2021

Volcan de Fuego: A Machine Learning Approach in Understanding the Eruptive Cycles Using Precursory Tilt Signals

Kay Sivaraj
Michigan Technological University, sivaraj@mtu.edu

VOLCAN DE FUEGO: A MACHINE LEARNING APPROACH IN UNDERSTANDING THE ERUPTIVE CYCLES USING PRECURSORY TILT SIGNALS

By

Kavya Sivaraj

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Geology

MICHIGAN TECHNOLOGICAL UNIVERSITY

2021

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

Department of Geological & Mining Engineering & Sciences

Thesis Advisor:	*Dr. Gregory Waite*
Committee Member:	*Dr. Simon Carn*
Committee Member:	*Dr. Roohollah Askari*
Department Chair:	*Dr. Aleksey Smirnov*

Contents

List of Figures

List of Tables

Preface

This thesis is written in partial fulfilment of the requirements for the Master's degree in Geology at Michigan Technological University. This is an original work carried out under the guidance of Dr. Gregory Waite from the Department of Geological & Mining Engineering & Sciences between August 2020-21.

Machine Learning is increasingly gaining importance in the field of volcano seismology and this is one of the earliest works to incorporate this technique to understand the mechanism of Volcan de Fuego. As such, I believe this work will prove useful to researchers who wish to further develop ML methods to monitor the volcano.

Abstract

Volcan de Fuego is an active stratovolcano located in the Central Guatemalan segment of the 1100 m long Central America Volcanic Arc System (CAVAS). Fuego-Acatenango massif consists of at least four major vents of which the Fuego summit vent is the most active and the youngest member. The volcano exhibits primarily Strombolian and Vulcanian behavior along with occasional paroxysms and pyroclastic flows. Historically, Fuego has produced basaltic-andesitic rocks with more recent eruptions progressively trending towards maficity. Several studies have used short-term deployments of broadband seismometers, infrasound, and long-term remote sensing techniques to characterize the mechanism of Fuego. In our study, we analyze the tilt derived from transient broadband seismometers and tiltmeter stationed over several days during 2009, 2012, and 2015 near the summit crater using unsupervised learning.

Unsupervised learning has the potential to play a significant role in monitoring volcanoes dominated by large, unlabeled datasets. In our study, we make use of dynamic time warping distance measure along with unsupervised classification methods to identify precursory tilt signals. The unsupervised classification revealed two types of tilt signals with opposite polarity, one of which confirms features identified in previous studies while the other signal has been previously unknown. Template matching implemented with the known signal identified 268 events between October 1, 2015

and January 13, 2016, the duration of which varied between 7 and 39 minutes. The temporal distribution of these events as well as the maximum amplitude of inflation showed clustering activity accompanied by intra-cluster waxing and waning. We created subsets of temporal clusters and calculated repose times between successive events. Auto-correlation functions were calculated for each subset and probability density functions were fitted which support survival/failure processes. The long-term tilt records provided a useful tool to characterize the activity and revealed a near-continuous cyclicity.

Chapter 1

Introduction

Volcan de Fuego also known as Chi Q'aq' (in Kaqchikel - one of the Mayan languages) is one of the most active volcanoes in Guatemala, and is located approximately 44 km from the capital Guatemala city. It is conjoined with another active volcano, Acatenango, to the north and is aptly nicknamed a "twin volcano". Together with Acatenango, it forms the dynamic volcanic complex of La Horquita. Fuego volcano is characterized by Strombolian/Vulcanian eruptions, frequent ash explosions leading up to paroxysms and occasional pyroclastic flows [*Lyons et al.*, 2009].

Since 1524, Fuego volcano has produced more than 60 VEI 2, explosive eruptions and on June 3, 2018 produced its most deadliest eruption in a 100 years accompanied by several pyroclastic flows [BBC News, 2018]. The eruptive plumes were shot into

the atmosphere as high as 10 km and could be seen from the space, and the ash falls reached as far as 40 km from the volcano. When the eruptive plumes collapsed, the pyroclastic flows inundated several villages resulting in hundred of casualties and destroyed more than half of the houses in San Miguel Los Lotes [*Wallace*, 2018]. Therefore, continued monitoring of the volcano is becoming increasingly essential to thwart future disasters.

1.1 Geological Setting

Volcan de Fuego is an active stratovolcano located in the northern part of the Central America Volcanic Arc System (CAVAS). CAVAS is a 1100 m long volcanic arc, parallel to the Pacific coast, formed by oblique subduction of the Cocos plate under the over-riding Caribbean plate, converging in N30E direction [*Gazel et al.*, 2021]. The volcanic arc is divided into several blocks - Western Guatemala, Central Guatemala, Eastern Guatemala, El Salvador, Western Nicaragua, Eastern Nicaragua, and Costa Rica- by transverse boundaries running perpendicular to the Middle-America trench which lies approximately 150 km south of the volcanic chain [*Carr*, 1984]. Each block is 100-300 km long with different strikes and volcanic lineaments, and houses several active volcanoes that exhibit striking inter-volcano variations in morphology, geochemistry, eruptive styles and patterns. The boundaries between the blocks are marked by major normal faulting and offset in volcanic lineaments. The blocks tear

at shallow depths and enter into the mantle at different dips which may explain the offsets in the positions of volcanoes between different segments. The boundaries are characterized by relatively higher shallow earthquakes, small volume, and violent eruptions. Between the depths of 100 km and 150 km is a zone characterized by an absence of earthquakes which corresponds with the location of volcanic front and thereby the position of underthrust crustal melting [*Stoiber and Carr*, 1973].

The Central Guatemala segment is the most active of all the Central America Volcanic Arc segments and accomodates the five active volcanoes: Santa Maria, Atitlan-Toliman, Fuego-Acatenango, Agua, and Pacaya. This segment is 145 km long and has the third largest volume-length ratio of 2.7. This segment is characterized by the manifestation of transverse volcanic vents which results in the construction of paired volcanoes. Two of the paired volcanoes - Atitlan-Toliman and Fuego-Acatenango are located away from the transverse boundaries and account for about 60 percent of the total volcanic volume of the segment [*Stoiber and Carr*, 1973].

The Fuego-Acatenango pair has consisted of four active vents located north - south [*Halsor and Rose*, 1988]. The Acatenango complex consists of Yepocapa and Acatenango vents and the Fuego complex consists of Meseta and Fuego vents. The activity at Fuego has been more frequent with more than 60 eruptions consisting of Al_2O_3 rich, basaltic lava whereas Acatenango has been less frequent with only three historical eruptions in the last century, consisting of andesitic rocks [Vallance et al.,

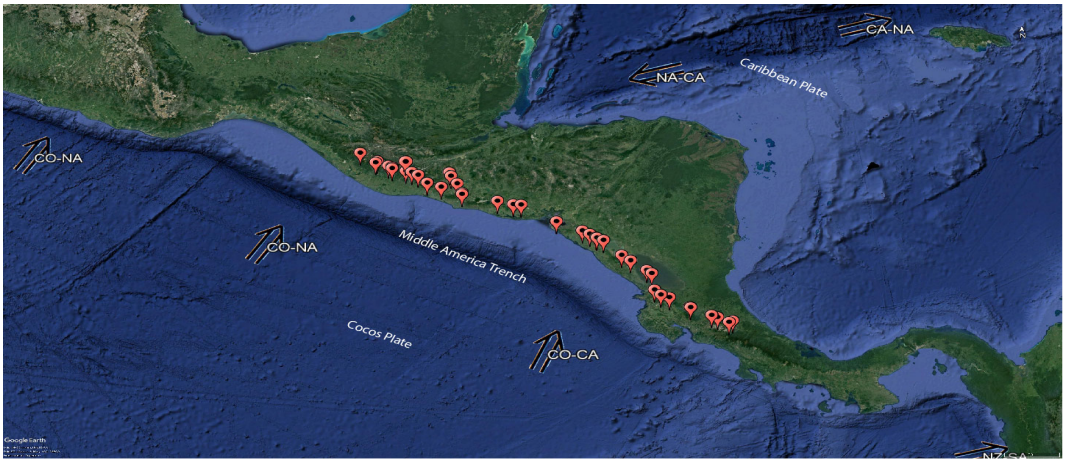

Figure 1.1: The figure shows the Central America Volcanic Arc System (CAVAS) created by the subduction of Cocos plate under Caribbean plate. The red symbols indicate the volcanic front extending from Guatemala in the west to Costa Rica in the east.

2001]. The major rock composition of Fuego ranges between 45-55 percent of silica whereas the composition of Acatenango is much more varied with 50-70 percent of silica [*Halsor and Rose*, 1988].

Figure 1.2: The figure shows the four major vents of La Horquita complex.

Chesner and Rose [1984] proposed that Fuego consists of two magma chambers -

a deep, primary magma chamber located between 8 km and 15 km at the base of the crust and shallow, dike-like magma chambers located between 2 km and 8 km below the surface. The primary magma chamber at the crustal base is shared by the two volcanic complexes. The magma undergoes initial fractionation at this chamber before starting its ascent towards the shallow magma chambers located beneath each of the volcanic complexes. The changes in chemical composition of the two complexes is attributed to the interaction between the primary magma and the country rock before it reaches the shallow magma chambers. The age estimates show that the Fuego complex is at least 17,000 years old with Fuego being the youngest vent and the older rocks of Fuego are more silicic compared to the younger rocks suggesting that the magma is trending towards mafic over time. This can be explained by the changes in magma chamber geometry caused by the destruction of Meseta cone. Although the major vents have been active coincidently, the activity seems to be shifting towards south. This is consistent with the other paired volcanoes in the region where the activity is younging towards the seaward volcanoes. This can be explained by a shift in the location of primary magma chamber towards the oceanic trench. As the magma chamber shifts, the seaward volcano, Fuego "steals" magma from the landward volcano, Acatenango thus diffusing its activity. This shift can be explained by a change in convergence rate and dip angle of the Cocos plate [*Halsor and Rose*, 1988].

1.2 Eruptive History

Fuego volcano has produced more than 60 VEI 2 eruptions and multiple VEI 4 eruptions (Fig: 1.3) ranging from Strombolian to sub-Plinian activity accompanied by occasional pyroclastic flows [*Lyons et al.*, 2009]. Historical observations of the volcano started after the Spanish settlement in 16th century. Prehistoric eruptive records (before 1524) have been largely obtained through stratigraphic studies and carbon-dating [*Naismith et al.*, 2019]. Throughout history, the volcano has undergone several eruptive cycles each lasting 20-70 years followed by periods of repose. Each cluster is characterized by low-level, background Strombolian eruptions, multiple ash-laden Vulcanian and sub-Plinian activity, and occasional pyroclastic flows. Several notable eruptions (VEI 4) have occurred in the years of 1581, 1717, 1737, 1857, 1880, 1932 and 1974 [*Martin and Rose*, 1981]. Although scoring relatively low (VEI 3) in the Volcanic Explosivity Index, the 2018 eruption is notorious due to its significant devastation [BBC News, 2018].

Hutchison et al. [2016] translated a Spanish document, "Autos Hechos Sobre el Lastimoso, Estrago y Ruina que Padecio esta Ciudad de Guatemala. . . " which catalogs first account statements from several eye-witnesses during the 1717 eruption of Fuego. The activity began on August 27 with small ash explosions that culminated in high plumes and pyroclastic flows. One interesting aspect of this eruptive cycle is the mudflows

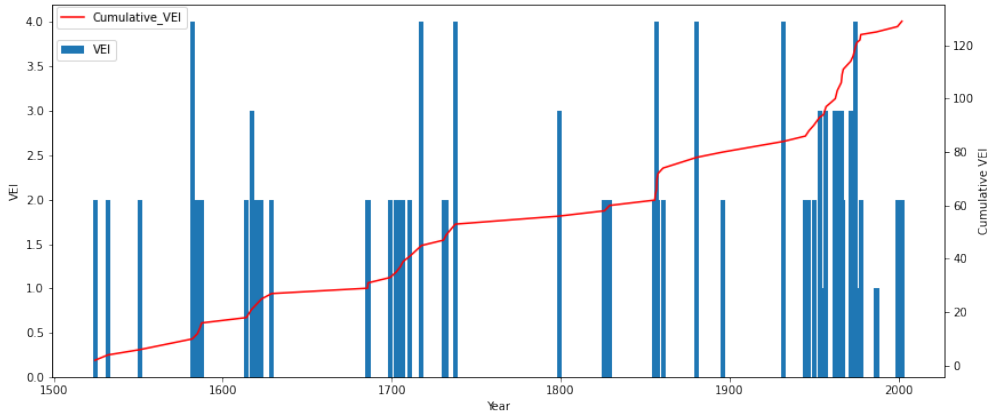

Figure 1.3: The figures show eruptive clusters consisting of VEI 2 or more eruptions recorded since 1524. The red line indicates cumulative VEI of the explosions. Data collected from Global Volcanism Program (GVP).

of Volcan de Agua preceded by an earthquake occurred in September, 1717. They hypothesized that it could have been influenced by Fuego's activity resulting in a collapse of Agua's hydrothermal system. *Martin and Rose* [1981] has given a detailed catalog of eruption intensities after 1932. Fuego erupted in 1932 after five decades of quiescence and continued its activity well through 1979 interrupted by occasional periods of silence within clusters. The 1974 eruption produced more ash (0.2 km^3) than the previous eruptions that occurred in 20th century. *Rose et al.* [1978] conducted a detailed study on the 1974 eruption using tephra deposits.The major activity occurred between October 10 and October 23 in four pulses each lasting less than 24 hours. They proposed a deep zone crystal fractionation possibly at depths greater than 5 km beneath the summit and suggested a shallow, dike-like conduit for the upward movement of magma influenced by tidal accelerations. After two decades of almost

no activity, the volcano entered its recent eruptive cycle in 1999 [*Lyons and Waite*, 2011; *Naismith et al.*, 2019]. This cycle is characterized by continuous, background lava effusion, intense explosive activity accompanied by ash fall, sustained plumes, and paroxysmal eruptions followed by lava-less degassing explosions. *Naismith et al.* [2019] used remote sensing data obtained from MODIS/MIROVA to demonstrate an increase in eruption intensities after late 2015 and proposed a new eruptive regime for the volcano. The most notable eruption in recent history occurred on June 3, 2018 producing multiple pyroclastic flows that inundated the Las Lajas ravine causing severe damage to the communities of San Miguel Los Lotes and El Rodeo [*Wallace*, 2018]. INSIVUMEH reported at least 110 casualties and a significant damage to the infrastructure [GVP, 2018]; and the unofficial death toll may exceed 2000 fatalities [ctv news, 2018].

1.3 VLP Seismicity

Very Long Period (VLP) seismic signals are common occurences in volcanoes throughout the world. These signals are generally repetitive and occur between 0.5 Hz and 0.01 Hz. These signals can be caused by agitations due to the movement of magma or volatile towards the surface [*Chouet and Matoza*, 2013]. As such, these signals give a glimpse into conduit geometries of volcanoes. *Lyons and Waite* [2011] used VLP signals recorded from Fuego volcano that accompanied ash-rich explosions. They

attributed the VLP to a decompression of the upper conduit following rupture of a pressurized seal near the top of the conduit. Using full-waveform moment-tensor inversion, they modelled the conduit geometry as a near vertical dike connected to a shallower dipping sill. The events that produce VLP can also produce measurable tilt below the low corner of instruments [*Lyons et al.*, 2012] . Hence, a catalog of VLP signals can provide a starting point to identify volcanic tilt events.

1.4 Unsupervised Learning

Volcano seismology is a data-driven field that relies on the ability to model and visualize interior of the Earth through seismic signals. With the deployment of broadband sensors all around the world, it is becoming increasingly possible to record these signals over a broad range of frequencies. Consequently, the plethora of information coming from these instruments is becoming ever so difficult to process by human expertise alone. To overcome this challenge, seismologists are actively co-opting and adapting machine learning techniques initially used in robotics and natural language processing [*Fiorini et al.*, 2020; *Nasukawa and Yi*, 2003; *Zagibalov and Carroll*, 2008].

Supervised and unsupervised algorithms are subsets of machine learning techniques that can be used to classify waveforms and identify new patterns respectively. Supervised classification uses pre-labeled datasets as input to search and classify similar

waveforms into their respective groups. This class of algorithms is more suitable when datasets that have been previously processed and labeled by human experts are available, whereas unsupervised classification does not require any pre-labeled datasets and can be used in places where such information is readily unavailable. Therefore, unsupervised learning is widely used to identify precursory patterns in volcanic eruptions [*Seydoux et al.*, 2020].

Seydoux et al. [2020] used deep scattering network in conjunction with Gaussian mixture model to cluster seismic data recorded over a period of 24 hours before the 2017 Nugaatsiaq landslide. Their cluster analysis revealed a repeating, precursory signal that started 9 hours before the landslide. *Witsil and Johnson* [2020] used unsupervised clustering techniques to analyze infrasound data recorded over a period of three daya at Stromboli volcano. Their analysis suggested an existence of common plumbing system shared by the six active vents. Ren et al. [2020] studied seismic signals recored at Piton de la Fournaise volcano over a long period of 6 years to identify signals associated with different eruptive behavior. *Anzieta et al.* [2019] used seismic records obtained over a period of nine months from Cotapaxi volcano. They employed spectral analysis along with distance measures to identify a seismic pattern associated with the upward movement of magma.

Following these studies, we have employed unsupervised clustering techniques on seismically-derived tilt signals recorded in the east component of stations F900 and

NW1 to identify precursors in an effort to better understand the mechanism of Fuego. The station F900 was operating in 2009 whereas the station NW1 was operational during 2012 and 2015.

Chapter 2

Data Collection and Methods

2.1 Data

The data used in this study were collected from broadband seismometers - located at F900 and NW1 - and a tiltmeter, installed towards north of the summit vent (Fig: 2.2). The station F900 housed a Guralp CMG40T-30 sec seismometer which was operating from January 10, 2009 to January 26, 2009. This station was recording data at the rate of 100 samples per second or 100 Hz. The station NW1 contained a Nanometric Trillium-120 sec seismometer and was functional from January 16, 2012 to February 02, 2012. This station was sampling the data at 100 Hz. The station NW1 stationed a Guralp CMG3ESP-60sec seismometer was operating from February

12, 2015 to February 26, 2015 and was sampling at a rate of 50 samples per second or 50 Hz. The tiltmeter was operational over a period of 105 days from October 1, 2015 to January 13, 2016 and was sampling at a rate of 1 sample per minute. (Fig: 2.1).

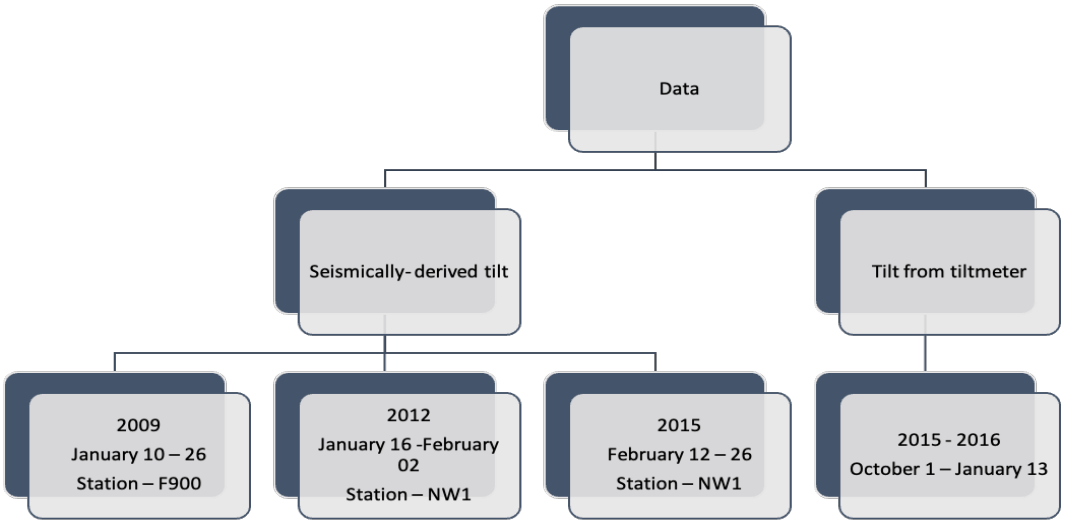

Figure 2.1: The figure shows a schematic representation of data as well as the operating time period of stations used in this study. The tiltmeter was operating over a period of 105 days providing a relatively longer-term record.

2.2 Methods

The analysis consists of six steps: (1) identifying VLP events from the seismic traces using STA/LTA algorithm; (2) computing tilt values associated with the VLP events; (3) creating a feature matrix using Dynamic Time Warping (DTW) distance measure; (4) applying clustering algorithms to identify meaningful groups; (5) identifying

Figure 2.2: The figure shows the top view of Fuego volcano and the locations of stations F900 and NW1. F900 was operating during 2009 and NW1 was operational during 2012 and 2015. The tiltmeter was collocated with the station F900 and was active during 2015 and 2016.

similar events in tiltmeter recorded tilt data using the meaningful groups for template matching; and (6) statistical analysis on events obtained through template matching.

2.2.1 STA/LTA Algorithm

Automatic trigger algorithms are gaining importance particularly when the situation demands dealing with large amounts of datasets. These algorithms range from simple triggering mechanism such as detecting amplitudes over a user-specified threshold to more advanced neural networks. The latter is widely used in studies that require high sensitivity and low false triggers, and is usually specific to waveforms and study area.

Therefore, the simple triggering algorithms are still widely used for their simplicity and overall satisfactory performance.

These algorithms include simple amplitude threshold trigger which detects for amplitude above a user-defined threshold, Root-Mean-Square (RMS) trigger which stores events that are above a particular RMS value, and the Short Term Average (STA)/ Long Term Average (LTA) detection algorithm which detects events based on two moving windows.

The STA/LTA trigger is suitable for both weak-motion and strong-motion seismology, and is not sensitive to natural seimic noise such as those produced by ocean waves as well as human-made seismic noise that are continuous. Hence, it is preferred over the other two trigger detection algorithms [*Trnkoczy*, 1970].

The STA/LTA detection algorithm keeps track of all the absolute amplitudes in two moving windows and takes average over these values. The ratio of these two moving windows is then calculated, and if the value exceeds a pre-defined threshold the trigger mechanism is activated. This algorithm requires four inputs: (1) the length of the STA window; (2) the length of the LTA window; (3) the upper threshold limit; and (4) the lower threshold limit. Care should be taken to find a leverage between good sensitivity and low false triggers.The trigger inputs are set following the procedure outlined by *Trnkoczy* [1970].

The STA gives the measure of instant amplitude within the specified window. The trigger sensitivity to local earthquakes can be increased by decreasing the window duration and to teleseismic earthquakes by increasing the duration parameter. In general, the STA window should be a little longer than the minimum period of potential earthquakes. Otherwise, the trigger will be influenced by individual periods rather than the whole signal. At the same time, it should be shorter than the minimum duration of individual events so as to capture most of the events. The LTA keeps track of amplitude of seismic noise. As such, it should be a little longer than the expected noise variations. Otherwise, the LTA window will adjust itself to the instant earthquakes thus decreasing the trigger sensitivity. At the same time, using too long a window may potentially increase false triggers.

2.2.2 Calculation of Tilt

The horizontal components of seismometers consist of a long boom whose one end is fixed while the other supports a mass. The length of the boom is long enough for its motion to be considered as translation. But, several studies have showed that the rotational effects produced by these sensors might not be as insignificant as previously thought [Rogers, 1968; *Pillet and Virieux*, 2007]. Rogers [1968] showed the response of horizontal seismometers to surface waves and demonstrated that long-period Rayleigh waves had significant tilting effects. *Pillet and Virieux* [2007] compared

the response of horizontal and vertical seismometers, and showed that the former had considerable tilt effects while the latter exhibited negligible influence. This tilting effect pollutes the ideal response of seismometers and is undesirable in seismic studies. Recent studies have taken advantage of this "noise" below corner frequency to characterize the mechanism of magmatic processes.

Aoyama and Oshima [2008] conducted an experimental investigation on tilt records and demonstrated similarity between the tilt measured from an actual tiltmeter and derived values from a broadband seismometer. *Genco and Ripepe* [2010] used derived tilt records of Stromboli volcano and identified inflation-deflation signals associated with continuous magma discharge and replenishment. *Ripepe et al.* [2009] used tilt records obtained from Stromboli volcano and identified a relationship between tilt amplitude and explosion intensity. They developed an early warning system based on their analysis which forecast a paroxysmal eruption and a tsunami successfully. *Lyons et al.* [2012] used an array of broadband seismometers installed on Fuego volcano as tiltmeters and modeled a pressure source with spherical geometry, located towards 120 m west and 40 m deep below the summit crater.

This study follows the procedure outlined by *Lyons et al.* [2012] to calculate tilt output from broadband seismometers: (1) the mean of the seimic trace is removed; (2) the demeaned trace is integrated with respect to time; (3) the trace is passed through a low-pass filter to remove the frequencies above the low corner of the instrument;

and (4) the trace is multiplied by the ratio of product of instrument sensitivity and corner frequency to the gravitational acceleration.

2.2.3 Dynamic Time Warping

Dynamic Time Warping (DTW) is a distance measure that computes similarity between two or more waveforms (Fig: 2.3). This method gained importance in the speech processing community in the 1960s. DTW minimizes the distance between points of two time series in a non-linear fashion and computes a distance matrix with quadratic complexity. This property reduces its susceptibility to pessimistic similarity or time axis distortion unlike Euclidean Distance (ED) measure and allows for comparing time series of similar shapes but different phases [*Ratanamahatana and Keogh*, 2004]. In the last decade, DTW has gained popularity in various fields such as data mining, chemical engineering, medical imaging, and geophysics. But due to its computational complexity, the application of DTW is limited to small to medium sized databases. To overcome this problem, several studies have introduced lower bounding methods to limit the warping path such as Sakoe Chiba band, Itakura parallelogram, and LB Keogh [*Ratanamahatana and Keogh*, 2004; *Salvador and Chan*, 2007]. These warping windows not only reduce the computational time but also limit pathological warping by constraining the number of samples to be mapped for each distance calculation [*Salvador and Chan*, 2007].

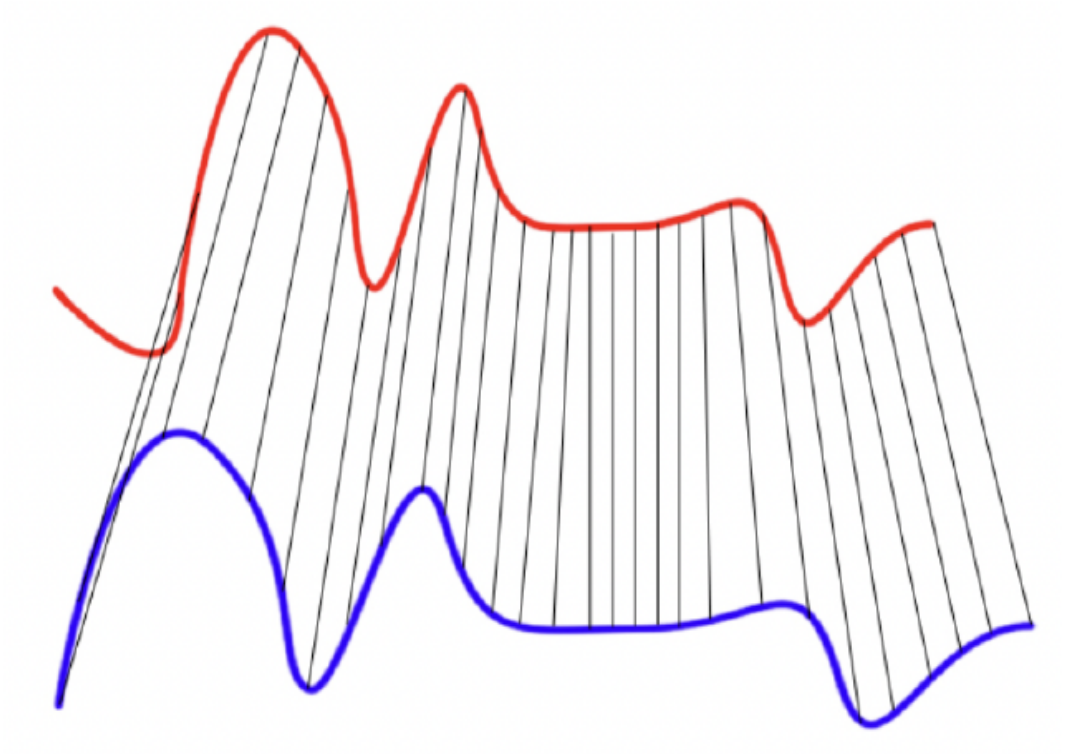

Figure 2.3: The DTW algorithm minimizes the distance between each point on time series 1 (blue) with multiple points on time series 2 (red) thus accounting for phase-shifts in time series data. Source: Wikimedia Commons: Euclidean vs DTW.jpg

2.2.4 Volcanogenic Radiant Flux

Volcanogenic Radiant Flux (VRF) measures the thermal radiance emitted by lava flows and therefore it is directly proportional to volcanic lava output. VRF is calculated and made available within a couple of hours of satellite over-passing by MODVOLC using MODIS instrument. Moderate Imaging Spectroradiometer (MODIS) consists of two sensors located each on National Aeronautical Space Administration's (NASA) Terra and Aqua satellites. These satellites together orbit the Earth every

two days and record information in 36 spectral bands. MODVOLC algorithm utilizes thermal infrared radiation (3.959 micrometer wavelength) captured by MODIS sensors to monitor thermal anomalies produced by volcanic eruptions. This algorithm calculates radiant flux for each hotspot by using equation 2.1 [*Wooster*, 2003].

$$\phi_e = 1.87 * 10^7 * (L_{3.959\ \mu\mathrm{m}} - L_{3.959\ \mu\mathrm{m,\ bg}}) \tag{2.1}$$

$L_{3.959\ \mu m}$ = spectral radiance of each pixel; $L_{3.959\ \mu m,\ bg}$ = spectral radiance of adjacent pixels.

The radiant flux for each pixel identified at a particular time is summed up to calculate the Volcanogenic Radiant Flux (VRF) for each observation time. Only the night time images were used to avoid errors due to solar radiation [*Wright et al.*, 2015].

2.2.5 Statistical Analysis

Statistical analyses can be used to identify underlying trends and patterns in particularly large datasets. The datasets can be divided into two types based on trends: stationary and non-stationary datasets. The type of dataset to be used is highly subjective and largely depends on the nature of data. In our study, we have used

a non-stationary dataset which contains no trends but periodicity. Due to paroxysmal nature of Fuego volcano, we believe the periodicity of eruptions constitute an important factor.

There are several methods to quantify randomness and clustering within datasets. Coefficient of Variation (CoV) is a simple but effective tool which calculates the ratio of standard deviation and mean of the data. A value of 1 indicates clustering whereas a value less than 1 signifies Poissonian process [*Bebbington and Lai*, 1996]. Similarly Auto-Correlation Function (ACF) is another statistical method commonly used in forecasting to identify suitable time series models. One of the advantages of ACF is that it can reveal periodicity within the time series data. It computes correlation between each time series and time lagged version of itself called lags. The ACF is usually represented using a plot between lags and ACF coefficient called correlogram. The significance of each correlation is determined using a confidence interval of 95 percent. If the correlation value exceeds the confidence interval, it is considered to be significantly larger than the mean value, otherwise the coefficient is considered insignificant. If the deviation of ACF coefficient shows gradual variation, it is indicative of periodicity or cyclicity in the data and an auto-regressive model (AR) can be used in forecasting. On the other hand, if the coefficient has abrupt cut offs around the mean, it is indicative of randomness and a moving-average (MA) method is best suited for the forecasting [*Varley et al.*, 2006]. The Probability Distribution Function (PDF) is commonly used to identify best fits for distribution of events within a

dataset. Depending on the best fit, it might be possible to identify if the data follow stochastic/renewal distributions or survival/failure models. The renewal process assumes that the events are stationary, time independent, and the probability of future eruption is dependent only on the time lapsed since the previous eruption (example: exponential or Poissonian distributions). Therefore, it is suitable for volcanoes where the magmatic system undergoes recharge between each eruption and no periodicity exists. There are other set of distributions called survival/failure models which incorporates the periodicity/cyclicity and these include gamma, log-logistic, lognormal, and Weibull distributions. The log-logistic distribution suggests that the data is dominated by competing sources while the lognormal and Weibull distributions suggest that the data exhibits quasi-periodic distribution [*Bebbington and Lai*, 1996]. In this study, we have used Poissonian distributions such as Gaussian and exponential as well as survival distributions to identify better fits for maximum amplitudes and repose times. The goodness of the fits is determined using Kolmogorov-Smirnov test with a 95 percent confidence. The KS test rejects or accepts the hypothesis that a given set of samples belongs to a particular distribution using a D-value and a P-value [Bebbington, 2013]. If the calculated D-value is less than the critical D-value (depends on the number of samples) then the null hypothesis is accepted. The P-value is calculated from the observed D-statistic and gives the probability that the critical D-value is larger than the observed. As such, for a good fit, the P-value should be as greater as possible than the significance level.

Chapter 3

Results

3.1 Tilt from Seismometers

The seismic records were collected over a period of several days during each year under investigation (see Section: 2.1) (Note: not all stations were operating at the same time). A previous study on Fuego volcano has showed that the tilt signals attenuate more rapidly with distance and the magnitude is particularly stronger along east channel [*Lyons et al.*, 2012]. Therefore, only the east channel from stations closest to the summit vent was taken for analysis.

The raw data were passed through a bandpass filter to limit the frequencies to VLP bands of 10 s to 60 s. A STA/LTA trigger mechanism was applied to the filtered

data. The trigger ratio was limited to 10/350 while the amplitude range was limited between 10 and 0.7 in order to capture maximum events with limited false triggers.

A standard 45 minute raw subset preceding each VLP onset time recorded in the east component was taken for further analysis. The raw data were passed through a lowpass filter to limit the frequencies higher than low corner of each instruments - data from the stations NE1, and NW1 were passed through a 60 s lowpass filter whereas that of station F900 was passed through a 30 s filter. Tilt records were calculated by multiplying traces of each subset with the respective instrument sensitivity following equation 3.1.

$$\tau = -S * \omega_0^2/g \int p(t)dt \tag{3.1}$$

where τ is the tilt; S is the seismometer sensitivity; Ω_0 is the natural or corner frequency of the instrument; g is the gravitational acceleration, and p(t) is the seismic trace.

Unsupervised classification yields better results when the dimensionality of datasets is reduced. This can be achieved by extracting suitable features such as distance matrices from data. This study utilized the DTW distance measure for creating a distance matrix between each time series. This procedure has a quadratic complexity

of A(n^2) and therefore parallel processing in C was adopted to speed up the process and a window constraint was applied to prevent pathological warping [*Ratanamahatana and Keogh*, 2004]. Finally, the calculated distance matrix was analyzed for groups using a bottom-top approach of hierarchical clustering which assumes each time series as initially independent and constantly groups the events until a predefined distance threshold is reached.

The choice of distance threshold depends on the nature of data and various methods are employed to select an appropriate value. In our analysis, we used cross-validation, dendogram, and trial & error approach [*Kumarsagar and Sharma*, 2014] to determine appropriate distance threshold and found that trial & error method yielded better results. This procedure yielded 5, 7, and 7 groups with distance thresholds of 1000, 600, and 1500 in data collected from 2009, 2012, and 2015 respectively. Further analysis of all these groups revealed two meaningful patterns with opposite polarity in each year (Fig: 3.4). The pattern 1 formed a well-defined, meaningful group that can be easily distinguished from other groups, while the pattern 2 was hidden among noisy time series (Fig: 3.1). The tilt is negligible on the vertical component of seismometers. Therefore, the vertical component of pattern 2 events were analyzed to rule out possible instrument errors and also the seismic traces of pattern 2 events were analyzed in higher frequency bands between 0.5 Hz and 10 Hz.

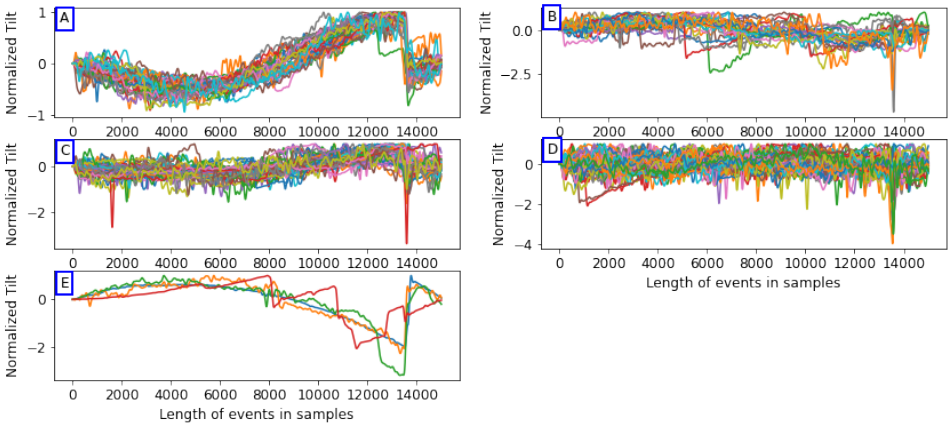

Figure 3.1: The figure shows the groups obtained through unsupervised learning on 2009 data recorded in the east component of the F900 seismic station. Figure A consists of 40 signals and shows well-defined pattern 1 while the figure E consists of only 4 signals and shows pattern 2 mixed with noise. The figures B, C, and D are mostly comprised of noise with 22, 39, and 53 signals.

3.2 Tilt from Tiltmeter

The tiltmeter was operating over a period of 105 days from October 1, 2015 to January 13, 2016 continuously and was sampling at a rate of one sample per minute. Although this sampling rate was very low, it was adequate enough to record events of longer duration. The tilt signals recorded in Fuego gradually vary over a period of several minutes to less than an hour. Therefore, an STA/LTA window would slowly adjust itself to the tilt signals thereby missing positive triggers. Hence, template-matching was chosen over STA/LTA to analyze tilt obtained from tiltmeter. The tiltmeter data was searched manually for an event that looked similar to the patterns

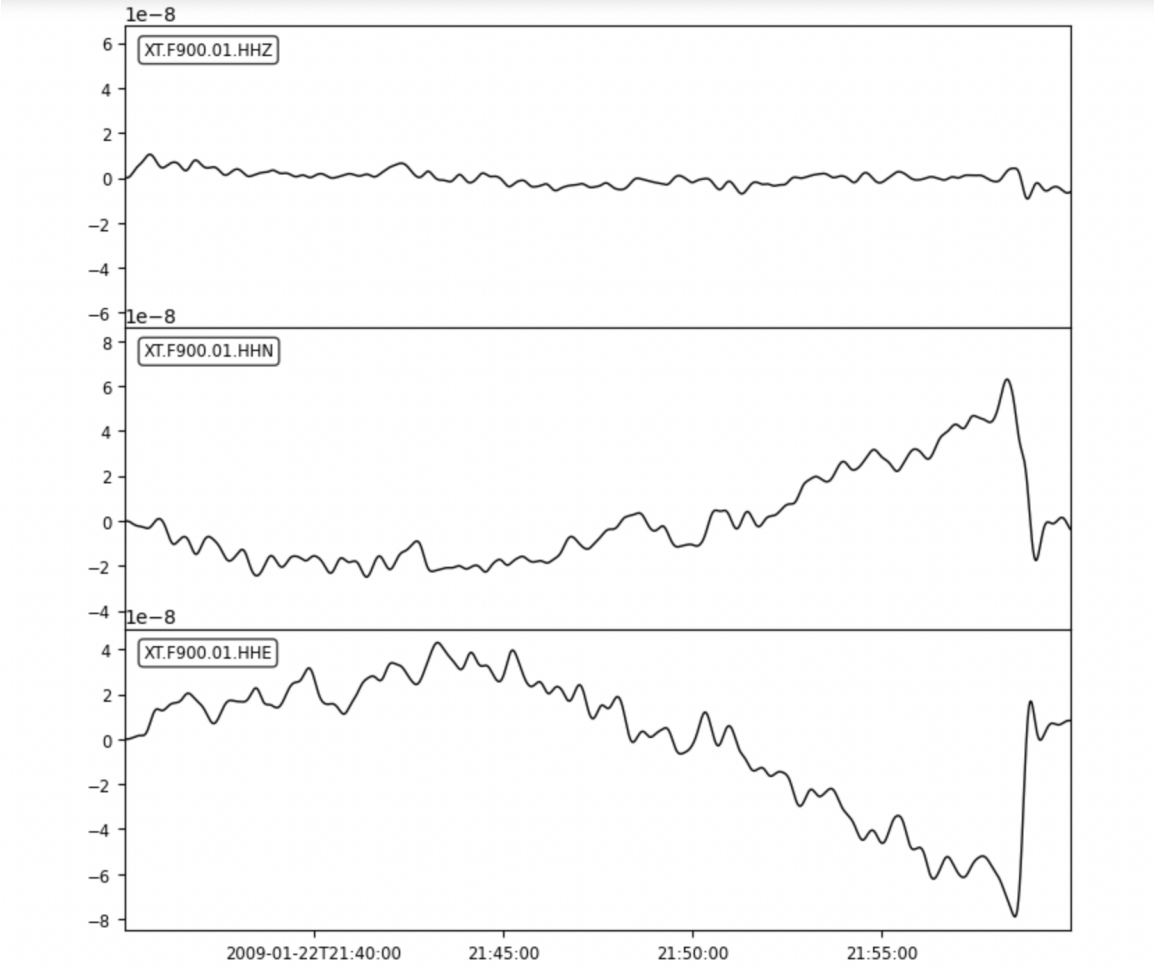

Figure 3.2: The figure shows three components (vertical, north, and east) containing tilt traces of a pattern 2 event. Instrument errors would affect all the components equally whereas the tilt signals only affect horizontal components. The very low amplitude signal in the vertical component confirms that the pattern 2 events were indeed produced by tilt.

obtained through unsupervised learning. An event that resembled pattern 1 was used as a template in the template-matching technique and we were unable to identify a template that resembled pattern 2. The meaningful patterns obtained from seismic derived tilt were not directly used as templates due to very low sampling rate of the

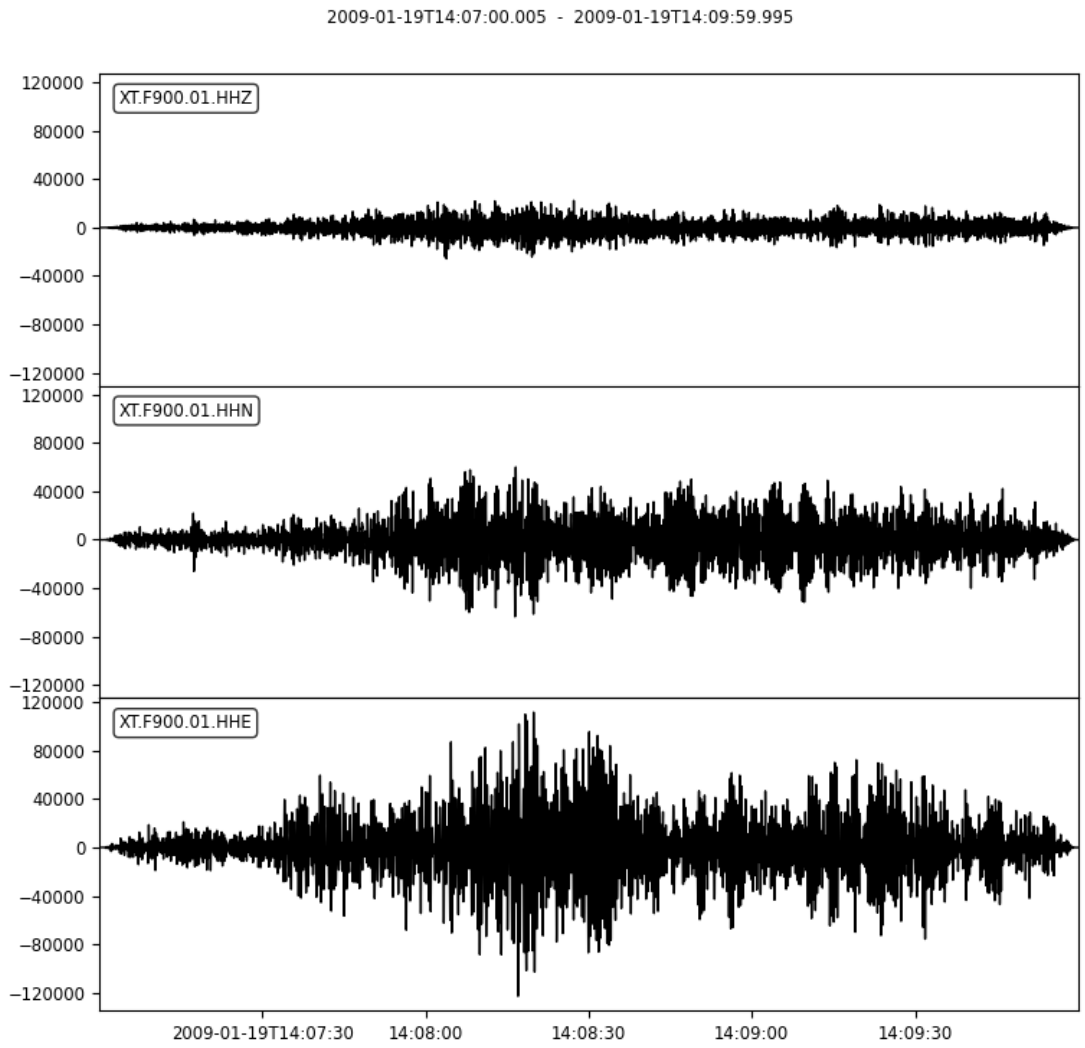

Figure 3.3: The figure shows examples of seismic traces associated with events of pattern 2 in higher frequency bands between 0.5 Hz and 10 Hz. All the three components (vertical, east, and north) show tremor signals.

tiltmeter. The template-matching technique was carried out to identify similar looking events with 75 percent similarity or above and this procedure yielded 268 events of high signal-to-noise ratio. A catalog was created containing onset time, duration, amplitude of inflation, amplitude of deflation, and repose interval between the events. (Appendix: A).

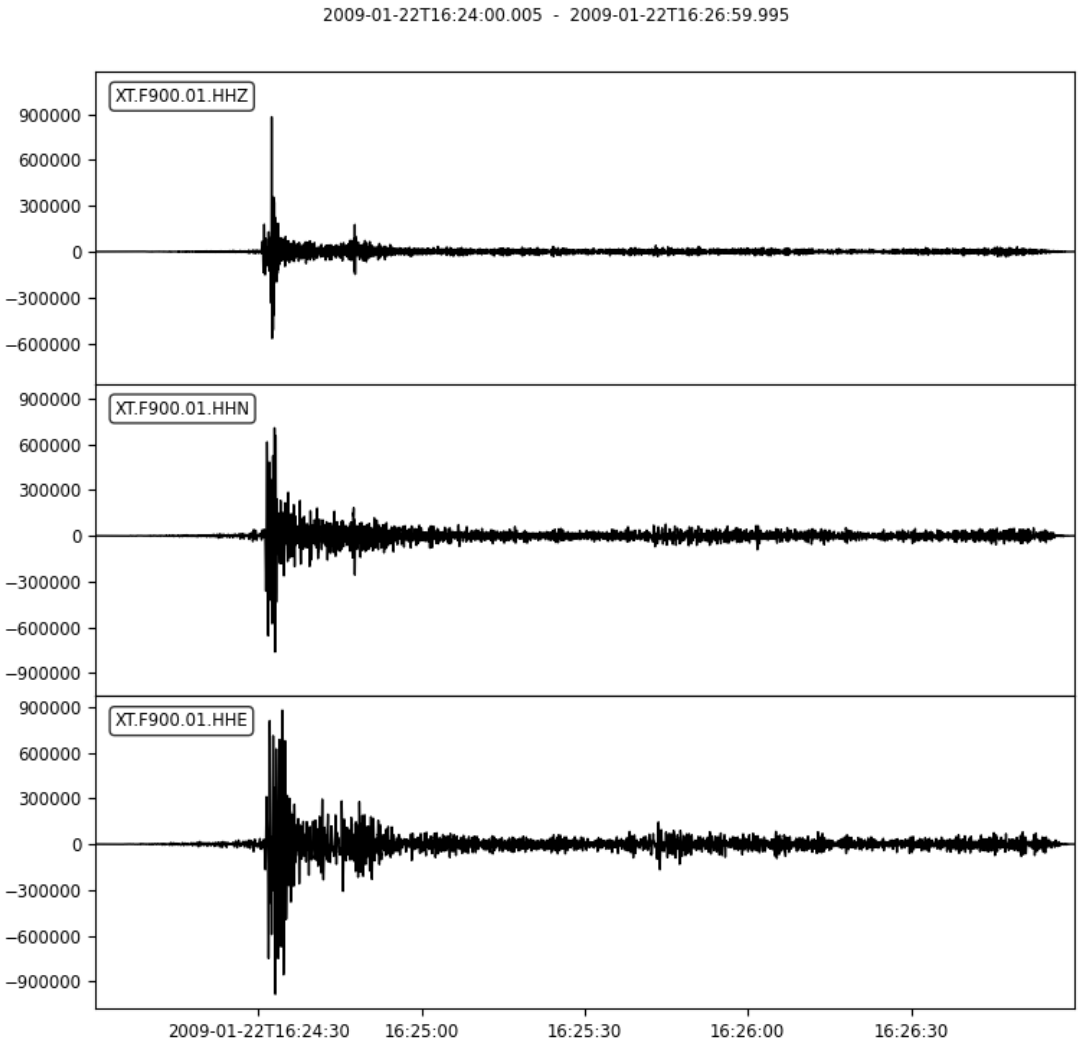

Figure 3.4: The figure shows examples of seismic traces associated with events of pattern 1 in higher frequency bands between 0.5 Hz and 10 Hz. All the three components (vertical, east, and north) show impulsive signals associated with explosions accompanied by long codas.

Statistical analysis was carried out on the data obtained from the catalog. A preliminary analysis of cumulative number of events (Fig: 3.6) revealed change in gradients suggesting a clustering of events in time. This hypothesis was further corroborated by a comparison between event onset time and number of events per day which revealed

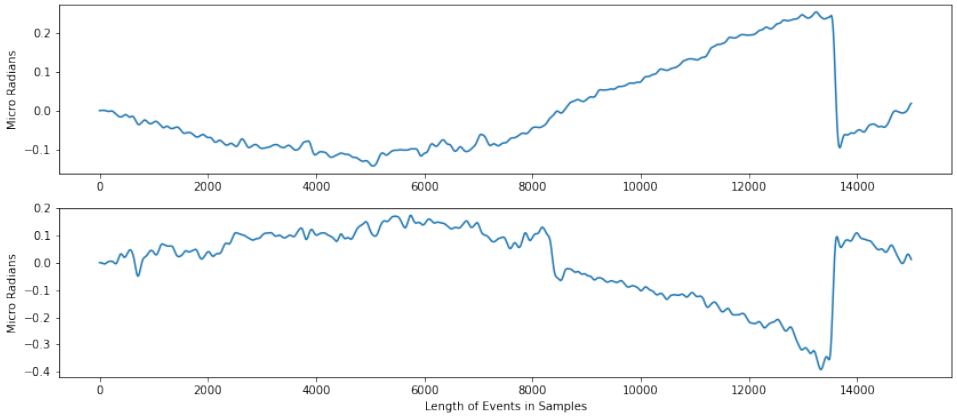

Figure 3.5: Examples of meaningful patterns - pattern 1 (top) and pattern 2 (bottom) recorded in the east component of the station F900 during 2009 - with opposite polarities obtained through unsupervised learning carried out on seismic derived tilt data.

seven individual clusters each lasting approximately 10 - 15 days (Fig: 3.7). A comparison with radiative power output obtained from MODVOLC revealed that these clusters always preceded high VRF which is directly proportional to lava flows (see subsection 2.2.4). Three of the seven clusters that occurred before October 28, 2015 were disregarded due to very low sampling size and only the remaining four clusters were taken for further analysis (hereafter referred to as clusters 1 through 4).

An analysis on the duration of inflation revealed that the events varied in length with a range of 32 minutes - the shortest event lasted up to 7 minutes while longest persisted for 39 minutes (Fig: 3.8). In cluster 2, the events with longer duration were distributed in the middle of the cluster while those with shorter duration were distributed towards the beginning and end of the cluster. In all other clusters the

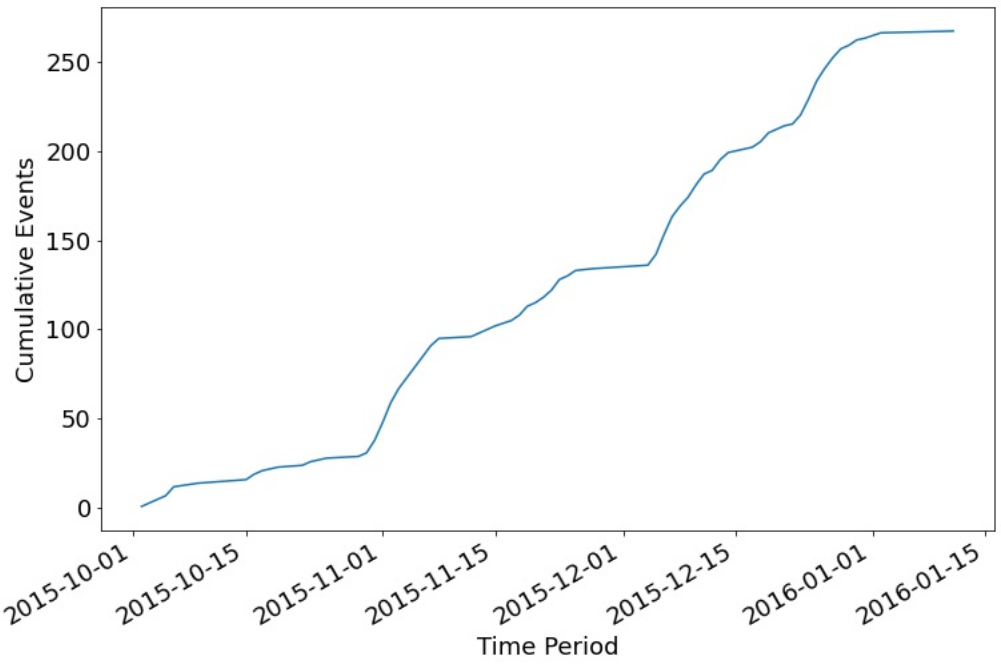

Figure 3.6: Cumulative number of tilt events recorded in tiltmeter, obtained through template matching. The varying gradients suggest a change in eruption frequency and coincide with eruptive clusters 1 to 4.

duration of events exhibited a random distribution.

The maximum amplitude of inflation within each clusters followed a waxing and waning pattern similar to the number of events per day (Fig: 3.9). The maximum amplitude increased gradually in the initial 3 - 5 days before culminating in the next 4 - 6 days and finally waning off. The Auto-Correlation Function (ACF) revealed a gradual change in ACF coefficient calculated for each lag. The gradual change suggests that the events exhibit clustering or periodicity. More than 5 percent of lags exceed the 95 percent confidence interval in the clusters 1, 3, and 4. Therefore,

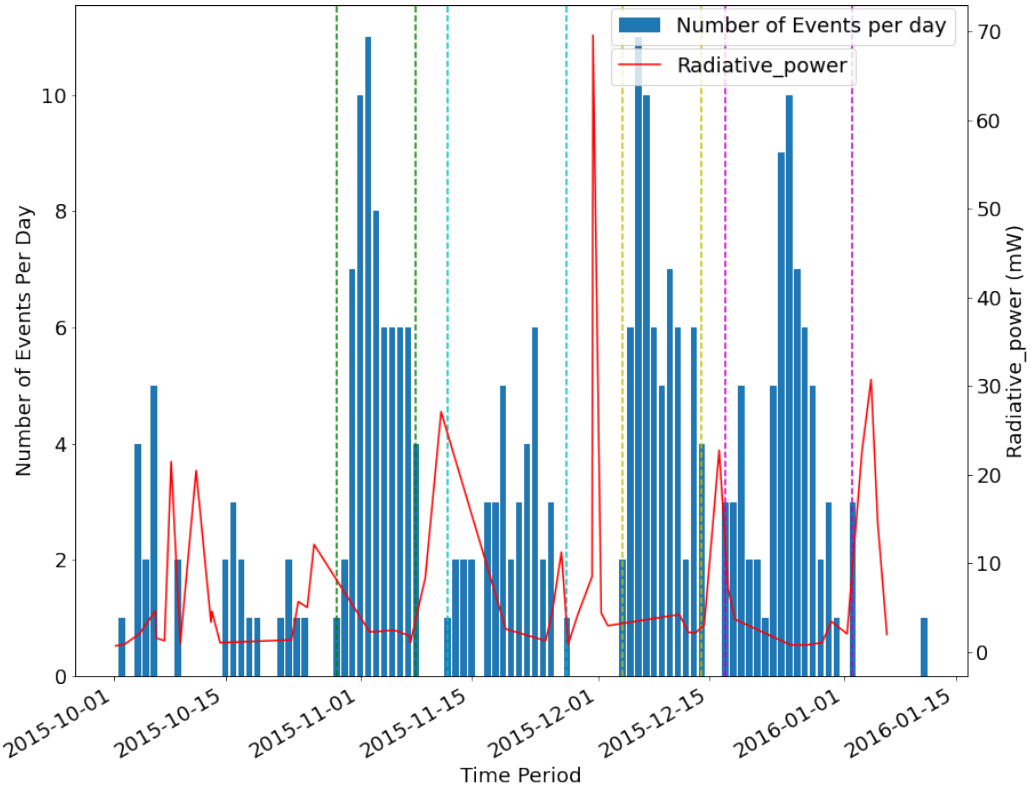

Figure 3.7: The figure shows number of tilt events recorded per day from October 1, 2015 to January 13, 2016. The red lines indicate Volcanic Radiative Flux (VRF) measured from MODIS night time images (Source: MODVOLC). The clusters 1 through 4 are represented by green, cyan, yellow, and magenta dashed lines respectively.

higher order auto-regressive models can be used to forecast the time-series. Higher order of the model suggests that the value of a future event depends on more than one event that immediately precedes it, therefore suggesting a clustering process (Fig: 3.10). The cluster 4 showed the highest correlation and periodicity while the cluster 2 exhibited the least correlation and periodicity, suggesting that the former was

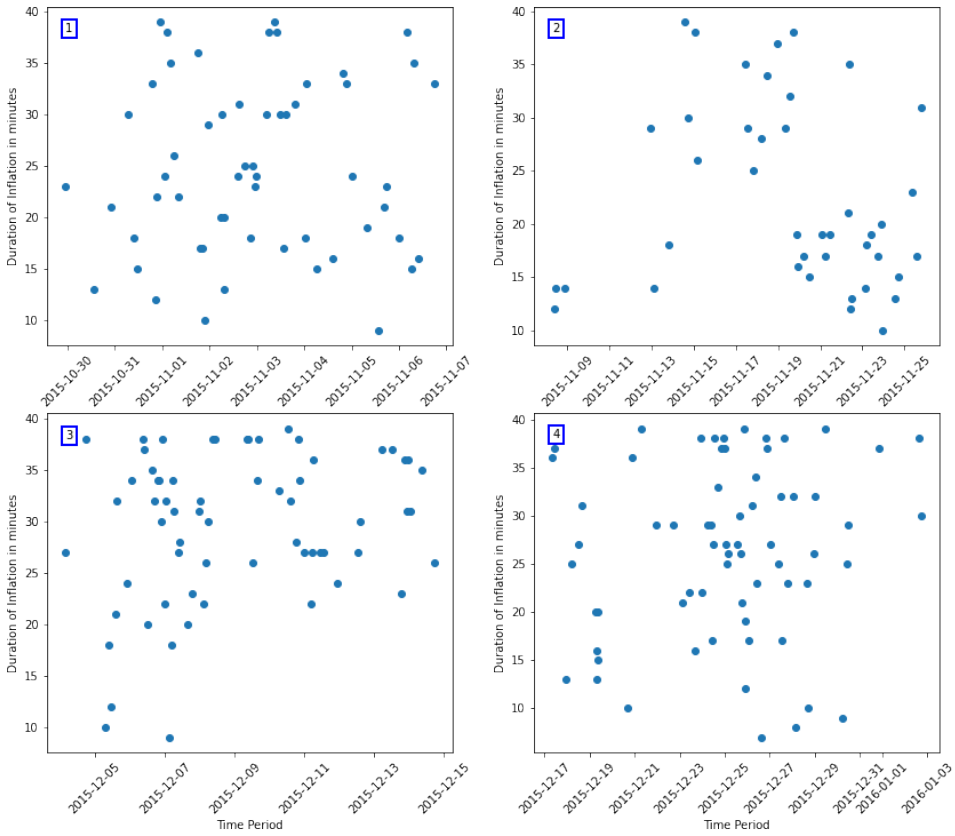

Figure 3.8: The figure shows the cluster of events obtained between October 1, 2015 to January 13, 2016, and the inflation duration of tilt events in each cluster. The duration is distributed randomly in all the clusters except cluster 2 in which the longer duration events are distributed in the middle while the events with shorter duration are spread towards the beginning and end of the cluster.

much more clustered than the latter. The Probability Density Function (PDF) was calculated for each cluster which showed that the clusters 3 and 4 exhibited gamma distribution while the clusters 1 and 2 displayed a normal distribution (Fig: 3.11). The gamma distribution belongs to the group of distributions called survival/failure

distributions which may suggest a periodicity or cyclicity within a dataset. Although first two clusters exhibit a normal distribution, we believe it might be caused by incompleteness of the dataset as both the ACF (Fig: 3.10) and time vs amplitude plot (Fig: 3.9) indicate clustering activity. The significant value of PDF is 0.05 for a 95 percent confidence interval. The P-value of distributions should be greater than this significant value in order to be a good fit. The P-values of all the distributions exceed 0.05, therefore all the fitted distributions passed the goodness of fit test. (Table: 3.1)

.

Clusters	Best Fit	Parameters			P value - KST
		Shape	Location	Scale	
Cluster 1	Norm		0.3695	0.171	0.46
Cluster 2	Norm		0.457	0.198	0.67
Cluster 3	Gamma	3.004	0.067	0.117	0.87
Cluster 4	Gamma	3.902	-0.028	0.083	0.51

Table 3.1
Parameters of Maximum Amplitude PDF

The repose times were calculated for events within each clusters. The repose times for maximum amplitudes are randomly distributed for all the clusters except cluster 2 in which the events with maximum amplitude show lesser repose times (Fig: 3.12). The ACF values calculated for repose times of each cluster hovered near zero with

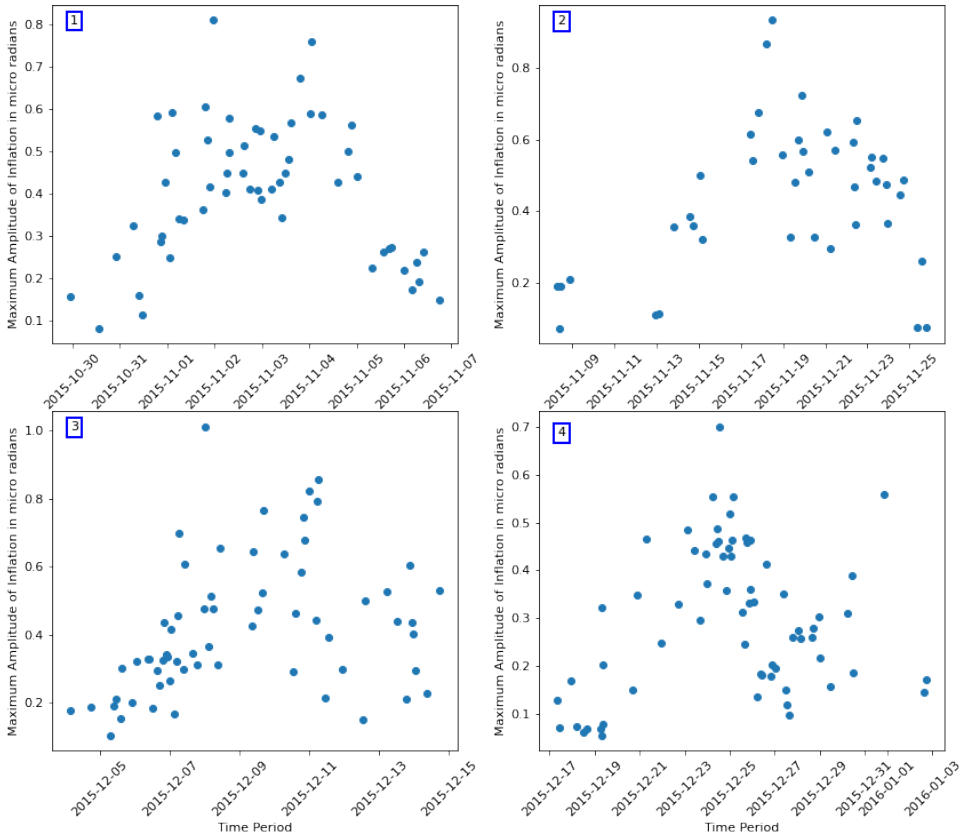

Figure 3.9: The figure shows the maximum amplitude of events in each cluster. The clusters exhibit a waxing and waning pattern where the amplitude starts to increase after each paroxysm and culminate before waning off to the original intensity. This is positively correlated with the number of events per day.

abrupt cut offs, suggesting very little in-clustering activity. Therefore, a moving-average (MA) model will better suit to forecast the future eruptions. This suggests that the probability of a future eruption cannot be determined only using repose times that immediately precede it (Fig: 3.14). The PDF calculated for clusters 1 and 3 exhibited exponential and gamma distributions while that of clusters 2 and 4

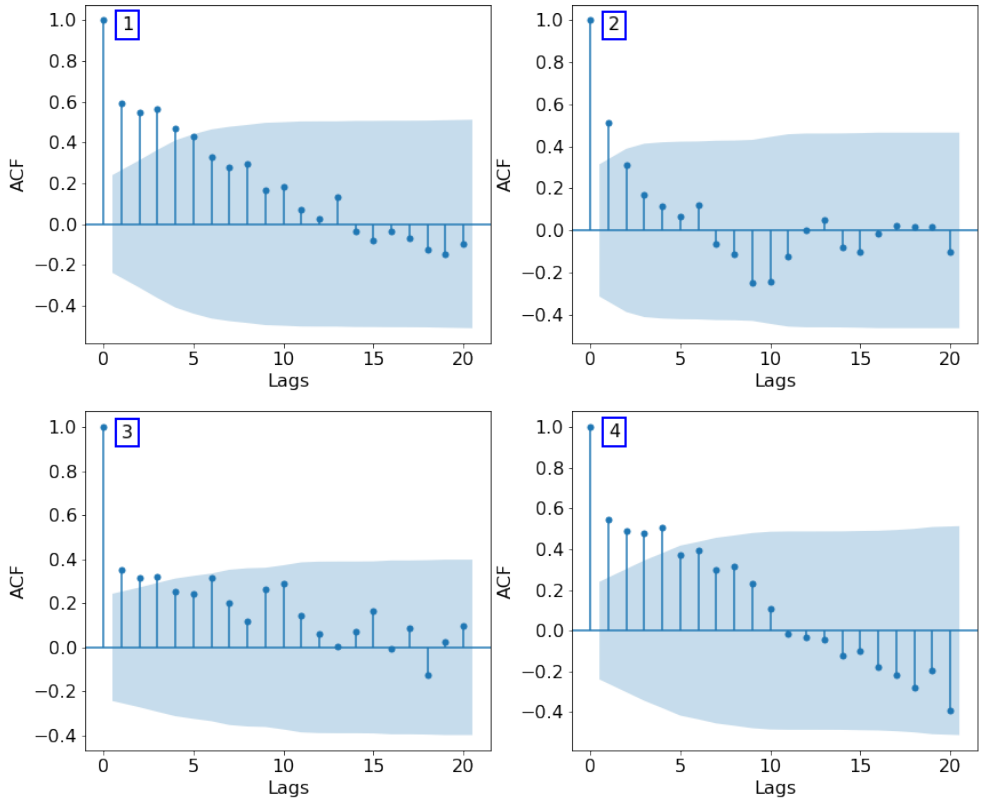

Figure 3.10: The figure shows correlogram of maximum amplitude calculated for each of the four clusters. The gradual variation of ACF and significance of coefficient (values that exceed the 95 confidence range are considered significant) suggest clustering of events.

exhibited a lognorm distribution, suggesting that the volcano goes through periods of survival and failure i.e. cyclicity (Table: 3.2) (Fig: 3.13).

Clusters	Best Fit	Parameters			P value - KST
		Shape	Location	Scale	
Cluster 1	Exponential	0.71	3	271	0.53

Table 3.2 continued from previous page

Cluster 2	Lognorm	1.33	41.50	234.13	0.83
Cluster 3	Gamma	0.62	7	263.91	0.70
Cluster 4	Lognorm	1.31	-4.12	158.88	0.97

Table 3.2
Parameters of Repose Time PDF

Clusters	Time Period	Coefficient of variation
	October 1, 2015 - January 13, 2016	2.51
Cluster 1	October 28, 2015 - November 7, 2015	1
Cluster 2	November 7, 2015 - December 1, 2015	1.21
Cluster 3	December 1, 2015 - December 15, 2015	1.25
Cluster 4	December 15, 2015 - January 5, 2016	1.40

Table 3.3
Coefficient of Variation of Repose Times

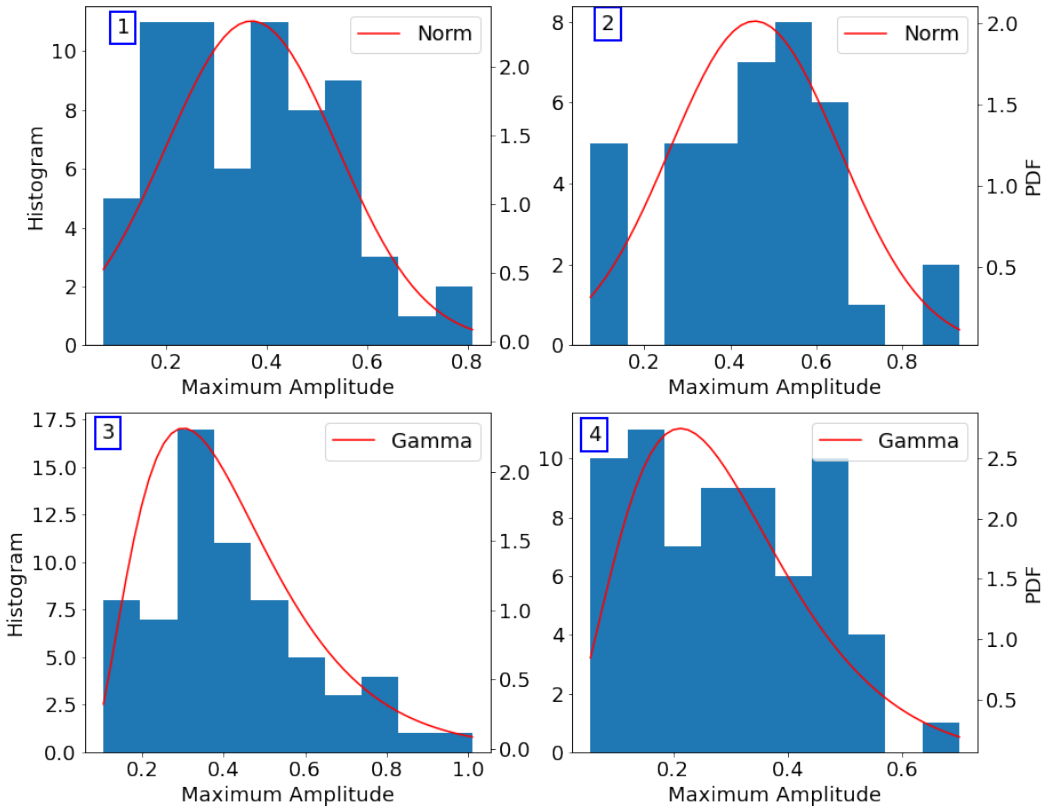

Figure 3.11: The figure shows Probability Distribution Functions (PDF) fitted for maximum amplitudes of inflation. The PDF fit shows variability between clusters which could be due to incompleteness of data.

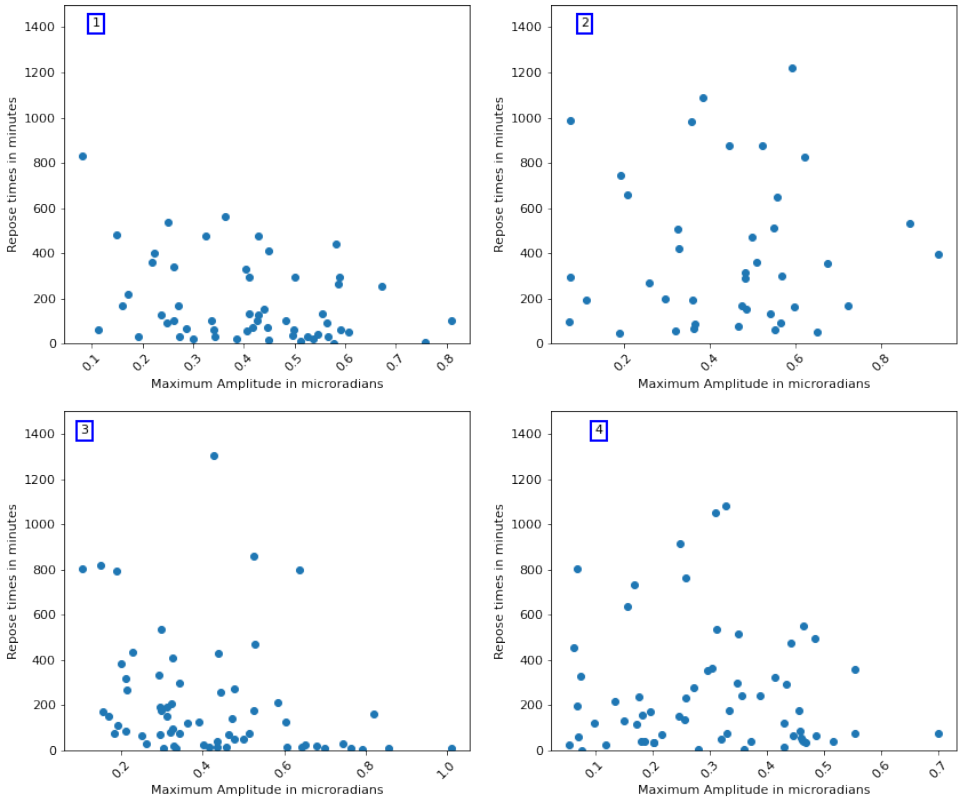

Figure 3.12: The figure shows maximum amplitude of events plotted against repose times. The values are distributed randomly for all the clusters except cluster 2 in which the events with maximum amplitude have lesser repose times.

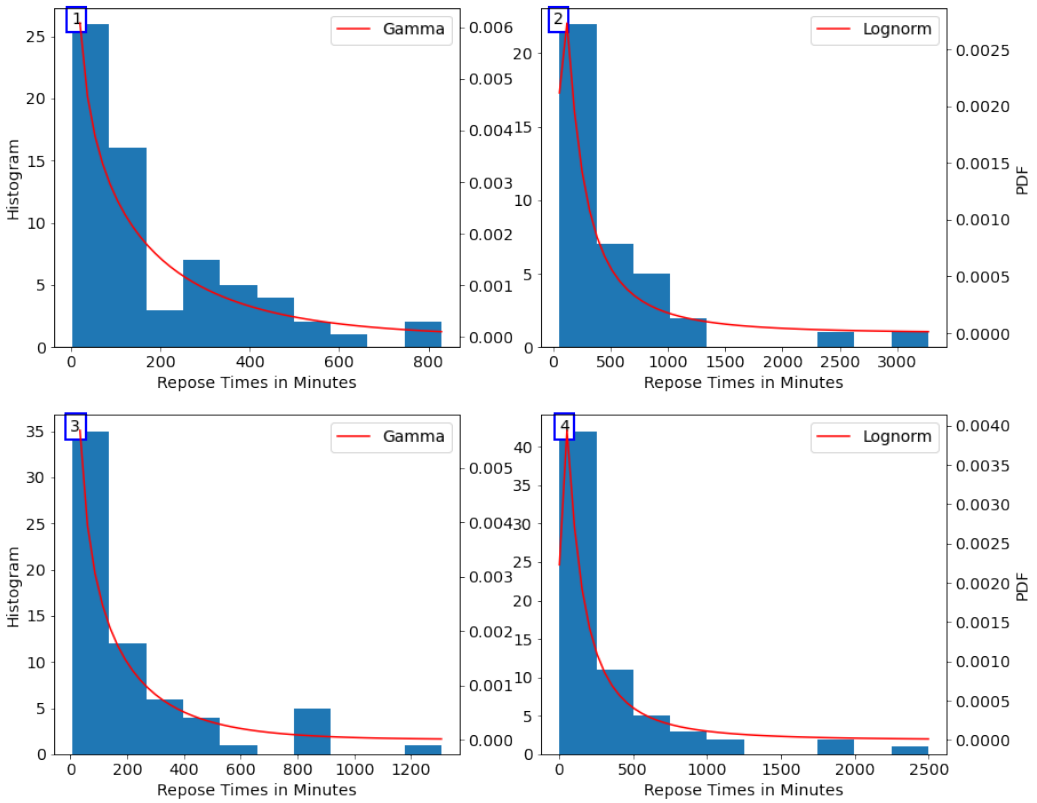

Figure 3.13: The figure shows Probability Distribution Functions (PDF) fitted for repose times of each of the four clusters. The clusters 2 through 4 follow survival/failure distributions suggesting a cyclic behaviour.

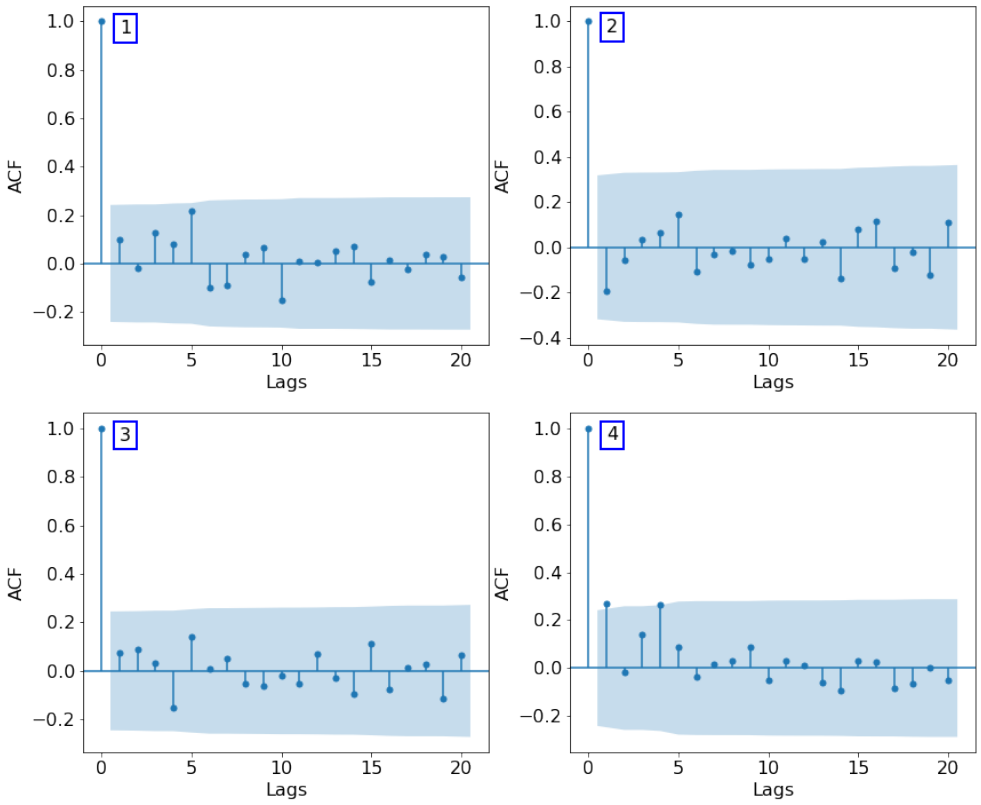

Figure 3.14: The figure shows correlogram calculated for repose times of each of the four clusters. The abrupt cut offs of coefficients around the mean indicate randomness of data.

Chapter 4

Discussion and Conclusion

4.1 Discussion

4.1.1 Unsupervised Learning on Seismically-Derived Tilt

The unsupervised learning applied to the seismically-derived tilt data revealed two meaningful patterns with seemingly opposite polarity (Fig: 3.4). We found 129 high signal-to-noise ratio events that resembled the pattern 1 whereas only 6 events showed similarity to the pattern 2 forming a very small sample size.

We focused on the east-west component of tilt because the amplitude of the tilt signal

was largest in this direction. In the absence of topography, we would expect little to no tilt on this component since it is nearly tangential to the source, and maximum tilt on the north component, which is nearly radial. But the fact that the east component has larger tilt reflects the influence of the north-south ridge that extends north from Fuego's summit.

The pattern 1 from the broadband data exhibited predominantly positive tilt in the east direction therefore tilting down towards east whereas the pattern 2 exhibited tilt down towards west. Predominant east and slight south downward tilt exhibited by pattern 1 events would suggest that the events undergo deflation followed by inflation just a few seconds prior to the eruptions. This behavior was observed in all the three years under study, and was also noted by *Lyons et al.* [2012] during their 2009 investigation of Fuego using seismically-derived tilt records. This behavior is contradictory to those observed in geophysical studies conducted on other basaltic as well as silicic volcanoes where inflation precedes deflation[*Maeda et al.*, 2017; *Nishimura et al.*, 2012; *Ripepe et al.*, 2021] and *Waite and Lanza* [2016]. *Lyons et al.* [2012] attribute this behavior to the shallow location of the source - source is located above the sensors - along with steep and uneven topography of the volcano, creating an illusion of reversed tilt (Fig: 4.1). The complex topography of the volcano especially the north - south ridge might contribute to the particularly strong tilt along east-west direction.

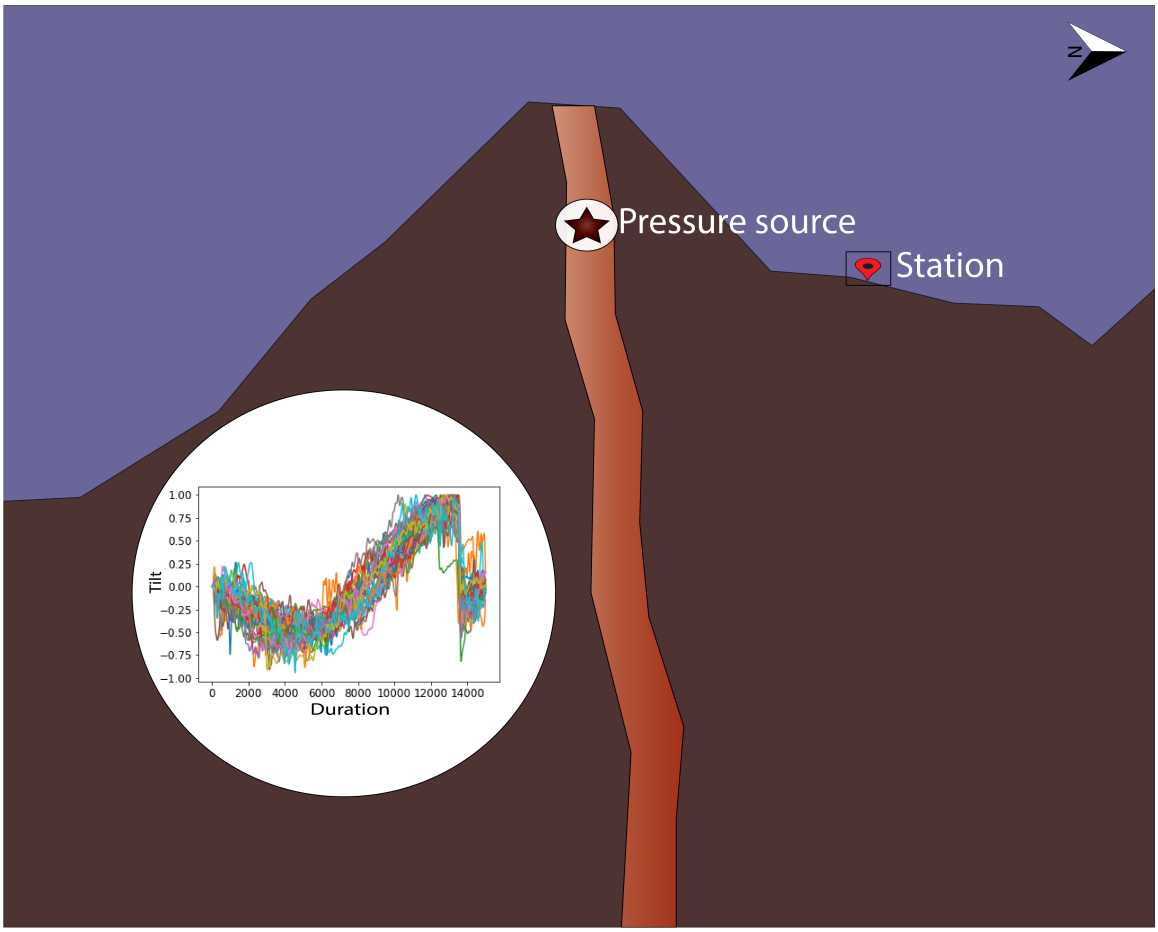

Figure 4.1: The figure shows a schematic representation of the pressure source located above the station location. The circle inset shows the pattern 1 events in the east component possibly associated with the above station level source. The positive tilt in east corresponds to deflation whereas the negative tilt corresponds to inflation. This seemingly reversed tilt might be attributed to the location of pressure source above the station coupled with irregular and steep topography.

Lyons et al. [2012] attribute these tilt signals to a pressurization/depressurization mechanism below a crystallized plug. When the Fuego magma reaches the surface, the water dissolved in it undergoes rapid exsolution leading to extensive plagioclase crystallization[*Lyons and Waite*, 2011]. The volcano undergoes rapid pressurization

beneath the plug until the overpressure exceeds the brittle strength of the plug resulting in deflation caused by ash-rich explosions.

The events of pattern 2 exhibit inflation followed by deflation. This might suggest that the source of these events is located below the level of stations and thereby invoking a varying pressure source along the conduit. But this is highly unlikely due to very low number of events. On the other hand, if the source location is assumed to be stationary and same as that of the pattern 1 events (above the level of stations), it would mean that the events of pattern 2 follow a deflation - inflation pattern. This behavior of deflation followed by inflation has only been mentioned in Kilauea and not in any other mafic or silicic systems. In Kilauea, this behavior is attributed to the convective flux in the lava lake created by sinking of gas-poor magma beneath gas-rich magma causing a blockage in the shallow conduit. The repeated creation and destruction of blockage produces these so called DI or deflation-inflation events [*Anderson et al.*, 2015]. But this mechanism is very different from what we observe in Fuego where the tilt events are caused by volatile overpressurization [*Lyons et al.*, 2012]. Alternately, these events of pattern 2 might be artifacts created by instrument errors. But we might expect this to reflect equally in all the three (vertical, north, and east) components of the seismometers. An analysis of three component tilt revealed no discernible tilt in the vertical (Fig: 3.2), therefore, we can rule out the possibility of artifacts. No significant difference was found in the maximum amplitude of these two patterns of events whereas the range of duration of pattern 2 events were slightly

less than that of pattern 1 events. The seismic traces associated with the events of patterns 1 and 2 were analyzed in higher frequency bands between 0.5 Hz and 10 Hz. While the events of pattern 1 revealed impulsive signals with long codas, the events of pattern 2 consist of tremor signals (Fig: 3.3, Fig: 3.4). Therefore, it is possible that these patterns are produced by different mechanisms. Temporal analysis of these pattern 2 events were not carried out due to very low sampling size and further research is required.

4.1.2 Tilt from Tiltmeter

The tilt data obtained from the tiltmeter was examined for an event similar to that of pattern 1 manually and this event was taken as a template for subsequent searching. The template matching using this template produced 268 events similar to that of pattern 1 but no events were identified similar to that of pattern 2. This could be due to low sampling rate of the tilt records obtained from the tiltmeter, and/or lower signal-to-noise ratio. The temporal distribution of the 268 events revealed seven clusters out of which four clusters were taken for further analysis. Each of these four clusters lasted for 10-15 days and exhibited a paroxysmal waxing and waning pattern (Fig: 3.6).

INSIVUMEH reported explosions accompanied by ash-rich eruptive plumes during

October 30, 2015 - November 3, 2015 which produced incandescent materials and ashfall as far as 12 km from the volcano. These explosions coincided with the start of cluster 1 marking the waxing period of the cycle. The number of explosions remained high until November 3 and started to wane but the amplitude of explosions remained high until November 5 (Fig: 3.8). By November 8, the explosive events subsided drastically followed by a paroxysmal eruption on November 9 and an increase in volcanic flux radiance indicative of new lava flows. This was further corroborated by INSIVUMEH reports [GVP, 2015].

The second cluster started immediately following the previous lava flow activity. The activity peaked during November 19 - November 24 during which large number of explosions were recorded. The activity started to wane off after November 25 producing a paroxysmal eruption on November 30, and new lava flows continued through November 29 and 30 indicated by high VRF and INSIVUMEH reports (Fig: 3.6) [GVP, 2015].

The third cluster started on December 3 and the number of events per day continued to increase up until December 6 but the maximum amplitude of inflation remained high until December 11. During this period, INSIVUMEH reported explosive eruptions accompanied by ash plumes 450-950 m in height and travelled as far as 12 km from the summit. The activity waned off after December 14 giving way to new lava flows and a paroxysmal eruption accompanied by block avalanches (Fig: 3.6) [GVP,

2015].

The fourth cluster started on December 16 immediately following the lava flows. The activity started to increase until December 18 and exhibited a decline for few days before resuming the intensity on December 22. The explosions remained high until December 25 and started to wane off giving rise to new lava flows. On December 30, INSIVUMEH reported that pyroclastic flows descended the SW drainages of Las Lajas and El Jute, but the number of explosions catalogued remained low during that time period whereas the maximum amplitude showed a relative increase. On January 3, paroxysmal eruption was reported and new lava flows were identified by high radiant flux which was also corroborated by INSIVUMEH reports (Fig: 3.6) [GVP, 2015].

The duration of these events showed a large variation ranging from 7 minutes to 39 minutes. But temporal variation of distribution within each cluster showed a random distribution (Fig: 3.7). The difference in duration between each events may be attributed to the rate of crystallization, extent of crystallization, and also the brittle strength of the crystallized plug.

The maximum amplitude of inflation followed a similar waxing and waning pattern and coincided with the number of events per day (Fig: 3.8). Several studies have reported this cyclic waxing and waning activity in Fuego volcano [*Lyons et al.*, 2009; *Naismith et al.*, 2019]. *Lyons et al.* [2009] proposed two models for this cyclic behavior

- collapsing foam model and magma rise-speed model. The collapsing foam model is based on experimental work carried out by *Jaupart and Vergniolle* [1988] and they showed that both effusive and explosive eruptions can be explained by the collapse of foam layer accumulated on the roof of a feeder structure. In their experiment, the authors noticed that the foam layer grew in response to increased gas flux and eventually collapsed into a large gas slug. They suggest that the rise of gas slug through the conduit could drive Strombolian like explosions. And if the viscosity of magma is high enough, it would give rise to cyclic gas slugs which may explain the explosive eruptions we see at Fuego. However, the increase in number of gas slugs and thereby explosions would allow less time for crystallization to occur in the top of the conduit. Therefore, we would observe lower amplitudes. But at Fuego, we observe an opposite trend where the increase in number of explosions is also accompanied by increase in amplitudes.

The magma rise-speed model was first proposed by *Parfitt and Wilson* [1995]. They related the change in magma rise speed to the change in eruptive behavior of volcanoes. When the magma speed is low, the bubbles formed would have more time to coalesce, giving rise to Strombolian like explosions. On the other hand, if the magma rise-speed is high enough, the speed differential between rising magma and volatile would be very low preventing the bubbles from coalescing into each other. Therefore, the ascending magma would reach the fragmentation threshold much earlier before it reaches the surface. The amount of fragmentation is directly related to the amount of

volatile present in the rising magma, and as the fragmentation increases so would the pressure in the conduit giving way to more frequent and explosive eruptions. Since our data shows a positive correlation between number of events and amplitudes indicative of increase in eruption frequency and explosivity, we favor the magma rise-speed model.

The distribution of amplitude showed wide variation similar to the duration of events and ranged as low as 0.5 micro-radian to almost 1 micro-radian (Fig: 3.7). *Varley et al.* [2006] analyzed the repose times of eruptions occured in Volcan de Colima during 2003, 2004, and 2005. They used ACF correlograms and identified periodicity or clustering of events in the volcano. A similar observation was made in Tungurahua volcano during 2004 eruptions [*Varley et al.*, 2006]. Similarly, the correlograms of each cluster showed a gradual decrease in ACF coefficient indicative of clustering of activity. Moreover, the significance of coefficients suggest a higher-order autoregressive (AR) model. This suggests that a future eruption at Fuego is dependent on more than one eruption that precede it. Although the PDF distributions fit both gamma and normal distributions, we believe this variation might be due to incompleteness of catalog. Therefore, the temporal distribution of amplitude, periodicity of ACF coefficients, and PDF fit with a survival model such as gamma strongly suggest a cyclic activity.

The ACF plot of repose time shows only minimal correlation and abrupt cut offs

suggesting less clustering of data (Fig: 3.13). The coefficient of variability also hovers a little over 1 for all the four clusters (Table: 3.3). The recorded number of events in the catalog is lower for the activity usually observed in Fuego. Therefore, we believe the incompleteness of data is masking the clustering of repose times. *Varley et al.* [2006] used PDF distributions to come up with potential models for Volcan de Colima and Tungurahua and found that none of the data fit renewal model but rather showed good fit towards survival models such as gamma, Weibull, and log-logistic. The repose times of Fuego volcano, although showing variability in fit between clusters, generally follow survival/failure models of gamma and lognorm. These survival models incorporate periodic behavior of the volcanoes and therefore is indicative of cyclic activities where volcanoes undergo repeated waxing and waning.

4.2 Conclusion

Unsupervised learning using DTW distance measure was used on tilt records derived from seismometers. This process revealed two meaningful patterns with opposite polarity. The seismic traces associated with pattern 1 tilt signals consist of impulsive, explosive signals with long codas whereas the seismic traces associated with pattern 2 contain tremor signals. Therefore, it is possible that these two tilt patterns are produced by different mechanisms. A template matching procedure was carried out on tilt data obtained from tiltmeter which produced 268 events similar to pattern

1. The temporal distribution of these events revealed four significant clusters each lasting 10 - 15 days. The number of events per day and maximum amplitude of inflation exhibited a positive correlation along with a waxing and waning pattern. This supports magma rise-speed model which proposes that an increase in the speed of ascending magma would give way to increased activity as well as intensity. The Auto-correlation function and probability distribution functions of maximum amplitude and repose times were calculated. The ACF coefficients of maximum amplitude followed a gradual change indicative of clustering activity, while the ACF of repose times exhibited abrupt cut offs around mean. We believe this variation might be attributed to incompleteness of the data. The PDF distributions fitted for maximum amplitude and repose times favored survival/failure distributions such as gamma and lognormal. These distributions are indicative of periodic activity in volcanic systems. The temporal and statistical analysis suggest that Fuego undergoes cyclic activity.

References

[1] Global volcanism program: Report on fuego (guatemala), 2015.

[2] Guatemala's fuego volcano: How the tragedy unfolded, *BBC News*, 2018.

[3] Global volcanism program: Report on fuego (guatemala) - august 2018, 2018.

[4] Guatemala ups number of missing to 332 in volcano eruption, 2018.

[5] Anderson, K., M. Poland, J. Johnson, and A. Miklius, *Episodic Deflation–Inflation Events at Kīlauea Volcano and Implications for the Shallow Magma System*, vol. 208, pp. 229–250, doi:10.1002/9781118872079.ch11, 2015.

[6] Anzieta, J. C., H. D. Ortiz, G. L. Arias, and M. C. Ruiz, Finding possible precursors for the 2015 cotopaxi volcano eruption using unsupervised machine learning techniques, *International Journal of Geophysics*, *2019*, 1–8, doi:10.1155/2019/6526898, 2019.

[7] Aoyama, H., and H. Oshima, Tilt change recorded by broadband seismometer

prior to small phreatic explosion of meakan-dake volcano, hokkaido, japan, *Geophysical Research Letters*, *35*(6), doi:10.1029/2007gl032988, 2008.

[8] Bebbington, M., and C. Lai, Statistical analysis of new zealand volcanic occurrence data, *Journal of Volcanology and Geothermal Research*, *74*(1-2), 101–110, doi:10.1016/s0377-0273(96)00050-9, 1996.

[9] Brill, K. A., and G. P. Waite, Characteristics of repeating long-period seismic events at fuego volcano, january 2012, *Journal of Geophysical Research: Solid Earth*, *124*(8), 8644–8659, doi:10.1029/2019jb017902, 2019.

[10] Carr, M. J., Symmetrical and segmented variation of physical and geochemical characteristics of the central american volcanic front, *Journal of Volcanology and Geothermal Research*, *20*(3-4), 231–252, doi:10.1016/0377-0273(84)90041-6, 1984.

[11] Chesner, C. A., and W. I. Rose, Geochemistry and evolution of the fuego volcanic complex, guatemala, *Journal of Volcanology and Geothermal Research*, *21*(1-2), 25–44, doi:10.1016/0377-0273(84)90014-3, 1984.

[12] Chouet, B. A., and R. S. Matoza, A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption, *Journal of Volcanology and Geothermal Research*, *252*, 108–175, doi:10.1016/j.jvolgeores.2012.11.013, 2013.

[13] Fiorini, L., G. Mancioppi, F. Semeraro, H. Fujita, and F. Cavallo, Unsupervised emotional state classification through physiological parameters for social robotics

applications, *Knowledge-Based Systems*, *190*, 105,217, doi:10.1016/j.knosys.2019.105217, 2020.

[14] Gazel, E., K. E. Flores, and M. J. Carr, Architectural and tectonic control on the segmentation of the central american volcanic arc, *Annual Review of Earth and Planetary Sciences*, *49*(1), 495–521, doi:10.1146/annurev-earth-082420-055108, 2021.

[15] Genco, R., and M. Ripepe, Inflation-deflation cycles revealed by tilt and seismic records at stromboli volcano, *Geophysical Research Letters*, *37*(12), doi:10.1029/2010gl042925, 2010.

[16] Halsor, S. P., and W. I. Rose, Common characteristics of paired volcanoes in northern central america, *Collected Reprint Series 1989, How Volcanoes Work*, p. 4467–4476, doi:10.1002/9781118782064.ch37, 1988.

[17] Hutchison, A., K. Cashman, C. Williams, and A. Rust, The 1717 eruption of volcán de fuego, guatemala: Cascading hazards and societal response, *Quaternary International*, *394*, 69–78, doi:10.1016/j.quaint.2014.09.050, 2016.

[18] Jaupart, C., and S. Vergniolle, Laboratory models of hawaiian and strombolian eruptions, *Nature*, *331*(6151), 58–60, doi:10.1038/331058a0, 1988.

[19] Kumarsagar, H., and V. Sharma, Error evaluation on k- means and hierarchical clustering with effect of distance functions for iris dataset, *International Journal of Computer Applications*, *86*(16), 1–5, doi:10.5120/15066-3429, 2014.

[20] Lyons, J. J., and G. P. Waite, Dynamics of explosive volcanism at fuego volcano imaged with very long period seismicity, *Journal of Geophysical Research*, *116*(B9), doi:10.1029/2011jb008521, 2011.

[21] Lyons, J. J., G. P. Waite, W. I. Rose, and G. Chigna, Patterns in open vent, strombolian behavior at fuego volcano, guatemala, 2005–2007, *Bulletin of Volcanology*, *72*(1), 1–15, doi:10.1007/s00445-009-0305-7, 2009.

[22] Lyons, J. J., G. P. Waite, M. Ichihara, and J. M. Lees, Tilt prior to explosions and the effect of topography on ultra-long-period seismic records at fuego volcano, guatemala, *Geophysical Research Letters*, *39*(8), doi:10.1029/2012gl051184, 2012.

[23] Maeda, Y., A. Kato, and Y. Yamanaka, Modeling the dynamics of a phreatic eruption based on a tilt observation: Barrier breakage leading to the 2014 eruption of mount ontake, japan, *Journal of Geophysical Research: Solid Earth*, *122*(2), 1007–1024, doi:10.1002/2016jb013739, 2017.

[24] Martin, D. P., and W. I. Rose, Behavioral patterns of fuego volcano, guatemala, *Journal of Volcanology and Geothermal Research*, *10*(1-3), 67–81, doi:10.1016/0377-0273(81)90055-x, 1981.

[25] Nadeau, P., J. Palma, and G. Waite, Linking volcanic tremor, degassing, and eruption dynamics via so2 imaging, *Geophysical Research Letters - GEOPHYS RES LETT*, *38*, doi:10.1029/2010GL045820, 2011.

[26] Naismith, A. K., I. M. Watson, R. Escobar-Wolf, G. Chigna, H. Thomas, D. Coppola, and C. Chun, Eruption frequency patterns through time for the current (1999–2018) activity cycle at volcán de fuego derived from remote sensing data: Evidence for an accelerating cycle of explosive paroxysms and potential implications of eruptive activity, *Journal of Volcanology and Geothermal Research*, *371*, 206–219, doi:10.1016/j.jvolgeores.2019.01.001, 2019.

[27] Nasukawa, T., and J. Yi, Sentiment analysis, *Proceedings of the international conference on Knowledge capture - K-CAP 03*, doi:10.1145/945645.945658, 2003.

[28] Nishimura, T., M. Iguchi, R. Kawaguchi, Surono, M. Hendrasto, and U. Rosadi, Inflations prior to vulcanian eruptions and gas bursts detected by tilt observations at semeru volcano, indonesia, *Bulletin of Volcanology*, *74*(4), 903–911, doi:10.1007/s00445-012-0579-z, 2012.

[29] Nishimura, T., M. Iguchi, H. Yakiwara, J. Oikawa, R. Kawaguchi, H. Aoyama, H. Nakamichi, Y. Ohta, and T. Tameguri, Mechanism of small vulcanian eruptions at suwanosejima volcano, japan, as inferred from precursor inflations and tremor signals, *Bulletin of Volcanology*, *75*(12), doi:10.1007/s00445-013-0779-1, 2013.

[30] Parfitt, E. A., and L. Wilson, Explosive volcanic eruptions-ix. the transition between hawaiian-style lava fountaining and strombolian explosive activity, *Geophysical Journal International*, *121*(1), 226–232, doi:10.1111/j.1365-246x.1995.

tb03523.x, 1995.

[31] Pillet, R., and J. Virieux, The effects of seismic rotations on inertial sensors, *Geophysical Journal International*, *171*(3), 1314–1323, doi:10.1111/j.1365-246x.2007.03617.x, 2007.

[32] Ramoni, M., P. Sebastiani, and P. Cohen, Unsupervised clustering of robot activities, *Proceedings of the fourth international conference on Autonomous agents - AGENTS 00*, doi:10.1145/336595.337081, 2000.

[33] Ratanamahatana, C. A., and E. Keogh, Everything you know about dynamic time warping is wrong, in *Third Workshop on Mining Temporal and Sequential Data*, 2004.

[34] Ren, C. X., A. Peltier, V. Ferrazzini, B. Rouet-Leduc, P. A. Johnson, and F. Brenguier, Machine learning reveals the seismic signature of eruptive behavior at piton de la fournaise volcano, *Geophysical Research Letters*, *47*(3), doi:10.1029/2019gl085523, 2020.

[35] Ripepe, M., D. D. Donne, G. Lacanna, E. Marchetti, and G. Ulivieri, The onset of the 2007 stromboli effusive eruption recorded by an integrated geophysical network, *Journal of Volcanology and Geothermal Research*, *182*(3-4), 131–136, doi:10.1016/j.jvolgeores.2009.02.011, 2009.

[36] Ripepe, M., et al., Ground deformation reveals the scale-invariant conduit dynamics driving explosive basaltic eruptions, *Nature Communications*, *12*, doi: 10.1038/s41467-021-21722-2, 2021.

[37] Rodgers, P. W., The response of the horizontal pendulum seismometer to rayleigh and love waves, tilt, and free oscillations of the earth, *Bulletin of the Seismological Society of America*, *58*(5), 1385–1406, doi:10.1785/bssa0580051385, 1968.

[38] Rose, W. I., A. T. Anderson, L. G. Woodruff, and S. B. Bonis, The october 1974 basaltic tephra from fuego volcano: Description and history of the magma body, *Journal of Volcanology and Geothermal Research*, *4*(1-2), 3–53, doi:10.1016/0377-0273(78)90027-6, 1978.

[39] Salvador, S., and P. Chan, Toward accurate dynamic time warping in linear time and space, *Intelligent Data Analysis*, *11*(5), 561–580, doi:10.3233/ida-2007-11508, 2007.

[40] Seydoux, L., R. Balestriero, P. Poli, M. D. Hoop, M. Campillo, and R. Baraniuk, Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning, *Nature Communications*, *11*(1), doi:10.1038/s41467-020-17841-x, 2020.

[41] Stoiber, R. E., and M. J. Carr, Quaternary volcanic and tectonic segmentation of central america, *Bulletin Volcanologique*, *37*(3), 304–325, doi:10.1007/bf02597631, 1973.

[42] Trnkoczy, A., Understanding and parameter setting of sta/lta trigger algorithm, 1970.

[43] Vallance, J., S. Schilling, O. Matías, W. I. Rose, and M. Howell, Volcano hazards at fuego and acatenango, guatemala, *Open-File Report*, doi:10.3133/ofr01431, 2001.

[44] Varley, N., J. Johnson, M. Ruiz, G. Reyes-Dávila, and K. Martin, *Applying statistical analysis to understanding the dynamics of volcanic explosions*, 2006.

[45] Waite, G. P., and F. Lanza, Nonlinear inversion of tilt-affected very long period records of explosive eruptions at fuego volcano, *Journal of Geophysical Research: Solid Earth*, *121*(10), 7284–7297, doi:10.1002/2016jb013287, 2016.

[46] Waite, G. P., P. A. Nadeau, and J. J. Lyons, Variability in eruption style and associated very long period events at fuego volcano, guatemala, *Journal of Geophysical Research: Solid Earth*, *118*(4), 1526–1533, doi:10.1002/jgrb.50075, 2013.

[47] Wallace, T., The guatemala volcano eruption: Before and after a deadly pyroclastic flow, 2018.

[48] Witsil, A. J., and J. B. Johnson, Analyzing continuous infrasound from stromboli volcano, italy using unsupervised machine learning, *Computers & Geosciences*, *140*, 104,494, doi:10.1016/j.cageo.2020.104494, 2020.

[49] Wooster, M., Fire radiative energy for quantitative study of biomass burning: derivation from the bird experimental satellite and comparison to modis fire products, *Remote Sensing of Environment*, *86*(1), 83–107, doi:10.1016/s0034-4257(03)00070-1, 2003.

[50] Wright, R., M. Blackett, and C. Hill-Butler, Some observations regarding the thermal flux from earths erupting volcanoes for the period of 2000 to 2014, *Geophysical Research Letters*, *42*(2), 282–289, doi:10.1002/2014gl061997, 2015.

[51] Zagibalov, T., and J. Carroll, Automatic seed word selection for unsupervised sentiment classification of chinese text, *Proceedings of the 22nd International Conference on Computational Linguistics - COLING 08*, doi:10.3115/1599081.1599216, 2008.

Appendix A

Catalog of Events - 2015

*Amp = Amplitude

Start Time	Inflation Duration	Initial Amp*	Max Amp*	Deflation Amp*
10/2/15 13:20	20	-0.0278	0.1138	-0.0463
10/4/15 1:41	13	-0.0301	0.0702	-0.0326
10/4/15 2:32	13	-0.0134	0.0947	-0.0442
10/4/15 14:57	15	-0.0112	0.054	-0.0546
10/4/15 15:43	10	-0.0078	0.0908	-0.0642
10/5/15 15:08	20	-0.0208	0.0636	-0.0216
10/5/15 18:53	38	-0.0405	0.0699	-0.0683
10/6/15 5:10	34	-0.0358	0.0838	-0.0777
10/6/15 7:05	8	-0.0213	0.0821	-0.0928

Table A.1 continued from previous page

10/6/15 8:20	30	-0.0549	0.0972	-0.0758
10/6/15 11:50	26	-0.0232	0.0922	-0.1236
10/6/15 12:52	23	-0.0111	0.0723	-0.0732
10/9/15 15:12	13	-0.0298	0.1134	-0.0451
10/9/15 22:55	8	-0.008	0.0908	-0.0358
10/15/15 21:00	25	-0.0465	0.144	-0.0804
10/15/15 21:28	32	-0.128	0.241	-0.209
10/16/15 11:22	17	-0.039	0.175	-0.021
10/16/15 12:26	24	-0.033	0.0969	-0.0284
10/16/15 18:23	22	-0.0318	0.1222	-0.0556
10/17/15 7:52	10	-0.0173	0.0729	-0.0247
10/17/15 10:13	37	-0.0261	0.0602	-0.0779
10/18/15 22:01	12	-0.0243	0.121	-0.0535
10/19/15 22:16	18	-0.045	0.159	-0.126
10/22/15 20:38	29	-0.027	0.0418	-0.0347
10/23/15 10:35	18	-0.0272	0.0921	-0.0685
10/23/15 14:46	28	-0.0291	0.1061	-0.0625
10/24/15 16:41	19	-0.0177	0.0781	-0.0409
10/25/15 13:19	17	-0.0195	0.0918	-0.0519
10/29/15 23:02	23	-0.0433	0.114	-0.0601

Table A.1 continued from previous page

10/30/15 13:14	13	-0.0175	0.064	-0.0445
10/30/15 22:22	21	-0.043	0.208	-0.097
10/31/15 6:40	30	-0.093	0.232	-0.222
10/31/15 9:58	18	-0.0645	0.095	-0.0969
10/31/15 11:20	15	-0.0291	0.0838	-0.0606
10/31/15 18:57	33	-0.238	0.345	-0.266
10/31/15 20:35	12	-0.055	0.232	-0.137
10/31/15 21:08	22	-0.082	0.217	-0.168
10/31/15 23:11	39	-0.217	0.21	-0.384
11/1/15 1:20	24	-0.083	0.165	-0.138
11/1/15 2:45	38	-0.257	0.334	-0.419
11/1/15 4:25	35	-0.208	0.29	-0.415
11/1/15 6:04	26	-0.138	0.202	-0.316
11/1/15 8:12	22	-0.103	0.234	-0.115
11/1/15 17:58	36	-0.126	0.237	-0.216
11/1/15 19:25	17	-0.169	0.437	-0.243
11/1/15 20:11	17	-0.146	0.38	-0.332
11/1/15 21:41	10	-0.087	0.33	-0.159
11/1/15 23:31	29	-0.249	0.561	-0.519
11/2/15 5:30	20	-0.135	0.268	-0.207

Table A.1 continued from previous page

11/2/15 6:05	30	-0.191	0.258	-0.272
11/2/15 7:10	13	-0.195	0.301	-0.236
11/2/15 7:26	20	-0.236	0.342	-0.275
11/2/15 14:36	24	-0.173	0.275	-0.489
11/2/15 15:13	31	-0.223	0.289	-0.251
11/2/15 17:57	25	-0.19	0.22	-0.088
11/2/15 20:34	18	-0.21	0.344	-0.434
11/2/15 21:50	25	-0.174	0.233	-0.494
11/2/15 22:55	23	-0.246	0.301	-0.224
11/2/15 23:40	24	-0.098	0.287	-0.245
11/3/15 5:00	30	-0.184	0.226	-0.575
11/3/15 5:52	38	-0.332	0.204	-0.397
11/3/15 8:37	39	-0.235	0.193	-0.39
11/3/15 9:49	38	-0.227	0.117	-0.327
11/3/15 11:37	30	-0.094	0.353	-0.149
11/3/15 13:47	17	-0.147	0.335	-0.271
11/3/15 14:34	30	-0.155	0.412	-0.379
11/3/15 19:19	31	-0.226	0.446	-0.345
11/4/15 0:42	18	-0.201	0.389	-0.293
11/4/15 1:06	33	-0.331	0.427	-0.395

Table A.1 continued from previous page

11/4/15 6:03	15	-0.211	0.375	-0.202
11/4/15 14:15	16	-0.111	0.317	-0.21
11/4/15 19:23	34	-0.16	0.34	-0.316
11/4/15 21:27	33	-0.154	0.409	-0.216
11/5/15 0:33	24	-0.115	0.324	-0.196
11/5/15 7:36	19	-0.071	0.152	-0.147
11/5/15 13:36	9	-0.049	0.212	-0.131
11/5/15 16:35	21	-0.075	0.196	-0.092
11/5/15 17:25	23	-0.099	0.174	-0.116
11/5/15 23:50	18	-0.0729	0.1455	-0.0864
11/6/15 3:49	38	-0.0484	0.1241	-0.0948
11/6/15 6:34	15	-0.07	0.167	-0.158
11/6/15 7:19	35	-0.0485	0.1427	-0.082
11/6/15 9:38	16	-0.027	0.234	-0.139
11/6/15 17:55	33	-0.0369	0.1125	-0.1323
11/6/15 23:08	27	-0.0266	0.0812	-0.0982
11/7/15 0:32	8	-0.007	0.142	-0.142
11/7/15 2:30	37	-0.095	0.178	-0.133
11/7/15 4:21	19	-0.0363	0.0834	-0.0988
11/7/15 8:44	28	-0.15	0.23	-0.159

Table A.1 continued from previous page

11/7/15 15:57	19	-0.031	0.179	-0.159
11/7/15 18:10	22	-0.02	0.217	-0.124
11/8/15 6:59	38	-0.047	0.145	-0.195
11/8/15 9:14	12	-0.0053	0.0686	-0.0661
11/8/15 10:10	14	-0.037	0.1524	-0.0398
11/8/15 21:21	14	-0.0367	0.1723	-0.0985
11/12/15 22:50	29	-0.0166	0.0934	-0.0394
11/13/15 2:33	14	-0.0221	0.091	-0.0487
11/13/15 19:11	18	-0.084	0.273	-0.24
11/14/15 13:38	39	-0.196	0.189	-0.306
11/14/15 17:28	30	-0.14	0.22	-0.277
11/15/15 1:48	38	-0.17	0.33	-0.408
11/15/15 3:24	26	-0.133	0.189	-0.324
11/17/15 10:23	35	-0.318	0.297	-0.566
11/17/15 13:10	29	-0.194	0.348	-0.54
11/17/15 19:32	25	-0.324	0.351	-0.593
11/18/15 4:49	28	-0.477	0.39	-0.525
11/18/15 11:52	34	-0.53	0.403	-0.568
11/18/15 23:13	37	-0.173	0.384	-0.474
11/19/15 8:15	29	-0.055	0.272	-0.278

Table A.1 continued from previous page

11/19/15 13:57	32	-0.104	0.378	-0.461
11/19/15 17:11	38	-0.33	0.268	-0.347
11/19/15 20:38	19	-0.197	0.526	-0.395
11/19/15 22:28	16	-0.14	0.426	-0.286
11/20/15 4:43	17	-0.14	0.371	-0.378
11/20/15 12:01	15	-0.072	0.257	-0.142
11/21/15 2:01	19	-0.097	0.524	-0.278
11/21/15 5:39	17	-0.06	0.237	-0.22
11/21/15 10:57	19	-0.154	0.415	-0.42
11/22/15 7:38	21	-0.222	0.369	-0.268
11/22/15 9:14	35	-0.12	0.348	-0.425
11/22/15 10:57	12	-0.103	0.259	-0.207
11/22/15 12:03	13	-0.209	0.443	-0.421
11/23/15 2:51	14	-0.092	0.43	-0.224
11/23/15 4:09	18	-0.199	0.352	-0.309
11/23/15 9:17	19	-0.109	0.374	-0.227
11/23/15 18:07	17	-0.153	0.396	-0.313
11/23/15 21:12	20	-0.132	0.343	-0.257
11/23/15 22:58	10	-0.074	0.293	-0.24
11/24/15 13:44	13	-0.147	0.299	-0.178

Table A.1 continued from previous page

11/24/15 16:29	15	-0.101	0.386	-0.273
11/25/15 9:10	23	-0.0181	0.0565	-0.0746
11/25/15 14:02	17	-0.019	0.241	-0.138
11/25/15 19:12	31	-0.0096	0.0674	-0.0412
11/27/15 12:48	7	-0.0062	0.0917	-0.0481
12/4/15 3:31	27	-0.0395	0.1375	-0.1053
12/4/15 17:14	38	-0.0835	0.1043	-0.1431
12/5/15 7:15	10	-0.0191	0.0857	-0.0879
12/5/15 9:18	18	-0.0366	0.1538	-0.0912
12/5/15 11:02	12	-0.025	0.185	-0.173
12/5/15 14:04	21	-0.0499	0.1049	-0.0861
12/5/15 14:37	32	-0.108	0.195	-0.212
12/5/15 21:33	24	-0.052	0.148	-0.134
12/6/15 1:22	34	-0.087	0.235	-0.068
12/6/15 8:45	38	-0.097	0.23	-0.26
12/6/15 9:45	37	-0.109	0.221	-0.236
12/6/15 11:40	20	-0.0423	0.1416	-0.1274
12/6/15 15:14	35	-0.113	0.181	-0.17
12/6/15 16:56	32	-0.096	0.154	-0.133
12/6/15 19:06	34	-0.138	0.188	-0.11

Table A.1 continued from previous page

12/6/15 19:57	34	-0.106	0.329	-0.299
12/6/15 21:47	30	-0.069	0.274	-0.157
12/6/15 22:28	38	-0.128	0.206	-0.334
12/6/15 23:38	22	-0.073	0.19	-0.126
12/7/15 0:15	32	-0.154	0.261	-0.213
12/7/15 3:20	9	0.007	0.175	-0.1
12/7/15 4:48	18	-0.068	0.253	-0.38
12/7/15 5:20	34	-0.128	0.329	-0.261
12/7/15 6:02	31	-0.205	0.494	-0.421
12/7/15 9:32	27	-0.09	0.209	-0.228
12/7/15 10:12	28	-0.165	0.442	-0.365
12/7/15 15:39	20	-0.101	0.243	-0.235
12/7/15 18:32	23	-0.075	0.237	-0.337
12/7/15 23:28	31	-0.188	0.289	-0.428
12/8/15 0:07	32	-0.714	0.297	-0.301
12/8/15 2:40	22	-0.097	0.267	-0.296
12/8/15 4:19	26	-0.159	0.354	-0.333
12/8/15 5:34	30	-0.198	0.279	-0.203
12/8/15 9:14	38	-0.1	0.212	-0.225
12/8/15 10:16	38	-0.285	0.368	-0.382

Table A.1 continued from previous page

12/9/15 8:41	38	-0.174	0.252	-0.321
12/9/15 9:32	38	-0.356	0.288	-0.605
12/9/15 12:30	26	-0.117	0.356	-0.292
12/9/15 15:54	34	-0.267	0.258	-0.364
12/9/15 16:39	38	-0.419	0.346	-0.737
12/10/15 6:37	33	-0.276	0.362	-0.523
12/10/15 12:43	39	-0.108	0.184	-0.473
12/10/15 14:35	32	-0.171	0.292	-0.594
12/10/15 18:38	28	-0.165	0.419	-0.489
12/10/15 19:37	38	-0.265	0.479	-0.697
12/10/15 20:35	34	-0.301	0.378	-0.595
12/10/15 23:52	27	-0.281	0.54	-0.617
12/11/15 4:39	22	-0.106	0.337	-0.34
12/11/15 5:08	27	-0.367	0.426	-0.533
12/11/15 5:43	36	-0.406	0.45	-0.494
12/11/15 10:47	27	-0.105	0.108	-0.199
12/11/15 13:19	27	-0.081	0.31	-0.235
12/11/15 22:42	24	-0.095	0.203	-0.136
12/12/15 12:45	27	-0.0459	0.1027	-0.1198
12/12/15 14:04	30	-0.162	0.337	-0.365

Table A.1 continued from previous page

12/13/15 4:54	37	-0.222	0.304	-0.527
12/13/15 12:40	37	-0.159	0.279	-0.376
12/13/15 18:38	23	-0.054	0.156	-0.268
12/13/15 21:07	36	-0.21	0.394	-0.485
12/13/15 22:23	31	-0.14	0.297	-0.429
12/13/15 23:18	36	-0.139	0.262	-0.408
12/14/15 1:06	31	-0.085	0.211	-0.259
12/14/15 8:53	35	-0.1272	0.101	-0.1687
12/14/15 17:21	26	-0.14	0.389	-0.256
12/14/15 19:06	24	-0.102	0.333	-0.187
12/17/15 8:14	36	-0.055	0.0729	-0.1017
12/17/15 9:49	37	-0.0374	0.0329	-0.0529
12/17/15 22:42	13	-0.0304	0.1387	-0.0394
12/18/15 4:22	25	-0.0197	0.0546	-0.0336
12/18/15 12:24	27	-0.0154	0.0463	-0.0282
12/18/15 16:08	31	-0.0121	0.0568	-0.0427
12/19/15 6:01	20	-0.0124	0.0565	-0.0406
12/19/15 7:10	16	-0.06	0.261	-0.164
12/19/15 7:53	13	-0.0233	0.0319	-0.0078
12/19/15 8:08	15	-0.007	0.0703	-0.0462

Table A.1 continued from previous page

12/19/15 9:00	20	-0.0484	0.1535	-0.1194
12/20/15 16:12	10	-0.0188	0.1313	-0.047
12/20/15 21:18	36	-0.086	0.263	-0.399
12/21/15 7:04	39	-0.203	0.262	-0.505
12/21/15 23:00	29	-0.095	0.153	-0.223
12/22/15 17:34	29	-0.089	0.239	-0.314
12/23/15 2:21	21	-0.101	0.383	-0.417
12/23/15 10:36	22	-0.08	0.362	-0.3
12/23/15 16:51	16	-0.066	0.23	-0.134
12/23/15 22:01	38	-0.215	0.219	-0.38
12/23/15 23:18	22	-0.114	0.259	-0.244
12/24/15 5:37	29	-0.171	0.383	-0.341
12/24/15 9:03	29	-0.158	0.299	-0.535
12/24/15 10:38	17	-0.082	0.405	-0.376
12/24/15 11:50	27	-0.102	0.359	-0.364
12/24/15 13:35	38	-0.275	0.425	-0.615
12/24/15 16:16	33	-0.127	0.303	-0.329
12/24/15 20:53	37	-0.072	0.285	-0.424
12/24/15 22:35	38	-0.138	0.309	-0.351
12/24/15 23:52	37	-0.093	0.424	-0.453

Table A.1 continued from previous page

12/25/15 0:43	27	-0.103	0.327	-0.23
12/25/15 1:50	25	-0.093	0.371	-0.141
12/25/15 3:30	26	-0.088	0.467	-0.36
12/25/15 12:50	27	-0.048	0.264	-1.24
12/25/15 15:47	30	-0.063	0.183	-0.232
12/25/15 16:50	26	-0.087	0.381	-0.275
12/25/15 18:41	21	-0.1	0.358	-0.371
12/25/15 20:20	39	-0.081	0.25	-0.284
12/25/15 21:43	19	-0.172	0.291	-0.046
12/25/15 22:08	12	-0.054	0.307	-0.389
12/26/15 1:18	17	-0.15	0.184	-0.164
12/26/15 5:14	31	-0.045	0.0898	-0.1243
12/26/15 8:20	34	-0.0502	0.1319	-0.1155
12/26/15 9:34	23	-0.0829	0.0981	-0.143
12/26/15 15:23	7	-0.112	0.302	-0.162
12/26/15 19:27	38	-0.061	0.1165	-0.1295
12/26/15 20:39	37	-0.041	0.162	-0.146
12/27/15 0:08	27	-0.06	0.136	-0.14
12/27/15 9:10	25	-0.067	0.284	-0.218
12/27/15 11:46	32	-0.072	0.078	-0.206

Table A.1 continued from previous page

12/27/15 12:44	17	-0.0091	0.1092	-0.1347
12/27/15 15:01	38	-0.0319	0.0661	-0.1262
12/27/15 19:30	23	-0.019	0.24	-0.173
12/28/15 0:30	32	-0.044	0.229	-0.175
12/28/15 3:20	8	-0.039	0.218	-0.191
12/28/15 16:11	23	-0.08	0.179	-0.087
12/28/15 16:39	10	-0.083	0.197	-0.188
12/28/15 22:53	26	-0.066	0.238	-0.13
12/29/15 0:30	32	-0.0677	0.1485	-0.1205
12/29/15 11:39	39	-0.0594	0.0964	-0.1417
12/30/15 5:50	9	-0.059	0.251	-0.09
12/30/15 10:04	25	-0.076	0.312	-0.227
12/30/15 11:07	29	-0.0661	0.12	-0.0708
12/31/15 20:50	37	-0.166	0.392	-0.452
1/2/16 15:04	38	-0.0604	0.0848	-0.0864
1/2/16 17:40	30	-0.0414	0.1303	-0.0757
1/2/16 22:14	36	-0.0331	0.0812	-0.0618
1/11/16 19:35	29	-0.0333	0.1183	-0.0771

Appendix B

Groups from Unsupervised Classification

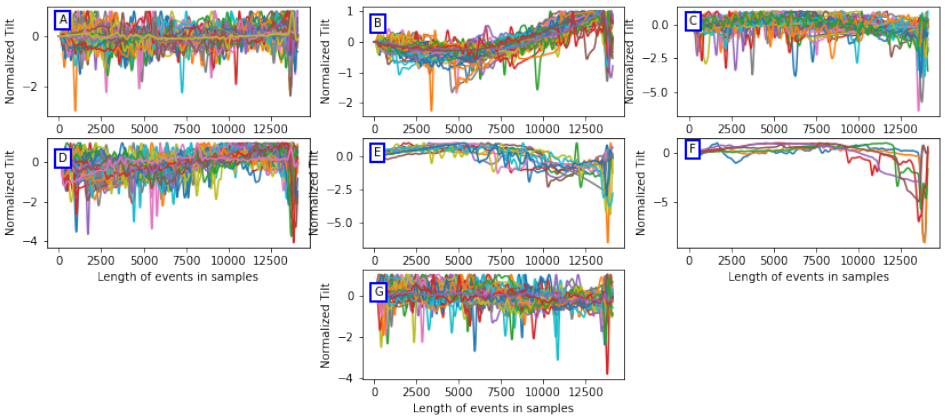

Figure B.1: The figures show groups obtained through unsupervised classification of 2015 data. The group B shows events of pattern 1 whereas the groups E and F consist of events of pattern 2 mixed with noise.

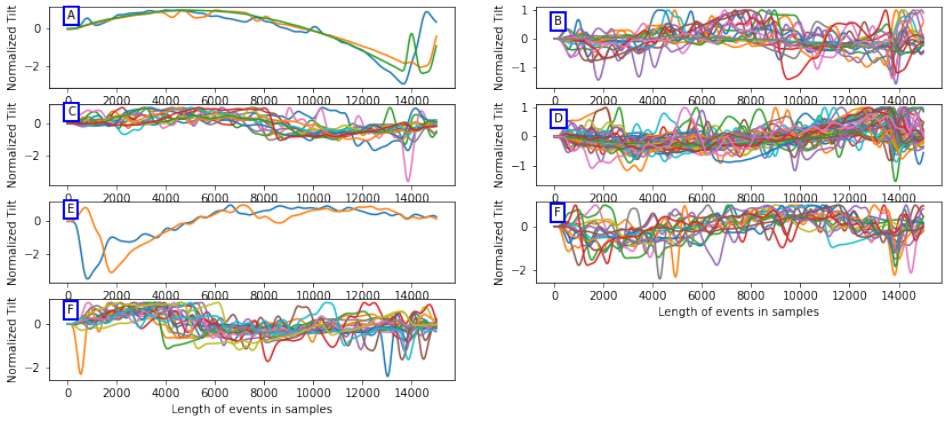

Figure B.2: The figures show groups obtained through unsupervised classification of 2012 data. The group D contains events similar to that of pattern 1 although exhibiting more noisy signals compared to other years under study. The events in group A contains events similar to that of pattern 2.