 ent moveout for Event 2 in Figure 5. The characteristic of two events occurring close in
 242 time is typical of blasting patterns at this site. The timing in the early evening, is also
 243 a typical characteristic of blasting, since permits are normally given for approximately
 244 5-6 pm once the site is clear and workers have left. Therefore we conclude that these events
 245 are from construction site blasting works.

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

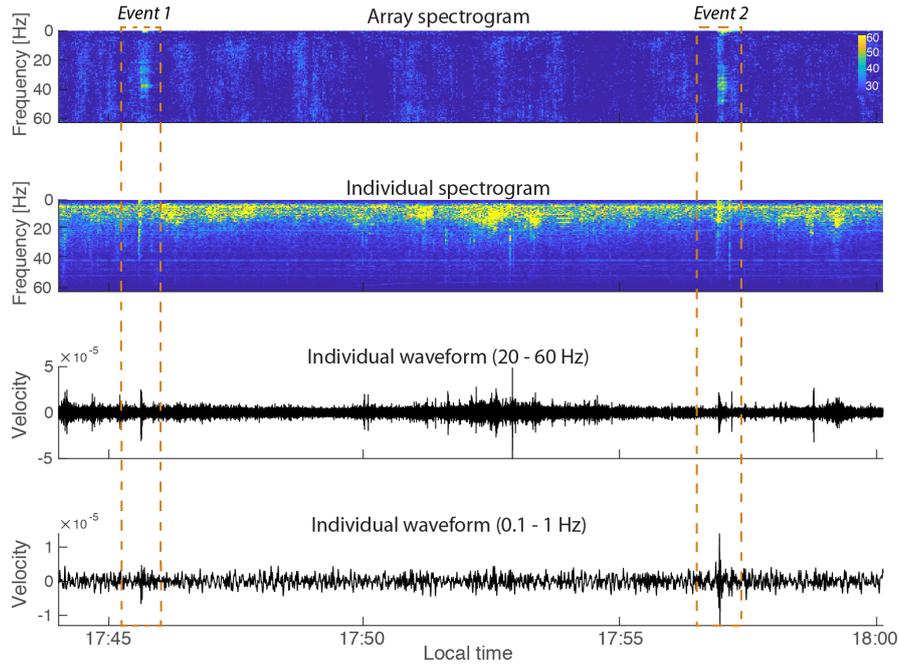


Figure 5. Two local events on 3rd March 2019 marked by orange boxes. Panels from top to bottom: array spectrogram; spectrogram at one station; waveform at same station filtered 20 - 60 Hz; same waveform filtered 0.1 - 1 Hz.

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

252 **6 Thunder quakes**

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

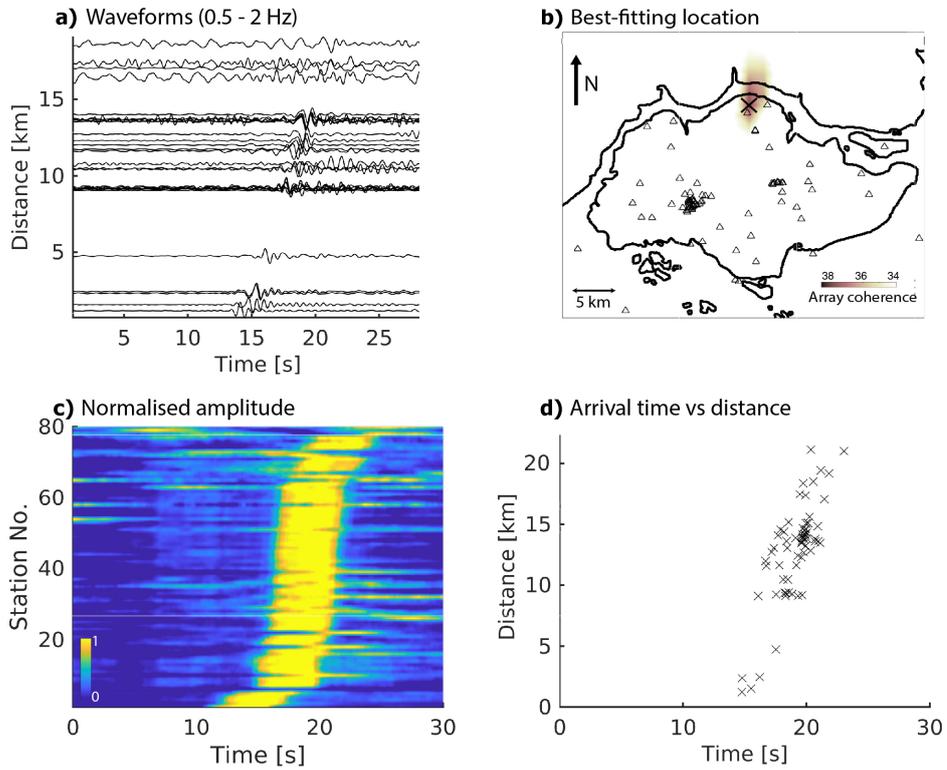


Figure 6. Local event on 3rd March 2019 (Event 2 in Figure 5). a) Waveforms recorded at stations in array, filtered 0.5 - 2 Hz. b) Map of Singapore with best-fitting location of source (cross) obtained by grid search of lag times. c) Moving mean of the absolute amplitude in 0.5 - 2 Hz frequency band 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).

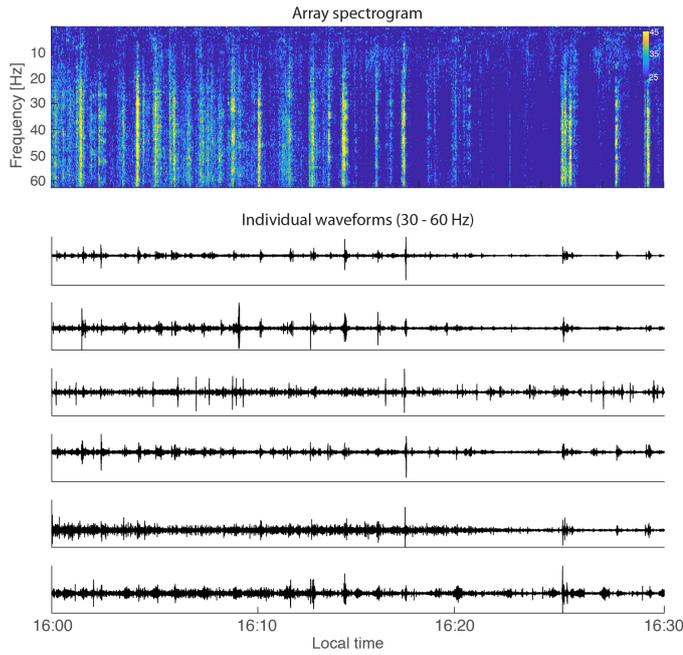


Figure 7. Seismic observations during a lightning storm on 11th March 2019. Top panel: array spectrogram. Bottom panels: representative waveforms at randomly selected stations around the island, filtered 30 - 60 Hz.

260 showed that microphone recordings of acoustic waves induced by lightning have frequen-
 261 cies ranging from 4 to 125 Hz, consistent with our observations.

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

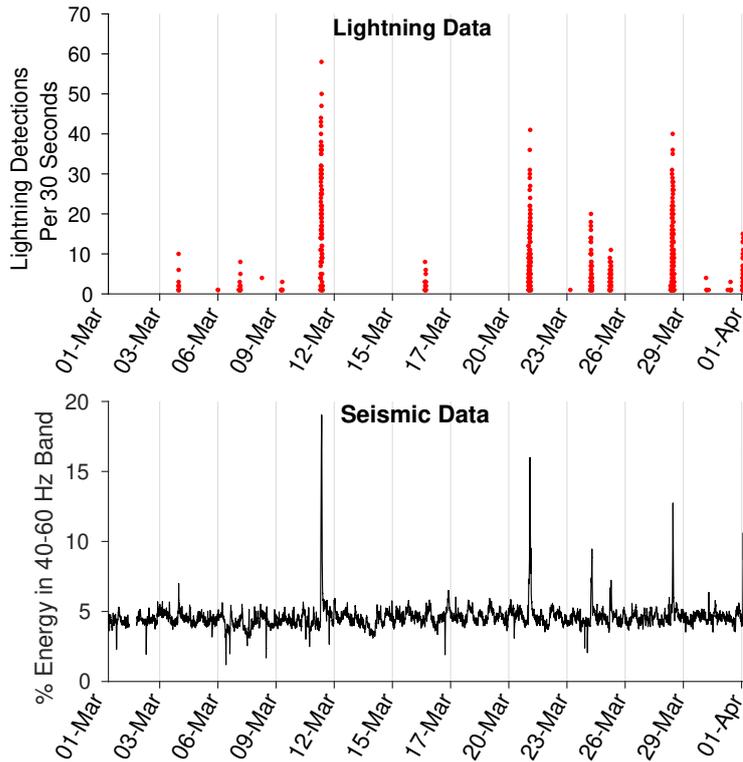


Figure 8. Correlation between high frequency seismic energy and frequency of lightning detected in Singapore. Seismic energy is calculated as the ratio of the average energy in the 40-60 Hz to the total seismic energy less than 60 Hz averaged for the whole array. Lightning strikes are cloud to ground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

275 2 and 6 pm due to the generation of storms by diurnal heating. However a night-time
 276 lightning storm on 1st April is evident from the larger percentage of high frequency seis-
 277 mic energy for the number of lightning strikes detected, due to quieter ambient seismic
 278 energy levels at night. The rare occurrences of night-time lightning storms are particu-
 279 larly valuable since they occur when cultural noise is a minimum and we use this for more
 280 detailed analysis of an individual event.

281 Figure 9 shows an example of a teleseismic earthquake arriving in Singapore dur-
 282 ing a lightning storm. The earthquake’s seismic waves are dominant at low frequencies
 283 (0.1 - 1 Hz). The signal from the earthquake is difficult to see in the spectrogram of an
 284 individual station, however they can be clearly identified on the array spectrogram. We
 285 note that for the thunder signals, the stacked signal from the entire array shows less low
 286 frequency energy (< 10 Hz) compared to higher frequencies, while the single station shows

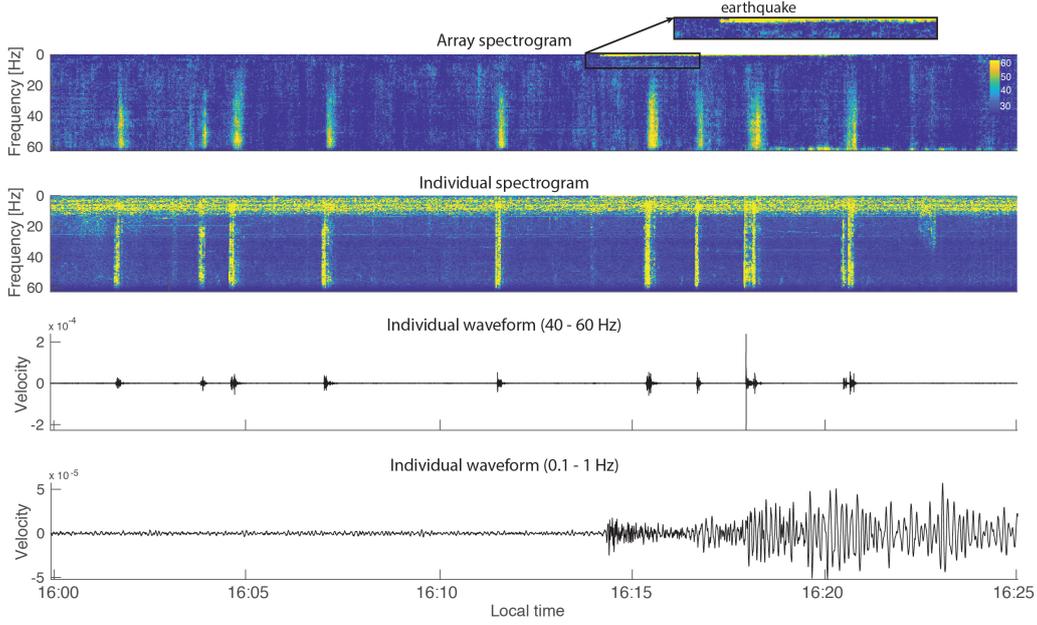


Figure 9. The concurrent arrival of seismic waves from an earthquake in Sumatra and a series of discrete lightning quakes. Seismic waves from the earthquake arrive at low frequencies, while the lightning quakes are across all frequencies.

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

290 One thunder event occurring at night is shown in Figure 10. The high frequency
 291 waveforms for the event are arranged by distance from the estimated source location in
 292 Figure 10a. We estimate the source location using a 3D grid search of hand picked first
 293 arrival times (Figure 10b). We also attempt to locate the source based on a grid search
 294 of differential lag times, as used to locate a local blasting event above (we re-calculate
 295 the coherence functions with a longer lag time to account for the slower moveout veloc-
 296 ity). However the best-fitting location from this method is incompatible with the first
 297 arrival times since the effectiveness of the waveform similarity method is limited due to
 298 low waveform coherence (Supplementary Figure 2). The energy of the event can be traced
 299 across the array (Figure 10c) although it is difficult to pick first arrivals at all stations
 300 - stations with a robust pick are shown as triangles in Figure 10b. The apparent move-
 301 out velocity of the event is 350 m/s (Figure 10d). We therefore use a constant velocity
 302 of 350 m/s in the grid search location and find the best-fitting location to be at an el-

303 evation of 4 km. Given that the moveout velocity is similar to the speed of sound, and
304 that the source is elevated by 4 km, we conclude that the seismic signal is due to the con-
305 version of energy from an acoustic wave in the atmosphere to an elastic wave in the ground.

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

320 7 Discussion

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

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

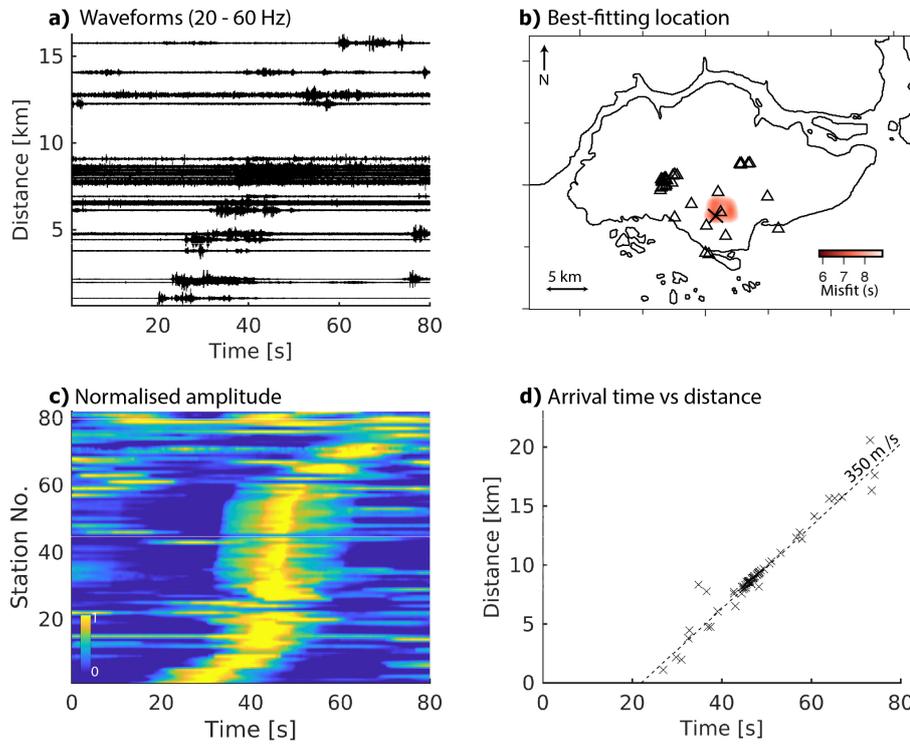


Figure 10. Seismic recording of a thunder quake. a) Waveforms recorded at stations in array, filtered 20 - 60 Hz. b) Map of Singapore with best-fitting location of source (cross) obtained by grid search of picked first arrival times. c) Moving mean of the absolute amplitude in 20 - 60 Hz frequency band 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.

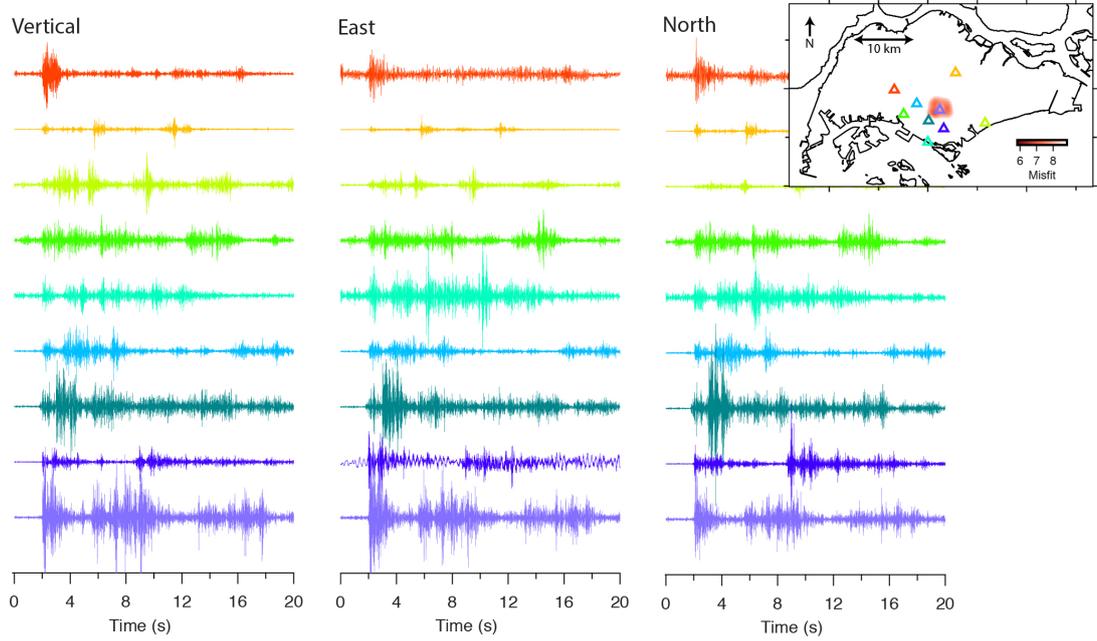


Figure 11. Waveforms induced by thunder at 9 nearby stations. Waveforms are aligned at their first arrival, bandpass filtered 20 - 60 Hz and are not normalised.

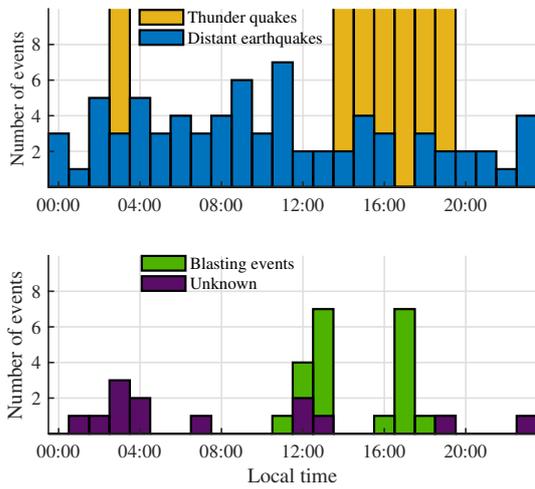


Figure 12. Total number of events per hour detected in each event classification. For comparison purposes the number of events is capped at 10, although hundreds of thunder quakes are detected (we do not count them all).

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

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

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

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

385 8 Conclusions

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

395 Acknowledgments

396 We thank the following Singapore agencies for assistance in data acquisition: Meteorological Service of Singapore, NParks, Ministry of Education and the Public Utilities Board.
397

398 We thank the Centre for Geohazard Observation at NTU, along with all other field as-
 399 sistants for their help with the field survey, particularly Jeffrey Encillo, Choong Yew Leong,
 400 Weiwen Chen and Mark Lim. Due to the sensitive nature of the data, not all of it can
 401 be shared. We make available datasets of the examples and events presented here which
 402 will be accessible from the NTU data repository (DOI to be provided). Data for review
 403 purposes is accessible via this link: [https://entuedu-my.sharepoint.com/:f:/g/personal/
 404 karen_lythgoe_staff_main_ntu_edu_sg/EtIpEEWWS8VGvBYp7CguTjkBhpNYbFfdjNhjnQ9va3
 405 _gXQ?e=TuzdIu](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.
 406 XXX. This research is partly supported by the National Research Foundation Singapore
 407 and the Singapore Ministry of Education under the Research Centres of Excellence ini-
 408 tiative (Project Code Number: XXXX). Karen Lythgoe is funded by an NTU Presiden-
 409 tial Postdoctoral Fellowship.

410 References

- 411 Albert, D. G., & Decato, S. N. (2017). Acoustic and seismic ambient noise measure-
 412 ments in urban and rural areas. *Applied Acoustics*, *119*, 135–143.
- 413 Chamarczuk, M., Nishitsuji, Y., Malinowski, M., & Draganov, D. (2020). Unsuper-
 414 vised learning used in automatic detection and classification of ambient-noise
 415 recordings from a large-n array. *Seismological Research Letters*, *91*(1), 370–
 416 389.
- 417 Díaz, J., Ruiz, M., Sánchez-Pastor, P. S., & Romero, P. (2017). Urban seismology:
 418 On the origin of earth vibrations within a city. *Scientific reports*, *7*(1), 1–11.
- 419 Gibbons, S. J., & Ringdal, F. (2006). The detection of low magnitude seismic events
 420 using array-based waveform correlation. *Geophysical Journal International*,
 421 *165*(1), 149–166.
- 422 Gradon, C., Moreau, L., Roux, P., & Ben-Zion, Y. (2019). Analysis of surface and
 423 seismic sources in dense array data with match field processing and Markov
 424 chain Monte Carlo sampling. *Geophysical Journal International*, *218*(2),
 425 1044–1056.
- 426 Green, D. N., Bastow, I. D., Dashwood, B., & Nippress, S. E. (2017). Characteriz-
 427 ing broadband seismic noise in central london. *Seismological Research Letters*,
 428 *88*(1), 113–124.
- 429 Hansen, S. M., & Schmandt, B. (2015). Automated detection and location of micro-

- 430 seismicity at mount st. helens with a large-n geophone array. *Geophysical Re-*
 431 *search Letters*, *42*(18), 7390–7397.
- 432 Hasselmann, K. (1963). A statistical analysis of the generation of microseisms. *Re-*
 433 *views of Geophysics*, *1*(2), 177–210.
- 434 Holmes, C., Brook, M., Krehbiel, P., & McCrory, R. (1971). On the power spec-
 435 trum and mechanism of thunder. *Journal of Geophysical Research*, *76*(9),
 436 2106–2115.
- 437 Kanamori, H., Mori, J., Anderson, D. L., & Heaton, T. H. (1991). Seismic excitation
 438 by the space shuttle columbia. *Nature*, *349*(6312), 781–782.
- 439 Leslie, G. A., Dodd, T. J., Gillespie, M. R., Kendall, R. S., Bide, T. P., Kearsey,
 440 M. R., Timothy I. and Dobbs, . . . Chiam, K. (2019). Ductile and brittle de-
 441 formation in singapore: A record of mesozoic orogeny and amalgamation in
 442 sundaland, and of post-orogenic faulting. *J. of Asian Earth Sciences.*, *180*.
- 443 Li, Z., Peng, Z., Hollis, D., Zhu, L., & McClellan, J. (2018). High-resolution seismic
 444 event detection using local similarity for large-n arrays. *Scientific reports*, *8*(1),
 445 1646.
- 446 Lin, F.-C. L., Li, D., Clayton, R. W., & Hollis, D. (2013). High-resolution 3d shal-
 447 low crustal structure in long beach, california: application of ambient noise
 448 tomography on a dense seismic array. *Geophysics*, *78*(4), 45–56.
- 449 Lin, T.-L., & Langston, C. A. (2007). Infrasonic from thunder: A natural seismic
 450 source. *Geophysical research letters*, *34*(14).
- 451 Lin, T.-L., & Langston, C. A. (2009a). Thunder-induced ground motions: 1. obser-
 452 vations. *Journal of Geophysical Research: Solid Earth*, *114*(B4).
- 453 Lin, T.-L., & Langston, C. A. (2009b). Thunder-induced ground motions: 2. site
 454 characterization. *Journal of Geophysical Research: Solid Earth*, *114*(B4).
- 455 Lythgoe, K. H., Qing, M. O. S., & Wei, S. (2020). Large-scale crustal structure be-
 456 neath singapore using receiver functions from a dense urban nodal array. *Geo-*
 457 *physical Research Letters*. doi: <https://doi.org/10.1029/2020GL087233>
- 458 Macpherson, K. A., Hidayat, D., Feng, L., & Goh, S.-H. (2013). Crustal thickness
 459 and velocity structure beneath singapore’s seismic network. *J. of Asian Earth*
 460 *Sciences*, *64*, 245–255.
- 461 Meng, H., & Ben-Zion, Y. (2018a). Characteristics of airplanes and helicopters
 462 recorded by a dense seismic array near anza california. *Journal of Geophysical*

- 463 *Research: Solid Earth*, 123(6), 4783–4797.
- 464 Meng, H., & Ben-Zion, Y. (2018b). Detection of small earthquakes with dense array
465 data: Example from the san jacinto fault zone, southern california. *Geophysical*
466 *Journal International*, 212(1), 442–457.
- 467 Meteorological Service of Singapore, M. (2020). *Climate of Singapore*. Re-
468 trieved 2020-02-17, from [http://www.weather.gov.sg/climate-climate](http://www.weather.gov.sg/climate-climate-of-singapore)
469 [-of-singapore](http://www.weather.gov.sg/climate-climate-of-singapore)
- 470 Miao, F., Carpenter, N. S., Wang, Z., Holcomb, A. S., & Woolery, E. W. (2020,
471 03). High-Accuracy Discrimination of Blasts and Earthquakes Using Neural
472 Networks With Multiwindow Spectral Data. *Seismological Research Letters*.
473 doi: 10.1785/0220190084
- 474 Pan, T. C., & Sun, J. (1996). Historical earthquakes felt in singapore. *Bull. Seis-*
475 *mol. Soc. Am.*, 86(4), 1173–1178.
- 476 Riahi, N., & Gerstoft, P. (2015). The seismic traffic footprint: Tracking trains, air-
477 craft, and cars seismically. *Geophysical Research Letters*, 42(8), 2674–2681.
- 478 Shuib, M. K., Manap, M. A., Tongkul, F., Bin Abd Rahim, I., Jamaludin, T. A.,
479 Surip, N., . . . Ahmad, Z. (2017). Active faults in peninsular malaysia with
480 emphasis on active geomorphic features of bukit tinggi region. *Mal. J. Geosci.*,
481 1(1), 13–26.
- 482 Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., & Trujillo, J.
483 (1998). A comparison of select trigger algorithms for automated global seismic
484 phase and event detection. *Bulletin of the Seismological Society of America*,
485 88(1), 95–106.
- 486 Yong, C. Z., Denys, P. H., & Pearson, C. F. (2017). Present-day kinematics of the
487 sundaland plate. *Journal of Applied Geodesy*, 11(3), 169–177.
- 488 Yoon, C. E., O’Reilly, O., Bergen, K. J., & Beroza, G. C. (2015). Earthquake de-
489 tection through computationally efficient similarity search. *Science advances*,
490 1(11), e1501057.
- 491 Zhao, J., Chen, C., & Cai, J. (2002). A hydrogeological study of the sembawang
492 hot spring in singapore. *Bulletin of Engineering Geology and the Environment*,
493 61(1), 59–71.
- 494 Zhou, Y., & Zhao, J. (2016). Assessment and planning of underground space use in
495 singapore. *Tunnelling and Underground Space Technology*, 55, 249–256.

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