Comprehensive study of the volcanic seismicity of the Klyuchevskoy volcano group

Nataliya Galina

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : STEP - Sciences de la Terre de l'Environnement et des Planètes
Spécialité : Sciences de la Terre et de l'Environnement
Unité de recherche : Institut des Sciences de la Terre

Etude approfondie de sismicité volcanique dans le groupe volcanique du Klyuchevskoy

Comprehensive study of the volcanic seismicity of the Klyuchevskoy volcano group

Présentée par :

Nataliya GALINA

Direction de thèse :

Nikolai SHAPIRO DIRECTEUR DE RECHERCHE, Université Grenoble Alpes	Directeur de thèse

Rapporteurs :

Zacharie DUPUTEL
CHARGE DE RECHERCHE HDR, CNRS DELEGATION ALSACE
Vadim LEVIN
FULL PROFESSOR, Rutgers, The State University New Jersey

Thèse soutenue publiquement le **14 novembre 2023**, devant le jury composé de :

Nikolai SHAPIRO DIRECTEUR DE RECHERCHE, CNRS DELEGATION ALPES	Directeur de thèse
Zacharie DUPUTEL CHARGE DE RECHERCHE HDR, CNRS DELEGATION ALSACE	Rapporteur
Vadim LEVIN FULL PROFESSOR, Rutgers, The State University New Jersey	Rapporteur
Diane RIVET PHYSICIENNE ADJOINTE, UNIVERSITE COTE D'AZUR	Examinatrice
Virginie PINEL DIRECTRICE DE RECHERCHE, IRD DELEGATION REGIONALE SUD EST	Examinatrice
Stéphane GARAMBOIS PROFESSEUR DES UNIVERSITES, UNIVERSITE GRENOBLE ALPES	Président

To my father,
Alexander V. Galin
1964–2021

Abstract

Nowadays there are various tools and techniques to monitor volcanoes in order to predict their unrest and eruptions. One of the most effective approaches is based on interpretation of seismic observations. Volcanoes are very dynamic systems hosting a whole variety of processes reflected in diverse seismic signals recorded by seismic stations. Analysis and interpretation of these signals is the principal task of seismo-volcanic monitoring.

This work is aimed at investigating the seismicity within the Klyuchevskoy volcano group (KVG) in Kamchatka. The KVG is one of the world's largest and most active subduction zone volcanic clusters whose aboundant seismo-volcanic activity make it a great "natural laboratory" to study and advance volcano seismology.

In the first part, we study in details the deep long period (DLP) earthquakes that systematically occur near the crust-mantle boundary beneath the Klyuchevskoy volcano. This seismicity is believed to be one of the earliest manifestation of the volcanic unrest. However, the source mechanism of the DLP earthquakes remains poorly understood. To start with, the statistical analysis of DLP events was performed. This investigation was done using a detailed catalog obtained after processing almost two years of continuous data with the sensitive template matching algorithm. At the next step, we tried to reconstruct the source of DLP earthquakes using comparison of S-to-P waves amplitudes ratios instead the full waveform or polarity inversions. As the main result of the first part, it was shown that the source mechanism of the DLP seismicity is different from the one of tectonic earthquakes. Also, the obtained observations turned out to be in agreement with previously reported connection of the deep magmatic reservoir of the KVG with its active volcanoes.

Second part of this thesis is dealing with methods of automatic signal analysis based on Machine Learning (ML) algorithms of classification and clustering. The ML exploration was started with the problem of separating two general classes of seismicity: tectonic and volcanic earthquakes. We show the results of events classification using the data recorded during the seismic crisis in 2018-2019. Different possible representations of signals and several ML algorithms were tested in order to obtain the best performance. Further, a possibility to identify various classes of volcanic seismicity from the simultaneously erupting volcanoes was considered. To do so, we clustered the seismic waveforms recorded during October-November 2022 when the KVG re-activated and the Shiveluch volcano has been active for several months. This work

allowed to distinguish a particular cluster of seismicity with the location unusual for the Shiveluch volcano.

The problems considered in this work are too broad and far from being solved within a thesis. However, the obtained results can contribute both in physical interpretation of observed signals and in development of the modern methods of volcano monitoring.

Résumé

De nos jours, il existe divers outils et techniques pour surveiller les volcans afin de prédire leurs manifestations de mécontentement et leurs éruptions. L'une des approches les plus efficaces repose sur l'interprétation des observations sismiques. Les volcans sont des systèmes très dynamiques abritant une grande variété de processus reflétés dans divers signaux sismiques enregistrés par des stations sismiques. L'analyse et l'interprétation de ces signaux constituent la tâche principale de la surveillance sismovolcanique.

Ce travail vise à étudier la sismicité au sein du groupe de volcans Klyuchevskoy (KVG) au Kamtchatka. Le KVG est l'un des plus grands et des plus actifs groupes de volcans en zone de subduction au monde, dont l'abondante activité sismovolcanique en fait un excellent "laboratoire naturel" pour étudier et faire progresser la sismologie des volcans.

Dans la première partie, nous étudions en détail les séismes profonds de longue période (DLP) qui se produisent systématiquement près de la limite croûte-manteau sous le volcan Klyuchevskoy. On pense que cette sismicité est l'une des premières manifestations du mécontentement volcanique. Cependant, le mécanisme source des séismes DLP reste mal compris. Pour commencer, une analyse statistique des événements DLP a été réalisée. Cette enquête a été menée à l'aide d'un catalogue détaillé obtenu après le traitement de près de deux ans de données continues avec l'algorithme sensible de correspondance de modèles. À l'étape suivante, nous avons essayé de reconstruire la source des séismes DLP en comparant les rapports d'amplitudes des ondes S-P plutôt que les inversions de forme d'onde complètes ou de polarité. En tant que principal résultat de la première partie, il a été montré que le mécanisme source de la sismicité DLP est différent de celui des séismes tectoniques. De plus, les observations obtenues se sont révélées être en accord avec la connexion précédemment rapportée du réservoir magmatique profond du KVG avec ses volcans actifs.

La deuxième partie de cette thèse traite des méthodes d'analyse automatique des signaux basées sur les algorithmes d'apprentissage automatique (ML) de classification et de regroupement. L'exploration du ML a commencé par le problème de séparation de deux classes générales de sismicité : les séismes tectoniques et volcaniques. Nous montrons les résultats de la classification des événements à l'aide des données enregistrées lors de la crise sismique de 2018-2019. Différentes représentations possibles des signaux et plusieurs algorithmes ML ont été testés afin d'obtenir les meilleures per-

formances. En outre, la possibilité d'identifier diverses classes de sismicité volcanique provenant de volcans entrant en éruption simultanément a été envisagée. Pour ce faire, nous avons regroupé les formes d'ondes sismiques enregistrées en octobre-novembre 2022, lorsque le KVG s'est réactivé et que le volcan Shiveluch est actif depuis plusieurs mois. Ce travail a permis de distinguer un cluster particulier de sismicité avec une localisation inhabituelle pour le volcan Shiveluch.

Les problèmes abordés dans ce travail sont trop vastes et loin d'être résolus dans le cadre d'une thèse. Cependant, les résultats obtenus peuvent contribuer à l'interprétation physique des signaux observés et au développement des méthodes modernes de surveillance des volcans.

Remerciements

The three and a half years it took to complete this thesis were the most difficult of my life, and not only because of the thesis itself. The experience I gained is mind changing in both professional and personal ways, and here is the moment to thank everybody who supported me through this stage of my life.

First of all I would like to thank my thesis supervisor, Nikolai, with whom we had been working for almost six years. It was incredible luck to start conducting the research in such collaboration, and I am very grateful that after my master's thesis I was invited to pursue the work as a graduate student. Thank you for believing in me and giving an incredible chance by bringing to Grenoble. And I cannot be more thankful for all the support and patience provided.

I have never seen people thanking their psychotherapist but this is exactly the case. Marina, I am happy I met you one day, thank you for literally saving my life.

Further, special thanks go to Tatiana. It was really nice to have a Russian buddy (with the shared past!) particularly given the background of these years. Thank you for all the adventures we had and returning me some taste and interest to life.

Also I would like to note Martina, my dear flatmate with whom we happily had been sharing a great apartment for quite a long time. Many thanks for our rare but profound discussions, I am happy to know you and all the people you introduced me to.

It would impossible to mention all *mes chers compatriots* here, but be sure that I will save the best memories about you and I infinitely appreciate the *moments de convivialité* we had in the lab and beyond it. Again, that was the mind changing experience.

I met an incredible amount of people during these years, and here are some final words of gratitude for the kindness and support which sometimes found me at very unexpected moments.

Спасибо!

Contents

Chapter 1

Introduction

Volcanoes are a vivid manifestation of the processes occurring in the Earth. Their eruptions are impressing and captivating, yet volcanoes impose tremendous hazard on the humanity: nowadays hundreds of millions of people live in close proximity to active volcanoes. The encouraging fact is that, unlike tectonic earthquakes, volcanic eruptions can be predicted, and we already know successful cases of forecasting [McNutt et al., 2015]. Thus, volcano monitoring is a task of the essential importance performing which will allow to answer several questions. To start with, we are interested in the temporal characteristics of an eruption: when will it start, how long will it last and when will it end? Furthermore, we would like to estimate the power of potential event and its form: for instance, will it be an effusive eruption or an explosive one? Finally, the volcanic risk for the population and infrastructure, potential economic losses, etc should be assessed as well. The mentioned problems are fairly general and broad, and while monitoring is being performed, the other fundamental questions should be addressed in order to understand the volcanoes behaviour better.

Nowadays there are various tools and techniques to monitor volcanoes: ground deformation measurements, gravimetry, satellite interferometry, gas emissions monitoring, etc. One of the most effective approaches is based on interpretation of seismic observations. Volcanoes are very dynamic systems hosting a whole variety of processes within their plumbing systems and edifices. This volcanic activity reflects itself in diverse seismic signals recorded by seismic stations in vicinity of volcanoes. Analysis and interpretation of these signals in order to understand a current state of a volcano and its magmatic system and, when possible, to predict its changes leading to eruptions is the principal task of seismo-volcanic monitoring.

The goal of this thesis was to test different approaches for the analysis of seismo-volcanic signals and to apply them to the data collected in the region of the Klyuchevskoy volcanic group in Kamchatka (Russia) that is known for its exceptional level and diversity of volcanic activity and provides very rich data for advancing volcano seismology.

1.1 Anatomy of volcanoes

The main goal of this thesis is a comprehensive study of deep long period earthquakes in order to reconstruct their physical mechanism. Nevertheless, before commencing a discussion on this specific subject a general image of volcanoes, their structure and driving forces should be introduced.

1.1.1 Magma generation

Current subsection expounds magma origin as this process is a crucial condition of volcanoes formation. *Magma* itself is a mixture containing the following components: *melt* (liquefied minerals), *solids* (crystals floating in the melt), *volatiles* (gases such as CO_2, SO_2 and H_2O vapor as well). A particular combination and a relative content of listed elements define the magma nature, and a volcano eruption style, consequently.

Two important definitions should be introduced here. They are *solidus*, the temperature of melting commencement, and *liquidus*, the temperature at which partially molten material transits into a liquid state (Figure 1.1). Thus, generally speaking magma is produced by mantle rocks melting but it can be done within different processes.

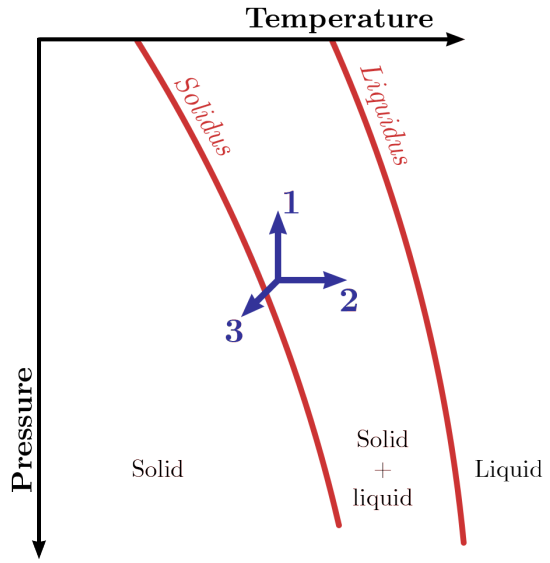

Figure 1.1: A schematic pressure-temperature diagram with solidus and liquidus temperatures (red lines). Blue arrows show possible ways of mantle melting by:
(1) pressure decrease, (2) temperature increase, and (3) changing the chemical composition. The image is modified after [Grove and Till, 2015]

Decompression melting

Convection occurring in the mantle generate the flow of the asthenosphere underlying the lithosphere. As the Earth's crust has low heat conductivity, there is no heat release and no temperature change therefore. This system is considered *adiabatically* isolated or, in other words, at this moment *adiabatic cooling* occurs. In this situation

the only way for the system to change its internal energy is to lower the pressure. Thus, a volume of hot mantle rising from the depths with higher temperatures finds itself in the area of lower pressure compared to the starting point. Under the resulting circumstances the material passes its melting point what means the start of melting process which leads to the more magma rising.

Heat-induced melting

This type of melting can be observed at *mantle plums*, anomalously hot columns of rock that ascend from the depthsc supplement the surrounding mantle with heat energy causing its temperature increase. As a result molten rocks start to rise and break through the lithosphere.

Flux melting

It is a process drastically different from the previous ones as it occurs due to changing composition of the mantle matter. The most common situation is when rocks become enriched with volatiles, for example H_2O. Melting occurs due to the lower solidus temperature of the new composition which remains in the same temperature conditions starts to melt and ascend.

1.1.2 Volcanism in different tectonic settings

Diverse topographic settings and corresponding processes, such as earthquakes and volcanism, are explained within the *plate tectonics theory* stating that the lithosphere, i.e. the outer Earth's layer, consists of several plates which move relatively to each other.

Thus, both tectonic and volcanic activities are concentrated in the regions of plates boundaries (Figure 1.2) which are generally separated into three groups. The first one is called *transform boundaries* where plates move along each other (Figure 1.3). Neither crust formation nor its destruction can be observed there, which explains another name for this type of boundaries, conservative. In such conditions, there are no triggers for magma generation and thus volcanoes do not form in these regions.

The case of plates moving towards each other is called *convergent* or *destructive boundaries*. Within this group one can distinguish two subgroups: collision and subduction. *Collisions* are formed by continental plates of approximately the same thickness and density, so there is no magma formation and no volcanism is possible. And on the contrary, the *subduction* zones are characterised by strong and sustained volcanic activity. If subduction is formed by a continental and an oceanic plates, the

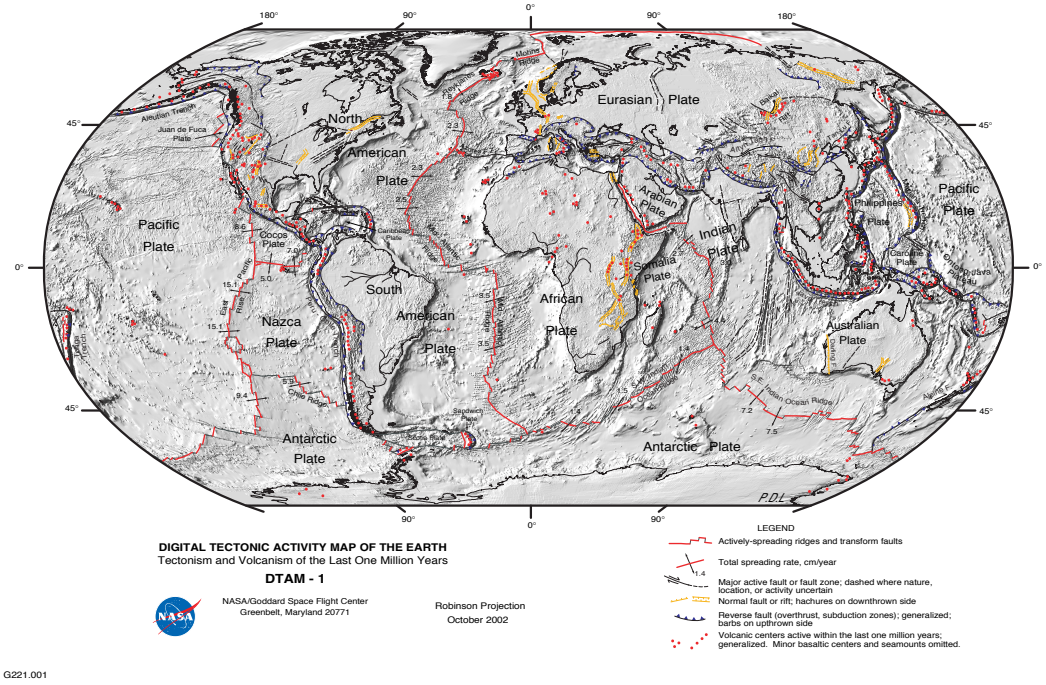

Figure 1.2: Map of tectonic plates. Colored lines represent different fault types, red dots show the volcanic centers. The image is the courtesy of NASA [Lowman et al., 1999]

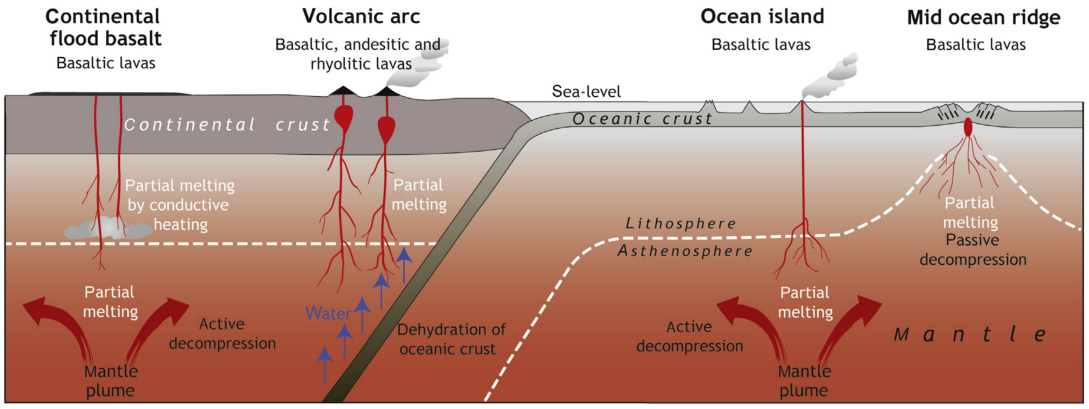

Figure 1.3: A cartoon of volcanism in different tectonic settings showing the crust, mantle, and lithosphere and corresponding melting mechanisms. The figure is adopted from [Rogers, 2015]

latter one which is colder and denser starts to dip beneath a continental plate. Subducting H_2O saturated oceanic lithosphere releases volatiles into the mantle causing flux melting described in the subsection 1.1.1.

Consequently, magma starts to rise through the denser continental lithosphere towards the surface and results in volcanism. Usually, long volcanic chains are formed

in these regions, and the Klyuchevskoy volcano group studied in this work as a part of the Eastern Kamchatkan range is an example of subduction volcanism. In case of two oceanic plates converging, one of them starts to dip in the same way as in continental-oceanic pair, forming an island arc.

The last option of tectonic plate movement is the case of moving them away from each other. This process is accompanied by formation of young crust therefore these plate boundaries are called *constructive*. Starting with a continental plate diverging followed by thinning and final splitting of it into separate plates, the continental lithosphere of a former rift zone transforms into the oceanic lithosphere. Thus, the process of plates diverging is not over, mid-ocean ridges are formed.

Volcanism described above is called *inter-plate* as in this case volcanoes emerge on the boundaries of tectonic plates. Besides it there is *intra-plate* volcanism formed by mantle plumes, anomalously hot columns of rock, resulting in magma rising and breaking through the lithosphere. Built volcanic edifices can be observed within both continental and oceanic lithospheres.

1.1.3 Plumbing systems and magmatic chambers

Probably when speaking about igneous systems of volcanoes the first association that comes to mind is a picture of an enormous reservoir filled with completely melted magma which flows through a well-defined conduit towards the surface. This paradigm was prevailing in the last century, and many investigations were carried out based on it. Yet, accumulating knowledge revealed inconsistencies between observations and this primitive model of a magmatic chamber. To start with, the existence and maintenance of such vast volumes of melt could be hardly explained by physical principles. Investigations in seismic tomography using local seismic data do not find evidences for vast chambers filled with fluid magma. Moreover, the results of seismic tomography revealed that the amount of melt in the medium beneath volcanoes is only of a few percent [Lees, 2007].

Thus, with the paradigm shifting a new concept has arose. As magma starts to solidify, the proportion of suspended crystals and thus the relative viscosity of the magma increases until a relatively immobile, continuous network of crystals and interstitial melt develops. In this crystal-dominated medium, so-called *"mush"*, melt is kept either between crystals, or in small pockets [e.g., Hildreth, 1981, 2004, Marsh, 2015, Glazner et al., 2016, Cashman et al., 2017]. The rheological transition from a magma to a crystal mush is partly dependent on its chemistry, but typically occurs abruptly when the particle volume increases above the 50–65% range.

Although the definition of mush was supposed decades ago [Marsh, 1981, Bach-

mann and Bergantz, 2004] it became widespread due to a new image of plumbing systems. Thus, within the modern concept magmatic chambers are represented by magma bodies such as sills or dykes when a plumbing system resembles a vertical mush column extended through the lithosphere [e.g., Annen et al., 2006, Cashman et al., 2017, Sparks et al., 2019]. With depth, the borders between components of this system, both physical and chemical, blur with the time. Comparison of old and new images of plumbing systems is presented in Figure 1.4.

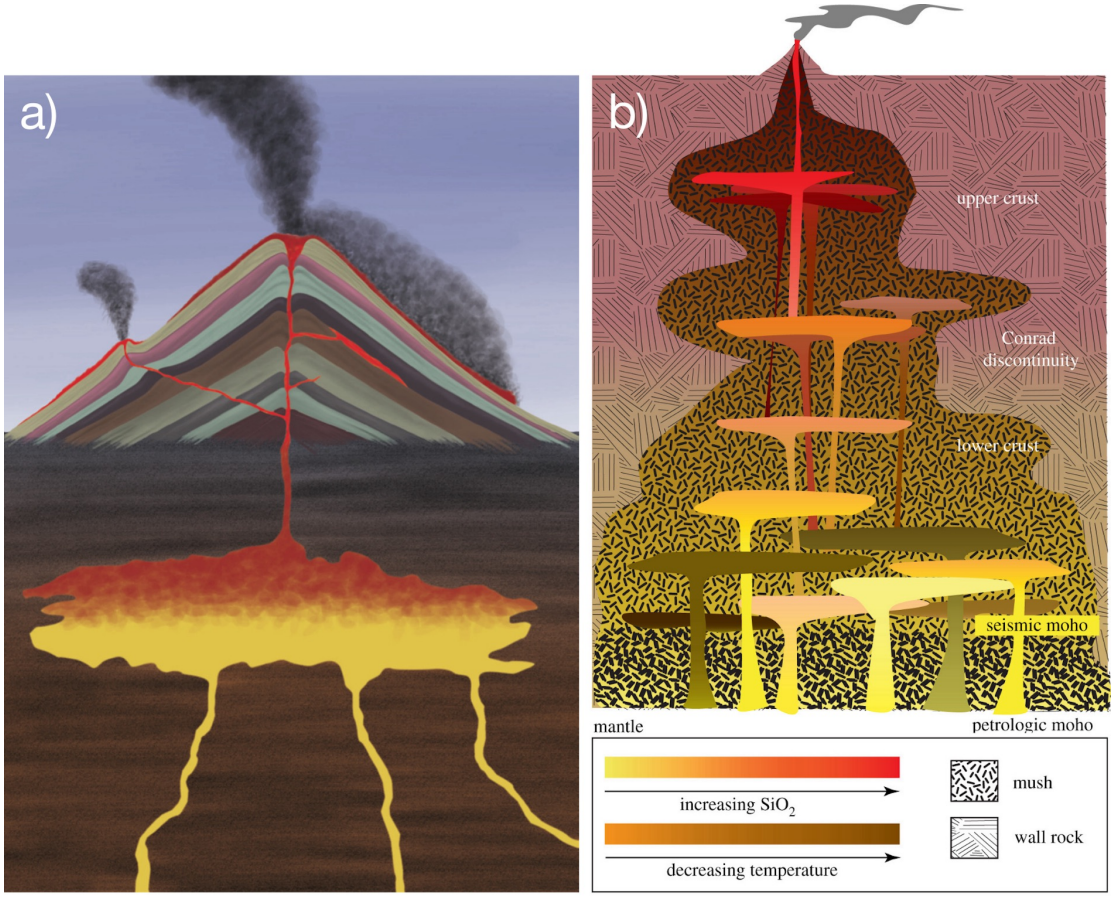

Figure 1.4: A cartoon demonstrating the old and the new concepts of volcanic plumbing systems: (a) a vast cavity filled with fully melted magma and (b) a mush column extended through the lithosphere with magma chambers represented by interconnected melt-rich lenses. The images are adopted from [Cashman et al., 2017] and [Sparks et al., 2019]

One more hypothesis regarding igneous systems of volcanoes is worth to be noted here. According to it magma is stored at "low" temperatures, i.e. below the solidus, and is mobilised during crisis period [Cooper and Kent, 2014]. This idea supports the observations that during some episodes of volcanic activation the erupted material was found to have been stored for hundreds of thousands of years. Thus, according

to [Cooper and Kent, 2014] a detection of a magma body with a high fraction of melt can signal an impending eruption.

1.2 Monitoring of volcanoes

Timely monitoring of volcanoes is a crucial task for understanding the processes within the igneous systems. In its turn, the advancement of our knowledge and solving the fundamental problems of volcanology inevitably leads to the more precise estimation of volcanic hazard and risks related to volcanic eruptions. Studies of volcanic and igneous plumbing systems (VIPS) bring together specialists of diverse disciplines (Figure 1.5).

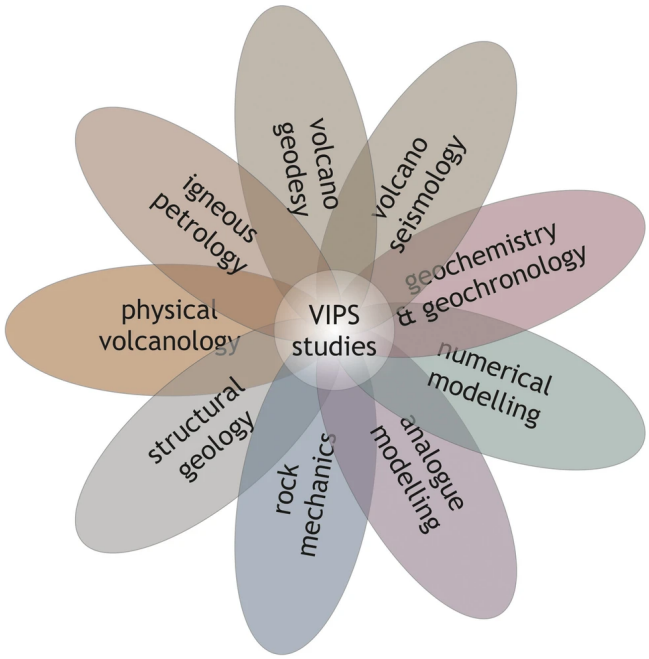

Figure 1.5: A diagram showing the various research fields and their overlaping as a comprehensive study of VIPS. The image is adopted from [Burchardt et al., 2022]

Data obtained within the framework of one or another monitoring method can be affected (for example, clouds or snow provide an impact on InSAR interferometry) or be unavailable at all (destruction of a seismic station, etc). The combination of various observations gives a more reliable picture of the state of a volcano and allows to conduct continuous monitoring and short term forecasting.

1.2.1 Seismovolcanic monitoring

Although monitoring of volcanoes includes measurements of various parameters such as ground deformations, gas emissions, heat transfer and so on, seismic investigations are likely the most important ones. Volcano seismology appeared in the last

century almost at the same time with the development and wide spreading of seismic networks. Since then, volcanic earthquakes and tremors as well as corresponding changes in their occurrence, rate and other parameters were noticed accompanying the unrest and eruptive stage of volcanoes in various regions and environments with different edifice structures and magma contents [e.g., Chouet, 2003, Chouet and Matoza, 2013]. Thus, seismic activity has shown itself to be a reliable precursor of forthcoming eruptions, [e.g., Shapiro et al., 2017a]. Nevertheless, as it was mentioned above, forecasting should be performed using various observations, and an example of a successful prediction based on diverse data is an eruption of the Soufriere Hills Volcano, Montserrat in 1995–1998, preceded by swarms of volcanic earthquakes started in 1992 [Sparks and Young, 2002].

The general definition of *seismology* can be given in the next form:

> *Studies of generation and propagation of mechanical waves within planets and stars to understand their internal structure and dynamics and to develop predictive models.*

Within this scientific field data is represented by *seismograms*, i.e. recordings of the ground motion as a function of time. Thus, one can define the main goal of seismology as obtaining inferences from seismograms. Given definitions are easily applicable to volcano seismology taking into account the fact that corresponding measurements and observations are done in the vicinity of volcanoes. It should be noted, that seismic stations deployed in volcanic regions register signals not only of volcanic origin but tectonic earthquakes of different scales (local, regional, teleseismic), ambient and anthropogenic noise, etc. Thus, the signal (displacement u) recorded by a station in the vicinity of a volcano can be expressed as:

$$u = u^{amb} + u^{ant} + u^{tect} + u^{volc} + u^{inst} \tag{1.1}$$

where u^{amb} and u^{ant} are the natural and the anthropogenic components of the ambient noise wavefield, respectively, u^{tect} are signals generated by tectonic sources not related to the volcano-magmatic activity, u^{volc} is the wavefield generated by different types of seismo-volcanic sources, and u^{inst} is the instrumental noise. Consequently, while performing research a scientist should solve the problem of defining and extracting the data useful in a particular investigation and corresponding to a phenomenon of interest.

So far, the seismovolcanic analysis starts with the "detection" step aimed at separating the "useful" part u^{volc} from other contributions (ambient, tectonic, instrumental). This task is not a fully trivial because different contributions in equation (1.1) can be act simultaneously and overlap in frequencies.

Then, the most important part of volcano-seismological analysis is interpreting u^{volc} in order to infer information about the state of the volcano and processes occurring in the volcano-plumbing system to eventually predict its evolution. In the most general case, a time (t) dependent displacement $u_i^{volc}(\mathbf{r}_r, t)$ associated with seismo-volcanic sources and recorded on the component i of a receiver located in $\mathbf{r}_r$ (the subscript r stands for receiver) can be written as [Kumagai, 2009]:

$$u_i^{volc}(\mathbf{r}_r, t) = \int_{-\infty}^{+\infty} \iiint_V f_j^{volc}(\mathbf{r}, \tau)\ G_{i,j}(\mathbf{r}_r, t-\tau, \mathbf{r}, 0)\ dV(\mathbf{r})\ d\tau \tag{1.2}$$

where $f_j^{volc}(\mathbf{r}, \tau)$ is a time and space dependent equivalent body force (j-component) and $G_{i,j}(\mathbf{r}_r, t-\tau, \mathbf{r}, 0)$ is the Green's function, i.e., the mathematical representation of the seismic wave propagation from the source to the station.

In a simplified case of a few seismic sources generated by different processes within a volcano and remaining stationary in space within the considered time window, this equation can be simplified as:

$$u^{volc}(\mathbf{r}_r, t) = \sum_k \int_{-\infty}^{+\infty} S^k(\mathbf{r}_s^k, \tau)\ \tilde{G}(\mathbf{r}_r, t-\tau, \mathbf{r}_s^k, 0)\ d\tau \tag{1.3}$$

where k indexes different acting tremor sources located in positions $\mathbf{r}_s^k$, $S^k(\mathbf{r}_s^k, \tau)$ is a generic representation of a source k time function that can be either a force vector or a moment tensor, and $\tilde{G}(\mathbf{r}_r, t-\tau, \mathbf{r}_s^k, 0)$ represents either the Green's function or its spatial derivative.

Equation 1.3 explains the complexity of seismovolcanic signals with their very variable time signature and frequency content [e.g., Kumagai, 2009, Chouet and Matoza, 2013] that is illustrated in Figure 1.6 showing one day of continuous data recorded by station KMN located in the region of the Klyuchevskoy volcanic group in Kamchatka (Russia). Multiple seismic events with very variable properties are observed and likely correspond to different sources (sometimes repetitive) related to different physical processes.

Ideally, we would like to achieve a full physical interpretation of the observed seismic signals. This would imply to identify all sources contributing to equation 1.3, to locate them, to characterize their source mechanisms and source time functions ($\mathbf{r}_s^k$, $S^k(\mathbf{r}_s^k, \tau)$), and finally to find a physical relationship between the emitted seismic waves and the underlying processes. Such a full physical interpretation is, however, far from being achieved within the modern state of the art in volcano seismology because of the multiple reasons, i.e., strong complexity of observed seismograms that contain sometimes thousands of events per day, intricacy of wave propagation or poor knowledge of the internal volcanic structure required to describe it (Green's functions:

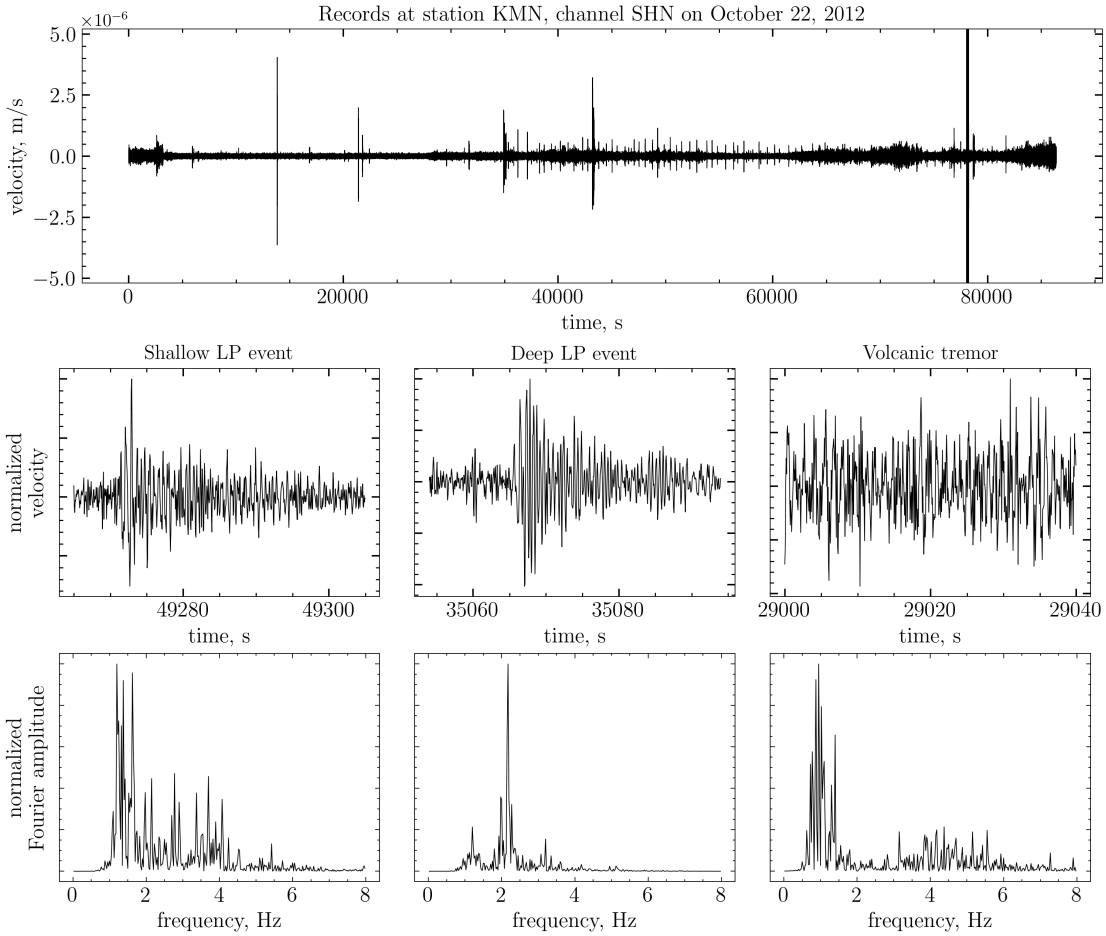

Figure 1.6: One day of seismic data recorded at station KMN located in the KVG region (top) and various signals and their spectra produced by different volcanic processes (bottom)

$\tilde{G}(\mathbf{r}_r, t-\tau, \mathbf{r}_s^k, 0)$), insufficient understanding of the physics of the source of most of the seismo-volcanic phenomena, etc.

Considering the impossibility of the full physical interpretation, empirical approaches are most often used to link the seismovolcanic observations with the volcanic unrest. Simplest methods are based on some approximate estimations of the level of seismic activity such as the earthquake occurrence rate or average level of the recorded signal estimated, for example, with the "Real-time Seismic Amplitude Measurement" (RSAM) [Endo and Murray, 1991]. More advanced methods of interpretation of seismovolcanic signals require classifying in several distinct categories.

1.2.2 Evolution of the volcanic seismicity classification and its present-day state

Seismic signals recorded at volcanoes are numerous and correspond to different processes. The first attempt to classify signals from volcanic earthquakes is described in [Minakami, 1960], where the author singled out the following groups of seismic signals based on the shape of waveforms and the hypocenters locations:

- A-type: earthquakes occurring at $1 - 10$ km depth beneath volcanic edifices, characterized by high frequencies of about 10 Hz and being observed before eruptions;
- B-type: shallow earthquakes with depths up to 1 km in the vicinity of craters with relatively low characteristic frequencies of $1 - 3$ Hz;
- explosive earthquakes: events accompanying explosions in volcanic craters, they have frequency composition similar to that of the B-type, but with a higher intensity;
- volcanic tremor: continuous process observed during lava or gas-ash flows, the frequency composition is identical to the composition of explosive events.

At the same time another classification was developed independently in Kamchatka, Russia [Tokarev, 1966]. According to it, volcanic events can be divided into 5 types (I–V), where significant attention is paid to the ratio of body to surface waves in the record and characteristic frequencies of these waves. In general, this classification is similar to that described in [Minakami, 1960], and volcanic events of types IV and V correspond to explosive earthquakes and volcanic tremors, respectively.

The classification problem has been puzzling scientists for many years, and nowadays volcanic seismicity is usually divided into two general groups: those occurring in fluid and those occurring in solid medium [Gordeev, 1998]. The latter one are called *volcano-tectonic* (VT) earthquakes and associated with the brittle failure of rocks in volcanic edifices. These events act as sensors that highlight the stress concentrations distributed throughout the volume of magma reservoirs. Figure 1.7a shows an example of VT earthquake waveforms which are characterised by bright and clear onsets of both P- and S-waves. These signals also have broad spectra with energy almost uniformly distributed over the entire range of frequencies. Although the source mechanism of VTs is similar to the one of regular earthquakes, volcanic events tend to occur in swarms rather than in mainshock-aftershock sequence [McNutt and Roman, 2015].

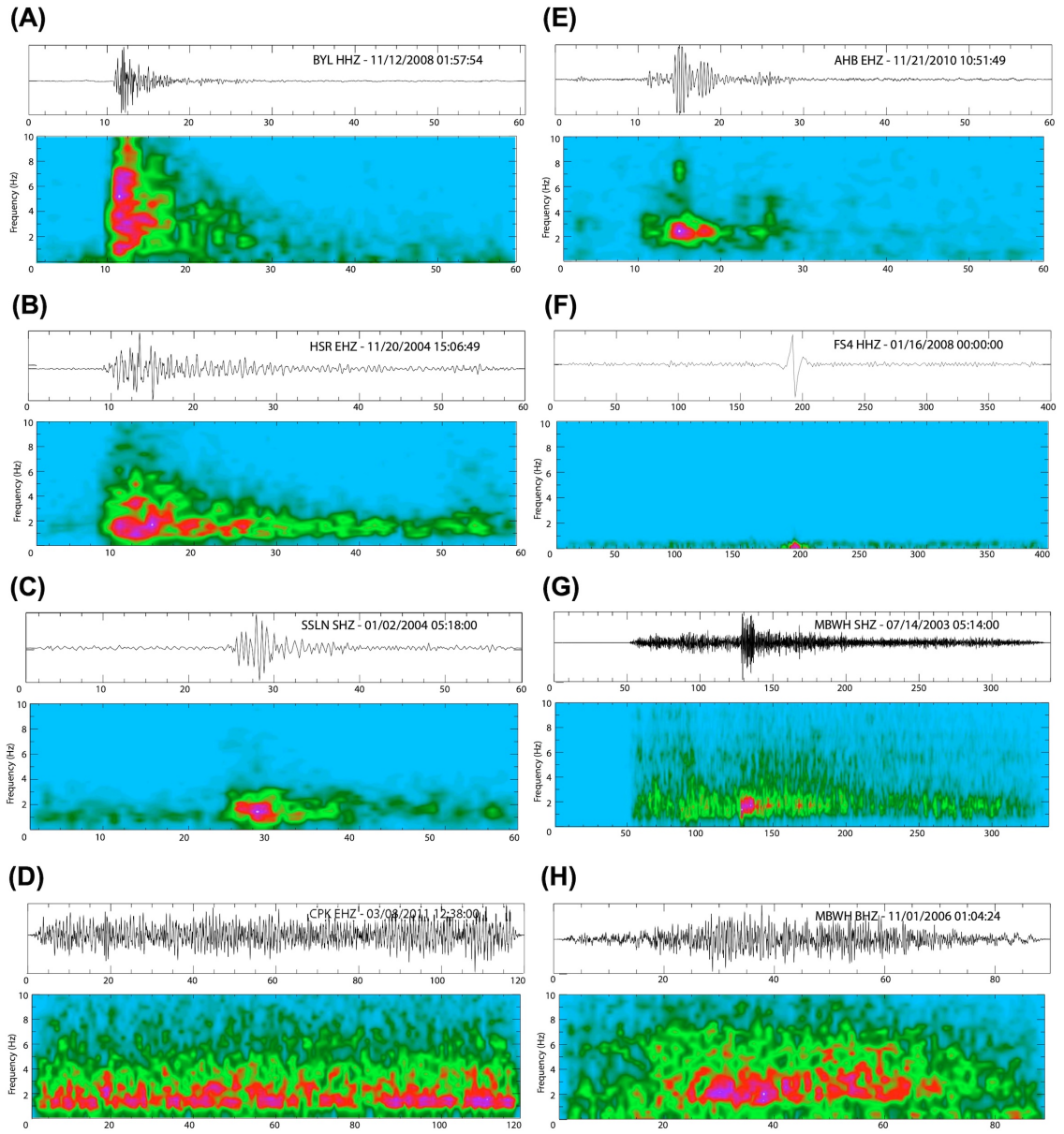

Figure 1.7: Examples of different signals of volcanic origin and their spectrograms: (a) Volcano-tectonic earthquake (Kilauea volcano, Hawaii). (b) Hybrid earthquake (Mount St Helens, Washington). (c) Long-period earthquake (Shishaldin volcano, Alaska). (d) Volcanic tremor (Kilauea volcano, Hawaii). (e) Deep long-period earthquake (Akutan volcano, Alaska) (f) Very long-period earthquake (Fuego volcano, Guatemala). (g) Explosion event (Soufriere Hills volcano, Montserrat). (h) Rockfall event (Soufriere Hills volcano, Montserrat). The images and the description are adopted from [McNutt and Roman, 2015]

The second big class of volcano seismicity called *long-period* (LP) is the result of unstable mass transfer and thermodynamics of fluid [Chouet, 1996b]. Their nature,

in particular the frequency of pressure fluctuations, provides direct information about the state of the fluid. This broad class includes a large variety of signals lasting from several seconds to days or even months (volcanic tremor).

Long-period earthquakes

As a rule, LP signals consist of emerging waves onsets followed by a harmonic coda (Figure 1.7c). Another feature of waveforms produced by LP events is their similarity within a recording channel that suggests a repeating excitation of a stationary source generating these earthquakes.

As it was described in previous sections, plumbing magmatic systems are likely to be sophisticated structures including magmatic bodies of different shapes and sizes through which magma propagates in order to reach the surface. Seismic signals of LP events as a result of this magma propagation often contain long lasting harmonic coda. This waveforms feature is commonly interpreted as resonance of fluid-filled bodies triggered by magmatic and hydrothermal activity. Thus, to reproduce a signal with its particular amplitude, dominating frequency and coda attenuation, one should reconstruct the shape and the characteristic size of the source as well as the parameters of the fluid filling the magmatic body. Also, a triggering source of mechanical energy is needed; this can be a small earthquake adjacent to a conduit, a flow transient or a pressure fluctuation within a conduit, gas bubbles expanding or contracting, a shock wave from choked flow, or other causes. Some researchers consider the source to be the mechanical energy alone, while others treat the source as the ensemble of the mechanical energy and the resonant response of the magma or water in conduits or dikes.

Hybrid earthquakes

The term hybrid was introduced for the first time by [Lahr et al., 1994] in order to describe earthquakes observed during the Redoubt volcano eruption. These events could not be assigned either to an LP class or a VT one as they combined the waveform characteristics of both of them. Thus, seismograms of hybrid events have a sharper and a clearer P-wave arrival than LP events but at the same time they have a harmonic coda atypical for VT (Figure 1.7b). Consequently, the general idea on the origin of hybrid earthquakes is that they are LP events triggered by shear faulting [Chouet and Matoza, 2013].

Volcanic tremor

As well as separate LP earthquakes, volcanic tremor is characterized by a harmonic waveform of a stable amplitude lasting from several minutes to days, and sometimes

several months or even longer. LP earthquakes and tremors differ in duration, while they have the same temporal and spectral components what suggests the existence of a common source underlying the two types of events [Chouet and Matoza, 2013]. Some authors [e.g., Latter, 1979, Fehler, 1983, Neuberg et al., 2000] propose an idea that volcanic tremor is the result of overlapping LP events, when their inter-event times become too short such that the waveforms are no longer distinguishable. Nevertheless, the controversial opinions can be found as well [e.g., Sciotto et al., 2022].

1.2.3 Limitations of classification schemes

The conventional approach of classification and further interpretation of volcanic seismicity is rather conditional and subjective. Perhaps, it had been reasonable when this part of volcano monitoring was at its birth stage, and instrumental and computational abilities were limited. Moreover, the assignment of an earthquake to one or another seismicity group is usually performed either manually by operators or automatically but using parameters defined from a human perspective. Also, generally there is no underlying physics and understanding of ongoing processes in a volcano, only visual characteristics. After all, usually no analysis of the obtained catalog is performed. This problem is especially acute in the case of volcanic tremors.

Some scientists have already pointed out the idea of a more sophisticated classification approach not limited by several groups of signals. For example, in [Neuberg et al., 2000] the results of spectral analysis of LP events observed at Montserrat are presented. The authors exceptionally note on the variety of spectral characteristics of signals, and reveal that this observation is due to a source effect and not a path one. Nevertheless, they stay in terms of two classes of seismicity, LP and hybrids, as end-members of the range with the other events falling in between.

In particular, ideas on improvement of the classification approach are widely spread in the works dedicated to the application of machine learning in seismology. There are two types of machine learning algorithms. The first class called “supervised learning” defines non-linear relationships between observations and labels (indicating that they belong to a particular class). Supervised approaches can be useful in detection and classification of already known classes of signals. However, these algorithms need to be “trained” on “labeled” datasets and strongly depend on the quality of the predefined labels, which in the case of volcano seismology are estimated in the obsolete manner. Therefore, they cannot be used for discovering new classes of seismic signals. To deal with this task some researchers call for the use of “unsupervised machine learning” [Bardainne et al., 2006, Köhler et al., 2009, Seydoux et al., 2020]. This type of algorithms recover patterns and structure of unlabeled data free from the bias of the

researchers perspective.

1.3 Motivations and content of this thesis

This work is aimed at investigating the seismicity beneath the Klyuchevskoy volcano group (KVG) in Kamchatka. The KVG is one of the world's largest and most active subduction zone volcanic clusters which comprises 13 closely spaced stratovolcanoes and covers a region with a diameter of 70 km. The extraordinary volcanic activity of this region is associated with the tectonic setting of the KVG (Figure 1.8) which is determined by the processes in the junction zone of the Kuril–Kamchatka and the Aleutian island arcs [Shapiro et al., 2017b]. This region is also host to the subduction of the Hawaiian–Emperor ridge, and the KVG in this framework is located above the edge of the sinking plate.

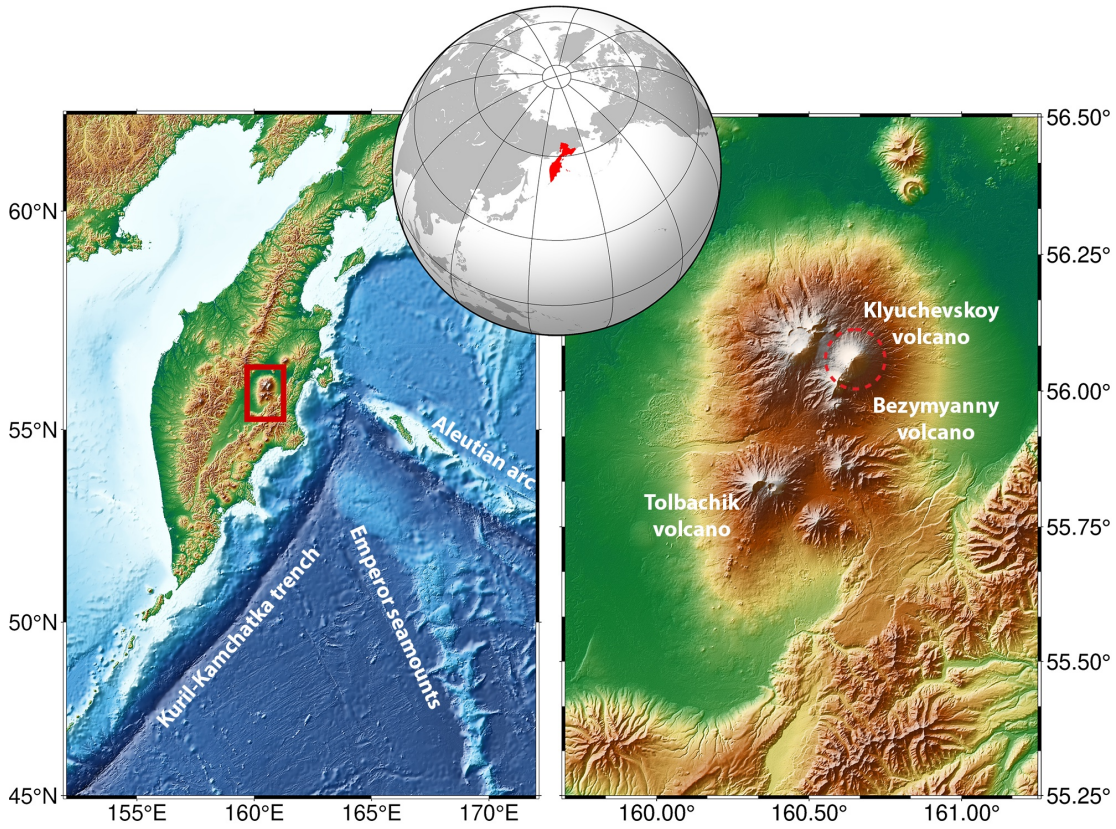

Figure 1.8: Map of Kamchatka peninsula with the main tectonic features and the KVG area shown with a red rectangle (left panel), and the detailed map of the KVG with the main active volcanoes (right panel), a red dotted line shows an approximate position of the cluster of DLP events (h = 30–35 km)

Thus, the geodynamic models seeking to explain the volcanic activity of this vol-

canic group are complex and incorporate many factors: fluid release from a thick and highly water-saturated crust of the Hawaiian–Emperor Ridge [Dorendorf et al., 2000], mantle flow around the edge of the Pacific Plate [Yogodzinski et al., 2001], or slab detachment from a subducting plate due to the recent structural rearrangement of the subduction beneath Kamchatka [Levin et al., 2002].

The feature of the KVG volcanoes is the different eruption types ranging from the Hawaiian type effusive eruptions as those during the two recent Tolbachik eruptions to the catastrophic explosive eruptions, for example, the Bezymyannyi eruption in 1956. Numerous eruptions and other volcanic manifestations generate an exceptional level of seismic activity including volcanic earthquakes of different types and tremors [Ozerov et al., 2007, Ivanov, 2008, Senyukov et al., 2009, Droznin et al., 2015] (Figure 1.6). The strength and variety of seismo-volcanic activity in the KVG region make it potentially one of the best "natural laboratories" in the world to study and advance volcano seismology.

Analysis of the seismograms from the KVG region is done together with the data of the whole Kamchatka seismic network by the operators of the Kamchatkan Branch of the Geophysical Survey of Russian Academy of Sciences (KB GS RAS). The data are processed in a "traditional" manner when all detections and travel time picks are verified "manually" by an operator. One of the problems of this approach is that it is not sufficient to efficiently monitor the enormous seismicity level of Kamchatka. So far, the KB GS RAS reports on average 20 regional and 90 volcanic (with only 30 of them located) earthquakes per day. Activity levels strongly increase during eruptions and aftershock sequences resulting in hundreds or sometimes thousands of events per day. Therefore, resulting catalogs of volcanic seismicity are often incomplete and lack details for advanced interpretation. Only basic parameters such as the average signal amplitude and the average earthquake occurrence rate (independent of their type) are routinely used for assessing the level of volcanic activity. Physical mechanisms linking most of the KVG seismic signals with the volcanic processes remain very poorly understood.

All this implies that additional research is needed to improve the seismic monitoring of the KVG volcanoes and to better explore the enormous potential of the observations collected in this region for better understanding of the seismo-volcanic processes. The main goal of this thesis was to contribute to such research and the work has been aimed both physical interpretation of some of the observed signals and development of the statistical analysis of the data with machine learning.

1.3.1 Part 1: Physical origin of deep long period volcanic earthquakes

Among various volcano-seismic phenomena we chose a particular class of seismicity, *deep long period earthquakes* (DLP), for a detailed inspection. This type of seismo-volcanic events is considered to be important because they are believed to represent one of the earliest manifestations of the volcanic unrest preceding the eruptions. Although the DLP earthquakes have been observed in various volcanic systems, their physical origin remains not fully understood. Part 1 includes Chapters 2 and 3.

Chapter 2 starts from the very basic but yet challenging step in volcano seismology: detecting and extracting signals of week DLP earthquakes recorded in noisy environment. In order to compile a seismic catalog in its traditional sense these waveforms should be quantified, and at this step a new approach of defining moment magnitude M_{W} is introduced as the classic technique of magnitude estimation applied in Kamchatka is not possible due to low signal-to-noise ratio of DLP earthquakes. The obtained detailed catalog is used further to verify if the DLP earthquakes satisfy the Gutenberg-Richter law, i.e. the empirical relationship between the number of earthquakes and their magnitude. After implementing the statistical investigation the deviation from the common distribution is found. This fact can point on a source mechanism different from the shear slip model. The results of this Chapter has been published as a **paper** in *Izvestiya, Physics of the Solid Earth* and also contributed to the development of a new physical model of the DLP earthquakes described in a co-authored paper [Melnik et al., 2020] (see Appendix).

In Chapter 3, we start to define the kinematic parameters of the source of DLP seismicity. To do so, various processes that can be observed in volcanoes, such as magma movement in a conduit, opening of a magmatic intrusion due to a rapid pressure change, etc., are considered. Thus, we take into account possible mass advection effects and add to the moment tensor a possible single force component [Takei and Kumazawa, 1994]. To constrain the source mechanism we use the fit to the observed amplitude ratios between P- and S-waves instead of performing a full inversion in the parameter space. This decision is taken because classic inversion procedure may be subject to a strong uncertainty. Alternatively, we proceed with testing a set of "plausible" hypotheses resulting in source mechanisms with a reduced number of parameters. Altogether, the first part of the thesis summarises the attempts to recover the physical mechanism of the DLP earthquakes and shows its clear difference from the classic model developed for regular tectonic seismicity. The results of this Chapter has been submitted as a **paper** in *Journal Volcanology and Geothermal Research.*

1.3.2 Part 2: Studying the seismicity in the vicinity of the Klyuchevskoy Volcanic Group with machine learning

The second part of this thesis is dedicated to the problem of classification of the seismic signals recorded in volcanic regions. Given the complexity and volumes of seismic data the machine learning (ML) approach is chosen. Chapter 4 gives an overview of the used ML methods.

In Chapter 5, we investigate if ML can help to separate volcanic and tectonic seismicity with using data only from one channel of a single station. In particular, the data recorded between December 2018 and March 2019 are analyzed. During this period, a long aftershock sequence of a strong M=7.3 Komandor Island earthquake was considered coincided in time with an eruption of Shiveluch and respective swarm of volcanic earthquakes. Various representations of seismic data were explored with both supervised and unsupervised algorithms.

In Chapter 6, the unsupervised ML is applied to identify different classes of seismicity during a time period when several volcanoes were active. The data from October-November 2022 when first Bezymianny and then Klyuchevskoy entered the activity during the ongoing eruption of Shiveluch are considered.

Part I

Physical origin of the deep long period volcanic earthquakes

Chapter 2

Recurrence of deep long period earthquakes beneath the Klyuchevskoy volcano group

Results presented in this Chapter have been published as a paper:

Galina, N.A., Shapiro, N.M., Droznin, D.V. et al. Recurrence of Deep Long-Period Earthquakes beneath the Klyuchevskoy Volcano Group, Kamchatka. Izv., Phys. Solid Earth 56, 749–761 (2020). https://doi.org/10.1134/S1069351320060026

2.1 Introduction

This thesis is dedicated to a very peculiar type of volcanic seismicity, namely deep long period earthquakes. The DLPs has been observed at various volcanic regions over the world. However, despite numerous studies of DLP seismicity there is no final solution on its generating mechanism. This problem and corresponding discussion is contained in detailed in Chapter 3. This chapter focuses on the statistical analysis of the DLP earthquakes.

Investigating the main parameters of tectonic earthquakes (magnitude, energy, seismic moment, source size and duration) with the statistical tools, it is possible to reveal the similarity laws reflecting the physical processes in the sources [e.g., Kanamori and Anderson, 1975]. Among the main statistical relationships characterizing the physics of seismic sources is the so called recurrence law which describes how the total number of the occurred earthquakes changes with the magnitude, i.e, well-known the Gutenberg–Richter frequency-magnitude relationship [Gutenberg and Richter, 1944]:

$$\lg N = a - b\,M \tag{2.1}$$

where N is the number of earthquakes with magnitude M; a is the parameter characterizing seismic activity; and parameter b referred to as the b-value is the slope of the frequency–magnitude graph. In fact, the b-value shows the proportion of large and small events in the catalogue. In case of regular tectonic earthquakes the b-value is very close to 1 [Frohlich and Davis, 1993]. This relationship is explained in the context

of a scale-invariant model where the earthquakes occur as slips on the fault planes and their probability is inversely proportional to the area of the involved fault segment [e.g., Stein and Wysession, 2009]. At the same time, the coseismic stress-drop is, on average, independent on the earthquake size [Shaw, 2009].

The results of statistical analysis in volcano seismology are not unequivocal and not as clearly systematized as in the case of the tectonic earthquakes. In other works, the attempts to describe the observations by the Gutenberg–Richter frequency–magnitude distribution lead to the b-value determinations substantially larger than 1 [McNutt, 2005, Varley et al., 2010, Roberts et al., 2015, Greenfield et al., 2022, Rodríguez-Pérez et al., 2022]. These observations show that the mechanisms of the volcanic LP earthquakes are not always fit by the scale-invariant model suitable for the tectonic earthquakes. At the same time, establishing a generally accepted form of the frequency–magnitude relationship for the volcanic LP earthquakes will require more detailed analysis of the data recorded from different volcanoes.

In particular, statistical analysis of DLP earthquakes recurrences can contribute to better understanding of the physical mechanism responsible for their generation. One of the world's most intense sources of volcanic DLP earthquakes is located beneath the Klyuchevskoy volcano in Kamchatka at the depths of 30–35 km corresponding to the crust–mantle boundary [Gorelchik and Storcheus, 2001, Gorelchik et al., 2004, Shapiro et al., 2017a]. Senyukov [2013] studied the dependence of the b-value on the depth beneath the KVG. The analysis revealed higher b-values in the region of a deep magma chamber located at the crust–mantle boundary and generating DLP earthquakes. Moreover, Gorelchik and Storcheus [2001] supposed that the frequency–magnitude relationship of the earthquakes at these depths could be better approximated with the normal distribution. Similar observations were reported at other volcanic regions including high b-values [Wiemer and McNutt, 1997, Murru et al., 2007] and evident deviations of the frequency-magnitude distribution from the Gutenberg–Richter relationship [Okada et al., 1981, Main, 1987, Lahr et al., 1994].

In this work, we analyze the frequency–magnitude distribution for the Klyuchevskoy DLP earthquakes on a more detailed basis. To this end, we apply a sensitive detection method based on a matched filter and use a magnitude scale based on the estimates of the scalar seismic moment from the records of S-waves. The sections below provide a detailed account of the analysis methods and the results of their application to the continuous seismic records for a period of 2011–2012.

2.2 Data

Seismic monitoring at KVG is conducted by a network of permanent seismic stations of Kamchatka Branch of Geophysical Survey of Russian Academy of Sciences (KB GS RAS) [Chebrov et al., 2013]. The data from all seismic network stations are transferred to Petropavlovsk-Kamchatskii and uploaded to the servers of the regional information processing center of KB GS RAS in close-to-real-time mode. The data acquisition and transfer system is organized on the basis of KB GS RAS local network using the Internet channels of two digital services providers (Rostelecom and InterKamService), RadioEthernet communication networks, OAO SetTelecom VSAT network, OAO Satis VSAT network based on "Idirect" technology with a hub in Petropavlovsk–Kamchatskii. The main file depot is a specialized seismic data archive servers based on two RAID 6 arrays. The data are stored in the form of daily files for each channel of each station [Chebrov et al., 2010, Senyukov et al., 2020].

In this work, we use radio telemetric seismic stations (RTSS) the signals from which are transmitted via radio channel with FM-FM modulation, directly or through a repeater, to the receiver centers in Kozyrevsk and Klyuchi villages where they are converted into digital records with a sampling frequency of 128 Hz. In this work, we used the records of the three-component CM3KV velocimeters with a frequency band of the recorded seismic signals of 0.7–20 Hz. The RTSS environment was developed in 1974–1982 for the purposes of on-line monitoring of active volcanoes [Gavrilov et al., 1987]. The seismic stations used in this study are shown in Figure 2.1.

2.3 Methods

2.3.1 Automatic earthquakes detection and seismic catalog compilation

To perform exhaustive statistical analysis a high quality seismic catalog should be obtained. In case of volcanic seismicity classic manual data processing is difficult and sometimes impossible due to several reasons. First of them is low signal-to-noise ratio (SNR) of waveforms of volcanic earthquakes with their emerging arrivals of P- and S-waves which are often indistinguishable. Moreover, the vast number of ongoing volcanic events can reach hundreds or even thousands during volcanic crisis. These circumstances inevitably led to the development of automated algorithms for seismic data processing.

At first, the described method received little attention from seismologists due to

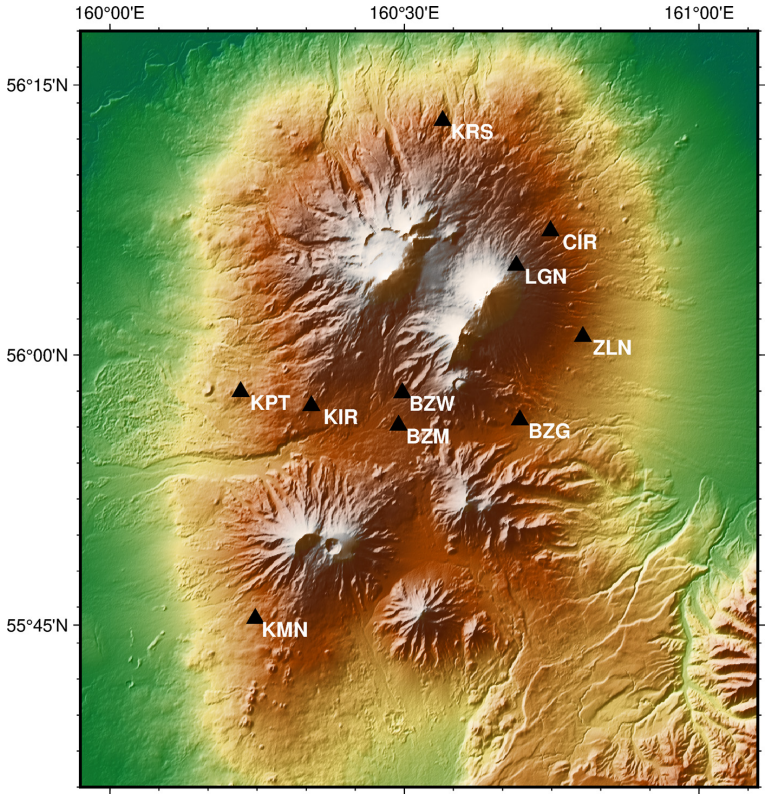

Figure 2.1: Map of the Klyuchevskoy volcano group, black triangles show the permanent stations data from which was used in this study

its high sensitivity to the input template waveform. However, the majority of recorded signals come from unknown sources, and consequently have unknown waveforms. In seismological practice, when registering an unknown signal, a method is often used where the signal energy in the short (STA) and long (LTA) time windows are compared for the waveform. Detection occurs when the STA/LTA ratio exceeds a predetermined threshold. The algorithm then defines whether the detections at multiple network stations are consistent with the seismic source. If a seismic event is registered at several stations, it is included in the catalog. The STA/LTA method confidently distinguishes earthquakes in which P- and S-wave arrivals are characterized by a high SNR. The advantage of this approach is the ability to register events without prior knowledge of the waveform or any information on the source. However, in more complex problems such as detecting volcanic earthquakes, sensitivity of the STA/LTA algorithm is too low [Yoon et al., 2015].

One of the most efficient methods for detecting a known signal in noisy time series is to compute the correlation function of the template waveform with successive sections of the data series. Parts of the signal that have a high degree of similarity to the template will result in a high value of the correlation function. Such a procedure is called *template matching* [Van Tees, 1968]. It allows to register events with very low SNR values, when the template waveform is presented with waveforms from several

stations, and the final decision on a detection is based on the average value of the correlation coefficient over a set of seismic stations [Gibbons and Ringdal, 2006a, Shelly et al., 2007]. It is a versatile and powerful tool that is capable of detecting events in a wide range of seismological studies: weak earthquakes of small magnitude [Schaff and Waldhauser, 2010], foreshocks [Kato and Nakagawa, 2014] and aftershocks [Peng and Zhao, 2009], earthquake swarms [Shelly et al., 2013], low frequency earthquakes in tectonic tremor [Shelly et al., 2007] and volcanic tremor itself [Tang et al., 2010]. This method is applicable to the study of induced seismicity [Skoumal et al., 2014] and microseismic monitoring [Plenkers et al., 2013, Song et al., 2010]. The degree of similarity between two waveforms has been repeatedly used to classify events according to the source location. Shapiro et al. [2017a] studied two groups of LP earthquakes: shallow, occurring directly under active volcanoes of the KVG, and at a depth of about 30 km, in the area of a deep magma chamber. Yet again, the main limitation of events detection using cross-correlation is the presence of a known template waveform.

Also a common method in practice of earthquakes detection is *autocorrelation* when the signal waveform of interest is unknown. In this approach continuous data is splitted into small overlapping windows, and then cross-correlation of all possible pairs of windows is calculated. When the correlation coefficient exceeds the threshold, corresponding signals are marked as "candidates". Subsequently, such events can be post-processed with additional cross-correlation calculation, grouped into "families" and form a set of less noisy template signals. Autocorrelation was applied in [Brown et al., 2008] where both known and previously unknown low-frequency earthquakes were successfully detected in a tectonic tremor.

Thus, autocorrelation is much more sensitive than STA/LTA and can detect unknown similar waveforms, but this method requires significant computational resources. This disadvantage makes it impossible to process very large volumes of seismological data. This algorithm does a lot of redundant work, since the vast majority of time windows are not correlated with each other. Autocorrelation is suitable for detecting repetitive earthquakes over several hours of continuous data [e.g., Brown et al., 2008].

Yoon et al. [2015] proposed a new detection algorithm that combines the advantages and eliminates the disadvantages of the previously described methods. It is based on the probabilistic dimension reduction of multidimensional data (LSH), and it is designed to compare not the original seismograms, but their "fingerprints", i.e. the key distinguishing features of the signals. Consequently, similar waveforms are expected to have similar "fingerprints".

Considering a particular feature of LP earthquakes, namely the similarity of their waveforms we finally selected the template matching algorithm [Van Tees, 1968]. The

main advantages of this method are (1) high sensitivity required for detecting weak LP earthquakes and (2) relatively low computational costs taking into account the amount of data to be processed. Thus, template matching algorithm searches for *a priori* known signal, template, in noisy time series by calculating the cross-correlation functions of the template signal with the successive segments of the incoming data stream. The fragments of the signal that have a high degree of similarity with the template waveform give a high value of the cross-correlation function. The detection sensitivity substantially increases in the case of simultaneous use of multicomponent seismic array records [Gibbons and Ringdal, 2006b]. Also, template matching algorithm has already shown its efficiency in various studies including LP seismicity investigations [e.g., Shapiro et al., 2017a].

The first step in processing seismological data is bandpass filtering from 1 to 5 Hz and decimation. Figure 2.2a shows one hour of record (vertical component at KMN station) containing several LP events. The template event used for detection and comparison of its waveform with the waveform of the detected earthquake is also shown (Figure 2.2b-c). The records of the template earthquake at all stations are presented in Figure 2.3.

According to the template matching algorithm, a time window of a template duration moves along the seismogram shifting at each step by one data count. At every step the correlation coefficient is calculated between the template and a signal segment of the same duration with the following equation:

$$CC(X,Y) = \frac{(X,Y)}{\|X\|\|Y\|} \tag{2.2}$$

where (X,Y) is the dot product of two vectors, X and Y, and $\|\mathrm{X}\|\|\mathrm{Y}\|$ are their norms correspondingly. If we represent vectors X and Y as sets of the counts $x_i, y_i,\ i = 1, ..., n$ in time:

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \tag{2.3}$$

the dot product and the norms of these vectors are calculated in the following way:

$$(X,Y) = \sum_{i=1}^{n} x_i y_i \tag{2.4}$$

$$\|X\| = \sqrt{\sum_{i=1}^{n} x_i^2},\ \|Y\| = \sqrt{\sum_{i=1}^{n} y_i^2} \tag{2.5}$$

For example, when applied to single component data, the algorithm yields a time series of the correlation coefficients for one day (top panel of Figure 2.4). In the case of

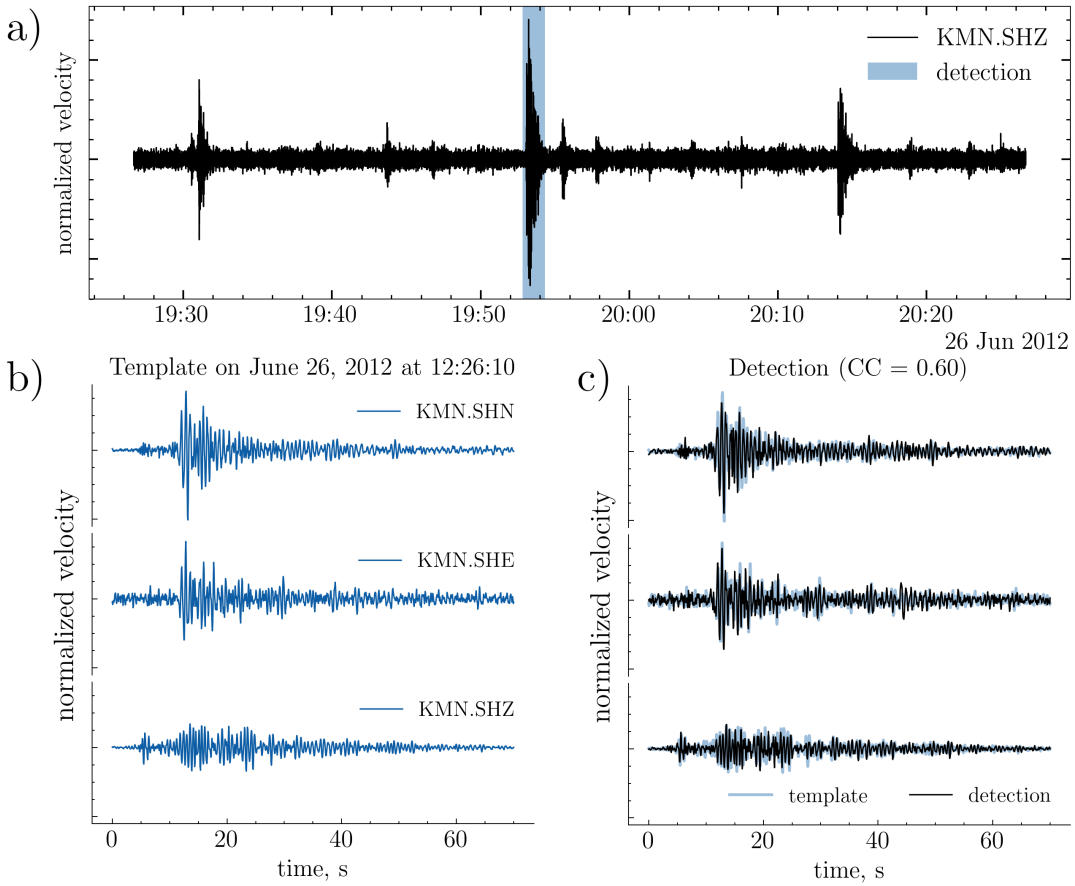

Figure 2.2: Example of an earthquake detection using template matching algorithm (all seismograms are band filtered from 1 to 5 Hz): (a) one hour of a continuous seismic record at vertical component of station KMN; (b) template used for detection (June 26, 2012 12:26:10), records at station KMN; (c) example of a detection with $CC = 0.60$ at station KMN (event is highlighted with blue in panel (a))

the analysis of multicomponent records, the time series of the correlation coefficients for individual components are averaged. In this work, we used the records from ten stations of KB GS RAS seismic network. Thus, the seismograms recorded with three channels of every station were processes with the template matching algorithm, and after that 30 obtained correlation functions were averaged. This procedure was repeated for every day of the study period. The example in the bottom panel of Figure 2.4 shows that simultaneous analysis of multicomponent data results in more precise detection and makes the algorithm more selective.

As a result, the algorithm detects the signals whose quite complex waveforms including both direct seismic waves and the coda highly similar to the initial template simultaneously at several channels (including the correction for time delays between stations). This waveform similarity is possible only in the case when the sources of all

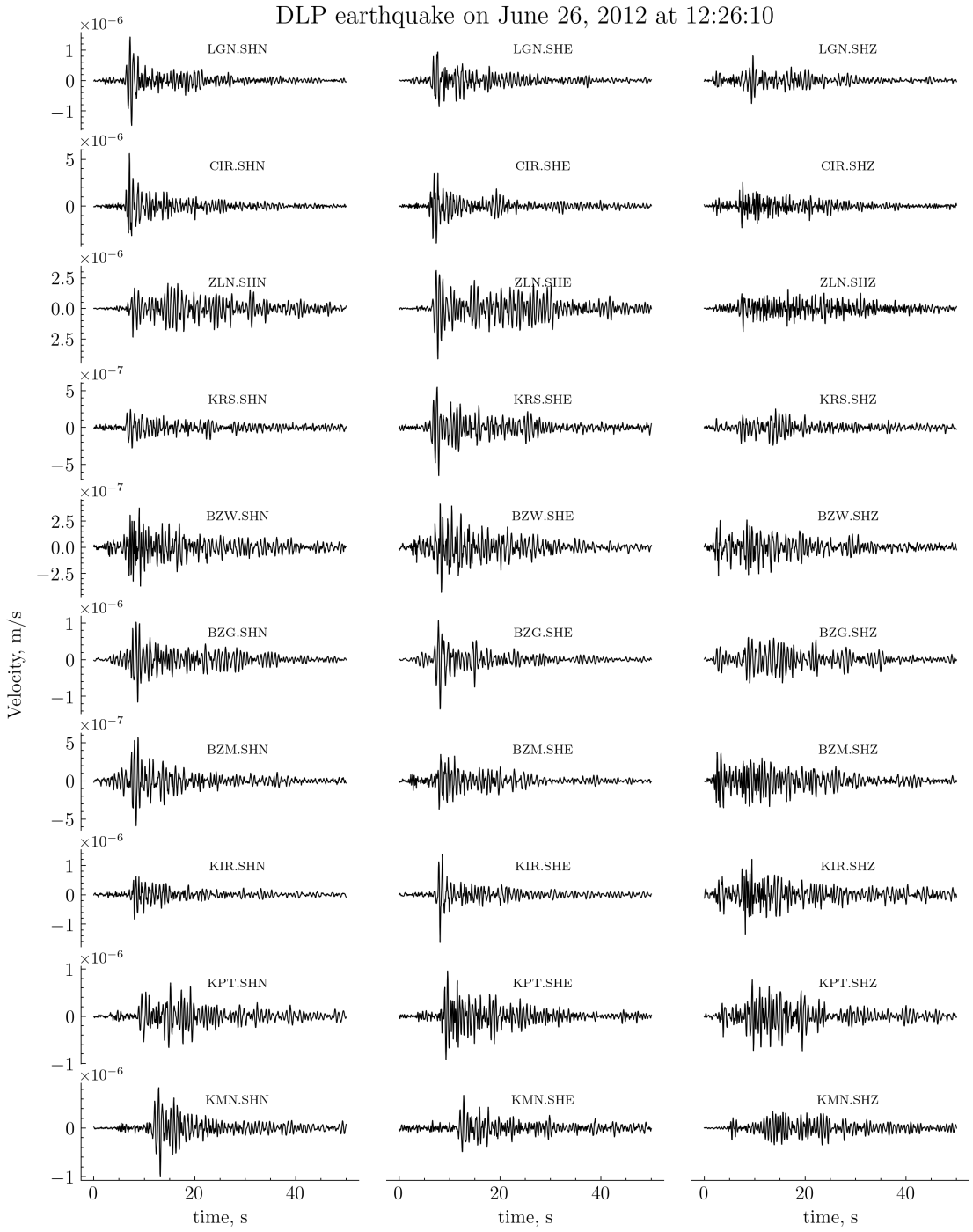

Figure 2.3: Seismograms of a template DLP event (June 26, 2012 at 12:26:10) recorded by the stations used in this study

detected signals are located very close to each other (at least, not further than half a wavelength) and have almost identical mechanisms. Such a group of earthquakes can be considered being a result of the action of a single source with very frequent recurrence.

The initial template used for detecting a group of earthquakes generated by the

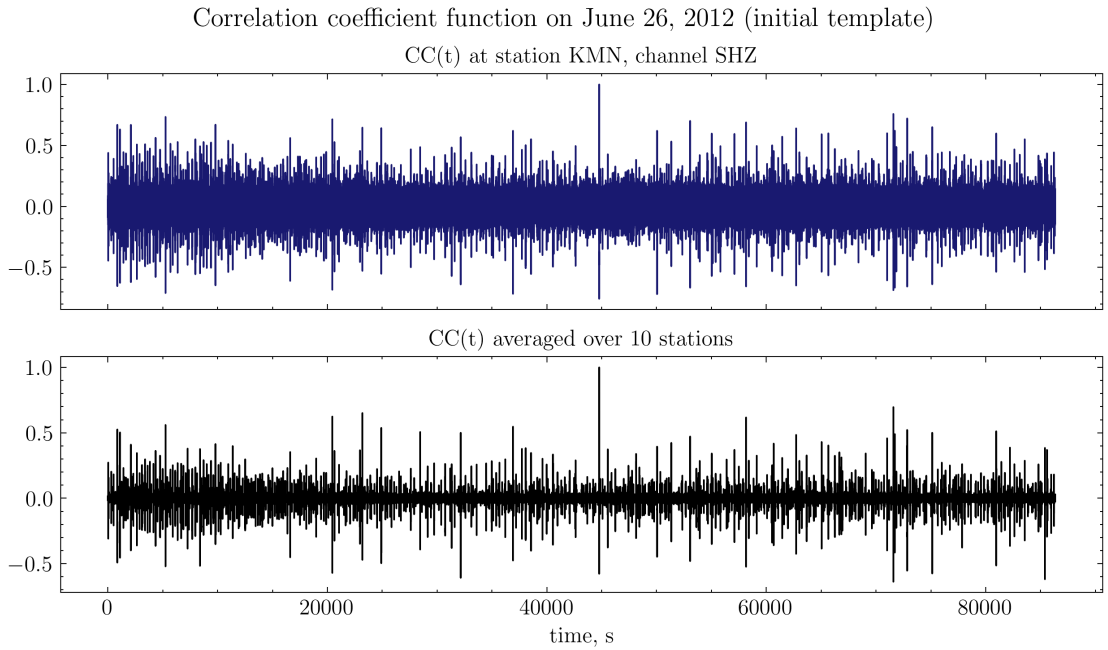

Figure 2.4: Time series of correlation coefficients for one day of observations. Single peak with $CC = 1$ corresponds to the record segment containing the template signal (June 26, 2012, 12:26:10, Figure 2.3): result for the vertical component of station KMN (top); correlation coefficient averaged over 10 stations and 3 components (bottom)

same recurrent source (or by a group of identical and very closely located sources) was selected, in a sense, at random. Therefore, for improving the detection quality, an stacked, i.e. averaged, template which is more representative of the entire group of the earthquakes was calculated. To this end, at the previous step, one should select all the detections which meet the condition $CC \geq 0.3$ and then stack their waveforms. This procedure allows to reduce the contribution of the incoherent noise and constructively sums the correlated signals. Figure 2.5 shows the waveforms of the original and stacked templates at one channel. After this, the procedure of correlation function calculation is repeated with the stacked template.

The detection criterion is that the averaged correlation coefficient over all stations must exceed a given threshold. At $CC_{stack} < CC_{threshold}$, the detection is considered false and the corresponding CC value in the series is zeroed (Figure 2.6). If the detection proves to be reliable, the moment magnitude of the detected event is calculated by the algorithm described in the section below.

2.3.2 Magnitudes calculation

Currently, in Kamchatka the classification of the regional earthquakes in terms of their energy class *Ks* is used [Rautian, 1960, 1964] (the S-wave class determined from

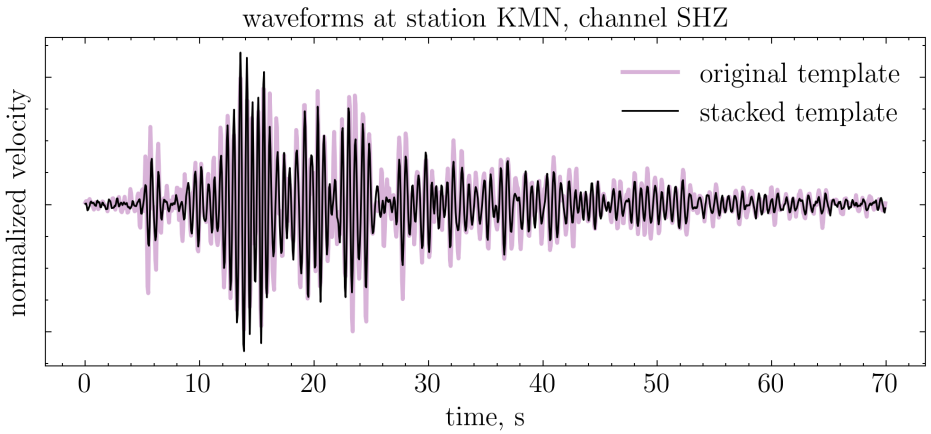

Figure 2.5: Examples of waveforms of the original (purple line) and the stacked (black line) templates

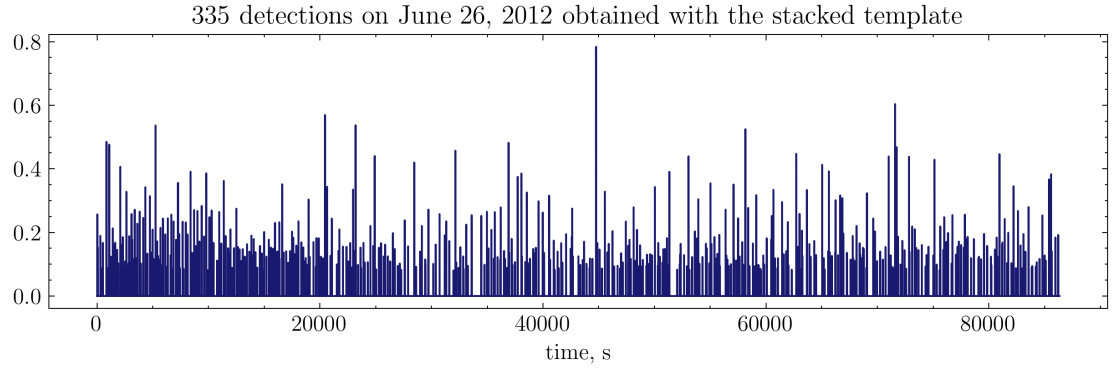

Figure 2.6: Result of event detection with the stacked template over 10 stations and 3 components; $CC_{average} \geq CC_{threshold} = 0.12$

S.A. Fedotov's nomogram [Fedotov, 1972]). This quantitative characteristic is defined as:

$$Ks = \lg E \tag{2.6}$$

where E is the energy (in joules) of seismic waves. The relationship between energy class and magnitude m^{CKM} is described by the following formula [Gusev and Melnikova, 1990]:

$$Ks = 2.00m^{\mathrm{CKM}} + 1.68 \pm 0.55 \tag{2.7}$$

where $m^{\mathrm{CKM}} \approx m\mathrm{B} + 0.18$, $m\mathrm{B} < 5.6$.

The energy class Ks is determined from the nomogram constructed for the tectonic earthquakes of a given region based on the maximum ratio of the amplitude in the S-wave to its period $(A/T)_{max}$ and the epicentral distance for the earthquakes with a source depth $h = 0 - 200$ km. The nomogram was calibrated using the earthquakes with $Ks = 10 - 11$ that corresponds to the magnitude range of $4.2 - 4.7$. Magnitudes

of volcanic earthquakes are substantially lower that challenges the applicability of this magnitude determination method to the studied type of seismic events. Therefore, in this work, we use the classical definition of the moment magnitude [Kanamori, 1977, Hanks and Kanamori, 1979] with the calculation of the seismic moment.

Despite the fact that the mechanism of the LP volcanic earthquakes is likely to differ from a pure shear slip along a fault, i.e. an idealized mechanism of tectonic earthquakes, and contains a significant volumetric component, it still can be described by a seismic moment tensor [Wech et al., 2020]. In some cases, a superposition of a moment tensor and a single force vector is considered [Chouet and Matoza, 2013, Wech et al., 2020]. In case of a tensor description of the source, it is possible to determine a scalar seismic moment. The latter will not necessarily be associated with the magnitude of the slip on the fault, as in the case of the tectonic earthquakes, but rather with a volume change. For example, considering propagation of magmatic melt through a system of microfaults, a probable geometry of a seismic source could be an opening crack. If so, the seismic moment tensor M is written out as [Aki and Richards, 1980]:

$$M = \begin{pmatrix} \lambda \mathrm{d}V & 0 & 0 \\ 0 & \lambda \mathrm{d}V & 0 \\ 0 & 0 & (\lambda + 2\mu)\mathrm{d}V \end{pmatrix} \tag{2.8}$$

where λ and μ are Lamé constants; $\mathrm{d}V$ is the change in the volume of a crack. In the case of the Poisson medium ($\lambda = \mu$), the scalar seismic moment $M_0 = 0.6\,\lambda \mathrm{d}V$. Sources with this type of the mechanism emit more S-waves than P-waves [Shi and Ben-Zion, 2009], it is consistent with the observed signals of DLP earthquakes [Shapiro et al., 2017a, Wech et al., 2020]. Based on these considerations, we proposed to use the moment magnitude scale for the DLP earthquakes:

$$M_\mathrm{w} = \frac{2}{3}\left(\lg M_0 - 9.05\right) \tag{2.9}$$

The hypocenters of the studied DLP earthquakes are localized within a small spatial domain beneath the Klyuchevskoy volcano at the crust–mantle boundary (at the depths from 30 to 35 km) [Shapiro et al., 2017a]. Thus, the distance from these hypocenters to the recording seismic stations is above ten wavelengths (for the frequencies of the order of 1.5 Hz used in this study). Therefore, the estimate of the seismic moment was obtained from the equations for ground surface displacements from S-waves in the far field [Aki and Richards, 1980]:

$$u(\mathbf{x}, t) = \gamma \frac{\dot{M}_0(t - r/\beta)}{4\pi\rho\beta^3 r} \tag{2.10}$$

where M_0 is the seismic moment; t is time; β is the S-wave velocity; ρ is the rocks density; r is the distance from the hypocenter to the observation point $\mathbf{x}$; γ is a factor associated with the radiation pattern of the source. In the practical calculations of magnitudes, the full seismic moment tensor and the corresponding radiation pattern are not estimated (largely due to the complexity and instability of the estimation procedure). Instead, it is assumed that the use of the records from multiple stations averages this factor. Therefore, in the subsequent analysis, we use the approximate value $\gamma = 1$. Thus, from equation 2.10 it follows:

$$\dot{M}_0 \sim 4\pi\rho\beta^3 r \cdot u^S \tag{2.11}$$

Calculating the derivative of the displacement in the Fourier domain, we obtain:

$$v^S \sim \dot{u}^S \sim 2\pi f \cdot u^S \tag{2.12}$$

where u^S and v^S are the amplitudes of the displacement and velocity records, respectively; f is the characteristic signal frequency of S-waves. The final estimate of the seismic moment is

$$M_0 \sim \frac{4\pi\rho\beta^3 r \cdot u^S}{2\pi f} = \frac{\rho\beta^3 r}{\pi f^2} \left|v^S_{max}\right| \tag{2.13}$$

where $\left|v^S_{max}\right|$ is the maximum amplitude in a DLP seismogram. The distances r are calculated from the known depth (h = 32 km) of the magma reservoir and the coordinates of seismic stations. In our calculations, we assumed the following parameters of the medium: ρ = 3000 kg/m^3, β = 3500 m/s.

Based on the results of visual inspection of the amplitude Fourier spectra at different stations and channels (example of station KMN in Figure 2.7), we decided to use the value of the characteristic frequency f = 1.5 Hz.

The maximum displacement velocities at an individual station were calculated from three components:

$$\langle v^S_{max} \rangle = \sqrt{\sum_{j=1}^{3} v_j^2} \tag{2.14}$$

where $v_{j_{max}}$ is the maximum displacement velocity in the j-th channel. From the obtained $\langle v^S_{max} \rangle$ the seismic moments (13) and moment magnitudes (8) are calculated at each station, and the obtained values are then averaged over all stations:

$$M_{\mathrm{w}} = \frac{1}{N} \sum_{i=1}^{N} M^i_{\mathrm{w}} \tag{2.15}$$

and this final value M_{w} is cataloged.

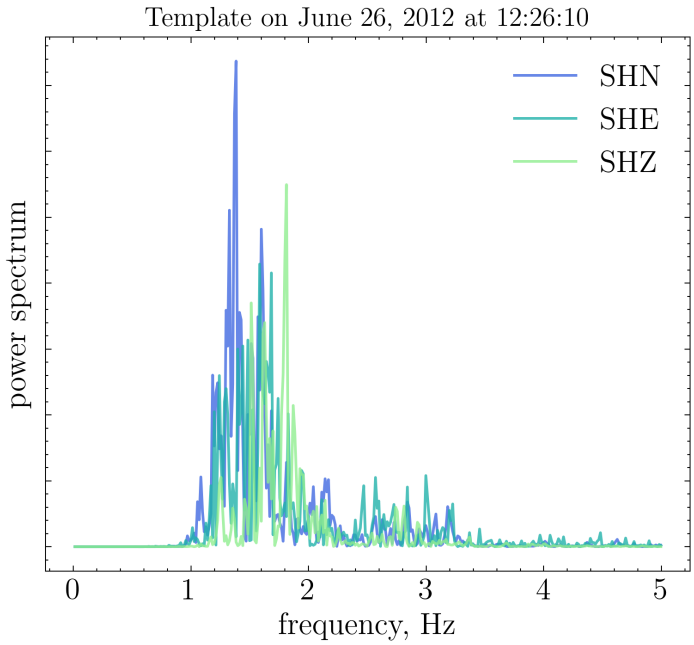

Figure 2.7: Examples of the amplitude Fourier spectra of a DLP earthquake (template on June 26, 2012 at 12:26:10) at various channels of station KMN

2.4 Results

We have processed the seismograms obtained by KB GS RAS stations network from January 1, 2011 to December 31, 2012 using more than 400 template earthquakes selected from the KB GS RAS catalog of volcanic earthquakes [Senyukov, 2013, Senyukov et al., 2014]. The processing result using one template is a seismic subcatalog where every seismic event has in correspondence the CC value between its waveform and the used template and calculated moment magnitude M. The minimum threshold value of the correlation coefficient for including an earthquake in a catalog is $CC_{min} = 0.08$. For each CC_i value, starting from CC_{min} with a step of 0.02, we compiled a subset of earthquakes whose correlation coefficients were higher than CC_i. Based on the results of the performed statistical analysis, for the subsequent study we decided to use the samples composed of the seismic events with $CC \geq 0.12$.

Figure 2.8 shows the frequency-magnitude distribution of DLP events from one of the largest subcatalogs obtained with the template shown in Figure 2.3, and this database includes 11627 events. Further we show the possible ways to approximate the obtained frequency–magnitude distribution. It is logical to begin with the power-law testing (Figure 2.8a); however, the b-value of the resulting linear relationship proves to be overestimated ($b = 3.4$), which is consistent with already existing results obtained with the catalogs of the volcanic earthquakes [e.g., Senyukov, 2013]. The b-value was estimated by the least square technique in the interval of magnitudes starting from $M_w = 1.40$. In this case, we do not intend to find the best approximation of the

frequency–magnitude plot but rather show that the b-value is strongly above 1 in any case (grey dashed line with the corresponding slope in Figure 2.8a). Taking into account some notes that the frequency–magnitude distributions of volcanic earthquakes clearly deviate from the Gutenberg–Richter relationship [e.g., Okada et al., 1981, Main, 1987, Lahr et al., 1994] we propose to describe the obtained plot with another type of distribution, for instance, the normal distribution:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{(x-\mu)^2}{2\sigma^2} \tag{2.16}$$

with the standard deviation $\sigma = 0.27$ and the mean $\mu = 1.35$. The right branch of the plot is approximated well with the normal distribution while the left branch is affected by the insufficient sensitivity of the network (or the effects of network blinding due to volcanic tremors).

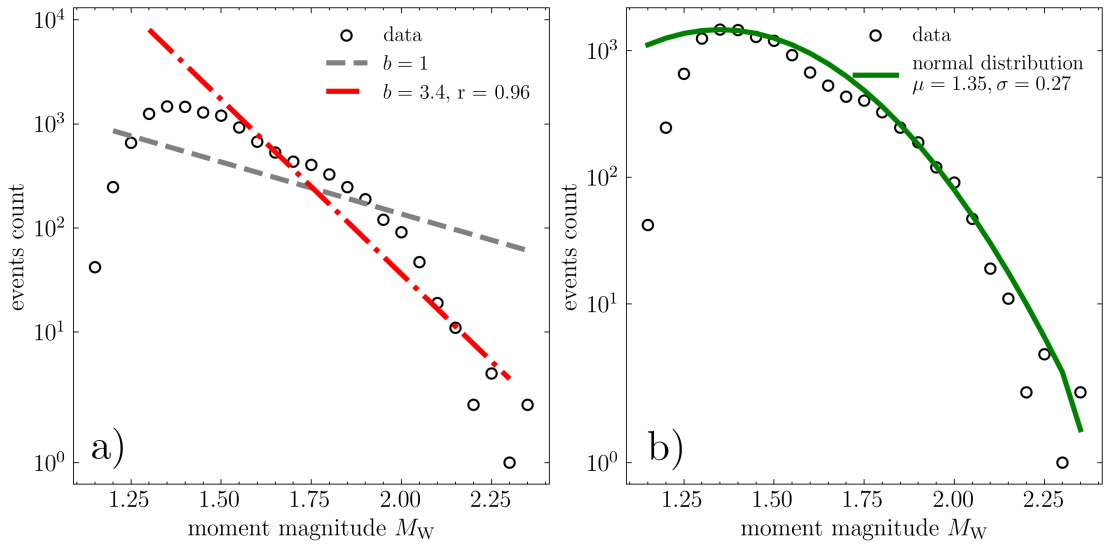

Figure 2.8: Frequency–magnitude distribution (open circles) for the subcatalog obtained with the template shown in Figure 2.3: (a) possible approximation with Gutenberg–Richter law and (b) possible approximation with the normal distribution.

As it was mentioned before the seismic data recorded during 2011–2012 was processed with more than 400 template events. The obtained subcatalogs for each template were merged to compile the final catalog of 48915 earthquakes. The frequency-magnitude distribution of this dataset is presented in Figure 2.9.

2.5 Discussion and Conclusions

The distributions obtained for the cluster of similar DLP earthquakes beneath the Klyuchevskoy volcano substantially differ from the Gutenberg–Richter relationship by

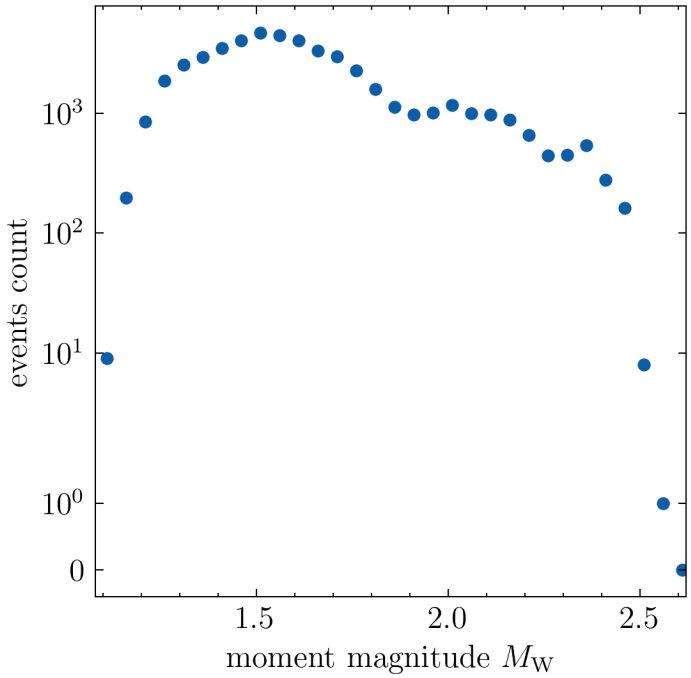

Figure 2.9: Frequency–magnitude distribution of the complete catalog of DLP earthquakes

the fact that for these earthquakes it is possible to determine a scale parameter that corresponds to a certain characteristic size of the generating source. This characteristic size scaling is fairly consistent with the generation of numerous recurrent earthquakes by a single source. The existence of these sources is described by the models considering pressure variations in magma. Seismic waves in these models can either be generated by a periodic pressure drop through mechanical barriers or valves [Shapiro et al., 2018, Wech et al., 2020] or by rapid avalanche-like degassing leading to a pressure increase [Melnik et al., 2020]. We note that for this type of the earthquakes, the "standard" analysis of the frequency–magnitude relationship with estimation of the b-value is meaningless as the obtained values will vary depending on the used magnitude range.

We built a full catalog of DLP earthquakes by merging subcatalogs obtained with the selected templates. The frequency–magnitude distribution for the complete catalog of DLP earthquakes beneath the Klyuchevskoy volcano has a complex shape with several local maxima. We interpret this effect as the result of overlapping of the distributions from the sources with several characteristic sizes.

Chapter 3

Source mechanisms of deep long period earthquakes beneath the Klyuchevskoy volcano group inferred from S-to-P amplitude ratios

Results presented in this Chapter have been submitted as a paper to Journal of Volcanology and Geothermal Research.

3.1 Introduction

Analysis of seismic signals recorded in the vicinity of volcanoes is one of the key elements of volcano monitoring and one of the important sources of information about active processes occurring at depths in the volcano plumbing systems [e.g., Chouet, 2003, Sparks, 2003, Nishimura and Iguchi, 2011, Zobin, 2011, Chouet and Matoza, 2013, McNutt and Roman, 2015, Matoza and Roman, 2022, Thelen et al., 2022].

Volcanic seismicity is often divided into two main classes. The first class of signals associated with volcanic activity is composed of volcano-tectonic (VT) earthquakes [e.g., Roman and Cashman, 2006]. These nearly impulsive signals are characterized by a wide spectral range (up to 15–20 Hz) and clear arrivals of P- and S-waves. As their name suggests, the origin of the VT earthquakes is believed to be similar to regular tectonic earthquakes and is associated with the brittle failure faulting of crustal rocks in the vicinity of volcanoes.

The second class of seismo-volcanic phenomena is called long-period (LP) seismicity and includes a large variety of signals ranging from relatively short LP earthquakes to long duration tremors. These signals have typical frequencies of a few Hz and their origin is often associated with pressure fluctuations within magmatic and hydrothermal fluids [e.g., Chouet, 1996a]. Thus, the LP seismicity is believed to be directly related to the processes of the magma motion and pressurisation and to be able to provide reliable precursors of volcanic eruptions.

While most of the observed LP seismicity originate at very shallow depths (less than 3-5 km), a particularly important sub-class known as the deep long period (DLP) earthquakes occurs at depths ranging from the middle crust to the crust-mantle boundary (25–40 km). The DLPs are usually linked with the processes occurring within the magmatic fluids at such large depths and can reflect the activation of deep-seated parts of the magmatic systems prior to eruptions. One of the first times, the swarms of DLP events were recorded beneath Mount Pinatubo [White et al., 1996], and linked to magma ascending from depth. Later DLPs were reported at different regions such as Cascadia [Nichols et al., 2011], the Aleutian arc [Power et al., 2004], Japan [Hasegawa and Yamamoto, 1994, Nakamichi et al., 2003, Aso et al., 2013, Niu et al., 2017, Yukutake et al., 2019, Kurihara et al., 2019, Ikegaya and Yamamoto, 2021, Kurihara and Obara, 2021], Kamchatka [Shapiro et al., 2017a], Eifel [Hensch et al., 2019], and Hawaii [Wech et al., 2020].

Despite these numerous studies, physical processes leading to generation of DLP earthquakes and their seismic mechanisms remain poorly understood. Several hypotheses have been suggested. As already mentioned, White et al. [1996] considered the DLP seismicity being the elastic manifestation of the injection of deep-seated basaltic fluids without, however, proposing any physical model. Aso and Tsai [2014] suggested that the DLP earthquakes can be caused by thermal stresses induced by cooling of deep magma bodies. The relationship between the possible focal mechanisms and stress orientations suggested in this work again remained only at a qualitative level. The DLP activity beneath Mauna Kea volcano in Hawaii has been also suggested to be related to magma bodies cooling by Wech et al. [2020]. These authors however relate the generation of the DLP earthquakes not with the thermal stresses but with the degassing caused by so called "secondary boiling", i.e., the exsolution of volatiles during the crystallization of cooling magma. Again, no quantitative model relating the degassing and the generation of seismic waves has been suggested. More generally, the hypotheses relating DLP activity with magma cooling do not seem appropriate for the cases when such activity is occurring beneath the active volcanoes and especially in association with eruptions and fresh magma supplies. For such situation, the model of Melnik et al. [2020] considered a possible DLP generating mechanism via the rapid growth of gas bubbles in response to the slow decompression of magma over-saturated with volatiles (CO_2–H_2O rich). These authors show that this model is compatible with the known composition of the basaltic magma emitted by the Klyuchevskoy volcano in Kamchatka (Russia) where a sustained DLP activity is observed [Shapiro et al., 2017a] and also can reasonably explain amplitudes and frequencies of the observed DLP signals.

Kinematic moment-tensor inversion of the DLP earthquakes [e.g., Nakamichi et al., 2003, Aso and Ide, 2014, Hensch et al., 2019, Ikegaya and Yamamoto, 2021] was carried out with fitting the amplitudes and waveforms of body waves and generally demonstrated a strong volumetric component in the seismic source, which would be in agreement with the generating process involving pressure variations in magma. However, the results of such source inversion are associated with strong uncertainties because of the combination of noisy data with poorly known wave propagation.

In this work, we study the kinematic source parameters of 29 DLP earthquakes beneath the Klyuchevskoy volcano group (KVG) in Kamchatka, Russia recorded by temporary stations of a recent seismic experiment [Shapiro et al., 2017b]. We use the fit to the observed amplitude ratios between P- and S-waves in order to constrain the source mechanism. We consider possible mass advection effects and add to the moment tensor a possible single force component [Takei and Kumazawa, 1994]. At the same time, instead of performing a full inversion in the parameter space that my be subject to a strong uncertainty, we rather proceed with testing a set of "plausible" hypotheses resulting in source mechanisms with a reduced number of parameters.

In the following sections, we start with describing the volcanic system and the dataset. We then discuss different hypotheses about the source mechanism and corresponding descriptions in terms of kinematic source parameters. The inversion method is then introduced and applied to the records of 29 selected DLP earthquakes.

3.2 Dataset

In this work we use the data of a joint Russian-German-French temporary seismic experiment KISS (Klyuchevskoy Investigation – Seismic Structure of an Extraordinary Volcanic System) [Shapiro et al., 2017b]. The catalog of earthquakes occurred during operation of this network (August 2015 – July 2015) has been compiled by the Kamchatka Branch of Russian Geophysical Survey [Senyukov et al., 2021] based on semi-automatic picking of arrivals of P- and S-waves [Droznin and Droznina, 2011]. Earthquakes with the hypocenters located in the vicinity of the crust-mantle boundaries beneath the Klyuchevskoy volcano were selected as potential DLPs. The first choice criterion was the spatial distribution: latitude $\varphi \in [56.2°\text{N}, 56.6°\text{N}]$, longitude $\lambda \in [160.5°\text{E}, 160.7°\text{E}]$ and depth $z > 25$ km. There were 136 events laying within given boundaries. The second criterion $M_L \geq 1.3$ allowed to select 50 strongest earthquakes. Finally, after visually verifying their frequency content and signal-to-noise ratios, we retained 29 earthquakes (Figure 3.1, Table 3.1) with waveforms suitable for the analysis of the source parameters recorded at 19 stations. Figure 3.2 presents an

example of a DLP event (No. 6 in Table 3.1)

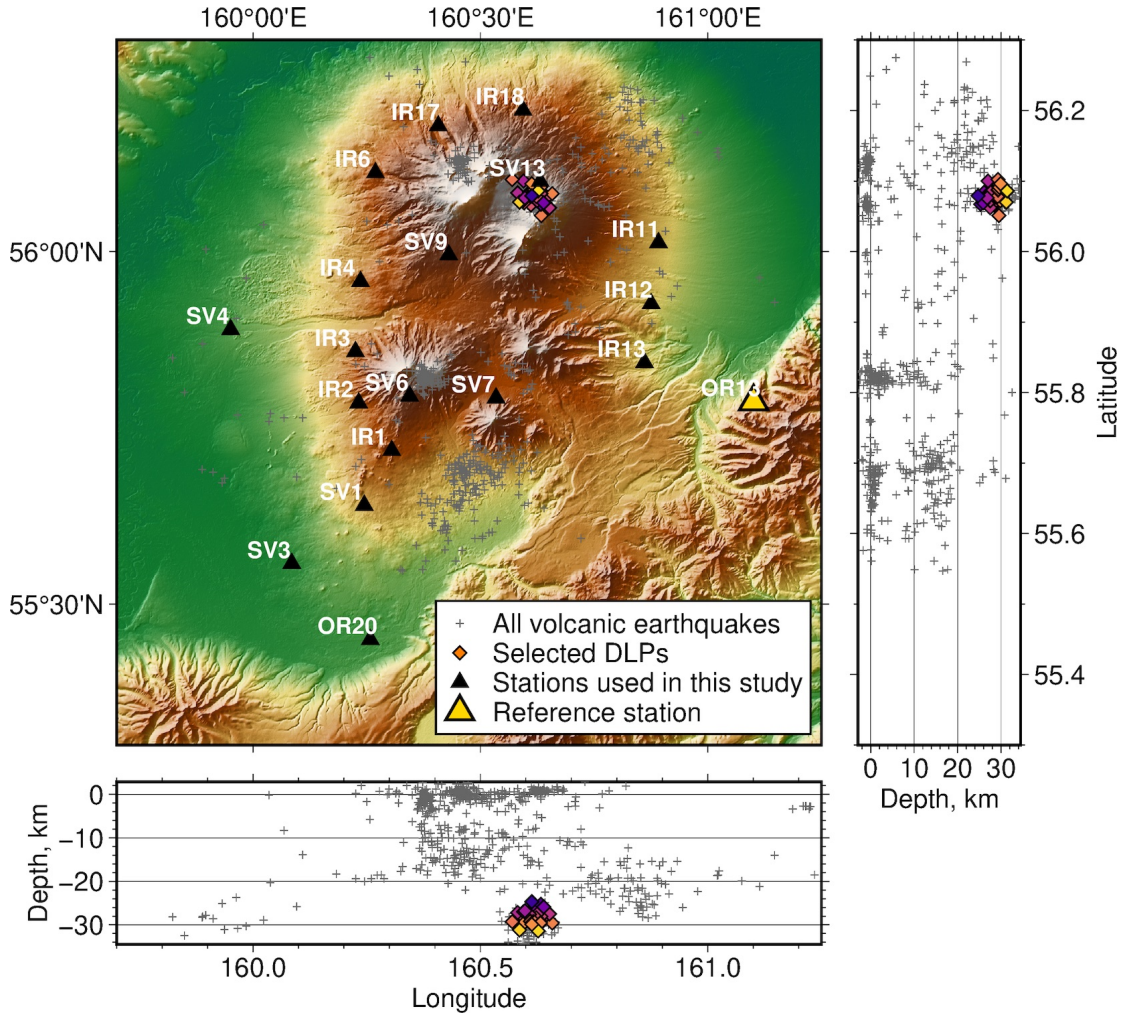

Figure 3.1: Map of the Klyuchevskoy volcano group. The stations used in this work are represented by black triangles. The "reference" station OR13 is additionally highlighted with yellow color. Grey crosses show the entire catalog of volcanic earthquakes with approximately crustal depth recorded beneath the KVG during the KISS experiment [Senyukov et al., 2021]. The diamonds show hypocenters of the DLP earthquakes selected for this study, the color of the diamonds represents the hypocenter depth.

No	Date	Time	Latitude	Longitude	Depth, km	M_L
1	2015-08-08	09:55:30	56.085	160.634	28.0	1.3
2	2015-08-09	23:24:57	56.077	160.628	27.4	1.4
3	2015-08-16	07:29:15	56.096	160.613	29.9	1.5
4	2015-08-16	12:14:57	56.066	160.611	28.4	1.3
5	2015-08-16	13:31:52	56.085	160.609	29.1	1.6

6	2015-08-20	12:23:54	56.084	160.616	26.1	1.6
7	2015-08-21	06:47:06	56.069	160.596	28.4	1.4
8	2015-08-25	10:36:26	56.083	160.582	27.2	1.4
9	2015-08-31	00:19:06	56.074	160.642	28.1	1.6
10	2015-09-02	02:09:31	56.051	160.634	29.5	1.4
11	2015-09-03	09:52:02	56.082	160.658	29.6	1.5
12	2015-09-04	16:25:07	56.100	160.595	27.0	1.4
13	2015-09-04	22:07:35	56.067	160.633	25.3	1.6
14	2015-09-10	23:28:12	56.062	160.652	27.5	1.5
15	2015-09-11	10:40:44	56.082	160.633	28.1	1.5
16	2015-09-14	14:47:30	56.072	160.625	27.5	1.6
17	2015-09-23	15:49:22	56.069	160.639	25.9	1.4
18	2015-09-29	07:15:15	56.078	160.623	28.4	1.3
19	2015-10-02	14:30:58	56.072	160.594	29.8	1.4
20	2015-10-02	17:44:59	56.070	160.586	31.2	1.4
21	2015-11-05	22:13:04	56.102	160.570	29.3	1.3
22	2015-11-23	03:18:38	56.084	160.608	27.6	1.8
23	2015-11-23	13:29:30	56.086	160.627	31.4	1.8
24	2015-11-26	06:28:15	56.078	160.610	29.2	1.5
25	2015-11-27	19:28:40	56.087	160.616	29.3	1.3
26	2015-11-27	22:05:23	56.079	160.613	24.7	1.3
27	2016-01-11	07:25:44	56.076	160.598	26.8	1.4
28	2016-05-21	11:40:28	56.088	160.619	29.4	1.7
29	2016-05-27	05:40:42	56.091	160.632	29.6	1.4

Table 3.1: List of selected deep long period earthquakes

3.3 Hypotheses about the origin and mechanisms of the DLP seismicity beneath the KVG

A general kinematic description of an earthquake source is based on a moment tensor representation [Aki and Richards, 1980]. This second-order symmetric tensor describes generally oriented and shaped discontinuities within the Earth, such as a slip across a fracture plane, or pressure variations within a volume of a nearly spherical shape, of a crack or of a pipe. In addition to moment tensor, the mass advection effects

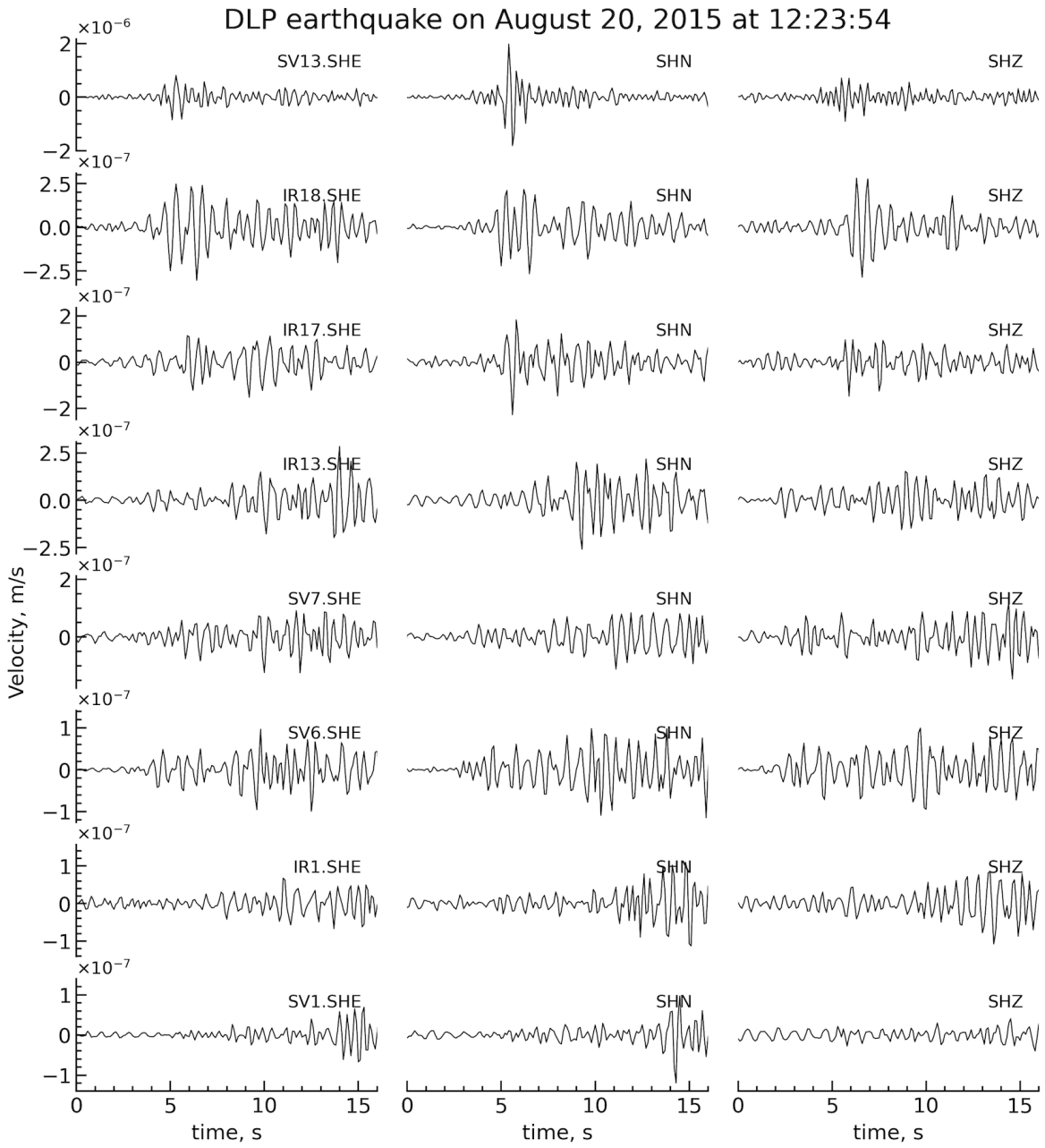

Figure 3.2: Example of seismograms recorded by multiple KISS stations during a DLP earthquake occurred on August 20, 2015 at 12:23:54.

[Takei and Kumazawa, 1994] in mechanically open systems such as volcanic conduits can result in a single force component of a seismic source. For example, acceleration of fluid would result in a changing viscous shear force acting on the conduit walls and oriented parallel to the flow [Ukawa and Ohtake, 1987, Shapiro et al., 2018].

Therefore, a most complete description of a seismo-volcanic source could be provided by a combination of a moment tensor and a single force [e.g., Kumagai, 2009]. This description contains nine independent parameters (three force components and

six independent components of the moment tensor) that can be simultaneously retrieved only in a case of very good data coverage. This situation can be achieved with inversion of very long-period (a few tens of seconds) waveforms recorded by multiple seismic stations [e.g., Ohminato et al., 1998, Chouet et al., 2003, 2005]. At such long wavelengths, the propagation of seismic waves (Green's functions) can be predicted with a sufficient accuracy based on 3D numerical simulation with accounting for topography and other effects.

At shorter periods, full waveform inversion becomes problematic because of the difficulty to characterize the wave propagation accurately. In this case, the source configurations are inferred from a reduced set of parameters measured from amplitudes and phases of the waveforms. Very often, kinematic source mechanisms are constrained with the polarities of first arriving P-waves. In the case of volcanic DLP earthquakes this approach is, however, problematic because the relatively weak DLP signals emerge from the noise only after band-pass filtering the raw seismograms. In particular, the strong microseismic noise is removed with high-pass filters with corner frequency near 1 Hz, which is very close to the dominant frequencies of LP waveforms. As a result, phases of filtered waveforms can be strongly distorted and their apparent polarities reversed.

Therefore, in this study we use the amplitude information and namely the distribution of S-to-P amplitude ratios. This type of observation has been shown to provide useful constraints on focal mechanisms and has been used in some previous studies of DLP earthquakes [e.g., Ukawa and Ohtake, 1987, Ohmi and Obara, 2002, Nakamichi et al., 2003, Ikegaya and Yamamoto, 2021]. At the same time it is associated with strong uncertainties [Hardebeck and Shearer, 2003] because of the noise and site amplifications. Taking these circumstances into consideration, we do not set up a full inversion for an arbitrary focal mechanism with 9 free parameters that would be highly unstable and non-unique. Instead, we test specific hypotheses about possible DLP generation processes. Each hypothesis is related with a "simplified" focal mechanism depending on a reduced number of parameters (maximum three).

One possibility is that DLP earthquakes are caused by the release of mechanical stresses in the crust beneath volcanoes that can be accommodated as a slip on faults. The shear-fault mechanism is described by three angles (Figure3.3a). Azimuth ϕ and dip angle δ define the fault plane orientation, and rake λ points the direction of the slip.

The DLP earthquakes can be caused by rapid magma pressure variations. The associated kinematic source mechanism will then depend on the shape of a magma filled volume. The most efficient mechanism of magma transport through the cold

crust is via formation of cracks along which magma flows in the form of dikes and sills [e.g., Rubin, 1995, Melnik et al., 2022]. Pressure variations in such planar intrusion can be approximated by a tensile crack mechanism. It can be oriented in space with a normal vector to a crack surface (Figure 3.3b), thus it is described with two angles: the azimuth ϕ and the polar angle (dip) θ. A similar kinematic mechanism will describe the opening of tensile cracks [e.g., Bean et al., 2014].

Pipe-shaped magmatic conduits are formed beneath eruptive vents of volcanoes. We consider a possibility of such conduits at depth. A pressure variation within such conduit is kinematically described as a cylindrical pipe whose orientation is also defined by two angles, azimuth ϕ and dip θ (Figure 3.3c).

Finally we consider a situation when acceleration of magma within a conduit can result in traction forces [Ukawa and Ohtake, 1987, Shapiro et al., 2018]. Such single-force mechanism is described by two angles, azimuth ϕ and dip θ, defining the force vector (Figure 3.3d).

3.4 Estimation of the DLPs source mechanism with S-to-P amplitude ratios

In this approach, we first compute theoretical ratios between the amplitudes of P and S waves and thereafter compare them with those measured from the real seismograms.

3.4.1 Calculation of the theoretical S-to-P amplitudes ratios

We start with theoretical expressions of far-field P- and S-wave displacements in a homogeneous isotropic media characterized by density ρ, P-wave velocity α, and S-wave velocity β [Aki and Richards, 1980]. For a source described by a single force $\mathbf{F}$, the wave amplitudes on component i are:

$$u_i^P = \frac{F_j R_{i,j}^{FP}}{4\pi\rho\alpha^2 r} \tag{3.1}$$

$$u_i^S = \frac{F_j R_{i,j}^{FS}}{4\pi\rho\beta^2 r} \tag{3.2}$$

where F_j is the force projection on direction j, r is the distance between the source and the receiver, and $R_{i,j}^{FP}$ and $R_{i,j}^{FS}$ are the force-type radiation patterns for P- and S-waves that are expressed via the directional cosines of a vector pointing from the source to the station γ:

$$R_{i,j}^{FP} = \gamma_i \gamma_j \tag{3.3}$$

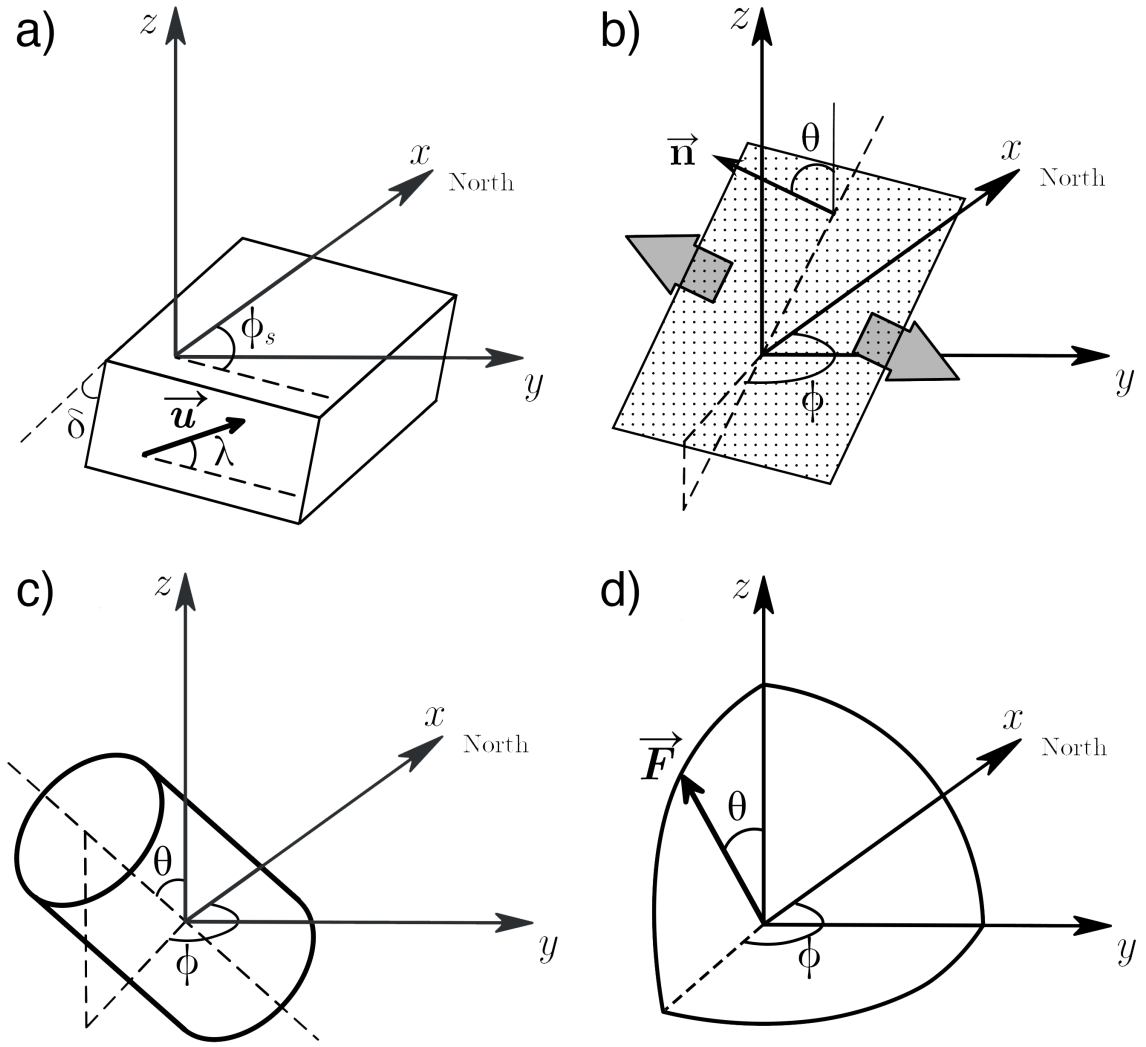

Figure 3.3: Considered "elementary" source mechanisms and respective angles required for their description: (a) a shear slip on a fault (strike $\phi_s \in [0°, 360°]$, dip (polar angle) $\delta \in [0°, 90°]$, rake $\lambda \in [-180°, 180°]$) (b) a tensile crack (azimuth $\phi \in [0°, 360°]$, dip $\theta \in [0°, 90°]$) (c) a cylindrical pipe (azimuth $\phi \in [0°, 360°]$, dip $\theta \in [0°, 90°]$) (d) a single force (azimuth $\phi \in [0°, 360°]$, dip $\theta \in [0°, 90°]$)

$$R_{i,j}^{FS} = \delta_{i,j} - \gamma_i \gamma_j \tag{3.4}$$

For a source described by a seismic moment rate tensor $\dot{\mathbf{M}}$, these expressions become:

$$u_i^P = \frac{\dot{M}_{j,k} R_{i,j,k}^{MP}}{4\pi\rho\alpha^3 r} \tag{3.5}$$

$$u_i^S = \frac{\dot{M}_{j,k} R_{i,j,k}^{MS}}{4\pi\rho\beta^3 r} \tag{3.6}$$

where $R^{MP}_{i,j,k}$ and $R^{MS}_{i,j,k}$ are the moment-type radiation patterns that are also expressed via the directional cosines γ [Aki and Richards, 1980]. Details about representations of different types of elementary sources and computation of radiation patterns are given in 3.A.

For known source and station positions the direction from the source to the station (azimuth φ and polar angle i_ξ) can be evaluated (i.e., with ray tracing). Based on this, directional cosines and radiation patterns can be computed and displacement components predicted with equations (3.3-3.4) or (3.5-3.6).

In some studies [e.g., Ukawa and Ohtake, 1987, Ikegaya and Yamamoto, 2021], a correction of the free surface effect is also introduced. It is computed separately for P, SH and SV waves and requires assumption about the velocity models and estimations of the incident angles at different stations, which can be highly uncertain. Therefore, considering the relatively deep location of the source and the significant velocity gradient in the crust, we assume the incidence of waves is close to vertical at all stations. Another suggested correction [e.g., Ukawa and Ohtake, 1987] is aimed to account for the difference in the attenuation for P- and S-waves. This difference is rather pronounced at relatively low frequencies where the S-wave quality factor is significantly lower that the one for the P-waves. However, for waves with frequencies higher than 1 Hz propagating in the Earth's crust, the seismic attenuation is dominated by the scattering resulting in a reverse situation with the quality factor of P-waves being lower than one of S-waves [e.g., Sato, 1984]. Measurements of the high-frequency P- and S-wave quality factors in the crust [e.g., Campillo and Plantet, 1991, Yoshimoto et al., 1993] produced a wide range of values showing that on average their ratio is not very different from the V_P/V_S ratio. In this situation, the distance dependent attenuation of P- and S-waves remains nearly identical. This scattering regime in the crust corresponds well with the DLP emitted waves. Therefore, we do not apply the attenuation correction.

With the considerations described above, full displacements of P and S waves are expressed as:

$$\begin{aligned} A_P^{calc} &= \sqrt{(u_x^P)^2 + (u_y^P)^2 + (u_z^P)^2} \\ A_S^{calc} &= \sqrt{(u_x^S)^2 + (u_y^S)^2 + (u_z^S)^2} \end{aligned} \tag{3.7}$$

and their ratios can be simply computed and compared with the observations.

3.4.2 Measuring S-to-P amplitudes ratios from real seismograms

Before measuring real observed amplitudes, seismograms were preprocessed with removing instrument response and filtering them in 1–4 Hz frequency band (Figure 3.2). Instead of measuring amplitudes of P and S waves separately at different channels, these values were obtained after calculating the 3-component amplitude as:

$$u^{3C} = \sqrt{u_N^2 + u_E^2 + u_Z^2} \tag{3.8}$$

Then, we smooth the amplitude with a 1 s moving window (Figure 3.4a). The resulting envelope was manually processed in order to limit time intervals of P- and S-waves arrivals (black and red dashed lines in Figure 3.4b correspondingly) and then pick A_P and A_S as maximum values within defined time intervals.

3.4.3 Correction for the site amplifications

Raw A_S/A_P measurements shown in Figure 3.4b contain some very elevated values. The reason for this is that they are strongly affected by site amplifications. Because of the geological structure of the studied area with the majority of its territory covered by soft sediments of very recent volcanic deposits [Green et al., 2020], most of sites on which the stations of the KISS experiment were installed are prone to strong amplification of seismic waves. This amplification does not affect equally the P- and the S-waves, which can introduce a significant bias into raw A_S/A_P measurements.

Therefore, we estimated the site amplifications for P- and S-waves separately and used these estimations to correct the amplitude ratio measurements. Details of this analysis are described in 3.B. The S-wave amplification factors were estimated from coda of relatively strong regional earthquakes. For P-waves, we used a few earthquakes occurred approximately beneath the network. The site amplification measurements are summarized in Table 3.B.2. After computing the average amplification factors for each station AF_S^i and AF_P^i, we correct the raw measurements as:

$$\lg\left(\frac{A_S}{A_P}\right)_i^{obs} = \lg\left(\frac{A_S}{A_P}\right)_i^{raw} - \lg\left(\frac{AF_S^i}{AF_P^i}\right) \tag{3.9}$$

An example of corrected measurements at a single station is shown in Figure 3.4. Logarithms of amplitude ratios before and after removing site amplification at all stations are presented in left and right panels of Figure 3.5 correspondingly.

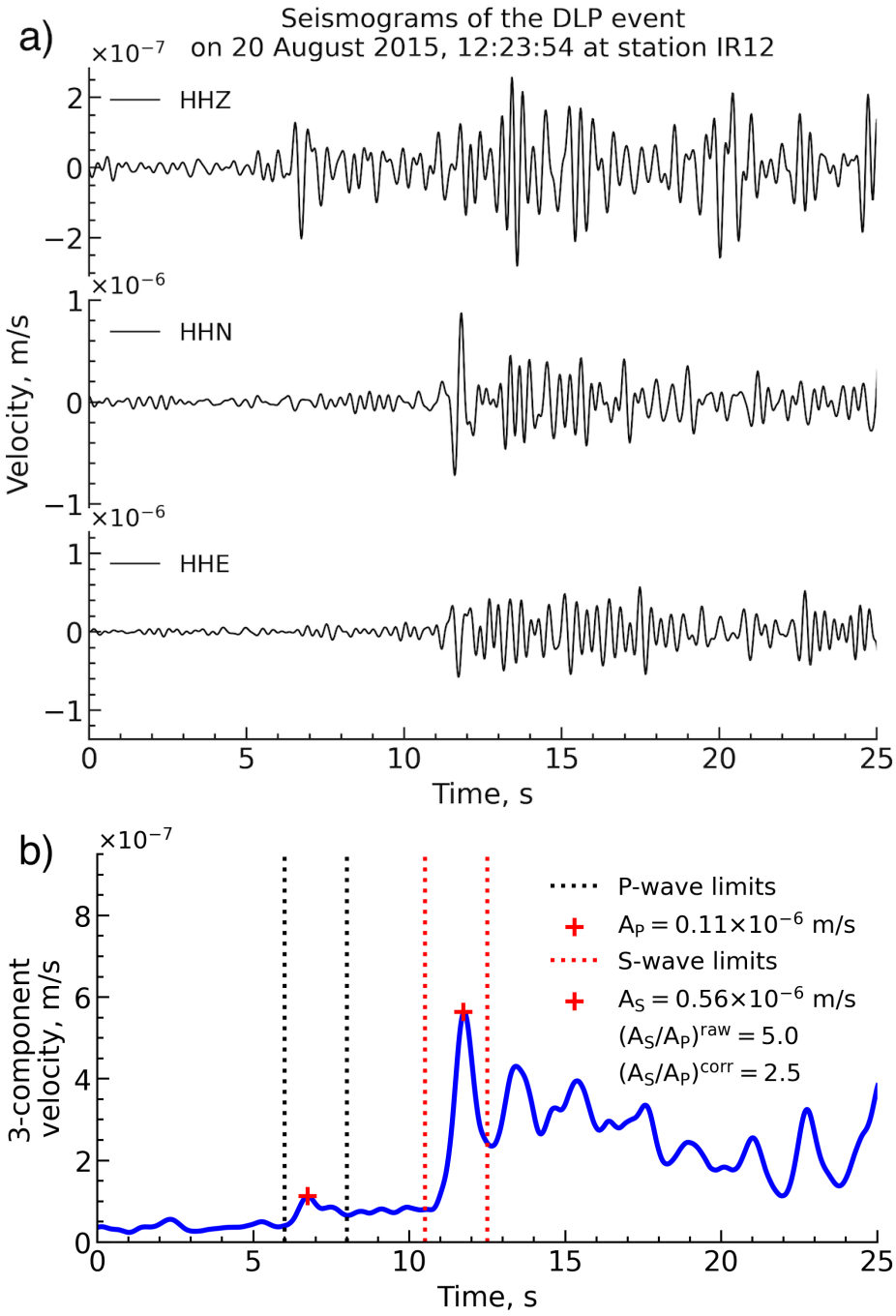

Figure 3.4: Example of measurements of P- and S-waves amplitudes for the DLP event on August 20, 2015 at 12:23:54 at station IR12. a) Seismograms after removing the instrument response and band-pass filtering (1–4 Hz). b) A three-component amplitude envelope smoothed with a 1 s window. Black and red vertical dotted lines indicate windows used for P and S waves. Maxima for P- and S-waves are indicated with red crosses.

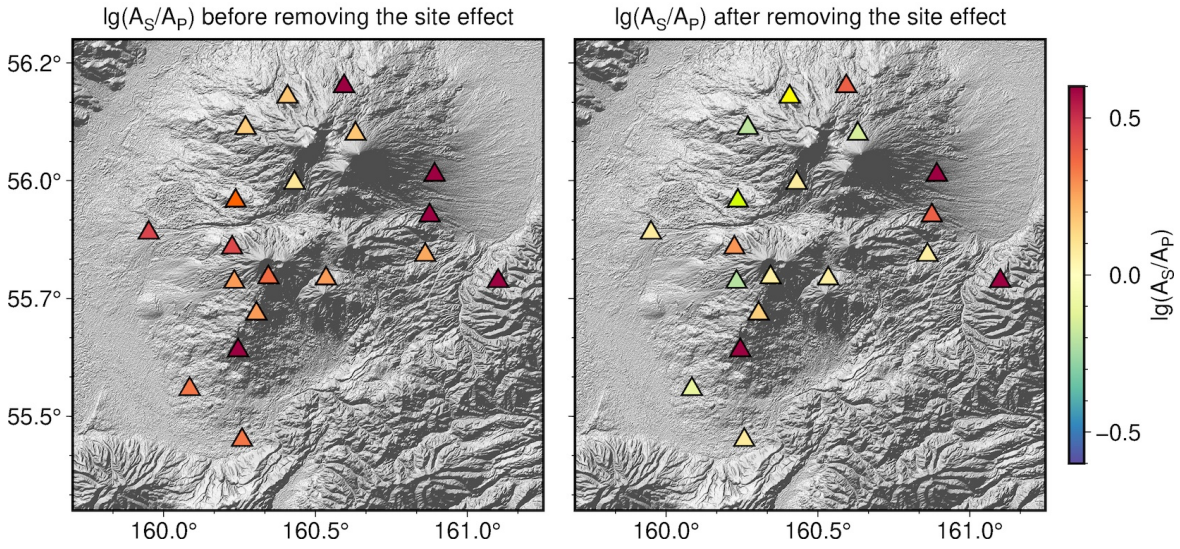

Figure 3.5: Logarithms of S-to-P amplitude ratios measured at various stations for a DLP earthquake occurred on August 20, 2015 at 12:23:54. (a) Raw measurements (b) Values after correcting for the site amplification.

3.4.4 Comparison of the observed and calculated amplitude ratios

We study source mechanisms of the earthquakes located by Senyukov et al. [2021]. For a known source location and a fixed mechanism orientation ξ (that can be described by two or three angles depending on the considered source type, 3.A), a residual between observations and a model at station i is estimated as L_1 norm:

$$\Delta_i(\xi) = \left| \lg \left(\frac{A_S}{A_P} \right)_i^{obs} - \lg \left(\frac{A_S(\xi)}{A_P(\xi)} \right)^{calc} \right| \tag{3.10}$$

And the overall misfit function over the entire network of stations is:

$$M_{L_1}(\xi) = \frac{1}{N_{st}} \sum_{i=1}^{N_{st}} \Delta_i(\xi) \tag{3.11}$$

To find the minimum of the misfit function, we perform a grid search in the domain of the mechanism orientations ξ. For mechanisms determined with two angles, the grid is designed to distribute the points homogeneously over the hemisphere (3.C, Figure 3.C.1). The third parameter (rake angle in the case of a shear-slip source and component ratio in the case of source represented by a combination of a horizontal crack and a force) was sampled homogeneously to construct a three dimensional grid (3.C, Figure 3.C.2).

3.5 Results

3.5.1 An example of a single DLP earthquake

We start with showing the source inversion results for a DLP earthquake occurred on August 20, 2015 at 12:23:54. Its seismograms are presented in Figure 3.2 and distribution of observed amplitude ratios in Figure 3.5. Figure 3.6 shows the comparison of the observed amplitude ratios with those computed for the best-fitting orientations for the different mechanisms. The corresponding distributions of the misfits as a function of orientation angles are shown in Figure 3.7.

The results of the inversion for the considered earthquake show that the shear-slip mechanism can be reasonably excluded because of its relatively large misfit (Figure 3.7) comparing to other types of "elementary" mechanisms. This difference is especially significant considering that the slip mechanism is described by three free parameters while all other only by two.

Best misfit values are found for mechanisms with volumetric changes, with a slight advantage for the tensile crack over the pipe. However, the misfit distribution for the crack mechanisms is very irregular. Additionally, all minima of this distribution are rather far from the center, i.e., from the horizontally oriented sill. Such horizontally oriented structures could expected beneath Klyuchevskoy in association with the near-Moho magamatic reservoir [Levin et al., 2014] likely formed by underplating [Annen et al., 2006]. Therefore, we also test a configuration when a sudden pressure increase within a horizontal sill is released in connected conduit through which it pushes the magma resulting in a viscous drag force [Ukawa and Ohtake, 1987]. Such "complex" source modeled as a combination of a horizontal tensile crack and arbitrary oriented single force results in the overall best misfit among all considered mechanisms. At the same time its description requires three free parameters: two force angles and a scaling coefficient between the force and the crack components (Equation 3.12), making its apparent advantage less significant.

$$r = \frac{A_{crack}}{A_{crack} + A_{force}} \tag{3.12}$$

Overall, we can conclude the DLP mechanisms contain a significant volumetric and/or single-force component. At the same time, the difference between different mechanisms with such components is not sufficient to unambiguously select one of the scenarios. Moreover, the distributions of the misfit values in the source parameters space may contain several minima (Figure 3.7) making the final inference highly uncertain. This confirms the assessment from some previous studies [Hardebeck and

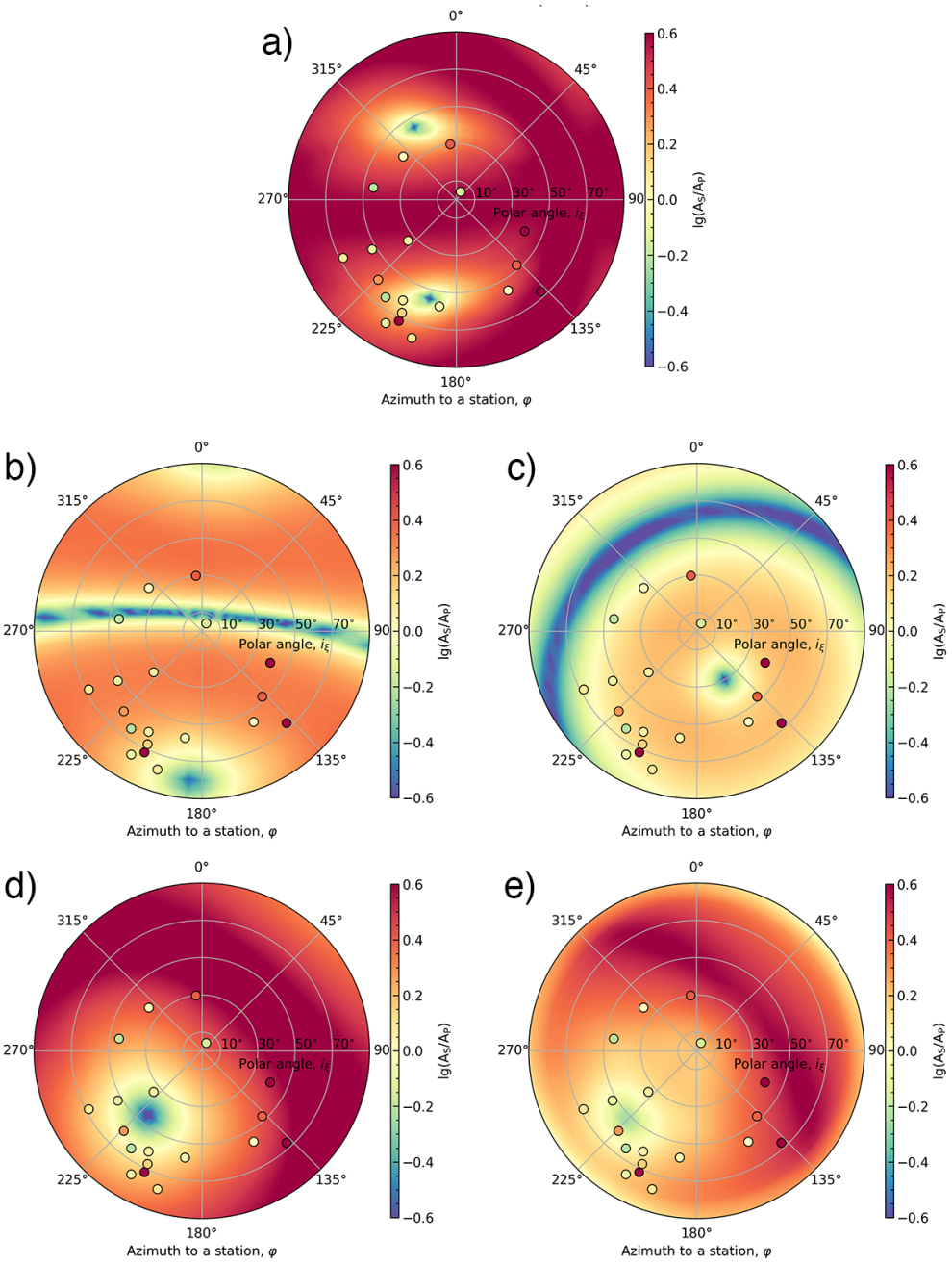

Figure 3.6: Distribution of computed logarithms of amplitude ratios and observed values at stations (shown in circles) for best-fit orientations of considered source mechanisms (an example on a DLP on August 20, 2015 at 12:23:54). (a) Shear fault: strike $\phi_s = 85°$, dip $\delta = 84°$, rake $\lambda = 66°$; (b) Tensile crack: azimuth $\phi = 184°$, dip $\theta = 81°$; (c) Cylindrical pipe: azimuth $\phi = 150°$, dip $\theta = 30°$; (d) Single force: azimuth $\phi = 220°$, dip $\theta = 45°$; (e) Combination of a horizontal tensile crack and a single force with orientation: azimuth $\phi = 228°$, dip $\theta = 51°$;

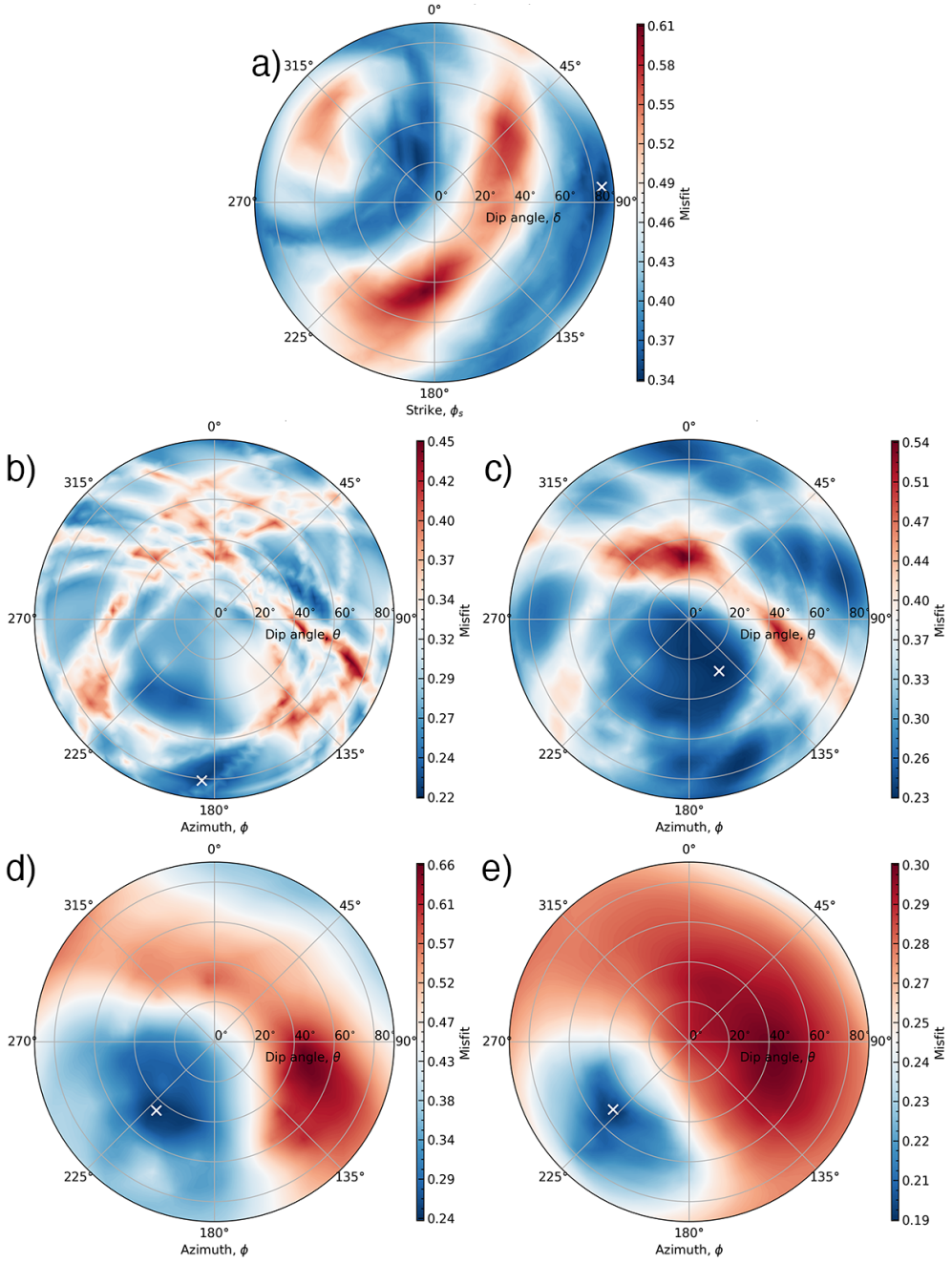

Figure 3.7: Distribution of misfits for considered source mechanisms (example of a DLP earthquake on August 20, 2015 at 12:23:54). Minima of misfits are shown with white crosses and have the next values: (a) Shear fault $M_{L_1} = 0.34$; (b) Tensile crack $M_{L_1} = 0.22$; (c) Cylindrical pipe $M_{L_1} = 0.23$; (d) Single force $M_{L_1} = 0.24$; (e) Combination of a horizontal tensile crack and a single force $M_{L_1} = 0.19$

Shearer, 2003] that the information contained in the S-to-P amplitude ratios is too uncertain to constrain the details of individual source mechanisms. Therefore, we try to make inferences based on the analysis of all sufficiently strong DLP earthquakes recorded during the KISS experiment.

3.5.2 Application to all selected DLP earthquakes

29 selected DLP earthquakes (Table 3.1) were recorded by minimum 9 and maximum 19 stations (Figure 3.8) depending on the signal-to-noise ratio of signals and data availability.

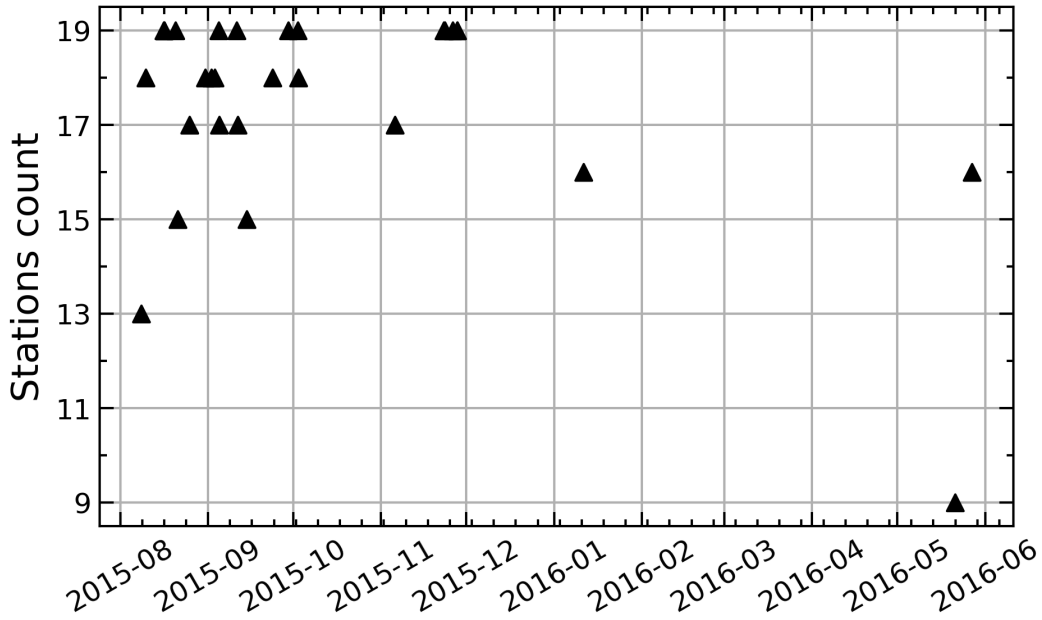

Figure 3.8: The number of stations used in processing of each of the selected DLPs

Figure 3.9 shows the values of misfit functions estimated for different types of mechanisms for the selected DLP earthquakes. The average values for each mechanism type are shown in the box of Figure 3.9.

Similar to the example of a single event, the pure shear slip mechanisms result in the worst overall misfit and can, therefore, be excluded. In the next step we analyze the consistence of the inferred source parameters over the ensemble of studied earthquakes. For a single force mechanism, both angles defining its orientation remain quite stable in time as shown in Figure 3.10. On average, they correspond to a force directed to the south-southwest and inclined ∼40 degrees relative to the vertical.

While these angles are more scattered for the pipe mechanism (Figure 3.11), its average southward orientation can still be deduced.

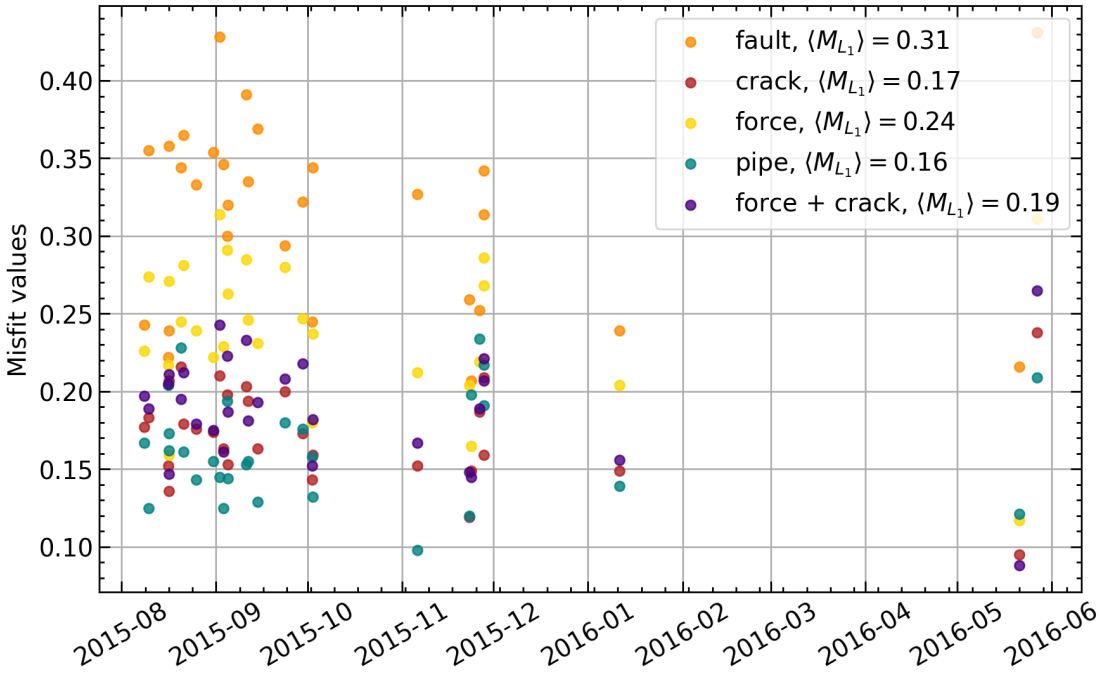

Figure 3.9: Distributions of misfits in time depending on the chosen source mechanism (shown with different colors as indicated in the legend).

The inferred angles of tensile cracks vary very strongly between different individual DLP earthquakes (Figure 3.12) and no preferred average orientation could be traced.

For the "combined" source (Figure 3.13), the force south-southwest azimuth remain highly stable while the polar angle exhibits considerable scattering.

Variations of the DLP source parameters in time may reflect the variability of the underlain physical process. Alternatively, these variations could be related to the unstable inversion procedure. As shown in the previous section (Figure 3.7), positions of the misfit minima in the parameter space are poorly constrained for a single DLP earthquake for such mechanisms as the tensile crack or the pipe which could be the cause of the observed "apparent" variation of the source parameters in time.

3.5.3 Stacked misfit distributions

If the generation of the DLP earthquake is related to the preferential magma pathways, these might be expected to be relatively stable and not varying strongly over short times. Such time stationarity of the DLP generating mechanism is partially confirmed by the high level of similarity of their waveforms over series of many events, i.e., the multiplet behavior [Shapiro et al., 2017a]. To test the hypothesis of stationary processes of generation of DLP earthquakes, we decided to compute "stacked" misfit

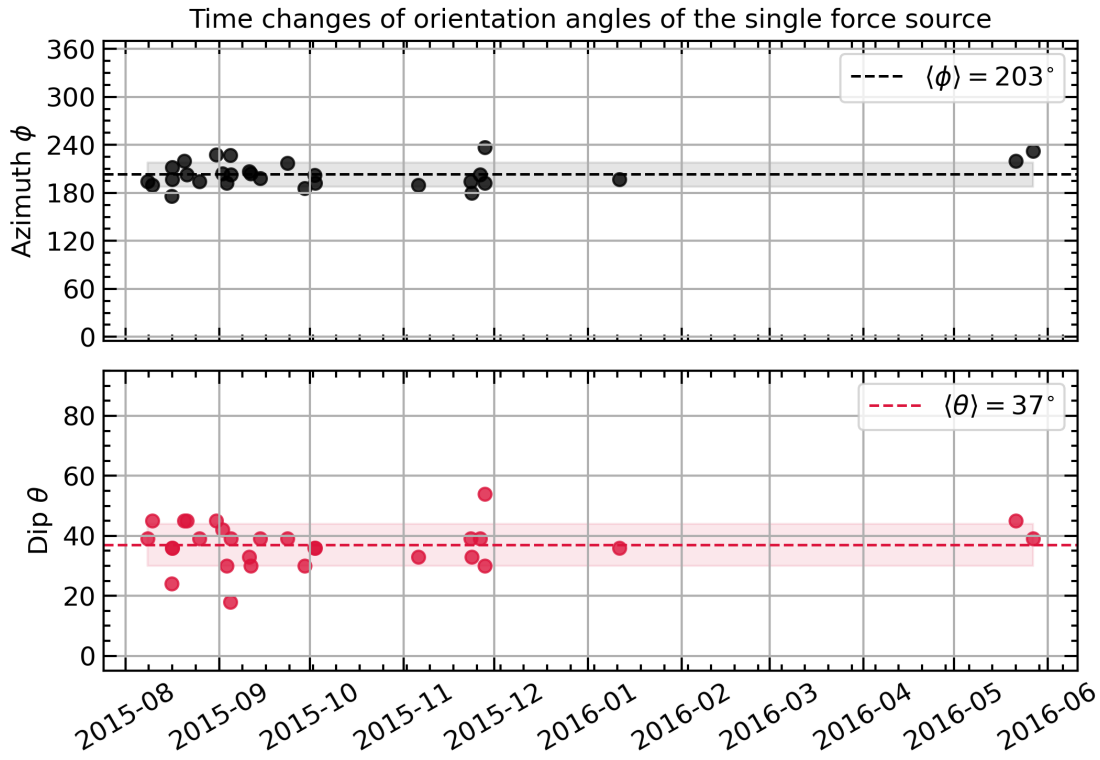

Figure 3.10: Temporal changes of the orientation of a single force vector in space: (a) Azimuth ϕ (b) Dip angle θ. Average values and standard deviations are shown by dashed lines and colored areas, respectively.

distributions for all selected events ($N_{DLP} = 29$):

$$M_{L_1}^{stack}(\phi,\theta) = \frac{1}{N_{st}^{cum}} \sum_{i=1}^{N_{DLP}} N_{st}^{i} \cdot M_{L_1}^{i}(\phi,\theta) \tag{3.13}$$

where $M_{L_1}^{i}(\phi,\theta)$ is a misfit distribution in the azimuth-polar angle plane corresponding to one of 29 selected DLPs (i.e., Figure 3.7). N_{st}^{i} is a number of stations used for DLP i. $N_{st}^{cum} = \sum_i N_{st}^{i}$ is the total number of used seismic records. For the "elementary" source mechanisms, equation 3.13 implies a simple stacking in the polar coordinate plane. For the combined sill-force source, an additional degree of freedom is still present because the third parameter r (equation 3.12) is allowed to vary for different event and angles. Physically, this would correspond to a situation when the geometry of the DLP generating part of the plumbing system remains fixed and the partitioning of energy between the pressure increase in the sill and the viscous drag in the conduit can be variable between different DLP earthquakes.

Figure 3.14 shows the distributions of "stacked" misfits for different source mechanisms. In comparison with the results for single DLP earthquakes (Figures 3.7 and 3.9) the minimum misfit values remain similar, implying that the consistency of whole

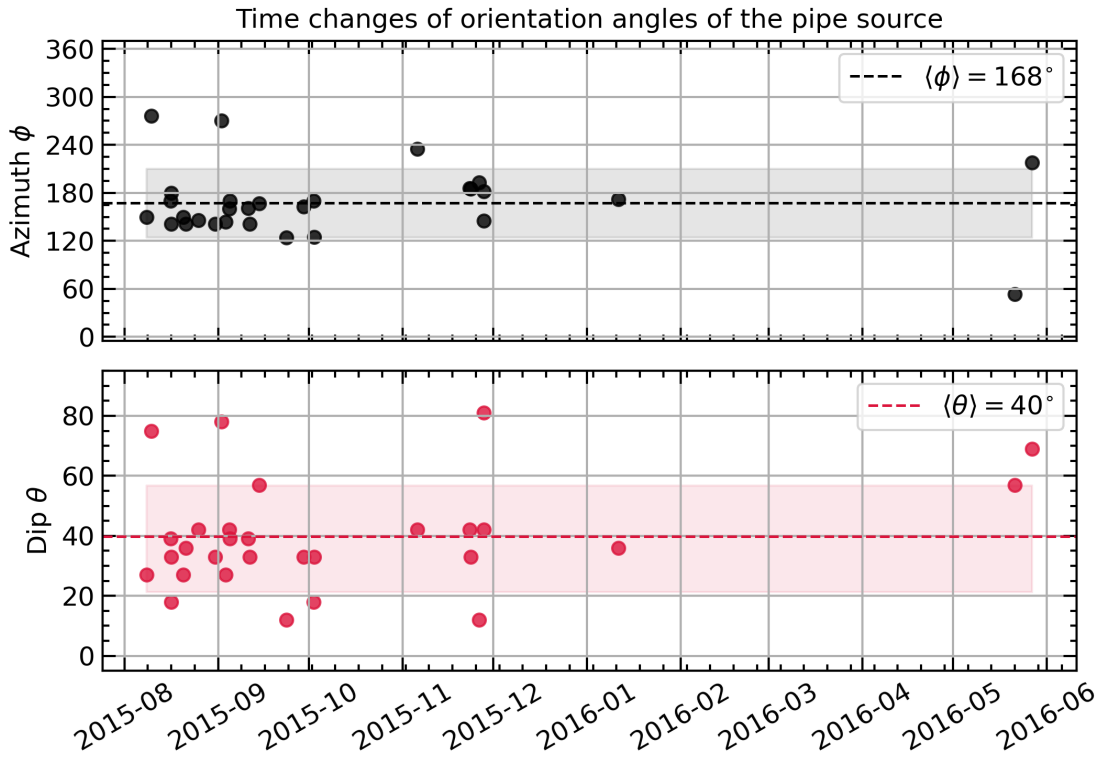

Figure 3.11: Temporal changes of the orientation of a cylindrical pipe axis: (a) Azimuth ϕ (b) Dip angle θ. Average values and standard deviations are shown by dashed lines and colored areas, respectively.

ensemble of observations is at the same level as for individual events.

The misfit distributions did not change significantly for the single force and "combined" mechanism indicating the preferential south-southwest direction. The distribution has been "stabilized" (the minima more clearly defined) for the pipe mechanism and shows more clearly the southward orientation. The "tensile crack" solution still remains unstable showing several misfit minima with close values. In terms of the absolute misfit minima values, the best solutions are obtained with the pipe and combined source mechanisms.

3.6 Discussion and Conclusions

Understanding physical processes leading to generation of DLP earthquakes remains an elusive task. Inferring their kinematic source mechanisms based on observed seismic waveforms could provide a key constraint to discriminate between different existing hypotheses. With this in mind, we explored the records of 29 DLP earthquakes recorded by broadband stations of the KISS temporary seismic experiment in the region of the Klyuchevskoy volcanic group in Kamchatka, Russia [Shapiro et al., 2017b].

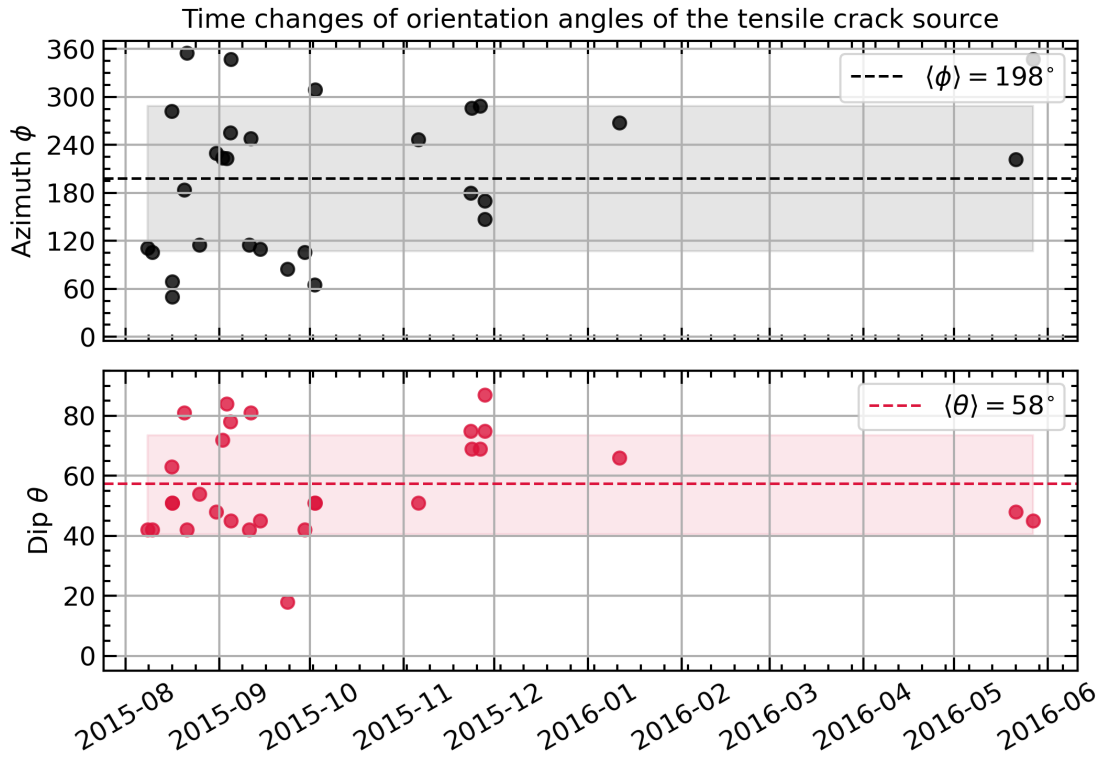

Figure 3.12: Temporal changes of the orientation of the normal to a tensile crack surface: (a) Azimuth ϕ (b) Dip angle θ. Average values and standard deviations are shown by dashed lines and colored areas, respectively.

We first argued that the methods based on the phase of seismograms (inversion of full waveforms or polarities of first arrivals) are not practical in the case of DLP earthquakes because their relatively weak signals emerge from the noise only after band-pass filtering. In particular, the strong microseismic noise is removed with high-pass filters with corner frequency near 1 Hz, which is very close to the dominant frequencies of LP waveforms. As a result, phases of filtered waveforms can be strongly distorted and their apparent polarities reversed.

Consequently, we explored a method based on the amplitude measurements. Namely, as has been suggested in some previous studies, we use the amplitude ratios of P- and S-waves [e.g., Ukawa and Ohtake, 1987, Ohmi and Obara, 2002, Nakamichi et al., 2003, Ikegaya and Yamamoto, 2021]. This type of measurement can be biased by possibly different surface site amplification of P- and S-waves. We implemented a correction for this effect with measuring S-wave and P-wave site amplifications (3.B) from a few regional earthquakes well recorded by the whole network. The estimated correction factors result on average decrease of the raw S-to-P amplitude ratios and are subject to significant uncertainties. Additional uncertainties in the inversion of the amplitude data arise from the only approximately known model of the wave propagation.

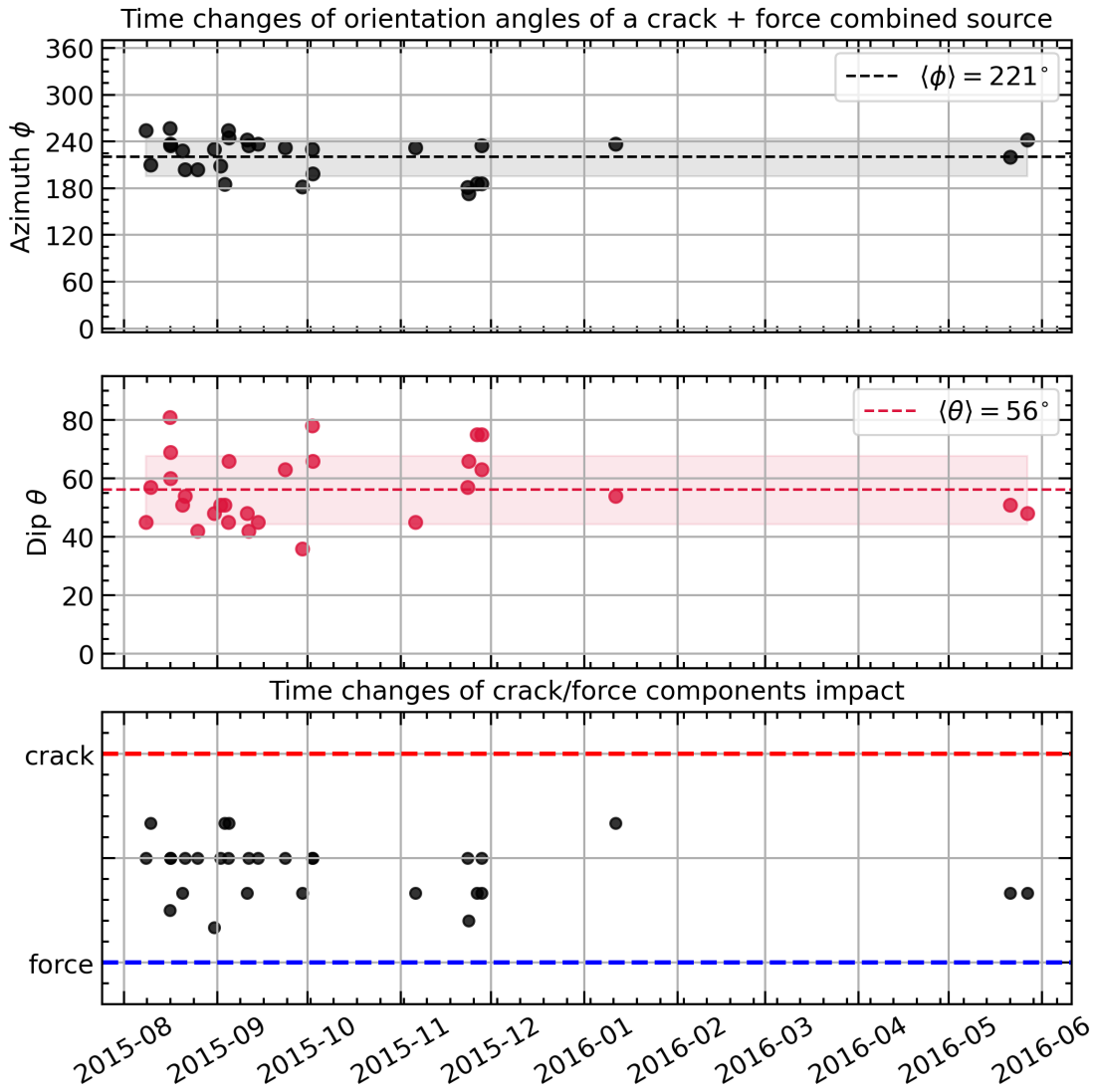

Figure 3.13: Temporal changes of the vector orientation of a single force compound of a "combined" source in space: (a) Azimuth ϕ (b) Dip angle θ. Average values and standard deviations are shown by dashed lines and colored areas, respectively. (c) Temporal changes of a crack to force ratio (Equation 3.12)

The mentioned uncertainties in the data and in the forward model contribute to the non-uniqueness of the inversion procedure. To mitigate this non-uniqueness, we decided to consider source models with a limited number of free parameters. So far, a complete kinematic description of a volcanic earthquake source would be obtained with a combination of a moment tensor and a single force [e.g., Kumagai, 2009], which would require nine independent parameters (three force components and six independent components of the moment tensor). Considering the limitations of our amplitude measurements, instead of trying to resolve all these parameters simultaneously, we test a set of simplified source mechanisms (involving 2 or 3 parameters) that represent

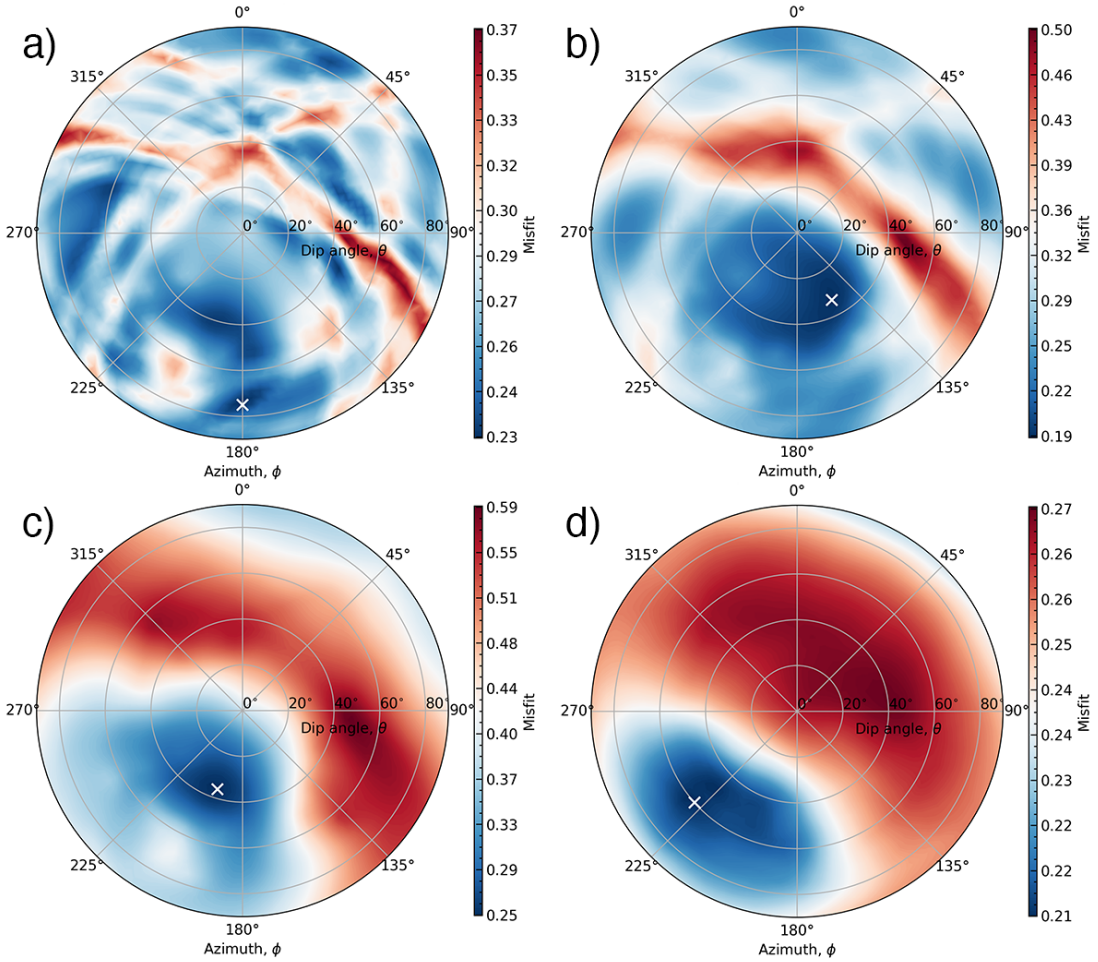

Figure 3.14: Misfit distributions "stacked" for all selected DLP events for different source mechanisms. The misfits minima are shown with white crosses and have the following values: (a) Tensile crack $M_{L_1}^{stack} = 0.23$; (b) Cylindrical pipe $M_{L_1}^{stack} = 0.19$; (c) Single force $M_{L_1}^{stack} = 0.25$; (d) Combination of a horizontal tensile crack and a single force $M_{L_1}^{stack} = 0.21$

"elementary" physical processes that may generate seismic waves: (1) slip on faults, (2) magma pressure variation within sills, dykes, or pipe-shaped conduits, and (3) viscous drag force caused by acceleration of magma flow in a conduit.

Even with such a reduced number of parameters, the inversion of individual DLP earthquakes remains non-unique for most of types of mechanisms which is manifested by multiple minima in the misfit distributions (Figure 3.7). To further stabilize the inversion, we consider the hypotheses of time stationarity of the DLP generating processes based on which we invert measurements for all studied DLP earthquakes simultaneously (via stacking the misfit distributions obtained for individual events). As a result of this procedure, the best-misfit levels are not deteriorated indicating the

reasonable likelihood of a single mechanism explaining all observations. The misfit distributions show relatively well defined minima for all mechanisms except the tensile crack.

Comparison of the minimum misfit levels for different types of studied mechanisms allows us to conclude that the generation of DLP earthquakes beneath the Klyuchevskoy volcano by slip on faults is not consistent with the observed S-to-P amplitude ratios whose explanation requires mechanisms with volumetric or single force components. For such mechanisms (tensile crack, pipe, and single force) the difference in misfit to seismic data is not very strong. This implies that seismic data alone is not sufficient to unambiguously constrain the details of the source mechanisms. At the same time, the plausibility of results of different seismic inversions can be interpreted based on some additional hypotheses about the functioning of the KVG volcano plumbing system.

So far, there are two reasons to consider that solutions obtained with a pure tensile crack mechanism are less likely (even if they cannot be definitely excluded). First, the strong difference in solutions for individual events (Figure 3.12) is not consistent with the time stationarity that may be expected for the geometry of plumbing system. Second, the average and most individual best-fit solutions correspond to significantly dipping structures while in the vicinity of the crust-mantle boundary nearly horizontal sills are more likely.

Results of inversion for the three remaining tested mechanisms show some consistency. Namely, they all indicate a south-southwest dipping magmatic conduit generating either pipe-shaped pressure perturbations or an along-conduit force. This leads to a possible interpretation shown in Figure 3.15. The DLP earthquakes are generated near the crust-mantle boundary (Moho) approximately beneath the Klyuchevskoy volcano, where the main deep magmatic reservoir feeding the KVG volcanoes is located [e.g., Levin et al., 2014]. This magma storage is likely shaped as a complex of underplated sills. From this deep reservoir magma penetrates the crust through a south-southwest dipping conduit (or a system of conduits). This orientation of dominant magma pathways would be in agreement with the overall interconnection of the transcrustal KVG plumbing system when the pressure perturbations in the deep reservoir are transmitted not only to Klyuchevskoy but also to the active volcanoes located south of it, Bezymianny and Tolbachik [Fedotov et al., 2010, Shapiro et al., 2017a, Coppola et al., 2021, Journeau et al., 2022].

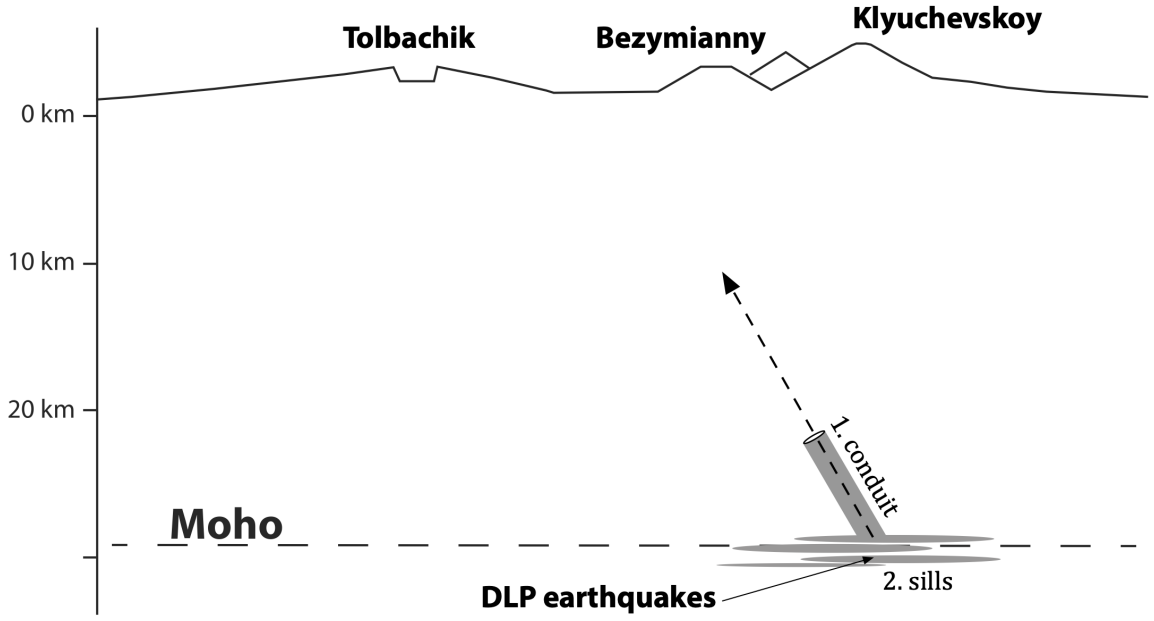

Figure 3.15: Possible interpretation based on the obtained results schematically shown on a SSW-NNE oriented vertical cross-section indicating positions of three active KVG volcanoes.

Appendices to Chapter 3

3.A Theoretical radiation patterns of P and S waves

Radiation patterns appearing in Equations 3.1-3.6 can be expressed via the direction cosines γ (from a source to a receiver):

$$R_{i,j}^{FP} = \gamma_j \gamma_j \tag{3.A.14}$$

$$R_{i,j}^{FS} = \gamma_j \gamma_j - \delta_{i,j} \tag{3.A.15}$$

$$R_{i,j,k}^{MP} = \gamma_j \gamma_j \gamma_k \tag{3.A.16}$$

$$R_{i,j,k}^{MP} = (\gamma_j \gamma_j - \delta_{i,j})\gamma_k \tag{3.A.17}$$

We use Cartesian coordinates with X axis directed toward the North, Y - toward the East, and Z - upward (Figure 3.A.1). Direction cosines can be calculated with known direction from a source to a receiver expressed via two radial angles: azimuth $\varphi \in [0°, 360°]$ measured clockwise from the North and inclination $i_\xi \in [0°, 180°]$ measured from the upward vertical:

$$\begin{cases} \gamma_x = \sin i_\xi \cos \varphi \\ \gamma_y = \sin i_\xi \sin \varphi \\ \gamma_z = \cos i_\xi \end{cases} \tag{3.A.18}$$

Single force

The orientation of a single force described with two angles: azimuth $\phi_f \in [0°, 360°]$ measured clockwise from the North and dip angle $\theta_f \in [0°, 90°]$ measured from the upward vertical (Figure 3.3a). Three components of a unit force can be expressed as projections on three axes:

$$\begin{cases} F_x = \sin \theta_f \cos \phi_f \\ F_y = \sin \theta_f \sin \phi_f \\ F_z = \cos \theta_f \end{cases} \tag{3.A.19}$$

Sources described by a moment tensor

Cylindrical pipe. Similar to a single force, its orientation is described by two angles: azimuth $\phi_p \in [0°, 360°]$ measured clockwise from the North and dip angle $\theta_p \in [0°, 90°]$ measured from the upward vertical (Figure 3.3b). The corresponding components of the normalized moment rate tensor are [Kumagai, 2009]:

$$\begin{aligned}
M_{xx} &= M_0\,(1 + \cos^2\theta_p \cos^2\phi_p + \sin^2\phi_p) \\
M_{xy} &= -M_0\,\sin^2\theta_p \sin\phi_p \cos\phi_p \\
M_{xz} &= -M_0\,\sin\theta_p \cos\theta_p \cos\phi_p \\
M_{yy} &= M_0\,(1 + \cos^2\theta_p \sin^2\phi_p + \cos^2\phi_p) \\
M_{yz} &= M_0\,-\sin\theta_p \cos\theta_p \sin\phi_p \\
M_{zz} &= M_0\,(1 + \sin^2\theta_p)
\end{aligned} \tag{3.A.20}$$

Shear slip is defined by three angles: ϕ_s, δ and λ

$$\begin{aligned}
M_{xx} &= -M_0\,(\sin\delta\cos\lambda\sin 2\phi_s + \sin 2\delta\,\sin\lambda\,\sin^2\phi_s) \\
M_{xy} &= M_0\,(\sin\delta\cos\lambda\cos 2\phi_s + \frac{1}{2}\sin 2\delta\sin\lambda\sin 2\phi_s) \\
M_{xz} &= -M_0\,(\cos\delta\cos\lambda\cos\phi_s + \cos 2\delta\sin\lambda\sin\phi_s) \\
M_{yy} &= M_0\,(\sin\delta\cos\lambda\sin 2\phi_s - \sin 2\delta\sin\lambda\cos^2\phi_s) \\
M_{yz} &= -M_0\,(\cos\delta\cos\lambda\sin\phi_s - \cos 2\delta\sin\lambda\cos\phi_s) \\
M_{zz} &= M_0\,(\sin 2\delta\sin\lambda)
\end{aligned} \tag{3.A.21}$$

Tensile crack is oriented by two angles: ϕ_c and θ_c

$$\begin{aligned}
M_{xx} &= M_0\,(1 + 2\sin^2\theta_c \sin^2\phi_c) \\
M_{xy} &= -M_0\,(\sin^2\theta_c \sin 2\phi_c) \\
M_{xz} &= M_0\,\sin 2\theta_c \sin\phi_c \\
M_{yy} &= M_0\,(1 + 2\sin^2\theta_c \cos^2\phi_c) \\
M_{yz} &= -M_0\,(\sin 2\theta_c \cos\phi_c) \\
M_{zz} &= M_0\,(1 + 2\cos^2\theta_c)
\end{aligned}$$

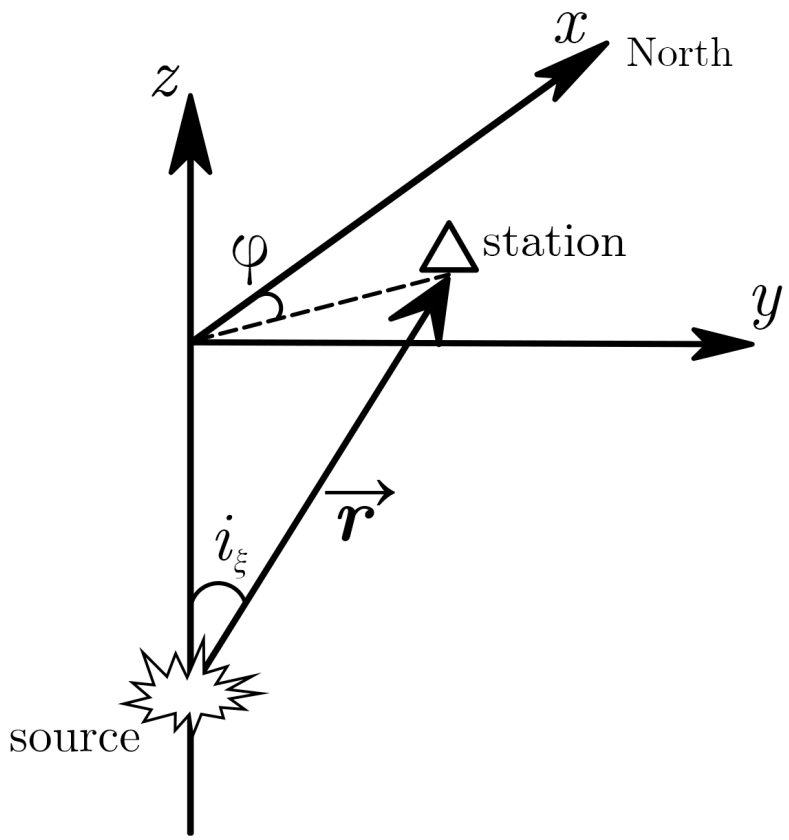

Figure 3.A.1: The coordinate system used in the study and definition of polar angles: azimuth $\varphi \in [0°, 360°]$ and inclination $i_\xi \in [0°, 180°]$

3.B Site amplification factors

3.B.1 Estimation of S-wave site amplification based on codas of seismograms

To estimate the S-wave site amplification we used the method based on coda of seismograms [Husker et al., 2010] that consist waves scattered at random heterogenities in the Earth. Because of the preferential P-to-S conversion druring the scattering of elastic waves, the seismic codas are dominated by S-wave energy [e.g., Margerin et al., 2000, Shapiro et al., 2000]. The coda envelope recorded at station i during event k can be expressed as [Aki and Chouet, 1975, Rautian and Khalturin, 1978]:

$$A_{ik}(f,t) = S_k(f)\, R_i(f)\, I_i(f)\, G_i(f,t) \tag{3.B.22}$$

where f is the frequency, t is the lapse time, $S_k(f)$ is the source spectrum, $R_i(f)$ is the frequency-dependent site amplification, $I_i(f)$ is the instrument response, $G(f,t)$

is the term describing the wave propagation (Green's function). At times significantly greater than the arrival of direct S-waves, the energy of scattered seismic becomes uniformly distributed in the media and the term $G(f,t)$ becomes approximately equal at all stations. As a consequence, in the late coda, the ratio of the envelopes (computed after removing the instrument response) stabilized in time and becomes equal to the ratio of the site amplification factors:

$$\frac{A_{ik}(f,t)}{A_{jk}(f,t)} = \frac{R_i(f)}{R_j(f)} \tag{3.B.23}$$

We apply this equation to three-component envelopes computed from seismograms filtered between 1 and 4 Hz and smoothed with 50 s long moving window to estimate average amplification coefficients in this frequency range. An example for a pair of stations is shown in Figure 3.B.1. The envelope ratios are computed within a window of 120 s length. Its starting time was chosen with visual inspection but it should be at least two times larger than the arrival of direct S waves. For the example presented in Figure 3.B.1 the starting and ending time were chosen as 110 s and 230 s correspondingly.

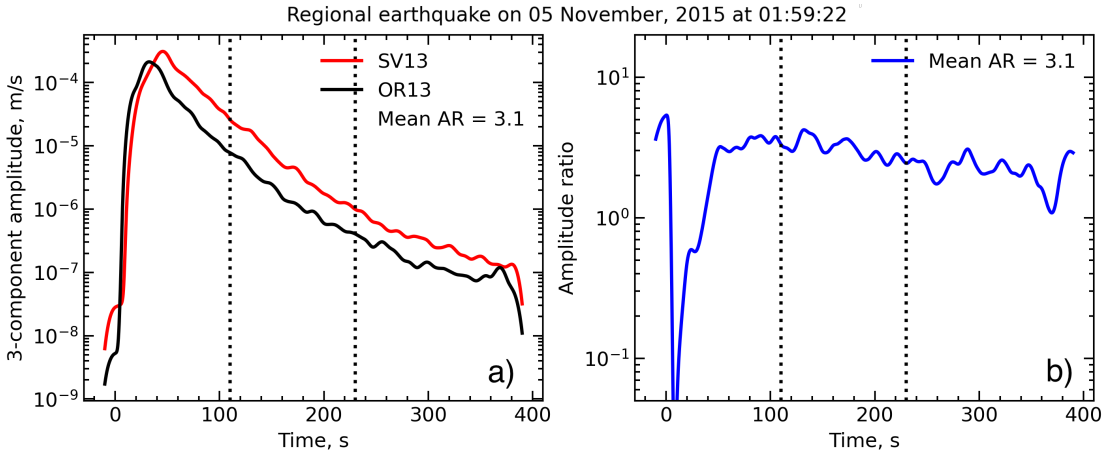

Figure 3.B.1: Example of measuring a relative amplification of S-waves between two stations from the coda of seismograms (a) Smoothed three-component envelopes. (b) Amplitude ratios between two stations. Dotted lines show the beginning and the end of 120 s long coda window.

We select 7 sufficiently strong regional earthquakes (Table 3.B.1) recorded by the majority of the KISS network. Overall, we used 53 stations. By computing amplitude ratios for all possible pairs of stations and events, we have a set of measurements that we use to constrain average site amplifications in a least square sense (after taking logarithms of equation 3.B.23). We also chose station OR13 located on a hard rock

site well outside the volcanic edifices and sedimentary layers as a reference and all site amplification coefficients are computed relative to it.

3.B.2 Site amplifications of P-waves

P-wave amplification can be estimated from fist arriving P waves not contaminated by S-wave energy rapidly appearing in the p-wave coda because of the scattering. Telesemic P-waves can be used for this purpose at relatively low frequencies. In our spectral range of interest, 1-4 Hz, all teleseismic P waves recorded suring the KISS experiment were attenuated. Therefore, we decided to use nearly vertically incident P waves from relatively deep subduction earthquakes located approximately beneath the network. We found 4 such events listed in Table 3.B.1 and shown in Figure 3.B.3. The process of calculating P-wave amplification is illustrated in Figure 3.B.2. For this example, a regional earthquake on October 16, 2015 at 06:38:23 was used. A significant source of uncertainty arises from the unknown radiation pattern of the used earthquakes. We can only assume, because of the nearly vertical incidence, the ray parameters do not differ strongly across the network and are close to the maxima of the P-wave radiation patterns expected from the subduction geometry.

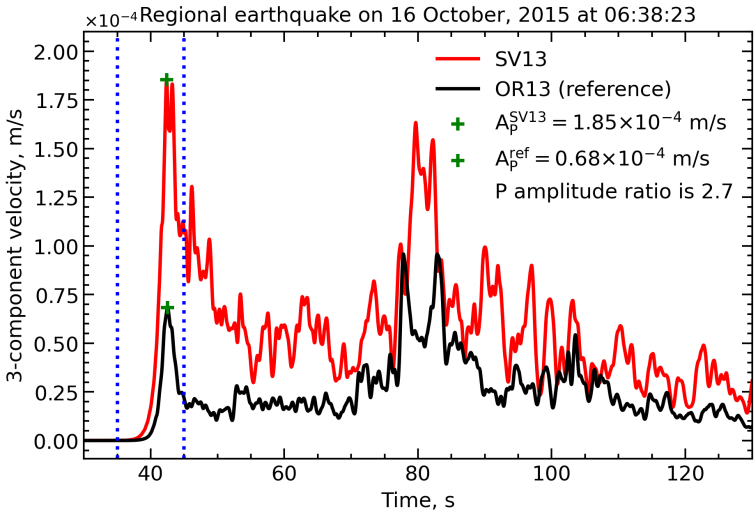

Figure 3.B.2: Measuring P-wave site amplification at station OR13, green crosses show the defined P-wave amplitudes

3.B.3 Final results

During the study period we found 4 and 7 regional tectonic events to estimate averaged amplification coefficients for P- and S-waves. Their parameters are presented in Table 3.B.1 and the epicenters are shown in Figure 3.B.3. The site amplification coefficients were estimated stations used for the source mechanism inversion for both P- and S-waves are shown in Table 3.B.2.

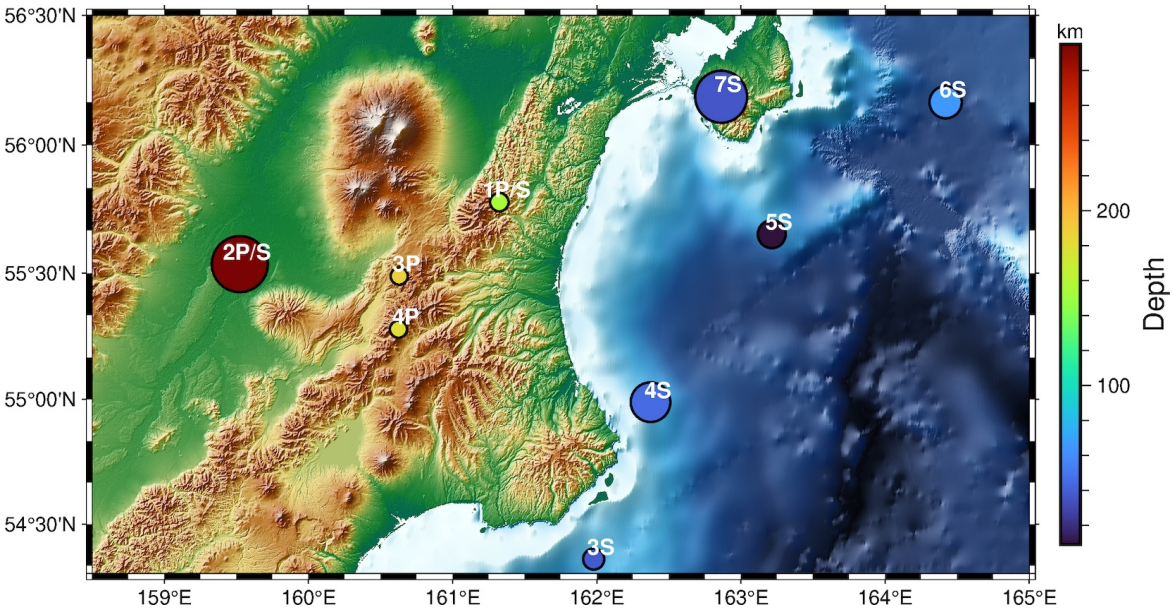

Figure 3.B.3: Map of the epicenters of the earthquakes used for measuring amplifications of P- and S-waves, their codes correspond to the codes in Table 3.B.1. The circle size represents the corresponding event magnitude.

	P-waves	S-waves
1	2015-11-02 20:02:48 (M = 4.3, z = 156 km)	
2	2015-10-16 06:38:23 (M = 5.9, z = 295 km)	
3	2015-11-18 22:37:27 (M = 4.2, z = 187 km)	2015-11-29 23:45:55 (M = 4.5, z = 38 km)
4	2016-02-08 04:25:03 (M = 4.2, z = 180 km)	2016-02-01 22:47:51 (M = 5.4, z = 45 km)
5		2015-09-29 04:33:27 (M = 4.9, z = 9 km)
6		2015-08-24 11:50:55 (M = 5.1, z = 65 km)
7		2015-11-05 01:59:22 (M = 5.8, z = 36 km)

Table 3.B.1: Regional earthquakes used for estimation of amplification coefficients.

	P-waves amplification	S-waves amplification
IR1	1.50 ± 0.81	1.98 ± 0.39
IR11	2.27 ± 0.74	4.51 ± 1.25
IR12	2.05 ± 0.82	4.20 ± 1.12
IR13	2.20 ± 0.87	3.91 ± 0.75
IR17	0.91 ± 0.66	1.40 ± 0.27
IR18	0.60 ± 0.48	1.61 ± 0.37
IR2	1.20 ± 0.53	3.33 ± 0.91
IR3	2.16 ± 2.34	3.15 ± 0.82
IR4	1.01 ± 0.48	2.62 ± 0.65
IR6	0.58 ± 0.58	1.35 ± 0.31
OR13	1.00 ± 0.00	1.00 ± 0.00
OR20	1.93 ± 0.49	3.44 ± 0.62
SV1	1.35 ± 0.69	1.76 ± 0.33
SV13	1.47 ± 0.83	3.13 ± 0.61
SV3	0.90 ± 0.39	2.30 ± 0.60
SV4	1.58 ± 2.06	3.64 ± 1.04
SV6	0.71 ± 0.47	1.42 ± 0.27
SV7	1.56 ± 0.89	3.22 ± 0.68
SV9	1.02 ± 0.82	1.07 ± 0.22

Table 3.B.2: Average amplification coefficients and their standard deviations.

3.C Discretization of the mechanism orientation space for the grid search

In order to calculate the misfit function a grid of possible source orientations should be defined.

Two parameters case

For source mechanisms whose orientation depends on two angles (a single force, a tensile crack, a cylindrical pipe) the grid of parameters is compiled by dividing a disk or a hemisphere into cells of equal area. Firstly, the range of values of the dip angle $\theta \in [0°, 90°]$ was uniformly discretized into N_θ points with a step $\Delta\theta$. At the next step the same parameter $\Delta\theta$ was used to define a number of intervals N_ϕ making up the outer circle, i.e. the one that corresponds to $\theta = 90°$:

$$N_\phi^{outer} = \frac{360°}{\Delta\theta} \tag{3.C.24}$$

After all, the number of points in every circle of θ_i was defined using a two-step procedure. Equation 3.C.25 allowed to calculate an aspect ratio a_i for a circle corresponding to θ_i.

$$a_i = \sin\,\theta_i \tag{3.C.25}$$

Then, the obtained aspect ratio was used for calculating a number of points in a circle corresponding to θ_i:

$$N_\phi^i = a_i\, N_\phi^{outer} \tag{3.C.26}$$

In this work $\Delta\theta$ was taken as 3° what provides $N_\theta = 31$ and $N_\phi^{outer} = 70$ (Figure 3.C.1). This configuration allowed to perform the grid search thorough enough but not computationally expensive at the same time.

Three parameters case

The classic shear slip model requires three angles to orient a source in space. Thus, besides strike ϕ_s and dip δ the grid of parameters becomes three dimensional due to rake angle $\lambda \in [0°, 180°]$. The range of possible rake values was discretized with a step of 3° as well. Then, the combined mechanism included an additional parameter responsible for the contribution of one or another source: a force or a horizontal crack. This ratio can be expressed as:

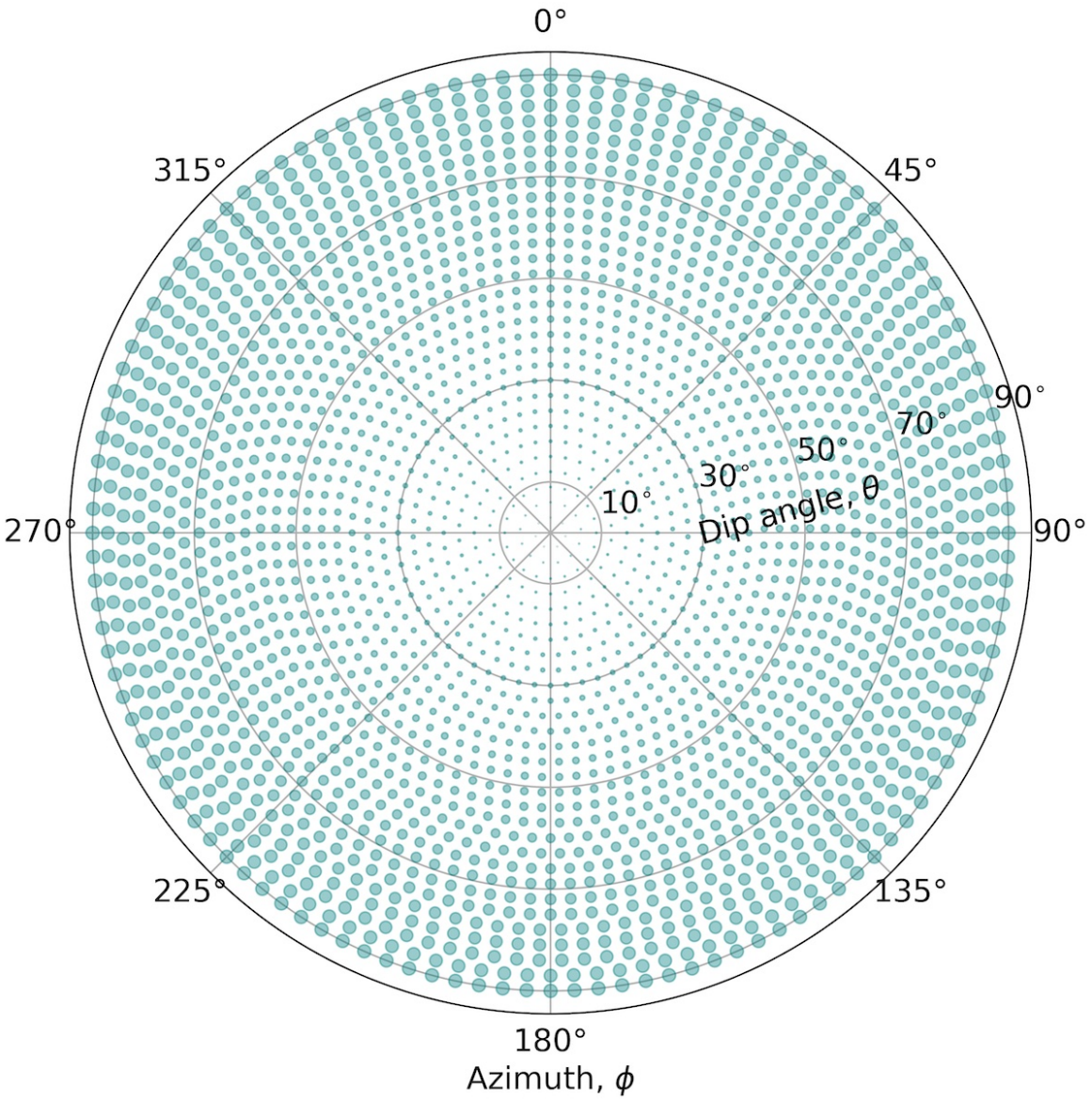

Figure 3.C.1: Two dimensional grid in polar coordinates

$$r = \frac{A_{crack}}{A_{crack} + A_{force}} \tag{3.C.27}$$

If $r \to 1$, a horizontal crack is dominating, while when $r \to 0$ the dominating mechanism is a single force in both cases. Parameter r lies in the range $[1/6, 5/6]$ or, in other words, ratio $A_{crack} : A_{force}$ varies from 1:5 to 5:1.

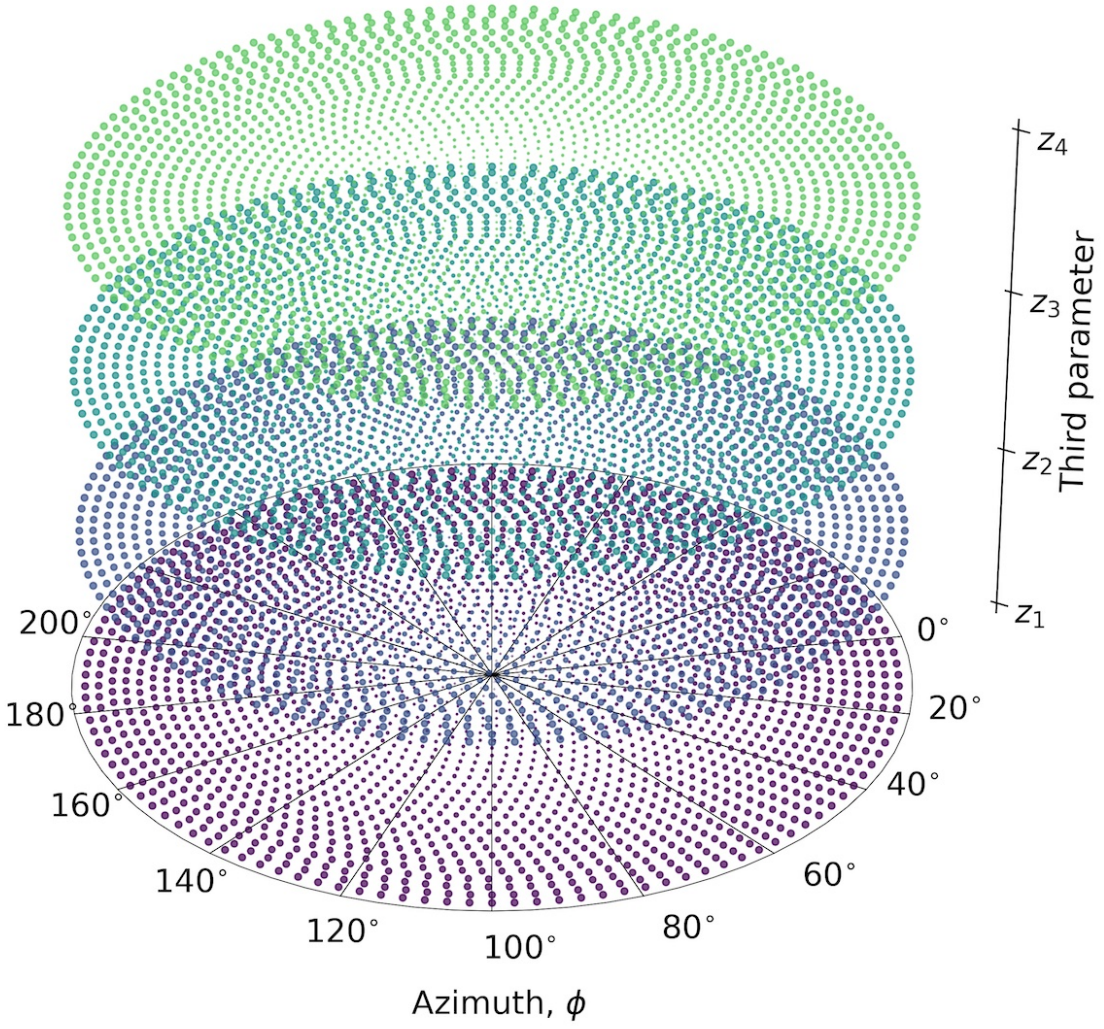

Figure 3.C.2: Three dimensional grid in polar coordinates. Vertical axis is obtained by discretizing a third parameter: either a rake angle λ or a ratio r which defines a dominating mechanism of a combined source

Part II

Studying the seismicity in the vicinity of the Klyuchevskoy Volcanic Group with machine learning

Chapter 4

Machine learning for analysis of the seismo-volcanic data

4.1 Introduction

4.1.1 Machine learning in seismology

The accelerating evolution of digital technologies gives us new tools to investigate volumes of data that grow rapidly at the same time. This phenomenon was called *big data* and one of its main consequences is the need for *machine learning* (ML) that allows to analyze these vast volumes of data. Active development of ML started in the middle of the XX century for problems such as computer vision, text and speech recognition [Turing, 1950] but during latter decades its methods have been successfully adopted to different scientific fields.

Geosciences in general and seismology in particular are data-driven disciplines. Figure 4.1 can give an idea about the amount of data in seismology and its growth rate according to the IRIS Data Data Management Center. The shown numbers are impressive and make application of machine learning promising for solving various problems. Early ML application in seismology concerned the analysis of seismic catalogs in the frameworks of the "earthquake prediction" based on identifying patterns of clustering of small- and intermediate-scale seismicity before large earthquakes [e.g., Keilis-Borok et al., 1988]. More recently, especially during the last decade, the development of seismological ML applications rapidly accelerated with many of them being applied directly to the "raw data", i.e., to the digital seismograms. A good review of the recent ML applications in seismology is provided by Mousavi and Beroza [2023]. These authors discuss applications to various problems that are grouped in four main categories (Figure 4.2).

The largest category concerns the automation of the seismological data processing. This implies training the ML algorithms to execute and accelerate different "standard" tasks of the seismic event analysis workflow such as phase picking, phase association, event detection, event discrimination, event classification, etc. Next two categories concern solving different forward and inverse problems in seismology. The last category

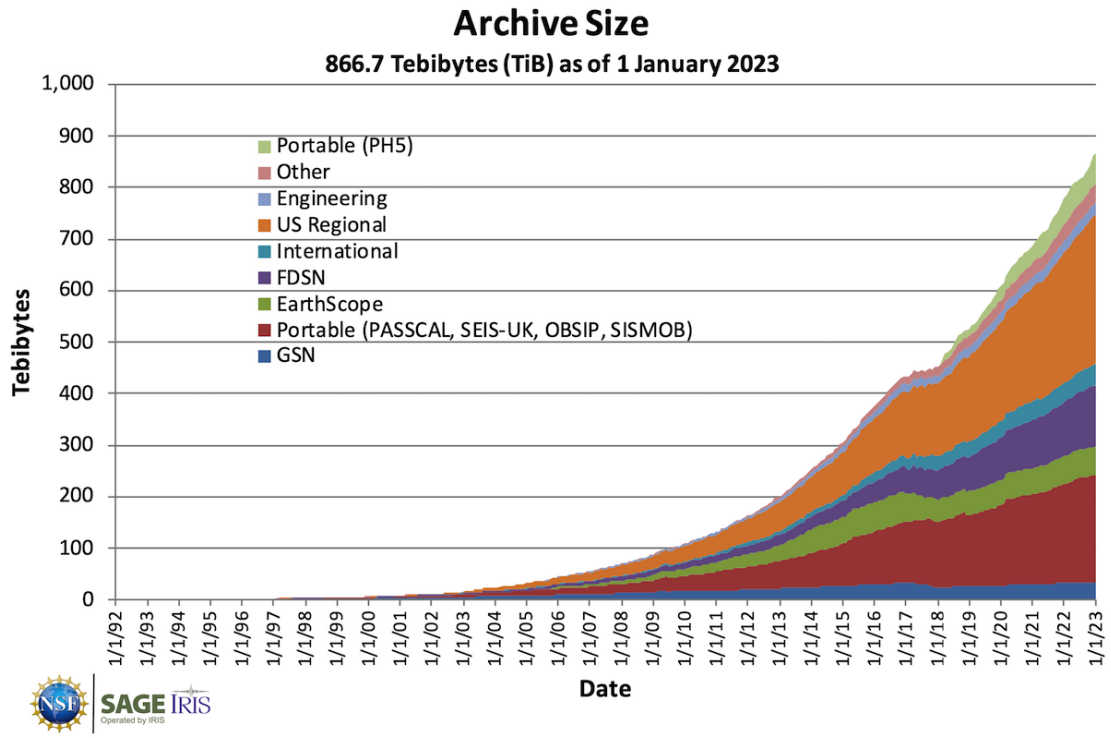

Figure 4.1: IRIS Data Management Center archive size since its inception. The image is courtesy of IRIS (https://www.iris.edu/).

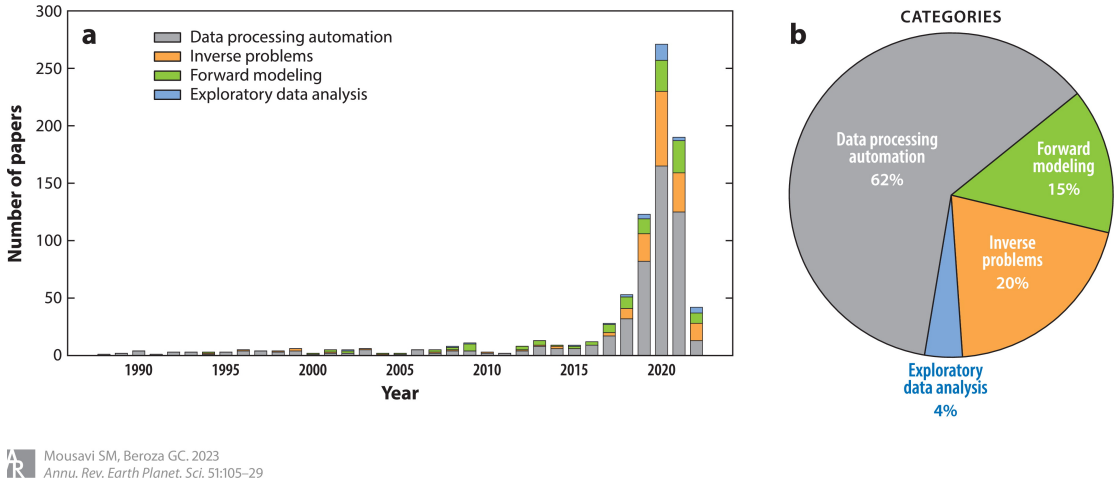

Figure 4.2: (*copied from Mousavi and Beroza [2023]*). Journal publications developing machine learning (ML) methods for a seismological task (including both earthquake and exploration seismology), published between January 1988 and May 2022.(a) Each bar shows the number of publications per year. Bars are color coded based on category of seismological tasks. (b) A pie chart shows the share of seismological tasks for ML applications.

is called "Machine Learning for Exploratory Research" and contains a smallest number of studies published to date. This implies using the ML as a research tool for finding patterns directly from seismic data and/or as an analytic framework to study and explore hypotheses or relationships.

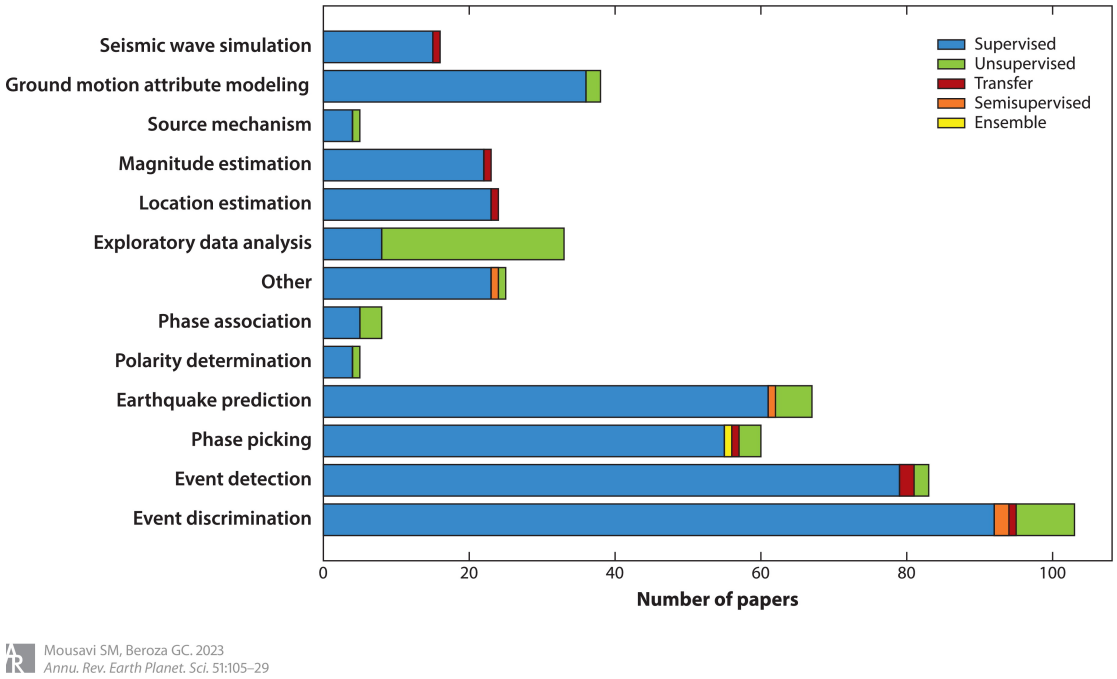

Figure 4.3: (*copied from Mousavi and Beroza [2023]*). The ratio of various machine learning approaches used for each seismological task.

Machine learning methods are aimed at understanding and developing predictive capabilities by learning relationships embedded in data. Most of the seismological ML applications developed to date are based on on "supervised learning" (Figure 4.3) when a set of training data and their corresponding "labels" or targets are used to build (train) a model to find the connection between the data and labels. Such approaches are very efficient for seismological tasks (e.g., earthquake detection and phase picking) where ample "labeled" data exist. At the same time, the reliability of labels in many available data sets is variable and uncertain and building training data sets with controlled quality remains a challenging task. In absence of reliable labeled data sets, "unsupervised learning" algorithms help to find the hidden patterns and structures in the data.

4.1.2 Specifics of the volcano seismology

This PhD thesis mainly deals with the volcano seismology, namely with the analysis of seismograms recorded in the vicinity of volcanoes in Kamchatka, Russia. This

branch of seismology has an extraordinary potential for being improved with the ML approaches both because of the high level of activity of many volcanic systems and because of the diversity of the seismo-volcanic signals. A desired goal of ML applications in volcano-seismology would be training algorithms to identify specific patterns in seismo-volcanic data that can be related to the volcanic and eruptive activity and ideally can be used to predict the latter. Despite the large amount of the available seismic data on volcanoes, this goal is however far from being achieved.

One of the main reasons for this is the difficulty to "label" the volcanic eruptions and other types of volcanic activity. There are more than 1,350 potentially active volcanoes worldwide 500 of which have erupted in historical time (`https://www.usgs.gov/faqs/how-many-active-volcanoes-are-there-earth`). Many of these volcanoes are located in remote areas and do not have a dedicated monitoring system. As a result, the global eruption database [Venzke, 2013] is far from being complete and even for catalogued eruptions important information such as exact timing is often lacked. Additionally, unfolding of volcanic eruption together with their preparation may be very complex with variable types of activity and/or eruptive styles. Training of supervised ML algorithms would require accurately cataloguing of this type of variability which is today missing even for very closely monitored volcanic systems.

Accurate cataloguing the seismo-volcanic data is another difficulty. During the periods of volcanic activity, these systems can generate tremendous amount of seismic signals, sometimes more than one thousand events per day. These signals are well known to be very variable in terms of their time and frequency signatures. Multiple classification schemes has been developed to account for this variability. Most frequently used categories are: volcano-tectonic (VT) earthquakes, long-period (LP) events (alternatively called low-frequency (LF) events), very-long-period (VLP) events, hybrid events, explosions, and rockfalls. These different classes of signals are supposedly related to different types of seismogenic sources and volcano-magmatic processes that are responsible for their excitation.

One of the most common directions for ML applications in volcano seismology is the automation of the signal analysis in order to improve the ability of volcano observatories to process the overwhelming volumes of data. Supervised ML algorithms are mostly used in this case [e.g., Malfante et al., 2018b, Titos et al., 2019]. Their training requires reasonably well label seismo-volcanic catalogs that are available only at a few volcanoes. At the same time, algorithms successfully trained in a particular volcano, cannot be simply applied on a different volcano because of the difference in the signal properties.

For the above mentioned reasons, analysis of most of seismo-volcanic datasets can-

not be within training the supervised ML algorithms. In this case, useful patterns in the data can be identified based on exploratory data analysis (Figure 4.3) that is often based on unsupervised algorithms [e.g., Cortés et al., 2021].

It is also important to mention that most published works on ML applications in volcano-seismology consider a simplified or idealized situation when the analysis is bases on data from a single station located very close to a single studied volcano. Consequently, the observed variations in the signal properties, event rates, etc are attributed to the activity of the studied volcano. However, many observations of seismic signals associated with volcanic activity are obtained in less "favorable" situations. Sometimes, available seismic stations are located relatively far from a volcano that enters in a period of activity. As a consequence, the signals related to the processes occurring within this volcano can be masked by environmental and anthropogenic noises, by tectonic seismicity (especially in subduction zones), or sometimes by signals emitted by other closely located volcanoes.

4.1.3 Context and goals of this work

Kamchatka peninsula belonging to the Pacific "ring of fire" is one of the most active regions of the World's subduction-related volcanism. 29 Kamchatka active volcanoes generate a tremendous amount of seismo-volcanic signals that are recorded by a network of seismic stations operated by Kamchatka Branch of Geophysical Survey of Russian Academy of Sciences (KB GS RAN) [Chebrov et al., 2013]. This network is used to monitor both the volcanic activity and the regional tectonic seismicity, mostly originating in the Kuril-Kamchatka and Aleutian subduction zones.

Nowadays, the processing of seismograms in Kamchatka is performed by KB GS operators, despite the exceptionally high seismic activity of the region. Thus, the problem of automatic earthquake detection and classification is acute. On average the operators process about 20 regional tectonic earthquakes and 90 volcanic ones per day. This situation is aggravated during aftershock sequences and activation of volcanoes. Moreover, most volcanic earthquakes are too weak to pick up waves onsets correctly what makes locating their hypocenters impossible. As a result, these earthquakes are catalogued without locations and sometimes are roughly classified in four classes following an approach suggested by Tokarev [1966]. This classification scheme has not been revised since late sixties and is not in accordance with modern knowledge about the origin of seismo-volcanic signals.

Overall, because of the very high seismicity rates, the catalogues of seismo-volcanic activity produced by the KB GS are for from being complete, despite their best efforts. Further improving the analysis of the seismo-volcanic data could be eventually

achieved with the implementation of appropriate ML approaches. At the same time, the existing catalogues cannot be used as labels to simply train the automatic algorithms. Therefore, the ML applications for the seismo-volcanic analysis in Kamchatka should be started as the "exploratory data analysis".

We focus our analysis on the region of the Klyuchevskoy volcano group [Shapiro et al., 2017b] where the volcanic activity has an exceptional level with several eruptions of different volcanoes occurring every year. Also strong tectonic earthquakes are regularly occurring in the Kuril-Kamchatka and Aleutian subduction zones close to the KVG. These multiple geological phenomena create a very interesting geodynamic environment with possible interactions. At the same time, they result in very complex seismic records when signals of many different origins might be superimposed. Therefore, to develop efficient ML approaches for analysing the seismicity, we need to start with attempts to reasonably "decipher" it, i.e., to correctly detect, identify, and separate the signals of different origin such as tectonic and volcanic earthquakes and those produced by different volcanoes. Only after such "source separation" the patterns of seismicity corresponding to different geological phenomena and objects can be studied.

To explore the capacity of the ML approaches to separate seismic signals of different origin, we focus on two particular strong episodes of intensive seismicity. The fist one started on December 20, 2018 with a M=7.3 earthquake occurred near the Kamchatka coast at the western termination of the Aleutian trench. Nearly simultaneously with this "Komandorskiy island" earthquake, the Shiveluch volcano, closest to its epicenter among the Kamchatka's active volcanoes, erupted. This coincidence raises a possibility of interaction between these two strong phenomena. The strong earthquake was followed by a long aftershock sequence lasting for weeks and simultaneously, many volcanic earthquakes were generated beneath Shiveluch. Additionally, the KVG region was activated in the beginning 2019 with increasing seismicity followed by two eruptions of Bezymianny volcano in January and March. When analyzing the seismicity occurred during this 2018-2019 episode we mostly focused on a problem of separation of volcanic and tectonic earthquakes. The second studied episode involved activation of Bezymianny and Klyuchevskoy volcanoes during an ongoing eruption of Shiveluch in October-November 2022. We used this data to explore the separation of earthquakes generated by different volcanoes.

4.2 Machine Learning for seismic signal analysis

4.2.1 Representing seismic signals with features

Modern seismological data are represented in digital form as discrete time series with relatively high sampling rates (100 sample per second). In this works we mainly focus on analysis of "impulsive" signals extracted from long continuously recorded signals. Even after selecting portion of the data corresponding to such signals, they contain many samples that in most of cases are not practical for analyzing and comparing different events. Therefore, we need to measure some characteristics which are referred as *features*. The latter can be divided into several groups:

- Boolean
- nominal (or categorical)
- ordinal
- quantitative

A set of several features is called a feature vector and describes one object, in our case, a seismic waveform. Thus, a group of objects can be represented with a matrix of features which is used as an input for machine learning algorithms at the next step.

It should be noted that the choice of features is of crucial importance as they should provide enough information about objects to be classified. Also, a higher number of features does not mean better accuracy of the results. On the contrary, a very detailed input set leads to *"overfitting"* in problems of supervised learning. Generally speaking, the built model works perfectly on the test set but fails when using unknown data. Also, some features may turn out to be linearly dependent and not affect the result but increase the dimension of data and computational cost consequently.

4.2.2 Supervised and unsupervised Machine Learning

There are two main classes of ML methods, which differ in their problem formulation. *Supervised* ML deals with *classification* of objects into several preset classes. In other words, they define a non-linear mapping between data and "labels". Another class of ML algorithms is called *unsupervised*, and it solves the *clustering* problem of reconstruction of the natural distribution or structure of the input data.

Clustering differs from classification in that it is not trained with "labeled" objects and should divide an input set into subsets (clusters) in the way that each cluster

consists of similar objects, and the objects of different clusters differ significantly. To do this, it is necessary to set the distance function on a set of objects. Usually, an algorithm has to define the number of clusters but in some cases, this value can be preset too.

In current work both supervised and unsupervised approaches are tested. For classification problem we consider *Support Vector Machine* (SVM) and *Random Forest* (RF), and clustering task is performed with *K-means*, *Agglomerative Clustering* and *Gaussian mixture model* (GMM). The next subsections explain each of those methods in details, however for more rigorous formulations and mathematical descriptions one should refer to further cited literature.

Supervised machine learning

At the end of the last century the method described in [Vapnik and Chervonenkis, 1974] was modified and generalized to the version currently known as SVM [Cortes and Vapnik, 1995]. This algorithm gained its popularity due to its easy implementation, fast computation and high efficiency even in big, high dimensional datasets. The main task of this algorithm is to build a hyperplane between classes of objects in order to maximize the margin width between these classes [e.g., Burges, 1998, Schölkopf and Smola, 1998]. Figure 4.1 illustrates the concept of the SVM algorithm. It can be seen that the position of the optimal hyperplane depends only on a few training objects which are called *support vectors*.

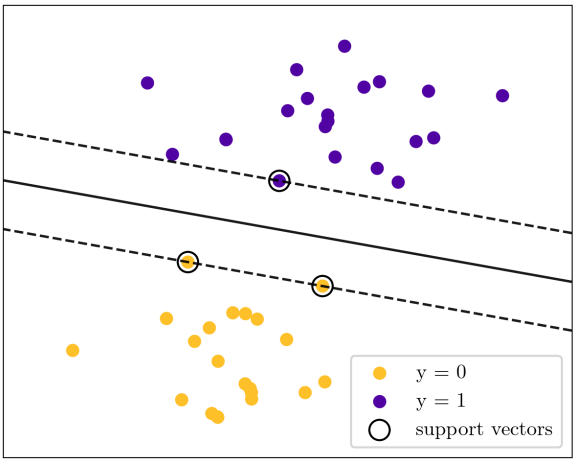

Figure 4.1: The main idea of the SVM algorithm: objects of different (two in this example) classes are separated with a hyperplane (solid line) in order to maximize the margin width (dashed lines). Support vectors are additionally circled with black lines. Modified from [Pedregosa et al., 2011]

In reality, most of the datasets are not linearly separable, but SVM deals with

these problems in another way by transforming an original feature space X to a new space H which can be linearly separable. Then objects are described not by original feature vectors x_i but vectors $\psi(x_i)$, and scalar product $\langle x, x' \rangle$ is then replaced with $\langle \psi(x), \psi(x') \rangle$. Thus, function K is called a *kernel function* if it can be represented as $K(x, x') = \langle \psi(x), \psi(x') \rangle$ for some mapping $\psi : \mathrm{X} \to \mathrm{H}$, where H is the space with scalar product. In this work the following kernels are considered:

- Linear: $K(x, x') = \langle x, x' \rangle$
- Polynomial of a degree d: $K(x, x') = (\gamma \langle x, x' \rangle + 1)^d$
- Radial basis functions (RBF): $K(x, x') = \exp(-\gamma \|x - x'\|)$

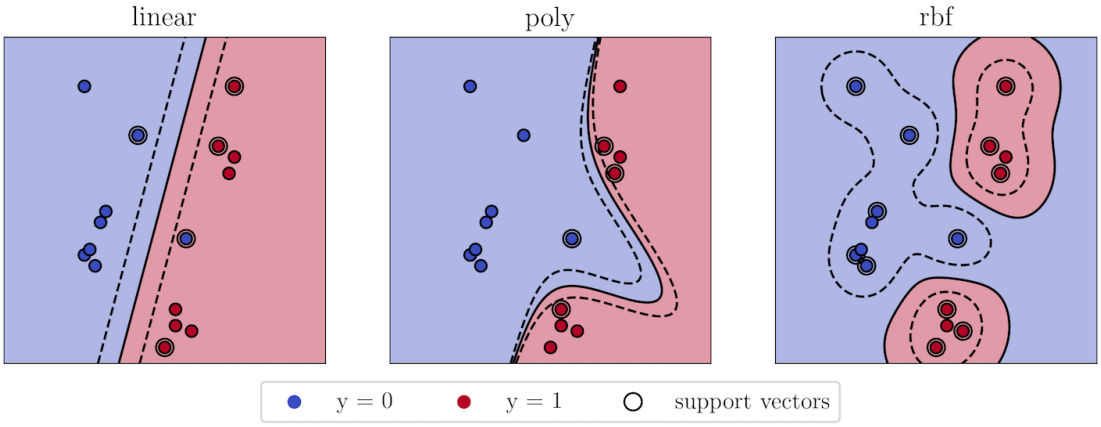

Figure 4.2: Three different types of SVM kernels, and the classification result based on the kernel choice. Points of different colors show the objects of two classes, solid line and dashed lines are the hyperplane and the margin borders correspondingly. Support vectors are additionally circled with black lines. Modified from [Pedregosa et al., 2011]

The next supervised algorithm considered in this work is *Random Forest* (RF). For the first time it was introduced in [Breiman, 2001], and it is based on the idea that the prediction based on several estimators is more reliable that the prediction of a single estimator. In case of RF these estimators are called *decision trees* which are schematically illustrated in Figure 4.3a. Decision nodes shown with diamonds represent features used at one step to make a decision based on some rule. Following a decision rule (black arrows) the algorithm moves on to the next node which can be a decision node again or a leaf node (green ellipse), i.e. the outcome. In the classification problem which is considered in this work the outcome is the assignment of an object to one of the classes. RF generates a large number of trees, and after obtaining their individual results the voting for the most popular class is carried out (Figure 4.3b).

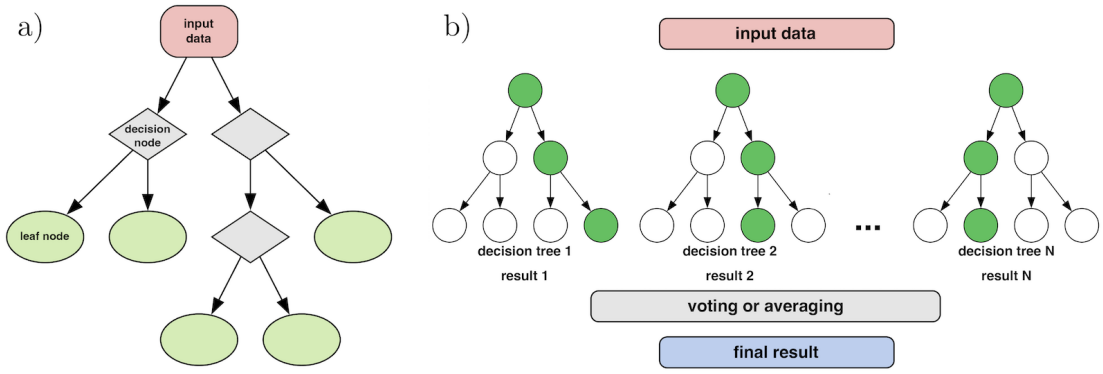

Figure 4.3: Schematic diagrams of (a) a decision tree estimator and (b) the principal of Random Forest algorithm [Breiman, 2001]

Unsupervised machine learning

In this work the investigation of clustering problem starts with the *K-means* algorithm [MacQueen, 1962, McQueen, 1967]. It has shown itself as a good first approach due to simplicity of application and computational efficiency [Har-Peled and Sadri, 2005, Arthur and Vassilvitskii, 2006]. Given a space of N objects algorithm should assign every object x_i from to one of k clusters. At the same time, every cluster j has its own centroid μ_j that is chosen by k-means in the way to minimize the sum of square distances between the objects and the cluster centroid:

$$\sum_{i=1}^{N} \min\left(\|x_i - \mu_j\|^2\right) \tag{4.1}$$

where N is the number of objects in a cluster j. After one iteration the centroids are redefined, and the objects are assigned to the new centroids. In fact, within an iteration this approach does not attribute objects to a particular class, but instead moves the centroids while objects will be ascribed automatically to them. These two steps, objects assignment and centroid calculation, are repeated until the process stabilizes. The above description emphasizes the high sensitivity of the k-means algorithm to the choice of initial centroids. One of the approaches to choose them is the random initialization [Lloyd, 1982], nevertheless this can lead to poor clustering results as the algorithm can to get stuck in a local minimum. Within scikit-learn Python package [Pedregosa et al., 2011] which is used in this work the method *k-means++* [Arthur and Vassilvitskii, 2007] is implemented to define the initial centers. The main idea of this approach is that the first set of centroids is chosen uniformly at random and each successive center is chosen with probability proportional to the square of its proximity to the nearest existing center.

Another unsupervised ML approach considered in this work is *Agglomerative Clus-*

tering which is a representative of hierarchical clustering [Lance and Williams, 1967]. Agglomerative clustering operates from the bottom to the top, i.e. at the very begging all objects of a dataset are considered as separate clusters that are being merged pair by pair while the algorithm is running. This strategy is well illustrated by *dendrograms* (Figure 4.4). These diagrams illustrate hierarchical relationships between objects of a dataset. The way the agglomerative clustering is used the dendrogram should be read from bottom to top. Thus, separate objects start to merge making clusters which merge further. The height of the vertical lines shows the degree of difference (or similarity) between branches.

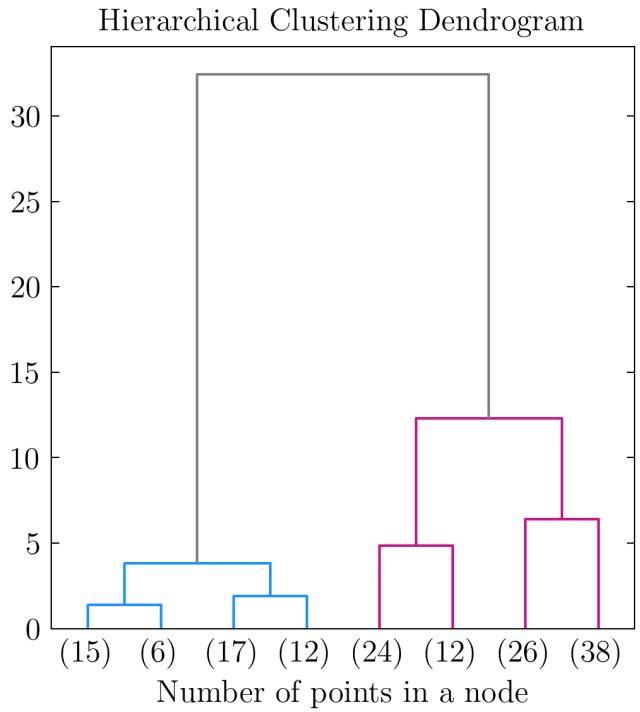

Figure 4.4: An example of a dendrogram visualising the result of hierarchical clustering. The image is modified from [Pedregosa et al., 2011]

The last chapter dedicated to clustering of volcanic seismicity uses one more clustering algorithm called *Gaussian mixture model* (GMM). In general, the problems of mixture separation are used when the "shape" of a class cannot be described by any distribution, and instead it may be described by a mixture of distributions. In other words, it is supposed that a sample from some population described by a probability density function. This density function is characterized by a parameterized model taken to be a mixture of component density functions; each component density describes one of the clusters. This model is then fit to the data by maximum likelihood [Hastie et al., 2009]. Within the GMM approach, components of a mixture of distributions have n-dimensional normal (Gaussian) distributions $\phi(x;\theta_j) = \mathcal{N}(x;\mu_j,\Sigma_j)$ with parameters $\theta_j = (\mu_j, \Sigma_j)$, where $\mu_j \in \mathbb{R}^n$ is the expectation vector, $\Sigma_j \in \mathbb{R}^{n\times n}$ is the

covariance matrix. Thus, the Gaussian mixture model has the following form:

$$f(x) = \sum_{j=1}^{k} w_j \phi(x; \theta_j) \tag{4.2}$$

with mixing proportions or weights w_j, $\sum_j w_j = 1$. An example of clustering with the GMM algorithm is presented in Figure 4.5.

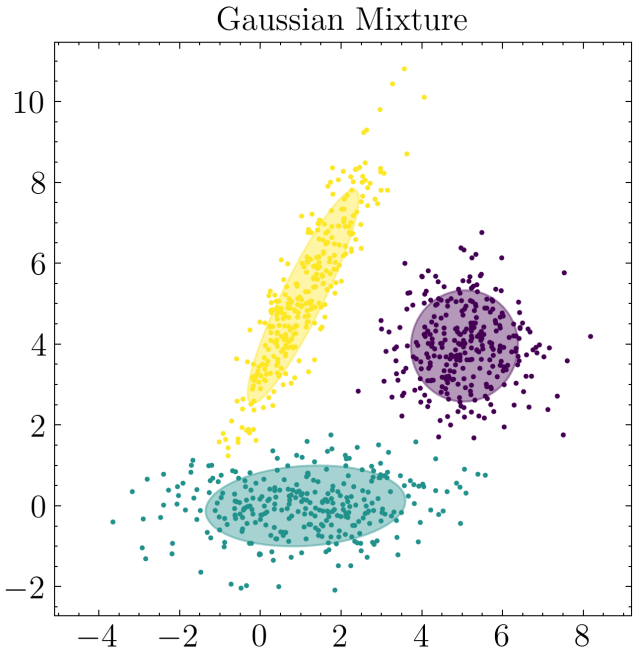

Figure 4.5: Illustration of the Gaussian mixture model clustering

4.3 Methods used in this study

This section contains the general workflow that was performed to process raw seismic data into the form suitable for machine learning application.

4.3.1 Extraction of impulsive signals from continuous data

Seismograms were band filtered in 0.5–10 Hz range, decimated and finally their absolute values were smoothed with a 30 seconds window (Figure 4.6). The maxima of the smoothed envelopes were identified as "peaks" and their signal-to-noise ratios were calculated as:

$$SNR = \frac{v_{max}}{v_{min}} \tag{4.3}$$

where $v_{min} = \max((v^1_{min},\ v^2_{min}))$ (Figure 4.6). The threshold value for SNR was set as 1.3, thus the detections with $SNR \geq 1.3$ were included into the catalog.

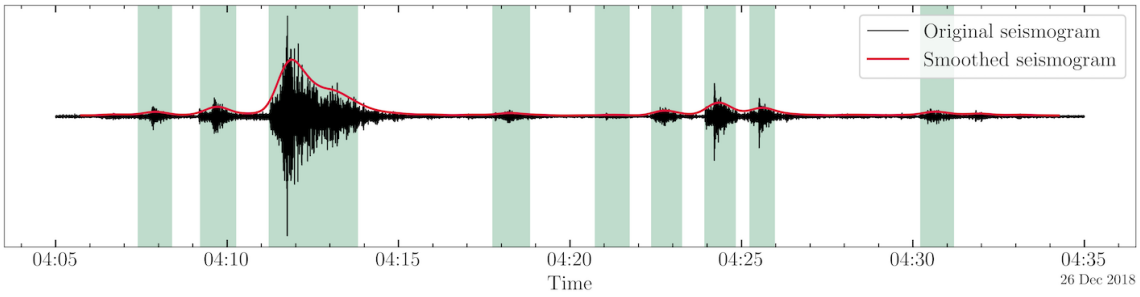

Figure 4.6: One hour of seismic data and its smoothed envelope (red line). Detected events with $SNR \geq 1.3$ are highlighted with green.

4.3.2 Features extraction

The correct choice of features is one of the key elements for the successful classification. In many pioneering investigations on classification of seismicity the waveforms themselves were used as features [e.g., Falsaperla et al., 1996, Scarpetta et al., 2005, Esposito et al., 2008]. This approach seems natural and obvious, however, use of these initial representation has shown no spectacular results, and it has been mentioned that other parameters of signals should be explored. Thus, much attention was paid to spectral representations of seismic signals. Classification studies in volcano seismology [e.g., Curilem et al., 2009, Malfante et al., 2018a] are dedicated to distinguishing various volcanic signals according to the common classification (see Section 1.2) which is mostly based on frequency content of waveforms. Similar attempts but with tectonic earthquakes also considered spectral representations of waveforms [e.g., Maurer et al., 1992, Sick et al., 2015]. Besides this, there are more parameters of seismic signals that were used in the studies, such as polarization [Köhler et al., 2009, Hammer et al., 2012], autocorrelation function [Falsaperla et al., 1996, Langer et al., 2003], etc. Quite exhaustive lists of features that can be extracted from signals and used in investigations can be found, for instance, in [Hammer et al., 2012, Malfante et al., 2018a].

Alongside with the experience of the other researchers, inspection of data used in this work shows significant difference between tectonic and volcanic earthquakes in their spectral domain (Figure 4.7).

As there is no final solution on what features give better classification result, this work explores both temporal and spectral waveform attributes which are listed below, while the process of their extraction is illustrated in Figure 4.8 as well.

1. Peak frequency $f_{peak} = \text{argmax}_f P(f)$

2. Mean frequency is defined as $f_{mean} = \frac{\sum_i P(f_i) f_i}{\sum_i P(f_i)}$

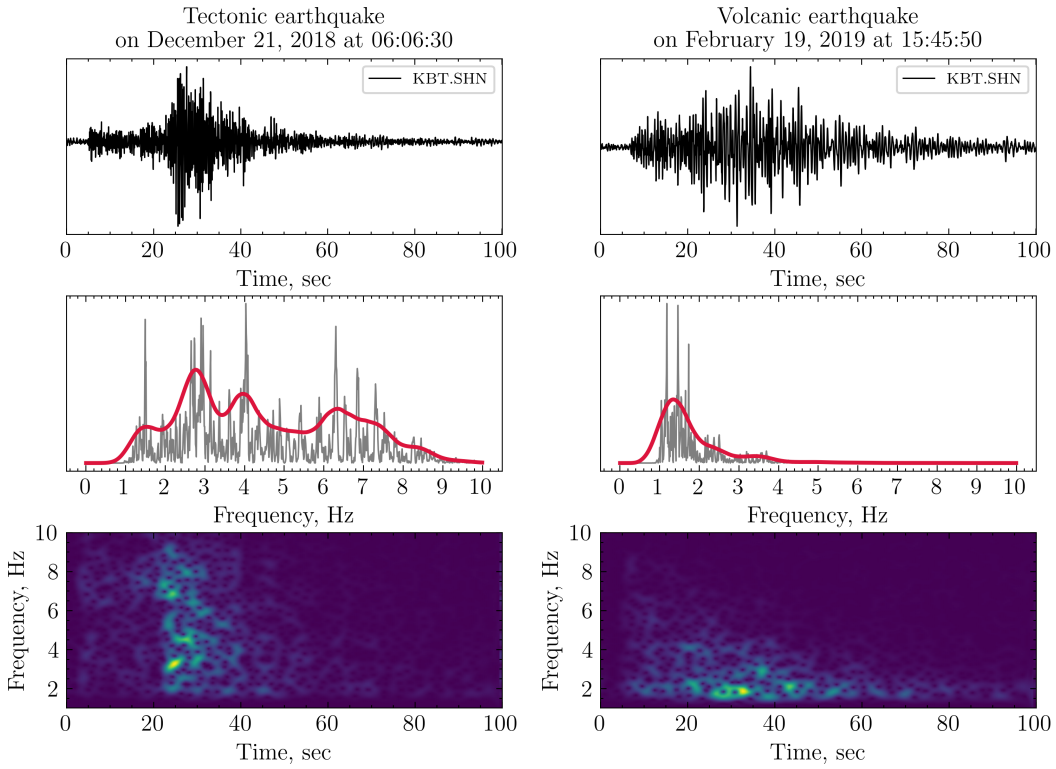

Figure 4.7: Waveform examples (top panels) and their spectral representations (unsmoothed and smoothed spectra (grey and red lines in the middle panels correspondingly), spectrograms (bottom panels)) of a tectonic and volcanic earthquakes.

3. Median frequency f_{median} is calculated using the following condition:

$$\sum_{j=1}^{j_{MDF}} P(f_i) = \sum_{j=j_{MDF}}^{M} P(f_i) = \frac{1}{2}\sum_{j=1}^{M} P(f_i)$$

   where j_{MDF} is the index of the median frequency

4. Bandwidth was taken as the difference between the first and last frequencies, where

$$P(f) \geq P_{max}$$

5. Standard deviation σ from the mean frequency f_{mean}

6. Signal duration

7. Signal maximum absolute amplitude

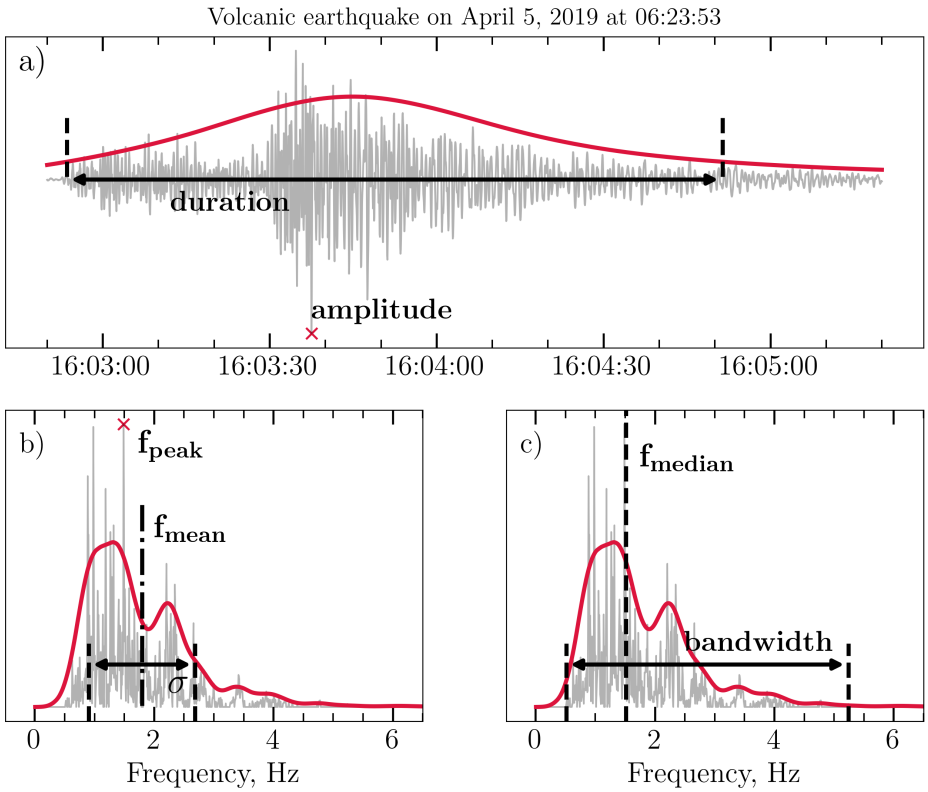

Figure 4.8: Extraction of "regular" features: a) Duration of a signal and its amplitude (red cross); red line shows the smoothed waveform b-c) Spectral characteristics of a signal such as a peak frequency f_{peak} (red cross); a mean frequency f_{mean} and a standard deviation σ from it; median frequency f_{median} and a spectrum bandwidth (an original spectral form and a smoothed one are shown with grey and red colours respectively)

Chapter 5

Application: Seismic crisis in 2018-2019

5.1 Motivation

A vivid example of seismic crisis started in November 2018 when a strong (M=7.3) earthquake occurred and was followed by an aftershock sequence. This tectonic activity was accompanied by the Shiveluch volcano eruption with its seismic manifestation. The major tectonic events and the active volcano are presented in Figure 5.1 as well as several seismic stations in this region.

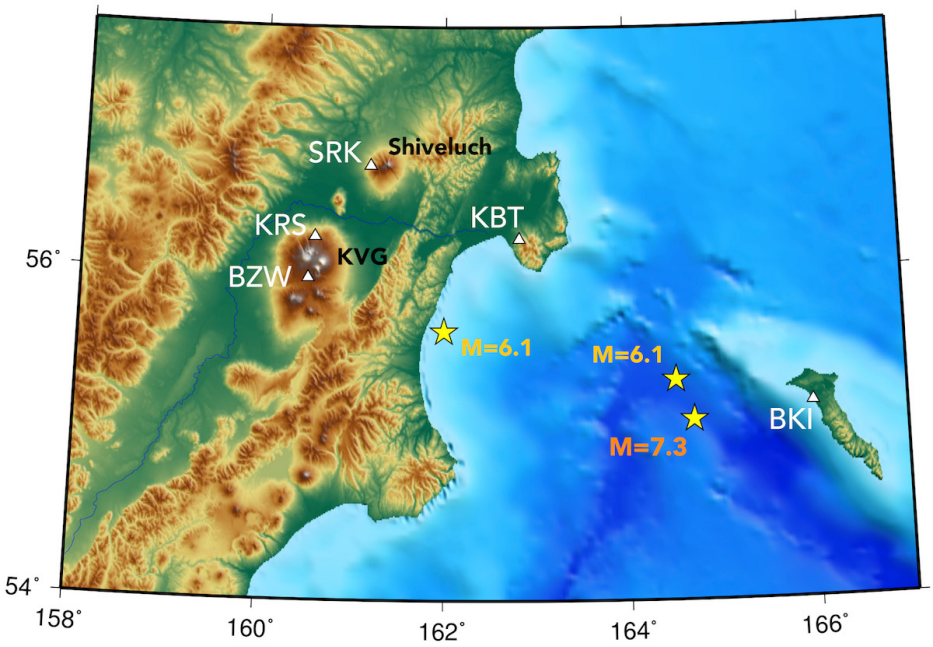

Figure 5.1: Map of the Kamchatka coast with the most active volcanoes (Shiveluch and Klyuchevskoy volcano group (KVG)) are shown. Yellow stars represent the mainshock and its strongest aftershocks, white triangles are the stations used in the study.

It can be seen that station KBT is almost equidistant from both Shiveluch and the area of tectonic activity. Signals from two events are presented in the middle panel at Figure 5.2, and without additional information it is difficult to define their origin. Yet using records from station SRK (located on the slope of the Shiveluch volcano) and station BKI (far from the coast) it is clear that the first waveform corresponds to

volcanic activity, and the second one is a regional tectonic earthquake.

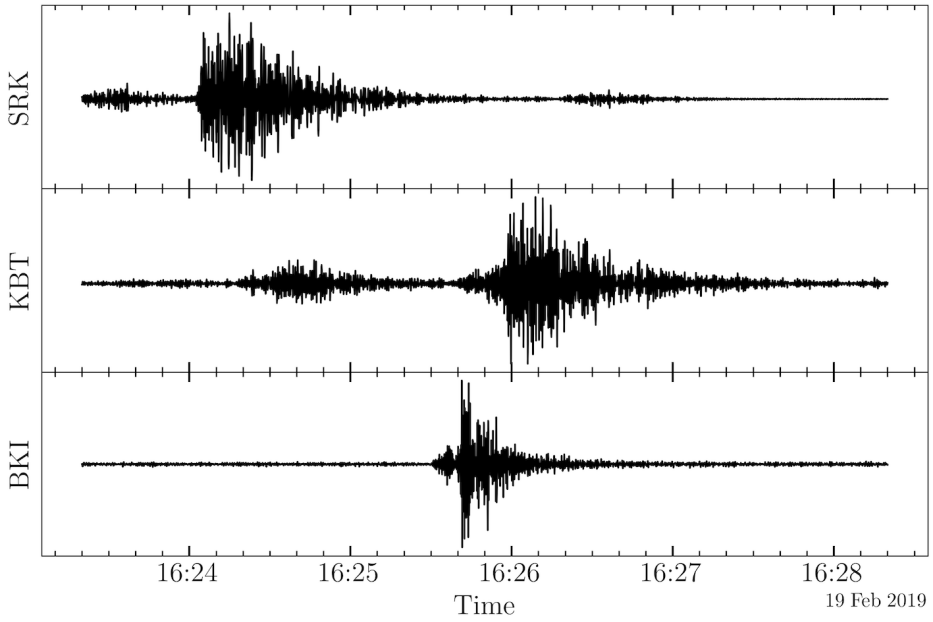

Figure 5.2: Seismograms from different stations (SHN channel): stations SRK and BKI recorded a volcanic earthquake and a tectonic one correspondingly, while in the records at KBT both events can be seen clearly.

Figure 5.3 shows the daily number of earthquakes detected by stations BKI and KBT. The highest peak of activity corresponds to the mainshock on 20 December 2018. The following aftershock swarm can be approximated with Omori law that describes the frequency of earthquakes with time:

$$n(t) = \frac{k}{(c+t)^p} \tag{5.1}$$

where k and c are empirical constants, which vary for different earthquake sequences, and p is a third parameter which modifies the decay, its typical range is 0.7–1.5. In case of station BKI, Omori law is preserved with time while at station KBT the increase in seismic activity can be clearly seen (red dashed line in Figure 5.3).

Thus, this section is dedicated to dividing seismicity in Kamchatka region into two general classes, tectonic and volcanic, in order to reveal the reason of the deviation from the Omori law in the seismic activity registered by station KBT (Figure 5.3) using data only from one channel of a single station.

5.2 Compilation of a labeled dataset

As it was mentioned earlier, unsupervised algorithms do not demand *a priori* knowledge to perform clustering as they reconstruct the structure of input data. On

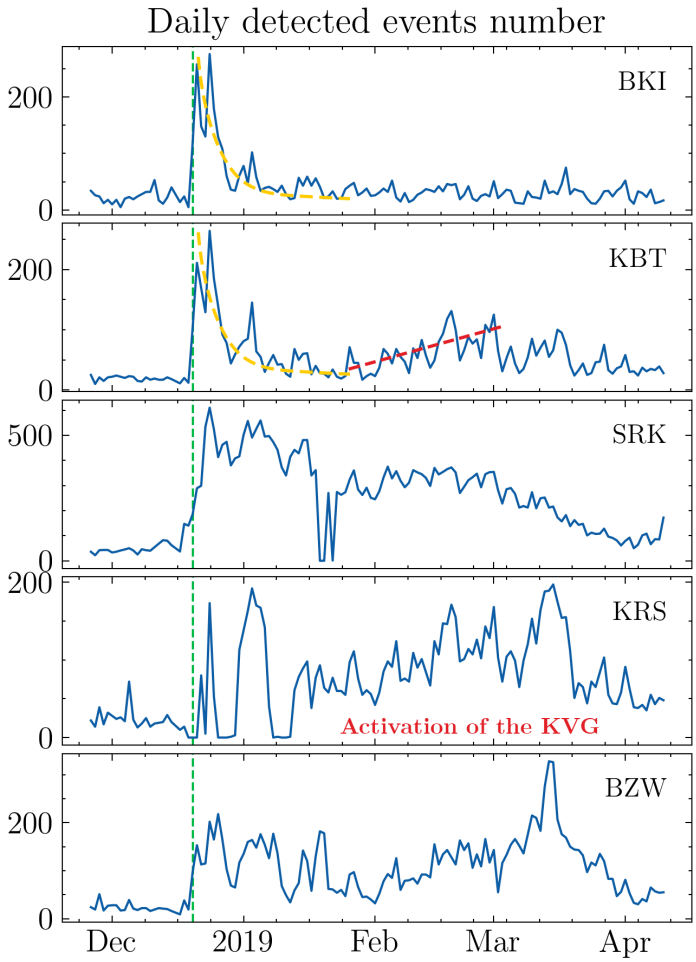

Figure 5.3: The daily number of earthquakes (blue lines) detected at stations located in different settings (Figure 5.1). The mainshock date is shown by a green vertical line; a yellow curve is the approximation of the aftershoks decay with the Omori law (equation 5.1), a red dashed line shows the deviation from it

the other hand, supervised (and semi-supervised) machine learning methods solve the problem of classification or regression by building a model from the data that has already been labeled. In common practice, a dataset with labels is called a *training* dataset. Thus, this subsection presents the process of creating the training set in order to classify the seismicity.

Compilation of a training set was done in a semi-automatic manner. At the first step, potentially "pure" tectonic or volcanic events were selected using reference stations SRK and BKI. Obviously, station SRK (Figure 5.1) located on the slope of the Shiveluch reliably recorded volcanic earthquakes. Station BKI in its turn is less affected by volcanic activity but can detect weaker aftershocks that started on November 20, 2018, as well as regional tectonic earthquakes. At the first step, seismic data from these stations was processed in the same manner as data recorded with KBT station:

the seismograms were smoothed, and a detection is included into a catalog if its SNR is higher than a threshold of 1.3. After that, pairs of catalogs, SRK vs. KBT and BKI vs. KBT, were compared. Considering the epicenter distances to these stations and corresponding waves arrivals, the next conditions were set for selecting a "pure" volcanic event:

$$\Delta t = t_{KBT} - t_{SRK} \in [15, 30]\, s \tag{5.2}$$

$$f^{SRK}_{median} \geq f^{KBT}_{median} \tag{5.3}$$

The first condition is derived from the fact that seismic waves emitted by a volcanic earthquake arrive at station SRK earlier than at station KBT. And the second condition reflects an idea of attenuation of high frequency energy during wave propagating. Thus, median frequency of signals recorded by SRK station is assumed to be higher than median frequency observed at KBT. Regarding tectonic events, only one condition was used:

$$\Delta t = |t_{KBT} - t_{BKI}| \leq 60\, s \tag{5.4}$$

Given the distance of 230 km between stations BKI and KBT, 60 seconds is the approximate maximum time difference between waves arrivals at the stations. Manual revision followed the automatic compilation of a preliminary training set. It allowed to eliminate very noisy waveforms, falsely detected signals of non-seismic origin (e.g., calibration signals). In addition to these unwanted signals, unexpectedly, the strongest tectonic earthquakes had to be removed although they were assumed to be the best examples of the tectonic seismicity in the training set. In Figure 5.4 it can been seen that these strong events are not recorded properly but with saturation effect which resulted in uncharacteristic spectral form. Consequently, the final labeled set contained tectonic events of moderate magnitudes.

5.3 Results: Supervised machine learning

5.3.1 Support Vector Machine

The first algorithm considered in this section is the SVM. Its principle and advantages were described in Subsection 4.2.2. This part of work is performed with scikit-learn Python package [Pedregosa et al., 2011], and within this SVM realization the performance of classification strongly depends on two parameters. First of them is $C > 0$ which defines the smoothness of a separating hyperplane. Another parameter γ is responsible for the influence of each object from a training dataset on other nearby objects. Grid search over possible values of these parameters can help to obtain the

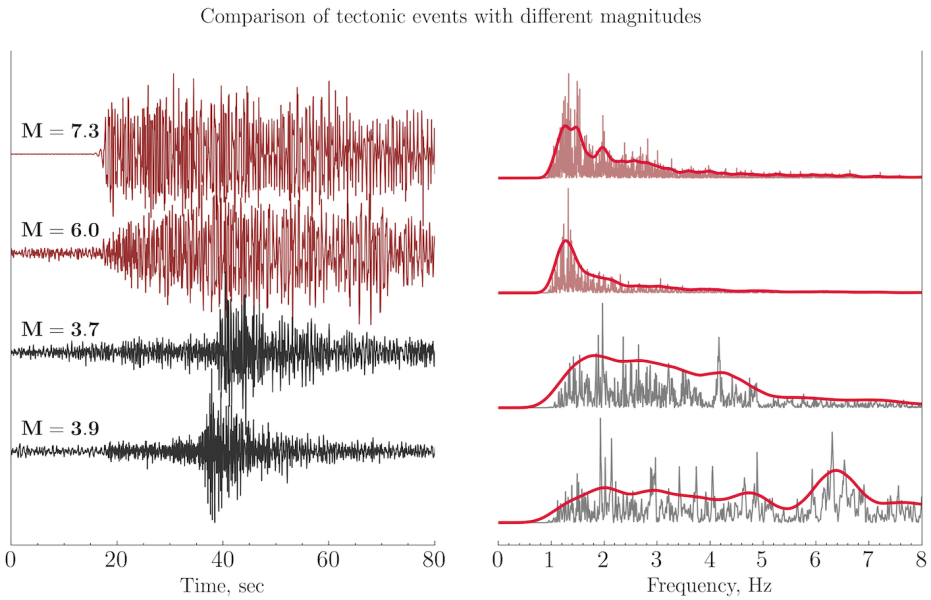

Figure 5.4: Recorded tectonic events with different magnitudes and their spectra. Very strong events with waveforms similar to the first two (red color) were eliminated from the training set. Moderate earthquakes (two bottom waveforms in black) remained in the set.

best combination. At the same time, choice of an optimal kernel for a given problem is an additional challenge. The detailed investigation for obtaining the optimal combination of parameters C and γ with a kernel type is presented in Appendix 5.A. Table 5.1 summarizes the result of this inspection, and the obtained parameters will be used further.

	Kernel type		
	Linear	Polynomial	RBF
C	10	0.01	10
γ	0.001	1	0.1

Table 5.1: Best pairs of parameters C and γ for different kernels revealed with the grid search (Appendix 5.A).

In supervised ML practice a labeled dataset is usually split into two parts. The first one, the *train set* is used to build a classifier, and commonly this dataset contains 2/3 of the available labeled data. The performance of the obtained classifier can be validated with a *test set* which is unknown data for a classifier. If the validation result is satisfactory the entire dataset can be introduced for its classification. Figures 5.5 – 5.7 present the results of the SVM classification with different kernels.

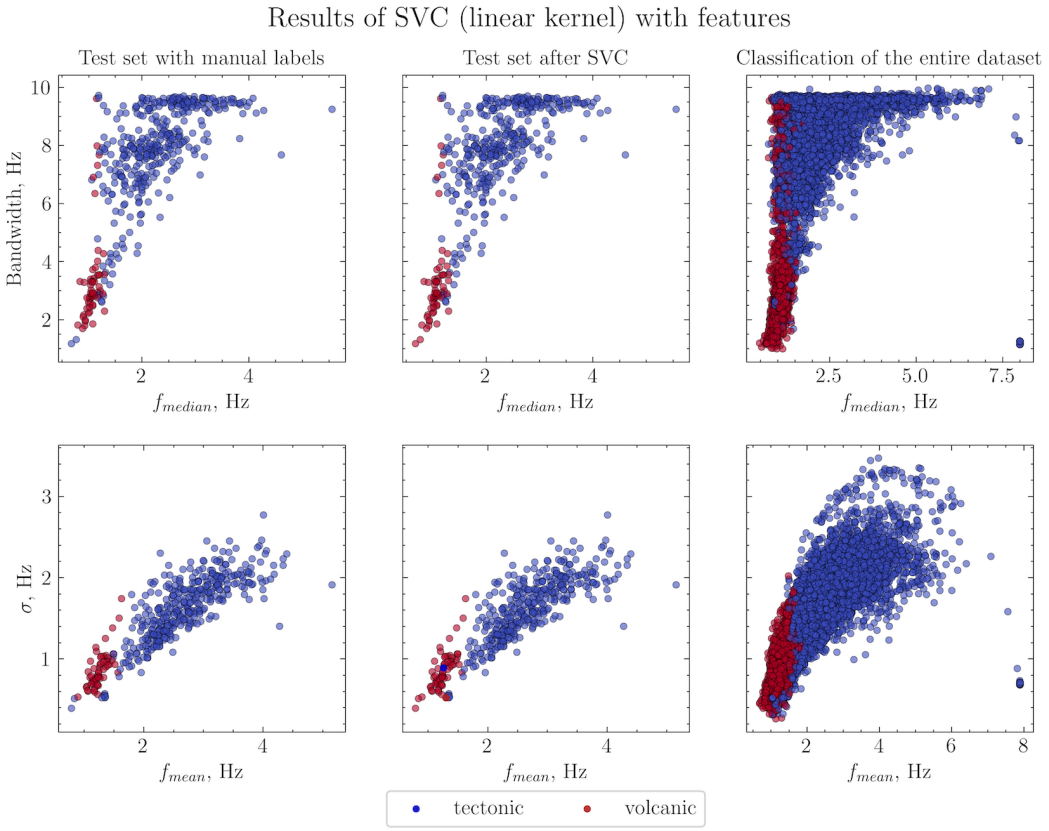

Figure 5.5: The test dataset with labels defined manually (first column) and the results of the SVM classification with a linear kernel on the test set and the entire data (second and third columns). Blue dots and red dots show the tectonic and volcanic events correspondingly.

The quantitative evaluation of the obtained results can be done using statistical measures. The first one is called *precision*:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} = \frac{True\ Positive}{Total\ Predicted\ Positive} \tag{5.5}$$

In our case, *True Positive* means a correct assignment of an event to a particular class, tectonic or volcanic. Thus, these numbers are diagonal values in a confusion matrix. Then *Total Predicted Positive* is a number of events within one class after classification, i.e. a sum of events in a row of a confusion matrix.

Actually, the precision alone is not enough to adequately estimate the quality of classification results. For example, a classifier correctly assigned labels to all objects of one class but at the same time falsely added plenty of objects from another one. Although the precision will be maximum it is clear that the general performance is not satisfying. In results evaluation it is useful to use an additional metric, for example,

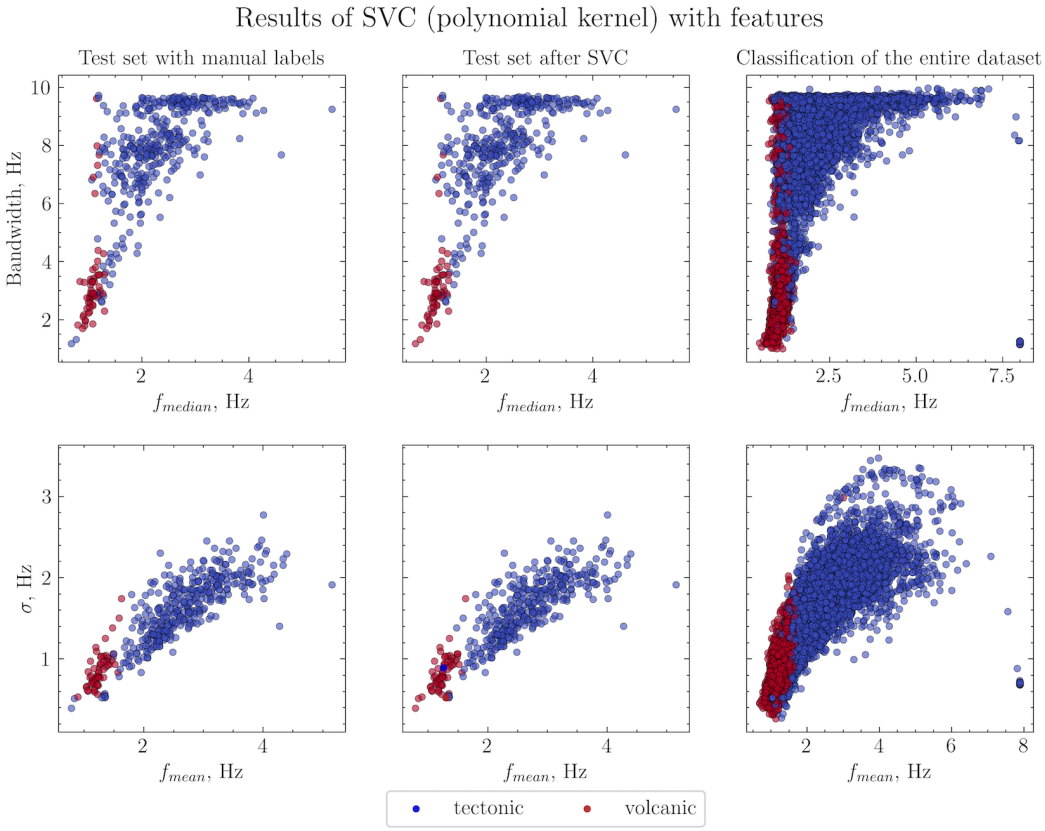

Figure 5.6: The test dataset with labels defined manually (first column) and the results of the SVM classification with a polynomial (3^{rd} degree) kernel on the test set and the entire data (second and third columns). Blue dots and red dots show the tectonic and volcanic events correspondingly.

recall:

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} = \frac{True\ Positive}{Total\ Actuall\ Positive} \tag{5.6}$$

Again, given that *True Positive* is a number of correct predictions for a particular class then *False Negative* is the amount of objects of the same group erroneously assigned to another one. Consequently, *Total Actual Positive* is a result of the summation over a column.

Although all classifiers show high performance on the test set, the preference will be given to an RBF kernel. The training and test datasets have a clear structure that can be easily divided by a classifier with a linear or a polynomial kernel. Unlike them, the whole dataset is more complex with no clear clusters, thus SVM-RBF algorithm is more reliable in defining complicated borders.

Finally, using the SVM-RBF classifier with the parameters C = 10 and $\gamma = 0.1$, defined with the grid search, the entire dataset was separated into two classes. Tem-

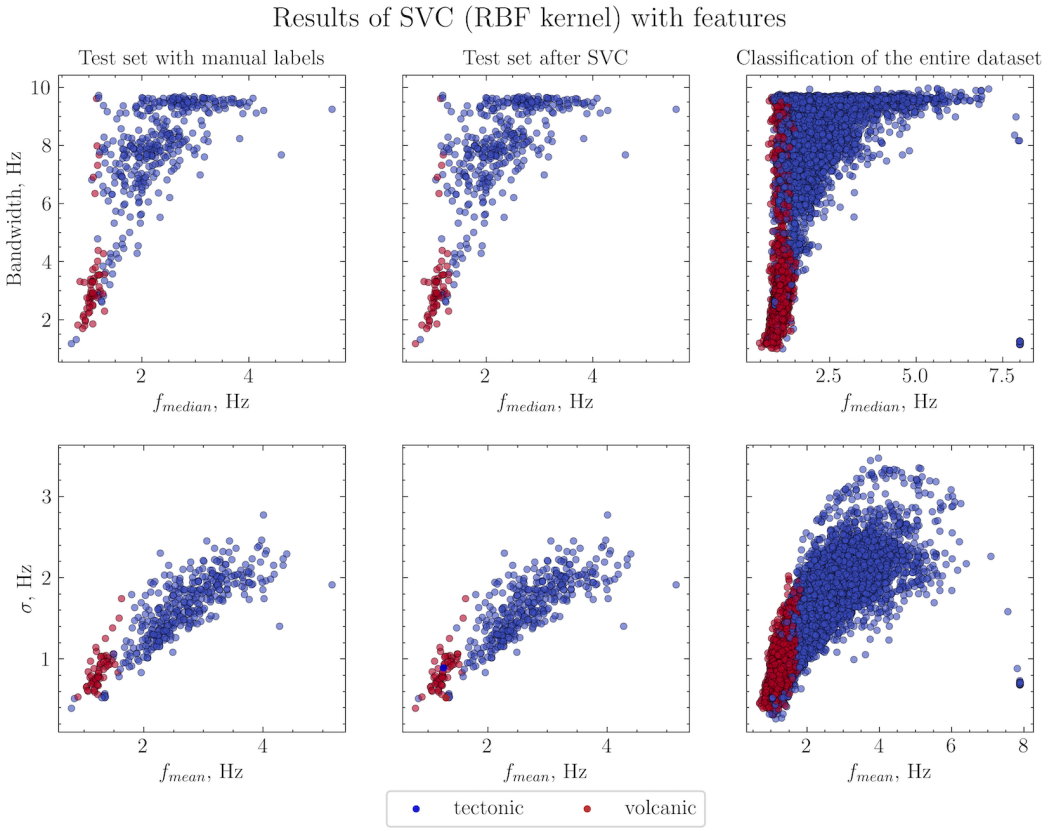

Figure 5.7: The test dataset with labels defined manually (first column) and the results of SVM classification with an RBF kernel on the test set and the entire data (second and third columns). Blue dots and red dots show the tectonic and volcanic events correspondingly.

poral changes of tectonic and volcanic seismicity are presented in Figure 5.8. The obtained results show the main peculiarities of seismic activity regime in 2018-2019. Two highest peaks correspond to the mainshock with its strongest aftershocks followed by activation of the Shiveluch volcano. However, some correlation in two plots can be noticed what implies that there are errors in classification, however this supposition can be revealed only after manual check.

5.3.2 Random Forest

This subsection presents the classification results obtained with Random Forest (RF) algorithm [Breiman, 2001]. RF classifiers are prone to overfitting which can be avoided by adjusting various hyperparameters necessary to build a model. The detailed description of possible parameters and their estimation is contained in Appendix 5.B. After obtaining the optimal hyperparameters values, RF classification of data was

Linear kernel

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**364**	1	1.00
	Volcanic	3	**57**	0.95
	Recall	**0.99**	**0.98**	

Polynomial kernel

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**361**	1	1.00
	Volcanic	6	**57**	0.90
	Recall	**0.98**	**0.98**	

RBF kernel

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**364**	1	1.00
	Volcanic	3	**57**	0.95
	Recall	**0.99**	**0.98**	

Table 5.2: Confusion matrices of the SVM classification with different kernels obtained on the test set.

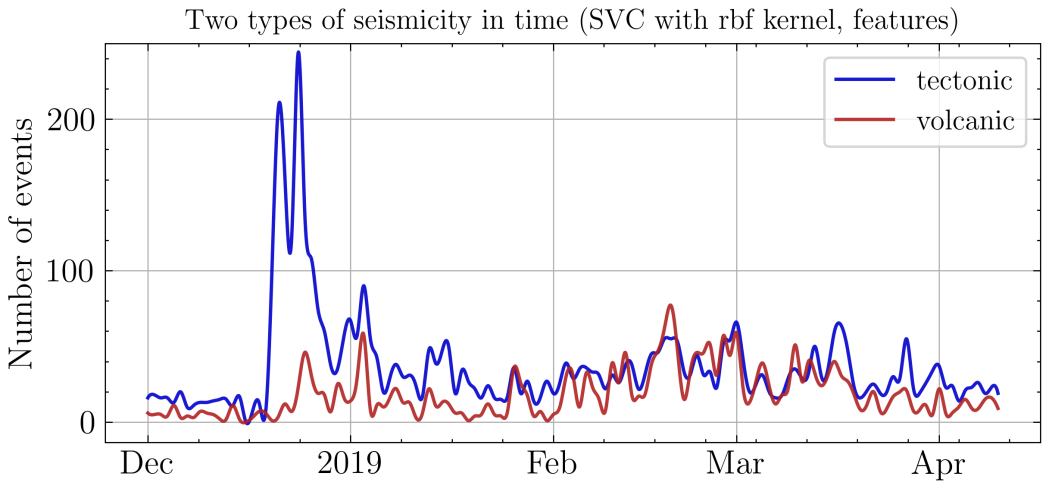

Figure 5.8: Temporal changes of tectonic (blue line) and volcanic (red line) activity based on the SVM classification

performed, and the results are shown in Figure 5.9. The confusion matrix is presented in Table 5.3 as well.

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**363**	2	0.99
	Volcanic	4	**56**	0.93
	Recall	**0.99**	**0.97**	

Table 5.3: Confusion matrix for RF classifier.

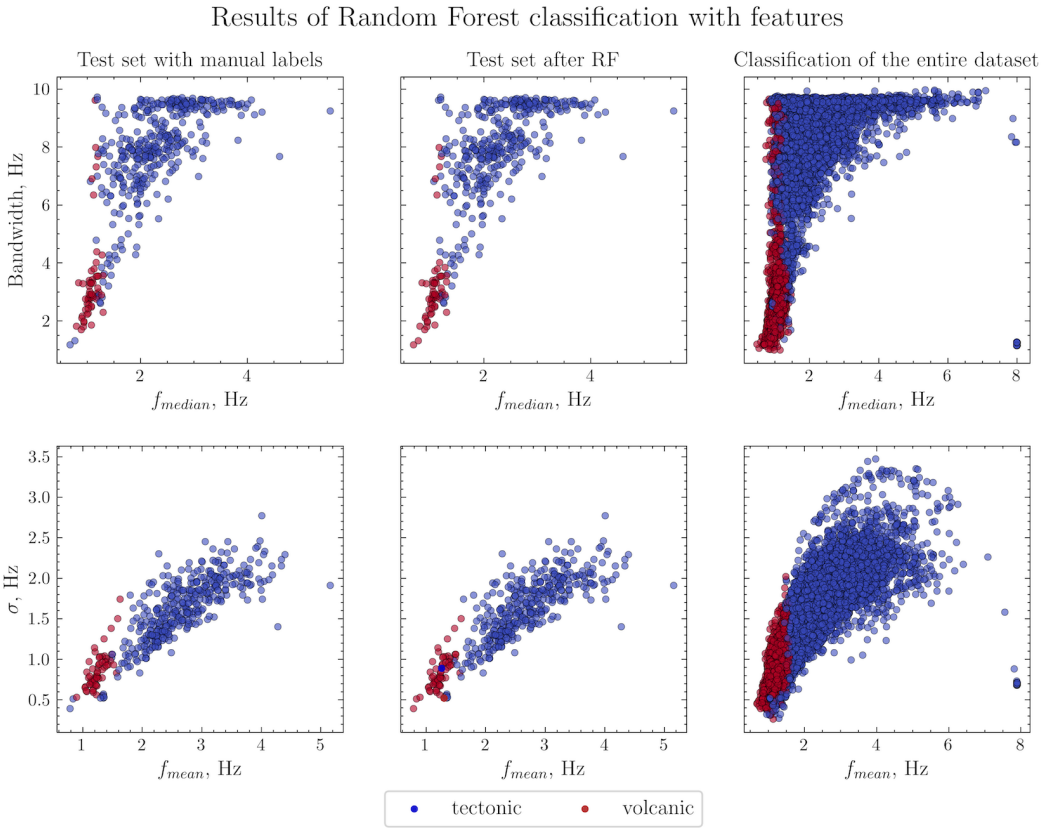

Figure 5.9: The test dataset with labels defined manually (first column) and the results of RF classification on the test set and the entire data (second and third columns). Blue dots and red dots show the tectonic and volcanic events correspondingly.

Based on the results of RF classification temporal changes in seismic activity were plotted (Figure 5.10). As well as in the case of SVM classifier one can trace the main features of seismic regime. Nevertheless, the target part of the plot (after February 2019) does not allow to confidently estimate the origin of the deviation from the Omori law. Again, correlation of two plots can likely reflects mistakes in classification that

can be verified solely manually.

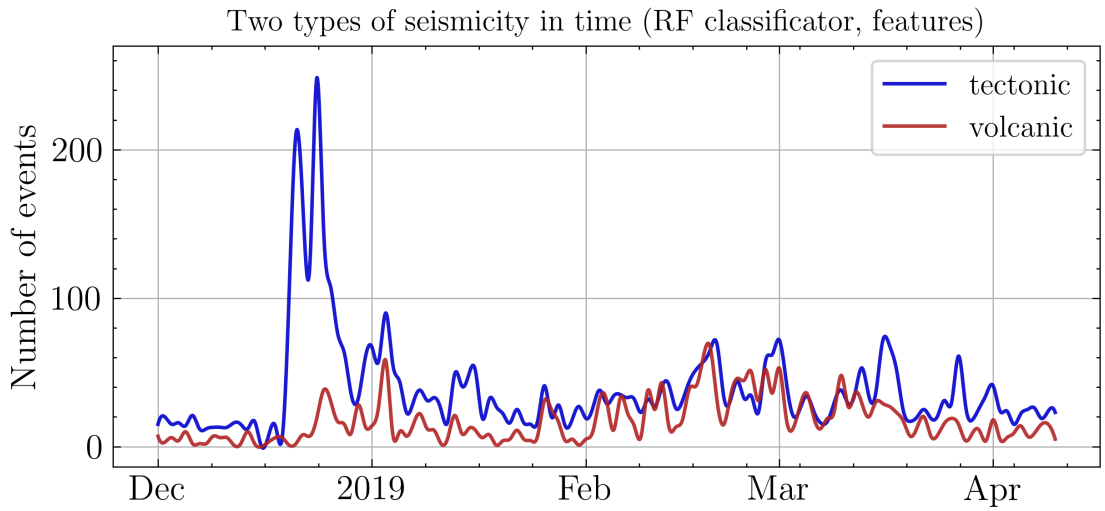

Figure 5.10: Temporal changes of tectonic (blue line) and volcanic (red line) activity based on RF classification.

5.4 Results: Unsupervised machine learning

5.4.1 K-means

For this work a couple of the most popular clustering techniques were chosen as well. Figure 5.11 presents the results of application of k-means to study data. Although here we consider unsupervised method that does not assume *a priori* knowledge about data it is useful to test the algorithm first on the labeled dataset created for supervised algorithms. This procedure will allow us to understand if the chosen methods are generally applicable to our data and how adequate are the results.

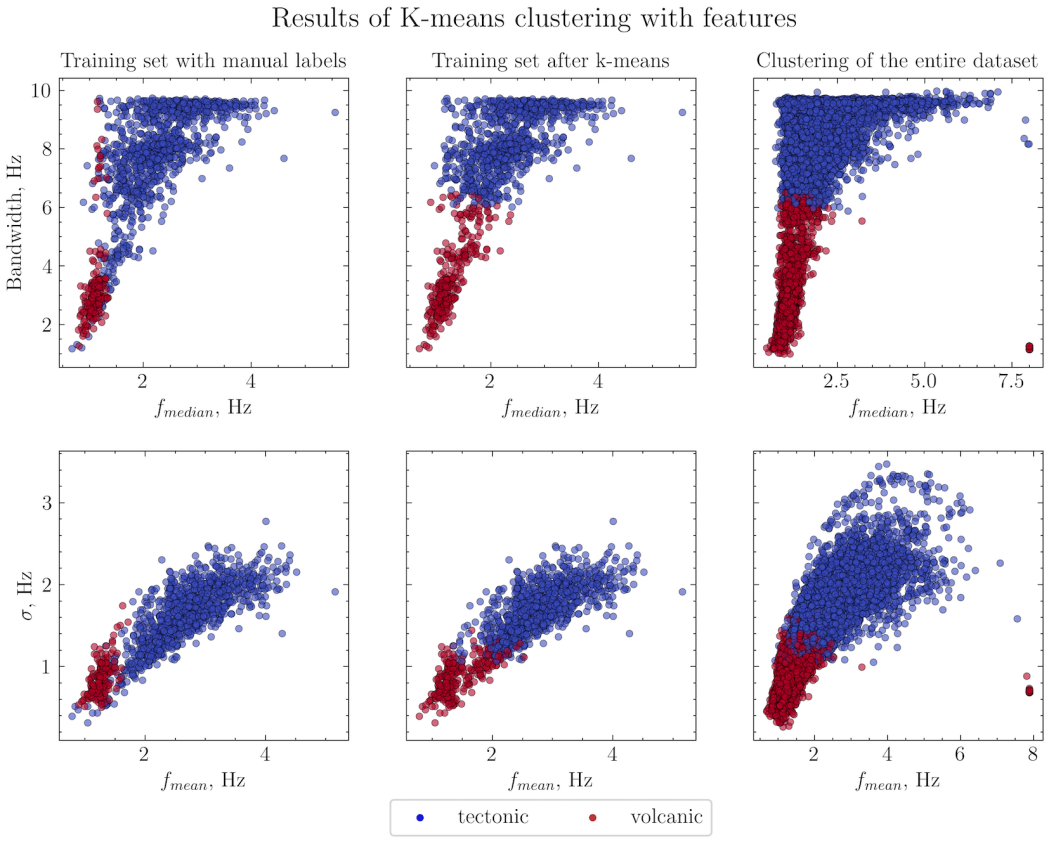

Figure 5.11: The labeled dataset with labels defined manually (first column) and the results of k-means clustering on the labeled set and the entire data (second and third columns). Blue dots and red dots show the tectonic and volcanic events correspondingly.

Comparison of first two columns in Figure 5.11 shows that the k-means algorithm satisfactorily divides the labeled dataset but a numerous errors can be already seen. Obviously that after performing clustering on the entire dataset the final result is not

K-means				
		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**979**	13	0.99
	Volcanic	134	**170**	0.56
Recall		**0.88**	**0.93**	

Table 5.4: Confusion matrix for k-means clustering.

able to answer the main question regarding the deviation from the Omori law. (Figure 5.12)

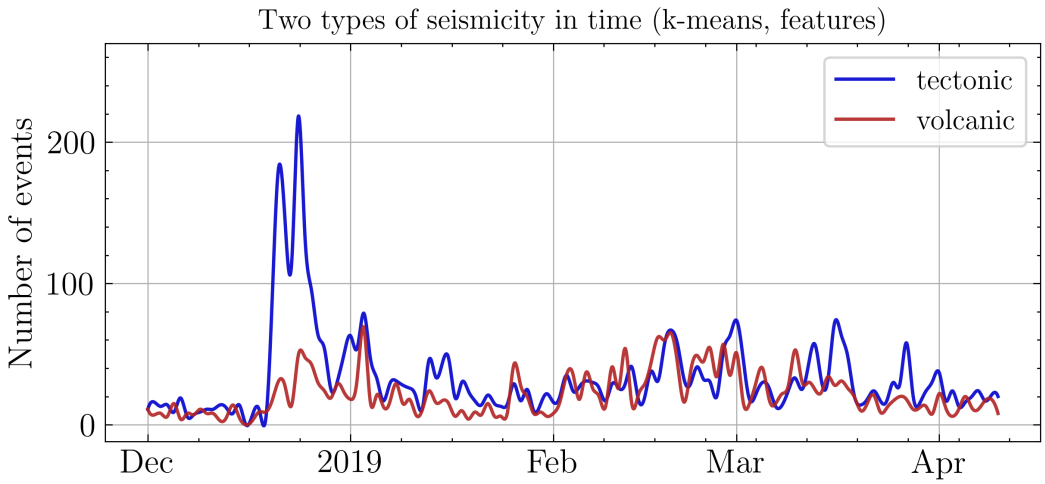

Figure 5.12: Temporal changes of tectonic (blue line) and volcanic (red line) activity based on k-means clustering.

5.4.2 Agglomerative Clustering

Figure 5.13 shows the results of agglomerative clustering of the datasets. The performance on the label data is better than one of k-means (Table 5.5, second column in Figure 5.13), although the border area between two clusters is challenging for the algorithm as well.

Thus, Agglomerative Clustering also cannot explain the deviation from the Omori law in the tectonic activity plot (Figure 5.14).

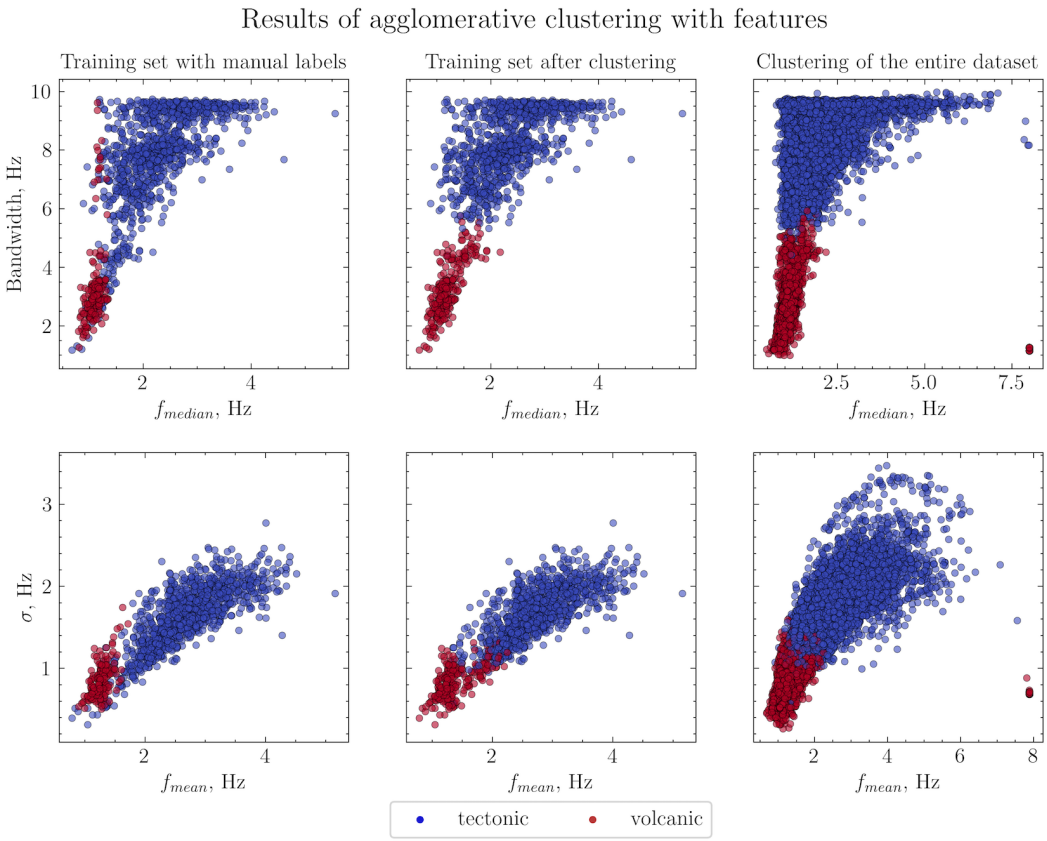

Figure 5.13: The labeled dataset with labels defined manually (first column) and the results of Agglomerative Clustering on the labeled set and the entire data (second and third columns). Blue dots and red dots show the tectonic and volcanic events correspondingly.

Agglomerative Clustering				
		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**1022**	15	0.99
	Volcanic	91	**168**	0.65
Recall		**0.92**	**0.92**	

Table 5.5: Confusion matrix for Agglomerative Clustering.

5.5 Results: Classification and clustering using spectral representations of data

Besides "regular" features, i.e. some parameters of the signals that are used by specialists in order to classify signals, several spectral representations of seismic signals

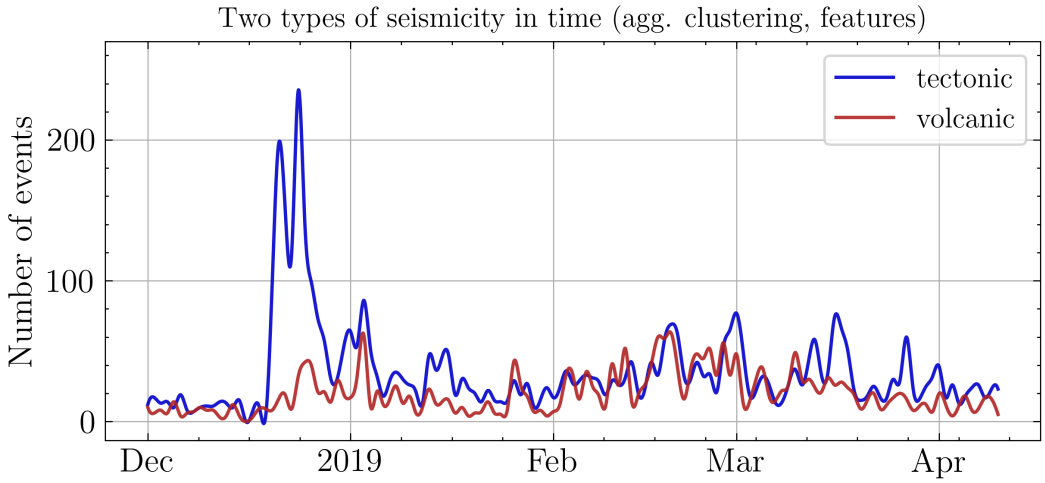

Figure 5.14: Temporal changes of tectonic (blue line) and volcanic (red line) activity based on Agglomerative Clustering.

were used as feature vectors. The reasonableness of this approach was commonly proposed in Section 4.3.2. Moreover, spectral representations are mentioned as the best ones for the classification of signals [e.g., Maurer et al., 1992, Falsaperla et al., 1996, Köhler et al., 2009, Soto et al., 2018].

For every event in the entire dataset the following information was extracted with number N of points, i.e. features in a feature vector:

- Smoothed spectra (N = 145)
- Unsmoothed spectra (N = 500)
- Spectrograms (N = 3600)

Besides spectral envelopes the time-frequency representation of seismic signals were considered as the general frequency characteristics of a signal are preserved but their temporal changes are also taken into account.

After extracting new features the same ML algorithms of classification (SVM, Random Forest) and clustering (K-means, Agglomerative Clustering) were tested using spectral representations of data as feature vectors. The results of classification and clustering in the form of confusion matrices are presented in Tables 5.C.1 – 5.C.4 contained in Appendix 5.C, and according to them the best performance was achieved by SVM-RBF classifier in a combination with smoothed spectra as features (Table 5.C.1). Figures 5.15 and **??** illustrate the best result.

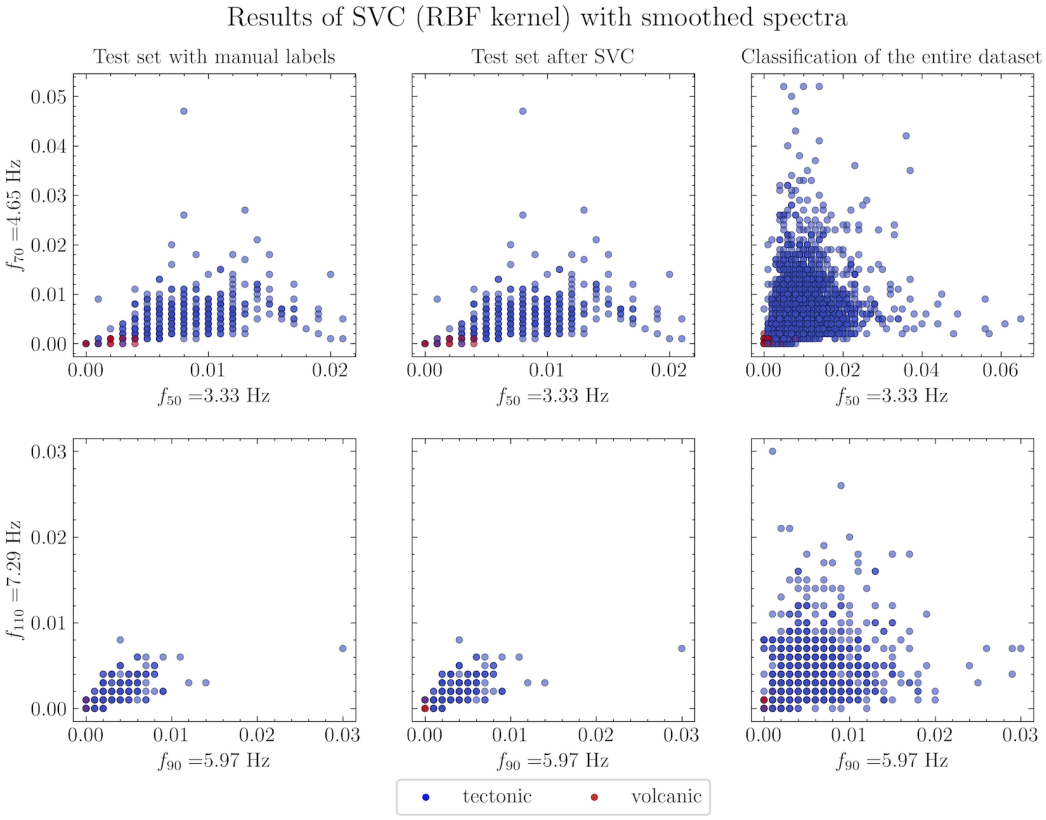

Figure 5.15: The labeled dataset with labels defined manually (first column) and the results of the SVM-RBF classifier on the labeled set and the entire data (second and third columns). Blue dots and red dots show the tectonic and volcanic events correspondingly.

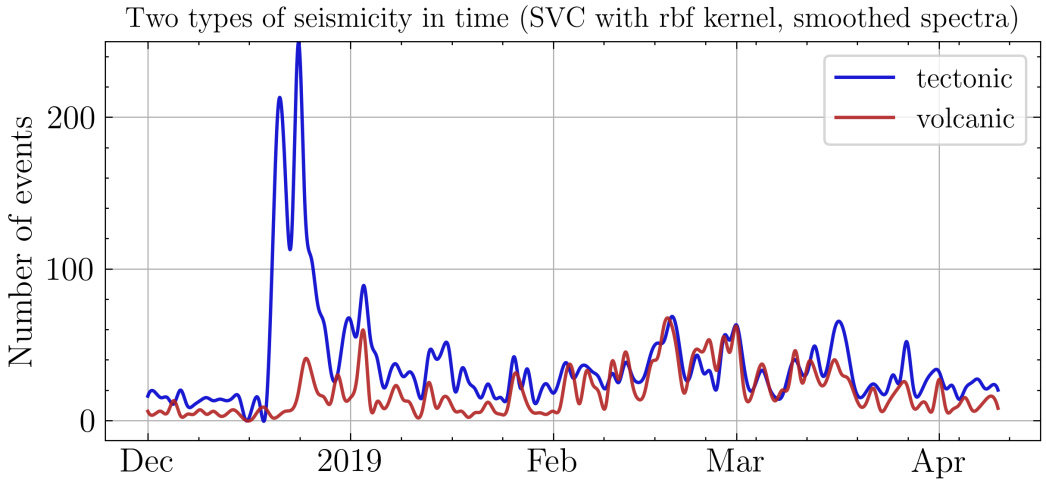

Figure 5.16: Temporal changes of tectonic (blue line) and volcanic (red line) activity based on SVM-RBF classification using smoothed spectra.

5.6 Results: Classification into three classes

The initial idea of this work was classification of seismicity into two general groups, volcanic and tectonic. Thus, the latter class includes all regional earthquakes that could be recorded. But at the same time the tectonic part of the labeled dataset contains mostly aftershocks of the strongest event of considered time period. In fact, we tried to separate earthquakes based on a source with a fixed location: the volcanic system of Shiveluch or the activated part of the subduction zone. However, the territory of Kamchatka region is too vast, and tectonic earthquakes are observed within the whole subduction zone on the regular basis, for example, there is was a relatively strong earthquake ($M_{\mathrm{L}} = 6.25$) with following aftershocks in the south of the peninsula.

Different hypocenters and consequently path effect should be reflected in the features of recorded signals. Thus, one can argue on inappropriateness of merging tectonic earthquakes in one big class instead of introducing smaller ones. For example, the simplest option here is to define (1) a class of local seismicity that includes the mainshock with its aftershocks, and (2) a class consisting of other regional events. Figure 5.17 shows events that represent these groups.

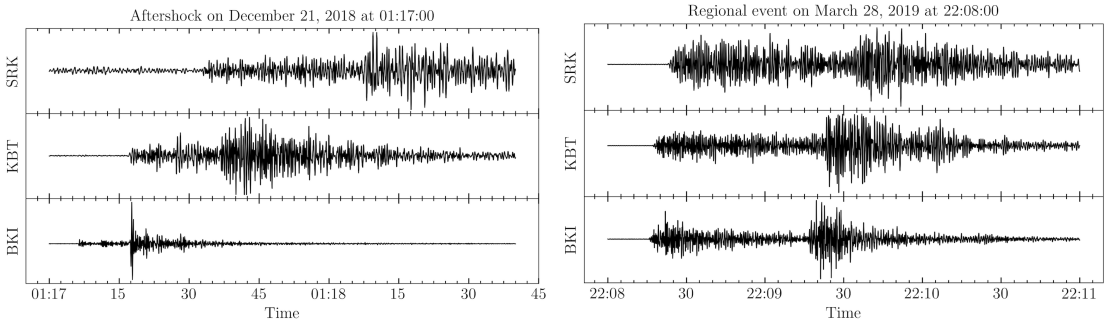

Figure 5.17: Examples of waveforms of tectonic earthquakes: an aftershock (left) and a regional event (right)

The selection of aftershock sequences from background seismicity was extensively studied for seismic catalogs declustering. These investigations started with window declustering approaches which discriminate aftershocks based on temporal and spatial windows. Here the simplest window algorithm from [Gardner and Knopoff, 1974] was used where the aftershock boundaries in space and time have the following form:

$$d = 10^{0.1238M+0.983}\,\mathrm{km} \tag{5.7}$$

$$t = \begin{cases} 10^{0.032M+2.7389}, \text{ if } M \leq 6.5 \\ 10^{0.5409M-0.547}, \text{ if } M \geq 6.5 \end{cases} \quad \text{days} \tag{5.8}$$

Given $M_{mainshock} = 7.2$, one can calculate that $d = 75$ km and $t = 932$ days. The obtained temporal window covers the whole study time period, so only the spatial parameter will regulate the type of an earthquake.

Figure 5.18 shows the map of Kamchatka with the distribution of tectonic earthquakes after selecting the aftershocks. Namely, the new labeled dataset contains in total 1291 events with SNR ≥ 1.35, of which 1035 aftershocks, 71 regional event and 185 volcanic earthquakes.

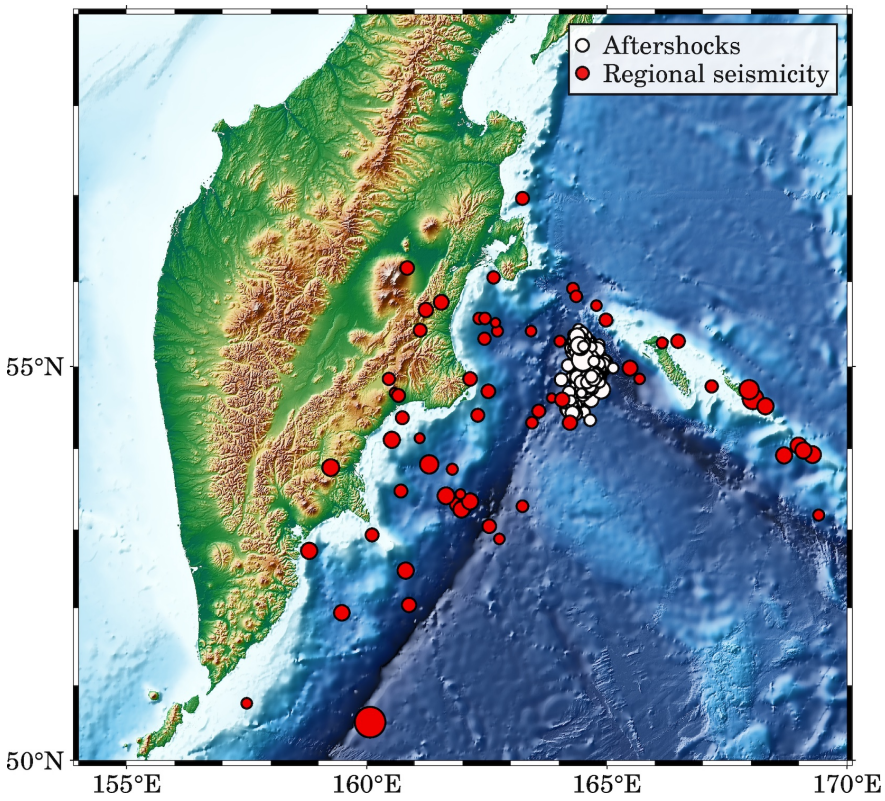

Figure 5.18: Two classes of tectonic activity: regional (red circles) and aftershocks (white circle). The circle size corresponds to the earthquake magnitude.

Figure 5.19 shows the result of the labeled dataset recompilation. It can be seen that many regional events were found before the mainshock and its aftershock. Moreover, with aftershocks decrease more background tectonic earthquakes are observed.

The procedure for classification or clustering the seismicity was the same as described in previous sections. Various signals representations ("regular" features and all spectral representations) were introduced to ML algorithms: k-means and agglomerative clustering as unsupervised methods, and SVM and Random Forest as supervised ones. There will be only summarized results presented.

Three classes clustering is too complicated for chosen algorithms due to complexity of objects distributions. As it can be seen from two dimensional projections of data

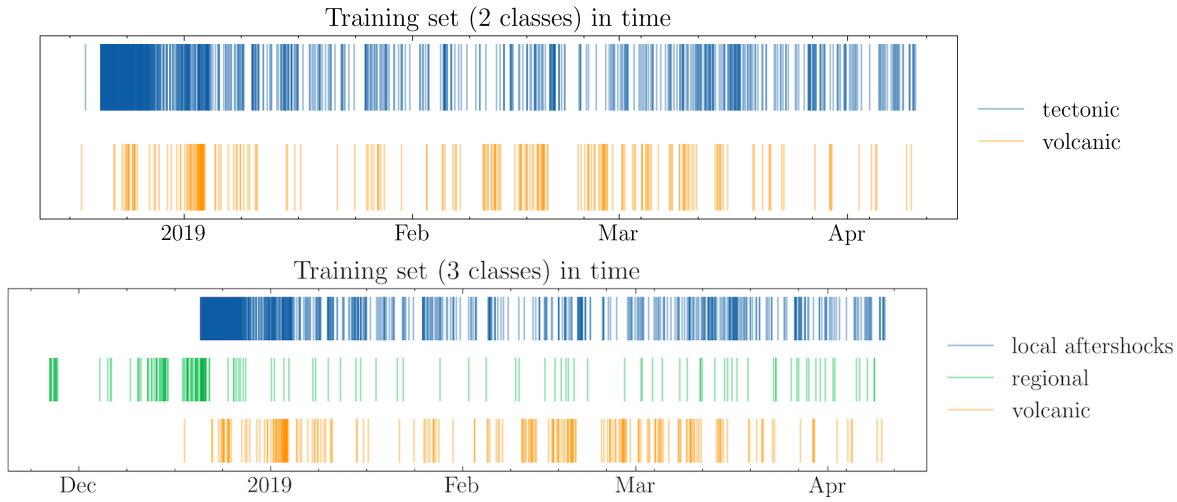

Figure 5.19: The labeled dataset before (top panel) and after (bottom panel) recompilation, i.e. splitting the big class tectonic activity into aftershock sequence and regional seismicity. Blue, green and orange bars show the aftershocks, regional and volcanic earthquakes correspondingly.

(first column in, e.g., Figure 5.5), there are no clearly defined clusters of objects even for a two class problem. Furthermore, the new class of regional seismicity is spread over two clusters instead of forming a separate cluster as it was expected. Increasing dimension of feature space did not help, and poor-quality results are obtained using spectral representations of signals.

Satisfactory classification results were obtained with SVM-RBF algorithm using "regular" features (Figure 5.20). After building a classifier it managed to distinguish a small cluster of regional seismicity among volcanic events (second column in Figure 5.20), although the accuracy of this result is poor: only 5 out of 12 regional events were classified correctly. After introducing the whole dataset to the algorithm, it has defined some events as regional (right column in Figure 5.20).

Comparing to the clustering results, there is an opportunity to plot the rate of three classes of seismicity in time (Figure 5.21). Some regional seismicity can be traced as a background of two other prominent classes. Nevertheless, the obtained result does not provide the explanation for the main question regarding the deviation of seismicity rate from the Omori law.

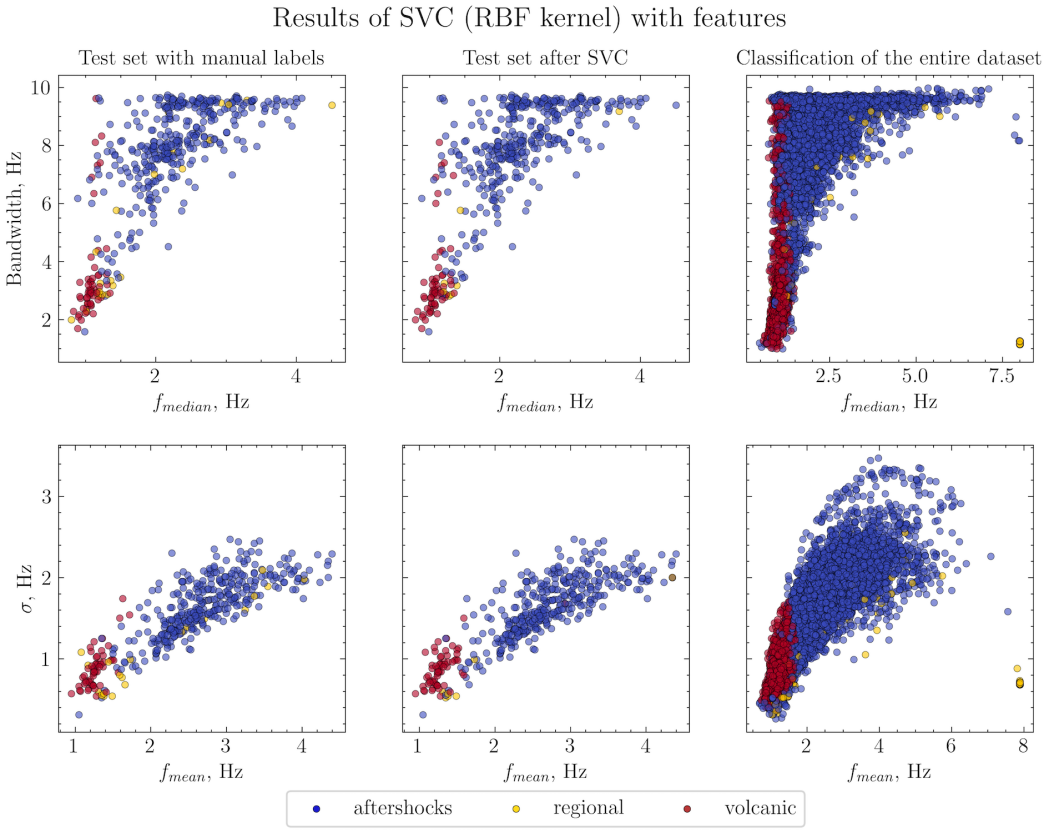

Figure 5.20: The test dataset with labels defined manually (first column) and the results of SVM-RBF classification into 3 classes of the test set and the entire data (second and third columns). Blue, yellow and red dots show the aftershocks, regional and volcanic earthquakes correspondingly.

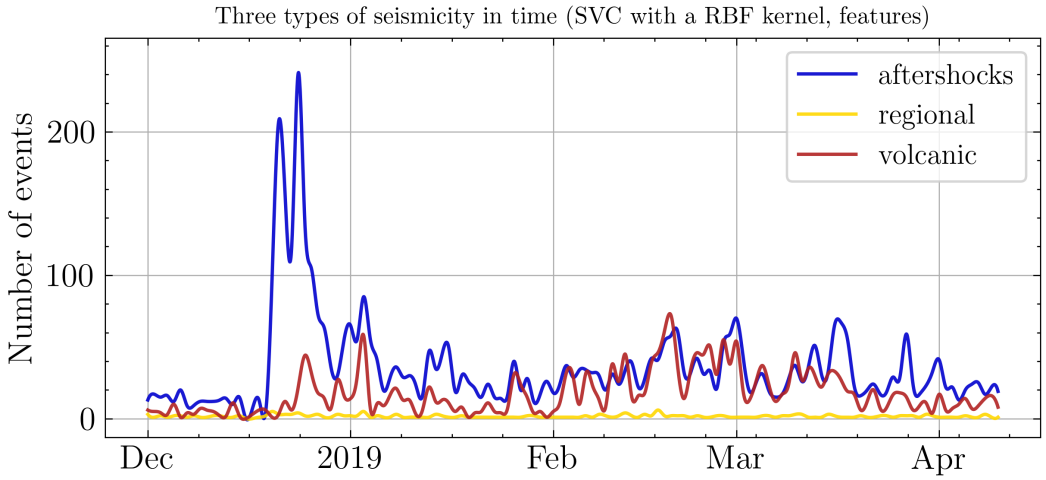

Figure 5.21: Temporal changes of aftershock (blue line), regional (yellow line) and volcanic (red line) activity according to SVM-RBF classification using "regular" features.

5.7 Discussion

5.7.1 On choice of features

One of the main steps in data exploration with machine learning algorithms is the right choice of features, i.e. the parameters of objects in a dataset. In this work we explore different types of signal representations with different number of points in a feature vector. In any case, any of used representations reflect the spectral characteristics of signals which are used by specialists to classify data in a conventional way. Nowadays abstract features are widely exploited with deep learning methods such as convolutional neural networks [e.g., Scarpetta et al., 2005, Perol et al., 2018, Seydoux et al., 2020]. These signal representations are extracted in unsupervised way and useful to discover new classes of seismicity. The problem of this work considers only two well-known classes of seismic activity, so we stuck to less elaborated representations of waveforms.

The obtained results support the idea that larger sets of features (or higher dimension of a feature space) do not necessarily provide more accurate classification or clustering. In ML practice this situation is referred as the *curse of dimensionality* [Bellman, 1957]. Indeed, the best performance was achieved using "regular" features with number N = 8 values in a feature set and smoothed spectra with N = 145. They allowed to divide prominent examples of volcanic and tectonic events while tectonic earthquakes of moderate and small magnitude are still mixed. Thus, some additional features that define the class of regional seismicity in the feature space should be found out. Again, considering parameters with more explicit physical meaning one can draw attention to some features explored in [Hammer et al., 2012, Malfante et al., 2018a], for example.

5.7.2 Clustering versus classification

In this work quite basic and easy to implement methods of machine learning were used: k-means and Agglomerative Clustering as unsupervised approach, and Support Vector Machine and Random Forest as supervised ones.

In the feature space of the lowest dimensionality where groups of objects do not have a well defined border between the them, clustering algorithms completely fail to separate data. At the same time, the supervised methods showed good performance on the test set (Tables 5.2 and 5.3), and after applying them to the entire dataset one can definitely distinguish the main characteristics of the seismic regime in the studied time period (Figures 5.8 and 5.10).

Regarding the feature space of a higher dimension built with smoothed spectra values, clustering techniques start to show better results (Appendix 5.C). Although k-means yet struggles with this unclear border between the clusters, Agglomerative Clustering provides the performance that can be compared with one of the supervised algorithms.

Nevertheless, as it was mentioned above the main obstacle for all the algorithms is tectonic events of moderate magnitude. It is unlikely that choosing more sophisticated ML algorithms would solve this problem. SVM and Random Forest have already shown their possibilities and strength in classification of various seismic signals [e.g., Giacco et al., 2009, Curilem et al., 2014, Hibert et al., 2017, Maggi et al., 2017, Malfante et al., 2018a].

5.7.3 On the one station and one channel approach

In this work data from only one channel of a single station was used. Starting with the amount of stations, it should be noted that the main interest of this work was the classification of data where the deviation from the Omori law was detected. Consequently, we aimed to reveal the origin of this deviation. The seismicity recorded at other stations can be interpreted with ease and data could be used complementary to the target station. However, the seismic network is sparse, and even if some signals are detected at several stations at the same time, the observed waveforms may have different origins. Thus, there is no clear advantage in using several stations.

Furthermore, we extracted information only from one of three available channels. Experience of using data of one trace or three components can be found ([e.g., Esposito et al., 2006, Maggi et al., 2017]). Intuitively it would seem that the information derived from three components could provide more information about waveforms. Indeed, some features regarding polarization can be obtained. However, it is unlikely that the signal-to-noise ratio of data used in this work would provide reliable information on the particle motion occurred during volcanic earthquakes. Also it should be mentioned that the seismograms of the vertical component of the same station were investigated absolutely in the same manner, and the obtained results did not turn out to be better.

5.8 Conclusion

In this chapter the problem of dividing seismicity in two groups was investigated. At the end of 2018 both tectonic and volcanic seismic activity was recorded with the stations of the permanent seismic network of the KB GS RAS. At the station located

further from the active Shiveluch volcano, a clear aftershocks decay fitting the Omori law was observed. At the same time, another station equidistant from the epicenters and the volcano, showed the same decay initially but then deviated from it. Thus, the main aim of this work was revealing the origin of this deviation. The additional condition was the use data from only one channel of a single seismic station.

A catalog of seismic events was automatically compiled based on signal-to-noise ratio of signals. Then, with the use of two "reference" stations the labeled dataset was built in a semi-automated manner. Then features, i.e. parameters carrying the information about signals, were extracted for each event in the resulting catalog.

We explored and compared the results of several supervised and unsupervised ML methods using different representations of seismic signals. In our case, classification algorithms provided more accurate results even with the smaller number of features. In some cases Agglomerative Clustering showed satisfactory results.

Surprisingly, this problem turned out to be more complicated that it was expected. We managed to build classifiers that can define the type of an earthquake when the events are large so the strong tectonic and volcanic activity at the beginning of the study period were distinguishing with satisfactory confidence. However, when borders between two general classes of seismicity become more obscure even proven algorithms do not manage this. Finally, the answer for the principal question on the reason of the Omori law deviation could not be obtained based on ML analysis of one-channel at one station.

Appendices to Chapter 5

5.A Grid search of parameters and kernel choice for the SVM classifier

As it was mentioned in Section 5.3.1 the performance of the SVM classifier depends on the parameter C and, in case of a polynomial and RBF kernels, parameter γ as well. The correct tuning of these hyperparameters allows to avoid overfitting. Also, the choice of a kernel is an additional task to be explored. Therefore, this appendix presents the results of the grid search of parameters using different kernels: linear, polynomial and RBF. To perform the grid search the labeled dataset was split into two parts: a train and a test sets. The first one was necessary to build a classifier and the second set is used to validate its performance. The grid was obtained by uniform distribution of parameters C and γ in a logarithmic scale from 10^{-3} to 10^{3}. For each node the classifier was built, and its score is represented with *accuracy* metric:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{5.A.1}$$

where TP is a True Positive, TN is a True Negative, FP is a False Positive and FN is a False Negative. In other words, accuracy is a fraction of predictions which are correct.

Figure 5.A.1 shows the results of the grid search, and the obtained values C and γ will be used further.

5.B Hyperparameters grid search for Random Forest classifier

Before building a Random Forest classifier one should tune various parameters to improve its performance and obtain more reliable results. Scikit-learn Python package [Pedregosa et al., 2011] used in this work offers the following options:

- max_depth is the maximum number of levels in each tree;
- min_samples_split is the minimum number of data points placed in a node before the node is split;
- max_leaf_nodes is the maximum number of leaf nodes a tree can have;

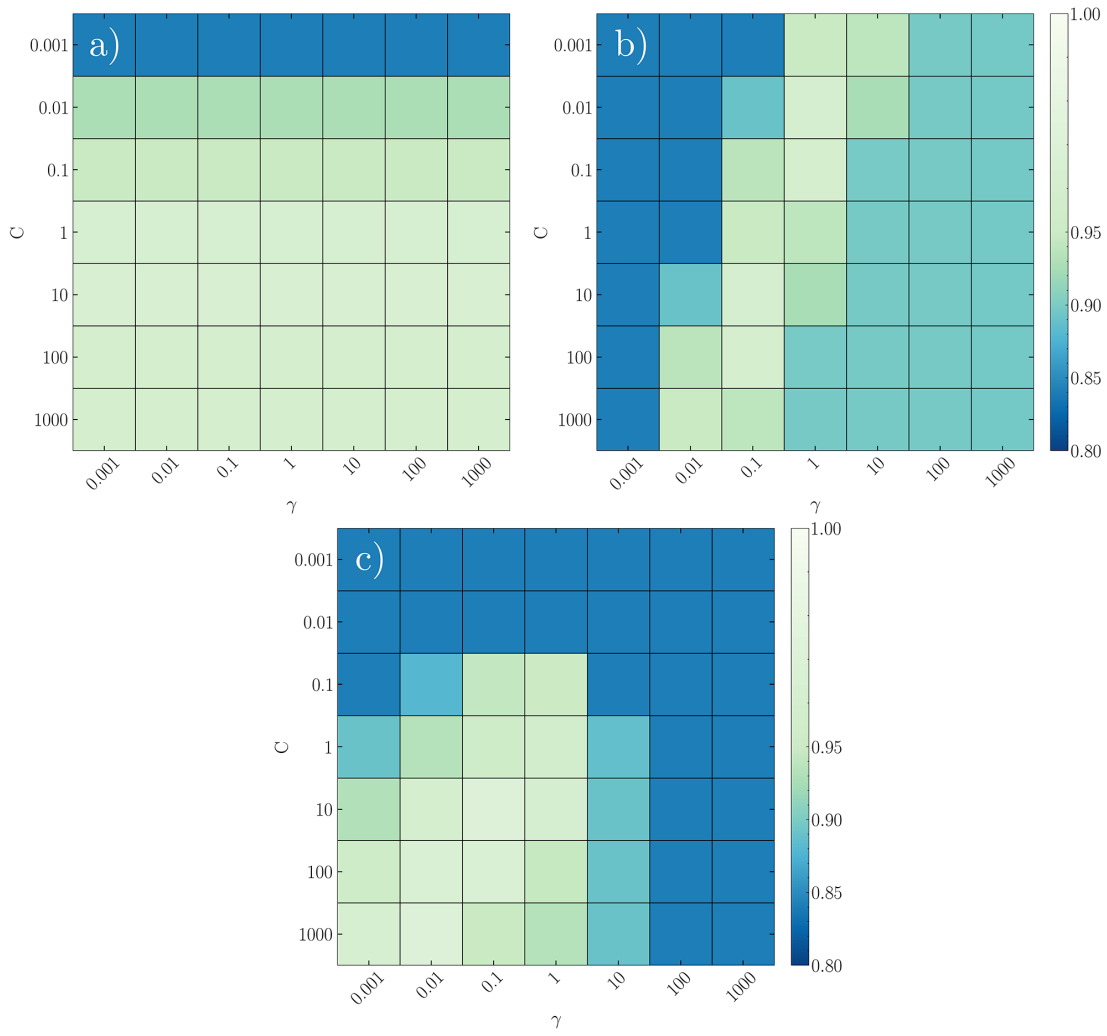

Figure 5.A.1: Results of grid search for different kernels and best parameters C and γ: a) linear kernel, C = 10; b) polynomial kernel (3^{rd} degree), C = $0.01, \gamma = 1$; c) RBF kernel, C = $10, \gamma = 0.1$.

- min_samples_leaf is the minimum number of samples required to be in a leaf node;
- n_estimators is the number of trees in the forest;
- max_features is the number of features randomly chosen and considered by the classifier when looking for the best split.

There are different ways to assess the performance of a model:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} = \frac{True\ Positive}{Total\ Predicted\ Positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} = \frac{True\ Positive}{Total\ Actual\ Positive}$$

It can be seen that precision reflects accuracy of a model. Therefore, it is a good measure when the costs of *False Positive* are high. On the contrary, recall shows how many of the *Actual Positive* were labeled as *True Positive* by the built model. In fact, recall is useful when there is a high cost associated with *False Negative*.

Another widespread measure in ML practice is *F1-score*, which is a function of precision and recall:

$$F_1 = 2 \cdot \frac{Presicion \times Recall}{Presicion + Recall}$$

Thus, F1-score provides balance between precision and recall, and this measure will be used further to tune the classifier hyperparameters.

To estimate the best parameters considering the peculiarities of data multiple runs were performed with one of the parameters changing over a set range of values while the others were fixed. Figure 5.B.1 presents the investigation results in the form of dependency of $F1$-score on the variation of one of the hyperparameters.

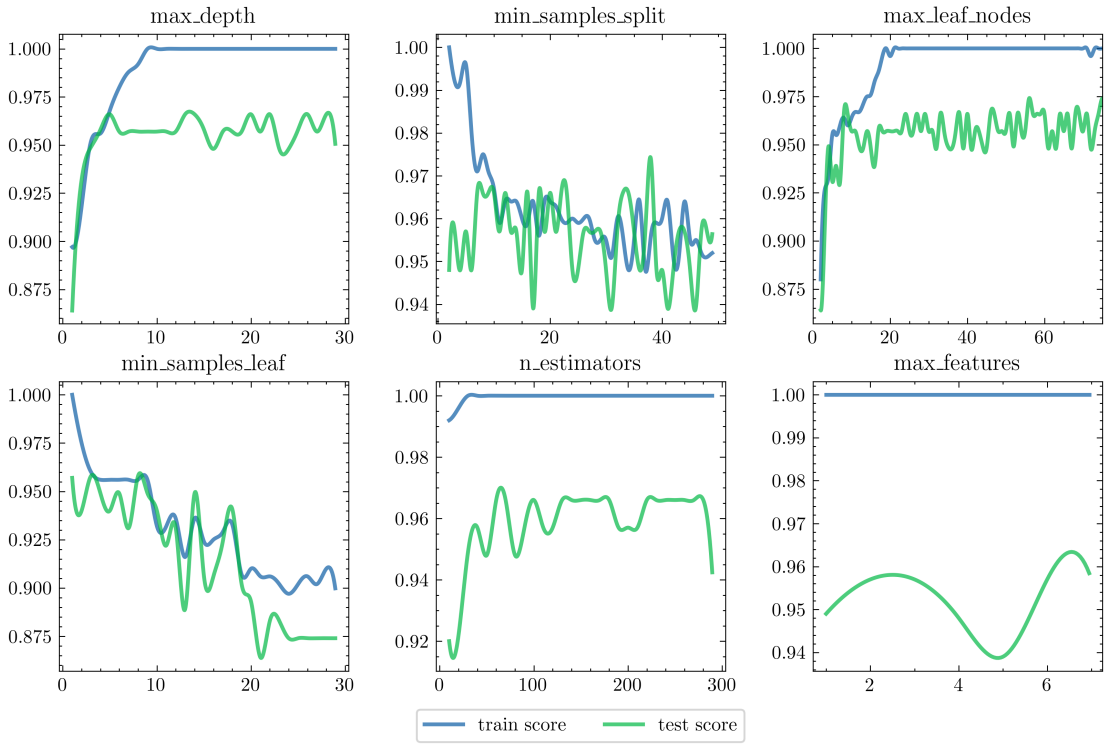

Figure 5.B.1: Performance of the Random Forest classifier on the train (blue) and test (green) set depending on values of its hyperparameters

It can be clearly seen from Figure 5.B.1 that increasing the values of parameters such as a tree depth or a number of trees does not lead to higher scores but obviously increases the computational costs. Thus, the final hyperparameters values (Table 5.B.1) were selected based on the compromise between these two factors.

max_depth	5
min_samples_split	10
max_leaf_nodes	10
min_samples_leaf	5
n_estimators	20
max_features	5

Table 5.B.1: Final values of RF hyperparameters used in further classifier building.

5.C Confusion matrices for classification using spectral representations of data

Smoothed spectra

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**361**	1	1,00
	Volcanic	6	**57**	0.90
Recall		**0.98**	**0.98**	

Unsmoothed spectra

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**361**	3	0.99
	Volcanic	6	**55**	0.90
Recall		**0.98**	**0.95**	

Spectrograms

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**347**	3	0.99
	Volcanic	20	**55**	0.73
Recall		**0.95**	**0.95**	

Table 5.C.1: Results of SVM classification (with an RBF kernel) using spectral representations of data

Smoothed spectra

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**360**	5	0.99
	Volcanic	7	**53**	0.88
Recall		**0.98**	**0.91**	

Unsmoothed spectra

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**359**	6	0.98
	Volcanic	8	**52**	0.87
Recall		**0.98**	**0.90**	

Spectrograms

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**347**	3	0.99
	Volcanic	20	**55**	0.73
Recall		**0.95**	**0.95**	

Table 5.C.2: Results of Random Forest classification using spectral representations of data

Smoothed spectra

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**995**	0	1.00
	Volcanic	118	**183**	0.61
Recall		**0.89**	**1.00**	

Unsmoothed spectra

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**858**	0	1.00
	Volcanic	255	**183**	0.42
Recall		**0.77**	**1.00**	

Spectrograms

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**775**	121	0.86
	Volcanic	338	**62**	0.16
Recall		**0.70**	**0.34**	

Table 5.C.3: Results of k-means clustering using spectral representations of data

Smoothed spectra

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**1022**	15	0.99
	Volcanic	91	**168**	0.65
Recall		**0.92**	**0.92**	

Unsmoothed spectra

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**1018**	0	1.00
	Volcanic	95	**183**	0.66
Recall		**0.91**	**1.00**	

Spectrograms

		True class		
		Tectonic	Volcanic	Precision
Predicted	Tectonic	**1037**	3	1.00
	Volcanic	76	**180**	0.70
Recall		**0.93**	**0.98**	

Table 5.C.4: Results of Agglomerative clustering using spectral representations of data

Chapter 6

Application: Activation of the Klyuchevskoy volcano group and the Shiveluch volcano eruption in 2022

6.1 Motivation

Another period interesting for analysis occurred in October-November 2022 due to reactivation of the Klyuchevskoy volcano group. At the beginning of the study period the Shiveluch volcano had been already permanently active for several months. The Bezymianny volcano was episodically active and had already erupted in March and May 2022. Its next eruption occurred on October 23 within the studied period. On November 19, 2022 KB GS RAS reported a strong (Ml = 5.7) tectonic earthquake (yellow star in Figure 6.1) followed by increased seismicity in the Klyuchevskoy volcano edifice and its dome glowing. Consequently, the seismic stations in the KVG region should have recorded signals from several volcanoes at the same time.

Thus, if Chapter 5 deals with two big classes of seismicity, tectonic and volcanic, this part of work considers only volcanic seismicity generated by various sources. Also here we consider only unsupervised ML approach, as we seek to explore the changes in the seismic regime and possibly discover new types of signals rather than fit waveforms into the conventional classification scheme.

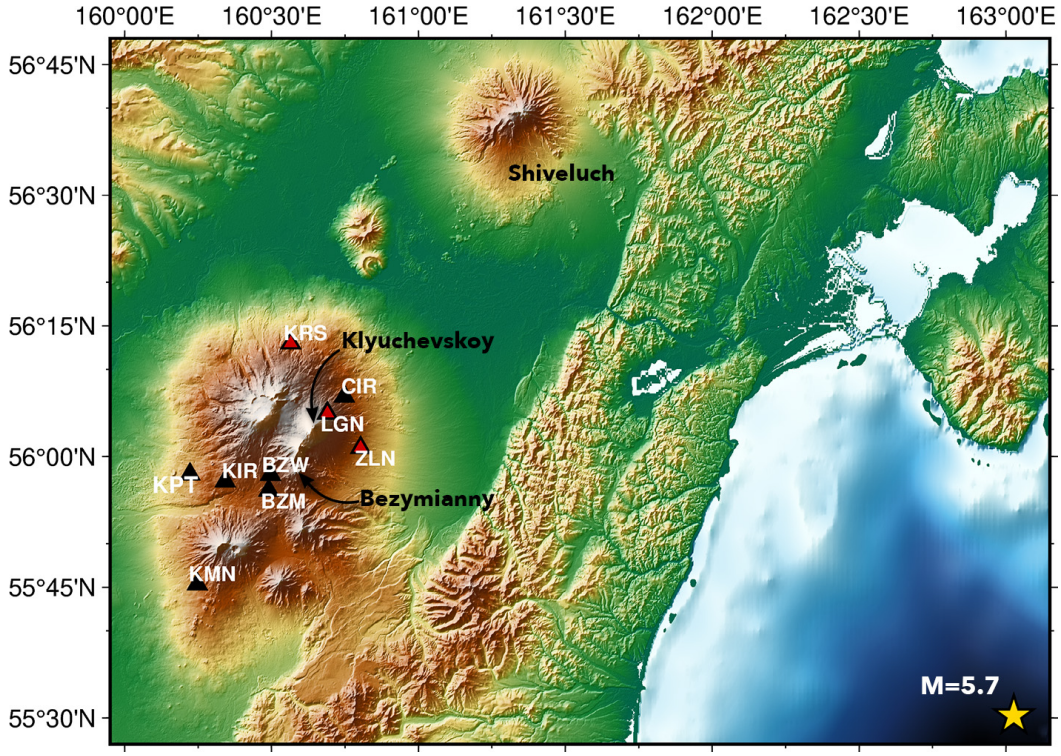

Figure 6.1: Map of the KVG region and the Shiveluch volcano. Stations of the permanent network data from which was used in this study are shown with triangles and corresponding codes, seismic data used for clustering is taken from stations marked with red triangles

6.2 Methods

6.2.1 Data preprocessing and visual inspection

For seismic data exploration we chose the stations of the KB GS RAS permanent network that could confidently register signals originated in different volcanoes, these stations are shown with red triangles in Figure 6.1. The seismic data processing was performed in the manner described in section 4.3 and included the next steps: preprocessing with bandfiltering (1–5 Hz), detection of signals using their signal-to-noise ratio and extraction of features from the reliable detections. For the ML application we used parameters that are referred as "regular" features, i.e. the characteristics of the signals that have clear physical meaning, such as amplitude, peak frequency of a spectrum, etc.

At the next step the final dataset, i.e. the feature space was plotted in all possible two dimensional projections. This allows us to visualize the structure of the dataset

and chose the appropriate ML algorithm or make corrections if necessary. Sets of all possible projections are contained in Appendix 6.A, and Figure 6.2 shows three of them based on data recorded at station KRS, SHN channel. These plots were chosen because some characteristic distributions of objects can be seen. Besides clear dense clusters, the color of dots representing the detection time shows some temporal changes in seismic regime. The very dense cluster in the middle panel of Figure 6.2 attracts the most attention, and hereinafter we focus on these events.

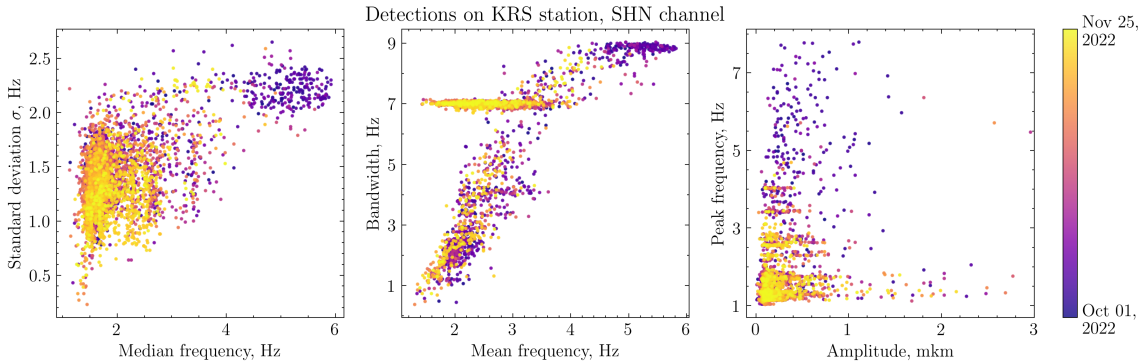

Figure 6.2: Projections of the feature space of the waveforms recorded at station KRS, SHN channel (dot color corresponds to the detection time)

6.2.2 GMM clustering

Observed distribution of objects in Figure 6.2 encourages the use of the Gaussian mixture model (GMM) algorithm explained in details in subsection 4.2.2. This clustering algorithm requires a predefined number of clusters, and this value N_{clust} was set as 4 after visual inspection. Nevertheless, Appendix 6.B contains results of clustering with different values of N_{clust}, and it can be seen that the cluster of the main interest is confidently distinguished from other clusters in any case, and from this perspective the result almost does not depend on N_{clust}.

Figure 6.3 shows the result of clustering the data from station KRS. Particularly interesting is cluster 1. Although it looks very small in the feature space it turns out to have the largest number of events if one plots the seismicity rate in time. Figure 6.4 shows both daily and cumulative count of detections alongside with the major events in the KVG region.

The same results can be traced using data from the other stations chosen for the study, LGN and ZLN (shown as red triangles in Figure 6.1). Both of them recorded some distinct class of seismo-volcanic activity which is confidently clustered by the GMM algorithm.

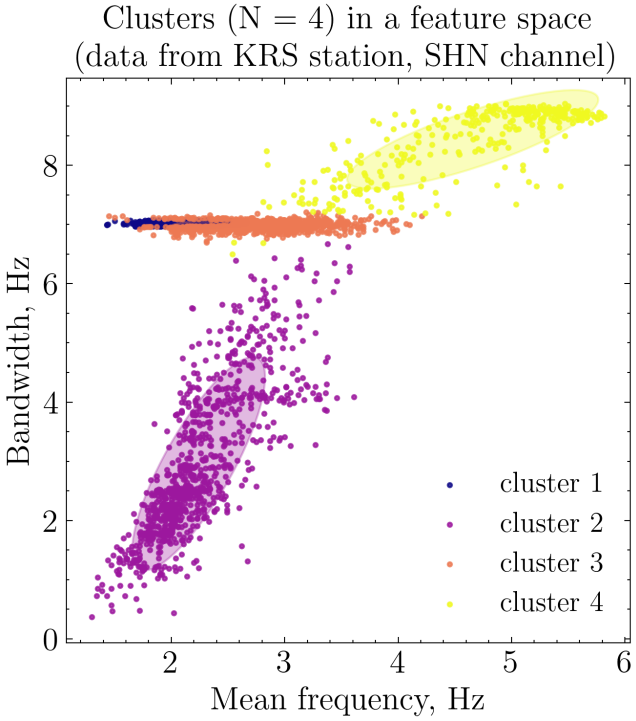

Figure 6.3: Results of the GMM clustering ($N_{clust} = 4$, data obtained from station KRS, channel SHN)

6.2.3 Multiplet detection with cross-correlations of waveforms

The high density of objects in the selected cluster suggests a very repetitive and stationary source producing these signals [Shapiro et al., 2017a]. If so, this observation should be reflected in the similarity of waveforms within the cluster of interest, and can be verified by the computing correlation coefficients between all pair of waveforms.

To carry out such analysis, 8 stations of the permanent network of KB GS RAS were chosen based on data availability and SNR level, they are shown with triangles in Figure 6.1. After band-filtering (1–5 Hz) and decimating, the waveforms of 100 s duration were cross-correlated with the variable time shift in order to find the highest correlation coefficient between signals. Thus, for every day of the study period a matrix of correlation coefficients was obtained. Figure 6.5a shows a daily cross-correlation matrix for November 16, 2022.

Multiplet families of highly correlated signals were observed almost every day during the studied period, and for each group within one day a master event was obtained in the following way. Summing the values over every row or column of a cross-correlation matrix reveals the events with the highest correlation coefficients. Then, using a threshold value of $CC = 0.55$ other signals were reassigned to master events. An example in Figure 6.5b shows an event whose family includes 39 earthquakes.

The first step of cross-correlation analysis was performed day by day due to computational efficiency. After extracting the multiplet families and obtaining corresponding

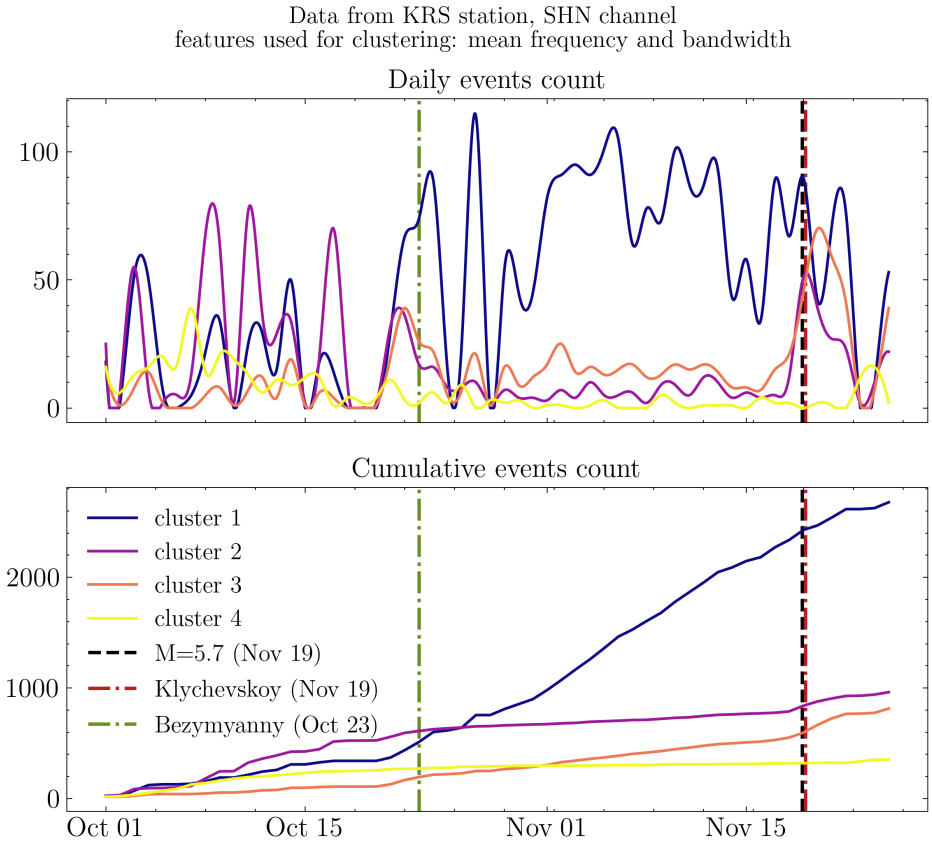

Figure 6.4: Seismic activity rate according to the GMM clustering: daily (top) and cumulative (bottom) count. Lines colors correspond to colors of clusters in Figure 6.3. Vertical lines show the main events in the region: eruption of Bezymianny on October 23 (green dashed line), Klyuchevskoy activation (red dashed line) after a tectonic earthquake with M=5.7 (black dashed line) on November 19

master events, it is possible to find out the relationships between these groups by correlating the waveforms of master events in the same way. Figure 6.6 shows the clear presence of a stable source active over the entire study period that can be traced at different stations.

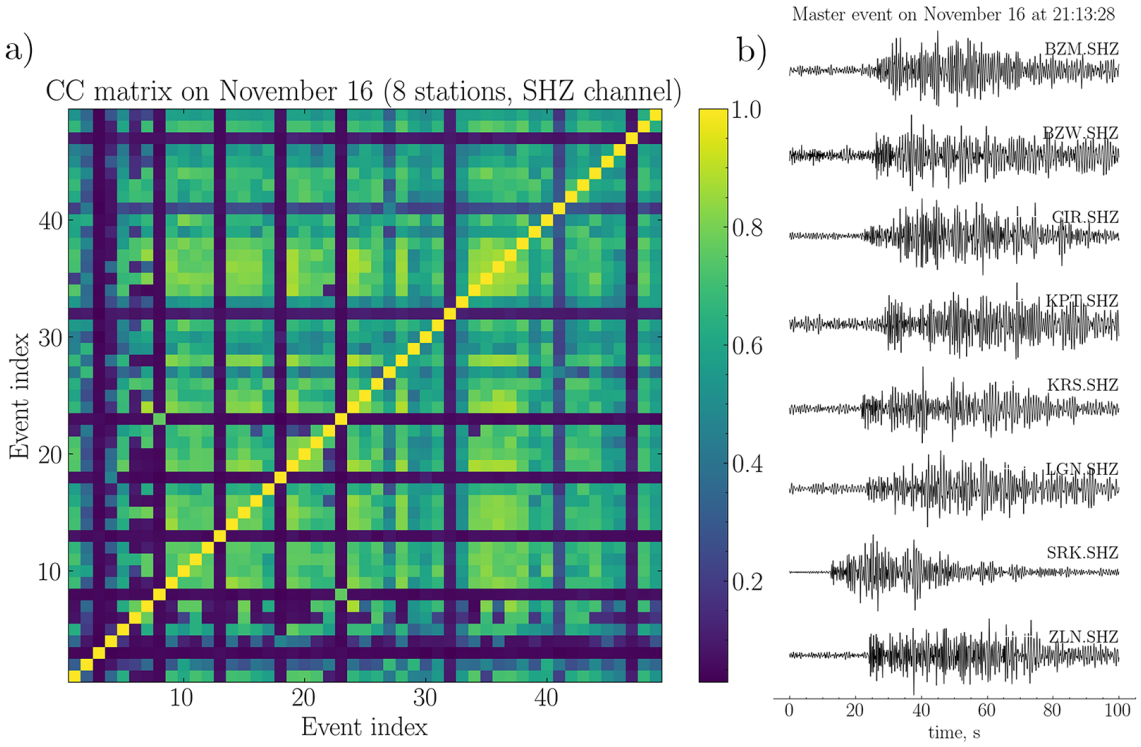

Figure 6.5: (a) Cross-correlation matrix of 49 events detected on November 16, 2022. (b) Waveforms of the master event at the stations used in the study.

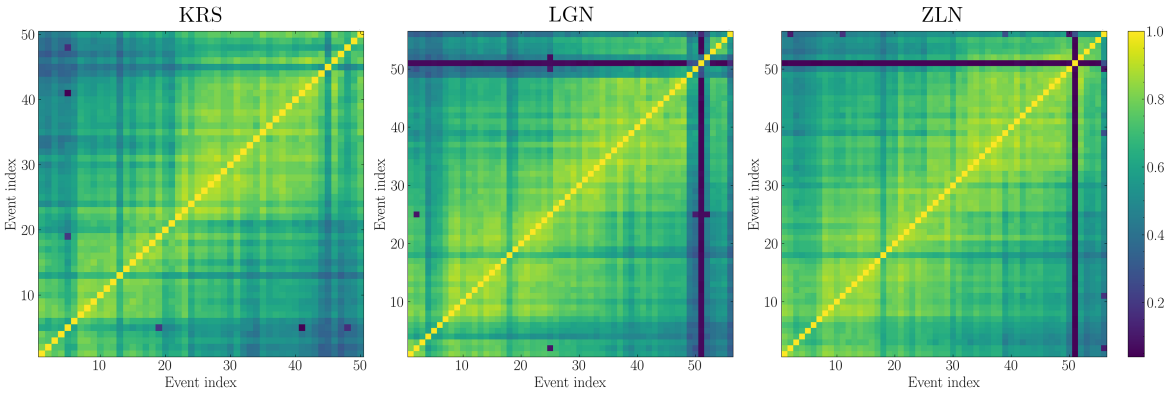

Figure 6.6: Cross-correlation matrix of master events at stations KRS, LGN and ZLN

6.3 Results

6.3.1 Comparison of the GMM clustering and the waveforms cross-correlation

This chapter presents investigations of seismic data recorded during reactivation of the Klyuchevskoy volcano group and the Shiveluch volcano eruption in October-November 2022. After compilation of a catalog of seismic events and their features two approaches were applied independently to the obtained dataset. Firstly, the Gaussian mixture model algorithm revealed a dense and populous cluster of signals in the feature space. Then cross-correlation analysis has shown persistent seismicity during entire study period with waveforms of high similarity. Here a question arises: are these two observations describing the same phenomenon? To answer it, we calculated the daily rates of seismicity revealed by two independent algorithms.

Every panel of Figure 6.7 contains two lines. Black lines show the daily event rate in the dominant cluster found with the GMM. Green lines show the event rate detected with the multiplet cross-correlation analysis by calculcating the number of events in the most numerous multiplet family of a day. Again, this is not the template matching detections done with cross-correlation of a template waveform with the continuous seismograms.

Coming back to the results in Figure 6.7, there are some deviations between the plots, however the clear correlation can be traced in all cases. These observations speak in favor of a presence of a persistent, stationary source of seismic signals that is dominating the overall seismicity during the period of re-activation of the KVG.

6.3.2 Location of the source of the dominating seismicity

Waveforms of five master events revealed through cross-correlation analysis were chosen for manual processing based on their SNR and data availability, their origin times are contained in Table 6.1. To estimate their hypocenters, HYPO71 software was used [Lee and Lahr, 1972]. This algorithm requires following input data: P- and S-waves arrival times, stations locations and a 1D velocity model. Figures 6.8–6.9 show the example of the result of manual seismograms processing in the form of picked P- and S-waves arrivals. Due to poor coverage, as all the stations are located from one side of potential hypocenters and the great hypocentral distance, the locations were obtained based mostly on P-wave arrival times, and S-wave could be distinguished only at the closest station SRK.

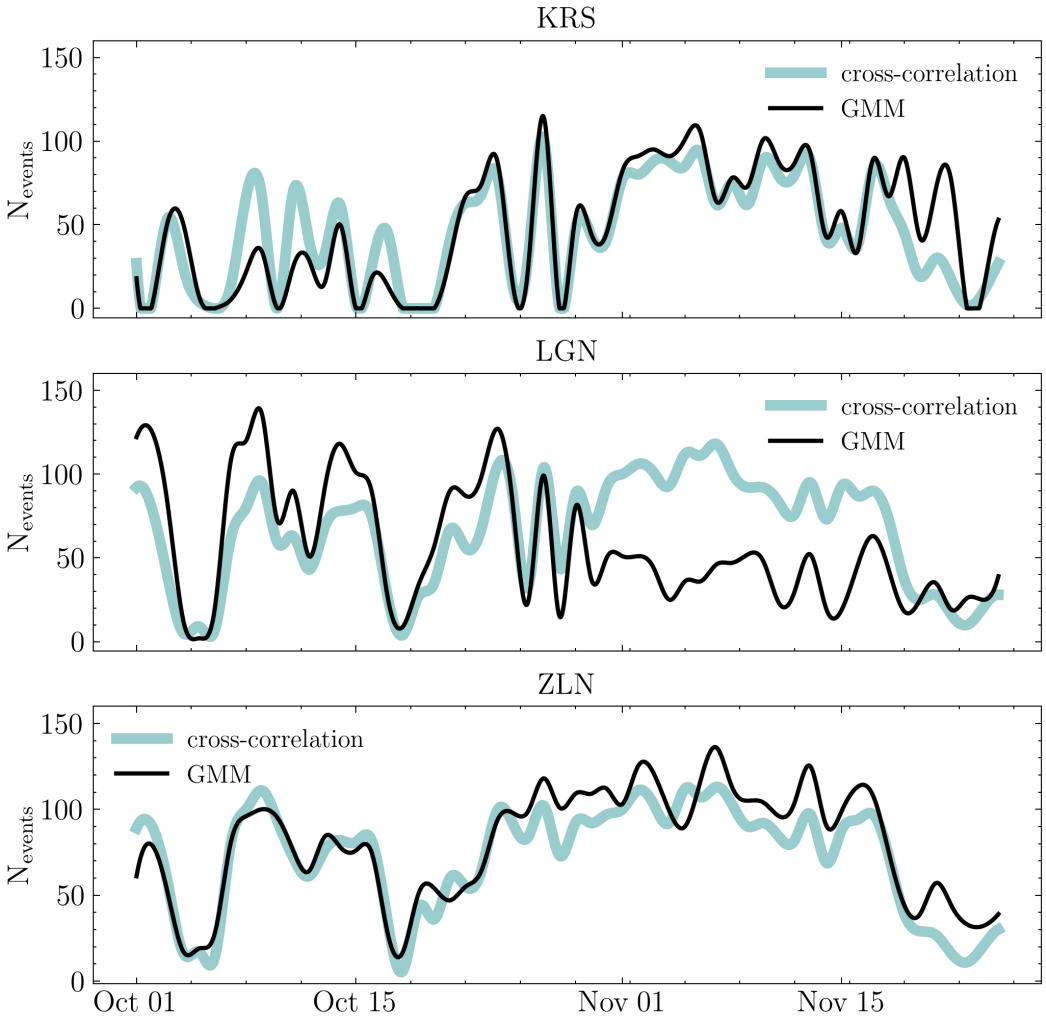

Figure 6.7: Comparison of the results obtained by clustering data with Gaussian mixture model (black line) and cross-correlation analysis (green line)

After processing the selected events, we found that the epicenters lie to the northwest of the main edifice of Shiveluch. The final solutions are presented in Table 6.1 and are shown in Figure 6.10 with red diamonds, uncertainties of the locations are shown with black bars as well. The errors in hypocenters locations can be significantly large due to poor coverage with stations despite their number (from 7 to 9 stations were used for inspection).

The obtained locations lay slightly aside from the edifice of Shiveluch at middle depths of 12 km. These depths are quite unusual for this cluster of volcanic seismicity: commonly earthquakes beneath Shiveluch are reported at very shallow depths of a few kilometres. However, the depth errors are very large and can be underestimated with the location procedure based on 1D model. A really distinctive property of the cluster of seismicity proximate to the Shiveluch volcano is that its intensification started nearly

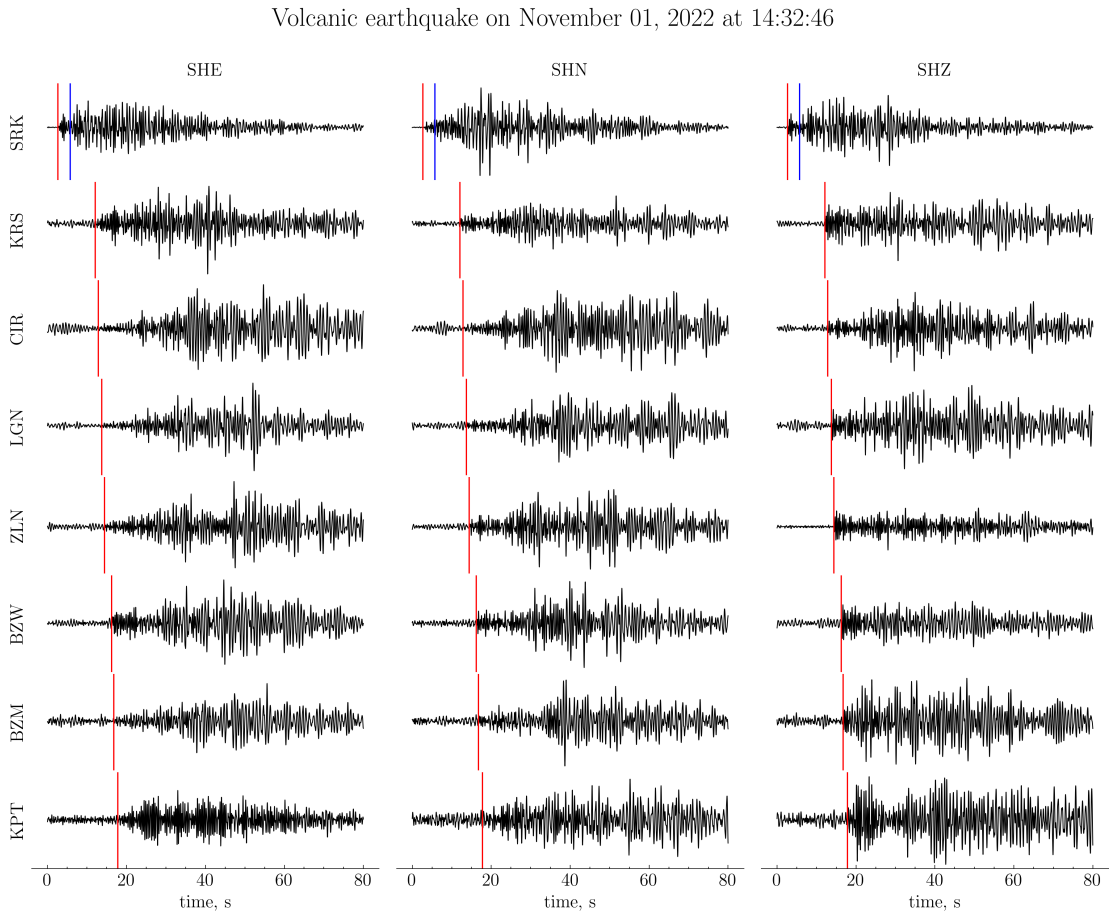

Figure 6.8: Manual definition of P- (red line) and S-waves (blue line) arrivals for the earthquake on November 1 at 14:32:52 (Event No. 4 in Table 6.1)

simultaneously with the Bezymianny eruption and continued up to the reactivation of Klyuchevskoy. This possibly indicates a change of the Shiveluch eruptive dynamics during the activity of KVG and might reflect an interaction between different volcanoes of Kamchatka similar to what has been previously suggested [e.g., Coppola et al., 2021].

No	Date	Time	Latitude	Longitude	Depth
1	2022-10-01	20:55:16.22	56.700	161.175	11.9
2	2022-10-11	08:32:34.18	56.705	161.300	12.3
3	2022-10-14	03:38:07.48	56.726	161.264	12.9
4	2022-11-01	14:32:52.74	56.784	161.267	13.3
5	2022-11-09	14:02:59.57	56.711	161.146	11.5

Table 6.1: Selected master events and their hypocenters.

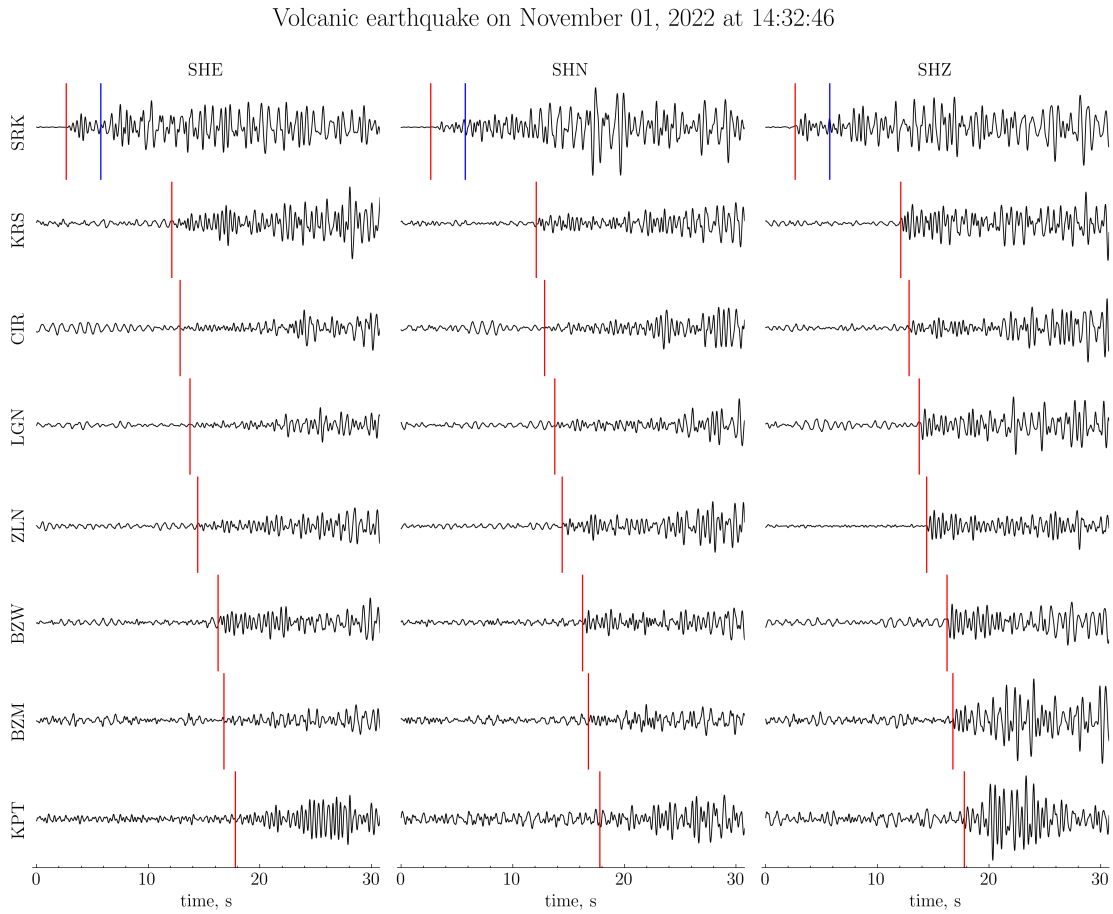

Figure 6.9: Manual definition of P- (red line) and S-waves (blue line) arrivals for the earthquake on November 1 at 14:32:52 (Event No. 4 in Table 6.1), zoomed

6.4 Conclusion

The main goal of volcano monitoring is the recognition of the state of a volcano using various observations. This chapter shows the results of processing seismic data recorded during reactivation of the Klyuchevskoy volcano group and the Shiveluch volcano eruption in 2022. The analysis was performed using two different techniques independently. First, we identified a class of stationary seismicity with extracting features from the signals and then clustering obtained data with the Gaussian mixture model algorithm. Further, cross-correlation of waveforms revealed a big multiplet family of earthquakes occurring almost every day during the studied period. Needless to mention, that presented approaches are quite rough and are subject to elaboration. However, using these methods and data from few stations it was is possible to discover an interesting cluster of seismicity that may reflect interaction between different volcanoes.

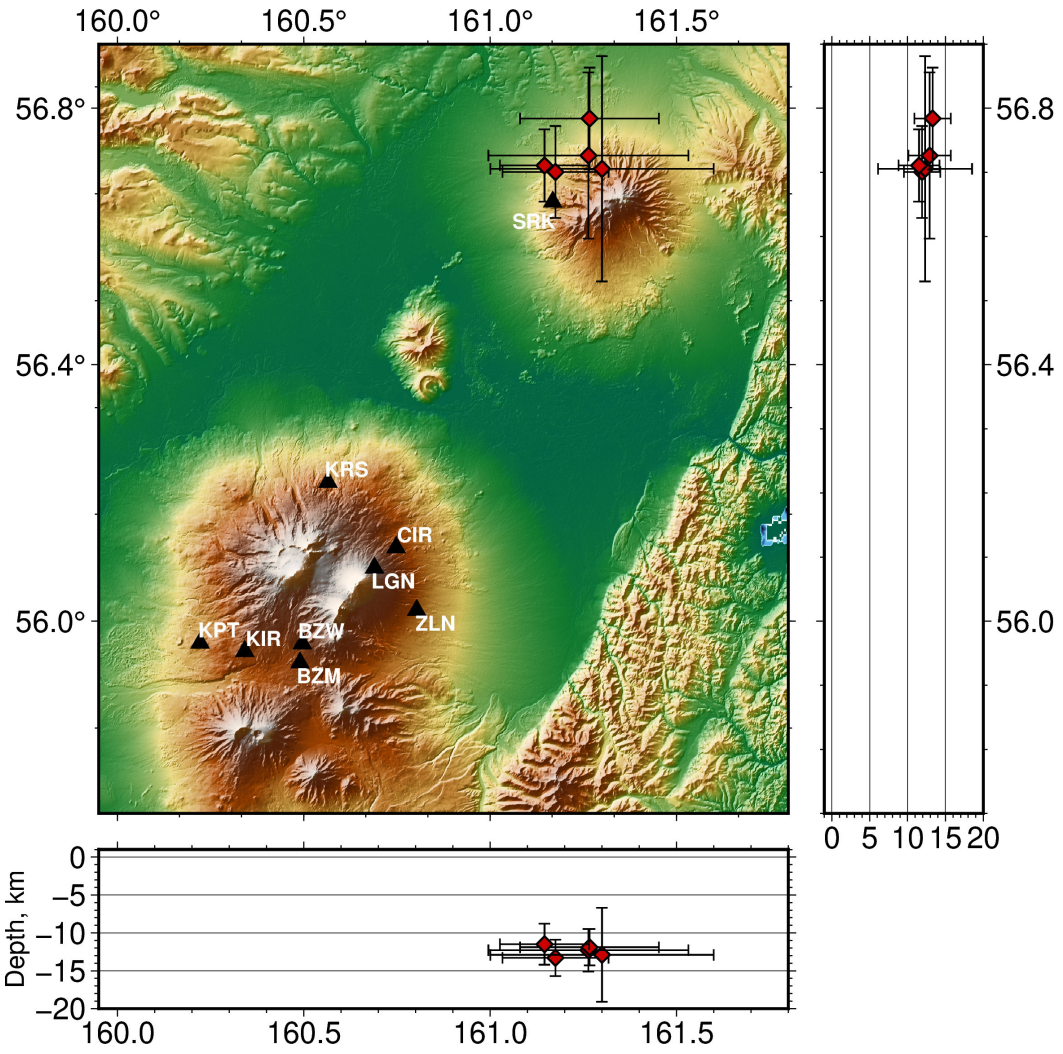

Figure 6.10: Map of the KVG and Shiveluch with defined hypocenters of selected master events (red diamonds). Seismic stations data from which was used are shown with black triangles, black bars represent the uncertainties of locations.

Appendices to Chapter 6

6.A 2D projections of features space

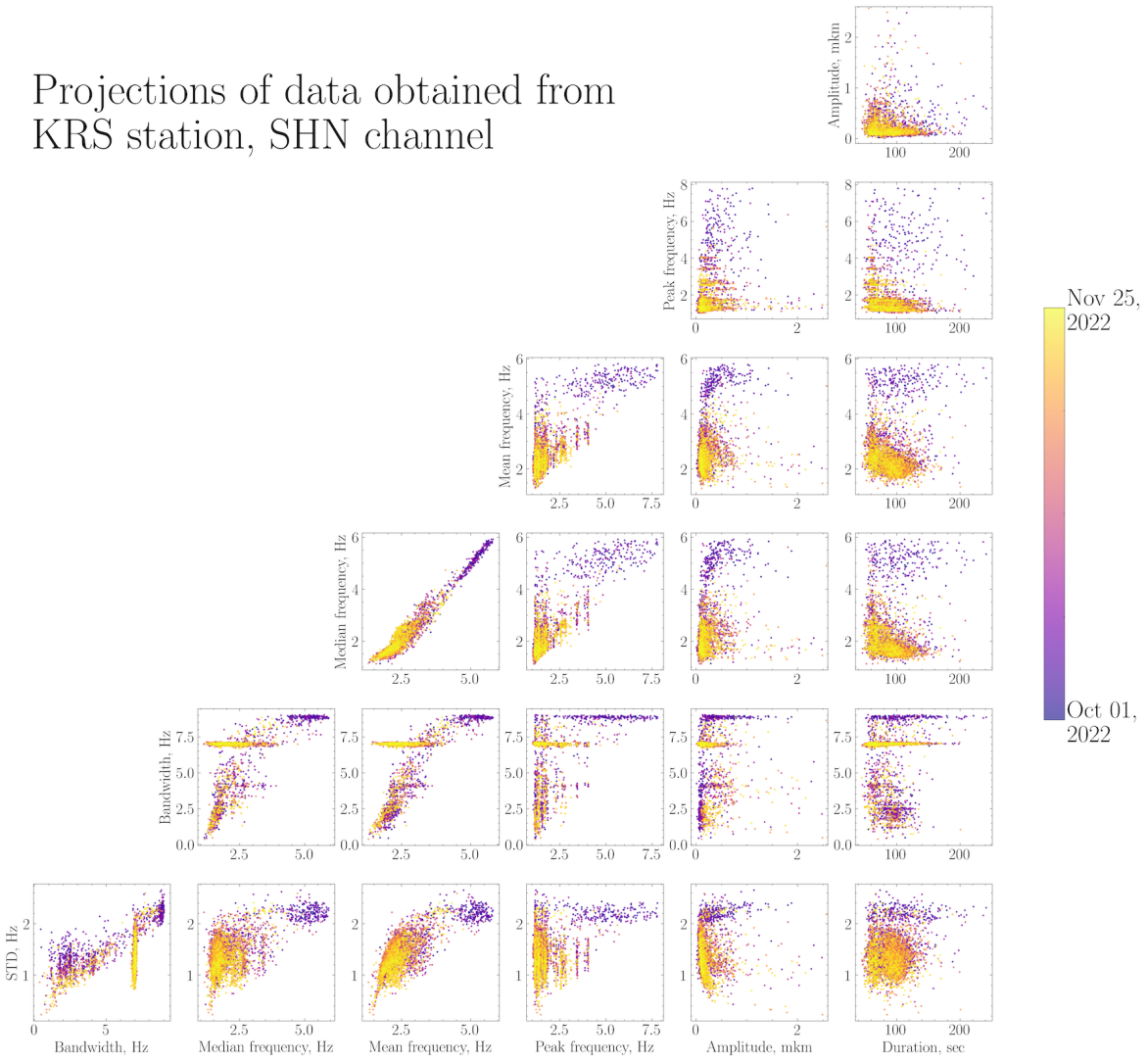

Figure 6.A.1: Projections of a feature space for the data recorded at station KRS, SHN channel

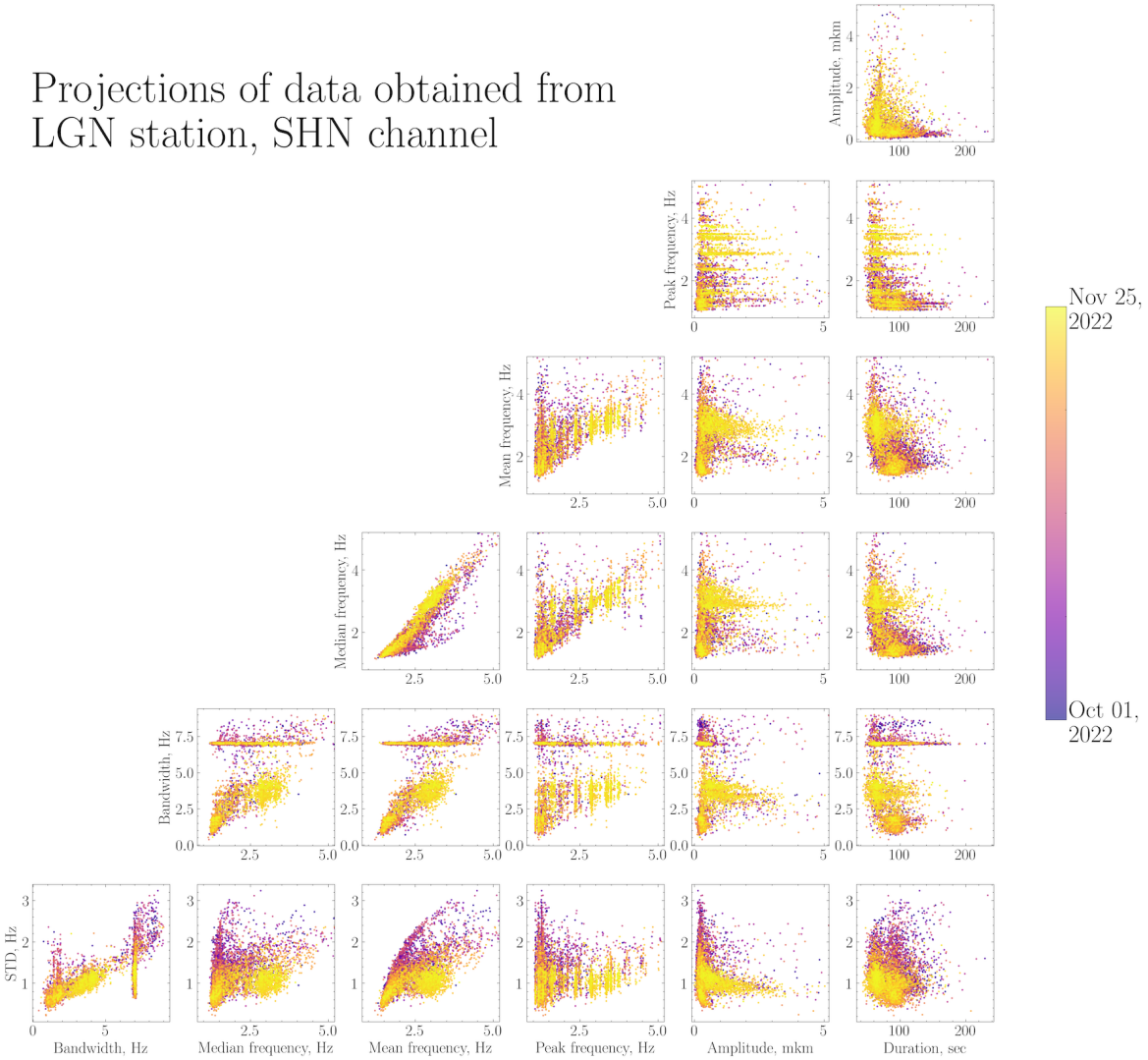

Figure 6.A.2: Projections of a feature space for the data recorded at station LGN, SHN channel

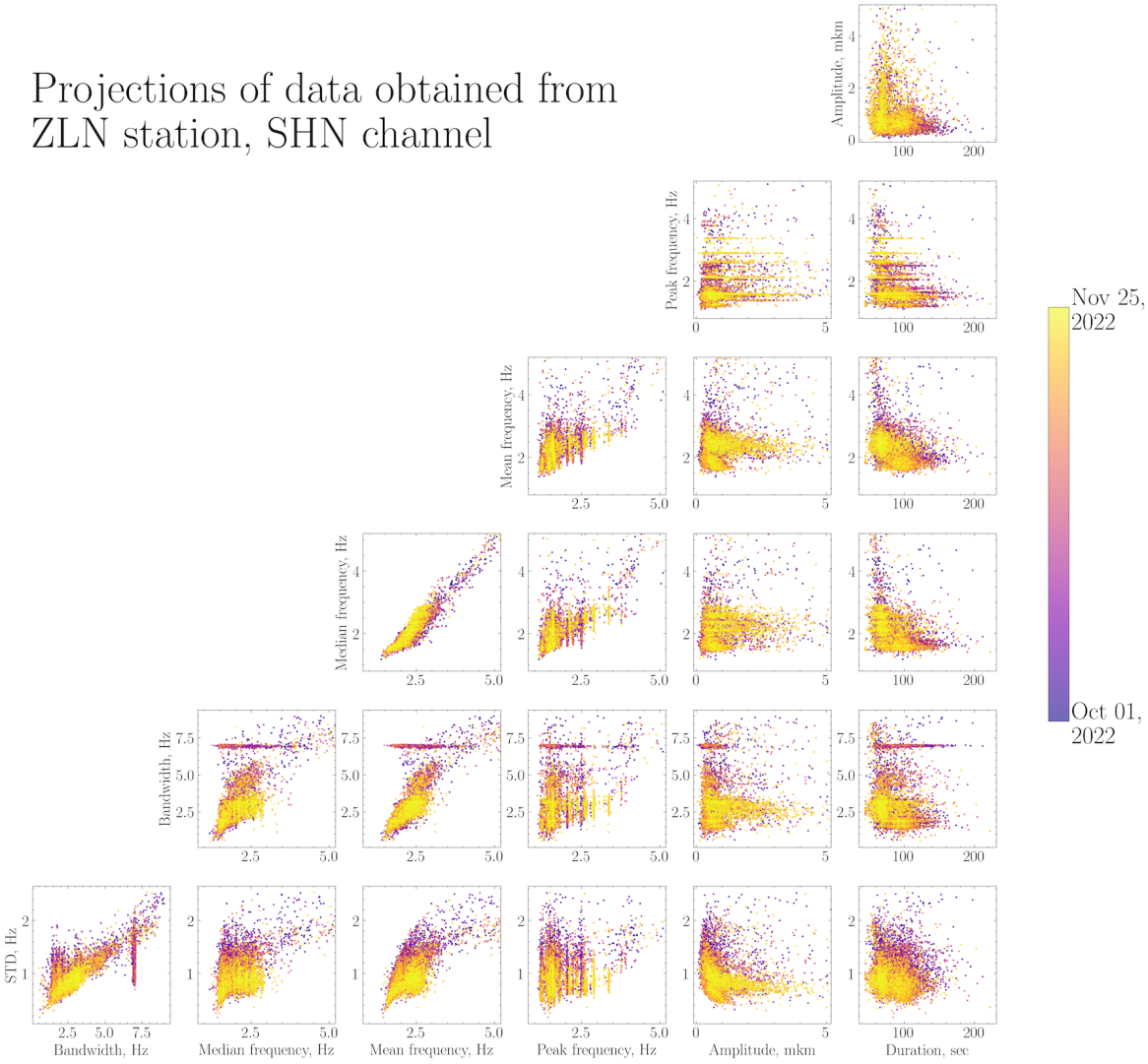

Figure 6.A.3: Projections of a feature space for the data recorded at station ZLN, SHN channel

6.B Results of the GMM clustering with different number of clusters

6.B.1 Station KRS

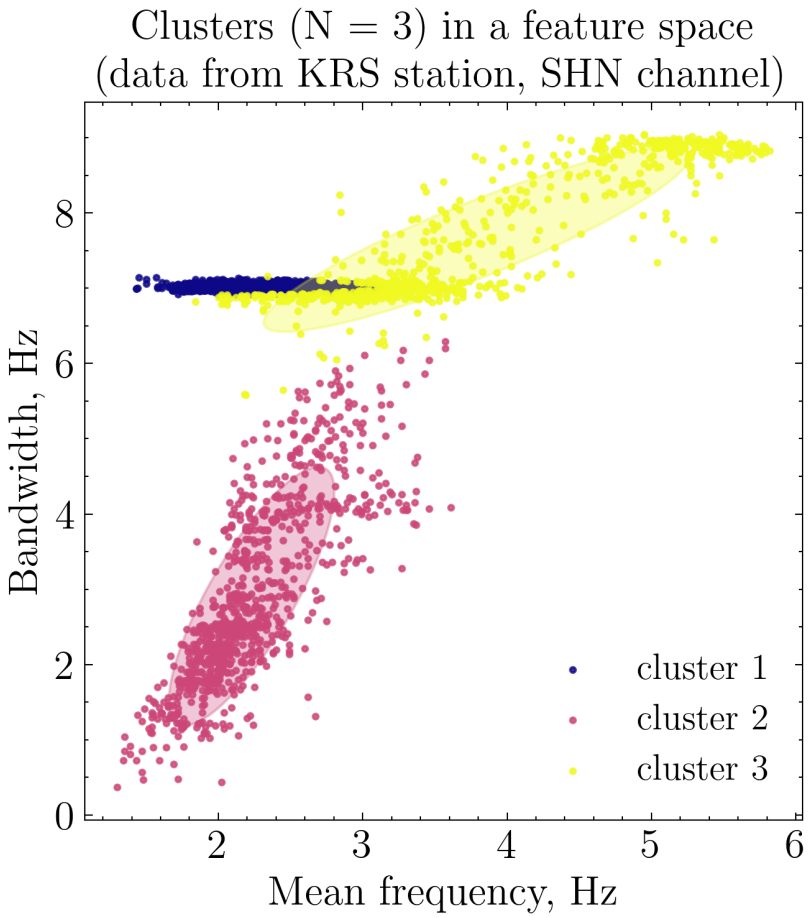

Figure 6.B.4: Results of GMM clustering ($N_{clust} = 3$, data obtained from station KRS, channel SHN)

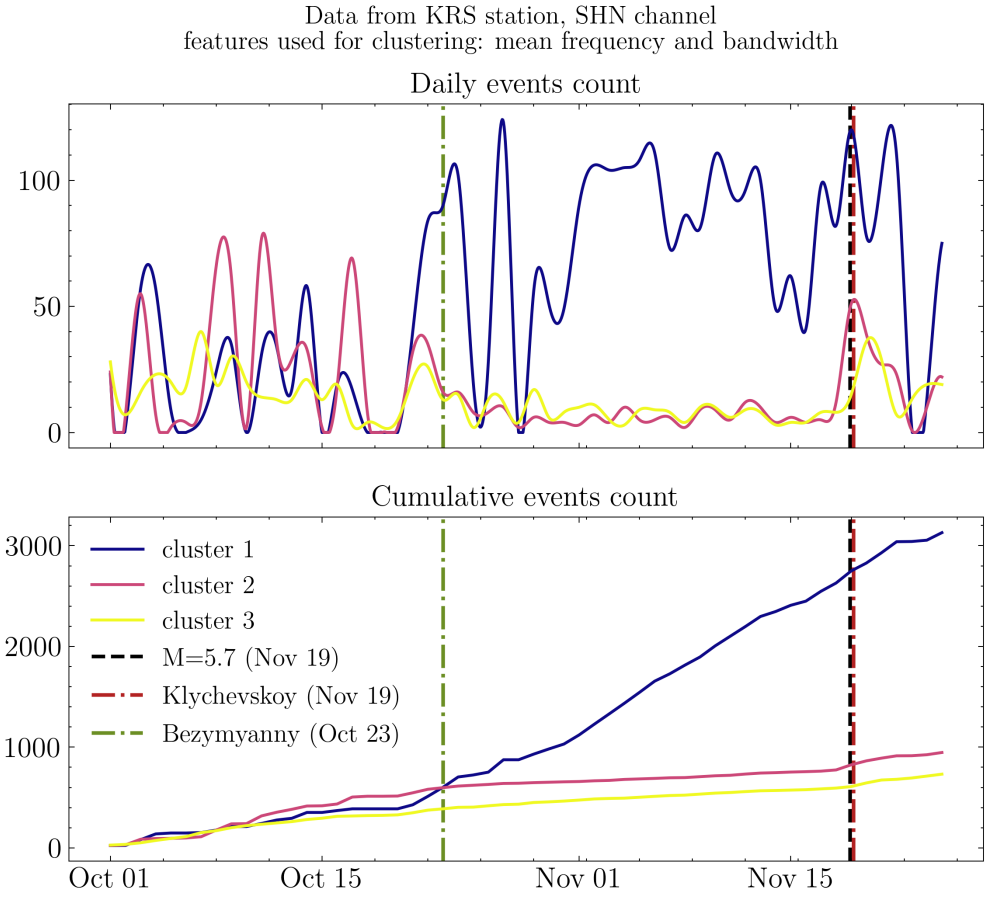

Figure 6.B.5: Seismic activity rate according to GMM clustering: daily (top) and cumulative (bottom) count. Lines colors correspond to colors of clusters in Figure 6.B.4. Vertical lines show the main events in the region such as volcanic eruptions and strong tectonic earthquakes (black dashed lines)

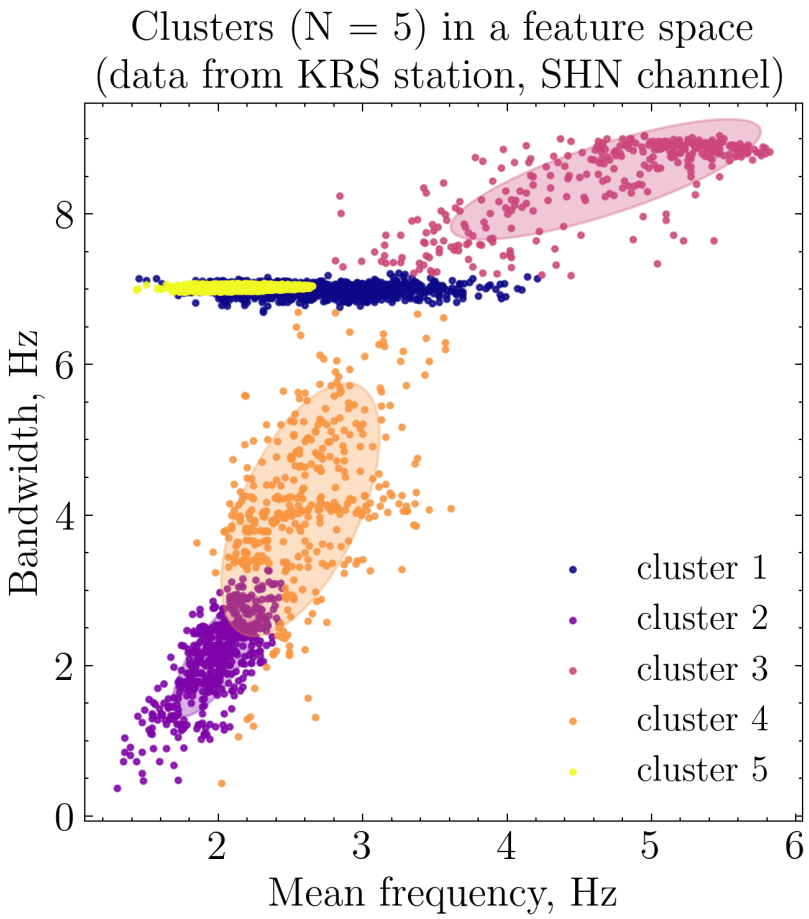

Figure 6.B.6: Results of GMM clustering ($N_{clust} = 5$, data obtained from station KRS, channel SHN)

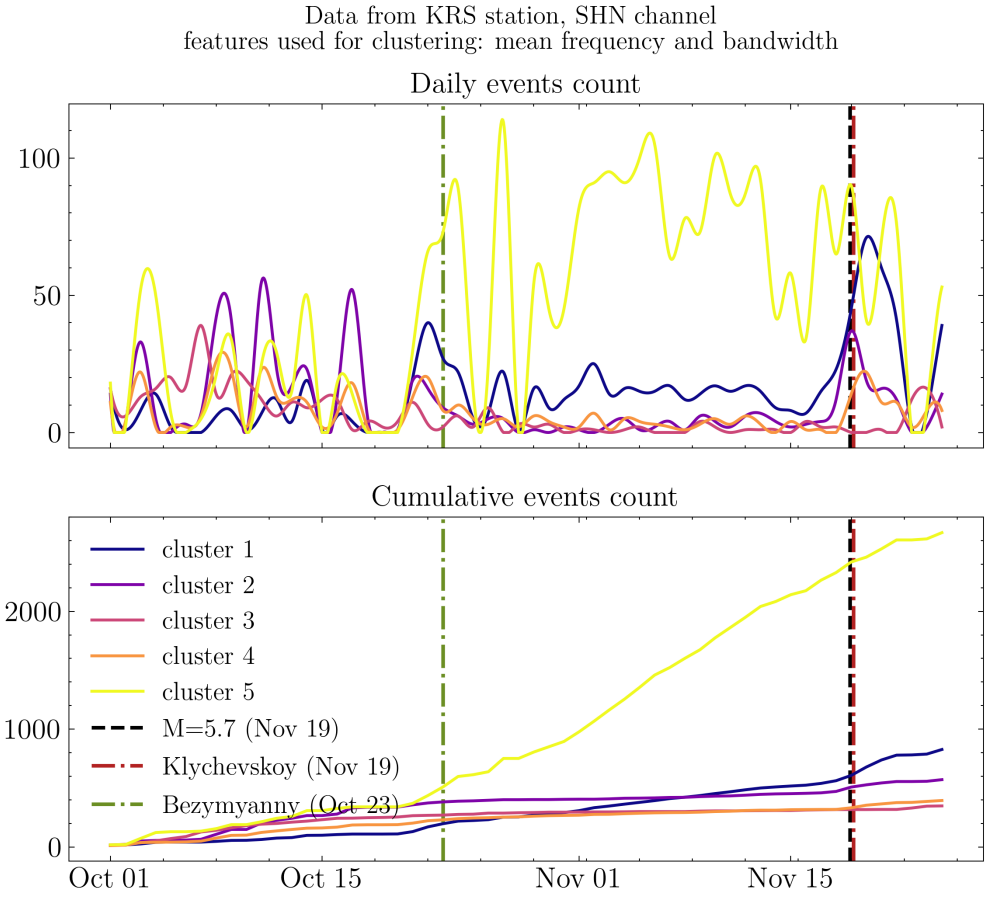

Figure 6.B.7: Seismic activity rate according to GMM clustering: daily (top) and cumulative (bottom) count. Lines colors correspond to colors of clusters in Figure 6.B.6. Vertical lines show the main events in the region such as volcanic eruptions and strong tectonic earthquakes (black dashed lines)

6.B.2 Station LGN

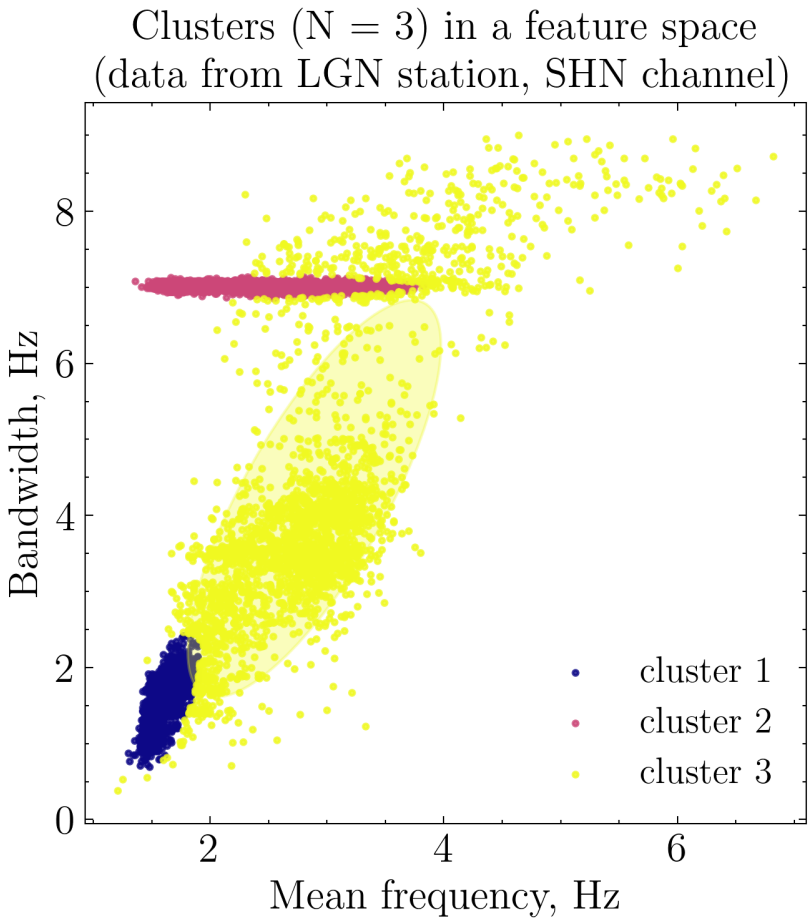

Figure 6.B.8: Results of GMM clustering ($N_{clust} = 3$, data obtained from station LGN, channel SHN)

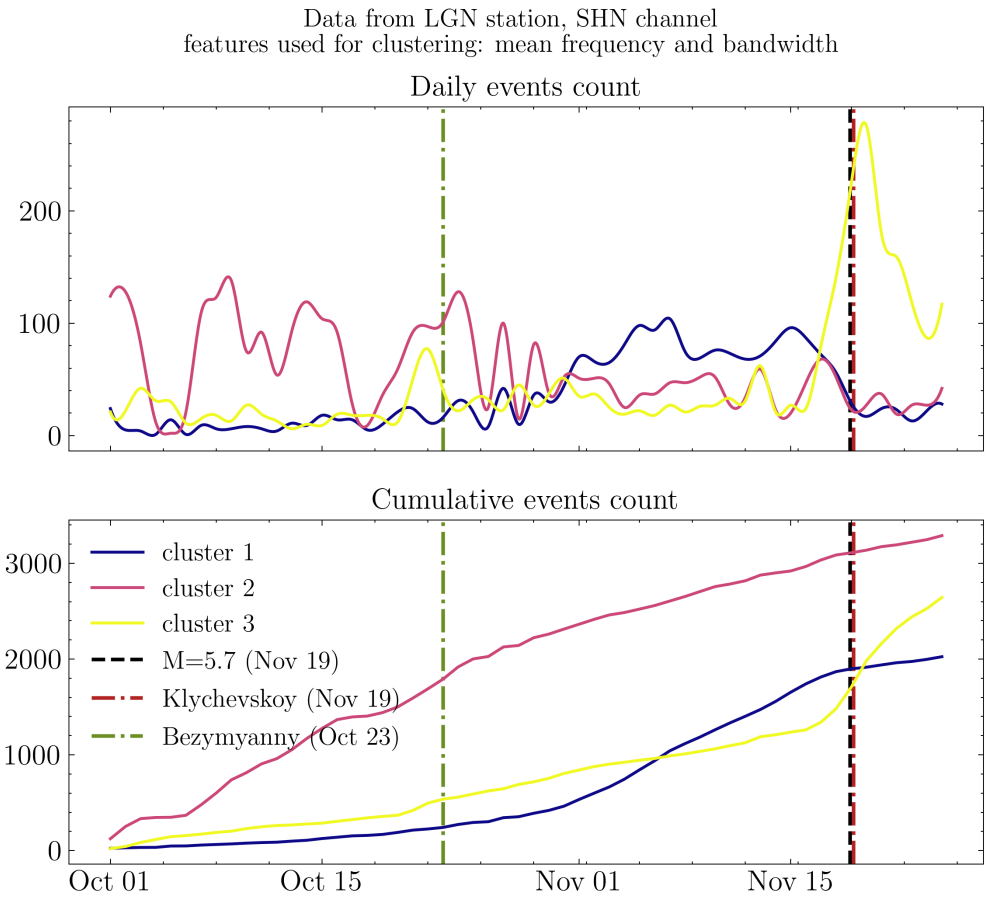

Figure 6.B.9: Seismic activity rate according to GMM clustering: daily (top) and cumulative (bottom) count. Lines colors correspond to colors of clusters in Figure 6.B.8. Vertical lines show the main events in the region such as volcanic eruptions and strong tectonic earthquakes (black dashed lines)

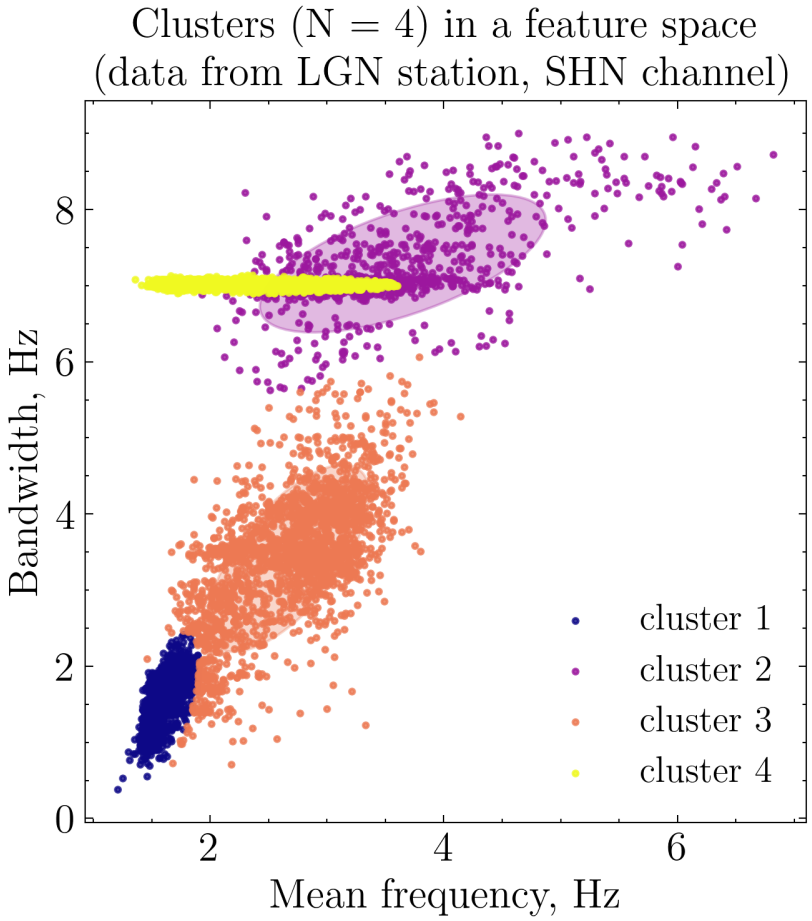

Figure 6.B.10: Results of GMM clustering (N_{clust} = 4, data obtained from station LGN, channel SHN)

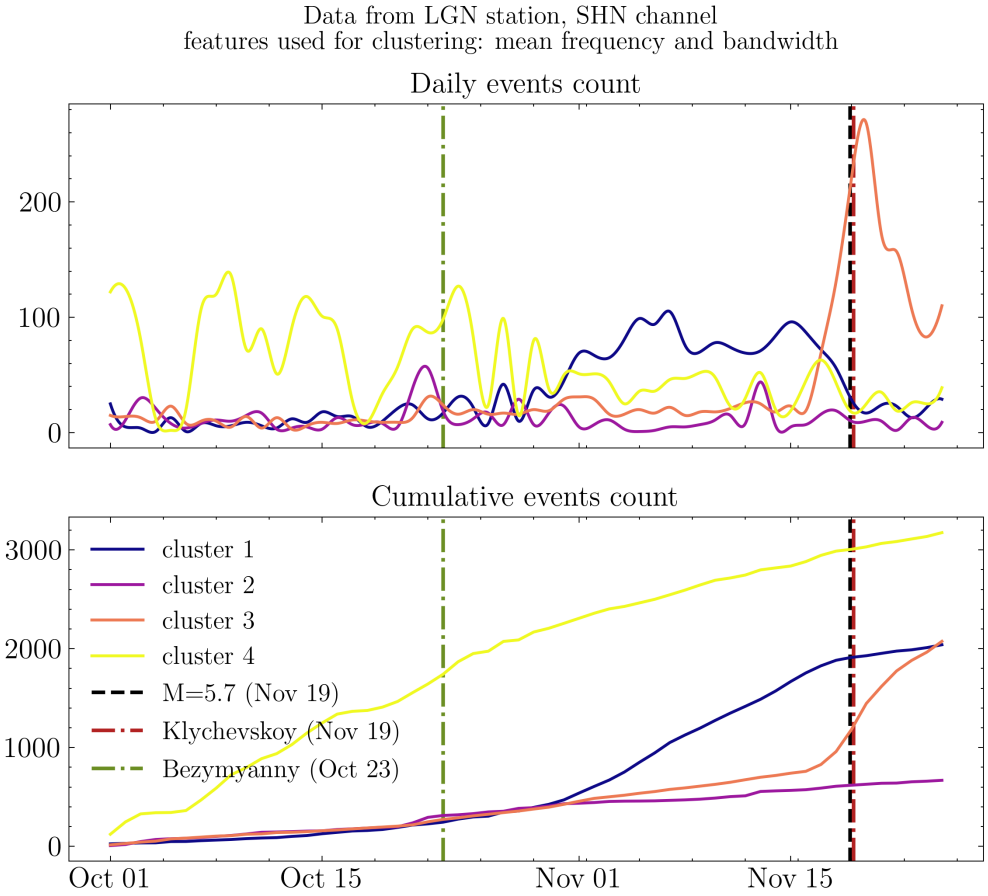

Figure 6.B.11: Seismic activity rate according to GMM clustering: daily (top) and cumulative (bottom) count. Lines colors correspond to colors of clusters in Figure 6.B.10. Vertical lines show the main events in the region such as volcanic eruptions and strong tectonic earthquakes (black dashed lines)

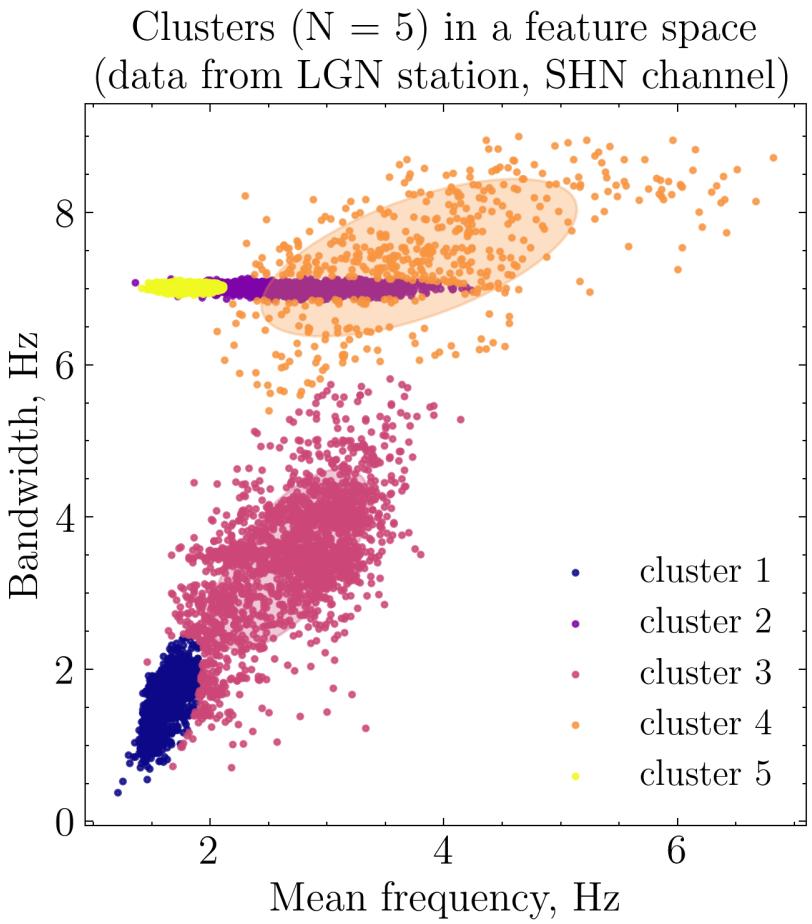

Figure 6.B.12: Results of GMM clustering (N_{clust} = 5, data obtained from station LGN, channel SHN)

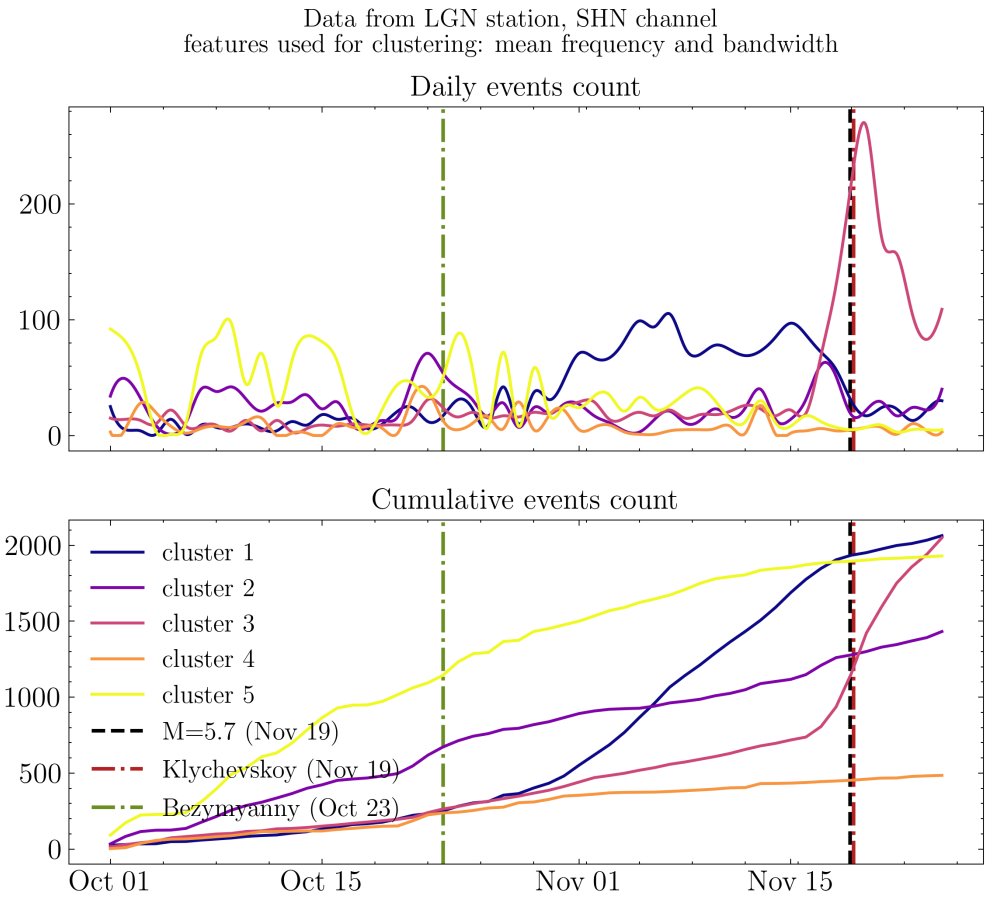

Figure 6.B.13: Seismic activity rate according to GMM clustering: daily (top) and cumulative (bottom) count. Lines colors correspond to colors of clusters in Figure 6.B.12. Vertical lines show the main events in the region such as volcanic eruptions and strong tectonic earthquakes (black dashed lines)

6.B.3 Station ZLN

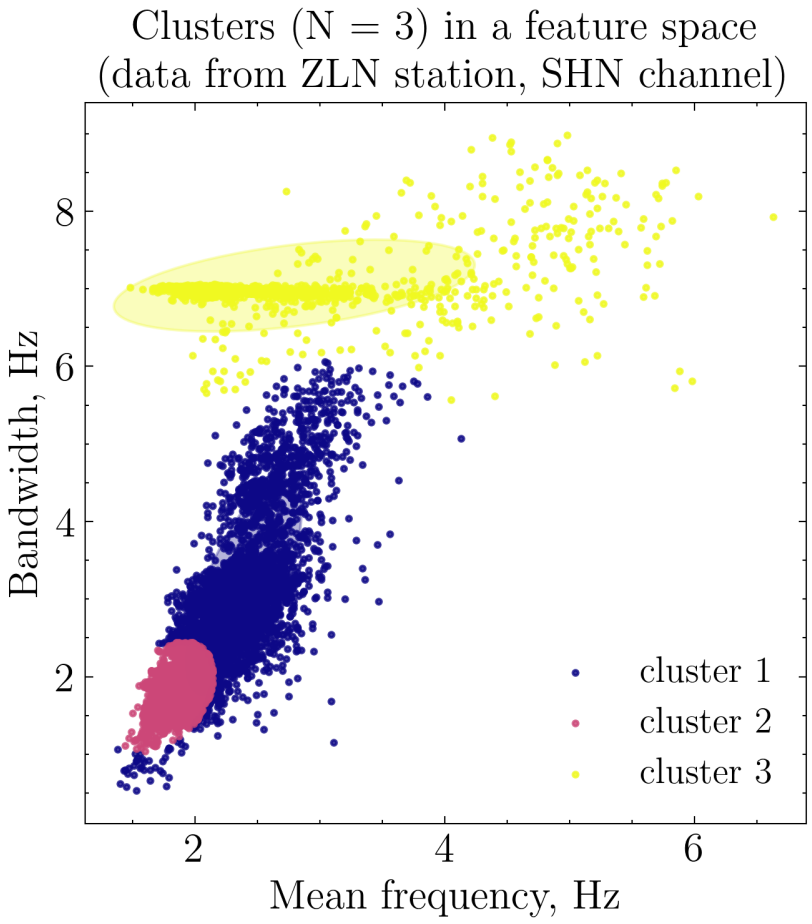

Figure 6.B.14: Results of GMM clustering ($N_{clust} = 3$, data obtained from station ZLN, channel SHN)

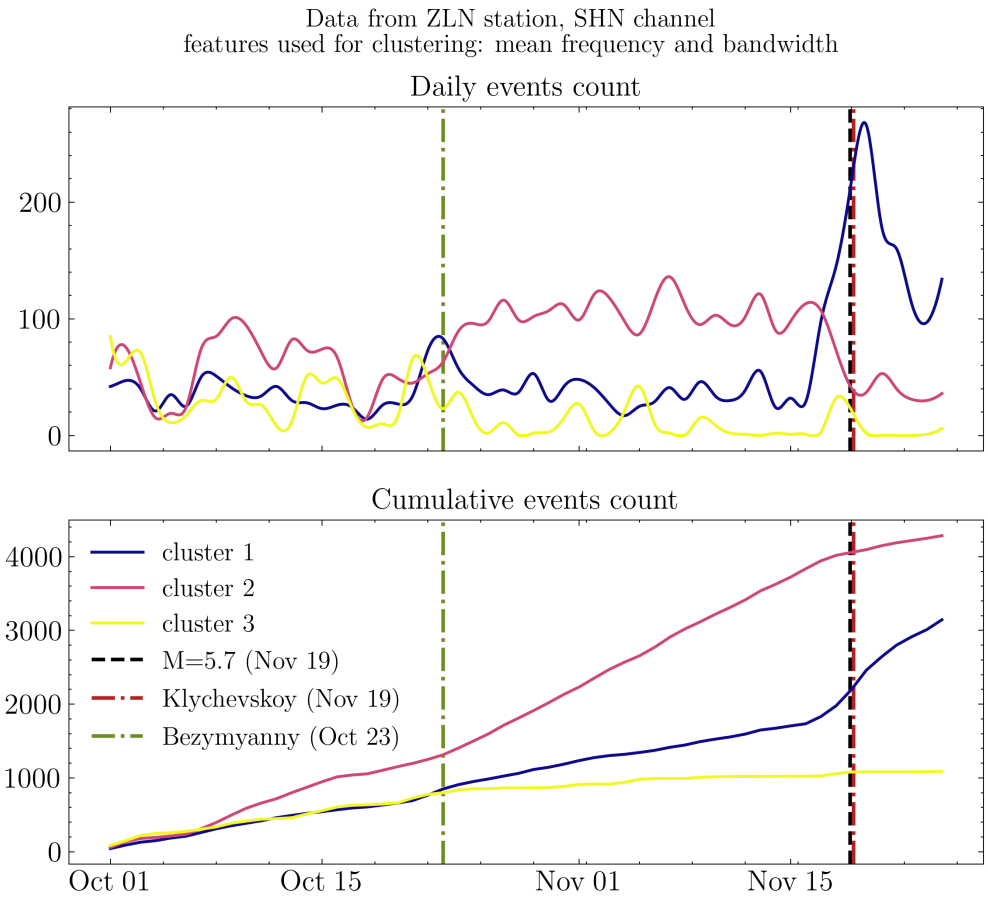

Figure 6.B.15: Seismic activity rate according to GMM clustering: daily (top) and cumulative (bottom) count. Lines colors correspond to colors of clusters in Figure 6.B.14. Vertical lines show the main events in the region such as volcanic eruptions and strong tectonic earthquakes (black dashed lines)

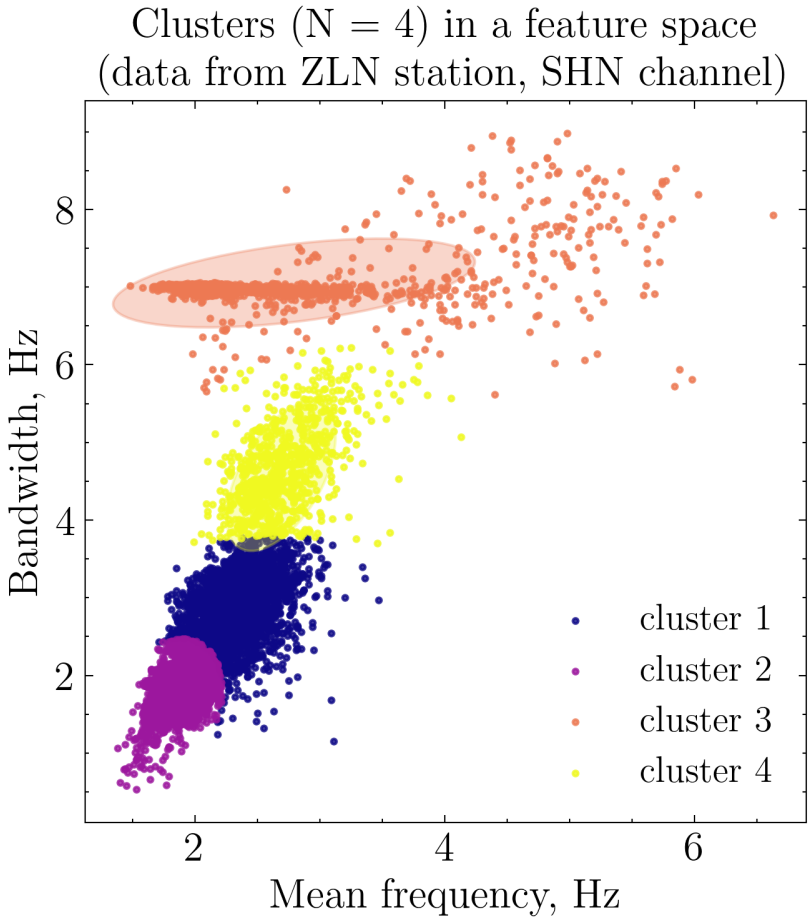

Figure 6.B.16: Results of GMM clustering ($N_{clust} = 4$, data obtained from station ZLN, channel SHN)

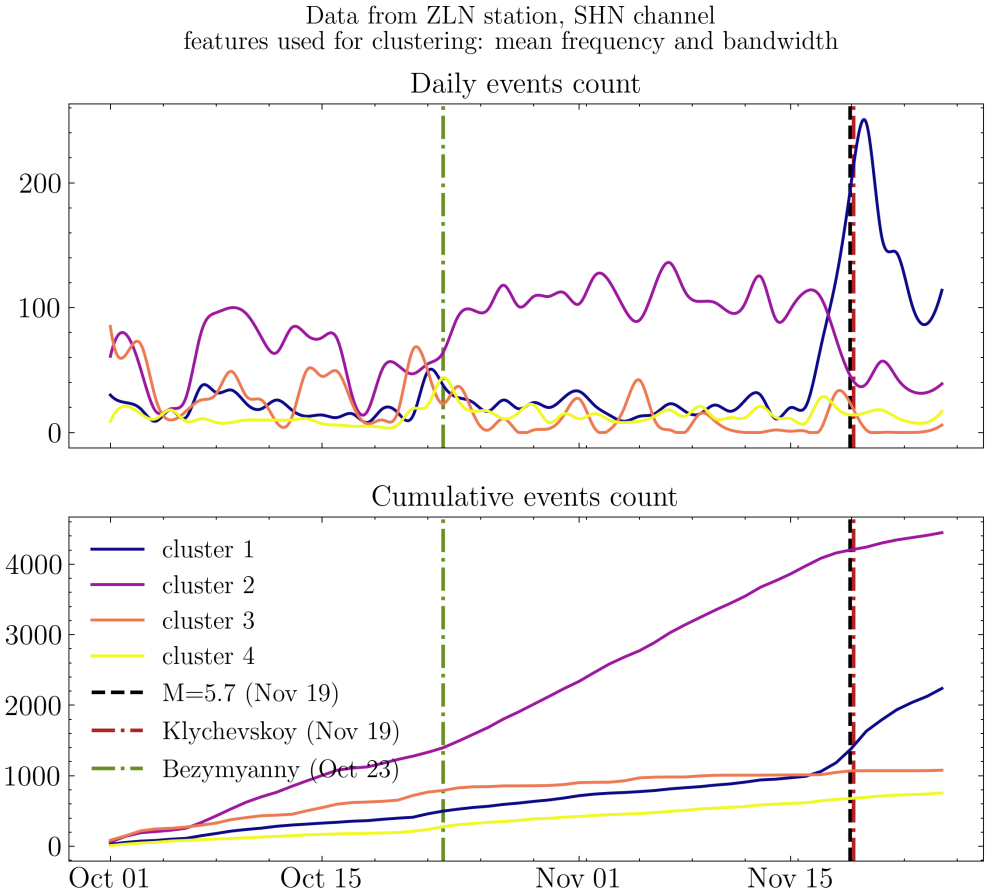

Figure 6.B.17: Seismic activity rate according to GMM clustering: daily (top) and cumulative (bottom) count. Lines colors correspond to colors of clusters in Figure 6.B.16. Vertical lines show the main events in the region such as volcanic eruptions and strong tectonic earthquakes (black dashed lines)

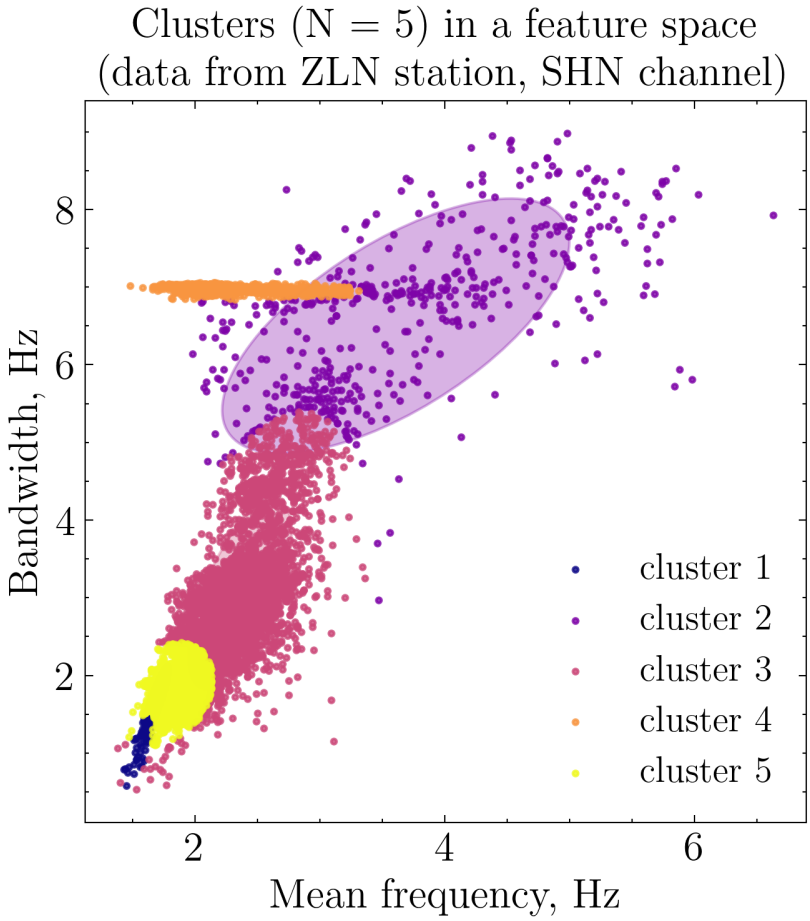

Figure 6.B.18: Results of GMM clustering (N_{clust} = 5, data obtained from station ZLN, channel SHN)

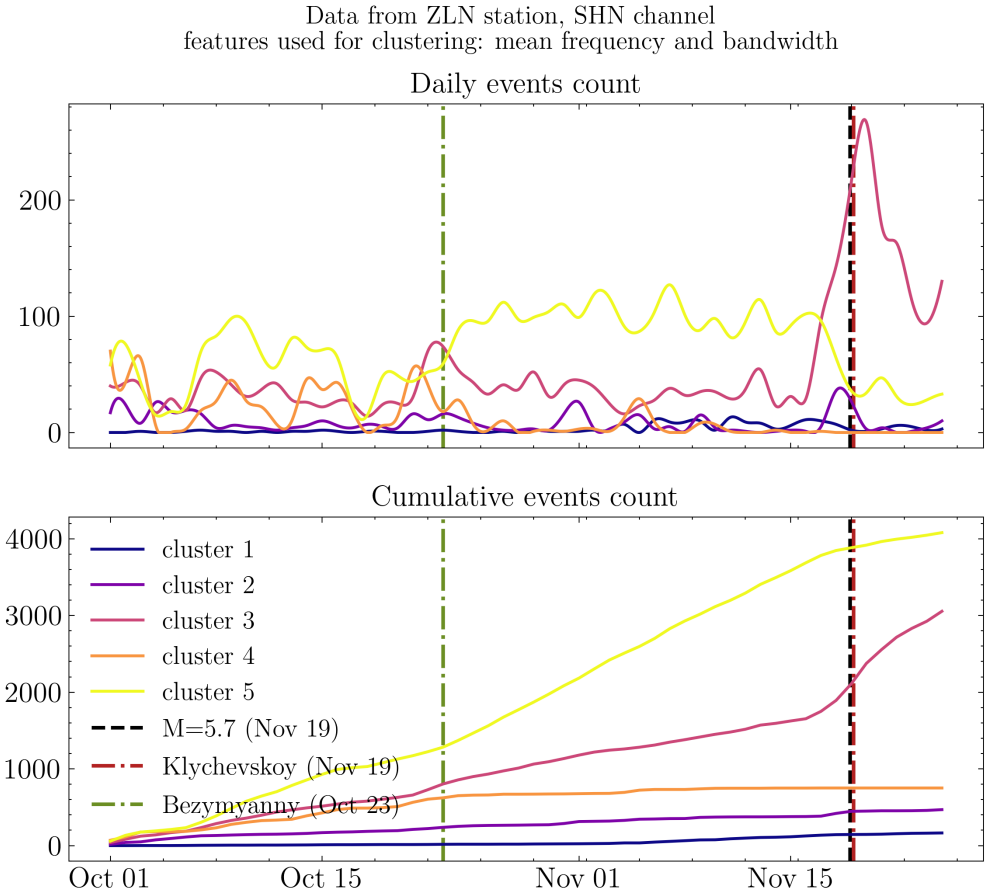

Figure 6.B.19: Seismic activity rate according to GMM clustering: daily (top) and cumulative (bottom) count. Lines colors correspond to colors of clusters in Figure 6.B.18. Vertical lines show the main events in the region such as volcanic eruptions and strong tectonic earthquakes (black dashed lines)

Chapter 7

Concluding remarks

The goals of this PhD thesis included developing new approaches to the analysis of seismovolcanic data with particular applications in the region of the Klyuchevskoy volcanic group in Kamchatka, Russia. This region was selected because of exceptional level of volcanic activity with four very active volcanoes: Klyuchevskoy, Bezymianny, Tolbachik, and Shiveluch, often erupting simultaneously. Combined with the numerous tectonic earthquakes mostly originated from the Kuril-Kamchatka and Aleutian subduction zones these volcanoes produce very intense and diverse seismicity making this region one of the best "natural laboratories" where the seismo-volcanic phenomena can be studied and new methods of analysis can be tested.

In the first part, we studied the Deep Long-Period (DLP) earthquakes that systematically occur near the crust-mantle boundary beneath the Klyuchevskoy volcano. This type of seismicity is particularly interesting because it might represent one of the earliest manifestation of the pre-eruptive volcanic unrest. The goal of this study was to use abundant Kamchatka data to try to better constrain the source mechanism and the physical origin of these DLP earthquakes that, despite a relatively strong interest, remain poorly understood.

We started with the statistical analysis of the DLP earthquakes, namely with the examination of their frequency-magnitude distribution. To perform this task, we explored almost two years (2011–2012) of continuous data recorded prior the Tolbachik volcano eruption. The dataset obtained from 10 seismic stations of the permanent network of KB GS RAS was processed using template matching algorithm (Chapter 2). Thus, applying this sensitive algorithm allowed us to obtain a detailed catalogs of DLP earthquakes. Additionally, we introduced an alternative way to estimate the moment magnitude M_{W} that is more adequate for volcanic earthquakes then the regional magnitude scale used by the the regional monitoring agency (KB GS RAS). Overall, we used 400 template events to detect respective multiplet families that were merged into one final catalog including 48915 events.

The combined catalog has been used to estimate the frequency-magnitude distribution of the Klyuchevskoy DLP earthquakes. An attempt to approximate it with a simple Gutenberg-Richter law resulted in a high b-value (more than 3). Such values that strongly deviate from the $b = 1$ characteristic for tectonic earthquakes have been

also reported in other volcanic regions. A further and more detailed analysis of the obtained frequency-magnitude plot has shown that it can be better approximated with a normal distribution that has some characteristic scale and in this sense is very different from the scale-free Gutenberg-Richter low. This observation points at a generating processes different from the one of tectonic earthquakes.

Calculated values of seismic moments have been used in modeling the process of magma degassing and consequent growth of bubbles as the source mechanism of DLP seismicity [Melnik et al., 2020]. Taken the characteristic parameters of magma content of the Klyuchevskoy volcano, we have shown that the elastic deformation of surrounding rocks cased by the bubble-induced magma expansion can be fast enough to generate seismic waves. Seismic moments estimations allowed us to show that amplitudes and frequency content of DLP earthquakes can be predicted with the proposed model of magma degassing.

After statistical analysis of magnitudes, we tried to constrain the kinematic source parameters of DLP earthquakes. In this part of the work we used data from a denser seismic networks than the one used at a previous step. Namely, we used seismic data recorded during a joint Russian-German-French temporary seismic experiment KISS in 2015–2016. We argued that this task can be addressed with a simple inversion because of the strong complexity of the data and the relatively high-frequency wave propagation. Therefore, we tried to test different plausible simplified scenarios of the DLP generation that can be described with source models including not more than 3 free parameters. As for methodology of source reconstruction, we prefer comparison of S-to-P waves amplitudes ratios to the full waveform or polarity inversions because of the phase uncertainty in the bandpassed data.

The obtained results showed that the shear-slip mechanism attributed to "regular" tectonic (and volcano-tectonic) seismicity does not explain the amplitudes measured from the KVG DLP waveform. This strongly suggested that processes other then shear faulting are responsible for their generation. In particular, mechanisms with significant volumetric and/or single-force components better explain the observed data. The former can be associated with the pressure perturbation withing magmatic reservoirs or conduits and the latter with the sudden acceleration of the magma movement. At the same time, the amplitude ratio analysis has been found not accurate enough to unambiguously choose the preferential mechanism among different tested scenarios. Such strong uncertainties associated with the inversion of the amplitude data for focal mechanisms have been already reported in previous studies. Therefore, the available seismological observations should be combined with additional arguments and/or data to constrain the origin of the DLP earthquakes. We concluded that the ensemble of our

observations is compatible with the configuration when the magma is stored in nearly horizontal sills near the crust-mantle boundary and penetrates into the crust through conduits dipping south-southwest, in agreement with previously reported connection of the deep magmatic reservoir with the Bezymianny and Tolbachik volcanoes [Coppola et al., 2021, Journeau et al., 2022].

Second part of this thesis is dealing with methods of automatic signal analysis, and namely with the Machine Learning (ML) based classification and clustering of earthquakes recorded in the vicinity of KVG. As discusses in Chapter 1, one of the main tasks of seismological volcano monitoring is the timely detection of signals and extracting ones of interest for further investigation and interpretation. The KVG region with its very intense and diverse seismicity is well suited for testing new methods.

We start the ML exploration of seismic data with the problem of separating two general classes of seismicity: tectonic and volcanic earthquakes. This interesting case is discussed in Chapter 5 where we show the results of events classification using the data recorded by KB GS RAS permanent stations during the seismic crisis in 2018–2019. A strong tectonic earthquake with M=7.3 occurred at the coast of Kamchatka which was followed by intense aftershock swarm. These tectonic sequence occurred on the background of the Shiveluch eruption which was accompanied by seismo-volcanic activity. Eventually, all phenomena reflected themselves in complex seismograms. In particular, we consider data from only one station located equidistant from major tectonic earthquakes and Shiveluch which confidently recorded both types of seismicity.

Despite apparent simplicity of the problem, even the choice of data representation turned out to be an ambiguous task. We considered several possible representations of signals but all of them contained information on spectral content. However, the higher dimension of the feature space does not provide more reliable results. Here we likely face the common problem in ML practice called the curse of dimensionality. Furthermore, different tested ML methods showed different performance using one or the other feature set. Commonly speaking, significant performance could be achieved in classification problem using lower dimension dataset, while clustering methods required more features for higher satisfactory results. Nevertheless, this was difficult to unequivocally point out a pair of features and a ML method that provides the most reliable result.

In the next chapter, we investigated a possibility to identify several classes of volcanic seismicity appearing when several volcanoes were simultaneously active. We chosen two months of continuous data, October-November 2022 when KVG reactivated and Shiveluch volcano have been active for several months. At this stage we used only unsupervised ML approach because the data were difficult to labeled considering that

we did not have good *a priori* ideas about different recorded classes of seismicity.

After applying an unsupervised clustering with GMM, we focus on a dominating cluster of seismicity standing out in several 2D projections of the feature space. Furthermore, this cluster has been retrieved in the data recorded at different stations. The selected seismicity cluster has the very high density of objects what can suggest a repetitive and stationary source that generates these signals. To verify this assumption we calculate the pair-wise correlations of signals detected during the entire study period. The obtained results showed the multiplet families of signals occurring almost every day of the studied period. Choosing the master events from every family we perform their cross-correlation and found that they belong to one big cluster of seismicity.

Finally, to compare the results of the GMM clustering and waveforms cross-correlation we calculate the number of events in the cluster detected with one or the other method. The comparison shows the strong correlation between the results obtained independently with two different techniques. Thus, we can speak about a persistent seismic source active during the re-activation of KVG.

At the final step we choose five earthquakes with best signal-to-noise ratios to estimate the source location of the identified cluster. To calculate the hypocenters the seismic data was processed manually to pick the P- and S-waves arrivals. According to our estimations, the hypocenters lay next to Shiveluch at middle depths of 12 km. The one-side coverage of stations and the very basic 1D velocity model bring high errors in the calculations. Based on the obtained results it is difficult to say if we see a new process occurring beneath Shiveluch or it exists for a long time. But an important observation can be made on the activity reinforcement almost immediately after the Bezymianny volcano eruption and lasting up to the reactivation of Klyuchevskoy. Thus, the main result of Chapter 6 is the demonstration of the ability to define the structure and anomalies in seismic data with the very basic approaches. Further developing the tools for volcano seismic monitoring will allow us to follow the state of the magmatic systems, understand the ongoing processes, and, ideally, predict the eruptions.

Appendix

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-17759-4 OPEN

Deep long period volcanic earthquakes generated by degassing of volatile-rich basaltic magmas

Oleg Melnik[1], Vladimir Lyakhovsky[2], Nikolai M. Shapiro[3,4]✉, Natalia Galina[3,4] & Olga Bergal-Kuvikas[5]

Deep long-period (DLP) earthquakes observed beneath active volcanoes are sometimes considered as precursors to eruptions. Their origin remains, however, unclear. Here, we present a possible DLP generating mechanism related to the rapid growth of gas bubbles in response to the slow decompression of over-saturated magma. For certain values of the gas and bubble content, the elastic deformation of surrounding rocks forced by the expanding bubbly magma can be fast enough to generate seismic waves. We show that amplitudes and frequencies of DLP earthquakes observed beneath the Klyuchevskoy volcano (Kamchatka, Russia) can be predicted by our model when considering pressure changes of ~10^7 Pa in a volume of ~10^3–10^4 m^3 and realistic magma compositions. Our results show importance of the deep degassing in the generation of volcanic seismicity and suggest that the DLP swarms beneath active volcanoes might be related to the pulses of volatile-rich basaltic magmas rising from the mantle.

[1] Institute of Mechanics, Moscow State University, 1 Michurinskiy prospekt, 119192 Moscow, Russia. [2] Geological Survey of Israel, 32 Yesha'ayahu Leibowitz st, 9692100 Jerusalem, Israel. [3] Institut de Sciences de la Terre, Université Grenoble Alpes, CNRS (UMR5275), CS 40700, 38058 Grenoble Cedex 9, France. [4] Schmidt Institute of Physics of the Earth, Russian Academy of Sciences, Bolshaya Gruzinskaya str., 10-1, 123242 Moscow, Russia. [5] Institute of Volcanology and Seismology, FEB RAS, 9 Piip Boulevard, 683006 Petropavlovsk-Kamchatsky, Russia. ✉email: nikolai.shapiro@univ-grenoble-alpes.fr

Deep Long Period (DLP) earthquakes occurring in middle to lower crust and uppermost mantle beneath volcanoes[1–9] remain enigmatic and in some cases, are believed to have connection with magmatic activity. Similar to volcanic long-period (LP) seismicity in general[10], the DLP earthquake has been considered to be generated by rapid pressure variations within magmatic plumbing systems. Alternatively, the effect of thermal stresses within cooling magma bodies has been considered[11]. The cooling magma stalled beneath the crust can also generate DLP earthquakes by so called "second boiling" or repeated pressurization of volatiles exsolved through crystallization, as has been recently suggested for dormant hot-spot Mauna Kea volcano in Hawaii[9]. However, such cooling-related mechanisms are unlikely for DLP events occurring beneath active volcanoes in association with eruptions. Different possible origins of pressure variations resulting in LP seismicity have been considered[12] including the unsteady magma motion, breaking of mechanical "barriers"[9,13], rapid degassing, etc. In any case, a reasonable model must provide a physical mechanism generating pressure variation *dP(t)* consistent with observed seismic waves. This implies that the time scale of these variations must by rather short, i.e., comparable with typical frequencies/periods of seismic waves (e.g., ~1 s). Second condition is that the fluid pressure variation should be strong enough and well coupled with the elastic media. This coupling may imply resonances of fluid-filled cracks or cavities[10] that under certain conditions can result in nearly monochromatic and very long duration signals. At the same time, such "strongly resonant" features are not observed for DLP signals that are characterized by rather short durations.

Here we propose that rapid changes of magmatic pressure near the crust-mantle boundary can be caused by nucleation and growth of gas bubbles in response to the slow decompression of over-saturated magma[14]. A volume of magma saturated with H_2O–CO_2 volatiles is subjected to slow de-pressurization because of its slow upwelling. This magma first reaches the saturation level and then achieves the critical supersaturation after which the gas bubbles nucleate (Fig. 1a) and grow very fast (Fig. 1b). Fast expansion of the bubbly magma deforms the surrounding rocks which respond elastically on the time scale associated with the bubble growth and magma pressure variations. As a result of this elastic rock deformation, seismic waves are radiated (Additional information provided in Methods) and can be recorded by seismographs installed in vicinity of volcanoes.

The pressure variation in the bubbly magma is simulated using the model that accounts for multiple dissolved volatiles (H_2O–CO_2) and diffusive gas transfer from magma into the growing bubbles. It is based on the full solution of advection-diffusion equation instead of quasi-static approach that was used before (Additional information provided in Methods)[15]. The bubble growth model is adopted to the case of bubble nucleation in basaltic magma[16].

We compare the results of our modeling with DLP earthquakes observed beneath the Klyuchevskoy volcanic group (KVG) in Kamchatka, Russia. This volcanic group is one of the largest and most active clusters of subduction-zone volcanoes in the World[17]. KVG eruptions and their precursory periods are accompanied by sustained seismovolcanic activity including volcanic earthquakes[7,18,19] and tremors[20]. We particularly focus on a persistent cluster of DLP earthquakes that occur in a small volume located at ~30 km depth beneath the Klyuchevskoy volcano[7,19,21]. The moment magnitudes (Additional information provided in Methods) of these DLP events range between 1.1 and 2.5 with maximum of their distribution at 1.4 (Supplementary Fig. 1).

Initial data on volatiles in Klyuchevskoy[22] suggested that primary magmas content 2.2–2.9 wt.% of water. Later, a detailed study of melt inclusions in olivines[23] has shown that parental magma has ~3.5 wt% H_2O and 0.35–0.9 wt% CO_2. Large increase of water content for some melt inclusions (up to 7 wt% H_2O) was explained by de-compressional crystallization, accumulation volatiles in the melt phase and consequent slow degassing[23,24]. However, recent experimental data shows that the volatile content of Klyuchevskoy magma is much larger than the one previously directly measured in melt inclusions due to coupled SiO_2-H_2O loss[25], suggesting that primary magma may contain more than 4 wt% of H_2O. Single H_2O volatile phase will result in a small saturation depth, but the addition of ~0.6 wt% of CO_2 increases volatile solubility dramatically so that magma becomes supersaturated at pressures of 800 MPa (~30 km depth) that alternatively requires ~10 wt% of pure H_2O.

We perform a parametric study to investigate the influence of volatiles content on the dynamics of bubble nucleation and growth. Our results show that the time scale of the bubble growth is mainly controlled by the gas and bubble content in the magma and under certain conditions can be sufficiently fast to generate seismic waves. In particular, we show that amplitudes and frequency content of DLP earthquakes observed beneath the Klyuchevskoy group of volcanoes can be predicted by our model when considering pressure changes of a few tens of MPa in a volume of ~10^3–10^4 m^3 and magmas containing ~4 wt% of H_20 and ~0.6 wt% of CO_2. Our results provide evidence for the role of the deep degassing in the generation of long-period volcanic seismicity and suggest that the DLP swarms observed beneath active volcanoes

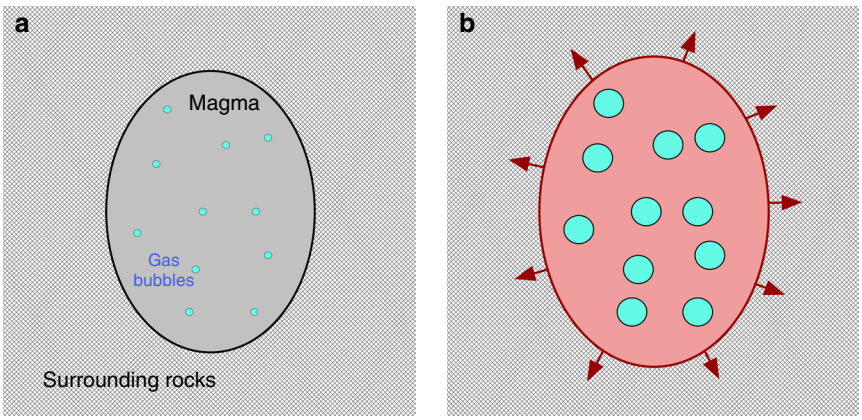

Fig. 1 Conceptual model of fluid-related source of long-period earthquakes. a Bubble nucleation in a volume of magma saturated with H_2O–CO_2 volatiles. **b** Bubble and pressure growth deforming the surrounding rocks.

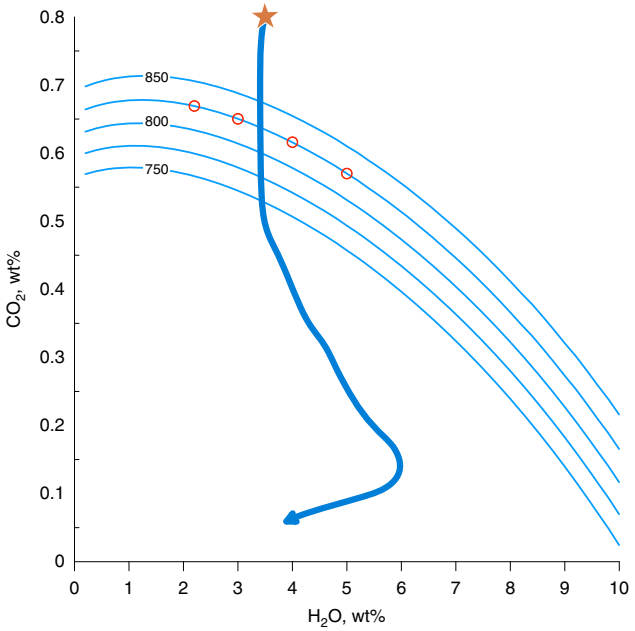

Fig. 2 Gas saturation isobars as function of CO_2–H_2O content. Thin blue lines show saturation isobars for different pressures (indicated values in MPa). Thick solid line indicates the decompression path of the Klyuchevskoy magmas[24] from initial state at 1 GPa shown with a star. Red circles show compositions along the 828 MPa isobar with 2, 3, 4, and 5 wt% of H_2O tested with numerical modeling (results shown in Fig. 3).

might be related to the pulses of fresh CO_2–H_2O rich basaltic magmas rising from the mantle.

Results

Volatiles content and the depth of degassing unset. Figure 2 shows how the solubility of the CO_2–H_2O mixture varies with pressure. Water and carbon dioxide concentrations were parametrized by polynomial functions of pressure and CO_2 content in the bubble at a fixed temperature of 1230 °C estimated from reversed crystallization of Klyuchevskoy melts from atmospheric conditions to 800 MPa pressure (Additional information provided in Methods) and extrapolated to 1000 MPa (Supplementary Table 1). The saturation point is reached at depths of the crust-mantle transition (~30 km; $P \approx 825$ MPa) for magmas with the volatile content typical for the Klyuchevskoy volcano, implying that degassing may start at such large depths.

Parameters controlling the time scale of bubble growth. Figure 3a, b shows typical evolution of bubble size, gas and melt pressure in the basaltic magma with density[26] of 2800 kg m^{-3}, viscosity of 10 Pa s and containing 10^{13} bubbles m^{-3} and for four different concentrations of H_2O in initial magma. The CO_2 contents (red circles in Fig. 2) were computed for initial pressure of 828 MPa that corresponds to the lithostatic pressure at depth of ~30 km for the average crustal density[27] of 2830 kg m^{-3}. Based on the experimental observations we adopt that the critical supersaturation for the bubble nucleation corresponds to the over-pressure of $\Delta P = 40$ MPa[14,28]. This means that after nucleation the pressure in the gas bubble will be equal to its saturation value and the melt pressure is lower by ΔP. Due to rapid bubble expansion gas pressure decreases extremely fast while melt pressure starts to increase as the volume of the magma increases. Initially gas pressure drop due to bubble expansion dominates pressure increase due to volatile influx into the growing bubble. After reaching minimum value P_g starts to increase, concentration gradients in the melt become smoother and volatile flux decreases. At later stages of the growth the difference between melt and gas pressures becomes small and bubble growth is controlled by the diffusion of volatiles.

The water diffusion coefficient is 1–2 orders of magnitude larger than the diffusion coefficient for CO_2. Thus, larger water content of magma for a fixed pressure require smaller amount of dissolved CO_2 and bubble during growth will suck more H_2O. Adding more water into initial magma results in H_2O enriched gas and more vigorous bubble and pressure grows (Fig. 3). The effect of water content is enhanced even stronger in predicted seismograms (Fig. 3d) with H_2O depleted magmas resulting in very weak signals. We compare amplitudes of synthetic seismograms computed for a magma source volume of 30,000 m^3 (linear dimension of a few tens of meters) with a real seismogram (Fig. 3e) recorded during DLP earthquake with a magnitude $M_W \approx 2$ at station LGN located nearly above the source region (Supplementary Fig. 1). Amplitudes and the frequency content (Fig. 3f) are reasonably well predicted with a model based on 4 wt% water in basaltic magma typical for the Klyuchevskoy volcanic group[25].

We then perform a sensitivity study of several other parameters on the pressure evolution in the growing bubbles and resulting melt (Supplementary Fig. 2). Critical supersaturation[28] that is required for bubble nucleation does not change melt-pressure recovery time significantly but will affect the amplitude of the source signal. We consider the melt viscosity range 10–10^5 Pa s[29]. If viscosity is smaller than some threshold its influence on resulting pressure is negligible. Only larger melt viscosities typical for more silica reach melts (10^5 Pa s) introduce some delay in pressure recovery. We assume instantaneous bubble nucleation in the whole batch of magma (Additional information provided in Methods). Thus, the size of the cell from which the bubble is growing is controlled by bubble number density (*BND*). We consider the *BND* range[30] between 10^{11} and 10^{15} m^{-3}. Increase in *BND* results in smaller cell sizes as $S_0 \sim BND^{-1/3}$. Melt pressure grows faster for smaller S_0.

Discussion

While the presented comparison of the observed and model-predicted seismograms (Fig. 3d–f) is based on significant simplifications of the source (ignoring realistic geometry and possible resonant behavior[10]) and the propagation effects (ignoring attenuation and wave scattering[31]), it shows that the amplitudes and the spectral content of the DLP signals observed at the Klyuchevskoy volcanic group can be explained to the order of magnitude by the bubble nucleation and growth in basaltic magmas according to the performed numerical simulation (Fig. 3d–f). Results of the presented modeling show that in the CO_2–H_2O rich basaltic magmas the degassing starts at large depths and is vigorous enough to produce strong and rapid pressure variations that can generate seismic radiation with amplitudes and frequency content comparable with those observed by seismographs during DLP earthquakes. Our results suggest that the DLP swarms observed beneath active volcanoes might be related to the intensification of the deep degassing caused by pulses of fresh CO_2–H_2O rich basaltic magmas rising from the mantle. This mechanism supports that the DLP earthquakes are early seismic manifestations of activation of deep parts of the Klyuchevskoy volcano plumbing systems. Similar behavior might be expected in other open and very active volcanic systems (with adjusting the model parameters based on their magma compositions and volatile contents). At

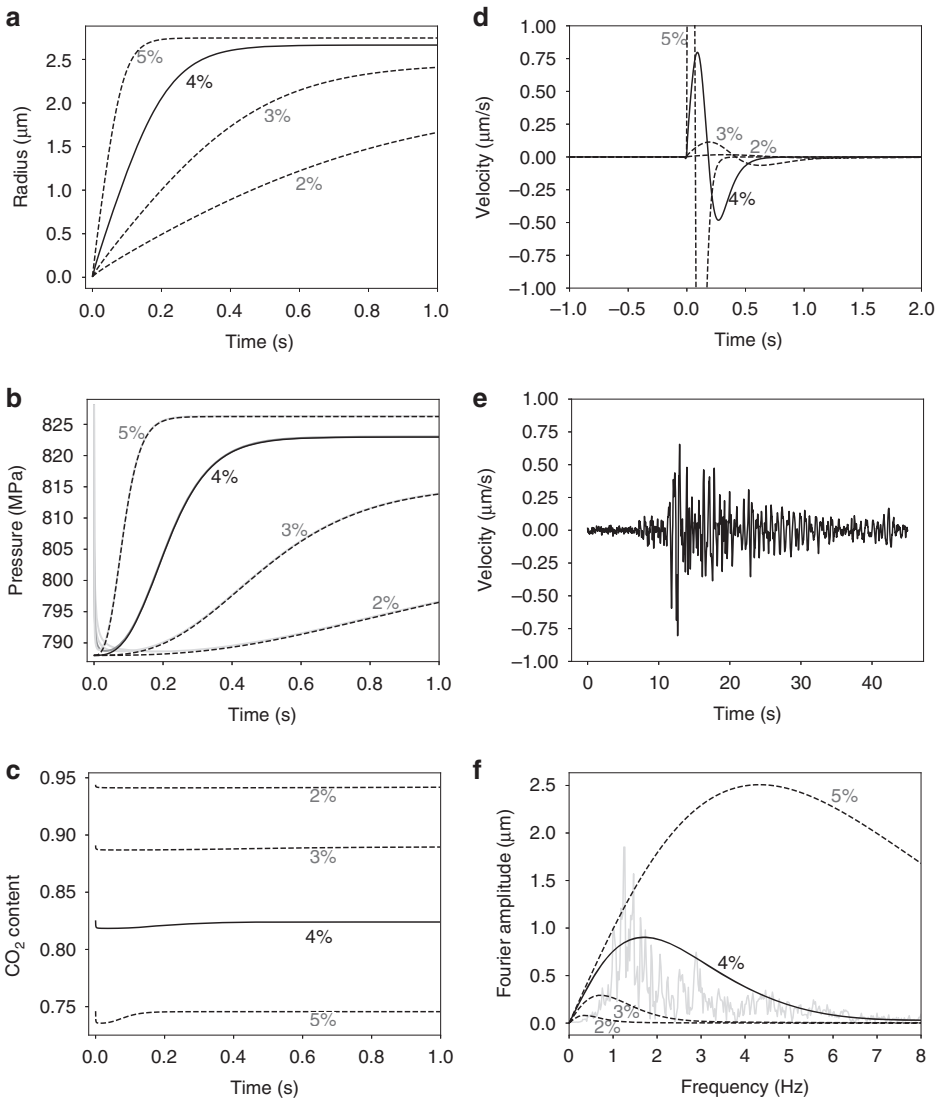

Fig. 3 Modeled dynamics of the bubble grows and magma pressure change. Results are shown for the bubble number density of 10^{13} m^{-3}, four different water contents indicated with wt% values in respective plots, and for CO_2 content computed for 828 MPa (red circles in Fig. 2). **a** Evolution of the bubble radius. **b** Evolution of magma pressure P_m (P_g values are shown with gray lines). **c** Evolution of the CO_2 content in bubbles. **d** Ground velocities estimated for a source located at a 30 km distance from the receiver (Additional information provided in Methods). **e** Example of real seismogram (east-west component at station LGN, Supplementary Fig. 1). **f** Fourier amplitudes computed from synthetic and real (gray line) signals.

the same time, magma-cooling-related DLP mechanisms can dominate beneath nearly closed or dormant volcanoes.

One of the key features of our model is that the depth of occurrence of DLP earthquakes is related to the CO_2 content in magmas. This is especially interesting considering that global volcanic CO_2 fluxes in modern Earth remain poorly known[32–38] and are often estimated indirectly based on CO_2/SO_2 or other ratio proxies, with direct CO_2 observations at volcanoes being technically challenging. Our results suggest that studies of the DLP volcanic seismicity provide additional constraints on the magmatic CO_2 content in the deep roots of volcanoes.

Methods

Mathematical model of gas bubble growth. We consider growth of an individual bubble in the center of a spherical cell of melt that expands with the bubble and supports it with volatiles. The spherically symmetric model includes equations of mass conservation of the melt in a cell Eq. (1), diffusion equations for volatiles (H_2O–CO_2) Eq. (2), Rayleigh–Lamb equation for bubble growth with negligibly small inertia terms and the equation for the melt pressure evolution due to expansion of the surrounding elastic host rock Eq. (3), mass balances for volatiles in the bubble Eq. (4), and equations that describe physical properties of the components Eq. (5):

$$\frac{\partial}{\partial r}\left(r^2\nu_r\right)=0;\nu_r|_{r=R}=\frac{dR}{dt} \tag{1}$$

$$\frac{\partial c_s}{\partial t}+\nu_r\frac{\partial c_s}{\partial r}=\frac{1}{r^2}\frac{\partial}{\partial r}\left(D_s r^2\frac{\partial c_s}{\partial r}\right);s=CO_2(c),H_2O(w) \tag{2}$$

$$P_g-P_m=\frac{2\sigma}{R}+4\mu\frac{dR}{dt}\left(\frac{1}{R}-\frac{R^2}{S^3}\right);P_m=P_m^0+\frac{4}{3}G\left(\frac{S^3-S_0^3}{S_0^3}\right) \tag{3}$$

$$\begin{gathered}\frac{4\pi}{3}\frac{d}{dt}\left(R^3\rho_g x_{CO2}^b\right)=4\pi R^2J_c;\frac{4\pi}{3}\frac{d}{dt}\left(R^3\rho_g\left(1-x_{CO2}^b\right)\right)=4\pi R^2J_w;\\ J_c=-D_c\rho_m\left(\frac{\partial c_c}{\partial r}\right)_{r=R};J_w=-D_w\rho_m\left(\frac{\partial c_w}{\partial r}\right)_{r=R}.\end{gathered} \tag{4}$$

$$\begin{gathered}\rho_g=\left(\frac{x_{CO2}^b}{\rho_{CO2}\left(P_g,T\right)}+\frac{1-x_{CO2}^b}{\rho_{H2O}\left(P_g,T\right)}\right)^{-1};\\ \rho_{CO2}=(0.371+0.13\times10^{-3}T)\cdot P_g+1194.65-0.4665T;\\ \rho_{H2O}=(0.22+0.13\times10^{-3}T)\cdot P_g+892.2-0.357T;\\ D_w=c_w\cdot\exp\left(-8.56-\frac{19110}{T}\right);\\ D_c=\exp\left(-13.99-\frac{(17367+1.945P_g)}{T}+\frac{c_w\cdot(855.2+0.271P_g)}{T}\right)\end{gathered} \tag{5}$$

Here t is time, r is the radial coordinate, R is the radius of the bubble, v_r is the radial velocity, c_c and c_w are the mass concentrations of CO_2 and H_2O in the melt, D_c and D_w are the volatile diffusion coefficients[39], P_g is the pressure of the gas inside the bubble, P_m is the melt pressure, σ is the surface tension, μ is the magma viscosity, S is the radius of the cell, G is the shear modulus of the host rock, ρ_g is the density of the gas in the bubble that depends on the pressure, temperature T and bubble volatile composition x^b_{CO2}. The densities of pure CO_2 (ρ_{C02})and H_2O (ρ_{H2O}) are approximated at a limited P-T range using tables produced by NIST Chemistry WebBook (https://webbook.nist.gov/chemistry/).

Equation (2) is subjected to two boundary conditions: concentration gradients are equal to zero at the outer surfaces of the cell mimicking symmetry of the system. At $r = R(t)$ volatiles in magma are in chemical equilibrium with the bubble. Thus, $c_s = c_s^{eq}\left(p, T, x^b_{CO2}\right)$. We use D-compress software[40] in order to calculate equilibrium concentrations.

The nucleation time of bubbles t_n from a supersaturated melt is related to the bubble number density BND via the nucleation rate $I(\mathrm{m}^{-3}\ \mathrm{s}^{-1})$: $t_n = BND/I$ According to classical nucleation theory[41], I increases extremely fast with oversaturation pressure ΔP: $I \sim exp(-1/\Delta P^2)$. It depends on the temperature and a number of melt properties including surface tension, volume and concentration of water molecules in the melt, as well as distance between them, diffusion coefficient of volatiles at the bubble-melt interface, probability that a nucleus at the top of the barrier will go on to form the new phase, rather than dissolve (Zeldovich factor), and others. With a huge uncertainty of these parameters and difficulties in their experimental constrain, the estimated nucleation rate values vary by orders of magnitudes. For a basaltic melt with an overpressure about 40 MPa a value of $I \sim 10^{26}\ \mathrm{m}^{-3}\mathrm{s}^{-1}$ has been suggested[42]. With this I-value, nucleation time for $BND = 10^{13}\ \mathrm{m}^{-3}$ (values preferred in our study) is $\sim 10^{-13}$ s, which is many orders of magnitude below the typical time scale of the simulated bubble growth and of the observed periods of seismic waves (~1 s). These estimations were obtained assuming that magma degassing is dominated by homogeneous nucleation. In the presence of crystals, their interfaces serve as a preferable location for the heterogeneous nucleation which takes roughly the same time, but produces significantly lower number of bubbles. In the case of heterogeneous nucleation, a pressure perturbation, induced by a limited number of new created bubbles, propagates through a magma-filled cavity providing a trigger for the homogeneous nucleation in the whole volume of magma. Such combination of heterogeneous and homogeneous nucleations is often assumed for many natural systems[28]. The duration of this process is controlled by a propagation time of a pressure pulse across a volume of over-saturated magma. With typical dimensions of a few tens of meters and sound speed being of the order of a few km/s, the combined heterogeneous and homogeneous nucleation will take less than 0.01 s, i.e., two orders of magnitude below typical bubble growth times. Therefore, we consider instantaneous nucleation in the whole volume.

Numerical method. Equation (1) can be integrated analytically and gives the following velocity distribution in the melt phase: $v_r = \frac{dR}{dt}\frac{R^2}{r^2}$. In order to solve Eq. (2) in a fixed domain we use front-fixing method[43]. A coordinate transformation $\xi = \frac{r - R(t)}{S(t) - R(t)}$ gives extra advective term in Eq. (2). The resulting equation is discretized on an irregular 1D mesh with a decrease of the step size towards the growing bubble boundary ($\xi = 0$). The resulting system of equations with three-diagonal matrix is solved by means of Thomas algorithm[44]. The forward step starts at the outer domain boundary ($\xi = 1$). The linear relation of volatile concentration on the bubble boundary and in the nearest mesh points together with discretized Eqs. (3) and (4) allows to calculate all parameters on the bubble-melt interface. Then, concentration distribution in the whole domain is calculated during backward substitution. We found this method stable and computationally efficient in comparison with explicit methods that require extremely small timesteps for stability reasons.

Estimation of magma composition. In order to estimate magma compositions in the deep magma reservoir we used "Petrolog" software[45]. Reverse crystallization from a more evolved magma (sample 12KY-108-1, 1987 AD eruption[46]) was performed. The starting pressure is set to atmospheric level and magma is H_2O-saturated. Our simulations reveal total amount of mineral phase of 20% for the starting composition, which is in a good agreement with the measurements on the samples[47]. Incremental increase in pressure to 800 MPa leads to the change in composition presented in Supplementary Table 1. These values were obtained considering the volatile component composed only of H_2O, resulting in a 800 MPa magma containing almost 11 wt% of dissolved water. Adding even a small amount of CO_2 affects significantly the water solubility that can be reduced to a few wt% as shown in Fig. 2 along the 800 MPa isobar. Based on data about Klyuchevskoy magma volatile content[22–25,48], we retain for our modeling a composition with ~4 wt% of H_20 and ~0.6 wt% of CO_2.

Estimation of magnitudes of deep low-frequency earthquakes. The DLP signals whose energy is concentrated in a narrow spectral band between 1 and 2 Hz are dominated by S-waves (Fig. 3e, f). The seismic moment can be approximately estimated from maximal signal amplitude in the following way. We start with an expression of the far-field (hypocenter distances exceeding 10 wavelengths) S-wave displacement[49] u^S and ignore the radiation pattern assuming that it approximately averages to 1. Based on this we can relate the time derivative of seismic moment with the observed S-wave displacement:

$$\dot{M}_0(t) \sim 4\pi\rho\beta^3 r \cdot u^S(t) \tag{6}$$

where t is time M_0 is seismic moment, ρ is density, β is S-wave speed, and r is the hypocentral distance. The observed ground velocity v^S is the derivative of the displacement that for a nearly monochromatic signal can be approximately estimated via multiplication by $2\pi f^{\max}$:

$$v^S(t) \sim \dot{u}^S(t) \sim 2\pi f^{\max} \cdot u^S(t) \tag{7}$$

where $f^{\max}$ is the dominant signal frequency. Integration of Eq. (6) to obtain the whole seismic moment can be also approximately estimated with dividing by $2\pi f^{\max}$. This leads to a final expression used to approximately estimate the seismic moment from one station:

$$M_0 \sim \frac{4\pi\rho\beta^3 r \cdot u^S}{2\pi f^{\max}} = \frac{\rho\beta^3 r}{\pi f^{\max 2}}\left|v^s_{\max}\right| \tag{8}$$

where $v^s_{\max}$ is the maximum amplitude of velocity seismograms (taking into account all three components). The final estimate is averaged from several stations that recorded the earthquake. We use $f^{\max} = 1.5$ Hz and typical crustal values for density[27], $\rho = 2830\ \mathrm{kg\ m}^{-3}$, and seismic velocity[50,51], $\beta = 3500\ \mathrm{m\ s}^{-1}$. The moment magnitude M_W is then computed as:

$$M_w = \frac{2}{3}(\lg(M_0) - 9.05) \tag{9}$$

Estimation of seismic radiation emitted by expanding magma volume. For simplicity, we start with considering a volume with a perfectly spherical shape embedded in an infinite elastic space with bulk modulus K. In response to the magma pressure change $dP(t)$, the volume will be modified by $dV(t)$:

$$dV(t) = \frac{dP(t)V}{K} \tag{10}$$

For a perfectly spherical magma body, the volume change can be related to the seismic moment as[49]:

$$M_o(t) = KdV(t) = dP(t)V \tag{11}$$

A spherically symmetric source would radiate in the far field only P waves. At the same time, signals from real DLP earthquakes are dominated by S waves. A simple explanation of this observations can be related to the deviation of the magma body shape from a perfect sphere. In this case, the change of the magma pressure will induce a significant amount of shear stress in the surrounding rocks resulting in a strong S-wave radiation[52]. A possible example is a pure tensile crack mechanism for which the seismic moment tensor can be written as[49]:

$$M(t) = \begin{pmatrix} \lambda dV(t) & 0 & 0 \\ 0 & \lambda dV(t) & 0 \\ 0 & 0 & (\lambda + 2\mu)dV(t) \end{pmatrix} \tag{12}$$

where λ and μ are Lamé constants that for most of elastic solids are nearly equal and have the same order of magnitude as bulk modulus ($K = \lambda + 2/3\mu$) implying that to the order of magnitude the relationship (11) between seismic moment (observed amplitudes of waves), pressure variations, and volume of affected fluid remain valid. Seismic radiation from such source for many directions is dominated by S-waves[53].

At this stage, we do not consider detailed description of seismic radiation from a non-spherical source that would vary significantly depending on the exact magma volume shape. We rather make an order of magnitude estimation and consider that Eq. (11) describes the relationship between the magma pressure change and the seismic moment observed in the far field (hypocenter distances exceeding 10 wavelengths). Based on Eq. (6), the ground displacement can be expressed as:

$$u(t) \approx \frac{\dot{M}_0(t)}{4\pi\rho\beta^3 r} = \frac{\dot{P}(t)V}{4\pi\rho\beta^3 r} \tag{13}$$

and the ground velocity is computed as its time derivative.

Data availability

The seismological time series used for the analysis were provided by the Kamchatka Branch of the Geophysical Survey of Russian Academy of Sciences (GS RAS) and are available on request (http://www.emsd.ru). The data are not publicly available due to the internal regulation of the GS RAS.

Received: 5 March 2020; Accepted: 10 July 2020;
Published online: 06 August 2020

References

1. Pitt, A. M. & Hill, D. P. Long-period earthquakes in the Long-Valley Caldera region eastern California. *Geophys. Res. Lett.* **21**, 1679–1682 (1994).
2. White, R. A. in *Fire and mud: eruptions and lahars of Mount Pinatubo, Philippines* (eds Newhall, C. G. & Punongbayan, R. S.) 307–326 (Univ. Washington Press, 1996).
3. Pitt, A. M., Hill, D. P., Walter, S. W. & Johnston, M. J. S. Mid-crustal, long-period earthquakes beneath northern California volcanic areas. *Seismol. Res. Lett.* **73**, 144–152 (2002).
4. Power, J. A., Stihler, S. D., White, R. A. & Moran, S. C. Observations of deep long-period (DLP) seismic events beneath Aleutian arc volcanoes; 1989–2002. *J. Volcanol. Geotherm. Res.* **138**, 243–266 (2004).
5. Nichols, M. L., Malone, S. D., Moran, S. C., Thelen, W. A. & Vidale, J. E. Deep long-period earthquakes beneath Washington and Oregon volcanoes. *J. Volcanol. Geotherm. Res.* **200**, 116–128 (2011).
6. Aso, N., Ohta, K. & Ide, S. Tectonic, volcanic, and semi-volcanic deep low-frequency earthquakes in western Japan. *Tectonophysics* **600**, 27–40 (2013).
7. Shapiro, N. M. et al. Deep and shallow long-period volcanic seismicity linked by fluid-pressure transfer. *Nat. Geosci.* **10**, 442–445 (2017).
8. Hensch, M. et al. Deep low-frequency earthquakes reveal ongoing magmatic recharge beneath Laacher See Volcano (Eifel, Germany). *Geophys. J. Int* **216**, 2025–2036 (2019).
9. Wech, A. G., Thelen, W. A. & Thomas, A. M. Deep long-period earthquakes generated by second boiling beneath Mauna Kea volcano. *Science* **368**, 775–779 (2020).
10. Chouet, B. A. Long-period volcano seismicity: its source and use in eruption forecasting. *Nature* **380**, 309–316 (1996).
11. Aso, N. & Tsai, V. C. Cooling magma model for deep volcanic long-period earthquakes. *J. Geophys. Res.* **119**, 8442–8456 (2014).
12. Chouet, B. A. & Matoza, R. S. A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. *J. Volcanol. Geotherm. Res.* **252**, 108–175 (2013).
13. Shapiro, N. M., Campillo, M., Kaminski, E., Vilotte, J. -P. & Jaupart, C. Low-frequency earthquakes and pore pressure transients in subduction zones. *Geophys. Res. Lett.* **45**, 11,083–11,094 (2018).
14. Lensky, N. G., Niebo, R. W., Holloway, J. R., Lyakhovsky, V. & Navon, O. Bubble nucleation as a trigger for xenolith entrapment in mantle melts. *Earth Planet. Sci. Lett.* **245**, 278–288 (2006).
15. Lyakhovsky, V., Hurwitz, S. & Navon, O. Bubble growth in rhyolitic melts: experimental and numerical investigations. *Bull. Volcanol.* **58**, 19–32 (1996).
16. Gonnermann, H. M. & Manga, M. Nonequilibrium magma degassing: results from modeling of the ca. 1340 AD eruption of Mono Craters, California. *Earth Planet. Sci. Lett.* **238**, 1–16 (2005).
17. Shapiro, N. M. et al. Understanding Kamchatka's extraordinary volcano cluster. EOS 98, https://doi.org/10.1029/2017eo071351 (2017).
18. Senyukov, S. L. *Forecasting of the eruptions of volcanoes Klyuchevskoy and Bezymianny at Kamchatka [in Russian]* (Lambert Academic, 2013).
19. Senyukov, S. L. et al. Studies in the activity of Klyuchevskoi volcano by remote sensing techniques between January 1, 2001 and July 31, 2005, Volcanol. *Seismol* **3**, 50–59 (2009).
20. Droznin, D. V. et al. Detecting and locating volcanic tremors on the Klyuchevskoy group of volcanoes (Kamchatka) based on correlations of continuous seismic records. *Geophys. J. Int.* **203**, 1001–1010 (2015).
21. Gorelchik, V. I., Garbuzova, V. T. & Storcheus, A. V. Deep-seated volcanic processes beneath Klyuchevskoi volcano as inferred from seismological data. *J. Volcanol. Seismol.* **6**, 21–34 (2004).
22. Khubunaya, S. A. & Sobolev, A. V. Primary melts of calc–alkaline magnesian basalts from Klyuchevskoy Volcano, Kamchatka, Dokl. *Akad. Nauk* **360**, 100–102 (1998).
23. Portnyagin, M., Hoernle, K., Plechov, P., Mironov, N. & Khubunaya, S. Constraints on mantle melting and composition and nature of slab components in volcanic arcs from volatiles (H_2O, S, Cl, F) and trace elements in melt inclusions from the Kamchatka Arc. *Earth Planet. Sci. Lett.* **255**, 53–69 (2007).
24. Mironov, N. L. & Portnyagin, M. V. H_2O and CO_2 in parental magmas of Kliuchevskoi volcano inferred from study of melt and fluid inclusions in olivine. *Russian Geol. Geophys.* **52**, 1353–1367 (2011).
25. Portnyagin, M. et al. Dehydration of melt inclusions in olivine and implications for the origin of silica-undersaturated island-arc melts. *Earth Planet. Sci. Lett.* **517**, 95–105 (2019).
26. Stolper, E. & Walker, D. Melt density and the average composition of basalt. *Contrib. Mineral. Petrol.* **74**, 7–12 (1980).
27. Christensen, N. I. & Mooney, W. D. Seismic velocity structure and composition of the continental crust: A global view. *J. Geophys. Res. Solid Earth* **100**(B6), 9761–9788 (1995).
28. Shea, T. Bubble nucleation in magmas: a dominantly heterogeneous process? *J. Volcanol. Geotherm. Res.* **343**, 155–170 (2017).
29. Shaw, H. R. Viscosities of magmatic silicate liquids; an empirical method of prediction. *Am. J. Sci.* **272**, 870–893 (1972).
30. Sable, J. E., Houghton, B. F., Del Carlo, P. & Coltelli, M. Changing conditions of magma ascent and fragmentation during the Etna 122 BC basaltic Plinian eruption: evidence from clast microtextures. *J. Volcanol. Geotherm. Res.* **158**, 333–354 (2006).
31. Aki, K. & Chouet, B. Origin of coda waves: source, attenuation, and scattering effects. *J. Geophys. Res.* **80**, 3322–3342 (1975).
32. Allard, P. et al. Eruptive and diffuse emissions of CO_2 from Mount Etna. *Nature* **351**, 387–391 (1991).
33. Edmonds, M. New geochemical insights into volcanic degassing. *Philos. Trans. R. Soc. A* **366**, 4559–4579 (2008).
34. Burton, M. R., Sawyer, G. M. & Granieri, D. Deep carbon emissions from volcanoes. *Rev. Mineral. Geochem.* **75**, 323–354 (2013).
35. Hartley, M. E., Maclennan, J., Edmonds, M. & Thordarson, T. Reconstructing the deep CO_2 degassing ehavior of large basaltic fissure eruptions. *Earth Planet. Sci. Lett.* **393**, 120–131 (2014).
36. Kelemen, P. B. & Manning, C. E. Reevaluating carbon fluxes in subduction zones, what goes down, mostly comes up. *Proc. Natl Acad. Sci. USA* **112**, E3997–E4006 (2015).
37. Taran, Y. et al. Gas emissions from volcanoes of the Kuril Island arc (NW Pacific): geochemistry and fluxes. *Geochem. Geophys. Geosyst.* **19**, 1859–1880 (2018).
38. Aiuppa, A. et al. CO_2 flux emissions from the Earth's most actively degassing volcanoes, 2005–2015. *Sci. Rep.* **9**, 5442 (2019).
39. Zhang, Y., Ni, H. Diffusion of H, C, and O Components in Silicate Melts. *Reviews in Mineralogy and Geochemistry.* **72**, 171–225 (2010).
40. Burgisser, A., Alletti, M. & Scaillet B. *D-Compress* https://vhub.org/resources/3791 (2015).
41. Hirth, G., Pound, G. M. & St Pierre, G. R. Bubble nucleation. *Metall. Trans.* **1**, 939–945 (1970).
42. Navon, O. & Lyakhovsky, V. Vesiculation processes in silicic magmas. *Geol. Soc., Lond. Spec. Pub* **145**, 27–50, https://doi.org/10.1144/GSL.SP.1996.145.01.93 (1998).
43. Crank, J. *Free and moving boundary problems* (Oxford Science Publications, 1987).
44. Mooney, D. D. & Swift, R. J. *A course in mathematical modeling* (Cambridge University Press, 1999).
45. Danyushevsky, L. V. & Plechov, P. Petrolog3: Integrated software for modeling crystallization processes. *Geochem. Geophys. Geosyst.* **12**, Q07021 (2011).
46. Bergal-Kuvikas, O. et al. A petrological and geochemical study on time-series samples from Klyuchevskoy volcano, Kamchatka arc. *Contrib. Mineral. Petrol.* **172**, 35 (2017).
47. Bergal-Kuvikas, O. *Geochemical studies of volcanic rocks from the northern part of Kuril-Kamchatka arc: tectonic and structural constraints on the origin and evolution of arc magma*. Doctoral dissertation. P.190. (Hokkaido University, 2015).
48. Auer, S., Bindeman, I., Wallace, P., Ponomareva, V. & Portnyagin, M. The origin of hydrous, high-delta O-18 voluminous volcanism: diverse oxygen isotope values and high magmatic water contents within the volcanic record of Klyuchevskoi volcano, Kamchatka, Russia. *Contrib. Mineral. Petrol.* **157**, 209–230 (2009).
49. Aki, K. & Richards, P. G. *Quantitative seismology* (University Science Books, 2002).
50. Levin, V., Droznina, S., Gavrilenko, M., Carr, M. J. & Senyukov, S. Seismically active subcrustal magma source of the Klyuchevskoy volcano in Kamchatka, Russia. *Geology* **42**, 983–986 (2014).
51. Droznina, S. et al. S-wave velocity model for several regions of the Kamchatka Peninsula from the cross correlations of ambient seismic noise. *Izvestiya Phys. Solid Earth* **53**, 341–352 (2017).
52. Eshelby, J. D. The elastic field outside an ellipsoidal inclusion. *Proc. R. Soc. A* **252**, 561–569 (1959).
53. Shi, Z. & Ben-Zion, Y. Seismic radiation from tensile and shear point dislocations between similar and dissimilar solids. *Geophys. J. Int* **179**, 444–458 (2009).

Acknowledgements

We thank Yuri Taran, Patrick Allard, Roberto Moretti, and Daniel Neuville for helpful discussions as well as Brian Shirofor his very helpful revision of the manuscript. This study was supported by the Russian Ministry of Education and Science (Grant no. 14. W03.31.0033), and by the European Research Council (ERC) under the European Union Horizon 2020 Research and Innovation Programme (Grant agreement 787399-SEISMAZE).

Author contributions

O.M. and V.L. designed the model of the gas bubble grows in the magma and made the computation. N.S. participated in the discussion of the model design and analyzed seismological observations. N.G. analyzed seismological observations. O.B. analyzed the composition of the Klyochevskoy basaltic magmas. All of the authors contributed to interpretation of the data, discussions of the results, and preparation of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-17759-4.

Correspondence and requests for materials should be addressed to N.M.S.

Peer review information *Nature Communications* thanks Brian Shiro and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Bibliography

K. Aki and B. Chouet. Origin of coda waves: source, attenuation, and scattering effects. *J. Geophys. Res.*, 80(23):3322–3342, 1975. doi: 10.1016/S0031-8914(53)80099-6.

K. Aki and P. G. Richards. *Quantitative Seismology: Theory and Methods*, volume 842. San Francisco, CA: Freeman, 1980. doi: 10.1002/gj.3350160110.

C. Annen, J. Blundy, and R. Sparks. The genesis of intermediate and silicic magmas in deep crustal hot zones. *Journal of Petrology*, 47(3):505–539, 2006.

D. Arthur and S. Vassilvitskii. How slow is the k-means method? In *Proceedings of the twenty-second annual symposium on Computational geometry*, pages 144–153, 2006.

D. Arthur and S. Vassilvitskii. K-means++ the advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035, 2007.

N. Aso and S. Ide. Focal mechanisms of deep low-frequency earthquakes in Eastern Shimane in Western Japan. *Journal of Geophysical Research: Solid Earth*, 119(1):364–377, 2014. doi: https://doi.org/10.1002/2013JB010681.

N. Aso and V. C. Tsai. Cooling magma model for deep volcanic long-period earthquakes. *J. Geophys. Res. Solid Earth*, 119(11):8442–8456, 2014. doi: 10.1002/2014JB011180.

N. Aso, K. Ohta, and S. Ide. Tectonic, volcanic, and semi-volcanic deep low-frequency earthquakes in Western Japan. *Tectonophysics*, 600:27–40, 2013. doi: https://doi.org/10.1016/j.tecto.2012.12.015.

O. Bachmann and G. W. Bergantz. On the origin of crystal-poor rhyolites: extracted from batholithic crystal mushes. *Journal of Petrology*, 45(8):1565–1582, 2004.

T. Bardainne, P. Gaillot, N. Dubos-Sallée, J. Blanco, and G. Sénéchal. Characterization of seismic waveforms and classification of seismic events using chirplet atomic decomposition. Example from the Lacq gas field (Western Pyrenees, France). *Geophysical Journal International*, 166(2):699–718, 2006.

C. J. Bean, L. De Barros, I. Lokmer, J.-P. Métaxian, G. O'Brien, and S. Murphy. Long-period seismicity in the shallow volcanic edifice formed from slow-rupture earthquakes. *Nature Geoscience*, 7(1):71–75, 2014. doi: 10.1038/ngeo2027.

R. Bellman. Dynamic programming. *Press Princeton, New Jersey*, 1957.

L. Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

J. R. Brown, G. C. Beroza, and D. R. Shelly. An autocorrelation method to detect low frequency earthquakes within tremor. *Geophysical Research Letters*, 35(16), 2008.

S. Burchardt, C. J. Annen, J. L. Kavanagh, and S. Hilmi Hazim. Developments in the study of volcanic and igneous plumbing systems: outstanding problems and new opportunities. *Bulletin of Volcanology*, 84(6):59, 2022.

C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167, 1998.

M. Campillo and J. Plantet. Frequency dependence and spatial distribution of seismic attenuation in France: experimental results and possible interpretations. *Physics of the Earth and Planetary Interiors*, 67(1):48–64, 1991. doi: https://doi.org/10.1016/0031-9201(91)90059-Q.

K. V. Cashman, S. J. Sparks, and J. D. Blundy. Vertically extensive and unstable magmatic systems: a unified view of igneous processes. *Science*, 355(6331):eaag3055, 2017.

D. Chebrov, S. Droznina, S. Senyukov, Y. Shevchenko, S. Mityushkina, V. Chebrov, D. Droznin, V. Sergeev, and E. Pantyukhin. A system for acquisition, processing, storing and representing seismological data and the results of their processing in the sp spts, hardware, algorithms and software (in russian). In V. Chebrov, editor, *Problemy kompleksnogo geofizicheskogo monitoringa Dal'nego Vostoka Rossii: Tr. vtoroi reg. nauchno-tekh. konf. (Problems of Integrated Geophysical Monitoring of the Russian Far East: Proc. 2nd Reg. Sci. Tech. Conf.)*, pages 332–336. Petropavlovsk-Kamchatskii: GS RAN, 2010.

V. N. Chebrov, D. V. Droznin, Y. A. Kugaenko, V. I. Levina, S. L. Senyukov, V. A. Sergeev, Y. V. Shevchenko, and V. V. Yashchuk. The system of detailed seismological observations in Kamchatka in 2011. *J. Volcanol. Geotherm. Res.*, 7(1):16–36, 2013. doi: 10.1134/S0742046313010028.

B. Chouet. Long-period volcano seismicity: its source and use in eruption forecasting. *Nature*, 380:309–316, 1996a. doi: 10.1038/380309a0.

B. Chouet. Volcano seismology. *Pure and applied geophysics*, 160(3):739–788, 2003.

B. Chouet and R. S. Matoza. A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. *J. Volcanol. Geotherm. Res.*, 252: 109–175, 2013. doi: 10.1016/j.jvolgeores.2012.11.013.

B. Chouet, P. Dawson, T. Ohminato, M. Martini, G. Saccorotti, F. Giudicepietro, G. De Luca, G. Milana, and R. Scarpa. Source mechanisms of explosions at Stromboli Volcano, Italy, determined from moment-tensor inversions of very-long-period data. *Journal of Geophysical Research: Solid Earth*, 108(B1):ESE 7–1–ESE 7–25, 2003. doi: https://doi.org/10.1029/2002JB001919.

B. Chouet, P. Dawson, and A. Arciniega-Ceballos. Source mechanism of Vulcanian degassing at Popocatépetl volcano, Mexico, determined from waveform inversions of very long period signals. *Journal of Geophysical Research: Solid Earth*, 110(B7), 2005. doi: https://doi.org/10.1029/2004JB003524.

B. A. Chouet. Long-period volcano seismicity: its source and use in eruption forecasting. *Nature*, 380(6572):309, 1996b.

K. M. Cooper and A. J. Kent. Rapid remobilization of magmatic crystals kept in cold storage. *Nature*, 506(7489):480–483, 2014.

D. Coppola, M. Laiolo, F. Massimetti, S. Hainzl, A. V. Shevchenko, R. Mania, N. M. Shapiro, and T. R. Walter. Thermal remote sensing reveals communication between volcanoes of the Klyuchevskoy volcanic group. *Scientific Reports*, 11(1):13090, 2021. doi: 10.1038/s41598-021-92542-z.

C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20:273–297, 1995.

G. Cortés, R. Carniel, P. Lesage, M. Á. Mendoza, and I. Della Lucia. Practical volcano-independent recognition of seismic events: Vulcan.ears project. *Frontiers in Earth Science*, 8, 2021. doi: 10.3389/feart.2020.616676.

G. Curilem, J. Vergara, G. Fuentealba, G. Acuña, and M. Chacón. Classification of seismic signals at Villarrica volcano (Chile) using neural networks and genetic algorithms. *Journal of volcanology and geothermal research*, 180(1):1–8, 2009.

M. Curilem, J. Vergara, C. San Martin, G. Fuentealba, C. Cardona, F. Huenupan, M. Chacón, M. S. Khan, W. Hussein, and N. B. Yoma. Pattern recognition applied to seismic signals of the Llaima volcano (Chile): An analysis of the events' features. *Journal of Volcanology and Geothermal Research*, 282:134–147, 2014.

F. Dorendorf, U. Wiechert, and G. Wörner. Hydrated sub-arc mantle: a source for the Kluchevskoy volcano, Kamchatka/Russia. *Earth and Planetary Science Letters*, 175(1):69–86, 2000. doi: https://doi.org/10.1016/S0012-821X(99)00288-5.

D. Droznin, N. Shapiro, S. Y. Droznina, S. Senyukov, V. Chebrov, and E. Gordeev. Detecting and locating volcanic tremors on the Klyuchevskoy group of volcanoes (Kamchatka) based on correlations of continuous seismic records. *Geophysical Journal International*, 203(2):1001–1010, 09 2015. doi: 10.1093/gji/ggv342.

D. V. Droznin and S. Y. Droznina. Interactive DIMAS program for processing seismic signals. *Seismic Instruments*, 47(3):215, 2011.

E. T. Endo and T. Murray. Real-time seismic amplitude measurement (RSAM): a volcano monitoring and prediction tool. *Bulletin of Volcanology*, 53(7):533–545, 1991.

A. Esposito, F. Giudicepietro, S. Scarpetta, L. D'auria, M. Marinaro, and M. Martini. Automatic discrimination among landslide, explosion-quake, and microtremor seismic signals at Stromboli volcano using neural networks. *Bulletin of the Seismological Society of America*, 96(4A):1230–1240, 2006.

A. M. Esposito, F. Giudicepietro, L. D'Auria, S. Scarpetta, M. G. Martini, M. Coltelli, and M. Marinaro. Unsupervised neural analysis of very-long-period events at Stromboli volcano using the self-organizing maps. *Bulletin of the Seismological Society of America*, 98(5):2449–2459, 2008.

S. Falsaperla, S. Graziani, G. Nunnari, and S. Spampinato. Automatic classification of volcanic earthquakes by using multi-layered neural networks. *Natural Hazards*, 13:205–228, 1996.

S. A. Fedotov. *Energy classification of Kuril-Kamchatka earthquakes and the magnitude problem (in Russian)*. Nauka, 1972.

S. A. Fedotov, N. A. Zharinov, and L. I. Gontovaya. The magmatic system of the Klyuchevskaya group of volcanoes inferred from data on its eruptions, earthquakes, deformation, and deep structure. *J. Volcanol. Geotherm. Res.*, 4(1):1–33, 2010. doi: 10.1134/S074204631001001X.

M. Fehler. Observations of volcanic tremor at Mount St. Helens volcano. *Journal of Geophysical Research: Solid Earth*, 88(B4):3476–3484, 1983.

C. Frohlich and S. D. Davis. Teleseismic b values; or, much ado about 1.0. *Journal of Geophysical Research: Solid Earth*, 98(B1):631–644, 1993.

J. Gardner and L. Knopoff. Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the seismological society of America*, 64(5):1363–1367, 1974.

V. Gavrilov, V. Voropaev, I. Golovshchikova, Y. Liannik, A. Pudov, and G. Torosian. System of radio telemetry instruments tesi–2 (in russian). *Seismic Instruments*, 19: 5–16, 1987.

F. Giacco, A. M. Esposito, S. Scarpetta, F. Giudicepietro, and M. Marinaro. Support vector machines and MLP for automatic classification of seismic signals at Stromboli volcano. In *IOS Press*, volume 204, page 116, 2009.

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

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

A. Glazner, J. M. Bartley, and D. S. Coleman. We need a new definition for "magma". *Eos*, 97, 2016.

E. I. Gordeev. Origin of seismic signals in active volcanoes (Dissertation for the Doctor of Science degree), 1998.

V. I. Gorelchik and A. V. Storcheus. Deep long-period earthquakes beneath Klyuchevskoy volcano, Kamchatka (in Russian). *Geodinamika i vulkanizm Kurilo-Kamchatskoi ostrovoduzhnoi sistemy*, pages 173–189, 2001.

V. I. Gorelchik, V. T. Garbuzova, and A. V. Storcheus. Deep-seated volcanic processes beneath klyuchevskoi volcano as inferred from seismological data. *Journal of Volcanology and Seismology*, 6:21–34, 2004.

R. G. Green, C. Sens-Schönfelder, N. Shapiro, I. Koulakov, F. Tilmann, J. Dreiling, B. Luehr, A. Jakovlev, I. Abkadyrov, D. Droznin, and E. Gordeev. Magmatic and

sedimentary structure beneath the Klyuchevskoy volcanic group, Kamchatka, from ambient noise tomography. *Journal of Geophysical Research: Solid Earth*, 125(3): e2019JB018900, 2020. doi: https://doi.org/10.1029/2019JB018900.

T. Greenfield, T. Winder, N. Rawlinson, J. Maclennan, R. S. White, T. Ágústsdóttir, C. A. Bacon, B. Brandsdóttir, E. P. Eibl, E. Glastonbury-Southern, et al. Deep long period seismicity preceding and during the 2021 fagradalsfjall eruption, iceland. *Bulletin of Volcanology*, 84(12):101, 2022.

T. L. Grove and C. B. Till. Melting the Earth's upper mantle. In *The Encyclopedia of Volcanoes*, pages 35–47. Elsevier, 2015.

A. A. Gusev and V. N. Melnikova. Relations between magnitudes: Global and kamchatka data (in russian). *Vulkanologiya i Seismologiya*, 6:55–63, 1990.

B. Gutenberg and C. F. Richter. Frequency of earthquakes in california. *Bulletin of the Seismological Society of America*, 34(4):185–188, 1944.

C. Hammer, M. Beyreuther, and M. Ohrnberger. A seismic-event spotting system for volcano fast-response systems. *Bulletin of the Seismological Society of America*, 102 (3):948–960, 2012.

T. C. Hanks and H. Kanamori. A moment magnitude scale. *J. Geophys. Res.*, 84(5): 2348–2350, 1979.

S. Har-Peled and B. Sadri. How fast is the k-means method? *Algorithmica*, 41: 185–202, 2005.

J. L. Hardebeck and P. M. Shearer. Using S/P amplitude ratios to constrain the focal mechanisms of small earthquakes. *Bulletin of the Seismological Society of America*, 93(6):2434–2444, 2003. doi: 10.1785/0120020236.

A. Hasegawa and A. Yamamoto. Deep, low-frequency microearthquakes in or around seismic low-velocity zones beneath active volcanoes in northeastern Japan. *Tectonophysics*, 233(3):233–252, 1994. doi: https://doi.org/10.1016/0040-1951(94)90243-7.

T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.

M. Hensch, T. Dahm, J. Ritter, S. Heimann, B. Schmidt, S. Stange, and K. Lehmann. Deep low-frequency earthquakes reveal ongoing magmatic recharge beneath Laacher See Volcano (Eifel, Germany). *Geophysical Journal International*, 216(3):2025–2036, 01 2019. doi: 10.1093/gji/ggy532.

C. Hibert, F. Provost, J.-P. Malet, A. Maggi, A. Stumpf, and V. Ferrazzini. Automatic identification of rockfalls and volcano-tectonic earthquakes at the Piton de la Fournaise volcano using a Random Forest algorithm. *Journal of Volcanology and Geothermal Research*, 340:130–142, 2017.

W. Hildreth. Gradients in silicic magma chambers: implications for lithospheric magmatism. *Journal of Geophysical Research: Solid Earth*, 86(B11):10153–10192, 1981.

W. Hildreth. Volcanological perspectives on Long Valley, Mammoth Mountain, and Mono Craters: several contiguous but discrete systems. *Journal of Volcanology and Geothermal Research*, 136(3-4):169–198, 2004.

A. Husker, S. Peyrat, N. Shapiro, and V. Kostoglodov. Automatic non-volcanic tremor detection in the Mexican subduction zone. *Geofis. Int.*, 49(1):17–25, 2010.

T. Ikegaya and M. Yamamoto. Spatio-temporal characteristics and focal mechanisms of deep low-frequency earthquakes beneath the Zao volcano, northeastern Japan. *Journal of Volcanology and Geothermal Research*, 417:107321, 2021. doi: https://doi.org/10.1016/j.jvolgeores.2021.107321.

V. V. Ivanov. Current cycle of the Kluchevskoy volcano activity in 1995–2008 based on seismological, photo, video and visual data (in russian). In *Proceedings of Conference Devoted to Volcanologist Day, Petropavlovsk-Kamchatsky, 27–29 March*, pages 100–109. Petropavlovsk-Kamchatskii, 27–29 March 2008.

C. Journeau, L. Shapiro, N. M.and Seydoux, J. Soubestre, I. Y. Koulakov, A. V. Jakovlev, I. Abkadyrov, E. I. Gordeev, D. V. Chebrov, D. V. Droznin, C. Sens-Schönfelder, B. G. Luehr, F. Tong, G. Farge, and C. Jaupart. Seismic tremor reveals active trans-crustal magmatic system beneath Kamchatka volcanoes. *Sci. Adv.*, 8(5):eabj1571, 2022. doi: 10.1126/sciadv.abj1571.

H. Kanamori. The energy release in great earthquakes. *Journal of geophysical research*, 82(20):2981–2987, 1977.

H. Kanamori and D. L. Anderson. Theoretical basis of some empirical relations in seismology. *Bulletin of the seismological society of America*, 65(5):1073–1095, 1975.

A. Kato and S. Nakagawa. Multiple slow-slip events during a foreshock sequence of the 2014 Iquique, Chile Mw 8.1 earthquake. *Geophysical Research Letters*, 41(15):5420–5427, 2014.

V. I. Keilis-Borok, L. Knopoff, I. M. Rotwain, and C. R. Allen. Intermediate-term prediction of occurrence times of strong earthquakes. *Nature*, 335(6192):690–694, 1988. doi: 10.1038/335690a0.

A. Köhler, M. Ohrnberger, and F. Scherbaum. Unsupervised feature selection and general pattern discovery using self-organizing maps for gaining insights into the nature of seismic wavefields. *Computers & Geosciences*, 35(9):1757–1767, 2009.

H. Kumagai. Volcano seismic signals, source quantification of. In R. A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*, pages 9899–9932. Springer New York, New York, NY, 2009. ISBN 978-0-387-30440-3. doi: 10.1007/978-0-387-30440-3_583.

R. Kurihara and K. Obara. Spatiotemporal Characteristics of Relocated Deep Low-Frequency Earthquakes Beneath 52 Volcanic Regions in Japan Over an Analysis Period of 14 Years and 9 Months. *J. Geophys. Res. Solid Earth*, 126(10): e2021JB022173, 2021. doi: 10.1029/2021JB022173.

R. Kurihara, K. Obara, A. Takeo, and Y. Tanaka. Deep low-frequency earthquakes associated with the eruptions of Shinmoe-dake in Kirishima volcanoes. *J. Geophys. Res. Solid Earth*, 124(12):13079–13095, 2019. doi: 10.1029/2019JB018032.

J. C. Lahr, B. A. Chouet, C. D. Stephens, J. A. Power, and R. A. Page. Earthquake classification, location, and error analysis in a volcanic environment: Implications for the magmatic system of the 1989–1990 eruptions at Redoubt Volcano, Alaska. *Journal of Volcanology and Geothermal Research*, 62(1-4):137–151, 1994.

G. N. Lance and W. T. Williams. A general theory of classificatory sorting strategies: 1. hierarchical systems. *The computer journal*, 9(4):373–380, 1967.

H. Langer, S. Falsaperla, and G. Thompson. Application of artificial neural networks for the classification of the seismic transients at Soufriere Hills volcano, Montserrat. *Geophysical research letters*, 30(21), 2003.

J. H. Latter. *Volcanological Observations at Tongariro National Park: 2. Types and Classification of Volcanic Earthquakes, 1976-1978*. Geophysics Division, Department of Scientific and Industrial Research, 1979.

W. H. K. Lee and J. C. Lahr. *HYPO71: A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes*. US Department of the Interior, Geological Survey, National Center for Earthquake Research, 1972.

J. M. Lees. Seismic tomography of magmatic systems. *Journal of Volcanology and Geothermal Research*, 167(1-4):37–56, 2007.

V. Levin, N. Shapiro, J. Park, and M. Ritzwoller. Seismic evidence for catastrophic slab loss beneath kamchatka. *Nature*, 418(6899):763, 2002.

V. Levin, S. Y. Droznina, M. Gavrilenko, M. J. Carr, and S. L. Senyukov. Seismically active subcrustal magma source of the Klyuchevskoy volcano in Kamchatka, Russia. *Geol.*, 42(11):983–986, 2014. doi: 10.1130/G35972.1.

S. Lloyd. Least squares quantization in PCM. *IEEE transactions on information theory*, 28(2):129–137, 1982.

P. Lowman, J. Yates, P. Masuoka, B. Montgomery, J. O'Leary, and D. Salisbury. A digital tectonic activity map of the Earth. *Journal of Geoscience Education*, 47(5): 428–437, 1999.

J. MacQueen. The classification problem. *Western Management Science Institute Working Paper*, (5), 1962.

A. Maggi, V. Ferrazzini, C. Hibert, F. Beauducel, P. Boissier, and A. Amemoutou. Implementation of a multistation approach for automated event classification at Piton de la Fournaise volcano. *Seismological Research Letters*, 88(3):878–891, 2017.

I. G. Main. A characteristic earthquake model of the seismicity preceding the eruption of Mount St. Helens on 18 May 1980. *Physics of the earth and planetary interiors*, 49(3-4):283–293, 1987.

M. Malfante, M. Dalla Mura, J. I. Mars, J.-P. Métaxian, O. Macedo, and A. Inza. Automatic classification of volcano seismic signatures. *Journal of Geophysical Research: Solid Earth*, 123(12):10–645, 2018a.

M. Malfante, M. Dalla Mura, J.-P. Metaxian, J. I. Mars, O. Macedo, and A. Inza. Machine learning for volcano-seismic signals: Challenges and perspectives. *IEEE Signal Processing Magazine*, 35(2):20–30, 2018b. doi: 10.1109/MSP.2017.2779166.

L. Margerin, M. Campillo, and B. Van Tiggelen. Monte Carlo simulation of multiple scattering of elastic waves. *Journal of Geophysical Research: Solid Earth*, 105(B4): 7873–7892, 2000. doi: https://doi.org/10.1029/1999JB900359.

B. D. Marsh. On the crystallinity, probability of occurrence, and rheology of lava and magma. *Contributions to Mineralogy and Petrology*, 78:85–98, 1981.

B. D. Marsh. Magma chambers. In *The encyclopedia of volcanoes*, pages 185–201. Elsevier, 2015.

R. S. Matoza and D. C. Roman. One hundred years of advances in volcano seismology and acoustics. *Bulletin of Volcanology*, 84(9):86, 2022. doi: 10.1007/s00445-022-01586-0.

W. J. Maurer, F. U. Dowla, and S. P. Jarpe. Seismic event interpretation using self-organizing neural networks. In *Applications of Artificial Neural Networks III*, volume 1709, pages 950–958. SPIE, 1992.

S. R. McNutt. Volcanic seismology. *Annu. Rev. Earth Planet. Sci.*, 32:461–491, 2005.

S. R. McNutt and D. C. Roman. Volcanic seismicity. In *The encyclopedia of volcanoes*, pages 1011–1034. Elsevier, 2015.

S. R. McNutt, G. Thompson, J. Johnson, S. De Angelis, and D. Fee. Seismic and infrasonic monitoring. In *The encyclopedia of volcanoes*, pages 1071–1099. Elsevier, 2015.

J. McQueen. Some methods for classification and analysis of multivariate observations. In *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1967*, pages 281–297, 1967.

O. Melnik, V. Lyakhovsky, N. M. Shapiro, N. Galina, and O. Bergal-Kuvikas. Deep long period volcanic earthquakes generated by degassing of volatile-rich basaltic magmas. *Nature Communications*, 11(1):3918, 2020. doi: 10.1038/s41467-020-17759-4.

O. E. Melnik, I. S. Utkin, and I. N. Bindeman. Magma chamber formation by dike accretion and crustal melting: 2d thermal model with emphasis on zircon record. *Authorea*, 2022. doi: 10.1002/essoar.10505594.1.

T. Minakami. Fundamental research for predicting volcanic eruptions, Part 1. *Bull. Earthq. Res. Inst. Univ. Tokyo*, 38:497–544, 1960.

S. M. Mousavi and G. C. Beroza. Machine learning in earthquake seismology. *Annual Review of Earth and Planetary Sciences*, 51(1):105–129, 2023. doi: 10.1146/annurev-earth-071822-100323.

M. Murru, R. Console, G. Falcone, C. Montuori, and T. Sgroi. Spatial mapping of the b value at mount etna, italy, using earthquake data recorded from 1999 to 2005. *Journal of Geophysical Research: Solid Earth*, 112(B12), 2007.

H. Nakamichi, H. Hamaguchi, S. Tanaka, S. Ueki, T. Nishimura, and A. Hasegawa. Source mechanisms of deep and intermediate-depth low-frequency earthquakes beneath Iwate volcano, northeastern Japan. *Geophysical Journal International*, 154 (3):811–828, 2003. doi: https://doi.org/10.1046/j.1365-246X.2003.01991.x.

J. Neuberg, R. Luckett, B. Baptie, and K. Olsen. Models of tremor and low-frequency earthquake swarms on Montserrat. *Journal of Volcanology and Geothermal Research*, 101(1-2):83–104, 2000.

M. Nichols, S. Malone, S. Moran, W. Thelen, and J. Vidale. Deep long-period earthquakes beneath Washington and Oregon volcanoes. *J. Volcanol. Geotherm. Res.*, 200(3):116–128, 2011. doi: 10.1016/j.jvolgeores.2010.12.005.

T. Nishimura and M. Iguchi. *Volcanic earthquakes and tremor in Japan*. Kyoto University Press, 2011.

X. Niu, D. Zhao, and J. Li. Precise relocation of low-frequency earthquakes in Northeast Japan: new insight into arc magma and fluids. *Geophys. J. Int.*, 212(2):1183–1200, 10 2017. doi: 10.1093/gji/ggx445.

S. Ohmi and K. Obara. Deep low-frequency earthquakes beneath the focal region of the Mw 6.7 2000 Western Tottori earthquake. *Geophys. Res. Lett.*, 29(16):54–1–54–4, 2002. doi: 10.1029/2001GL014469.

T. Ohminato, B. A. Chouet, P. Dawson, and S. Kedar. Waveform inversion of very long period impulsive signals associated with magmatic injection beneath Kilauea volcano, Hawaii. *Journal of Geophysical Research: Solid Earth*, 103(B10):23839–23862, 1998. doi: https://doi.org/10.1029/98JB01122.

H. Okada, H. Watanabe, H. Yamashita, and I. Yokoyama. Seismological significance of the 1977–1978 eruptions and the magma intrusion process of Usu volcano, Hokkaido. *Journal of volcanology and geothermal research*, 9(4):311–334, 1981.

A. Ozerov, P. Firstov, and V. Gavrilov. Periodicities in the dynamics of eruptions of klyuchevskoi volcano, kamchatka. In *Volcanism and Subduction: The Kamchatka region Geophysical Monograph Series*, volume 172, pages 283–291. American Geophysical Union, 2007.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Z. Peng and P. Zhao. Migration of early aftershocks following the 2004 Parkfield earthquake. *Nature Geoscience*, 2(12):877, 2009.

T. Perol, M. Gharbi, and M. Denolle. Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2):e1700578, 2018.

K. Plenkers, J. R. Ritter, and M. Schindler. Low signal-to-noise event detection based on waveform stacking and cross-correlation: Application to a stimulation experiment. *Journal of seismology*, 17(1):27–49, 2013.

J. Power, S. Stihler, R. White, and S. Moran. Observations of deep long-period (DLP) seismic events beneath Aleutian arc volcanoes; 1989–2002. *J. Volcanol. Geotherm. Res*, 138(3):243–266, 2004. doi: 10.1016/j.jvolgeores.2004.07.005.

T. G. Rautian. Energy of earthquakes. *Methods for the detailed study of seismicity (in Russian)*, pages 75–114, 1960.

T. G. Rautian. On the determination of energy of earthquakes at distances up to 3000 km (in russian). *Trans. Inst. Fiz. Zemli Akad. Nauk USSR*, 199:88–93, 1964.

T. G. Rautian and V. I. Khalturin. The use of the coda for determination of the earthquake source spectrum. *Bull. Seismol. Soc. Am.*, 68(4):923–948, 1978. ISSN 0037-1106. doi: 10.1785/BSSA0680040923.

N. S. Roberts, A. F. Bell, and I. G. Main. Are volcanic seismic b-values high, and if so when? *Journal of Volcanology and Geothermal Research*, 308:127–141, 2015.

Q. Rodríguez-Pérez, F. R. Zúñiga, and C. M. Valdés-González. Statistical earthquake characterization from relocated seismicity at volcan de fuego, colima western mexico. *Journal of Volcanology and Geothermal Research*, 431:107662, 2022.

N. Rogers. The composition and origin of magmas. In *The encyclopedia of volcanoes*, pages 93–112. Elsevier, 2015.

D. C. Roman and K. V. Cashman. The origin of volcano-tectonic earthquake swarms. *Geology*, 34(6):457–460, 06 2006. doi: 10.1130/G22269.1.

A. M. Rubin. Propagation of magma-filled cracks. *Annual Review of Earth and Planetary Sciences*, 23(1):287–336, 1995. doi: 10.1146/annurev.ea.23.050195.001443.

H. Sato. Attenuation and envelope formation of three-component seismograms of small local earthquakes in randomly inhomogeneous lithosphere. *Journal of Geophysical Research: Solid Earth*, 89(B2):1221–1241, 1984. doi: https://doi.org/10.1029/JB089iB02p01221.

S. Scarpetta, F. Giudicepietro, E. C. Ezin, S. Petrosino, E. Del Pezzo, M. Martini, and M. Marinaro. Automatic classification of seismic signals at Mt. Vesuvius volcano, Italy, using neural networks. *Bulletin of the Seismological Society of America*, 95 (1):185–196, 2005.

D. P. Schaff and F. Waldhauser. One magnitude unit reduction in detection threshold by cross correlation applied to Parkfield (California) and China seismicity. *Bulletin of the Seismological Society of America*, 100(6):3224–3238, 2010.

B. Schölkopf and A. Smola. Support vector machines. *Encyclopedia of Biostatistics*, pages 978–0471975762, 1998.

M. Sciotto, A. Cannata, G. Di Grazia, and P. Montalto. Volcanic tremor and long period events at Mt. Etna: Same mechanism at different rates or not? *Physics of the Earth and Planetary Interiors*, 324:106850, 2022.

S. Senyukov, I. Nuzhdina, and V. Chebrov. Seismic monitoring of volcanic regions of Kamchatka (in Russian). In A. Malovichko, editor, *Ezhegodnik "Zemletryaseniya Rossii v 2012 godu" (Yearbook "Earthquakes of Russia in 2012")*, pages 77–81. Obninsk: GC RAS, 2014.

S. Senyukov, I. Nuzhdina, and V. Chebrov. Kamchatka and commander islands (in russian). In A. Malovichko, editor, *Ezhegodnik "Zemletryaseniya Rossii v 2018 godu" (Yearbook "Earthquakes of Russia in 2018")*, pages 71–81. Obninsk: GC RAS, 2020.

S. Senyukov, D. Droznin, S. Droznina, N. Shapiro, I. Nuzhdina, T. Kozhevnikova, O. Sobolevskaya, Z. Nazarova, A. Dolzhikova, S. Toloknova, and E. Karpenko. Catalog of earthquakes from the data of the KISS seismic network in 2015-2016. In *Proceedings of the Eighth All-Russian Scientific and Technical Conference: Problems of integrated geophysical monitoring of seismically active regions*, pages 216–2020, Petropavlovsk-Kamchatsky, Russia, 2021.

S. L. Senyukov. Monitoring and prediction of volcanic activity in Kamchatka from seismological data: 2000–2010. *Journal of Volcanology and Seismology*, 7(1):86–97, 2013. doi: 10.1134/S0742046313010077.

S. L. Senyukov, S. Y. Droznina, I. N. Nuzhdina, V. T. Garbuzova, and T. Y. Kozhevnikova. Studies in the activity of Klyuchevskoi volcano by remote sensing techniques between January 1, 2001 and July 31, 2005. *Journal of Volcanology and Seismology*, 3(3):191–199, 2009. doi: 10.1134/S0742046309030051.

L. Seydoux, 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):3972, 2020.

N. M. Shapiro, M. Campillo, L. Margerin, S. K. Singh, V. Kostoglodov, and J. Pacheco. The energy partitioning and the diffusive character of the seismic coda. *Bulletin of the Seismological Society of America*, 90(3):655–665, 06 2000. doi: 10.1785/0119990021.

N. M. Shapiro, D. V. Droznin, S. Y. Droznina, S. L. Senyukov, A. A. Gusev, and E. I. Gordeev. Deep and shallow long-period volcanic seismicity linked by fluid-pressure transfer. *Nat. Geosci.*, 10(6):442–445, 2017a. doi: 10.1038/ngeo2952.

N. M. Shapiro, C. Sens-Schonfelder, B. G. Luhr, M. Weber, I. Abkadyrov, E. I. Gordeev, I. Koulakov, A. Jakovlev, Y. A. Kugaenko, and V. A. Saltykov. Understanding Kamchatka's extraordinary: volcano cluster. *Eos*, 98(7):12–17, 2017b.

N. M. Shapiro, M. Campillo, E. Kaminski, J.-P. Vilotte, and C. Jaupart. Low-frequency earthquakes and pore pressure transients in subduction zones. *Geophysical Research Letters*, 45(20):11,083–11,094, 2018. doi: https://doi.org/10.1029/2018GL079893.

B. E. Shaw. Constant stress drop from small to great earthquakes in magnitude-area scaling. *Bulletin of the Seismological Society of America*, 99(2A):871–875, 2009.

D. R. Shelly, G. C. Beroza, and S. Ide. Non-volcanic tremor and low-frequency earthquake swarms. *Nature*, 446(7133):305, 2007.

D. R. Shelly, D. P. Hill, F. Massin, J. Farrell, R. B. Smith, and T. Taira. A fluid-driven earthquake swarm on the margin of the Yellowstone caldera. *Journal of Geophysical Research: Solid Earth*, 118(9):4872–4886, 2013.

Z. Shi and Y. Ben-Zion. Seismic radiation from tensile and shear point dislocations between similar and dissimilar solids. *Geophysical Journal International*, 179(1):444–458, 2009.

B. Sick, M. Guggenmos, and M. Joswig. Chances and limits of single-station seismic event clustering by unsupervised pattern recognition. *Geophysical Journal International*, 201(3):1801–1813, 2015.

R. J. Skoumal, M. R. Brudzinski, B. S. Currie, and J. Levy. Optimizing multi-station earthquake template matching through re-examination of the Youngstown, Ohio, sequence. *Earth and Planetary Science Letters*, 405:274–280, 2014.

F. Song, H. S. Kuleli, M. N. Toksöz, E. Ay, and H. Zhang. An improved method for hydrofracture-induced microseismic event detection and phase picking. *Geophysics*, 75(6):A47–A52, 2010.

R. Soto, F. Huenupan, P. Meza, M. Curilem, and L. Franco. Spectro-temporal features applied to the automatic classification of volcanic seismic events. *Journal of Volcanology and Geothermal Research*, 358:194–206, 2018.

R. Sparks. Forecasting volcanic eruptions. *Earth and Planetary Science Letters*, 210 (1):1–15, 2003. doi: https://doi.org/10.1016/S0012-821X(03)00124-9.

R. S. J. Sparks and S. R. Young. The eruption of Soufriere Hills Volcano, Montserrat (1995-1998): overview of scientific results. *Memoirs-Geological Society of London*, 21:45–70, 2002.

R. S. J. Sparks, C. Annen, J. D. Blundy, K. V. Cashman, A. C. Rust, and M. D. Jackson. Formation and dynamics of magma reservoirs. *Philosophical Transactions of the Royal society A*, 377(2139):20180019, 2019.

S. Stein and M. Wysession. *An introduction to seismology, earthquakes, and Earth structure*. John Wiley & Sons, 2009.

Y. Takei and M. Kumazawa. Why have the single force and torque been excluded from seismic source models? *Geophysical Journal International*, 118(1):20–30, 07 1994. doi: 10.1111/j.1365-246X.1994.tb04672.x.

C.-C. Tang, Z. Peng, K. Chao, C.-H. Chen, and C.-H. Lin. Detecting low-frequency earthquakes within non-volcanic tremor in southern Taiwan triggered by the 2005 Mw8.6 Nias earthquake. *Geophysical Research Letters*, 37(16), 2010.

W. A. Thelen, R. S. Matoza, and A. J. Hotovec-Ellis. Trends in volcano seismology: 2010 to 2020 and beyond. *Bulletin of Volcanology*, 84(3):26, 2022. doi: 10.1007/s00445-022-01530-2.

M. Titos, A. Bueno, L. García, M. C. Benítez, and J. Ibañez. Detection and classification of continuous volcano-seismic signals with recurrent neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(4):1936–1948, 2019. doi: 10.1109/TGRS.2018.2870202.

P. I. Tokarev. *Eruptions and seismic regime of volcanoes of the Klyuchevskoy volcanic group in 1949-1963 (in Russian)*. Nauka, Moscow, 1966.

A. M. Turing. I.—Computing machinery and intelligence. *Mind*, LIX(236):433–460, 10 1950. ISSN 0026-4423. doi: 10.1093/mind/LIX.236.433.

M. Ukawa and M. Ohtake. A monochromatic earthquake suggesting deep-seated magmatic activity beneath the Izu-Ooshima Volcano, Japan. *J. Geophys. Res. Solid Earth*, 92(B12):12649–12663, 1987. doi: 10.1029/JB092iB12p12649.

H. L. Van Tees. Detection, estimation and modulation theory. *Part I, NY: Wiley*, 1968.

V. N. Vapnik and A. Y. Chervonenkis. *Pattern recognition theory: statistical problems of learning (in Russian)*. Nauka,, 1974.

N. Varley, R. Arámbula-Mendoza, G. Reyes-Dávila, R. Sanderson, and J. Stevenson. Generation of vulcanian activity and long-period seismicity at volcán de colima, mexico. *Journal of Volcanology and Geothermal Research*, 198(1-2):45–56, 2010.

E. Venzke. Global volcanism program: Volcanoes of the world, v. 4.7. 2. *Smithsonian Institution, https://doi. org/10.5479/si. GVP. VOTW4-2013*, 2013.

A. G. Wech, W. A. Thelen, and A. M. Thomas. Deep long-period earthquakes generated by second boiling beneath Mauna Kea volcano. *Science*, 368(6492):775–779, 2020. doi: 10.1126/science.aba4798.

R. A. White, C. G. Newhall, and R. S. Punongbayan. Precursory deep long-period earthquakes at Mount Pinatubo: Spatio-temporal link to a basalt trigger. *Fire and mud: Eruptions and lahars of Mount Pinatubo, Philippines*, pages 307–328, 1996.

S. Wiemer and S. R. McNutt. Variations in the frequency-magnitude distribution with depth in two volcanic areas: Mount st. helens, washington, and mt. spurr, alaska. *Geophysical research letters*, 24(2):189–192, 1997.

G. Yogodzinski, J. Lees, T. Churikova, F. Dorendorf, G. Wöerner, and O. Volynets. Geochemical evidence for the melting of subducting oceanic lithosphere at plate edges. *Nature*, 409(6819):500–504, 2001. doi: 10.1038/35054039.

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

K. Yoshimoto, H. Sato, and M. Ohtake. Frequency-dependent attenuation of P and S waves in the Kanto area, Japan, based on the coda-normalization method. *Geophysical Journal International*, 114(1):165–174, 07 1993. doi: 10.1111/j.1365-246X.1993.tb01476.x.

Y. Yukutake, Y. Abe, and R. Doke. Deep low-frequency earthquakes beneath the Hakone volcano, Central Japan, and their relation to volcanic activity. *Geophysical Research Letters*, 46(20):11035–11043, 2019. doi: https://doi.org/10.1029/2019GL084357.

V. M. Zobin. *Introduction to Volcanic Seismology*. Elsevier, 2011.

List of Figures

List of Tables