

# Volcanic seismicity of Villarrica Volcano

Dissertation zur Erlangung des Doktorgrades der  
Mathematisch-Naturwissenschaftlichen Fakultät der  
Christian-Albrechts-Universität zu Kiel

vorgelegt von  
Johanna Marijke Lehr  
Kiel, 2021



---

Erster Gutacher: Prof. Dr. Wolfgang Rabbel  
Zweiter Gutachter: Prof. Dr. Matthias Hort  
Tag der mündlichen Prüfung: 01.07.2021



# Contents

<b>Eidesstattliche Erklärung</b>	<b>9</b>
<b>Acknowledgements</b>	<b>11</b>
<b>Summary</b>	<b>12</b>
<b>Zusammenfassung</b>	<b>14</b>
<b>I. Introduction</b>	<b>16</b>
<b>1. Relevance &amp; Aim of the thesis</b>	<b>17</b>
1.1. Motivation . . . . .	17
1.2. Villarrica - a well studied yet little understood volcano . . . . .	19
1.3. Aim of the thesis . . . . .	24
1.4. Structure and contribution . . . . .	25
<b>2. Background</b>	<b>27</b>
2.1. Volcano seismology . . . . .	27
2.2. Villarrica volcano . . . . .	32
2.2.1. Geology . . . . .	32
2.2.2. Volcanology . . . . .	33
<b>3. Data</b>	<b>35</b>
3.1. Seismic network . . . . .	35
3.2. Seismic activity . . . . .	37
3.3. Volcanic activity . . . . .	42
3.4. Nomenclature . . . . .	42
<b>II. Publications &amp; manuscripts</b>	<b>44</b>
<b>4. Source location &amp; seismic velocities</b>	<b>45</b>
4.1. Introduction . . . . .	46
4.2. Source location by probabilistic slowness analysis (“array analysis”) . . . . .	48
4.2.1. Beamforming . . . . .	48

4.2.2.	Epicenter location . . . . .	51
4.2.3.	DOA determination . . . . .	53
4.2.4.	S-wave velocity profiles from phase velocity dispersion . . . . .	55
4.3.	Source location by amplitude decay analysis (“amplitude source location”) . . . . .	56
4.4.	Results . . . . .	61
4.4.1.	Location . . . . .	61
4.4.2.	Velocity profiles . . . . .	64
4.5.	Discussion . . . . .	67
4.5.1.	Source location . . . . .	67
4.5.2.	Attenuation and scattering . . . . .	68
4.5.3.	Velocity structure . . . . .	69
4.5.4.	Implications on the source process? . . . . .	70
4.6.	Conclusion . . . . .	71
4.7.	Appendix . . . . .	72
4.7.1.	Probabilistic intersection . . . . .	72
4.7.2.	Site amplification estimation . . . . .	73
<b>5.</b>	<b>Magnitudes &amp; interevent times of transients</b>	<b>78</b>
5.1.	Introduction . . . . .	80
5.2.	Flow regimes and statistical distributions of bubble sizes . . . . .	81
5.3.	Event detection . . . . .	83
5.3.1.	STA/LTA network trigger . . . . .	83
5.3.2.	Amplitude and magnitude . . . . .	84
5.4.	Magnitude of completeness . . . . .	84
5.4.1.	Monte-Carlo simulation of trigger . . . . .	85
5.4.2.	Sections of similar amplitude . . . . .	89
5.5.	Statistical analysis of event catalog . . . . .	89
5.5.1.	Time series . . . . .	89
5.5.2.	Magnitudes . . . . .	91
5.5.3.	Interevent times . . . . .	95
5.5.4.	Interdependence of magnitude and interevent times . . . . .	97
5.6.	Discussion . . . . .	100
5.6.1.	Magnitudes . . . . .	100
5.6.2.	Interevent times . . . . .	100
5.6.3.	Interdependence . . . . .	102
5.6.4.	Flow regime . . . . .	102
5.7.	Conclusion . . . . .	103
<b>6.</b>	<b>Waveform families of transient events</b>	<b>105</b>
6.1.	Introduction . . . . .	106

6.2. Identification of waveform families . . . . .	106
6.2.1. Cluster analysis . . . . .	106
6.2.2. Template matching . . . . .	112
6.3. Centroid waveforms of families . . . . .	112
6.4. Analysis of individual families . . . . .	116
6.4.1. Spectral properties and waveforms . . . . .	116
6.4.2. Lag between stations . . . . .	116
6.4.3. Source location . . . . .	117
6.5. Results . . . . .	118
6.5.1. Waveforms and spectral properties . . . . .	118
6.5.2. Relative travel times . . . . .	118
6.5.3. Location . . . . .	127
6.6. Discussion . . . . .	128
6.7. Conclusion . . . . .	134
<b>III. Concluding remarks</b>	<b>136</b>
<b>7. Overarching discussion &amp; conclusions</b>	<b>137</b>
<b>8. Recommendations for future seismic installations</b>	<b>144</b>
<b>9. Outlook</b>	<b>146</b>
<b>Bibliography</b>	<b>148</b>
<b>A. Supporting Information Chapter 4</b>	<b>173</b>
<b>B. Supporting Information Chapter 5</b>	<b>187</b>
<b>C. Independent component analysis</b>	<b>197</b>
<b>CV</b>	<b>208</b>





# Eidesstattliche Erklärung

Hiermit erkläre ich durch meine Unterschrift an Eides statt:

1. Die eingereichte Dissertation mit dem Titel "Volcanic seismicity of Villarrica Volcano" habe ich selbstständig und ohne unzulässige fremde Hilfe verfasst. Hierbei habe ich weder Textstellen von Dritten oder aus eigenen Prüfungsarbeiten, noch Grafiken oder sonstige Materialien ohne Kennzeichnung übernommen.
2. Es sind ausschließlich die von mir angegebenen Quellen und Hilfsmittel verwendet worden.
3. Sämtliche wörtliche und nicht wörtliche Zitate aus anderen Werken sind gemäß den wissenschaftlichen Zitierregeln kenntlich gemacht.
4. Die von mir vorgelegte Arbeit ist bisher noch in keiner Form als Bestandteil einer Prüfungs-/Qualifikationsleistung vorgelegt worden.
5. Die von mir eingereichte Dissertation habe ich unter Beachtung der Grundsätze zur Sicherung guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft erstellt.
6. Mir wurde noch nie ein akademischer Grad entzogen.
7. Ich versichere an Eides Statt, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.



# Acknowledgements

This work could not have been accomplished without the help and support of many people.

I would like to thank my supervisor Wolfgang Rabbel for his skilled insights and for his financial support.

I also would like to recognize the scholarship that was awarded by Kiel University and which allowed me to start this project in the first place.

I thank my colleagues for their insights, comments and support. In particular, I want to thank the Freundeskreis Litmod for making the home office during the Corona pandemic a little less isolated.

I am very grateful to Felix Gross for his kind offer to take a look at my thesis in the final stage, for encouraging me to finish and helping me to develop a structure.

I would like to thank my friends and family for their moral support. In particular, I have to thank my Dad for pushing me into finally writing up and developing a concept and structure for the thesis.

Finally, I am deeply grateful for my partner Wolfgang who supported me unconditionally through all the ups and downs of this time and who did a final proof reading of the work.

# Summary

This thesis analyzes the nature and origin of the low-frequent seismicity of Villarrica Volcano in Chile. Villarrica is an active volcano with an intermittent lava lake in its central vent. The typical volcanic activity consists of persistent degassing and usually mild Strombolian explosions. The activity is accompanied by a notorious seismic tremor overlain by swarms of transient events in approximately 1-min intervals. Both types of signals display frequencies between 0.5 and 5 Hz. They are usually attributed to the volcanic activity with the transients being interpreted as the bursting of slugs at the surface of the lava lake (Strombolian explosions). However, except for one study, no attempts have been made to locate the source of the signals using seismic methods. Moreover, reports on whether the seismic transients are related to slug bursting have been ambiguous.

The base for this work are about 12 days of seismic data recorded by a dense local seismic network in early March 2012. Despite the age of the data, the results are most likely applicable to today's conditions since Villarrica is generally a stable system. Even the latest eruption phase in 2015 did not cause substantial changes in the behavior of the volcano.

In the first part of this work, three different location methods - one using array techniques and two exploiting the decay of amplitude with distance - were applied to the tremor signal including the transients. All three approaches conclusively located the origin in the summit crater area, which is in agreement with earlier studies. Each of the methods however had a considerable uncertainty, which on the one hand demonstrated the difficulties and limitations associated with location techniques of seismic signals in a complex medium. On the other hand, it emphasized the benefit of combining different approaches for confirmation.

The array analysis showed that the wave propagation velocity varied with the frequency of the signal. A dispersion curve could be derived, which was inverted for S wave velocities. These results provide the first models of the near-surface seismic structure of the volcanic edifice at two sites 4-5 km north and west, respectively, of the summit.

A second focus was on the nature of the transient events. An event catalog was compiled using an automatic detection method and the occurrences of magnitudes and waiting times between two events (interevent time) were analyzed statistically. The interevent times were approximately log-normally distributed. Using a stochastic simulation of the noise conditions in the network, it could be shown that the number of events reduced towards low magnitudes after reaching a peak, resulting in an approxi-

---

mately log-normal distribution of amplitudes. In contrast to most common methods in seismology to find the magnitude of completeness of a catalog, the approach here allowed to estimate the number of undetected events without presumptions on the distribution of the magnitudes. Log-normal distributions have also been found for the sizes and spacings of gaseous slugs rising in a liquid-filled pipe. Hence, the estimated distributions of amplitudes and interevent times strongly support the interpretation of the transients as signatures of gaseous slugs and suggest a slug flow regime in the upper magma column..

In addition, at least 67 families of similar waveforms were found in the population of transients, the two largest of which comprised several thousand events. Both families displayed a less than 5s long, almost Ricker-wavelet-like waveform at the crater, but differed slightly in frequency content. Composite waveforms were generated for each family by stacking the events that showed the highest similarity. Two findings resulted from correlating the onsets with a reference station which complemented the previous results: 1) almost identical relative lag times suggest a similar source region of all families 2) velocity variations around the edifice could be identified. The origin was located in the crater area by fitting the spatial decay of the amplitudes but the results were too ambiguous to obtain a precise location.

# Zusammenfassung

Diese Arbeit liefert eine grundlegende Analyse bezüglich der Art und des Ursprungs der niederfrequenten Seismizität des Vulkans Villarrica in Chile. Villarrica ist ein aktiver Vulkan mit einem zeitweiligen Lavasee in seinem zentralen Schlot. Die typische vulkanische Aktivität besteht aus anhaltender Entgasung und meist leichten strombolianischen Explosionen. Die vulkanische Aktivität wird von einem seismischen Tremor begleitet, der von Schwärmen transienter Ereignisse in etwa 1-minütigen Intervallen überlagert ist. Beide Arten von Signalen weisen Frequenzen zwischen 0,5 und 5 Hz auf. Sie werden üblicherweise der vulkanischen Aktivität zugeschrieben, wobei die Transienten als das Zerplatzen von Gasblasen an der Oberfläche des Lavasees interpretiert werden (strombolianische Explosionen). Mit Ausnahme einer Studie wurden jedoch keine Versuche unternommen, die Quelle der Signale mit seismischen Methoden zu lokalisieren. Darüber hinaus ist unklar, ob die seismischen Transienten direkt mit dem Ausbruch von Slugs zusammenhängen.

Die Grundlage für diese Arbeit sind etwa 12 Tage seismischer Daten, die von einem dichten lokalen seismischen Netzwerk Anfang März 2012 aufgezeichnet wurden. Obwohl dies schon einige Zeit zurückliegt, sind die vorgestellten Ergebnisse höchstwahrscheinlich auch auf die heutigen Bedingungen anwendbar, da Villarrica ein sehr stabiles System ist. Selbst die jüngste Eruptionsphase im Jahr 2015 hat keine wesentlichen Änderungen im Verhalten des Vulkans bewirkt.

Im ersten Teil dieser Arbeit werden drei verschiedene Ortungsmethoden - eine unter Verwendung von Array-Techniken und zwei unter Ausnutzung des Abklingens der Amplitude mit der Entfernung - auf das Tremorsignal einschließlich der Transienten angewendet. Alle drei Ansätze lokalisierten den Ursprung eindeutig im Bereich des Gipfelkraters und bestätigten damit die Vermutungen aus früheren Studien. Jede der Methoden ist jedoch mit einer erheblichen Unsicherheit behaftet, was einerseits die Schwierigkeiten und Grenzen aufzeigt, die mit Ortungstechniken von seismischen Signalen in einem komplexen Medium verbunden sind. Andererseits betont es den Vorteil, mehrere Methoden zur gegenseitigen Bestätigung zu verwenden.

Aus der Array-Analyse wurden oberflächennahe Geschwindigkeitsmodelle an zwei Standorten 4-5 km nördlich bzw. westlich des Gipfels bestimmt. Die Langsamkeit (der Kehrwert der scheinbaren Wellenausbreitungsgeschwindigkeit) variierte deutlich mit der Frequenz des Signals. Daraus konnte eine Dispersionskurve abgeleitet werden, die für S-Wellen-Geschwindigkeiten invertiert wurde. Diese Ergebnisse liefern erste Hinweise auf die oberflächennahe seismische Struktur des Vulkangebäudes.

---

Ein zweiter Schwerpunkt lag auf der Natur der transienten Ereignisse. Mit Hilfe einer automatischen Erkennungsmethode wurde ein Ereigniskatalog erstellt und das Auftreten von Magnituden und Wartezeiten zwischen zwei Ereignissen (interevent time) statistisch ausgewertet. Die Wartezeiten waren annähernd lognormalverteilt. Mit Hilfe einer stochastischen Simulation des Hintergrundrauschens im Netzwerk konnte gezeigt werden, dass die Anzahl der Ereignisse nach Erreichen eines Peaks zu niedrigen Magnituden hin abnimmt, was zu einer annähernd lognormalen Verteilung der Amplituden führt. Im Gegensatz zu den in der Seismologie üblichen Methoden, die Vollständigkeitsmagnitudes eines Katalogs zu finden, erlaubt der hier verwendete Ansatz, die Anzahl der unentdeckten Ereignisse abzuschätzen, ohne Annahmen über die Verteilung der Magnituden zu treffen. Lognormal-Verteilungen werden auch für die Größen und Abstände von Gasblasen gefunden, die in einem flüssigkeitsgefüllten Rohr aufsteigen. Daher unterstützen die gefundenen Verteilungen der Amplituden und der Zwischenereigniszeiten die Interpretation der Transienten als Signaturen gasförmiger Slugs und deuten auf ein Slug-Flow-Regime in der oberen Magmasäule hin.

Darüber hinaus wurden mindestens 67 Familien ähnlicher Wellenformen in der Population der Transienten gefunden, von denen die beiden größten mehrere tausend Ereignisse umfassten. Beide Familien zeigten eine weniger als 5 s lange, fast Ricker-Wavelet-ähnliche Wellenform am Krater, unterschieden sich aber leicht im Frequenzgehalt. Für jede Familie wurden zusammengesetzte Wellenformen erzeugt, indem die Ereignisse mit der größten Ähnlichkeit gestapelt wurden. Aus der Korrelation der Ersteinsätze mit einer Referenzstation ergaben sich zwei Erkenntnisse, die die vorherigen Ergebnisse ergänzten: 1) fast identische relative Laufzeiten lassen auf eine ähnliche Quellregion aller Familien schließen 2) Geschwindigkeitsvariationen um das Vulkangebäude herum konnten identifiziert werden. Der Ursprung der Signale wurde durch Anpassung des räumlichen Abklingens der Amplituden im Kraterbereich lokalisiert, aber die Ergebnisse waren zu mehrdeutig, um eine genaue Lokalisierung zu erhalten.

Part I.  
Introduction



# 1. Relevance & Aim of the thesis

## 1.1. Motivation

Without doubt, volcanoes cause some of the most fascinating natural spectacles on this planet (Fig. 1.1). For geoscientists, they also provide a unique window into its interior. Unfortunately, they are also a significant natural hazard [Schmincke, 2000, Tilling, 2005, Sheets, 2015, Stovall et al., 2019]. Luckily, even though eruptions can not really be prevented by humans nor their force be controlled, they can nonetheless be reasonably well predicted, so that the loss of infrastructure and - above all - lives can be minimized [Gregg et al., 2015]. Geophysical and geochemical methods help us to detect the precursory signals like subtle earthquakes or changes in gas composition before the volcano emits a new lava flow or tons of ash [Caracausi et al., 2003, Gambino et al., 2014, Carn, 2015, Freymueller et al., 2015, McNutt and Roman, 2015, Sigmundsson et al., 2018]. However, for an effective hazard monitoring, good knowledge of the volcano is required and an effective instrumentation is crucial to provide useful data [Pallister and McNutt, 2015]. New developments and improvements of existing techniques facilitate multi-parameter monitoring. The ever growing fleet of satellites adds remote sensing data which is particularly useful to monitor very remote locations [Massonnet and Sigmundsson, 2000, Wright et al., 2004, Delgado et al., 2017, Ebmeier et al., 2018, Pritchard et al., 2018, Coppola et al., 2020].

Yet, although more data indubitably offer the possibility of more accurate predictions and a better understanding, there are (at least) two things to consider: First, the richest data set is worth nothing if the data are not effectively analyzed and boiled down into relevant information that help authorities to take actions. Therefore we need methodologies and strategies how to analyze and manage the data streams from the many available sources. Secondly, monitoring volcanoes (and natural hazards in general) places a significant financial burden on society [Chester, 2005]. Therefore we need affordable solutions that provide maximal information at low cost.

Seismology provides very efficient tools to monitor volcanic activity [McNutt et al., 2015]. Even before the advent of seismometers, it was well known to the local populations that major eruptions were (among other things of course) heralded by a perceivable seismic unrest and increased numbers of earthquakes [Cronin and Cashman, 2008, Zobin, 2017d]. Before the actual eruption, the magma has to form its path to the surface by creating new fractures or reviving old ones. These brittle failures in the Earth's crust produce small earthquakes [Stovall et al., 2019]. In addition, a number of the dynamical



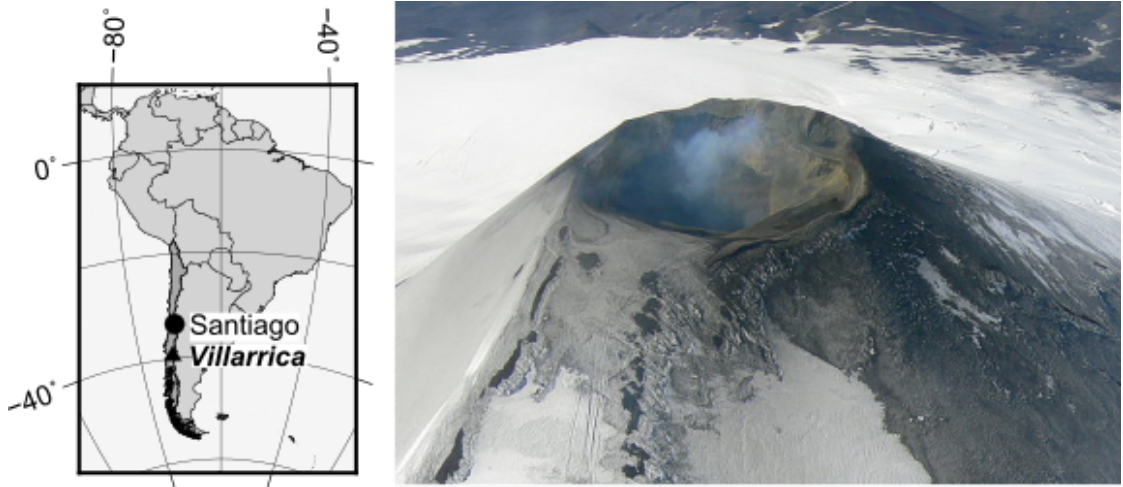
**Figure 1.1.:** a) Eruption of Kilauea in 2018 [US Geological Survey, Public Domain, 2020b] b) Plume of ash and gas reflecting the glow from the lava of Kilauea’s summit crater in 2020 [Hawaii Volcanoes National Park, Public Domain, 2021] c) Ash cloud by icelandic Eyjafjallajökull in 2010 [David Karnå, CC BY 3.0, 2020] d) Lava flow from Kilauea Volcano destroying a street in Hawaii [US Geological Survey, Public Domain, 2020a] e) House destroyed by a lava flow from Mt. Etna [Official U.S. Navy Page, 2021] f) Abandoned Plymouth, Montserrat, West Indies, 20 years after being destroyed by lahars and pyroclastic flows in 1997 [Andrew Shiva / Wikipedia / CC BY-SA 4.0, 2007]

processes inside the volcano cause distinct seismic signals, that can help to decipher the ongoing activity [Métaxian et al., 1997, Chouet et al., 1998, Kawakatsu and Yamamoto, 2007, Zobin, 2011, Wassermann, 2012, Chouet and Matoza, 2013, McNutt and Roman, 2015]. For example, a gas bubble passing through narrowings or widenings in the plumbing system can cause significant reverberations that couple into the ground as seismic waves [James et al., 2006, O'Brien and Bean, 2008]. Flowing or convecting magma is known to produce sustained tremor signals. Explosions may be too small to be seen from save distance but can still be detected seismically as can surface activities like lahars, pyroclastic density currents or lava flows. Beside identifying ongoing processes in the volcano, seismology also provides means to image the interior of the edifice - either by tomographic methods that reveal the structure in terms of seismic parameters like wave propagation velocity [Brennguier et al., 2007, Yeguas et al., 2011, Prudencio et al., 2013, Mora-Stock, 2015, Zuccarello et al., 2016, García-Yeguas et al., 2017] or by identifying source locations and mechanisms that indicate active features of the plumbing system [Lahr et al., 1994, Kawakatsu et al., 2000, Almendros et al., 2001a, 2002, Patanè et al., 2008, Waite et al., 2008, De Barros et al., 2011, Eyre et al., 2013, Chouet and Matoza, 2013, Sigmundsson et al., 2014]. Understanding the structure of the volcano is particularly important for modeling the propagation of seismic waves in a structurally complex medium, which in turn is crucial for deciphering the seismic signals. For example, the so-called "long-period" events were explained for a long time as a resonance phenomenon [Chouet, 1996]. However O'Brien and Bean [2009] demonstrated by numerical modeling, that similar waveforms can be produced solely by scattering and path effects in a complicated medium.

Similarly, the different classes of signals may look very different at any two volcanoes owing to their individual composition and structure. In order to understand and reliably predict a volcano's behavior, it is - and probably always will be - necessary to study each volcano individually. Nevertheless, these analysis advance our general understanding not only of volcanoes but also of the dynamic process in the Earth.

## 1.2. Villarrica - a well studied yet little understood volcano

Villarrica Volcano is a basaltic to basaltic-andesitic, glacier-covered strato-volcano in South Central Chile (Fig. 1.2). Due to its high level of activity, its hazard potential and its intermittent lava lake, Villarrica is among the more frequently investigated volcanoes in South America. At least 30 previous studies address in about equal parts the stratigraphic investigation of past and prehistoric eruptions, the present-day degassing activity, and the seismic and infrasonic signals that accompany the volcanic activity. Nevertheless, most of these studies stand on their own and introduce their own models rather than building upon each other. As a consequence, an overarching concept of what drives the volcanic activity at Villarrica and how the dynamics are expressed in monitoring signals

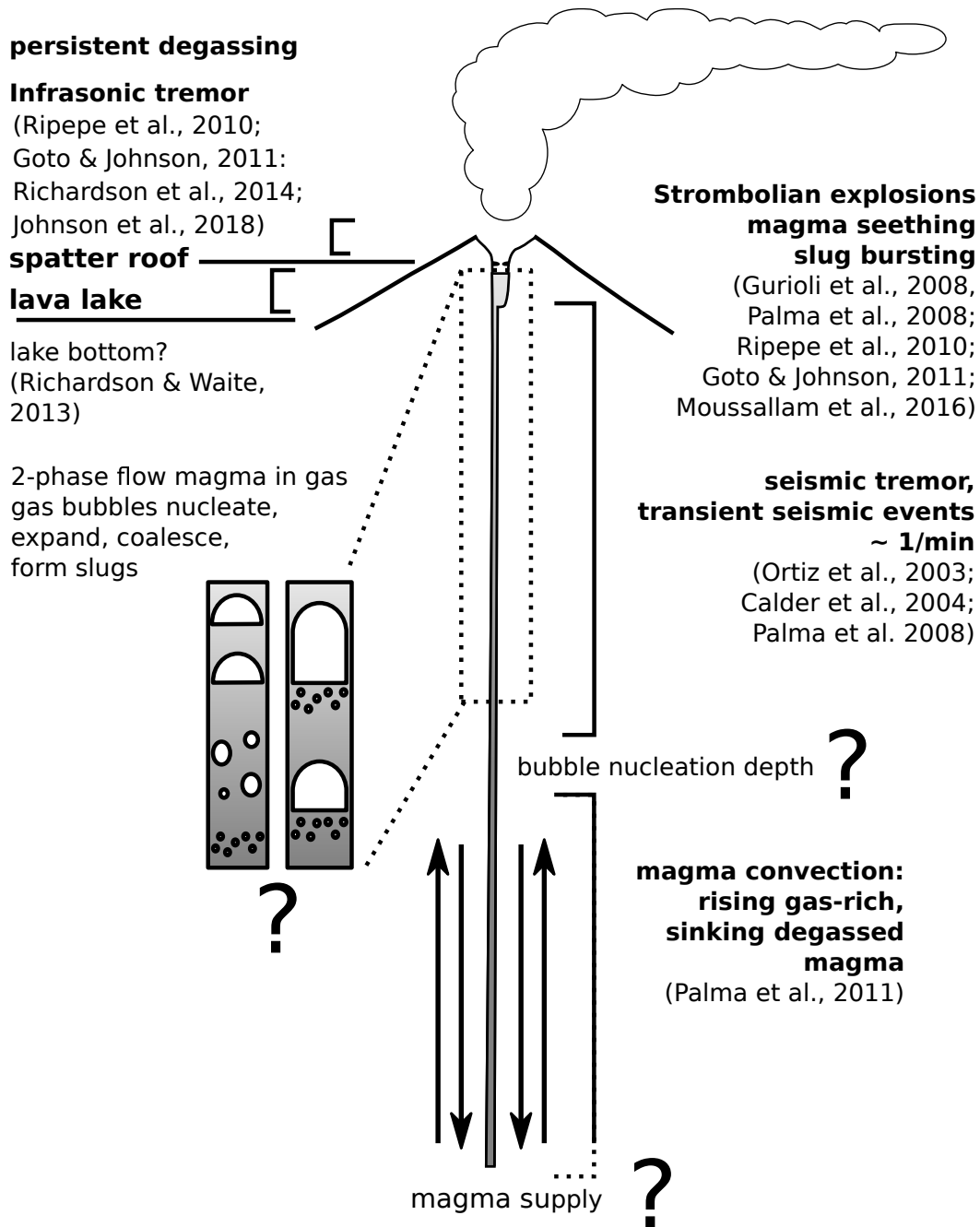


**Figure 1.2.:** Location of Villarrica Volcano in Chile (South America) and aerial view of the summit crater area [ Sarah and Iain, CC BY 2.0, 2007]

- such as seismicity, infrasound or gas flux - is still lacking.

Palma et al. [2011] depicted the degassing and magma dynamic model of Villarrica based on an open-vent basaltic, i.e. low-viscosity, system (Fig.1.3): A long narrow conduit connects a magma reservoir at depth with the outlet at the summit crater. It was suggested that the conduit is nearly vertical, at least 2000 m long [Palma et al., 2008] and has a radius of 5.8 m [Palma et al., 2011]. The enormous degassing rates of 15 kg/s  $\text{SO}_2$  and above [Palma et al., 2011, Bredemeyer and Hansteen, 2014] despite little magma ejection are likely achieved by an efficient overturn of rising gas-rich magma and sinking degassed magma [Palma et al., 2008, 2011, Moussallam et al., 2016]. This two-phase liquid-in-liquid convection takes place up to a depth at which the pressure is low enough for the gas to leave the magma and bubbles start to nucleate. The details of the proposed two-fluid convection process were elaborated by Palma et al. [2011] based on analog modeling experiments and measurements of degassing rates.

The fluid dynamic regime above this bubble nucleation depth is largely unknown. Presumably, there is a two-phase flow regime, consisting of gas bubbles or slugs in liquid magma. This would agree with the observation that the infrasonic and seismic tremor signal resemble the signals of bubbly gas flow in technical applications [Ripepe et al., 2010]. At times when the lava lake is visible, the wide range of bubble bursting styles also suggests a vivid convection [Moussallam et al., 2016]. The activity ranges from a gentle rollover and the bursting of small bubbles - reminiscent of boiling water - via larger and more bubbles to small lava fountains and more violent Strombolian explosions [Palma et al., 2008, Gurioli et al., 2008, Moussallam et al., 2016]. Similar to Palma et al. [2011], Moussallam et al. [2016] argue for a rapid, turbulent convection of degassed



**Figure 1.3.:** Schematic illustration of the conduit system of Villarrica. The observed geophysical and geochemical phenomena might be linked to a number of fluid dynamic processes.

gas-rich magma based on their video footage of the lake surface and the lack of periodic variations in gas flux and composition. The depth of the the lava lake varies in irregular cycles of days to weeks between 100 m and 250 m [Richardson et al., 2014]. Higher levels are associated with increased seismic activity. Richardson et al. [2014] suggested that this indicates the arrival of a batch of new, gas-rich magma at depth.

Its persistent infrasonic tremor made Villarrica an interesting target for infrasonic studies. A first explanation for the nature of this signal was offered by Ripepe et al. [2010] who attributed it to the gas flux through the magma. In this model, higher surficial gas velocities (i.e. an increase in gas flux) would cause a higher tremor frequency. Later, contrasting models were proposed in which the upper conduit and crater opening act as acoustic resonator. Goto and Johnson [2011] invoked a Helmholtz resonator formed by the volume between the lake surface and the spatter roof (basically a temporary lid above the free magma surface that forms from ejected and cooled sheets of magma) whereas Richardson et al. [2014] proposed that the funnel-shaped crater and vent form a Bessel horn with the lake surface as lower boundary. In both cases, a change in tremor frequency would simply be achieved by a varying lake depth which essentially changes the geometry of the resonator. Richardson et al. [2014] demonstrated that the lake depths predicted from the frequency of the infrasonic tremor using the Bessel horn model were in good agreement with those obtained independently from the delay between seismic and acoustic signals. Johnson et al. [2018] later generalized the concept by modeling the resonator and the acoustic damping numerically.

The use of seismic data at Villarrica has largely been limited to the descriptive analysis of the tremor level and the number of events in conjunction with the eruptive activity. Its seismicity is commonly described as a notorious tremor signal between 0.5 and 3 Hz which is sometimes overlain by transient events [Ortiz et al., 2003, Calder et al., 2004, Palma et al., 2008, Gurioli et al., 2008, Ripepe et al., 2010, Goto and Johnson, 2011, Richardson et al., 2014, Mora-Stock, 2015, Lehr et al., 2019]. These transients are commonly attributed to Strombolian explosions due to correlation with visual observations [Palma et al., 2008, Gurioli et al., 2008] and occur approximately once per minute. Palma et al. [2008] and Gurioli et al. [2008] provided detailed accounts of the visible activity of the lava lake including classification schemes. However, although a strong association with seismic signals was noted, they did not elaborate whether the classifications could be extended to seismic (or infrasonic) waveforms - that is, whether each of the activity classes was associated with a specific waveform. According to Gurioli et al. [2008], the waveforms resembled those observed during gas bursting events (puffing) at Stromboli. Some events however lacked a seismic or acoustic signature. From the strong coupling of infrasonic and seismic waveforms Gurioli et al. [2008] inferred a shallow source depth.

Richardson and Waite [2013] provided the first location and inversion for the source mechanism of one prominent repetitive long-period transient waveform. They found a shallow source with a strong horizontal east-west component. The result was interpreted as a magma dragging at a hypothetical bottom of the lava lake while filling the void that

was left by the release of a gas slug. In addition, Richardson and Waite [2013] proposed that the tremor could result from a superposition of closely spaced transient events.

Mora-Stock [2015] presented the first image of the seismic structure of the edifice, which was obtained by travel-time tomography using a swarm of volcano-tectonic earthquakes. Her work was based on the same data set as the present thesis but used exclusively volcano-tectonic events. Her results suggest a deep magma source slightly east of the current summit which would imply a somewhat oblique feeding system of the lava lake. This is in contrast to the vertical conduit proposed earlier by Palma et al. [2008] although this was merely based on the consideration that the conduit must allow efficient, unhindered convection.

The interplay between degassing activity and magma dynamics as well as seismic and infrasonic signals remains unclear. Richardson and Waite [2013] located a single, persistent waveform but most of the transients and tremor signals have not been located yet. Similarly the natures of the seismic tremor and transients as well as their relation are still unclear. Without confirmation by visual observations of the volcanic activity, it is difficult to relate the transients unambiguously to explosions based on the seismic waveforms. However, long-term visual observations are difficult to come by. Therefore, it might generally be desirable to be able to relate seismic signals to degassing activity without needing visual observations for confirmation. The study by Mora-Stock [2015] provided a first model of the deeper velocity structure of the volcanic edifice. However studies on e.g. the source mechanism of seismic signals or models of the wave propagation might greatly benefit from information of the near-surface seismic velocities. Finally, if the seismic transients are indeed signatures of bubble burstings, their analysis might help to elucidate the gas and magma dynamics in the upper conduit.

In summary, the following questions arose from the literature and in view of the available data:

- Where do the tremor and transients originate?
- How are the tremor and the transients related? Is the tremor just a superposition of closely spaced transients or is it a different signal? Could it be magma convection in the conduit?
- What is the nature of the transients? Are they indeed bursting gas bubbles or could they have a tectonic nature? What is the driving process?
- Are the waveforms of the transients repetitive or are they unique?
- What is the seismic structure of the edifice? How does it affect wave propagation and waveforms?
- How does the magma convection take place in the conduit? And what signals are associated with that process?

The present thesis attempts to tackle some of these aspects. The database for this work was collected in early March 2012 by Rabbel and Thorwart [2019] using an extensive local seismic network of 75 stations. The network consisted of a perimeter of stations around the base of the volcano, three stations at the rim of the summit crater and three mini-arrays. The data were acquired using geophones instead of seismometers. The former are considerably cheaper and easier to deploy albeit at the cost of reduced sensitivity. Nevertheless, the results presented here demonstrate that a substantial amount of information on the volcanic structure and dynamics can be gained from such a low-cost solution.

### 1.3. Aim of the thesis

Out of the many open questions regarding the seismicity of Villarrica Volcano, this thesis concentrates on three major objectives:

1. Locating the source of the seismic tremor
2. Elucidating the shallow seismic velocity structure of the edifice
3. Characterizing the occurrence of the transient events

The first objective lays the base for the more detailed analysis of the seismicity. Both the tremor as well as the typical transient events associated with the volcanic activity lack clear onsets which precludes the use of classic earthquake location techniques. Moreover these methods require a robust knowledge of the seismic velocity structure which was not available for Villarrica. Numerous alternative approaches have been developed in the last decades to locate volcanic seismic signals. Two methods were particularly suited for the network layout: Beamforming - which makes use of seismic arrays to determine the direction of arrival of a signal, and the amplitude location method in which the decay of the seismic amplitude with distance is traced back to the source of the signal. These two independent location techniques were applied to corroborate the origin of the signals.

The second objective developed as a byproduct of the array analysis. Since the apparent velocities clearly depended on frequency, dispersion curves of Rayleigh surface waves could be determined and inverted for S-wave velocities. This provided the first indications of the near-surface velocity structure of Villarrica so far.

Finally, the transient events, which make up a considerable part of the seismicity, are investigated more closely. A catalog of automatically detected events was compiled and its magnitude of completeness determined. The statistical distribution of magnitudes and interevent times help to distinguish the events from tectonic seismicity. Instead, they could be related to eruption dynamics at other volcanoes and slug flow processes in two-phase flow.



Complementary findings regarding all three objectives were obtained by identifying families of transient events. Each family consisted of similar waveforms, which allowed to obtain stacked waveforms which were clearer and less noisy. The stacked waveforms could be used to locate the source position of the transient events. Differences in travel times of the onsets between stations helped to constrain the source region and revealed differences in the wave propagation velocity towards the different stations. The clearer waveforms also facilitated comparison with examples from the literature and thus an interpretation of the waveform.

## 1.4. Structure of the thesis and contributions by others

Considerable parts of this thesis were taken from one published scientific article and one manuscripts written by the thesis's author as first author. Both are based on the same set of seismic data and share similar introductory parts on Villarrica. To maintain the logic structure of the present work and avoid repetition, these sections were taken out from the original works and combined into Section 2.2 on the geology and seismicity of Villarrica in general and Chapter 3 on the seismic data base and the seismic and volcanic activity specifically during the campaign.

Since each topic is largely self-contained, the chapters each follow the conventional scientific structure of introduction - methods - results - discussion.

The origin of the seismicity, notably the tremor (Aim 1), and the velocity model (Aim 2) are subject of Chapter 4 and were published in:

*Lehr, J.; Eckel, F.; Thorwart, M. & Rabbel, W.*

***Low-Frequency Seismicity at Villarrica Volcano: Source Location and Seismic Velocities,***

*Journal of Geophysical Research: Solid Earth, American Geophysical Union (AGU), 2019 , 124 , 11505-11530*

The text was written by Johanna Lehr, who also performed the beamforming analysis and velocity modeling. The amplitude location method was based on results from the Bachelor thesis of Felix Eckel. Martin Thorwart and Wolfgang Rabbel served in advisory capacity. Since the modeling of the velocity structure emanated directly from the location procedure it was included in the paper.

The recurrence of the seismic transients (Aim 3) is analyzed in Chapter 5. It is essentially manuscript which is in reply at the time of writing the thesis:

*Lehr, J, & Rabbel, W.*

*Statistical aspects of volcanic seismicity at Villarrica Volcano in March 2012, submitted to Geophysical Research: Solid Earth, American Geophysical Union (AGU), In Reply*

Text and all results were provided by Johanna Lehr. Wolfgang Rabbel supervised the work.

An additional analysis of the transient events is presented in Chapter 6. It addresses the similarity of the transient waveforms and thus their nature. The text is structured in a similar manner as a scientific article but is not yet submitted for publication. Analysis, text and concept are by Johanna Lehr.

All findings are reviewed together in an overall conclusion and outlook in Chapter III. The chapter also discusses indications to the remaining questions even though they were not specifically addressed by the objectives.

With the two manuscripts, documents with supporting information were additionally submitted. These are appended here as Appendices A and B. Moreover, Appendix C presents an additional analysis, in which it was tried to separate the transients and tremor using Independent component analysis.

## 2. Background

### 2.1. Volcano seismology

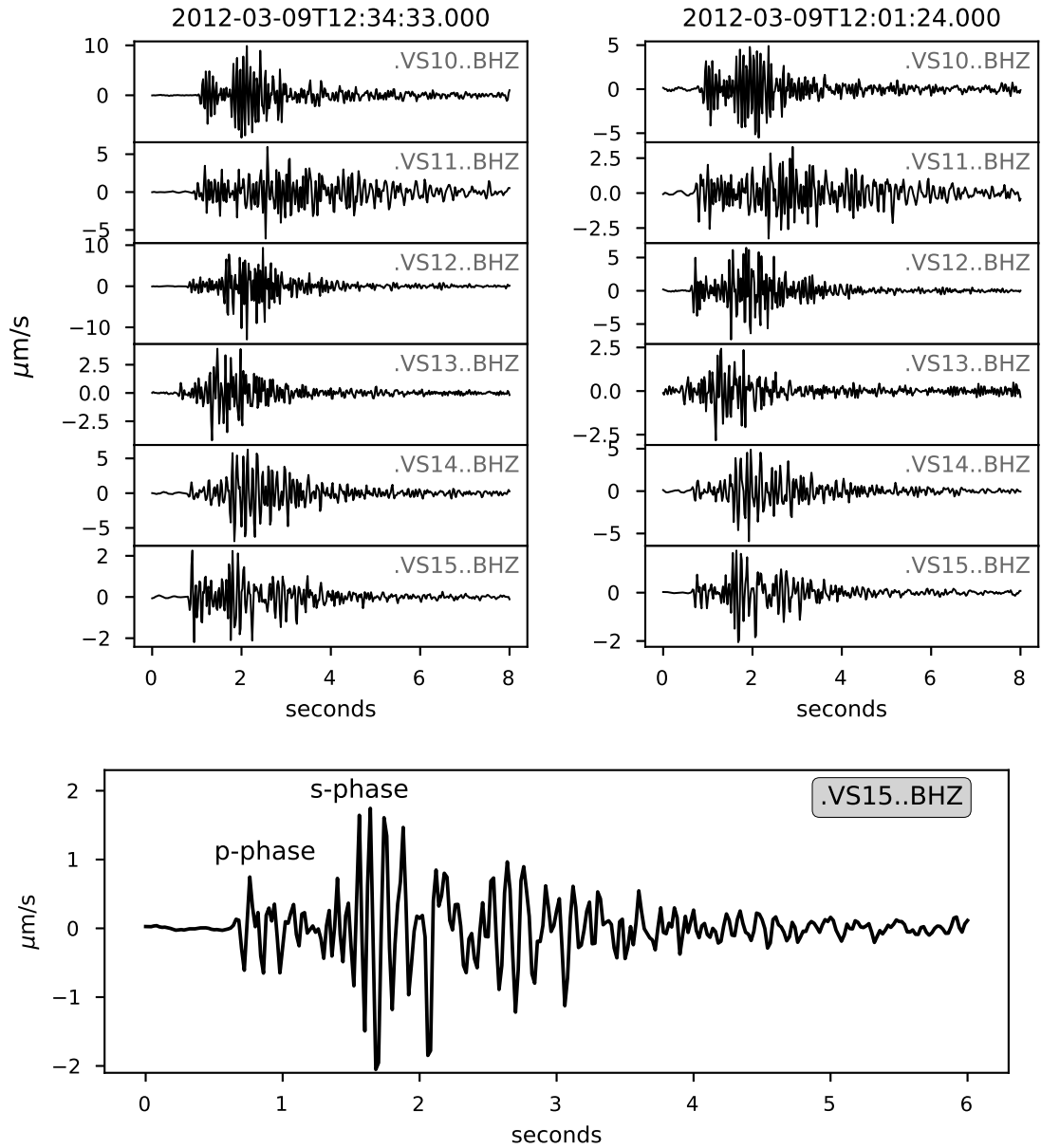
Seismology is a powerful tool to both monitor volcanic activity and image the interior of volcanoes. The fundamental difference to classic tectonic seismology is that volcanoes produce a number of different seismic signals besides the familiar brittle failure of normal tectonic earthquakes. Good overviews on the peculiarities of volcano seismology were given e.g. by Kawakatsu and Yamamoto [2007], Zobin [2011], Wassermann [2012] and Chouet and Matoza [2013]. The location of the seismic sources indicate key features of the volcanic system while the waveforms reflect the source processes and therefore the dynamics of the volcano. In addition, the waveform is modulated by reflection, refraction, scattering and attenuation during propagation and thus reflects the seismic structure of the edifice.

Normal earthquakes are the result of brittle shear failure. In volcanic context, such events are referred to a **volcano-tectonic (VT)** events. They are characterized by relatively high frequencies ( $\geq 5$  Hz), a short duration of 5 to 10 s and clear onsets of P- and S-phase (Fig. 2.1). They usually originate from the response of the surrounding rock to changes in the stress field due to volcanic processes. For example in Iceland, the propagation of a magmatic dyke was heralded by a myriad of small VT-events [Sigmundsson et al., 2014]. Hence, VT events can indicate the presence or migration of volcanic fluids.

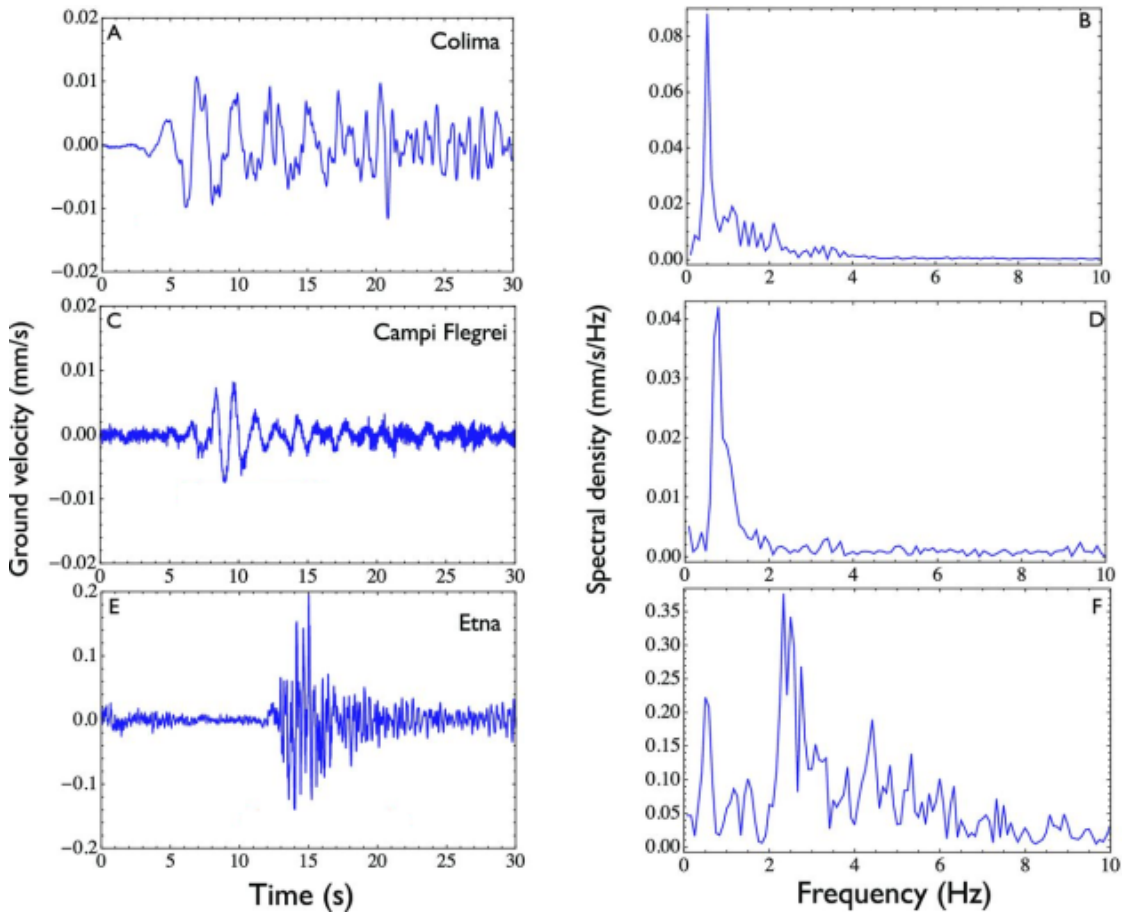
The so-called **long-period (LP)** events are commonly associated with fluid-dynamic processes in the plumbing system [Chouet, 1996] and their mechanism is fundamentally different from the one of VTs. LP-events (Fig. 2.2) are typically characterized by frequencies of 0.5-5 Hz, spindle-shaped, elongated waveforms and emergent onsets without clear P- and S-phases [Chouet, 1996]. The frequency spectrum may be very narrow or harmonic. Often the waveforms are very similar, which indicates a stable, repeatable source mechanism. Their reverberating character is frequently explained as the result of resonating interface waves that develop at the boundary of a fluid-filled reservoir, such as a conduit, dike or crack [Chouet, 1996, Neuberg, 2000, Sturton and Neuberg, 2006, Chouet and Matoza, 2013]. While resonance effects explain the oscillating coda of LPs, the actual trigger can be many things. Chouet and Matoza [2013] proposes five broad categories of trigger mechanisms: 1) self-sustained fluid oscillations, 2) magma-hydrothermal interactions, 3) magmatic degassing, 4) brittle fracture of melt, and 5) solid extrusion dynamics and plug stick-slip.

In contrast, Bean et al. [2014] suggested a slowly rupturing brittle failure mechanism

## 2. Background



**Figure 2.1.:** Samples of volcano-tectonic (VT) events from Villarrica, filtered between 2 and 20 Hz. They display clear onsets and - at most stations - distinguishable p- and s-phases. The relative importance of p- and s-phases is determined by the radiation pattern of the source and site effects of the stations. The VT-source was located 2-3 km east of the summit.



**Figure 2.2.:** Samples of LP-waveforms and their frequency spectra from three different volcanoes. From Pezzo et al. [2013]

that allows for “dry” LP-events. Moreover, a shallow source depth and strong scattering effects along the propagation path may cause significant distortion of a waveform such that even a VT-event may be mistaken as LP-waveform [Wassermann, 2012, O’Brien and Bean, 2009, Bean et al., 2014].

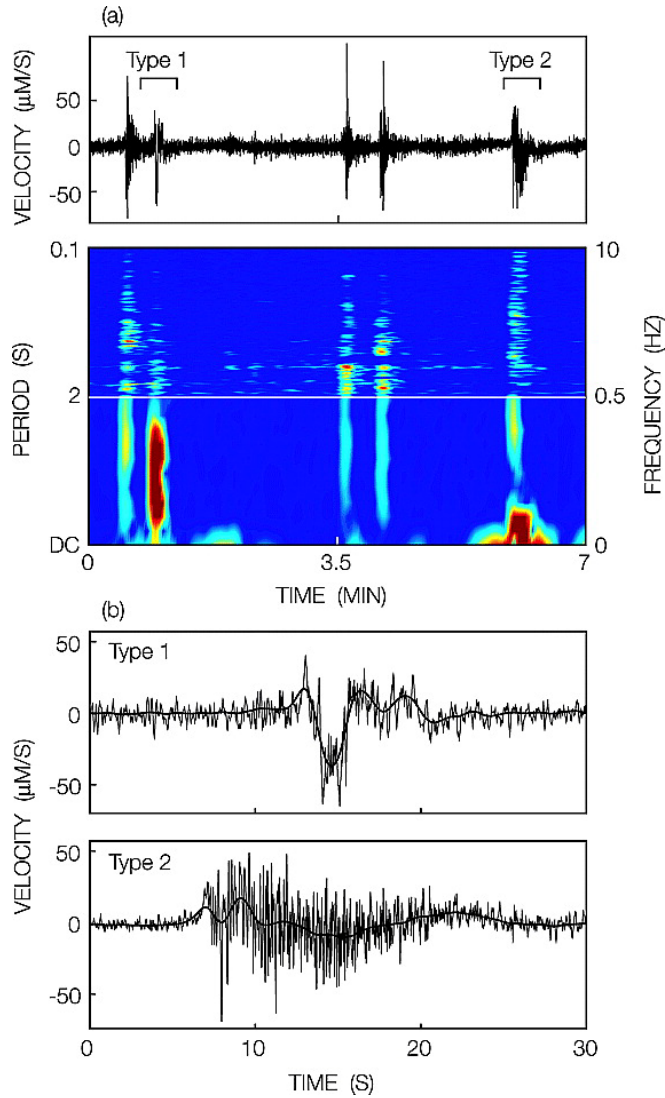
Events with frequencies of 0.01-0.5 Hz are referred to as **very-long period (VLP)** [Chouet and Matoza, 2013]. They are attributed to pressure perturbations in volcanic fluids and therefore show a volumetric source mechanism. Such perturbations occur e.g. at branching points and other changes in the fluid pathway or by the opening or closing of passages. Similar to LPs, the waveforms are highly repetitive.

**Hybrid events (HYB)** start with high frequencies and transition into a low-frequent coda. Hence, they display characteristics of both VT- and LP-events. The high-frequent onset is usually interpreted as a trigger event for the LP-like reverberations [Wassermann, 2012, Chouet and Matoza, 2013]. Chouet and Matoza [2013] emphasized that only events with a shear-failure component - indicated by mixed first polarities - should be classified as hybrids.

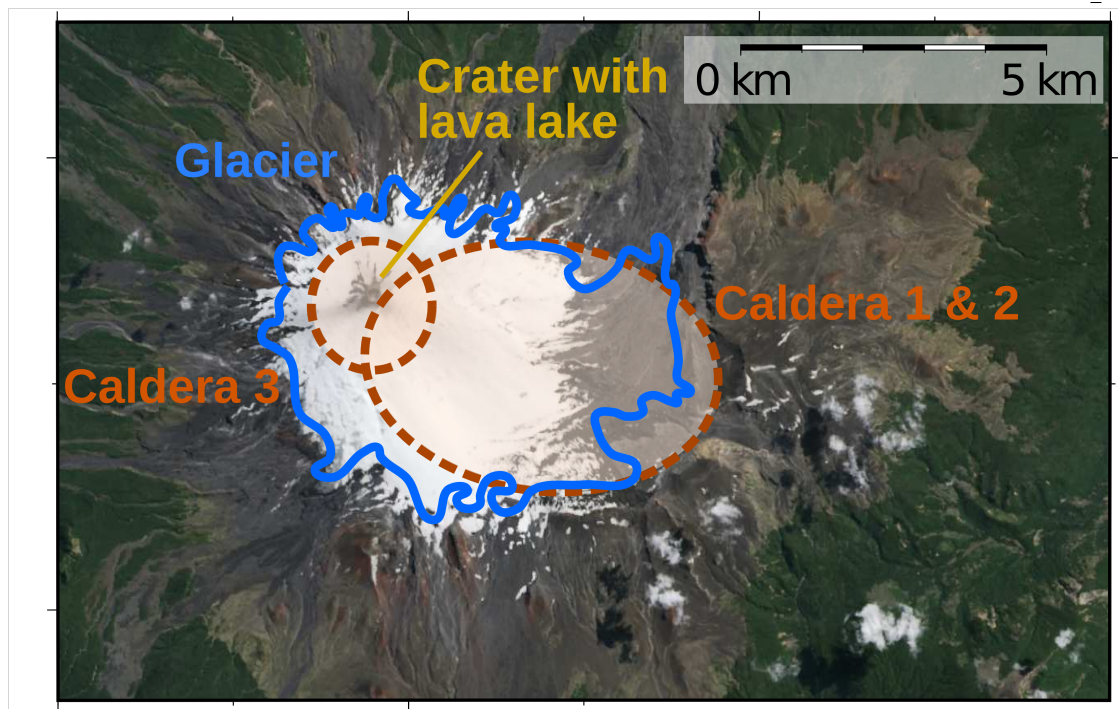
**Volcanic tremor** is a continuous signal that lasts for several seconds up to days and has similar spectral properties as LP-events. Typical frequencies range between 0.5 and 5 Hz with a narrow spectral peak [Konstantinou and Schlindwein, 2003, Wassermann, 2012, McNutt and Roman, 2015]. Different models of tremor generation have been invoked. Sometimes, tremor signals unambiguously show overtones in their spectrum suggesting a resonance phenomenon similar to LP signals. The flow of magma through a conduit or dike could produce a sustained vibration [Konstantinou and Schlindwein, 2003, Kawakatsu and Yamamoto, 2007, Ripepe et al., 2010, Chouet and Matoza, 2013]. The extended character may also be achieved by the rapid succession of repeating seismic events Neuberg et al. [2000], Eibl et al. [2017].

**Volcanic eruptions** are usually accompanied by a complex seismic signal that reflects the transport and emission of material from depth to the surface [Zobin, 2017c]. Other sources of seismic signals at volcanoes include surface activity such as rockfall, lahars and pyroclastic density currents or interaction with ground water [Wassermann, 2012, McNutt and Roman, 2015, Zobin, 2017d].

**Strombolian activity** is a spectrum of mild volcanic activity named after the typical activity of Stromboli volcano. It ranges from bursting of smaller gas bubbles (puffing) without material ejection to the explosion of large slugs that involve emission of bombs and ash [Gaudin et al., 2017a]. A number of studies demonstrated in analogue and numerical experiments that the ascent and bursting of gaseous slugs through liquid-filled pipes produces signals similar to those that typically accompany Strombolian activity [Ripepe et al., 2001, James et al., 2004, 2008, O’Brien and Bean, 2008, Chouet et al., 2010, Spina et al., 2019]. Moreover, James et al. [2006] showed that similar signals could also be produced if the slug passed through a sudden widening in the tube. The signals show characteristics of LP-events such as an emergent onset, and usually have a VLP-component [Neuberg et al., 1994, Chouet, 2003, Wassermann, 2012]. They may



**Figure 2.3.:** *a)* E-W component and corresponding spectrogram of two types of Strombolian explosions at Stromboli volcano, recorded  $\approx 400$  m from the active vent during a swarm of explosion in September 1997. The period axis of the spectrogram is partitioned into linear segments by above and below the white line at 2 s. *b)* Close-up of Waveforms of the two types: they show an emergent onset and a strong VLP-component. Figure from Chouet [2003]



**Figure 2.4.:** Satellite image of Villarrica Volcano with outlines of major features of the edifice.

also display a high-frequency component which corresponds to an acoustic wave in the atmosphere and coincides or follows shortly after the visible eruption [Ripepe et al., 2001].

## 2.2. Villarrica volcano

### 2.2.1. Geology

Villarrica Volcano is a 2847 m high stratovolcano located about 700 km south of Santiago in Southern Central Chile (Fig. 3.1, top left). It is situated above the subducting oceanic Valdivia fracture zone, the fluid release from which is thought to feed the magmas of the Villarrica system [Dzierma et al., 2012]. The volcano is composed of basaltic to basaltic-andesitic rocks. Its edifice consists of two larger, collapsed calderas of about 6 km diameter (Fig. 2.4) which formed during the Pleistocene [Moreno and Clavero, 2006]. At their western rim, the current cone is situated in a younger (3.5 ka, [Global Volcanism Program, 2013]), smaller caldera (about 2 km diameter). It hosts a summit crater of about 200 m diameter and an active lava lake with a varying depth of 50-250 m below the



crater rim [Richardson et al., 2014]. A glacier covers the upper 1000 m of the volcano.

Significant deposits were produced by three violent explosive eruptions during the Holocene. From the Lícán eruption (13.9 ka), up to 25 m thick, exposed layers of ignimbrites were found [Lohmar et al., 2007, 2012], while the Pucón ignimbrites (3.5 ka) reach up to 70 m [Parejas et al., 2010]. The Chaimilla eruption (3.1 ka) deposited up to 2.5 m of material [Costantini et al., 2011]. These ignimbrites consist mostly of alternating layers of ash, scoria, lapilli and volcanic bombs, deposited by fallout or pyroclastic flows. The Lícán [Lohmar et al., 2012] and Chaimilla [Costantini et al., 2011] ignimbrites are found predominantly to the west and northwest of the volcano, respectively, while the Pucón deposits [Parejas et al., 2010] are widespread around the edifice.

### 2.2.2. Volcanology

Volcanic activity consists of mild Strombolian eruptions and lava effusion, occasional mild-to-moderate explosions and ash/scoria ejections as well as continuous degassing. Its high activity and the proximity to human settlements make Villarrica Volcano one of the most dangerous volcanoes in Chile.

Studies by Witter et al. [2004], Mather et al. [2004], Palma et al. [2008], Gurioli et al. [2008], Palma et al. [2011], Moussallam et al. [2016], Aiuppa et al. [2017] on gas flux rates, gas and magma composition indicate that vigorous convection of a two-phase system (gas bubbles in liquid magma) takes place in the conduit. Convective two-phase flow could also explain the notorious seismic and infrasonic unrest [Ripepe and Marchetti, 2002]. Between 2000 and 2011 the daily means of typical degassing rates of SO<sub>2</sub> at Villarrica ranged between 0.5 and 20 kg/s with an average at 5 kg/s and rarely exceeded 50 kg/s during periods of enhanced activity [Witter et al., 2004, Mather et al., 2004, Palma et al., 2008, Bredemeyer and Hansteen, 2014].

Two studies by Moussallam et al. [2016] and Liu et al. [2019] investigated periodicities in gas parameters at sampling rates of 0.125-1.0 Hz. Although both studies measured the SO<sub>2</sub> flux at comparable locations of the plume (slightly above the crater rim and approximately 200 m above the magma surface [Moussallam et al., 2016]) Moussallam et al. [2016] showed periodicities at 30-380 s while Liu et al. [2019] found cycles of 345-714 s. However, Moussallam et al. [2016] themselves were reluctant about their findings, since contemporaneously measured gas concentration and temperature lacked any periodicity. Interestingly, Liu et al. [2019] observed cycles on a similar scale (30-50 s) but in SO<sub>2</sub> concentration within the plume (using a drone). These differences are possibly caused by an exceptionally low SO<sub>2</sub> flux during Moussallam's campaign. From the largely lacking periodicities Moussallam et al. [2016] deduced an efficient mixing of raising gas-rich and sinking degassed magma in the conduit resulting in a steady gas composition and flux rate. Liu et al. [2019] in contrast reported notable, audible bursts before the peaks in the SO<sub>2</sub> concentration. Moreover, they found a significant lack of correlation between the SO<sub>2</sub> concentration measured inside the plume directly above the crater and that

## 2. Background

---

measured by an instrument positioned approximately 100 m downwind at the crater rim. From the former finding, they concluded that the structure of the gas plume was predominantly formed by the (active) degassing process of the magma whereas from the latter, they inferred a nevertheless considerable influence of atmospheric effects (variable wind speed, turbulences, etc.). Due to a low  $\text{CO}_2/\text{SO}_2$  molar ratio of around 1:1, they also suggested that gas bubbles remain coupled to the magma until reaching shallow depths and being actively released. Periodicities on time scales of hours to weeks were reported in  $\text{SO}_2$  degassing rates [Bredemeyer and Hansteen, 2014] as well as seismic amplitude [Palma et al., 2008, Richardson et al., 2014]

## 3. Data

### 3.1. Seismic network

A network of 75 short-period stations was deployed in an area of approximately 63 km x 55 km around Villarrica Volcano between 1 and 14 March 2012. It was designed for exploring the  $P$  wave velocity structure in a previous study [Mora-Stock, 2015]. The 58 stations used in this paper for determining the origin of the seismicity are shown in Fig. 3.1 in black, blue and orange - depending on the method.

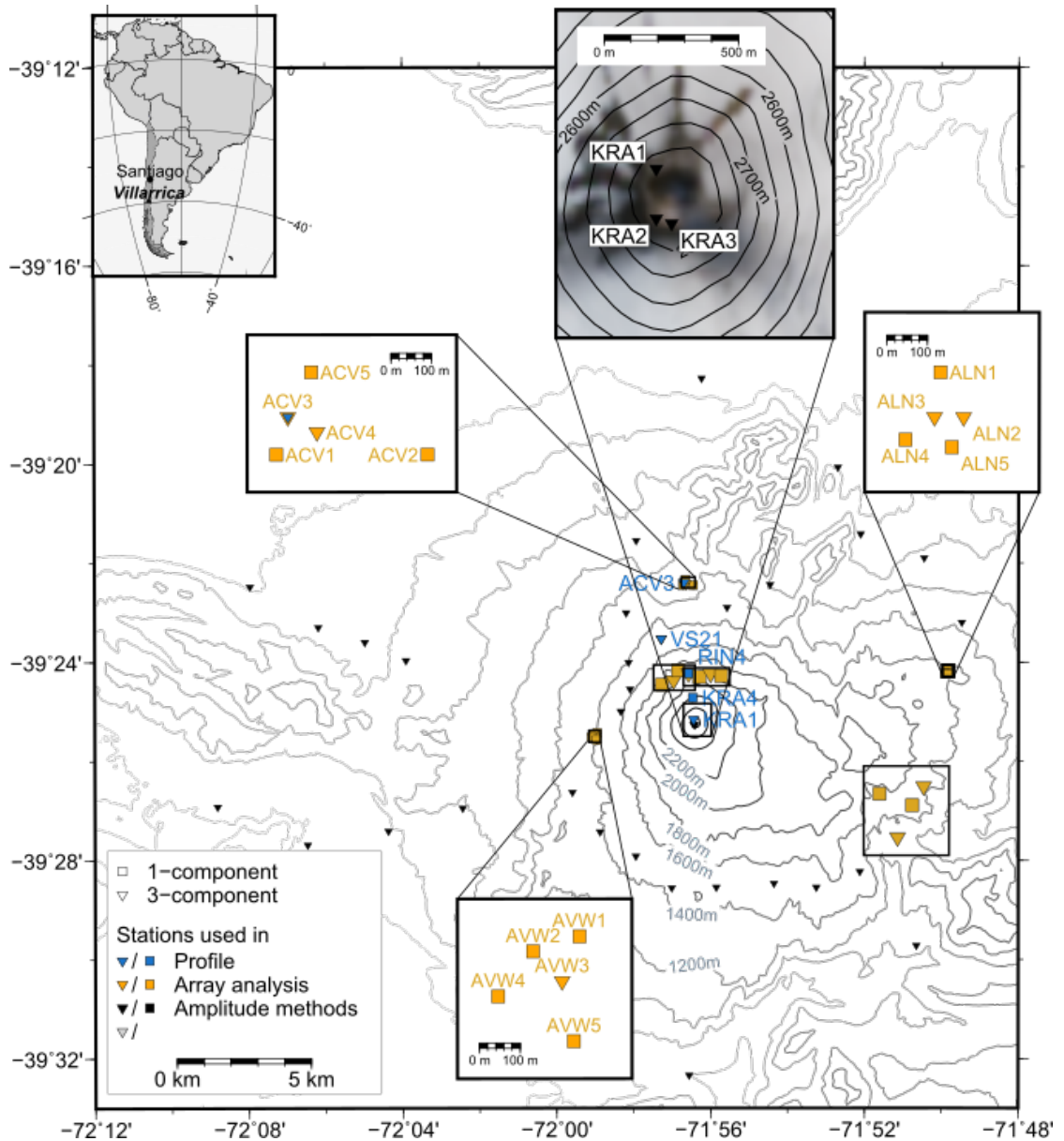
The stations consisted of a DSS-Cube and a SM-6/U 4.5 Hz-geophone from the Geophysical Instrument Pool Potsdam of the GeoForschungsZentrum Potsdam (Germany). Thirty stations were equipped with 3-component (3C) and 45 with 1-component (1C) sensors. The instrument type is represented by the symbol shape in Fig. 3.1. The operation time of each station is depicted in Fig. A.11 in the Appendix. The instrument response was removed using a 4-corner-bandpass pre-filter of (0.3, 0.4, 30, 45 Hz) and a waterlevel of 40. It was verified in a laboratory experiment - following Havskov and Alguacil [2004], Pavlis and Vernon [1994], Wielandt and Bormann [2002] - that the bandwidth of these sensors can reliably be extended to one or even two decades below their nominal natural frequency by removing the instrument response [Hilbert, 2015]. Examples from the laboratory study are included in Appendix A together with a filter example using field data from this study. All data were filtered using a strict lower cut-off frequency of 0.5 Hz before any further analysis.

About 50 stations were situated within a radius of 10 km around the summit (Fig. 3.1). Three stations (KRA1 - KRA3) were installed at the crater rim 2800 m above sea level (inlet Fig. 3.1). They are used for the automatic detection of transient events in Chapter 5. A third station (KRA2) operated only the first half of the observation period and is used to choose the frequency range for the analysis.

Seven stations (RIN1 - RIN7) were deployed on the northern flank at a height of approximately 1800 m a.s.l. and 24 were distributed around the volcano on the 1500 m-contour line. The glacier is roughly outlined by the 1800-contour line showing that only stations KRA1-4 and RIN1-7 were installed within or near the glacier.

For locating the source of the tremor and the velocity modeling (chapter 4), three smaller subarrays to the west (AVW), north (ACV) and east (ALN) of the summit (insets in Fig. 3.1) were used as antennas for beamforming together with three spatially larger subgroups of seismometers located at the rim and southeast of the crater (RINW, RINE and VSE). Stations belonging to the subarrays are colored in orange in Fig. 3.1 whereas

### 3. Data



**Figure 3.1.:** Topography and seismic station set up. Symbol shapes represent the instrument type while black, blue and orange colors indicate what the station was used for in Chapter 4. Subarrays ACV, ALN and AVW are enlarged in extra boxes. The satellite image shows the crater area with topography and stations overlain. Topography: Jarvis et al. [2008], Satellite image: NASA Earth Observatory [2015]. See Fig. A.12 for all station names.

groups are surrounded by black boxes.

The three subarrays AVW, ACV and ALN consisted of five stations each with a maximum aperture of 300 m, 300 m and 200 m, respectively, and a mean station spacing of approximately 100 m. They were situated to the west, north and east of the summit in distances of 4 km, 6 km and 10 km respectively. The RIN-stations were divided into two linear arrays RINW and RINE, comprising the stations RIN1-4 and RIN4-7. The distance to the summit was about 2 km and the spacing between the stations was about 400 m. This resulted in a maximum aperture of 1.5 km for each subarray towards the summit. The last subarray consisted of the stations VS12-15 which were located to the southeast of the summit in a distance of 8-9 km. These stations also had a fairly large spacing of 900-1500 m, resulting in a maximum aperture of 2.5 km. The records of the small arrays AVW, ACV and ALN show coherent signals at frequencies up to 4 Hz, whereas RINE, RINW and VSE show coherent records only up to about 2 Hz because of the larger station spacing.

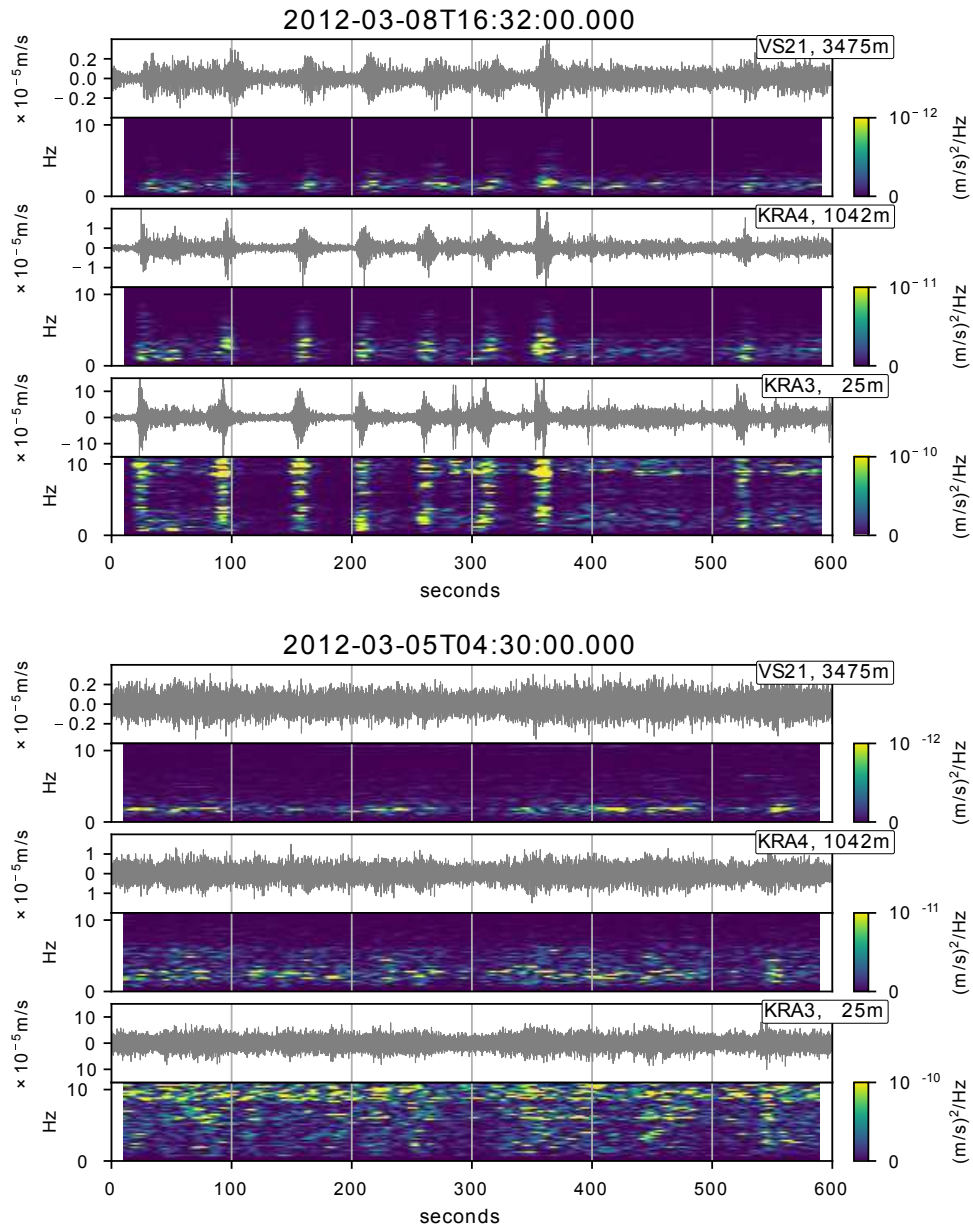
RINE and RINW were very close to the expected source region, so that the plane wave front assumption was no longer strictly fulfilled. However, test computations with plane and circular wavefronts showed that the resulting deviation of beam directions were less than  $3^\circ$ , which appeared tolerable.

## 3.2. Seismic activity

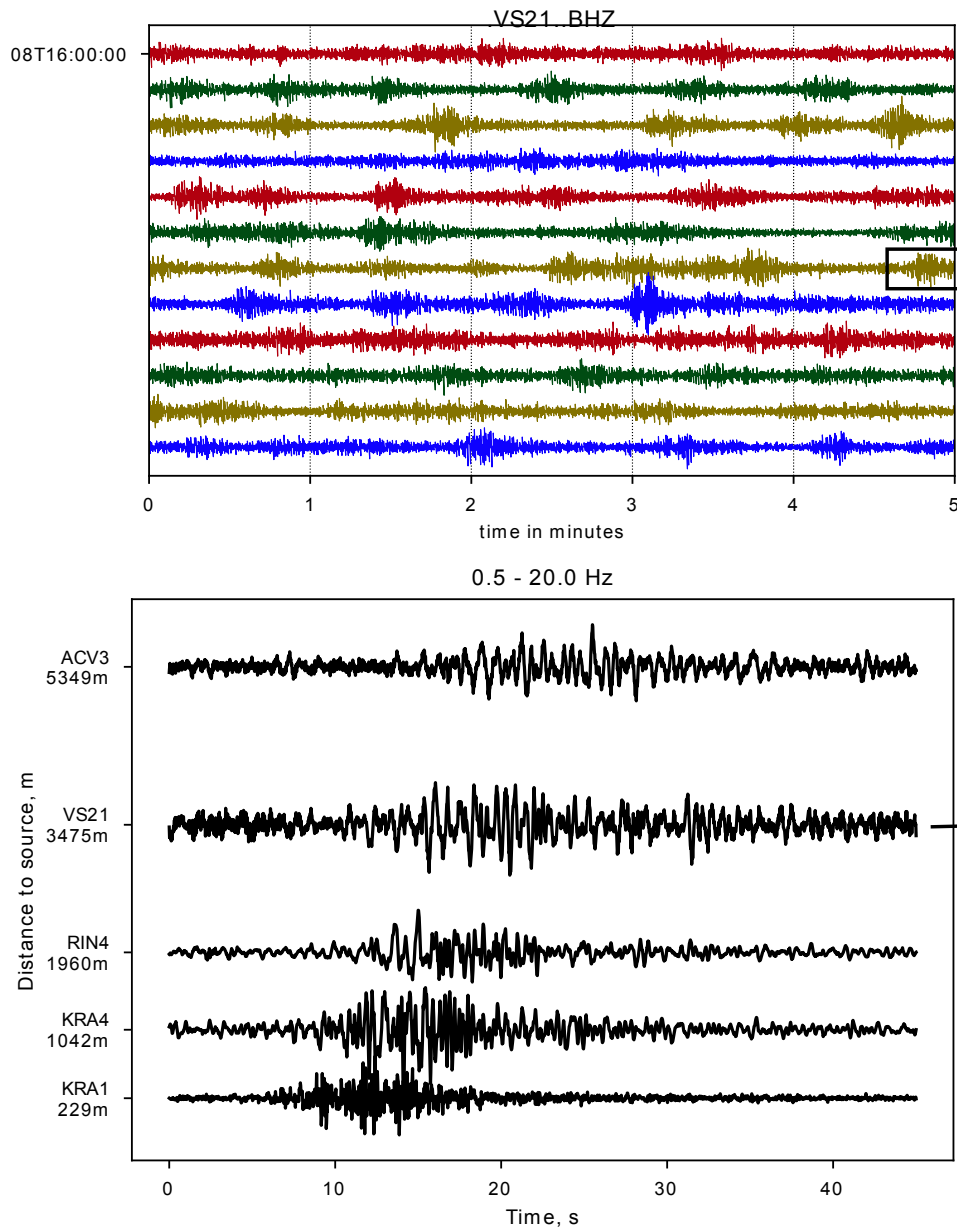
The seismicity was dominated by a persistent low-frequency signal throughout the whole observation time. The waveforms varied between a diffuse tremor of relatively uniform intensity over longer time spans ( $> 60$  s) and sections with high-amplitude transient signals. Examples of these two extreme cases are shown in the first 400 s of Fig. 3.2, top and Fig. 3.2, bottom, respectively. For lack of detailed knowledge about the true nature of the events and tremor, the rather generic terms "tremor" and "transient events" are employed as suggested by Palma et al. [2008]. Most of the time, the signal consisted of a mixture thereof, that is periodic transients separated by a relatively high level of background activity (Fig. 3.2, bottom,  $t \geq 400$  s). The corresponding spectrograms show that the frequency range varies with distance to the crater. At the crater rim, the signal energy is distributed between 0.5 and 10 Hz, whereas it is concentrated to the 0.5 to 2 Hz band at distances larger than 3 km. Station VS21 was situated similarly to the stations used in the earlier studies of Ortiz et al. [2003], Calder et al. [2004], Palma et al. [2008]. Here, the energy peak lies at about 2 Hz which is consistent with previously reported ranges. Fig. 3.3 (top) shows transient events in a 1 h-long record section of station VS21. The comparison to the earlier studies (data from October 2000, March 2000 and February 2005, respectively) revealed a quite similar character of the seismicity with transients occurring about every 60 s.

As shown in Fig. 3.3, bottom, these bursts showed emergent onsets and spindle-shaped,

### 3. Data



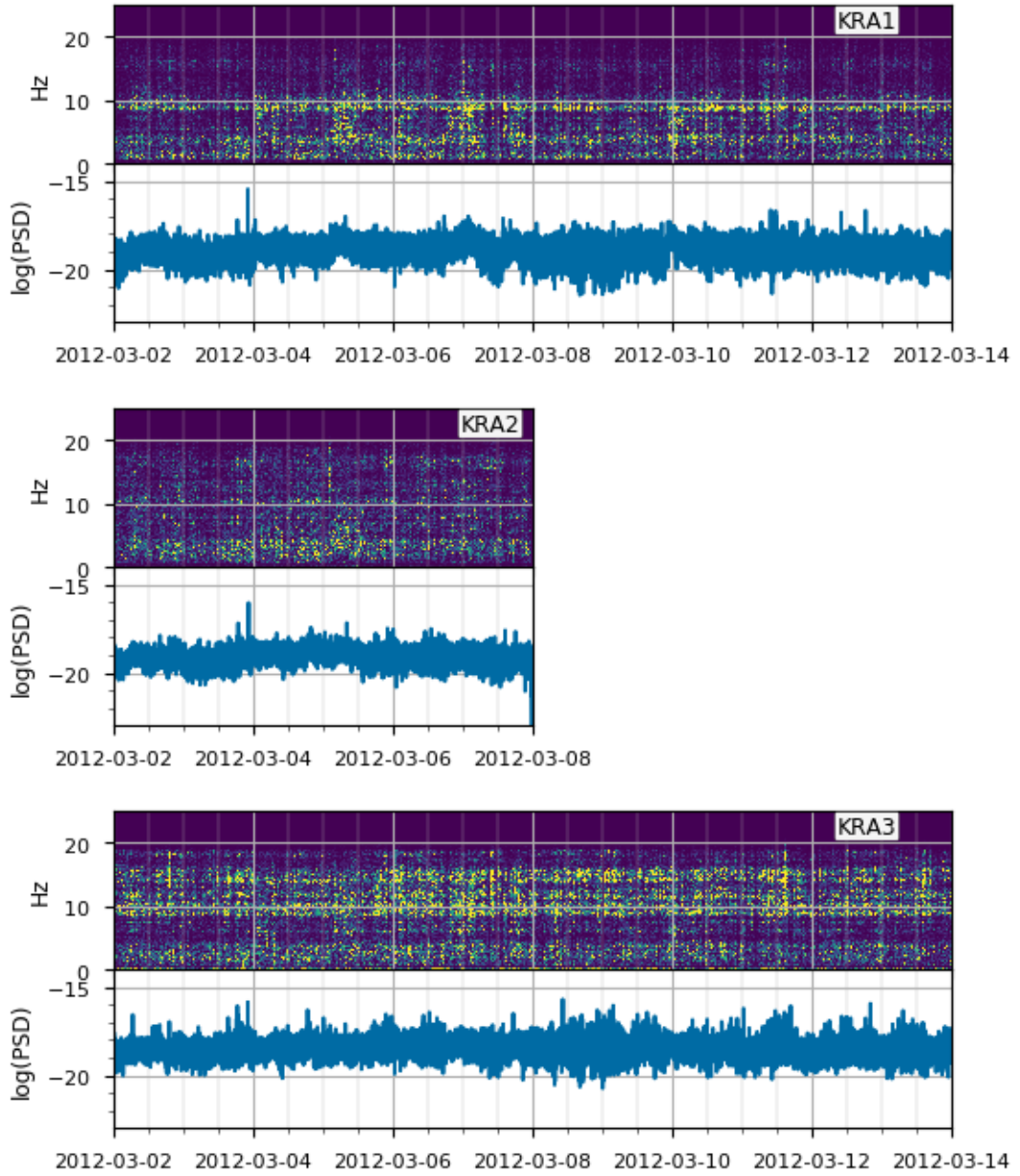
**Figure 3.2.:** Samples of the two end-members of the tremor/event mix: Top: distinct, impulsive transients, embedded in a low background tremor level during the first 400 s. Bottom: High tremor level with no or barely visible events. Frequency: 0.5-25.0 Hz, sampling rate: 50 Hz, spectrogram: 20.48 s with 20 s overlap, indicated distance is to source location at the summit.



**Figure 3.3.:** Transient events; signal is filtered between 0.5 and 20.0 Hz. top: occurrences in 1-hour section of station VS21. bottom: Evolution of a single event with distance to the source on vertical component: Amplitudes are normalized to the respective trace maximum. The spacing between the traces corresponds to horizontal distance between stations. The waveform originates at the summit and propagates along the northern flank.

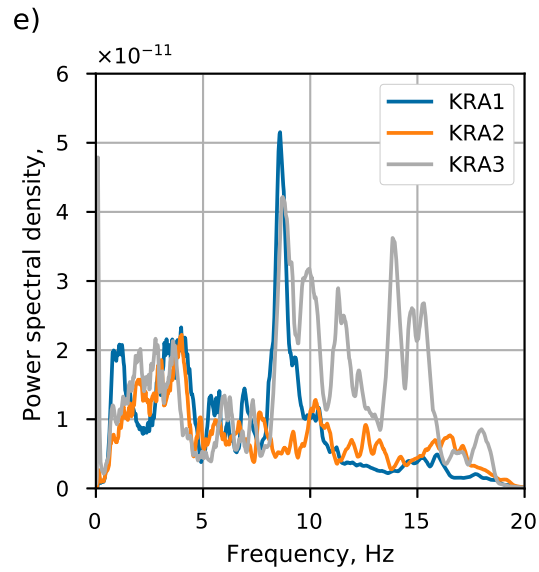
### 3. Data

---



**Figure 3.4.:** Spectrograms of seismic signal at the crater rim. Power spectral densities were computed in 40.96 s-long (2048 samples) windows with 50% overlap. The lower trace shows the summed psd over all frequencies.





**Figure 3.5.:** Mean power spectral density at crater stations over the whole available time period (12 days for KRA1, KRA3; 7 days for KRA2). Individual spectra were computed for 40.96 s-long segments, overlapping by 50%. Three bands can roughly be determined: 0.5-5.0 Hz with medium to high amplitudes, 5.0-7.5 Hz with low amplitudes and  $f > 7.5$  with dominating amplitudes at KRA3 and KRA1.

elongated waveforms at most stations (upper four traces). These features are frequently associated with LP events [Chouet, 1996]. Near the crater ( $< 500$  m) however, the waveforms were significantly shorter and more impulsive. Additionally, they contained much higher frequencies up to 10 Hz or more. Thus, the event waveforms showed a significant alteration and elongation along their travel path, indicating a strong influence of scattering effects. The same phenomenon was already observed by Richardson and Waite [2013]. In particular, they noted the possibility of falsely interpreting the discrete events as tremor if they were observed only at late stations.

The frequency content was generally stable throughout the period of observations (Fig. 3.4). Near the vent, the frequency range extended up to 16 Hz and can be divided roughly into 3 bands: 0.5-5.0 Hz, 5.0-7.5 Hz and 7.5-16 Hz (Fig. 3.5). Frequencies beyond 7.5 Hz even dominated the signal at stations KRA1 and KRA3 but showed very different patterns at the two stations. At KRA2 however amplitudes were reduced in that range. Furthermore, the correlation of the signal between the stations was poorer at the high frequency band than for the lowest one. All in all, it was concluded that the frequencies beyond 7.5 Hz were too much influenced by site effects. For the detection of transients the lowest frequency band of 0.5-5.0 Hz was selected.

### 3.3. Volcanic activity

The seismic campaign took place during a period of relative quiescence in volcanic activity [Global Volcanism Program, 2015]. The last Strombolian activity before the campaign was reported on 26 September 2011, followed by a period with almost no explosive activity until March 2012. During our campaign, two small ash emissions were observed on 7 March, incandescence from the lava lake on 7 and 8 March and lava spattering from 7-9 March. Shortly afterwards, another four small ash emissions occurred during 13-14 March, succeeded by a large white plume on 20 March and an 50 m high ash plume on 19 April. Incandescence ceased during April and was not observed until at least November. Analysis of infrared satellite data by MODVOLC [Wright et al., 2004] revealed an elevated level of thermal radiance, that started in early 2010 and ended shortly after the previously described activity in 2012. Incandescence and thermal anomalies indicate, that the lava lake is active. Significant activity was not resumed until the end of September 2014.

### 3.4. Nomenclature

For lack of detailed knowledge about the true nature of the events and tremor, we employ the rather generic terms *tremor* and *transient events* as suggested by Palma et al. [2008]. The transients are usually attributed to degassing activity such as explosions, bubble bursting and seething of magma in the lava lake Calder et al. [2004], Palma et al. [2008], Richardson and Waite [2013] but it is unclear whether this is true for all of these

waveforms. Therefore the term explosion might be too specific. Richardson and Waite [2013] used the term long-period events (LPs) in the context of a repetitive waveform. The official monitoring agency in Chile OVDAS also employs this term. At first sight, the expression merely suggests a description of the frequency range of the waveform, which - at least at the stations further away from the source - is even accurate. However, after more than three decades of research on the nature of LP events, this name seems to imply a bit more than that. Notably, it is associated with resonance effects of fluid-filled cavities.

Part II.

## Publications & manuscripts

## 4. Low-frequency seismicity at Villarrica Volcano: Source location and seismic velocities

Parts of this chapter have been published as:

*Lehr, J.; Eckel, F.; Thorwart, M. & Rabbel, W.*

*Low-Frequency Seismicity at Villarrica Volcano: Source Location and Seismic Velocities*

*Journal of Geophysical Research: Solid Earth, American Geophysical Union (AGU), 2019, 124, 11505-11530*

### **Main results / conclusions:**

- Tremor originates in the crater area
- Shallow velocity models at arrays AVW, ACV

**Abstract** *Villarrica Volcano (Chile) is one of the most active volcanoes in South America. Its low-frequency ( $\leq 5$  Hz) seismicity consists of a continuous notorious tremor, overlain by impulsive transient events of higher amplitude in 60 s-intervals. This signal was recorded in March 2012 by an extensive local network, comprising 75 stations and including six subarrays. It allowed us to apply and compare three techniques to locate the origin of the seismicity: intersection of propagation directions determined by array analysis, mapping amplitudes and modeling of amplitude decay. All methods yield almost identical, temporally stable, epicenters inside the summit crater, which confirms earlier attributions of the seismicity to volcanic activity inside the conduit. The discrete transients and the inter-event tremor share the same source location. From the dominance of surface waves and the obvious scattering, we infer a source near the surface. For two arrays at the northern and western flank, a dispersion relation was derived which allowed to determine  $S$  wave velocity-depth functions. At both locations, the velocity structure can be modeled by three layers with interfaces at 100 m and 400 m depth. The velocities (300 to 3000 m/s) correspond to pyroclastic material at different states of consolidation. The modeling of the amplitude decay reveals a quality factor around 50.*

## 4.1. Introduction

Villarrica is a 2847 m high, glacier-covered stratovolcano of basaltic to basaltic-andesitic composition in the Chilean Andes. It is considered one of the most active and dangerous volcanoes in South America with lahars posing the largest threat to the local population ( $\approx 35,000$  persons within 30 km) and the numerous tourists of the national park. The last thirty years have been marked by a sustained mild level of activity including an active lava lake which feeds mild strombolian eruptions and lava effusion, occasional explosions and scoria ejections as well as continuous degassing. This activity mostly affects only the crater region. The latest larger eruption occurred in March 2015. It caused several lahars at the northwestern flank, an ash plume of 3 km height and a lava fountain reaching 1.5 km [Global Volcanism Program, 2015].

The volcanic activity is accompanied by a persistent tremor of 0.5-5.0 Hz. For extended periods of days to weeks, it may be overlain by transient, high amplitude events, which are attributed to strombolian explosions [Ortiz et al., 2003, Calder et al., 2004, Palma et al., 2008]. A similar signal is observed in the infrasonic spectrum [Ripepe et al., 2010, Goto and Johnson, 2011, Richardson et al., 2014]. The entirety of the signal has been denoted by previous authors as “tremor” [Ortiz et al., 2003, Palma et al., 2008, Ripepe et al., 2010] while the impulsive bursts were termed “strombolian explosions” [Ortiz et al., 2003, Calder et al., 2004], “discrete, higher-amplitude bursts”, “high-amplitude transients” [Palma et al., 2008] or “long-period events” [Richardson and Waite, 2013]. This mix of terminology reflects the lack of detailed knowledge concerning the nature of the seismicity. For example, few studies specifically address the issue of locating the source of the seismic signal. Ripepe et al. [2010] used a small array of infrasonic sensors to attribute the infrasonic tremor to the crater. From the similarity of infrasonic and seismic signal they deduced a common origin. Richardson and Waite [2013] applied a waveform inversion to a single, repetitive “LP”-event and located its source at shallow depth inside the conduit.

The continuity of the tremor and the lack of clear event onsets inhibit the application of conventional location methods. To overcome these typical limitations for volcanic signals, a number of alternative approaches have been developed. Array-based techniques have been successfully applied to locate discrete, emergent events - such as long-period events (LPs) [Chouet et al., 1997, Almendros et al., 2001b, Saccorotti et al., 2001a] - as well as continuous tremor [Métaxian et al., 1997, 2002, Di Lieto et al., 2007, Eibl et al., 2017]. Multiple small arrays allow to determine the epicenter without a velocity model by intersecting the backprojected backazimuths (e.g. Métaxian et al. [2002]). With an accurate velocity model at hand, more sophisticated methods, which refine the determination of the slowness vector by comparison with simulated parameters, significantly reduce the uncertainty of the location and even enable the determination of the source depth [Wassermann, 1997, Almendros et al., 2001b]. Inherently, array methods provide information on the subsurface velocity structure. If the frequency-dependent phase velocity of surface waves can be determined, this dispersion relation can be inverted for a

shallow velocity model. In early attempts by Chouet et al. [1998] and Métaixian et al. [1997] regularly spaced, semi-circle arrays were deployed to make use of the SPAC-method introduced by Aki [1957], whereas Zuccarello et al. [2016] and Saccorotti et al. [2001b] used irregular arrays and a modified version of SPAC.

An entirely different approach was proposed by Battaglia [2003]. They located long-period and other events at Kilauea (Hawaii) and Piton de la Fournaise (La Reunion) firstly by interpolating the amplitudes within the sensor network and secondly by inverting for the amplitude-distance relation of the decaying amplitude. The decay is controlled by geometrical spreading and intrinsic attenuation (see e.g. Aki and Richards [2002], Stein and Wysession [2009]). Since the latter has a considerable effect on the result [Kumagai et al., 2010], the method provides also an estimate of the attenuation coefficient of the medium. The method is applicable to surface waves as well as body waves

Numerous variations of these techniques have been applied in the last two decades - along with different approaches such as e.g. the combination of particle motion with semblance analysis for very long-period events [Almendros and Chouet, 2003, Zuccarello et al., 2013] or the use waveform inversion (e.g. Ohminato et al. [1998]) or cross-correlation (e.g. De Barros et al. [2009]).

The main objective of this paper is to locate the origin of the low-frequency seismicity at Villarrica Volcano. To that purpose, we use data of a dense temporary sensor network, which was deployed in March 2012 - originally intended for a velocity tomography [Mora-Stock, 2015]. The abundance of stations on and around the volcano make the network predestined for the application of the amplitude source location method. In addition, the network comprised six sub-groups which we used for array analysis. In the present study, we compare the location results obtained by both approaches. We do not distinguish between events and tremor but look at the seismicity in its entirety. However, we examine the temporal variation of the epicentral zone. As pointed out above, both location techniques provide also information on  $S$  wave velocities and average attenuation of the subsurface. Contrary to other authors, who used the SPAC method to derive a dispersion relation for surface wave inversion, we directly use the slowness-frequency functions obtained from the array analysis. As of today, the knowledge about the seismic structure of Villarrica is sparse. Hence, our results may provide a valuable base for future more detailed work including numerical simulations.

The paper is organized as follows. In the next sections, we introduce the geology and the technical details of the data acquisition campaign, followed by an overview of the seismic and volcanic activity during that time. Subsequently, we explain the array and the amplitude-based methods including the determination of the velocity structure. Sections 4.4, 4.5 and 4.6 contain the results, discussion and conclusion, respectively.

## 4.2. Source location by probabilistic slowness analysis (“array analysis”)

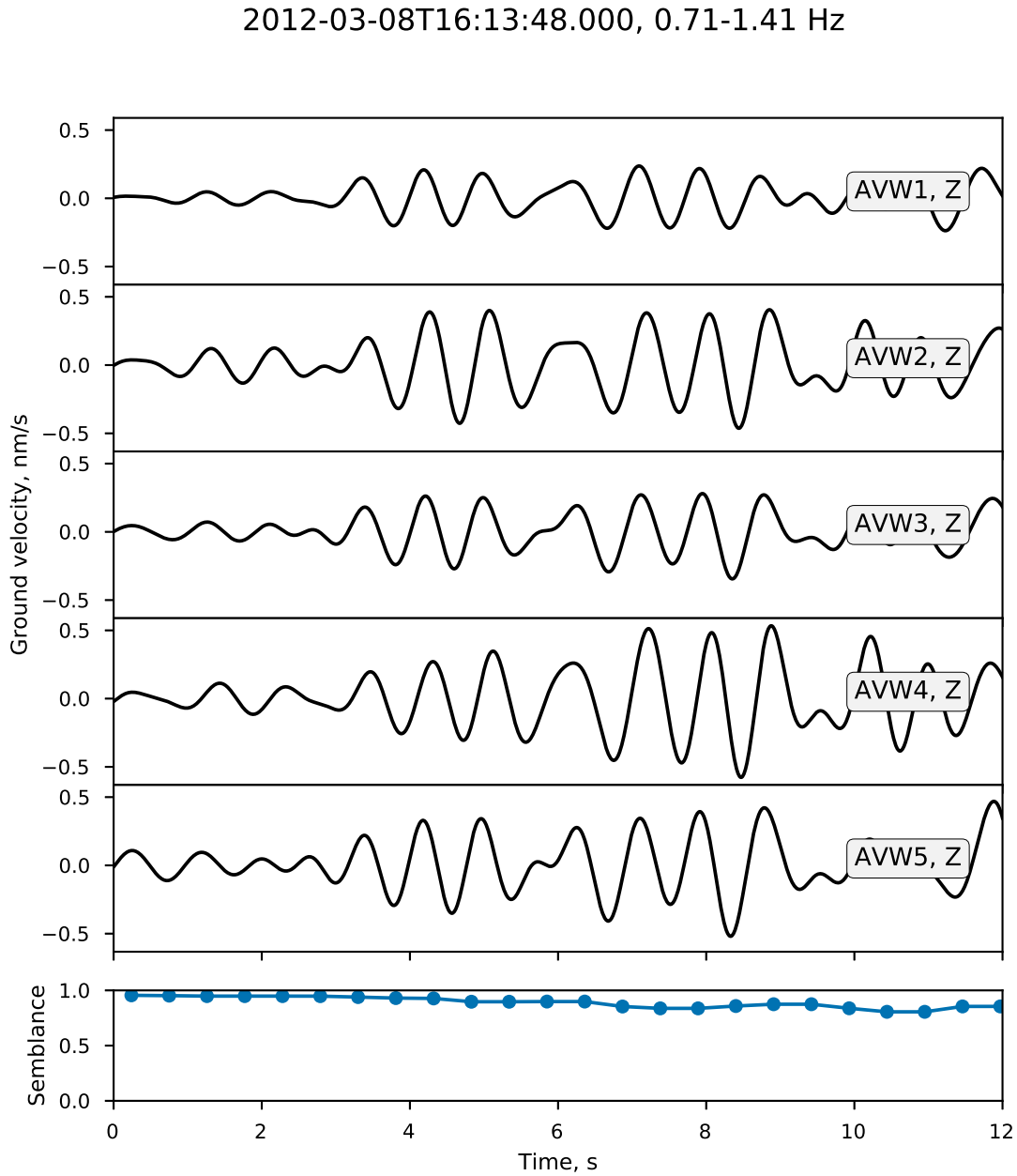
### 4.2.1. Beamforming

A group of at least three relatively closely spaced stations can serve as an antenna to detect the direction of arrival (DOA) and the horizontal slowness (inverse horizontal velocity) of an incoming wavefield. A plane wave, that crosses this group, produces the same waveform at each station, only slightly shifted in time (Fig. 4.1). The set of pairwise time shifts is determined by the DOA and the propagation velocity of the wavefront. As long as the wavefield maintains its coherency across the antenna, the time shifts are easy to detect. The corresponding slowness vector is usually found by a systematic grid search, which can be implemented in the time or frequency domain. The waveforms are shifted in time according to the test vector and stacked (beamforming or beampacking). The most suitable vector yields the highest power. Detailed explanation on the principles of array analysis and implementation of how to determine the slowness vector can be found e.g. in Schweitzer et al. [2012].

We applied the frequency-wavenumber analysis implemented in ObsPy [The ObsPy Development Team, 2017]. The quality of the stacked signal is measured by the semblance coefficient, which is a dimensionless quantity between 0 (no coherency) and 1 (perfect coherency). The analysis was applied to 1-hour-sections, which were divided into successive time windows of 5.12s with an overlap of 90%. The arrays operated from 2012-03-04T00:00:00 to 2012-03-13T00:00:00. For subarrays ALN and ACV two hours are missing due to numerical issues with the data. We tested 61 slowness values between 0.05 and 3.0s/km (corresponds to 0.3-20 km/s). DOAs were sampled in intervals of  $2^\circ$  between  $0^\circ$  and  $360^\circ$ . The output consisted of time series of the maximum semblance as well as the corresponding DOA and slowness values. In addition, we computed the absolute energy content of the stacked signal. Following Almendros et al. [2001a], we quantified the uncertainty of the DOA and slowness by determining the extent of the 95%-of-the-maximum-region in both slowness dimensions. We chose five frequency bands of one octave between 0.5 Hz and 4.0 Hz, overlapping each other by one half octave.

Fig. 4.2 shows the time series of all three parameters for three frequency bands during two consecutive episodic events. Evidently, high semblance values are not only achieved during events but also during inter-event tremor. The DOA oscillates around a stable mean value which is mostly similar at all frequencies. Strong, erratic deviations occur (e.g. at 21.5 min) but their amount is sparse compared to the periods of stable DOA. Moreover, the DOA during these episodes depends stronger on the frequency band. Such fluctuations may occur if reflections or signals from other sources yield more coherent phases in an analysis window. The general stability of the DOA is maintained throughout the whole observation time. Fig. 4.2 also indicates, that the DOA is independent of the frequency while the slowness clearly increases with frequency.

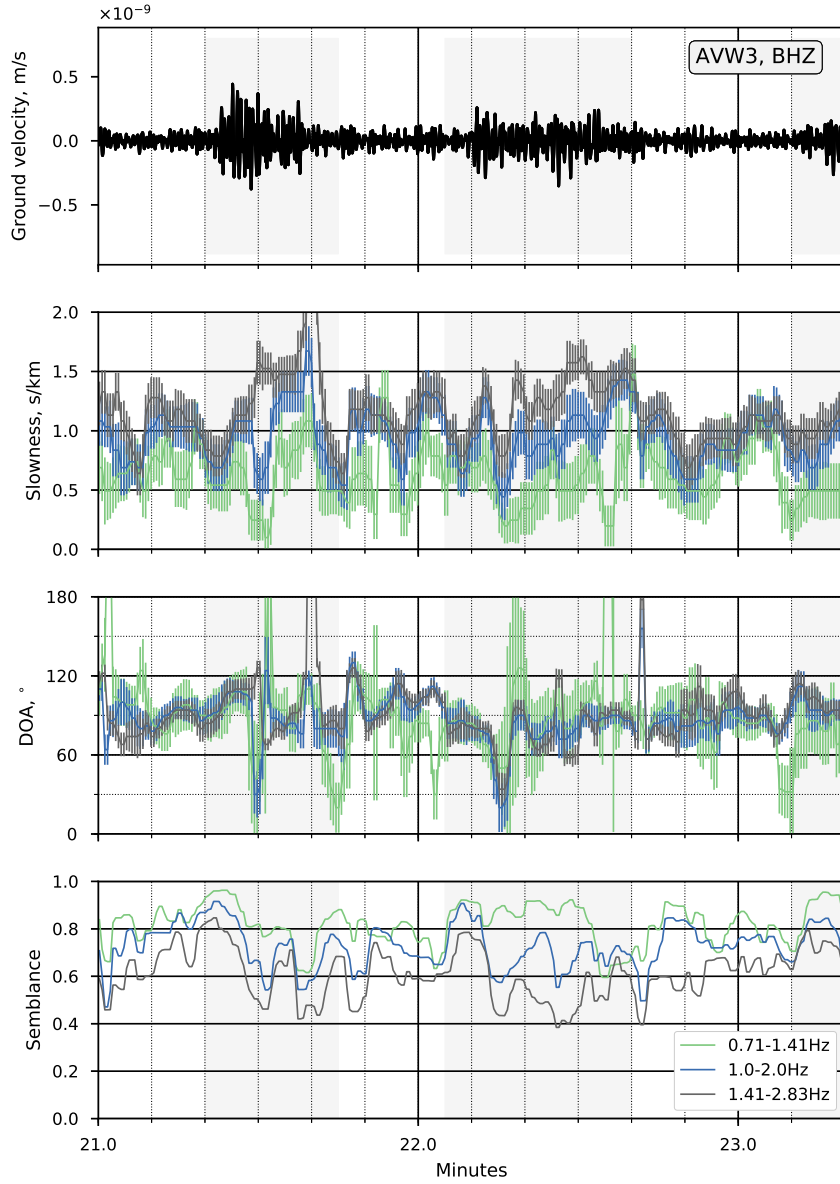




**Figure 4.1.:** Coherent waveforms during a transient event within subarray AVW. Data were filtered between 0.71 and 1.41 Hz. Corresponding semblance values (lowest panel) are drawn at the beginning of each 5.12 s analysis window.

#### 4. Source location & seismic velocities

2012-03-08T16:00:00.000, 0.71-2.83 Hz



**Figure 4.2.:** Time series of beamforming results for signal (top) at AVW: transient events are shaded in gray. Parameters vary independently of event occurrences.

### 4.2.2. Epicenter location

We located the seismic source area by determining the intersection of the beams from our six subarrays. This approach is similar to those described by Almendros et al. [2001b] and Métaixian et al. [2002]. The basic idea is that the origin of the signal lies at the intersection of the rays determined by the DOA from multiple antennas positioned around the source. Given the DOA measurement of an antenna  $A_i$  at  $(x_i, y_i)$  and an estimate of its uncertainty, we could derive a continuous probability density function (PDF)  $f_i(\varphi)$ . From this, we determined the likelihood of any direction interval to be the DOA of the detected signal. These likelihoods were backprojected onto a searchgrid. The joint probability of the intersecting beams is given as

$$p(x_j, y_j) = \frac{\prod_i^M f_i(\varphi_{ij})}{\sum_{l=1}^N (\prod_i^M f_i(\varphi_{il}))} \quad (4.1)$$

with  $\varphi_{ij}$  being the DOA at  $A_i$  for a signal starting at a location  $L_j = (x_j, y_j)$  of the searchgrid. Instead of the location quality factor proposed by Almendros et al. [2001b] and Métaixian et al. [2002], we normalized the joint probability density (the dividend in Eq. 4.1) by the sum over the entire searchgrid. This allowed us to interpret the result as an actual probability. The derivation of Eq.4.1 is explained in detail in Appendix 4.7.1.

In principle, any PDF can serve as  $f_i$ . We applied the von-Mises distribution which is an easy-to-use, normalized approximation of the normal distribution for periodic data [Mardia and Jupp, 1999]:

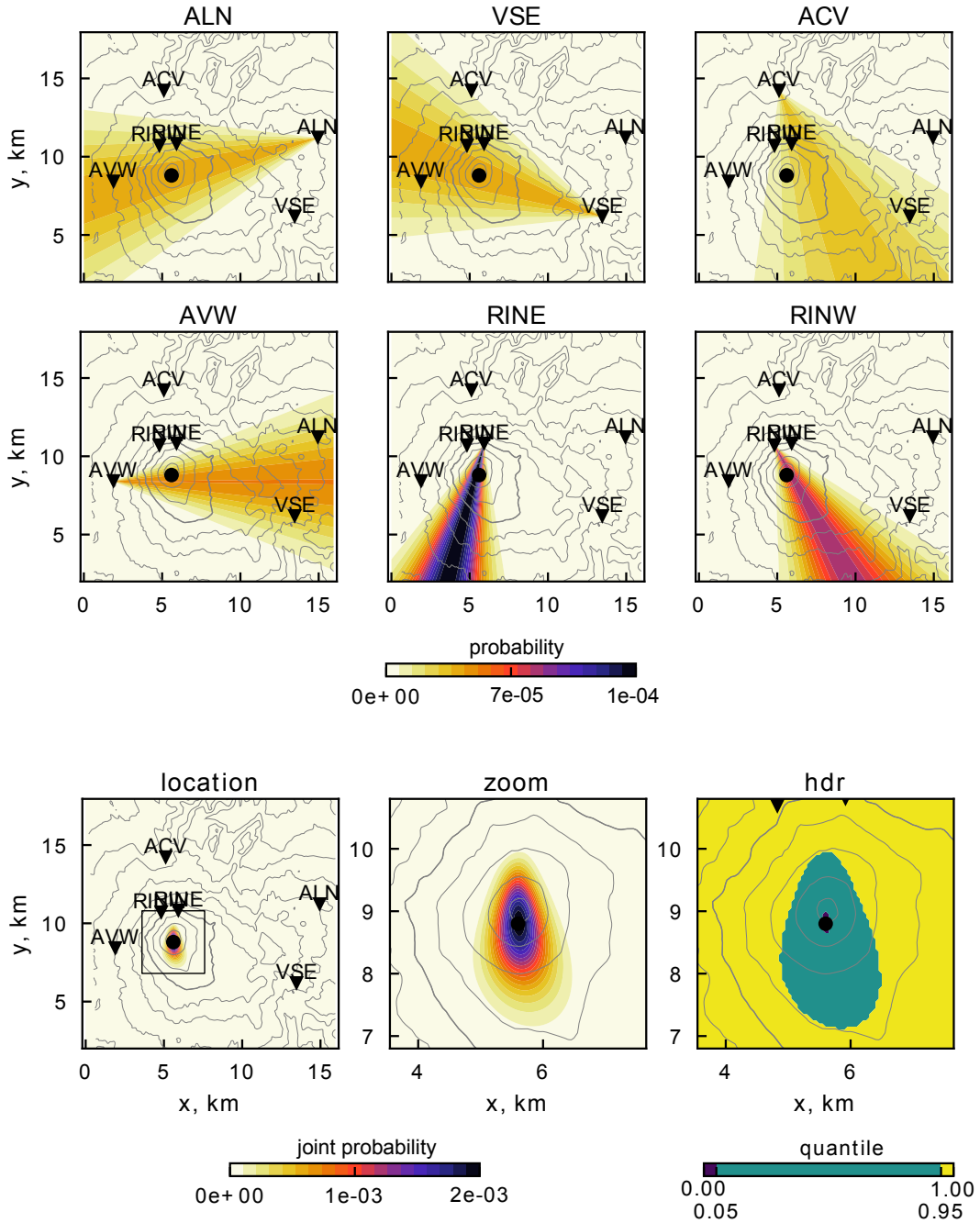
$$f(\varphi | \mu, \kappa) = \frac{e^{\kappa \cos(\varphi - \mu)}}{2\pi I_0(\kappa)}, \quad \varphi \in [0, 2\pi[ \quad (4.2)$$

with  $I_0$  being the modified Bessel function of order 0 and  $\mu$  and  $1/\kappa$  corresponding to mean and variance ( $\sigma^2$ ) of the standard normal distribution.

We evaluated Eq. 4.1 on a grid of  $16 \text{ km} \times 16 \text{ km}$  with a sample spacing of  $\Delta x = \Delta y = 0.05 \text{ km}$ , including the crater and all arrays. The grid point, which yields the highest probability, is considered as the source.

In order to assess the significance of the outcome of Eq.4.1, we used the concept of “highest density regions” (HDR) [Hyndman, 1996, Kruschke, 2015a]. In Bayesian literature it is also referred to as “highest density interval” (notably for 1D cases) or “credible interval/region”. This region comprises those grid points, which are bound by lines of equal probability such that their sum contributes  $100 \cdot q\%$ ,  $q \in [0, 1]$  of the total probability over the whole grid. In other words, the HDR is defined by a contour line of the data, which represents a  $q$ -quantile of the probability distribution. The sought-after source location lies with a probability of  $100 \cdot q\%$  within that region, given the estimated uncertainties of the measurements. For the computation of the HDR, we used a Python implementation of the algorithm proposed by Kruschke [2015b] ([www.github.com/aloctavodia/Doing\\_bayesian\\_data\\_analysis](http://www.github.com/aloctavodia/Doing_bayesian_data_analysis)).

4. Source location & seismic velocities



**Figure 4.3.:** Probability for epicenter locations from each single array (top and middle rows) and normalized superposition of probability density functions (PDFs) indicating the probability of the epicenter from combined beams arrays. The location procedure can be understood as a superposition of PDF "beams". Topography is indicated in 200-m-intervals, 2000 m-contour is highlighted.

The projections of the probability density functions and the result of their intersection are shown in Fig. 4.3. As described below in detail, the probability density functions are representative of the entire recording period. The highest probability is achieved at grid coordinates (5.6 km, 9.1 km), corresponding to 71.940°W and 39.4196°S. The source location is thus situated at the northern rim of the summit crater. The area which contributes 95% of the joint probability extends approximately 1000 m to the north and south of the maximum location and about 500 m to the east and west.

### 4.2.3. DOA determination

The stability of the DOA estimate, regarding frequency and time, is most evident after sorting the DOAs into a histogram (Fig. 4.4). Sections of one hour or one day always yielded similar distributions of the directions as when using the full time range. The spread of the distributions accommodates the episodic fluctuations of the DOA. The means are independent of the frequency band as shown in Fig. 4.4 (top). Therefore, the data of the available frequency ranges were merged before determining the parameters of Eq. 4.2. This minimizes the contribution of the strong perturbations of the DOA (as seen in Fig. 4.2) because they are less consistent regarding frequency. The parameters were then found by fitting the von-Mises distribution to the histogram using a nonlinear least-squares solver. The histogram was converted into relative frequencies through division by the number of total observations and normalized so that the sum between  $[0, 2\pi[$  is one. The histograms and fitted von-Mises distributions for the raw DOA data are shown in gray in Fig. 4.4.

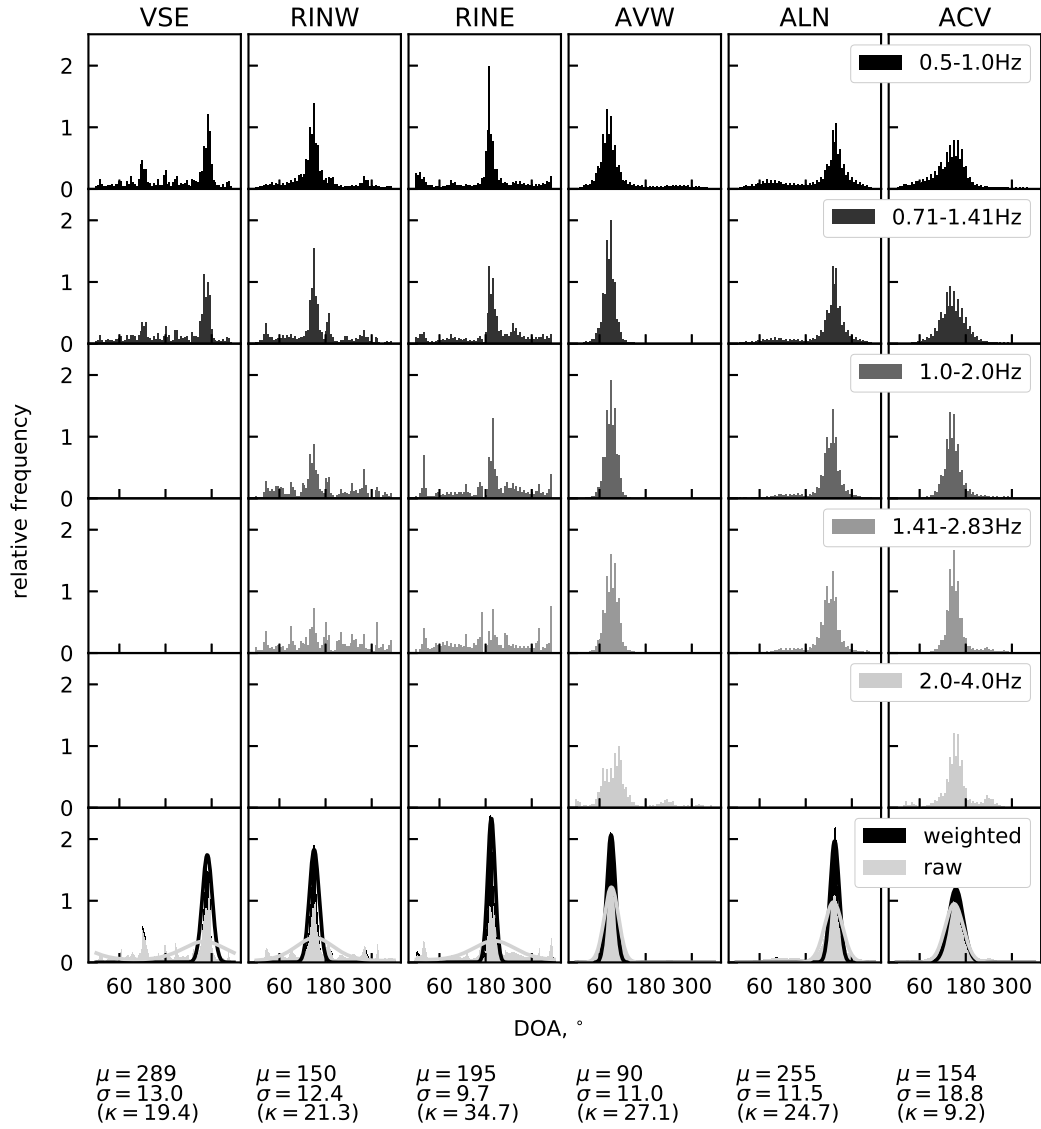
We found, that the variance of the parameter estimate reduced significantly when the contributions to the histogram were weighted by semblance and sample uncertainty using:

$$w = S^n \left(1 - \frac{\epsilon}{\epsilon_{max}}\right)^m \quad (4.3)$$

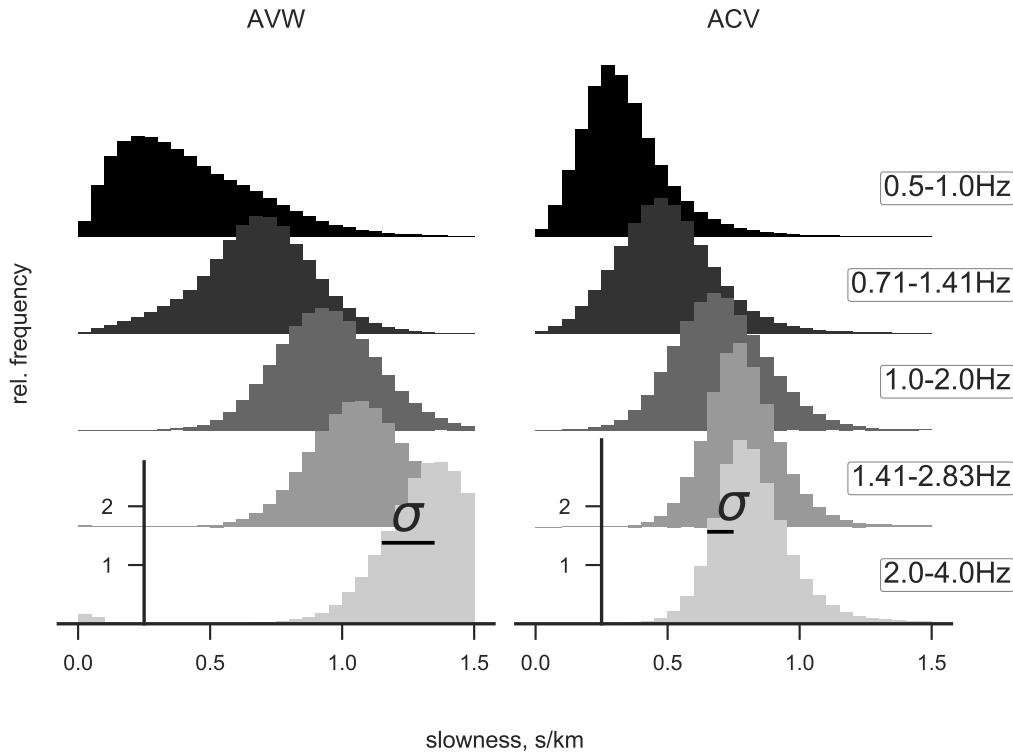
with  $S$  being the semblance,  $\epsilon_{max} = \pi$  being the largest expectable error and  $n, m$  being scaling factors. We chose  $n = 10, m = 10$  which effectively muted data with  $S \leq 0.8$  when higher semblance values were available but still allowed the method to be applied when the overall semblance was low. The resulting distributions are shown in black in the last row of Fig. 4.4.

Eq. 4.3 is an adaption of the weighting functions used by [Almendros et al., 2001b]. The semblance was included in order to enhance sequences with high semblances without using an absolute threshold for preselection.

In this way, we extracted DOAs using data of 1-hour and 24-hour sections as well as the the full time period of 9 days. In order to elucidate the relation between the episodic, high-amplitude transients and the inter-event tremor, we separated the data into event and tremor portions based on the energy content of the signal. The histogram analysis yielded the same DOA for both data sets.



**Figure 4.4.:** Histograms of DOAs. Rows 1-6 show unweighted data for each frequency range available at the corresponding antenna. Last row shows histograms and fitted von-Mises distribution of weighted (black) and unweighted (grey) merged data sets. Shape parameters of the von-Mises distribution are given for weighted merged data set underneath the respective column.



**Figure 4.5.:** Histograms of slowness values determined by beamforming using the full time range of data. Frequencies are normalized to the total number of observations in the histogram. The slowness clearly increases with signal frequency.

#### 4.2.4. S-wave velocity profiles from phase velocity dispersion

At arrays ACV and AVW the slowness shows a distinct dependence on frequency: the maximum in the histogram representation clearly moves towards higher slownesses for increasing frequency (Fig. 4.5), which is a characteristic of surface wave dispersion. Since only vertical components were used in the beamforming analysis, we interpret the detected slowness-frequency relationship as the phase velocity dispersion of Rayleigh waves.

The histogram peaks were attributed to the center frequency of the respective frequency band (the geometric mean of the upper and lower cutoff frequency) while the half-widths at half-maximum served as uncertainty  $\sigma_i$ . From these dispersion curves,  $S$  wave velocity-depth functions were determined using the neighborhood algorithm provided by Wathelet [2008]. It is a stochastic direct search method where the random generation of new models is guided by the results of previous runs. The quality of a model is assessed by the misfit between the modeled and the observed dispersion curve. It corresponds to the square-root

of the chi-squared test statistic:

$$\chi = \sqrt{\sum_{i=0}^{n_f} \frac{(c_{oi} - c_{mi})^2}{\sigma_i^2 n_f}} \quad (4.4)$$

where  $n_f$  is the number of sampled frequencies and  $c_{oi}$  and  $c_{mi}$  are the observed and modeled velocities in the dispersion relation.

For the inversion, the dispersion curves were resampled to 15 samples between 1.0 and 2.8 Hz on a logarithmic scale. These boundaries correspond to the center frequencies of the second and fifth frequency band. The lowest frequency band of 0.5-1.0 Hz and a center frequency of 0.71 Hz was removed since unrealistically high velocities were required to achieve a reasonable fit. The low-frequency DOAs, however, are consistent with those at higher frequencies indicating that the low-frequency arrivals were emitted from the same source. Their slowness of 0.25 km/s, corresponding to a velocity of 4 km/s, suggest that these arrivals are low-frequency  $P$  waves.

The formal variables of the inversion are  $S$  wave and  $P$  wave velocity, linked by Poisson's ratio, and density [Wathelet et al., 2004, Wathelet, 2008]. The model is organized in layers, and layering can be introduced for each of the parameters separately. Each layer is assigned a range for the respective parameters and the thickness. Since the influence of the density is negligible, it was fixed at 2700 kg/m<sup>3</sup> without layering. For the Poisson's ratio, we restricted the layering to one layer on top of a half-space, with the depth of the top layer coupled to  $v_S$ . Expecting unconsolidated material at the surface, we allowed a wider and higher range of 0.25-0.4 for Poisson's ratio and 0.2-0.3 for the half-space [Gercek, 2007]. Number and depth of layers for  $v_P$  were always analogous to  $v_S$ .

To find the basic layer structure we started the inversion with a model consisting of two layers of constant velocity. Only three constant-velocity layers were needed to fit the observed dispersion curve. We also tested models where the constant layer velocities were replaced by velocities with a constant depth gradient. However, given the uncertainties of the measured dispersion curve, no major improvement of the fit could be gained in this way, and the resulting velocity-depth models showed the same basic structure as the velocity models with constant layer velocities. Each input parameter set was tested with five separate runs, with each run converging after 10200 to 40800 models.

### 4.3. Source location by amplitude decay analysis ("amplitude source location")

Battaglia et al. [2003] introduced two methodical variants how the decay of amplitude of a signal along its travel path can be exploited for locating its source: 1) determining the coordinates of the maximum of the amplitude field by interpolating the amplitudes



observed at the receivers of a network and 2) performing a grid search for the location, for which the amplitude-distance relation fits a theoretical decay curve best.

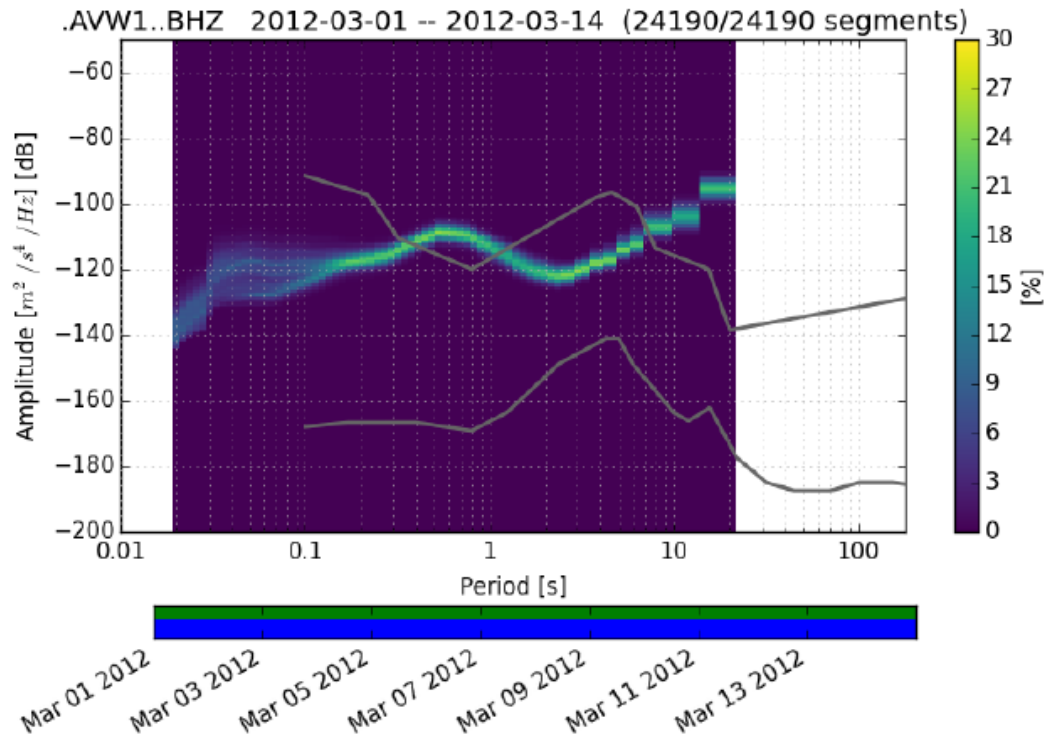
For each station, we determined the mean amplitude by computing the probabilistic power spectral density (PPSD) of the signal [McNamara and Buland, 2004]. For this purpose, the power spectral density (PSD) is computed for sequential time windows. The resulting large number of PSD estimates allows for a probabilistic evaluation. At each frequency, a histogram of the amplitudes is computed. Hence, the PPSD provides information about how often a combination of frequency and amplitude occurs and therefore allows an interpretation regarding the stability of the frequency content in time.

PPSDs were computed using velocity amplitudes and time windows of 100 s. For each station, all available data were used. The operation time of each station is given in Fig. 13 of the Supporting Information. As shown in the Supporting Information, we expect correct amplitude determinations for periods between 0.02 s and 2.5 s (frequencies of 0.4-50 Hz) after correction for the instrument response.

Fig. 4.6 shows the PPSD at station AVW1. The prevalent frequency-amplitude-distribution appears as a distinct, narrow band which indicates a highly stable signal during the whole period of observation. This stability is maintained throughout the network. Therefore, we consider the respective PSDs as representative for the station despite the different record lengths. The broad peak around 0.8 s appears at all stations and its amplitude decays with increasing distance to the summit, suggesting a volcanic origin. Based on this interpretation, we averaged the amplitudes between 1.25 Hz and 3.3 Hz to obtain a representative amplitude of the volcanic signal for the whole recording period. Most stations within 20 km distance to the summit showed a sufficiently high signal-to-noise ratio to be included in the analysis. The 58 stations used for the amplitude source location are marked in Fig. 3.1 in black (or blue and orange, if used before). Site amplification effects were quantified by analyzing the *S* wave amplitudes of regional events. The procedure is described in Appendix 4.7.2.

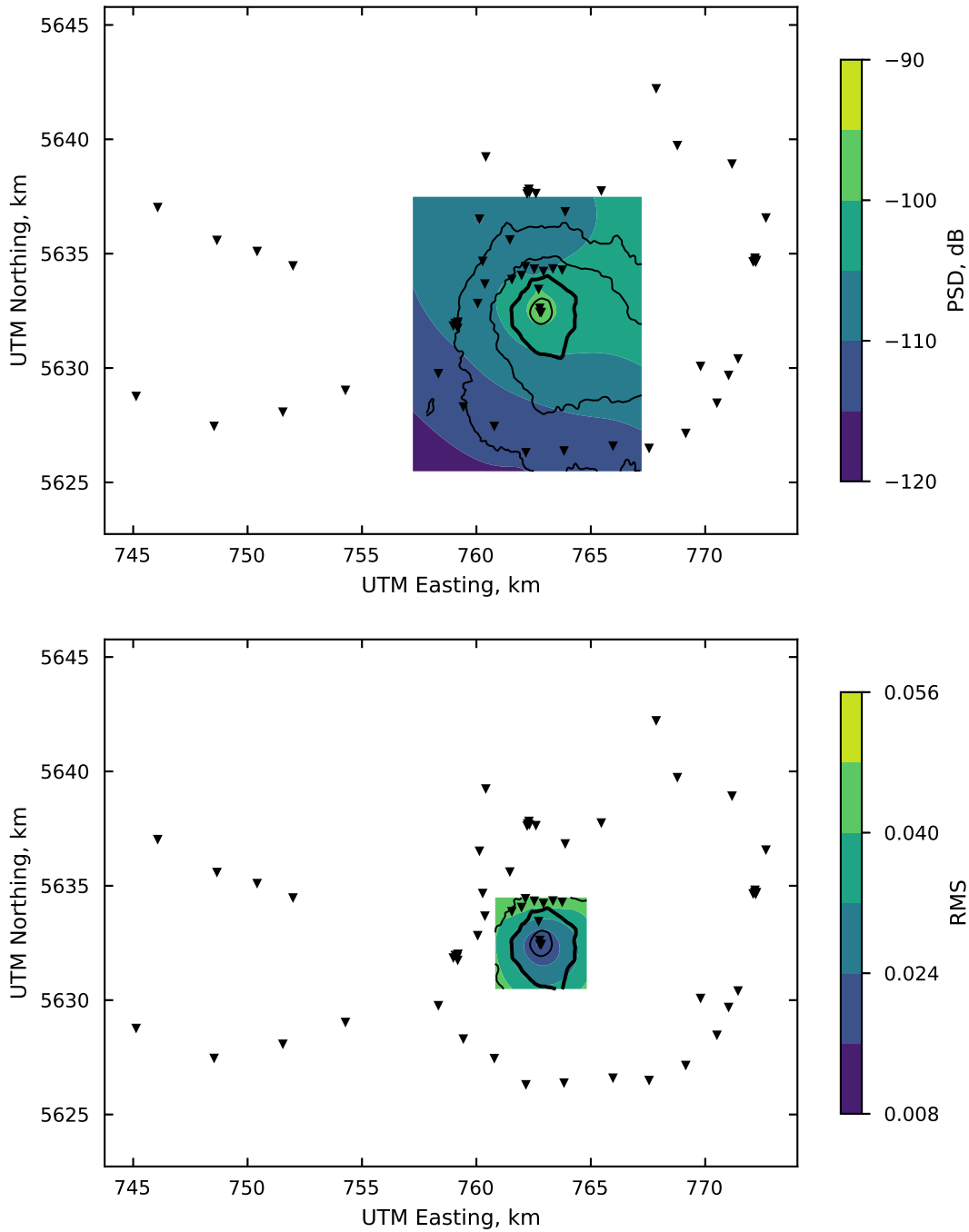
For the amplitude interpolation, we used a minimum curvature spline interpolation as proposed by Smith and Wessel [1990] to obtain a map of the amplitudes (Fig. 4.7, top) on a 50 m×50 m grid. For a point source in a homogeneous medium, amplitude isolines would form concentric circles around a maximum, which represents the epicenter. Deviations from the circular shape reflect either uncertainties in the determined amplitude or spatial variations of geometric spreading or attenuation of the seismic wave. However, even in this case, the location of the amplitude maximum can still be considered as an approximation of the epicenter if the source is situated close to the surface. For deep sources, mislocation due the projection of topographic information on a plane might become problematic if the source is not located underneath the summit. While Battaglia et al. [2003], Battaglia [2003] interpreted the amplitude maps only qualitatively, we determined the coordinates of the maximum. This enabled us to analyze the accuracy of the result using the jackknife method [Efron and Stein, 1981] as explained below.

The second variant uses the theoretical description of amplitude decay based on geo-



**Figure 4.6.:** Probabilistic power spectral density for station AVW1. The color shows the occurrence of each period/amplitude pair in per cent. The lower bar plot shows the available data in green and the coverage by the single 100 s long segments in blue. The gray lines in the diagram represent the low-noise and the high-noise model.

4.3. Source location by amplitude decay analysis (“amplitude source location”)



**Figure 4.7.:** Results of amplitude-based methods. Triangles mark contributing stations. Topography contours are given in black in intervals of 400 m. The 2200 m contour is highlighted. Top: Amplitude mapping. The maximum marks the epicenter. Bottom: Fitting of amplitude decay. The minimum marks the epicenter.

metrical spreading and absorption (e.g. Aki and Richards [2002], Stein and Wysession [2009]):

$$A(r) = \frac{A_0}{r^p} \exp(-Cr) \quad (4.5)$$

The exponential term describes the anelastic attenuation where the attenuation coefficient is composed of:

$$C = \frac{\pi f}{Qv} \quad (4.6)$$

Here,  $f$  is the frequency,  $Q$  is the quality factor and  $v$  is the velocity. The geometrical spreading is given by  $r^p$  with  $p = 0.5$  for surface waves and  $p = 1$  for body waves. Thus, one can perform a spatial grid search to seek the hypocenter where the difference between the observed and theoretical amplitude-distance relation is minimal.

The array analysis (Section 4.2.4, Fig. 4.5) had shown that the considered wave portions basically consist of Rayleigh waves. This implies that the waves propagate along the earth's surface and that  $p = 0.5$ , accordingly. We considered the effect of topography on the travel path by approximating the path by the 3-dimensional Euclidean distance between grid point and receiver. Thus, we neglected details of topographic variation between source and receiver. To ensure that this is a reasonable approximation, we computed the path lengths along the topography between the stations and a source at the crater and compared it to the corresponding 3- and 2-dimensional Euclidean distances. The median of the differences of all stations is 1% and 5% for the 2d and 3d case, respectively. Another consequence of the surface wave assumption is that we implicitly assume that the sources are located near the surface. For the inversion, we used a linearized version of Eq. 4.5 with respect to  $r$  by multiplying by the spreading factor  $r^p$  and taking the natural logarithm [Grazia et al., 2006]:

$$\ln(A(r) r^p) = \ln A_0 - Cr \quad (4.7)$$

The slope corresponds to the attenuation coefficient  $C$  and the intercept to the logarithm of the initial amplitude  $A_0$ . For their determination we used a linear least-square solver which minimizes the Euclidean L2-norm. The residual at any given grid point is the root-mean-square sum of the differences between the observed and modeled left-hand side of Eq. 4.7. By this approach, we avoided using a predefined quality factor, as e.g. in Battaglia et al. [2003] or Morioka et al. [2017].

The search grid covered an area of 4 km × 4 km, centered on the summit. The distance between the nodes was 50 m and the vertical source position was given by the topography data. We used a Digital Elevation Model of 90 m resolution, based on SRTM satellite data and provided by CGIAR-CSI [Jarvis et al., 2008]. The resulting map of minimum residuals is shown in Fig. 4.7 (bottom). The epicenter is defined by the minimum.

In order to estimate the accuracy of the solution we applied the jackknife resampling technique [Efron and Stein, 1981] on both methods. Coordinates of the maximum amplitude (interpolation) or minimum residual (grid search) were determined repeatedly by

systematically leaving out the  $i$ -th station. In doing so, we obtained a set of  $N$  locations with coordinates  $(x_{1i}, x_{2i})$  for the amplitude interpolation and the amplitude decay fit, respectively, of which the mean and variance was computed.  $N$  is the number of stations.

## 4.4. Results

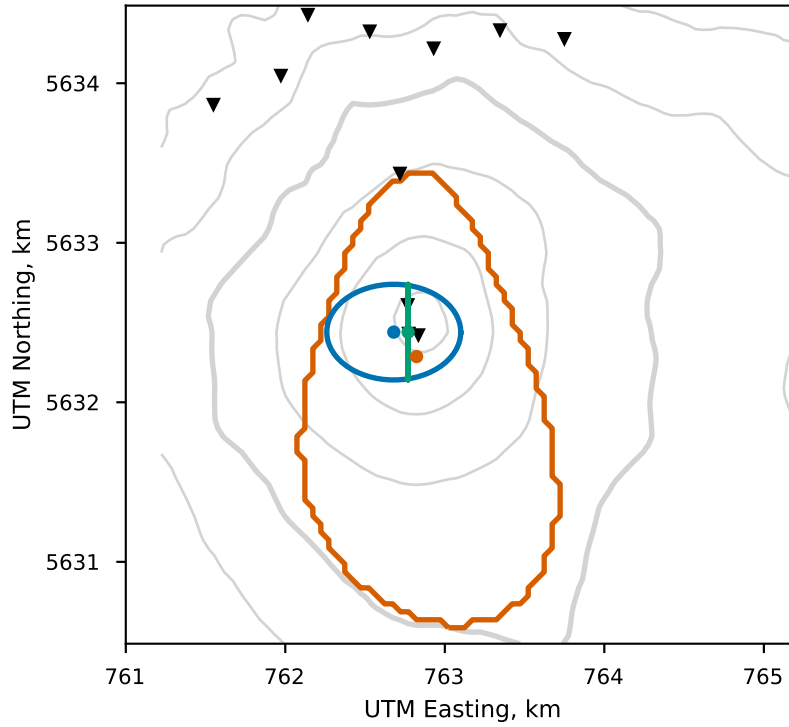
### 4.4.1. Location

The source locations and their uncertainties resulting from all three methods are presented in Fig. 4.8. The exact UTM coordinates of the optimum points are listed in Table 4.1. For the beamforming analysis, these coordinates correspond to the grid cell with the highest single probability. The figure shows the outline of the 95%-HDR, while the table lists its approximate extents. For the amplitude methods, the optimum point is the jackknife estimator and the uncertainty is given as twice the standard deviation, derived from the jackknife-variance. This corresponds to the extents of the 95%-confidence interval of the x- and y-coordinates, respectively. The differences between the results of all three methods is less than 200 m. The mean epicenter is situated at the geographical coordinates (71.94058°W, 39.42129°S). The epicentral areas determined by the amplitudes methods are situated on the southwestern crater rim, whereas the array analysis positions the source slightly outside the summit crater. Most importantly, the uncertainty intervals overlap completely.

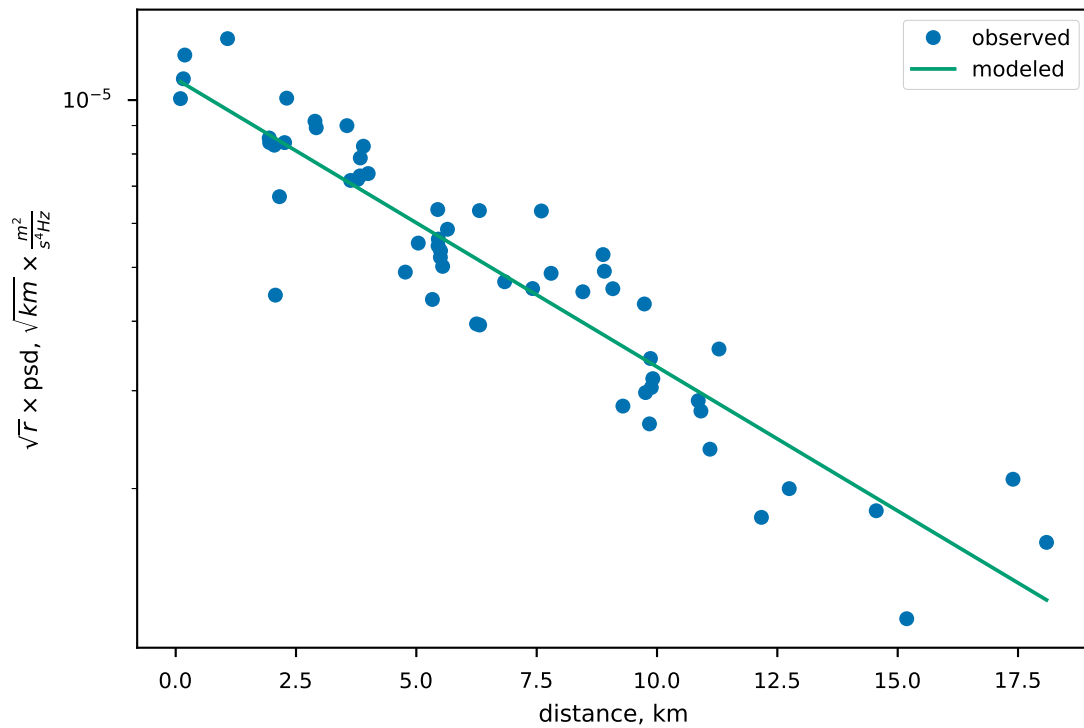
The strong path effects (Fig. 3.3), observed at single event waveforms, and the dependence of slowness on frequency (Fig. 4.5), indicate a prevalence of surface waves in the wavefield. Therefore, we could not deduce a potential source depth.

Fig. 4.9 shows the fit between the observed and modeled amplitudes at the optimum location according to the jackknife test. The mean relative error of the fit is around 2% with a maximum of about 5%. The slope corresponds to an attenuation coefficient of  $C = 0.12 \text{ km}^{-1}$ . The array analysis yielded slownesses for frequencies around 2 Hz between 0.8 km/s and 1.1 km/s, corresponding to phase velocities of 1.4 km/s and 0.8 km/s, respectively. Using these in Eq. 4.6 with  $f = 2.0 \text{ Hz}$  results in quality factors  $Q$  of 37 and 58.

The DOA values turned out to be very stable during the observation period. Fluctuation occurs on a time scale below hours since histogram analysis on hour- or day-long subsets reveal an equally stable mean DOA. When applied to time segments of one hour, the resulting optimum directions  $\mu$  mostly oscillate within  $5^\circ$  at most around a mean value. Therefore, the corresponding locations vary only slightly within a few grid cells (Fig. 4.10, top, center). The variation in North-South is somewhat larger than in East-West direction which is in good accordance with the elliptic shape of the HDR. There is no perceptible trend with time, that might be attributed to a moving source. We see the least variation of the coordinates as well as the area of the 95%-HDR with time between 7 and 9 March (Fig. 4.10). Similarly, separating the data into “event” and “tremor”



**Figure 4.8.:** Comparison of the optimum epicenter locations and their confidence intervals based on three different methods (cf. sections 4 and 5): Jackknife-estimator and 95%-confidence region for amplitude interpolation (Green) and amplitude decay fitting (Blue); Probability maximum and 95%-highest density region for array analysis (Orange). Inverted triangles represent stations. Topography contours are given in grey in intervals of 200 m. 2200 m-contour is highlighted.



**Figure 4.9.:** Fit of theoretical and observed amplitude decay for the source coordinates determined by jackknife resampling.

portions based on the energy content of the signal results in the same optimum location.

#### 4.4.2. Velocity profiles

The slowness histograms, in contrast to the DOA, vary strongly with frequency. They show a tendency of higher slownesses (lower velocities) towards higher frequencies - a characteristic that indicates surface wave dispersion. The effect is most pronounced at antennas AVW and ACV (Fig. 4.5). Since they are sensitive to the widest frequency range, they are the most suitable for the determination of dispersion curves and velocity profiles.

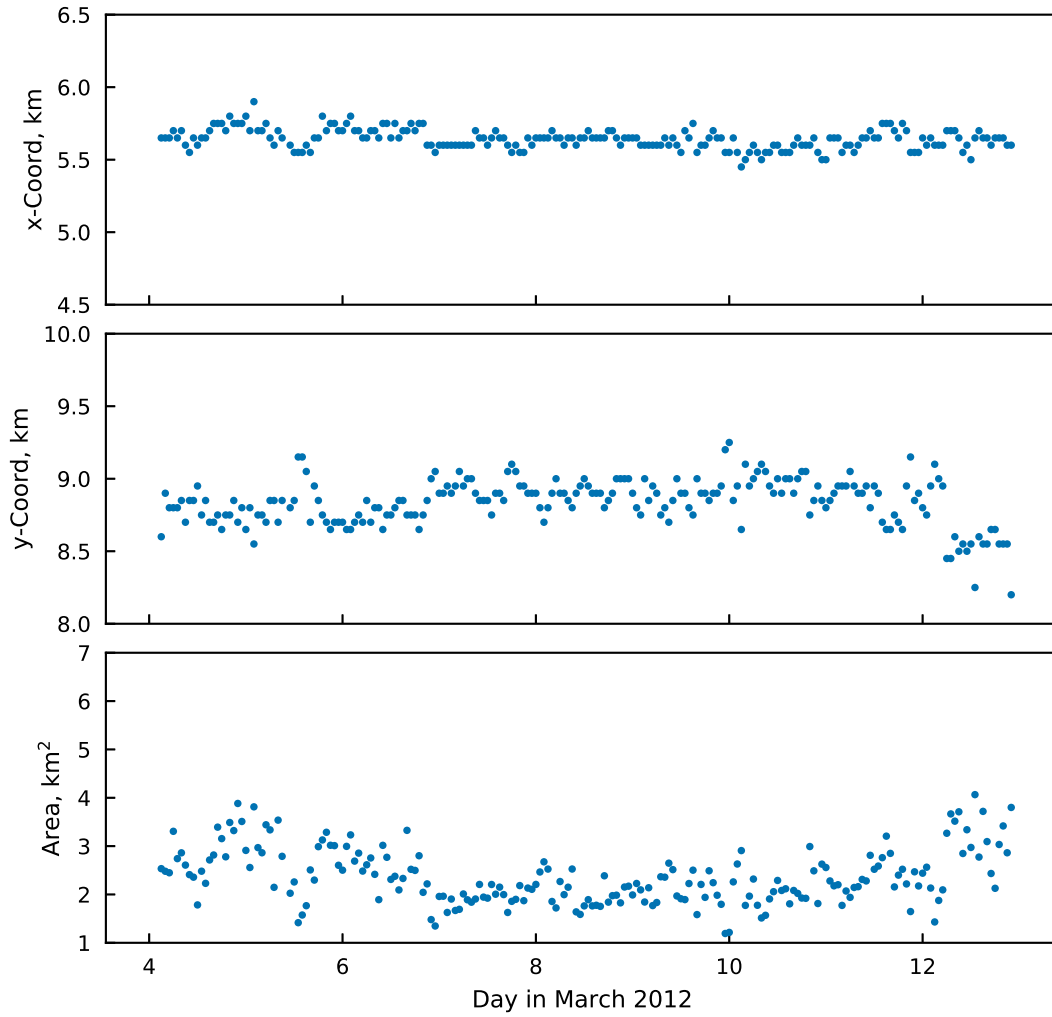
As representative examples, Fig. 4.11 shows the results for every tenth model that yields a misfit  $\leq 0.1$ . These ensembles comprise the models generated during the five separate, fully independent inversion runs, that yielded the lowest misfits. The velocity profiles (top row) and corresponding dispersion curves (bottom row) are color-coded by the misfit. The original dispersion curve and its uncertainties, derived by histogram analysis of the slowness values, is included in black. The interpolated target function used in the inversion is shown in gray. In the following, we primarily refer to the properties of the best-fitting models, while the ensemble of possible solutions gives an impression of the spread along the depth and velocity axis, given the uncertainty of the slowness. The models are cut at an arbitrarily chosen depth of 800 m. Note, that this deepest layer is the top of the half-space model.

For ACV, a minimum misfit around 0.028 is achieved for a 3-layer model. Most remarkably, the best-fitting models exhibit a shallow top layer with faster velocities than the one underneath. The ensemble, however, also offers solutions where the lowest velocities are in the surface layer. The inversion converges reliably for five of five independently generated starting models and after 20400 iterations.

In contrast, AVW exhibits a purely increasing velocity profile, while maintaining the 3-layer structure. However, the parameter set does not converge as easily. Only five out of twenty inversion runs converged at a final misfit around 0.0145 after 20400 iterations. Yet, within these five successful runs, the properties of the top layer remain ambiguous. About 30% of the best models feature a 50 m thick layer with a velocity of about 370 m/s, whereas the majority yields a 100 m thick layer with  $v_s \approx 700$  m/s.

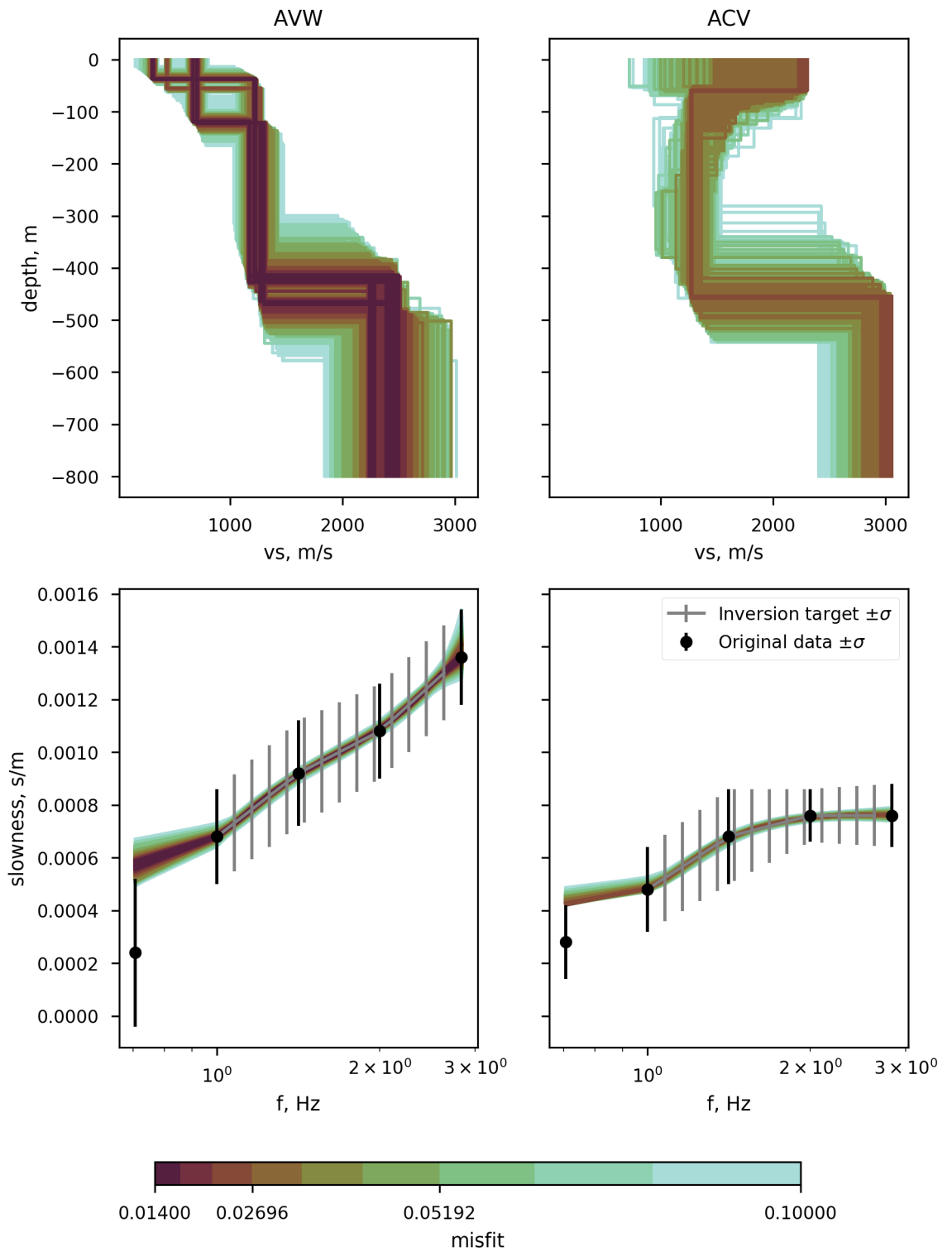
While the velocities in the uppermost 100 m differ greatly, both locations, ACV and AVW, exhibit a layer of 1100-1500 m/s in a depth between approximately 100 m and 400 m below the surface. The underlying halfspace is again slightly different, with velocities around 2500 m/s for AVW and 3000 m/s for ACV.





**Figure 4.10.:** Variation of optimum location parameters with time, determined by array analysis. Parameters were computed for 1-hour-segments.

4. Source location & seismic velocities



**Figure 4.11.:**  $S$  wave velocity profiles (top) based on inversion of dispersion curves (bottom) of surface waves at antennas ACV (right) and AVW (left). For the sake of clarity, only every tenth model with a misfit  $\leq 0.1$  is shown.

## 4.5. Discussion

### 4.5.1. Source location

We applied three different methods to locate the origin of the persistent low-frequency seismicity of Villarrica Volcano in March 2012. All of them yielded an almost identical source region inside the summit crater. This result is in good agreement with the study of Richardson and Waite [2013] who located the source of a repetitive LP event over a time span of three years, including 5-11 March 2012, inside the summit crater. In contrast to their study, we did not locate a selected single event but analyzed the signal as a whole over the full time interval of nine days. Therefore it can be concluded that the entire permanent unrest representing the seismicity of Villarrica in the considered time span originated at the summit crater. The way, we determined the amplitudes and DOAs for the array method parameters, may implicitly enhance the transients due to their higher energy content and slightly better correlation. However, a deliberate separation of high- and low-amplitude portions using the beamforming parameters indicated identical epicenters for the background tremor and the transient events.

The applied location methods imply that the seismic source is at or near the earth's surface. This assumption is supported by our finding that the observed waveforms are surface waves, from the dispersion of which geologically plausible  $S$  wave velocity-depth functions could be determined. To further evaluate the assumption of a shallow source depth, we tentatively assumed that the wavefield consists of body waves and repeated the location by amplitude decay fitting. For this test, the distance exponent  $p$  in Eq. 4.5 was set to 1 and the search space was extended to 3D. It turned out that the amplitude decay could also be fitted with body waves while the source location remained close to the surface. We see this as a numerical confirmation of a shallow source depth. Furthermore, the epicentral coordinates determined for the body wave solution deviated by several 100 m from those obtained consistently by array analysis and amplitude interpolation. Therefore, we conclude that the assumptions of surface waves and a shallow source depth are valid. Besides, a shallow source depths favors a strong scattering of the wavefield and, hence, explains the observed significant elongation and alteration of waveforms with source distance better than body waves radiated from a deep source. However, it should be noted that at the lowest frequencies (0.5-1.0 Hz), the velocities obtained by the array analysis were too fast to be surface wave velocities while the corresponding DOAs - and hence their origin - were similar to those at higher frequencies. Although the slowness would correspond to a plausible  $P$  wave velocity, we can neither rule out that it results from slower body waves impinging at a steeper angle which would imply a deeper source.

Regarding the spatial resolution, the probabilistic diagrams (Figs. 4.3, 4.7, 4.8) show that the applied location methods would not be able to separate closely spaced sources ( $\leq 100$  m). This applies also to a possibly varying source depth inside the conduit. Within these uncertainty limits, we showed that the horizontal coordinates of the signal origin

were stable in time and position during our period of observation. Minor fluctuations occurred only on hourly basis. These may be generated numerically by changes in the surrounding noise which affects the DOA determination but could also be caused by actual variations of the active source.

Our results are strictly valid only for a limited period of observations of nine days. Therefore, inferences about the source at other time periods should be made with care. However, as we pointed out in Section 3.2, comparisons with data from other observation periods suggests that the seismic activity during our campaign was similar in character and frequency to that reported in the previous studies (Ortiz et al. [2003], Calder et al. [2004], Palma et al. [2008], Mora-Stock et al. [2014]). Besides, the aforementioned repetitive LP event in the study of Richardson and Waite [2013] occurred over a time span of three years, including March 2012, which indicates a certain stability of the volcanic system during that time.

##### 4.5.2. Attenuation and scattering

In contrast to other authors using the amplitude source location method (e.g. Battaglia et al. [2003], Kumagai et al. [2011], Morioka et al. [2017]), we did not need to predetermine the quality factor due to the linearization of Eq. 4.6. Instead, we deduced the damping coefficient from the data during the fitting of the amplitude decay which yielded a quality factor between 37 and 58 for velocities at the frequency of 2 Hz.

Battaglia et al. [2003] used a fixed  $Q$  of 50 in their grid search at Kilauea (Hawaii). Using data from Tungurahua and Cotopaxi (both Ecuador), Kumagai et al. [2011] investigated the dependence of the misfit on various given  $Q$  and found the lowest residual for  $Q = 30$  in the frequency band of 1-6 Hz. Kumagai et al. [2011, 2010], Morioka et al. [2017] point out, that the quality factor used in the amplitude source location method is composed of scattering ( $Q_s$ ) as well as intrinsic ( $Q_i$ ) attenuation by  $Q^{-1} = Q_s^{-1} + Q_i^{-1}$ . Assuming that the wavefield is strongly affected by scattering,  $Q_s$  should be low and therefore dominate the total  $Q$ . Wegler and Lühr [2001] provided a method to separate scattering and intrinsic attenuation by inverting seismogram envelopes based on the energy diffusion equation. At Merapi (Indonesia), they found  $Q_s \approx 2$  and  $Q_i \approx 100$  for frequencies as low as 4 Hz. Using the same method, Prudencio et al. [2013] presented mean quality factors of  $Q = 35$ ,  $Q_i = 119$  and  $Q_s = 45$  for frequencies around 6 Hz at Tenerife Island. Especially the latter results are fairly similar to our findings. Strong scattering effects become further manifest in the massive alteration of waveforms along their propagation path.

According to Kumagai et al. [2010] and Kumagai et al. [2011] in order for Eq. 4.5 to be valid, one must assume isotropic radiation from the source, which is not the case for typical volcanic sources, notably cracks. In order to solve this contradiction, they suggested that scattering effects overprint the initial radiation pattern. To ensure this effect, the amplitude method should be carried out at high frequencies since they are

more affected by scattering. At Villarrica, Richardson and Waite [2013] also determined a nearly horizontal, single force source mechanism for their studied LP event, resulting in a strong horizontal component of the signal. Since their observation period partially overlapped with ours, we may assume that at least parts of our data were produced by a similar (or indeed the same) non-isotropic source. Therefore, this problem has to be considered here as well. There are basically three reasons indicating that our results are not affected seriously by radiation patterns: First, as already stressed above, we see indeed evidence of strong scattering effects at the analyzed frequencies. Second, we used a wide and dense station network, which would have revealed systematic distortions of the wavefield due to radiation patterns if they were significant. This is in contrast to Kumagai et al. [2010] and Kumagai et al. [2011] who only used five stations. Third, we used the mean amplitude of a period of nine days instead of single events. This averaging will compensate radiation bias if the source orientation is rather random than constant over time. Indeed, exemplary particle motion diagrams of the 3-component stations revealed that source radiation is variable in time, indicating variation of source types or orientations (see Supporting Information).

### 4.5.3. Velocity structure

The velocity-depth profiles, despite the considerable distance between their locations, are remarkably similar below 100 m. The  $P$  wave tomography of Mora-Stock [2015] proposed a  $P$  wave velocity of 4.5 km/s at both sites. This value would be consistent with the  $S$  wave velocity of 1.5 km/s in the middle layer and a Poisson's ratio of 0.25. Métaxian et al. [1997] presented vertical  $S$  wave velocity models from surface wave inversion for Masaya Volcano (Nicaragua) which are comparable in structure and velocities to ours. They found four layers with thicknesses of around 50-70 m, 130-170 m and 550 m. The  $S$  wave velocities increase from less than 1.0 km/s at the surface to around 2 km/s above the half-space, starting at 800 m. Velocity-depth functions were obtained at opposite sites of the volcano.

The relatively low velocity values of around 1500 m/s between 100 m and 400 m are typical for moderately consolidated and heterogeneous materials such as deposits from tephra fallouts or pyroclastic density currents. Indeed, exposed deposits from the Chaimilla (3.1 ka), Pucón (3.5 ka) and Licán (13.9 ka) eruptions reveal up to several tens of meters thick layers of ignimbrites, pyroclastic material and tephra fallout [Lohmar et al., 2007, Parejas et al., 2010, Costantini et al., 2011]. These older deposits are widespread around the western and northern flanks [Parejas et al., 2010]. Therefore, we may expect a similar velocity structure below the juvenile deposits of recent eruptions.

Velocities around 2000-3000 m/s, as found below 400 m, may be attributed to older and more consolidated pyroclasts or ignimbrites.

In contrast, the velocities in the surface layer are notably different between the two sites. The low velocities at the surface at AVW indicate loose and coarse material such

as young/fresh pyroclastic deposits while the high velocities at ACV point to more solid rocks. Corresponding outcrops are found at the respective sites: AVW was placed on rather loose pyroclastic deposits whereas ACV was situated on a solidified lava flow, known as the Cuevas Volcánicas.

The colors in Fig. 4.11 represent the misfit between the target dispersion curve and the one yielded by the respective model. Obviously, the dispersion curves even for the least fitting models lie well inside the error bars such that they are to be considered as of similar statistical value. Yet, the emanating structure is still well represented by the best fitting models. A dubious case are the velocities of the top layer at ACV, which could also be lower than the ones in the subjacent layer. The fast velocities however are backed by the visible geological site properties.

#### 4.5.4. Implications on the source process?

Together with the mild volcanic activity during the campaign, the epicenter inside the summit crater strongly suggests an intimate relation between seismicity and volcanic activity. Despite the considerable number of available studies, little is known about the actual cause of the seismic unrest and the event-like transients at Villarrica. Calder et al. [2004] attributed the transients to small strombolian explosions. Palma et al. [2008] stated a generally good correlation between SO<sub>2</sub> flux and intensity of the seismicity. However, their study is based on only a small number of daily gas observations. They also observed a temporal correlation between bursting bubbles at the lava lake surface and transient seismic events in 2006. Both findings support the general idea, that tremor but also LP events are caused by migrating fluids [Chouet, 1996]. On the other hand, Goto and Johnson [2011] reported a lack of correlation between the transients in the infrasonic tremor signal and video observations of lake activity in 2010.

The study of Richardson and Waite [2013] used the longest observation time series, in which they identified a repetitive “LP” waveform that originated in the crater. They interpreted its source-time function as a magma dragging at a horizontal lake bottom. These authors used the term “long-period” events but attributed the investigated signal to seismic and infrasonic emission close to or at the surface of the lava lake. Hence, it is not entirely clear whether their studied event is fundamentally different to our transient events. In the following discussion, we use quotation marks to denote the terminology of Richardson and Waite [2013]. Nevertheless, our study likewise located the origin of the seismicity inside the crater. Assuming that the “LPs” of Richardson and Waite [2013] are similar to our transient, our result is in good accordance with theirs.

Moreover, Richardson and Waite [2013] already pointed out that signals may be strongly affected by path effects, and they found an empirical correlation between tremor intensity and the number of detected events. Based on these two findings, they suggested an intimate link between single “LP-waveforms” and tremor, such that the tremor signal consists at least partially of overlapping “LP”-codas.

Our results support this idea inasmuch, as we can confirm the observation that signals evolve from short bursts at the crater into extremely elongated waveforms at a few kilometers distance. Hence, superposition of the signals becomes more and more of an issue up to the point that the discrete impulsive character of the transient events is lost. Signals recorded only at some distance to the crater might be misinterpreted as uniform tremor or LPs.

In addition, we demonstrated that the entirety of the low-frequency signal of Villarrica can be attributed to a confined source region. This could be seen as another indicator for a common source of tremor and events. Yet, we have to point out, that our methods would probably not resolve too closely spaced sources ( $\leq 100$  m). This is especially true for different source depths inside the conduit, due to the lack of accurate depth determination.

Finally, not only the location but also the frequency content showed little variation during the period of observation. Nevertheless, we observed significant fluctuations in the character of the seismicity as reported in section 3.2. This suggests at least a close relationship between tremor and transient events. In any case, the stability of the location and frequency imply a sustained generation process of the signal.

On the other hand, Palma et al. [2008] described a complex and not always straightforward interplay of average amplitude level (RSAM) as an indicator of tremor, the number of transient events, degassing and visible activity of the lava lake. Most importantly, these authors reported a lack of correlation between the event rate and RSAM for the period from November 2004 to April 2005. Calder et al. [2004] even reported the absence of strombolian explosion quakes during a tremor period in August-December 1999. Both observations rather indicate two different source processes for the tremor and the transient events. A more detailed investigation of individual events might provide a better insight into this issue.

## 4.6. Conclusion

In early March 2012, Villarrica Volcano showed a continuous low-frequency tremor signal with overlain, transient high-amplitude bursts in intervals of approximately 60 s. Using three independent methods (beamforming with intersection, modeling of amplitude decay, amplitude mapping) we could demonstrate that the epicenters of the signal components as well as of the composite signal as a whole were located in the same area inside the summit crater and close to the earth's surface. By dispersion analysis it could be shown that the wavefield produced by the seismicity was composed mainly of surface waves. This is consistent with a shallow source depth. The uncertainty of the epicentral coordinates is of the order of some 100 m. Using a moving-window analysis, we could show that the source position was stable during the nine days of observation. The probabilistic evaluation of the power spectral density revealed an equally stable frequency

content of the seismic signal. Together with the identical epicenters of transient events and tremor this may indicate an intimate link between the different characteristics of the low-frequency seismic signal. Yet, whether this indicates a common (or even identical) source process remains disputable.

Aside from the origin, we obtained  $S$  wave velocities at the northern and western flank and attenuation parameters of the Villarrica volcano from the location methods. The velocities of the surface layer are in accordance with the site structures: the low velocities at the western flank correspond to the fresh, poorly consolidated ashes, while the high velocities at the northern location are caused by the lava flow that forms the Cuevas Volcanicas. Between 100 m and 400 m, both locations exhibit very similar velocities around 1200 m/s, corresponding to pyroclastic deposits and ignimbrites from earlier eruptions. The average seismic quality factor  $Q$  is of the order of 50 at around 2 Hz in the uppermost 200 to 300 m.

In comparison with earlier reports on the seismic and volcanic activity of Villarrica ([Ortiz et al., 2003, Calder et al., 2004, Palma et al., 2008, Richardson et al., 2014]), we believe that our data represent more or less a common behavior of the volcano. Therefore, the source location, that was found in this study, may be considered representative for this type of activity. Nevertheless, we point out, that our observation period comprises merely two weeks, and hence inferences to other periods should be made with care. The medium parameters on the other hand are less affected by this limitation unless significant changes happen to the volcanic edifice. They may provide a solid base for a closer look at individual events which could reveal interesting details about the dynamics inside the conduit.

## 4.7. Appendix

### 4.7.1. Probabilistic intersection

We realize this concept using a probabilistic, Bayesian framework. Given the DOA measurement of a signal by an antenna  $A_i$  at  $(x_i, y_i)$  and an estimate of its uncertainty, we may derive a continuous probability density function (PDF)  $f_i(\varphi)$ . Thus, we know the likelihood of any direction to be the DOA of the detected signal. A signal from a source location  $L_j = (x_j, y_j)$  impinges at  $A_i$  from a direction

$$\varphi_{ij} = g(x_i, y_i, x_j, y_j) = \arctan\left(\frac{x_j - x_i}{y_j - y_i}\right) \quad (4.8)$$

This relation allows us to quantify the likelihood that  $L_j$  is the source of the signal detected at  $A_i$  given that  $f_i(\varphi)$  represents the distribution of the results of the DOAS measurements.



Bayes theorem describes the probability of hypothesis  $H$ , given an observation  $O$ . We postulate that  $H : L_j$  is a source of the signal detected at antenna  $A_i$  while  $O$  : we observed an incoming signal with  $f_i$  at  $A_i$ .

Eventually, we are going to compare a number of  $L_j$  values drawn from a set of  $N$  possible locations. We hereby assume that these hypotheses are exclusive (there can be only one source) and exhaustive (one  $L_j$  in the set must be the source), so that we need the generalized form of Bayes' theorem (e.g. Bernardo [2011]):

$$P(H_j | O) = \frac{P(O | H_j) P(H_j)}{\sum_{l=1}^N P(O | H_l) P(H_l)} \quad (4.9)$$

Instead of probabilities one may also use probability densities.

The probability to observe a DOA  $\varphi_{ij}$  if  $H_j$  were true, is given by  $f_i$ , thus:

$$p(O | H_j) = f_i(\varphi_{ij}) \quad (4.10)$$

Concerning the hypotheses, we may assume  $P(H_j) = \text{const.}$  because without any further knowledge all  $L_j$  are equally likely to be the source. Consequently,  $P(H_j)$  is eliminated in Eq. 4.9.

Now we want to exploit that we actually have multiple observations  $O_i$  of the signal. These observations are statistically independent since one does not require another antenna to detect a signal at  $A_i$ . Therefore, the joint probability for making all observations at the same time is given by the product of the individual probabilities and we may extend the above to [Bernardo, 2011]:

$$p(O_1 \cap \dots \cap O_M | H_j) = \prod_i^M P(O_i | H_j) = \prod_i^M f_i(\varphi_{ij}) \quad (4.11)$$

Moreover, we may treat the set of observations  $\{O_i\}$  equivalently to a single observation, so that Eq. 4.9 can be extended by induction and combined with Eq. 4.11 to:

$$\begin{aligned} P(H_j | O_1 \cap \dots \cap O_M) &= \frac{P(O_1 \cap \dots \cap O_M | H_j)}{\sum_{l=1}^N P(O_1 \cap \dots \cap O_M | H_l)} \\ &= \frac{\prod_i^M P(O_i | H_j)}{\sum_{l=1}^N (\prod_i^M P(O_i | H_l))} \\ &= \frac{\prod_i^M f_i(\varphi_{ij})}{\sum_{l=1}^N (\prod_i^M f_i(\varphi_{il}))} \end{aligned} \quad (4.12)$$

#### 4.7.2. Site amplification estimation

Site amplification factors are needed to correct the amplitudes of the volcanic signals for the influence of near-surface geological layering. We determined a relative site amplification factor for each station following the procedure of Bonilla et al. [1997] for S-waves.

However, since most of our instruments were single-component sensors we considered only the vertical component. The procedure was applied to the frequency band between 1 Hz and 3 Hz.

The amplitude spectrum  $A(\omega)$  of the observed ground motion of an earthquake can be expressed by:

$$A(\omega) = E(\omega) \cdot P(\omega) \cdot S(\omega) \quad (4.13)$$

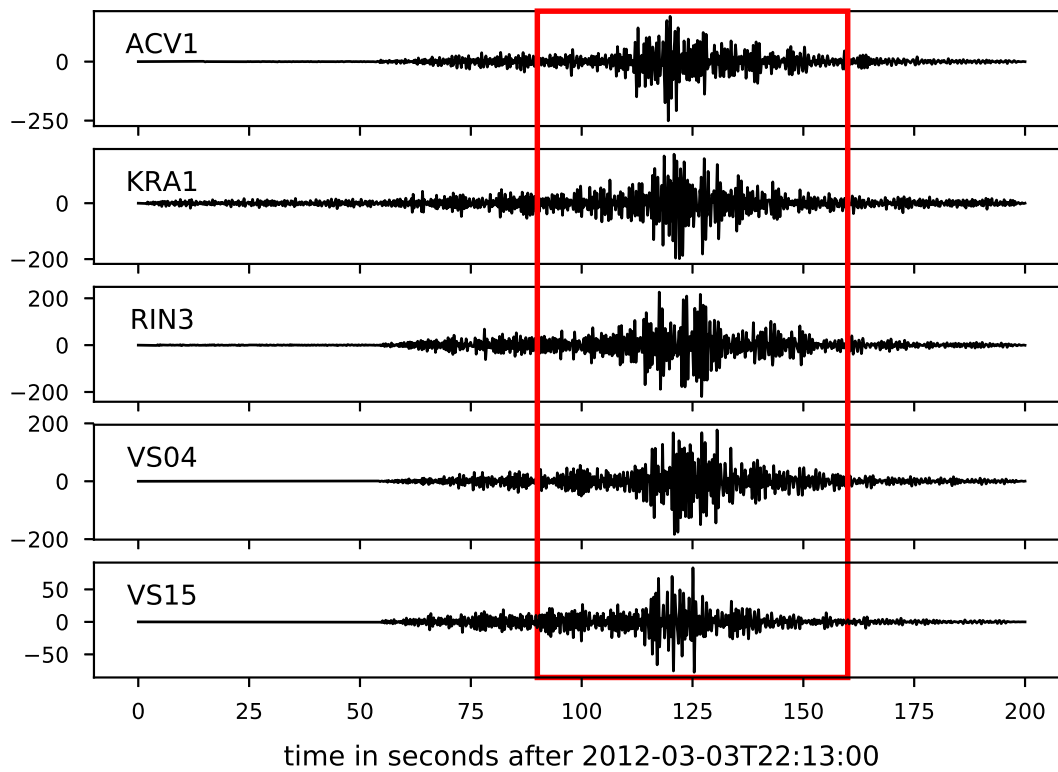
$E(\omega)$  is the amplitude spectrum emitted from the hypocenter of the earthquake (moment, size, radiation pattern),  $P(\omega)$  is the alteration of the amplitude spectrum due to the travel path (e.g. geometrical spreading, damping) and  $S(\omega)$  is the amplification due to local structure at the recording station (site effect). The site amplification factor relative to a reference station is:

$$\frac{S(\omega)}{S_{ref}(\omega)} = \frac{E_{ref}(\omega)}{E(\omega)} \cdot \frac{P_{ref}(\omega)}{P(\omega)} \cdot \frac{A(\omega)}{A_{ref}(\omega)} \quad (4.14)$$

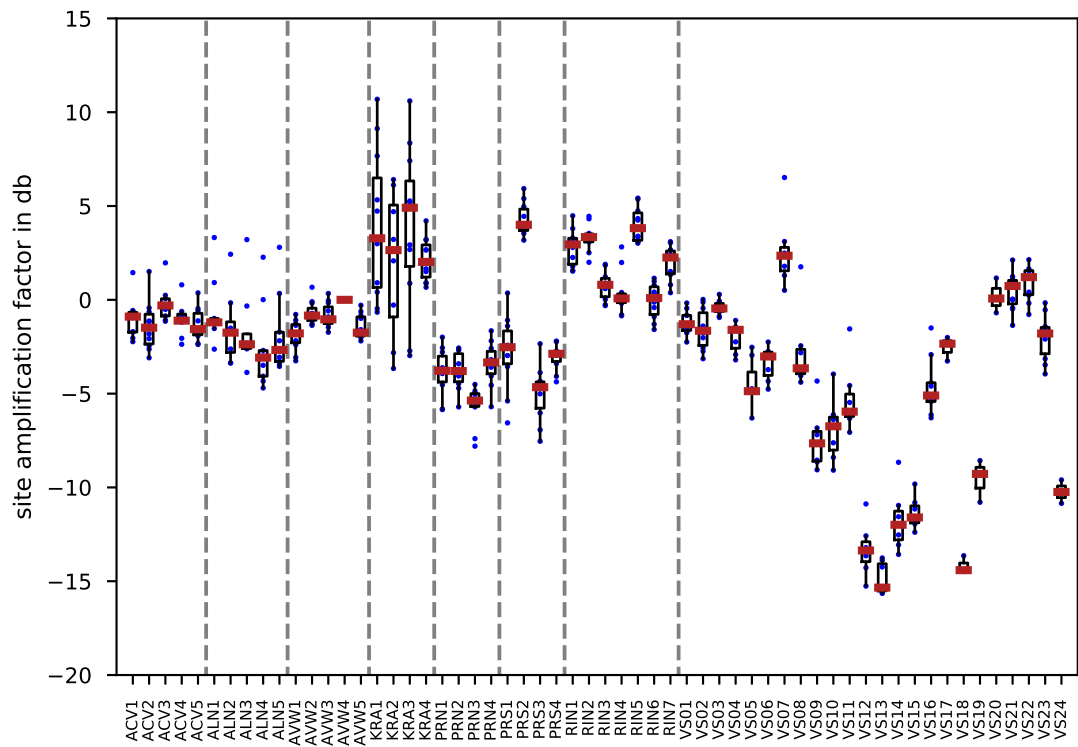
To derive a relative site amplification factor we selected eleven regional earthquakes (Table 4.2) showing a strong S-wave arrival. In the analysis, the following steps were performed: First, the waveform of the vertical component was corrected for the instrument response to obtain the ground motion of the seismic signal. Secondly, the data were filtered with a bandpass between 1 Hz and 3 Hz and the signal energy was determined for a time window around the S-wave arrival (Fig. 4.12). The energy values were normalized by the signal energy of a reference station (in our case station AVW4) and expressed in dB.

We assumed that the amplitude spectrum of the incoming wave  $E(\omega)$  is approximately the same for all stations, since the radiation angles differ by a maximum of only 3 degrees. This is because lateral extent of the network of 20 km is small compared to the epicentral distance of more than 200 km (Table 4.2). Therefore, the correction for differences in the emitted source spectrum  $E_{ref}(\omega)/E(\omega)$  can be neglected. For the same reason we assumed that the differences in the alteration  $P(\omega)$  of the amplitude spectrum caused by differences in the travel path are so small that they can be neglected.

The site amplification factor of each station was then determined by computing the median of the normalized energy values derived from the selected eleven regional events (Fig. 4.13). For comparison we also determined the site amplification factors by the more popular coda wave method [Aki and Chouet, 1975, e.g.] but found no significant differences between both methods for our data set.



**Figure 4.12.:** Regional event  $M_L=4.9$  on 2012-03-03 at 22:12:56.11. Waveforms are instrument corrected to ground velocity (in  $\mu\text{m/s}$ ) and filtered with a bandpass between 1 Hz and 3 Hz. Red rectangle indicates the time window for the energy determination.



**Figure 4.13.:** Site amplification factors in db. Blue dots indicate every observation, red stars and the lightblue box the median and the 25%-75%-quantiles.

**Table 4.1.:** Source location in UTM zone 19S

Method	x	y
Array analysis	246901.7 $\pm$ 500 m	5632237.7 $\pm$ 905 m
Ampl. interpolation	246851.7 $\pm$ 0 m	5632387.7 $\pm$ 150 m
Ampl. decay	246761.7 $\pm$ 210 m	5632387.7 $\pm$ 150 m

**Table 4.2.:** Regional events

Event-ID	Date	Time	Latitude	Longitude	Depth	ML	Distance
600794002	2012-03-01	06:44:26.53	-38.2115	-73.6291	25.8	4.2	199 km
604845792	2012-03-01	20:33:28.90	-36.782	-73.57	22.7	3.8	326 km
604845802	2012-03-02	23:11:19.30	-37.746	-74.745	20.0	3.9	307 km
600954421	2012-03-03	00:52:38.92	-36.5418	-72.47	51.3	3.9	323 km
600760711	2012-03-03	22:12:56.11	-35.6975	-72.8917	20.6	4.9	422 km
600954498	2012-03-03	23:20:14.65	-37.9104	-75.1229	35.0	3.9	324 km
600760714	2012-03-03	23:43:04.58	-35.7061	-72.8923	20.2	4.8	421 km
604845822	2012-03-05	08:06:54.37	-36.215	-73.2807	10	3.9	375 km
600791120	2012-03-09	00:43:33.40	-34.7027	-72.8732	2.6	4.3	531 km
604845863	2012-03-11	03:30:02.80	-37.562	-74.11	40.4	3.6	280 km
600778273	2012-03-12	19:37:36.19	-34.961	-71.6532	66.1	4.9	496 km



---

## 5. Magnitude and interevent time statistics of transient events and inference regarding the flow regime

Parts of this chapter have been submitted as:

*J. Lehr & W. Rabbel*

*Magnitude and interevent time statistics of Strombolian activity of Villarrica Volcano and inference regarding the flow regime*

*Journal of Geophysical Research: Solid Earth, 2021*

### **Main results / conclusions:**

- Magnitudes of explosions do not obey GR-law, they have a peak magnitude
- Interevent times indicate a slightly periodic recurrence.
- Comparison of the statistical distributions with size distributions of gas bubbles flowing in liquid suggest a sustained slug flow regime for the upper conduit.

### **Abstract**

*The different flow regimes of two-phase gas flow - as it occurs in the shallow conduit of basaltic volcanoes - are characterized by distinct frequency-size distributions of the liquid and gaseous slugs. Assuming that the ascent of gaseous bubbles is indicated by seismic events, we explore the possibility to infer the flow regime from the frequency distributions of magnitudes and interevent times. Our data set consists of 20,000 volcanic seismic events recorded in early March 2012 at Villarrica Volcano (Chile). They are attributed to Strombolian activity. One crucial factor is the completeness of the catalog in terms of detectable amplitudes which we assess using a stochastic simulation of the network output based on statistical properties of the ambient seismicity.*

*Magnitudes show an exponential occurrence in the complete part. Yet, the simulation approach indicates that low magnitude events occur indeed more sparsely than expected for an exponential distribution, and that the magnitude distribution does not obey the Gutenberg-Richter law. Interevent times between events in the complete part of the catalog suggest a log-normally distributed occurrence of events which implies a preferred recurrence interval. The interevent time possibly correlates with the preceding magnitude. Our findings suggest a sustained slug flow regime, although more factors should be taken into account for a definite answer.*

## 5.1. Introduction

Magma flow and degassing in open-vent volcanoes is the result of a complex interaction between liquid rock and the gas phase. When magma rises inside a conduit, gas bubbles start to nucleate once the lithostatic pressure is low enough and a separated two-phase flow develops consisting of gas bubbles in liquid magma [Sparks, 1978, Jaupart and Vergnolle, 1988, Parfitt and Wilson, 1995, Seyfried and Freundt, 2000, Parfitt, 2004, Shinohara, 2008, Burgisser and Degruyter, 2015, Pering et al., 2017, Spina et al., 2019].

These bubbles become visible at volcanoes when they burst at the surface and produce a Strombolian-type activity. In addition, the gas release can be measured directly with chemical detectors and optical sensors, also indicating intermittent degassing [Shinohara, 2008, Nadeau et al., 2011, Carn, 2015, Pering et al., 2015, Gaudin et al., 2017a, Pering et al., 2019a]. Understanding how these surface observations relate to the processes inside the conduit could be beneficial for volcano monitoring and ultimately hazard management.

Since direct observation of the conditions in the magma column is impossible, analogues with technical studies have typically been used to characterize the flow regime. Laboratory experiments suggest specific mechanisms by which the gas pockets rise inside the liquid phase [Jaupart and Vergnolle, 1988, Vergnolle and Jaupart, 1986, Seyfried and Freundt, 2000, James et al., 2004, 2006, Spina et al., 2019]. Different flow regimes display distinct statistical distributions of bubble size and eruption frequency [Barnea and Sh emer, 1989, Barnea and Taitel, 1993, Liu and Bankoff, 1993, Dukler and Fabre, 1994, Mudde, 2005]. Thus, comparing the statistical distribution of gas bursting events at a volcano to the theoretical expectation might suggest or contradict a flow regime inside the conduit.

Typically Strombolian explosions, i.e. gas bursts are accompanied by seismic events [Neuberg et al., 1994, Ripepe et al., 2001, Chouet, 2003, Zobin, 2017a,b, Spina et al., 2019, Ishii et al., 2019]. Therefore, seismology might provide alternative or additional information about the flow regime to surface observations. However, similar waveforms may occur without visible activity as well as bubble bursts without seismic expression were observed, suggesting that this relation might be more complicated [Spina et al., 2019].

Statistical analysis of the inter-event time distribution of seismic transients and their magnitudes might be a more robust way of comparing seismic activity to theoretical flow models. The main question is, whether seismic transients exhibit ‘random’ behaviour or whether they repeat with certain preferred frequencies.

Previous investigations of inter-event times of volcanic events found some evidence for periodicity Ripepe and Gordeev [1999], Ripepe et al. [2001, 2002], Varley et al. [2006], Martino et al. [2012], Taddeucci et al. [2013], Pering et al. [2015], Dominguez et al. [2016], Bell et al. [2017, 2018], Pering et al. [2019a]. However, the validity of these results is sometimes questionable because the events might not have been detected correctly. The



incompleteness of seismic event catalogues towards low magnitudes is a well-known problem in tectonic seismology and a number of methods have been established to determine the magnitude of completeness of a catalogue [Mignan and Woessner, 2012]. However, most approaches assume that magnitudes follow the Gutenberg-Richter relation, which is a power-law. While this is a well-established relation for tectonic earthquakes it might be unsuited to magma dynamics in a volcano.

In this work, we characterize the occurrence of seismic transient events at Villarrica volcano in terms of the statistical distribution of interevent times and magnitudes. Villarrica is an active, basaltic to andesitic strato-volcano in Chile. Its central summit crater hosts an active lava lake which is the source of persistent degassing and frequent mild Strombolian explosion. The seismicity consists of a notorious tremor overlain by transient events which are attributed to the explosions [Palma et al., 2008, Gurioli et al., 2008]. We compiled a catalogue of these transients using an STA/LTA trigger from 12 days of continuous seismic records from two stations at the crater rim. We then applied a Monte-Carlo simulation to estimate the detection limit of our seismic stations [Ringdal, 1975, von Seggern and Blandford, 1976]. Our method even allows estimating the true distribution of magnitudes.

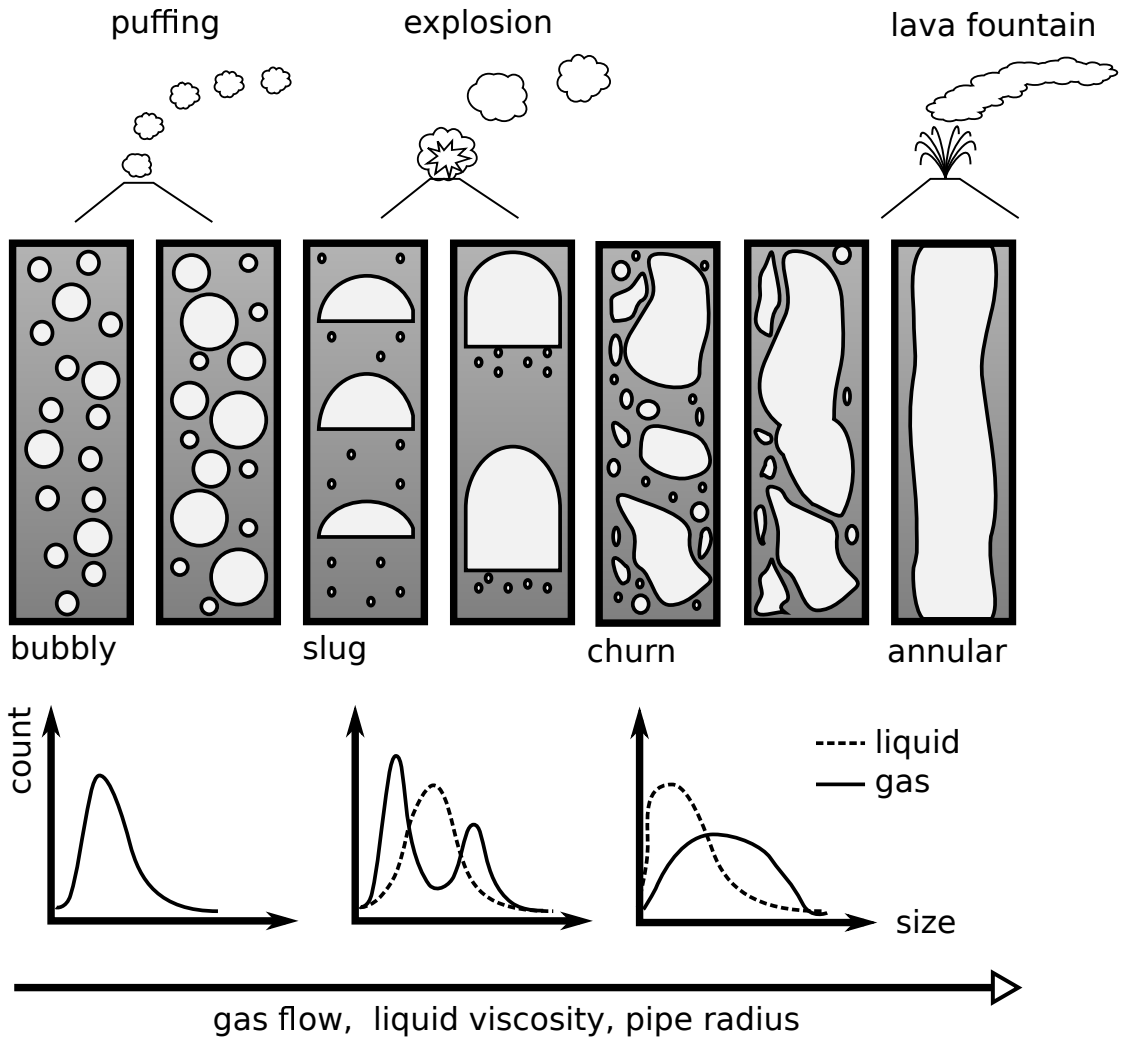
The paper is structured as follows: At first, we review some relevant aspects of gas flow and volcanic activity. After introducing the volcano and the seismic network, we explain the event detection process and the determination of magnitudes, which provides the data base for our analysis. Next, we present how we estimated the magnitude completeness. Subsequently, we analyze the magnitudes and interevent times of the complete part of the event catalog. Finally, the results are discussed in comparison to studies at other volcanoes and in the context of flow regime experiments.

## 5.2. Flow regimes and statistical distributions of bubble sizes

Laboratory experiments of vertical two-phase flow showed that the size distribution of the gas bubbles depends on the flow regime [Barnea and Shémer, 1989, van Hout et al., 1992, Barnea and Taitel, 1993, Dukler and Fabre, 1994, van Hout et al., 2001]. In these experiments, small gas bubbles are released at the bottom of a liquid-filled pipe and their velocity and size during their way to the top of the column is measured. Depending on the gas flow rate and viscosity of the liquid different flow patterns arise.

We adopted the technical terminology in which the term “bubble” is used generically for gas pockets of any size. A “slug” in contrast has approximately the same diameter as the pipe and may be longer than wide in which case it is sometimes denoted as “Taylor bubble”. Sometimes the liquid sections between the gas slugs are referred to as “liquid slugs”.

Fig. 5.1 summarizes characteristics of the different flow regimes. Bubbles in **bubbly flow** are quite similar in size resulting in a narrow unimodal distribution. **Slug flow**



**Figure 5.1.:** Flow regimes with associated distributions of bubble sizes and liquid bridges as well as typical volcanic activity. Dark and light colors indicate liquid and gaseous phase, respectively.

exhibits a bimodal distribution. The peak at larger sizes represents the gas slugs while the peak around smaller sizes originates from the small bubbles which trail in the wake of the slug and are dispersed in the liquid [Barnea and Shémer, 1989]. The lengths of the slugs (Taylor bubbles) and the liquid packages between them can both be modeled e.g. by a log-normal distribution [van Hout et al., 1992, 2001]. At the transition to **churn flow**, the distribution of liquid slug lengths becomes much more right-skewed [van Hout et al., 1992] while the peaks of the gas slug size distribution merge into a broad, unimodal distribution [Barnea and Shémer, 1989]. In **annular flow**, the gas flows continuously through the annularly displaced liquid.

The transition from bubbly over slug to churn and finally annular flow occurs with increasing viscosity of the liquid, higher superficial gas velocities (which is the volumetric gas flux divided by the cross-sectional area of the pipe) and the pipe diameter [Barnea and Shémer, 1989, van Hout et al., 1992, Ohnuki and Akimoto, 2000, Spina et al., 2019].

It is quite well-established that the different aspects of Strombolian activity are the result of differently sized gas releases [Jaupart and Vergnolle, 1988, Ripepe and Gordeev, 1999, Seyfried and Freundt, 2000, Parfitt, 2004, Pering et al., 2017, Gaudin et al., 2017a, Spina et al., 2019]. Puffing - the intermittent release of small amounts of gas - corresponds to relatively small bubbles, while explosions are produced by larger slugs. Sustained lava fountains are produced by temporary annular flow. While the dynamic and eruption process of single gas bubbles are reasonably well understood, the connection between the continual eruption process and the gas flow regime in the conduit is subject of current research [Pioli et al., 2012, Pering and McGonigle, 2018, Spina et al., 2019, Ishii et al., 2019].

## 5.3. Event detection

### 5.3.1. STA/LTA network trigger

Transient events were detected using an short-term-average/long-term-average (STA/LTA) trigger [Withers et al., 1998] on stations KRA1 and KRA3. Both stations needed to trigger to declare an event (network coincidence trigger). The trigger function at each station consisted of the ratio between a short-term and a long-term moving mean of the squared amplitude. The ratio at each sample was computed from the LTA window terminating and the STA window starting at that sample to ensure statistical independence between the two windows.

The window lengths were chosen based on the recommendations by Amadej [2009]. The STA window should be long enough to capture a few periods of the lowest frequency. We chose 4s which yields two periods for the lowest frequency of 0.5 Hz. The LTA window determines whether the trigger is more sensitive to emergent or sharp onsets. The longer the LTA window the more sensitive is the trigger to emergent onsets. The seismicity contained transient events with both a very sharp onset and short duration as well as

longer lasting, emergent events. To capture this broad variety, we used multiple LTA window lengths of 10, 12, 16, 24, 32, 48, 64s and combined the catalogs afterwards. This was done by a modified algorithm of the network coincidence trigger in which each LTA-catalog was treated as a station. We required a detection for at least 3 LTA window sizes to include the event in the final catalog. A data sample demonstrating the different catalogs is shown in Fig. S2. The trigger onset threshold was set to 2.0 and the dettrigger threshold to 1.0.

Six regional tectonic events, visibly affected the data of the network. Therefore, any transient events occurring during these earthquakes were removed from the catalog. The final catalog contained 23505 events.

### 5.3.2. Amplitude and magnitude

For the final catalog, peak-to-peak velocity amplitudes of the z-component were extracted for each event at both stations. Magnitude scales in general link the amplitude, measured at a station and the distance to the event to provide a station-independent measure of the strength of an event. This is generally a linear relationship between the logarithm of the observed maximum displacement amplitude  $\log(A)$  and distance  $r$  of type  $M_L = \log(A) + a \log(r) + br + c$  (Havskov, 2010). Typically, the half peak-to-peak amplitude  $A^{(PTP)}$  is used. Note, that we used amplitudes of particle velocity as  $A^{(ptp)}$  rather than displacement amplitudes. Factors  $a, b, c$  are specific to each station and can be determined by least-squares inversion. In our case, we assumed that the events all originated from pretty much the same location, namely the lava lake [Lehr et al., 2019]. Therefore, the distance-related terms could be omitted and a makeshift local magnitude was determined as:

$$m_j = \log\left(\frac{1}{2}A_i^{(ptp)}\right) - c_i$$

We inverted:

$$\log\left(\frac{1}{2}A_i^{(ptp)}\right) = m_j + c_i$$

To keep the computational load at bay,  $m_j$  and  $c$  were determined from randomly selected 1500 events. This was repeated 50 times. If an event was selected multiple times the mean over all available values was taken as the "inverted magnitude". The site parameters  $c$  were determined from the mean over all runs. Eventually, we computed the magnitudes of all events as the mean of the converted amplitudes over  $N$  stations:  $m_j = \frac{1}{N} \sum_i^N (\log(A_{ij}^{(ptp)}) - c_i)$  with  $c_i = -10.29$  for station KRA1 and  $c_i = -10.35$  for KRA3.

## 5.4. Magnitude of completeness

No statistical analysis of event amplitudes is complete without knowing the magnitude of completeness (MoC) of the catalog which is the lowest magnitude at which all events

were detected. Many methods of determining the MoC work solely on the data in the catalog and rely on the assumption of exponentially distributed magnitudes. However, we preferred to not exclude other possibilities at this stage. Instead, we resolved to a stochastic simulation of the STA/LTA trigger and the network trigger. The result was corroborated by comparing the magnitude distribution (Gutenberg-Richter plots) of events at different levels of background activity.

#### 5.4.1. Monte-Carlo simulation of trigger

Ringdal [1975] and von Seggern and Blandford [1976] suggested to use a stochastic simulation in order to determine the detection capabilities of a network given the statistical properties of the data and a magnitude distribution model. Many practical details are explained in von Seggern [2004]. For this purpose, the stochastic properties of STA and LTA and their relation between each other and to the magnitude of an event were determined empirically from the real data. The STA/LTA trigger was then simulated by creating random pairs of LTA and STA values for a random set of magnitudes for each station. The network was simulated from a logic combination of the stations. The detection probability of any given magnitude resulted from the ratio of detected to simulated events of that magnitude. In addition, we applied a logistic regression to determine the detection curve.

This approach allows to represent relatively complicated networks and statistical parameters. Remember that we combined the catalogs of a series of LTA windows while each catalog being the result of a physical 2-station network. For the simulation experiment, we treated this situation as a network of virtual stations, that represent a combination of a real station with an LTA window length.

#### LTA samples

A representative sample of the entire observation period is needed to derive a statistical description of the LTA amplitudes (=“noise level”). For each LTA window size, 5-min long samples were extracted at the beginning of every hour from the LTA-processed data (i.e. mean-squared amplitudes over the LTA window).

For each LTA window, the sample was approximately log-normally distributed, even though we note that the Kolmogorov-Smirnov test for similarity failed. The log-normal data model is however frequently used because it allows to incorporate the correlation between stations and LTAs. Hence, the mean and standard deviation of the logarithmic data were determined for each window size.

The event catalog was composed of events that were detected at different stations and at different LTA window lengths, mimicking a network coincidence trigger. Obviously, the noise level correlates between stations and also between the different average lengths, which should be reflected in the simulation. The generation of correlated normal random

variables is fairly easy and requires merely the correlation matrix between all involved variables (Fig. B.4).

Let  $\mathbf{R}$  be the matrix containing all pairwise correlations  $r_{ij}$ . The correlated samples  $Y$  can be generated using the transformation:

$$y = \mathbf{A}x$$

with  $x$  being an uncorrelated multivariate normal distribution and  $\mathbf{A}\mathbf{A}^t = \mathbf{R}$ .  $\mathbf{A}$  can be determined from the Cholesky decomposition of  $\mathbf{R}$ .

For our log-normal data, the actual LTA values were then obtained by transforming

$$y_i = \exp(\sigma_i z_i + \mu_i)$$

### STA samples

The magnitude of an event converts to an STA value at each of the stations. We found that the magnitudes correlated approximately linear with the logarithmic maximum STA value ( $a_{ij}$ ) at any given station during the event. Thus the magnitude determines the observed STA as:

$$a_{ij}^{sta,max} = m_j q_i + b_i$$

### Detection curves

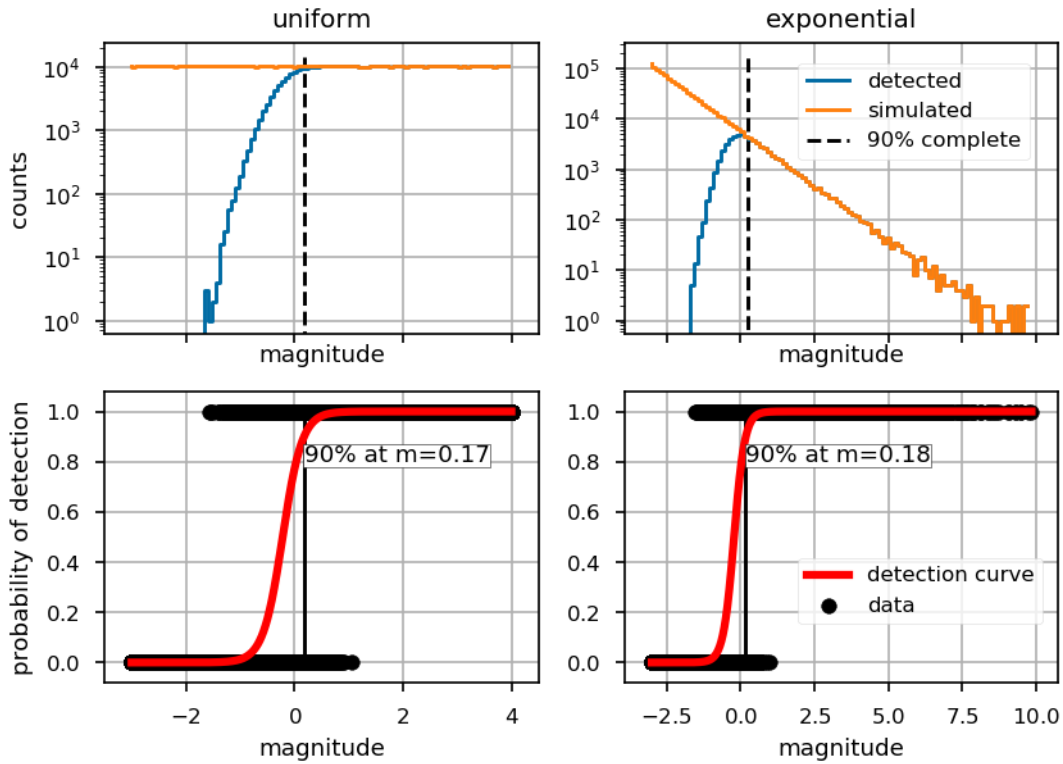
The detection capability of the network was tested on two different sets of simulated events: one with an exponential and one with a uniform distribution of the magnitudes. Each set consisted of 100000 events (Fig. 5.2). For the determination of the threshold, we applied a logistic regression to obtain a detection curve, that gives the probability of detecting an event of magnitude  $m$  as:

$$p(m) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 m)}}$$

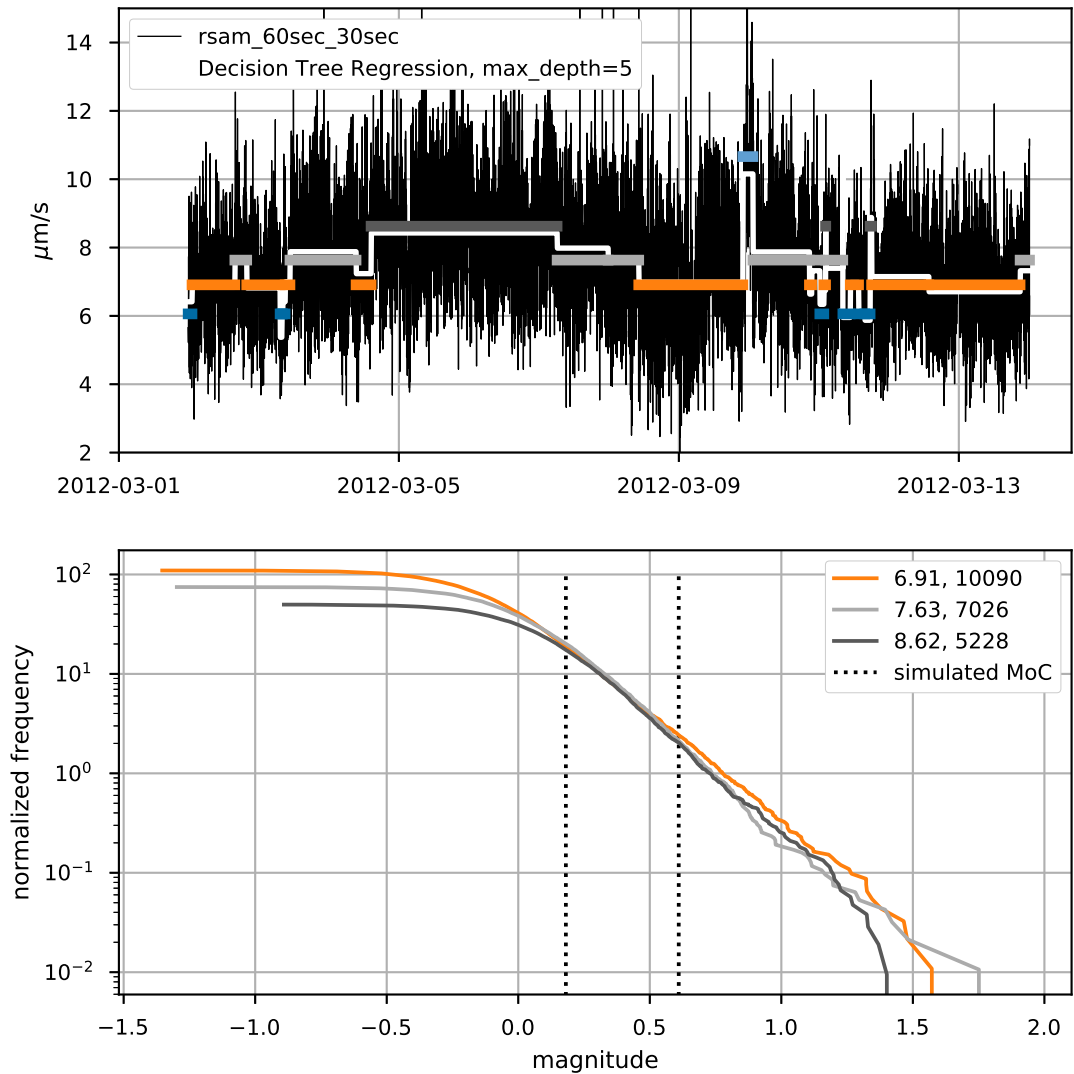
Using the definition of the odds  $O = \frac{p}{1-p}$ , the inverse of the above equation gives the magnitude at which a certain level of completeness is reached:

$$m(p) = \frac{\log(O(p)) - \beta_0}{\beta_1}$$

Each model was repeated 100 times resulting in a mean and standard deviation of the logistic parameters as:  $\beta_0 = 1.194 \pm 0.010$  and  $\beta_1 = 5.532 \pm 0.026$  for the uniform magnitude model and  $\beta_0 = 1.197 \pm 0.013$  and  $\beta_1 = 5.535 \pm 0.025$  for the exponential model. Hence, on average the parameters are identical for both models. Using the values of the exponential model, one obtains threshold magnitudes of  $m = 0.181$  and  $m = 0.614$  for 90% and 99% completeness of the catalog.



**Figure 5.2.:** Stochastic simulation of event detection using a uniform (left) and exponential (right) distribution of simulated magnitudes. Top panels give histograms of all simulated (orange) and detected (blue) magnitudes. Bottom panels show detection curves (red) derived by logistic regression between detected and undetected events (black).



**Figure 5.3.:** Magnitude-frequency distributions for events in sections of similar amplitude. *Top:* The thick white line marks segments of constant RMS-amplitude (black) found by Decision Tree Regression. Colored lines represent the clusters of similar amplitudes. *Bottom:* Gutenberg-Richter plots of the magnitude-frequency relation for events in sections of similar amplitude with more than 1000 events. Number of events is given in legend. Colors correspond to those in top panel. Counts are normalized to the number of events in  $m = [0.6, 0.7]$ . Dotted lines indicate the threshold magnitudes from the MC simulation. The distribution at the highest amplitude level (dark grey) branches off at approximately the simulated 90%-completeness magnitude.



### 5.4.2. Sections of similar amplitude

We also derived a magnitude of completeness directly from the catalog (Fig. 5.3). To this purpose, we identified sections of similar amplitude levels using the root-mean-square amplitude in 60s-intervals shifted by 30s. At first, a Decision Tree Regression (scikit-learn) was used to find discrete sections of constant amplitude. The resulting finite set of unique amplitudes was then further simplified by finding similar amplitudes using a KMeans-algorithm.

## 5.5. Statistical analysis of event catalog

We investigated mainly two parameters of the event catalog for their statistical distribution: the magnitude of the events and the difference between the start times of the two consecutive events - the interevent time.

The coefficient of variation is a measure of the dispersion of a distribution the standard deviation of the data divided by its mean:  $C_v = \sigma/\mu$ . Applied to the interevent times, it is a common measure in earthquake statistics to discriminate between a rather periodic ( $C_v < 1$ ) process, a rather clustered ( $C_v > 1$ ) or a completely random one [Bottiglieri et al., 2005, De Lauro et al., 2009, Martino et al., 2012, Bell et al., 2017]. The latter is known as Poisson process. We computed  $C_V$  of the interevent time between events within 12-h long, adjacent time frames. Using the same binning, we also counted the number of events and computed the mean interevent time, the inverse of which is the rate.

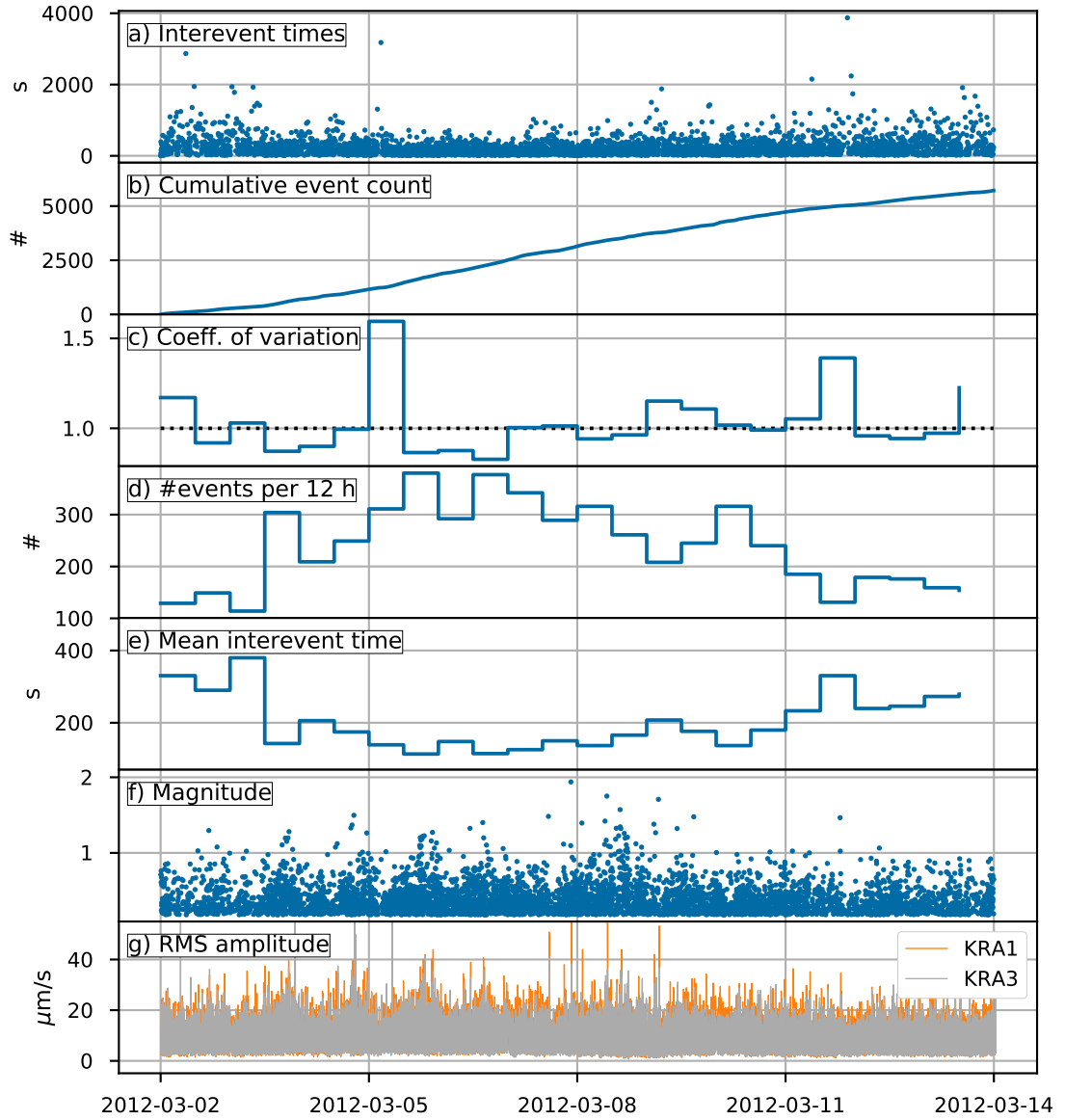
We first look at the occurrence of the parameters in time before we move to the analysis of the statistical distribution of magnitudes and interevent times.

### 5.5.1. Time series

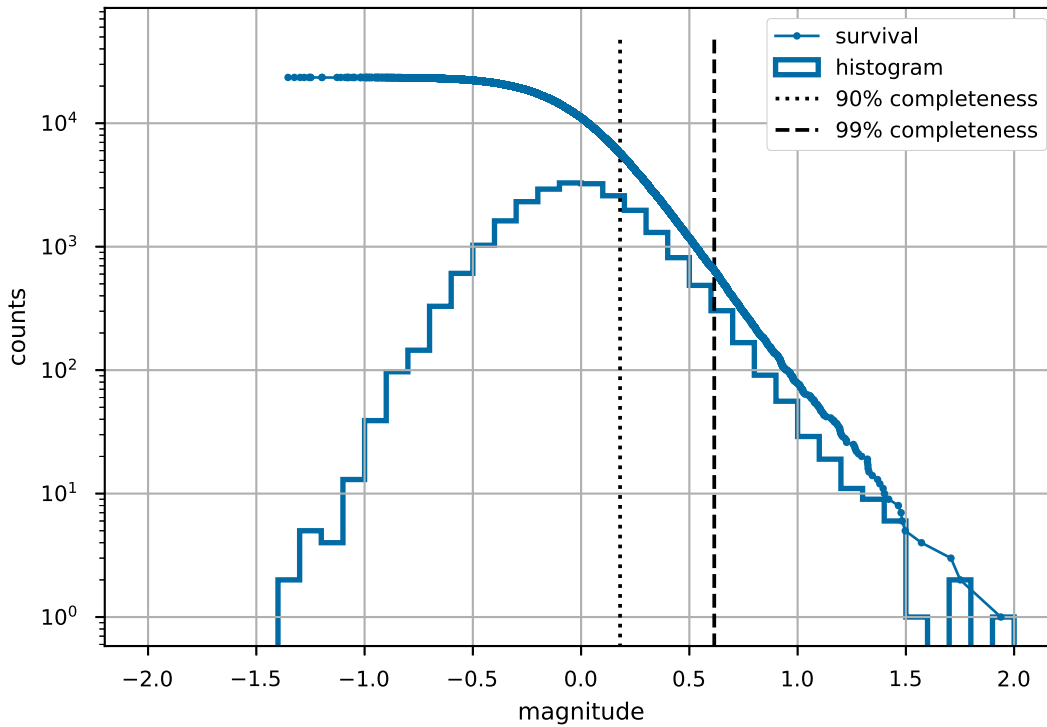
The time series are shown in Fig. 5.4. The occurrence of higher magnitudes is reflected in the RMS-amplitude at the crater stations (Fig. 5.4f,g). The coefficient of variation oscillates around 1 with a mean of 1.03 and a standard deviation of 0.17. Extreme values occur solely in the clustered regime. High  $C_V$  are associated with interevent times close to or larger than 2000s. Their occurrences are not associated with one of the removed regional earthquake which could possibly limit event detection due to their coda.

The first two and last three days seem to have seen fewer events than the days in between, resulting in larger interevent times toward the start and end of the observation period (Fig. 5.4a,d,e). This is reflected in the slight changes in the slope of the accumulated number of events (Fig. 5.4). The  $C_V$  seems to be unaffected.

These changes might possibly indicate changes in the eruption process and non-stationarity in the data.



**Figure 5.4.:** Time series representation of magnitude, interevent time and related parameters. Parameters in c)-e) were computed in 12 h bins. The RMS amplitude in g) was computed every 2s in 4s-windows. It is clipped at  $49 \mu\text{s}/\text{m}$  to highlight the similarity with the magnitudes. The dotted line in c) marks a  $C_V = 1$  which indicates a Poisson process.



**Figure 5.5.:** Frequency-magnitude distributions of detected transients. “survival” shows the number of events with  $M > m$  against  $m$  and is the traditional representation in earthquake statistics [Orfanogiannaki et al., 2010]. It corresponds to an empirical survival function. “histogram” provides an approximation of the probability density function.

### 5.5.2. Magnitudes

The magnitudes of the transient events at Villarrica range from -1.5 to 2.0 (Fig. 5.5). The distribution resembles those of tectonic earthquakes with an almost horizontal part towards small magnitudes, a linear slope towards higher magnitudes and the roll-off where the two parts connect. For tectonic earthquakes, the roll-off and flat part result from the incompleteness of the catalog at low magnitudes. The slope of the linear part corresponds to the actual exponential distribution of the magnitudes (Gutenberg-Richter law, [Gutenberg and Richter, 1956]) and in fact continues at small magnitudes. The completeness threshold of the catalog lies at higher magnitudes than the roll-off and beyond the mode, respectively. Hence, for the complete catalog, the magnitudes are exponentially distributed. The catalog contains 5697 events above  $m = 0.181$  (90% completeness) and 643 events above  $m = 0.615$  (99% completeness). The histogram indicates a unimodal distribution, reflecting the lack of events at small magnitudes and

fewer events at high magnitudes.

### Discarding the exponential model

The complete part of the catalog suggests an exponential distribution of magnitudes. We tested this hypothesis by deriving the exponential distribution that fits the complete part of the catalog and using it in the detection simulator to see whether the decline in detected events could be reproduced. If the exponential distribution continued at low magnitudes, the simulated detected events should correspond to the observed distribution. The rate of the exponential distribution was determined using the maximum-likelihood estimate  $1/\lambda = \frac{1}{N} \sum (m - m_0) = 5.507$  with  $m_0 = 0.615$  being the completeness magnitude and  $N$  being the number of events with  $m \geq m_0$ . Since the number of events that needed to be simulated to match the observed complete part was too large to be run at once, the simulation was run repeatedly until the number of simulated events with  $m > M_0$  was larger than that of the observed ones. The simulation clearly detected more events than were observed (Fig. 5.6) suggesting that the true magnitude-frequency distribution of the transients tends to smaller values at the low magnitude end than the exponential Gutenberg-Richter distribution.

### Deriving a magnitude model

The detection curve  $h(x) = 1/(1 + \exp(-kx - x_0))$  tells us how many events were missed at any given magnitude by the automatic detection method. Hence, we can use it also to estimate the true distribution of magnitudes based on the model of the observed magnitudes  $m_{obs}(x)$  as  $m_{true}(x) = m_{obs}(x)/h(x)$ .

To find the latter, a log-normal, log-logistic, generalized-gamma, gamma, normal and Weibull distribution were fitted to the observed magnitude distribution using maximum-likelihood estimation (using `scipy.stats`). The quality of the models was assessed using the Akaike Information Criterion (AIC, Akaike [1974]), which gives the relative goodness-of-fit, and the Kolmogorov-Smirnov-Test (KS-Test), which provides an absolute fit. The KS-test tests for the null-hypothesis that two samples come from the same distribution.

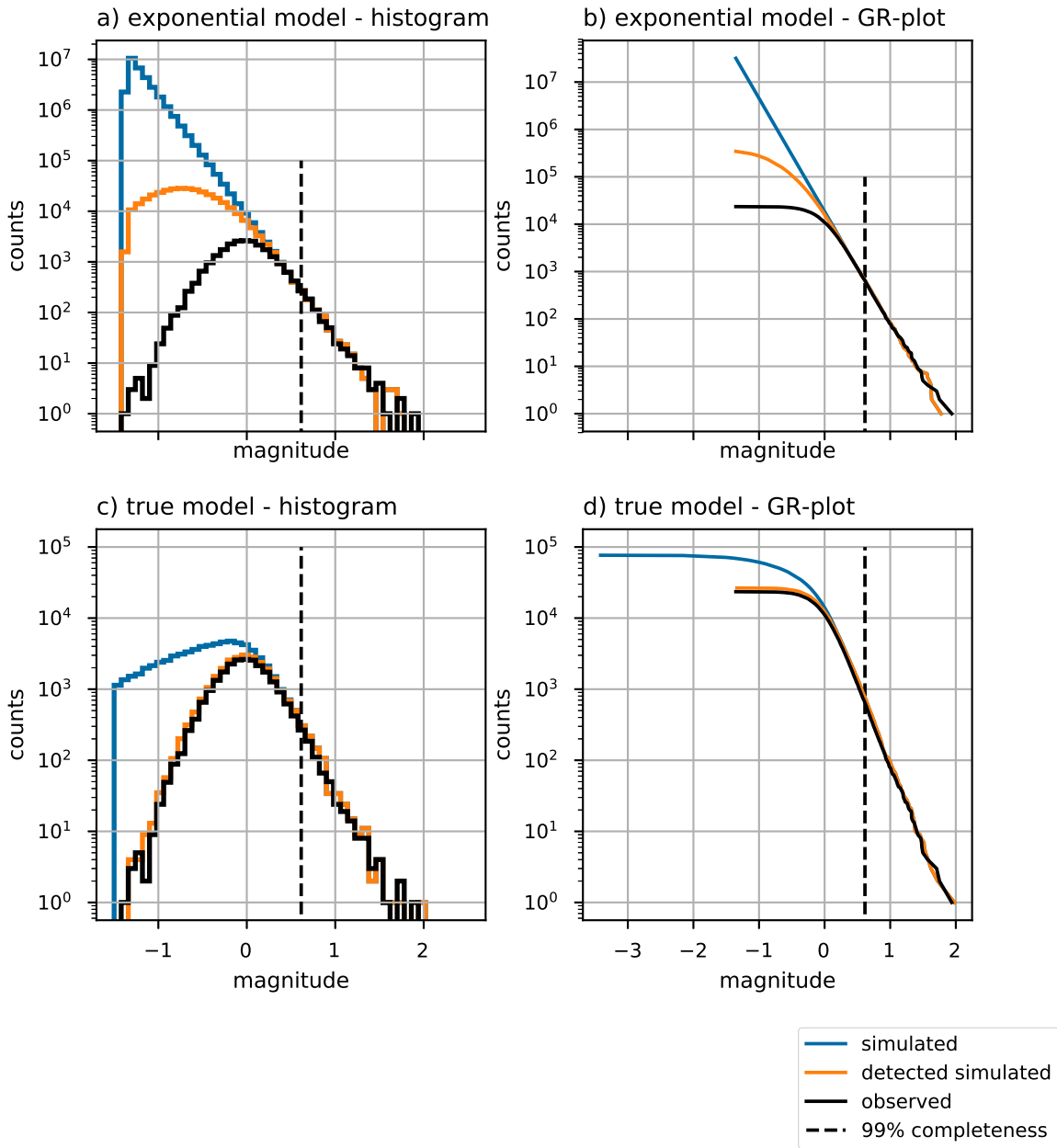
The best model was clearly the log-logistic distribution

$$f(x(x_0), \beta, \alpha) = \frac{\beta}{\alpha} \frac{x^{-\beta-1}}{\alpha} \left(1 + \frac{x^{-\beta}}{\alpha}\right)^{-2} \quad \text{with } x = x' - x_0, \quad x' > 0$$

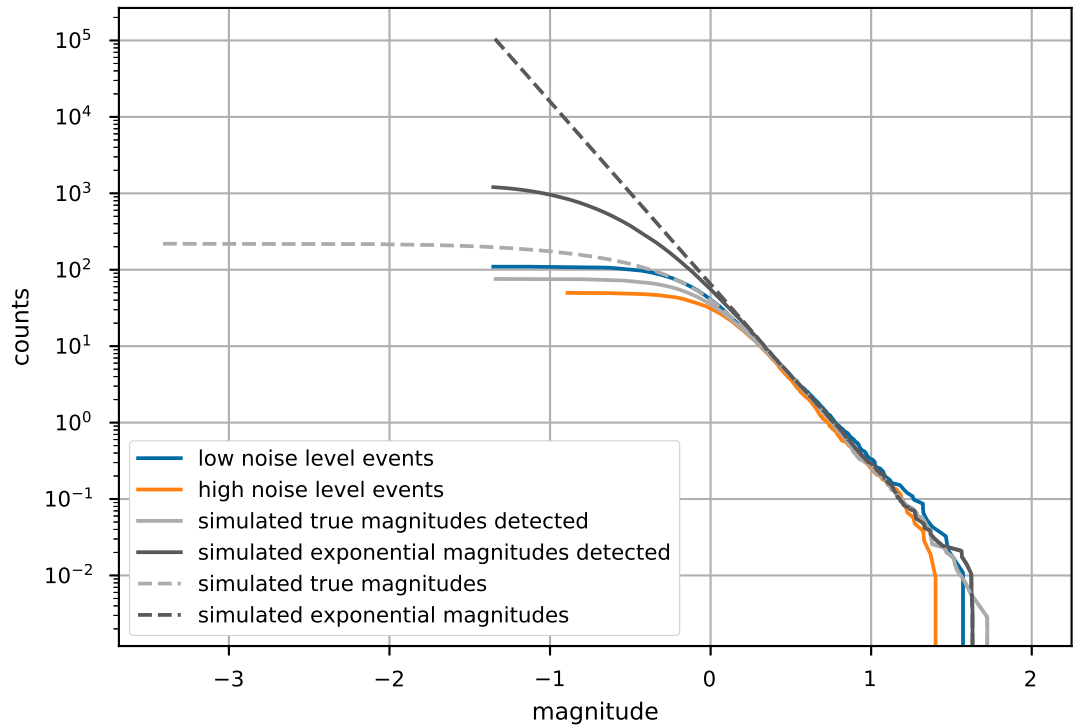
with shape  $\beta = 35.552054$ , scale  $\alpha = 6.120555$  and location  $x_0 = -6.142437$ . It yielded the lowest AIC and was the only model which yielded a p-value above 0.1 (0.42) in the KS-Test, thus did not reject the null hypothesis (Fig. S8, Table S3).

The combination of the log-logistic distribution function and the detection curve gives:

$$q(x) = \frac{\frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1}}{\left(1 + \left(\frac{x}{\alpha}\right)^{\beta}\right)^2} (1 + \exp(-kx - x_0))$$



**Figure 5.6.:** Histogram (left) and empirical survival function (right) for simulated event detection. Black solid lines represent observed data, dashed lines indicate completeness magnitude. *a, b*) Events with exponentially distributed magnitudes derived from the complete catalog. *c, d*) Magnitudes distributed according to “true” model (Eq. 5.1).



**Figure 5.7.:** Comparison of cumulative magnitude distributions of observed events during high (orange) and low (blue) noise level (sections of similar amplitude) and simulated detections from an exponential (solid dark gray) and “true” (solid light gray) distributions. The input distributions to the simulations are shown as dashed lines. The distributions are vertically aligned to a common count at  $m=0.6$

This function still needs to be normalized in order to serve as a probability density function. Since the integral could not be found analytically, we integrated the function numerically within the interval  $[-2, 2]$  and divide by the found value  $n = 2.806$ . So the probability density of the “true” magnitude distribution is given as:

$$f(x) = \frac{1}{n}q(x) = 0.356 \frac{\frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1}}{\left(1 + \left(\frac{x}{\alpha}\right)^{\beta}\right)^2} (1 + \exp(-kx - x_0)) \quad (5.1)$$

Running the detection simulation with this distribution resulted in the observed distribution (Fig. 5.6).

As shown in Fig. 5.7 the magnitude distribution of events observed during low noise levels continues the one of simulated detections from the “true” distribution much better than the one from the exponential model. In addition, the magnitude distribution of

low-noise events is identical to the “true” distribution down to magnitudes well below 0. Since the determination of sections of similar amplitude was independent from the MC-simulation, we see this as a confirmation that the simulation provides adequate information on the detection capabilities of low magnitude events.

Since the definition of “true” distribution in Eq. 5.1 is somewhat artificial, we fitted a variety of standard models to the simulated magnitudes and assessed their relative fit using the AIC. The KS-Test always rejected the null-hypothesis of similarity which is however little surprising given the high number of samples and the fact that they were derived from a different pdf. The Weibull distribution performs best followed by the Gumbel and Gompertz distribution. The Gumbel distribution of magnitudes is interesting inasmuch as it indicates an exponential distribution of amplitudes. The Weibull distribution is part of Eq.5.1 if  $k = \beta$ , and given its success in the fitting may be the dominating term. Interestingly, the log-logistic distribution performed worst. The generalized gamma, normal and inverse Gaussian were mediocre.

### 5.5.3. Interevent times

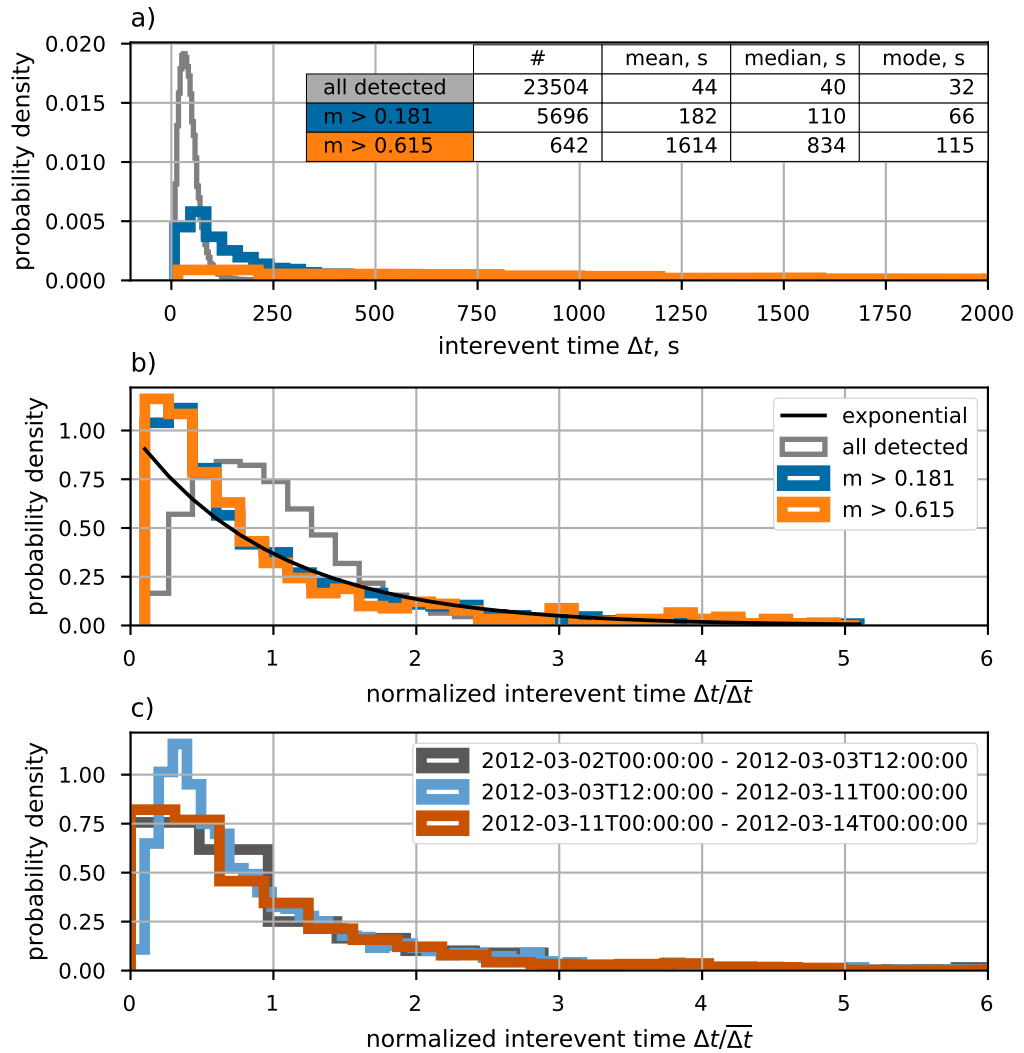
#### Dependence on threshold magnitude

At first, we show the distributions of interevent times between events of magnitudes above the 90%- and 99%-completeness threshold found by the MC simulation. The 90%-completeness provides a more robust estimate due to the almost tenfold number of events. We used the data from the entire period of observation hence assuming stationarity. We regard this as justified due to the rather gentle changes in the time series, notably in the cumulative event count. For better comparison of the distributions, the interevent time is normalized by its mean. For the (almost) complete part of the catalog, the distributions are strongly left-skewed with a mode close to 0. For the higher threshold magnitude the mode shifts closer to 0. The right tail is well represented by an exponential distribution. Hence, the occurrence of events in time is approximately a Poisson process which is corroborated by the mean coefficient of variation of  $\approx 1$ .

To demonstrate how strongly the interevent time distribution depends on the completeness of the catalog, we also plotted the distribution for the entire catalog (i.e. all detected events). It clearly has a mode  $> 0$ , indicating a preferred and thus more regular recurrence time.

#### Dependence on time

The time series of total event count, mean interevent time and occurrence rate suggests changes in the process after 1.5 days (2012-03-02T12:00) and after 9 days (around 2012-03-11T00:00), albeit very subtle ones (dashed lines in Fig. 5.4). The interevent time distributions of events with  $m > 0.181$  within these three time segments are quite different (Fig. 5.8 bottom). Especially the center part exhibits a clearly unimodal distribution with



**Figure 5.8.:** *a)* Distributions of interevent times between events at different threshold magnitudes. The gray curve comprises the entire catalog while the blue and orange curves only comprise events with magnitudes larger than the simulated 90%- and 99%-completeness threshold. Table gives group size, mean, median and mode. *b)* Same as *a)* but interevent times are normalized by their means for better comparison. A standard exponential distribution (solid black line) is given for orientation. *c)* Distributions of normalized interevent times for events with  $m > 0.181$  (90%-completeness) within time segments indicated in Fig. 5.4 (dashed lines). All histograms are normalized to represent probability density.



a steep decrease towards 0. A smaller bin width indicates such a decrease also for the last segment but leads to an otherwise very bumpy histogram. For the interevent times of the early segment, no such decline towards 0 could be revealed by different binning.

### Interevent time models

In order to find a model for the occurrence of events in time, we fitted different probability density functions to the interevent time distributions of events above the 90%-completeness magnitude. Fitting of parameters and model selection was done in the same way as for the magnitudes (see section 5.5.2) except that the location parameter was fixed to 0. We tested the log-normal, log-logistic, generalized-gamma, gamma, normal, Weibull and exponential distribution on the data of the whole period as well as the sub-periods analyzed earlier.

The log-normal distribution yielded the lowest AIC (best relative fit) for all sub-periods but the first one. The latter was better fit by a generalized gamma (GG) distribution although followed by a log-normal distribution. For the other three data sets, the GG distribution scored second on the AIC. Based on the p-values of the KS-test and using a significance level of 1%, for the entire period all models had to be rejected, for the first sub-period all models were acceptable, for the second sub-period only the log-normal model was just not rejected and for the last period the log-normal, log-logistic and generalized gamma model were not rejected.

It is noteworthy though that also the GG-model of the first sub-period declines towards 0 and that therefore all models indicate a preferred recurrence time. This suggests a slight periodicity in the temporal occurrence of the transients.

#### 5.5.4. Interdependence of magnitude and interevent times

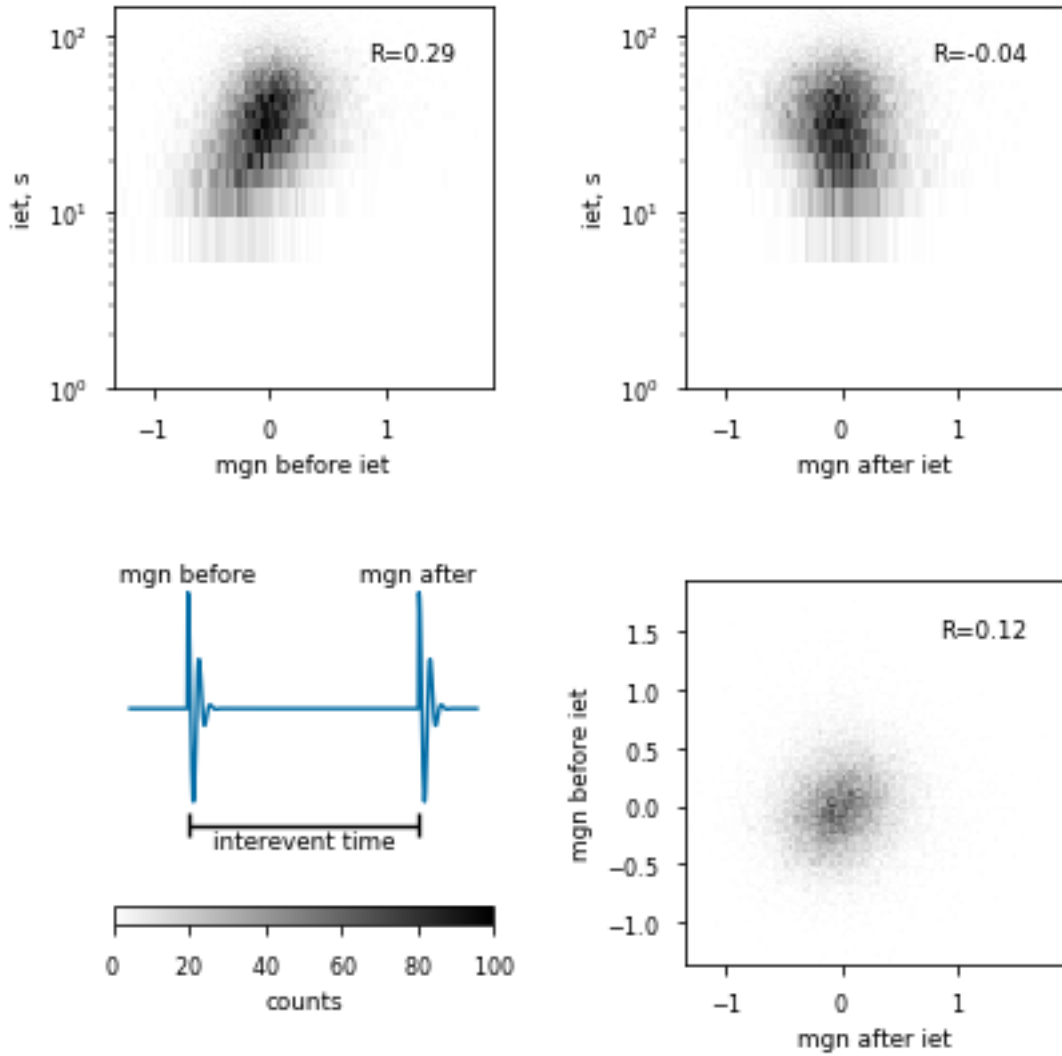
To illuminate the relationship between the magnitudes of consecutive events and the interevent time between them, we computed 2D-histograms of the corresponding pairings of variables (Fig. 5.9). We used the data from all detected events because we think that it gives a more realistic representation of the relationships between adjacent events despite the uneven detection probability. If a high threshold magnitude is selected, interjacent smaller events are almost certainly missed, which leads to clearly false interevent times. Using the entire data increases the chance to analyse intact sequences of events. Nevertheless, the following results are possibly affected by the decreasing detection rate toward lower magnitudes.

The magnitudes of two consecutive events were uncorrelated. Similarly, no correlation was found between the interevent time and the magnitude of the event after the repose interval. In contrast, the magnitude appears to be correlated to the logarithm of the subsequent interevent time, although the correlation coefficient is only  $\approx 0.3$ .

5. Magnitudes & interevent times of transients

2012-03-02T00:00:00 - 2012-03-14T00:00:00						
	AIC	KS_pval	loc	scale	shape_0	shape_1
lognorm	34811.5	7.67e-04	0.0	115.50	0.94	-
fisk	34912.7	2.64e-05	0.0	113.70	1.83	-
gengamma	34886.9	5.11e-08	0.0	0.06	11.15	0.31
gamma	35265.1	3.84e-33	0.0	146.70	1.24	-
weibull	35329.4	9.24e-26	0.0	186.15	1.05	-
expon	35342.5	1.79e-37	0.0	181.95	-	-
2012-03-02T00:00:00 - 2012-03-03T12:00:00						
	AIC	KS_pval	loc	scale	shape_0	shape_1
lognorm	2648.8	2.23e-01	0.0	204.44	1.04	-
fisk	2655.3	2.72e-01	0.0	211.40	1.67	-
gengamma	2646.0	6.90e-01	0.0	0.60	8.11	0.35
gamma	2655.8	1.73e-01	0.0	281.53	1.18	-
weibull	2657.9	2.04e-01	0.0	338.96	1.06	-
expon	2658.9	1.78e-01	0.0	331.07	-	-
2012-03-03T12:00:00 - 2012-03-11T00:00:00						
	AIC	KS_pval	loc	scale	shape_0	shape_1
lognorm	25570.5	1.00e-02	0.0	101.27	0.88	-
fisk	25645.4	9.98e-04	0.0	100.04	1.96	-
gengamma	25614.4	3.61e-05	0.0	0.04	12.81	0.32
gamma	25852.5	1.75e-23	0.0	105.32	1.42	-
weibull	25937.2	3.36e-18	0.0	157.82	1.14	-
expon	26000.8	3.22e-43	0.0	149.75	-	-
2012-03-11T00:00:00 - 2012-03-14T00:00:00						
	AIC	KS_pval	loc	scale	shape_0	shape_1
lognorm	6381.6	2.00e-01	0.0	164.17	1.00	-
fisk	6402.8	6.19e-02	0.0	165.30	1.71	-
gengamma	6385.3	4.31e-01	0.0	0.07	11.06	0.30
gamma	6431.2	5.68e-04	0.0	223.13	1.19	-
weibull	6438.1	1.88e-03	0.0	269.75	1.04	-
expon	6439.7	1.18e-04	0.0	264.65	-	-

**Table 5.1.:** Tested distributions as interevent time models for entire observation period (top) and the three sub-periods. The last for columns give the parameters of the pdf determined by MLE. The first two columns give the criteria for model selection. A lower Akaike Information Criterium (AIC) indicates a better relative fit. Models yielding the lowest AIC are highlighted. If the p-value of the KS-test is above a chosen significance level, e.g. 0.01 the data and the model can be considered as similar.



**Figure 5.9.:** Interdependence between magnitudes of consecutive events and the interevent time in between for the unfiltered catalog ( $M > -\infty$ ). Each panel shows a 2D-histogram computed on a  $100 \times 100$  linear grid. The correlation coefficient  $R$  in the upper 2 panels was calculated using  $\log_{10}(iet)$ .

## 5.6. Discussion

We investigated frequency distributions of magnitudes and interevent times for transient seismic events at Villarrica which are commonly attributed to Strombolian explosions. Based on the statistical properties of the seismic noise, we derived a detection curve for our network which indicates the probability to detect a given magnitude and the magnitude of completeness for our event catalog. Additionally, we could reconstruct a possible distribution of magnitudes which we refer to as the “true” distribution. In the following, we compare the distributions to tectonic earthquakes and to results from other volcanoes. We discuss possible implications for the flow regime, assuming that the events reflect slugs or bubbles of gas in the magma column.

### 5.6.1. Magnitudes

The distributions of the magnitudes and the interevent times are clearly unimodal, which sets these events apart from normal shear fracture earthquakes. Even though the magnitude-frequency relation seems to be exponential for larger magnitudes, we could demonstrate that the occurrence of small magnitudes decreases again, given their observed frequency and their detection probability in our network. In other words, the Gutenberg-Richter law [Gutenberg and Richter, 1956] does not apply here. The GR-law predicts a power-law distribution of seismic amplitudes and therefore an exponential distribution for the whole range of magnitudes with magnitudes being proportional to the logarithm of the amplitude.

Deviations from the GR-law were already observed for Strombolian activity in other basaltic systems [Cauchie et al., 2015, Nishimura et al., 2016, Pering et al., 2015]. Nishimura et al. [2016] proposed an exponential decay of amplitudes (not magnitudes!) based on seismic data. If the amplitudes were exponentially distributed, the magnitudes would follow a Gumbel distribution which was among the better performing models for the simulated “true” distribution. Therefore, we do not want to exclude the possibility of exponentially distributed amplitudes.

However many other studies suggest a rather unimodal (that is having a mode and a decline towards low sizes) distribution of volcanic explosion sizes measured as seismic amplitude, energy, gas mass or jet height [Martino et al., 2012, Zobin, 2017a, Cauchie et al., 2015, Pering et al., 2015, Taddeucci et al., 2013, Tamburello et al., 2012]. In those cases in which the distributions were fitted by probability density functions, log-normal [Martino et al., 2012] or Weibull [Taddeucci et al., 2013] distributions were found.

### 5.6.2. Interevent times

Unlike the magnitudes, the interevent times between undetected events can not easily be reconstructed from the detection curve. Therefore, we discuss the interevent times be-

tween the events in the 90%-complete catalog. Their occurrence in time is best described by a log-normal distribution. This result discards in particular the simple Poisson process model which is defined by exponentially distributed interevent times and represents a completely random occurrence. The unimodal shape of the log-normal distribution suggests a preferred recurrence time of the events. Although, if only very large events were taken into account, the distribution of interevent times might become exponential.

A generally non-Poissonian occurrence of explosions was already noted for Strombolian and Vulcanian activity at other volcanoes. Varley et al. [2006] found distributions of monotonically decaying as well as unimodal shape, although eruption sequences classified as Strombolian only showed the latter. Dominguez et al. [2016] on the other hand found only unimodal shapes for both types of activity, which were all reasonably well fitted by log-logistic distributions. Other studies proposed Weibull [Cauchie et al., 2015] or Gamma [Bell et al., 2017] distributions. Preferred recurrence times of Strombolian events were not only found in seismic data but also using thermal or SO<sub>2</sub> imaging techniques or infrasound [Gaudin et al., 2017b, Harris and Ripepe, 2007, Pering et al., 2015, Ripepe et al., 2002, Taddeucci et al., 2013].

While most studies of explosion sequences used relatively short periods of observations (especially those relying on video footage comprise only a few hours at most), Martino et al. [2012] presented interevent time distributions of three 0.5-1 year long sequences from Stromboli. They found periodic recurrence of explosions only during the comparatively short periods of increased activity while quiet periods showed nearly Poissonian behavior. In this regard, we note, that our data stems from a relatively short period of observation during which the volcano was considered very active by the local monitoring agency OVDAS. Obviously only an investigation of a longer time period could reveal a possibly different behavior during quieter periods.

The mean interevent times in the 90%-complete catalog were between 100 s and 400 s. Even for the unfiltered catalog, the mean time between events was around 40 s. In contrast, observations of bubble bursting activity by camera on basaltic systems (Stromboli, Etna) revealed median interevent times of less than 10 s [Dominguez et al., 2016, Pering et al., 2015]. For Villarrica, Gurioli et al. [2008] reported an average spacing of 6-7 s between all types of degassing events. The average time between more violent events (which involves mass ejection) is around 18 s. At Masaya Volcano, Pering et al. [2019b] counted 5-25 bubbles within 5 s at the surface of the lava lake. Using only seismic observations, it is difficult, or even impossible to achieve such a detailed resolution because the noise level inhibits the detection of small events, an extended coda masks the following event and simultaneous occurring events simply appear as a single waveform. Nevertheless, we want to point out, that the number of simulated “true” events (based on the magnitude distribution) yields a mean spacing of 13.5 s which is in the range of the values reported by Gurioli et al. [2008].

### 5.6.3. Interdependence

Our results on the relationship between magnitudes and interevent times resemble the findings by Pering et al. [2015] for degassing events at Mt. Etna. They found a linear correlation between the logarithms of erupted slug mass and time to the next event but no correlation between mass and time to the preceding event. Video footage of slug trains in subsequent lab experiments by Pering et al. [2017] suggest that this “repose gap” forms because of coalescence of closely spaced slugs. During coalescence, the trailing slug is accelerated toward the leading one, resulting in a larger slug with a larger gap towards the following slug.

### 5.6.4. Flow regime

In terms of flow regimes, the magnitude of events corresponds to the size of the slugs, i.e. notably their length. With  $V = CL$  being the volume of a slug with cross-sectional area  $C$  and length  $L$ , from the seismic amplitude  $A \propto V$  one obtains  $M \propto \log A \propto \log V = \log(CL)$  for the magnitude. The interevent time is proportional to the length of the liquid bridge between the gaseous slugs. The distributions of magnitudes and interevent times both display a relatively concise peak. The distribution of interevent times is approximately log-normal similar to the distribution of the liquid bridge lengths in slug flow. The estimated “true” distribution of magnitudes may be interpreted as normal in which case slug lengths would be log-normally distributed. Arguing that the population of smaller bubble sizes in slug flow can not be resolved seismically, the observed and estimated magnitudes could correspond to the size distributions of the gaseous slugs in slug flow.

In the light of the experiments by Pering et al. [2017] on the ascent of slug trains, the (albeit weak) correlation between preceding magnitude and interevent time may be seen as further evidence of a sustained slug flow regime. The sustained flow is in contrast with the models of formation and ascent of single slugs, which do not interact with each other. We speculate that in this case the occurrence of slugs would be random rather than slightly periodic.

A differentiation from bubble or churn flow based on qualitative comparison of size-frequency distributions alone is difficult. Transition criteria are commonly defined by the volumetric gas flow rate or superficial gas velocity which however are unknown in our case. We argue though that the size of bubbles in bubbly flow is probably too small and their number too high to produce distinct seismic signals. Instead, it might produce a tremor-like signal [Mudde, 2005, Ripepe et al., 2010]. Churn flow on the other hand indicates a more unstable and turbulent system which perhaps better reflects the vigorous, boiling appearance of the lava lake surface. If it is still reasonably close to slug flow, the interevent time and magnitude distributions are probably indistinguishable from slug flow.

A strict comparison with the distributions from the flow experiments requires that each slug produces exactly one event only. Field observations and lab experiments however suggest that seismic events may be generated by bursting of a slug at the surface or deeper inside the conduit [Spina et al., 2019]. The former additionally generates an acoustic signal while the latter does not, leading to a higher number of seismic than acoustic events. Hence, one slug possibly causes multiple seismic signals.

In this context, we envision one particular scenario for Villarrica in which one slug could produce multiple seismic events. The difference between the diameter of  $\approx 12$  m, calculated for the lower conduit from degassing rates [Palma et al., 2011] and the width of the upper lake reservoir of 20-30 m [Moussallam et al., 2016] requires a substantial widening of the conduit somewhere. Richardson and Waite [2013] interpreted the source force of a frequent, repetitive transient waveform as a drag force at the nearly horizontal lake bottom and depicted a step-like termination of the conduit to the lake reservoir. James et al. [2006] showed in lab experiments that the passage of a slug through a tube widening leads to pressure and displacement signals which bear strong resemblance with the transient waveforms found at Villarrica. Larger slugs may break up into two or more daughter bubbles, each of which could produce an additional signal when finally bursting at the free magma surface. The reported mismatches between visible bubble bursting and seismic/acoustic signals [Gurioli et al., 2008, Goto and Johnson, 2011, Ripepe et al., 2010] at Villarrica at least suggests that the relation between single slugs and seismic events is not straightforward. However a better knowledge about the source mechanisms and origin depths of the events would be required to further explore this aspect and refine the frequency distributions.

We see potential that the statistical analysis of the occurrence of seismic events can help in monitoring the flow regime in volcanic conduits over longer time periods. Valuable complementary information could be gained by adding infrasound and degassing rate monitoring. However, clearly more experimental data on the relation between gas flow regime and seismic and acoustic signals is needed to provide a solid reference for the identification of flow patterns.

## 5.7. Conclusion

We investigated the statistical distributions of magnitudes and interevent times of transient seismic events from the lava lake of Villarrica Volcano which are commonly associated with Strombolian activity. By taking into account the detection limits of our seismic network we could demonstrate that the amplitudes do not follow the power-law of tectonic earthquakes. Instead the distribution of magnitudes declined again towards small values. The distribution of interevent times was strongly influenced by the selection of the threshold magnitudes with means of 44 s and 182 s for all detected events and those above the 90%-completeness magnitude, respectively. The corresponding modes were

## *5. Magnitudes & interevent times of transients*

---

32 s and 115 s. In any case the distributions were approximately log-normal, suggesting a preferred recurrence time, i.e. a weak periodicity. The time between two events appeared to correlate with the magnitude of the first event. Based on comparison with analogous gas flow experiments, we tentatively concluded that the flow regime in the conduit of Villarrica is slug flow.

Further corroboration for this hypothesis could be gained from 1) the superficial gas velocities or volumetric gas flow rate in the conduit and 2) better knowledge of the origin depth and nature of the seismic events, which are currently however unavailable.



## 6. Waveform families of transient events

This chapter is a first draft for a scientific article.

### Main results / conclusions:

- For the transient events, at least 67 families of similar waveforms could be identified, the two largest having several thousand members.
- Similar relative travel times suggest a similar source region of all families
- Preliminary attempts to locate the source by fitting the decay of amplitude suggest a source at or near the crater.

### Abstract

*Volcanic seismic events, that are related to fluid-dynamic processes, are often the result of a non-destructive source. As a consequence, waveforms can be very similar. From a catalog of >20000 transient events, 67 templates were generated through a cluster analysis of the largest 132 events ( $M > 0.9$ ) in a frequency range of 0.45 to 2 Hz. For almost all templates, more than one instance of the waveform was found by matched filter detection. The two by far largest families comprised  $\approx 2500$  and  $> 7000$  events, respectively.*

*Composite waveforms for most stations were generated by stacking the 50 events with the highest similarity to the template. Near the source, the waveforms of the two largest families resembled a Ricker wavelet, while most other families showed more complex waveforms. All families developed much more complex, elongated waveforms within  $\approx 1$  km of the first station.*

*Because of the stacking, the onsets of the waveforms became much clearer even at distant stations. By cross-correlation, relative travel times between the stations were found to be very similar for all families, suggesting a similar source region. This analysis also indicated regions of higher and lower wave propagation velocities.*

*The composite waveforms were located by fitting the spatial decay of amplitudes. However, relatively large uncertainties and discrepancies between the observed relative travel times and those for the obtained locations impeded a reliable determination of the source locations.*

## 6.1. Introduction

It is quite common at volcanoes that seismic waveforms occur repeatedly over time [Chouet, 2003, Green and Neuberg, 2006, Waite et al., 2008, Matoza and Chouet, 2010, Cauchie et al., 2015, Bell et al., 2017]. This is usually interpreted as the result of a non-destructive source [Chouet, 1996, Chouet and Matoza, 2013]. Therefore it is particularly common in the group of LP events which are thought to result from interaction of volcanic fluids and the plumbing system. Determination of the location and source mechanism can thus help to identify key elements of the volcanic system.

Repeating instances of an event can be stacked together to emphasize the actual waveform [Waite et al., 2008, Richardson and Waite, 2013]. This is particularly useful for weak events in a noisy environment as the noise is eliminated during the stacking through destructive interference.

Richardson and Waite [2013] already showed the presence of a very persistent event family at Villarrica which produced stable waveforms over the course of three years. After stacking these waveforms into a composite event, they were able to apply a moment tensor inversion to the event. The abundance and variety of events found in Chapter 5 however suggests that there could be more families than the one found by Richardson and Waite [2013]. Identifying additional families could facilitate the analysis of the signals and help to shed light on the cause and nature of the transients and hence the processes inside the volcano.

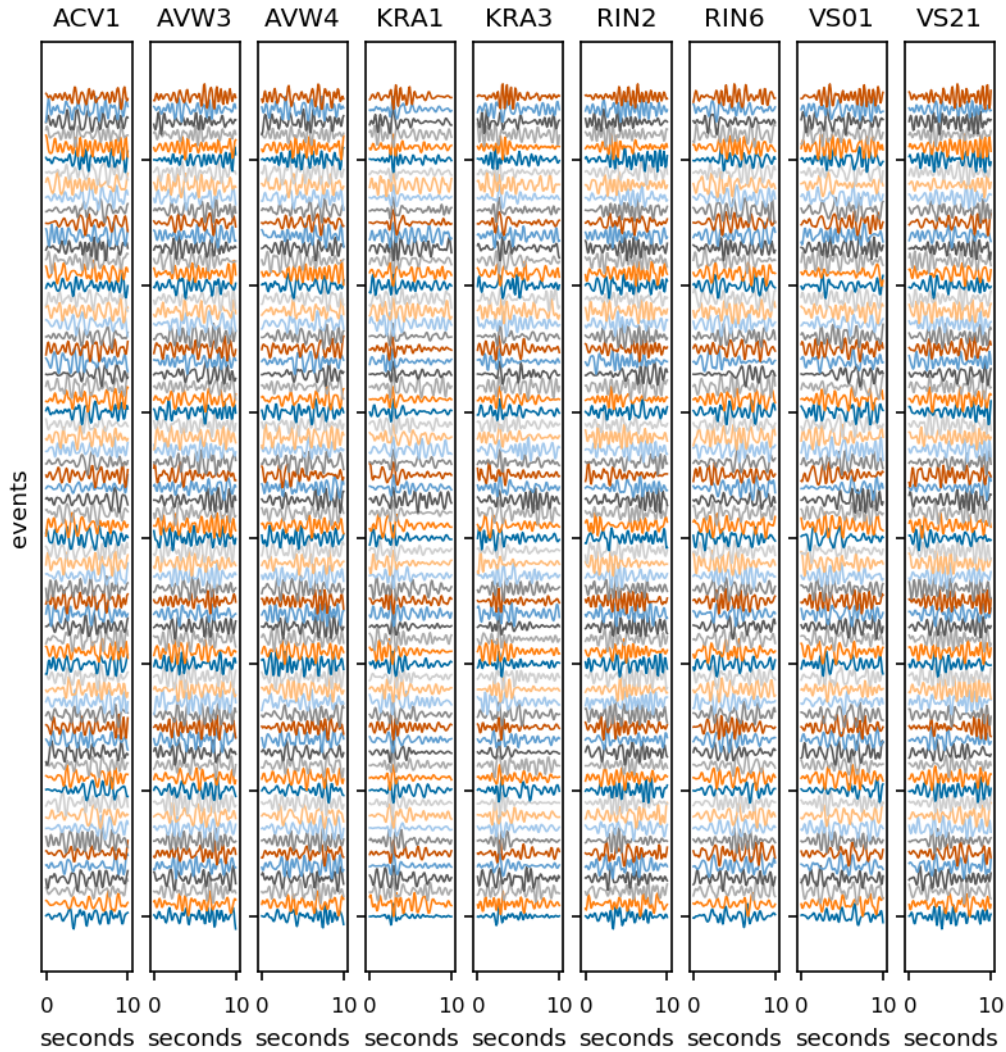
This chapter explains how a number of waveform families could be identified from cluster analysis on the event catalog obtained in Chapter 5 and subsequent template matching. The detection process is explained in the following section. Subsequently, the amplitude decay location method from Chapter 4 is applied to the composite events.

## 6.2. Identification of waveform families

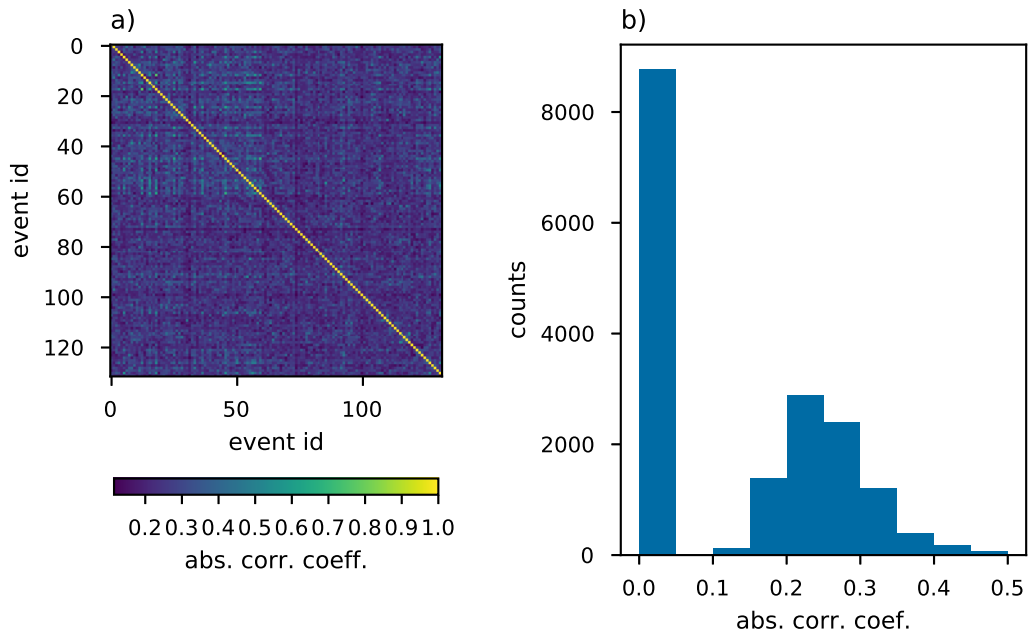
### 6.2.1. Cluster analysis

A common way to find waveform families consists of computing cross-correlations between all events in a catalog and running a clustering algorithm on the matrix of correlation coefficients. Here, the event catalog obtained in Chapter 5 served as basis. However, due to the large number of events, correlating all events with each other was not feasible. Therefore, only events with a magnitude  $>0.9$  were selected. To ensure that the waveform characteristics of a template were stable throughout the network 9 stations were included: ACV1, AVW3, AVW4, KRA1, KRA3, RIN2, RIN6, VS01, VS21. Of the 134 events with magnitude  $>0.9$ , 132 were recorded by all selected stations, thus were used in the template generation. The similarity between two events was determined as follows:

1. Cross-correlation functions (CCF) were computed between corresponding stations.



**Figure 6.1.:** Waveforms selected for template creation. Each panel shows the waveforms at the indicated station. Waveforms within a row (having the same color) correspond to the same event. Time axis are relative to the starttime  $t_{i,j}$  at each station (see text for details) which is aligned to the maximum at KRA1. For clarity only every second station and event is shown.



**Figure 6.2.:** *a)* Matrix of correlation coefficients between events. Values are the absolute maximum of the cross-correlation function between two events, averaged over corresponding stations. *b)*: Histogram of pairwise correlation coefficients (lower triangle in matrix in *a)*).

2. A mean CCF was computed by summing the CCFs from all stations and dividing by the number of stations
3. The position and value of the absolute maximum of the mean CCF gave the time lag and measure of similarity

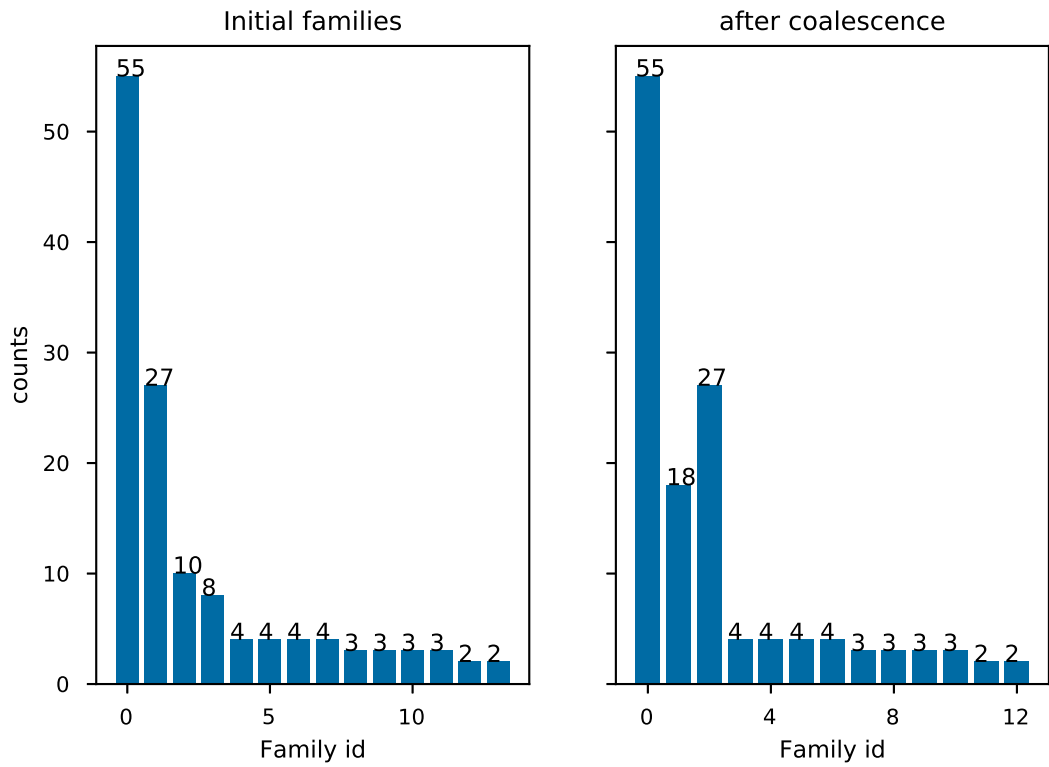
The waveforms were first selected based on the trigger times in the catalog. Note that the trigger times were determined using only the stations KRA1 and KRA3. The waveforms were then aligned to the maximum at the earliest station KRA1 (Fig. 6.1) by adjusting the start time to 3s before the maximum at station KRA1 within the time window in the catalog ( $T_{i,0} = t_{i|max} - 3$ ). Travel time differences between the stations were compensated using  $\Delta t_j = r_j/v$  with  $r_j$  being the distance between the (supposed) source location at the summit crater (71.94°W, 39.419°S, 2692,m a.s.l) and the station. The velocity  $v$  was set to 1000,m/s. Hence waveforms were cut at each station to  $t_{i,j} = T_{i,0} + \Delta t_j$ . The samples were 10s long at each station. In order to reduce the computational load, the data were downsampled from 50 Hz to 10 Hz and filtered between 0.45 and 2.0 Hz.

The maximum absolute mean correlation coefficients (CC) between events ranged between 0.0 and 0.6 (Fig. 6.2). Clusters of similar events were determined in two stages [Green and Neuberg, 2006, Bell et al., 2017]. At first, individual events were grouped into families by [Bell et al., 2017]:

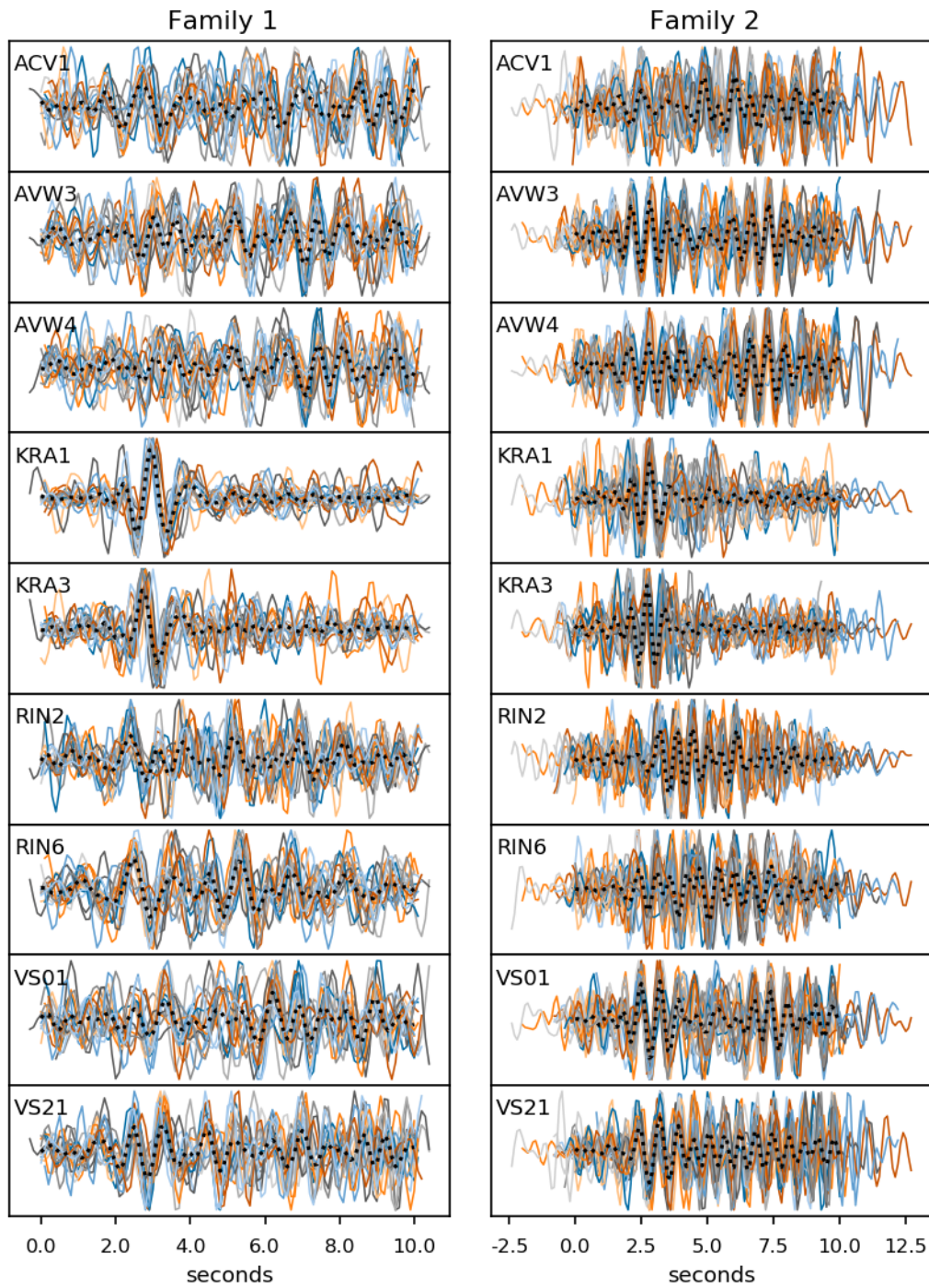
1. select event with the highest  $CC > thr_1$
2. Assign all events that have  $CC > thr_1$  with this event to the family and remove from matrix
3. Define new families by repeating these steps until all remaining  $CC < thr$

The second stage involves a family coalescence during which the families were tested for similarity. To this end, average waveforms were computed from the members of each family. These were clustered again using the steps described above, but with a higher threshold  $thr_2$ . If the threshold was exceeded, the individual events were reassigned into one family, hence reducing the total number of families (Fig. 6.3). The thresholds were set to  $thr_1 = 0.4$  and  $thr_2 = 0.5$ . The chosen values were a trade-off between a comparably high CCs, a number of families around 10 and most families consisting of more than two events.

The chosen CC values are very low compared to other studies. This is likely because here multiple stations were used rather than only one. Tentative use of only one station in the cross-correlation yielded much higher CCs. Using only a single station also allows to better tailor the waveform to the relevant part, and hence increase the quality of the correlation. Despite the low CCs, the quality of the waveform families was found acceptable (Fig. 6.4). The outcome of the clustering process was quite sensitive to the selection of the thresholds.



**Figure 6.3.:** Family sizes after the two stages of the clustering process. Family 0 contains all events that were not well correlated with any other event.



**Figure 6.4.:** Members of the two largest families 1 and 2, shifted according to correlation with a reference event in the family. Lag times between stations are not included here. The average waveform is shown as a dotted black line.

### 6.2.2. Template matching

The stacked waveforms of each family as well as the remaining individual events were used as templates in a matched filter detection algorithm (Python API EQCorrscan [Chamberlain et al., 2017]). Here, the template is cross-correlated with the continuous data. If the CC exceeds a given threshold, an instance of the template is registered. Similarly to the previous template creation, an average cross-correlation function over all stations is used. If multiple templates exceed the trigger threshold within a given time window, the one with the highest CC is selected. The threshold was set to 0.35 and the time window in which no overlapping templates were allowed was set to 10s. In total, 23306 events were detected in this way. Most of the repeaters belonged to one of eleven families found during the clustering (Fig. 6.5) with Families 0 and 1 getting by far the most matches. Despite the low threshold for the correlation coefficient, the waveforms within a family clearly share common features at each station (Fig. 6.6). The similarity between family members also applies to stations that were not included in the matched filter detection (Fig. 6.7).

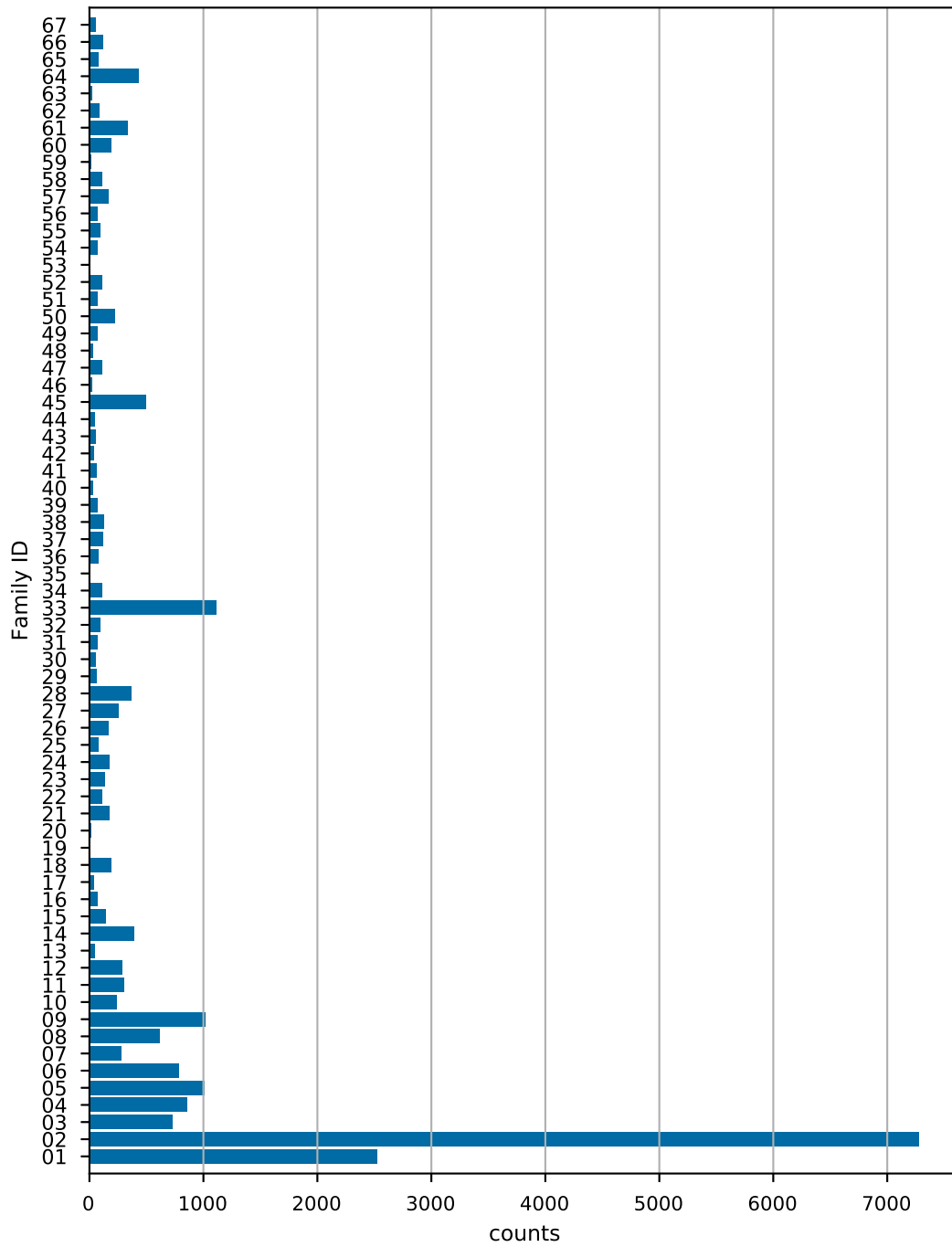
In summary, the following processing steps were performed:

1. Select events from STALTA catalog with  $M > 0.9$  for which all stations have data
2. Extract their waveforms such that we have 15 s long samples centered around the maximum within the nominal duration  $\pm 2$  s of the event
3. Compute pairwise CC between corresponding stations. Use average cross-correlation function over all stations to determine absolute maximum CC
4. Cluster this matrix and compute family stacks using  $thr_1 = 0.35$
5. Coalesce the families, i.e. cluster them again with  $thr_2 = 0.5$ ; if some stacks correlate better now their events are merged into one family.
6. The family stacks and the solitary events (family id=0) are used as templates.
7. Run matched filter detection on the whole data stream.
8. Decluster the detections such that a minimum distance of 10 s is maintained.

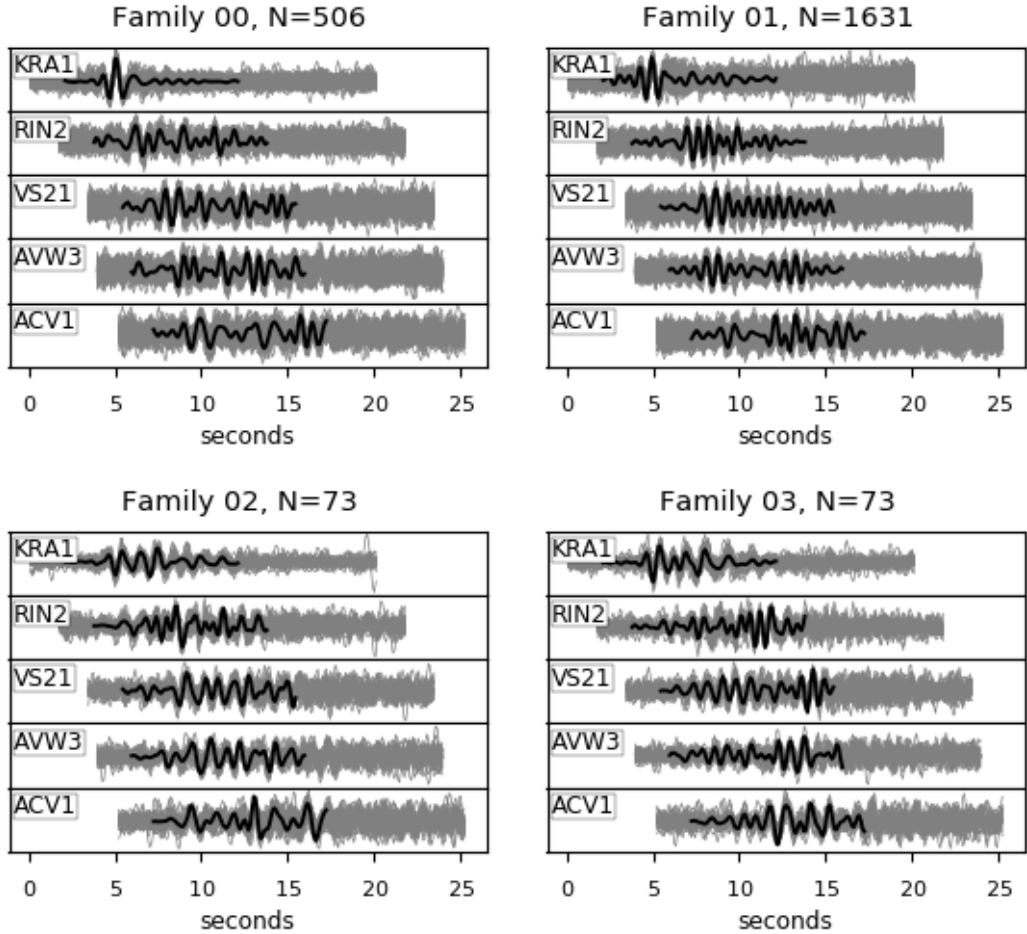
### 6.3. Centroid waveforms of families

For each event and station, centroid waveforms were computed from the members detected by the template matching. To this purpose, the members of each family were sorted by their correlation coefficient with the template. A composite stack was computed from the best 50 members. In families with less than 50 members, all members

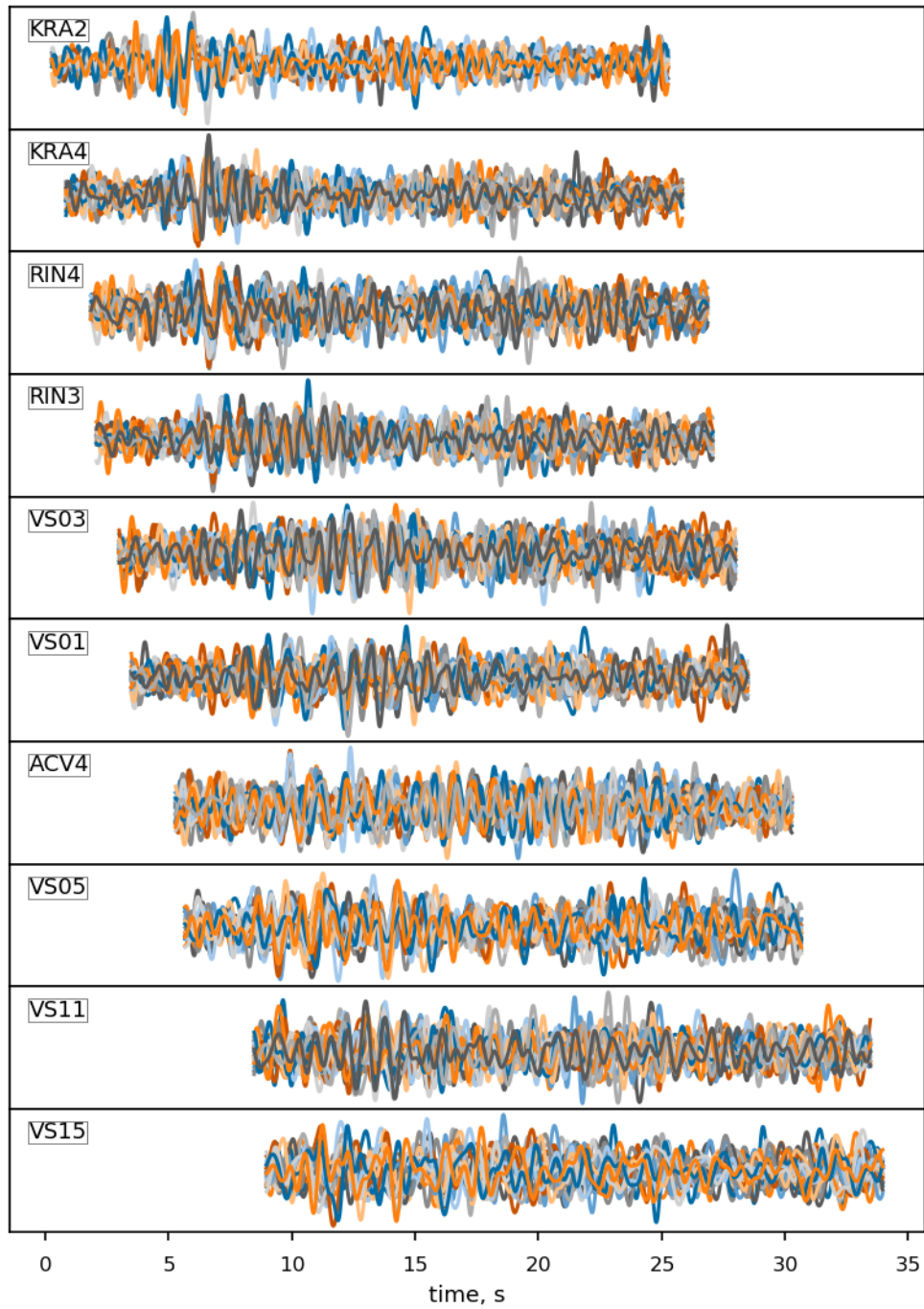




**Figure 6.5.:** Number of repeating events by family. The first 11 families correspond to the stacked waveforms of the families found by clustering. The other templates were individual events. For some of them no repeaters were found.



**Figure 6.6.:** Members of first 4 largest families detected by matched filtering (grey) and corresponding template (black) at selected stations. Waveforms were cut to 2s before the onset + lag time  $\Delta t_j$  and 10s after the end of the template. Only waveforms of events with an average  $CC > 0.45$  (number given above each panel) are shown. See main text for computation of lag times. Waveforms are normalized to RMS-amplitude



**Figure 6.7.:** Waveforms of Family 0 with average  $CC > 0.5$  including stations that were not used in the template matching. In addition, here the original data at sampling rate of 50 Hz is shown.

contributed to the stacked waveform. Here, the “normal” data, which was sampled at 50 Hz was used.

The majority of the stations did not operate during the entire time period to which the matched filter detection was applied. This means, not all of the family members were recorded by all stations. Hence, the number of events contributing to the stacked waveform of a station varies within the family with 50 being the maximum number.

## 6.4. Analysis of individual families

The families were investigated in terms of their spectral content and their source location.

### 6.4.1. Spectral properties and waveforms

Power spectral densities were computed using Welch’s method (Implementation of Python API Scipy [Virtanen et al., 2020] after Welch [1967], Bartlett [1950]). Spectrograms were computed using Fourier transforms in overlapping time windows. The waveforms were cut to 25 s-long windows. The start time at each station was computed using a homogeneous velocity of 1300 m/s and line-of-sight distances between a source in the crater and the stations. The selected stations were located at increasing distance to the crater in a northward direction (blue stations in Fig. 3.1).

### 6.4.2. Lag between stations

Differences in arrival times between stations could indicate a different source location. Therefore, for each family  $j$ , the time difference  $\Delta t_{ij}$  between the  $i$ -th station and KRA1 was determined using cross-correlations of 5 s-long windows around the onsets. The position of the window for each station and family was determined as follows: 1) The crude onsets of Family 01 were picked manually to obtain a set of absolute pick times  $t_{i,pick}$ . 2) For each family, the data of KRA1 were correlated with those from Family 01 to determine the start of each family relative to Family 01 ( $\Delta t_{KRA1,j}$ ). 3) For each family, the manual picks were used, positioned according to this time shift ( $t'_{ij} = t_{i,pick} + \Delta t_{KRA1,j}$ ). The eventual start time resulted in  $t_{ij} = t_{i,pick} + \Delta t_{KRA1,j} + \Delta t_{ij}$ .

The cross-correlation function was computed for a shift of  $\pm 0.8$  s. This relatively small window precludes substantial shifts in arrival times and therefore source positions. In other words, the source location was assumed to be relatively similar for all families. Visual control of the aligned waveforms of randomly selected families suggested, that this assumption was justified. When larger shifts were allowed, the lag times were more scattered, resulting in poor alignment of the waveforms. In particular, neighboring stations showed implausible differences in their arrival times. The positive maximum of the cross-correlation function was used to determine the lag, meaning that the polarity was assumed to be the same at all stations. Similar polarities would be found e.g. for

volumetric source such as explosions. This assumption would be false if the source mechanism was predominantly a double-couple, which represent the shear failures of normal tectonic events. The data were filtered between 0.5 and 1.0 Hz for this step.

At each station a histogram of the obtained lag times was computed, weighted by the coefficient of correlation, the modes of which yielded an average lag time curve.

### 6.4.3. Source location

A major goal of the analysis of the families was to locate the origin of the transients and most importantly determine their source depth if possible. To this end, the grid search for the best fit of the amplitude decay with distance (amplitude source location method, ASL) from Chapter 4 was used. In detail, the approach had to be slightly adapted, since here finite waveforms were targeted rather than the quasi-stationary, continual tremor. The energy of the waveforms was computed between the onset (determined by the cross-correlation approach described above) and the time at which the envelope of the amplitude fell below  $1/e$  of its maximum. The energy was calculated by integrating over the squared amplitude and corrected by the site factors determined in Chapter 4.

Surface waves were assumed for the decay of energy with distance. Therefore, only a 2D search grid for the lateral position was used. The grid was a  $2\text{ km} \times 2\text{ km}$ -square at 100 m intervals, draped over the topography and centered to the crater. The grid point, that yielded the lowest RMS-error was selected as source location of the family. The occurrence of locations over all families was assessed by computing a 2D-histogram of the results using the grid points as bins.

In addition, an energy decay for body waves was tested using the energy in the first 2.6 s after the onset. In this case, the search grid was extended to 3D to include the vertical source position. The depth was also sampled at 100 m intervals between 2800 and 1800 m above sea level (a.s.l.).

In order to examine the uncertainty of the locations, statistical Bootstrapping [Efron, 1979] was performed for Families 01, 02 and 06: The ASL using the surface wave assumption was performed repeatedly with random sets of stations. Per set, the stations were selected randomly with replacement, hence some stations could occur several times while others were left out. The size of the set corresponded to the original set size (44 stations); the sampling was repeated 500 times per family.

Moreover, for the same families, ASL was applied to selected individual members, rather than the stacked waveforms. The members were selected such that their magnitude was  $> 0.18$ , the average correlation coefficient in the matched filtering  $> 0.44$  and their mean signal-to-noise ratio  $\text{SNR} > 2$ . The SNR was computed from the ratios between the energies during the waveform and 10 s before and after it. The ratios were averaged over all stations. 25 such events were used for Families 01 and 02 and 23 for Family 06. Again, for the ASL, surface waves were assumed and the energy computed over the window between the onset and  $1/eE_{max}$  at the respective stations. The start and end

times of the windows were taken from the stacked waveforms. However, the data were filtered between 0.45 and 5.0 Hz.

## 6.5. Results

### 6.5.1. Waveforms and spectral properties

Waveforms and spectra were distinct for each family. For most families however, the spectral content within a family was relatively similar at all stations although the waveforms were very different (Fig. 6.8).

The largest families 1 and 2 featured very short, impulsive waveforms at the crater while most of the other families displayed more complex, elongated waveforms. The contrast in waveform shape between Families 1 and 2 and the rest was less pronounced at more distant stations. Families 1 and 2 differed in their spectral content with the dominant frequency of Family 1 being around 1. Hz and that of Family 2 around 1.5 Hz. Higher frequencies were not taken into account here, since the waveforms were identified only at frequencies of 0.45-2.0 Hz.

Other families showed up to three spectral peaks. Their waveforms often showed characteristics of hybrid events even at the crater with high frequencies in the beginning and low frequencies in the latter part (e.g. Family 04, Family 09 in Fig. 6.8).

The spectrograms showed that Family 01 maintained an impulsive beginning towards later stations with the energy concentrating mostly in the first 5 s of the event (Fig. 6.9). All frequencies set in at about the same time. Only at station VS21, low frequencies appeared slightly earlier.

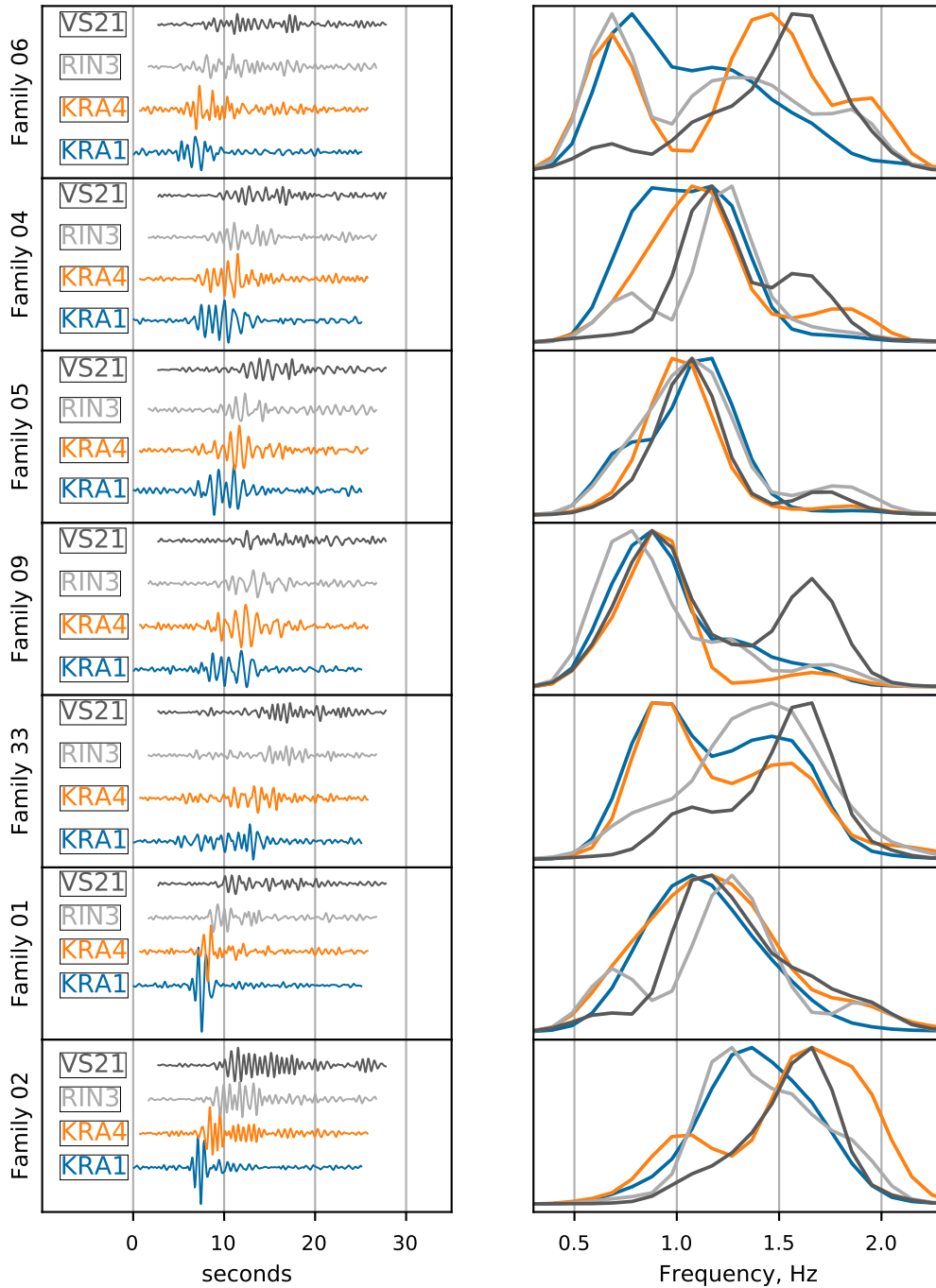
In contrast, the waveform of Family 02 developed a coda already at station KRA4 (Fig. 6.9). In addition, Family 02 clearly displayed a low-frequency onset, followed by higher frequencies. This was most pronounced at KRA4.

The spectrograms of Family 04 emphasized its hybrid character: at station KRA1, KRA4 and RIN3, the lower frequencies started 2-3 s after the onset (Fig. 6.10). Moreover, the low- and high-frequency portions appeared as separate bands. Only at station VS21 the pattern changed.

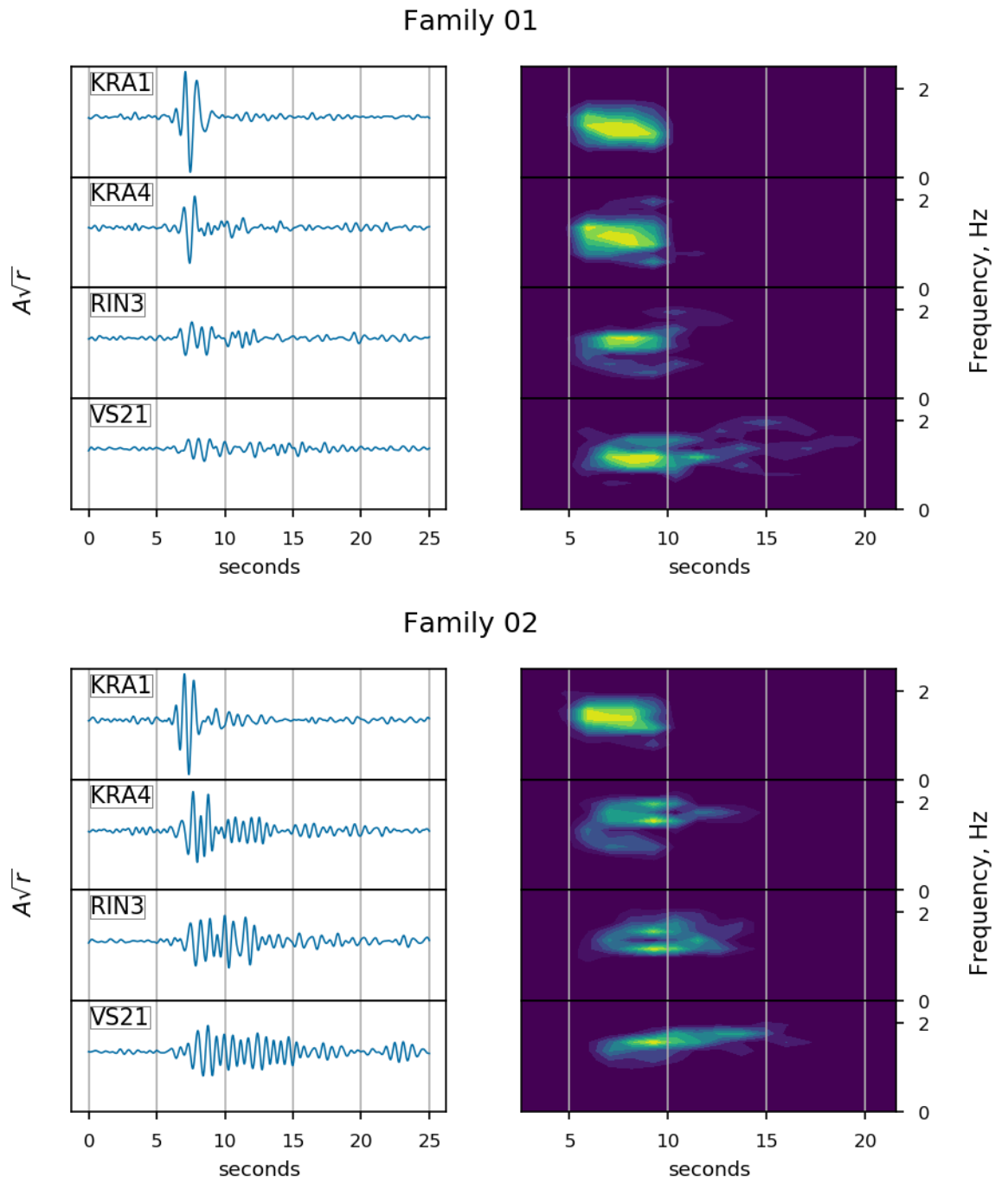
Family 06 displayed an even more complex pattern (Fig. 6.10). At stations KRA4 and RIN3, 2-3 separate frequency bands could be identified while at stations KRA1 and VS21, they seemed to be more intertwined. The low frequencies appeared to start slightly after the high ones, but the difference is less pronounced than for Family 04.

### 6.5.2. Relative travel times

The differences in lag times of the onsets to station KRA1 were investigated to detect possible variations in the source location. The obtained lag times clustered narrowly within 1 s around the average curve, which was defined by the modes of the weighted

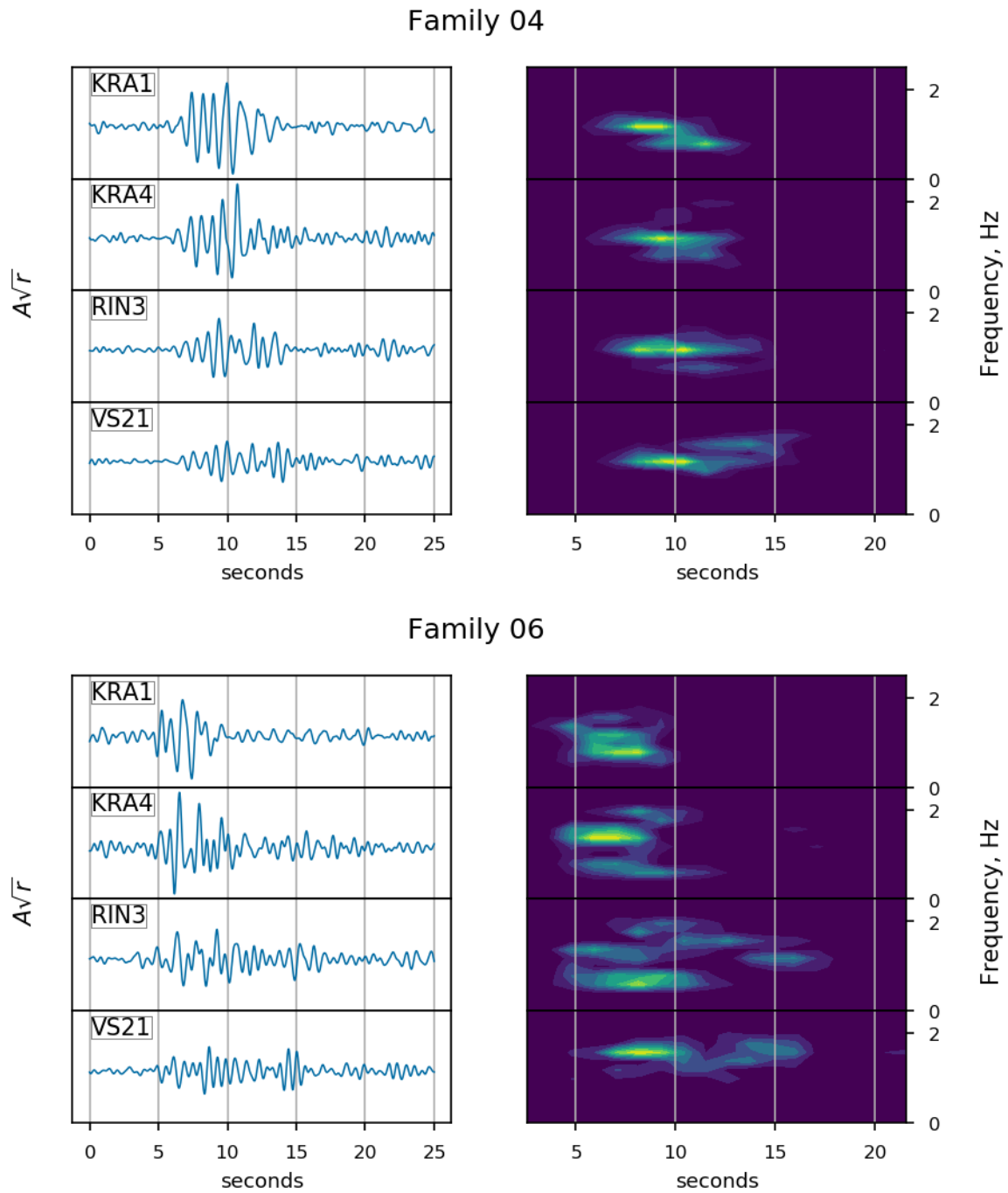


**Figure 6.8.:** Waveforms (left) and power spectral density (right) of the 7 largest families at 4 selected stations. Stations are sorted by increasing distance to the crater. All curves are normalized to their maximum. The spectra vary significantly between the families while they are relatively similar between stations for each family.



**Figure 6.9.:** Waveforms and spectrograms at 4 selected stations of the two largest Families 01 and 02. Stations are sorted by distance to the source near KRA1.





**Figure 6.10.:** Waveforms and spectrograms at 4 selected stations of Families 04 and 06. Stations are sorted by distance to the source near KRA1.

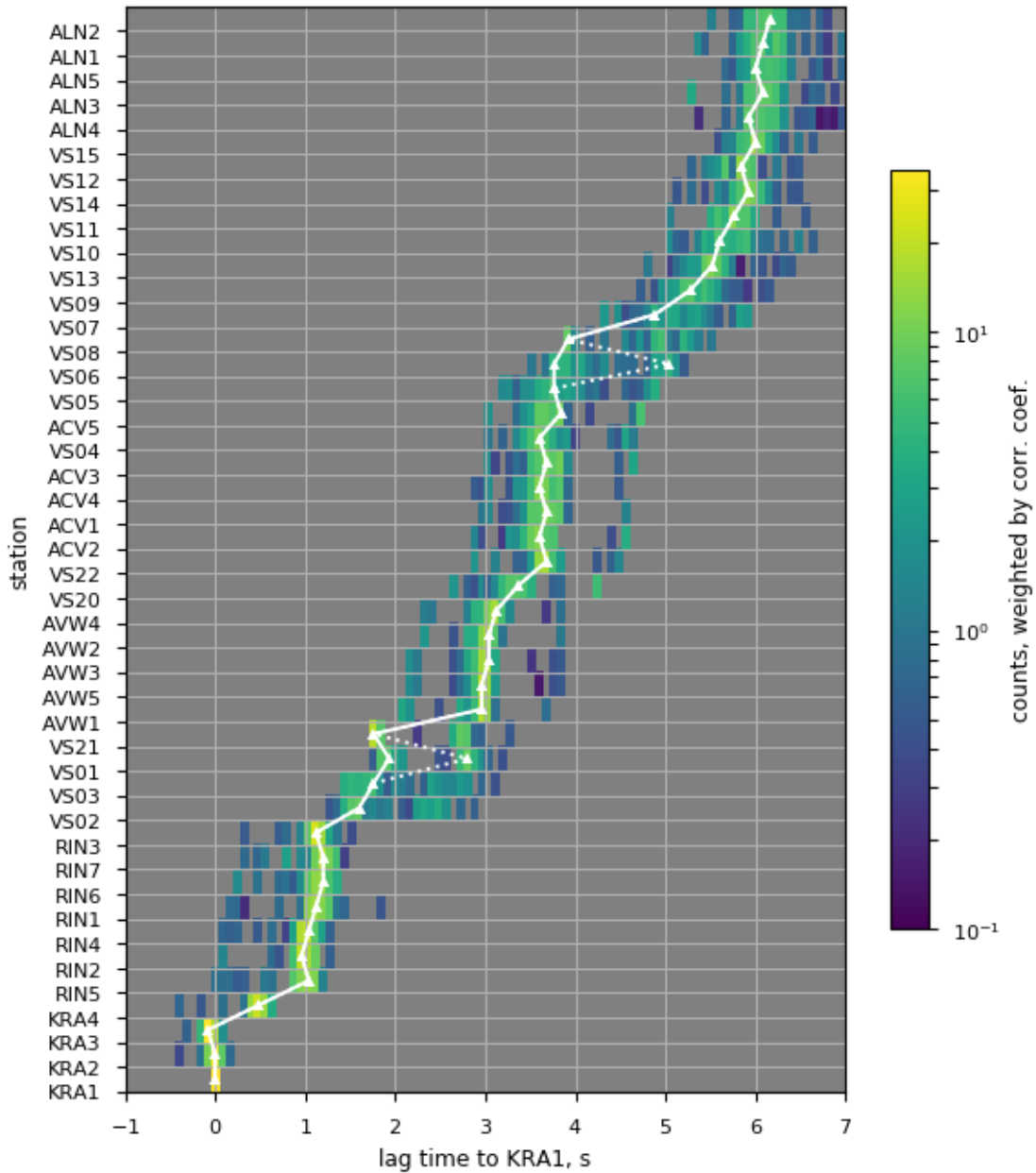
histograms at each station (Fig. 6.11). It must be noted though, that the maximum shift for the cross-correlation was restricted to  $\pm 0.8$  s. Differences larger than 1 s represented outliers of single stations rather than a systematic shift of one family. Firstly, no family showed such a large shift systematically at a larger number of stations (Fig. 6.12). Especially at stations close to the source, where the signal was still very similar to the one of KRA1 (and hence the identification of the onsets by cross-correlation very reliable), only very small variations were found. Secondly, the shift suggested by the result of the cross-correlation could not be confirmed upon visual inspection of these cases. Fig. 6.13 shows examples for three families. Stations, for which the lag time differed significantly from the average, clearly displayed a less distinct signal shape which obviously hampered a reliable cross-correlation result. This problem generally aggravated with increasing distance to the source.

The initial average curve (dotted lines in Figs. 6.11 and 6.13 showed implausible deviations at stations VS01 and VS06 with respect to the neighboring stations. The lag times of Family 6 were used to correct these outliers. At the stations in question, the alignment of the waveforms of Family 6 was more consistent than that obtained for other families but with lag times closer to the initial average.

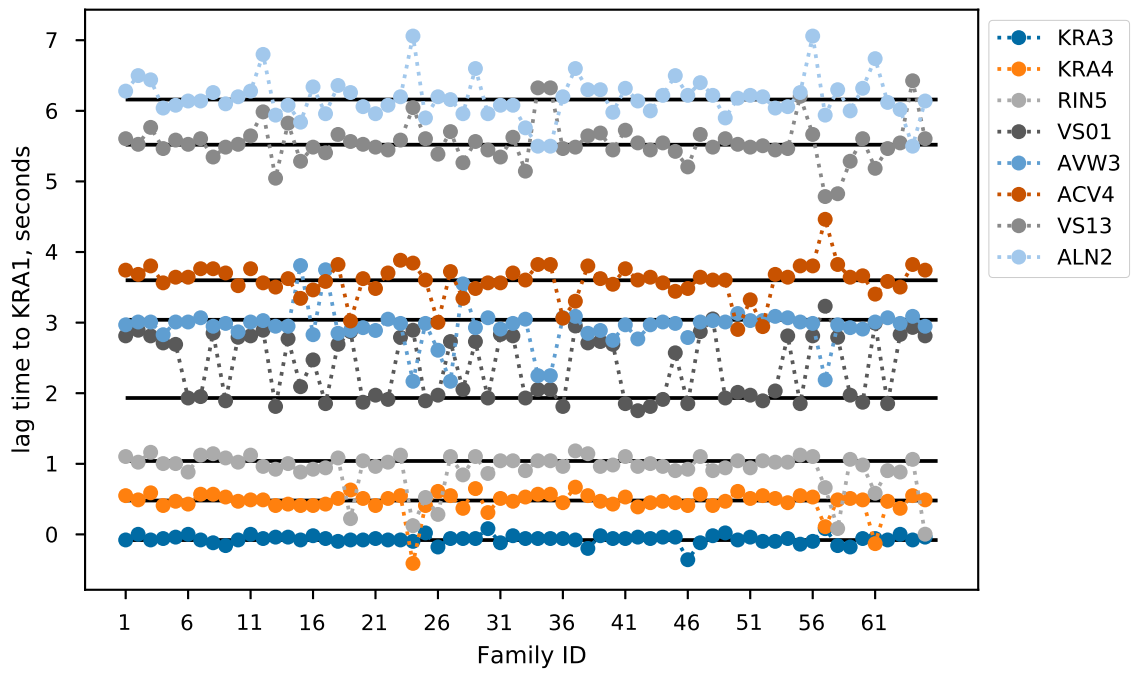
Fig. 6.13 also demonstrates, that the lag times for the three shown families were basically identical. Thus, in particular the two largest families seem to share a common source location.

The average lag times were compared to the theoretical lag times computed for a simple, homogeneous velocity model and wave propagation along the line-of-sight path between a hypothetical source at the crater and the stations (Fig. 6.14). The plot demonstrates the trade-off between unknown source depth and unknown velocity model. A deeper source and a faster velocity both decrease the lag times between stations. Especially, at most of the stations, the average lag times could equally be explained by a source at 2000 m a.s.l. and a velocity of 1.2 km/s and a source at 2800 m a.s.l. and a velocity of 1.5 km/s. Nevertheless, assuming that the sources were bound to the center of the crater, the signal seemed to originate from the upper 800 m. The fit for the shallower source (2700 m a.s.l., black dashed line) might be slightly better than for the deeper one. A shallower source would also be more consistent with the general observation, that the transients occur together with slug bursts in the lava lake.

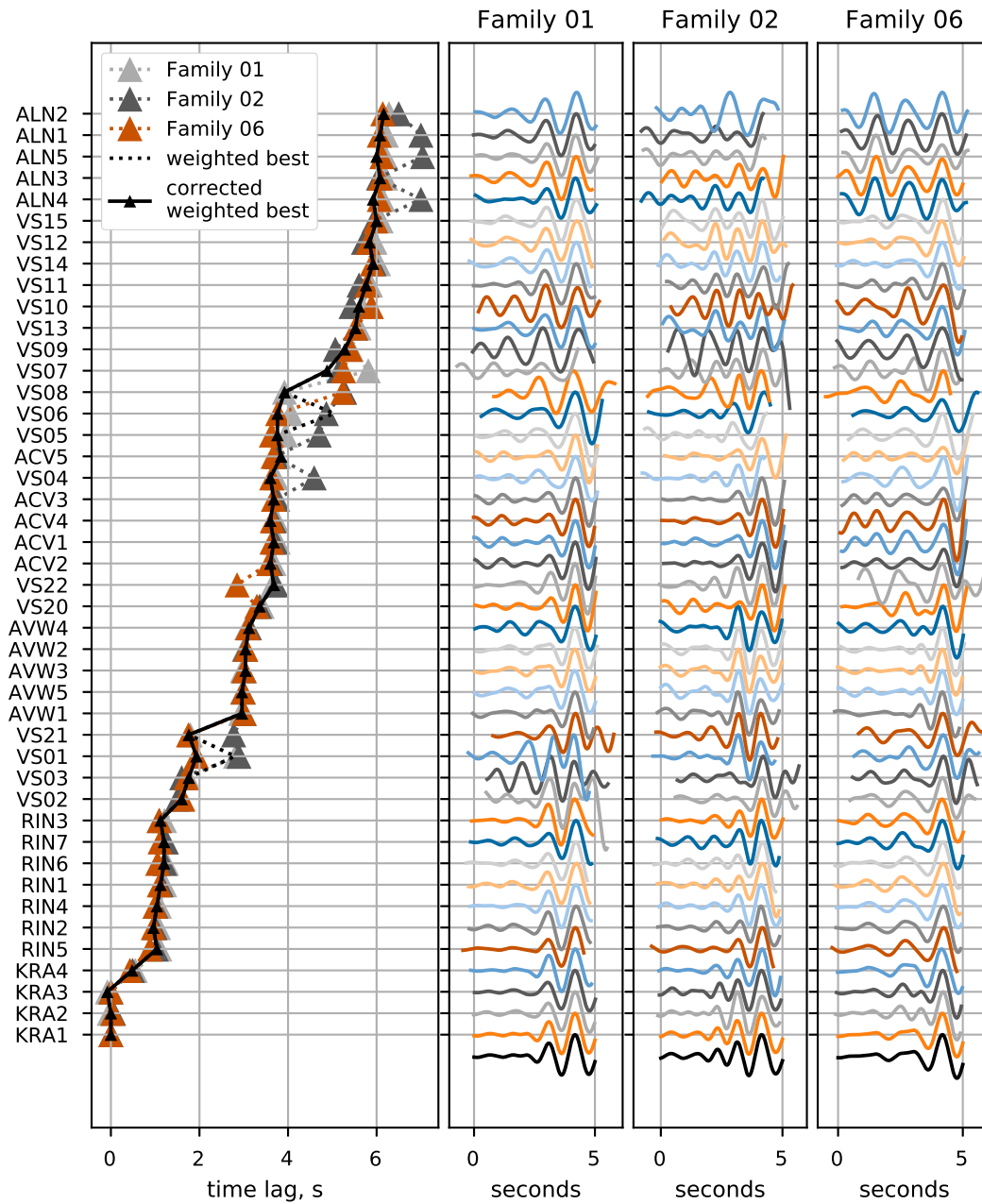
The onsets propagated with a velocity of 1-2 km/s which corresponds roughly to the S wave velocities found at the two arrays AVW and ACV. For the shallow source, the average velocity is around 1.5 km/s. Based on this model, wave propagation seems to be slower than average towards the west and southeast while faster towards the east. Closer to the summit (stations between KRA1 and VS21), the faster models tend to be more appropriate, suggesting velocities of 1.5-2.0 km/s.



**Figure 6.11.:** Occurrence of time shifts to KRA1. Stations are sorted by distance to the summit crater. For each station a weighted histogram was computed over the lag times of all families, weighted by the correlation coefficient. The dotted white line (mostly identical with the solid one) indicates the mode of this histogram at each station. Two outliers were corrected manually (see also Fig.6.13), resulting in the solid white line which was taken as the average lag time per station.

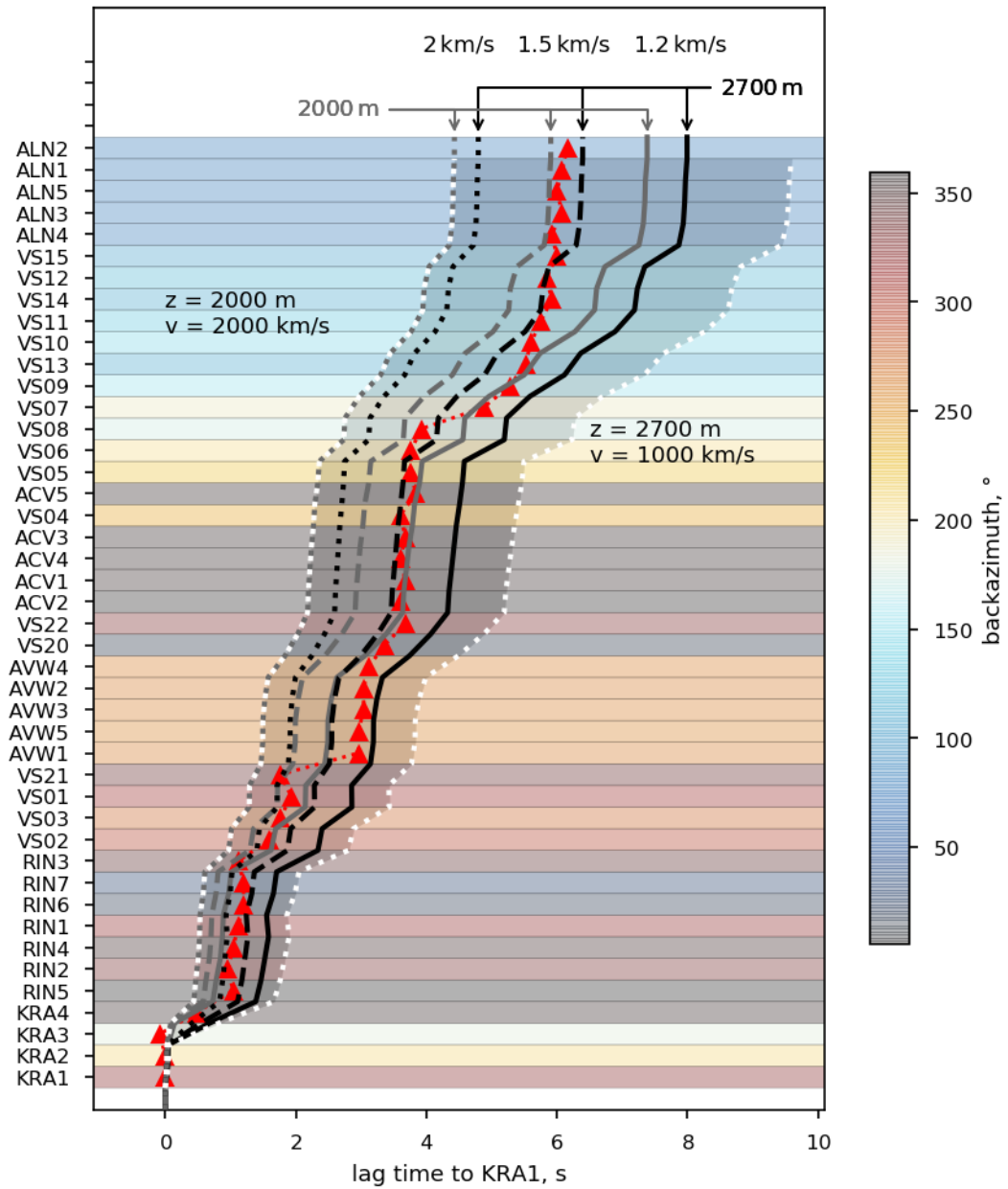


**Figure 6.12.:** Travel time lags to KRA1 for selected stations for all families. Solid black lines indicate the "mean" lag time. Most stations show very small fluctuations for most families. Deviations  $> 1$  s are rather a consequence of poor data quality rather than a true shift. VS01 is particularly affected, requiring manual adjustment of the "mean" lag time to make it consistent with neighboring stations.



**Figure 6.13.:** Lag times (left panel) for Families 1, 2 and 6 with corresponding data window (smaller panels) used for correlation. The samples are shifted by lag time, thus aligned to the signal at KRA1 (black). Stations are sorted by distance to the summit. Lag times of Family 6 were used to correct outliers in the average lags at VS01 and VS06. The resulting alignment fitted optically better than e.g. Family 2 which fell on the initial curve. Significant deviations from the average (mostly Family 2 (dark-grey)) result from poor signal quality at these stations, which hinders a meaningful correlation.

6. Waveform families of transient events



**Figure 6.14.:** Average lag times relative to KRA1 (red triangles) in comparison with lag times for homogeneous velocity models. The source was placed within the crater at depths of 2000 m (gray) and 2700 m (black) a.s.l.. Dotted, dashed and solid lines indicate different velocities of 2, 1.5 and 1.2 km/s, respectively. The shaded area between the two white dotted lines indicates the range of plausible models. Background colors represent the backazimuth of the station. Note the ambiguity between unknown source depth and velocity (compare black dashed with solid grey line).

### 6.5.3. Location

The analysis of the lag times with distance suggested wave propagation velocities for the onsets of 1000-2000 m/s. This range corresponds roughly to the S wave velocities found at the two arrays AVW and ACV. The initial energy could therefore be transported by body waves. The waveforms however became increasingly elongated with distance from the source, suggesting that the later parts traveled much slower and were probably subject to dispersion and scattering. Hence, the majority of the waveform and thus its energy propagated as surface waves.

Using the energy over the entire waveform and assuming surface waves, the ASL method yielded source locations for the transients mostly around the NW half of the summit (Fig. 6.16). Most sources were found close to station KRA1. Interestingly, no sources were found right at the center of the crater.

A similar circular distribution pattern as for the surface wave assumption was found using only the energy over the first 2.6s of the waveforms and assuming body waves (Fig. 6.17). Again, the sources distributed around the crater rather than inside of it. In addition, they were more spread out. Two depth layers clearly stood out: 2500 m a.s.l. and 2700 m a.a.l. - the latter effectively indicating a source at the surface. The relative similarity between the patterns of the source locations for the two wave types seems to confirm the initial assumptions, that the onset of the waveforms contain body waves while the bulk are surface waves. To ensure that the result for the body waves is unique for the onset, the same procedure was performed for a 2.6 s window starting 10s before the onset and a 5 window starting at the onset. Both resulted in a distribution of the locations that was clearly less similar to the result for the surface wave assumption.

The range of locations is largely consistent with the observed range of travel time lags between stations. Hence, the variations in lag time could be the result of variations of the horizontal position by some 100 m. A more detailed examination of Families 01, 02 and 06 however revealed some inconsistencies (Fig. 6.18). The body and surface wave location for Family 01 were found near station KRA1 at the northwestern rim of the crater while those of Families 02 and 06 were found to the southwest. Using body waves and the energy of the onset, a shallow source depths was obtained. Lag times were computed using a velocity of 1.5 km/s and compared to those obtained by cross-correlation of the onsets. At most stations the differences in the theoretical lag times between the locations were only minor which would correspond to the only small variations in the observed lag times. At stations AVW1-5 however, the theoretical travel times showed substantial differences which were not found in the observed lag times. This discrepancy suggests that the uncertainty of the source locations should be more carefully examined.

The Bootstrapping-analysis for Families 01, 02 and 06 demonstrated already a significant variation of several 100 m if different station sets were used (Fig. 6.19, top). Interestingly, if the ASL was applied to individual members of the family and higher frequencies were included, the source locations moved into the crater or at least closer

towards it (Fig. 6.19, bottom).

## 6.6. Discussion

67 families of transient waveforms and in total 23306 events were found using cluster analysis and matched filter detection. The majority of events belonged to one of two families. Their waveforms were very similar but they differed slightly in their frequency content. At the crater stations, their waveforms were very short, almost resembling a Ricker wavelet. The smaller families in contrast showed more complicated waveforms with high frequencies in the beginning, followed by a low-frequent part. A common interpretation is, that the high-frequent beginning marks an event that triggers a resonance effect which forms the later part of the waveform [Wassermann, 2012]. This could be, for example, a small explosion that sets the conduit in resonance.

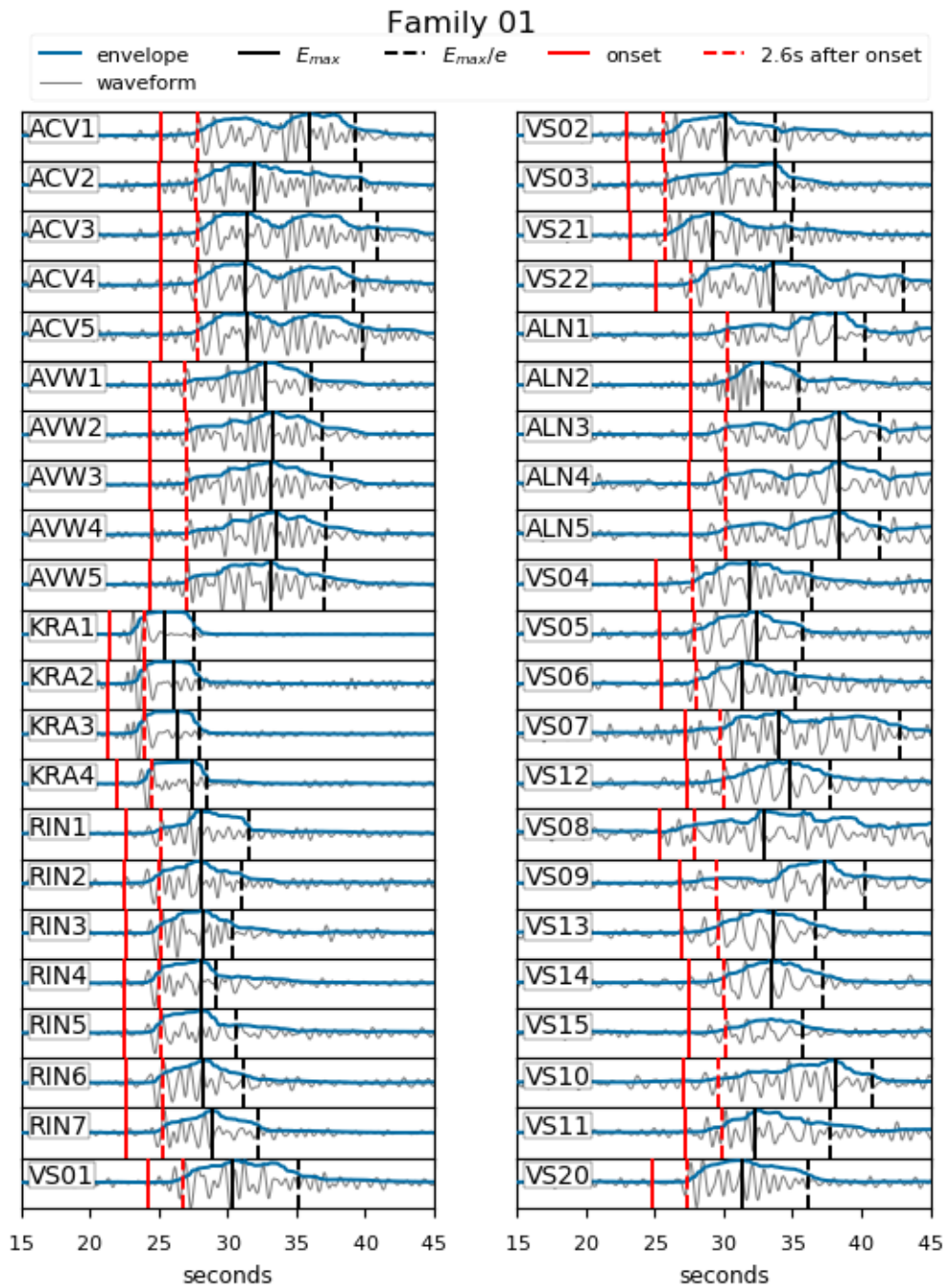
At Kilauea, Hawaii [Chouet et al., 2010] and Stromboli [Neuberg et al., 1994, Ripepe et al., 2001, Chouet et al., 2003], degassing bursts were associated with LP- and VLP-components. Similar combinations were also found in the seismic response of analogue and numerical models of degassing processes [James et al., 2004, 2006, 2008, Chouet et al., 2010]. Hence finding such a component might be a good indicator for a degassing process. Therefore it is somewhat unfortunate, that the limited frequency range of the deployed sensors precludes the analysis of the very-low frequency range.

The sources of the signals were located around the crater but the locations have a high uncertainty. The lag times between stations showed little variations between the families, suggesting a common source region. The reports from earlier studies that the transients occur in combination with Strombolian activity in the lava lake [Calder et al., 2004, Palma et al., 2008, Gurioli et al., 2008] were strongly supported by the statistical properties of the magnitudes and interevent times in the population of the transients (see Chapter 5). In this light, there is little reason to assume that the source of at least most of the transients lies outside of the crater. In this respect, it seems particularly odd, that the sources were found exclusively at the rim of the crater and not inside of it.

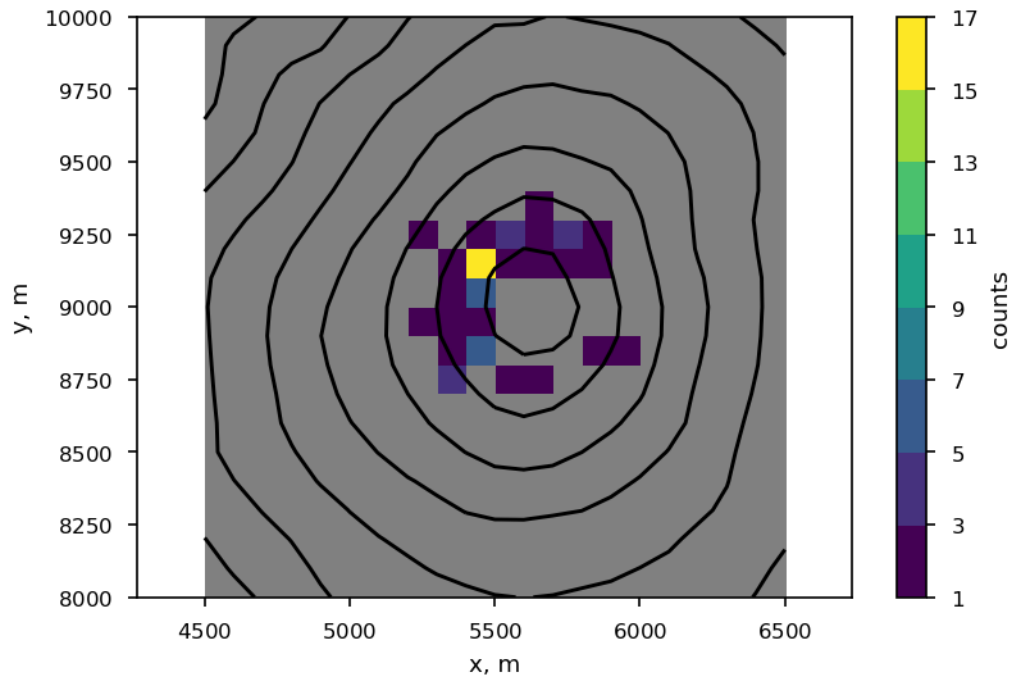
Nevertheless, other options should be taken into account. In particular, Lamb et al. [2020] raised awareness for the possibility that low-frequent volcanic events could be mistaken for icequakes at glaciated volcanoes. At Villarrica, this could play a role since the glacier extends up to the crater. Yet, in previous studies this has never seriously been considered.

More stations closer to the source and distributed at many different distances from and directions towards the crater could help to improve the results of the ASL. In contrast to the approach of Chapter 4, fewer stations were used. In particular, the two westernmost lines running towards the volcano were omitted because the events were mostly attenuated at this distance. In addition, the site factors also need to be very well known since they can also strongly influence the final location. For example, omitting them leads to

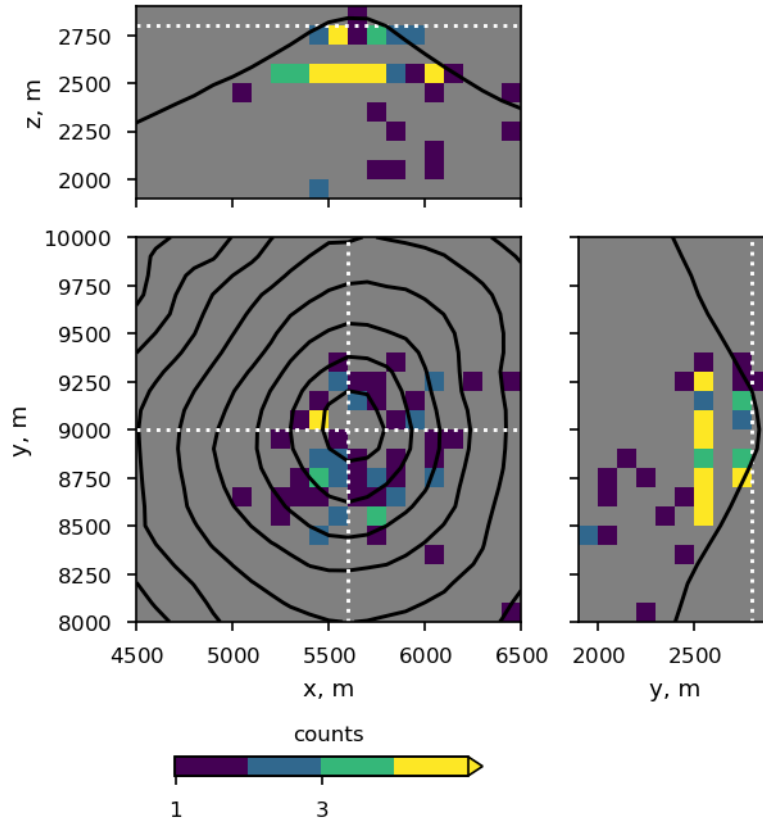




**Figure 6.15.:** Waveforms and envelopes of Family 01. Vertical lines indicate onset (red), energy maximum (black, solid) and time at which the energy dropped to  $1/e$  of the maximum. Energy for the amplitude source location method was computed by integrating the envelope between onset and  $E_{max}/e$

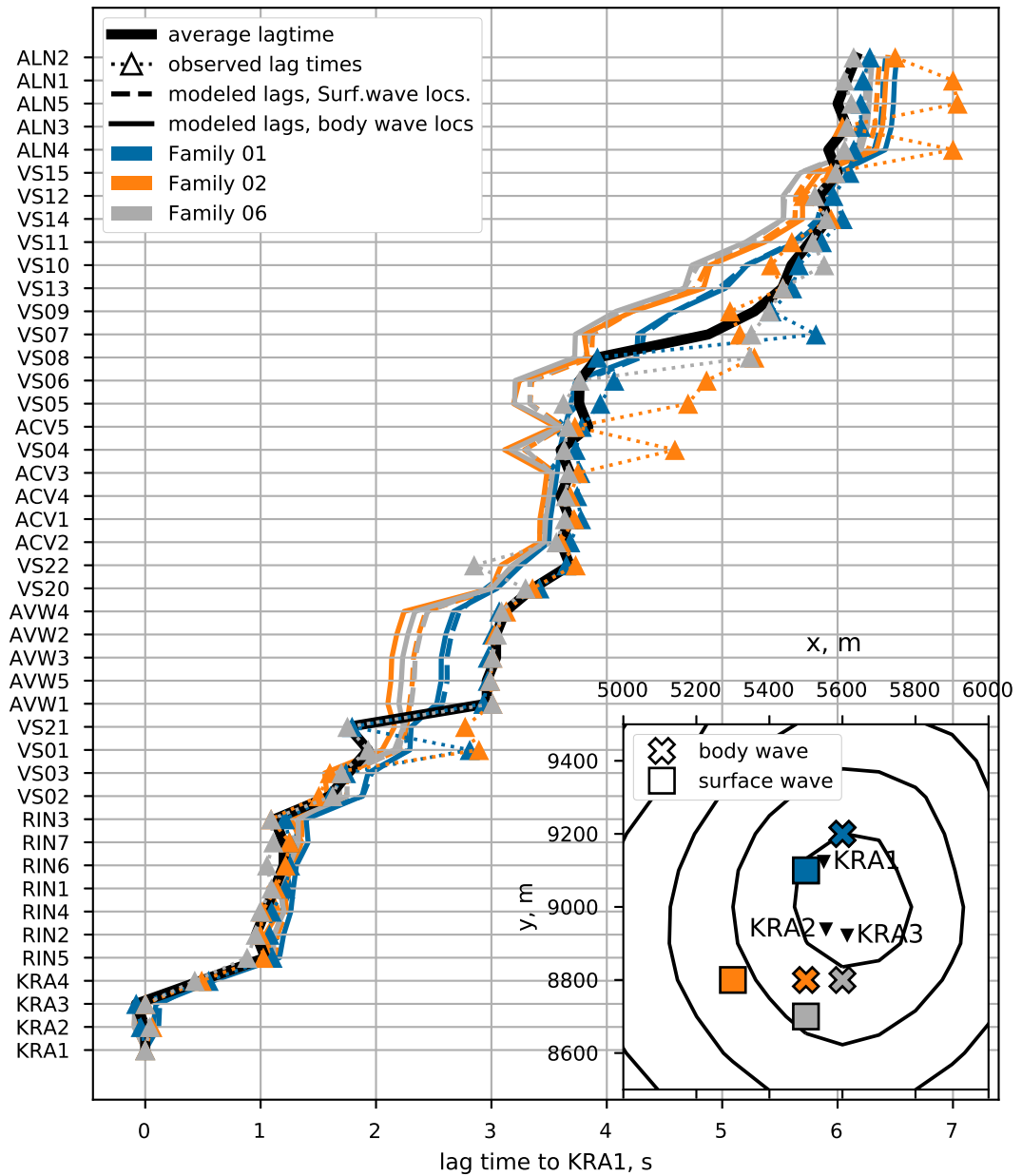


**Figure 6.16.:** 2D-histogram of source locations obtained by ASL on family centroids assuming surface waves. Therefore, no depths were obtained. Energy was computed between the onset and the time at which the energy dropped below  $1/e$  of the maximum. The highest number of events locations in one bin of the horizontal grid is 15.

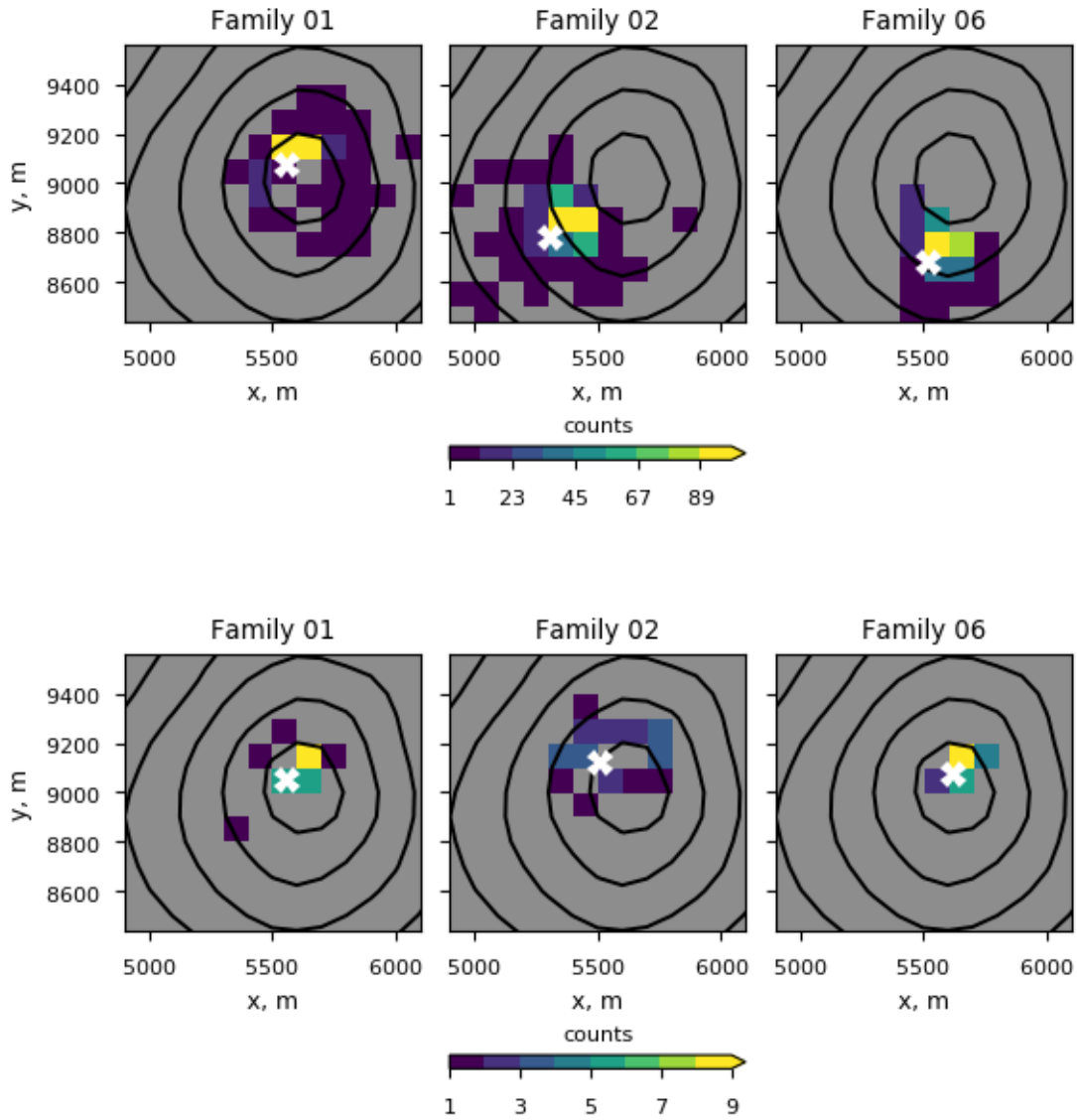


**Figure 6.17.:** 2D-histograms of source locations obtained by ASL on family centroids. In this case, the amplitude was computed over the first 2.6s of the waveform and body waves were assumed- Therefore it was also searched for a depth. Right and top panel show locations projected onto N-S and W-E cross-section through the summit. Bins correspond to search grid. The highest number of events locations in one bin of the horizontal grid is 12. Since surface waves were forced, the deep locations are physically implausible

6. Waveform families of transient events



**Figure 6.18.:** Observed and computed lag times for the source locations obtained by ASL of Families 01, 02 and 06 (color-coded). Lag times were computed for a velocity of 1500 m/s. Observed lag times are shown as triangles. The average observed lag time is shown as thick black lines. Thin and dashed colored lines correspond to the lag times computed for the respective locations for body (crosses) and surface (squares) waves shown in the inset.



**Figure 6.19.:** *top row:* Frequency of location results using a Bootstrap-selection of stations. *bottom:* Frequency of location results for individual family members instead of the stacked waveforms. The white cross marks the mean of x- and y-coordinates.

shifts by some 100 m.

The average lag times suggest a propagation velocity for the onsets of around 1500 m/s. Since this value is well in the range of the S wave velocity models determined for the two array sites north and west of the summit, this could be seen as an indication that the early part of the waveform is made up of body waves. Surface waves generally travel at slower speed although they can reach up to about 90% of the S wave velocity. Body waves, in principle, would be more helpful in determining the source depth.

For the inference of velocity variations from the relative travel times, it is particularly important that the onsets - which were identified by the initial crude manual picks and subsequent cross-correlation - are determined correctly. At the early groups of stations - KRA and RIN - the result was relatively unambiguous. At later stations however, alternative associations of the phases might be possible for some families. In these cases, the relative lag time to KRA1 would shift by  $\pm 1$  s (=1 period at 1 Hz). These alternatives are indicated by the parallel stripes in Fig. 6.11 at stations VS01-03, VS21, AVW1-5 and ACV1-5. In addition, if different polarities were allowed more ambiguities might arise. Therefore a more careful inspection of all families individually might be necessary to resolve these uncertainties.

### 6.7. Conclusion

More than 20000 events could be identified as members of one of 67 families the two largest of which comprise  $\approx 2500$  and  $>7000$  events, respectively. Hence a huge number of transients is produced by a stable, repetitive source. Stacked, average waveforms were obtained, which were subjected to further analysis in terms of waveform, travel time and location. The two largest families showed short, impulsive waveforms near the source but developed an elongated waveform within 1-2 km of the crater. Most of the other larger families displayed more complicated waveforms even at the crater which might be interpreted as a high-frequent trigger event that causes a rather low-frequent resonance response for example of the conduit.

The origin was located at or near the crater but several inconsistencies were found, suggesting a still significant uncertainty. In particular, most sources were actually located on the crater rim which contradicts the common assumption, that the events are explosion quakes. Hence the source was expected to be within the crater and not outside of it. However, if the preliminary results turned out to be true, alternative source processes would need to be taken into account, for example icequakes in the glacier.

The presented results were obtained in a sort of batch processing manner. In this way, the large number of families could be analyzed very quickly. However, a more careful and individual analysis of the event families - which should in particular include a meticulous quality control of the events taken into account - might significantly improve the reliability of the results. In this regard, the methods and results shown here, should

be seen more as a proof of concept which indicate the potential gain of information. Using particularly strong events and stacked waveforms, it seems realistic to locate the source of the signals. The stacked waveforms allow to identify the onsets of the signals and thus the determination of travel time differences. These may be used to infer a velocity model. Another aspect, which was not tested yet, could be the analysis of the damping and scattering parameters of the medium.

Part III.

Concluding remarks



## 7. Overarching discussion & conclusions

This work presented analysis on  $\sim 12$  days of seismic data from Villarrica Volcano, recorded by an extensive local network in early March 2012. The data were analyzed in terms of the seismic activity and the seismic structure of the volcano. The following section discusses the main conclusions regarding the objectives of the thesis.

### 1. Locating the source of the seismicity

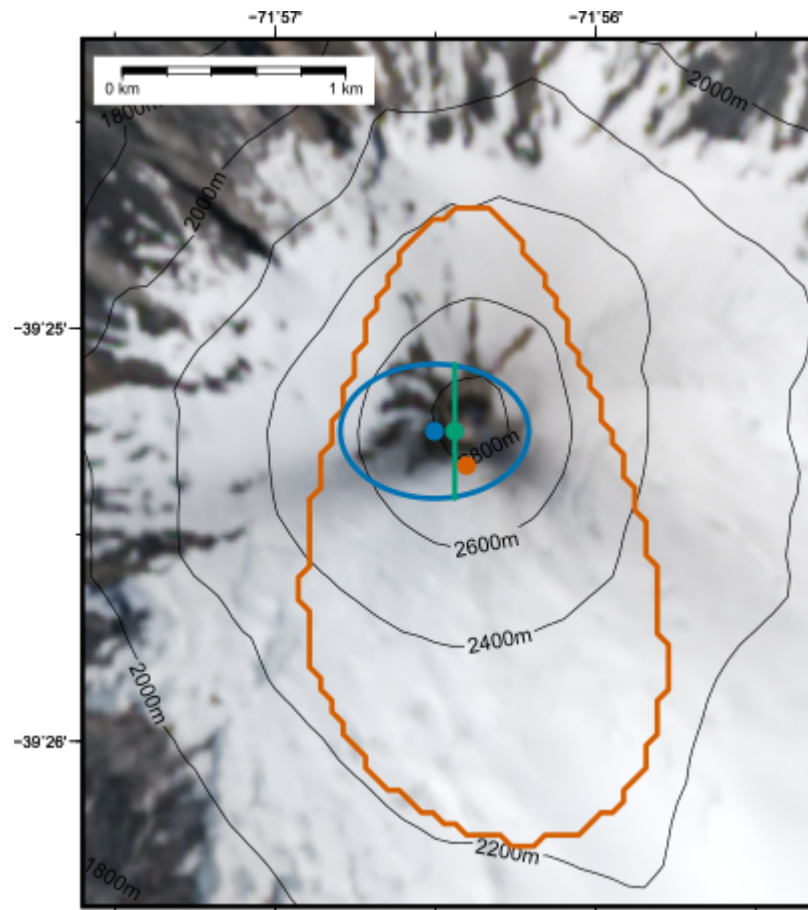
The origin of the seismic unrest of Villarrica volcano could clearly be located in the summit crater confirming the hypothesis of earlier studies (Fig. 7.1). Three different approaches were presented, two of which were based on the decay of amplitude with distance and one on array analysis. No distinction was made between tremor and transient events. Instead the respective parameters for each method - backazimuths for the array analysis, peak power spectral density for the amplitude methods - were obtained from the probabilistic distributions over the observation period.

Possible small variations in source location could not be resolved in this manner. The horizontal location derived by the array approach varied over time by a few hundred meters but since these variations fell well within the uncertainty range of the method, they were not considered as true changes in location.

Both, the array analysis as well as the fitting of the amplitude decay to distance, suggested that the analyzed signal was composed predominantly of surface waves. Unfortunately, this impeded the determination of a source depth, since surface waves are only excited once the signal impinges at the surface. Consequently, even if the source was located deeper it could only be traced back to the epicenter not its depth. The strong surface waves and scattering of the transient waveforms with distance however suggested a rather shallow source [Cesca et al., 2008].

Plausible candidates for source mechanisms of the tremor are superimposed codas from transients (and thus probably the bubble bursting and magma seething at the surface of the lava lake) as well as frictional forces exerted by magma convection on the conduit walls. Although, in the latter case, the assumption of a point source would actually not be very accurate. Alternatively, or additionally, resonance effects along the conduit or sections of it may play a role.

Using the fitting of amplitude decay to distance (amplitude source location - ASL), the individual families of transient waveforms could also be located to the summit area. Again, the predominance of surface waves impeded an estimation of the source depth.



**Figure 7.1.:** Locations of the tremor with uncertainty. Colors indicate the different location approaches: orange - Beamforming, blue: fit of amplitude decay with distance, green: mapping of amplitudes. Locations of transient families are not shown because of they were too ambiguous.

---

Lateral variations in location between event families were found but their uncertainty was too large to reliably identify differences in their source location.

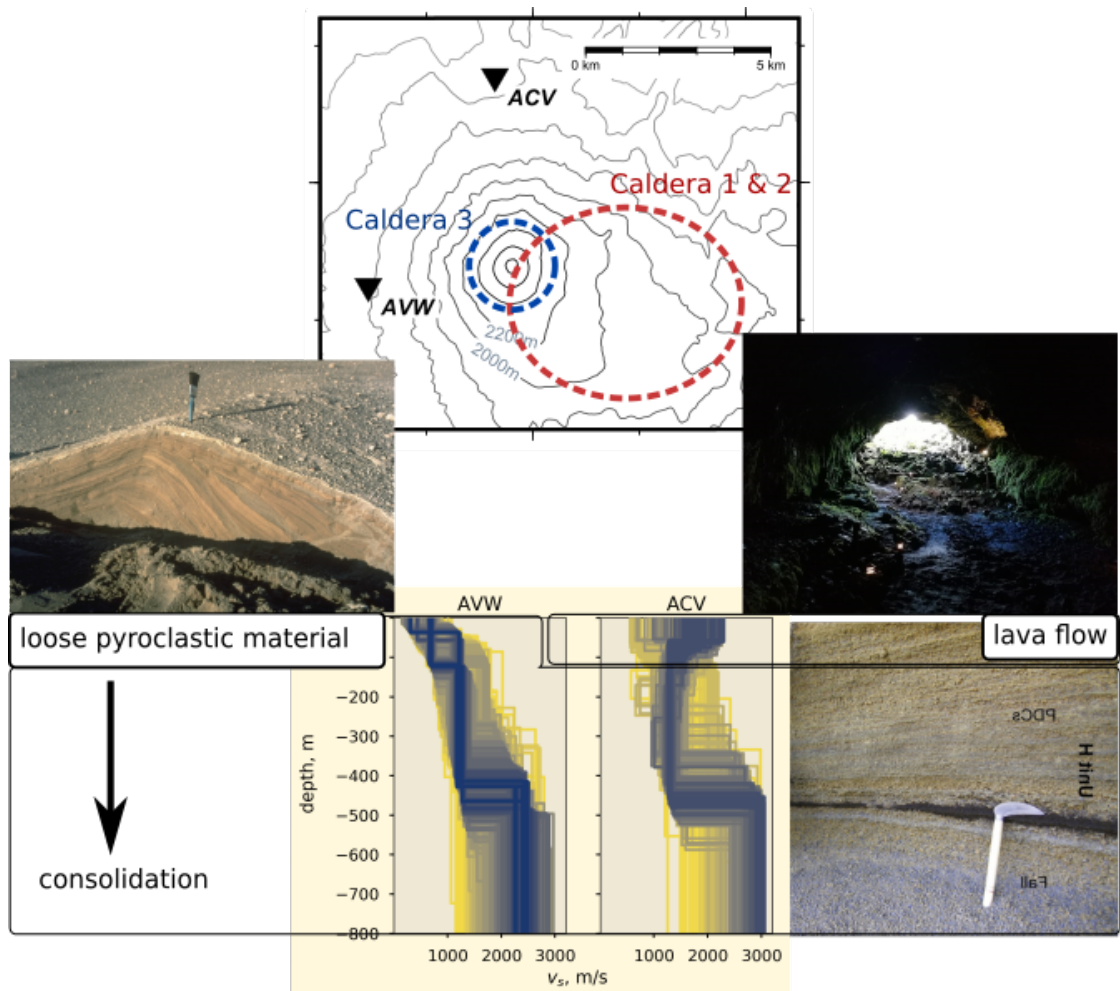
The similar source region of tremor and transients indicates a close relationship between both signal types. Yet, the slight differences in lag times relative to station KRA1 for the waveform families suggest that there could be variations in source locations of the transients, most likely the depths. Identifying these locations could greatly improve our understanding of the upper conduit system since they might indicate changes in the conduit system. Unfortunately, with the present data set, precise locations may be difficult or even impossible to achieve. Due to the large distance to most stations and notably the arrays, the detected signals are more likely to be affected by path effects. The amplitude method would also benefit from more stations at different and most importantly shorter radii around the summit since the method relies on the distance to the source.

Some of the sources were located outside of the crater area. Under the assumption that the transients are caused by explosions, these results were considered as erroneous locations. Given the volcanic activity in the lava lake and the limited resolution power of the network, this assumption is very well justified. Nevertheless it should be kept in mind, that the volcano is covered by a glacier and that icequakes can display similar characteristic traits as the transient events. Hence, some of them were possibly icequakes and thus originated from outside the crater. For a reliable differentiation, precise location methods and simultaneous monitoring of the volcanic activity by non-seismic methods would be required.

## **2. Elucidating the shallow seismic velocity structure of the edifice**

At two sites - 4-5 km west (AVW) and north (ACV) of the summit - dispersion curves of the wave propagation velocity could be gained from the array analysis. The dispersion curves were inverted for seismic velocity models of the upper 500 m at the respective locations (Fig.7.2). At both sites a 2-Layer structure was found, with boundaries at similar depths of 50-100 m and 400-500 m. S wave velocities for the deeper layer and the subjacent half-space ranged from 1000-3000 m/s and were interpreted as consolidated pyroclastic material. The surface layers differed considerably, although both could be explained by the local surface structure: AVW was installed on young deposits of pyroclastic density currents which have very low velocities of less than 1000 m/s, whereas ACV was situated on an old lava flow leading to the high velocities of  $\approx 2000$  m/s.

These results provide the first substantial evidence for the shallow seismic structure of Villarrica volcano. They may serve as a valuable base for future studies, that involve modeling of the seismic wave propagation, e.g. the inversion for moment tensors. A good velocity model also significantly improves the results of conventional source location methods which can be applied especially to VT-events and possibly in some cases to other event types.



**Figure 7.2.:** Interpretation of the shallow velocity models at array sites AVW (4 km west of summit) and ACV (5 km north of the summit), obtained from inversion of dispersion curves. The upper ~100 m are covered by loose pyroclastic material (AVW) and solid lava (ACV). Lower layers are made up of consolidated pyroclastic deposits. Extrapolation to other locations are generally difficult. Especially the old calderas might introduce a significant change in the velocity structure. Pictures (clockwise from upper left): US Geological Survey [1980], Vito0783, CC BY-SA 4.0 [2019], Costantini et al. [2011]

---

Nevertheless, extrapolation of the models to other areas of the volcano should be made only very carefully. Radial and azimuthal variations in the velocity structure with respect to the summit are indicated by the lag times between different stations for the transient event families (Fig. 6.12). The relative similarity of the two models underneath the surface layer suggests that they are to some degree representative for the northwestern sector. However, the arrays were also located at comparable radial distance from the crater of 4-5 km, so there is no information on whether the model can be extended radially inwards to the summit. Major geological features that likely introduce a change in the velocity structure, are the glacier and the older calderas. Since the youngest cone of the volcano is situated at the western rim of the older calderas, the lower edifice is strongly asymmetric with the glacier and old caldera extending far more to the east than to the west of the summit. The velocity models are therefore probably not easily transferable to the eastern half of the volcano.

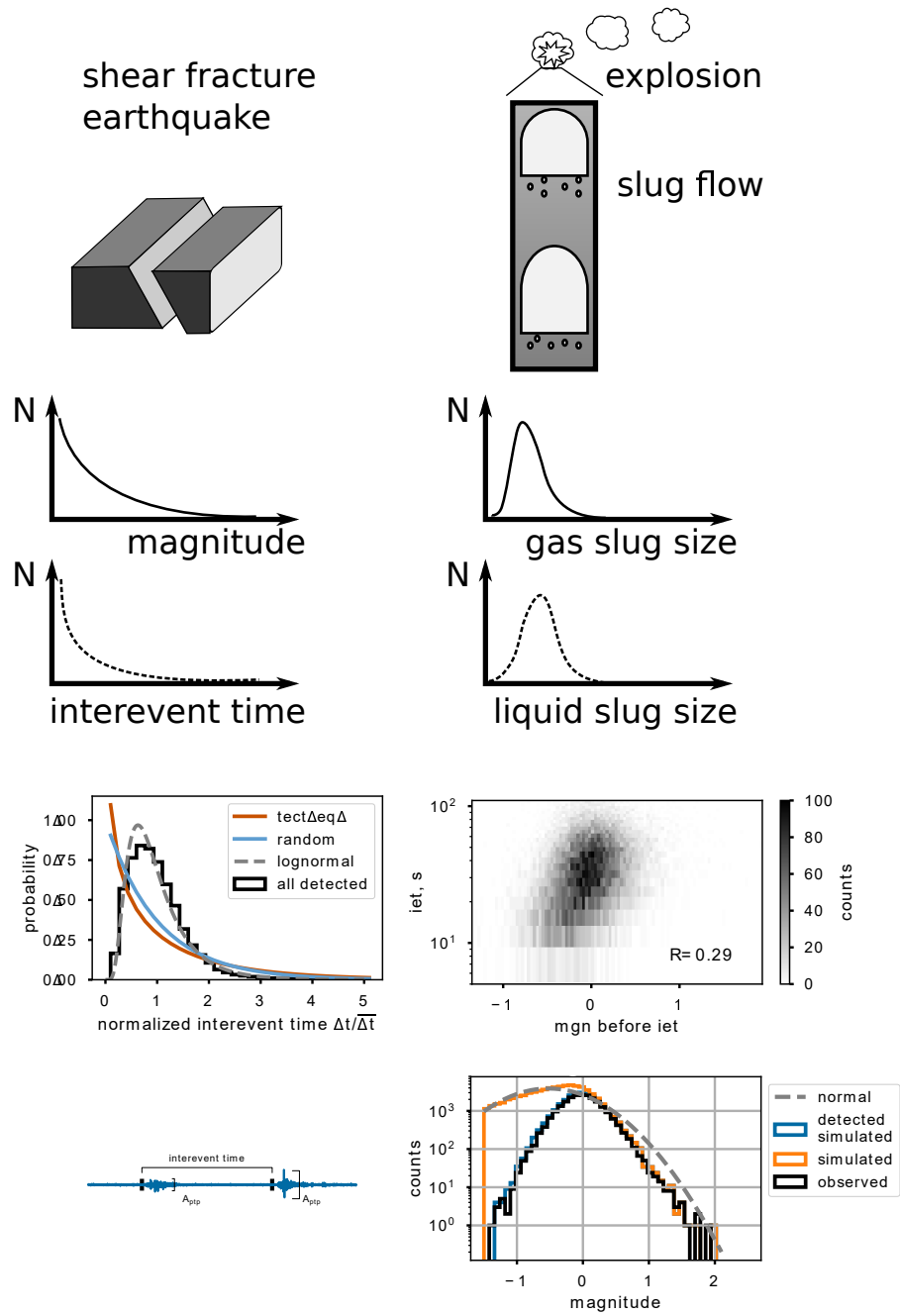
### 3. Characterizing the occurrence of transient events

A catalog of automatically detected transient events was investigated in terms of the statistical properties of magnitudes and interevent times between two consecutive events (Fig 7.3).

Based on the stochastic properties of the noise, obtained from the data, a detection curve was determined that indicated the probability to detect a given magnitude. This function in turn allowed to reconstruct a model, that reproduced the observed distribution of magnitudes. This distribution decayed towards low magnitudes after reaching a peak. If interpreted as a normal distribution, it would indicate log-normally distributed amplitudes. The distribution of interevent times resembled a log-normal distribution, indicating a preferred recurrence interval, in other words a slight periodicity. The interevent times could however only be determined between observed events. Although the magnitude model suggests the presence of additional events, it does not give a timing of them. Nevertheless, for the actually observed data, a weak correlation was found between the logarithm of the interevent time and the preceding magnitude.

The frequency distributions of both parameters clearly differ from those of tectonic earthquakes which are characterized by exponential magnitudes and a (nearly exponential) Gamma distribution of interevent times. Moreover, a random occurrence - signaled by exponential interevent times - of the transients could be discarded. Instead, the distributions could be the result of a slug flow in which the transients reflect the gaseous slugs that move in the liquid magma. The magnitudes would reflect the size of these slugs and the interevent times would indicate the separation distance. In slug flow, the sizes of the gas pockets and the liquid bridges between them are log-normally distributed. The interpretation as slug flow is further supported by the correlation between interevent time and preceding magnitude.

While the results strongly suggests slugs as the origin of the transients, they do not



**Figure 7.3.:** Summary of the statistical analysis of the transient events in terms of magnitudes and interevent times.

---

indicate whether the transients are the results of the bursting of a slug at the surface or merely a signal of a slug passing an obstacle in the plumbing system at depth. Both processes are known to produce similar waveforms as the transients. In principle, it may even be possible that the same slug produces two (or even more) transients: one at depth and one upon bursting in at the free magma surface. This case however was not considered in the statistical considerations since there was no way of discriminating possibly shallow and deep sources.

The detection simulator was a key element in the analysis because it allowed inferences on the statistical model of the magnitudes beyond the magnitude of completeness of the catalog. Otherwise, the decline in low magnitudes - which is a strong indication against exponential, and thus tectonic, magnitudes - could not have been found. Inferring a model for the interevent times in a similar manner is slightly more complex (and therefore was not done yet) but should in principle be possible.

The data set is relatively old for such a dynamic system as an active volcano. Therefore it is worth noting that there is little indication of drastic changes in Villarrica's behavior or edifice in the past few years, even though the volcano saw a major eruptive phase in 2015. The characteristics of the seismic and infrasonic tremor were remarkably stable throughout the past 20 years. During the same period, the occurrence interval of the transient events was always reported as approximately 1 min [Ortiz et al., 2003, Calder et al., 2004, Palma et al., 2008].

In particular the results obtained in the context of objectives 1 and 3 largely corroborated postulations, hypothesis and findings from other studies that 1) located the origin of the seismicity of Villarrica at the crater and 2) attributed the transients to gas slugs. The thesis focused however on a data-driven, methodically sound analysis rather than relying on anecdotal evidence from field observations. In particular the relation between slugs and transients was based on the latter. Regarding objective 2, this thesis provided first results regarding the near-surface velocity structure of Villarrica volcano, albeit very sparse. Potentially, the approach to use the correlations of onsets of the stacked waveforms families, to determine relative travel times could be extended to gain a more widespread picture of the velocity variations around the edifice.

## 8. Recommendations for future seismic installations

The network, as is was installed at that time, provided a valuable data set of unprecedented detail on the seismic activity of Villarrica volcano. It was designed (and well suited) to detect VT-events to use them for a travel time tomography [Mora-Stock, 2015]. As shown in this work, it also served well for analyzing the volcanic activity from the crater area. However, several shortcomings were encountered while working on the data, which ruled out more in-depth analysis. The following list contains some suggestions and recommendations for future installations, notably if they were to focus on the activity in the crater area.

**Install more stations closer to the source (crater)** : Generally, this should minimize path effects and scattering on the waveform and wave propagation.

**Probe the radius and the azimuth:** The circular deployment of most stations around the edifice provided great azimuthal coverage albeit at a considerable distance from the source. If possible, several rings of stations at different radii would be desirable. A better sampling of the proximal distances could, in particular, improve the results of the amplitude decay methods.

**Use 3-component stations:** Proper 3-component station data would allow e.g. inversion for source mechanism or analysis of particle motion.

**Orient stations upon installation:** If 3-component stations are used, they must be oriented properly. Realignment of the components on the base of regional earthquakes in post-processing as it is common practice for data from Ocean Bottom Seismometers turned out to be virtually impossible: even the particle motion of a strong regional tectonic event was strongly distorted in the summit area, rendering the attempts to realign the components in post-processing impossible. Small deviations are tolerable, especially since the source is close but absolutely no certainty about the orientation at all clearly precluded most advantages of 3-component data.

**Use at least some broadband stations:** A significant portion of the seismicity seems to occur at frequencies high enough to be detected with geophones. Geophones are cheap, lightweight and easy to deploy which makes them good alternatives to seismometers for installations in difficult terrain if the frequency content of the targeted



---

signals is high enough. Nevertheless, these instruments are less sensitive and essentially deaf for lower frequencies (in our case 0.45 Hz). As a result, signals such as VLPs, if they were present, could not be observed. This is particularly unfortunate for the waveform analysis of the transient events. It would be very interesting to see whether there is a VLP-component to these signals that indicate the ground response to the migration of gas slugs as it was e.g. observed for Strombolian explosions at Stromboli [Neuberg et al., 1994, Chouet, 2003]. Moreover, better sensitivity to low frequencies might extend the dispersion curves, derived by the array analysis, to lower frequencies, and thus enhance sensitivity to lower velocity layers.

**Run longer observation periods:** Of course, an extensive network for a detailed analysis can only be operated for a relatively short time. However, it should be considered, to leave a reduced installation in place for longer to provide more context for the analyzed activity.

**Use mini-arrays:** The mini-arrays proved very useful in terms of source location as well as velocity analysis. However, in order to serve both purposes, the stations must be distributed in both, radial and azimuthal direction. E.g. the RIN-arrays had good sensitivity for the backazimuth but poor sensitivity to the slowness due to the stations being aligned basically circular around the source. More such arrays could be used to probe the velocity structure of the edifice. Additional stations to reduce the spacing could improve resolution at higher frequencies. Broadband stations could be used to detect longer periods and thus provide information on deeper velocity layers.

**Go interdisciplinary:** Especially the volcanic activity, such as degassing, explosions, and the dynamic of the lava lake can and should be monitored with additional methods, like SO<sub>2</sub>, thermal or visual cameras, gas composition sensors or infrasound. The imaging methods can be very costly to operate, so they might provide data only for a few days. Nevertheless, this could give valuable insights into the processes associated with the seismic activity.

## 9. Outlook

The most fundamental question regarding the low-frequent seismicity at Villarrica are:

- Which processes cause the tremor and transient events?
- Are they related?

This thesis provides some clarity regarding the location of the source of both signals and the nature of the transients. In this thesis the origin of the tremor and transients could clearly be located in the summit region and the transients could be attributed to slug flow and bursting. However, the similar horizontal source position and the lack of information on the depth do not exactly make it easier to untangle the two signal types. It is still unclear if the tremor is the result of superimposed transient waveforms or a different signal.

An imminent question regarding the seismic transients remains, at what depth they occur. Are they solely the result of bubble bursting at the surface of the lava lake or can they also originate at depth? In this work, the determination of the source depth was impeded by the wavefield being dominated by surface waves. Although this could mean that the source is indeed a shallow one, it could also "just" be an effect of the medium. Even a shallow source could be at the surface (like a slug surfacing in the lava lake) or only near the surface (like a bubble pushing through the shallow plumbing system). Here, infrasonic records and imaging method of the degassing and lava lake activity (e.g. visual, thermal or SO<sub>2</sub> cameras) could help to discriminate between surficial and deeper seismic sources. This in turn would help to better understand whether the statistical distributions of interevent times and magnitudes represent a single process (such as surficial Strombolian explosions) or multiple processes (events at the surface as well as depth).

An extensive multi-parametric study could also reveal details of the explosion process. Measurements of high-resolution gas flux, gas-composition, temperature of the plume, infrasound and seismicity were all deployed at some point in time at Villarrica. However, in order to study the interplay between all of them, they must be available simultaneously at comparable quality and time-resolution. Seismology can play a key role in tracking the processes through time. Imaging techniques of the plume and lava lake (temperature, gas, visual,...) can provide great details of the process but are costly to deploy. Such data can thus be acquired only for short time periods. In contrast, seismology is a well established monitoring tool which delivers almost continuous data over relatively long

---

periods of time. If it was possible to identify the seismic signatures of the activity in the vent such as the different styles of degassing (bubble bursting, seething, lava fountains,...), one could track this activity over longer time periods and from save distance.

With the analysis from the mini-arrays, some information on the seismic structure of the edifice were gained. However, in order to understand the signal propagation from the crater region more details must be gathered to build an adequate model of the crater area. It was demonstrated that minimalist arrays suffice to determine the near-surface velocity structure. Wave propagation in volcanic areas is still modeled using relatively simple models. On the one hand, the details are simply unknown; on the other hand, the computational cost for a complex model is still relatively high. Nevertheless, several studies already demonstrated that seismic waveforms, particle motion or moment tensors are strongly influenced by the topography and complex medium of a volcano [Neuberg and Pointer, 2000, Cesca et al., 2008, Bean et al., 2008, O'Brien and Bean, 2009]. Therefore, the structural complexity should be taken into account when analyzing seismic waveforms from volcanoes.

Undoubtedly, extensive, interdisciplinary data sets and modeling efforts are necessary for a detailed picture of the processes associated with a volcano. They require however an immense effort to congregate. Therefore, their availability is limited to a only a few locations and time periods. Based on their analysis, generalizations and simplifications must be made that allow investigations and in particular long-term monitoring of other volcanoes with reduced resources. The statistical analysis of seismic events to determine their nature, which was introduced in this works, might be a potential candidate for such a simplified analysis tool. The necessary data can be measured using only a single instrument and the analysis can be almost fully automated. However, the gained information so far is relatively basic. Further studies could investigate for example, whether the statistics could also be used to infer rheological properties of the magma. In might also be necessary to be able to distinguish between different processes. On the one hand, the seismic events could be classified based on their waveforms before entering the statistical analysis. On the other hand, more elaborate statistical methods might be able to identify different populations of events and thus separate them into different processes. The outcome of such attempts must initially of course be carefully evaluated against information gained by independent methods before being used as sole source.

## Bibliography

- Sarah and Iain, CC BY 2.0. File:Villarica Volcano (aerial view)1.jpg — Wikimedia Commons, the free media repository. Creative Commons Attribution 2.0 Generic, Jan. 2007. URL [https://commons.wikimedia.org/w/index.php?title=File:Villarica\\_Volcano\\_\(aerial\\_view\)1.jpg&oldid=455394616](https://commons.wikimedia.org/w/index.php?title=File:Villarica_Volcano_(aerial_view)1.jpg&oldid=455394616). [Online; accessed 25-March-2021].
- A. Aiuppa, M. Bitetto, V. Francofonte, G. Velasquez, C. B. Parra, G. Giudice, M. Liuzzo, R. Moretti, Y. Moussallam, N. Peters, G. Tamburello, O. A. Valderrama, and A. Curtis. A CO<sub>2</sub>-gas precursor to the march 2015 villarrica volcano eruption. *Geochemistry, Geophysics, Geosystems*, 18(6):2120–2132, June 2017. doi: 10.1002/2017gc006892.
- H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, Dec. 1974. doi: 10.1109/tac.1974.1100705.
- K. Aki. Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bull. Earthq. Res. Inst.*, 35:415–456, 1957.
- K. Aki and B. Chouet. Origin of coda waves: Source, attenuation, and scattering effects. *Journal of Geophysical Research*, 80(23):3322–3342, Aug. 1975. doi: 10.1029/jb080i023p03322.
- K. Aki and P. G. Richards. *Quantitative Seismology, 2nd Ed.* University Science Books, 2002.
- K. Aki, M. Fehler, and S. Das. Source mechanism of volcanic tremor: fluid-driven crack models and their application to the 1963 kilauea eruption. *Journal of Volcanology and Geothermal Research*, 2(3):259–287, sep 1977. doi: 10.1016/0377-0273(77)90003-8.
- J. Almendros and B. Chouet. Performance of the radial semblance method for the location of very, long period volcanic signals. *Bulletin of the Seismological Society of America*, 93(5):1890–1903, 2003.
- J. Almendros, B. Chouet, and P. Dawson. Spatial extent of a hydrothermal system at Kilauea Volcano, Hawaii, determined from array analyses of shallow long-period seismicity: 2. Results. *Journal of Geophysical Research: Solid Earth*, 106(B7):13581–13597, July 2001a. ISSN 2156-2202. doi: 10.1029/2001JB000309. URL <http://onlinelibrary.wiley.com/doi/10.1029/2001JB000309/abstract>.

- J. Almendros, B. Chouet, and P. Dawson. Spatial extent of a hydrothermal system at Kilauea Volcano, Hawaii, determined from array analyses of shallow long-period seismicity: 1. Method. *Journal of Geophysical Research: Solid Earth (1978-2012)*, 106 (B7):13565–13580, 2001b. doi: 10.1029/2001jb000310.
- J. Almendros, B. Chouet, P. Dawson, and T. Bond. Identifying elements of the plumbing system beneath Kilauea Volcano, Hawaii, from the source locations of very-long-period signals. *Geophysical Journal International*, 148(2):303–312, 2002.
- T. Amadej. Understanding and parameter setting of STA/LTA trigger algorithm. *New Manual of Seismological Observatory Practice (NMSOP)*, 2009. doi: 10.2312/GFZ.NMSOP\_R1\_IS\_8.1.
- Andrew Shiva / Wikipedia / CC BY-SA 4.0. File:Soufrière Hills volcanic aftermath (Aerial views, Montserrat, 2007) 02.jpg — Wikimedia Commons, the free media repository. Creative Commons Attribution-Share Alike 4.0 International, Apr. 2007. URL [https://commons.wikimedia.org/w/index.php?title=File:Soufri%C3%A8re\\_Hills\\_volcanic\\_aftermath\\_\(Aerial\\_views,\\_Montserrat,\\_2007\)\\_02.jpg&oldid=474710920](https://commons.wikimedia.org/w/index.php?title=File:Soufri%C3%A8re_Hills_volcanic_aftermath_(Aerial_views,_Montserrat,_2007)_02.jpg&oldid=474710920). [Online; accessed 25-March-2021].
- D. Barnea and L. Shémer. Void fraction measurements in vertical slug flow: applications to slug characteristics and transition. *International Journal of Multiphase Flow*, 15(4): 495–504, July 1989. doi: 10.1016/0301-9322(89)90049-9.
- D. Barnea and Y. Taitel. A model for slug length distribution in gas-liquid slug flow. *International Journal of Multiphase Flow*, 19(5):829–838, Oct. 1993. doi: 10.1016/0301-9322(93)90046-w.
- M. S. Bartlett. Periodogram analysis and continuous spectra. *Biometrika*, 37(1-2):1–16, 1950. doi: 10.1093/biomet/37.1-2.1.
- J. Battaglia. Location of seismic events and eruptive fissures on the piton de la fournaise volcano using seismic amplitudes. *Journal of Geophysical Research*, 108(B8), 2003. doi: 10.1029/2002jb002193. URL <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2002JB002193>.
- J. Battaglia, J.-L. Got, and P. Okubo. Location of long-period events below Kilauea Volcano using seismic amplitudes and accurate relative relocation. *Journal of Geophysical Research: Solid Earth (1978-2012)*, 108(B12), 2003. doi: 10.1029/2003jb002517.
- C. Bean, I. Lokmer, and G. O’Brien. Influence of near-surface volcanic structure on long-period seismic, signals and on moment tensor inversions: Simulated examples from, Mount Etna. *Journal of Geophysical Research: Solid Earth*, 113(B8):B08308, Aug. 2008. ISSN 2156-2202. doi: 10.1029/2007JB005468. URL <http://onlinelibrary.wiley.com/doi/10.1029/2007JB005468/abstract>.

- 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, Jan. 2014. ISSN 1752-0894. doi: 10.1038/ngeo2027. URL <https://www.nature.com/ngeo/journal/v7/n1/full/ngeo2027.html>.
- A. F. Bell, S. Hernandez, H. E. Gaunt, P. Mothes, M. Ruiz, D. Sierra, and S. Aguaiza. The rise and fall of periodic ‘drumbeat’ seismicity at Tungurahua volcano, Ecuador. *Earth and Planetary Science Letters*, 475:58–70, Oct. 2017. ISSN 0012-821X. doi: 10.1016/j.epsl.2017.07.030. URL <http://www.sciencedirect.com/science/article/pii/S0012821X17304144>.
- A. F. Bell, N. Mark, H. Stephen, I. G. Main, H. E. Gaunt, M. Patricia, and R. Mario. Volcanic Eruption Forecasts From Accelerating Rates of Drumbeat Long-Period Earthquakes. *Geophysical Research Letters*, 45(3):1339–1348, Feb. 2018. ISSN 0094-8276. doi: <https://doi.org/10.1002/2017GL076429>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL076429>.
- J. M. Bernardo. Bayesian Statistics. In *International Encyclopedia of Statistical Science*, pages 107–133. Springer, Berlin, Heidelberg, 2011. doi: [https://doi.org/10.1007/978-3-642-04898-2\\_139](https://doi.org/10.1007/978-3-642-04898-2_139). URL [https://link.springer.com/referenceworkentry/10.1007/978-3-642-04898-2\\_139](https://link.springer.com/referenceworkentry/10.1007/978-3-642-04898-2_139).
- M. A. Biot. Propagation of elastic waves in a cylindrical bore containing a fluid. *Journal of Applied Physics*, 23(9):997–1005, sep 1952. doi: 10.1063/1.1702365.
- L. F. Bonilla, J. H. Steidl, G. T. Lindley, A. G. Tumarkin, and R. J. Archuleta. Site amplification in the San Fernando Valley, California: variability of site-effect estimation using the S-wave, coda, and H/V methods. *Bulletin of the Seismological Society of America*, 87(3):710–730, 1997.
- M. Bottiglieri, S. D. Martino, M. Falanga, C. Godano, and M. Palo. Statistics of intertime of Strombolian explosion-quakes. *Europhysics Letters (EPL)*, 72(3):493–498, Nov. 2005. doi: 10.1209/epl/i2005-10258-0.
- S. Bredemeyer and T. H. Hansteen. Synchronous degassing patterns of the neighbouring volcanoes Llaima, and Villarrica in south-central Chile: the influence of tidal forces. *International Journal of Earth Sciences*, 103(7):1999–2012, 2014. ISSN 1437-3262. doi: 10.1007/s00531-014-1029-2. URL <http://dx.doi.org/10.1007/s00531-014-1029-2>.
- F. Brenguier, N. M. Shapiro, M. Campillo, A. Nercessian, and V. Ferrazzini. 3-d surface wave tomography of the piton de la fournaise volcano using seismic noise correlations. *Geophysical Research Letters*, 34(2), jan 2007. doi: 10.1029/2006gl028586.

- A. Burgisser and W. Degruyter. Magma Ascent and Degassing at Shallow Levels. In *The Encyclopedia of Volcanoes*, pages 225–236. Elsevier, 2015. doi: 10.1016/b978-0-12-385938-9.00011-0.
- G. Cabras, R. Carniel, and J. Wassermann. BLIND SOURCE SEPARATION: AN APPLICATION TO THE MT. MERAPI VOLCANO, INDONESIA. *Fluctuation and Noise Letters*, 08(03n04):L249–L260, dec 2008. doi: 10.1142/s0219477508005124.
- E. S. Calder, A. J. L. Harris, P. Peña, E. Pilger, L. P. Flynn, G. Fuentealba, and H. Moreno. Combined thermal and seismic analysis of the Villarrica volcano lava lake, Chile. *Revista geológica de Chile*, 31(2):259–272, 2004. doi: <http://dx.doi.org/10.4067/S0716-02082004000200005>. URL <http://www.andeangeology.cl/index.php/revista1/article/view/V31n2-a05/html>.
- A. Caracausi, R. Favara, S. Giammanco, F. Italiano, A. Paonita, G. Pecoraino, A. Rizzo, and P. M. Nuccio. Mount etna: Geochemical signals of magma ascent and unusually extensive plumbing system. *Geophysical Research Letters*, 30(2), Jan. 2003. doi: 10.1029/2002gl015463.
- S. A. Carn. Chapter 65 - Gas, Plume, and Thermal Monitoring. In H. Sigurdsson, editor, *The Encyclopedia of Volcanoes (Second Edition)*, pages 1125–1149. Academic Press, Amsterdam, second edition edition, 2015. ISBN 978-0-12-385938-9. doi: <https://doi.org/10.1016/B978-0-12-385938-9.00065-1>. URL <https://www.sciencedirect.com/science/article/pii/B9780123859389000651>.
- L. Cauchie, G. Saccorotti, and C. J. Bean. Amplitude and recurrence time analysis of LP activity at Mount Etna, Italy. *Journal of Geophysical Research: Solid Earth*, 120(9):2015JB011897, Sept. 2015. ISSN 2169-9356. doi: 10.1002/2015JB011897. URL <http://onlinelibrary.wiley.com/doi/10.1002/2015JB011897/abstract>.
- S. Cesca, J. Battaglia, T. Dahm, E. Tessmer, S. Heimann, and P. Okubo. Effects of topography and crustal heterogeneities on the source estimation of LP event at Kilauea volcano. *Geophysical Journal International*, 172(3): 1219–1236, Mar. 2008. ISSN 0956-540X. doi: 10.1111/j.1365-246X.2007.03695.x. URL <https://academic.oup.com/gji/article/172/3/1219/2140417/Effects-of-topography-and-crustal-heterogeneities>.
- C. J. Chamberlain, C. J. Hopp, C. M. Boese, E. Warren-Smith, D. Chambers, S. X. Chu, K. Michailos, and J. Townend. EQcorrscan: Repeating and near-repeating earthquake detection and analysis in Python. *Seismological Research Letters*, 89(1):173–181, Dec. 2017. doi: 10.1785/0220170151.
- D. K. Chester. Volcanoes, society, and culture. In *Volcanoes and the Environment*, pages 404–439. Cambridge University Press, Oct. 2005. doi: 10.1017/cbo9780511614767.015.

- B. Chouet. Excitation of a buried magmatic pipe: A seismic source model for volcanic tremor. *Journal of Geophysical Research*, 90(B2):1881, 1985. doi: 10.1029/jb090ib02p01881.
- B. Chouet. Dynamics of a fluid-driven crack in three dimensions by the finite difference method. *Journal of Geophysical Research*, 91(B14):13967, 1986. doi: 10.1029/jb091ib14p13967.
- B. Chouet. Resonance of a fluid-driven crack: Radiation properties and implications, for the source of long-period events and harmonic tremor. *Journal of Geophysical Research: Solid Earth*, 93(B5):4375–4400, May 1988. ISSN 2156-2202. doi: 10.1029/JB093iB05p04375. URL <http://onlinelibrary.wiley.com/doi/10.1029/JB093iB05p04375/abstract>.
- B. Chouet. Long-period volcano seismicity: its source and use in eruption forecasting. *Nature*, 380(6572):309–316, 1996.
- B. Chouet. Volcano seismology. *Pure and Applied Geophysics*, 160(3-4):739–788, 2003.
- B. Chouet, G. Saccorotti, M. Martini, P. Dawson, G. De Luca, G. Milana, and R. Scarpa. Source and path effects in the wave fields of tremor and explosions at Stromboli Volcano, Italy. *Journal of Geophysical Research: Solid Earth*, 102(B7):15129–15150, July 1997. ISSN 2156-2202. doi: 10.1029/97JB00953. URL <http://onlinelibrary.wiley.com/doi/10.1029/97JB00953/abstract>.
- B. Chouet, G. De Luca, G. Milana, P. Dawson, M. Martini, and R. Scarpa. Shallow velocity structure of Stromboli volcano, Italy, derived from small-aperture array measurements of Strombolian tremor. *Bulletin of the Seismological Society of America*, 88(3):653–666, 1998. URL <https://pubs.geoscienceworld.org/ssa/bssa/article/88/3/653/102728/shallow-velocity-structure-of-stromboli-volcano>.
- B. Chouet, P. Dawson, T. Ohminato, M. Martini, G. Saccorotti, F. Giudicepietro, G. D. 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, Jan 2003. doi: 10.1029/2002jb001919.
- B. A. Chouet and R. S. Matoza. A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. *Journal of Volcanology and Geothermal Research*, 252(Supplement C):108–175, Feb. 2013. ISSN 0377-0273. doi: 10.1016/j.jvolgeores.2012.11.013. URL <http://www.sciencedirect.com/science/article/pii/S0377027312003435>.



- B. A. Chouet, P. B. Dawson, M. R. James, and S. J. Lane. Seismic source mechanism of degassing bursts at kilauea volcano, hawaii: Results from waveform inversion in the 10-50 s band. *Journal of Geophysical Research*, 115(B9), sep 2010. doi: 10.1029/2009jb006661.
- A. Ciaramella, E. D. Lauro, S. D. Martino, B. D. Lieto, M. Falanga, and R. Tagliaferri. Characterization of strombolian events by using independent component analysis. *Nonlinear Processes in Geophysics*, 11(4):453–461, oct 2004. doi: 10.5194/npg-11-453-2004.
- D. Coppola, M. Laiolo, C. Cigolini, F. Massimetti, D. D. Donne, M. Ripepe, H. Arias, S. Barsotti, C. B. Parra, R. G. Centeno, S. Cevuard, G. Chigna, C. Chun, E. Garaebiti, D. Gonzales, J. Griswold, J. Juarez, L. E. Lara, C. M. López, O. Macedo, C. Mahinda, S. Ogburn, O. Prambada, P. Ramon, D. Ramos, A. Peltier, S. Saunders, E. de Zeeuw-van Dalfsen, N. Varley, and R. William. Thermal remote sensing for global volcano monitoring: Experiences from the MIROVA system. *Frontiers in Earth Science*, 7, jan 2020. doi: 10.3389/feart.2019.00362.
- L. Costantini, L. Pioli, C. Bonadonna, J. Clavero, and C. Longchamp. A Late Holocene explosive mafic eruption of Villarrica Volcano, Southern Andes: the Chaimilla deposit. *Journal of Volcanology and Geothermal Research*, 200(3):143–158, 2011. doi: <https://doi.org/10.1016/j.jvolgeores.2010.12.010>. URL <https://www.sciencedirect.com/science/article/pii/S0377027310003938>.
- S. J. Cronin and K. V. Cashman. Volcanic oral traditions in hazard assessment and mitigation. In J. Gratton and R. Torrence, editors, *Living under the shadow: Cultural impacts of volcanic eruptions*, pages 175–202. Left Coast Press Inc. Walnut Creek, CA, USA, 2008. URL <https://research-information.bris.ac.uk/en/publications/volcanic-oral-traditions-in-hazard-assessment-and-mitigation>.
- David Karnå, CC BY 3.0. File:Eyjafjallajökull ash cloud 20100417.jpg — Wikimedia Commons, the free media repository. Creative Commons Attribution 3.0 Unported, 2020. URL [https://commons.wikimedia.org/w/index.php?title=File:Eyjafjallaj%C3%B6kull\\_ash\\_cloud\\_20100417.jpg&oldid=489431007](https://commons.wikimedia.org/w/index.php?title=File:Eyjafjallaj%C3%B6kull_ash_cloud_20100417.jpg&oldid=489431007). [Online; accessed 25-March-2021].
- L. De Barros, C. J. Bean, I. Lokmer, G. Saccorotti, L. Zuccarello, G. S. O’Brien, J.-P. Métaxian, and D. Patane. Source geometry from exceptionally high resolution long period event, observations at Mt Etna during the 2008 eruption. *Geophysical Research Letters*, 36(24), 2009.
- L. De Barros, I. Lokmer, C. J. Bean, G. S. O’Brien, G. , Saccorotti, J.-P. Métaxian, L. Zuccarello, and D. Patanè. Source mechanism of long-period events recorded by

- a high-density, seismic network during the 2008 eruption on Mount Etna. *Journal of Geophysical Research: Solid Earth*, 116(B1):B01304, Jan. 2011. ISSN 2156-2202. doi: 10.1029/2010JB007629. URL <http://onlinelibrary.wiley.com/doi/10.1029/2010JB007629/abstract>.
- E. De Lauro, S. De Martino, M. Falanga, and M. Palo. Modelling the macroscopic behavior of Strombolian explosions at Erebus volcano. *Physics of the Earth and Planetary Interiors*, 176(3-4):174–186, Oct. 2009. doi: <https://doi.org/10.1016/j.pepi.2009.05.003>.
- F. Delgado, M. E. Pritchard, S. Ebmeier, P. González, and L. Lara. Recent unrest (2002-2015) imaged by space geodesy at the highest risk Chilean volcanoes: Villarrica, Llaima, and Calbuco (Southern Andes). *Journal of Volcanology and Geothermal Research*, 344:270–288, sep 2017. doi: 10.1016/j.jvolgeores.2017.05.020.
- B. Di Lieto, G. Saccorotti, L. Zuccarello, M. L. Rocca, and R. Scarpa. Continuous tracking of volcanic tremor at Mount Etna, Italy. *Geophysical Journal International*, 169(2):699–705, May 2007. ISSN 0956-540X. doi: 10.1111/j.1365-246X.2007.03316.x. URL <https://academic.oup.com/gji/article/169/2/699/2011975>.
- L. Dominguez, L. Pioli, C. Bonadonna, C. B. Connor, D. Andronico, A. J. L. Harris, and M. Ripepe. Quantifying unsteadiness and dynamics of pulsatory volcanic activity. *Earth and Planetary Science Letters*, 444:160–168, June 2016. doi: 10.1016/j.epsl.2016.03.048.
- A. E. Dukler and J. Fabre. GAS-LIQUID SLUG FLOW:.. *Multiphase Science and Technology*, 8(1-4):355–469, 1994. doi: 10.1615/multscientechn.v8.i1-4.80.
- Y. Dzierma, W. Rabbel, M. Thorwart, I. Koulakov, H. Wehrmann, K. Hoernle, and D. Comte. Seismic velocity structure of the slab and continental plate in the region of the 1960 Valdivia (Chile) slip maximum - Insights into fluid release and plate coupling. *Earth and Planetary Science Letters*, 331-332:164 – 176, 2012. ISSN 0012-821X. doi: <http://dx.doi.org/10.1016/j.epsl.2012.02.006>. URL <http://www.sciencedirect.com/science/article/pii/S0012821X12000817>.
- S. K. Ebmeier, B. J. Andrews, M. C. Araya, D. W. D. Arnold, J. Biggs, C. Cooper, E. Cottrell, M. Furtney, J. Hickey, J. Jay, R. Lloyd, A. L. Parker, M. E. Pritchard, E. Robertson, E. Venzke, and J. L. Williamson. Synthesis of global satellite observations of magmatic and volcanic deformation: implications for volcano monitoring & the lateral extent of magmatic domains. *Journal of Applied Volcanology*, 7(1), feb 2018. doi: 10.1186/s13617-018-0071-3.
- B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), Jan. 1979. doi: 10.1214/aos/1176344552.

- B. Efron and C. Stein. The Jackknife Estimate of Variance. *The Annals of Statistics*, 9 (3):586–596, 1981. ISSN 0090-5364. doi: <https://www.jstor.org/stable/2240822>. URL <http://www.jstor.org/stable/2240822>.
- E. P. S. Eibl, C. J. Bean, K. S. Vogfjörð, Y. Ying, I. Lokmer, M. Möllhoff, G. S. O’Brien, and F. Pálsson. Tremor-rich shallow dyke formation followed by silent magma flow at Bárðharbunga in Iceland. *Nature Geoscience*, 2017.
- T. S. Eyre, C. J. Bean, L. De Barros, G. S. O’Brien, F. Martini, I. Lokmer, M. M. Mora, J. F. Pacheco, and G. J. Soto. Moment tensor inversion for the source location and mechanism of long period (LP) seismic events from 2009 at Turrialba volcano, Costa Rica. *Journal of Volcanology and Geothermal Research*, 258:215–223, May 2013. ISSN 0377-0273. doi: 10.1016/j.jvolgeores.2013.04.016. URL <http://www.sciencedirect.com/science/article/pii/S0377027313001273>.
- V. Ferrazzini and K. Aki. Slow waves trapped in a fluid-filled infinite crack: Implication for volcanic tremor. *Journal of Geophysical Research*, 92(B9):9215, 1987. doi: 10.1029/jb092ib09p09215.
- J. T. Freymueller, J. B. Murray, H. Rymer, and C. A. Locke. Ground Deformation, Gravity, and Magnetism. In *The Encyclopedia of Volcanoes*, pages 1101–1123. Elsevier, 2015. doi: 10.1016/b978-0-12-385938-9.00064-x.
- S. Gambino, G. Falzone, A. Ferro, and G. Laudani. Volcanic processes detected by tiltmeters: A review of experience on sicilian volcanoes. *Journal of Volcanology and Geothermal Research*, 271:43–54, feb 2014. doi: 10.1016/j.jvolgeores.2013.11.007.
- A. García-Yeguas, J. Ledo, P. Piña-Varas, J. Prudencio, P. Queralt, A. Marcuello, J. M. Ibañez, B. Benjumea, A. Sánchez-Alzola, and N. Pérez. A 3D joint interpretation of magnetotelluric and seismic tomographic models: The case of the volcanic island of Tenerife. *Computers & Geosciences*, 109:95–105, Dec. 2017. ISSN 0098-3004. doi: 10.1016/j.cageo.2017.08.003. URL <http://www.sciencedirect.com/science/article/pii/S0098300416308512>.
- D. Gaudin, J. Taddeucci, P. Scarlato, E. del Bello, T. Ricci, T. Orr, B. Houghton, A. Harris, S. Rao, and A. Bucci. Integrating puffing and explosions in a general scheme for Strombolian-style activity. *Journal of Geophysical Research: Solid Earth*, 2017a. doi: 10.1002/2016jb013707.
- D. Gaudin, J. Taddeucci, P. Scarlato, A. Harris, M. Bombrun, E. D. Bello, and T. Ricci. Characteristics of puffing activity revealed by ground-based, thermal infrared imaging: the example of Stromboli Volcano (Italy). *Bulletin of Volcanology*, 79(3), Feb. 2017b. doi: 10.1007/s00445-017-1108-x.

- H. Gercek. Poisson's ratio values for rocks. *International Journal of Rock Mechanics and Mining Sciences*, 44(1):1–13, Jan. 2007. ISSN 1365-1609. doi: 10.1016/j.ijrmms.2006.04.011. URL <http://www.sciencedirect.com/science/article/pii/S136516090600075X>.
- Global Volcanism Program. Villarrica. In E. Venzke, editor, *Volcanoes of the World, v. 4.4.1*. Smithsonian Institution, <http://dx.doi.org/10.5479/si.GVP.VOTW4-2013>, 2013. URL <http://dx.doi.org/10.5479/si.GVP.VOTW4-2013>. Downloaded 20 Nov 2015.
- Global Volcanism Program. Report on Villarrica (Chile). In S. K. Sennert, editor, *Weekly Volcanic Activity Report, 25 February-3 March 2015*. Smithsonian Institution and US Geological Survey, 2015.
- A. Goto and J. B. Johnson. Monotonic infrasound and Helmholtz resonance at Volcan Villarrica (Chile). *Geophysical Research Letters*, 38(6):L06301, Mar. 2011. ISSN 1944-8007. doi: 10.1029/2011GL046858. URL <http://onlinelibrary.wiley.com/doi/10.1029/2011GL046858/abstract>.
- G. D. Grazia, S. Falsaperla, and H. Langer. Volcanic tremor location during the 2004 Mount Etna lava effusion. *Geophysical Research Letters*, 33(4), 2006. doi: 10.1029/2005gl025177.
- D. N. Green and J. Neuberg. Waveform classification of volcanic low-frequency earthquake swarms, and its implication at Soufrière Hills Volcano, Montserrat. *Journal of Volcanology and Geothermal Research*, 153(1-2):51–63, May 2006. ISSN 0377-0273. doi: 10.1016/j.jvolgeores.2005.08.003. URL <http://www.sciencedirect.com/science/article/pii/S0377027305003835>.
- C. E. Gregg, B. Houghton, and J. W. Ewert. Volcano Warning Systems. In *The Encyclopedia of Volcanoes*, pages 1173–1185. Elsevier, 2015. doi: 10.1016/b978-0-12-385938-9.00067-5.
- L. Gurioli, A. J. L. Harris, B. F. Houghton, M. Polacci, and M. Ripepe. Textural and geophysical characterization of explosive basaltic activity at Villarrica volcano. *Journal of Geophysical Research: Solid Earth*, 113(B8):B08206, Aug. 2008. ISSN 2156-2202. doi: 10.1029/2007JB005328. URL <http://onlinelibrary.wiley.com/doi/10.1029/2007JB005328/abstract>.
- B. Gutenberg and C. F. Richter. Earthquake magnitude, intensity, energy, and acceleration: (Second paper). *Bulletin of the Seismological Society of America*, 46(2):105–145, Apr. 1956. ISSN 0037-1106. URL <https://pubs.geoscienceworld.org/ssa/bssa/article/46/2/105/115777>.

- A. Harris and M. Ripepe. Temperature and dynamics of degassing at Stromboli. *Journal of Geophysical Research: Solid Earth*, 112(B3):B03205, Mar. 2007. ISSN 2156-2202. doi: 10.1029/2006JB004393. URL <http://onlinelibrary.wiley.com/doi/10.1029/2006JB004393/abstract>.
- J. Havskov and G. Alguacil. *Instrumentation in earthquake seismology*, volume 358. Springer, 2004.
- J. Havskov and L. Ottemöller. Instruments and Waveform Data. In *Routine Data Processing in Earthquake Seismology*, pages 53–81. Springer Netherlands, 2010. doi: 10.1007/978-90-481-8697-6\_3.
- Hawaii Volcanoes National Park, Public Domain. File:Kīlauea volcano eruption 20201220.jpg — Wikimedia Commons, the free media repository. Public Domain, 2021. URL [https://commons.wikimedia.org/w/index.php?title=File:K%C4%ABlauea\\_volcano\\_eruption\\_20201220.jpg&oldid=540357121](https://commons.wikimedia.org/w/index.php?title=File:K%C4%ABlauea_volcano_eruption_20201220.jpg&oldid=540357121). [Online; accessed 25-March-2021].
- H.-S. Hilbert. Die Übertragungsfunktion von Geophonen: Eigenschaften und Restituierbarkeit. Bachelor’s thesis, Kiel University, 2015.
- R. J. Hyndman. Computing and Graphing Highest Density Regions. *The American Statistician*, 50(2):120–126, 1996. ISSN 0003-1305. doi: 10.2307/2684423. URL <http://www.jstor.org/stable/2684423>.
- K. Ishii, A. Yokoo, T. Kagiya, T. Ohkura, S. Yoshikawa, and H. Inoue. Gas flow dynamics in the conduit of Strombolian explosions inferred from seismo-acoustic observations at Aso volcano, Japan. *Earth, Planets and Space*, 71(1), Feb. 2019. doi: 10.1186/s40623-019-0992-z.
- M. R. James, S. J. Lane, B. Chouet, and J. S. Gilbert. Pressure changes associated with the ascent and bursting of gas slugs in liquid-filled vertical and inclined conduits. *Journal of Volcanology and Geothermal Research*, 129(1-3):61–82, Jan 2004. doi: 10.1016/s0377-0273(03)00232-4.
- M. R. James, S. J. Lane, and B. A. Chouet. Gas slug ascent through changes in conduit diameter: Laboratory insights into a volcano-seismic source process in low-viscosity magmas. *Journal of Geophysical Research: Solid Earth*, 111(B5):n/a–n/a, May 2006. doi: 10.1029/2005jb003718.
- M. R. James, S. J. Lane, and S. B. Corder. Modelling the rapid near-surface expansion of gas slugs in low-viscosity magmas. *Geological Society, London, Special Publications*, 307(1):147–167, 2008. doi: 10.1144/sp307.9.

- A. Jarvis, N. A. Reuter, H.I., and E. Guevara. Hole-filled SRTM for the globe Version 4. SRTM 90m Database, International Centre for Tropical Agriculture (CIAT), 2008. URL <http://srtm.csi.cgiar.org>. available from <http://srtm.csi.cgiar.org/>.
- C. Jaupart and S. Vergnolle. Laboratory models of Hawaiian and Strombolian eruptions. *Nature*, 331(6151):58–60, Jan. 1988. doi: 10.1038/331058a0.
- J. B. Johnson, L. M. Watson, J. L. Palma, E. M. Dunham, and J. F. Anderson. Forecasting the Eruption of an Open-Vent Volcano Using Resonant Infrasound Tones. *Geophysical Research Letters*, 45(5):2213–2220, Mar. 2018. doi: 10.1002/2017gl076506. URL <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017GL076506>.
- P. Jousset, J. Neuberg, and S. Sturton. Modelling the time-dependent frequency content of low-frequency volcanic, earthquakes. *Journal of Volcanology and Geothermal Research*, 128(1-3):201–223, #nov# 2003. ISSN 0377-0273. doi: 10.1016/S0377-0273(03)00255-5. URL <http://www.sciencedirect.com/science/article/pii/S0377027303002555>.
- P. Jousset, J. Neuberg, and A. Jolly. Modelling low-frequency volcanic earthquakes in a viscoelastic medium with topography. *Geophysical Journal International*, 159(2):776–802, #nov# 2004. ISSN 0956-540X. doi: 10.1111/j.1365-246X.2004.02411.x. URL <https://academic.oup.com/gji/article/159/2/776/2067222/Modelling-low-frequency-volcanic-earthquakes-in-a>.
- H. Kawakatsu and M. Yamamoto. Volcano seismology. *Treatise on geophysics*, 4:389–420, 2007.
- H. Kawakatsu, S. Kaneshima, H. Matsubayashi, T. Ohminato, Y. Sudo, T. Tsutsui, K. Uhira, H. Yamasato, H. Ito, and D. Legrand. Aso94: Aso seismic observation with broadband instruments. *Journal of Volcanology and Geothermal Research*, 101(1):129–154, 2000.
- K. I. Konstantinou and V. Schlindwein. Nature, wavefield properties and source mechanism of volcanic tremor: a review. *Journal of Volcanology and Geothermal Research*, 119(1):161–187, #jan# 2003. ISSN 0377-0273. doi: 10.1016/S0377-0273(02)00311-6. URL <http://www.sciencedirect.com/science/article/pii/S0377027302003116>.
- J. K. Kruschke. Chapter 4 - What is This Stuff Called Probability? In *Doing Bayesian Data Analysis (Second Edition)*, pages 71–97. Academic Press, Boston, 2015a. ISBN 978-0-12-405888-0. doi: <https://doi.org/10.1016/B978-0-12-405888-0.00004-0>. URL <https://www.sciencedirect.com/science/article/pii/B9780124058880000040>.
- J. K. Kruschke. Chapter 25 - Tools in the Trunk. In *Doing Bayesian Data Analysis (Second Edition)*, pages 721–736. Academic Press, Boston, 2015b. ISBN 978-0-12-

- 405888-0. doi: <https://doi.org/10.1016/B978-0-12-405888-0.00025-8>. URL <https://www.sciencedirect.com/science/article/pii/B9780124058880000258>.
- H. Kumagai, M. Nakano, T. Maeda, H. Yepes, P. Palacios, M. Ruiz, S. Arrais, M. Vaca, I. Molina, and T. Yamashima. Broadband seismic monitoring of active volcanoes using deterministic and stochastic approaches. *Journal of Geophysical Research: Solid Earth*, 115(B8), Aug. 2010. ISSN 0148-0227. doi: 10.1029/2009JB006889. URL <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2009JB006889>.
- H. Kumagai, T. Saito, G. O'Brien, and T. Yamashina. Characterization of scattered seismic wavefields simulated in heterogeneous media with topography. *Journal of Geophysical Research: Solid Earth*, 116(B3), Mar. 2011. ISSN 0148-0227. doi: 10.1029/2010JB007718. URL <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2010JB007718>.
- 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):137–151, Aug. 1994. ISSN 0377-0273. doi: 10.1016/0377-0273(94)90031-0.
- O. Lamb, J. Lees, L. F. Marin, J. Lazo, A. Rivera, M. Shore, and S. Lee. Investigating potential icequakes at Llaima volcano, Chile. *Volcanica*, 3(1):29–42, 2020. doi: 10.30909/vol.03.01.2942.
- J. Lehr, F. Eckel, M. Thorwart, and W. Rabbel. Low-frequency seismicity at villarrica volcano: Source location and seismic velocities. *Journal of Geophysical Research: Solid Earth*, 124(11):11505–11530, Nov. 2019. doi: 10.1029/2018jb017023.
- E. J. Liu, K. Wood, E. Mason, M. Edmonds, A. Aiuppa, G. Giudice, M. Bitetto, V. Francofonte, S. Burrow, T. Richardson, M. Watson, T. D. Pering, T. C. Wilkes, A. J. S. McGonigle, G. Velasquez, C. Melgarejo, and C. Bucarey. Dynamics of outgassing and plume transport revealed by proximal unmanned aerial system (UAS) measurements at Volcán Villarrica, Chile. *Geochemistry, Geophysics, Geosystems*, 20(2):730–750, Feb. 2019. doi: 10.1029/2018gc007692.
- T. J. Liu and S. G. Bankoff. Structure of air-water bubbly flow in a vertical pipe—II. void fraction, bubble velocity and bubble size distribution. *International Journal of Heat and Mass Transfer*, 36(4):1061–1072, Mar. 1993. doi: 10.1016/s0017-9310(05)80290-x.
- S. Lohmar, C. Robin, A. Gourgaud, J. Clavero, M. A. Parada, H. Moreno, O. Ersoy, L. López-Escobar, and J. A. Naranjo. Evidence of magma-water interaction during the 13,800 years BP, explosive cycle of the Licán Ignimbrite, Villarrica Volcano, (Southern Chile). *Revista geológica de Chile*, 34:233 – 247, 2007.

- ISSN 0716-0208. URL [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0716-02082007000200004&nrm=iso](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0716-02082007000200004&nrm=iso).
- S. Lohmar, M. Parada, F. Gutiérrez, C. Robin, and M. C. Gerbe. Mineralogical and numerical approaches to establish the pre-eruptive conditions of the mafic licán ignimbrite, villarrica volcano (chilean southern andes). *Journal of Volcanology and Geothermal Research*, 235-236:55 – 69, 2012. ISSN 0377-0273. doi: <https://doi.org/10.1016/j.jvolgeores.2012.05.006>. URL <http://www.sciencedirect.com/science/article/pii/S0377027312001448>.
- K. V. Mardia and P. E. Jupp. Basic Concepts and Models. In *Directional Statistics*, pages 25–56. John Wiley & Sons, Inc., 1999. ISBN 978-0-470-31697-9. URL <http://onlinelibrary.wiley.com/doi/10.1002/9780470316979.ch3/summary>. DOI: 10.1002/9780470316979.ch3.
- S. D. Martino, M. Palo, and G. B. Cimini. A statistical study of the stromboli volcano explosion quakes before and during 2002-2003 eruptive crisis. *Journal of Geophysical Research*, 116(B4), Apr. 2011. doi: 10.1029/2010jb008047.
- S. D. Martino, A. Errico, M. Palo, and G. B. Cimini. Explosion swarms at Stromboli volcano: a proxy for nonequilibrium conditions in the shallow plumbing system. *Geochemistry, Geophysics, Geosystems*, 13(3):n/a–n/a, Mar. 2012. doi: 10.1029/2011gc003949.
- D. Massonnet and F. Sigmundsson. Remote sensing of volcano deformation by radar interferometry from various satellites. In *Remote Sensing of Active Volcanism*, pages 207–221. American Geophysical Union, 2000. doi: 10.1029/gm116p0207.
- T. A. Mather, V. I. Tsanev, D. M. Pyle, A. J. S. McGonigle, C. Oppenheimer, and A. G. Allen. Characterization and evolution of tropospheric plumes from Lascar and Villarrica volcanoes, Chile. *Journal of Geophysical Research: Atmospheres*, 109(D21), Nov. 2004. doi: 10.1029/2004jd004934.
- R. S. Matoza and B. A. Chouet. Subevents of long-period seismicity: Implications for hydrothermal dynamics during the 2004–2008 eruption of Mount St. Helens. *Journal of Geophysical Research: Solid Earth*, 115(B12):B12206, Dec. 2010. ISSN 2156-2202. doi: 10.1029/2010JB007839. URL <http://onlinelibrary.wiley.com/doi/10.1029/2010JB007839/abstract>.
- D. E. McNamara and R. P. Buland. Ambient Noise Levels in the Continental United States. *Bulletin of the Seismological Society of America*, 94(4):1517–1527, Aug. 2004. ISSN 0037-1106. doi: 10.1785/012003001. URL <https://pubs.geoscienceworld.org/ssa/bssa/article/94/4/1517/121021/ambient-noise-levels-in-the-continental-united>.



- S. R. McNutt and D. C. Roman. Volcanic Seismicity. In *The Encyclopedia of Volcanoes*, pages 1011–1034. Elsevier, 2015. doi: 10.1016/b978-0-12-385938-9.00059-6.
- S. R. McNutt, G. Thompson, J. Johnson, S. D. Angelis, and D. Fee. Chapter 63 - Seismic and Infrasonic Monitoring. In H. Sigurdsson, editor, *The Encyclopedia of Volcanoes (Second Edition)*, pages 1071–1099. Academic Press, Amsterdam, second edition edition, 2015. ISBN 978-0-12-385938-9. doi: <https://doi.org/10.1016/B978-0-12-385938-9.00063-8>. URL <https://www.sciencedirect.com/science/article/pii/B9780123859389000638>.
- J.-P. Métaixian, P. Lesage, and J. Dorel. Permanent tremor of Masaya Volcano, Nicaragua: Wave field analysis and source location. *Journal of Geophysical Research: Solid Earth*, 102(B10):22529–22545, Oct. 1997. ISSN 2156-2202. doi: 10.1029/97JB01141. URL <http://onlinelibrary.wiley.com/doi/10.1029/97JB01141/abstract>.
- J.-P. Métaixian, P. Lesage, and B. Valette. Locating sources of volcanic tremor and emergent events by seismic triangulation: Application to Arenal volcano, Costa Rica. *Journal of Geophysical Research: Solid Earth*, 107(B10):2243, Oct. 2002. ISSN 2156-2202. doi: 10.1029/2001JB000559. URL <http://onlinelibrary.wiley.com/doi/10.1029/2001JB000559/abstract>.
- A. Mignan and J. Woessner. Estimating the magnitude of completeness for earthquake catalogs. *Community Online Resource for Statistical Seismicity Analysis*, 2012. doi: 10.5078/CORSSA-00180805.
- C. Mora-Stock. *Seismic Structure and Seismicity of the Villarrica Volcano, (Southern Central Chile)*. Phd Thesis, Christian-Albrechts-Universität zu Kiel, 2015. URL [http://macau.uni-kiel.de/receive/dissertation\\_diss\\_00017649](http://macau.uni-kiel.de/receive/dissertation_diss_00017649).
- C. Mora-Stock, M. Thorwart, T. Wunderlich, S. Bredemeyer, T. H. Hansteen, and W. Rabbel. Comparison of seismic activity for Llaima and Villarrica Volcanoes, prior to and after the Maule 2010 earthquake. *International Journal of Earth Sciences*, 103(7):2015–2028, 2014.
- H. Moreno and J. Clavero. Mapa geológico del área del volcán Villarrica, Regiones de la, Araucanía y de los Lagos (escala 1:50.000). Carta Geológica de Chile, Série Geología Básica, No., 98, 35, Servicio Nacional de Geología y Minería, 2006. URL <http://www.sernageomin.cl/volcan-villarrica/>.
- H. Morioka, H. Kumagai, and T. Maeda. Theoretical basis of the amplitude source location method for volcano-seismic signals. *Journal of Geophysical Research: Solid Earth*, 122(8):6538–6551, Aug. 2017. ISSN 2169-9313. doi: 10.1002/2017JB013997. URL <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017JB013997>.

- Y. Moussallam, P. Bani, A. Curtis, T. Barnie, M. Moussallam, N. Peters, C. I. Schipper, A. Aiuppa, G. Giudice, Á. Amigo, G. Velasquez, and C. Cardona. Sustaining persistent lava lakes: Observations from high-resolution gas measurements at Villarrica volcano, Chile. *Earth and Planetary Science Letters*, 454:237–247, Nov. 2016. doi: 10.1016/j.epsl.2016.09.012.
- R. F. Mudde. GRAVITY-DRIVEN BUBBLY FLOWS. *Annual Review of Fluid Mechanics*, 37(1):393–423, Jan. 2005. doi: 10.1146/annurev.fluid.37.061903.175803.
- P. A. Nadeau, J. L. Palma, and G. P. Waite. Linking volcanic tremor, degassing, and eruption dynamics via SO<sub>2</sub> imaging. *Geophysical Research Letters*, 38(1):n/a–n/a, jan 2011. doi: 10.1029/2010gl045820.
- NASA Earth Observatory. Eruption of Villarrica Volcano - OLI (Landsat-8) on February 22, 2015. Image of the Day for March 12, 2015, Mar. 2015. URL <https://earthobservatory.nasa.gov/images/85465/eruption-of-villarrica-volcano>. [Online; accessed 16-April-2021].
- J. Neuberg. Characteristics and causes of shallow seismicity in andesite volcanoes. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 358(1770):1533–1546, #may# 2000. ISSN 1364-503X, 1471-2962. doi: 10.1098/rsta.2000.0602. URL <http://rsta.royalsocietypublishing.org/content/358/1770/1533>.
- J. Neuberg and C. O’Gorman. A model of the seismic wavefield in gas-charged magma: application to Soufrière Hills Volcano, Montserrat. *Geological Society, London, Memoirs*, 21(1):603–609, 2002. doi: 10.1144/gsl.mem.2002.021.01.29.
- J. Neuberg and T. Pointer. Effects of volcano topography on seismic broad-band waveforms. *Geophysical Journal International*, 143(1):239–248, 2000. doi: 10.1046/j.1365-246x.2000.00251.x.
- J. Neuberg, R. Luckett, M. Ripepe, and T. Braun. Highlights from a seismic broadband array on Stromboli Volcano. *Geophysical Research Letters*, 21(9):749–752, 1994. ISSN 1944-8007. doi: 10.1029/94GL00377. URL <http://dx.doi.org/10.1029/94GL00377>.
- 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, #aug# 2000. ISSN 0377-0273. doi: 10.1016/S0377-0273(00)00169-4. URL <http://www.sciencedirect.com/science/article/pii/S0377027300001694>.
- T. Nishimura, M. Iguchi, M. Hendrasto, H. Aoyama, T. Yamada, M. Ripepe, and R. Genco. Magnitude–frequency distribution of volcanic explosion earthquakes. *Earth, Planets and Space*, 68(1), July 2016. doi: 10.1186/s40623-016-0505-2.

- G. S. O'Brien and C. J. Bean. Seismicity on volcanoes generated by gas slug ascent. *Geophysical Research Letters*, 35(16), aug 2008. doi: 10.1029/2008gl035001.
- G. S. O'Brien and C. J. Bean. Volcano topography, structure and intrinsic attenuation: Their, relative influences on a simulated 3D visco-elastic wavefield. *Journal of Volcanology and Geothermal Research*, 183(1):122–136, 2009.
- Official U.S. Navy Page. File:Mount Etna 2021 eruption.jpg — Wikimedia Commons, the free media repository. CC BY 2.0, 2021. URL [https://commons.wikimedia.org/w/index.php?title=File:Mount\\_Etna\\_2021\\_eruption.jpg&oldid=540381716](https://commons.wikimedia.org/w/index.php?title=File:Mount_Etna_2021_eruption.jpg&oldid=540381716). [Online; accessed 25-March-2021].
- 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, Oct. 1998. doi: 10.1029/98jb01122.
- A. Ohnuki and H. Akimoto. Experimental study on transition of flow pattern and phase distribution in upward air-water two-phase flow along a large vertical pipe. *International Journal of Multiphase Flow*, 26(3):367–386, Mar. 2000. doi: 10.1016/S0301-9322(99)00024-5.
- K. Orfanogiannaki, M. Naylor, and D. Harte. Exploratory data analysis: magnitude, space and time. *Community Online Resource for Statistical Seismicity Analysis*, 2010. doi: 10.5078/CORSSA-92330203.
- R. Ortiz, H. Moreno, A. García, G. Fuentealba, M. Astiz, P. Peña, N. Sánchez, and M. Tárraga. Villarrica volcano (Chile): characteristics of the volcanic tremor and forecasting of small explosions by means of a material failure method. *Journal of Volcanology and Geothermal Research*, 128(1):247–259, 2003. doi: [https://doi.org/10.1016/S0377-0273\(03\)00258-0](https://doi.org/10.1016/S0377-0273(03)00258-0). URL <https://www.sciencedirect.com/science/article/pii/S0377027303002580>.
- J. Pallister and S. R. McNutt. Synthesis of volcano monitoring. In *The Encyclopedia of Volcanoes*, pages 1151–1171. Elsevier, 2015. doi: 10.1016/B978-0-12-385938-9.00066-3. URL <https://www.sciencedirect.com/science/article/pii/B9780123859389000663>.
- J. L. Palma, E. S. Calder, D. Basualto, S. Blake, and D. A. Rothery. Correlations between SO<sub>2</sub> flux, seismicity, and outgassing activity at the open vent of Villarrica volcano, Chile. *Journal of Geophysical Research: Solid Earth*, 113(B10):B10201, Oct. 2008. ISSN 2156-2202. doi: 10.1029/2008JB005577. URL <http://onlinelibrary.wiley.com/doi/10.1029/2008JB005577/abstract>.

- J. L. Palma, S. Blake, and E. S. Calder. Constraints on the rates of degassing and convection in basaltic open-vent volcanoes. *Geochemistry, Geophysics, Geosystems*, 12(11):n/a–n/a, Nov. 2011. doi: 10.1029/2011gc003715.
- C. S. Parejas, T. H. Druitt, C. Robin, H. Moreno, and J.-A. , Naranjo. The Holocene Pucón eruption of Volcán Villarrica, Chile: deposit architecture and eruption chronology. *Bulletin of Volcanology*, 72(6):677–692, 2010. doi: <https://doi.org/10.1007/s00445-010-0348-9>. URL <https://link.springer.com/article/10.1007/s00445-010-0348-9>.
- E. A. Parfitt. A discussion of the mechanisms of explosive basaltic eruptions. *Journal of Volcanology and Geothermal Research*, 134(1):77–107, June 2004. ISSN 0377-0273. doi: 10.1016/j.jvolgeores.2004.01.002. URL <http://www.sciencedirect.com/science/article/pii/S0377027304000137>.
- E. A. Parfitt and L. Wilson. Explosive volcanic eruptions-IX. The transition between Hawaiian-style lava fountaining and Strombolian explosive activity. *Geophysical Journal International*, 121(1):226–232, Apr. 1995. doi: 10.1111/j.1365-246x.1995.tb03523.x.
- D. Patanè, G. Di Grazia, A. Cannata, P. Montalto, and E. , Boschi. Shallow magma pathway geometry at Mt. Etna volcano. *Geochemistry, Geophysics, Geosystems*, 9(12):n/a–n/a, 2008. ISSN 1525-2027. doi: 10.1029/2008GC002131. URL <http://dx.doi.org/10.1029/2008GC002131>. Q12021.
- G. L. Pavlis and F. L. Vernon. Calibration of seismometers using ground noise. *Bulletin of the seismological society of America*, 84(4):1243–1255, 1994. URL <https://pubs.geoscienceworld.org/ssa/bssa/article/84/4/1243/119878/calibration-of-seismometers-using-ground-noise>.
- Pering, Ilanko, and Liu. Periodicity in Volcanic Gas Plumes: A Review and Analysis. *Geosciences*, 9(9):394, Sept. 2019a. doi: 10.3390/geosciences9090394.
- T. Pering and A. McGonigle. Combining spherical-cap and Taylor bubble fluid dynamics with plume measurements to characterize basaltic degassing. *Geosciences*, 8(2):42, Jan. 2018. doi: 10.3390/geosciences8020042.
- T. D. Pering, G. Tamburello, A. J. S. McGonigle, A. Aiuppa, M. R. James, S. J. Lane, M. Sciotto, A. Cannata, and D. Patanè. Dynamics of mild strombolian activity on Mt. Etna. *Journal of Volcanology and Geothermal Research*, 300:103–111, July 2015. doi: 10.1016/j.jvolgeores.2014.12.013.
- T. D. Pering, A. J. S. McGonigle, M. R. James, A. Capponi, S. J. Lane, G. Tamburello, and A. Aiuppa. The dynamics of slug trains in volcanic conduits: Evidence for expan-

- sion driven slug coalescence. *Journal of Volcanology and Geothermal Research*, 348: 26–35, Dec. 2017. doi: 10.1016/j.jvolgeores.2017.10.009.
- T. D. Pering, T. Ilanko, T. C. Wilkes, R. A. England, S. R. Silcock, L. R. Stanger, J. R. Willmott, R. G. Bryant, and A. J. S. McGonigle. A rapidly convecting lava lake at Masaya Volcano, Nicaragua. *Frontiers in Earth Science*, 6, feb 2019b. doi: 10.3389/feart.2018.00241.
- E. D. Pezzo, F. Bianco, and I. Borgna. Magnitude scale for LP events: a quantification scheme for volcanic quakes. *Geophysical Journal International*, 194(2):911–919, apr 2013. doi: 10.1093/gji/ggt126.
- L. Pioli, C. Bonadonna, B. J. Azzopardi, J. C. Phillips, and M. Ripepe. Experimental constraints on the outgassing dynamics of basaltic magmas. *Journal of Geophysical Research: Solid Earth*, 117(B3), Mar. 2012. doi: 10.1029/2011jb008392.
- M. E. Pritchard, J. Biggs, C. Wauthier, E. Sansosti, D. W. D. Arnold, F. Delgado, S. K. Ebmeier, S. T. Henderson, K. Stephens, C. Cooper, K. Wnuk, F. Amelung, V. Aguilar, P. Mothes, O. Macedo, L. E. Lara, M. P. Poland, and S. Zoffoli. Towards coordinated regional multi-satellite InSAR volcano observations: results from the Latin America pilot project. *Journal of Applied Volcanology*, 7(1), jun 2018. doi: 10.1186/s13617-018-0074-0.
- J. Prudencio, E. Del Pezzo, A. Garca-Yeguas, and J. M. Ibanez. Spatial distribution of intrinsic and scattering seismic attenuation in active volcanic islands - I: model and the case of Tenerife Island. *Geophysical Journal International*, 195(3):1942–1956, 2013. ISSN 0956-540X. doi: 10.1093/gji/ggt361. URL <https://academic.oup.com/gji/article/195/3/1942/626933/Spatial-distribution-of-intrinsic-and-scattering>.
- W. Rabbel and M. Thorwart. Villarrica Tomography (VITO). Dataset, GFZ Data Services, 2019.
- J. P. Richardson and G. P. Waite. Waveform inversion of shallow repetitive long period events at Villarrica, Volcano, Chile. *Journal of Geophysical Research: Solid Earth*, 118(9):4922–4936, 2013. doi: <https://doi.org/10.1002/jgrb.50354>. URL <http://onlinelibrary.wiley.com/doi/10.1002/jgrb.50354/full>.
- J. P. Richardson, G. P. Waite, and J. L. Palma. Varying seismic-acoustic properties of the fluctuating lava lake, at Villarrica Volcano, Chile. *Journal of Geophysical Research: Solid Earth*, 119(7):5560–5573, 2014. doi: <https://doi.org/10.1002/2014JB011002>. URL <https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2014JB011002>.

- F. Ringdal. On the estimation of seismic detection thresholds. *Bulletin of the Seismological Society of America*, 65(6):1631–1642, Dec. 1975. ISSN 0037-1106. URL <https://pubs.geoscienceworld.org/ssa/bssa/article-abstract/65/6/1631/117495/On-the-estimation-of-seismic-detection-thresholds?redirectedFrom=fulltext>.
- M. Ripepe and E. Gordeev. Gas bubble dynamics model for shallow volcanic tremor at Stromboli. *Journal of Geophysical Research: Solid Earth*, 104(B5):10639–10654, May 1999. ISSN 2156-2202. doi: 10.1029/98JB02734. URL <http://onlinelibrary.wiley.com/doi/10.1029/98JB02734/abstract>.
- M. Ripepe and E. Marchetti. Array tracking of infrasonic sources at Stromboli volcano. *Geophysical Research Letters*, 29(22):33–1, Nov. 2002. ISSN 0094-8276. doi: 10.1029/2002GL015452. URL <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2002GL015452>.
- M. Ripepe, S. Ciliberto, and M. D. Schiava. Time constraints for modeling source dynamics of volcanic explosions at Stromboli. *Journal of Geophysical Research: Solid Earth*, 106(B5):8713–8727, May 2001. doi: 10.1029/2000jb900374.
- M. Ripepe, A. J. L. Harris, and R. Carniel. Thermal, seismic and infrasonic evidences of variable degassing rates at Stromboli volcano. *Journal of Volcanology and Geothermal Research*, 118(3-4):285–297, Nov. 2002. doi: 10.1016/S0377-0273(02)00298-6. URL <https://www.sciencedirect.com/science/article/pii/S0377027302002986>.
- M. Ripepe, E. Marchetti, C. Bonadonna, A. J. L. Harris, L. Pioli, and G. Ulivieri. Monochromatic infrasonic tremor driven by persistent degassing and convection at Villarrica Volcano, Chile. *Geophysical Research Letters*, 37(15):L15303, Aug. 2010. ISSN 1944-8007. doi: 10.1029/2010GL043516. URL <http://onlinelibrary.wiley.com/doi/10.1029/2010GL043516/abstract>.
- G. Saccorotti, B. Chouet, and P. Dawson. Wavefield properties of a shallow long-period event and tremor at Kilauea Volcano, Hawaii. *Journal of volcanology and geothermal research*, 109(1):163–189, 2001a. doi: 10.1016/S0377-0273(00)00310-3. URL <http://www.sciencedirect.com/science/article/pii/S0377027300003103>.
- G. Saccorotti, R. Maresca, and E. D. Pezzo. Array analyses of seismic noise at Mt. Vesuvius Volcano, Italy. *Journal of Volcanology and Geothermal Research*, 110(1-2): 79–100, Sept. 2001b. doi: 10.1016/S0377-0273(01)00204-9.
- H.-U. Schmincke. *Vulkanismus: mit 307 Farbabbildungen*. Wissenschaftliche Buchgesellschaft, 2000.

- J. Schweitzer, J. Fyen, S. Mykkeltveit, and T. Kværna. Seismic arrays. In P. Bormann, editor, *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, pages 1–80. Deutsches GeoForschungsZentrum GFZ, Potsdam, 2012. doi: [http://doi.org/10.2312/GFZ.NMSOP-2\\\_ch9](http://doi.org/10.2312/GFZ.NMSOP-2\_ch9). URL <http://gfzpublic.gfz-potsdam.de/pubman/item/escidoc:43213>.
- R. Seyfried and A. Freundt. Experiments on conduit flow and eruption behavior of basaltic volcanic eruptions. *Journal of Geophysical Research: Solid Earth*, 105(B10): 23727–23740, Oct. 2000. doi: 10.1029/2000jb900096.
- P. Sheets. Volcanoes, Ancient People, and Their Societies. In *The Encyclopedia of Volcanoes*, pages 1313–1319. Elsevier, 2015. doi: 10.1016/b978-0-12-385938-9.00076-6.
- H. Shinohara. Excess degassing from volcanoes and its role on eruptive and intrusive activity. *Reviews of Geophysics*, 46(4):RG4005, #dec# 2008. ISSN 1944-9208. doi: 10.1029/2007RG000244. URL <http://onlinelibrary.wiley.com/doi/10.1029/2007RG000244/abstract>.
- F. Sigmundsson, A. Hooper, S. Hreinsdóttir, K. S. Vogfjörð, B. G. Ófeigsson, E. R. Heimisson, S. Dumont, M. Parks, K. Spaans, G. B. Gudmundsson, V. Drouin, T. Árnadóttir, K. Jónsdóttir, M. T. Gudmundsson, T. Högnadóttir, H. M. Fridriksdóttir, M. Hensch, P. Einarsson, E. Magnússon, S. Samsonov, B. Brandsdóttir, R. S. White, T. Ágústsdóttir, T. Greenfield, R. G. Green, Á. R. Hjartardóttir, R. Pedersen, R. A. Bennett, H. Geirsson, P. C. L. Femina, H. Björnsson, F. Pálsson, E. Sturkell, C. J. Bean, M. Möllhoff, A. K. Braidon, and E. P. S. Eibl. Segmented lateral dyke growth in a rifting event at Bárðarbunga volcanic system, Iceland. *Nature*, 517(7533):191–195, dec 2014. doi: 10.1038/nature14111.
- F. Sigmundsson, M. Parks, R. Pedersen, K. Jónsdóttir, B. G. Ófeigsson, R. Grapenthin, S. Dumont, P. Einarsson, V. Drouin, E. R. Heimisson, Á. R. Hjartardóttir, M. Gudmundsson, H. Geirsson, S. Hreinsdóttir, E. Sturkell, A. Hooper, Þ. Högnadóttir, K. Vogfjörð, T. Barnie, and M. J. Roberts. Magma Movements in Volcanic Plumbing Systems and their Associated Ground Deformation and Seismic Patterns. In *Volcanic and Igneous Plumbing Systems*, pages 285–322. Elsevier, 2018. doi: 10.1016/b978-0-12-809749-6.00011-x.
- W. H. F. Smith and P. Wessel. Gridding with continuous curvature splines in tension. *GEOPHYSICS*, 55(3):293–305, Mar. 1990. ISSN 0016-8033, 1942-2156. doi: 10.1190/1.1442837. URL <http://library.seg.org/doi/10.1190/1.1442837>.
- R. S. J. Sparks. The dynamics of bubble formation and growth in magmas: A review and analysis. *Journal of Volcanology and Geothermal Research*, 3(1-2):1–37, mar 1978. doi: 10.1016/0377-0273(78)90002-1.

- L. Spina, A. Cannata, D. Morgavi, and D. Perugini. Degassing behaviour at basaltic volcanoes: New insights from experimental investigations of different conduit geometry and magma viscosity. *Earth-Science Reviews*, 192:317–336, May 2019. doi: 10.1016/j.earscirev.2019.03.010.
- S. Stein and M. Wysession. *An introduction to seismology, earthquakes, and earth structure*. John Wiley & Sons, 2009.
- W. K. Stovall, C. L. Driedger, E. G. Westby, and L. M. Faust. Living with volcano hazards, 2019.
- S. Sturton and J. Neuberg. The effects of a decompression on seismic parameter profiles in a gas-charged magma. *Journal of Volcanology and Geothermal Research*, 128(1-3): 187–199, nov 2003. doi: 10.1016/s0377-0273(03)00254-3.
- S. Sturton and J. Neuberg. The effects of conduit length and acoustic velocity on conduit resonance: Implications for low-frequency events. *Journal of Volcanology and Geothermal Research*, 151(4):319–339, mar 2006. doi: 10.1016/j.jvolgeores.2005.09.009.
- J. Taddeucci, D. M. Palladino, G. Sottili, D. Bernini, D. Andronico, and A. Cristaldi. Linked frequency and intensity of persistent volcanic activity at Stromboli (Italy). *Geophysical Research Letters*, 40(13):3384–3388, July 2013. doi: 10.1002/grl.50652.
- G. Tamburello, A. Aiuppa, E. P. Kantzas, A. J. S. McGonigle, and M. Ripepe. Passive vs. active degassing modes at an open-vent volcano (Stromboli, Italy). *Earth and Planetary Science Letters*, 359(Supplement C):106–116, Dec. 2012. ISSN 0012-821X. doi: 10.1016/j.epsl.2012.09.050. URL <http://www.sciencedirect.com/science/article/pii/S0012821X12005468>.
- The ObsPy Development Team. Obspy 1.1.0, Oct. 2017. URL <https://doi.org/10.5281/zenodo.165135>.
- R. I. Tilling. Volcano hazards. In *Volcanoes and the Environment*, pages 55–89. Cambridge University Press, oct 2005. doi: 10.1017/cbo9780511614767.003.
- US Geological Survey. Pyroclastic Flow deposit Mt. St. Helens. Global Volcanism Program, 1980. URL <https://volcano.si.edu/gallery/ShowImage.cfm?photo=GVP-02908>. [Online; accessed 31-March-2021].
- US Geological Survey, Public Domain. File:USGS Kilauea multimediaFile-1955.jpg — Wikimedia Commons, the free media repository. Public Domain, 2020a. URL [https://commons.wikimedia.org/w/index.php?title=File:USGS\\_K%C4%ABlauea\\_multimediaFile-1955.jpg&oldid=520943410](https://commons.wikimedia.org/w/index.php?title=File:USGS_K%C4%ABlauea_multimediaFile-1955.jpg&oldid=520943410). [Online; accessed 25-March-2021].



- US Geological Survey, Public Domain. File:Kilauea Fissure 8 cone erupting on 6-28-2018.jpg — Wikimedia Commons, the free media repository. Public Domain, 2020b. URL [https://commons.wikimedia.org/w/index.php?title=File:Kilauea\\_Fissure\\_8\\_cone\\_erupting\\_on\\_6-28-2018.jpg&oldid=491832848](https://commons.wikimedia.org/w/index.php?title=File:Kilauea_Fissure_8_cone_erupting_on_6-28-2018.jpg&oldid=491832848). [Online; accessed 25-March-2021]25.
- R. van Hout, L. Shemer, and D. Barnea. Spatial distribution of void fraction within a liquid slug and some other related slug parameters. *International Journal of Multiphase Flow*, 18(6):831–845, Nov. 1992. doi: 10.1016/0301-9322(92)90062-1.
- R. van Hout, D. Barnea, and L. Shemer. Evolution of statistical parameters of gas–liquid slug flow along vertical pipes. *International Journal of Multiphase Flow*, 27(9):1579–1602, Sept. 2001. doi: 10.1016/s0301-9322(01)00016-7.
- N. Varley, J. Johnson, M. Ruiz, G. Reyes, K. Martin, F. de Ciencias, and M. Colima. Applying statistical analysis to understanding the dynamics of volcanic explosions. In *Statistics in Volcanology*, volume 1, pages 57–76. Citeseer, 2006. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.521.1265&rep=rep1&type=pdf>.
- S. Vergnolle and C. Jaupart. Separated two-phase flow and basaltic eruptions. *Journal of Geophysical Research: Solid Earth*, 91(B12):12842–12860, nov 1986. doi: 10.1029/jb091ib12p12842.
- P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, feb 2020. doi: 10.1038/s41592-019-0686-2.
- Vito0783, CC BY-SA 4.0. File:Cuevas volcánicas, Villarrica.jpg — Wikimedia Commons, the free media repository. Creative Commons Attribution-Share Alike 4.0 International, Mar. 2019. URL [https://commons.wikimedia.org/w/index.php?title=File:Cuevas\\_volc%C3%A1nicas,\\_Villarrica.jpg&oldid=521424504](https://commons.wikimedia.org/w/index.php?title=File:Cuevas_volc%C3%A1nicas,_Villarrica.jpg&oldid=521424504). [Online; accessed 31-March-2021].
- D. H. von Seggern. Seismic Background Noise and Detection Threshold in the Southern Great Basin Digital Seismic Network. *Bulletin of the Seismological Society of America*, 94(6):2280–2298, Dec. 2004. doi: 10.1785/0120040003.

- D. H. von Seggern and R. Blandford. Seismic threshold determination. *Bulletin of the Seismological Society of America*, 66(3):753–788, June 1976. ISSN 0037-1106. URL <https://pubs.geoscienceworld.org/ssa/bssa/article-abstract/66/3/753/117618/Seismic-threshold-determination?redirectedFrom=fulltext>.
- G. P. Waite. An integrated analysis of low-frequency seismicity at villarrica volcano, chile, 2010.
- G. P. Waite, B. A. Chouet, and P. B. Dawson. Eruption dynamics at Mount St. Helens imaged from broadband seismic waveforms: Interaction of the shallow magmatic and hydrothermal systems. *Journal of Geophysical Research: Solid Earth*, 113(B2):B02305, #feb# 2008. ISSN 2156-2202. doi: 10.1029/2007JB005259. URL <http://onlinelibrary.wiley.com/doi/10.1029/2007JB005259/abstract>.
- J. Wassermann. Locating the sources of volcanic explosions and volcanic tremor at Stromboli volcano (Italy) using beam-forming on diffraction hyperboloids. *Physics of the Earth and Planetary Interiors*, 104(1-3):271–281, Nov. 1997. doi: 10.1016/S0031-9201(97)00041-1. URL <https://www.sciencedirect.com/science/article/pii/S0031920197000411?via%3Dihub>.
- J. Wassermann. Volcano Seismology. In P. Bormann, editor, *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, pages 1–77. Deutsches GeoForschungsZentrum GFZ, Potsdam, 2012. URL [http://doi.org/10.2312/GFZ.NMSOP-2\\_ch13](http://doi.org/10.2312/GFZ.NMSOP-2_ch13).
- M. Wathelet. An improved neighborhood algorithm: Parameter conditions and dynamic scaling. *Geophysical Research Letters*, 35(9), May 2008. doi: 10.1029/2008gl033256.
- M. Wathelet, D. Jongmans, and M. Ohrnberger. Surface-wave inversion using a direct search algorithm and its application to ambient vibration measurements. *Near Surface Geophysics*, 2(4):211–221, aug 2004. doi: 10.3997/1873-0604.2004018.
- U. Wegler and B. G. Lühr. Scattering behaviour at Merapi volcano (Java) revealed from an, active seismic experiment. *Geophysical Journal International*, 145(3):579–592, 2001. doi: <https://doi.org/10.1046/j.1365-246x.2001.01390.x>.
- P. Welch. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73, jun 1967. doi: 10.1109/tau.1967.1161901.
- E. Wielandt and P. Bormann. Seismic sensors and their calibration. *New Manual of Seismological Observatory Practices*, 1:46, 2002.
- M. Withers, R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo. A comparison of select trigger algorithms for automated

- global seismic phase and event detection. *Bulletin of the Seismological Society of America*, 88(1):95–106, Feb. 1998. ISSN 0037-1106. URL <https://pubs.geoscienceworld.org/ssa/bssa/article-abstract/88/1/95/102726/A-comparison-of-select-trigger-algorithms-for>.
- J. B. Witter, V. C. Kress, P. Delmelle, and J. Stix. Volatile degassing, petrology, and magma dynamics of the Villarrica Lava Lake, Southern Chile. *Journal of Volcanology and Geothermal Research*, 134(4):303–337, July 2004. ISSN 0377-0273. doi: 10.1016/j.jvolgeores.2004.03.002. URL <http://www.sciencedirect.com/science/article/pii/S0377027304000630>.
- R. Wright, L. P. Flynn, H. Garbeil, A. J. L. Harris, and E. Pilger. MODVOLC: near-real-time thermal monitoring of global volcanism. *Journal of Volcanology and Geothermal Research*, 135(1-2):29–49, July 2004. doi: 10.1016/j.jvolgeores.2003.12.008.
- A. G. Yeguas, J. Almendros, R. Abella, and J. M. Ibáñez. Quantitative analysis of seismic wave propagation anomalies in azimuth and apparent slowness at Deception Island volcano (Antarctica) using seismic arrays. *Geophysical Journal International*, 184(2):801–815, #feb# 2011. ISSN 0956-540X. doi: 10.1111/j.1365-246X.2010.04864.x. URL <https://academic.oup.com/gji/article/184/2/801/594068/Quantitative-analysis-of-seismic-wave-propagation>.
- V. M. Zobin. *Introduction to volcanic seismology*. Elsevier, 2011.
- V. M. Zobin. Chapter 14 - Seismic Signals Associated With Volcanic Explosions. In *Introduction to Volcanic Seismology (Third Edition)*, pages 323–359. Elsevier, Oxford, 2017a. ISBN 978-0-444-63631-7. URL <http://www.sciencedirect.com/science/article/pii/B9780444636317000145>. DOI: 10.1016/B978-0-444-63631-7.00014-5.
- V. M. Zobin. Chapter 16 - Swarms of Microearthquakes Associated With Effusive and Explosive Activity at Volcanoes. In *Introduction to Volcanic Seismology (Third Edition)*, pages 389–416. Elsevier, Oxford, 2017b. ISBN 978-0-444-63631-7. URL <http://www.sciencedirect.com/science/article/pii/B9780444636317000169>. DOI: 10.1016/B978-0-444-63631-7.00016-9.
- V. M. Zobin. Chapter 11 - Origin of Eruption Earthquakes. In *Introduction to Volcanic Seismology (Third Edition)*, pages 243–262. Elsevier, Oxford, 2017c. ISBN 978-0-444-63631-7. URL <http://www.sciencedirect.com/science/article/pii/B978044463631700011X>. DOI: 10.1016/B978-0-444-63631-7.00011-X.
- V. M. Zobin. Chapter 2 - Seismicity at Volcanoes. In *Introduction to Volcanic Seismology (Third Edition)*, pages 13–34. Elsevier, Oxford, 2017d. ISBN 978-0-444-63631-7. URL <http://www.sciencedirect.com/science/article/pii/B9780444636317000029>. DOI: 10.1016/B978-0-444-63631-7.00002-9.

- L. Zuccarello, M. R. Burton, G. Saccorotti, C. J. Bean, and D. Patanè. The coupling between very long period seismic events, volcanic tremor, and degassing rates at Mount Etna volcano. *Journal of Geophysical Research: Solid Earth*, 118(9):4910–4921, Sept. 2013. doi: 10.1002/jgrb.50363. URL <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/jgrb.50363>.
- L. Zuccarello, M. Paratore, M. La Rocca, F. Ferrari, A. Messina, S. Branca, D. Contrafatto, D. Galluzzo, S. Rapisarda, and L. García. Shallow velocity model in the area of Pozzo Pitarrone, Mt. Etna, from single station, array methods and borehole data. *Annals of Geophysics*, 59(4):0433, 2016. doi: <https://doi.org/10.4401/ag-7086>. URL <https://www.annalsofgeophysics.eu/index.php/annals/article/view/7086>.

## A. Supporting Information Chapter 4

The following section describes how the bandwidth of the deployed geophones can be extended up to at least one decade below their natural frequency of 4.5 Hz. It is illustrated in Figs. S1-4 and Table S1 provides technical parameters of the sensors.

Figs. 5-10 show additional data examples to give a better impression of the recorded seismicity. Processing steps are given in the captions.

Fig. 11 shows the operation time of the stations used for the amplitude decay methods. The positions of these stations are shown in Fig. 12.

### Instrument correction of 4.5Hz-geophones

#### Laboratory experiment

Our group tested geophones, that are typically used in exploration seismology for their suitability in passive seismology. These sensors have relatively high natural frequencies (e.g. 1 Hz, 4.5 Hz, 10 Hz). In general, it is not recommended to use the frequency range below this natural frequency since the signal may be affected by filter effects that result from the sensor system itself. However, if this effect (the transfer function) is known, it can be - to some extent - corrected. In order to determine how much this bandwidth can be increased below the natural frequency, our group investigated some sensors under laboratory conditions [Hilbert, 2015]. Amongst others, a OYO Geospace 4.5Hz 3-components GS-11D/3Y and a Iongoe SM-6 10Hz sensor have been tested according to Havskov and Alguacil [2004], Pavlis and Vernon [1994], Wielandt and Bormann [2002]. In the basement of the institute building, the geophones were placed next to two calibrated Guralps Systems CMG-3TD broadband seismometers which served as reference. Three days of noise were recorded by this set up and used for calibration, verification of the transfer function and comparison of the low-frequent signals.

Here, we present samples of the vertical displacement after instrument correction, filtered at different frequency ranges (Figs. A.1-A.3). Between 0.5-1.0 Hz, even the recording of the 10 Hz geophone is remarkably similar to that of the Guralp seismometers. At lower frequencies (0.1-0.2 Hz, Fig. A.3) however, the amplitude can not be reconstructed anymore for this sensor. The 4.5 Hz-geophone on the other hand still shows almost identical amplitudes to the seismometers.

**Table A.1.:** Instrument parameters

sensor	Generator const.	damping	samp.rate	poles	sensitivity
1C	28.8 Vs/m	0.56	100 Hz	-20.07+19.91j, -20.07-19.91j	$0.1180 \times 10^{10}$
3C	28.8 Vs/m	0.7	100 Hz	-15.83+23.43j, -15.83-23.43j	$0.1891 \times 10^{10}$

### Field data from Villarrica

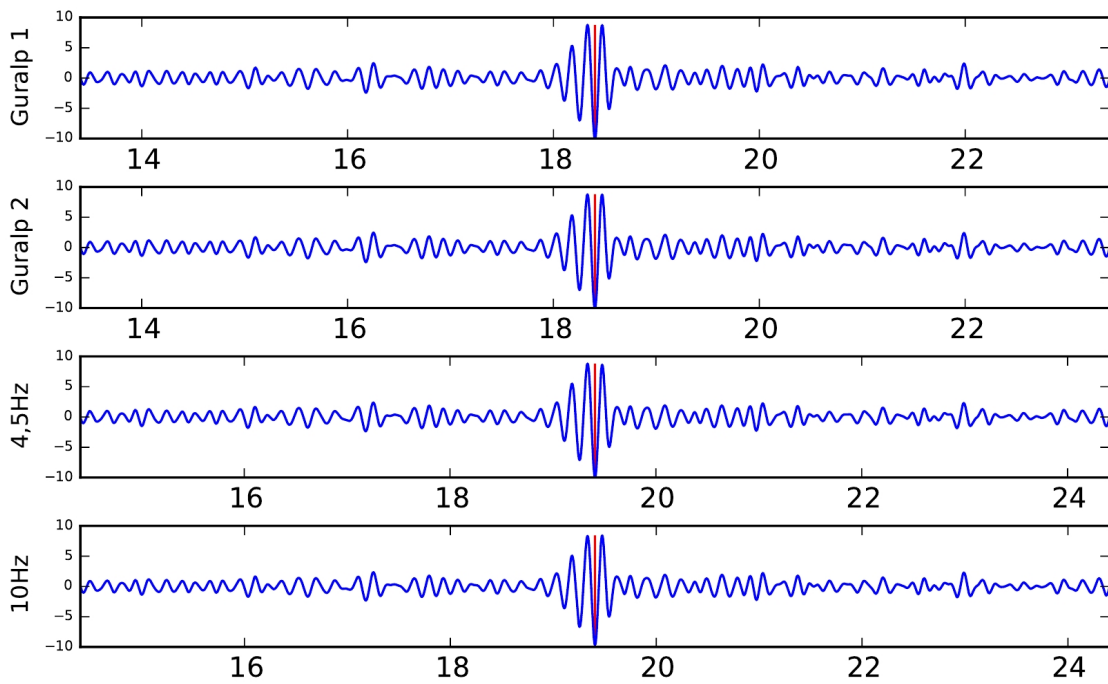
The sensors used in the experiment above were not exactly identical to those used at Villarrica. Therefore we show the effect of a bandpass filter with constant upper and decreasing lower cut-off frequency using samples from array AVW. This approach is inspired by an example in Havskov and Ottemöller [2010]. The data is corrected for the velocity response using a water level of 40. The prefilter, that was used in the actual data processing by default, is not applied here. If frequencies below the resolution capability of the system (sensor, digitizer and applied transfer function) are included, the instrument correction starts to enhance the electronic noise of the instrument. This is usually a low-frequency signal. In order to stress, that it is indeed an artificial signal, we compare the seismograms of the five stations in array AVW. The mean distance between the stations is around 100 m. Thus at the critical frequencies ( $\leq 4.5$  Hz), the spacing is small enough that we may expect a passably coherent signal at all stations (e.g. for  $v = 500$  m/s and  $f = 4.5$  Hz the wavelength is  $\lambda \approx 100$  m). Nevertheless, we point out, that this is not a laboratory experiment, but a site with very heterogeneous structure. Therefore, we can neither expect perfect similarity.

The data is filtered using a bandpass with a constant upper cut-off frequency of 8.0 Hz and different lower cut-off frequencies  $f_{min}$  of 4.5, 2.0, 1.0, 0.5, 0.1, 0.01 and 0.001 Hz (Fig. A.4).

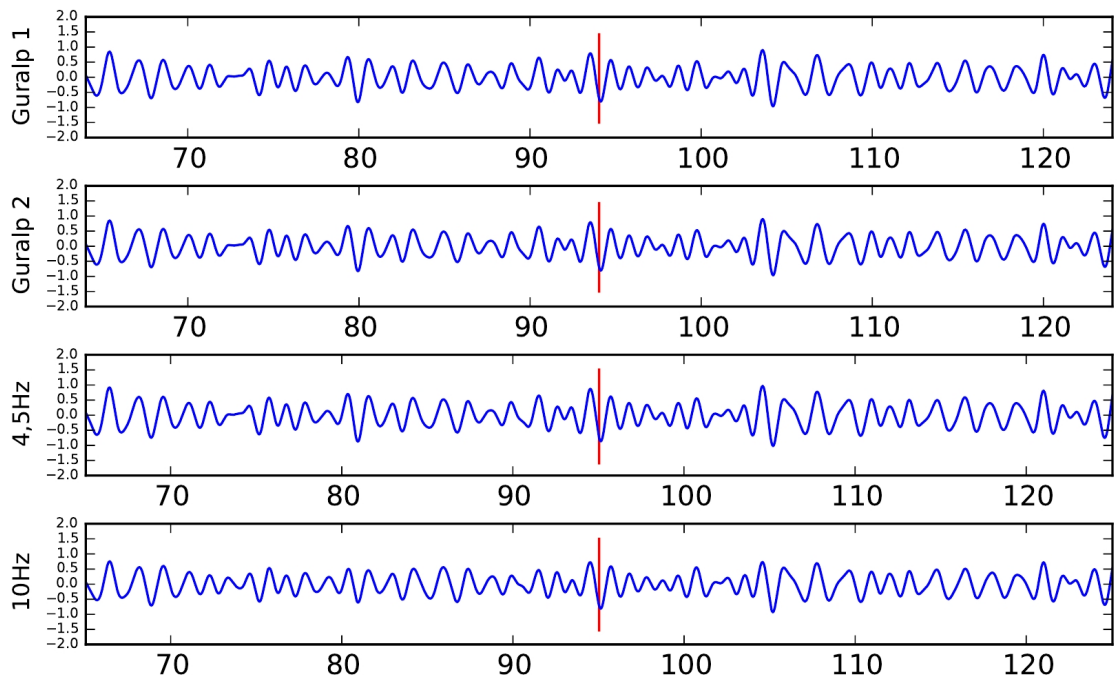
For  $f_{min} \leq 0.1$  Hz, incoherent, low-frequency phases emerge. These become more evident at even lower  $f_{min}$ . Above 0.5 Hz however, no such artifacts are visible. We acknowledge, that the coherency between the stations is not perfect at higher frequencies. Nevertheless, the main similarities between the signals are maintained above 0.5 Hz, whereas below 0.1 Hz this is undeniably not the case anymore. Therefore, we think, the incoherency at higher frequencies is caused by a complicated site structure rather than by the sensor unit.

All in all, we are therefore confident, that the data of these sensors can safely be interpreted for frequencies as low as 0.5 Hz.

### Data samples

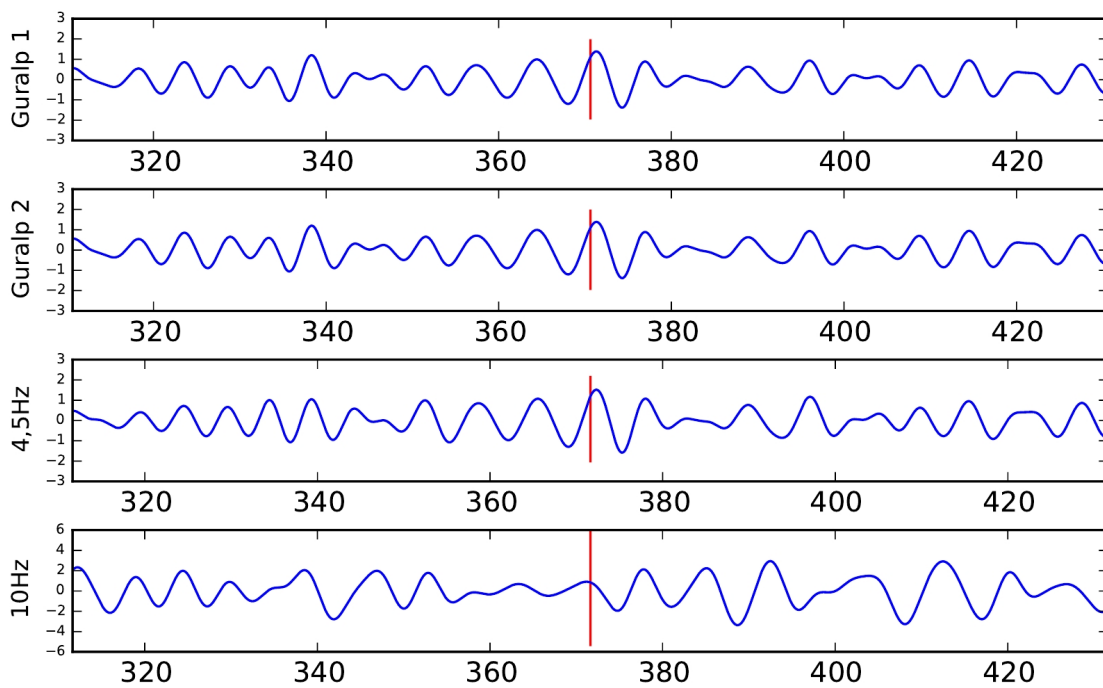


**Figure A.1.:** Signal filtered between 5 Hz and 10 Hz (traffic noise). Y-axis is displacement in nm, x-axis is time in s.

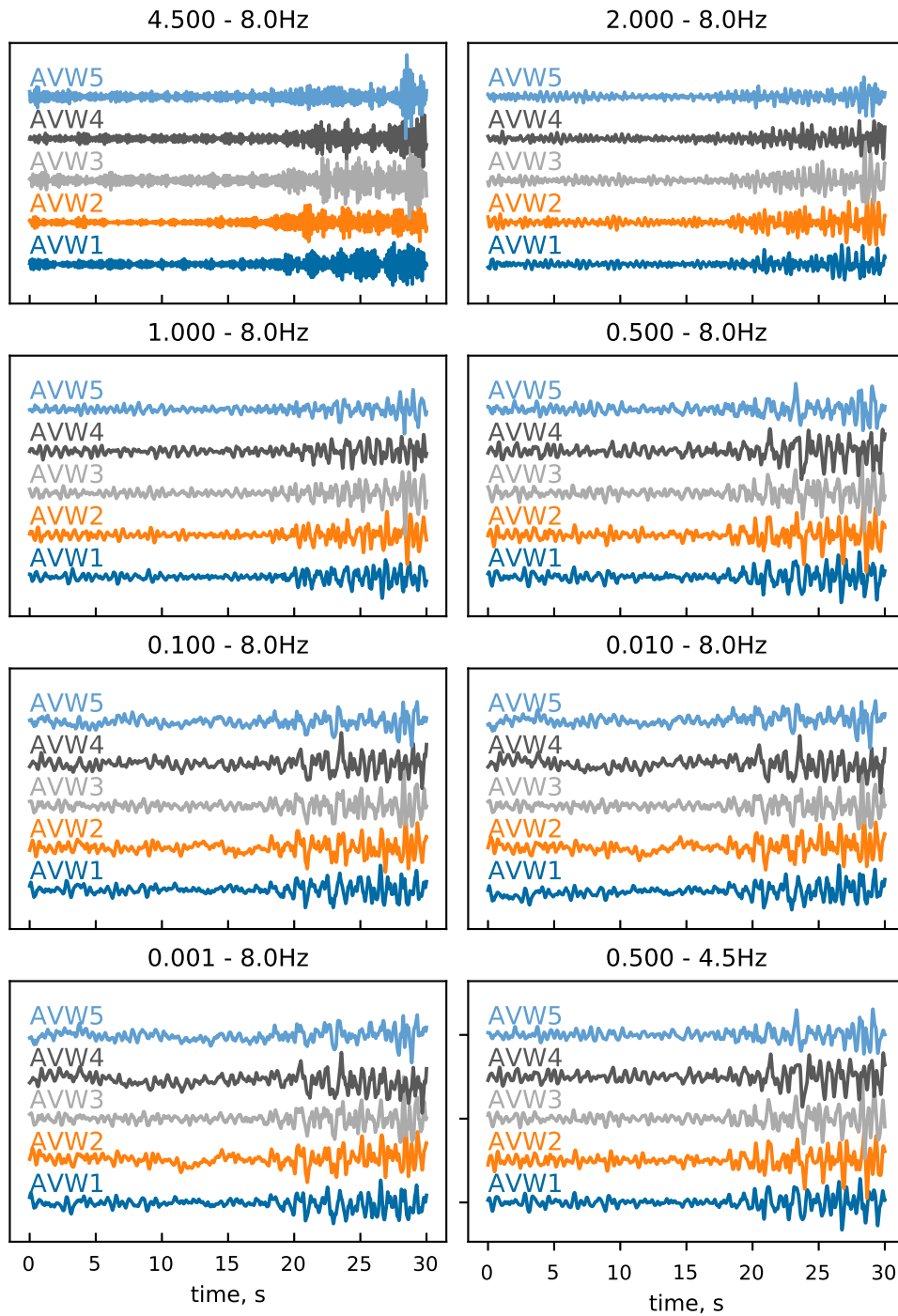


**Figure A.2.:** Signal filtered between 0.5 Hz and 1.0 Hz (ocean noise). Y-axis is displacement in nm, x-axis is time in s.

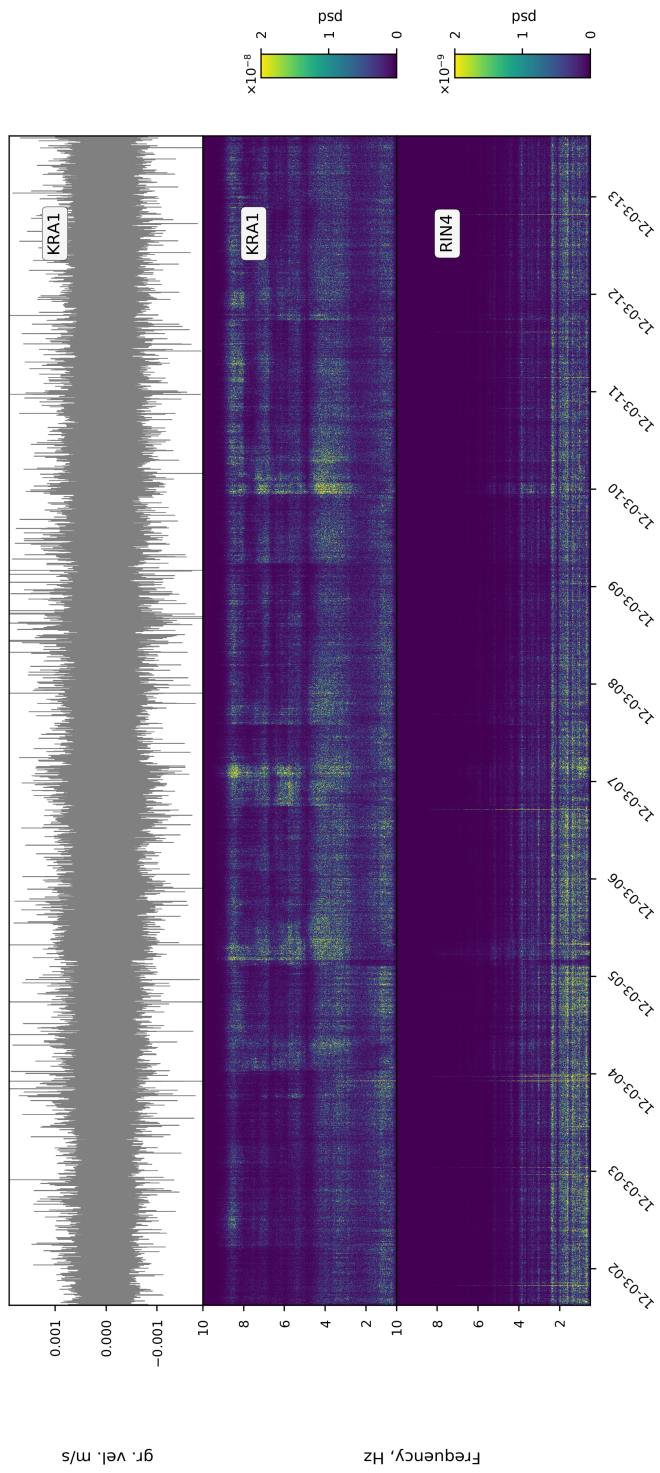




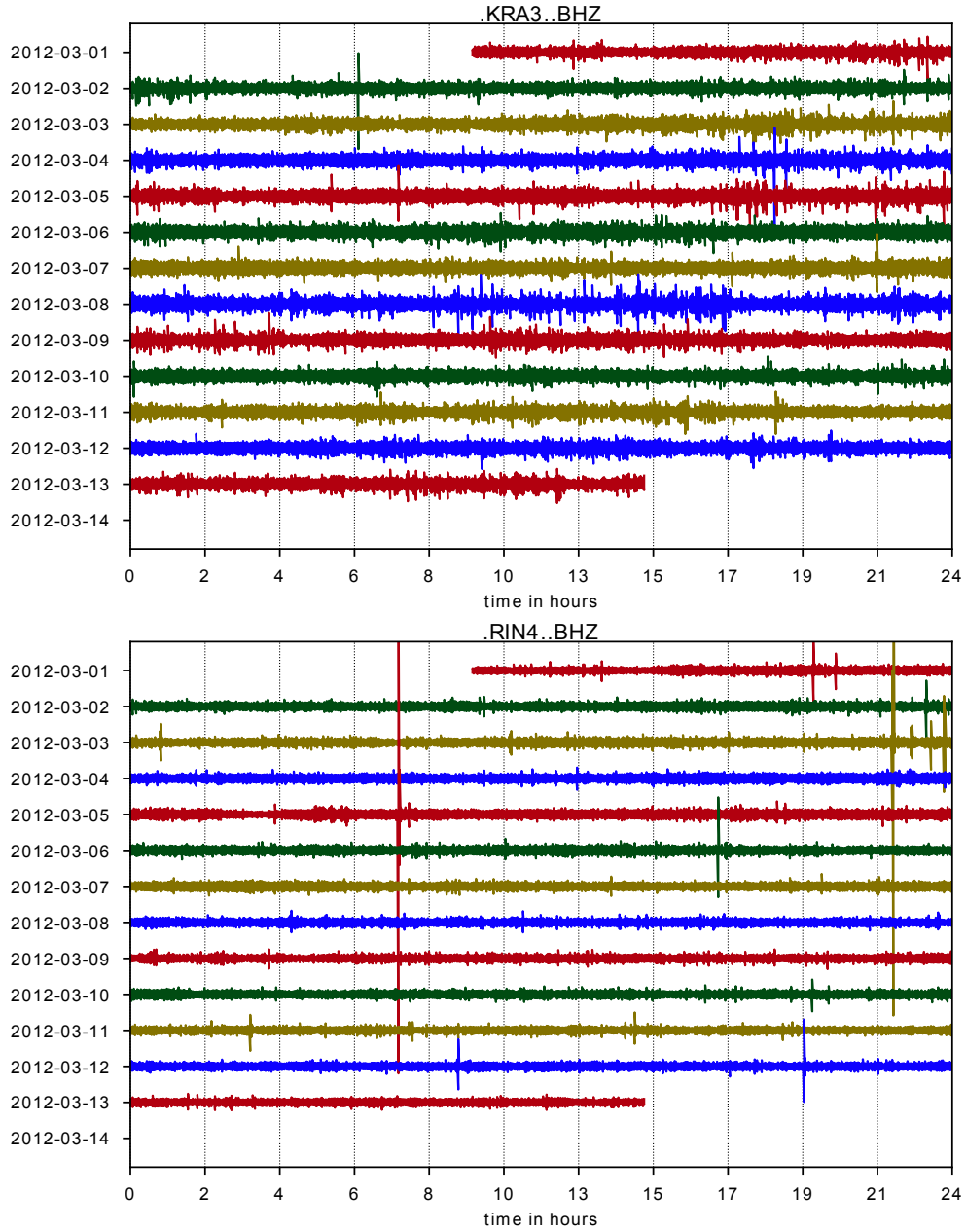
**Figure A.3.:** Signal filtered between 0.1 Hz and 0.2 Hz (ocean noise). Y-axis is displacement in nm, x-axis is time in s.



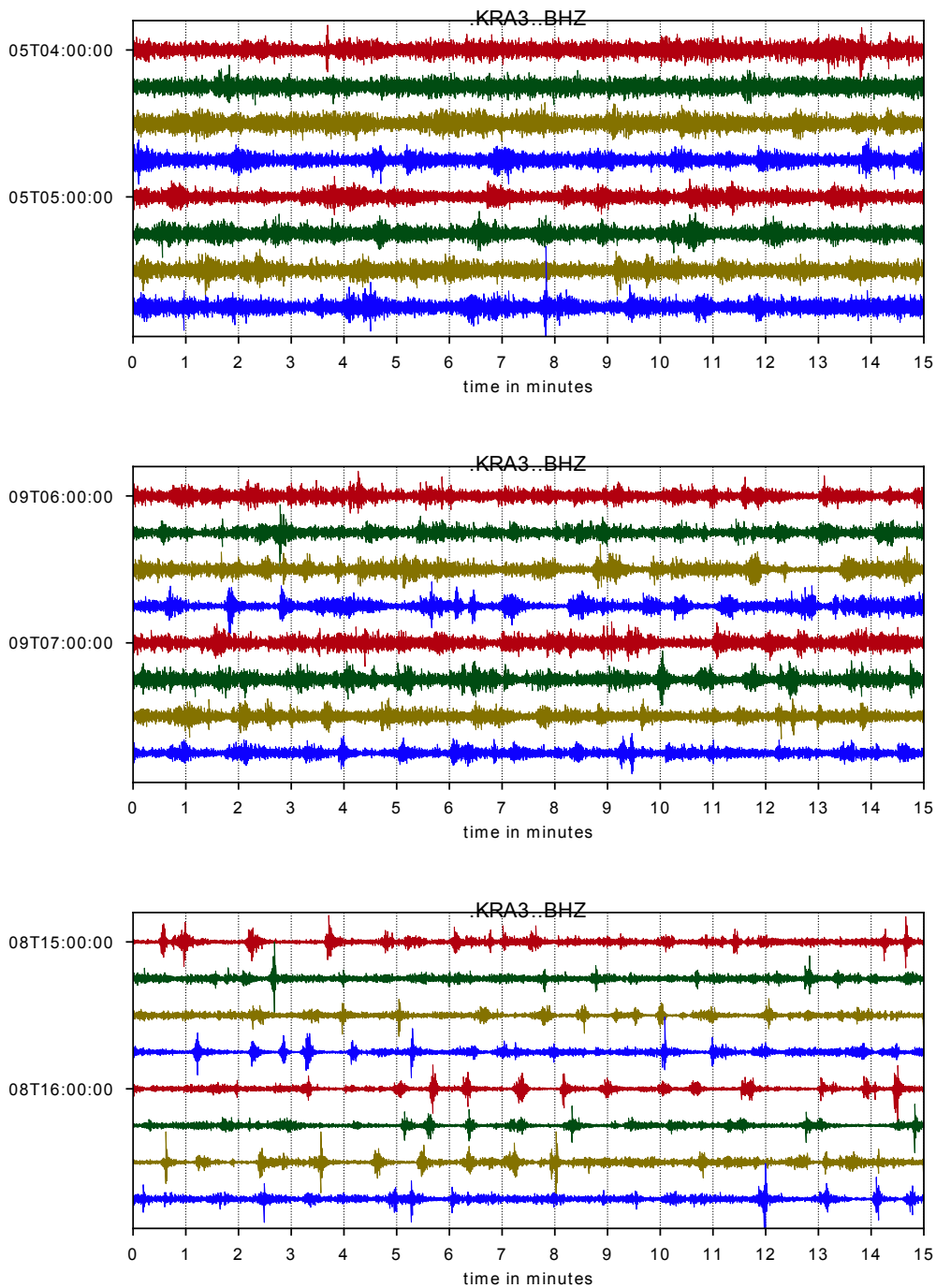
**Figure A.4.:** Velocity seismograms from array AVW, filtered using different lower cut-off frequencies. Lower right panel shows recoverable low-frequency part to demonstrate waveform similarity.



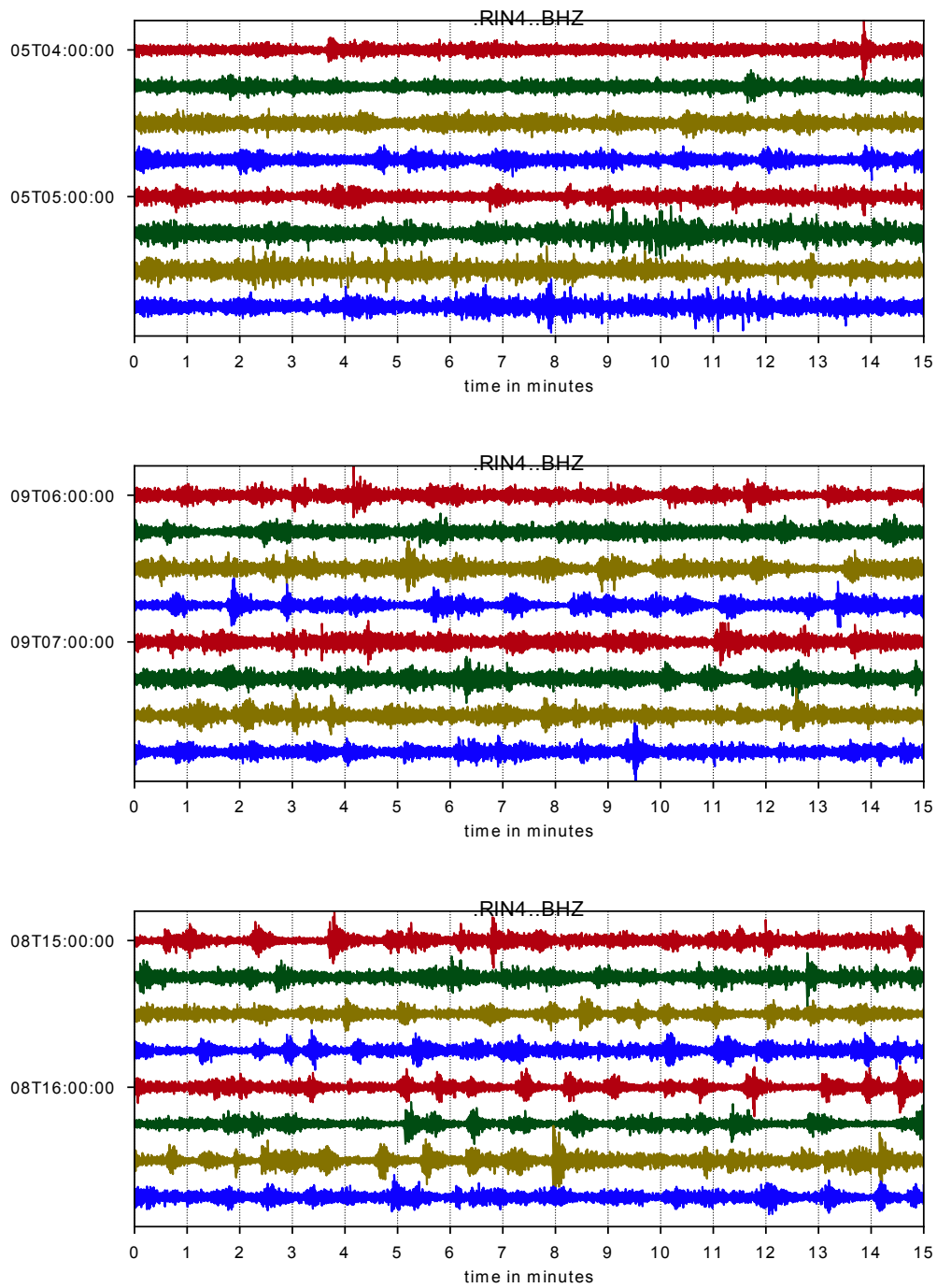
**Figure A.5:** Ground velocity at KRA1 and corresponding spectrograms at KRA1 and RIN4. Filter: 0.5-12.5 Hz, sampling rate: 25 Hz, amplitudes corrected for geometrical spreading by  $\sqrt{r}$ , spectrogram: 5.5 min with 2.7 min overlap



**Figure A.6.:** Ground velocity during full observation period at the summit (KRA3, top) and in approx. 2 km distance. High-amplitude spikes as RIN4 represent regional tectonic earthquakes. Filter: 0.5-12.5 Hz, sampling rate: 25 Hz, amplitudes corrected for geometrical spreading by  $\sqrt{r}$

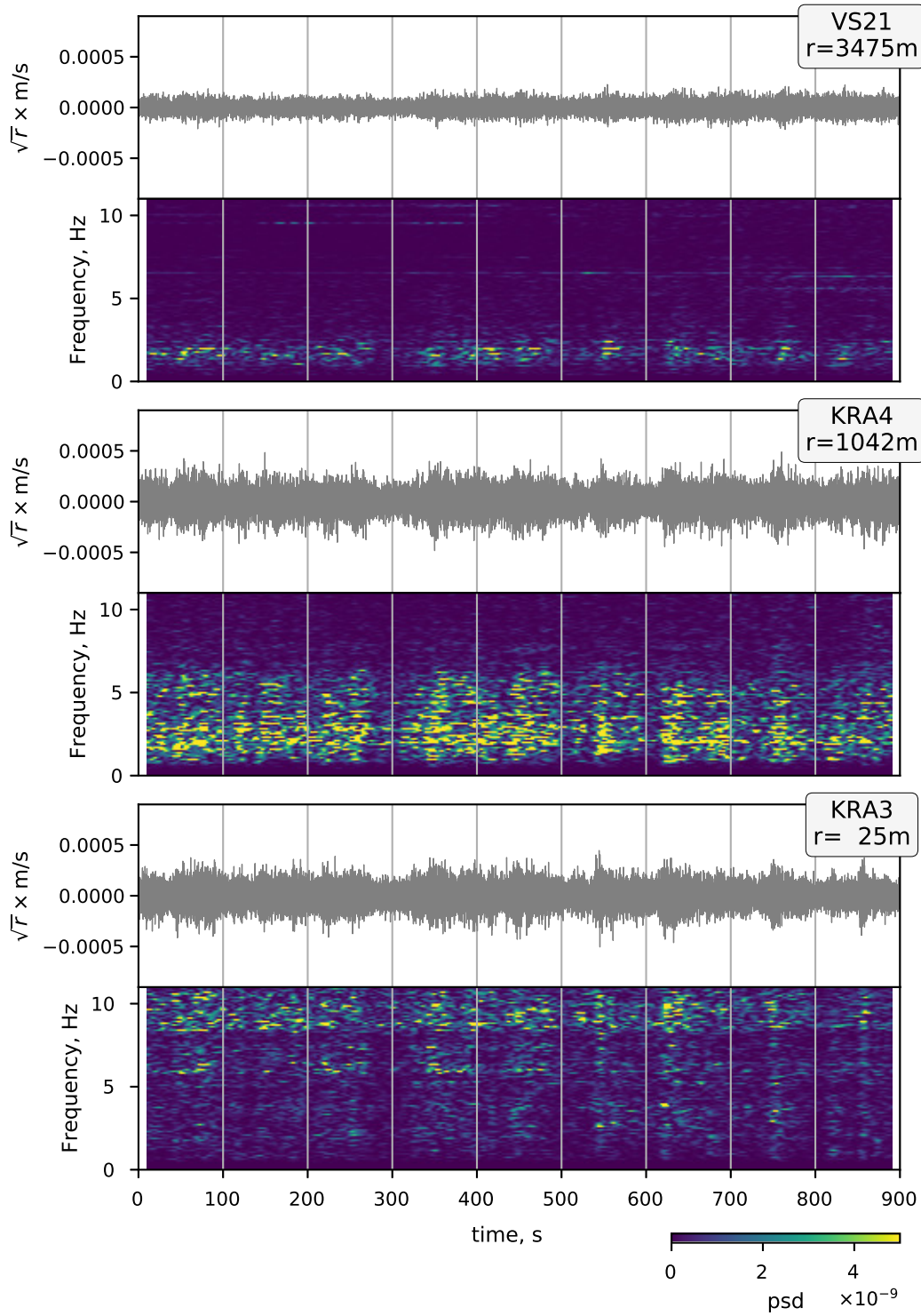


**Figure A.7.:** Samples of tremor dominated (top), mixed (middle) and event dominated (bottom) sequences at KRA3. Same samples as in Fig. A.8. Filter: 0.5 - 12.5 Hz, amplitudes corrected for geometrical spreading by  $\sqrt{r}$

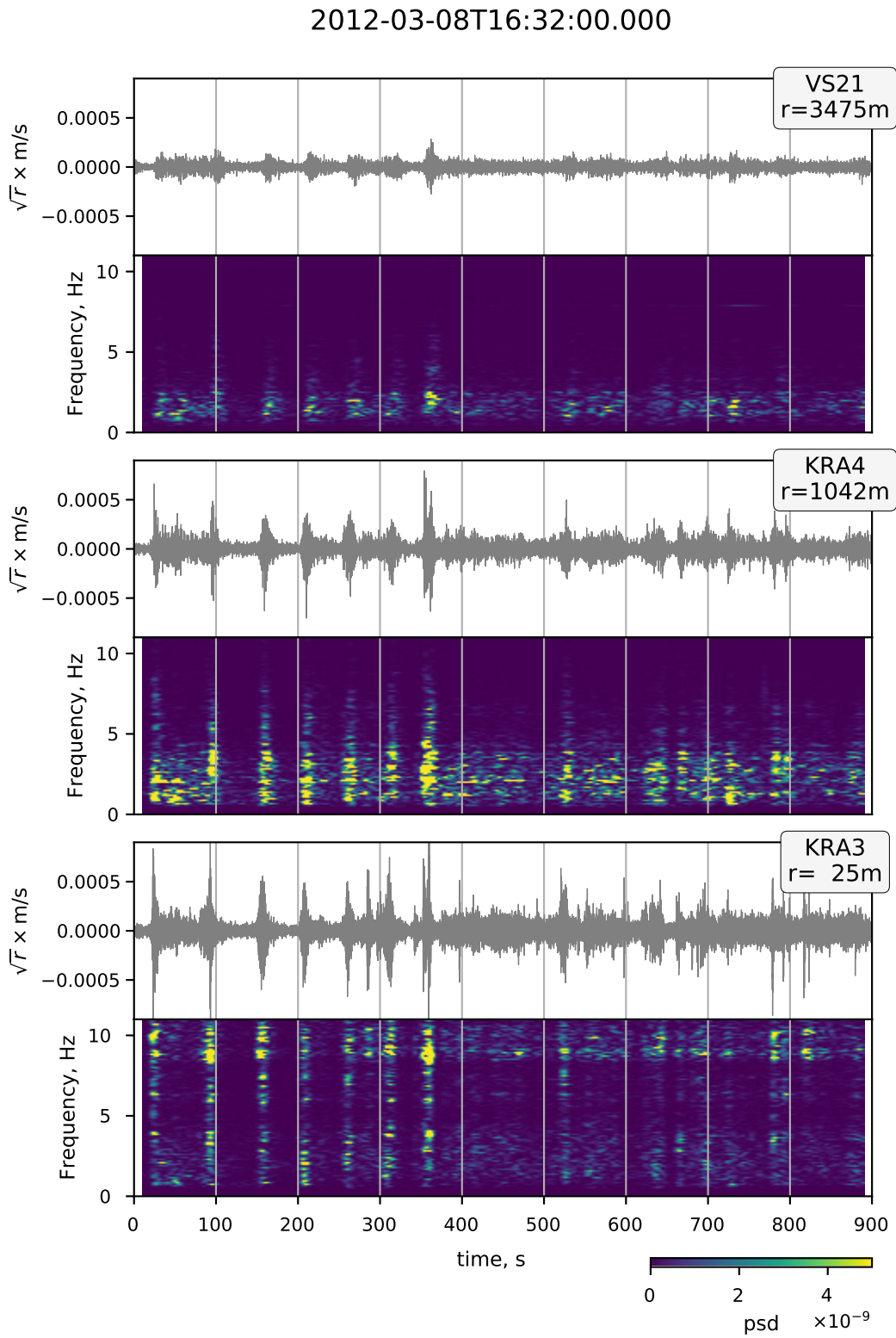


**Figure A.8.:** Samples of tremor dominated (top), mixed (middle) and event dominated (bottom) sequences at RIN4. Same samples as in Fig. A.7. Filter: 0.5 - 12.5 Hz amplitudes corrected for geometrical spreading by  $\sqrt{r}$

2012-03-05T04:30:00.000



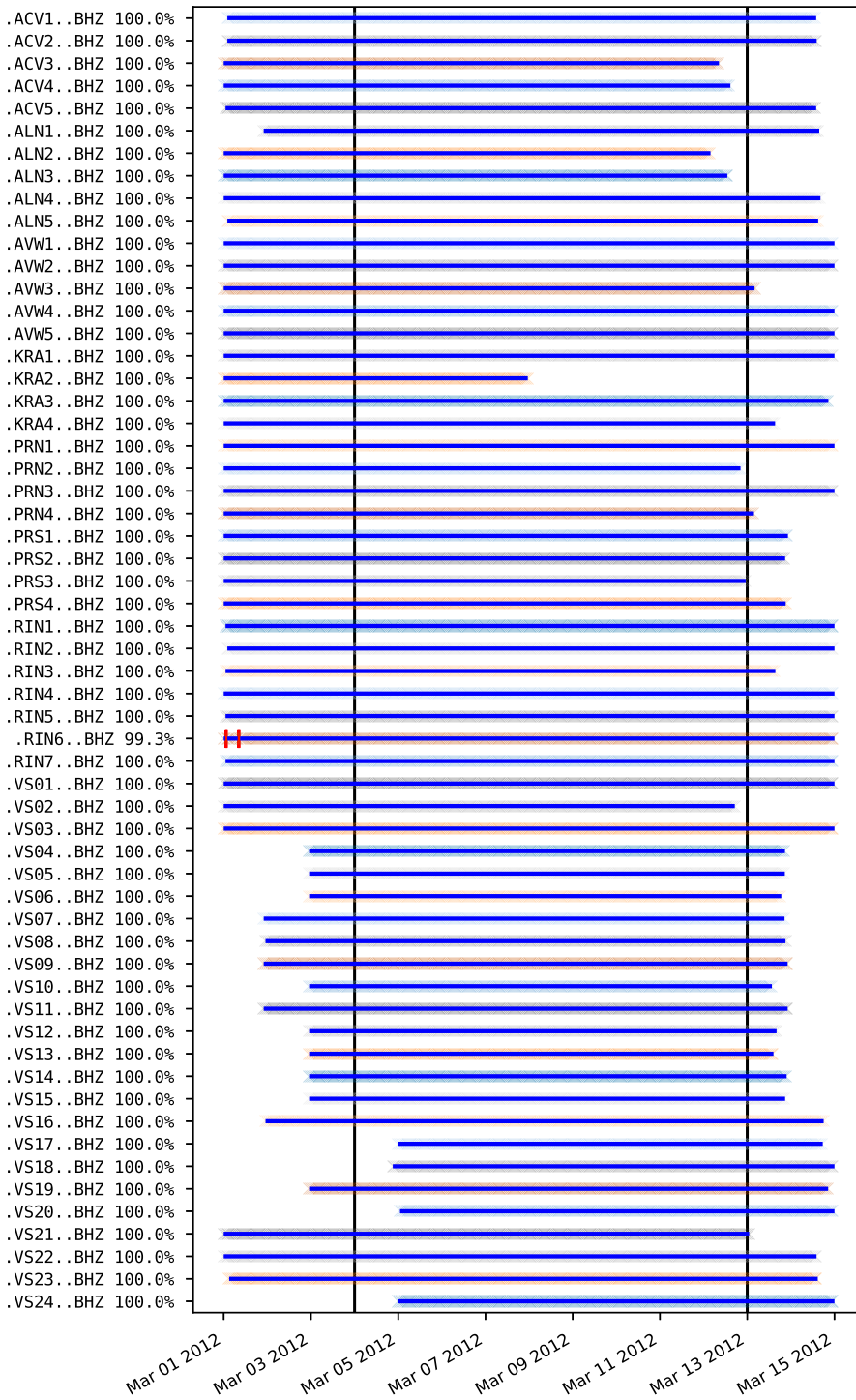
183  
**Figure A.9.:** Tremor dominated sequence. Variations in amplitude could be caused by embedded events. Frequency: 0.5 - 25.0 Hz, sampling rate: 50 Hz, spectrogram: 20.48 s with 20 s overlap, amplitudes are compensated for distance.



184

**Figure A.10.:** Event dominated sequence. Transients are particularly prominent in the first 400s. Afterwards inter-event tremor level increases and obscures events. Frequency: 0.5 - 25.0 Hz, sampling rate: 50 Hz, spectrogram: 20.48s with 20s overlap, amplitudes are compensated for distance.





**Figure A.11.:** Data availability at stations used in the paper. Black lines mark time window used for array analysis. Horizontal channels, if available, are not shown.

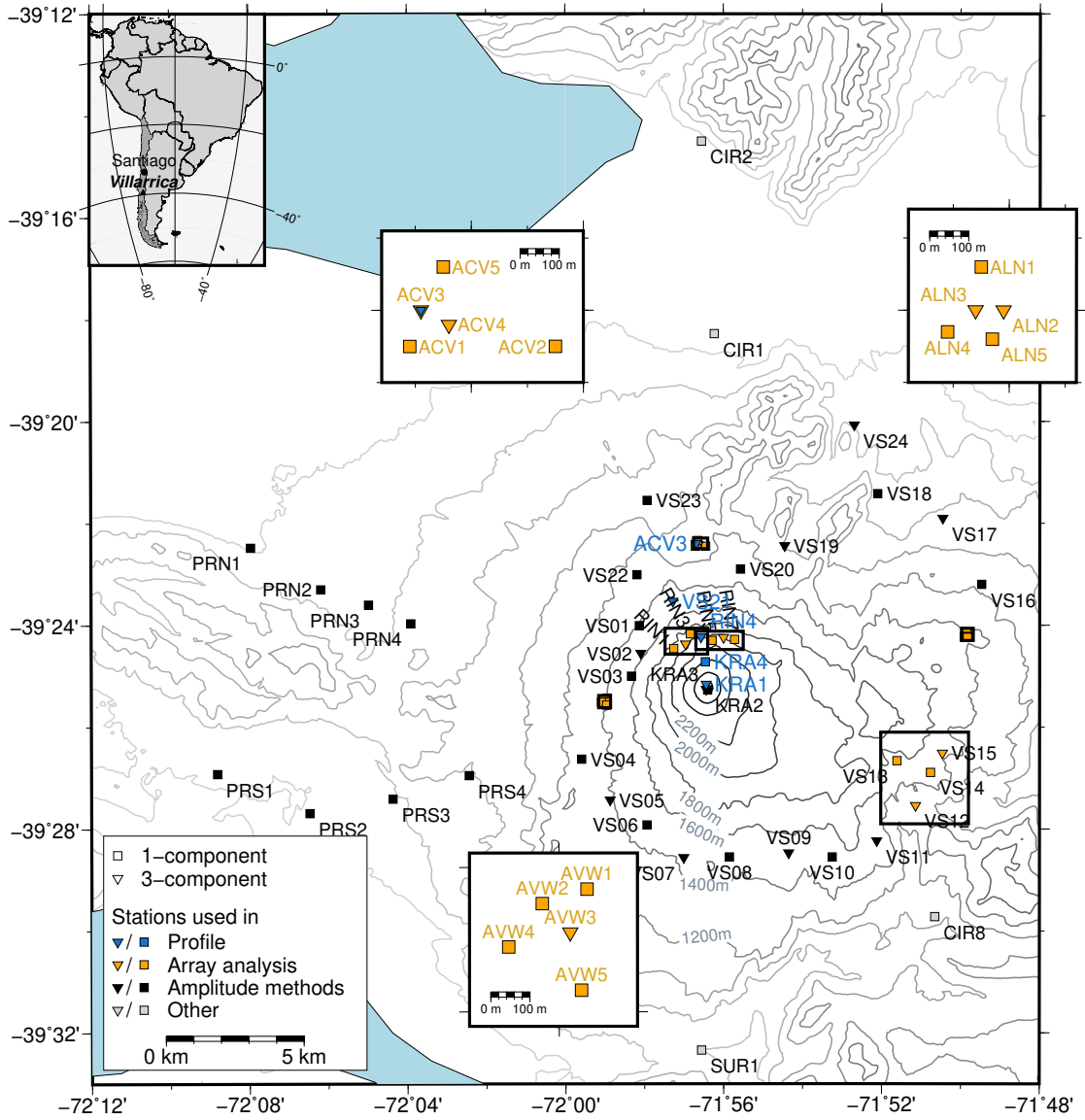


Figure A.12.: Map of network including station names.

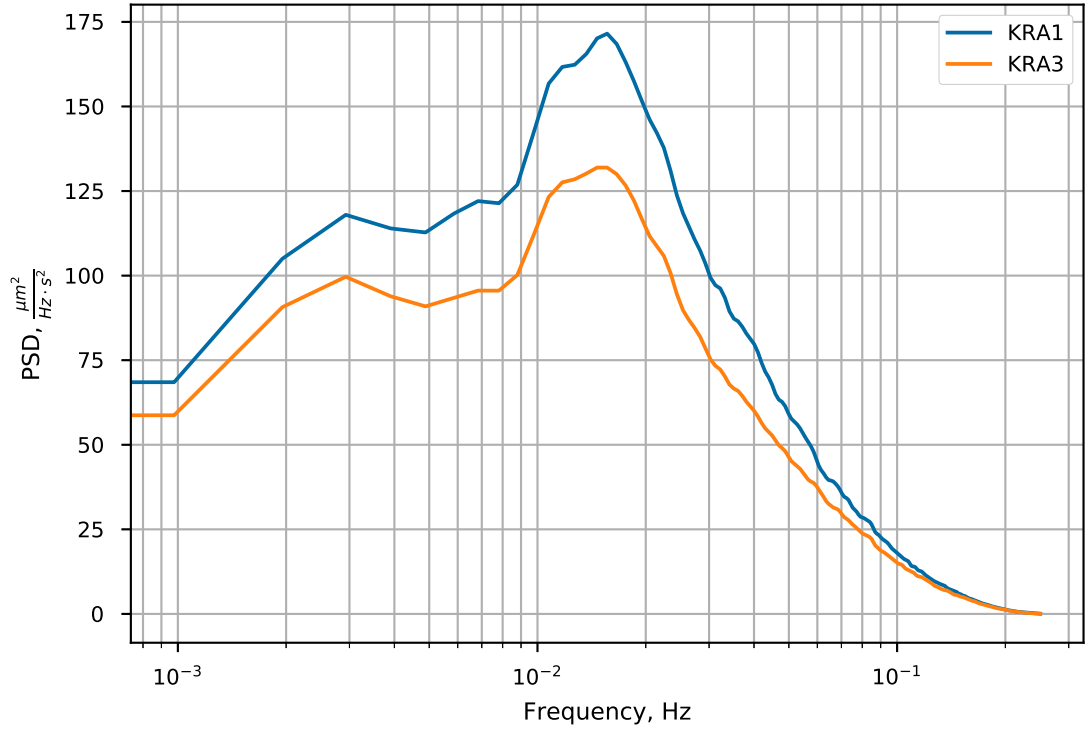
## B. Supporting Information Chapter 5

**Table B.1.:** These six regional events in South America visibly affected the data in our network. From USGS.

---

2012-03-12T19:38:36.000Z
2012-03-05T07:48:40.040Z
2012-03-03T23:44:04.000Z
2012-03-03T23:43:59.040Z
2012-03-03T22:46:40.000Z
2012-03-03T22:13:55.000Z

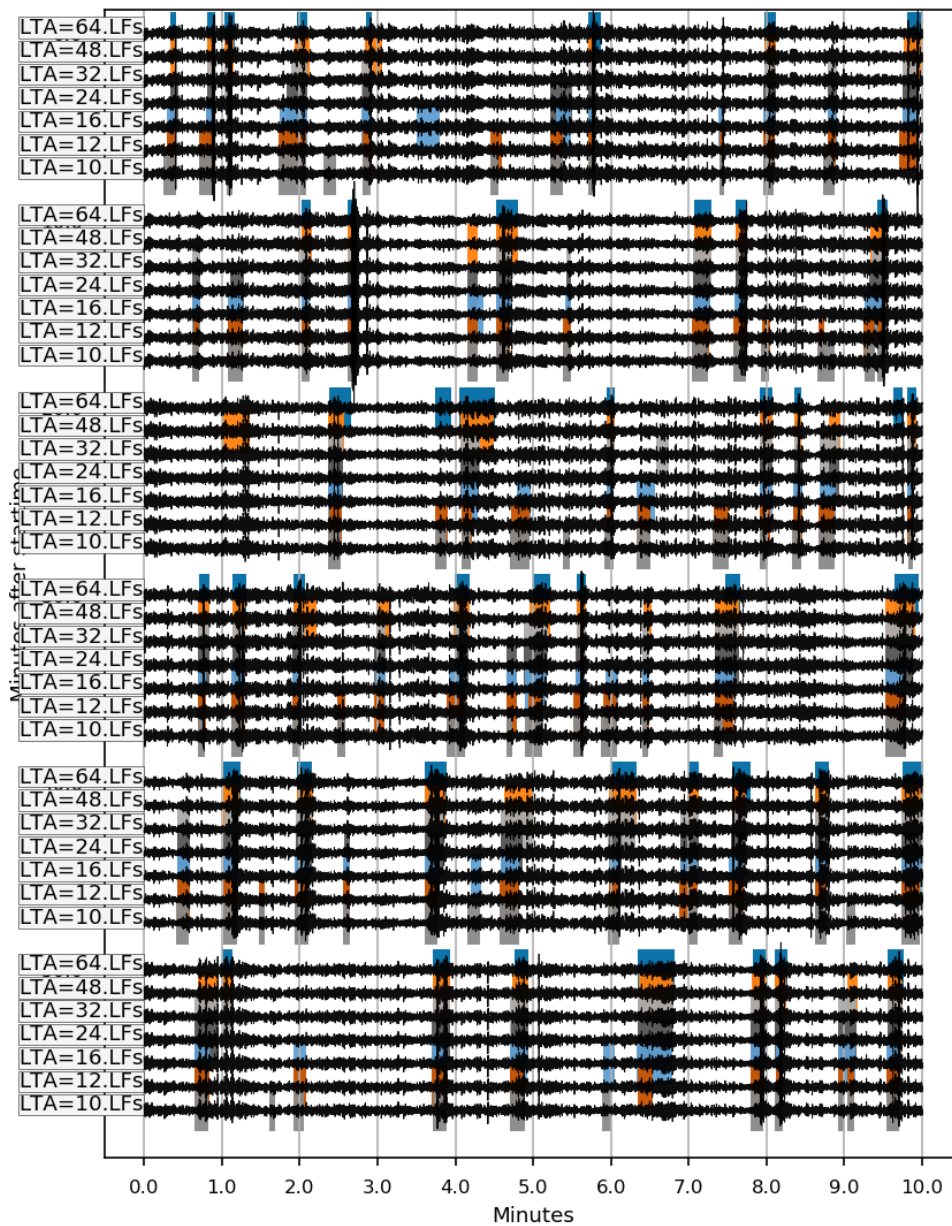
---



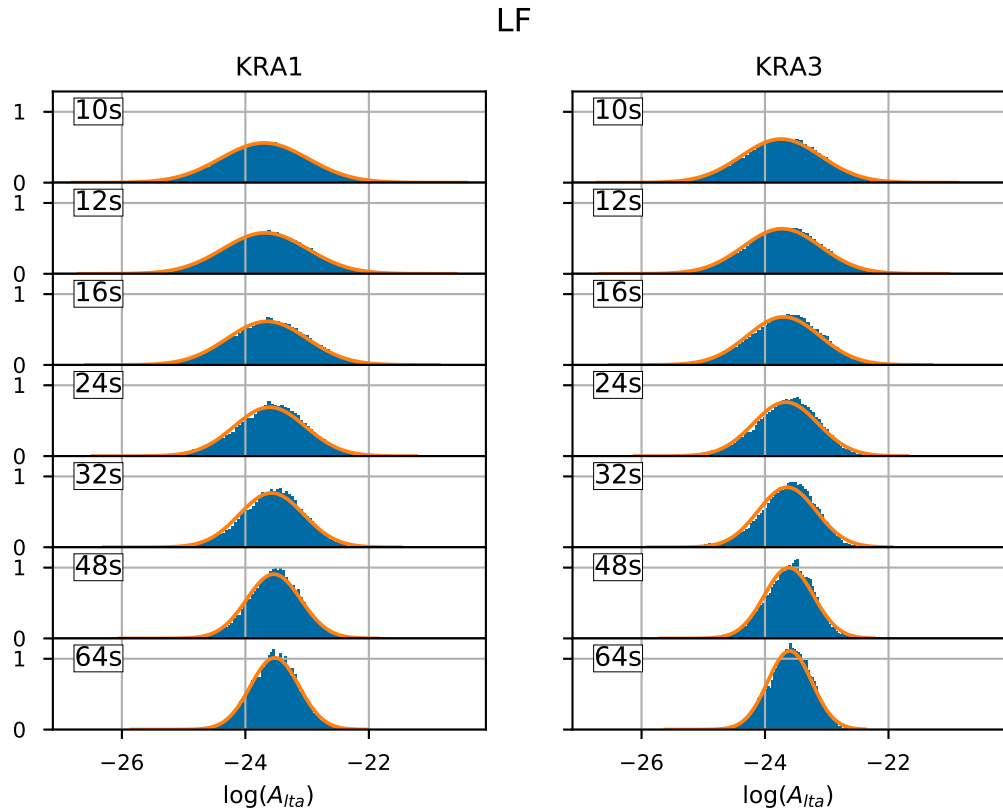
**Figure B.1.:** Power spectral density of mean amplitude in 4 s windows every 2 s at 0.5-5.0 Hz. There is a spectral peak at  $f=0.0146$  Hz corresponding to a period of 68 s. This is similar to the data presented by Ripepe et al. [2010].

**Table B.2.:** Parameters of linear relation between magnitude and maximum short-term moving mean squared amplitude of an event. Numbers are rounded to the third significant digit

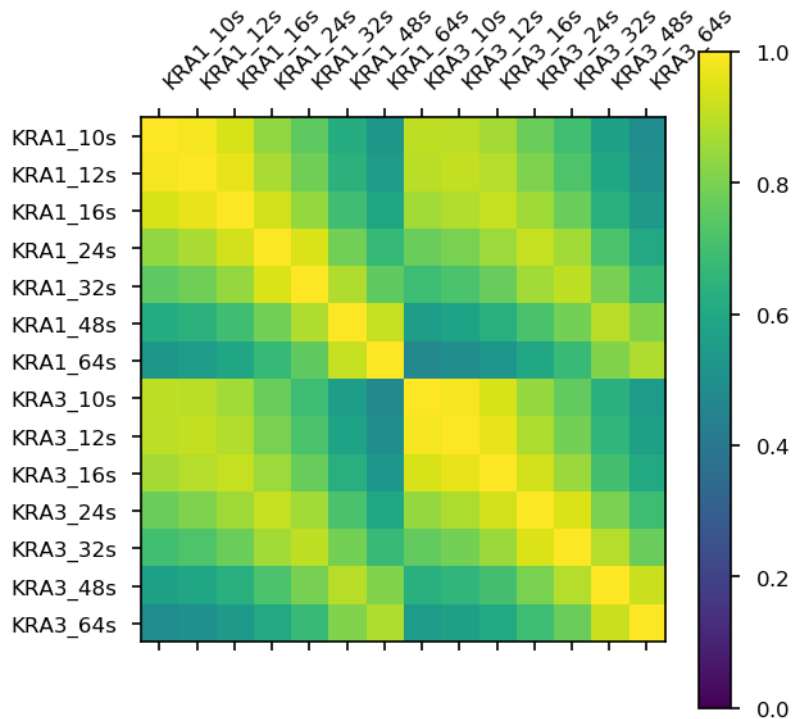
	slope	intercept	mean_error	std_error
KRA1	1.81	-22.5	1.86e-15	0.239
KRA3	1.6	-22.6	7.47e-15	0.22



**Figure B.2.:** STA/LTA triggered events using different LTA window sizes. Each stack of seven traces shows consecutive 10 minutes of seismic data. Within each stack, all traces are identical but show detections at different LTA window lengths. The window size is given at the beginning of each trace and represented by different colors. Seismic data are the vertical velocity component of station KRA3, filtered at 0.5-0.5 Hz.



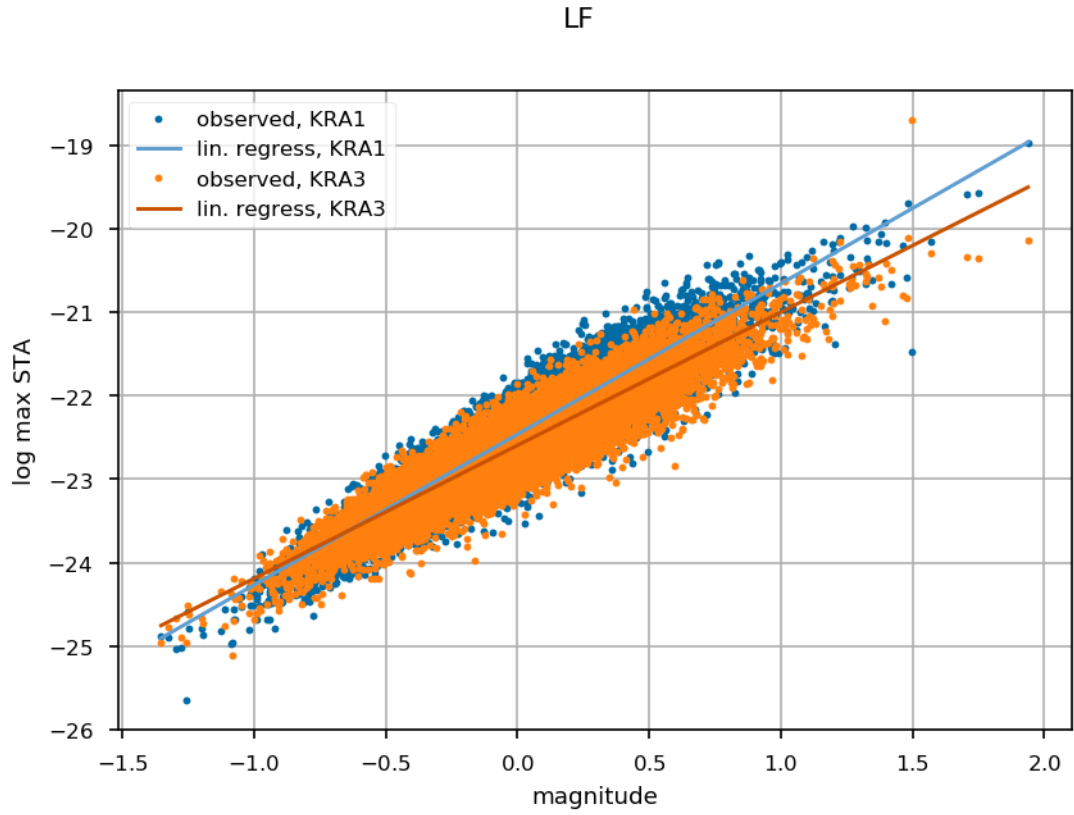
**Figure B.3.:** Distribution of logarithmic mean square amplitudes of the noise samples. The window size in seconds is indicated in the panels. Columns correspond to stations while rows correspond to window sizes. Mean and standard deviation of the logarithmic data served as input in the MC-simulation for the distributions of the LTA-amplitudes. Values are given in Data Set S1. The data were generated by computing a moving mean-square amplitude over 5-min long samples which were extracted at the beginning of every hour over the entire period of observation.



**Figure B.4.:** Correlation coefficient between LTA-data of both stations and all window sizes. Visualization of Data Set S1.

**Table B.3.:** Results of maximum likelihood estimation of probability density functions to model the observed magnitude distribution. Calculations were performed with scipy 1.4.1

pdf	AIC	KS_pval	loc	scale	shape_0	shape_1
lognorm	5748.4	5.31e-08	-6.2	6.13	0.05	-
loglogistic	5578.3	4.19e-01	-6.1	6.12	35.55	-
gengamma	5795.0	3.96e-13	-2.0	0.81	7.78	2.28
invgauss	5903.7	2.75e-17	-2.6	167.90	0.02	-
norm	5815.2	2.61e-11	-0.0	0.31	-	-
weibull	6464.8	3.08e-40	-1.4	1.46	4.56	-
gumbel	7548.0	6.30e-90	-0.2	0.31	-	-



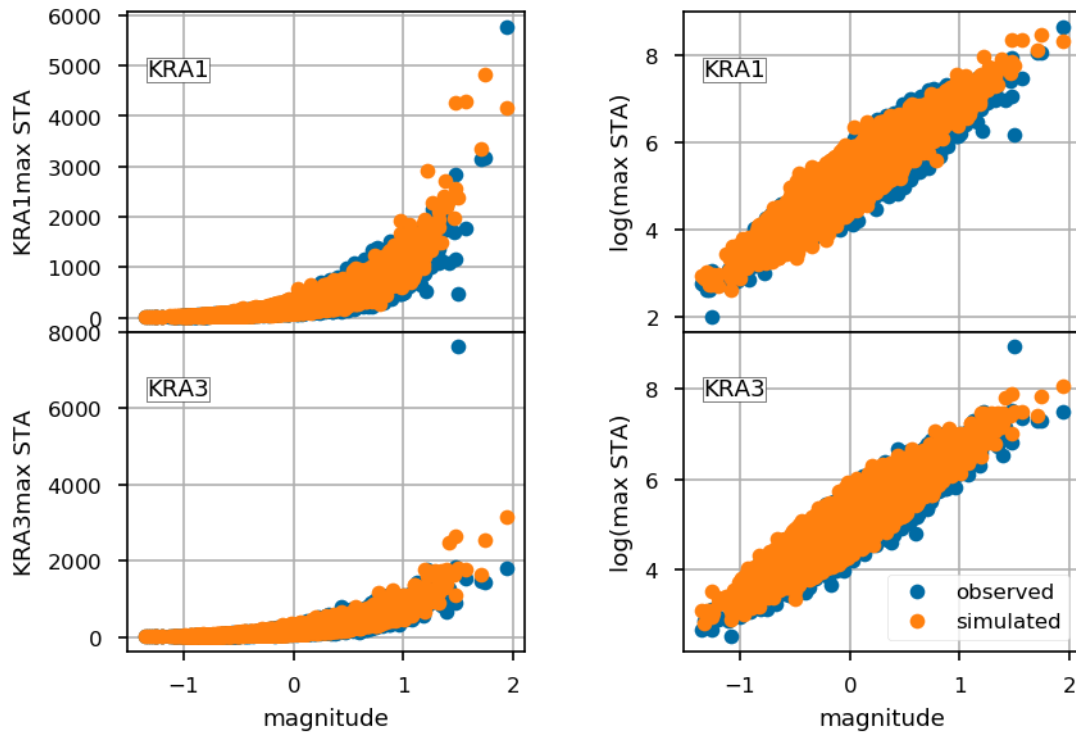
**Figure B.5.:** Linear relationship between observed magnitudes and maximum short-term mean squared amplitudes (STA). Averages were computed over 4 s long windows.

**Table B.4.:** Parameters and AICs for fitting different probability density functions to the simulated “true” distribution of magnitudes using MLE. Calculations were performed with scipy 1.4.1

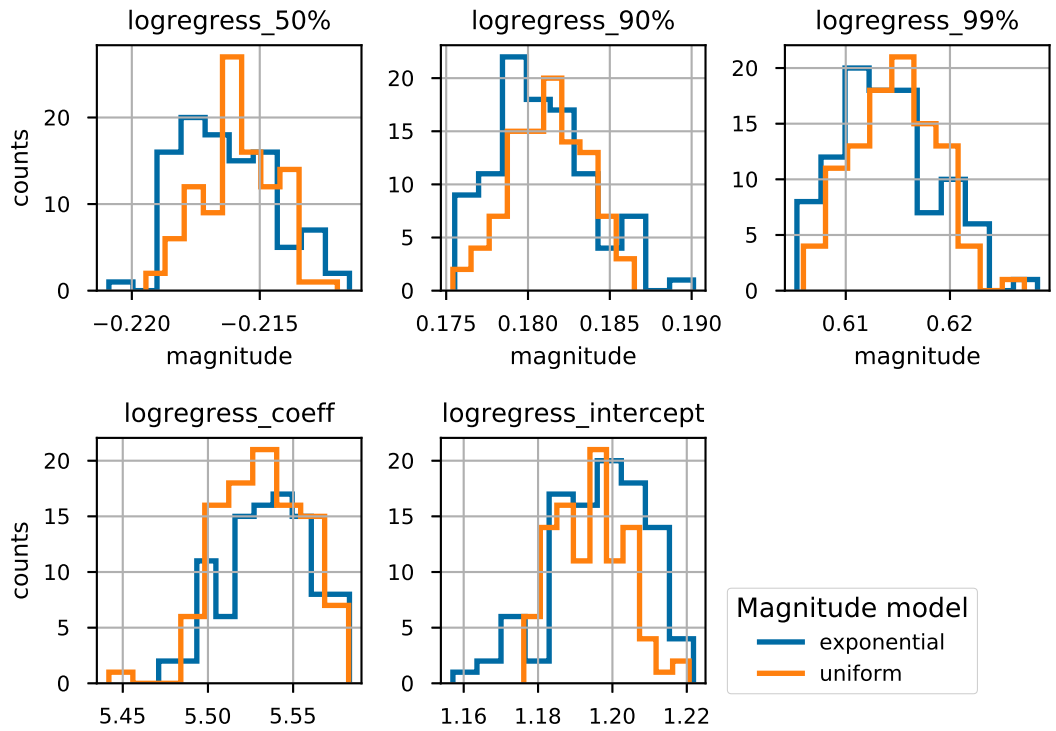
pdf	AIC	KS_pval	loc	scale	shape_0	shape_1
gengamma	68307.9	6.85e-208	-21.9	9.76	52.29	5.04
gompertz	67262.2	3.04e-28	-3.4	0.52	0.00	-
weibull	66256.3	1.00e-28	-4.8	4.53	8.64	-
gumbel	67430.9	5.83e-27	-0.2	0.52	-	-
genextreme	71234.9	5.77e-303	-0.8	0.63	0.23	-
invgauss	71977.5	0.00e+00	-6.2	467.14	0.01	-
norm	68717.0	1.75e-229	-0.5	0.59	-	-
loglogistic	107995.4	0.00e+00	-3646817375.3	3646817374.80	3923209659.94	-



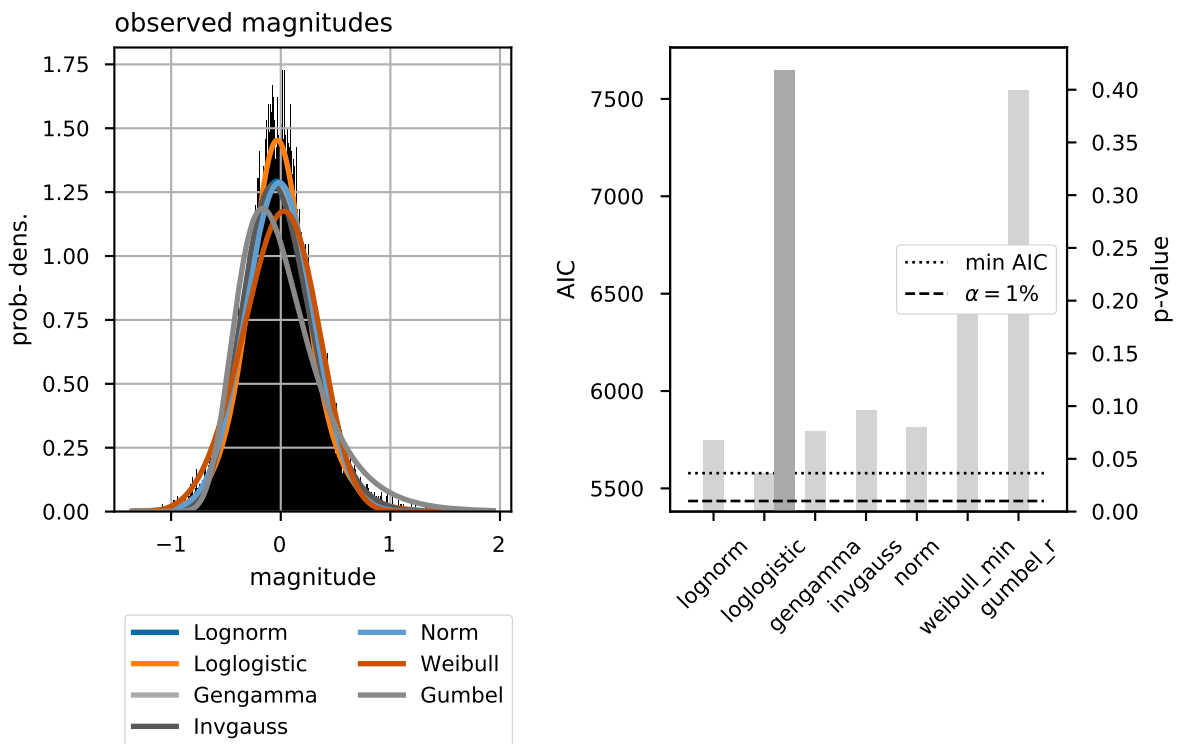
LF



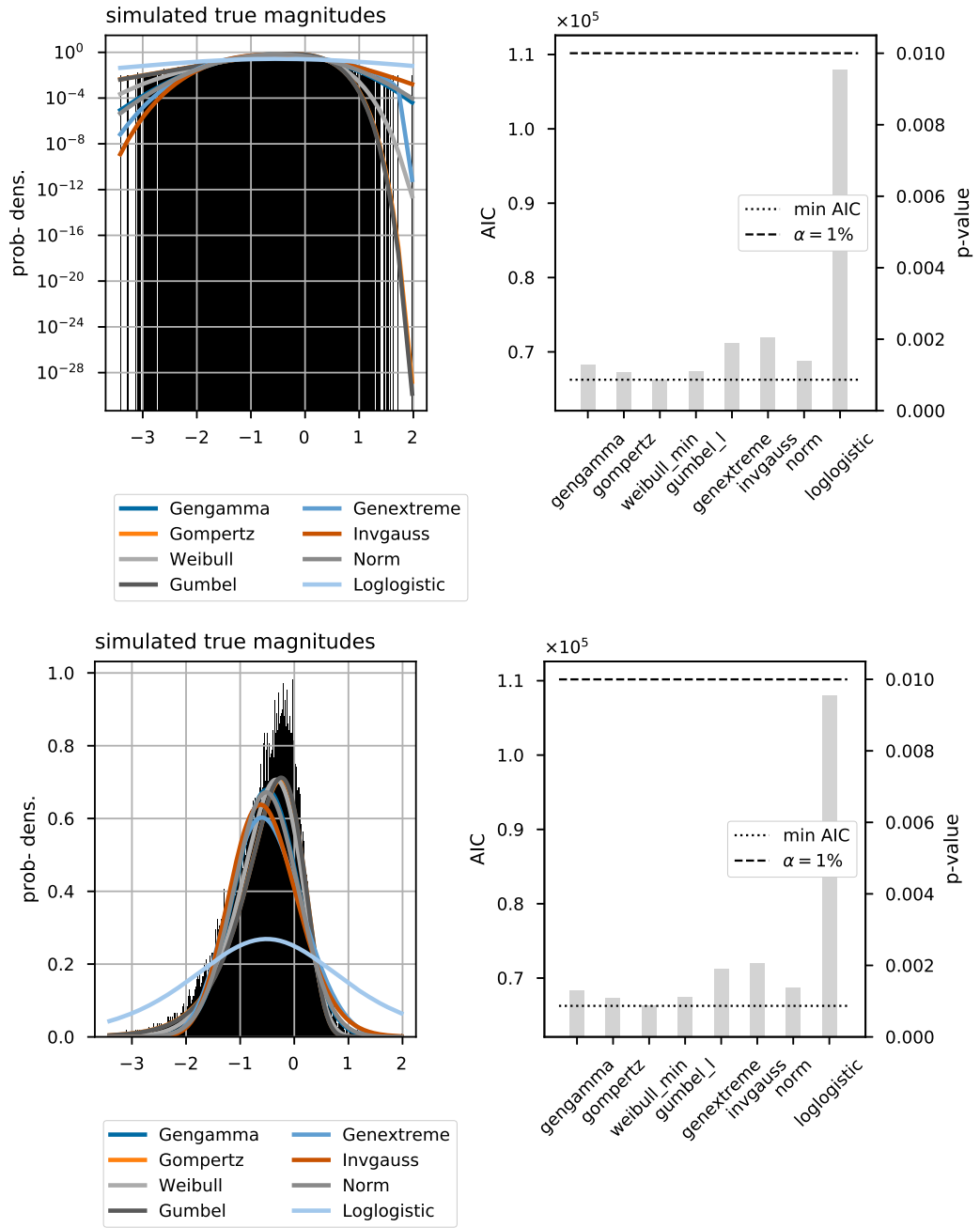
**Figure B.6.:** Simulated and actually observed STA values for observed magnitudes. The simulated STA values were obtained from the observed magnitudes using the linear relationship and its associated Gaussian error.



**Figure B.7.:** Magnitudes of completeness at different completeness levels from 100 simulations of 1 million events and coefficients of the logistic detection curves. Colors indicate the two tested magnitude distributions.



**Figure B.8.:** Results for maximum-likelihood estimation of probability density functions to model the observed magnitude distribution.



**Figure B.9.:** Fitting different probability density functions to the simulated “true” distribution of magnitudes using MLE. Same data in log (top) and linear (bottom) y-scale.

# C. Independent component analysis

## Introduction

Thanks to two concurrent measurement campaigns in March 2012 (Rabbel and Thorwart [2019], Waite [2010]), the seismicity was recorded by 6 instruments positioned at the rim of the central summit crater for 15 h simultaneously. These instruments surrounded the source, which is somewhere within the magma column of the lava lake and recorded the seismicity from different angles. Despite the seemingly similar conditions with respect to distance to the source and site the signals were surprisingly different in both time and frequency domain. This indicated a strong influence of site effects, possibly masking the common properties of the source signal. In the following, we use independent component analysis to separate the signal into 5 mixture components which brings out what could be the fundamental resonance frequency of the conduit system.

## Method

### Independent component analysis

Independent component analysis (ICA) is a method of blind signal separation. One assumes that a recorded signal is the (possibly non-linear) mixture of different source signals. With ICA, one attempts to separate these sources into statistically independent, non-Gaussian signals. For a proper reconstruction of the sources, at least as many recordings as sources are required. For the basic application of ICA, time differences between the recordings should be negligible.

The recorded signals  $x(t) = (x_1(t), \dots, x_M(t))$  is modeled as a mixture of the source signals  $s(t) = (s_1(t), \dots, s_N(t))$ :

$$x(t) = \underline{A} \cdot s(t)$$

ICA tries to recover the source signals by maximising the non-gaussianity of the input data

In volcano seismology, ICA has been used to separate volcanic tremor from oceanic micro-seismicity [Cabras et al., 2008] or isolate oscillation periods in volcanic explosions and tremor [Ciaramella et al., 2004, Martino et al., 2011].

In order to obtain a sufficient number of recordings two ways were established: 1) Using one station and repeating specimens of the signal or 2) Using multiple stations.

In the latter case the differences in arrival times at the stations should be negligible. Because of the favorable station set up at the crater, the second approach was used.

The FastICA algorithm (<https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.FastICA.html#sklearn.decomposition.FastICA>) of scikit-learn was used here, which obtains the source signal by maximizing their non-Gaussianity.

## Preprocessing

Processing of seismic data includes:

- Z-component
- instrument response removed
- filter: 0.45-16 Hz

Preprocessing for ICA includes demeaning and whitening of the data by rotation into orthogonal principal components and possibly a dimensionality reduction.

Principal component analysis (PCA) was applied to estimate the number of components. The relative weight of a PC (corresponding to the eigenvalue) can be interpreted as the amount of information that this component represents. Subsequently, FastICA was applied to the data including the preprocessing steps mentioned above.

For the results presented here, PCA and ICA were applied first to a 2-min long tremor sequence because of the much shorter computation time. The obtained mixing matrix was then applied to the whole 15-h sequence of data. The spectral properties of the independent components were almost identical for the sample and the long data sequence.

Finally, we compare the Welch power spectral densities of the original data and the ICs and analyze their frequency content.

## Quality control

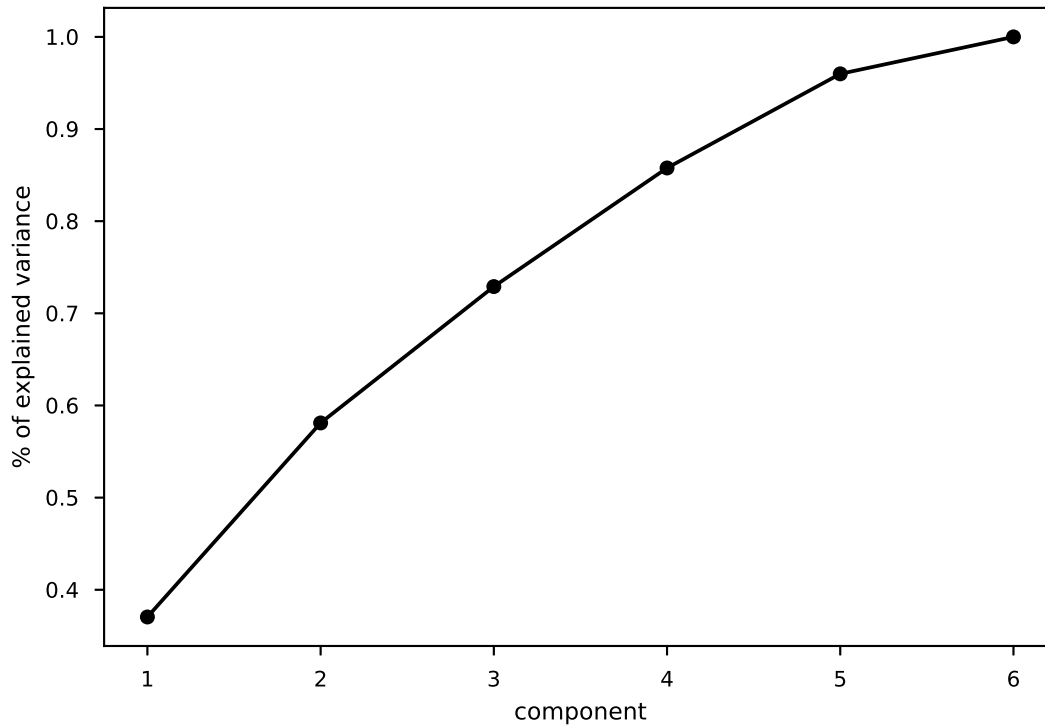
In order to verify, if the extracted independent components (ICs) are stable, the ICA was repeated 100 times and the obtained components matched to those of the initial ICA by correlation. The average normalized (Pearson) correlation coefficient between the corresponding components provided a measure of how reproducible the separation of the initial ICs was. Furthermore, the normalized correlation coefficients between the ICs and the original recordings were computed to check whether the ICs are truly new components and not basically identical with the original data.

## Results

For the 4-min-sample, the PCA indicated that 2 components provide already >70% of the information. Three and four components yield >80% and >90% of the information,

---

### Importance of principal components



**Figure C.1.:** Principle component analysis of 4-min-long tremor sample from the crater. The percentage of variance indicates how much of the data can be explained by the respective number of principal components.

respectively (Fig. C.1).

ICA was tested with 2-5 components.

- 2 ICs yielded 2 signals that were significantly different from the original data. The spectral peaks were around 3 and 8.9 Hz.
- with 3 ICs, one of them ends up being a copy of V220 most of the time
- With 4 IC, 2 are very different from all original traces and two are quite similar to KRA1 and KRA3.
- 5 ICs mostly end up being very similar to one of the stations. A minority solution is occasionally obtained in which the ICs are relatively different from the original data. The additional component has a peak around 6 Hz.

Figs. C.2 and C.3 show the results for 4 ICs.

- 2 spectral peaks at around 3 Hz and 8.6/9 Hz
- secondary peak at 6 Hz and perhaps 1 Hz

## Analysis and interpretation of spectral peaks

Assuming the lava lake reservoir is a resonating fluid-filled pipe and that  $f=3$  Hz is the resonance frequency  $f$ , its length  $L$  (i.e. the lake depth) would be

$$L = \frac{nv}{4f}$$

with  $v$  - velocity,  $n$  - number of harmonics,  $n = 1, 2, 3, \dots$  for pipe open or closed at both ends and  $n = 1, 3, 5, \dots$  for one open and one closed end.

For velocities of 1000-2000 m/s, the pipe would have a length of 75-150 m (Fig. C.4). Hence 3 Hz might be the natural resonance frequency of the lake reservoir. Interestingly, the presence of the 6-Hz peak indicates a pipe model with either two closed or open ends, rather than one open (the free magma surface) and one closed (the lake bottom) end. Possibly, the spatter roof - despite its usual opening - acts as lid.

Admittedly, it is very tempting to interpret the results in terms of such a simple model. However, it must be emphasized that the numerous authors showed that resonance effects in volcanoes are usually much more complex [Chouet, 1985, 1986, 1988, Neuberg, 2000, Sturton and Neuberg, 2003, 2006, Jousset et al., 2003]. The resonance, that is found in seismic data, is mostly the result of waves that live on the interface between solid conduit walls and the liquid magma inside the conduit. These interface waves may develop into standing waves and radiate energy into the medium mostly at the top and bottom end of the conduit. They propagate at a velocity that is slower than the acoustic velocity of the liquid. In contrast, in the resonating pipe model (organ-pipe model), a pressure perturbation travels as an acoustic wave within the fluid in the conduit. It was shown in numerical models [Jousset et al., 2003, 2004] that this wave hardly couples into the surrounding medium and is therefore unlikely to be detected seismically. An analytical solution exists for a fluid-filled infinite crack [Aki et al., 1977, Chouet, 1985, 1986, Ferrazzini and Aki, 1987] which is similar to the the Stonley waves found at fluid-filled bore holes [Biot, 1952]. Aki et al. [1977] found that the dynamics of the crack are controlled by the dimensionless stiffness  $C = \frac{bL}{\mu h}$  where  $b$  is the bulk modulus of the fluid,  $L$  and  $h$  are the length and with of the crack and  $\mu$  is the rigidity of the solid. From this parameter, it can be seen, that - at best - ratios about the medium properties and geometries can be obtained unless three of the quantities are well constrained.

In addition, the assumption of a homogeneous solid and liquid is most likely a gross oversimplification. The acoustic velocity and viscosity of the gas-magma mixture in



---

the conduit strongly depends on the gas content [Neuberg et al., 2000, Neuberg and O'Gorman, 2002, Sturton and Neuberg, 2003]. However, the amount of gas in the magma changes due to decompression, so assuming a constant velocity along the entire conduit is hardly realistic. Small differences in the assumed parameters such as density, gas fraction and viscosity can lead to significantly different seismic velocities and dimensions of the inferred conduit geometry. Due to this complex interplay and the resulting non-uniqueness of solutions, Neuberg et al. [2000] cautioned against inverting presumed resonance frequencies for structural geometries.

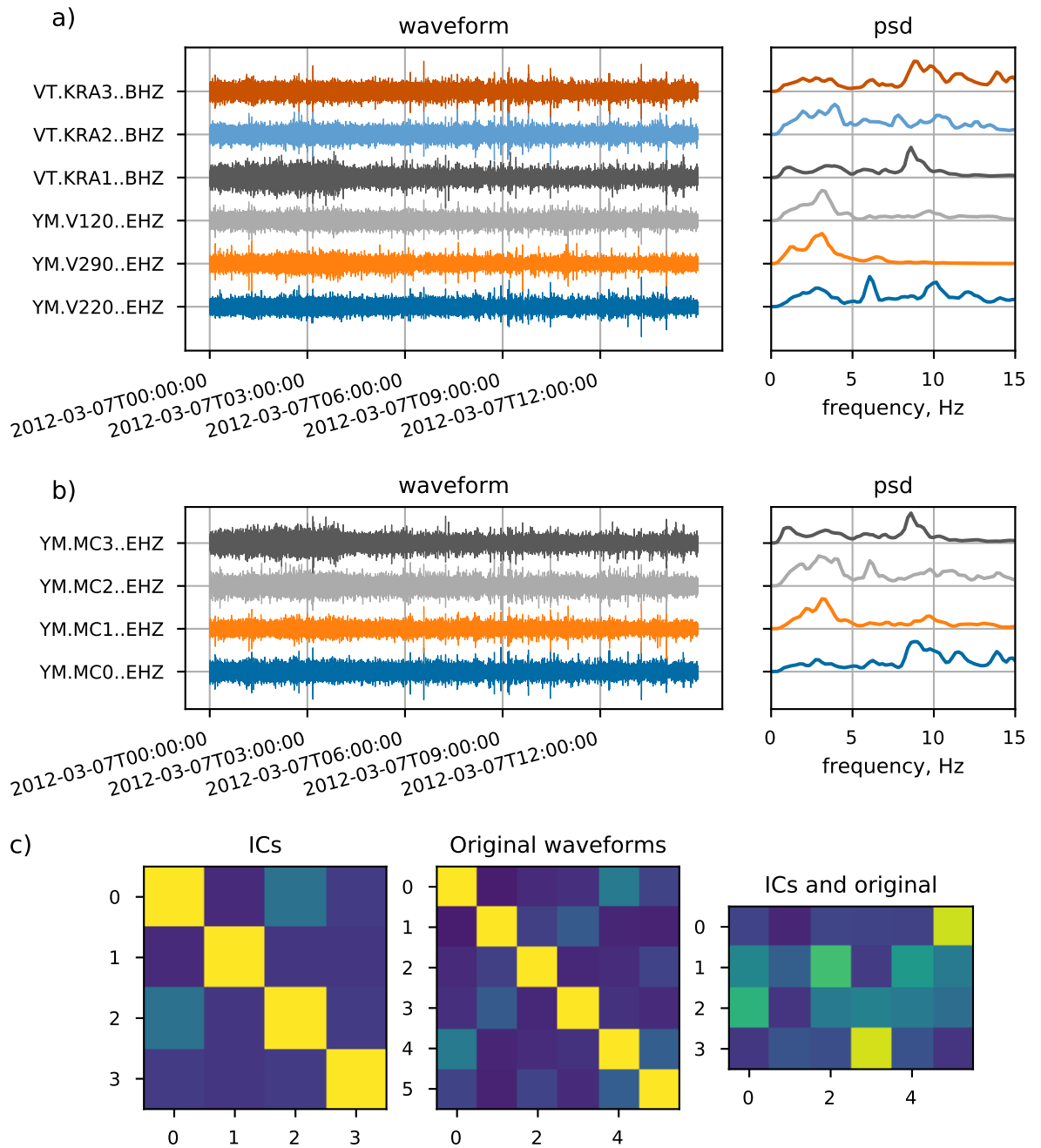
## Separation to transients and tremor

Another important question regarding the low-frequency seismicity of Villarrica is, whether the background tremor and the transient events are similar in nature or not. If they were different, it should be possible to separate the two signals using ICA. The results above showed however, that this was not the case for the stations at the crater rim. None of the independent components showed a clear preference for transient or tremor signals, respectively. Possibly, this is only because the stations at the crater were too close to the source of the signals or too much affected by site effects.

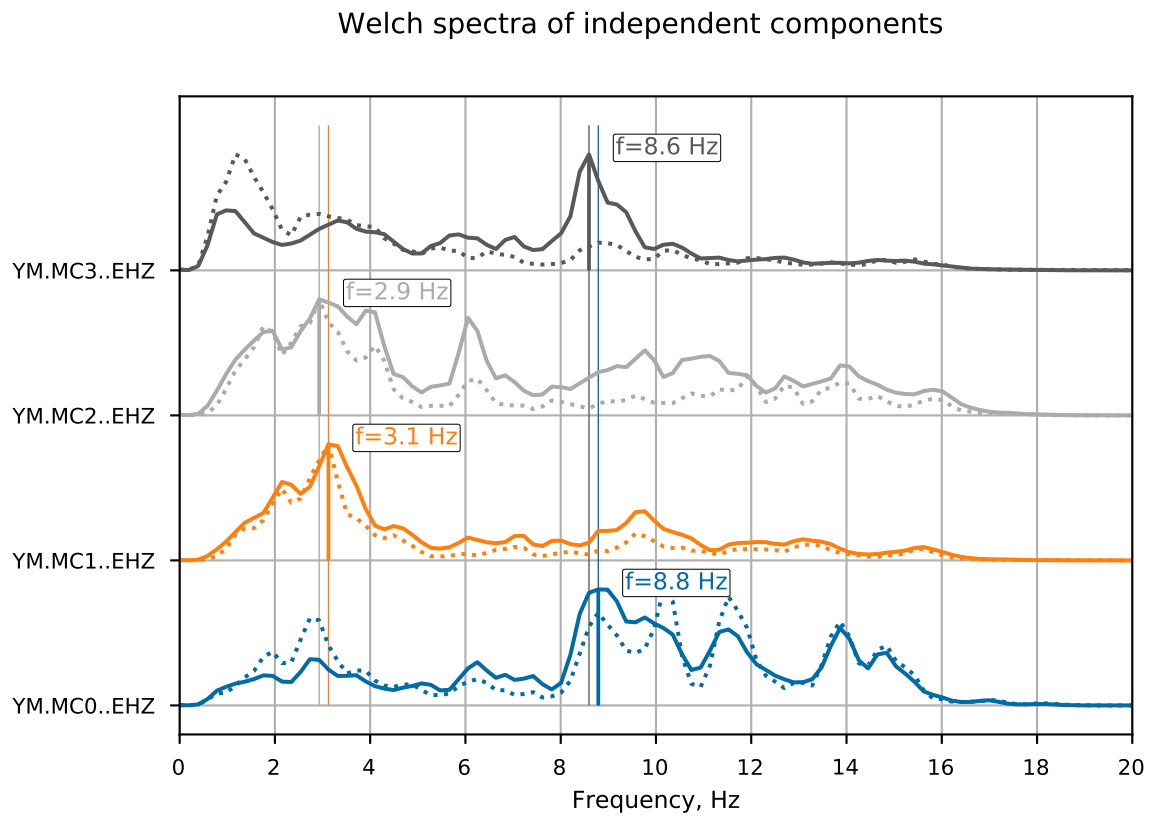
Therefore, the approach above was also applied using the 5 stations from array AVW, which is 4-5 km west of the summit crater. A 5 min-long data segment from 2012-03-08T15:00:00 was chosen, which displayed clearly separated transient events. Again 2 to 5 independent components were tested (Fig. C.5). The transients could always be identified on all components, suggesting that a separation into tremor and transients is not possible. The same occurs if only one event is used.

Nevertheless, the idea of the separation of signals could at least be successfully demonstrating for a sequence of three unknown, small signals recorded by array ALN (Fig. C.6).

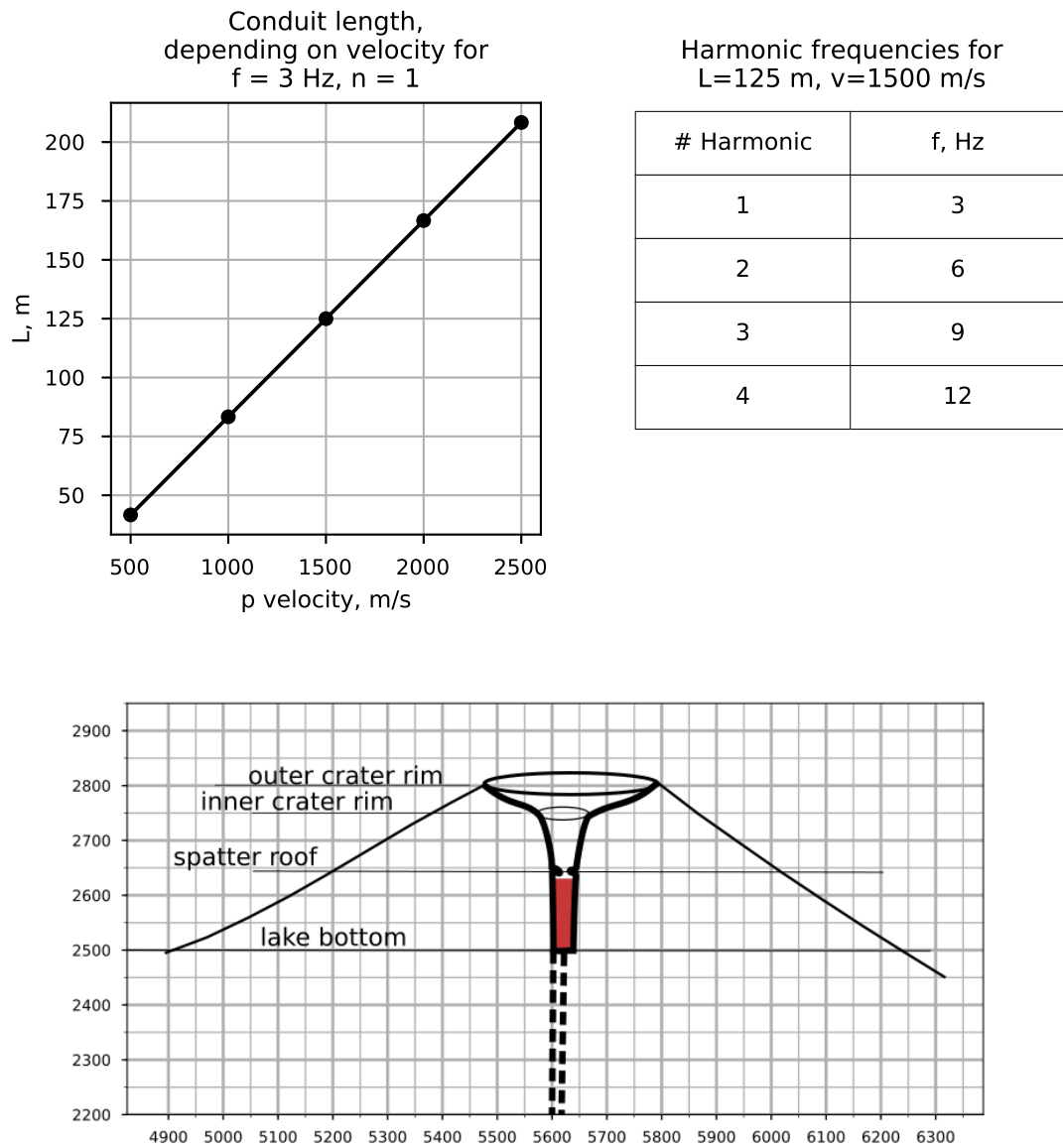
In the original data, these events are only discernible at some stations and even there they are quite mixed with the noise. Separation into 2 ICs however yielded one component with the events perfectly highlighted and another containing the background seismicity.



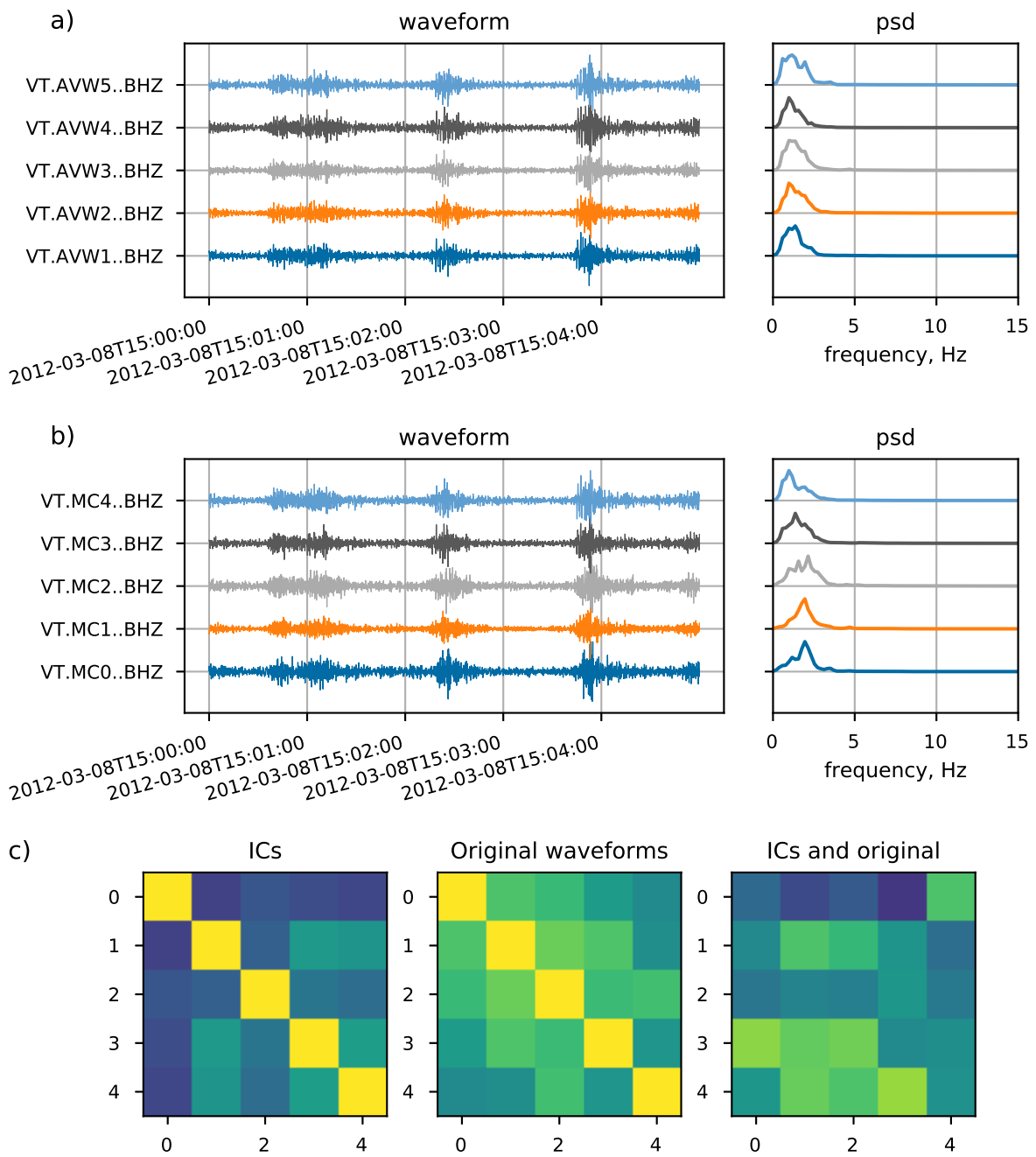
**Figure C.2.:** a) Original 15 h-signals and corresponding spectra b) independent components and spectra. The mixing matrix was determined from a 4 s-long sample. c) Correlation coefficients between traces. IC 0 and 3 correspond to stations KRA1 and KRA3



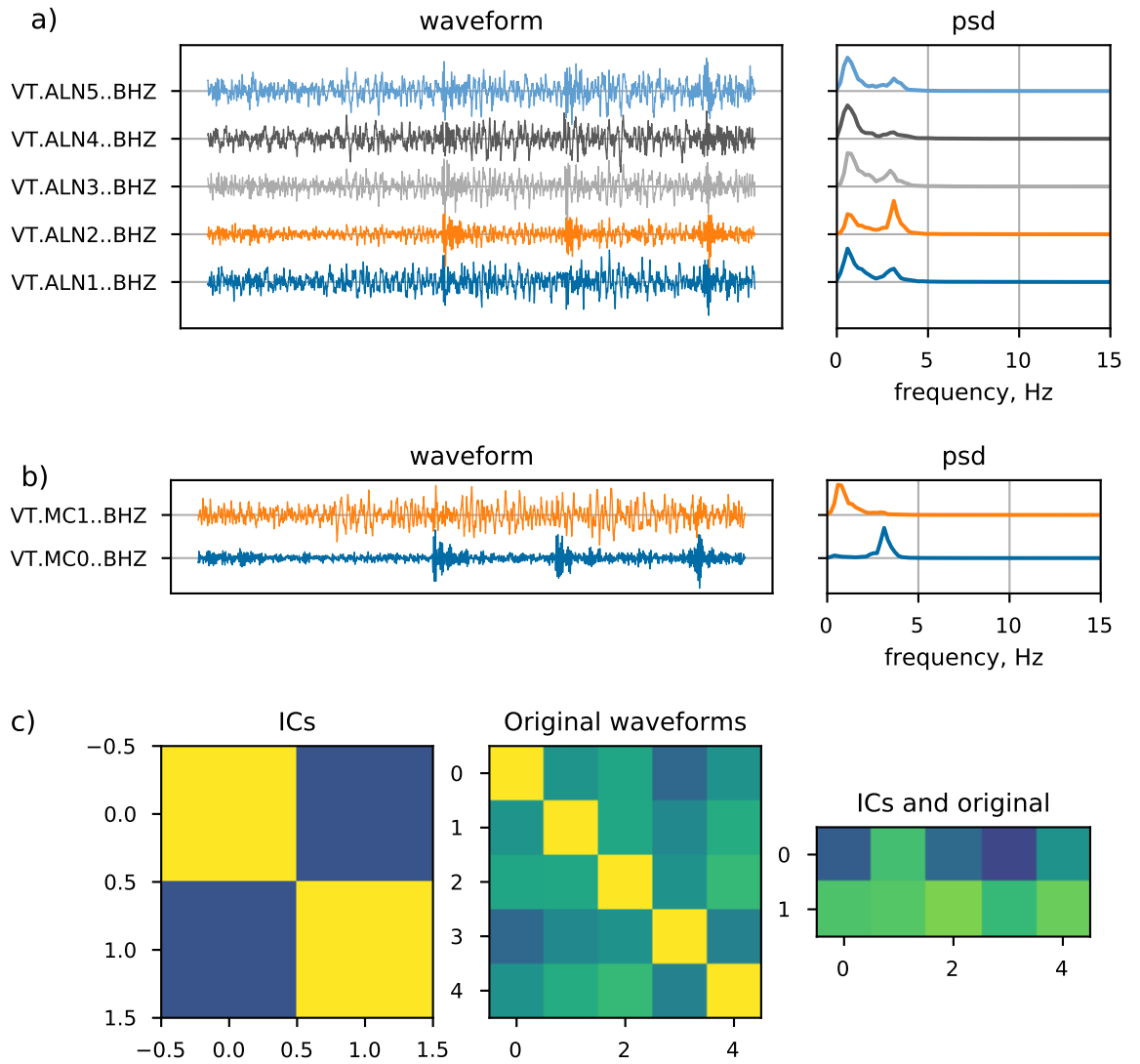
**Figure C.3.:** Power spectral density of ICs for the long data section. Dotted lines indicate spectra of the 4 min-sample. IC 0 and 3 are very similar to the original data of stations KRA1 and KRA3. Spectral peaks are found around 3 and 9 Hz. These peaks are also present if only 2 ICs are used. Secondary peaks emerge around 1-1.5 Hz, 6 Hz



**Figure C.4.:** *top:* Pipe lengths depending on acoustic velocity of the containing fluid and harmonic frequencies for a simple model of a resonating pipe. *bottom:* Sketch of the crater and upper conduit geometry based on Moussallam et al. [2016], Richardson and Waite [2013], Richardson et al. [2014]. Coordinates are in meters. Red color indicates the resonating reservoir of the lava lake.



**Figure C.5.:** ICA applied to a 5 min-long segment containing three transient events at array AVW. Even with 5 ICs the transients appear on all components, hence a separation into tremor and transient events was not possible.



**Figure C.6.:** ICA applied to a series of small signals at array ALN. While in the original data, the events are mixed into the noise, they appear clearly on the independent component.



# Johanna Lehr

## Curriculum Vitae

---

### Personal information

Last Name LEHR  
First Name Johanna Marijke  
Date of birth 12 november 1988  
Place of birth Leiden, The Netherlands  
Nationality German

---

### Research interests

My current interests centers around the geophysical investigation of volcanoes with a special focus of volcano seismology. Seismology provides a window into the internal structure of the volcanic edifice as well as into the dynamics of the volcanic activity. Furthermore, I am generally interested in computational modeling of physical processes and I enjoy programming.

---

### Education

- 2012-2016 **Msc. Geophysics**, *Kiel University, Kiel.*  
Thesis: *Location and polarization of long-period volcanic signals at Villarrica Volcano (Chile)*  
Supervisor: Prof. Wolfgang Rabbel
- 2008-2012 **Bsc. "Physics of the Earth system"**, *Kiel University, Kiel.*  
Thesis: *Interaction of gravitational and slope parallel sediment transportation processes at the continental slope off Uruguay - deduced from seismic data (in German)*  
Supervisor: Prof. Sebastian Krastel
- 1999-2007 **German-French university entrance diploma (Abitur/Baccalauréat)**, *Deutsch-Französisches Gymnasium/Lycée Franco-Allemand, Saarbrücken.*

---

### Research experience

- since 2017 **PhD student**, *Inst. of Geosciences, Kiel University.*  
Low-frequency seismicity of Villarrica Volcano, Chile
- 2016-2017 **Research assistant**, *Inst. of Geosciences, Kiel University, EU H2020 project Descramble.*  
Modeling of seismic reflectors
- 2014-2015 **Student research assistant**, *Inst. of Geosciences, Kiel University.*
- Processing of reflection seismic data
  - Modeling of seismic reflectors using a propagator algorithm (qseis)



Aug./Sept. 2013 **Participation in geophysical field campaign**, *Inst. of Geosciences*, Kiel University.  
Survey of various archaeological targets in Turkey by GPR, magnetic and electrical methods

---

## Conferences

- Posters **Lehr, J.**, Rabbel, W. and Bredemeyer, S. (2021). Inferring a shallow degassing model for Villarrica Volcano from seismic explosion signals and SO<sub>2</sub> flux. Presented at 81. Jahreshauptversammlung der Deutschen Geophysikalischen Gesellschaft.
- Lehr, J.** and Rabbel, W. (2020). Characteristics of seismic activity of Villarrica Volcano. Presented at EGU General Assembly.
- Lehr, J.**, Eckel, F., Seidel, A., Rabbel, W. and Thorwart, M.(2018). Seismicity at Villarrica Volcano (Chile): Characteristics, origin and implications on wave propagation. Presented at EGU General Assembly.
- Lehr, J.**, Bredemeyer, S., Rabbel, W. and Thorwart, M.(2018). Inter-event times of volcanic earthquakes in comparison with gas emission events of Villarrica (Chile). Presented at EGU General Assembly.
- Talks **Lehr, J.**, Rabbel, W. and Thorwart, M.(2019). Low-frequency seismicity at Villarrica Volcano: Results from a dense local network. Presented at 79. Jahreshauptversammlung der Deutschen Geophysikalischen Gesellschaft.

---

## Original publications in peer-reviewed journals

- Lehr, J.**, Rabbel, W. (2021). Magnitude and interevent time statistics of Strombolian activity of Villarrica Volcano and inference regarding the flow regime. *Journal of Geophysical Research: Solid Earth*. In Reply.
- Lehr, J.**, Eckel, F., Thorwart, M. and Rabbel, W. (2019). Low-frequency seismicity at Villarrica Volcano: Source location and seismic velocities. *Journal of Geophysical Research: Solid Earth*. doi.org/10.1029/2018JB017023

---

## Scholarships

2017-2019 **PhD scholarship of Kiel University.**

Kiel, April 12, 2021