THESIS

EVENT DETECTION AND ANALYSIS OF A DENSE THREE-COMPONENT NEAR-SUMMIT SEISMIC ARRAY DEPLOYED AT EREBUS VOLCANO

Submitted by
Erika Jaski
Department of Geosciences

In partial fulfillment of the requirements
For the Degree of Master of Science
Colorado State University
Fort Collins, Colorado
Fall 2022

Master's Committee:

Advisor: Richard C. Aster

Derek Schutt
Margaret Cheney

ABSTRACT

EVENT DETECTION AND ANALYSIS OF A DENSE THREE-COMPONENT NEAR-SUMMIT SEISMIC ARRAY DEPLOYED AT EREBUS VOLCANO

Erebus volcano on Ross Island, Antarctica has maintained an erupting phonolitic lava lake for at least five decades. During active periods, the lava lake hosts large (up to ~10-m diameter) gas slugs rising through the conduit that create impulsive Strombolian eruptions and produce very long period (VLP) signals on broadband seismograms. We combine near-summit broadband observations and reanalyze data from a 100-station three-component short-period (4.5 Hz geophones) network deployed in an approximately 3 by 3 km region around the Main Crater during December 2008. Lava lake eruption template events are identified on broadband seismograms from their characteristic and repeating VLP spectral signature of nonharmonic modes between 0.033 and 0.2 Hz. Multi-channel and multi-station waveform matched filter correlations are performed across the short-period network using template events and correlation values that are three or more standard deviations are extracted into a working Inner Crater event catalogue, yielding 819 event detections over 19 days. While 94% of the signals in this catalogue are unique, 17 "families" of repeating lava lake events can also be identified through similar waveforms determined by Ward clustering on 5 stations, which are further interpreted for trends in location, size, and occurrence. We observe time-varying quasi-Poissonian interevent times and an approximately power-law size-frequency distribution with an excess of small events. Investigating the various event families that transpire in the Inner Crater region contributes to improved characterization and understanding of the seismogenic behavior of the lava lake degassing system and assists in the creation of a workflow that can be applied in volcanic and other circumstances that generate prolific low-level impulsive seismicity.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Rick Aster, whose comprehensive support and guidance was an integral component of this project. I would also like to thank my committee members, Margaret Cheney and Derek Schutt, for their support and input on this research. Thank you to Dr. Ronni Grapenthin and Dr. Julien Chaput for meeting with our research group throughout the duration of this project to assist with development and providing feedback. I would like to thank the National Science Foundation for funding this collaborative research project. Lastly, I would like to thank my friends and family for their love and support while I did my research.

This project was funded by the National Science Foundation Antarctic Earth Sciences grant 1916978.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Mount Erebus, the southernmost active volcano on Earth, is located 35 km north of McMurdo Station on Ross Island, Antarctica (Figure 1.1). This stratovolcano is unique due to its shallowly convecting phonolitic lava lake and long-term open conduit state, which may have been continuously present since its discovery in 1841 [1]. This open vent system offers a rare opportunity to directly monitor the status of a degassing lava lake system. Mount Erebus has been monitored since 1972 using various techniques that include seismic, infrasound, GPS, magnetic, thermal, tiltmeters, meteorological, radar, gas chemistry, and video [9]. The lowest-level sustained seismic activity of the Erebus lava lake has never been studied in detail, therefore, this work aims to detect and analyze its smallest detectable impulsive signals. These high frequency events are much more common than larger events associated with known eruptions, and allow us to assess its short-term seismogenic activity. More generally, monitoring the location and type of seismic activity improves our understanding of the volcano's behavior by indicating areas of seismogenic deformation as well as eruptive signals [10]. This procedure is particularly important at Mount Erebus because of its proximity to McMurdo Station and its widespread logistical impact on Antarctic activity, as well as for future research on the volcano. Although Mount Erebus has been relatively peaceful in recent years, with most eruptive impacts limited to within its Main Crater, hazardous potential is evident from larger eruptions in the geologic and recent past. Tephra can be found throughout Ross Island, and as far as the Transantarctic Mountains [4]. A notably energetic Strombolian eruption period occurred from September 1984 to January 1985 during which ash approached 6000 meter altitudes. Phreatic explosions on October 19, 1993 were sufficiently powerful that a new vent formed near the edge of the Inner Crater, and the Inner Crater was enlarged. Future eruptions, similar or larger to the size of these, present significant risk to Ross Island aviation and to those working on the volcano [1].

Figure 1.1: Southern face of Mount Erebus. Photograph by Sylvian Pichat. From [1]

Mount Erebus experiences frequent Strombolian eruptions caused by rising gas bubbles that eject phonolite bombs and ash, but rare phreatomagmatic and phreatic eruptions have also been observed [7]. These eruptions over the past few decades have a Volcano Explosivity Index (VEI) between 0 and 1 [11]. Seismic and infrasonic activity were first recorded on a network of analogue short period (1 s) vertical component seismographs in 1974 [1]. Monitoring since then reveals that Mount Erebus characteristically has 1-10 eruptions a day, and has usually maintained this pattern since the early 70's, with the exception of reduced activity between 2001-2004 and elevated activity in 1983-1984 [1]. According to [12], these eruptions can occur up to many tens times per day within variable intervals. Variations in eruption rates and style are still poorly understood and may be associated with gas accumulation in the conduit system and changes in the shape of the magma conduit at show depths [5, 13].

Movement of magma, ice-quakes, and ocean tides are sources of seismic activity at Mount Erebus. Tremors can be observed at most volcanoes, however, at Mount Erebus, tremor signals are solely associated with nearby iceberg collisions. The lack of volcanic tremor likely arises because of low levels of magmatic input and an open conduit system that prevents the buildup of internal pressure and seismogenic magmatic transport [14]. An important discovery in the behavior of Mount Erebus is the identification of a very long period (VLP) signal that is associated with lava lake eruptions [7]. This prolonged (minutes-long) signal results from a gas slug ascending through the conduit, and its bursting at the lava lake surface. Following the explosive eruption, which discharges material from the lava lake as ash and bombs, a seismic VLP coda lasting several minutes is observed to persist until the lava lake resumes its pre-eruptive elevation. Although it is mainly a post-eruptive signal, it initiates slightly before the surface explosion with the pre-eruptive bulging of the lava lake. This repetitive VLP signal demonstrates the non-destructive and self-reconstructing quality of the erupting lava lake system [1]. Larger eruptive events can be identified in broadband near-crater seismic records by these unique signals [7] and leveraged to investigate second-order temporal changes [1]. Because the volcanic edifice is highly heterogeneous, there is an abundance of scattering in high frequency seismic signals that results in cigar-shaped scattering envelopes at all source-receiver distances that obscures source information. However, scattering has a stronger effect on high frequency waves, which allows VLP signals to remain observable and to preserve source characteristics [7]. Because of this, VLP signals are commonly used to identify lava lake eruptions at Mount Erebus.

Chapter 2

Background

2.1 Geologic Setting

Mount Erebus, an alkali-rich volcano, began forming ~1.3 Mya at the southern region of the Terror Rift, part of the more extensive West Antarctic Rift System (WARS) [3]. WARS includes a sequence of rift valleys that are the result of shear extension during the Cretaceous period [15]. This extension is roughly perpendicular to the Transantarctic Mountains, which separate East and West Antarctica. Just east of this mountain range lies Ross Island, which is connected to the West Antarctic Ice Sheet through McMurdo and Ross Ice Shelves. Ross Island is home to Mount Erebus and to three other large-volume inactive volcanoes: Mount Terror, Mount Bird, and Mount Terra Nova (Figure 2.1). With a height of 3794 m and a volume of 1670 km^3, Mount Erebus is the largest and highest of these volcanoes. In total, about 4520 km^3 of material has been erupted on Ross Island [1]. Lava flows and tephra indicate that Mount Erebus lava is phonolitic and has maintained this geochemical signature for thousands of years. At low degrees of mantle melting, CO^2 rich, silica undersaturated, basanitic magma forms and evolves to a phonolitic composition through fractional crystallization [1]. Late Cenozoic erupted material that lies on Ross Ice Shelf and within the Ross Sea make up the larger McMurdo Volcanic Group [16].

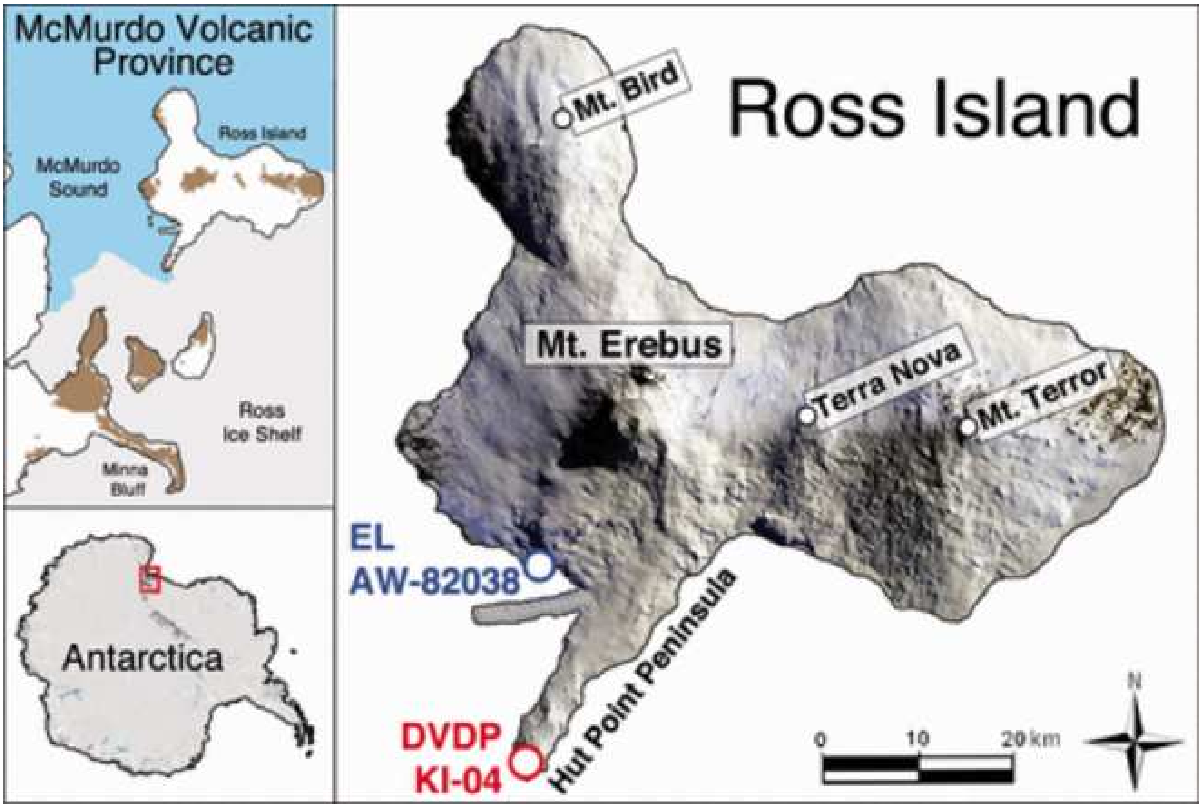
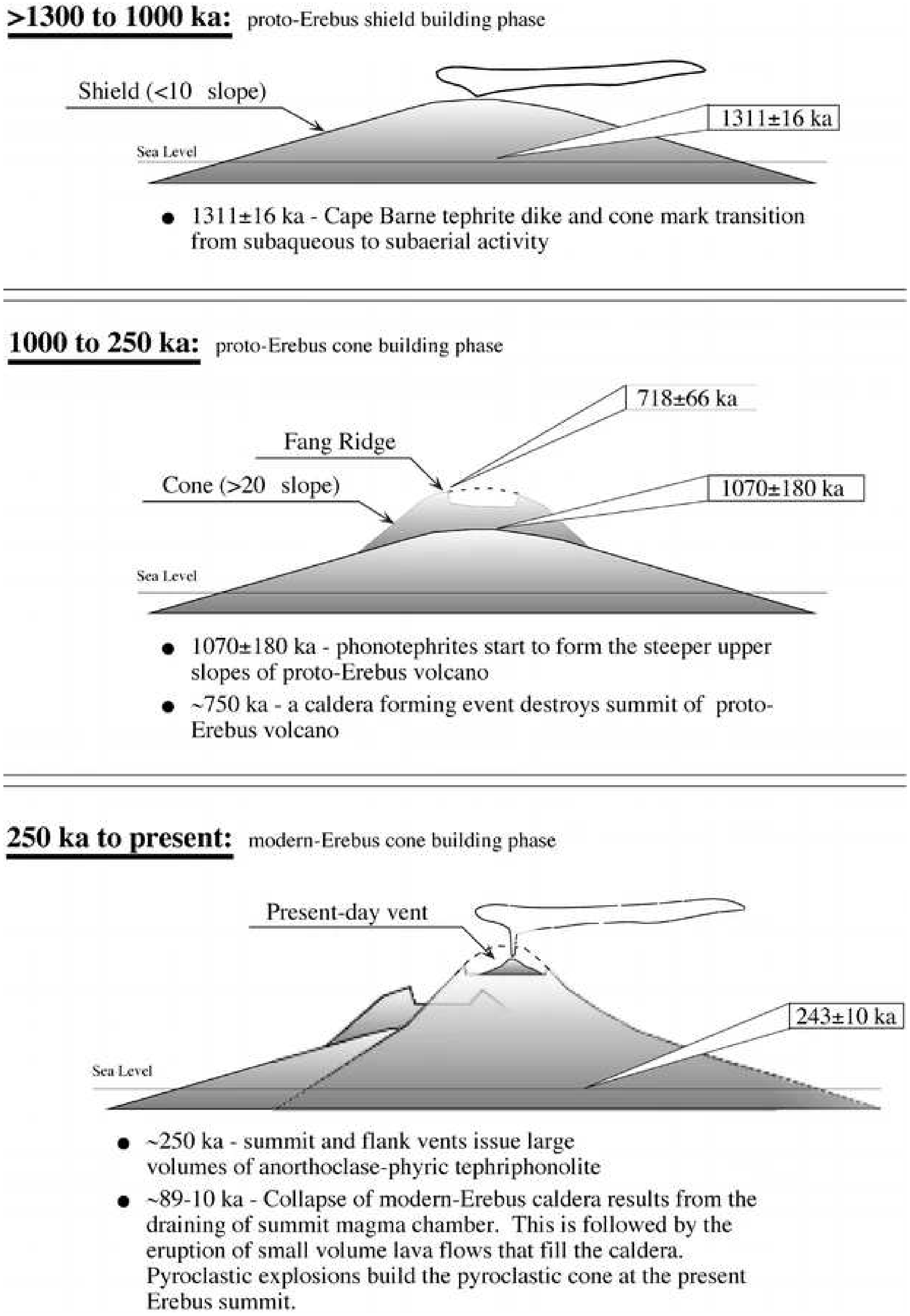


Figure 2.1: Map of Ross Island denoting the four principal volcanoes. The dark areas in the top left map indicate the McMurdo Volcanic Province. From [2]. The red and blue circles indicate locations that rock samples were collected for [2].

Three phases of Mount Erebus' history have been identified: proto-Erebus shield-building phase, proto-Erebus cone-building phase, and a modern Erebus cone-building phase (Figure 2.2) [3]. The proto-Erebus shield-building phase occurred between 1.3 and 1 Mya. This phase consisted of basanitic lava eruptions that transitioned from being erupted underwater, indicated by hyaloclastic deposits, to subaerially. This material built the original shield volcano [1]. The prominent Fang Ridge contains one of the last lava flows from this time [17]. The proto-Erebus cone-building phase transpired between 1 Ma and 250 ka. This phase is characterized by phonotephrite lava flowing out of the summit that now make up most of the volcano's exposed slopes. Based on the configuration of Fang Ridge, the active vent during this time is estimated to be ~2 km north of the present-day vent [3]. After a caldera collapse ended the second stage of Mount Erebus' life, the current phase began around 250 ka. Anorthoclase-phyric tephriphonolite and phonolitic lavas erupted from the summit region to build the present-day upper structure of the volcano which was largely formed by 90 ka [3]. During this time, there were two caldera building events and two caldera collapses that occurred between 80-25 ka, which may have been the result of a Plinian eruption, and 25-11 ka (Figure 2.3) [4]. Presently, volcanic activity consists of crater-limited lava flows and Strombolian eruptions that eject lava bombs within and less frequently, outside of, the crater.



>1300 to 1000 ka: proto-Erebus shield building phase

- 1311±16 ka - Cape Barne tephrite dike and cone mark transition from subaqueous to subaerial activity

1000 to 250 ka: proto-Erebus cone building phase

- 1070±180 ka - phonotephrites start to form the steeper upper slopes of proto-Erebus volcano
- ~750 ka - a caldera forming event destroys summit of proto-Erebus volcano

250 ka to present: modern-Erebus cone building phase

- ~250 ka - summit and flank vents issue large volumes of anorthoclase-phyric tephriphonolite
- ~89-10 ka - Collapse of modern-Erebus caldera results from the draining of summit magma chamber. This is followed by the eruption of small volume lava flows that fill the caldera. Pyroclastic explosions build the pyroclastic cone at the present Erebus summit.

Figure 2.2: Three phases of Mount Erebus. From [3].

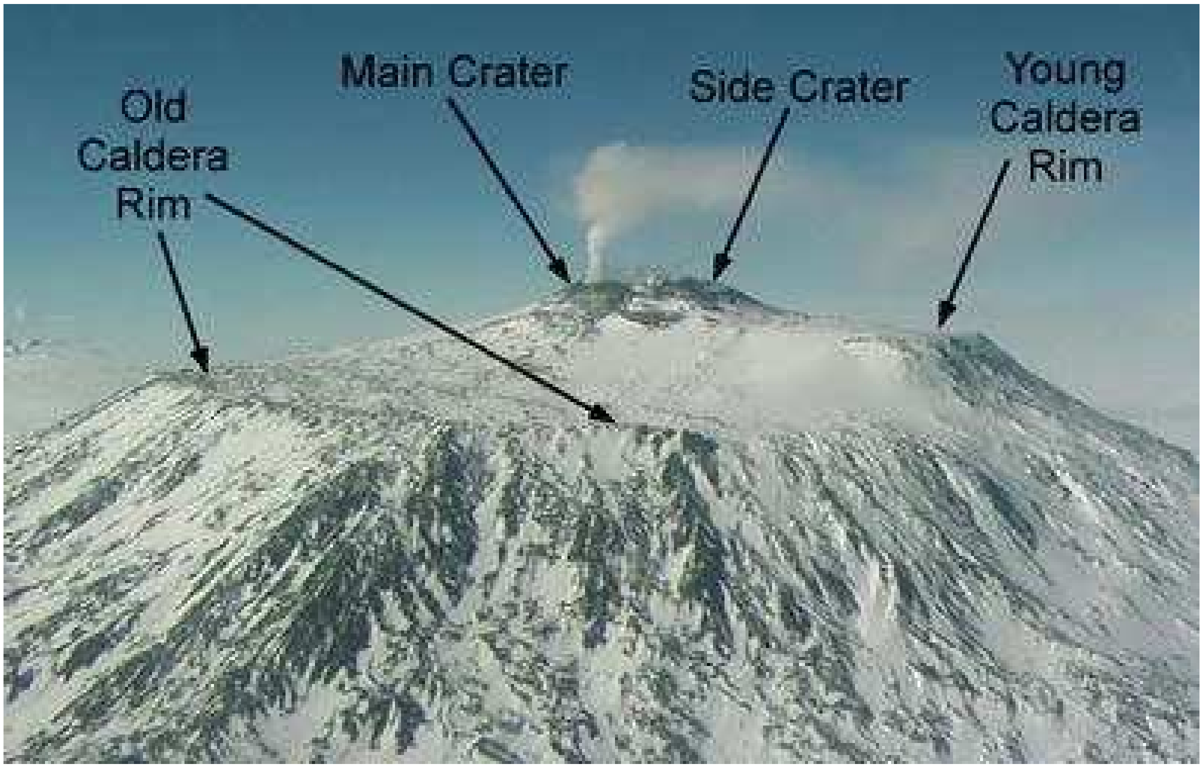


Figure 2.3: Structure of Mount Erebus. From [4].

Notable geographic features of Mount Erebus include Fang Ridge, Tramway Ridge, Ice Tower Ridge, and the Side Crater (Figure 2.4). Fang Ridge is what remains of a former cone of Mount Erebus on the northeast edge and is no older than 718 ± 66 ka [3]. This feature is at an elevation of approximately 3159 m and is ~7 km long. Tramway Ridge is an area of heated soil about 1.5 km northwest of the Main Crater. At an elevation of 3350 m, Tramway maintains an average near-surface temperature of 65°C [18]. Ice Tower Ridge is a series of ice towers that have aligned over degassing vents. As many as 43 of these towers form as water vapor escaping from vents condenses in the low temperatures, and these towers can be as tall as 15 m [19]. Southwest of the Main Crater is the Side Crater, which has a diameter of 250–350 m [20]. After the Main Crater formed during the Modern Erebus phase, a secondary vent became active where the Side Crater now lies. Although this once may have been part of a fissure system that runs northeast-southwest, in line with the vents of Ice Tower Ridge, it is now inactive [20].

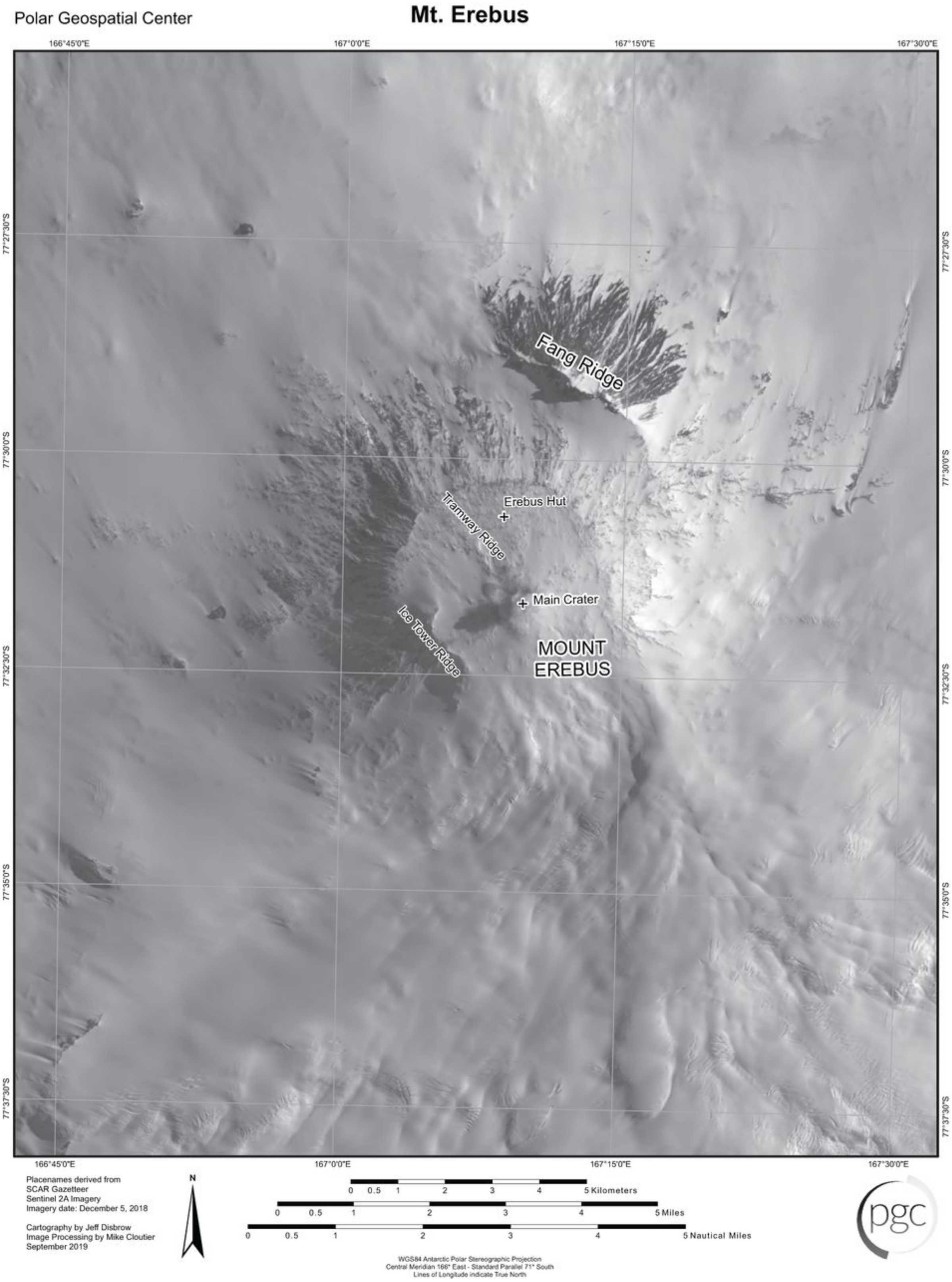


Figure 2.4: Satellite image of Mount Erebus's summit displaying the Main Crater and other geologic features. (Jeff Dishbrow, Polar Geospatial Center.)

2.2 Volcanic Features

Mount Erebus' volcanic activity is sourced from an upwelling mantle plume, suggested by a low velocity zone found up to 1200 km below Ross Island [1, 21]. The Main Crater, which is roughly 250 m in radius and 120 m deep, contains an Inner Crater (80 m radius) that is home to a lava lake of 5–15 m radius (Figure 2.5) [7]. The lava lake is a surface expression of a shallow magma body, which is formally called "Ray's Lava Lake." The lava lake maintains magmastatic equilibrium with the magma body through bidirectional flow [22]. There are two nearby degassing vents, the Active Vent and Werner's fumarole, which sometimes manifests itself as a second lava lake. These openings are the site of all past eruptions. The existence of these vents suggests that the conduit has a complex geometry at shallow depths [1].

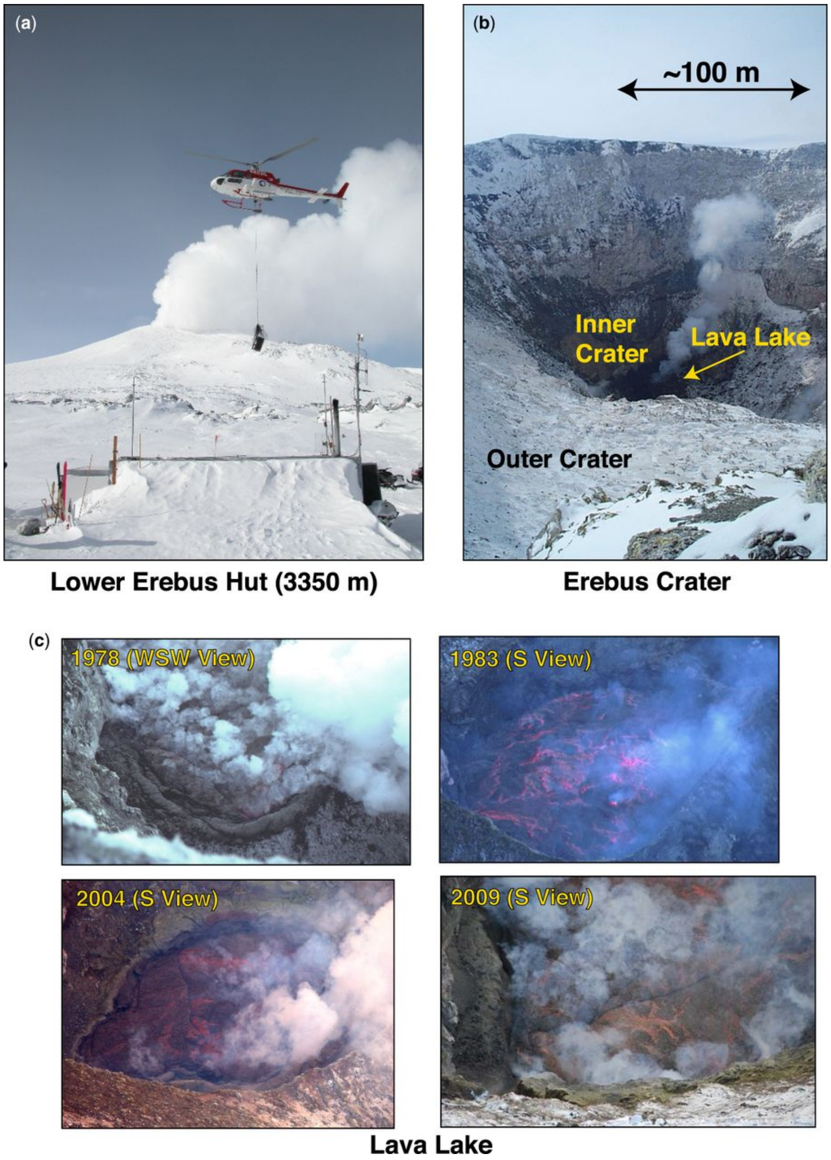


Figure 2.5: a) Southern view of Lower Erebus Hut. b) Location of lava lake as viewed from the southern rim. c) Images of the lava lake through various years. From [1].

The location of the magma body and VLP seismic source has been investigated using multiple methods. Moment rate tensor inversion of stacked seismograms demonstrates that the VLP source centroid is approximately 0.5 km northwest of the lava lake at a shallow top-surface depth near ~0.4 km (Figure 2.6) [13]. The couples and forces were concluded to be a superposition pressurization, minor shear component, and a north-dipping single force which suggests its presence in a north-dipping locale of constriction. Also of significance was the recognition of a progressively varying lag between short-period (SP) and VLP signals of ±1 s [5]. Each lag is interpreted as a representation of an elastic Stoneley boundary wave velocity change arising from small changes in the width of conduit width, which requires a narrow (<8 m) conduit. A near-summit shallow magma body is identified 500–1000 m northwest of the lava lake that reaches depths greater than approximately 500 m using backprojections of autocorrelated icequake signals [23] and inversion of P-wave arrival times [24]. These tomography results suggests that the top few hundred meters of the conduit is narrower than 10 m since it was not resolvable, which is consistent with the earlier findings and with the VLP-SP modeling of [5]. Low-seismic velocities associated with low gravimetric observations provide further evidence that the magma body is northwest of the lava lake [25]. Around 6000 meters below sea level, evidence of scattering indicate that there may be multiple deep magma reservoirs [26]. These details support two models of Mount Erebus' magma system: 1) a mantle source that is entirely attached to the lava lake or 2) ascending basanitic melt providing shallow reservoirs [27].

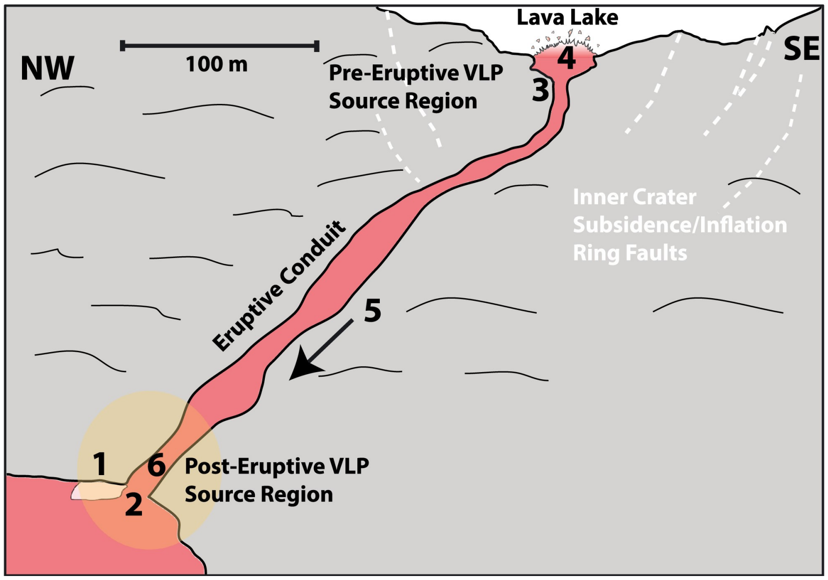


Figure 2.6: Depiction of conduit that connects the VLP source to the lava lake. Region 1 indicates where storage or transport of gas slugs occurs before eruption. 2) the beginning of the conduit that gas slugs rise through, 3) gas slug arrives at the lava lake causing preeruptive VLP signal, 4) surface explosion of gas slug which causes the short period signal, 5) elastic communication of surface explosion to onset of conduit, and 6) source region of VLP signal that occurs as the conduit establishes equilibrium. From [5].

Chapter 3

The Tomo Erebus Project

During the austral summer of 2007–2008, the ETB network was deployed, consisting of 23 intermediate-period 3-component seismometers to record passive seismic activity. This field season was useful in determining the location of the active experiment planned for December of 2008 where 100 3-component 4.5 Hz Sercel L-28 sensors were deployed for 19 days as a part of the Tomo Erebus project [9]. 21 of these seismometers were placed in a 90 km East-West line across Ross island to record 3 chemical blasts, with the goal of creating a 2-D image of the magmatic structure (Figure 3.1). The other 79 short-period seismometers were deployed approximately 300–500 m apart in a 3x3 km grid around the Main Crater to record 11 chemical blasts, with the goal of creating a 3-D model of the volcano [24]. Overall, the stations were deployed at distances between 0.46 and 48.7 km of the lava lake center. The sensors recorded 200 samples per second, and most of the stations successfully captured the shot records within their data with the exception of a few instrument malfunctions. The results demonstrated the heterogeneity of Mount Erebus through variations in waveforms and frequencies across the network [9].

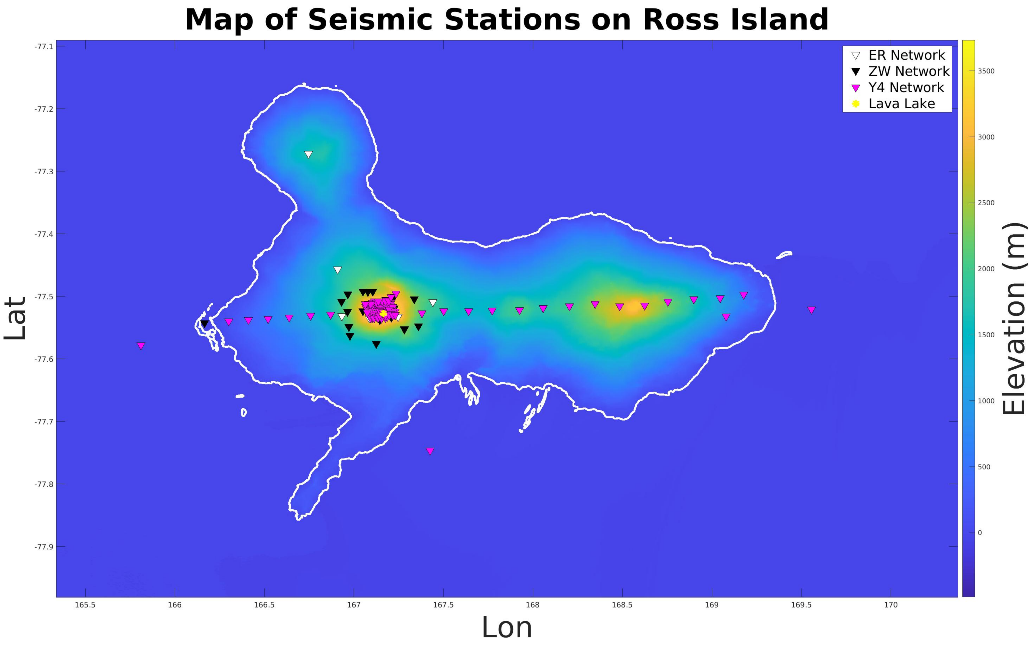


Figure 3.1: Map of TOMO Erebus era seismic stations on Ross Island. Includes the IRIS DMC FDSN code ER, ZW, and Y4 networks.

The Tomo Erebus experiment provided us with data from 99 stations that are all within 49 km of the lava lake, however, they did not all run continuously or simultaneously throughout the study period from December 10-28, 2008 (Figure 3.2). These data were retrieved from the IRIS data management system in hour-long segments. I consider a station to be operational when all three components work for the entire hour. Most stations recorded from the beginning of the study period until day 17 or 18 of the study. 61 stations were working at the start, and none worked until the very end of the study period, although station 91 came close by working until UTC hour 21 on the last day. At any given time, the most amount of stations recording was 93, which occurred during four isolated hours. Stations 10, 29, and 68 never turned on. The highest amount of working hours for a single station is 413, which was achieved by stations 33, 34, 36, 61, and 62.

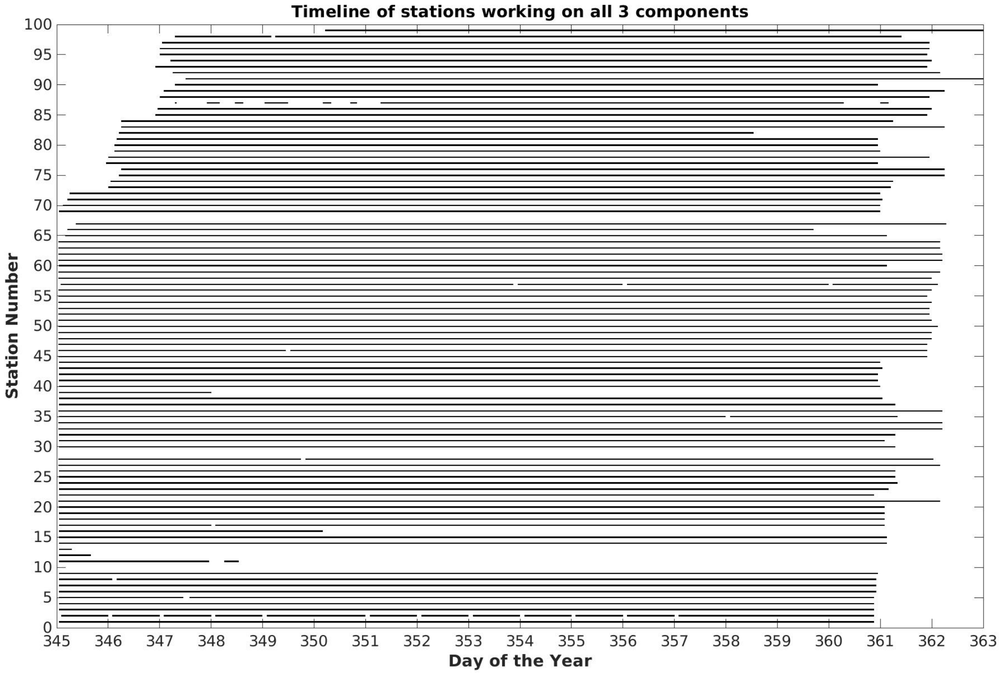


Figure 3.2: Timeline of when each station in the Tomo Erebus project was running on all three components.

Several studies have previously utilized these data. [24] used the experiment to obtain P-wave tomographic velocity images which, in conjunction with coda interferometry, identified the near-summit magmatic system 500 m northwest of the Inner Crater. [23] investigated seismic coda and found evidence of Rayleigh or Rayleigh-Gans scattering from strong frequency-dependent scattering of wavefields in the 1–4 Hz band. Because of this property, high-frequency seismograms recorded on Erebus are dominated by path, rather than source, effects. [26] was able to image the magmatic system at high resolution and depths using a matrix approach with coda, resulting in the identification of a reservoir at 2500 m, several cavities, and a chimney-like structure. In this study, the Tomo Erebus data will be analyzed for temporal and spatial patterns to categorize lava lake events into "families" and provide further characterization of lava lake eruptive activity during the deployment.

Chapter 4

Methods

This section details the theoretical background and process for each method utilized. Building a catalogue of events sourced from Mount Erebus' lava lake over 19 days requires a robust method for auto-detecting events. Our preferred technique, matched filtering, is discussed here, as well as the process for selecting templates. The next portion of our study is to analyze the catalogue for trends and similarities between events, so this section also explores the approach for clustering.

A recent similar study, executed on a larger temporal scale, was conducted by [28]. The goal of this research was to update the model of shallow conduit dynamics at Volcan de Fuego by creating a catalogue of seismic events from 2008-2015 using matched filtering, and analyzing trends in magnitude, waveform shape, and source location to determine the behavior and stability of the magma system. Like Mount Erebus, Volcan de Fuego is an open-vent volcano that experiences very-long period events due to pressurization in the conduit before an eruption [28]. Three categories of VLP events were identified during deployments in 2008 and 2009 based on the type of arrival on seismograms and gas emissions. The continuation of these events was explored in another deployment in 2012 using seismic data, time-lapse images, and infrasound records. For periods of time that an instrument was unusable, and for gaps in visual records in 2014-2015, matched filtering was used to broaden the catalogue. This method successfully yielded 79, 5, and 63 detected events in 2012, 2014, and 2015, respectively. This provided them with a catalogue on which cross-correlation was performed, indicating that event types did not come from different sources. A decrease in event magnitude over time, followed by an increase was also observed. However, the most important result comes from the sheer ability to detect these VLP events in each of the years studied, demonstrating that the source mechanism and location was continuously stable [28]. They find that the shallow system is quite complex and dynamic, while the deep conduit is stable. A key take-away from this study is that a change in the VLP waveforms could indicate instability of the deeper structure, which could present some seismic risk, however, long period signal's lack of sensitivity to small-scale structures may inhibit this application.

4.1 Matched Filtering

Matched filtering is a convenient and efficient tool for identifying seismic events that may be difficult to recognize among data with a low signal-to-noise ratio. With our abundance of data, an automated method is crucial. Matched filtering is achieved by finding cross correlation values between a template signal (x_j) and unknown signals found within seismic data (y_j) (eq. 4.1). This is done by determining the (usually energy-normalized) dot product of the two signals, where the result signifies the projection of one vector in the direction of the other in a space with the dimension of the time series length. This provides us with values that represent how similar the two signals are at any given point in time.

$$\mathbf{z_k} = \sum_{j=-\infty}^{\infty} x_j y_{j-k} \tag{4.1}$$

I efficiently implement this method by equivalently using convolution on the template signal and a time reversed unknown signal [29]. Convolution is used on two signals to compare them, and can be very easily implemented using the convolution theorem via Fourier transforms and spectral multiplication, where the time reversal of one of the time series results in cross-correlation values. To perform convolution in the time domain, we realize the discrete time counterpart of this equation:

$$\mathbf{c(t)} = x_j(t) * y_j(t) = \int_{-\infty}^{\infty} x_j(\tau) y_j(t-\tau) d\tau \tag{4.2}$$

Which is modified in the following way to get correlation values:

$$\mathbf{c(t)} = x_j(t) * y_j(-t) = \int_{-\infty}^{\infty} x_j(\tau) y_j(\tau - t) d\tau \tag{4.3}$$

To perform this operation in the frequency domain, we multiply two spectra. This allows us to compute the entire correlation between the two signals rather than performing time-domain multiplication and summation at each time lag [29]. I am looking for the presence of the template signal within the data, which is assumed to be a multiple of the template with random noise added [30]. If the two signals are similar, their convolution will become constructive, resulting in high correlation values at a specific time (Figure 4.1). Matched filtering is described in detail by [31], which suggests that it may have been used as early as 1943, and perhaps even earlier, to determine what affects signal to noise ratios in radar. By 1964, matched filtering was used in pulse-compression echo-ranging systems, communications, detecting periodic signals among noise, estimating spectra, filtering, determining impulse-response, and identifying and locating signal sources [29]. It was first used for earthquake detection in 1967 by [30] to identify surface waves at teleseismic distances, but the idea of using correlation for seismic signal detection has been considered in exploration seismology since at least 1957 [32].

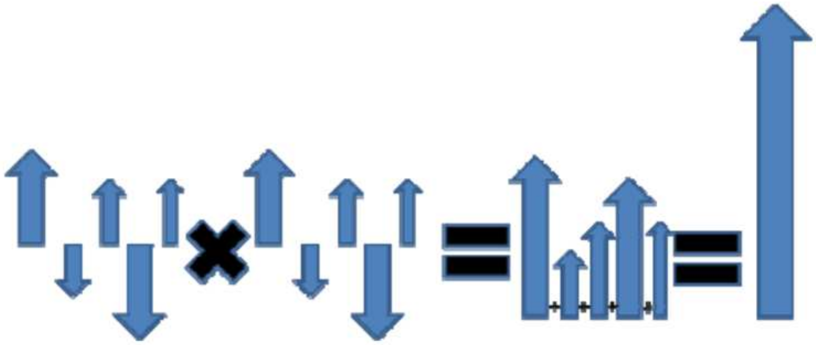


Figure 4.1: Illustration of high correlation values at a single lag value resulting from constructive signals. From [6].

Matched filtering is beneficial due to its relative location of events and resilience against noise. By comparing waveforms from the entire wavefield, similar signals can be identified regardless of noise [33]. The source of an earthquake will affect the waveform based on the heterogeneities it passes through, so matched filtering will reveal detections with the same source mechanism and location [30]. Matched filtering may in some cases allow for the extraction of further information such as radiation pattern, magnitude estimates, and energy content [30]. Additionally, matched filtering has a low rate of false detections, which reduces detection validation time greatly. This aspect can be important for disciplines such as nuclear explosion monitoring [34]. These advantages have made matched filtering a popular technique in various fields. Some examples include detecting microseismic events induced by hydraulic fracturing [6], detecting coronal loops in the sun's atmosphere [35], locating blood vessels [36], and detecting clusters of galaxies [37].

A common event detection method, short-term-algorithm/long-term-algorithm (STA/LTA), was not chosen for event detection because this method does a poor job with recovering small amplitude events. Although matched filtering is not applicable for all scenarios within seismology due to the need for a template signal, it has been shown in studies such as [38] that repeating signals may be more common than anticipated. Matched filtering was able to successfully identify 1301 of 14,000 seismic events listed in the Annual Bulletin of Chinese Earthquakes, many of which were not included in other databases because of their small magnitude [38]. Because we are interested in building a catalogue of seismic events coming from Mount Erebus' lava lake caused by small scale eruptions or other seismic activity, this makes matched filtering the perfect method for this study. The general workflow for using this method is to choose template events, filter the data, perform the convolution between template events and data on like channels, calculate median correlation values, determine event threshold, and remove duplicates from the catalogue.

4.1.1 Detecting VLP Template Events

Because we are interested in events that came from the lava lake, I ensure our template events come from the lava lake. As discussed earlier, VLP events characterize gas slugs rising through the conduit and exploding at the surface of the lava lake. The lava lake restabilizes within several minutes after an eruption [7], which appears in seismic data as a highly repeatable very long-period coda. Because of their distinct energy signature between 8 and 30 seconds, I can definitively search for these eruptive events in broadband data. Since 1999, broadband data have been collected from the volcano, with a majority of these temporary and longer-term stations clustered around the summit of Mount Erebus. Similarly to the Tomo Erebus data, hour-long segments of data are downloaded from the IRIS data management system during the Tomo Erebus experiment period. The data is detrended and divided by the sensitivity to convert the measurements to velocity. I searched for clear VLP signals using the vertical component of stations within 2.5 km of the lava lake, where a low-pass Butterworth filter is applied at 0.2 Hz. Velocity spectra are calculated using a short-time Fourier transform with windows of 900 samples and 855 sample overlap. A finite impulse response filter is applied with weights of 0.25 and the geometric mean of the spectrogram is displayed in units of dB at time-localized frequencies for each hour between December 10 and 28, 2008 (Figure 4.2). The times of VLP events indicated by the concentration of energy between 0.03 and 0.125 Hz are recorded. Although large lava lake events are identifiable in broadband data, we only use it to find templates because the short period data from the Tomo Erebus project allows us to expand our catalogue by including smaller events.

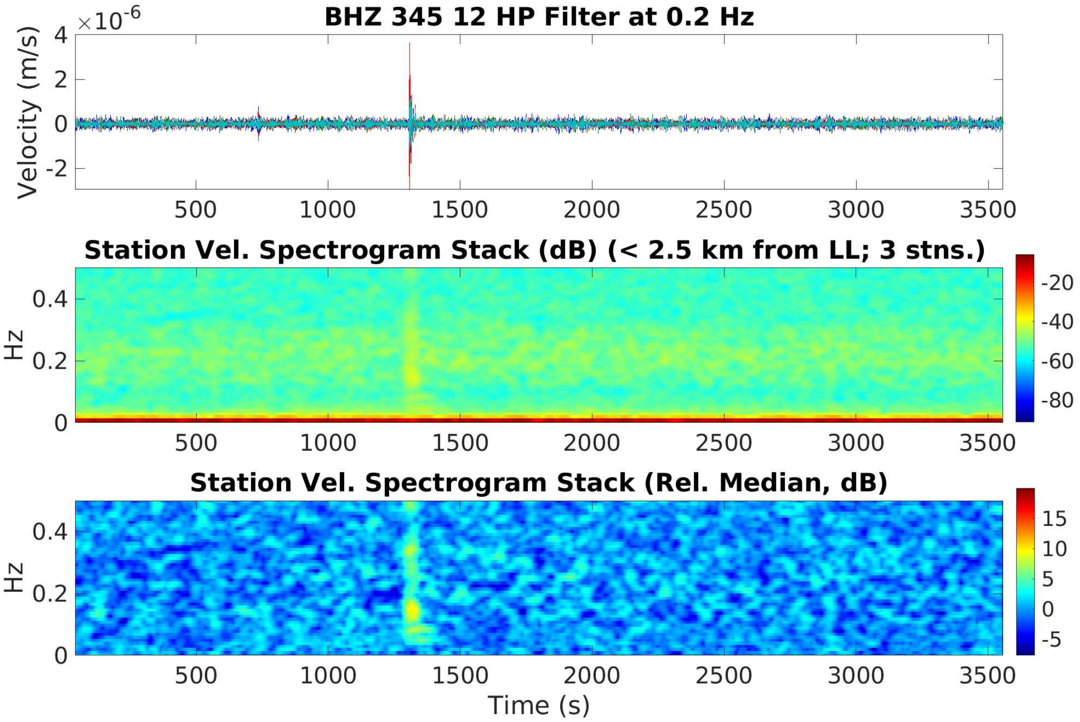


Figure 4.2: Example eruptive (as discriminated based on very long period energy; [7]) event recorded on broadband stations in ETB network. The top panel displays one hour of vertical velocity data from UTC hour 12 on day 345 high pass filtered at 0.2 Hz. This is the same event presented in Figure 4.3. The middle panel exhibits the broadband acceleration power spectral density spectrogram stack in dB (relative to 1 $(m/s^2)^2/Hz$) for these data, and the bottom panel is the broadband spectrogram relative to the median spectrum of the middle panel in dB. Spectrograms are created using the spectrogram function in MATLAB with a window length of 900 samples, an overlap of 95% (855 samples), and a sample rate of 10.

This procedure resulted in 17 detections, which are used to extract short-period templates for matched filtering to discover other seismic events coming from the lava lake. Figure 4.3 displays an example of a VLP event on a short period station. These events occurred between days 345 and 360 of 2008, so none were identified on broadband stations during the last three days of the study. Over 16 days, these events occur on average about once a day, with a maximum of four VLP events on day 350, during which two events occurred as closely as 4 hours apart. The longest time between detected eruptive events is 2.5 days.

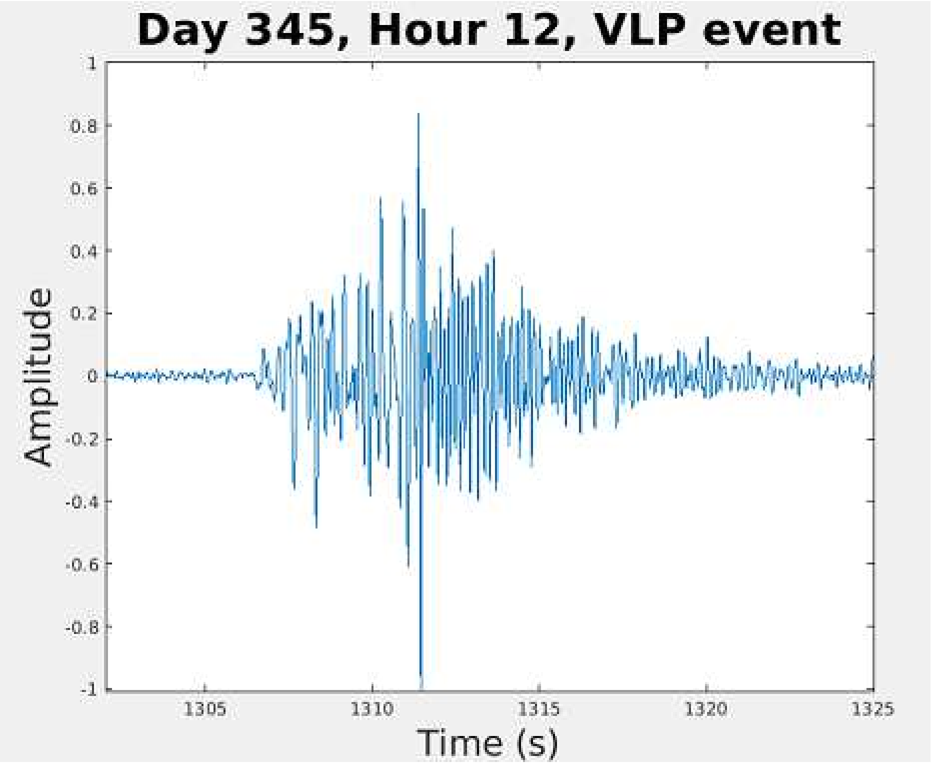


Figure 4.3: Example VLP event on ETS station 52, vertical component. This event occurred during day 345, UTC hour 12, and is displayed with a normalized velocity amplitude.

4.1.2 Auto-Detection of Lava Lake Events

Matched filtering is applied to 19 days of seismic data with each of the 17 VLP event templates on like channels by employing the *conv* function in MATLAB, which performs a fast convolution via spectral multiplication (utilizing the convolution theorem). VLP template events are searched for in hour-long segments of data, constituting the "unknown" signal. As mentioned before, using the spatial wavefield as sensed by the network, made up of 99 3-component short period seismic sensors, ensures that detected events come from the lava lake because the signal move-out with distance, and other features of the wavefield, must match that of the template event at many stations.

A 4^{th}-order Butterworth band-pass filter is applied to the data between 1 and 20 Hz. The data are demeaned and detrended. I then apply convolution on each hour of data where the vector is flipped to account for the time reversal, with a 20-s window of each VLP template event. A simplified version of this procedure is demonstrated for just one component on one station using only 30 minutes of data as the unknown signal (Figure 4.4). The template event in this example occurred on day 345, UTC hour 12, at ~1300 seconds, so the output correlation value with a red asterisk represents a successful self-detection of the template event.

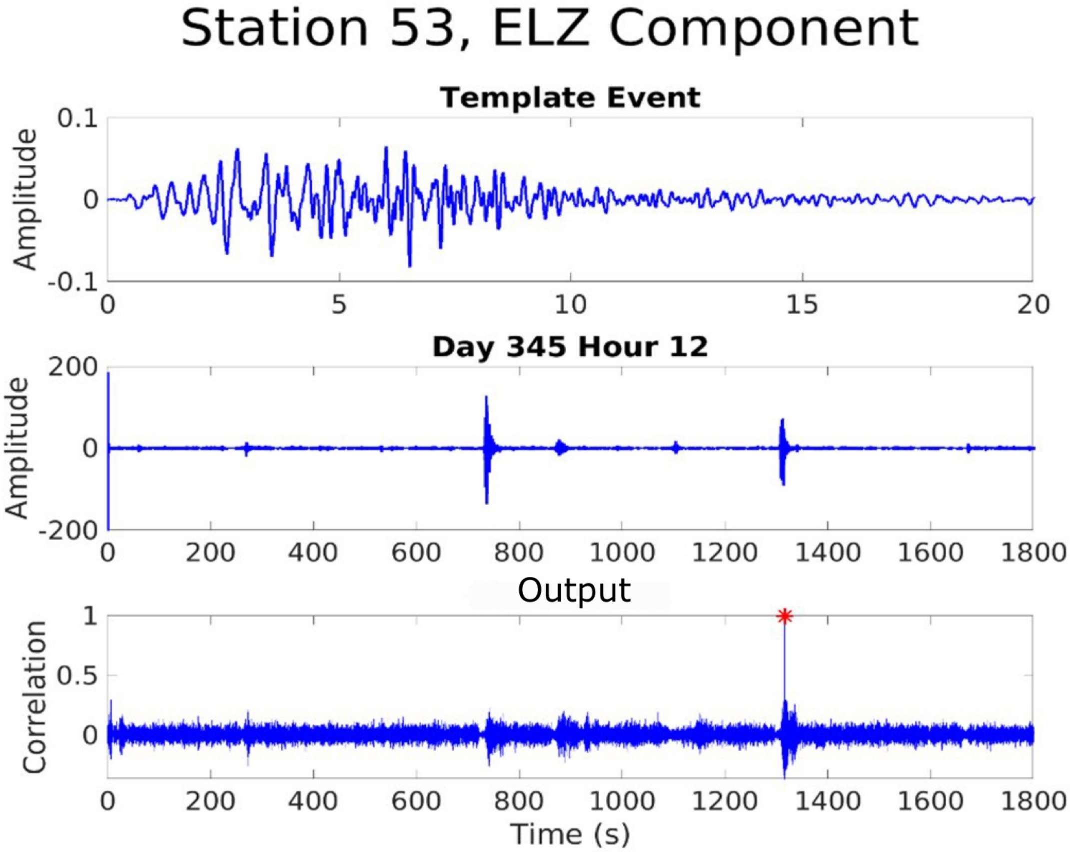


Figure 4.4: A single-station example of matched filtering using the vertical component of station ETS53. The 20-second template event is displayed on the top panel. This is time-reversed, and convolved using the *conv* function in MATLAB with 30 minutes of data from UTC hour 12 on day 345, shown in the middle panel. The resulting correlation values are displayed in the bottom panel. The template event occurred during day 345, UTC hour 12, at ~1300 seconds, so the red asterisk at a filter value of one signifies a self-detection.

This is performed on like channels for up to 99 stations and each of the three (V, N, E) components, so each channel has a "unique" template. Only channels that were operating during both the time of the template event and the entire specified hour are used. The number of stations used for detections ranges from 42 to 93. The moving averaged root mean square (eq. 4.5) is used to normalize the resulting correlation values (eq. 4.4) since the amplitude of events is not important. This tells us how similar a waveform is to the template waveform on a scale of 0 to 1. If the template event is convolved with itself, we will get a correlation value of 1 like we observe in the single station example (Figure 4.4).

$$\mathbf{y}_k = d_i * v_j \tag{4.4}$$

$$\mathbf{yn_k} = \frac{y_k}{\sqrt{{d_i}^2 * u}} \tag{4.5}$$

where d_i represents the column of unknown data that is recorded on the same channel as column j of v, which is the template data. Matrix d has 720,001 rows and matrix v has 4,001 rows, based on the number of samples recorded in the duration of these data. y_k is the output of correlation values, which has the same number of rows as d. The maximum number of columns, k, will be different for each hour's output depending on how many stations were running. u is a vector of ones that is the same length as the template event.

I take the median of normalized correlation values (yn_k) at each sample across all channels and find the average. I chose the event detection threshold to be 3 standard deviations above this value. Although this threshold may be objectively low, as long as correlation values are relatively higher than those around it, it is considered significant [34]. Events are determined by peaks in the median correlation value above the threshold with a prominence (height of peak relative to the height of the signal) of at least 0.1 and must be at least 10 seconds from another peak. To ensure that the results yielded real seismic events, I tested the threshold with higher values using 4, 5, and 6 standard deviations. However, this did not reduce the number of detections because each clear detection peak is an extreme outlier in that it is many standard deviations above the average (Figure 4.5). Although many samples for each detected event will be above the threshold, the local maximum is recorded as the onset time of the event. The day, hour, and start time of the event is documented.

The start time is in seconds of the specified hour, and is recorded as 20 seconds before the triggered detection time (when the maximum correlation between data and template occurs) to ensure the entire event is captured. The catalogue also includes the maximum correlation value between the event and template, the template number which triggered the detection (or the template number with the highest correlation value if detected by multiple templates), maximum amplitude of the event on station ETS52, number of stations used for detection, and the number of templates that detected the event. Events that are not entirely within one hour of data are not included.

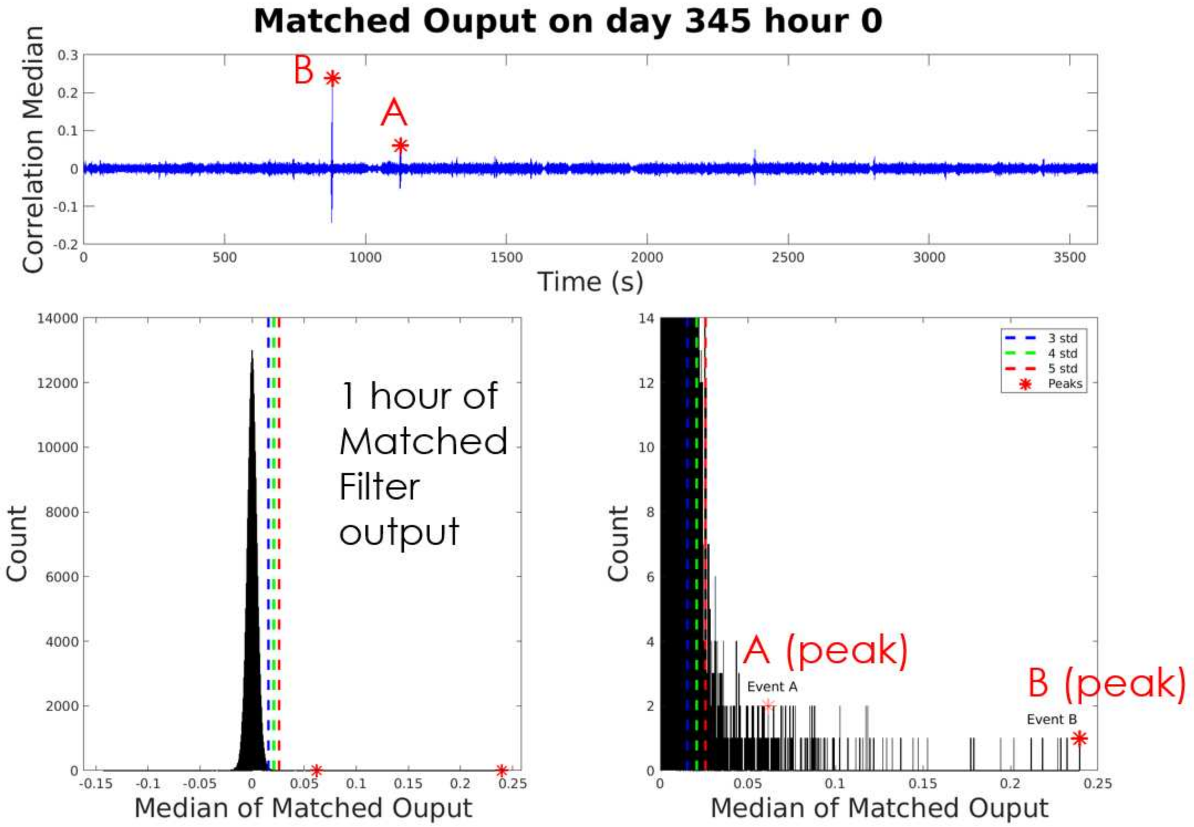


Figure 4.5: The top panel displays multichannel (61 stations, 182 channels total) median correlation values for one hour (day 345, UTC hour 0). The template used in this example is the same event used in figure 4.4, occurring on day 345, UTC hour 12. Peaks A and B indicate correlation value above the threshold that are identified as events. The bottom left panel shows the distribution of correlation values for the same hour. The blue, green, and red dashed line indicate 3, 4, and 5 standard deviations above the average, respectively. The bottom right panel is a zoomed in version, where the two detected events can be seen many standard deviations above the average. The other peaks present that are above the threshold are from the matched filter detection output side lobes.

After this process is repeated for each template event, a total of 10,332 events are identified. On average, each template detected 607 events. A template from a VLP event on day 356, UTC hour 10 detected only 299 events, which was the lowest out of all of the templates. The VLP template event on UTC hour 5 of day 350 identified 1,154 events, which is the most of all templates. Events that were detected by multiple template events are determined using a criterion of having start times within 5 seconds of each other. The resulting duplicate detections are removed. A catalogue is created with the information for each of the remaining unique events. This catalogue, which will be referred to as catalogue A, contains 1,741 events. 124 of these events were detected by all templates, and 488 events were detected by a single template. The average number of templates that identified each event in catalogue A is six. A record section of each event is plotted and examined manually to confirm the presence of a clear seismic event. Events were discarded if the signal was buried in an abundance of noise to the extent that its amplitude and other waveform characteristics were unable to be accurately measured. The final culled catalogue of higher signal-to-noise events from this process consists of 819 events. This catalogue, which will be referred to as catalogue B, has 47% of the events from all detected events in catalogue A. 111 of these events were detected by all templates, and 200 events were detected by just one template. On average, events from catalogue B were detected by eight templates. I experimented with repeating this process using matched filtering of (analytic signal-derived) waveform envelopes, but this technique was not discriminatory enough to produce additional useful lava lake event detections.

In earlier stages of the study, an STA/LTA algorithm was used to build catalogue B from catalogue A, rather than manual inspection. Since catalogue B is intended to be used for clustering and further waveform analysis, I thought it may be an adequate technique to find the largest amplitude events. A short-term window of 0.5 seconds, long-term window of 6 seconds, and triggering threshold at a short-term/long-term amplitude ratio of 3 was used. This yielded a catalogue (C) with only 385 out of the original 1,741 events, meaning that 434 clear, small amplitude events that were ultimately included in catalogue B had been missed.

4.2 Clustering of Similar Events

I utilize cross-correlation once again to examine how similar the detected lava lake events in catalogue B are with each other to assess the tendency for repetition of specific event types. Near-repeating seismograms at high frequencies would indicate repeating regions of activity within the lava lake system. My workflow is first developed using just the 17 template events. I compare each event to all others on 5 of the closest stations that ran continuously through the study period (Figure 4.6). This is once again performed on the three seismic components, each treated as individual channels.

The data are filtered using a second order butterworth low-pass filter at 4 Hz, and is decimated using a factor of 10, which lowers the sampling rate from 200 to 20 Hz. Decimation is another form of low-pass filtering that reduces the chance of aliasing using a low-pass Chebyshev Type I IIR filter of order 8. The mean is removed from the signal, and 60-second windows of data that begin 10 seconds before the recorded start time of events are captured. Our data are stored in hour-long segments, so if an event begins within the first 10 seconds of an hour, the window is the first minute of the hour. Similarly, an event beginning within the last 60 seconds of an hour operates with the last minute as the window.

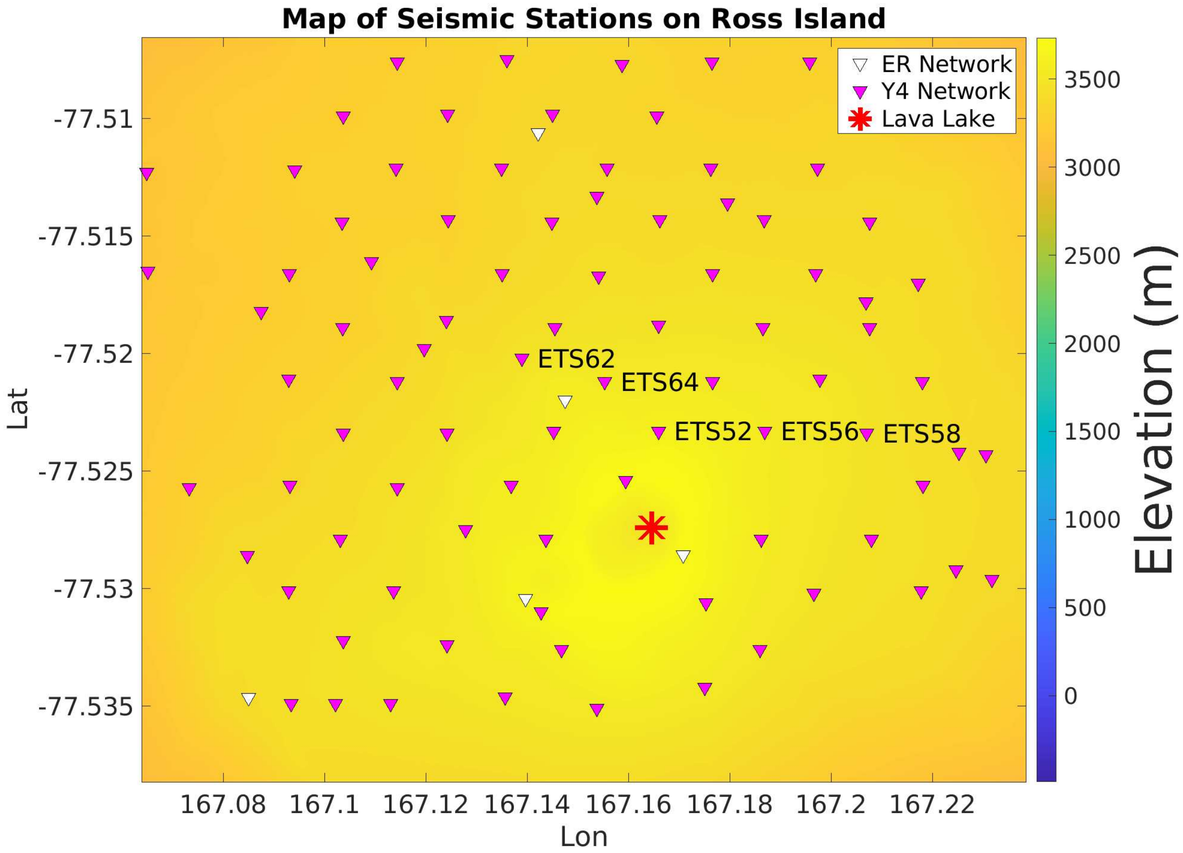


Figure 4.6: Zoomed in map of short-period stations on Mount Erebus. The 5 closest stations that ran continuously through the study period and are used in the clustering process are labeled.

I could immediately begin to correlate each event pair on like channels and form clusters based on low dissimilarity values, however, I want to improve our comparisons by lining up our waveforms sample-to-sample. I started by doing cross correlation with two events on one channel using the *xcorr* function in MATLAB, so multiple correlation values are obtained as the second time series continues to shift by one sample. The time lag that results in the maximum correlation value is saved, and weighted by the square of the correlation value. This process is repeated for each event pair on each channel. These weights are inserted into a system of equations. Low weight constraints (<0.2) are removed, and the best alignment shifts are solved for using iteratively reweighted least squares (which minimizes the L_1 norm metric [39]). This method allows us to find the best lag for each waveform without the solution being affected heavily by outliers. I circularly shift each time series, so if a waveform shifts to the left, then the data at the beginning will move to the end and we maintain the length of our vector (Figure 4.7).

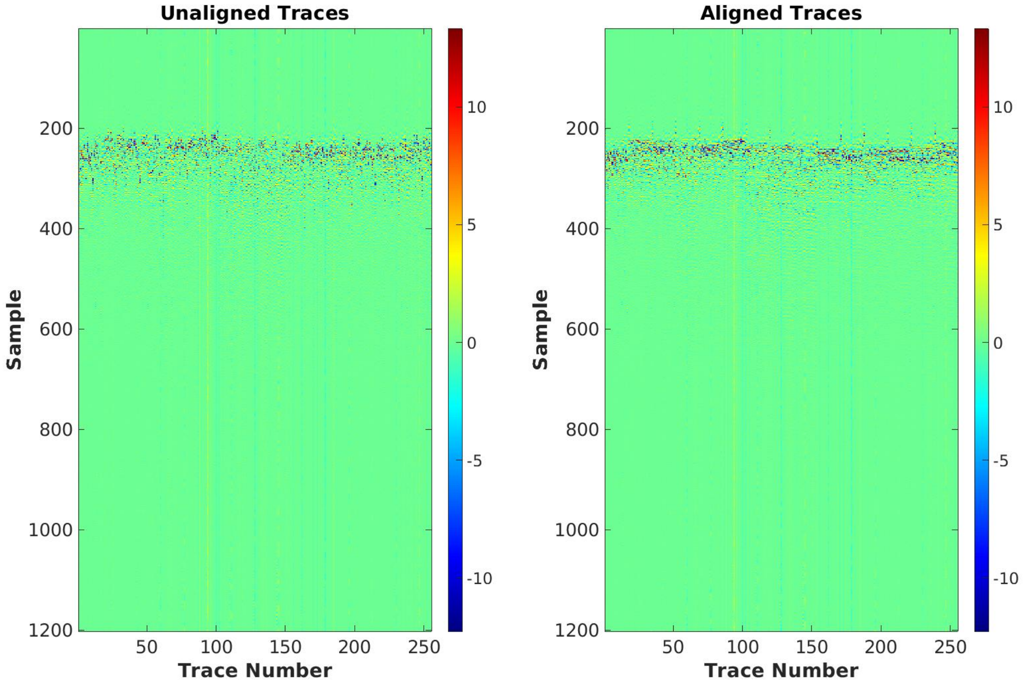


Figure 4.7: Display of seismic amplitude traces in nm/s for each VLP template event pre- and post-alignment. Note that high amplitudes for each trace line up rather than being scattered through time.

Now that the waveforms are aligned, I calculate linear correlation coefficients between all events on each channel. I take the average across all channels to get one correlation value for each event pair (Figure 4.8). A dissimilarity matrix, obtained by subtracting the average correlation matrix from one, is used to obtain an agglomerative hierarchical cluster tree (Figure 4.9). These clusters are determined using a Ward method to measure the Euclidean distance between clusters. The Ward method evaluates the inner squared distance (squared Euclidean distance between two points) between each pair of events, and links the pair with the lowest distance. This provides a reasonable measure for how similar (or dissimilar) the events are to each other, but the cutoff threshold will define which events fall into each cluster. I tested 191 threshold values between 0.1 and 2. I perform the clustering with each of these thresholds and document the number of groups that are formed with each one. The vector containing these quantities is smoothed twice, and then the twice differentiation is smoothed. I choose a final threshold at 0.05 above the maximum of this computation for assembling clusters. The clustering for template events is performed with a threshold of 0.97.

Once this clustering workflow was developed, it was applied to the event catalogue. The earliest implementation used catalogue C, which originally yielded 46 clusters. Because these events are linked through a hierarchical tree, some clusters may hold miscellaneous waveforms where each event may be slightly similar to only one other event in the cluster. The events from clusters that were more diverse were reclustered to create better correlated clusters, resulting in a total of 73 clusters. To prevent this from skewing our results with catalogue B, I only consider clusters to be those with an average interevent correlation greater than 0.8. The VLP events do not cluster very strongly. Five clusters were formed out of the 17 events, however, none of them had interevent correlations greater than 0.8. Each cluster had only two events, and the interevent correlations ranged from 0.06 to 0.513. Using this new criterion, I apply the clustering to catalogue B with a threshold of 1.64.

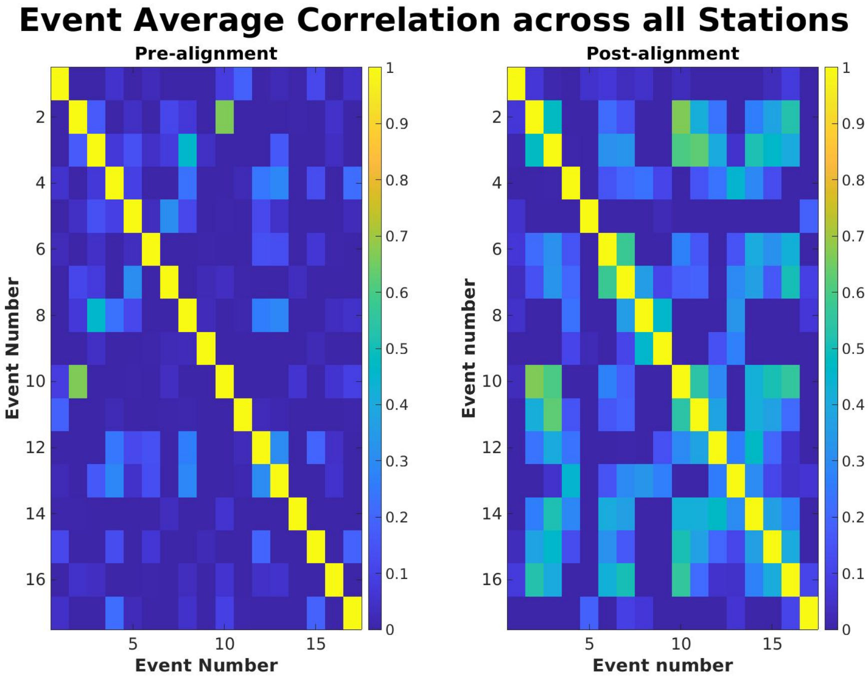


Figure 4.8: Comparison of average station correlation between VLP template events pre- and post-alignment. Correlation values increase after alignment is improved.

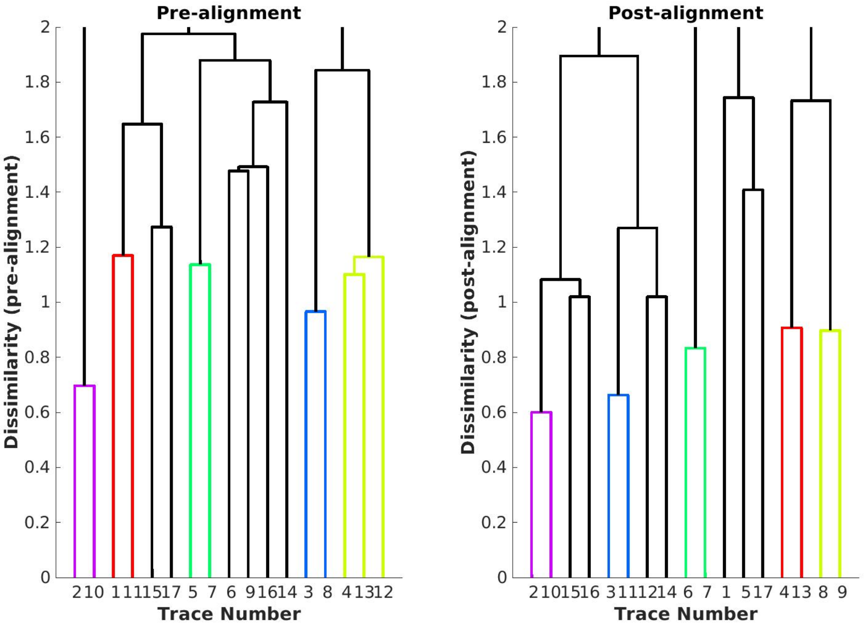


Figure 4.9: Comparison of dendrograms for VLP template events pre- and post-alignment. Proper alignment increases correlation of waveforms and their dissimilarity values decrease.

Chapter 5

Results of Lava Lake Activity Using Matched Filter-Detected Events

I analyze catalogue B for distribution of event size and timing and clusters of repeating events. I investigate these properties for trends in occurrence statistics such as the Gutenberg-Richter relationship and Poissonian distribution.

According to the Gutenberg-Richter law, the magnitudes of earthquakes are distributed exponentially [40]. This relationship is described as:

$$\mathbf{log}_{10} N(m) = a - bm \tag{5.1}$$

Where m is the magnitude, N is the number of earthquakes occurring at that magnitude, a is a constant which depends on the seismicity rate, and b is a scaling parameter describing a power law falloff of events with increasing magnitude [41]. Although the exact correlation between b values and its physical meaning has various interpretations, estimations using maximum likelihood for many seismic processes are typically around one, but can vary by location and with time [40]. For instance, in volcanic regions with small events, b is commonly between 1 and 2 [42]. This magnitude-frequency relationship is demonstrated in almost all cases, even with varying ranges of magnitude, locations, and time [42]. Here, I use velocity amplitudes as proxies for magnitude, and evaluate if the distribution of amplitudes of events in the Mount Erebus lava lake catalogue (B) obey the Gutenberg-Richter relationship.

The Poissonian distribution is a probability estimate of how many events will occur in a specified amount of time [43]. A Poissonian process occurs when events occur independently of each other at some rate (or equivalently, average time between events) [44]. Mathematically this can be represented as a probability density function characterizing the number of events observed in a unit time interval

$$\mathbf{P}(X=k)=\frac{\lambda t^k e^{-\lambda t}}{k!} \tag{5.2}$$

where the number of events observed in an interval of length t is a discrete random variable, and λ is the Poissonian rate parameter specifying the expected number of events per unit time interval [45]. The expected value of X is $E[N(t)] = \lambda t$. The time between events for a Poissonian processes has an exponential random distribution with mean $1/\lambda$. I compare the distribution of interevent times observed in our catalogue (B) to an exponential random distribution to determine if our seismically detected lava lake events at Mount Erebus can be approximately described by a Poissonian occurrence model.

5.1 Catalogue Results

The initial catalogue (A) contained 1,741 detected lava lake events based on matched filtering with VLP events. Our final catalogue (B) contains 819 inferred clear, high-signal-to-noise lava lake events. These small events all have vertical amplitudes of less than 10^{-4} m/s and last ~10–15 seconds (Figure 5.1). The peaks in amplitude demonstrate that I have detected true events. These events are between 6.5 and 172 (!) standard deviations above the triggering threshold, for an average of 23.5 standard deviations. The average maximum correlation values between events in catalogue B and template events is 0.139.

When examining the cumulative number of events in catalogue B throughout the study period, I notice that events are occurring at a relatively steady rate for the first ten days at an average of 49.4 events per day. During December 20 (day 355 of the year), there is a distinct speed-up to 198 events in one day (Figure 5.2). This is quickly followed by a slow-down in the average rate to 21.16 events per day for the last six days of the catalogue. No events were detected on the last two days of the study period. The average rate of events for the entire study period is 48.2 events per day, so the speed-up and slow-down periods essentially offset each other to maintain the original event rate.

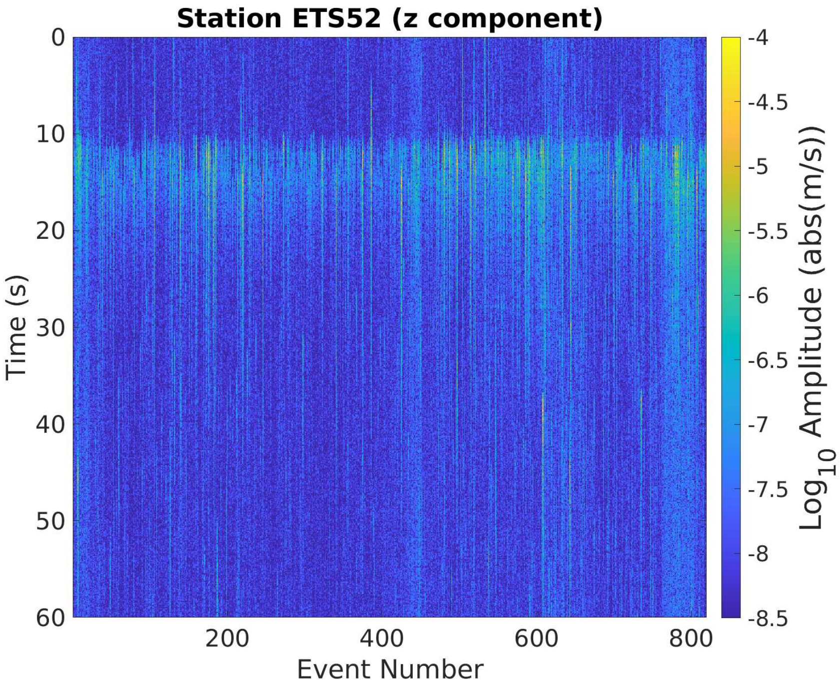


Figure 5.1: The amplitude of the z-component waveform for each event in the high signal-to-noise catalogue (B). Amplitudes are recorded on the ELZ component at station 52.

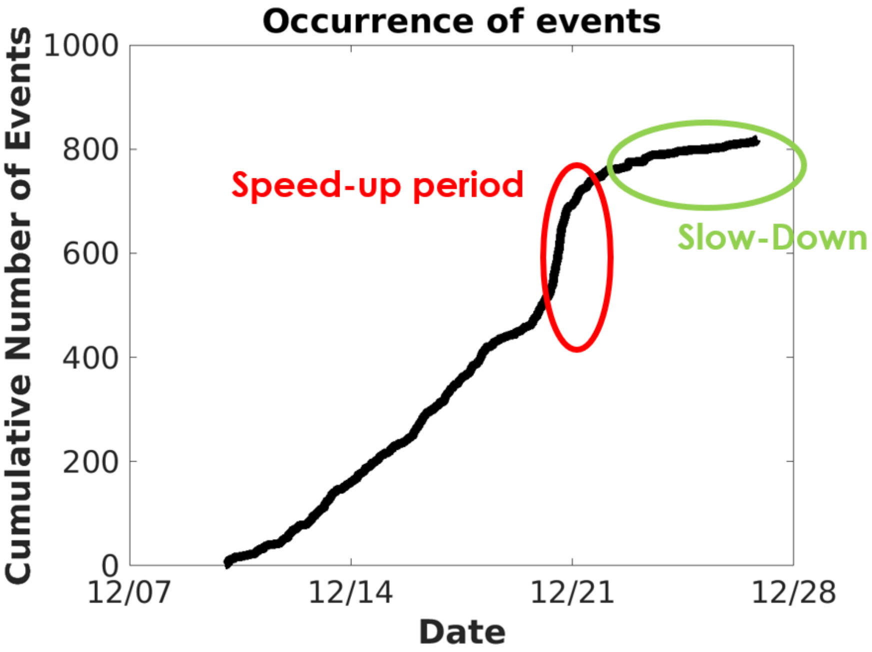


Figure 5.2: Cumulative plot of occurrence of events over time.

The total vertical-component "energy" released over the study period in our detected events is ~3.74x10^{-6} (m/s)^2 (Figure 5.3). This pseudo-energy is estimated by summing the square of the velocity recorded on the vertical component of station ETS52 during each event. On the day of the speed-up period, the greatest number of events per hour was observed (23 events during UTC hour 14). During the slow-down period, only 127 events occurred over six days. This is roughly half of the average rate of events for the entire study period. The minimum number of events recorded in one day was 9, which happened on both day 359 and 360. Throughout the slow-down period, ~3.37x10^{-7} (m/s)^2 of energy was released, which is about 9% of the total energy release (Figure 5.4).

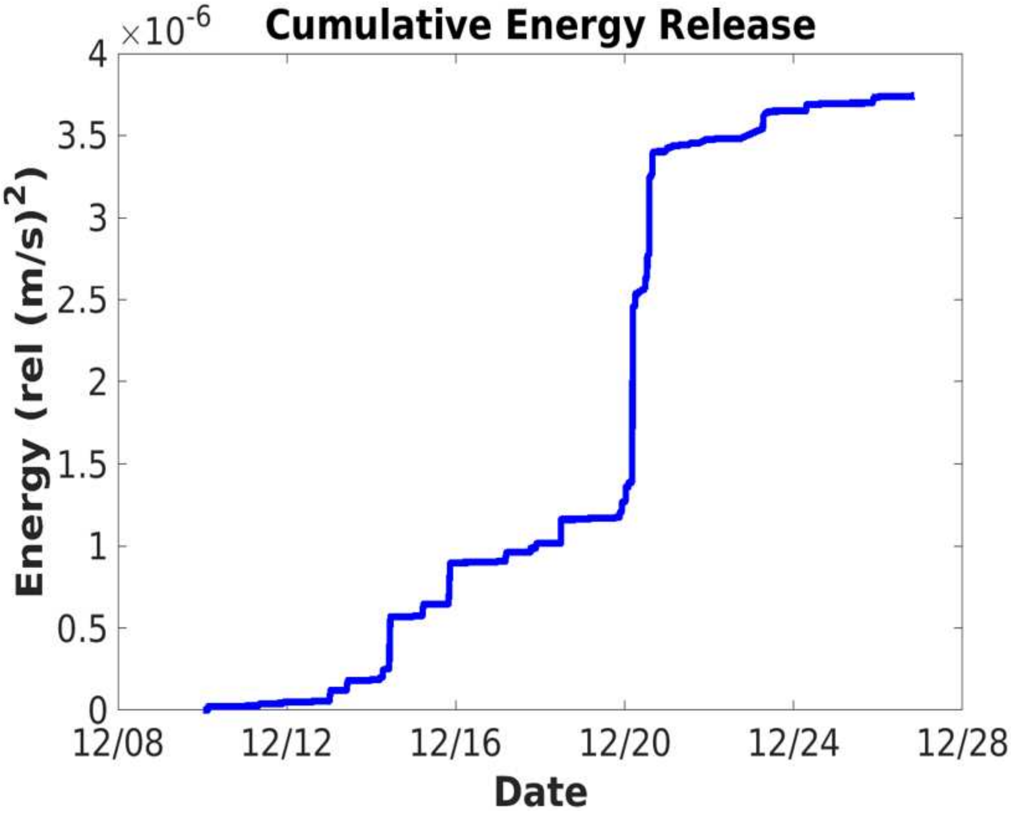


Figure 5.3: Cumulative plot of velocity seismogram "energy" released over time on the vertical component of station ETS52, calculated by summing the square of each event's velocity time series.

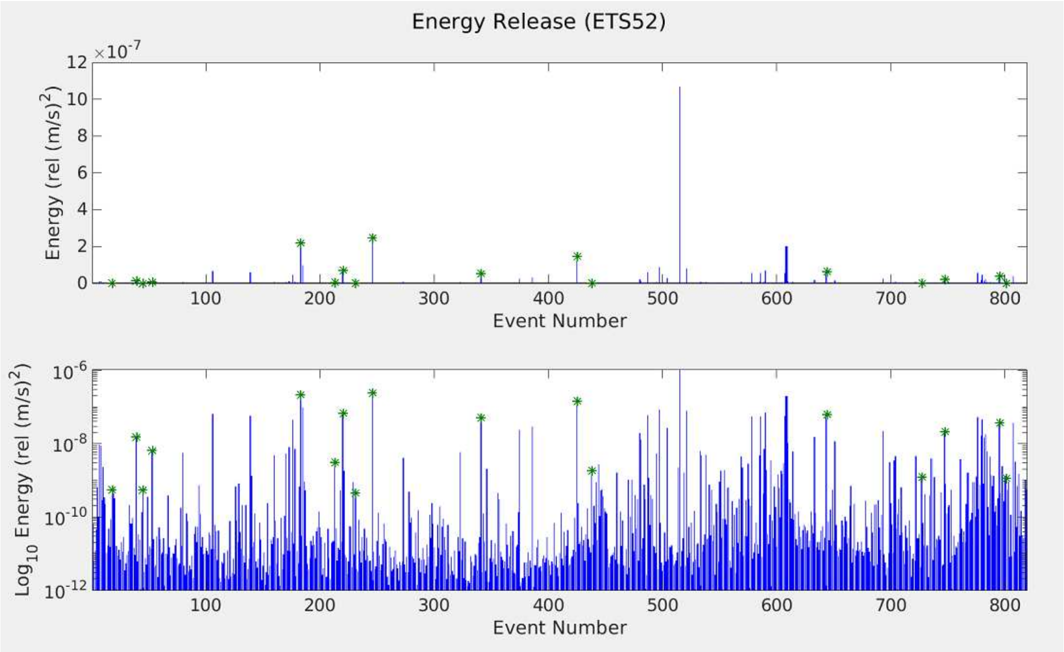


Figure 5.4: Event-by-event pseudo energy ($(\mathrm{m/s})^2$) released through time as observed in vertical component velocity seismograms at station 52. Top panel displays amplitudes plotted on a true scale, bottom panel is displayed on a logarithmic scale. The green stars represent VLP template events.

5.2 Event Clusters

I find that most of the events in catalogue B are unique below 4 Hz. There are originally 235 clusters, however, when considering "families" to be only those with interevent correlation > 0.8, there are only 17. Ten of these clusters have just two similar events, but there are families with up to 4 and 5 events. In total, just 46 events make up these clusters, which is 5.6% of catalogue B. Figures 5.5 and 5.6 demonstrate the similarity between waveforms of events in a cluster. The first cluster includes events with impulsive arrivals, and the second cluster is an example of longer period events with an emergent arrival.

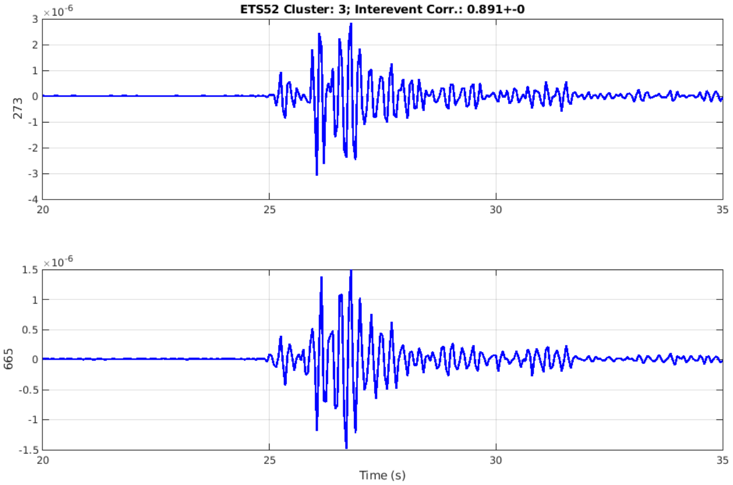


Figure 5.5: Cluster 3 contains two events with a small, impulsive first arrival and larger secondary arrivals.

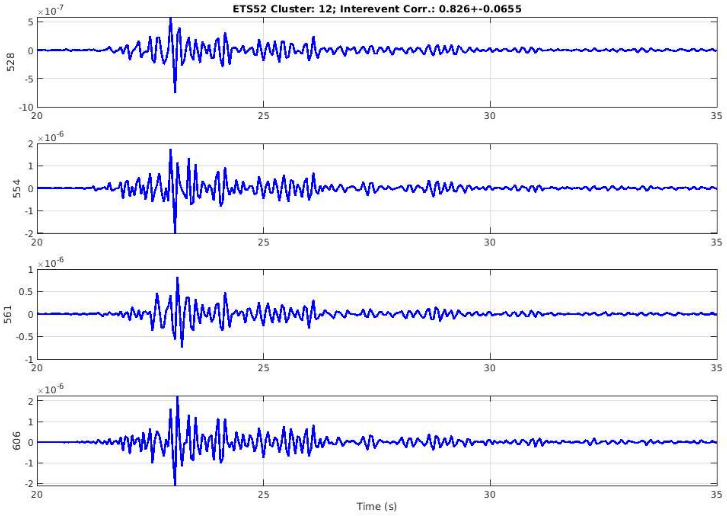


Figure 5.6: Cluster 12 contains four events with an emergent first arrival.

I investigate if the events in a particular cluster are caused by the same mechanism by reviewing the timing of the events. If they occurred around the same time and place, perhaps an ongoing process is causing the repeating events. I determine when each of the clustered events occurred through out the study period, and found that there was a not a strong correlation with time for most clusters. However, we notice that 7 of the families do cluster in time. Figure 5.7 zooms into a ~38 hour period where some of these clusters occur. All of the events in these temporal clusters transpire within less than 30 minutes of each other, but some events occur within as little as 25 seconds of each other. Clusters 1, 14, and 15 contain two events each that occur within 15 minutes, 29 minutes, and 27 seconds of each other, respectively. Clusters 9, 11, and 12 hold between 3 and 4 events each, but only 2 of each are clustered in time. Cluster 16, containing 5 events, is especially clustered with three events taking place within a 9 minute window on December 24^{th}.

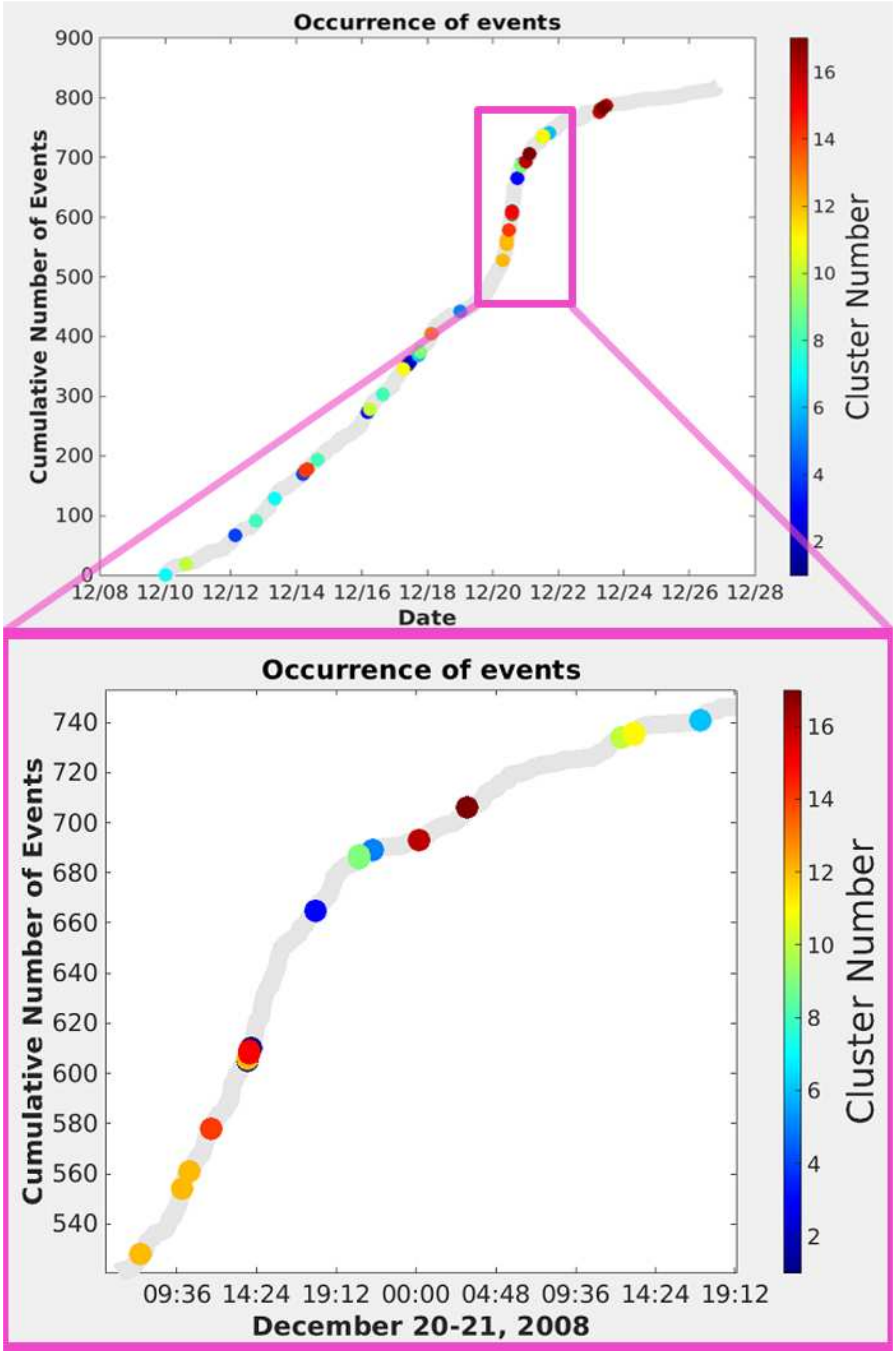


Figure 5.7: Cumulative plot of occurrence of events over time with each cluster of events color coded. The bottom panel is a closer look at the speed-up period where many of the cluster events occur.

5.3 Event Size Distribution

I examine the distribution of observed event size, estimated as the maximum amplitude (m/s) of each event on the vertical component at station ETS52. I compare our distribution to a Gutenberg-Richter power law process with a=-6.6 and b=1.6. The a value is arbitrarily set to align with the position of the observed amplitude distribution. The b value is computed using iteratively re-weighted least squares to solve for a minimum L_1-norm fit between the center values of each bin in the amplitude histogram and their counts. I find that our events generally follow this power-law size frequency distribution (Figure 5.8). I also observe an excess of small amplitude events. The average event size was 2.32x10^{-6} m/s. The smallest detected event was 6x10^{-8} m/s which occurred during hour UTC 18 on day 346. The largest occurred on day 355 during the speed-up period with an amplitude of 1.57x10^{-4} m/s.

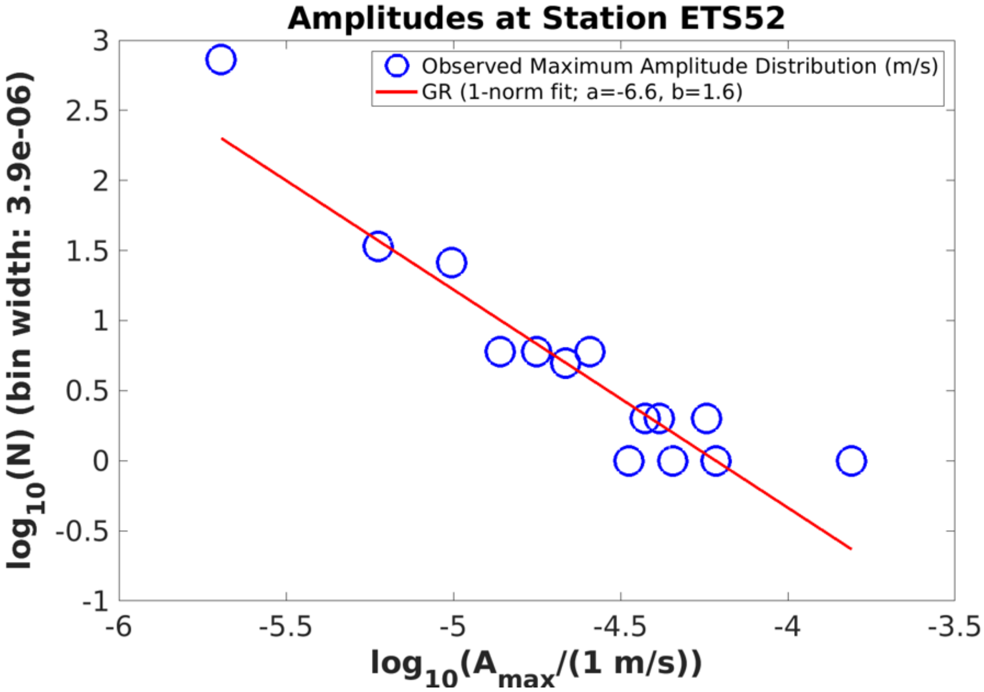


Figure 5.8: Distribution of maximum amplitudes (m/s) for all events in catalogue B recorded on the vertical component at station ETS52.

When analyzing the distribution of event amplitudes through time, I notice that there are not as many small events at the end of the study period (Figure 5.9). The observation of these bigger events agrees with the amount of energy that we observe being released during the slow-down period in spite of the decline in the number of events. The minimum event size detected changes around December 22nd. Prior to this, the smallest event detected was $6\text{x}10^{-8}$ m/s, and on average, the smallest event detected each day was $1.02\text{x}10^{-7}$ m/s. After December 22nd, the average daily minimum event size was $4.14\text{x}10^{-7}$ m/s, with no events below this magnitude. The smallest events detected after December 22nd are about 4 times larger than the smallest events detected before this date.

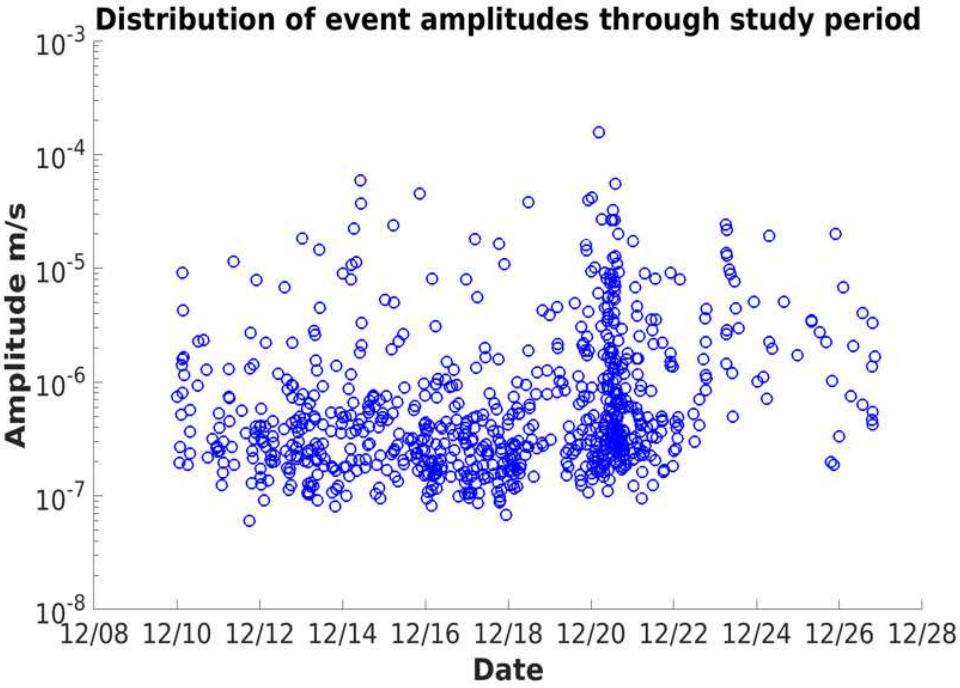


Figure 5.9: Distribution of maximum event amplitudes (m/s) for all events in catalogue B through time. The speed-up period is manifested as a condensed amount of points on December 20^{th}, and the slow-down period shows a decrease in small amplitude events.

5.4 Event Timing

I analyze the distribution of interevent times for the entirety of catalogue B, which spans 16.9 days. There is an excess of short (<20 minutes) interevent times and a small peak around 31 minutes (Figure 5.10). I compare the observed distribution to an exponential random distribution, which describes a Poissonian process. For 819 events over 16.9 days, the average number of events (λ) is 48.5 per day. The observed distribution is not as smooth as the exponential random distribution for this rate, demonstrating that our events occur at time varying rates, but is quasi-Poissonian. The shortest amount of time between two events is 5.4 seconds, and the longest amount of time between two events is 11.1 hours. The shortest interevent time surprisingly did not occur during the speed-up period, but on day 346. The longest interevent time did occur during the slow-down period, from UTC hour 18 on day 357 to hour 6 on day 358.

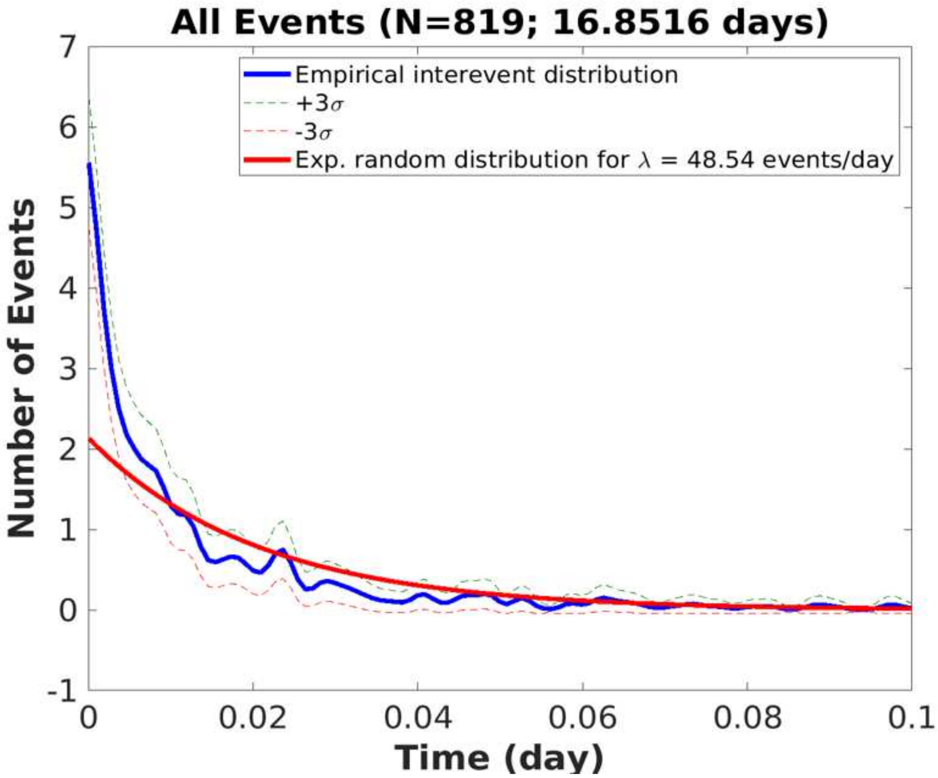


Figure 5.10: Empirical observed interevent distribution (blue) and exponential random distribution (red). For all events in catalogue B, the average Poissonian rate is 48.5 events per day. The dashed lines represent ±3 standard deviations.

I examine the Poissonian rate of events for each individual day. The speed-up period is once again visible in this illustration, with an average of 204 events in a one-day window (Figure 5.11). The second highest average occurred on day 351, at 69.2 events. The lowest average is observed on day 359, with a rate of 8.19. Because of the varying interevent time, when each day is analyzed individually, the empirical observed distribution has much better agreement with the exponential random distribution.

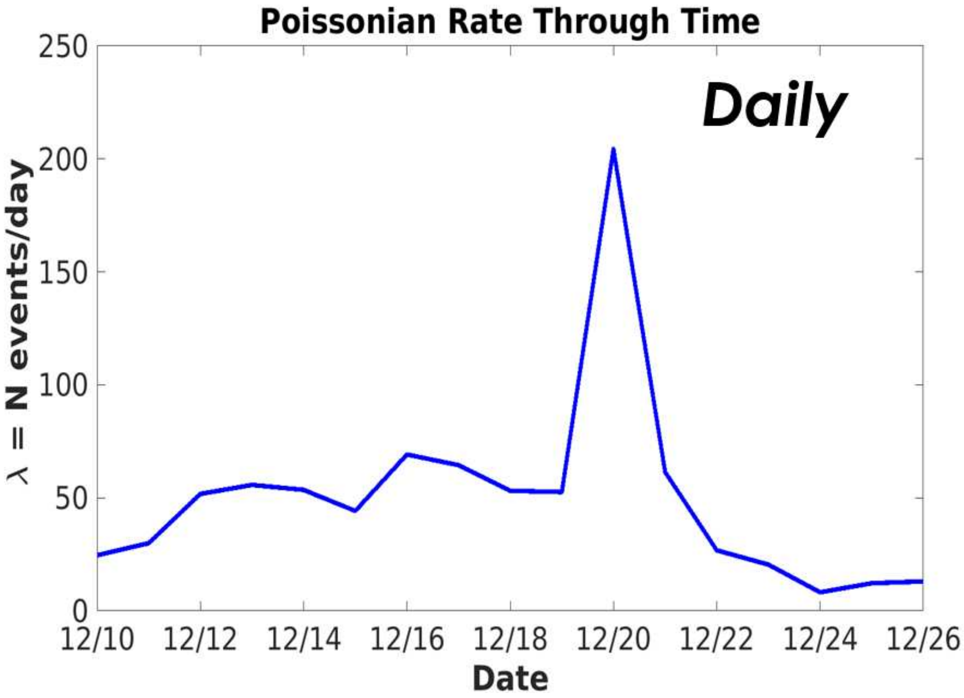


Figure 5.11: Average Poissonian rate of events for each day in the study period. The peak represents the speed-up period.

5.5 Speed-up Period

I take a closer look at the speed-up period to evaluate for a change in statistics. First looking at the size distribution, I noticed the preservation of the excess of small events (Figure 5.12). The average event size is larger than the average size of the entire study period, with an amplitude of 4x10^{-6} m/s. The smallest detected event was 1.1x10^{-7} m/s, which is larger than the smallest event during the entire study period, and the largest amplitude during this time is the largest during the entire study period. The slope (b) of the Gutenberg-Richter relationship is 0.83, whereas the slope for the entire period was 1.6. This indicates that small events make up less of catalogue B during this time. The energy that was released during the speed-up period was ~2.13x10^{-6} (m/s)^2, which is 57% of all energy released during the study period.

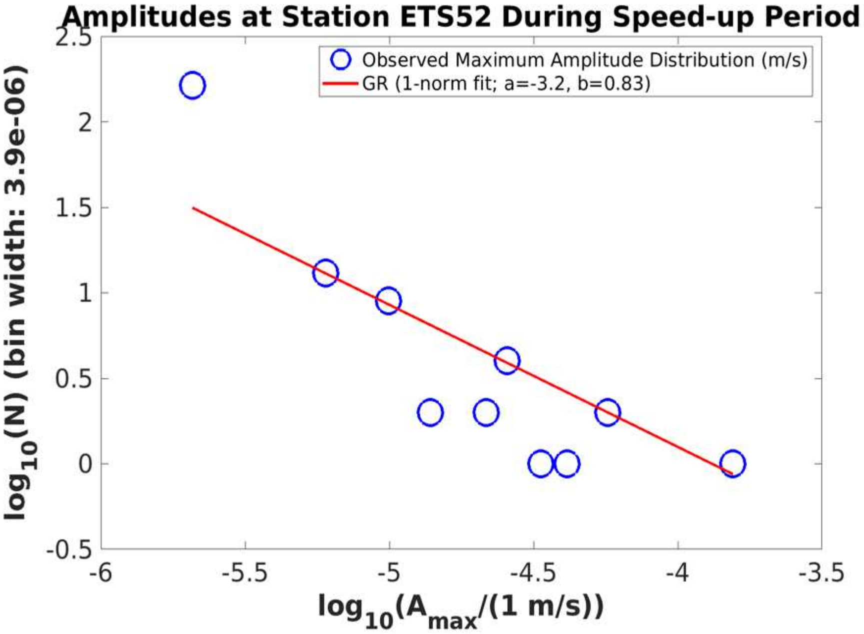


Figure 5.12: Distribution of maximum amplitudes (m/s) for events during the speed-up period recorded at station ETS52 on the vertical component.

The rate of events is generally Poissonian for the speed-up period (Figure 5.13). While there is still some variance, it has better correspondence with the exponential random distribution compared to catalogue B as a whole. There were 198 events during this time, which lasted 0.96 days, for an average rate of 204.4 events per day. This day accounts for 24.2% of all events in catalogue B. Throughout this time, the shortest amount of time between two events is 15.5 seconds, which is ~3 times longer than the shortest interevent time for the whole study period. The longest amount of time between two events is one hour and 43 minutes, which occurred at the end of day 355.

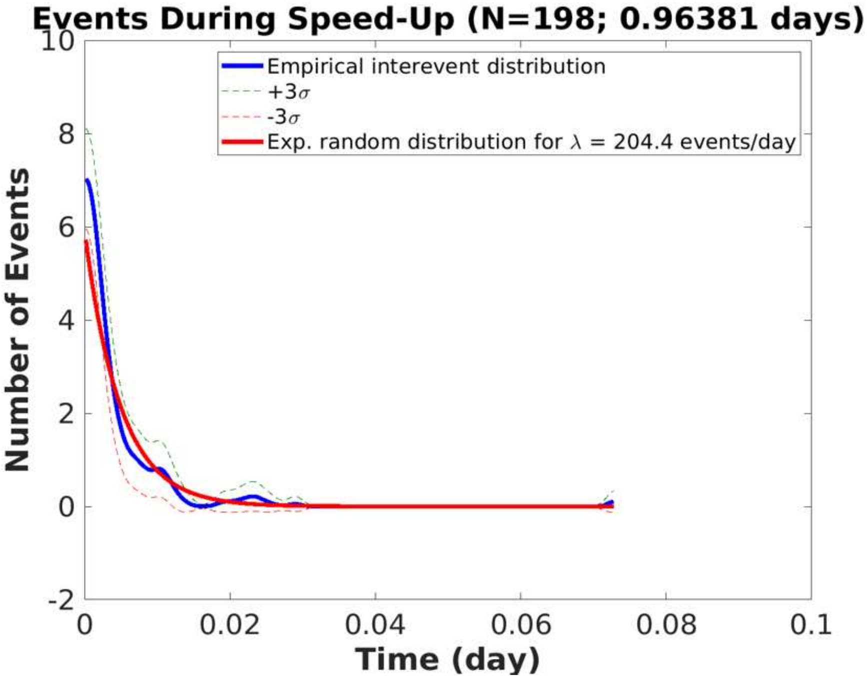


Figure 5.13: Empirical observed interevent distribution (blue) and exponential random distribution (red). For events during the speed-up period, the average Poissonian rate is 204.4 events per day. The dashed lines represent ±3 standard deviations.

Chapter 6

Discussion

6.1 Analysis of Results

6.1.1 Implications of Lava Lake Events

Our catalogue (B) indicates that the lava lake was consistently active between December 10–28 of 2008. This observation is in agreement with a hypothesized shallow magma chamber, which is usually associated with increased seismicity [10]. This continuous out-gassing maintains the stability of the magma system [22]. The continuous identification of VLP signals throughout the study period indicates that the deeper structure of Mount Erebus is stable during this time [28]. The waveforms of this catalogue, especially the VLP template events, can be compared to signals outside of the study period to determine when changes in the deep structure occur, which may indicate instability that can lead to increased seismic risk.

Because the VLP events were used as a template for matched filter detection, the events in the catalogue likely come from the same source, making them small burps at the surface of the lava lake. The largest detected event, which occurred on December 21st is larger than any of the VLP template events that were identified by their low frequency energy in broadband data. We assume that the largest amplitude events would have been detected during that process, however, despite the large short-period amplitude, there was no VLP signal. The record section of this event confirms that the event originated from the vicinity of the lava lake, so further scrutiny is required for the source of this unique event (Figure 6.1).

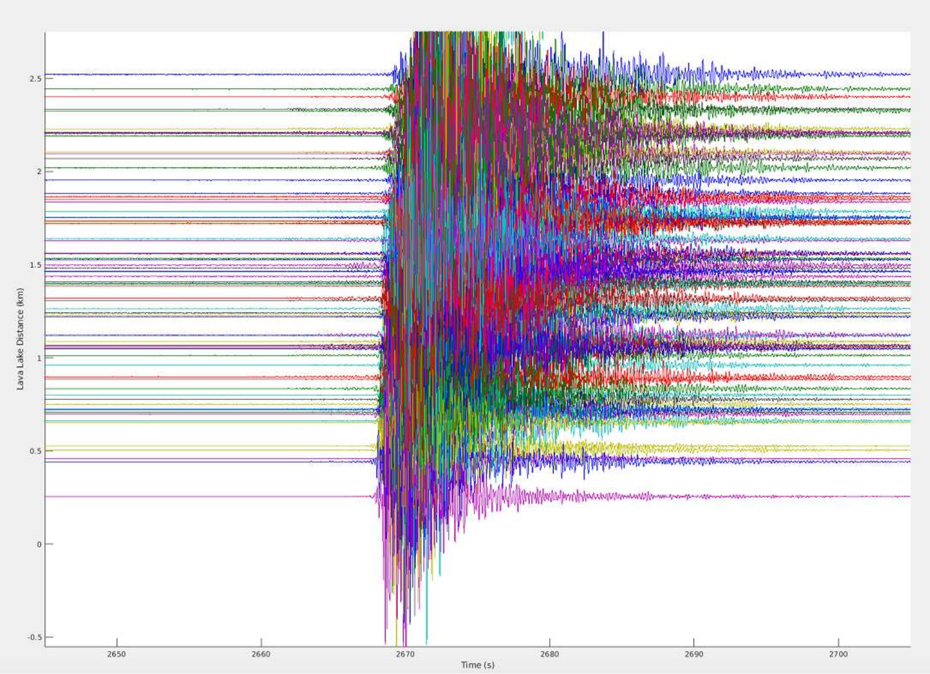


Figure 6.1: Record section of the largest event in the catalogue (event number 515 in catalogue B) for stations within 2.5 km of the lava lake. This event occurred on day 355 at 2567 seconds during hour 4. The move-out observed here confirms that the source location of the event is within (or very near) the lava lake.

6.1.2 Analysis of Event Size Distribution

Earlier studies also observed lava lake explosions with small amplitudes [12]. 364 events between October 1997 and July 1998 were discovered to possess a range of sizes from 10^{-7} to $10^{-5.7}$ m/s. The distribution of amplitudes did not follow a single trend line, but displayed two different slopes that reorients at an amplitude of $10^{-6.5}$. Events with amplitudes above this value have a slope of 1.7, which is comparable to the slope of 1.6 that I found for catalogue B. These values also align with the expected b value of 1-2 in volcanic regions [42]. While this slope solely defines the bigger events for [12], all of our events have amplitudes in this range (above $10^{-6.5}$ m/s and up to sizes of $10^{-3.7}$). This change in slope is hypothesized to be due to a lack of detection of smaller events, the result of two different sources, or a simple excess of medium-sized events, which could all be applicable scenarios to our catalogue. In our case, the lack of detection of smaller events is not the result of stations turning off, since the number of stations running per day did not affect the smallest detectable event (Figure 6.2).

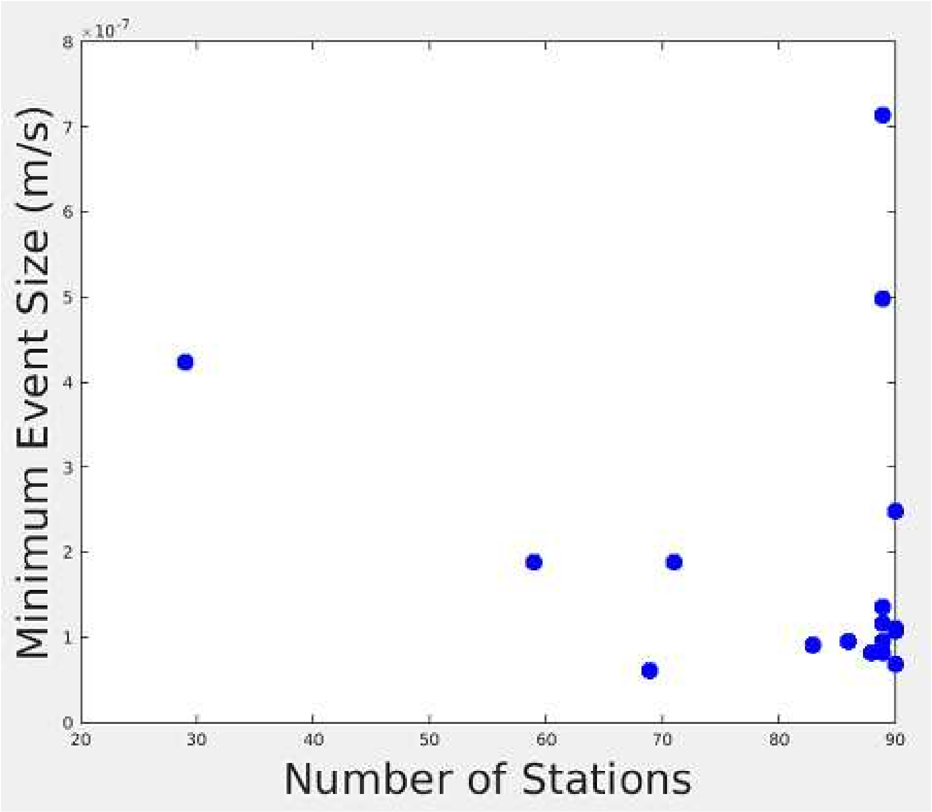


Figure 6.2: Number of stations running per day versus the smallest detected event on that day.

6.1.3 Analysis of Event Timing

Based on an average of 48.5 events per day, the average interevent time is 29.6 minutes. However, the observed interevent time varies greatly, which is a pattern that is frequently observed on short- and long-term scales in Mount Erebus' history [1]. If I consider all 1,741 of the events in catalogue A that were originally detected before throwing out the small, unclear events, the average over 17 days is 102.4 events per day. This equates to an average interevent time of 14 minutes. Like the interevent time distribution for catalogue B, there is an excess of interevent times with events occurring primarily less than 1 minute apart from each other. On average, there are 0-1 events every 10 minutes (Figure 6.3).

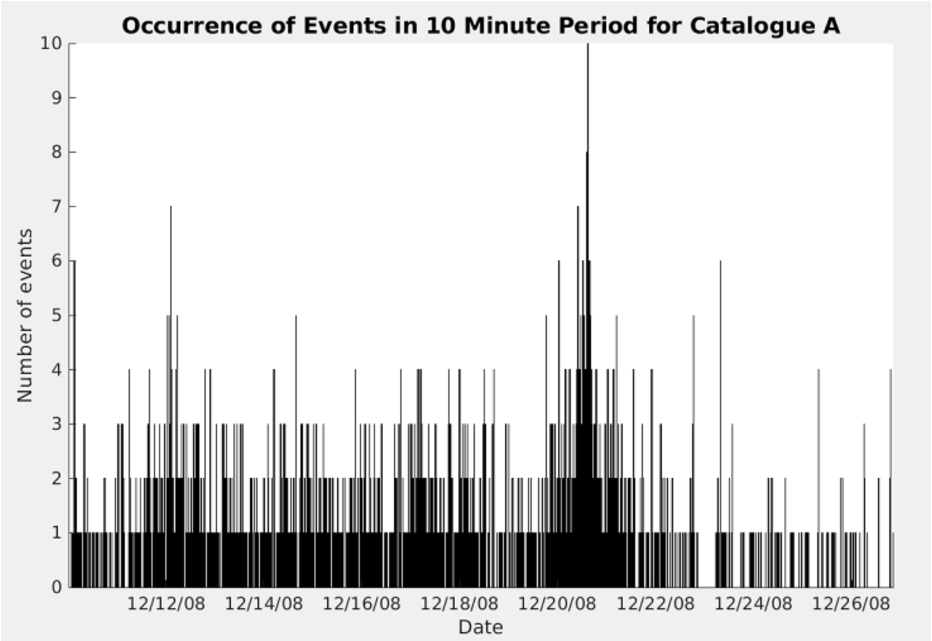


Figure 6.3: A count of the number of catalogue events that occur within each 10 minute window during the study period. Note that even during the December 20 event rate speed-up period, the maximum number of events per 10-minute window was only 10. This makes our catalogue marginally or inadequately dense for robustly assessing shorter duration eruptive behavior relative to the elevation periodicity range (e.g., 5 - 20 minutes) that has been sporadically noted during periods of direct lava lake observation [8], and spectral analysis of the counts per 10-minute, and other intervals, did not reveal any distinct periodicity peaks.

The Erebus lava lake has been sporadically observed to exhibit an elevation periodicity between 5 and 18 minutes, with the cycle predominantly occurring near a periodicity of 10 minutes [8]. Degassing periodicities at 40, 50, and 90 minutes have also been observed associated with varying rates of SO_2 emissions. [46] found that the average rate of sulfur dioxide released at Mount Erebus in December between 1992 and 2005 was 61 ± 27 Mg/day, but fluctuated greatly. The periodicity is generally thought to be related to dynamic processes that affect surface elevation, surface speed, and overall gas composition [47]. All of these properties are anticipated to be a result of changes in gas accumulation within the conduit and convection [46] [13].

Although the average interevent time for catalogue A coincides with these periodicities, the observed interevent times vary greatly and do not show a significant preference for 10 minutes (Figure 6.4). I was unable to search for the presence of the previously observed periodicities at Mount Erebus due to the lack of additional instrumentation on the summit during the time of our study such as thermal infrared cameras, ultraviolet cameras, spectrometers, video cameras, and terrestrial laser scanners. In previous work, it was noted that large seismic events were not the cause of the periodicity and did not change the cycles [47], so it is reasonable that the same pattern was not distinguished here.

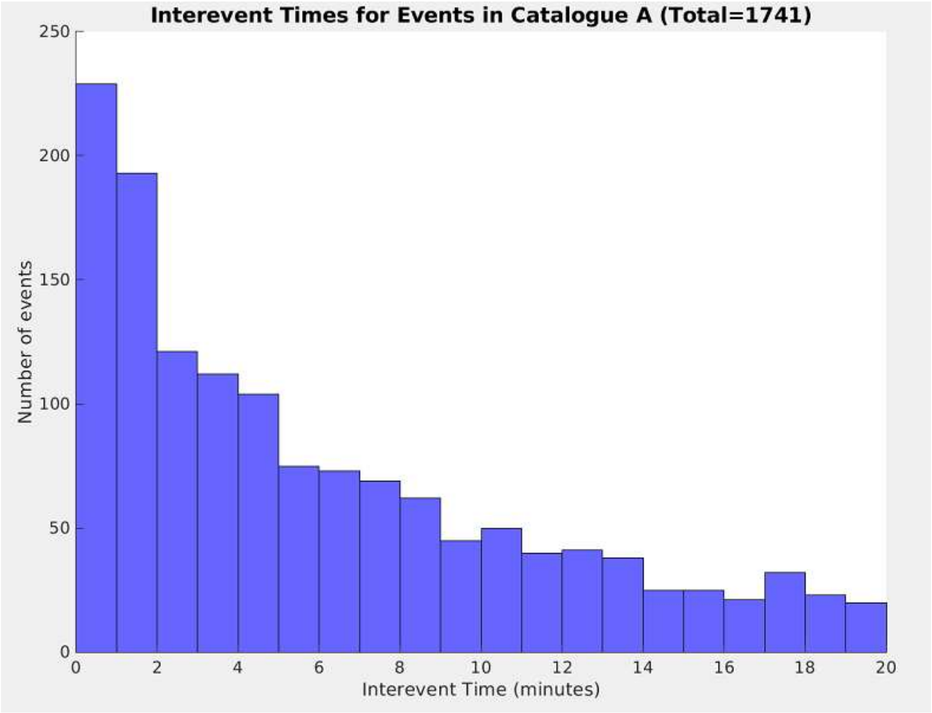


Figure 6.4: Distribution of interevent times for events in catalogue A. Note that only interevent times below 20 minutes are displayed in search for the existence of periodicities observed in [8].

6.1.4 Repeating Waveforms Hypothesis

Repeating waveforms that occur close together in time for events in clusters may be explained by weak spots in the lava lake. Waveforms with this degree of similarity are unexpected due to the heterogeneity within volcanoes, so it is likely that events in each cluster come from designated source locations and erupt at the same location at the surface of the lava lake. Gas bubbles taking identical paths through the conduit to the surface would allow for the recreation of waveforms. Once a gas bubble erupts at the surface, that area of the lava lake remains warm and weak for some time afterward as demonstrated by snapshots of an eruption captured on infrared video (Figure 6.5).

The gas bubbles that approach the surface of the lava lake shortly after an eruption may take advantage of this weak spot and erupt in the same location.

Perhaps the identification of various waveform types indicates that there are small scale changes to the magma structure or changes that are focused in the shallow structure [28]. [1] found that the upper structure of Mount Erebus is very complex, so this may be a contributing factor. Small-scale changes to the shallow magmatic structure may also promote repeating waveforms by affecting the path of gas bubbles.

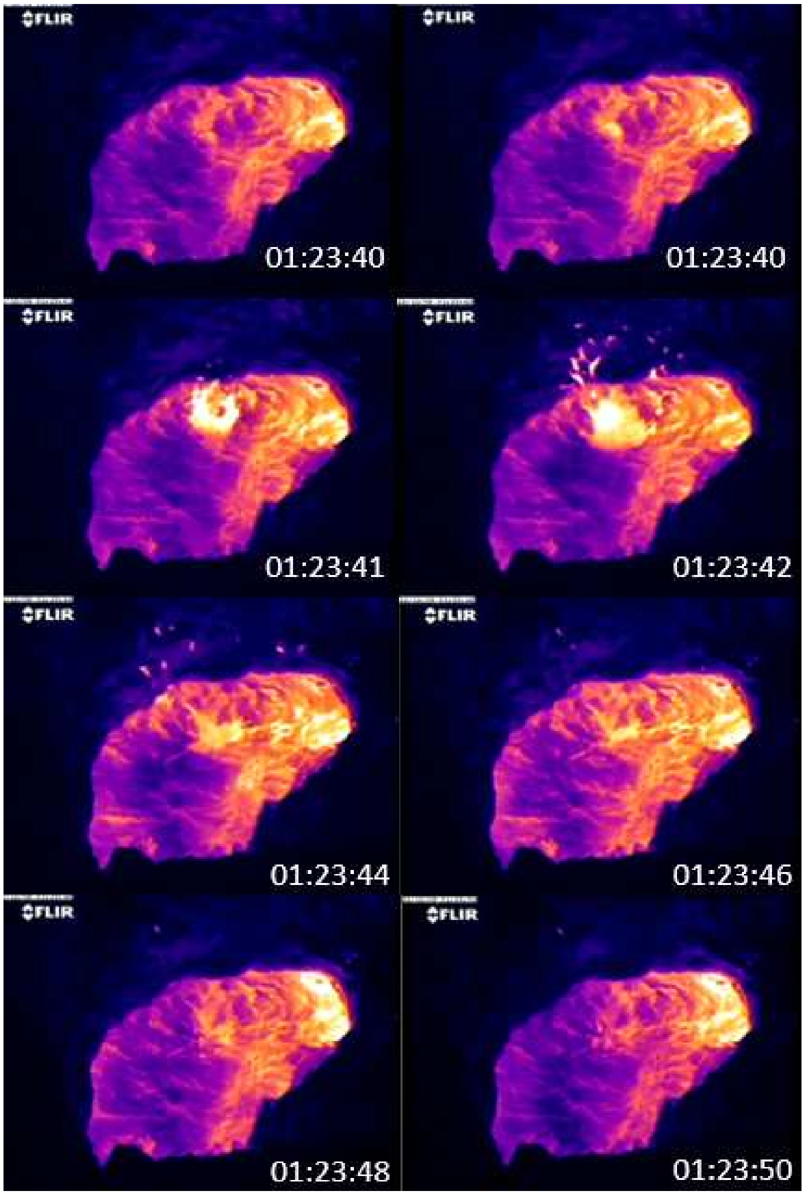


Figure 6.5: Screenshots of an infrared video of a small eruption in the lava lake on December 22, 2009. The video spans 10 seconds, with images displayed every 1-2 seconds. The eruption itself lasts about 6 seconds.

6.2 Future Work

To get a better idea of the source of the lava lake events, reverse-time imaging (e.g., back projection) would be a useful tool. This would be especially valuable for the clusters since they may represent a specific repetitive process within the larger lava lake system. It may also indicate different source areas that cause the excess of small event amplitudes. This method was used, for example, by [10] to locate the source of internal microseismicity at Mount St. Helens. This study utilized 904 geophones to record seismic activity over a course of 2 weeks, so similarly to our study, they were dealing with a large quantity of data. In this fully automated process, the data are continuously back-projected, making it a valuable technique for a plethora of events happening in a short amount of time [10]. The waveforms of events detected by an STA/LTA algorithm are aligned according to an assumed arrival time based on fast-marching ray tracing. This process is repeated for each location in the source image, which is plotted as the median of the traces. Once a catalogue of confirmed events is compiled, the hypocenters are determined by identifying the maximum amplitude within the trigger interval. However, the extreme seismic scattering present at Erebus [23] means that the back propagation process would only be straightforwardly implementable for the initial (and emergent) part of the seismograms for which deterministic ray paths and thus back projections could be estimated.

We might improve the number and diversity of detections in our catalogue (e.g., by bootstrapping the process using progressive detection as new events are discovered), as well as include smaller events, by using additional templates in the matched filtering process. This template could be a large event that is detected in the catalogue. It could also be the result of stacking either the events of the catalogue, or just the VLP events. Reprocessing the data could aid in further refining this step of the study. However, the significant detection overlap using the VLP detected template set suggests that even these limited templates likely already span the significant majority of high signal-to-noise events.

Along with this, the clustering process can be improved by increasing the number of stations that are used in the cross correlation. In this study, my clustering is run on 5 stations to provide meaningful results while making the run time tractable. This number of stations results in about 24 hours of run time with my MATLAB code on a remote multicore linux computer server (GEO3) in the CSU Geoscience Department utilizing parallel computing. Because each additional station requires correlation with 2,457 (819 events on 3 components) more seismograms with the existing seismograms, the run time gets (at least) quadratically longer with the number of events and at least linearly longer as the number of stations is increased. This process could potentially utilize 22 stations that were operational over the span of catalogue B on all 3 components, or a greater number if varying station constellations were utilized. Processing the data with a low pass filter at lower frequencies might also improve clustering process, although the instruments have limited low frequency energy on the ETS seismographs (due to their 4.5 Hz corner frequency geophones). More efficiencies may also be achievable using multichannel matched filtering sectioning and other optimized algorithms recently developed by the seismological community [48].

Chapter 7

Conclusions

I utilized multichannel matched filtering from a dense network of 99 continuously recording (asynchronously) three-component short period (4.5 Hz geophones) stations deployed in the near-summit region of Erebus volcano between December 10 and December 28, 2008 in the TOMO Erebus project [9] to resolve small scale impulsive seismic events from the lava lake. This process, based on waveform similarity with repeating very-long-period signals, yielded an initial catalogue (A) of 1,741 unique events that was visually inspected to produce a high-signal-to-noise catalogue (B) containing 819 clear events for waveform analysis. Events rarely cluster closely in their waveforms, with only 6% of events being so characterized using a Ward clustering algorithm. This small population of repeating events tends to encompass event groups that are also close in time. This discovery bears further scrutiny, but may indicate repeated small gas explosions at closely similar locations either defined by the boundaries of lava lake convective cells, by the edges of the lava lake, or by weak spots in the lava lake following an eruption. I examined the occurrence and size statistics of the high signal-to-noise catalogue (B) and noted event size statistics are consistent with a Gutenberg-Richter statistical characterization with an a-value of -6.6, an overall b-value of 1.6, and an excess of small events. I observed time-varying quasi-Poissonian occurrence statistics with Poissonian rates per day varying from 8.19 to 204.4 events and an excess of short interevent times. A notable speed-up in activity of approximately 23 hours occurred on December 20th, accounting for 24.2% of the detected events in catalogue B and 57% percent of the pseudo energy (calculated from the summed squared vertical seismic velocity observed at reference station ETS52). Determining the eruption rate for this time period improves our understanding of the characteristic behavior of Mount Erebus. Waveform analysis of the catalogue also aids in establishing the stability of the magmatic system, and the associated seismic risk. This study illustrates the value of operating large numbers of simultaneous seismic stations on active volcanoes in detecting and discriminating high frequency seismicity, and suggests that even short-period instruments, which can

can be designed to be small, cheaper, and at some level expendable, are worthwhile components for future volcano monitoring and research. The development and distribution of this workflow will be useful in application for other glacial volcanoes.

Bibliography

[1] Kenneth WW Sims, Richard C Aster, Glenn Gaetani, Janne Blichert-Toft, Erin H Phillips, Paul J Wallace, Glen S Mattioli, Dan Rasmussen, and Eric S Boyd. Mount Erebus. *Geological Society, London, Memoirs*, 55(1):695–739, 2021.

[2] Kayla Iacovino, Clive Oppenheimer, Bruno Scaillet, and Philip Kyle. Storage and evolution of mafic and intermediate alkaline magmas beneath Ross Island, Antarctica. *Journal of Petrology*, 57(1):93–118, 2016.

[3] Richard P Esser, Philip R Kyle, and William C McIntosh. 40Ar/39Ar dating of the eruptive history of Mount Erebus, Antarctica: volcano evolution. *Bulletin of Volcanology*, 66(8):671–686, 2004.

[4] Christopher J Harpel, Philip R Kyle, Richard P Esser, William C McIntosh, and David A Caldwell. 40Ar/39Ar dating of the eruptive history of Mount Erebus, Antarctica: summit flows, tephra, and caldera collapse. *Bulletin of Volcanology*, 66(8):687–702, 2004.

[5] HA Knox, JA Chaput, RC Aster, and PR Kyle. Multiyear shallow conduit changes observed with lava lake eruption seismograms at Erebus Volcano, Antarctica. *Journal of Geophysical Research: Solid Earth*, 123(4):3178–3196, 2018.

[6] Leo Eisner, David Abbott, William B Barker, Michael P Thornton, and James Lakings. Noise suppression for detection and location of microseismic events using a matched filter. In *2008 SEG Annual Meeting*. OnePetro, 2008.

[7] R Aster, S Mah, P Kyle, W McIntosh, N Dunbar, J Johnson, M Ruiz, and S McNamara. Very long period oscillations of Mount Erebus Volcano. *Journal of Geophysical Research: Solid Earth*, 108(B11), 2003.

[8] Nial Peters, Clive Oppenheimer, Philip Kyle, and Nick Kingsbury. Decadal persistence of cycles in lava lake motion at Erebus volcano, Antarctica. *Earth and Planetary Science Letters*, 395:1–12, 2014.

[9] Daria Zandomeneghi, Philip Kyle, Pnina Miller, Iris Passcal, Catherine Snelson, and Richard Aster. Seismic tomography of Erebus Volcano, Antarctica. *Eos, Transactions American Geophysical Union*, 91(6):53–55, 2010.

[10] Steven M Hansen and Brandon Schmandt. Automated detection and location of microseismicity at Mount St. Helens with a large-N geophone array. *Geophysical Research Letters*, 42(18):7390–7397, 2015.

[11] Richard Aster, W MacIntosh, Philip Kyle, Richard Esser, B Bartel, Nelia Dunbar, J Johnson, R Karstens, C Kurnik, M McGowan, et al. Real-time data received from Mount Erebus volcano, Antarctica. *Eos, Transactions American Geophysical Union*, 85(10):97–101, 2004.

[12] CA Rowe, RC Aster, PR Kyle, RR Dibble, and JW Schlue. Seismic and acoustic observations at Mount Erebus Volcano, Ross island, Antarctica, 1994–1998. *Journal of Volcanology and Geothermal Research*, 101(1-2):105–128, 2000.

[13] R Aster, D Zandomeneghi, S Mah, S McNamara, DB Henderson, H Knox, and K Jones. Moment tensor inversion of very long period seismic signals from Strombolian eruptions of Erebus Volcano. *Journal of Volcanology and Geothermal Research*, 177(3):635–647, 2008.

[14] Seelye Martin, Robert Drucker, Richard Aster, Fred Davey, Emile Okal, Ted Scambos, and Douglas MacAyeal. Kinematic and seismic analysis of giant tabular iceberg breakup at Cape Adare, Antarctica. *Journal of Geophysical Research: Solid Earth*, 115(B6), 2010.

[15] CS Siddoway. Tectonics of the West Antarctic Rift System: new light on the history and dynamics of distributed intracontinental extension. *Antarctica: A keystone in a changing world*, pages 91–114, 2008.

[16] PR Kyle. A. McMurdo volcanic group western Ross embayment. *Volcanoes of the Antarctic plate and southern Oceans*, 48:18–145, 1990.

[17] Peter J Kelly, Nelia W Dunbar, Philip R Kyle, and William C McIntosh. Refinement of the late Quaternary geologic history of Erebus volcano, Antarctica using 40Ar/39Ar and 36Cl age determinations. *Journal of Volcanology and Geothermal Research*, 177(3):569–577, 2008.

[18] Chelsea J Vickers, Craig W Herbold, S Craig Cary, and Ian R Mcdonald. Insights into the metabolism of the high temperature microbial community of Tramway Ridge, Mount Erebus, Antarctica. *Antarctic Science*, 28(4):241–249, 2016.

[19] LJ Wardell, PR Kyle, and AR Campbell. Carbon dioxide emissions from fumarolic ice towers, Mount Erebus volcano, Antarctica. *Geological Society, London, Special Publications*, 213(1):231–246, 2003.

[20] Kurt S Panter and Brian Winter. Geology of the Side Crater of the Erebus volcano, Antarctica. *Journal of Volcanology and Geothermal Research*, 177(3):578–588, 2008.

[21] Ph R Kyle, JA Moore, and MF Thirlwall. Petrologic evolution of anorthoclase phonolite lavas at Mount Erebus, Ross Island, Antarctica. *Journal of Petrology*, 33(4):849–875, 1992.

[22] Clive Oppenheimer, Alexandra S Lomakina, Philip R Kyle, Nick G Kingsbury, and Marie Boichu. Pulsatory magma supply to a phonolite lava lake. *Earth and Planetary Science Letters*, 284(3-4):392–398, 2009.

[23] J Chaput, M Campillo, RC Aster, P Roux, PR Kyle, H Knox, and P Czoski. Multiple scattering from icequakes at Erebus volcano, Antarctica: Implications for imaging at glaciated volcanoes. *Journal of Geophysical Research: Solid Earth*, 120(2):1129–1141, 2015.

[24] D Zandomeneghi, R Aster, P Kyle, A Barclay, J Chaput, and H Knox. Internal structure of Erebus volcano, Antarctica imaged by high-resolution active-source seismic tomography and coda interferometry. *Journal of Geophysical Research: Solid Earth*, 118(3):1067–1078, 2013.

[25] Kevin Mickus and P Kyle. A detailed gravity survey of the Erebus Volcano, Antarctica. In *Geological Society of America Annual Meeting*, volume 19, page 22, 2014.

[26] Thibaud Blondel, Julien Chaput, Arnaud Derode, Michel Campillo, and Alexandre Aubry. Matrix approach of seismic imaging: Application to the Erebus volcano, Antarctica. *Journal of Geophysical Research: Solid Earth*, 123(12):10–936, 2018.

[27] Clive Oppenheimer, Roberto Moretti, Philip R Kyle, Al Eschenbacher, Jacob B Lowenstern, Richard L Hervig, and Nelia W Dunbar. Mantle to surface degassing of alkalic magmas at Erebus volcano, Antarctica. *Earth and Planetary Science Letters*, 306(3-4):261–271, 2011.

[28] Kyle A Brill, Gregory P Waite, Simon A Carn, Amilcar Elias Roca Palma, and Gustavo Chigna. Long-term stability of conduit dynamics at Fuego volcano, Guatemala, 2008–2015. *Bulletin of Volcanology*, 84(4):1–15, 2022.

[29] Nigel Allister Anstey. Correlation techniques–a review. *Geophysical Prospecting*, 12(4):355–382, 1964.

[30] SS Alexander and DB Rabenstine. Detection of surface waves from small events at teleseismic distances. Technical report, TELEDYNE INDUSTRIES INC ALEXANDRIA VA EARTH SCIENCES DIV, 1967.

[31] George Turin. An introduction to matched filters. *IRE transactions on Information theory*, 6(3):311–329, 1960.

[32] Ben S Melton and Leslie F Bailey. Multiple signal correlators. *Geophysics*, 22(3):565–588, 1957.

[33] Peter M Shearer. Global seismic event detection using a matched filter on long-period seismograms. *Journal of Geophysical Research: Solid Earth*, 99(B7):13713–13725, 1994.

[34] Steven J Gibbons and Frode Ringdal. The detection of low magnitude seismic events using array-based waveform correlation. *Geophysical Journal International*, 165(1):149–166, 2006.

[35] Song Zhiming, Yan Xiaoli, Qu Zhongquan, and Li Hong-Bo. Automatic detection and extraction algorithm of coronal loops based on match filter and oriented directivity. *Monthly Notices of the Royal Astronomical Society*, 490(4):5567–5584, 2019.

[36] AD Hoover, Valentina Kouznetsova, and Michael Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical imaging*, 19(3):203–210, 2000.

[37] Jeremy Kepner, Xiaohui Fan, Neta Bahcall, James Gunn, Robert Lupton, and Guohong Xu. An automated cluster finder: the adaptive matched filter. *The Astrophysical Journal*, 517(1):78, 1999.

[38] David P Schaff and Paul G Richards. Repeating seismic events in China. *Science*, 303(5661):1176–1178, 2004.

[39] Richard C Aster, Brian Borchers, and Clifford H Thurber. *Parameter estimation and inverse problems*. Elsevier, 2018.

[40] C Godano, E Lippiello, and L De Arcangelis. Variability of the b value in the Gutenberg–Richter distribution. *Geophysical Journal International*, 199(3):1765–1771, 2014.

[41] Rune Olsson. An estimation of the maximum b-value in the Gutenberg-Richter relation. *Journal of Geodynamics*, 27(4-5):547–552, 1999.

[42] John B Rundle. Derivation of the complete Gutenberg-Richter magnitude-frequency relation using the principle of scale invariance. *Journal of Geophysical Research: Solid Earth*, 94(B9):12337–12342, 1989.

[43] SK Katti and A Vijaya Rao. Handbook of the Poisson distribution, 1968.

[44] Thalia Anagnos and Anne S Kiremidjian. A review of earthquake occurrence models for seismic hazard analysis. *Probabilistic Engineering Mechanics*, 3(1):3–11, 1988.

[45] Roy D Yates and David J Goodman. *Probability and stochastic processes: a friendly introduction for electrical and computer engineers*. John Wiley & Sons, 2014.

[46] Dawn Sweeney, Philip R Kyle, and Clive Oppenheimer. Sulfur dioxide emissions and degassing behavior of Erebus volcano, Antarctica. *Journal of Volcanology and Geothermal Research*, 177(3):725–733, 2008.

[47] Nial Peters, Clive Oppenheimer, Drea Rae Killingsworth, Jed Frechette, and Philip Kyle. Correlation of cycles in Lava Lake motion and degassing at Erebus Volcano, Antarctica. *Geochemistry, Geophysics, Geosystems*, 15(8):3244–3257, 2014.

[48] Nader Shakibay Senobari, Gareth J Funning, Eamonn Keogh, Yan Zhu, Chin-Chia Michael Yeh, Zachary Zimmerman, and Abdullah Mueen. Super-efficient cross-correlation (SEC-C): A fast matched filtering code suitable for desktop computers. *Seismological Research Letters*, 90(1):322–334, 2019.

Appendix A

Glossary

Alkaline: Containing alkali metals such as potassium and sodium.

Basanitic: Describing a volcanic rock with aphanitic (very fine minerals) to porphyritic (distinctly visible minerals) texture.

Broadband: Spans a wide range of frequencies and captures longer periods.

Butterworth filter: A signal processing filter that flattens the frequency response and achieves a more linear phase response. A second-order filter decreases at -12 dB per octave.

Coda: The incoherent signal following seismic events caused by scattering.

Corner frequency: The limit above which a seismic sensor does not accurately capture the frequency response from an event.

Finite impulse response filter: A filter used to produce an impulse response that approaches zero in a defined amount of time.

Infrasonic: Sound waves which cannot be heard by humans due to their low frequency.

Phonolite: An intermediate (between felsic and mafic) extrusive rock containing low amounts of silica. Forms when highly aluminous rocks undergo low degrees of partial melting.

Phreatic: Water or steam erupting after being warmed underground by magma.

Plinian: Volcanic eruptions characterized by large ash columns that spread out horizontally in the atmosphere.

Poissonian: A distribution that describes the probability of a given number of events occurring during an allotted time interval assuming the events occur at an average rate. Poissonian processes indicate that events occur independently of each other, and the timing of the last event does not affect the timing of the next event.

Shield Volcano: Formed by the eruption of low viscosity lava which allows the lava to travel farther and in thinner flows.

Stratovolcano: A volcano that is formed by multiple layers (strata) of erupted material.

Strombolian Eruption: Moderately explosive eruptions of basaltic magma. The explosive bursts are capable of ejecting fragments of rock.

Volcanic Explosivity Index: A logarithmic scale that uses the magnitude and intensity of an eruption to convey the size of the eruption. The scale ranges from 0 to 8.

Appendix B

List of Events

B.1 VLP Template EVents

Table B.1: List of VLP lava lake events discriminated by very long period frequency content on broadband stations, and used as templates for matched filter detection on the ETS network. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The maximum amplitude is at all stations within 60 seconds of the start time of the event, recorded in (m/s) (x10^6).

Template	Day	Hour	Start time	End time	Max. Amplitude
1	345	12	1306	1326	2.58
2	346	8	2729	2749	12.8
3	346	18	3024	3044	4.1
4	346	22	421	441	9.1
5	349	10	1696	1716	65
6	350	0	2008	2028	6.2
7	350	5	2053	2073	31.3
8	350	11	580	600	2.78
9	350	20	3082	3102	49.3
10	352	4	2483	2503	21.7
11	353	11	3008	3028	41
12	353	19	3288	3308	4.43
13	355	15	2423	2443	27.9
14	356	10	3072	3092	4.1
15	356	22	74	94	21.2
16	359	7	1056	1076	28.6
17	360	7	3334	3354	6.38

B.2 Lava Lake Event Catalogues

An Excel spreadsheet with catalogue A and catalogue B are included in the supplementary material. Catalogue A, which includes all 1,741 events detected by matched filtering from December 10-28, is on sheet one, and catalogue B, which is the high signal-to-noise events from catalogue A, is on sheet two. The first column is the day of the year in 2008. Day 345 corresponds to December 10th. Column two is the UTC hour of the day. The start times in column three are in seconds. Column four is the maximum correlation value between the event and template that detected the event. Since multiple events could have detected each event, the template that had the highest correlation is used for column four, and that template number is recorded in column five, which corresponds to the numbering in table B.1. Column six (Amplitude) is the maximum amplitude recorded on station ETS52 within 60 seconds of the start time of the event, recorded in (m/s) (x10^6). The seventh column displays the number of stations that were used in the matched filtering process. The eighth column displays the number of templates that detected each event. In sheet one, if this event was clear enough to be included in catalogue B, the last column includes the corresponding event number in catalogue B. Otherwise, the last column contains a zero. For sheet two, the last column contains the corresponding event number in catalogue A.

Appendix C

Clustering Results

This section includes a list of events for each cluster and their waveforms. These clusters are identified by cross correlation with interevent correlation values greater than 0.8. Each figure includes a 15 second segment of data that captures the event. Event number is on the left. The amplitudes are the ELZ component recorded on station ETS52 in m/s.

C.1 Cluster 1

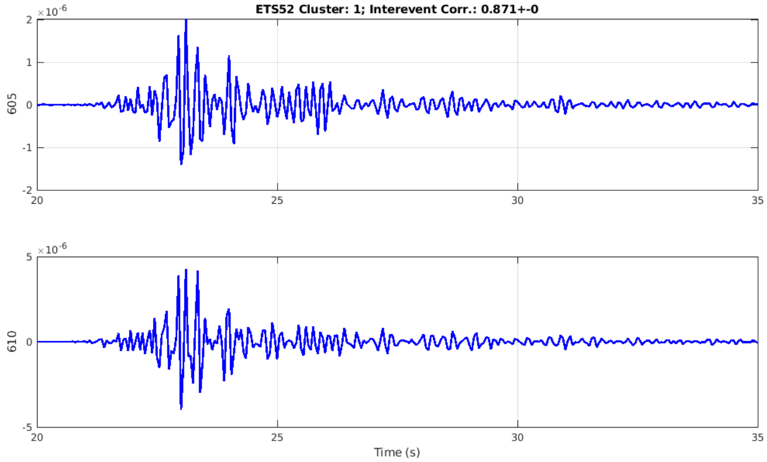


Figure C.1: Waveforms of events in cluster 1.

Table C.1: List of events that are identified as cluster 1. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
605	355	13	2991	3051	91
610	355	14	326.1	386.1	91

C.2 Cluster 2

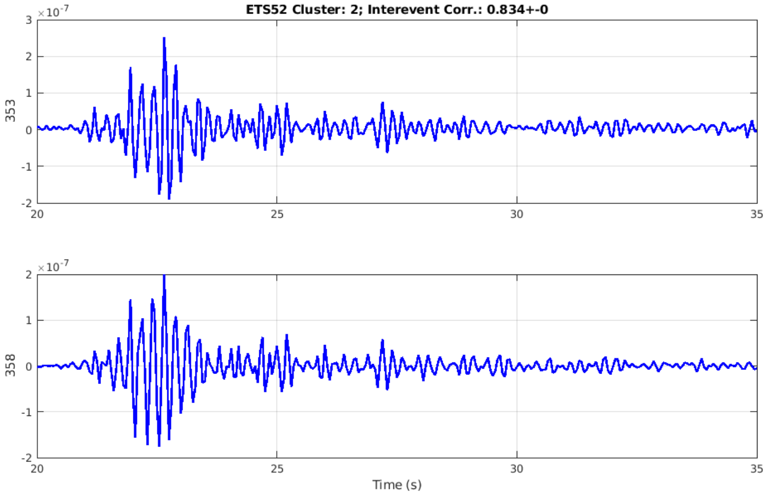


Figure C.2: Waveforms of events in cluster 2.

Table C.2: List of events that are identified as cluster 2. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
353	352	9	2488	2548	91
358	352	11	1279	1339	91

C.3 Cluster 3

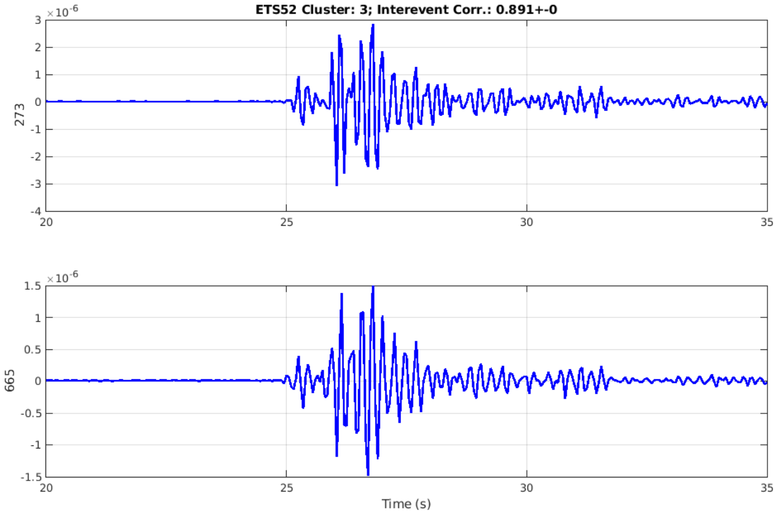


Figure C.3: Waveforms of events in cluster 3.

Table C.3: List of events that are identified as cluster 3. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
273	351	4	105.5	165.5	90
665	355	17	3364	3424	91

C.4 Cluster 4

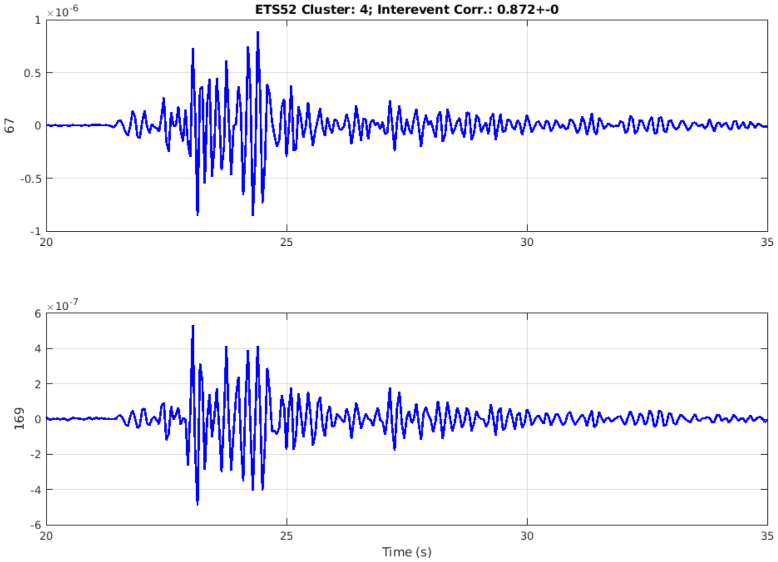


Figure C.4: Waveforms of events in cluster 4.

Table C.4: List of events that are identified as cluster 4. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
67	347	3	2211	2271	87
169	349	4	3052	3112	90

C.5 Cluster 5

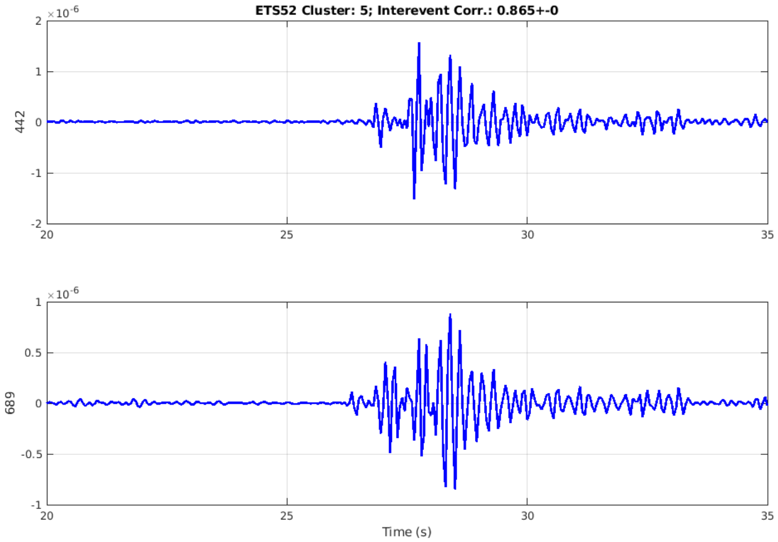


Figure C.5: Waveforms of events in cluster 5.

Table C.5: List of events that are identified as cluster 5. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
442	354	0	376.7	436.7	90
689	355	21	1400	1460	91

C.6 Cluster 6

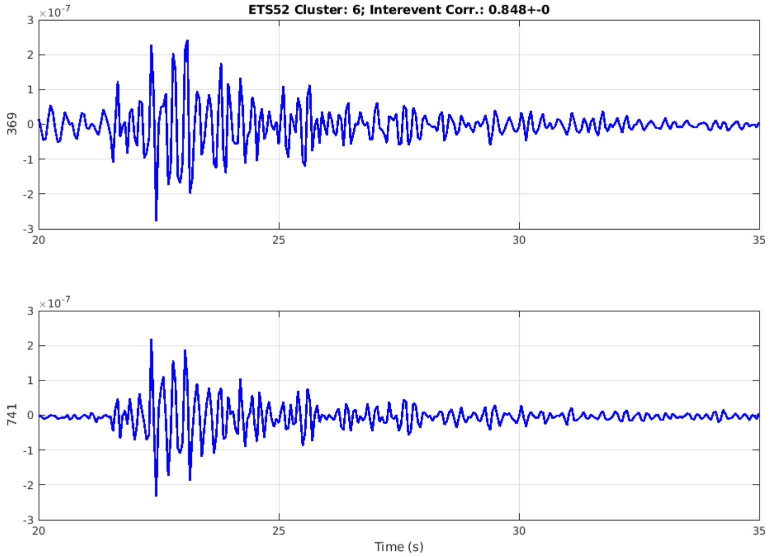


Figure C.6: Waveforms of events in cluster 6.

Table C.6: List of events that are identified as cluster 6. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
369	352	17	1151	1211	91
741	356	17	380.3	440.3	91

C.7 Cluster 7

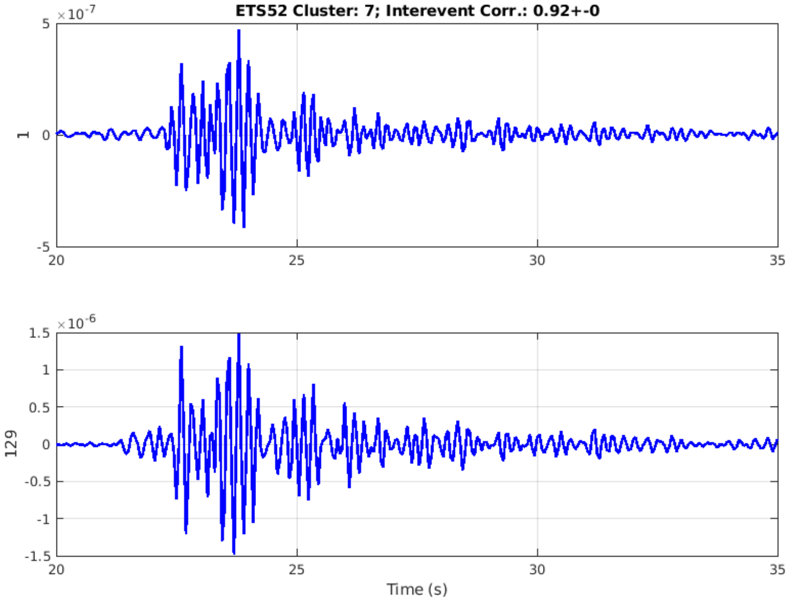


Figure C.7: Waveforms of events in cluster 7.

Table C.7: List of events that are identified as cluster 7. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
1	345	0	861.1	921.1	61
129	348	8	803.7	863.7	92

C.8 Cluster 8

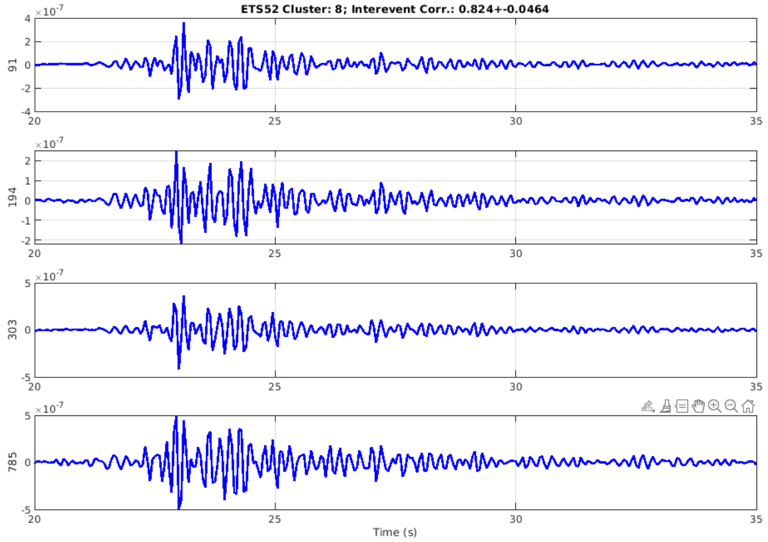


Figure C.8: Waveforms of events in cluster 8.

Table C.8: List of events that are identified as cluster 8. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
91	347	18	2204	2264	92
194	349	15	2028	2088	90
303	351	15	2342	2402	91
785	358	9	3016	3076	91

C.9 Cluster 9

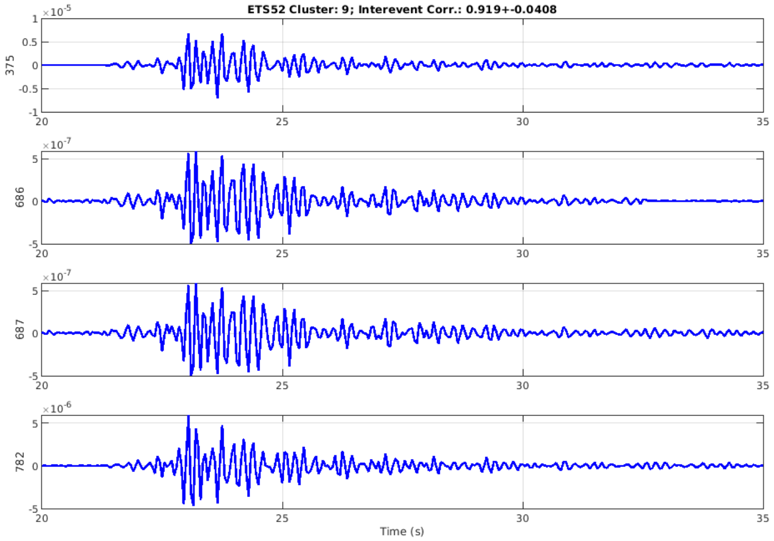


Figure C.9: Waveforms of events in cluster 9.

Table C.9: List of events that are identified as cluster 9. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
375	352	18	2014	2074	91
686	355	20	1993	2053	91
687	355	20	2021	2081	91
782	358	6	2959	3019	91

C.10 Cluster 10

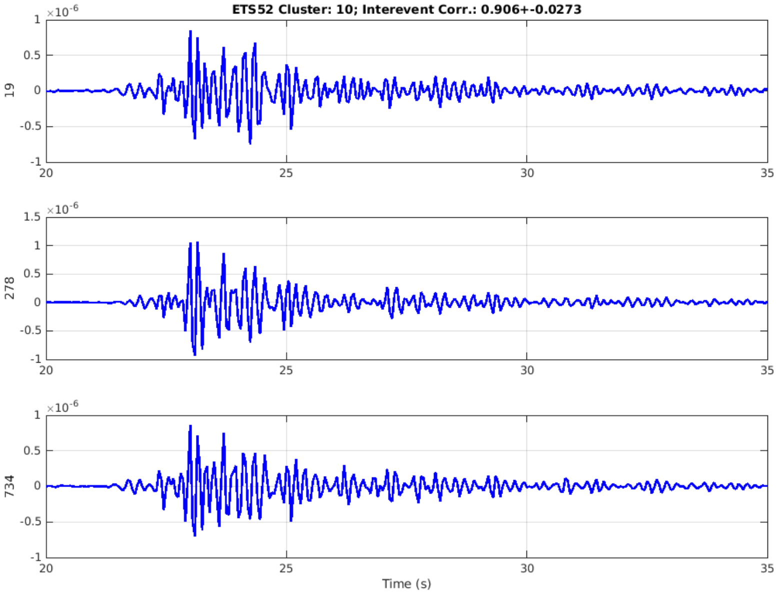


Figure C.10: Waveforms of events in cluster 10.

Table C.10: List of events that are identified as cluster 10. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
19	345	15	1630	1690	68
278	351	5	2874	2934	90
734	356	12	1227	1287	91

C.11 Cluster 11

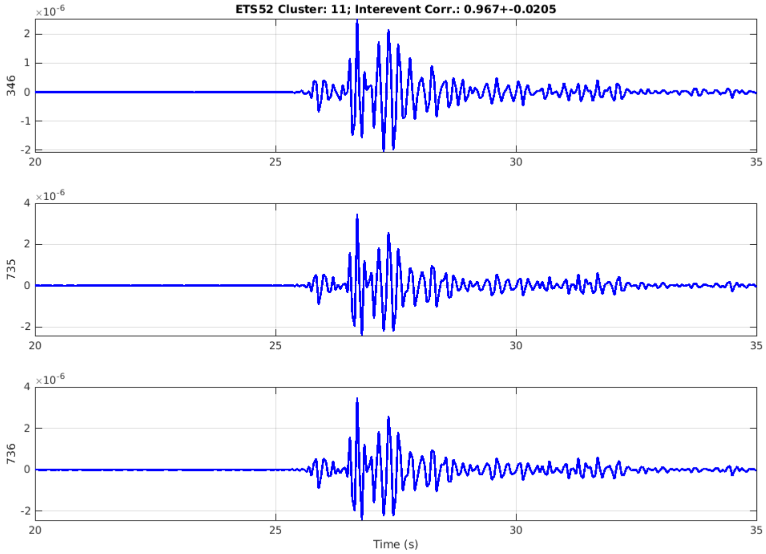


Figure C.11: Waveforms of events in cluster 11.

Table C.11: List of events that are identified as cluster 11. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
346	352	6	652.8	712.8	91
735	356	13	424.2	484.2	91
736	356	13	450.3	510.3	91

C.12 Cluster 12

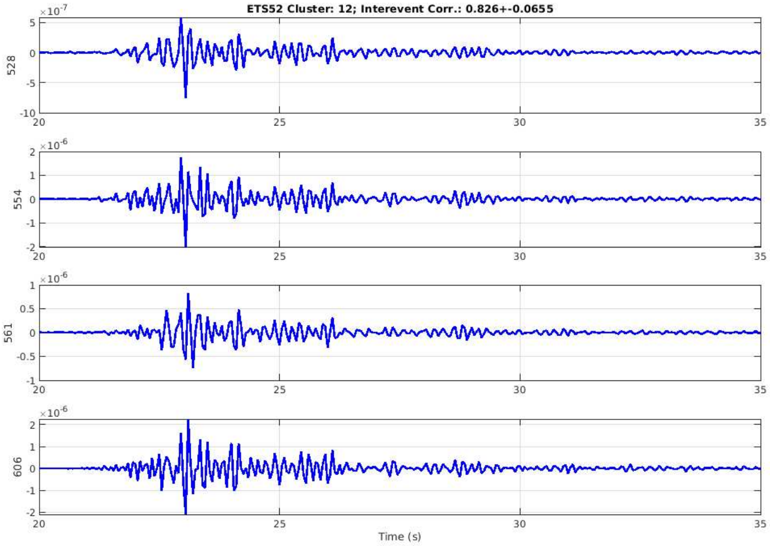


Figure C.12: Waveforms of events in cluster 12.

Table C.12: List of events that are identified as cluster 12. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
528	355	7	1414	1474	91
554	355	9	3395	3455	91
561	355	10	1268	1328	91
606	355	13	3096	3156	91

C.13 Cluster 13

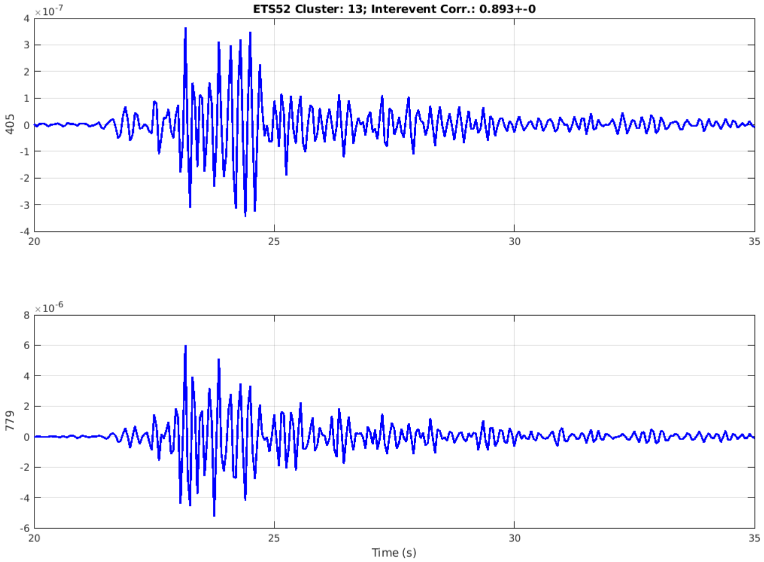


Figure C.13: Waveforms of events in cluster 13.

Table C.13: List of events that are identified as cluster 13. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
405	353	3	1104	1164	91
779	358	6	2134	2194	91

C.14 Cluster 14

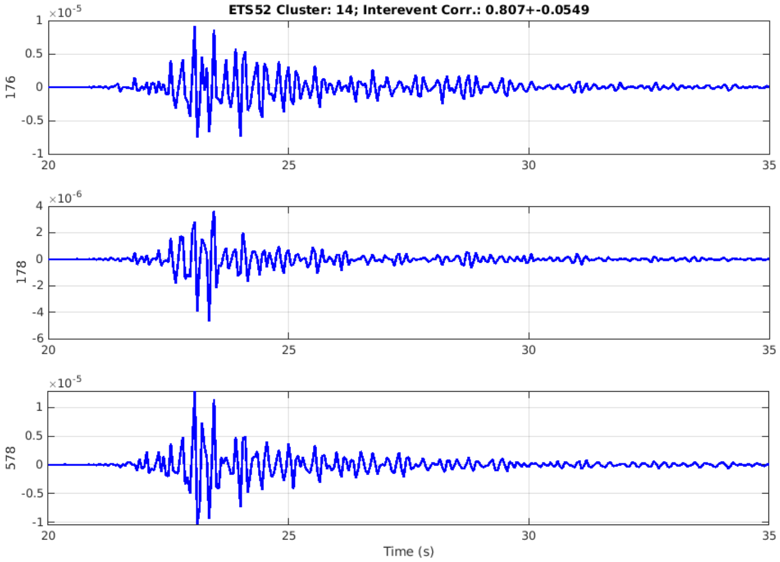


Figure C.14: Waveforms of events in cluster 14.

Table C.14: List of events that are identified as cluster 14. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
176	349	6	2266	2326	91
178	349	8	436.4	496.4	91
578	355	11	2499	2559	91

C.15 Cluster 15

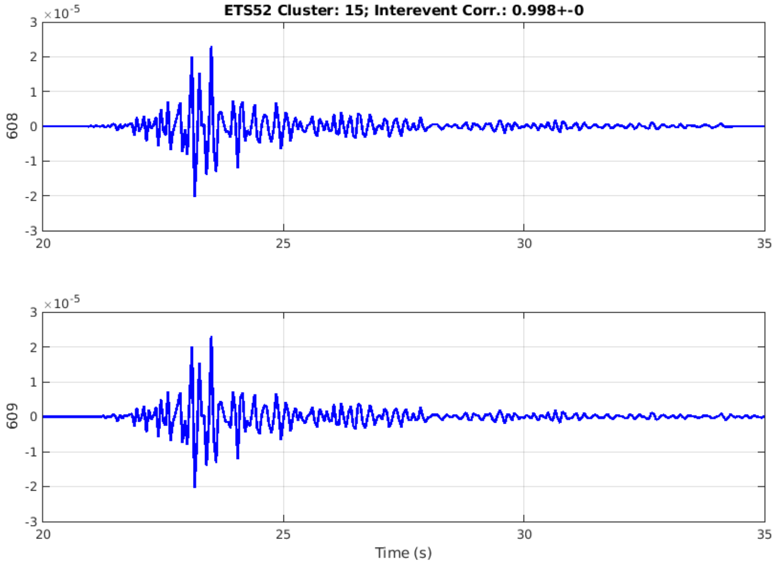


Figure C.15: Waveforms of events in cluster 15.

Table C.15: List of events that are identified as cluster 15. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
608	355	13	3501	3561	91
609	355	13	3528	3588	91

C.16 Cluster 16

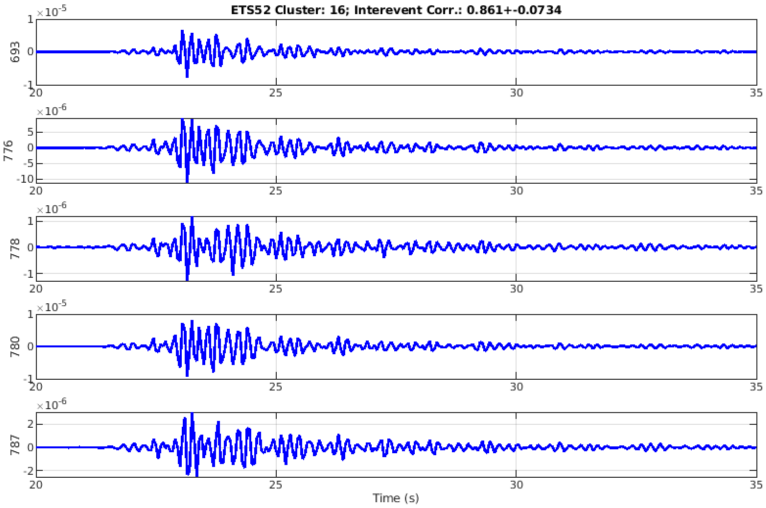


Figure C.16: Waveforms of events in cluster 16.

Table C.16: List of events that are identified as cluster 16. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
693	356	0	660	720	89
776	358	6	167.3	227.3	91
778	358	6	1969	2029	91
780	358	6	2515	2575	91
787	358	11	841.9	901.9	91

C.17 Cluster 17

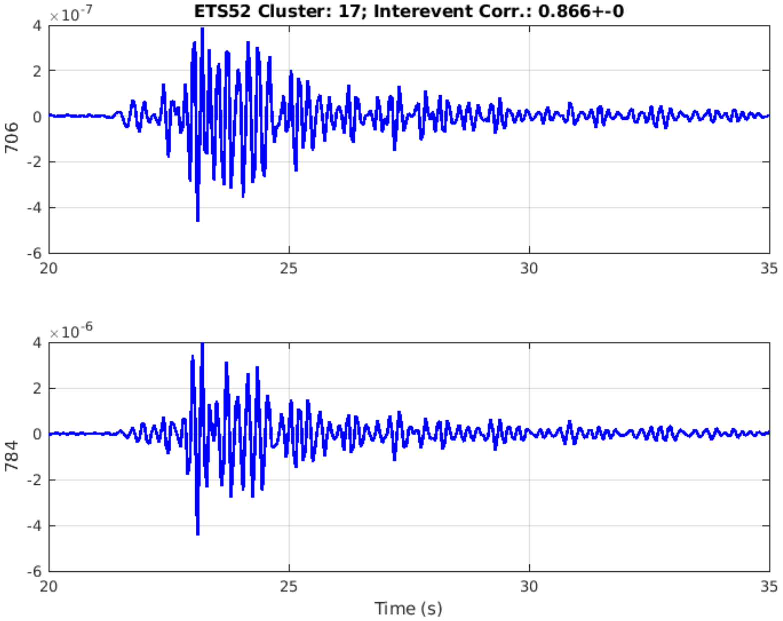


Figure C.17: Waveforms of events in cluster 17.

Table C.17: List of events that are identified as cluster 17. The first column in the event number from catalogue B. The second column is the day of the year in 2008. Day 345 corresponds to December 10th. Column three is the UTC hour of the day. The start and end times are in seconds. The last column is the number of stations used in matched filter detection.

Event	Day	Hour	Start time	End time	Stations
706	356	3	240.7	300.7	91
784	358	8	2571	2631	91

Appendix D

Description of Code

Table D.1: List of filenames for code included in supplementary material and a description of each. The files are listed in order of use following the workflow of this study. The last file can be used as needed to visualize and inspect results. Files with .m extension can be executed in MATLAB.

Number	Filename	Description
1	download_data	Downloads Tomo Erebus ETS data from IRIS dmc.
2	vlp_spectra.m	Displays spectra of broadband data to identify VLP events.
3	create_template.m	Formats identified VLP event times for use in times_series_run_workflow.m.
4	timeseries_run_workflow.m	Conducts the matched filtering process.
5	plot_corr_dist.m	Visualize the distribution of correlation values between detected events and template event. This will help you determine if you need to adjust your detection threshold for timeseries_run_workflow.m.
6	waveform_catalog.m	Compiles list of events for each template.
7	detect_comparison.m	Creates a catalouge of unique events from each template's catalogue.
8	plot_cat_record_sec.m	Plot record sections of the events in the original catalogue to determine clearest events. Plot signal on a single station and save maximum amplitude.
9	examine_clusters_fast.m	Performs clustering of events in catalogue B.
10	amp_dist.m	Analyze amplitude and estimate of energy distribution for events in the catalogue.
11	poisson_test.m	Analyze timing of events and clusters.
12	create_cat_excel.m	Combine all important information for the final catalogue and export as Excel spreadsheet.
13	record_section.m	Plot the record section for any specified time.