Michigan Technological University

Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's Reports

2023

COMPREHENSIVE ANALYSIS OF SEISMIC SIGNALS FROM PACAYA VOLCANO USING DEEP LEARNING EVENT DETECTION

Jessica L. DeVlieg
Michigan Technological University, jldevlie@mtu.edu

COMPREHENSIVE ANALYSIS OF SEISMIC SIGNALS FROM PACAYA
VOLCANO USING DEEP LEARNING EVENT DETECTION

By

Jessica L. DeVlieg

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Geophysics

MICHIGAN TECHNOLOGICAL UNIVERSITY

2023

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Geophysics.

Department of Geological and Mining Engineering and Sciences

Thesis Advisor: *Gregory Waite*

Committee Member: *Roohollah Askari*

Committee Member: *Thomas Oommen*

Department Chair: *Aleksey Smirnov*

Table of Contents

Preface

The work contained herein was undertaken by Jessica DeVlieg in partial fulfillment of the requirements of a MS in Geophysics at Michigan Technological University. This included data analysis from seismometers, interpretation, and preparations of this manuscript. Gregory Waite assisted with the planning and implementation of all phases of this project, especially the data analysis and interpretation.

Acknowledgements

The success of this project started long before the project even began when I went through the challenge and was given the opportunity to attend Tech. I want to thank Luke Bowman and Blaire Orr for giving me an opportunity and working with me to open a path which eventually had led to working on a volcano in Guatemala. I am grateful for the time, energy, and kindness they showed throughout this process.

I want to thank INSIVUMEH and all the employees for sharing their data to help support this project. I also want to thank Federica Lanza for doing the great work that I was able to build upon in this thesis. I want to thank my teacher and advisor, Greg Waite, for being patient and kind and always instilling excitement about the process even when it includes hours of trying to figure out why the code wasn't working. Greg was a great advisor to work with and always took extra time to work through problems with me to help me understand.

I want to thank the National Service Graduate Fellowship for their financial support with this project and classes while I attended this school.

Lastly, I want to thank the support of my family and friends for listening to me and encouraging me throughout this entire process. I would not have been at this point if it weren't for you.

Abstract

Pacaya volcano located 30 km SW of Guatemala City, Guatemala, has been erupting intermittently since 1961. Monitoring of seismicity is crucial to understanding current activity levels within Pacaya. Traditional methods of picking these small earthquakes in this noisy environment are imprecise. Pacaya produces many small events that can easily blend in with the background noise. A possible solution for this problem is a machine learning program to pick first arrivals for these earthquakes. We tested a deep learning algorithm (Mousavi et al., 2020) for fast and reliable seismic signal detection within a volcanic system. Data from multiple deployments were used, including permanent and temporary arrays from 2015 to 2022. Initially over 12,000 independent events were detected although most were unlocatable. A predetermined 1D velocity model calculated by Lanza & Waite (2018) was initially used to locate the earthquakes. This velocity model was updated using VELEST and the locations were calculated using new 1D P-wave and S-wave velocity models. This resulted in 512 events after a quality control filtering process. These events ranged in depths from -2.5 km (summit of Pacaya) to 0 km (sea level) all located directly beneath the vent. The detection process took about 2-3 hours per 15 days on each 3-component broadband seismometer. The method shows promise in providing an efficient and effective method to pick volcano tectonic seismic events, and it did well identifying the emergent arrivals in these datasets; however, it has shortcomings in detecting some low-frequency event types. This could be addressed through additional training of the algorithm. The very low speeds in our new P-wave and S-wave velocity models highlight the poor consolidation of the young MacKenney cone. Further study is encouraged to better understand the accuracy and type of earthquakes picked, especially the increased level of activity during or leading up to an eruption at Pacaya volcano.

Introduction

   Seismic events produced by volcanoes can tell us an increasing amount of information about conditions within the volcano, such as magma movement, level of activity, and fluid transport. In recent years with the increased number of seismometers along with the associated mountains of data, traditional methods of detecting earthquakes are proving to be too slow to efficiently process the data. The most common method for detecting earthquakes has been short-term average/long term average (STA/LTA). Although this method is computationally simple and faster than manually picking the events, it still takes time and performs poorly at picking events with a low signal-to-noise ratio (SNR) (LeCun, Bengio, & Hinton, 2015). Low SNR is especially common on volcanoes where events are small and background signals from volcanic activity produce enough noise to drown out small seismic activity. Earthquake detection methods which utilize machine learning techniques have been increasingly employed due to their ability to process data much faster and pick earthquakes with a low false positive rate even in regions with high levels of noise (LeCun et al., 2015). The purpose of this study is to analyze the feasibility of a machine learning process for monitoring volcanoes such as Pacaya volcano, Guatemala. A portion of the data analyzed here has been previously processed by using STA/LTA and template matching by (Lanza & Waite, 2018). A comparison of the findings will be discussed. In addition, we use the newly identified event catalog to analyze the velocity structure of the cone and attempt to model source mechanisms from first motions.

Background

History of Pacaya Volcano

   Pacaya volcano, Guatemala is a relatively young active composite stratovolcano that formed 23,000 years ago. It is located 30 km south of Guatemala City, Guatemala. The volcano is 2550 m high and consists of several cones, Cerro Chino, Cerro Grande, Cerro Chiquito, and the MacKenney cone (14.381° N, 90.601° W) which is the only cone currently active (Rose, Palma, Escobar Wolf, & Matías Gomez, 2013). The development of the Pacaya Volcanological Complex has grown in 4 main stages (Bardintzeff & Deniel, 1992). Most of the initial ancestral volcano has been eroded and subsequently covered by new lava flows. The composition of the flows range from basaltic to basaltic andesite and have been relatively consistent over the history of the complex (Bardintzeff & Deniel, 1992). Recent eruptions at Pacaya began in 1961 and have been continuing intermittently since this time. The eruptions have been a combination of strombolian and plinian with consistent degassing (Rose et al., 2013). Pacaya experiences intermittent flank movement (Gonzalez-Santana, Wauthier, & Burns, 2022) the west side of the cone where the majority of lava flows have occurred in the past 70 years (fig 1). These 'a'ā lava flows have traveled as far as 2.5 km from the cone (Rose et al., 2013). The last large eruption was categorized by lava flows, strombolian eruptions, and ash plumes. The lava flows erupted from a fissure which opened on the southwest flank. During this time over 1,500 farmers were evacuated, and many surrounding buildings were destroyed. This

level of activity (as of writing this paper) lasted from January to August of 2021. Pacaya volcano is currently monitored by National Institute of Seismology, Volcanology, Meteorology and Hydrology, Guatemala (INSIVUMEH), primarily with seismic and visual observations.
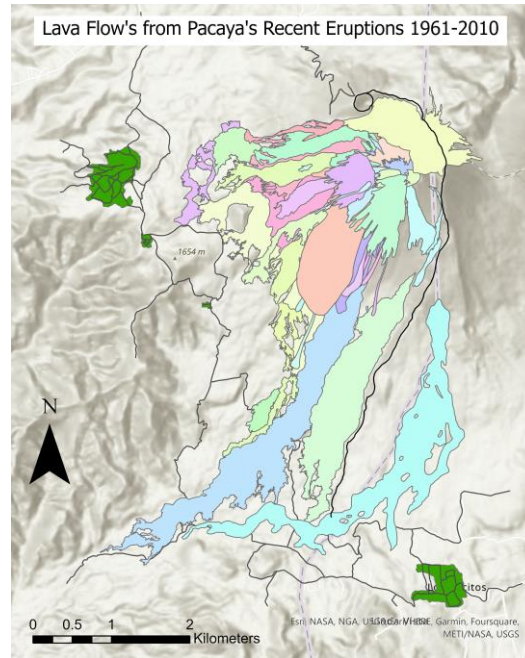

Figure 1: Historic lava flow of Pacaya's Eruptions from 1961 to 2010. Rock from recent eruptions contain large air pockets and are loose which result in large amount of noise in the seismic signal (Matias Gomez, Rose, Palma, & Escobar Wolf, 2012).

Plate Tectonic Structure

Pacaya volcano sits along the Central American Volcanic Arc (CAVA) and has formed as a result of the subduction of the Cocos plate underneath the Caribbean plate (Mann, Gahagan, & Rogers, 2007). The region has been tectonically active for the past few million years. A transform fault runs NE through Guatemala where the North American Plate and the Caribbean Plate meet. Extensional forces in Guatemala have resulted in a graben which is located directly underneath Guatemala City (fig 2). The structure of Guatemala and the surrounding plates has been limitedly mapped, but the various forces contribute to the stresses and volcanic activity in the region (Schaefer et al., 2013).
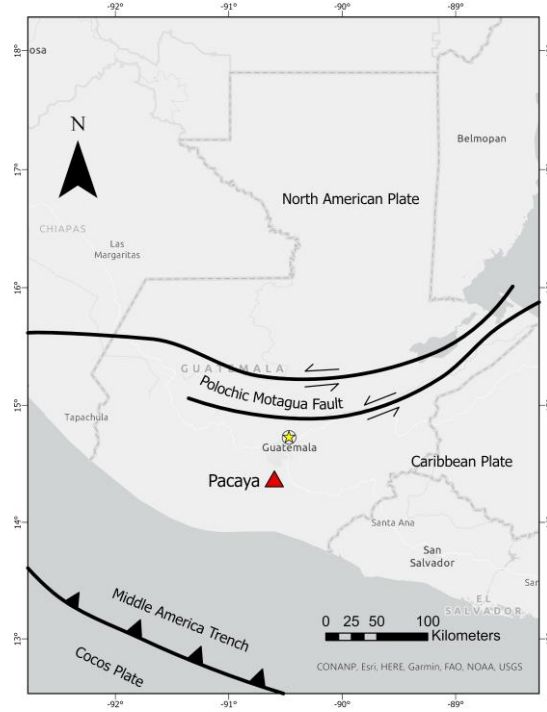
2

Figure 2: Geographical relationship of Pacaya Volcano to Guatemala City and the Cocos Plate subduction zone (Schaefer et al., 2013).

Review of Previous Work

A main goal of this paper is to compare the method of deep learning earthquake detection to traditional methods of earthquake detection. Two temporary station networks used in this study were originally installed and analyzed by (Lanza & Waite, 2018) using STA/LTA and waveform stacking methods. Processing included an initial detection using STA/LTA and was further narrowed by stacking similar waveform picks to amplify picks to detect the lower SNR events. These events were then located through VELEST. A seismic signal was derived from stacking thousands of these events. A subset of events was then inverted to find a source moment tensor. The final locations were found using a nonlinear waveform inversion that used a grid search to find the source type and assure reliability. This resulted in a model which allowed a better understanding of the magmatic conduit structure of Pacaya volcano.

Overview of STA/LTA

STA/LTA has long been used to detect weak motion earthquakes. This method looks at two moving windows, and the average amplitude is calculated over both windows. The shorter window is sensitive to rapid change in amplitude while the longer window is better representative of the background seismic level. When a ratio exceeds a specified threshold, the result is declared an event (Trnkoczy, 1999). STA/LTA has been used for decades to detect earthquakes. The method is good in regions where there is a high SNR ratio. STA/LTA becomes much less effective in regions with a low SNR ratio. STA/LTA is often used in correlation with other detection methods such as cross correlation to find similar events that were not picked up by STA/LTA (Lanza & Waite, 2018). In situations such as volcano monitoring where the noise is extremely high and there is a large quantity of data, researchers are looking towards other methods, such as machine learning algorithms. These could have the benefit of drastically decreasing the processing time and producing a higher number of event picks.

History of machine learning codes

Machine learning technologies are used everywhere in our everyday lives ranging from web searches to voice recognition on your smartphone. The questions we are trying to solve using machine learning have evolved from simple recognition to higher levels of abstraction such as categorizing complex pieces of data within images or speech. The abstraction of data being processed and recognized has led to 'deep learning' systems where data is transformed at different levels which allows the original raw data to represent abstract information. More abstract information can be obtained and classified from more layers of transformations (LeCun et al., 2015).

The two main categories of machine learning are supervised and unsupervised. Supervised learning consists of a training dataset with labeled data which has been hand-picked and labeled by individuals. This process is good for quickly identifying known information within a large dataset. Unsupervised machine learning is trained on an unlabeled dataset and identifies classes based on differentiation of the data. This is best for identifying new classes which might have been previously unknown (Seydoux 2020). Most success thus far has been achieved with supervised learning since there is more control over the outcome and the accuracy can easily be measured.

Machine learning algorithms have been used in seismology for detecting faults (Y. Chen, Verschuur, Guan, & Qu, 2020; Xiong et al., 2018), geological hazard analysis (Ma & Mei, 2021), earthquake detection (Bergen & Beroza, 2018; Y. Chen, 2018; Li et al., 2022; Perol, 2018; Yoon, O'Reilly, Bergen, & Beroza, 2015), classification of seismic signals (Duque et al., 2020; Falcin et al., 2021; Johnson, Ben-Zion, Meng, & Vernon, 2020; Lara et al., 2020; Malfante et al., 2018), volcanic deformation (Anantrasirichai, Biggs, Albino, Hill, & Bull, 2018; Biggs, Anantrasirichai, Albino, Lazecky, & Maghsoudi, 2022), and phase detection (Ross, Meier, Hauksson, & Heaton, 2018). There

is no consensus on the best method or a singular algorithm that has been primarily used. Due to this, many people program their own machine learning algorithm which is specialized for their use case. In this paper we will analyze a relatively widely-used algorithm. The capabilities of these algorithms are growing as we learn more about the best ways to use them for each case. This also leads to more problems that have the potential to be solved using machine learning.

Deep learning algorithm – EQTransformer

EQTransformer was picked for use in this study (S. M. Mousavi et al., 2020) due to its applicability in a variety of geologic scenarios, the tested background, and its ability to rapidly assess multiple stations at once. EQTransformer was developed by a team at Stanford University and Georgia Institute of Technology. The development of this code focused on picking precise and accurate earthquake detection and phase picking. Phase picking, in this case, is measuring the arrival time of P-wave and S-wave within the earthquake detection to be used to accurately locate earthquakes. These need to be very accurate since a very small-time error can result in the earthquake location to be 10s of meters off. Accurate detection and phase picking is also important to reduce the number of false positives.

The development of this code is based on the concept that specific parts of a seismic signal are more important to determining the features. This is done by applying an attention mechanism which is based on how humans regulate incoming signals into important information i.e., important information is in focus while less important information is fuzzy. EQTransformer can utilize this mechanism at a global level and at the local level to simultaneously pick the earthquake, P-wave arrival, and S-wave arrival. In order to keep the processing time down, a down-sampling layer was added (S. M. Mousavi et al., 2020). This resulted in a 56 layered deep neural network with about 372K trainable parameters, which results in a high number of accurate event picks. EQTransformer is a supervised deep learning algorithm and was trained by the developers on STanford EArthquake Dataset (STEAD), a global earthquake dataset with a combination of noise and seismic signals (fig 3). There is an option to train the model with a different set of seismic signals, but for the purpose of this study the initial training set was used. Training the model specifically for volcanic seismic signals produced at Pacaya would likely result in a higher number of detections, but we wanted to utilize the initial trained model. EQTransformer was compared to traditional earthquake detection methods and other deep learning codes and was found to detect more earthquakes at a lower SNR than the comparative earthquake detection methods (S. M. Mousavi et al., 2020). It was also able to detect lower magnitude earthquakes down to -1 Mw (S. M. Mousavi et al., 2020). Situations with a lower SNR, larger distance between seismic stations and smaller magnitude earthquake are output with a lower probability. These situations can be easily removed by an increased probability which is required to classify the detection as an event (S. M. Mousavi et al., 2020). EQTransformer does not include the capability to classify and distinguish between types of seismic signals, and the analysis of such is not covered in this paper.

EQTransformer has been used in a variety of studies including cataloging earthquakes in the southern Mariana subduction zone (H. Chen et al., 2022), seismicity analysis in Ghana (Mohammadigheymasi et al., 2023), foreshock sequencing of the 6.4 MS of Yangbi earthquake (Zhu, Yang, Tan, et al., 2022), and seismicity at extinct mid-ocean ridge (Zhu, Yang, Yang, & Zhang, 2022).

Training Dataset – STEAD

STEAD was developed to make the process of training seismic detection and classification machine learning algorithms easy and result in an accurate product with a robust set of high-quality labeled data. This dataset includes a combined group of seismic signals and noise which amount to about 1.2 million time series or over 450k events. These earthquakes range globally and have a variety of source mechanisms. Additionally, a test dataset is included with seismic data from Japan which is not included in the original training set. This test set allows the user to validate their model after it was trained using STEAD. With the introduction of new methods of earthquake detection, the ability to validate methods against one another is necessary for accurate comparison. This gives the user the ability to isolate the earthquake detection code for deficiencies. The global dataset also allows for many possible types of earthquakes and seismic signals. This includes regions with low SNR and low magnitude earthquakes (S. Mostafa Mousavi et al., 2019). STEAD was used to train EQTransformer for these features, with the goal of having an overarching algorithm which could be used in many situations.
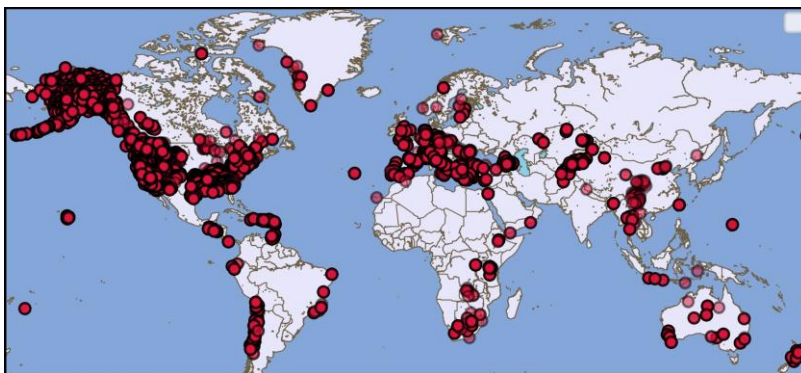

Figure 3: Denotes the stations where data was used for the training EQTransformer. This included labeled seismic signals and labeled noise (S. Mostafa Mousavi, Sheng, Zhu, & Beroza, 2019).

Methodology

Datasets

There were two groups of datasets collected from Pacaya in which the EQTransformer detection method was applied. The first was a set of temporary networks

(dataset 1 and 2) which were deployed in 2013 and 2015 respectively (Lanza & Waite, 2018). These were also used in the original analysis performed by (Lanza & Waite, 2018). Dataset 3 is a permanent network with a larger aperture. This network was installed and is monitored by INSIVUMEH.

Dataset 1: Temporary Network - 2013

The first temporary network included 4 stations deployed from October 31st to November 11th 2013 (Lanza & Waite, 2018). These include three Guralp CMG-ESPC 3-component broadband seismometers (60s corner period) and one Guralp CMG-40T sensor (30s corner period). The seismometers range from 500m to 1600m from the summit and surround the vent (Figure 5b). Stations P01 and P02 contained 3-element equilateral triangular infrasound arrays with 30m between elements. All data was recorded on Reftek 130 digitizers collecting 125 samples per second as they ran in continuous mode with GNSS clocks for timing (Lanza & Waite, 2018). Weak strombolian style explosive eruptions dominated the activity at Pacaya during this time.

Dataset 2: Temporary Network - 2015

Temporary network 2 was a larger network of short-period stations installed in 2015 (Lanza & Waite, 2018). This included a network of 19 Sercell L22 3-component sensors which were distributed around the active vent on MacKenney Cone (figure 5c). The seismometers were located between 100 m and 1500 m away from the vent. These operated in 2015 from January 10th to the 22nd. The data was recorded on Reftek 130 digitizers for each station which operated in continuous mode at 125Hz with GPS. Activity at Pacaya during this time can be characterized as passive outgassing. There was no visible magmatic activity at the crater vent.

Dataset 3: Permanent Network – 2019-2022

The last dataset is from continuous data from a network of 5 stations (fig 4). These are maintained and monitored by INSIVUMEH in Guatemala. Their main purpose on Pacaya is volcanic hazard monitoring which means long term records are not routinely stored for analysis. Outages are common and often only 1 to 2 stations were active at a time. The data in this study runs from October 2019 to October 2022. Volcanic activity varied from passive outgassing to strombolian explosions to lava flows during this time.

Figure 4: Past and present seismic stations operated by INSIVUMEH.

Station	Sampling Rate (hz)	Seismometer
PCG	40	Sercel/Mark Products L-4C
PCG2	50	OSOP Sixaola
PCG4	50	Trillium 120
PCG5	50	Trillium 120/ Sercel/Mark Products L-4C
PCGT0	50	OSOP Sixaola

Figure 5: (a) geographic distribution of Pacaya Volcano Complex in relation to Guatemala. (b) geographic distribution of dataset 1. (c) geographic distribution of dataset 2. (d) geographic distribution of dataset 3. Not all these stations were running simultaneously, but at least one ran continuously through the time of this study.
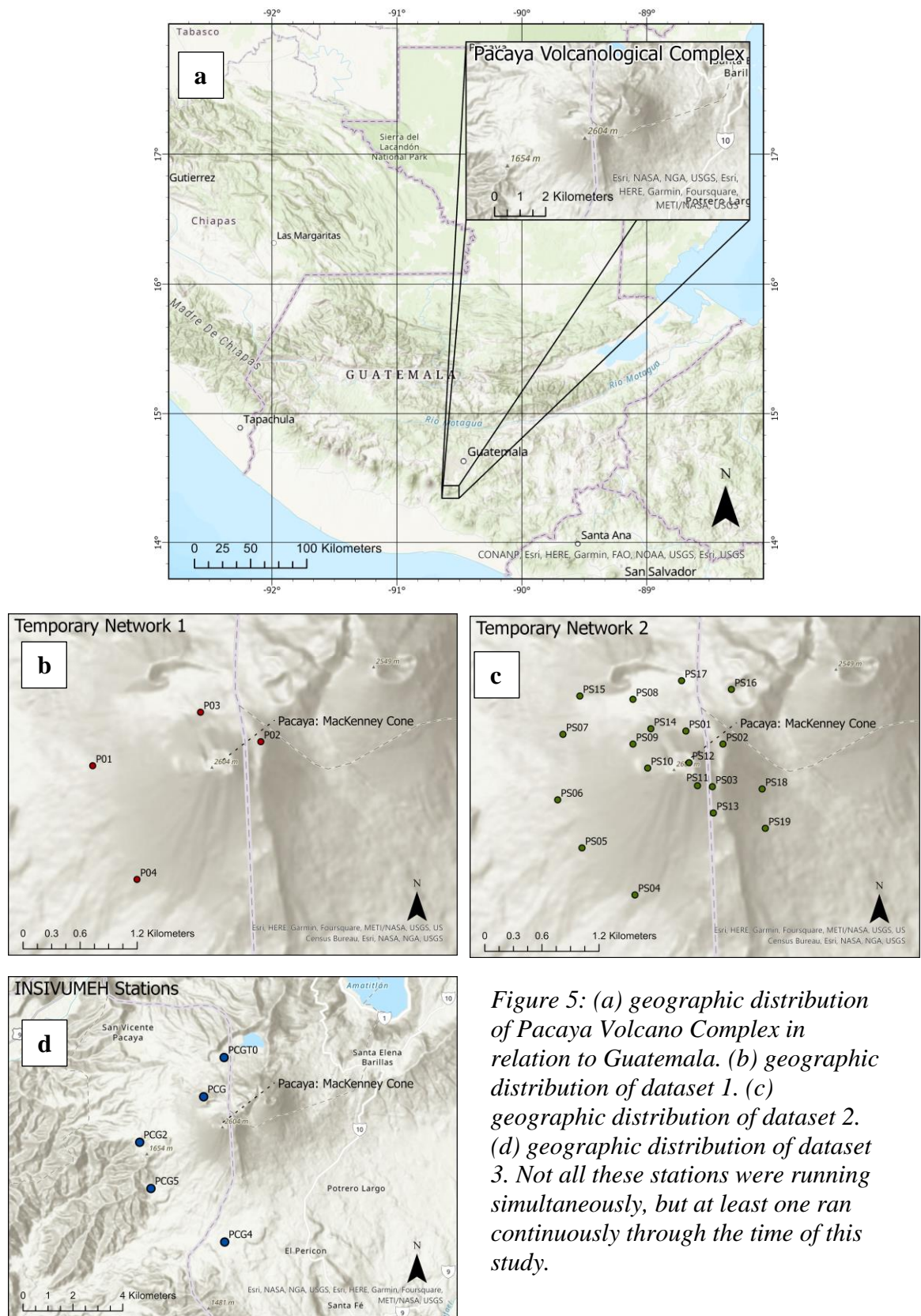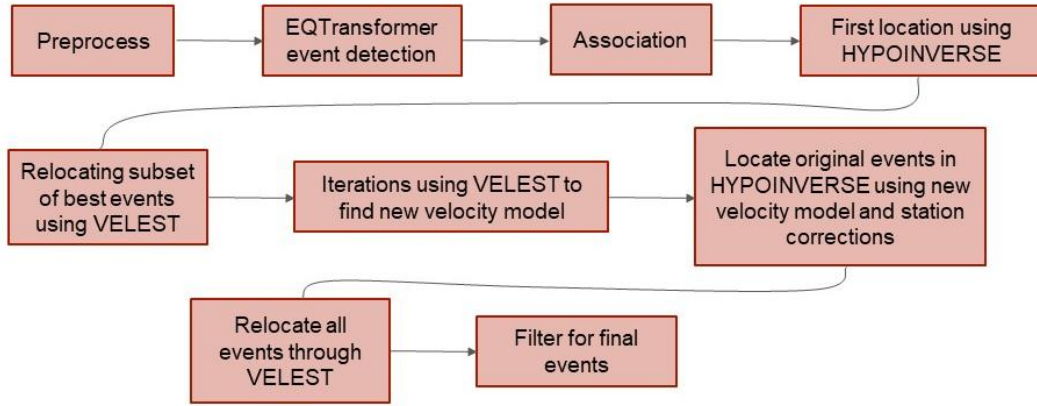
Figure 5: (a) geographic distribution of Pacaya Volcano Complex in relation to Guatemala. (b) geographic distribution of dataset 1. (c) geographic distribution of dataset 2. (d) geographic distribution of dataset 3. Not all these stations were running simultaneously, but at least one ran continuously through the time of this study.

Preprocessing and picking

Preprocessing and picking were completed through EQTransformer with the initial setup. Preprocessing included separating the waveform into small overlapping data chunks and saving them in a .hdf5 and a .csv file. This allowed easy access to the raw data and prepared it for detection. EQTransformer can pick events for multiple stations, but these detections are run independently. The data was split into 15-day segments to make the detection process more manageable by a personal computer. All the processing was completed using an AMD Ryzen 7 5700G with Radeon Graphics 3.80 GHz CPU with 32GB of ram.

EQTransformer is trained to result in almost no false positives at a 0.1 probability detection threshold, but since there is a high level of noise within a volcanic system, a probability threshold of 0.3 was used for an event detection. The probability threshold for P-wave time was 0.3 and for S-wave the probability was also 0.3. Since S-waves in volcanic seismic signals can be weak, distorted, or nonexistent, a detection was included if there was only a P-wave pick and no associated S-wave pick. An event would only be included if there was a P-wave pick. Initial detection tests were performed with a high number of output detection figures that show the pick within the waveform and the uncertainty of the P-wave and S-wave probabilities. The uncertainty is a range of the probability certainty for the duration of each pick. Since these both added little value beyond verification of accuracy and increased the processing time between 2 to 5 times, these parameters were removed in the later data processing.

Phase Association

The phase association algorithm built into EQTransformer was used to link detections recorded at multiple stations for a single event. A minimum of 4 P-wave or S-wave picks were required to define an event. This gave us a rough idea of the number of events, since during most of the data, no more than 3 stations were running congruently. This would later be increased to 4 stations when locating the earthquakes. This is a simple phase association algorithm that utilizes the time and distance from stations to determine an event.

Joint 1D Velocity Model and Inversion

An inversion process utilizing HYPOINVERSE was used to initially locate the earthquakes. HYPOINVERSE2000 was originally developed in 1978 and can use a variety of velocity models to locate earthquakes. The 1D velocity model option was utilized for this study. The 1D velocity model was created by combining two existing velocity models in the region. The first was from (Lanza & Waite, 2018) wherein iterations of the best fit velocity were calculated using VELEST. The locations used to calculate this model originated by (Lanza & Waite, 2018) with the temporary networks 1 and 2 analyzed in this study. This provided an estimate of the velocity model from the vent at Pacaya (-2.5 km) down to 5.5 km below sea level. This information was combined with the velocity model of the Motagua Fault region in Guatemala calculated by (Franco et al., 2009) to estimate the deeper layers. The resulting velocity model along with the associated earthquakes were inputted into HYPOINVERSE with a requirement of 4 stations per earthquake for an event. Due to sparse data retention, this filtered the majority of the time from dataset 3 from Feb 15th, 2021 to August 30th 2021. A minimum of 4 stations were running during the length of both temporary networks, 1 and 2. The phase association resulted in 11,740 events; however, initial location resulted in around 450 events likely due to difficulty picking first arrivals and multiple picks at one station (S-wave pick and P-wave pick) which were filtered out to have less than 4 stations through HYPOINVERSE.

Due to the large increase in locations up from 250 events when compared to (Lanza & Waite, 2018) results, we decided to update the velocity model using the new locations. This would analyze the possibility that the deep locations were an artifact of a low-resolution velocity model and create a clearer understanding of the structure beneath Pacaya volcano.

Updated Velocity Model and Relocation

To update the velocity model, an alternative program VELEST was used (Kissling, Ellsworth, Eberhart-Phillips, & Kradolfer, 1994). The best picks were run through VELEST along with iterations of different velocity models. All event picks were filtered

by rms of <0.25 and an azimuthal gap of <180 degrees. VELEST updates the velocity model to best fit the locations and updates the locations to fit the new velocity model. Station PS12 was chosen as the anchor due to its central location near the vent and high number of picks. The picks were further refined through VELEST by running iterations using the original velocity model and throwing out events where the average residual was greater than 1 second. Next, if a station within a pick had a high residual, the individual station was removed from that event. This process continued until no events outside of these parameters were produced.

The updated velocity model along with the station corrections were then used to rerun the original associated event picks through HYPOINVERSE to produce a new set of locations. These events were then processed through VELEST in single event mode to update the locations. These were filtered using the same parameters to find the final locations.

Focal Mechanisms

High quality events were chosen to analyze focal mechanisms. These were events that had low errors and were centered within the seismic array and were detected by a large majority of the stations. A variety of depths were analyzed. This process was completed manually by analyzing the waveform and picking the first motion. The first motion was defined to be significantly larger than the background noise. The amplitude of the first motion was compared against the surrounding stations to find the expected amplitude. This was used as a base to compare the expected amplitude size of the surrounding stations. Stations located along the east side of the vent consistently showed an up motion as the first motion. First motions along the eastern vent were indistinguishable. The common pattern showed a slow build up to a larger amplitude, likely due to the random path these waves took within the geologic substrate.

Results and Discussion

Initial Detection and Phase Association

The process of running the detection for a single station with continuous data for 15 days under these parameters varied from 2 to 3 hours. This increased to 10 to 18 hours when the parameters of P-wave and S-wave uncertainty were included. Initial detection and association from EQTransformer produced approximately 29,000 events. Due to the limited amount of time when more than 4 stations were running continuously in the permanent network (dataset 3), only dates from February 15th 2021 to August 15th 2021 were associated. Half of the total events (14,500) were from dataset 2. This was due to the higher density and increased number of stations during this period.

EQTransformer Visualizations

Visualizations were made using the initial data to assess accuracy. This feature is built into EQTransformer. Figure (6) shows the station locations of dataset 2 and the corresponding number of detections. Stations closer to the summit (PS01, PS09, PS12) show a higher number of detections, while stations furthest from the summit (PS04, PS05, PS19, PS18) show the lowest number of detections. The number of detections on the stations furthest from the summit are about half of those closest. Figure (7) shows examples of earthquake picks from dataset 2. These range in amplitude but resemble expected volcanic seismic signals.
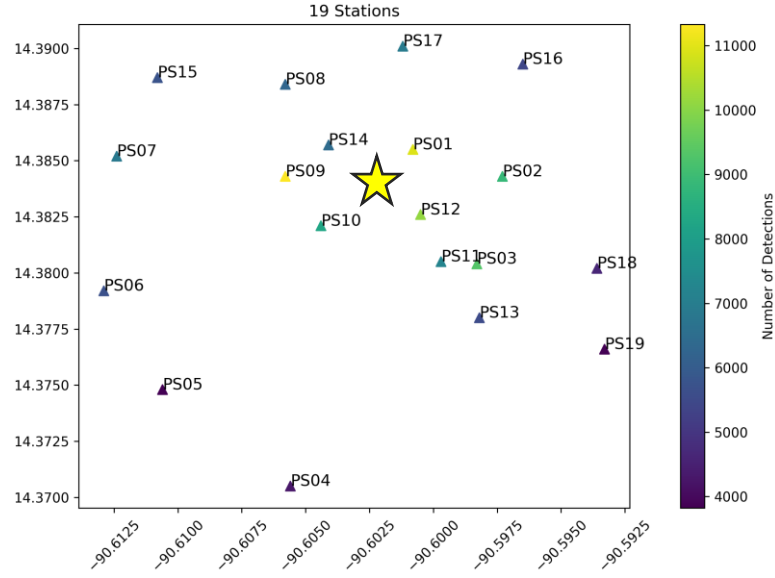

Figure 6: Geographic distribution and associated number of initial earthquake detections from dataset 2. Stations PS14, PS01, PS02, PS12, PS10, and PS09 all circle the vent. For the most part these recorded the highest number of detections while stations much farther from the vent (PS04 and PS05) recorded a relatively low number of detections.
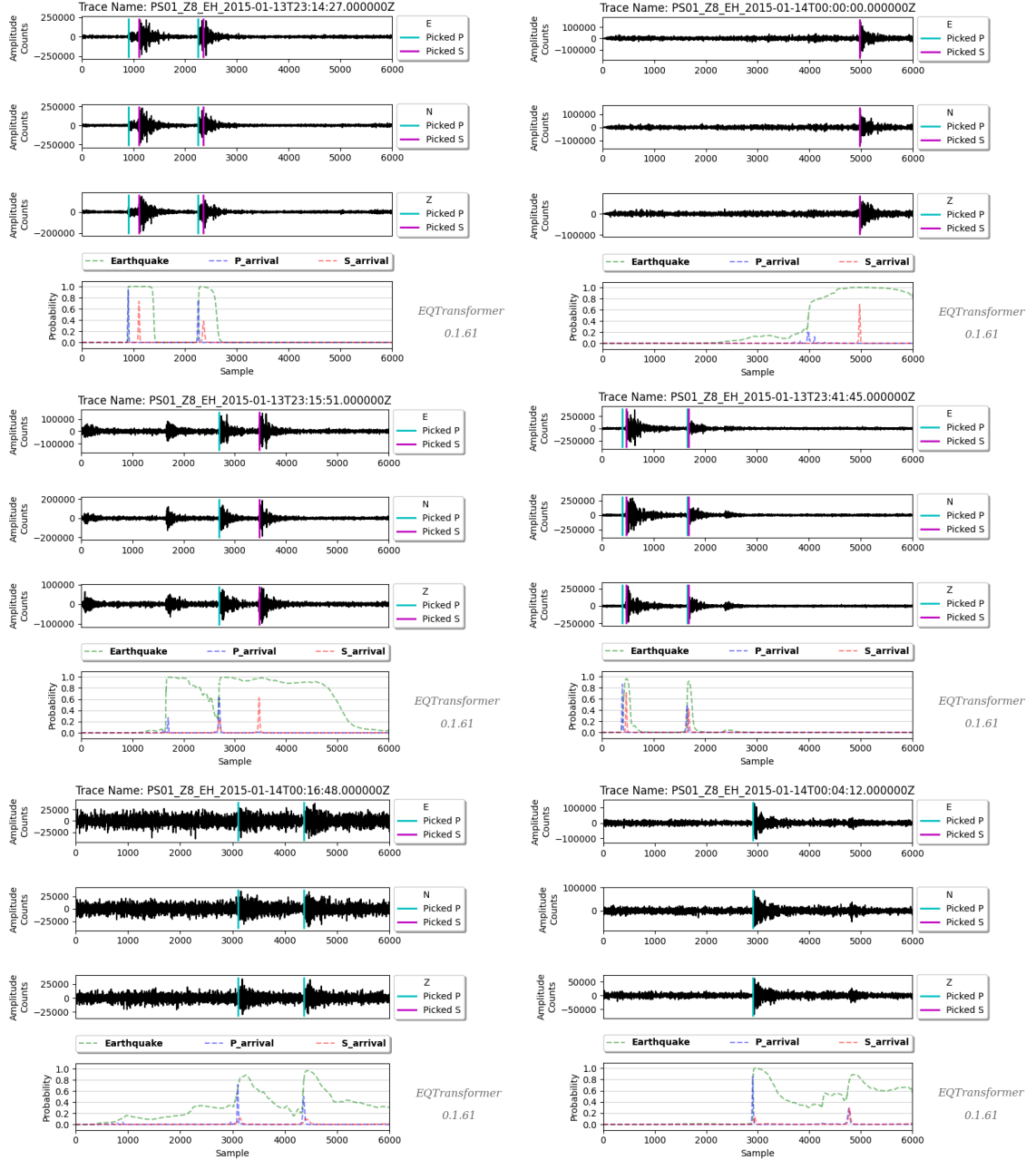
Figure 7: Figures associated with detections from dataset 2. These show picks ranging in amplitude, but all with a relatively high probability of the overall earthquake (initial pick 0.8 or higher). Many events show a P-wave pick, with no associated S-wave pick. Due to the window length, there are times when multiple events appear within the same window.

Initial Location using 1D Velocity Model

The location process using HYPOINVERSE resulted in about 11,000 events from the initial 29,000 associations. This reduction was due to many events not being locatable. However, many of these events had large location uncertainties likely resulting in poor phase picks. In order to produce a subset of the best located events, we applied the following filter criteria: (a) azimuthal gap of <180, (b) rms of <0.5, vertical error of <5 km. This filter procedure left 822 events to be used for computing VP and VS models. All events with an azimuthal gap greater than 180 degrees were outside of the station array and therefore location error was too high. All events had a rms that was relatively high due to the limited stations and difficulty in picking the first arrival. The events with a rms between 0.5 and 0.25 were focused much deeper. These events were left in with a higher error to show an estimation of the deeper events. The vertical error followed the same pattern. Most events with a vertical error of <5 km tended towards a depth of 2 km or less. The higher error vertical error events represented the deeper events. S-picks were also more common on these events. The S-picks were able to be deciphered due to the distance allowing the first arrivals to be further apart. Due to the distance from the stations and the narrow array, vertical error this region is expected. After these filters, associated These events were located mostly directly beneath the volcano from the vent (2.5 km above sea level) to 5 km below sea level. The focus of these events was from sea level up to the vent. Below these was a narrow column of events that went down from the main cluster to about 5 km deep. This column later disappeared when the new velocity model and station corrections were used and was likely an artifact of a low-resolution velocity model. The lack of station corrections during this location process could also account for the apparent vent that traveled down to 5 km. Specifically stations located along the western flank were significantly slower than the velocity model. The events were not necessarily deep, they just traveled extremely slowly in the upper portion of the recent lava flows.

Velocity Model

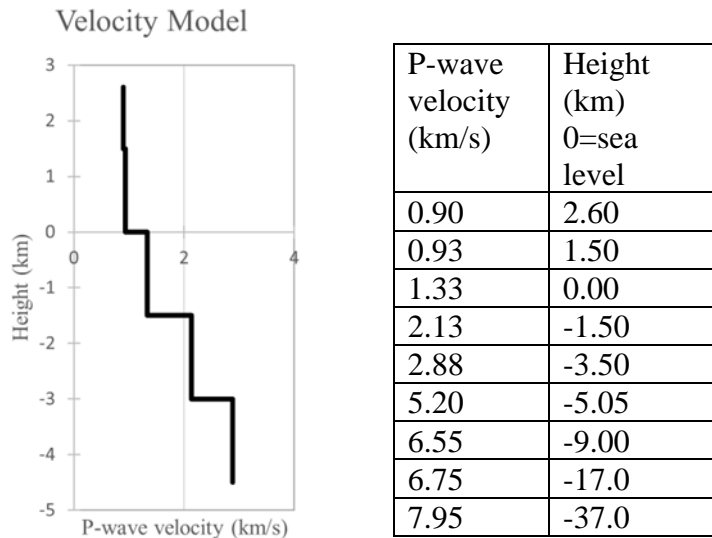| P-wave velocity (km/s)	Height (km) 0=sea level
0.90	2.60
0.93	1.50
1.33	0.00
2.13	-1.50
2.88	-3.50
5.20	-5.05
6.55	-9.00
6.75	-17.0
7.95	-37.0

Figure 8: 1D crustal velocity model estimated using a grid search over all possible velocities and the model stopped once the variation in velocities reach an equilibrium (Lanza & Waite, 2018). This was combined with data of deeper crustal velocity from the associated Motagua Fault region (Franco et al., 2009).

Pick Error

Even though EQTransformer is trained on over 400k events including noise, there is a high level of error associated with picking first arrivals within this dataset. This issue also arose during (Lanza & Waite, 2018) study where STA/LTA and waveform matching techniques were implemented. Stations located along the western flank of Pacaya, specifically those from dataset 2, contained high levels of background noise. This is likely from the tephra from the recent lava flows (fig 1). The large air bubbles and unconsolidated, blocky lava is a slow medium for seismic waves and causes the waves to scatter so that little direct energy reaches the stations. Because the events are so frequent, the arrival of the scattered waves produces noise within the waveform that can obscure subsequent event arrivals. Stations located or east of the vent recorded higher amplitudes since the medium was denser and faster, the seismic waves experienced less attenuation making first arrival picks more accurate. The accuracy of final locations will then rely heavily on stations located on the eastern side of the vent which increases the level of error. Most events were located in the center of the network, which means they need the information from stations located on the western flank. Due to this all locations will contain a high level of error and minimal interpretation will be done on the shape of the final event cloud. The uncertainty of the final picks was reduced through the high filtration and inversion process.

Updated Velocity Model and Relocation

To update the velocity model, multiple iterations of different velocity models were run though VELEST with the best initial picks found using HYPOINVERSE. The slowest velocity model and faster velocity model were the same limits used by (Lanza & Waite, 2018) to predict the original velocity model. The velocity models did not change dramatically with each iteration. All models tended slower until the chosen model (fig 9).
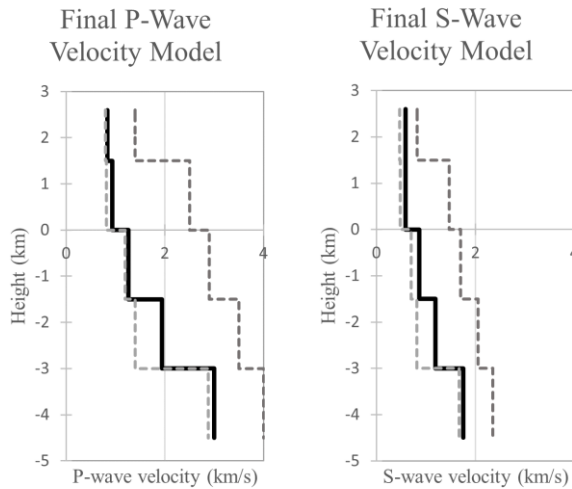
Velocity Model	RMS
Original P-wave	0.362
Final P-wave	0.214

Figure 9: Final 1D crustal velocity models both P-wave and S-wave. Iterations of different velocity models were tested, and the common merge point was the best estimation for an updated velocity model. The S-wave velocity model was created through this process, the initial S-wave 1D starting models were created from the Vp/Vs ratio of √3.

15

This could be due to the poor picks from the high noise level which created local minima in the velocity model.

The average rms of the tested velocity models ranged from 0.214 to 0.362 s. The final model (fig 9) was like the original model except the first two layers from the vent at Pacaya to sea level slightly slower. The remainder of the layers were not altered since there were no earthquakes detected at lower depth. Previously, only the P-wave velocity model was calculated due to the limited number of S-picks, and the S-wave velocity model was estimated using the normal ratio (Vp/Vs) of approximately 1.73 (Lanza & Waite, 2018). In our model, the S-wave velocity model also increased with the P/S ratio averaging 1.5 for the top two layers. Since most of the earthquakes were located from 1 km depth to the vent, this is the location where the velocity model changed the most dramatically. The extremely slow velocity model likely results from the unconsolidated tephra and lava flows that make up the young volcano.

Station corrections should reflect the superficial geology with positive delays associated with slower velocities beneath the stations. The P-wave station corrections from this model ranged from -0.36 to 0.48 s with the distribution shown in (fig 10). The S-wave station corrections ranged from -0.23 to 0.7 s with all but two stations showing a positive (slow) station correction. The corrections were compared to the geology of Pacaya. Stations positioned along the western flank showed a slowing of wave which is consistent with the recent and loose lava flows which predominately flow west of the vent (Lanza, Kenyon, & Waite, 2016). Stations which were located further away and on the eastern side of the vent had faster station corrections. This was consistent with the more competent rock that makes up the older portions of the Pacaya complex. The S-wave corrections showed a similar pattern although they are skewed from 0 suggesting that the 1D S model may still be too fast.
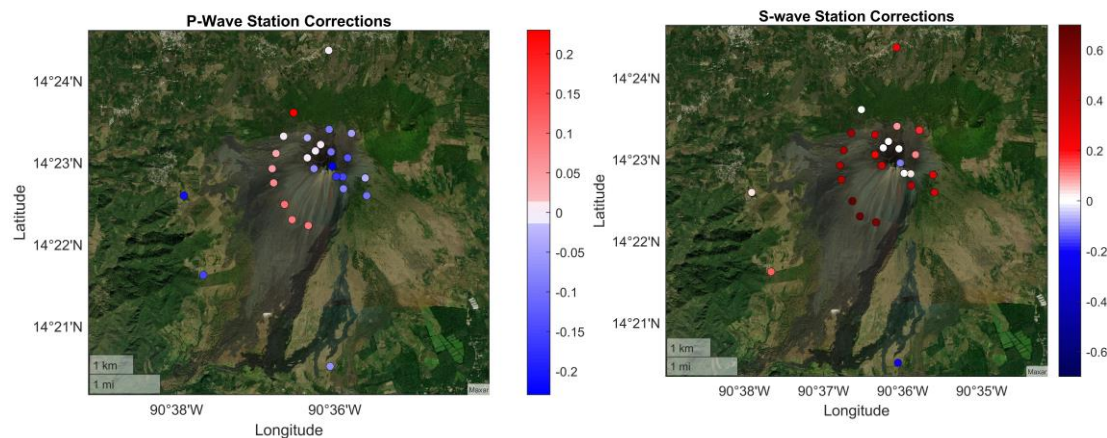

Figure 10: Station delays for final velocity model. Slow stations are indicated with a red marker, while fast stations are indicated with a blue marker. Stations on the edifice on new lava flows are red while stations further away located on more dense material are blue.

The relocation and filtration process resulted in events that were still directly underneath the vent, but the depth was much shallower ranging from 0 (sea level) to -2.5 km (vent). A total of 512 events were located here with a higher density around -1.5 km. Of the original number of events detected only 4.2% were accurately located. There were few events that were scattered from 0 to 2 km depth (fig 11). We explored the event depth sensitivity to the starting location in HYPOINVERSE and found that it had a large impact on the final locations and the associated error. When the starting depth was set to -2.5 km the final events located were few and the vertical error was on the range of 99 km. The starting depth of -1 km resulted in the most events after the filtration process. These events also were associated with the smallest vertical and horizontal error and the lowest rms. This was likely due to the high error associated with the picks.
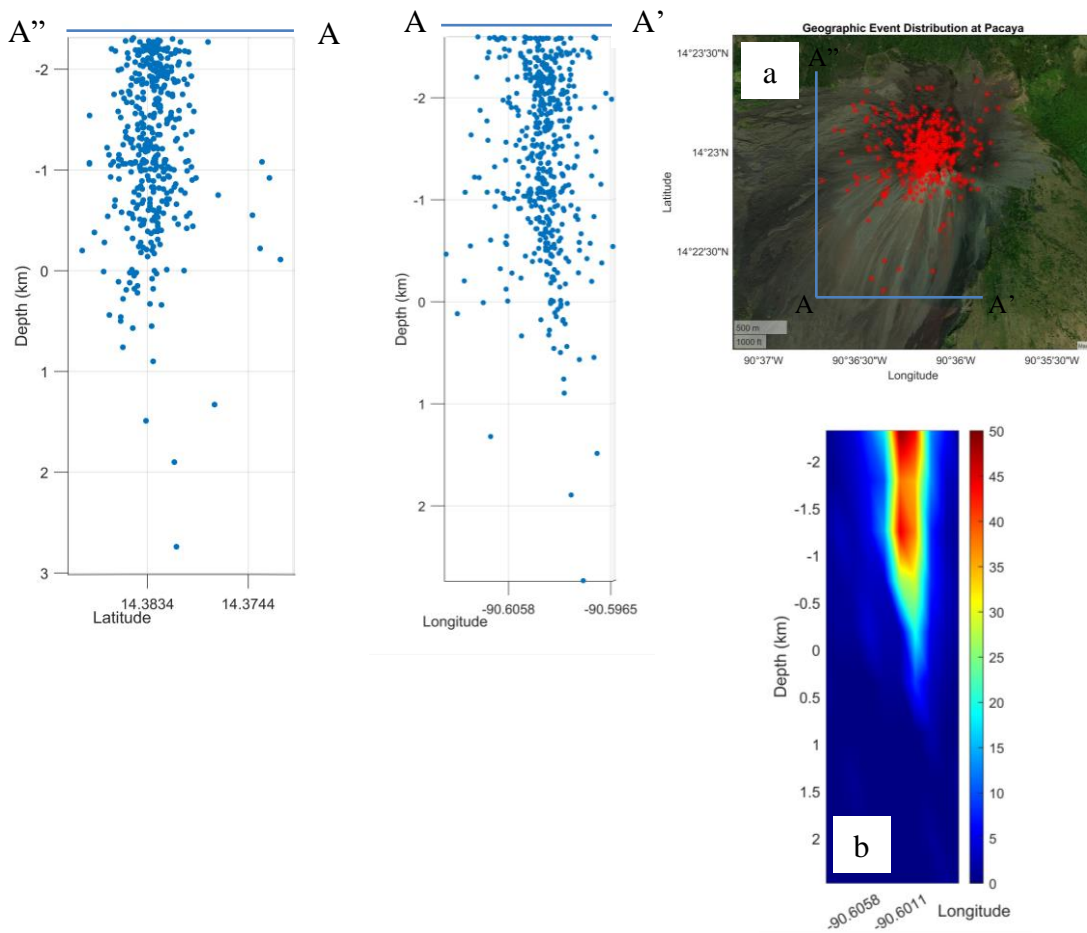

Figure 11: (a) Final location of events at Pacaya volcano. The events were in a roughly symmetrical cone that bent slightly NE around a depth of -1 km. This is beneath the depth of the edifice and below the surrounding topography. (b) Density map of events. The focus of the events is from -1.5 km to the summit. The highest density location reflects that found in (Lanza & Waite, 2018) with an additional high density zone around -1.25km. The slight bend in the events to the NE is reflected below 0 km.

Events located by (Lanza & Waite, 2018) were also focused directly underneath the vent but were shallower ranging from -2.5 km to -1.7 km depth. These events are within the edifice. There were fewer final events, around 200 events. The events in that study formed a cylinder shape that abruptly stopped at -1.7 km. The final event cloud found in this study showed an elongated column that followed a similar pattern of a high density of events directly underneath the vent with addition events that decreased density with depth. Even though events were picked in the same location as (Lanza & Waite, 2018), there were only 24 overlapping events which was 13% of the original catalog (fig 12). The speed at which these events were picked was faster. The approximate time to pick events using the STA/LTA and waveform matching was multiple months for 1 month of data. This included time manually looking through the waveform and picking events that represent usual events during this time. This process likely resulted in more representative events than using EQTransformer, since they were specific to Pacaya. This process also resulted in less overall picks, since Pacaya is a noisy environment with small seismic activity that is difficult to differentiate from the noise.
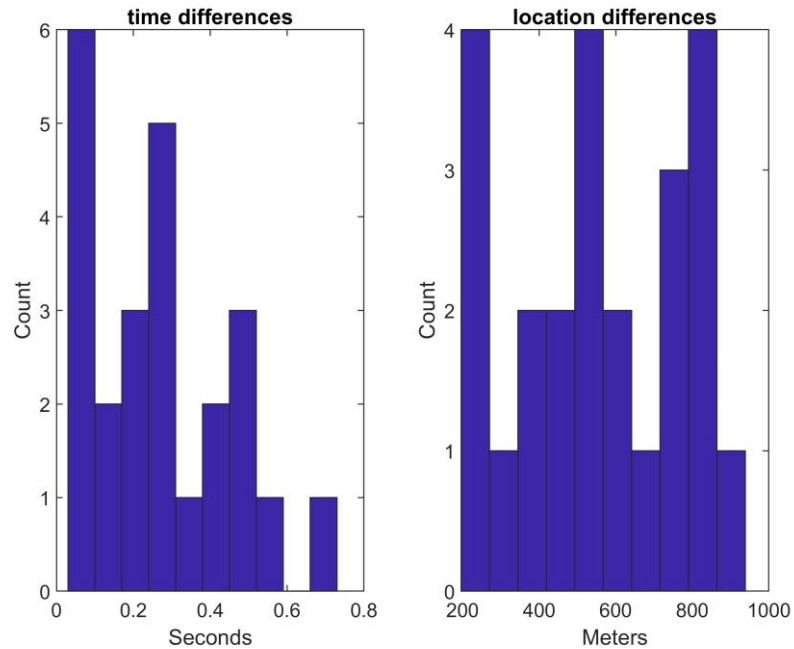

Figure 12: Overlapping events from the catalog produced in this study and the original event catalog produced by (Lanza & Waite, 2018). The difference in events is likely due to the types of events picked by the deep learning algorithm and those picked using cross-correlation.

EQTransformer was not trained specifically for volcanic events, but quickly picked out events representative of tectonic activity which often occur within a volcanic system. Events that were located followed the pattern of high frequency with large initial amplitude spikes that gradually attenuated. EQTransformer was trained on a wide variety of tectonic seismic signals and associated noise but had no training dataset that contained volcanic specific signals such as tremors, long period (LP) or hybrid events. Many events picked did not include low frequency events. These events are not necessarily absent at Pacaya Volcano, but since EQTransformer is trained exclusively on tectonic events, these are not picked as events. This can be remedied in the future by training EQTransformer on a catalog of events that contains a large portion of volcanic events.
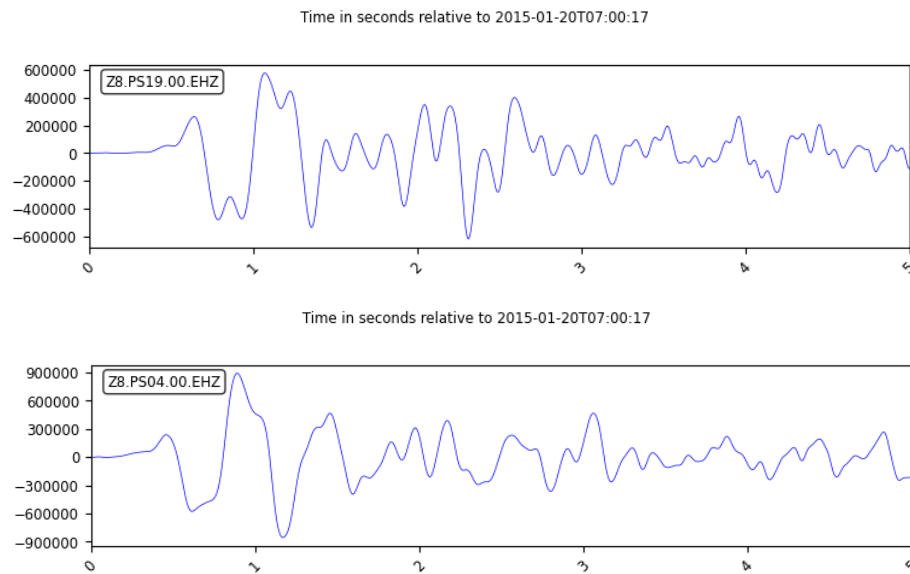

Figure 13: Manual pick examples for low frequency events, which are not included in the final event log. EQTransformer can be trained to pick low frequency events and other volcanic specific earthquakes.

A subset of events was chosen manually to compare first motions. Focal mechanisms were not plotted, due to the difficulty in picking first motions. There was a mix in the first arrival motions between "up" and "down" with the majority showing down. First motions were also clearest at stations PS09 and PS10 (fig 14). Stations located along the western flank consistently did not show clear first motions and were not able to be picked for most events. This is likely due to the tephra and new lava new lava flows within this region. Given that events picked were high frequency, this could have led to a highly level of attenuation and scattering within this catalog. Low frequency events could have different results for the first motion picks.
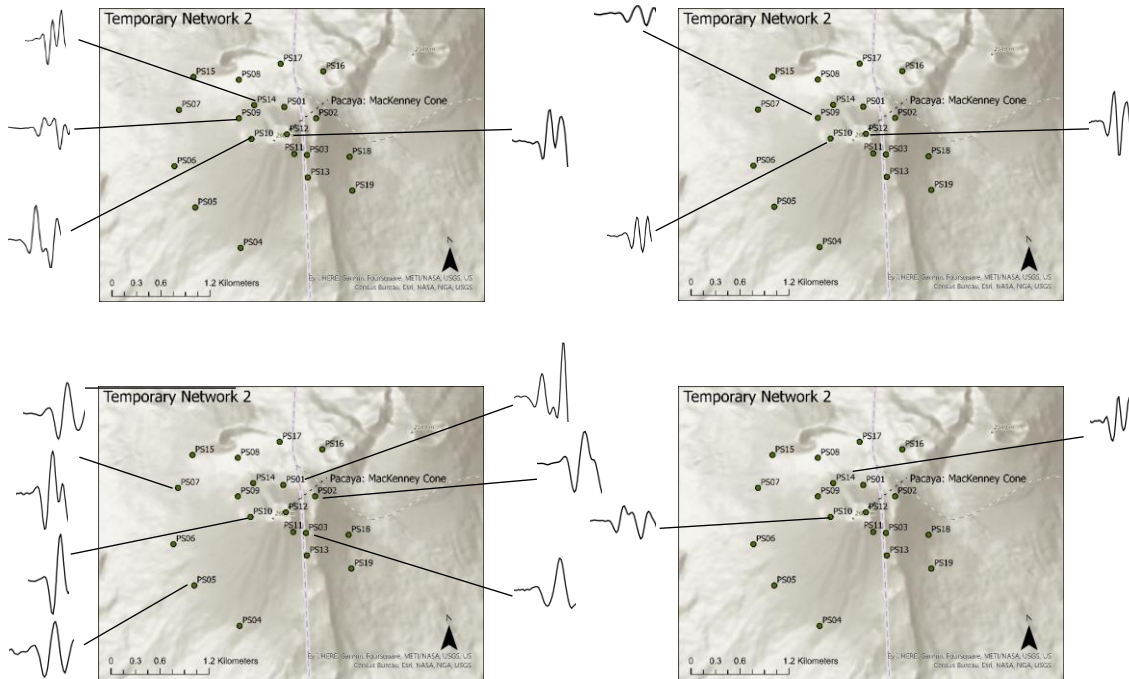
Figure 14: Example first motion picks from various events detected by dataset 2. These first motions are not to scale but represent the pattern of "down" within the first motion. This analysis was conducted on randomized picks that contained low error and low azimuthal gap. Stations PS09 and PS10 contained the clearest first motions, while stations PS04-PS07 showed less clear first motions on the same events.

Geologic implications

Beyond the work mapping the seismicity at Pacaya volcano (Lanza & Waite, 2018), imaging techniques have been limited to GPS and InSAR which have high temporal and spatial resolution, but only can map shallow features. Seismic events at depths below -1.5 km have not been mapped before at Pacaya. The majority of the events in the final catalog are from dataset 2 (January 2015). Although the activity at this time was associated with degassing, in June of 2015 the most recent eruptive period began and continued through November of 2021. This activity started with ash plumes and incandescence inside the vent. Growth of a new pyroclastic cone was confirmed inside the vent later in 2015. Magmatic activity in January of 2015 was likely a precursor

related to the following activity and the events cataloged here could be representative of activity in the magmatic chamber at Pacaya.

Although the locations of specific events changed with multiple iterations of the velocity model and there was some difficulty picking the first arrival at some stations, the overall event cloud remained relatively similar throughout the iterations. This final event cloud could likely represent part of the magma system under Pacaya.

Conclusion

EQTransformer is a viable method for quickly analyzing large quantities of waveform data to pick volcanic tectonic seismic events. The program picks a high percentage of false or unlocatable events, but these can be filtered out through adjusting the probability of an event when processing and the inversion process. The accuracy for picking within a volcanic system can also be increased with training EQTransformer on a set of volcanic seismic events. Smaller amplitude events were picked than with previous STA/LTA and waveform matching methods. More studies need to be conducted on the types of events that EQTransformer is equipped to pick and test this process with volcanoes that contain a different magmatic composition and are eruptive. Further studies also need to analyze the viability of using EQTransformer as a monitoring method to continually monitor active volcanoes such as Pacaya. Further analysis using EQTransformer could also be used to analyze the western flank underlying Pacaya that has led to collapses in recent history.

Reference List

Anantrasirichai, N., Biggs, J., Albino, F., Hill, P., & Bull, D. (2018). Application of Machine Learning to Classification of Volcanic Deformation in Routinely Generated InSAR Data. *Journal of Geophysical Research: Solid Earth*. doi:10.1029/2018jb015911

Bardintzeff, J.-M., & Deniel, C. (1992). Magmatic evolution of Pacaya and Cerro Chiquito volcanological complex, Guatemala. *Bulletin of Volcanology, 54*, 17. doi:10.1007/BF00301482

Bergen, K. J., & Beroza, G. C. (2018). Earthquake Fingerprints: Extracting Waveform Features for Similarity-Based Earthquake Detection. *Pure and Applied Geophysics, 176*(3), 1037-1059. doi:10.1007/s00024-018-1995-6

Biggs, J., Anantrasirichai, N., Albino, F., Lazecky, M., & Maghsoudi, Y. (2022). Large-scale demonstration of machine learning for the detection of volcanic deformation in Sentinel-1 satellite imagery. *Bull Volcanol, 84*(12), 100. doi:10.1007/s00445-022-01608-x

Chen, H., Yang, H., Zhu, G., Xu, M., Lin, J., & You, Q. (2022). Deep Outer-Rise Faults in the Southern Mariana Subduction Zone Indicated by a Machine-Learning-Based High-Resolution Earthquake Catalog. *Geophysical Research Letters, 49*(12). doi:10.1029/2022gl097779

Chen, Y. (2018). Fast waveform detection for microseismic imaging using unsupervised machine learning. *Geophysical Journal International, 215*(2), 1185-1199. doi:10.1093/gji/ggy348

Chen, Y., Verschuur, E., Guan, Z., & Qu, S. (2020). Automatic high-resolution microseismic event detection via supervised machine learning. *Geophysical Journal International, 222*(3), 1881-1895. doi:10.1093/gji/ggaa193

Duque, A., González, K., Pérez, N., Benítez, D., Grijalva, F., Lara-Cueva, R., & Ruiz, M. (2020). Exploring the unsupervised classification of seismic events of Cotopaxi volcano. *Journal of Volcanology and Geothermal Research, 403*. doi:10.1016/j.jvolgeores.2020.107009

Falcin, A., Métaxian, J.-P., Mars, J., Stutzmann, É., Komorowski, J.-C., Moretti, R., . . . Lemarchand, A. (2021). A machine-learning approach for automatic classification of volcanic seismicity at La Soufrière Volcano, Guadeloupe. *Journal of Volcanology and Geothermal Research, 411*. doi:10.1016/j.jvolgeores.2020.107151

Franco, A., Molina, E., Lyon-Caen, H., Vergne, J., Monfret, T., Nercessian, A., . . . Requenna, J. (2009). Seismicity and Crustal Structure of the Polochic-Motagua Fault System Area (Guatemala). *Seismological Research Letters, 80*(6), 977-984. doi:10.1785/gssrl.80.6.977

Gonzalez-Santana, J., Wauthier, C., & Burns, M. (2022). Links between volcanic activity and flank creep behavior at Pacaya Volcano, Guatemala. *Bulletin of Volcanology, 84 (9)*. doi:10.1007/s00445-022-01592-2

Johnson, C. W., Ben-Zion, Y., Meng, H., & Vernon, F. (2020). Identifying Different Classes of Seismic Noise Signals Using Unsupervised Learning. *Geophysical Research Letters, 47*(15). doi:10.1029/2020gl088353

Kissling, E., Ellsworth, W. L., Eberhart-Phillips, D., & Kradolfer, U. (1994). Initial reference models in local earthquake tomography. *Journal of Geophysical Research, 99*, 19635-19646.

Lanza, F., Kenyon, L. M., & Waite, G. P. (2016). Near-Surface Velocity Structure of Pacaya Volcano, Guatemala, Derived from Small-Aperture Array Analysis of Seismic Tremor. *Bulletin of the Seismological Society of America, 106*(4), 1438-1445. doi:10.1785/0120150275

Lanza, F., & Waite, G. P. (2018). Nonlinear Moment-Tensor Inversion of Repetitive Long-Periods Events Recorded at Pacaya Volcano, Guatemala. *Frontiers in Earth Science, 6*. doi:10.3389/feart.2018.00139

Lara, P. E. E., Fernandes, C. A. R., Inza, A., Mars, J. I., Metaxian, J.-P., Dalla Mura, M., & Malfante, M. (2020). Automatic Multichannel Volcano-Seismic Classification Using Machine Learning and EMD. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13*, 1322-1331. doi:10.1109/jstars.2020.2982714

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444. doi:10.1038/nature14539

Li, W., Chakraborty, M., Sha, Y., Zhou, K., Faber, J., Rümpker, G., . . . Srivastava, N. (2022). A study on small magnitude seismic phase identification using 1D deep residual neural network. *Artificial Intelligence in Geosciences, 3*, 115-122. doi:10.1016/j.aiig.2022.10.002

Ma, Z., & Mei, G. (2021). Deep learning for geological hazards analysis: Data, models, applications, and opportunities. *Earth-Science Reviews, 223*. doi:10.1016/j.earscirev.2021.103858

Malfante, M., Dalla Mura, M., Mars, J. I., Métaxian, J.-P., Macedo, O., & Inza, A. (2018). Automatic Classification of Volcano Seismic Signatures. *Journal of Geophysical Research: Solid Earth, 123*(12), 10,645-610,658. doi:10.1029/2018jb015470

Mann, P., Gahagan, L., & Rogers, R. (2007). Overview of plate tectonic history and its unresolved tectonic problems. In *Central America*.

Matias Gomez, R., Rose, W. I., Palma, J. L., & Escobar Wolf, R. (2012). Map of the 1961-2010 eruption of Pacaya Volcano, Guatemala. *GSA Digital Map and Chart Series, 10*(10). doi:doi:10.1130/2012.DMCH010

Mohammadigheymasi, H., Tavakolizadeh, N., Matias, L., Mousavi, S. M., Silveira, G., Custódio, S., . . . Moradichaloshtori, Y. (2023). Application of deep learning for seismicity analysis in Ghana. *Geosystems and Geoenvironment, 2*(2). doi:10.1016/j.geogeo.2022.100152

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer-an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat Commun, 11*(1), 3952. doi:10.1038/s41467-020-17591-w

Mousavi, S. M., Sheng, Y., Zhu, W., & Beroza, G. C. (2019). STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI. *IEEE Access, 7*, 179464-179476. doi:10.1109/access.2019.2947848

Perol, T. G., Michael; Denolle, Marine. (2018). Convolutional neural network for earthquake detection and location. *Science Advanced*.

Rose, W. I., Palma, J. L., Escobar Wolf, R., & Matías Gomez, R. O. (2013). A 50 yr eruption of a basaltic composite cone: Pacaya, Guatemala. In *Understanding Open-Vent Volcanism and Related Hazards*.

Ross, Z. E., Meier, M. A., Hauksson, E., & Heaton, T. H. (2018). Generalized Seismic Phase Detection with Deep Learning. *Bulletin of the Seismological Society of America, 108*(5A), 2894-2901. doi:10.1785/0120180080

Schaefer, L. N., Oommen, T., Corazzato, C., Tibaldi, A., Escobar-Wolf, R., & Rose, W. I. (2013). An integrated field-numerical approach to assess slope stability hazards

at volcanoes: the example of Pacaya, Guatemala. *Bulletin of Volcanology, 75*(6). doi:10.1007/s00445-013-0720-7

Trnkoczy, A. (1999). *Understanding and parmeter setting of STA/LTA trigger algorithm*. Retrieved from

Xiong, W., Ji, X., Ma, Y., Wang, Y., AlBinHassan, N. M., Ali, M. N., & Luo, Y. (2018). Seismic fault detection with convolutional neural network. *Geophysics, 83*(5), O97-O103. doi:10.1190/geo2017-0666.1

Yoon, C., O'Reilly, O., Bergen, K., & Beroza, G. (2015). Earthquake detection through computationally efficient similarity search. *Science Advanced*, 13.

Zhu, G., Yang, H., Tan, Y. J., Jin, M., Li, X., & Yang, W. (2022). The cascading foreshock sequence of the Ms 6.4 Yangbi earthquake in Yunnan, China. *Earth and Planetary Science Letters, 591*. doi:10.1016/j.epsl.2022.117594

Zhu, G., Yang, H., Yang, T., & Zhang, G. (2022). Along-Strike Variation of Seismicity Near the Extinct Mid-Ocean Ridge Subducted Beneath the Manila Trench. *Seismological Research Letters*. doi:10.1785/0220220304