Towards accurate monitoring of explosive volcanoes using seismic noiseinterferometry

Alexander Yates

UNIVERSITÉ SAVOIE MONT BLANC

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ SAVOIE MONT BLANC

Spécialité : **Sciences de la Terre et de l'Environment**

Arrêté ministériel : 25 Mai 2016

Présentée par

Alexander YATES

Thèse dirigée par **Philippe LESAGE** et
codirigée par **Corentin CAUDRON, Aurélien MORDRET**

préparée au sein du **l'Institut des Sciences de la Terre**
dans **l'École Doctorale Sciences de la Terre de l'Environment et des Planètes**

Vers une surveillance précise des volcans explosifs par interférométrie du bruit sismique

Towards accurate monitoring of explosive volcanoes using seismic noise interferometry

Thèse soutenue le 27/11/2023 devant le jury composé de :

Philippe Lesage Maître de Conférences Émérite, Université Savoie Mont Blanc	Directeur de thèse
Jean-Luc Got Professeur, Université Savoie Mont Blanc	Président du Jury
Gregor Hillers Professeur, University of Helsinki	Rapporteur
Silvio De Angelis Professeur, University of Liverpool	Rapporteur
Alicia Hotovec-Ellis Research Geophysicist, USGS California Volcano Observatory	Examinatrice
Léonard Seydoux Assistant Professor, Institut de Physique du Globe de Paris (IPGP)	Examinateur
Corentin Caudron Assistant Professor, Université Libre de Bruxelles	Membre Invité
Aurélien Mordret Research Scientist, Université Grenoble Alpes	Membre Invité

Abstract

The technique of seismic noise interferometry is gaining popularity for monitoring volcanoes. By utilizing the cross-correlation of continuously recorded ambient noise, changes in seismic velocity can be recovered that provide insight into the behavior of volcanic systems. However, applying coda wave interferometry to volcanoes presents unique challenges due to the variable nature of the seismic wavefield. Furthermore, the sensitivity of changes to both internal (volcanic) and external (environmental) forcings complicates the interpretations. This thesis explores these challenges, focusing on explosive volcanoes, which are often more complex than effusive ones.

The first objective of the thesis is to characterize seismic datasets by understanding the seismic wavefield coherence through network covariance matrix analysis and determining the characteristics of cross-correlation functions in response to observed features of the seismic wavefield. This analysis is carried out at five volcanoes with varied eruptive activity to better constrain the expectations of seismic datasets intended for coda wave interferometry.

Volcanic tremor regularly produced higher signal to noise ratio cross-correlation functions relative to non-tremor periods. This suggests potential for monitoring if the long-term stability can be verified. To this end, an unsupervised machine-learning approach of agglomerative hierarchical clustering is applied to cross-correlation functions. This allows data with similar characteristics to be grouped without requiring a defined reference period. The approach is tested at Piton de la Fournaise volcano (Réunion Island, France) and Mt Ruapehu (New Zealand), where distinct clusters related to changes in the seismic source and the medium following volcanic eruptions are identified. At Ruapehu, the method confirms relative stability in cross-correlation functions computed during individual tremor episodes.

The thesis then focuses on two case studies of explosive volcanoes: Mount Ruapehu and Stromboli (Italy). In both cases, volcanic tremor is used to compute seismic velocity changes at high frequencies (above 1 Hz).

At Ruapehu, analysis of the eruption chronology suggests a higher probability of small eruptions in spring. Measured seismic velocity changes show substantial seasonal variations that differ on and off the volcano. Analysis suggests an influence of snow cycles on the volcano, with reduced water infiltration during winter the most likely mechanism dominating. A smaller contribution from snow-loading/unloading is also suggested, and considered the most plausible trigger for spring-time eruptive activity. These findings highlight the need to examine further the relationships between external forcings and the modulation of eruptive activity towards improved

hazard assessment.

At Stromboli, velocity changes are computed over nine years encompassing effusive activity, major explosions, and paroxysms. Two paroxysms in the summer of 2019 are of particular interest, occurring two months apart with minimal warning. While seasonal processes dominate shallow velocity changes (1–4 Hz), they also reveal mid-to-long-term precursors in the years prior to the paroxysms. The closest station to the active crater shows a steady increase of 0.2–0.3% relative to other stations. This is accumulated over three years of heightened activity, peaking in the months before the first paroxysm. A long-term decrease is also observed in deeper velocity changes (0.5-1.0 Hz) during the same time interval. It is hypothesized that these changes represent increased magma overpressure from greater volatile input from depth. The different response in the shallow subsurface may reflect an anisotropic distribution of fractures and microcracks in proximity to the crater terrace. These findings illustrate how coda wave interferometry can provide meaningful insights into the evolving dynamics of open-conduit basaltic volcanoes.

Résumé

La technique d'interférométrie de bruit sismique gagne en popularité pour la surveillance des volcans. En utilisant le bruit ambiant enregistré en continu, il est possible d'estimer des changements de vitesse sismique, ce qui donne un aperçu du comportement des systèmes volcaniques. Cependant, l'application de l'interférométrie d'ondes de coda aux volcans présente des défis uniques en raison de la nature variable du champ d'ondes sismiques. De plus, la sensibilité des changements aux forçages internes et externes complique les interprétations. Cette thèse explore ces défis, en se concentrant sur les volcans explosifs, qui sont souvent plus complexes que les volcans effusifs.

Le premier objectif de la thèse est de caractériser un jeu de données sismiques en prenant en compte la cohérence du champ d'ondes sismiques grâce à l'analyse de la matrice de covariance du réseau et en déterminant les caractéristiques des fonctions de corrélation croisée. Cette analyse est réalisée sur cinq volcans à activité éruptive variée afin de mieux contraindre le potentiel des jeux de données sismiques destinés à l'interférométrie à ondes coda.

Les trémors volcaniques produisent régulièrement des fonctions de corrélation croisée dont le rapport signal sur bruit est plus grand que pendant les périodes sans trémor. Cela suggère un potentiel pour la surveillance si la stabilité à long terme peut être vérifiée. À cette fin, une approche d'apprentissage automatique non supervisée d'agglomération hiérarchique est appliquée aux fonctions de corrélation croisée. Cela permet de regrouper des données présentant des caractéristiques similaires. L'approche est testée au volcan Piton de la Fournaise (Île de la Réunion, France) et au Mont Ruapehu (Nouvelle-Zélande), où des « clusters » distincts liés aux changements de la source sismique et du milieu sont identifiés. Au Ruapehu, la méthode confirme la stabilité relative des fonctions de corrélation croisée calculées au cours des épisodes de trémor.

La thèse se concentre ensuite sur deux études de cas de volcans explosifs : le Mont Ruapehu et le Stromboli (Italie).

Au Ruapehu, l'analyse de la chronologie des éruptions suggère une probabilité de petites éruptions plus élevée au printemps. Les changements de vitesse sismique mesurés montrent des variations saisonnières substantielles qui diffèrent sur et hors du volcan. L'analyse suggère une influence des cycles nivologiques sur le volcan, la réduction de l'infiltration d'eau en hiver étant le mécanisme dominant le plus probable. Une contribution plus faible du chargement/déchargement de neige est également suggérée et considérée comme le déclencheur le plus plausible de l'activité

éruptive printanière. Ces résultats soulignent la nécessité d'examiner plus en détail les relations entre les forçages externes et la modulation de l'activité éruptive en vue d'une meilleure évaluation des risques.

Au Stromboli, les changements de vitesse sont calculés sur neuf ans englobant de l'activité effusive, des explosions majeures et des paroxysmes. Deux paroxysmes survenus en 2019 sont particulièrement intéressants, survenant à deux mois d'intervalle avec un minimum de précurseurs. Alors que les processus saisonniers dominent les changements de vitesse peu profonds (1 à 4 Hz), la station la plus proche du cratère actif présente une augmentation continue de 0,2 à 0,3 % par rapport aux autres stations. Ceci représente le cumul sur trois années d'activité intense, avec un pic dans les mois précédant le premier paroxysme. Une diminution à long terme est également observée dans les vitesses plus profondes (0,5 à 1,0 Hz) au cours du même intervalle de temps. On suppose que ces changements résultent d'une surpression accrue du magma due à un apport plus important de volatiles provenant des profondeurs. Ces résultats illustrent comment l'interférométrie d'ondes de coda peut fournir des informations significatives sur l'évolution de la dynamique des volcans basaltiques à conduit ouvert.

Acknowledgements

Firstly, I extend my thanks to the great team around me. My supervisors Corentin Caudron, Philippe Lesage and Aurélien Mordret, who supported me throughout this project. Also Virginie Pinel, who would certainly be an official co-supervisor if it wasn't for admin restrictions! I thank you all for the various insights and wisdom you've shared, and also the trust you had in me from the start of this project. It made for a highly enjoyable PhD experience.

Thanks also to the jury members Jean-Luc Got, Gregor Hillers, Silvio De Angelis, Alicia Hotovec-Ellis, and Léonard Seydoux (especially Gregor and Silvio who had the tough task of reading the thesis in four weeks!). Thanks also to Florent Brenguier and Jean-Luc (again) for participating in my thesis committee. A big thanks also to 'Le ministère de l'Enseignment supérieur, de la Recherche et de l'Innovation' (MESRI) who funded this PhD thesis.

There are also many others who have been involved in this project at various stages. I'd like to extend my thanks to our Italian colleagues (Andrea, Flavio, Lucia) at INGV and our New Zealand colleagues at Geonet (Oliver, Craig, Nico and Steve) who were involved in the two main case studies of this project. Also colleagues in Belgium at the Royal Observatory (Thomas, Raphael), and those with the ISTremor project (Kristin, Yesim), who hosted me for three weeks in Brussels and funded a visit to Iceland. Finally, Jean and Silvana for your aid with the Tungurahua dataset.

I also need to thank my friends and colleagues around Chambéry and Grenoble, who were the source of my day-to-day happiness during my three years in France. Thanks to my office mates, Pauline versions one and two, the ever-present Shan, Roger, and for a briefer, much colder, moment, Patricia. A big thanks to all the other folks around Chambéry, Camila, Louise, Anna, Léa, Anne-Clo, Kristina, Mathilde, Jean, Fabien, Baptiste, Jacques, Camille, Thibault. Camila, in particular, I appreciated immensely our little getaways during the final summer of my PhD! Less so always having to eat two meals because you over-ordered. And Louise, always a great source of late-afternoon/evening procrastination at the office. From Grenoble, Andres, Estelle, and Giuseppe. I'm sorry/not sorry to have given us all covid after AGU; though, the tasteless christmas eve fajitas in New Orleans are certainly a highlight of the PhD ;).

Thanks also to the GTM forum crew who provided immense support both at the start of my PhD during the covid lockdown and also in the final couple months when I was stuck inside writing. It's hard to imagine how i would have got through this period without seeing your faces online each day!

A big thank you also to my long-time friend Sascha, who was instrumental in the completion of my Masters in New Zealand. We may now be apart geographically, but our continued friendship and your endless support means the world to me.

And finally, the biggest thank you to my family, to whom I certainly owe many of the opportunities I have been afforded in life. Your constant encouragement and support to follow my goals and aspirations is more than I could have ever asked for. I am eternally grateful.

Contents

Abstract i

Résume ii

Acknowledgements v

List of Figures xii

List of Tables xxii

1 Introduction **1**

1.1 Motivation and Primary Objectives 1

1.2 Monitoring volcanoes through passive seismic interferometry 4

1.2.1 Extraction of medium properties using ambient noise 5

1.2.2 Coda-wave inteferometry for monitoring temporal changes . . 7

1.2.3 From raw data to seismic velocity changes 9

1.2.4 Application to volcanoes 14

1.3 Additional notes on thesis 17

1.3.1 Chapter contributions 17

1.3.2 Additional projects 19

2 Characterizing seismic datasets for coda wave interferometry at volcanoes: seismic sources and their influence on cross-correlation functions **21**

2.1 Introduction 21

2.2 Covariance matrix analysis of the seismic wavefield 22

2.2.1 Background theory 23

2.2.2 Data and processing 25
2.2.3 Testing different network configurations 28
2.3 Quantitative approaches to assessing cross-correlation quality 32
2.3.1 SNR vs time 32
2.3.2 Phase stack amplitude 34
2.3.3 Amplitude ratio 36
2.4 Characteristics of cross-correlation functions 36
2.4.1 Piton de la Fournaise 37
2.4.2 Kilauea 42
2.4.3 Mt. Ruapehu 46
2.4.4 Stromboli 53
2.4.5 Whakaari 58
2.5 Discussion 61
2.5.1 Influence of volcanic tremor 61
2.5.2 Influence of oceanic noise sources 63
2.5.3 Influence of eruptive activity 65
2.5.4 Considerations towards coda wave inteferometry 67
2.6 Conclusions 68
2.7 Supplementary Material 70

3 Assessing Similarity in Continuous Seismic Cross-Correlation Functions using Hierarchical Clustering: Application to Ruapehu and Piton de la Fournaise Volcanoes 75

3.1 Abstract 75
3.2 Introduction 76
3.3 Methods 79
3.3.1 Seismic data 79
3.3.2 Cross-correlation functions 79
3.3.3 Hierarchical clustering 80

3.3.4 Network covariance matrix spectral width 82
3.3.5 Seismic Velocity changes 84
3.4 Results 84
3.4.1 Piton De La Fournaise 85
3.4.2 Ruapehu 93
3.5 Discussion 97
3.5.1 Choice of algorithm and distance measure 97
3.5.2 Clustering cross-correlation functions as a similarity measure 102
3.5.3 Additional insights 105
3.5.4 Application in monitoring and future direction 107
3.6 Conclusions 109
3.7 Acknowledgements 110
3.8 Data Availability 110
3.9 Supplementary Material 111
S3.1 Single-station clustering 111
S3.2 Lower-distance-threshold Piton de la Fournaise (1.0–2.0 Hz) . 114

4 Seismic velocity changes over nine years at Stromboli volcano, Italy 117
4.1 Introduction 117
4.2 Methods 120
4.2.1 Seismic sources at Stromboli 121
4.2.2 Computing cross-correlation functions and velocity changes . . 123
4.2.3 Hierarchical clustering of cross-correlation functions 124
4.2.4 Complementary datasets 125
4.3 Results 126
4.3.1 Hierarchical clustering within tremor 126
4.3.2 Velocity changes 131
4.4 Discussion 137

4.4.1 Stability of cross-correlation functions in time 137
4.4.2 Seasonality in velocity changes 143
4.4.3 Deviating velocity at near-summit station (STRA) 147
4.4.4 Interpretation of velocity changes 150
4.4.5 Implications for monitoring 155
4.5 Conclusions 156
4.6 Supplementary Material 158
S4.1 Methods 158
S4.2 Results 163
S4.3 Variation of diffusivity through time 169

5 Exploring the influence of seasonal snow cycles on seismic velocity changes and eruptive activity at Mt Ruapehu, New Zealand 175

5.1 Introduction 175
5.2 Data and methods 177
5.2.1 Seismic data and cross-correlation function preparation 177
5.2.2 Tremor as a seismic source 178
5.2.3 Measuring velocity changes using wavelet transform 180
5.2.4 Complementary datasets 184
5.3 Results 185
5.3.1 Measured velocity changes 185
5.3.2 Location in depth 189
5.4 Discussion 196
5.4.1 Reliability of velocity changes computed from tremor 197
5.4.2 Mechanism of velocity changes 201
5.4.3 Possible links to eruptive activity 216
5.5 Conclusions 221
5.6 Supplementary material 223

S5.1 Methods 224
S5.2 Snow-load modeling 230
S5.3 Declustering of eruption catalog 234
S5.4 Statistical Analysis against non-uniformity 238

6 Discussion and Conclusions **245**

6.1 Introduction 245
6.1.1 Ruapehu: Real-time monitoring in 2022 247
6.1.2 Tungurahua 2014–2018 251
6.2 Coda wave inteferometry within tremor frequency band 258
6.3 External sources of seismic velocity changes 263
6.4 Future perspectives towards monitoring explosive volcanoes 267
6.4.1 Improved instrumentation 268
6.4.2 Improved methodologies 269
6.4.3 Targetted studies of meteorological processes in volcanic settings 270
6.4.4 Further multidisciplinary studies 270
6.5 Conclusions 271

List of Figures

1.1 Demonstration of temporal normalization techniques 10

1.2 Demonstration of spectral normalization techniques 11

2.1 Map of seismic network at Piton de la Fournaise 23

2.2 Covariance matrix spectral width at Piton de la Fournaise non-normalized vs normalized. 27

2.3 Comparison between spectrogram and spectral width measurement . 28

2.4 Spectral width measurements at Piton de la Fournaise using different network configurations. 30

2.5 Comparison between spectral width computed using two single-stations at Piton de la Fournaise. 31

2.6 Example of SNR computation for cross-correlation function 34

2.7 Example of phase stack amplitude computation from cross-correlation functions 35

2.8 Characteristics of cross-correlation functions for station-pair UV06-UV15 at Piton de la Fournaise. 38

2.9 Examining visual character of cross-correlations in response to changing SNR, phase similarity, and amplitude ratio. 39

2.10 Characteristics of cross-correlation functions for station-pair FJS-FOR at Piton de la Fournaise. 40

2.11 Characteristics of cross-correlation functions for single-station UVO5 at Piton de la Fournaise. 41

2.12 Phase similarity as a function of frequency and lag time at Piton de la Fournaise. 42

2.13 Map of seismic network at Kilauea volcano. 43

2.14 Spectral width measurements at Kilauea volcano 44

2.15 Characteristics of cross-correlation functions for station-pair HAT-OBL at Kilauea. 45

2.16 Phase similarity as a function of frequency and lag-time for station-pair HAT-OBL 46

2.17 Station map for Mount Ruapehu. 46

2.18 Spectral width measurements at Ruapehu. 48

2.19 Characteristics of cross-correlation functions for station-pair FWVZ-WPVZ at Ruapehu. 49

2.20 Phase similarity as a function of frequency and lag-time for station-pair FWVZ-WPVZ. 49

2.21 Phase similarity measurements for single-station FWVZ. 50

2.22 Amplitude ratios for single-station FWVZ. 51

2.23 Characteristics of cross-correlation functions for station-pair NGZ-TUVZ at Ruapehu. 52

2.24 Comparison of phase similarity for different station-pairs at different frequencies during tremor. 53

2.25 Seismic network at Stromboli 54

2.26 Spectral width measurements at Stromboli 55

2.27 Characteristics of cross-correlation functions for station-pair IST3-ISTR at Stromboli 56

2.28 Average phase similarity with time at low and high frequencies 57

2.29 Phase similarity as a functon of frequency and lag-time at Stromboli . 57

2.30 Seismic network at Whakaari volcano. 58

2.31 Spectral width measurements at Whakaari 59

2.32 Measurements of phase similarity at single-station WIZ. 60

2.33 Phase similarity as a function of frequency and lag-time for single-station WIZ. 60

2.34 Tests for convergence of cross-correlation functions at five volcanoes. . 65

S2.1 Example of SNR computation with normalization of cross-correlation functions. 71

S2.2 Spectral width at Stromboli during two-month period in 2019. 72

S2.3 Amplitude ratios for single-station UV05 at Piton de la Fournaise. . . 72

S2.4 Spectral width measurement between two stations at Ruapehu. 73

S2.5 Spectral width measurement between four broadband stations at Mount Ruapehu in 2022. 73

3.1 Station map at Piton de la Fournaise and Mt Ruapehu 78

3.2 Demonstration of hierarchical clustering approach. 82

3.3 Cluster output using 0.1–1.0 Hz cross-correlation functions (CCFs) for station-pair UV05-UV12 at Piton de la Fournaise 87

3.4 Clustering output using 0.1–1.0 Hz cross-correlation functions (CCFs) at station-pairs UV01-UV07. 88

3.5 Clustering single-station cross-component correlation functions computed using station UV07 at Piton de la Fournaise (0.1–1.0 Hz). . . . 89

3.6 Clustering output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pair UV05-UV12 at Piton de la Fournaise. 91

3.7 Clustering output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pairs UV01-UV07. 92

3.8 Clustering single-station cross-component correlation functions computed using station UV12 at Piton de la Fournaise (0.1–1.0 Hz). . . . 92

3.9 Clustering output using 0.2–0.7 Hz cross-correlation functions (CCFs) for station-pair FWVZ-WPVZ at Ruapehu. 94

3.10 Clustering output using 1.8–4.0 Hz cross-correlation functions (CCFs) for station-pair FWVZ-WPVZ at Ruapehu. 96

3.11 Summary of the cluster positions in time at both Piton de la Fournaise and Ruapehu. 97

3.12 Different clustering algorithms applied to group 0.1–1.0 Hz cross-correlation functions computed from Piton de la Fournaise seismic data. 98

3.13 Different clustering algorithms applied to group 1.0–2.0 Hz cross-correlation functions computed from Piton de la Fournaise seismic data. 99

3.14 Synthetic cross-correlation creation demonstration. 100

3.15 Demonstration of different distance measures used in clustering when velocity change is introduced to synthetic cross-correlation functions (1.0–2.0 Hz). 101

3.16 Demonstration of different distance measures used in clustering low SNR cross-correlation functions when velocity change is introduced to synthetic cross-correlation functions (1.0–2.0 Hz). 102

S3.1 Clustering single-station auto-correlation functions computed using station UV07 at Piton de la Fournaise (0.1–1.0 Hz). 111

S3.2 Clustering single-station auto-correlation functions computed using station UV12 at Piton de la Fournaise (0.1–1.0 Hz). 112

S3.3 Clustering 0.5–1.5 Hz cross-correlation functions computed for station pair FWVZ-WPVZ at Ruapehu. 113

S3.4 Clustering output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pairs UV01-UV07 (a-d) and UV07-UV09 (e-h), with lower distance thresholds. 114

S3.5 Comparison between using 1-day (a–c) and 10-day (d–f) stacks for clustering 0.1–1.0 Hz cross-correlation functions. 114

S3.6 Comparison between using 1-day (a–c) and 10-day (d–f) stacks for clustering 1.0–2.0 Hz cross-correlation functions. 115

4.1 Map of stromboli volcano, with broadband seismic stations used in this study shown. 119

4.2 Activity at Stromboli over nine years. 120

4.3 Covariance matrix spectral width over nine years at Stromboli. 122

4.4 Average spectral width split into seasons. 123

4.5 Hierarchical clustering output for 1.0-2.0 Hz cross-correlation functions (CCFs) computed for station-pair IST3-ISTR. 127

4.6 Hierarchical clustering output for 2.0-4.0 Hz cross-correlation functions (CCFs) computed for station-pair IST3-ISTR. 128

4.7 Comparing time and frequency domain character of average cross-correlation functions within three main clusters at 1–2 Hz and 2–4 Hz. 129

4.8 Comparing time and frequency domain character of average cross-correlation functions for individual years within main 2015–2019 cluster. 130

4.9 Hierarchical clustering output for 0.5-1.0 Hz cross-correlation functions (CCFs) computed for station-pair IST3-ISTR. 131

4.10 Velocity changes computed in three different frequency bands using stretching technique. 133

4.11 Spectra of velocity changes (stretching technique) at different frequencies. 133

4.12 Smoothed 0.5–1.0 Hz velocity changes for both single-station and station-pairs (stretching technique). 134

4.13 Single-station and station-pair seismic velocities (stretching technique) for each year plotted against calendar day. 135

4.14 Comparing 2.0–4.0 Hz velocities (stretching technique) recorded at STRA station with other four stations. 136

4.15 Closer look at relative difference between 2–4 Hz velocity changes at STRA station and other four stations (IST3, ISTR, STR1, STRE). . 137

4.16 Comparison of spectral width measurements over five individual months. 139

4.17 Comparing time and frequency domain character of average cross-correlation functions for main 2015–2019 cluster with smaller cluster during heightened activity within same time period. 141

4.18 Single-station velocity changes for three stations closest to summit (STR1, STRA, STRE) from period of renewed activity in 2017 to post-2019 paroxysms. 142

4.19 Comparing cross-correlation functions computed during periods of heightened activity at 1–2 Hz. 143

4.20 Single-station velocity changes at 2.0–4.0 Hz compared with different meteorological data. 144

4.21 Single-station velocity changes at 0.5–1.0 Hz compared atmospheric pressure changes. 145

4.22 Comparison of observed single-station velocity changes and synthetic changes. 146

4.23 Lateral sensitivity kernels based on 10 s central lag time in the coda. 149

4.24 Comparison of low frequency and high frequency velocity changes. . . 151

S4.1 Comparison of spectral width measurements before and after normalizing by the maximum value. 158

S4.2 Zoom of spectral width measurements computed during two months in 2019. 158

S4.3 Comparison of temperature and sea level pressure measurements at Stromboli with those from other meteorological stations. 159

S4.4 Measurements of cross-correlation characteristics at Stromboli for a single station-pair (IST3-ISTR). 160

S4.5 Phase similarity as a function of frequency and lag-time for station-pair (IST3-ISTR). 161

S4.6 Hierarchical clustering output for 1.0-2.0 Hz cross-correlation functions (CCFs) computed for station-pair IST3-ISTR. 162

S4.7 Comparison of velocity changes computed using full reference period versus reference period corresponding to main cluster between 2017–2019. 163

S4.8 Velocity changes computed in three different frequency bands using moving-window cross-spectral technique. 164

S4.9 Comparing 2.0–4.0 Hz velocities (using moving-window cross-spectral technique) recorded at STRA station with other four stations. 165

S4.10 Comparing 1.0–2.0 Hz velocities (stretching technique) recorded at STRA station with other four stations. 166

S4.11 Closer look at relative difference between 1–2 Hz velocity changes at STRA station and other four stations (IST3, ISTR, STR1, STRE). . 167

S4.12 Sensitivity kernels for Rayleigh waves between 0.5–4.0 Hz. 167

S4.13 Comparing velocities measured at STRA station and all other stations at different lag times in the coda. 168

S4.14 Comparison between lateral sensitivity kernels computed using 10 s central lag time and sensitivity kernels using 20 s central lag time. . . 169

S4.15 Normalized velocity changes (single-station, 2–4 Hz) during four different one-year windows (black-line) plotted with normalized pore pressure curves for three different values of diffusivity c. 170

S4.16 Diffusivity c estimated in one-year moving window comparing single-station velocity changes (2–4 Hz) with pore pressures. 171

S4.17 Comparison of the optimal diffusivity c and corresponding residual computed for 2–4 Hz single-station velocities with different detrending options 172

S4.18 Comparison of the optimal diffusivity c and corresponding residual computed for 2–4 Hz station-pair velocities with different detrending options. 173

5.1 Map of Ruapehu compared, snow cover, and eruption history by month. 177

5.2 Hierarchical clustering output for 1.0-4.0 Hz cross-correlation functions (CCFs) computed for station-pair MTVZ-TRVZ. 179

5.3 Comparison of the averaged cross-correlation functions for individual clusters. 181

5.4 Example of delay-time measurements using wavelet transform for a cross-correlation function during coherent tremor (2007-03-01) at DRZ station. 183

5.5 Snow coverage and snow thickness datasets plotted against months of the year. 185

5.6 Rainfall measurements recorded at the Chateau station and corresponding mean pore pressure changes. 186

5.7 Velocity changes measured at two stations located close to summit (DRZ and FWVZ) and one station at distance from the volcano (MWVZ). 186

5.8 Averaged velocity changes between 1–2.5 Hz for seismic stations DRZ, FWVZ, and MTVZ. 187

5.9 Velocity changes for all nine stations plotted against calendar day. . . 188

5.10 Comparison of the timing of seasonal velocity high (observed at 2 Hz, black lines) against station. 189

5.11 Rayleigh wave depth sensitivity kernels (fundamental mode) for selected frequencies. 190

5.12 L-curve plots for various choices of regularization parameters. 192

5.13 Comparison between observed (black points) and predicted (red curve) velocity changes (based on resolved model parameters) for select days at DRZ station. 193

5.14 Comparison between observed (black points) and predicted (red curve) velocity changes (based on resolved model parameters) for select days at FWVZ station. 193

5.15 Resolved velocity changes with depth compared with restitution index.195

5.16 Summary of inversion solution for upper 500 m at DRZ and FWVZ stations. 196

5.17 Comparison of yearly velocity changes at DRZ, FWVZ, and MTVZ for different minimum values of yearly coherence. 198

5.18 Comparison of yearly velocity changes at DRZ, FWVZ and MTVZ after removing individual years from average. 199

5.19 Yearly amplitude ratio at MTVZ-TRVZ 200

5.20 Model inputs for axisymmetric snow load (with sigmoid shape) applied at the surface. 204

5.21 Pressure change calculations at different radius r between June 20 and October 21 for sigmoid-shaped snow-load and the estimated velocity changes. 205

5.22 Model parameters used in converting pressure changes with depth to shear velocity changes. 206

5.23 Observed and depth-inverted velocity changes compared with estimated velocity changes from sigmoid shaped snow-load at DRZ and FWVZ station. 208

5.24 Comparison between velocity changes recorded at FWVZ station and temperatures recorded at nearby Whakapapa ski station. 209

5.25 Comparison between velocity changes recorded at DRZ and snow thicknesss at both Whakapapa and Turoa ski stations. 212

5.26 Comparison between measurements made using dynamic coda window and static lag times windows. 213

5.27 Classification of de-clustered eruptions using descriptions from catalog of Scott (2013). 218

5.28 Pressure change calculations at different radius r between June 20 and October 21 for sigmoid-shaped snow-load and corresponding pressure gradient for depth. 219

5.29 Violin plot of SO2 flux by month recorded at Ruapehu. 220

5.30 Eruptions from de-clustered catalog through time. 221

S5.1 orrelogram of station-pair MTVZ-TRVZ (top) and FWVZ-WPVZ (bottom). 224

S5.2 Hierarchical clustering output for 1.0-4.0 Hz cross-correlation functions (CCFs) computed for station-pair FWVZ-WPVZ. 225

S5.3 Example of delay-time measurements using wavelet transform for a cross-correlation function during period of reduced tremor (2006-07-09) at DRZ station. 226

S5.4 Example of delay-time measurements using wavelet transform for a cross-correlation function during period where tremor characteristics differ from longer-term average (2008-01-01) at DRZ station. 227

S5.5 Average yearly coherence measured at each station. 228

S5.6 Average yearly velocity changes at DRZ, FWVZ, and MTVZ stations after removing eruption years. 228

S5.7 Slope measurements along a N-S and W-E profile at Mt Ruapehu, using Google Earth. 230

S5.8 Model inputs for axisymmetric snow load (with uniform shape) applied at the surface. 230

S5.9 Pressure change calculations at different radius r between June 20 and October 21 for uniform shaped snow-load. 231

S5.10 Model inputs for axisymmetric snow load (with cone shape) applied at the surface. 232

S5.11 Pressure change calculations at different radius r between June 20 and October 21 for cone shaped snow-load. 233

S5.12 All events of magnitude 0-5 at Mt Ruapehu in catalog of Scott (2013). 235

S5.13 Cumulative count of events of different magnitude, after removing 1945 and 1995/96 eruptive activity following Scott (2013). 236

S5.14 All events of magnitude 2-5 at Mt Ruapehu in catalog of Scott (2013), excluding events pre-1960 and removing 1995/96 eruption sequences. 236

S5.15 Number of eruptions plotted again de-clustering interval for magnitude 2+ events in Scott (2013) eruption catalogue, after removing both events pre-1960 and the 1995/96 eruption sequence. 237

S5.16 De-clustered catalog of magnitude 2+ events at Mt Ruapehu from catalog of Scott (2013), excluding events during 1995/96 eruption sequence, and any events occurring within 14-days of a previous eruption. 238

S5.17 De-clustered catalog of events at Mt Ruapehu from catalog of Scott (2013) (see Section S5.3), with only events post-1960, compared with expected events assuming a uniform distribution. 239

S5.18 De-clustered catalog of events at Mt Ruapehu from catalog of Scott (2013) represented as a circular dataset. 241

S5.19 De-clustered catalog of events at Mt Ruapehu from catalog of Scott (2013), with only events post-1940, compared with expected events assuming a uniform distribution. 242

S5.20 Event counts by magnitude between 1940–1960 (top) and 1960–1960 (bottom). 243

6.1 Seismic network around Ruapehu in 2022. 247

6.2 Covariance matrix spectral width measurements (normalized) computed using four broadband stations (MAVZ, FWVZ, WHVZ, TRVZ) on the Ruapehu edifice. 248

6.3 Seismic drum recording at FWVZ station on 5 May, 2022. 248

6.4 Measurements of cross-correlation characteristics at Ruapehu for a single station-pair (FWVZ-MAVZ). 249

6.5 Correlogram of 1-day cross correlation functions computed for station-pair FWVZ-MAVZ. 250

6.6 Single-station cross-component (SC) subdaily velocity changes (6-hour cross-correlation functions) computed between 2–4 Hz. 251

6.7 Map of Tungurahua volcano. 252

6.8 Covariance matrix spectral width computed using vertical component stations at Tungurahua. 253

6.9 Measurements of cross-correlation characteristics at Tungurahua for a single station-pair (T04–T12). 254

6.10 Phase similarity as a function of frequency and lag-time for station-pair T04-T12. 255

6.11 Velocity changes computed at different frequencies at Tungurahua. . . 256

6.12 Zoomed in velocity changes from Sept 2014 to April 2016 compared with spectral width. 257

6.13 Cluster output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pair T03-T04 at Tungurahua. 258

6.14 Comparison of dominant frequency of cross-correlation functions with velocity changes at Tungurahua. 260

6.15 Cluster output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pair FWVZ-MAVZ at Ruapehu 2022. 262

6.16 Velocity changes recorded at different single stations (average of EN, EZ, NZ) during period of strongest tremor. 264

6.17 Velocity changes corresponding to four rainfall events. 265

6.18 Plotting velocity change against elevation for the four different rainfall events. 266

List of Tables

S2.1 Filters used for analysing cross-correlation function characteristics. . . 70

5.1 Comparison of correlation coefficients computed between different pairs of clusters identified in Figure 5.2 with the corresponding distance they would be agglomerated in the dendrogram. Averages of all individual cross-correlation functions within each cluster are used for computing the correlation coefficients (between -20 and -10 s as with clustering). Since the distance is computed as 1 minus the correlation coefficient, we show the value of 1 minus the distance for easier comparison. 180

5.2 Velocity model used for velocity-depth inversion. 190

5.3 Summary of arguments for and against the different mechanisms that could influence seasonal velocity changes on Ruapehu. 215

S5.1 Bandpass filters used for measuring amplitude ratio. 229

S5.2 Semi-quantitative estimate of Ruapehu eruptions based on observed effects. 234

S5.3 Contingency table used for Fisher's exact test, magnitudes 2+, grouping events between September-November. 240

S5.4 Summary of p-value results for different tests and subsets of the data. 240

S5.5 Summary of p-value results for different tests and subsets of the data when using a cut-off year of 1940. 242

Chapter 1

Introduction

This chapter covers the primary motivations and objectives for the thesis (Section 1.1), including a brief description of each chapter. This is followed by an introduction to relevant background literature (Section 1.2). Finally, additional notes relevant to the thesis are given in Section 1.3.

1.1 Motivation and Primary Objectives

The need for monitoring volcanoes is clear. During the 20th century, volcanic eruptions were responsible for approximately 80,000 deaths (Sigurdsson et al., 2015). Recent eruptions at Nyiragongo (Congo) in 2002, Mount Merapi (Indonesia) in 2010, Volcán de Fuego (Guatemala) in 2018, and Anak Krakatoa (Indonesia) in 2018 were responsible for at least 1200 deaths combined (Tedesco et al., 2007; Rahman et al., 2016; Naismith et al., 2019; Ye et al., 2020). Economical impacts can also be significant. For example, the eruption of Eyjafjallajökull (Iceland) in 2010 caused major diruption due closure of the European airspace following the transport of fine ash across Europe (Gudmundsson et al., 2010; Budd et al., 2011). Total loses to the airline industry were over one billion dollars (Budd et al., 2011). Thus it is evident that volcanic eruptions have far-reaching impacts on human life, both from loss of live and substantial economic consequences.

Volcano monitoring can be considered to serve two primary functions. The first, and perhaps most challenging, is to improve our ability to forecast volcanic eruptions. This ideally requires an understanding of when, where, and what style of eruptive activity will occur (Caudron et al., 2020). Furthermore an understanding of the duration of the eruption and degree to which populations will be affected is also highly sought (Sparks, 2003; Caudron et al., 2020). Some success has been achieved. For example, a correct forecast of eruptive activity at Mount Pinatubo in 1991 led to tens of thousands of lives being saved after large evacuation zones were established within days of eruption (Harlow et al., 1996). Similarly, an eruption at Hekla (Iceland) in 2000 occurred following a warning of an impending eruption provided to civil defense authorities only 50 minutes earlier, with a public radio announcement 15 minutes before the onset (Agustsson et al., 2000). Despite these

successes though, there are many examples of volcanic eruptions that occur with minimal warning. For example, phreatic eruptions at both Mount Ontake (Japan) in 2014 and Whakaari (New Zealand) in 2019 were responsible for 60 and 22 deaths respectively. Both occurred with minimal warning, despite real-time monitoring systems in place. Improving the ability to provide short-to-medium term forecasting is therefore highly sought. Additionally, the recent catastrophic eruption of the volcano Hunga Tonga Ha'apai highlights the need for improved monitoring of submarine volcanoes (Poli and Shapiro, 2022; Matoza et al., 2022). This eruption was among the largest ever recorded by modern geophysical instrumentation, with tsunami-like waves generated by the resulting atmospheric pressure waves (Matoza et al., 2022). The unprecedented magnitude of this event, combined with the difficulties of monitoring underwater volcanoes, highlights that continued advancements and innovations in volcanic monitoring techniques are neccessary to better predict and mitigate the impacts of such unforeseen and potentially devastating events

The second primary function of monitoring is to develop our understanding of the structure and dynamics of volcanoes. Many different techniques are now applied in the pursuit of understanding volcanic systems and pre-cursory phases of activity. Common data acquired includes measures of seismicity, deformation, gas emissions, and temperature. These can be combined towards the development of conceptual models of volcanic systems that can be used to inform forecasting. Thus, there is great value in the study of older datasets, at different time scales, towards improving our knowledge of volcanic systems and the life-cycle of volcanoes more broadly.

This thesis focuses on the application of an increasingly popular seismic method known as seismic noise interferometry (or coda-wave inteferometry). The technique takes advantage of continuously recorded seismic energy (noise) recorded at the earths surface to recover subtle variations in seismic velocity that can provide insights into the behaviour of volcanic systems. The continuous nature of measured velocity changes is a key strength of the technique, allowing information to be retrieved in the absense of seismicity. Furthermore, the method is sensitive to changes at depth. It is then possible to detect perturbations in the subsurface that are not detectable at the Earth's surface. For these reasons, the method has been increasingly applied in volcanic settings over the past two decades. Currently, the majority of usage is focused on the identification of velocity changes in hindsight. However, a few volcano observatories do now apply the technique in real-time (Duputel et al., 2009; Lecocq et al., 2014).

Despite the growing popularity of seismic noise interferometry, and the measurement of temporal seismic velocity challenges, key challenges remain. One of these relates to the highly variable nature of the seismic wavefield at volcanoes. Volcano seismicity (particularly volcanic tremor) can have a strong influence on the recovery of seismic velocity changes (Gómez-García et al., 2018; Caudron et al., 2021). The effect may be to limit our ability to extract velocity measurements or worse, could artifically alter velocities such that variations no longer reflect real subsurface changes. The latter can be more problematic because it adds an additional layer of difficulty when interpreting the results. This also closely relates to another key challenge in the application to volcanoes; that is, to distinguish between internal (volcanic) and external (environmental or tectonic) forcings. Earthquake-induced changes are rel-

atively straightforward to recognise in the short-term, usually recorded as sudden velocity decreases coinciding with the arrival of seismic waves (e.g. Brenguier et al., 2008a, 2014; Lesage et al., 2014; Yates et al., 2019). However, longer-term recovery of velocities may last multiple years, potentially masking other weaker effects. Similarly, meteorological effects (for example, rain and temperature variations) can induce velocity changes of comparable, or larger, amplitudes relative to velocity changes due to volcanic processes. Thus, determining whether recovered changes relate to an internal (volcanic) process, external (environmental or tectonic) process, or an artificial change due to variable noise sources is a significant challenge. Solving some of these issues would be highly benefitial towards wider use of the methodology.

We explore some of these themes further, with a particular focus on explosive volcanism. Unlike effusive activity, explosive activity can often occur suddenly with little warning (Cassidy et al., 2018; Stix and de Moor, 2018; Ardid et al., 2022). This is especially true for phreatic/phreato-magmatic eruptions originating from shallow depths; for example, within or just below the hydrothermal system. The eruptions at Mount Ontake (2014) and Whakaari (2019) are such examples of phreatic eruptions that occurred with minimal warning. The heightened complexity of explosive eruptions is reflected in the recovered velocity changes, where precursory activity is often absent (Lesage et al., 2018a) or requires more specialized data manipulation, or analysis techniques, to uncover meaningful changes (Caudron et al., 2022). In this regard, continued study of explosive volcanoes is needed to understand the ability of seismic noise interferometry to detect changes associated with different styles of explosive activity.

With these considerations in mind, the main objectives of thesis are:

1. Consider how seismic datasets might be characterized towards gaining a better understanding of the potential and limitations of monitoring through coda wave interferometry at any given volcano.

2. Extract seismic velocity changes from different explosive volcanoes towards improving our understanding of both forecasting potential and other insights that may be obtained by applying coda wave interferometry (for example, the influence of external forcings on volcanoes).

3. Identify key areas for future research and development, while also considering practical steps that could enhance our ability to monitor and forecast explosive activity.

The chapters are divided as follows:

The first two chapters (two and three) are focused on the first objective. In Chapter two, we consider steps that can be taken to characterize the background seismic wavefield at volcanoes and the effect observed features have on the data. Five volcanoes with varied eruptive activity are used as examples towards constraining the expectations when applying coda wave interferometry at volcanoes. These are Piton de la Fournaise (La Réunion, France), Kilauea (Hawaii), Mt Ruapehu (New Zealand),

Stromboli (Italy), and Whakaari (New Zealand). In Chapter three, we apply an unsupervised machine learning approach in the form of agglomerative hierarchical clustering to cross-correlation functions. This can be used to group data that share similar characteristics, where we test the approach at Piton de la Fournaise and Ruapehu. These results are published in Geophysical Journal International (Yates et al., 2023), with the methodology also later applied at Kamchatka volcano, Russia (Makus et al., 2023). It is hoped that methods discussed in both chapters can be used as a starting point towards understanding a new volcanic dataset intended for coda wave interferometry.

Chapters four and five focus on individual case studies at volcanoes known for explosive activity. These are Stromboli (Chapter four) and Mt Ruapehu (Chapter five). The combination of these two case studies is therefore focused on the second objective. At Stromboli, particular interest is given to identifying precursory changes at open-vent basaltic systems that mark the transition between mild activity and more significant explosive activity. Such was the case in 2019, when two paroxysms occurred in the space of two months with minimal warning (Giudicepietro et al., 2019; Inguaggiato et al., 2019; Giudicepietro et al., 2020). At Ruapehu, analysis of the eruption chronology suggests a higher probability of smaller, predominantly phreatic eruptions, in Spring. Thus we use coda wave interferometry to better understand the influence of seasonal processes on the volcano and possible trigger mechanisms for spring-time eruptions. Both chapters are structured such that they can be read in isolation.

In Chapter 6, the various insights from earlier chapters are discussed. As part of this, new datasets are briefly introduced for Tungurahua volcano (Ecuador) and a few months of heightened activity at Ruapehu in 2022. These are used in combination to earlier chapters to support the discussion. Finally, we conclude by outlining future perspectives towards monitoring explosive activity using coda wave interferometry, thus addressing the final (third) objective of the thesis.

1.2 Monitoring volcanoes through passive seismic interferometry

Seismic interferometry refers to the principle of cross-correlating seismic responses recorded at different receivers to approximate the impulse response between them (Wapenaar et al., 2010). The impulse response, also known as the Green's function, represents the seismic response that would be recorded at one receiver if the source were located at the other. In passive seismic inteferometry, the Green's function is estimated from passively recorded signals as opposed to an active/controlled source used in controlled-source interferometry (Wapenaar et al., 2010). This thesis focuses on passive seismic interferometry, with the different components of the passive seismic wavefield discussed further in Section 1.2.1.1.

In this Section, the foundation for monitoring subsurface properties is introduced, including the use seismic coda to monitoring time-lapse changes. Basic processing steps are then described towards converting raw seismic data into seismic veloc-

ity changes. Finally, the specific application of monitoring volcanoes is discussed, including considerations for explosive activity in particular.

1.2.1 Extraction of medium properties using ambient noise

1.2.1.1 Properties of the seismic noise wavefield

The Earth's surface is in a state of continuous motion, driven by energy originating from a variety of sources (Gutenberg, 1951, 1958). The complex wavefields generated by such sources are commonly refered to as 'seismic ambient noise'. The different sources that comprise this can usually be distinguished through their statistical and spectral properties.

At very long periods (more than 30 s), the free oscillations of the Earth are observed known as Earth 'hum' (Nawa et al., 1998; Suda et al., 1998). Such oscillations are thought to be excited by oceanic infragravity waves interacting with seafloor topography (Rhie and Romanowicz, 2004) and atmospheric forcing (Kobayashi and Nishida, 1998). Typically, this energy is too low in frequency to be of use in monitoring volcanoes.

Dominating the seismic noise wavefield, with periods between approximately 1–20 s are the oceanic microseisms. These are characterized by broad spectral peaks at around 14 s and 7 s (Hasselmann, 1963; Peterson, 1993; Stutzmann et al., 2000). These are often refered to as the primary and secondary microseisms respectively. Primary microseisms, consisting of both Love and Rayleigh waves (Friedrich et al., 1998), are generated through the interaction of ocean waves with the coastal seafloor (Hasselmann, 1963). Secondary microseisms, consisting predominantly of Rayleigh waes (Haubrich and McCamy, 1969; Friedrich et al., 1998) are generated through non-linear inteference of incoming and reflected ocean waves. This interaction can generate a standing wave that excites surface waves on the ocean floor (Longuet-Higgins, 1950; Hasselmann, 1963). The secondary microseism is typically stronger than the primary microseism, and characterised by a gentle slope for periods shorter than 7 s and a steep slope for periods longer than 7 s (Stutzmann et al., 2000).

At higher frequencies, below approximately 1 s period, sources are mostly associated with both natural and anthropogenic phenomena. Natural source, for example, include wind interacting with the topography (Hillers and Ben-Zion, 2011), and precipitation and runoff (Burtin et al., 2008). Anthropogenic sources can include energy from traffic, industrial machinary, and trains (e.g. Hillers et al., 2012; Boese et al., 2015; Riahi and Gerstoft, 2015; Lecocq et al., 2020). These are usually recognisable through their periodicity, for example temporal variations between day and night, or variations based on the day of the week.

At active volcanoes, it is common to have additional sources of seismic noise not observed in non-volcanic settings. Volcanic tremor, in particular, can be observed at many volcanoes with frequencies dominant between 0.5–5 Hz (Chouet and Matoza, 2013; Sigurdsson et al., 2015). This signal represents a more continuous vibration of the ground related to hydrothermal activity or to magma degassing (Chouet and

Matoza, 2013, and references therein). Unlike tectonic or volcano-tectonic activity, tremor can have durations of minutes to days, or months to years (Sigurdsson et al., 2015). Furthermore, there is a wide variability in the amplitude and frequency of tremor both at volcanoes globally (McNutt and Nishimura, 2008; Chouet and Matoza, 2013). Variability of tremor properties with time can also be observed at single volcanoes (Chouet and Matoza, 2013). Such variety indicates that a number of different processes may be driving volcanic tremor.

1.2.1.2 Green's Function Recovery

Aki (1957) was the first to show that the ambient noise wavefield can be used to extract valuable information on the propagation medium. Through the method known as the Spatial Auto Correlation method (SPAC), it was proposed that the nature of the propagation medium could be recovered from recorded noise. Claerbout (1968) later demonstrated that the auto-correlation of seismic energy recorded at the surface could approximate the reflection response of a 1D acoustic medium. An important advancement came in the twenty-first century, where Weaver and Lobkis (2001) and Lobkis and Weaver (2001) demonstrated under laboratory conditions that the Green's function could be obtained from cross-correlating diffuse wavefields. This key contribution paved the way for modern usage of seismic noise interferometry in the field of seismology.

The breakthrough in the field of seismology came with the demonstration that the seismic Green's function could be approximated by cross-correlating ambient noise recorded at pairs of seismometers (Shapiro and Campillo, 2004; Sabra et al., 2005; Shapiro et al., 2005). Further, the Green's function is demonstrated to be recoverable over different length scales, from interstation distances of a few kilometers (e.g. Sabra et al., 2005; Brenguier et al., 2007) to interstation distances of thousands of kilometers (e.g Yang et al., 2007; Bensen et al., 2008). Thus, the approach has become popular in the application of seismic tomography (e.g. Yao et al., 2006; Cho et al., 2007; Yang et al., 2007; Lin et al., 2007; Behr et al., 2010).

An important question is whether the recovery of the Green's function is possible in real-world conditions. Theories describing the relationship between the cross-correlation of seismic noise and the Green's function rely on the assumption that the noise sources are equipartitioned (Snieder, 2004; Sánchez-Sesma and Campillo, 2006). That is, the sources are uniformly distributed and independent with a zero net flux. In reality, such assumptions are rarely valid, with limited independent sources and azimuthal biases in the noise wavefield regularly observed (Tsai, 2010). However, Wapenaar (2006) showed that a one-sided distribution of noise sources can still accurately resolve the Green's function on one side of the cross-correlation function. Further, Tsai (2010) demonstrated that partial Green's function phase information can still be retrieved even if the wavefield is not fully partitioned. Thus, approximations of the Green's function can still be retrieved without satisfying theoretical assumptions of independant and equipartitioned noise sources.

Different sub-branches of seismic inteferometry are identified depending on the type of waves used in analysis (Wapenaar et al., 2010; Obermann and Hillers, 2019).

These are direct-wave inteferometry, reflected-wave interferometry, and coda-wave interferometry. Direct-wave interferometry (Campillo and Paul, 2003; Shapiro and Campillo, 2004; Sabra et al., 2005; Shapiro et al., 2005) focuses on the retrieval of ballistic waves travelling directly from source to receiver. Reflected-wave interferometry (Claerbout, 1968; Wapenaar et al., 2010), on the other-hand, aims to recover the reflection response at one receiver from a virtual source at another. Finally, coda-wave interferometry focuses on the seismic coda consisting of multiply scattered waves (Snieder et al., 2002; Snieder, 2006). This approach is most popular for monitoring applications and is the adopted approach in this thesis.

1.2.2 Coda-wave inteferometry for monitoring temporal changes

1.2.2.1 Seismic coda

Many different signals are recorded by a seismic stations following a nearby earthquake. The first arrivals on a seismogram are body waves (P and S) and surface waves (Rayleigh and Love) that travel along a direct path between source and receiver. These are followed by a long wave train characterized by slow amplitude decay (Aki and Chouet, 1975). These later arrivals are referred to as the 'seismic coda', consisting of multiply scattered waves resulting from subsurface heterogeneities. In cross-correlation functions, peaks associated with energy that has traveled a multiple-scattered path is also refered to as the coda. The first arrivals, in contrast, are typically refered to as direct (or ballistic) waves.

Several useful properties of coda waves makes them an ideal target for monitoring. First, the presence of multiple scatterers acts to make the seismic wavefield more isotropic (Snieder, 2004; Froment et al., 2010). That is, the distribution of arriving energy is no longer entirely dependant on the original seismic source location. Rather, it is influenced by the distribution of scatterers in the medium. Thus, the assumptions of equipartitioned noise sources can be relaxed due to the expectation that scattered waves propagate more isotropically near receivers (Snieder, 2004). Similarly, the reduced influence of non-isotropic source distribution means that the coda are less sensitive to biases in travel time estimation relative to direct arrivals (Froment et al., 2010). Furthermore, the longer travel times of the coda relative to direct arrivals means they spend longer sampling the sub-surface. Thus they can accumulate larger delay-times in the presence of velocity perturbations, increasing the sensitivity of measurements to subtle changes. Finally, it has been shown that monitoring weak changes in the coda is possible without reconstructing the Green's function (Hadziioannou et al., 2009). Combining these advantages underpins the wide usage of the methodology over the previous two decades towards monitoring the subsurface.

1.2.2.2 Time-lapse monitoring in the coda

Seismic waves traveling at the Earth's surface are sensitive to small perturbations in elastic properties (Aki, 1957; Sato et al., 2012). Specifically, variations in the

elastic moduli alter both the compressional (P-wave) and shear (S-wave) velocities of the medium. In the case of the coda, they are primarily sensitive to changes in shear velocity (Aki and Chouet, 1975; Snieder, 2006). Thus, monitoring changes in seismic velocity through time can provide valuable insights into the properties of the subsurface in response to evolving behaviour of volcanic systems.

Two types of changes are typically recorded in the seismic coda that are of interest in monitoring applications. These are changes in phase or changes in waveform shape. Changes in phase are used to estimate seismic velocity changes while changes in waveform shape are thought to reflect changes in the scattering properties of the medium (Obermann et al., 2013). Variability in the seismic wavefield can also induce changes in phase and waveform shape; however, these do not represent physical changes in the medium and thus are undesirable.

Poupinet et al. (1984) were the first to apply coda wave interferometry, using pairs of micro-earthquakes, or earthquake doublets. In a series of moving windows, the recorded seismograms for one earthquake were cross-correlated with seismogram of another to measure differences in arrival time (Poupinet et al., 1984). Fractional changes in shear velocity were measured by examining the time delay as a function of time within the coda. Similar techniques were then later used to extract seismic velocity changes from repeated controlled explosions (Nishimura et al., 2000) and earthquake multiplets (Peng and Ben-Zion, 2006).

Measuring velocity changes through examination of the time delay as a function of lag time in the coda assumes a homogenously distributed velocity change. In this case, a new arrival time of seismic energy $(t + \delta t)$ in response to a velocity change (δv) — assuming the same source location — is given by:

$$t + \delta t = \frac{d}{v + \delta v} \tag{1.1}$$

where d is the distance traveled. Rearranging for the distance, and multiplying out the terms gives:

$$d = vt + v\delta t + t\delta v + \delta v\delta t \tag{1.2}$$

If velocity changes are assumed to be small, i.e. $\delta v\delta t \approx 0$, this becomes:

$$\frac{\delta t}{t} = -\frac{\delta v}{v} \tag{1.3}$$

This relationship forms the basis of measuring seismic velocity changes in the coda. That is, the fractional change in seismic velocity change be recovered through analysis of the fractional change in travel time. In reality, it is unlikely that velocity changes are entirely uniform in space. Thus, it is common to refer to measured velocity changes as being an 'apparent' change.

Applications of coda wave interferometry using seismic noise followed soon after the

discovery that the Green's function could be approximated through this approach (Shapiro and Campillo, 2004; Sabra et al., 2005; Shapiro et al., 2005, e.g). Sens-Schönfelder and Wegler (2006) were the first to apply the method towards time-lapse monitoring. At Merapi volcano, they recorded seasonal changes in seismic velocity due to precipitation (Sens-Schönfelder and Wegler, 2006). Shortly after,Brenguier et al. (2008b) recorded the first changes of seismic velocity induced by volcanic processes. Over 18 months at Piton de la Fournaise, they identified velocity decreases of approximately 0.1% in the weeks preceding four volcanic eruptions. Piton de la Fournaise has since been the subject of many studies that confirm the connection between volcanic activity and changes in seismic velocity (e.g. Duputel et al., 2009; Clarke et al., 2013; Obermann et al., 2013; Sens-Schönfelder et al., 2014; Rivet et al., 2014; De Plaen et al., 2016).

Over the previous two decades the number of studies applying coda-wave interferometry towards monitoring has grown rapidly. From this, many natural phenomena has been found to produce measurable seismic velocity changes. These include short short and long-term volcanic changes (Brenguier et al., 2008b; Mordret et al., 2010; Obermann et al., 2013; Rivet et al., 2014), tectonic earthquake activity (Brenguier et al., 2008c; Chen et al., 2010; Minato et al., 2012; Froment et al., 2013; Taira et al., 2015; Heckels et al., 2018), precipitation (Sens-Schönfelder and Wegler, 2006; Hillers et al., 2014), atmospheric pressure loading (Silver et al., 2007; Niu et al., 2008), temperature induced thermoelasticity (Meier et al., 2010; Richter et al., 2014; Hillers et al., 2015a), snow-loading (Hotovec-Ellis et al., 2014; Wang et al., 2017; Makus et al., 2023), and tidal cycles (Yamamura et al., 2003; Takano et al., 2014). The magnitude of velocity changes recorded tends to be on the order of 0.1%, though larger changes of a few percent have been recorded (Minato et al., 2012; Taira et al., 2015; James et al., 2017). Discriminating between the different phenomena driving velocity changes at a single location remains a challenge in the monitoring applications.

1.2.3 From raw data to seismic velocity changes

While many different processing schemes exist, several steps have become relatively standard in the application of coda wave interferometry using seismic noise (e.g. Bensen et al., 2007). These are described below, with processing throughout the thesis using these steps (specific parameters can be found within individual chapters). The python package MSNoise (Lecocq et al., 2014) is used towards this purpose, unless otherwise stated. A more complete review of different steps available today can be found in Lesage et al. (In review).

1.2.3.1 Data Preparation and Preprocessing

Raw seismic data corresponding to a single day, at each station, are first grouped into continuous chunks, demeaned and tapered, and merged again to give 1-day long seismic traces. If small gaps exist (less than 10 seconds), they are filled with interpolated values. Larger gaps are retained (to be removed later). Traces are then

bandpassed and decimated to a new sampling frequency. If intending to compute velocity changes between pairs of stations with different instrument response, a correction is applied. The 1-day traces for each station are then split into smaller slices (30-min slices in this thesis) for further pre-processing and computation of cross-correlation functions. If gaps are still present within an individual slice of data (i.e. too large for interpolation), they are discarded.

Two forms of normalization are typically applied to seismic data prior to computing the cross-correlation function. These are temporal normalization and spectral normalization. Temporal normalization is applied towards reducing the influence of discrete signals, such as earthquakes. In MSNoise, the choices include the application of one-bit normalisation (e.g. Larose et al., 2004) — where only the sign of the signal is retained — or clipping the signal at a multiple of the root mean square (e.g. Sabra et al., 2005). The effect of these approaches on both the amplitude and frequency spectrum is shown in Figure 1.1. Spectral normalization, or whitening, is applied to broaden the spectrum of noise. This is typically achieved by setting the amplitude of the signal spectrum to 1 at all frequencies, thus reducing the influence of peaks associated with individual sources. Similarly, whitening results in cross-correlation functions that are not dominated by energy corresponding to the microseisms (Bensen et al., 2007). This process is demonstrated in Figure 1.2.

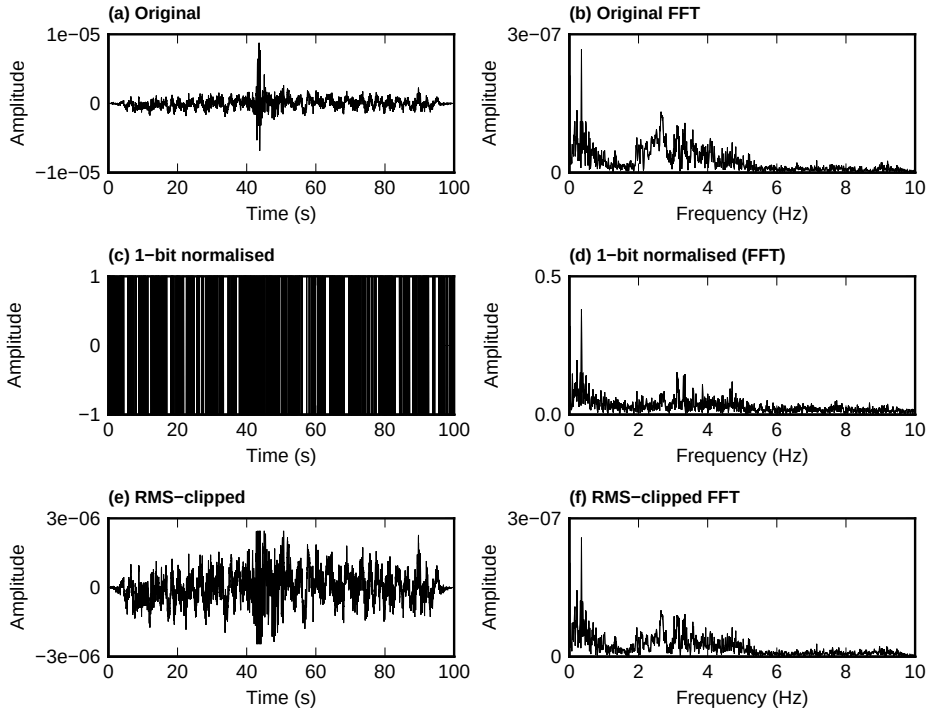

Figure 1.1: Different methods of temporal normalisation are demonstrated, showing the effects in both the amplitude and frequency spectrum. Taken from (Yates, 2018). (a) Original raw waveform, (b) FFT of raw signal, (c) 1-bit normalised raw signal, (d) FFT or 1-bit normalised signal, (e) raw signal after being clipped at 3 times the RMS, (f) FFT of RMS-clipped signal.

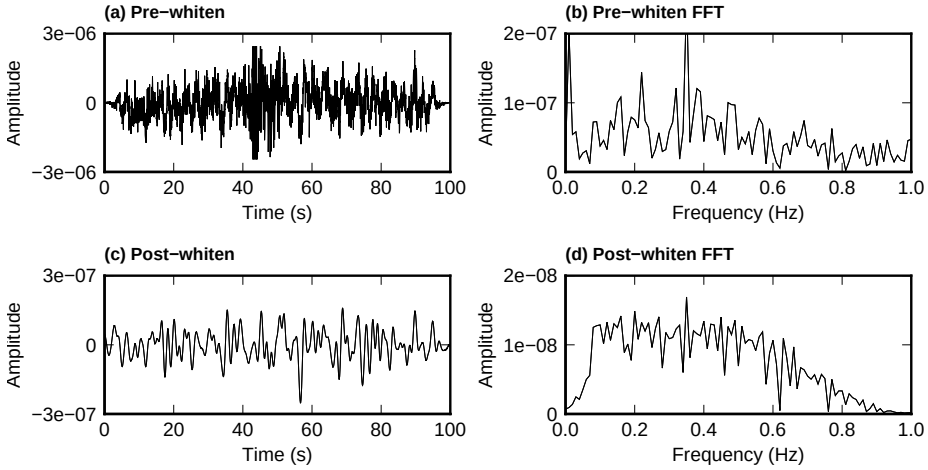

Figure 1.2: Comparison of RMS-clipped signal before and after whitening is applied between 0.1 - 0.4 Hz. Taken from (Yates, 2018). (a) Pre-whitened signal, (b) FFT of pre-whitened signal, (c) Post-whitened signal, (d) FFT of Post-whitened signal.

Note, in the case of auto-correlations (i.e. cross-correlation between the same component of a single-station), spectral whitening is commonly not applied as only the phase of the signal would be retained (Hobiger et al., 2014; Plaen et al., 2016). For auto-correlations this does not contain any information on the medium. Thus, in the absense of whitening, auto-correlations are more sensitive to changes in the seismic source relative to cross-component cross-correlations (Hobiger et al., 2014; Plaen et al., 2016; Yates et al., 2019). This issue can be solved however through the use of alternative approaches, such as the Phase Cross Correlation (PCC), where no normalization is required (Schimmel et al., 2011; De Plaen et al., 2019; Caudron et al., 2022).

1.2.3.2 Computing Cross-Correlation Functions and Temporal Stacking

The cross-correlation of seismic traces is most efficiently performed in the frequency domain (Lecocq et al., 2014). For two times series $x(t)$ and $y(t)$, and their corresponding Fourier transforms $X(f)$ and $Y(f)$, the cross-spectrum $C(f)$ is computed as:

$$C(f) = X^*(f) \cdot Y(f) \tag{1.4}$$

where $X^*(f)$ is the complex conjugate of $X(f)$. Taking the inverse Fourier transform then gives the cross-correlation function in the time domain $c(t)$, where t is the lag time.

After computing the cross-correlation function for individual time slices, it is typically necessary to perform temporal stacking. This acts to improve the signal-to-noise ratio, with the expectation that coherent energy will combine constructively while incoherent energy combines destructively. The end result is to have cross-

correlation functions converge towards a stable function that allows for accurate and repeatable measurements on near identical waveforms. This, however, comes at the cost of temporal resolution and therefore reduces the ability to detect short-term velocity changes. Thus the aim is to choose a stack size that is large enough to accurately measure velocity changes while being small enough to recover valuable temporal information.

The most common method is to take the time average of the individual cross-correlation functions following:

$$c(t) = \frac{1}{N} \sum_{i=1}^{N} c_i(t) \tag{1.5}$$

where $c_i(t)$ represents a individual cross-correlation functions being stacked, and N the number of signals. This is known as the linear stack. This can be modified through the use of weights to minimise the influence of low quality data. For example, through the use of the phase weighted stack (Schimmel and Paulssen, 1997). Throughout this thesis, however, we use the linear stack and apply quality control criteria at later stages.

1.2.3.3 Measuring Velocity Changes

Velocity changes are often computed between smaller (current) stacks of cross correlation functions and a longer reference stack. This is chosen to reflect the background state, with the criteria that $N_{ref} >> N_{curr}$. Here N_{curr} is the number of days stacked for the current stack and N_{ref} the number for the reference stack. The choice of reference function can impact the resulting velocity change measurements. In many studies, the reference is constructed using all available data (e.g. Brenguier et al., 2008b,a). Alternatively, it may be preferable to use time periods of relative quiescence (e.g. Obermann et al., 2013), or multiple references if cross-correlation functions change significantly in time (Sens-Schönfelder et al., 2014). Finally, it should be acknowledged that methods that allow velocities to be computed without a reference do exist (Brenguier et al., 2014; Gómez-García et al., 2018; Muzellec et al., 2023). For example, Brenguier et al. (2014) compute the relative velocity change between all available pairs of cross-correlation functions. They then perform a Bayesian least-squares inversion towards recovering velocity changes through time. This approach removes the need to define a reference period, while also being useful when cross-correlation functions are highly variable through time. In this thesis, however, we measure velocity changes using techniques that make use of a reference period. The two most commonly used methods are the stretching technique (Sens-Schönfelder and Wegler, 2006) and the moving-window cross-spectral technique (Poupinet et al., 1984; Ratdomopurbo and Poupinet, 1995; Clarke et al., 2011).

For the stretching technique, each current stack is compared to stretched or compressed versions of the reference stack. This stretching is controlled by a stretching factor ϵ, where a waveform $s[t]$ becomes $s[t(1-\epsilon)]$. The value ϵ that maximises the

correlation coefficient (ϵ_{max}) between the current and reference function then gives the apparent velocity change as $\delta v/v = -\epsilon_{max}$ (for the case where the reference stack is stretched).

For the moving-window cross-spectral (MWCS) technique, delay times are measured in the frequency domain in a series of moving windows within the coda. Specifically, the delay is obtained from the unwrapped phase of the cross-spectrum of the Fourier-transformed current and reference stacks (Clarke et al., 2011). The length of the moving windows is typically chosen to ensure accurate delay time measurements can be made in the frequency range of interest. Thus, most studies use window lengths equal to (Hadziioannou et al., 2011; Taira et al., 2018) or greater than (Hillers et al., 2015a; Mordret et al., 2016) the longest period of the cross-correlation functions. After measuring the delay at different points in the coda, the final apparent velocity can be calculated as the slope of delay times ($\delta t/t = -\delta v/v$).

Both methods have been found to produce similar results (Hillers et al., 2015b; Machacca-Puma et al., 2019; Caudron et al., 2021), and the choice of method may depend on author preference. The stretching technique has been found to produce more stable velocity changes (Hadziioannou et al., 2009), while also being less prone to cycle skipping (Mikesell et al., 2015). The moving-window cross-spectral technique on the other-hand, through the separation of the amplitude and phase spectrum, is theoretically less sensitive to changes in seismic source (Zhan et al., 2013). A more complete discussion on the advantages and disadvantages of each technique can be found elsewhere (e.g. Obermann and Hillers, 2019).

Finally, alternative algorithms have been proposed that may offer advantages to traditional approaches (e.g. Mikesell et al., 2015; Mao et al., 2020; Yuan et al., 2021). One that holds particular promise is the wavelet cross-spectrum technique (Mao et al., 2020). In this approach, the delay time between current stacks and the reference stack are retrieved in the time-frequency space. This is done by obtaining the phase of both functions by performing convolutions between the functions and a series of scaled and translated versions of a mother wavelet. The phase difference between the current and reference function can then be calculated, and measurements of delay time recovered by dividing by $2\pi f$ at each frequency. The final measurement of velocity change is then recorded as $\delta t/t = \delta v/v$, as with the moving-window cross-spectral technique. Further details can be found in Mao et al. (2020).

A key advantage of the wavelet cross-spectrum approach is to obtain velocity changes in the time-frequency space. This avoids the need to apply several different filters in order to assess the relationship between velocity changes and frequency using the stretching or moving-window cross-spectral technique. Furthermore, even when processing different filters using the doublet technique, Mao et al. (2020) find evidence that greater smearing of velocity changes computed between different frequency bands occurs relative to the wavelet cross-spectrum approach. The wavelet approach therefore holds much promise towards providing greater frequency resolution to velocity change estimates relative to traditional techniques.

1.2.4 Application to volcanoes

Despite the challenges associated with monitoring volcanoes, they can also be considered ideal locations for coda wave interferometry. For example, their highly heterogeneous structure should contribute to a strong scattering of seismic waves. The scatterers then act as secondary sources of seismic noise that can improve energy equipartition. Furthermore, volcanic rock should be more sensitive to stress variations relative to other rock types due to being strongly fractured and porous. As a result, these unique geological characteristics provide promising opportunities to monitor seismic velocity changes through time.

Identifying the mechanism / volcanic process driving velocity changes remains a challenging problem. However, since the first applications of coda wave interferometry towards monitoring volcanoes (e.g. Sens-Schönfelder and Wegler, 2006; Brenguier et al., 2008c), much effort has gone into uncovering the sources of velocity changes. These are discussed below, followed by a brief discussion of some of considerations specific to explosive volcanoes.

1.2.4.1 Physical mechanisms behind sub-surface changes

Early studies of seismic velocity changes at volcanoes considered the role of microcrack distribution (Ratdomopurbo and Poupinet, 1995; Brenguier et al., 2008b). Laboratory experiments have shown that when stresses are initially increased, seismic velocities in rock increase due to the closure of microcracks (Nur, 1971). As stresses are increased further, new cracks appear as the surrounding rock is deformed (Lockner et al., 1977). The response to an increase in crack density is then a seismic velocity decrease (Nur and Simmons, 1969; Lockner et al., 1977). These effects are enhanced in highly porous rocks (Eberhart-Phillips et al., 1989; Shapiro, 2003; Heap et al., 2014a). Thus volcanic rocks, which contain high vesicle densities and are intensely fractured, can be highly sensitive to changes in effective pressure (Rocchi et al., 2004; Stanchits et al., 2006; Heap et al., 2014b).

Several real-world observations support the interpretation of a relationship between seismic velocity changes and microcrack distribution at volcanoes. For example, velocity changes have been observed to coincide with edifice inflation and deflation events (Brenguier et al., 2008b; Duputel et al., 2009; Obermann et al., 2013; Rivet et al., 2014; Bennington et al., 2015; Cubuk-Sabuncu et al., 2021a). In this case, summit inflation is accompanied by increased fracturing in the surrounding rock (velocity decrease) while summit deflation acts to close microcracks (velocity increase). Furthermore, Donaldson et al. (2017) identified a strong correlation between seismic velocities and radial tilt deformation, at Kilauea volcano. In this case, however, the trend was opposite to those observed previously with inflation-deflation cycles causing velocity increases and decreases respectively (Donaldson et al., 2019). Similarly, velocity increases were observed at both Whakaari (Yates et al., 2019; Caudron et al., 2021) and Ontake (Caudron et al., 2022), inferred as related to shallow pressurization within the shallow hydrothermal system due to sealing.

Differences in sign of velocity changes at different volcanoes in response to defor-

mation highlights another challenge. That is, when both velocity increases and decreases are possible, it isn't straightforward to recognise what should be expected at a given volcano. One explanation for observing both velocity increases and decreases may relate to the previously described transition from elastic (accompanied by crack opening/closing) to plastic deformation (accompanied by new crack development) as strain increases (Nur and Simmons, 1969; Nur, 1971; Lockner et al., 1977). In this scenario, both velocity increases and decreases may be observed in different locations at the same time depending on the strain field resulting from a pressure source (i.e. decrease closer to the source, increase further from the source). Such an effect has been considered at Merapi volcano, where velocity decreases were observed in the upper edifice and velocity increases in the lower edifice (Budi-Santoso and Lesage, 2016a). Alternatively, they considered that the influence of the free surface, topography, and heterogeneity of volcanoes could result in both tensional and compressional stresses being produced in different parts of the volcano in response to a pressure source (Got et al., 2013; Budi-Santoso and Lesage, 2016a). Finally, at Kilauea, the presence of ring faults (with similarly orientated microcracks) is suggested to increase the sensitivity of the subsurface to radial strains rather than volumetric strains (Hotovec-Ellis et al., 2022; Muzellec et al., 2023). In this case, modeling suggests compression dominates over much of the volcano, promoting a velocity increase (Hotovec-Ellis et al., 2022; Muzellec et al., 2023). These observations highlight the continued development towards understanding the velocity changes observed at volcanoes, but also the difficulties in proposing a precise interpretation.

Volcanic fluids (both magmatic and hydrothermal) can also have an influence on seismic velocity changes. While shear wave velocities are theoretically not sensitive to fluid saturation (Nur and Simmons, 1969; Lumley, 2010; Clarke et al., 2020), increases in fluid content can increase pore pressures and thus reduce the effective stress (Grêt et al., 2006; Doetsch et al., 2018). Similarly, conversion of water to steam through interaction with high temperature fluids/volatiles could also lower the effective stress (Caudron et al., 2015). In either case, the response would be a velocity decrease (Grêt et al., 2006; Caudron et al., 2015; Doetsch et al., 2018).

Finally, previously-mentioned non-volcanic sources can have a significant influence on seismic velocity changes at volcanoes. Changes due to rainfall, for example, have been observed at multiple volcanoes (e.g. Sens-Schönfelder and Wegler, 2006; Rivet et al., 2015; Budi-Santoso and Lesage, 2016a). The effect is similar to that described above, with increasing pore fluid pressure resulting in velocity decreases. Snow cover can also influence seismic velocity changes at volcanoes, with velocity increases interpreted to result from pore space reduction and crack closing under increased surface loading (Hotovec-Ellis et al., 2014; Wang et al., 2017; Donaldson et al., 2019). However, velocity decreases may also be recorded in response to snow-loading in the presence of a large fluid body, where fluid diffusion and pore pressure increases may dominate (Silver et al., 2007; Taira and Brenguier, 2016). Finally, large tectonic earthquakes are regularly observed to induce sharp velocity decreases at volcanoes (e.g. Battaglia et al., 2012; Lesage et al., 2014; Machacca-Puma et al., 2019; Yates et al., 2019). The source of these changes is usually attributed to softening, or damage, of the shallow layers in response to seismic shaking (Guyer and Johnson, 1999; Johnson and Jia, 2005). While the initial velocity change (decrease) is often instantaneous, the long-term recovery due to progressive healing of the rock

can last for multiple years (Lesage et al., 2014; Yates et al., 2019).

The many different non-volcanic processes that can influence seismic velocity changes at volcanoes poses a major challenge in the interpretation. To address this challenge, several discriminative approaches can be employed towards distinguishing between volcanic and non-volcanic processes. For example, the use of long time series containing multiple seasonal cycles (e.g. many years) can be used to understand the expected patterns of velocity changes. From this, any strong deviations may be likely candidates for changes of volcanic origin. Alternatively, comparing seismic velocities recorded close to the volcano with those using seismic stations at greater distances (thus unlikely to be influenced by volcanic sources) could indicate whether changes are likely to be volcanic or non-volcanic in origin (e.g. Yates et al., 2019). This follows that environmental changes would be expected to effect a wider area than volcanic changes. Finally, multiple studies have attempted to model velocity changes resulting from environmental processes (such as rainfall and snow-loading) (e.g. Rivet et al., 2015; Wang et al., 2017; Donaldson et al., 2019; Caudron et al., 2022). Similarly, any strong deviations from the modeled velocity changes could then be candidates for changes of volcanic origin. This is assuming that a good fit between observed and modeled velocity changes is observed during times when activity is low. This can be similarly complicated by incomplete understanding of the contribution of different meteorological processes to the observed velocity changes.

In summary, while there have been great advances in our understanding of the processes driving velocity changes at volcanoes, there is still significant uncertainty in identifying the exact process responsible. This can apply both to changes of volcanic origin and non-volcanic origin. However, each application of coda wave interferometry in a different volcanic environment improves our understanding of the velocity changes one might expect in response to different volcanic processes. Furthermore, the integration of novel methodologies, combined with enhanced monitoring capabilities, holds promise towards further enhancing our ability to discern the underlying cause of velocity changes at volcanoes.

1.2.4.2 Explosive versus effusive volcanism

Explosive volcanism is generally more challenging to monitor than effusive activity. Whereas explosive volcanism can occur through a variety of processes — with or without the involvement of new magma — effusive activity can be generally considered to result from the ascent of magma towards the Earth's surface. Such ascent is typically associated with considerable ground deformation and therefore seismic velocity changes (e.g. Obermann et al., 2013; Donaldson et al., 2019). Furthermore, final ascent rates are much slower for effusive activity compared to explosive magmatic eruptions. Thus, much of the early success in the identification of seismic velocity changes due to volcanic processes was based on monitoring effusive activity at Piton de la Fournaise (e.g. Brenguier et al., 2008c; Duputel et al., 2009; Clarke et al., 2013; Obermann et al., 2013; Rivet et al., 2014; Sens-Schönfelder et al., 2014; Rivet et al., 2015).

In cases where explosive activity is preceded by the gradual injection of new magma,

with subsequent pressurization of the magmatic system, velocity changes may be expected over the wider volcanic edifice. For example, seismic velocity changes were recorded over the wider edifice in the weeks preceding explosive activity at Merapi (Budi-Santoso and Lesage, 2016a) and Ubinas (Machacca-Puma et al., 2019). However, explosive activity can also originate from shallow sources, with minimal involvement of magma. For example, recent eruptions at Ontake in 2014 and Whakaari in 2019 were gas-driven events originating in the shallow hydrothermal system. In both cases, seismic velocity increases were identified only by a seismic stations closest to the pressure source (Yates et al., 2019; Caudron et al., 2021, 2022). Those slightly further from the source showed no clear change. Furthermore, at Mt Ruapehu (New Zealand), only one of two gas-driven eruptions in 2006 and 2007 showed a velocity decrease (Mordret et al., 2010); in this case, the smaller of the two in 2006.

The unpredictability and rapid onset of explosive activity make it inherently more hazardous and challenging to forecast. This is especially true for both volcanoes prone to phreatic and phreatomagmatic activity, but also volcanoes that show persistent mild activity that can transition to more violent explosive activity with minimal warning (e.g. Stromboli). It is therefore paramount that future research efforts continue to explore the warning signs and precursors to explosive activity.

1.3 Additional notes on thesis

This section includes additional information that may be useful with regards to this thesis. This covers mention of other contributors, and various codes, used in this thesis (Section 1.3.1) and also co-authored work that was completed during this thesis but is not included (Section 1.3.2).

Finally, all data processing and writing/presenting of results in this thesis was done by myself (Alexander Yates).

1.3.1 Chapter contributions

In addition to the supervisory team (Corentin Caudron, Philippe Lesage, Aurélien Mordret), other contributers have been involved at various stages during the undertaking of this thesis. The contributions to each chapter from 2–6 are referenced below.

1.3.1.1 Chapter Two

Chapter Two makes use of network-based covariance matrix analysis and measures of SNR, phase similarity, and amplitude ratios of broadband cross-correlation functions towards characterizing a dataset towards use for coda wave interferometry.

The application of network based-covariance matrix analysis makes use of the Python package Covseisnet (Covseisnet, 2020). Example codes available on the Github were

adapted for this thesis towards computation of the spectral width over longer time periods. This code is used throughout the thesis.

Cross-correlation functions are computed using the python package MSNoise (Lecocq et al., 2014). This is the case throughout this thesis, though some modifications are made to the code to allow greater flexibility with parameter choices. I have also made contributions to the online github repository.

Measures of SNR, phase similarity, and amplitude ratios were prepared myself. The codes are available at the Github page , in addition to codes used to process long-term spectral width measurements. Jupyter notebooks are also provided.

1.3.1.2 Chapter Three

Chapter Three applies hierarchical clustering to cross-correlation functions at Piton de la Fournaise and Mt Ruapehu volcanoes. As with Chapter Two, spectral width measurements and cross-correlation functions are computed using Covseisnet and MSNoise respectively.

The codes used for hierarchical clustering of cross-correlation functions are based on the Scipy python package (Virtanen et al., 2020). My own codes were then developed towards application to cross-correlation functions output by MSNoise. These are available at , also with a Jupyter notebook.

The results of this chapter are published as Yates et al. (2023). The method of applying hierarchical clustering to cross-correlation functions was then also then used in the paper Makus et al. (2023), where I am a co-author.

In addition to the supervisory team, other contributors to the published paper are Thomas Lecocq and Jean Soubestre.

1.3.1.3 Chapter Four

Chapter Four applied coda wave interferometry to study velocity changes at Stromboli, Italy. As with Chapter Two spectral width measurements and cross-correlation functions are computed using Covseisnet and MSNoise respectively.

Codes used to compute the lateral sensitivity kernels were provided by Thomas Lecocq.

In addition to the supervisory team, other contributors to this work are Andrea Cannata, Flavio Cannavo, Thomas Lecocq, and Lucia Zaccarelli.

1.3.1.4 Chapter Five

Chapter Five applied coda wave interferometry to study velocity change at Mount Ruapehu, New Zealand. As with Chapter Two spectral width measurements and

cross-correlation functions are computed using Covseisnet and MSNoise respectively.

Codes used to apply the wavelet transform from cross-correlation functions output by MSNoise were provided by researchers at the Royal Observatory of Belgium. These codes were adapted in this thesis to allow for more flexible parameter choices.

Codes used to model the influence of a load on a homogeneous elastic half-space were provided by Alexandra Le Contellec and Virginie Pinel.

In addition to the supervisory team, other contributors to this work are Virginie Pinel and researchers at GNS Science, New Zealand, who have provided input (Nico Fournier. Craig Miller, Oliver Lamb).

1.3.1.5 Chapter Six

Chapter Six introduced two new datasets at Mt Ruapehu, New Zealand and Tungurahua, Ecuador. As with Chapter Two spectral width measurements and cross-correlation functions are computed using Covseisnet and MSNoise respectively.

In addition to the supervisory team, other contributors to this work (specifically for the Tungurahua dataset) are Jean Battaglia and Silvana Hidalgo.

1.3.2 Additional projects

During the undertaking of this thesis, I was involved in several other projects applying coda wave interferometry to monitor seismic velocity changes. This includes monitoring changes at volcanoes (Caudron et al., 2021; Makus et al., 2023), changes due to large tectonic earthquakes (Madley et al., 2022) and slow slip earthquakes (Wang et al., 2022). I was also involved in the preparation of a review chapter on coda wave interferometry at volcanoes, currently in review. The full list of co-author publications are below.

Coda wave interferometry for volcano monitoring
P. Lesage, C. Caudron, **A. S. Yates**.
IAVCEI Book Modern Volcano Monitoring. Accepted.

Deciphering the whisper of volcanoes: Monitoring velocity changes at Kamchatka's Klyuchevskoy group with fluctuating noise fields
P. Makus, C. Sens-Schönfelder, L. Illien, T. R. Walter, **A. S. Yates**, F. Tilmann.
Journal of Geophysical Research: Solid Earth, 128 (4), 2023.

Velocity changes around the Kaikōura earthquake ruptures from ambient noise cross-correlations
M. Madley, **A. S. Yates**, M. K. Savage, W. Wang, T. Okada, S. Matsumoto, Y. Iio, K. Jacobs.
Geophysical Journal International, 229 (2), 1357-1371, 2022.

Temporal velocity variations in the northern Hikurangi margin and the relation to slow slip
W. Wang, M. K. Savage, **A. S. Yates**, H. J. Zal, S. Webb, C. Boulton, E. Warren-Smith, M. Madley, T. Stern, B. Fry, K. Mochizuku, L. Wallace
Earth and Planetary Science Letters, 584, 117443, 2022.

A quest for unrest in multiparameter observations at Whakaari/White Island volcano, New Zealand 2007–2018
C. Caudron, T. Girona, A. Jolly, B. Christenson, M. K. Savage, R. Carniel, T. Lecocq, B. Kennedy, I. Lokmer, **A. S. Yates**, I. Hamling, I. Park, G. Kilgour, A. Mazot
Earth, Planets and Space, 73 (1), 1-21, 2021.

Chapter 2

Characterizing seismic datasets for coda wave interferometry at volcanoes: seismic sources and their influence on cross-correlation functions

2.1 Introduction

An outstanding problem in the wider use of coda wave interferometry relates to the many choices available in processing seismic velocity changes. Choices made in pre-processing can, for example, have a large influence on the cross-correlation functions (Bensen et al., 2007; Fichtner et al., 2017). While many choices exist in pre-processing, standard procedures discussed in Section 1.2.3 have been found, generally, to produce reasonable results. However, further decision making, such as the choice of frequency band, the reference period, and coda window, also require special consideration. Failure to justify these choices can lead to uncertainty in the validity of resolved seismic velocity changes.

At volcanoes, the influence that a variable seismic wavefield can have on seismic velocity changes is a significant concern (Ballmer et al., 2013; Gómez-García et al., 2018). Volcanic tremor, in particular, has been shown to contaminate cross-correlation functions and produce unreliable cross-correlation functions (Ballmer et al., 2013; Gómez-García et al., 2018; Caudron et al., 2021). However, tremor has also been demonstrated to be a useful seismic source if its relative stability in time and space can be verified (Donaldson et al., 2019). At the time of writing, few studies have examined the character of cross-correlation functions produced during tremor, or other seismic sources, present at volcanoes.

In this chapter, we consider steps that can be taken to characterize seismic datasets at volcanoes towards coda wave interferometry. These can then be used to inform better decision making in the processing of velocity changes. In the following

sections, we apply covariance matrix analysis to understand the coherent seismic sources propagating over a network, considering also how differences in network geometry may influence the results. Following this, we examine the characteristics of cross-correlation functions in response to the observed seismic sources at five different volcanoes. This sets the platform to determine how coda wave interferometry can be applied at each volcano, understanding both the possibilities and limitations based on seismic source distribution and network geometry.

2.2 Covariance matrix analysis of the seismic wavefield

The seismic sources recorded over a network may act to limit what can be achieved through coda wave interferometry (e.g. Gómez-García et al., 2018) or, alternatively, may allow for non-traditional treatment of a seismic dataset e.g. using tremor as the dominant seismic source (e.g. Donaldson et al., 2017). Thus, understanding the background seismic wavefield can aid in the identification of what is and isn't possible at a given volcano.

Common techniques regularly applied to understand the background seismic wavefield are based on analysing the spectral content of seismograms. One way of doing this is through Power Spectral Density (PSD) functions, where the frequency content recorded at individual seismic stations may be visualised (e.g. McNamara and Buland, 2004; Marzorati and Bindi, 2006). Similarly, a spectrogram can be used to visualise frequency content over time, and is therefore regularly used in the identification of volcano seismicity such as tremor or volcano-tectonic events (e.g. Chouet and Matoza, 2013; Olivier et al., 2019; Caudron et al., 2021). However, both can be limited by their reliance on using data from a single seismic station. Thus, features of power spectral density functions or spectrograms may reflect highly local sources (Soubestre et al., 2018). It may therefore be neccessary to analyse multiple stations in order to verify that sources are coherent over a network.

In this section, we introduce a recently developed network-based technique based on the analysis of the array covariance matrix computed from seismic spectra recorded by multiple stations (Seydoux et al., 2016a; Soubestre et al., 2018, 2019). Through this, we can identify key features of the seismic wavefield that are coherent over a network. We begin by introducing the theory behind the approach, and introducing the python package CovSeisNet (Covseisnet, 2020) which is used to perform the computations. The UnderVolc network at Piton de la Fournaise (Brenguier et al., 2012) is then used to test the potential of this technique to identify features of the seismic wavefield considering different network configurations, including a single-station approach that utilises the three different components (north, east, vertical).

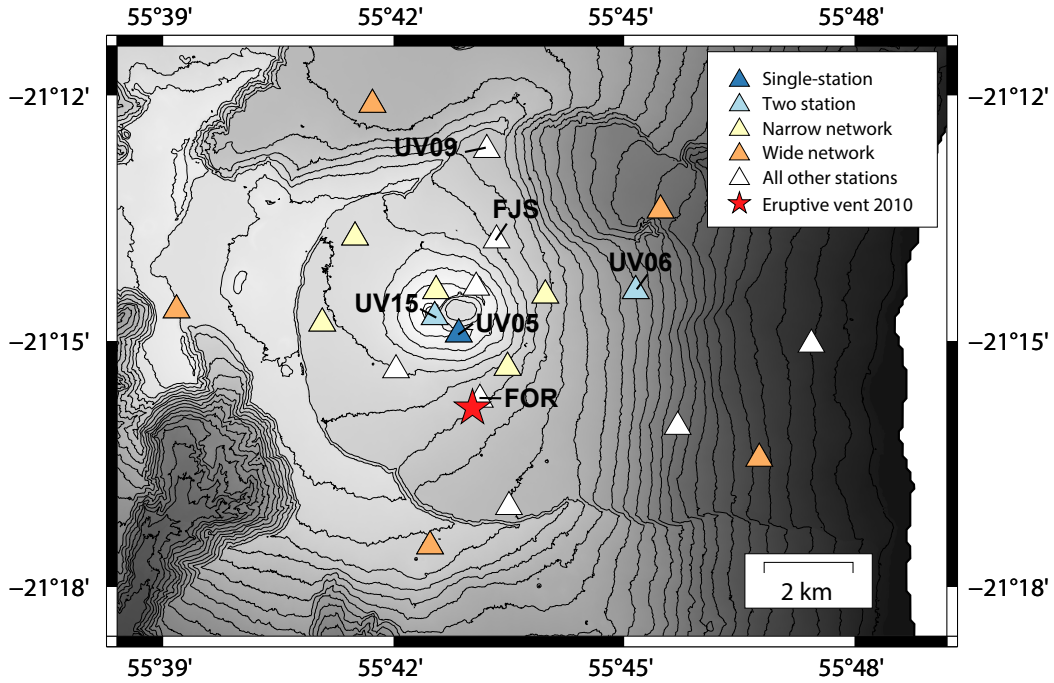

Figure 2.1: UnderVolc Network Seismic Stations (triangles) at Piton De La Fournaise (2009-2011). Different colours correspond to different subsets of stations used to compute the covariance matrix spectral width. Vent location for 2010 proximal eruption (red star) from Roult et al. (2012). Labeled stations are those used in later analysis.

2.2.1 Background theory

The approach introduced by Seydoux et al. (2016a) makes use of the array covariance matrix formed using interstation cross correlations represented in the Fourier domain. The basic idea is the number of independent signals making up the wavefield is reflected by the rank of the covariance matrix (Bartlett, 1954; Lawley, 1956). The more coherent the seismic wavefield, the larger the associated eigenvalues are, with smaller eigenvalues corresponding to background noise (Wagner and Owens, 1996). The distribution of the eigenvalues can be quantified through the calculation of the spectral width σ, defined by Seydoux et al. (2016a), as:

$$\sigma(f) = \frac{\sum_{i=1}^{N} (i-1)\lambda_i(f)}{\sum_{i=1}^{N} \lambda_i(f)} \tag{2.1}$$

where the spectral width is computed as a function of frequency, and $\lambda_i(f)$ represents the eigenvalues arranged in decreasing order. The spectral width then tends to zero when the seismic wavefield is dominated by a single source, and approaches a maximum when non-coherent noise is considered. The value of this maximum depends on a few factors, such as the relative amplitudes of coherent and incoherent noise, the frequency content of the data, and the interstation distances of the seismic array (Seydoux et al., 2016a). In any case, we can use the parameter to better understand the coherence of the seismic wavefield at different frequencies through

time.

The open-source Python Package Covseisnet (v0.4)(Covseisnet, 2020) is used to compute the spectral width of the array covariance matrix. The covariance matrix $C(f)$ is estimated by averaging Fourier cross-spectra matrices over M overlapping subwindows of length δt:

$$C(f) = \frac{1}{M} \sum_{m=1}^{M} u_m(f) u_m^{\dagger}(f) \tag{2.2}$$

where $u_m(f)$ is the vector of data Fourier spectra within each subwindow m and $\dagger$ signifies the Hermitian transpose. Specifically, $u(f)$ is a vector composing of data collected over N seismic stations, taking the form $u(f) = [u_1(f), u_2(f), ..., u_N(f)]^T$. The choices of the number of windows M to average and the length of each subwindow δt, depend on network geometry, frequency range of interest, and temporal resolution (Seydoux et al., 2016a). These are discussed further in the section following. The total duration of the time window used to estimate the array covariance matrix is then $\Delta t = Mr\delta t$ where r is the overlapping ratio. Finally, the array covariance matrix is decomposed into its eigenvalues and eigenvectors, taking the form:

$$C(f) = \sum_{n=1}^{N} \lambda_n(f) \nu_n(f) \nu_n^{\dagger}(f) \tag{2.3}$$

where ν_n are the complex eigenvectors. Note that, because the array covariance matrix is Hermitian, the eigenvalues λ_n are real and positive. The spectral width is then calculated at each frequency following Equation 4.2.

The spectral width measurement was first designed with signal detection in mind (Seydoux et al., 2016a). At volcanoes, lower values of spectral width can be used to detect volcano seismicity that is not easily identified by detection algorithms designed for signals with a sudden onset (Allen, 1982; Baillard et al., 2014; Seydoux et al., 2016a). The approach has, however, also been used towards obtaining a general understanding of the spatial coherence and isotropy of the background seismic wavefield (Seydoux et al., 2016b). For example, Seydoux et al. (2016b) demonstrate — using the U.S. Transportable Array Experiment — that the seismic wavefield cannot be considered equipartitioned even after the application of standard preprocessing techniques to homogenize the wavefield (e.g. temporal normalization).

While we do not neccessarily require exact timing of event onset, as detection algorithms often seek to achieve, understanding the volcano seismicity present at a volcano is useful for a number of reasons. First, volcano seismicity can contaminate cross-correlation functions (Ballmer et al., 2013), reducing the ability to accurately measure seismic velocity changes (Gómez-García et al., 2018). For this reason, studies applying coda wave interferometry at volcanoes may focus a frequency band that does not include volcano seismicity (e.g. Mordret et al., 2010; Cannata et al., 2017). On the other hand, stable volcanic tremor can provide an alternative source of seis-

mic noise in cross-correlation computation (e.g. Donaldson et al., 2017). Thus the identification of volcano seismicity through computation of the spectral width can assist in the decision making towards applying coda wave interferometry at volcanoes.

2.2.2 Data and processing

The Python package Covseisnet (Covseisnet, 2020) provides the tools needed to compute the array covariance matrix and perform spectral width calculations. Two main objects are defined: an ArrayStream object and a CovarianceMatrix object. The ArrayStream class inherits methods from the Obspy Stream class, and is responsible for synchronizing and pre-processing seismic traces recorded over a network. The CovarianceMatrix class is based on the Numpy multidimensional array 'ndarray', with some additional methods for estimating the covariance matrix and computing the spectral width (Covseisnet, 2020). It is within these two objects that the bulk of the workflow is carried out.

We begin by downloading seismic data recorded at Piton de la Fournaise using the temporary UnderVolc seismic network installed from September 2009 to May 2011 (Brenguier et al., 2012; Institut De Physique Du Globe De Paris (IPGP), 2021) using now standard Obspy tools (Krischer et al., 2015), saving the data as day-long MSEED files. These data are available through the French Seismologic and Geodetic Network data center (RESIF). No instrument response correction is performed, as all stations are broadband with a flat response in the frequency ranges of interest. A number of pre-processing steps are then performed prior to computing the covariance matrix. Day-long obspy stream objects containing all seismic traces over the network are split into 6-hour chunks. If an individual 6-hr trace has more than 5% of data missing, it is removed from the analysis. Otherwise, gaps are padded with zeros. The choice to process 6-hour slices instead of day-long traces was made to avoid full days being discarded if gaps were present during only part of the day. Data are then decimated to 25 Hz, detrended and demeaned, and tapered at 0.05% prior to performing spectral whitening and 1-bit temporal normalization. Both normalization steps can be performed using functions built into Covseisnet. Finally, we bandpass the data between 0.01 and 10 Hz, where frequencies above this are of limited use due to minimal depth sensitivity. The array covariance matrix and corresponding spectral width values are computed using functions within the CovarianceMatrix object. The output is a matrix with spectral width values corresponding to the different frequencies through time.

The choice of subwindow length δt and number of subwindows to average M is important in the computation of the array covariance matrix. The subwindow length controls the frequencies that can be resolved, and one should aim to have at least two wavelengths of the lowest frequency of interest (Seydoux et al., 2016a). Noting the primary ocean microseism has a period of approximately 16 seconds, we set $\delta t = 100$ s to ensure this seismic energy is well-recovered. For M, increasing the number of subwindows to average acts to increase the robustness of the covariance matrix. We choose a value of $M = 100$, indicated to be the optimal value for the network configuration at Piton de la Fournaise for the spectral width to converge

to maximum values at each frequency (Seydoux et al., 2016a). This gives a total window length of $\Delta t = 5000$ s, with a 50% overlap between each subwindow.

Different values of subwindow length δt and and number to average M were tested. Smaller subwindow lengths (down to 25s) resulted in less clearly resolved features at lower frequencies, while larger windows reduced the overall temporal resolution. Averaging a smaller number of windows resulted in noisier measurements of spectral width relative to those obtained averaging over 100 windows. At the time scales of interest (days to weeks for comparison with results of coda wave interferometry), no clear difference was observed averaging over less windows. Furthermore, the processing time was far greater for a smaller window number. Thus, a value of $M = 100$, combined with $\delta t = 100$ s, is used throughout this thesis.

An example of the spectral width computed over three months in 2010 at Piton De La Fournaise is shown in Figure 2.2a using the vertical component at all available stations. As indicated earlier, this time period has been analysed previously (Seydoux et al., 2016a; Journeau et al., 2020). Two features dominate the plot: 16 days of low spectral width above 1 Hz, starting mid-October, and a sustained low spectral width at approximately 0.2 Hz. The low spectral width from mid-October relates to eruptive tremor as a result of a proximal eruption beginning 14 October (Roult et al., 2012; Seydoux et al., 2016a). The location of the eruptive vent can be seen in Figure 2.1. At $\sim$0.2 Hz, the persistent low spectral width corresponds to the secondary ocean microseism (Seydoux et al., 2016a). Such strong coherence of the secondary microseism at Piton De La Fournaise is not surprising, being an island station in close proximity to noise sources originating from the Indian Ocean (Davy et al., 2015).

Reduced values of the spectral width across all frequencies can be observed during times where the number of seismic stations available is reduced (Fig. 2.2a). For increased clarity when distinguishing between real and artificial changes in wavefield coherence based on station availability, we normalise each measurement of the spectral width by the maximum value recorded across all frequencies (Fig. 2.2b). This is an imperfect solution, noting that the spectral width does not converge to the same value at all frequencies (Seydoux et al., 2016a). However, for our purposes, applying this normalization significantly reduces the fluctuations associated with changing station numbers while retaining the key features described. We can also now clearly see coherent signals below 0.1 Hz associated with teleseismic arrivals, for example the M7.0 Darfield earthquake on 4 September (local time) (Gledhill et al., 2010) and the M7.8 Mentawai earthquake on 25 October (Lay et al., 2011).

Finally, we compare the normalized covariance matrix spectral width measurements with a spectrogram computed using vertical component seismic data at UV05 station (Figure 2.3). For the spectrogram, seismic traces are resampled to 25 Hz, and merged to day-long traces (padding with zeros in presence of gaps). These are then sliced into 5000 s traces (matching window length of the covariance matrix estimation), de-trended and de-meaned, and a 5% hanning window taper applied at both sides. The Fast-Fourier Transform (Harris et al., 2020) is then applied and the amplitude spectrum recovered through taking the absolute value of the output. From this, we observe that the key features of the spectrogram are also revealed by the spectral width, namely the volcano seismicity and the microseism energy.

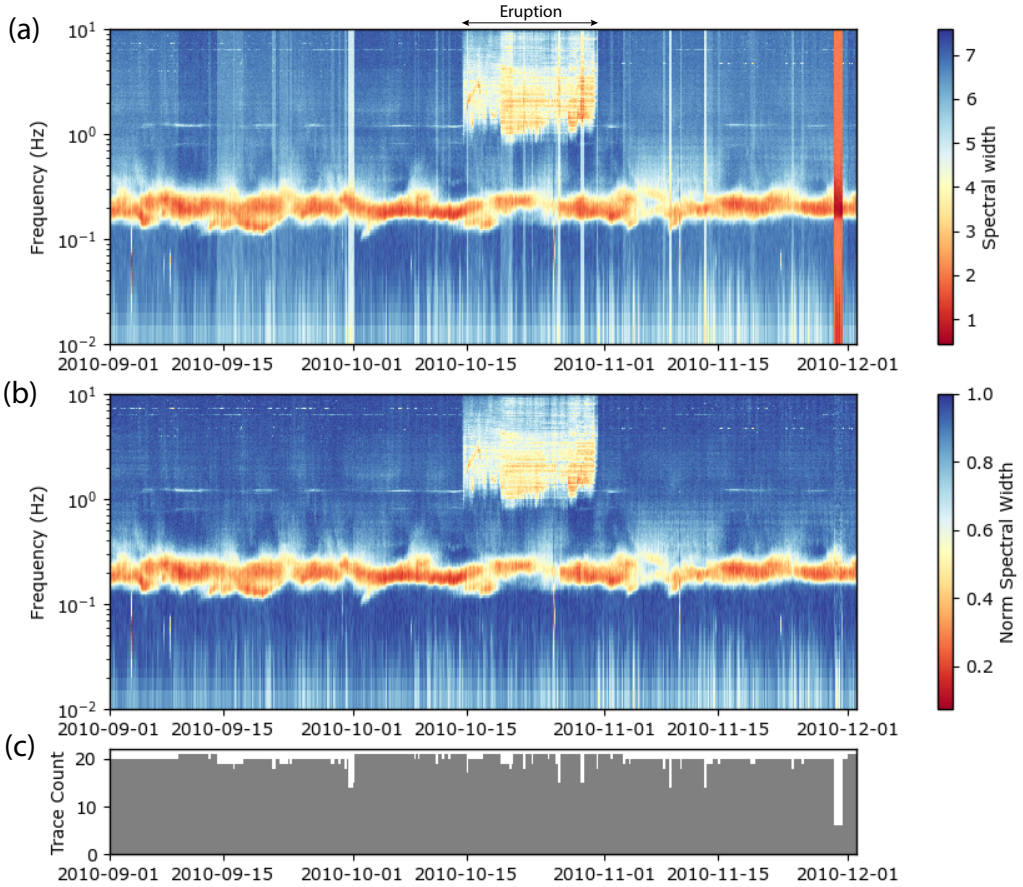

Figure 2.2: Covariance matrix spectral width computed at Piton De La Fournaise using all available Undervolc stations (Z-component only). (a) Spectral width computed between 2010-09-01 to 2010-12-01. (b) Normalized spectral width. (c) Trace count used for spectral width computation, where the trace represents the number of individual frequency spectra used to estimate the covariance matrix. Thus, if using Z-component data only, the trace count corresponds to the number of stations used.

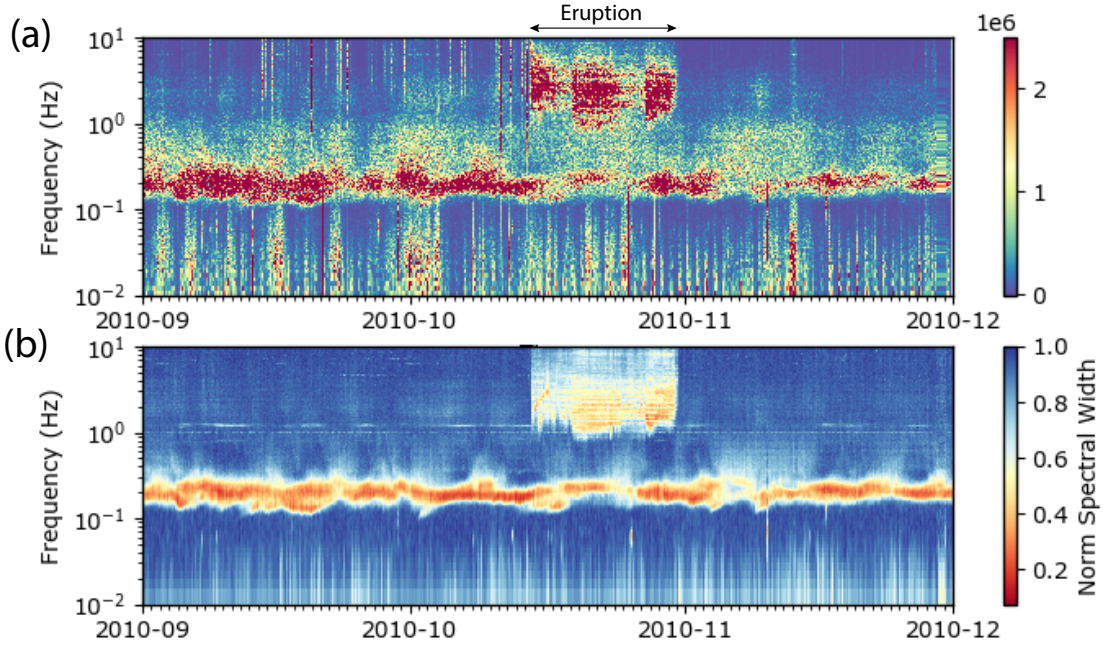

Figure 2.3: Comparison between spectrogram (a) for UV05 station, Z-component and normalized spectral width at Piton de la Fournaise (b) as shown in Figure 2.2.

2.2.3 Testing different network configurations

Different volcanoes have different network configurations, varying in both the number of seismic stations available and the location of sensors relative to volcanic activity. The goal of this section is to evaluate the consistency of the covariance matrix spectral width outputs when processing different subsets of the available stations at Piton de la Fournaise. Specifically, we divide the network into a 'narrow' network (five stations with interstation distances approximately 2–4 km) and a 'wide' network (five stations with interstation distances 6–14 km). Two-station and single-station processing is also tested. These configurations are shown in Figure 2.1.

Normalized spectral width outputs for both the narrow and wide network configurations are shown in Figures 2.4a and 2.4b respectively. In both cases, the previously observed volcano seismicity and secondary microseism are well-recovered by a smaller number of stations. Similarly, teleseismic earthquake activity is also recovered. Some differences are, however, observed in the value of the normalized spectral width for the wide network, with slightly higher spectral width values for both the secondary microseism and volcano seismicity (Fig. 2.4b). This is most clearly observed above 5 Hz, where the decrease in spectral width associated with volcano-seismicity is much less clearly resolved. This is not unexpected, with higher frequencies likely to be less coherent over larger distances due to seismic attentuation.

Finally, we re-compute the spectral width plots using two more examples: one with only two stations (just Z-component data) and the other a single-station using all three-components. These are shown in Figures 2.4c and 2.4d respectively. In the case of the single-station, we are no longer measuring spatial coherence of the wavefield, but rather the component-wise coherence. In essence, the eigenvalues of the covariance matrix formed from a single-station reflect the polarization of the wavefield (Wagner and Owens, 1996). In an ideal situation with isolated signals and white noise, this should be viable as a detection tool (Wagner and Owens, 1996). We once again observe that even with two stations, we recover the same key features as observed with the full network (Fig. 2.4c). The largest differences are observed when comparing single-station spectral width outputs (Fig. 2.4d) to the full network. Volcano seismicity is still clearly observed, whereas the ocean microseism is much less distinguishable from background. High coherence is also observed at the lowest frequencies. We also see component-wise coherent energy above 0.7–0.8 Hz throughout that was not present in previous spectral width plots. It may be that this seismic energy is very local to the station, and therefore not spatially coherent. This is perhaps supported by the observation of much less energy above 1 Hz at UV09 station (Fig. 2.5), located further north. Thus, if the purpose is to understand the spatial coherence of the seismic wavefield, using more than one station is demonstrated to be preferable.

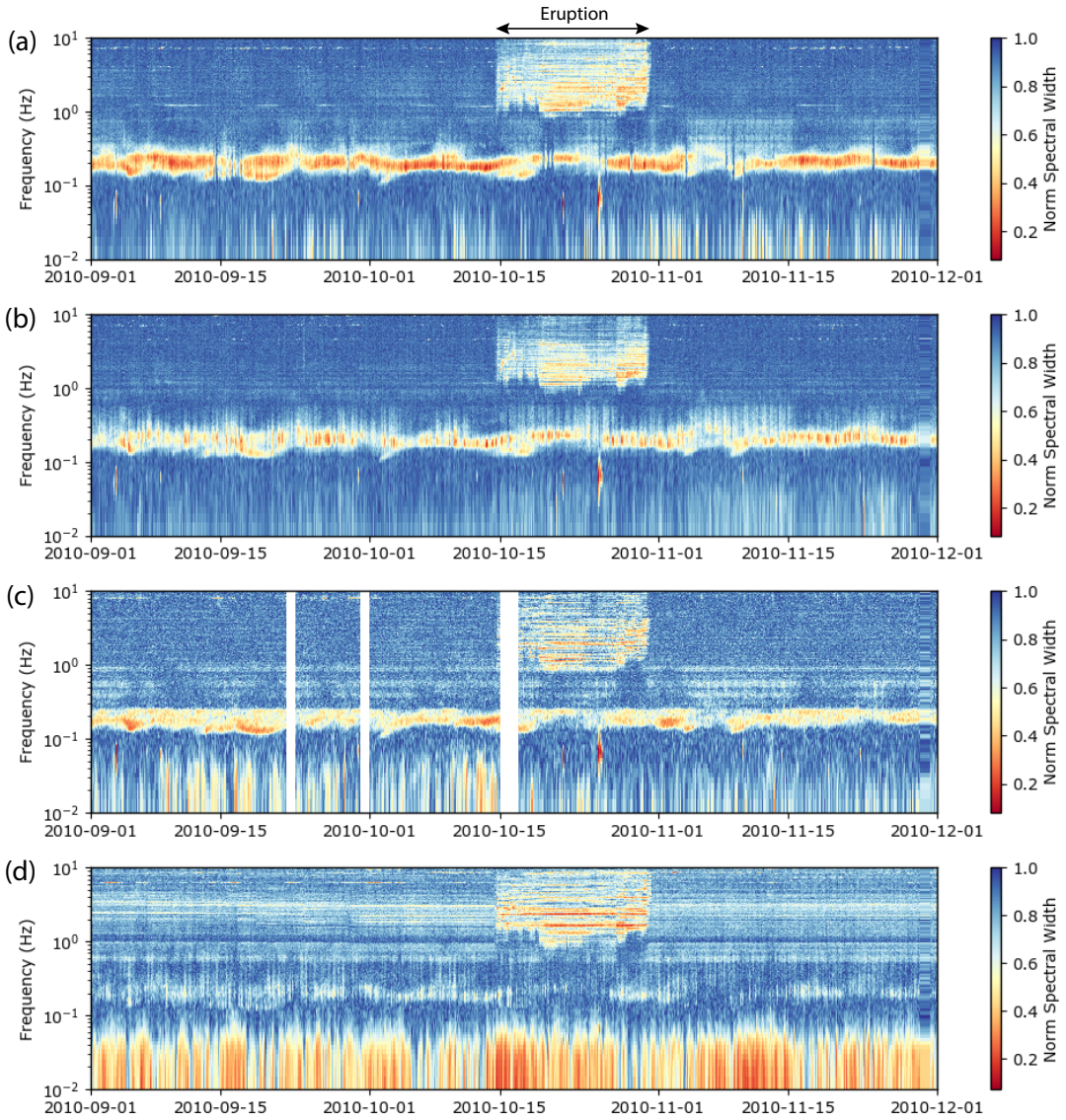

Figure 2.4: Normalized spectral width computed using different subsets of seismic stations, following legend in Figure 2.1, Z-component only. (a) Narrow network. (b) Wide network. (c) Two-station. (d) Single-station.

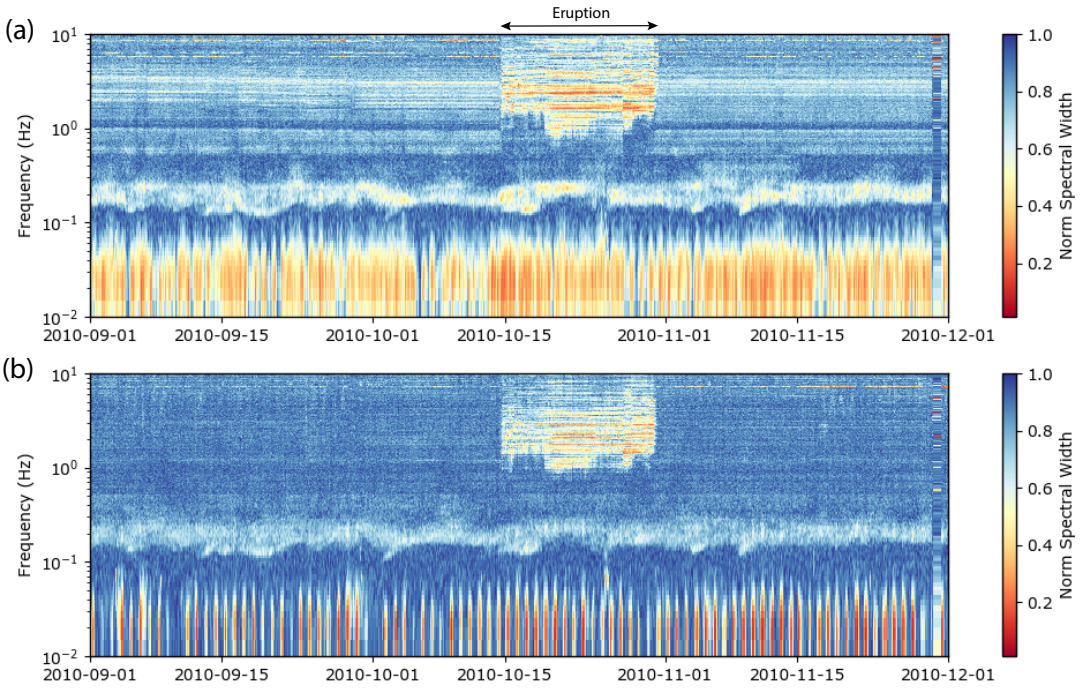

Figure 2.5: Normalized spectral width using three-components of single-station for stations (a) UV05 and (b) UV09, at Piton de la Fournaise volcano.

Lower values of spectral width are regularly observed at the lowest frequencies, with the effect stronger for two-station and single-station approaches. This could relate to both the low rank of the covariance matrix due to a smaller number of seismic traces and the larger wavelength of the signals relative to the interstation spacing. Regarding the rank of the covariance matrix, the smaller number of seismic traces and associated eigenvectors (equal to the number of traces) leads to less robust discrimination between coherent signals and noise. This may then be amplified at the lower frequencies, where the wavelength of the signal is long relative to the interstation distances. In this scenario, it may not be possible to measure oscillations of the wavefield, with noise being seen as coherent (Seydoux et al., 2016b). The reduced influence of low spectral width values at lower frequencies using the wider network (Fig. 2.4b) relative to the narrow network (Fig. 2.4a) may support this. In any case, frequencies below 0.1 Hz are generally of limited value in monitoring studies focused on near-surface processes. Thus, low spectral width values associated with noise at lower frequencies is not so problematic with regards to understanding the seismic wavefield properties above 0.1 Hz.

2.3 Quantitative approaches to assessing cross-correlation quality

In order to maximise the value that can be extracted from covariance matrix spectral width outputs towards coda wave interferometry, it is useful to understand how changes in the spectral width influence the character of cross-correlation functions.

In this section, we describe various techniques used to highlight characteristics of cross-correlation functions. Some of these have been previously applied towards quality control, while others are either new, or not previously applied with quality control in mind.

2.3.1 SNR vs time

The signal-to-noise ratio (SNR) is a common measure used to assess the quality of cross-correlation functions. However, to date, there is no agreed definition for the SNR in coda wave interferometry studies. A common approach uses the amplitude ratio between the direct arrivals and a 'noise' window at a much later lag time (Bensen et al., 2007). The larger this ratio, the higher the SNR. While this could be a useful indicator for studies that make use of the direct arrivals, e.g. ambient noise tomography, coda wave interferometry relies on scattered wave arrivals. Thus, a measure of SNR based on the direct arrival may not be good representation of data quality towards this purpose.

An alternative approach defines the SNR of a stacked cross-correlation function as a function of lag time t, following Larose et al. (2007) and Clarke et al. (2011). The level of the signal $s(N,t)$ is defined as the amplitude envelope of the N-day stacked cross-correlation function $cc(t)$ following:

$$s(N,t) = |cc(t) + iH(cc(t))| \tag{2.4}$$

where $H(.)$ the Hilbert transform. The level of noise $\sigma(t)$ is determined as the amplitude of residual fluctuations between each constituent 1-day cross-correlation function making up the N-day stack following:

$$\sigma(N,t) = \sqrt{\frac{\langle cc(t)^2\rangle - \langle cc(t)\rangle^2}{N-1}}. \tag{2.5}$$

where $< . >$ denotes the average over N 1-day cross-correlation functions. The SNR is then estimated as:

$$SNR(N,t) = \frac{\overline{s(N,t)}}{\overline{\sigma(N,t)}} \tag{2.6}$$

where the overlines indicate smoothing. Both the signal and noise function are smoothed using a Hanning window prior to computing the SNR. Here, we use a window length equal to two times the central frequency of the cross-correlation function in question. This definition of the SNR has been used previously to investigate the stack size required to produce reliable measurements (Clarke et al., 2011), to determine which frequencies to use in their analysis of velocity changes in permafrost (James et al., 2017), and to determine optimal parameters in pre-processing at Whakaari volcano (Yates et al., 2019).

One limitation of this approach is that sudden changes in amplitude of the cross-correlation function can result in reduced values of SNR. This will occur if the 1-day cross-correlation functions being compared have different amplitudes, even if the phases remain the same. One solution to this problem could be to normalize each of the 1-day cross-correlation functions between -1 and 1 prior to computing the SNR (Fig. S2.1). However, this can greatly amplify the SNR of direct arrivals, due to the maximum amplitude of each individual cross-correlation function being likely to occur at, approximately, the same point in time. Normalizing then reduces the variation between 1-day stacks at this specific lag time, lowering the noise signal (Eqn. 2.5) and, therefore, increasing the SNR (Fig. S2.1b and S2.1c).

Furthermore, differences in SNR at different lag times may not reflect the true quality of the signal towards coda-wave interferometry. For example, direct arrivals will usually be assigned a much larger SNR than scattered arrivals due to greater amplitudes in the stacked cross-correlation function. This would occur even if the phases in the coda remain equally consistent through time. Thus, differences between the SNR calculated for direct and scattered arrivals may not be a true indication of the quality of the coda toward monitoring seismic velocity changes.

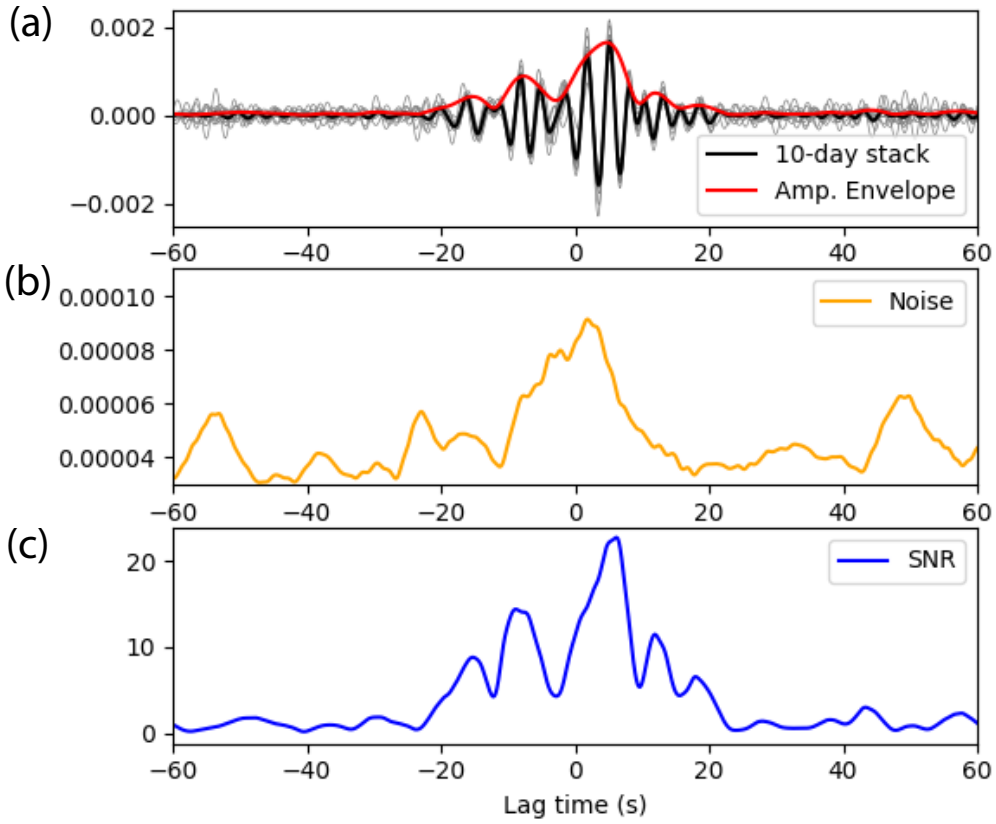

Figure 2.6: Example of SNR computation for 10-day cross-correlation function using station pair UV06-UV08. (a) Ten 1-day cross-correlation functions up to 2010-09-22 are filtered between 0.21-0.42 Hz (light grey lines), stacked (black line), and the amplitude envelope computed (red line). (b) Noise signal computed following Equation 2.5. (c) SNR computed as amplitude envelope divided by Noise.

2.3.2 Phase stack amplitude

An alternative approach to explore the quality of CCFs is to investigate the coherence of the instantaneous phases. We achieve this through computation of the phase stack, defined by Schimmel and Paulssen (1997) as:

$$c(t) = \frac{1}{N}\left|\sum_{j=1}^{N} \exp\left[i\phi_j(t)\right]\right| \tag{2.7}$$

where $\phi_i(t)$ represents the instantaneous phase of the $i = 1...N$ signals (Fig. 2.7). The phase stack in this form originates from the construction of the complex trace $S(t)$ from a seismic trace $s(t)$ following:

$$S(t) = A(t)\exp\left[i\phi(t)\right] \tag{2.8}$$

where $A(t)$ is the time-dependent amplitude. The exclusion of this term in Equation 2.7 means the phase stack being unbiased by amplitude (Schimmel and Paulssen, 1997).

The phase stack has commonly been used to weight the linear stack of cross-correlation functions (e.g. Schimmel and Paulssen, 1997; Schimmel et al., 2011; Zeng et al., 2017). By doing this, greater emphasis is given to phase coherence during stacking rather than just amplitudes. For our purposes, we appreciate the simplicity in its interpretation of the phase coherence of cross-correlation functions. Values of the phase stack lie between zero and one. Zero amplitudes indicate phases combine entirely destructively, while a value of one indicates waveforms are perfectly in phase. Thus the phase stack amplitude provides an easily interpretable measure of phase similarity at different lag times. This can then be combined with the amplitude-based SNR discussed above to have a more complete picture of how stable the cross-correlation functions are through time.

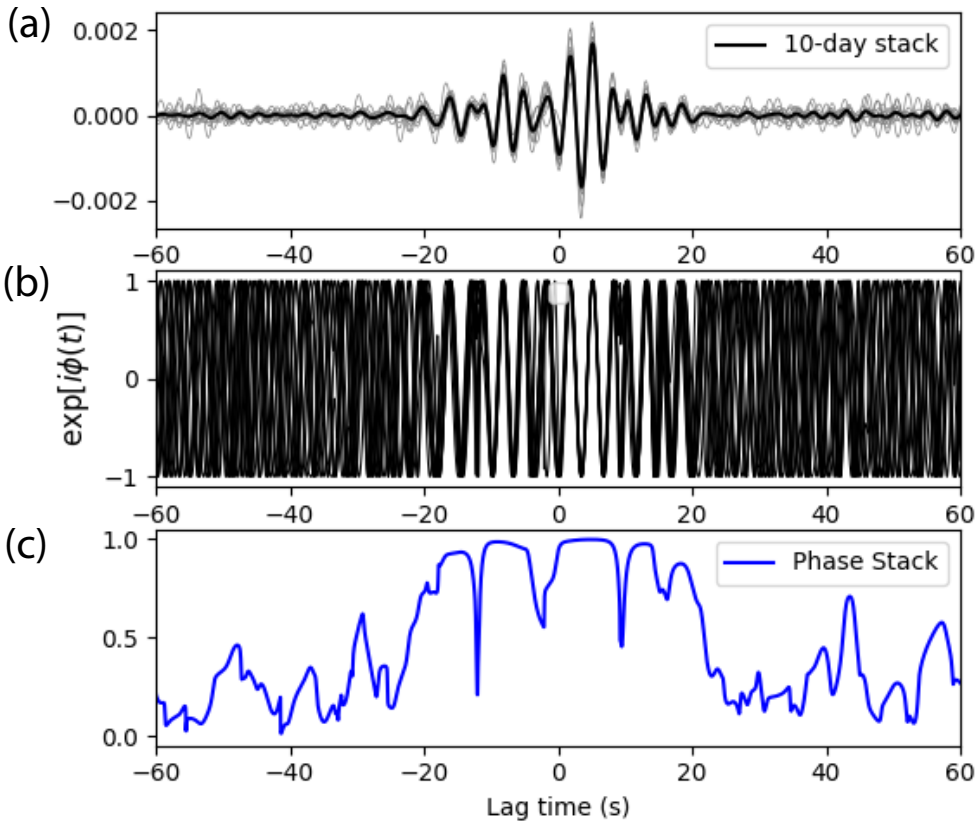

Figure 2.7: Example of Phase Stack computation for 10-day cross-correlation function using station pair UV06-UV08. (a) Starting one-day cross-correlation functions (light grey), with the linear stack of these. (b) Plot of instantaneous phase $\sigma(t)$ of each 1-day cross-correlation function in form $\exp[i\sigma(t)]$. (c) Phase stack, computed by taking the average of absolute values of the instantaneous phases following Equation 2.7.

2.3.3 Amplitude ratio

As discussed in Section 1.2.1, the noise field is rarely uniformly distributed in real-world conditions (Snieder, 2004; Sánchez-Sesma and Campillo, 2006). Instead, there are tendencies for azimuthal biases in the noise strength (Tsai, 2010). This is typically reflected through amplitude asymmetry in the resultant cross-correlation function.

Changes in the symmetry, or asymmetry, of cross-correlation functions may reflect changes in the dominant seismic sources or source location. To assess differences in the amplitude of each side of cross-correlation functions, we compute the ratio between positive (causal) and negative (acausal) sides following:

$$R = \log_2 \left(\frac{\max(cc(t^+))}{\max(cc(t^-))} \right). \tag{2.9}$$

where t^+ represents positive lag times and t^- negative lag times. Taking the ratio to log base 2 means that $R = 0$ corresponds to the same maximum amplitude either side of the cross-correlation function. Furthermore, an increase or decrease in R of 1 corresponds to a change in the amplitude ratio by a factor of 2. For example, R values of 1 or -1 indicate an amplitude difference of factor two, and R values of 2 or -2 a difference of factor four. Note, in order to exclude spurious arrivals or pre-cursory noise that arrives prior to the P-wave, the time before a defined minimum lag is excluded both sides of the cross-correlation function prior to finding the maximum.

2.4 Characteristics of cross-correlation functions

In the following section, the characteristics of cross-correlation functions are assessed at different volcanoes using the measures introduced in Section 2.3. These can be compared with the outputs of network covariance matrix analysis to determine how key features of the spectral width output influence resultant cross-correlation functions.

Our approach begins with broadband cross-correlation functions computed in MSNoise (Lecocq et al., 2014). Starting with raw seismic data, the data are decimated to 25 Hz and bandpassed between 0.005 and 12 Hz. Spectral whitening is then applied, followed by 1-bit temporal normalization. Cross-correlation functions are computed in 1800 second time slices, and stacked to 1-day functions using a linear stack.

From the broadband cross-correlation functions, we define a number of narrow bandpass filters based on a central frequency, with an octave between the low and high pass frequencies (Table S2.1). Then, for each filter, we bandpass the 1-day cross-correlation functions prior to computing the SNR, phase stack, and amplitude ratio in a moving window of ten cross-correlation functions. Thus the value recorded on a particular day, for N-day stacks, uses 1-day stacks of that day and the $N-1$ days before it.

To determine a single value for the SNR and phase stack amplitude for a given day and frequency, an average is computed over a defined lag time window. We define the length of the window to be fifteen cycles of the central period of each filter. The minimum lag is defined based on a minimum ballistic wave speed and interstation distance. Unless otherwise stated, we use a minimum wave speed of 0.5 km/s to ensure that measurements are within the coda.

Five volcanoes are analysed; these are (1) Piton de la Fournaise, (2) Kilauea, (3) Ruapehu, (4) Stromboli, and (5) Whakaari. For each, the spectral width and the various measures are computed during the same period.

2.4.1 Piton de la Fournaise

Focusing on the same time period in section 2.2.2, we first look at the characteristics of cross-correlation functions recorded at two pairs of seismic stations. The first of these (UV06-UV15, Fig. 2.8) corresponds to the pair of stations used for covariance matrix analysis in Figure 2.4c. The second pair (FJS-FOR, Fig. 2.10) is approximately perpendicular to this one.

Beginning with station-pair UV06-UV15, one of the key changes occurs during volcanic tremor concurrent with the October 2010 eruption (Fig. 2.8). This is most clearly represented by a sign change in the amplitude ratio of cross-correlation functions computed (Fig. 2.8d). Note, the shifted appearance of this feature relative to the spectral width output reflects the 10-day moving window used. Both the SNR (Fig. 2.8b)and phase stack amplitude (Fig. 2.8c) show an initial decrease above 1 Hz accompanying the start of the tremor. They then increase to approximately pre-tremor values before decreasing again momentarily at the cessation of tremor. The decreases either side of the tremor reflect differences between the cross-correlation functions computed inside and outside of volcanic tremor (Fig. 2.9). Specifically, different dominating seismic sources within the same 10-day window result in greater variation of 1-day functions used to compute the SNR and phase stack amplitude. Thus, both of these parameters decrease until the window contains only cross-correlation functions of similar character. This is seen clearly in Figure 2.9, where a full recovery of SNR and phase similarity is only recorded approximately 10-days after the tremor has ended (i.e. the length of the moving window).

Outside of the period of volcanic tremor, we observe consistently high SNR and phase stack amplitudes between approximately 0.7–1.5 Hz. This could reflect a coherent seismic source just above 1 Hz, within a very narrow frequency range (Fig. 2.8a). If so, the increase in SNR and phase stack coherence over a wider frequency band may reflect the use of overlapping bandpass filters (Table S2.1). We also observe that the increase in SNR above 0.7 Hz is accompanied by a change in the character of the amplitude ratio. Specifically, the ratio is almost entirely positive above 0.7 Hz relative to a mix of positive and negative ratios at lower frequencies.

The same measures are computed for the second station-pair (FJS-FOR) at Piton De La Fournaise (Fig. 2.10). The different orientation of these stations is reflected clearly by the change in sign of the amplitude ratio relative to the other pair (Fig.

2.10d). Many of the features that we observe in the SNR and phase stack amplitude are, however, similar to those observed at the previous station-pair. This suggests that we can assess the primary characteristics of the cross-correlation functions recorded over a network using a small number of station-pairs. Similarly, the same features are observed when examining the phase stack amplitude computed for single-station cross-components at UV05 station (Fig. 2.11). Note, however, that in the case of station-pair FJS-FOR, a slightly broader range of frequencies above 0.7 Hz have high phase similarity (Fig. 2.10c) relative to station-pair UV06-UV15 and single-station UV05 results.

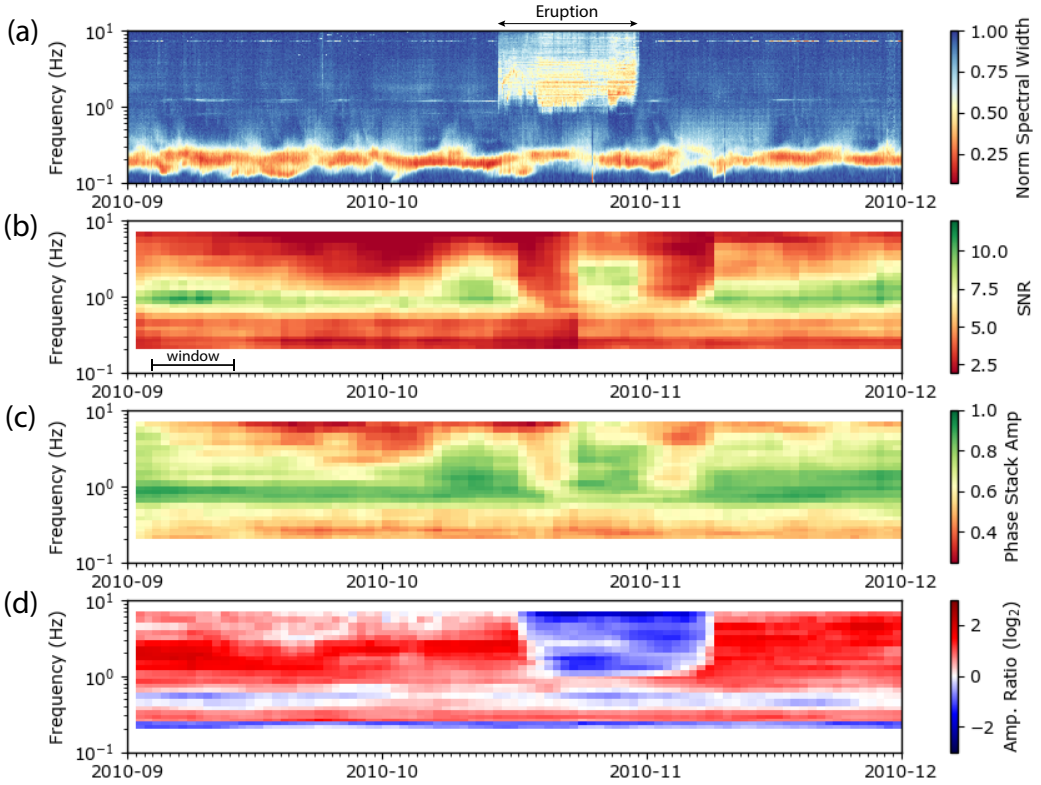

Figure 2.8: Measurements of cross-correlation characteristics at Piton De La Fournaise for a single station-pair (UV06-UV15), computed using 10-day moving window. For SNR and phase stack amplitude, the average value is taken between minimum lag defined by 0.5 km/s ballistic arrival speed and maximum lag equal to the minimum lag plus 15 cycles of central period of the bandpass filter. (a) Normalized covariance matrix spectral width as computed over full network. (b) SNR. (c) Phase stack amplitude. (d) Amplitude ratio.

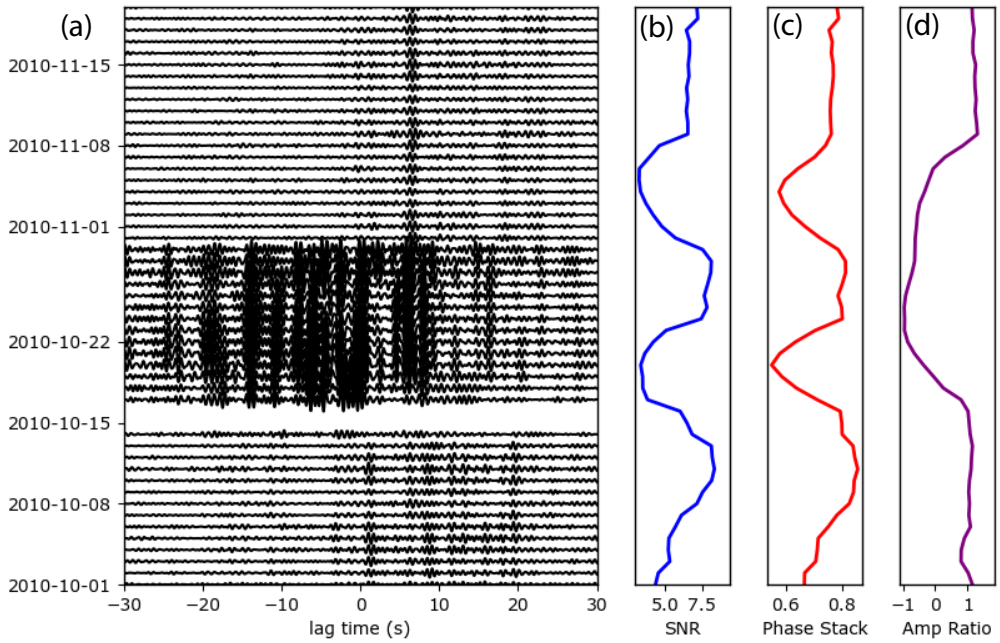

Figure 2.9: Comparing cross-correlation functions filtered between 1–2 Hz for station-pair UV06-UV15 with average SNR, phase stack amplitude, and amplitude ratio within same frequency band (a) 1-day stacked cross-correlation functions through time. (b) SNR. (c) Phase stack amplitude. (d) Amplitude Ratio.

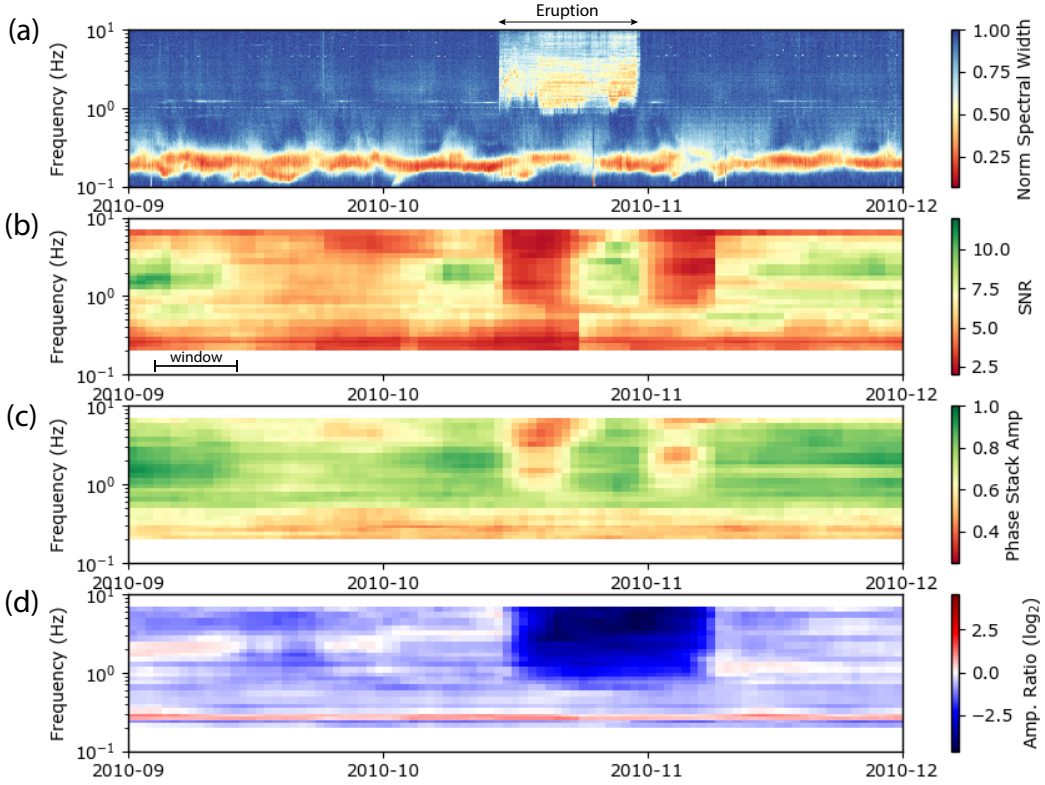

Figure 2.10: Measurements of cross-correlation characteristics at Piton De La Fournaise for a single station-pair (FJS-FOR), computed using 10-day moving window. For SNR and phase stack amplitude, the average value is taken between minimum lag defined by 0.5 km/s ballistic arrival speed and maximum lag equal to the minimum lag plus 15 cycles of central period of the bandpass filter. (a) Normalized covariance matrix spectral width as computed over full network. (b) SNR. (c) Phase stack amplitude. (d) Amplitude ratio.

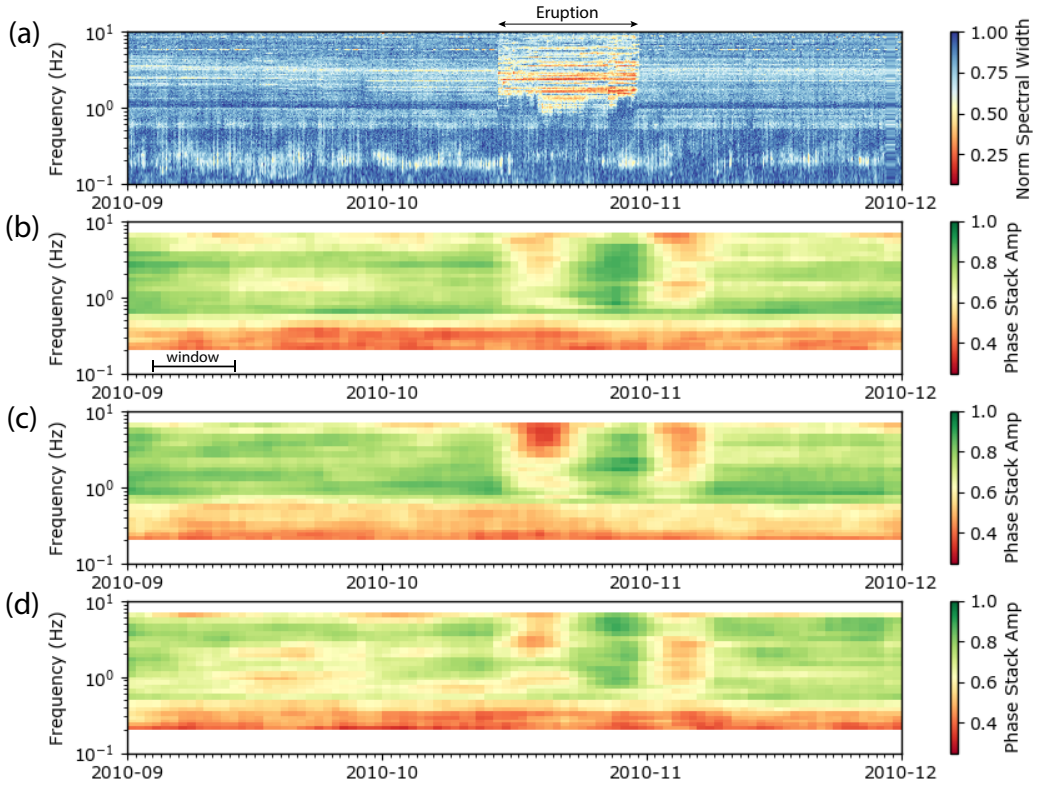

Figure 2.11: Measurements of cross-correlation characteristics for a single-station (UV05). (a) Normalized covariance matrix spectral width. (b) phase similarity EN component. (c) phase similarity EZ component. (d) phase similarity NZ component. All phase similarity measures represent average between window defined by minimum lag of 10 s and width of 15 cycles of the central period.

Finally, the phase similarity of cross-correlation functions are examined as a function of lag time and frequency for both station-pairs (Figure 2.12). The phase stack amplitude at a given lag time and frequency represents the average over a defined date range. Comparing pre-eruption with co-eruptive cross-correlation functions shows stronger phase similarity at later lag times during the eruption at frequencies above 1 Hz (Fig. 2.12). For the pair UV06–UV15, with an interstation path approximately perpendicular and offset from the eruption site, increases in phase similarity are observed at both positive and negative lag times (Fig. 2.12b). For the pair FJS-FOR, higher phase similarity is mostly observed at negative lag times, reflecting energy traveling from FOR to FJS station (Fig. 2.12d). This is in agreement with the close proximity of FOR station to the eruption site.

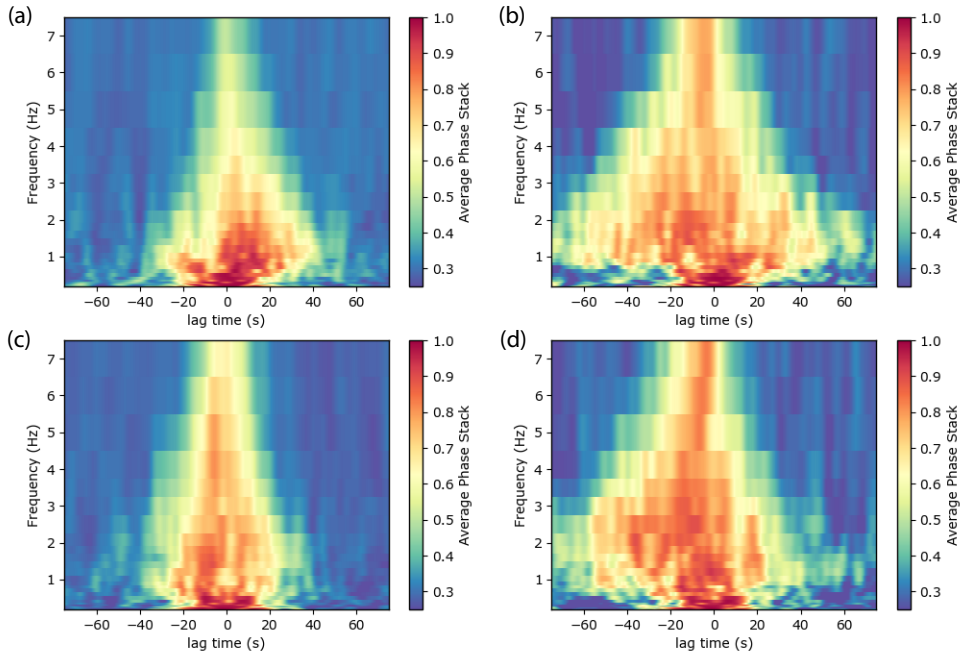

Figure 2.12: Phase similarity as a function of frequency and lag-time for two station-pairs shown in Figures 2.8 and 2.10. Values computed as average between defined date ranges representing tremor (2010-10-25 to 2010-11-01) and non-tremor (2010-09-01 to 2010-10-15) periods. (a) UV06-UV15, non-tremor period (b) UV06-UV15, tremor period (c) FJS-FOR, non-tremor period. (d) FJS-FOR, tremor period.

2.4.2 Kilauea

The dataset at Kilauea provides the opportunity to examine cross-correlation functions used previously to compute seismic velocity changes in the frequency band of volcanic tremor (Donaldson et al., 2017). As with Piton de la Fournaise, covariance matrix analysis of the seismic wavefield has been performed previously for this volcano (Soubestre et al., 2021). We repeat this analysis for the 2018 eruptive episode using ten broadband stations at the summit of Kilauea, following the processing described in Section 2.2.2. The final spectral width output is shown in Figure 2.14.

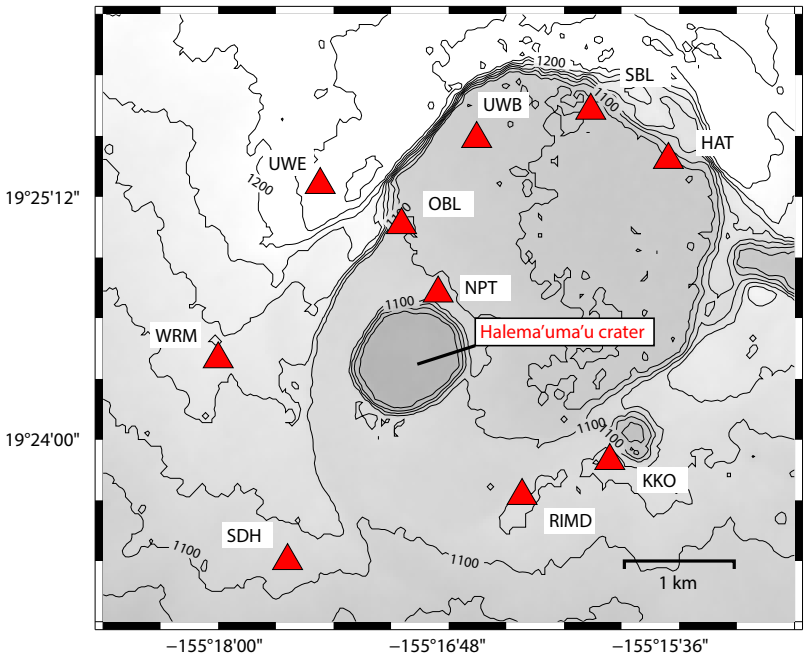

Figure 2.13: Map of seismic stations at Kilauea volcano, showing ten broadband (red triangles) stations used in processing spectral width.

Features of the seismic wavefield at Kilauea in 2018 can be approximately split into three phases (Soubestre et al., 2021). The first period, prior to the start of activity in May, corresponds to a period of relative stability. As with at Piton de la Fournaise, energy corresponding to the oceanic microseims is visible in the 0.1–0.3 Hz frequency band. Low spectral width values at the same frequency were interpreted by Soubestre et al. (2021) to be long-wavelength energy — coherent over the whole network — masking the oceanic microseism. However, the observation here of two distinct bands of low spectral width that vary with time (not seen in Soubestre et al. (2021)) suggests the low spectral width values more likely correspond to oceanic microseisms. Differences in pre-processing between our study and Soubestre et al. (2021) may be responsible for these observations (for example, they apply no 1-bit temporal normalization). Above approximately 0.5 Hz, we observe reduced spectral width values (Fig. 2.14) associated with stable volcanic tremor (Donaldson et al., 2017; Soubestre et al., 2021). This appears strongest within the 0.5–1.0 Hz frequency band, but exists also to higher frequencies.

The second phase, from May 5–27, is associated with disappearance of the lava lake and beginning of the caldera collapse (Soubestre et al., 2021). During this period, gliding tremor is observed over a broad frequency range (between approximately 0.2–3 Hz, Fig. 2.14). Finally, in phase three (from late-May/early-June), the previously clear tremor is absent. Further, slightly reduced spectral width values can be seen in the 0.1–0.5 Hz frequency band (most visible beteen 0.3–0.5 Hz). This may match the semi-regular collapse events within the Kilauea caldera observed by Soubestre et al. (2021). Further details for all three phases can be found in the many studies on this eruptive episode (e.g. Olivier et al., 2019; Soubestre et al., 2021; Hotovec-Ellis et al., 2022; Muzellec et al., 2023).

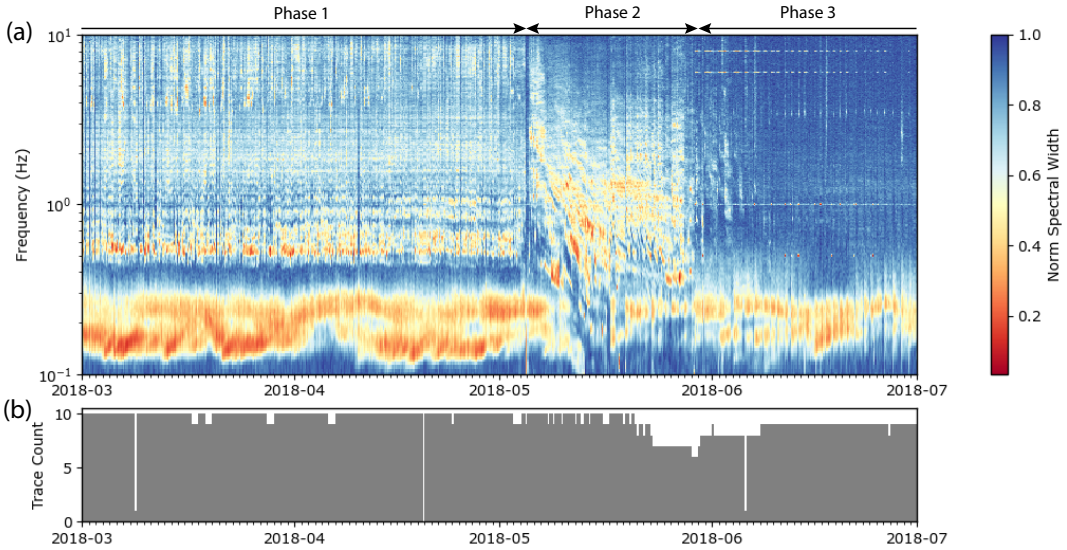

Figure 2.14: Normalized spectral width using ten stations in close proximity to Kilauea caldera (Z-component), shown in Figure 2.13. (a) Spectral width. (b) Trace count.

As before, the features of the seismic wavefield are compared with cross-correlations from a pair of stations (Figure 2.15). The highest values of SNR (Figure 2.15b) and phase stack amplitude (Figure 2.15c) are observed within the frequency band of volcanic tremor, between approximately 0.7–2.0 Hz during the pre-eruptive phase. Outside of tremor frequencies, relatively high values of phase stack amplitude (greater than approximately 0.7) are also observed, suggesting high similarity across a broad frequency range pre-eruption. Similarly, the amplitude ratio of cross-correlations remains relatively stable pre-eruption across all frequency bands (2.15d). Note, the amplitude ratio also clearly differentiates between the oceanic and tremor sources, with negative amplitude ratios above 0.4 Hz and positive amplitudes below.

The start of eruptive activity, and the period of gliding tremor, is accompanied by significant changes in the SNR, phase stack amplitude, and amplitude ratio. Drops in the SNR and phase stack amplitude are observed across all frequencies, with the previously high values associated with the tremor activity disappearing completely. Similarly, the previously stable pre-eruptive amplitude ratios are disrupted, with the original positive ratio below 0.4 Hz becoming negative for 1–2 weeks (Fig. 2.15d). Thus, the eruptive activity, and associated collapse events, clearly contributes to an unstable seismic wavefield which, in turn, results in reduced similarity in the cross-correlation functions.

The final phase, from late-May/early-June, shows a recovery in the phase amplitude at frequencies below approximately 0.5 Hz (Fig. 2.15c). Similarly, amplitude ratios progressively return to approximately pre-eruption values (Fig. 2.15d). This matches closely the oceanic microseism observed in spectral width plots, which appeared partially over-printed during the May activity (Fig. 2.15a). Above 0.5 Hz, we observe that the SNR and phase stack amplitudes remain low relative in the

absense of the original tremor source. Similarly, amplitude ratios do not return to pre-eruptive values. Above $\sim$ 1.5 Hz, very low ratios are observed indicating minimal asymmetry in cross-correlation functions. Between $\sim$0.7–1.5 Hz however, the amplitude ratio becomes positive. This represents a large change from pre-eruptive values and, thus, suggests a different seismic source is now dominating within this frequency band.

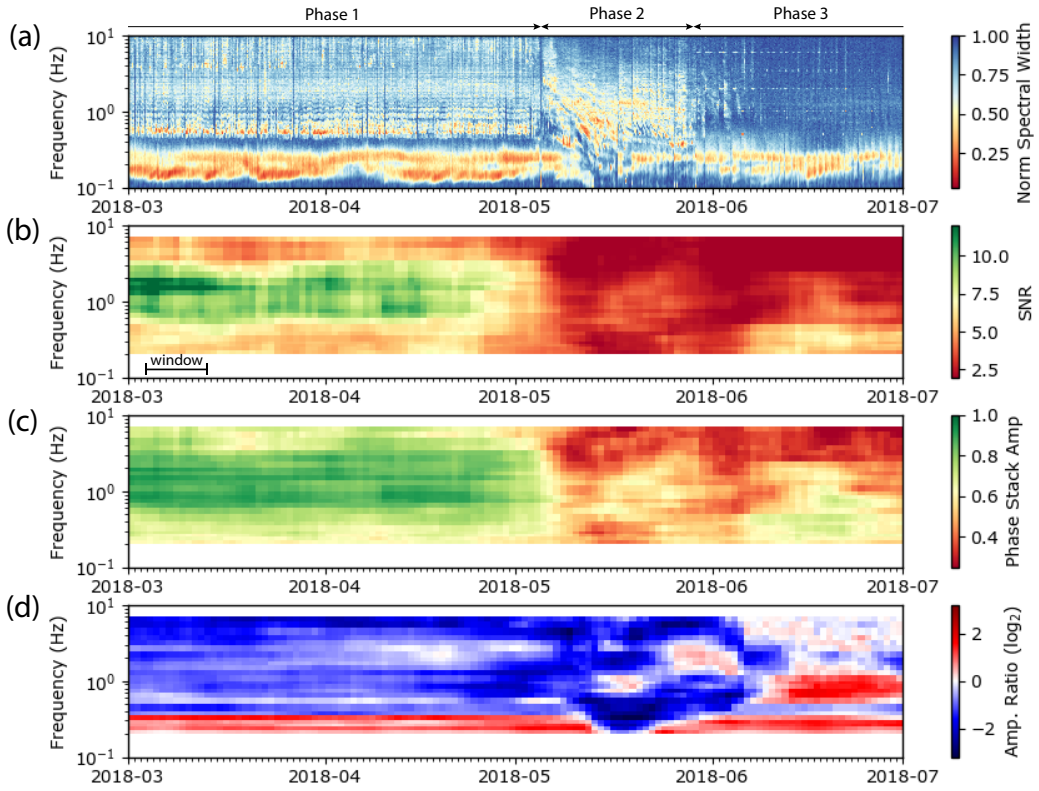

Figure 2.15: Measurements of cross-correlation characteristics at Kilauea for a single station-pair (HAT-OBL), computed using 10-day moving window. For SNR and phase stack amplitude, the average value is taken between minimum lag defined by 0.5 km/s ballistic arrival speed and maximum lag equal to the minimum lag plus 15 cycles of central period of the bandpass filter. (a) Normalized covariance matrix spectral width as computed over full network. (b) SNR. (c) Phase stack amplitude. (d) Amplitude ratio.

Differences during the three different phases are also clearly visible when examining the phase similarity of cross-correlation functions at different lag times (Fig. 2.16). During the tremor pre-eruption (Fig. 2.16a), we observe that cross-correlation functions above approximately 0.5 Hz are similar in phase out to much later lag times than when the tremor is absent (Fig. 2.15c). Below 0.5 Hz, however, differences are minimal pre- and post- eruption. This is consistent with the increase in phase similarity observed in Fig. 2.15c at the start of phase three. During phase two, where gliding tremor is observed, there is still some energy out to later lag times relative to post-eruption, but it is of a lower phase similarity (Fig. 2.16b). The disruption at lower frequencies is also observed, with reduced phase similarity below 0.5 Hz relative to other time periods.

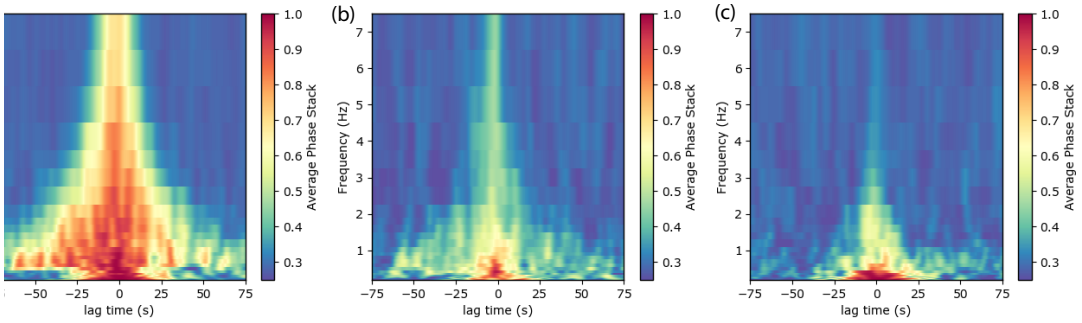

Figure 2.16: Phase similarity as a function of frequency and lag-time for station-pair (HAT-OBL), as shown in Figure 2.15. Values computed as average between defined date ranges corresponding to (a) pre-eruption (2018-03-01 to 2018-05-01). (b) co-eruption(2018-05-05 to 2018-05-27). (c) post-eruption (2018-06-01 to 2018-07-01).

2.4.3 Mt. Ruapehu

Ruapehu is another volcano, like Kilauea, that experiences regular volcanic tremor lasting for many months (Sherburn et al., 1998). In particular, tremor with a dominant spectral peak between 1.8 and 2.3 Hz is the most common at the volcano (Latter and Balsillie, 1984; Hurst and Steven, 1993; Hurst, 1998; Sherburn et al., 1998).

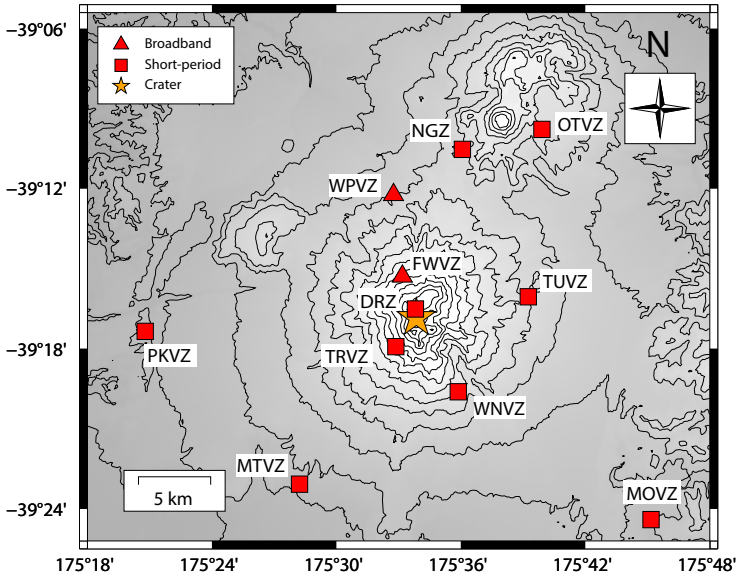

Figure 2.17: Network around Mt Ruapehu volcano, consisting of two broadband and nine short-period sensors.

The network around Ruapehu is more sparse relative to other volcanoes in this chapter, with larger interstation distances (Fig. 2.17). For example, some of the stations

are more than 10 km away from the volcano. It is therefore unclear whether it is preferable to include these more distant stations towards understanding the seismic wavefield in closer proximity to the volcano. Thus, we test multiple configurations when computing the spectral width. These are: (1) all nine short-period stations shown in Fig. 2.17, (2) the four short-period stations closest to the Mt Ruapehu (DRZ, TRVZ, TUVZ, WNVZ), and (3) DRZ single-station. The exclusion of broad-band stations FWVZ and WPVZ allows for measurements to be made without correcting for instrument response.

At frequencies above ∼1.0 Hz, we observe many periods of reduced spectral width likely associated with volcanic tremor (Fig. 2.18). This is observed using all network configurations, suggesting propagation over the wider network. However, clearly distinct peaks are better resolved using stations closer to the volcano (Fig. 2.18b) or DRZ single-station (Fig. 2.18c). Most of this energy is within 1–3 Hz, likely reflecting the usual tremor observed at the volcano. Unlike at Kilauea, however, the tremor does not appear to be continuous in time during the 2005–2009 period, consistent with usual observations at the volcano. For example, there appears to be reduced, or very little tremor, on multiple occasions in the second half of 2006. Similarly, the tremor appears to have reduced energy mid-2007.

At lower frequencies, we see near-continuous energy between approximately 0.1 and 0.7 Hz. While this appears less distinct relative to the previous examples (Figs. 2.2 & 2.14, this could relate to the short-period instrumentation, which has a natural frequency at 1 Hz. Furthermore, both Piton de la Fournaise and Kilauea are island volcanoes, whereas Ruapehu lies at a greater distance from the coast.

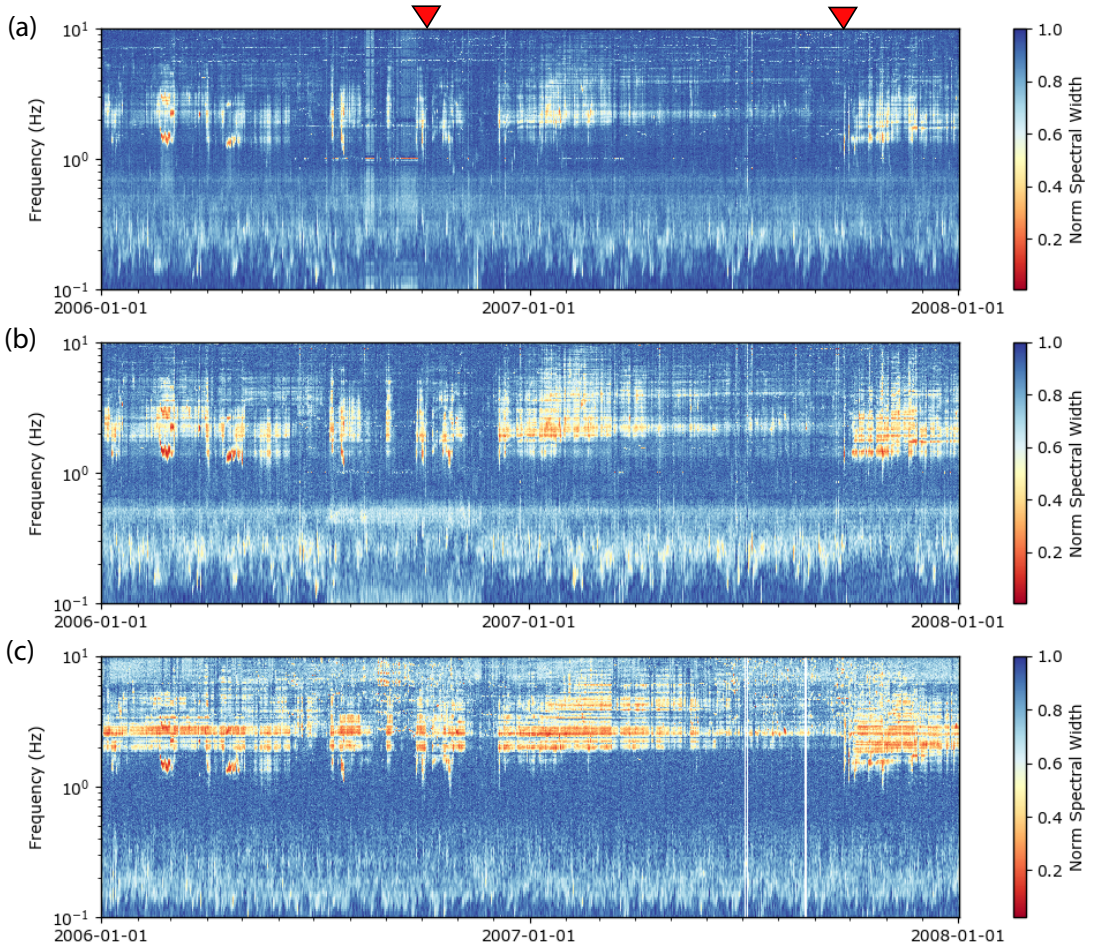

Figure 2.18: Spectral width computed at (a) nine short-period stations. (b) four short-period stations closest to the volcano/crater (Z-component). (c) three components of FWVZ station. All stations are shown in Figure 2.17.

Examining the characteristics of cross-correlation functions for the pair of broadband stations FWVZ-WPVZ again shows a consistency between the times and frequencies of lower spectral width and increases in the SNR and phase similarity (Fig. 2.19). In particular, volcanic tremor appears to contribute to larger values of phase similarity (Fig. 2.19c), with these episodes especially visible when examining the amplitude ratios (Fig. 2.19d). Similarly, we observe that there is higher phase similarity at later lag times during tremor relative to non-tremor episodes (Fig. 2.20). Below the frequency of tremor, we also see elevated values of SNR and phase similarity between approximately 0.4–0.8 Hz.

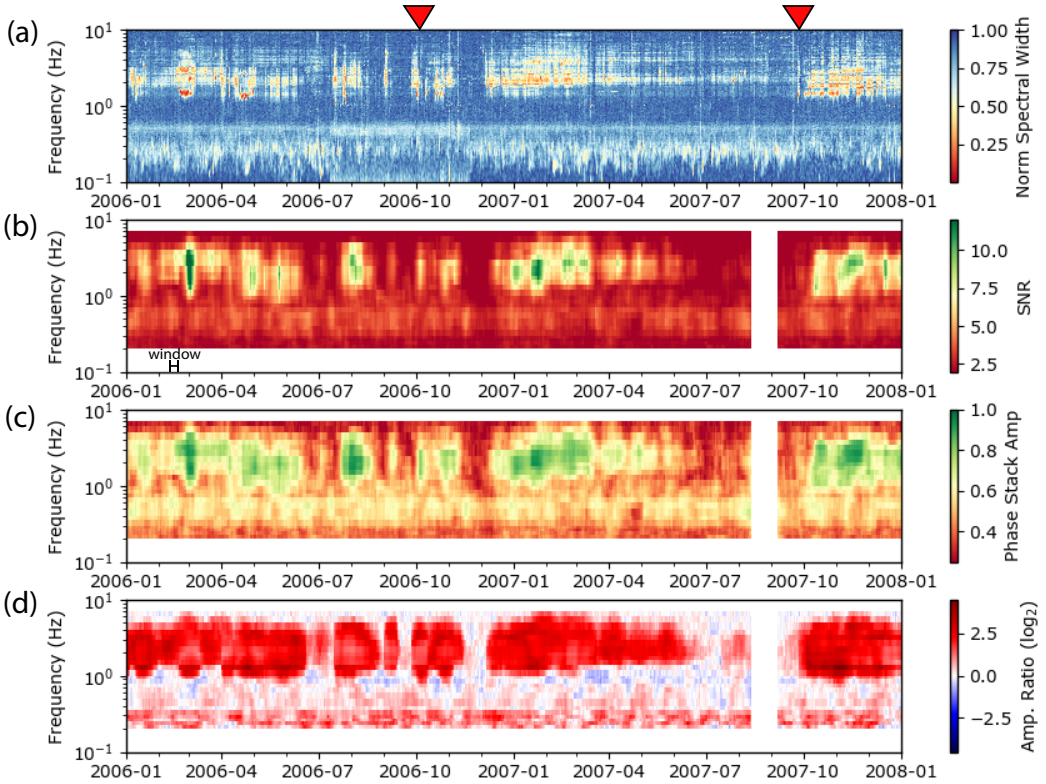

Figure 2.19: Measurements of cross-correlation characteristics at Mt Ruapehu for a single station-pair (FWVZ-WPVZ), computed using 10-day moving window. For SNR and phase stack amplitude, the average value is taken between minimum lag defined by 0.5 km/s ballistic arrival speed and maximum lag equal to the minimum lag plus 15 cycles of central period of the bandpass filter. (a) Normalized covariance matrix spectral width as computed over full network. (b) SNR. (c) Phase stack amplitude. (d) Amplitude ratio.

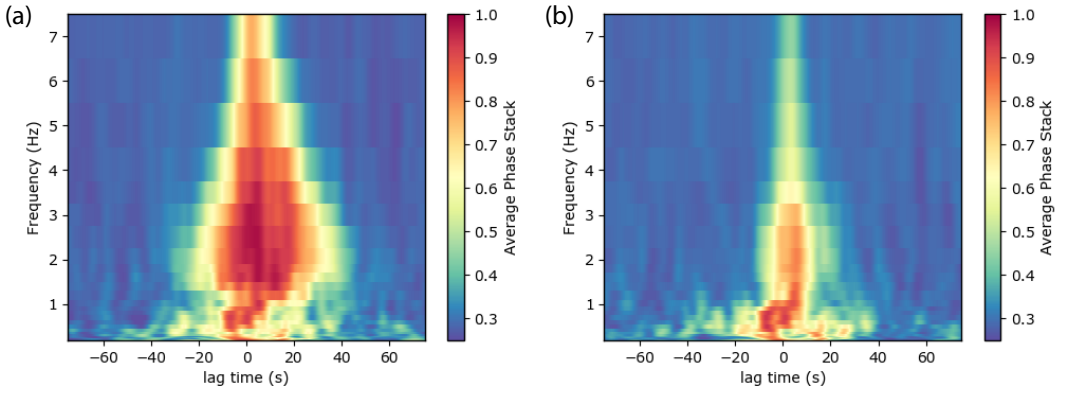

Figure 2.20: Phase similarity as a function of frequency and lag-time for station-pair (FWVZ-WPVZ), as shown in Figure 2.19. Values computed as average between defined date ranges corresponding to (a) tremor (2007-01-01 to 2007-03-01) (b) reduced tremor (2007-08-01 to 2007-10-01).

As with at Piton de la Fournaise, we can verify that the characteristics of single-station cross-components at Ruapehu reveal similar features. Figure 2.21 shows the phase similarity for the three cross-component cross-correlation functions at FWVZ station. Once again we observe high levels of similarity during times of volcanic tremor. Furthermore, the phase similarity remains elevated during weaker tremor episodes in comparison to those measured for cross-correlation functions at FWVZ-WPVZ. Of interest, when examining the amplitude ratios between the different components, we find there is less consistency through time than observed using the pairs of stations (Fig. 2.22). For example, within the frequency of tremor, component-pairs EN and EZ appear to have negative ratios below approximately 2 Hz and positive ratios between 2–4 Hz (Fig. 2.22b and 2.22c). Furthermore, from mid-2006, the ratio is predominantly negative during tremor episodes. This does not appear to be unique to Ruapehu, with similar variability also observed for Piton de la Fournaise single-station amplitude ratios (Fig. S2.3).

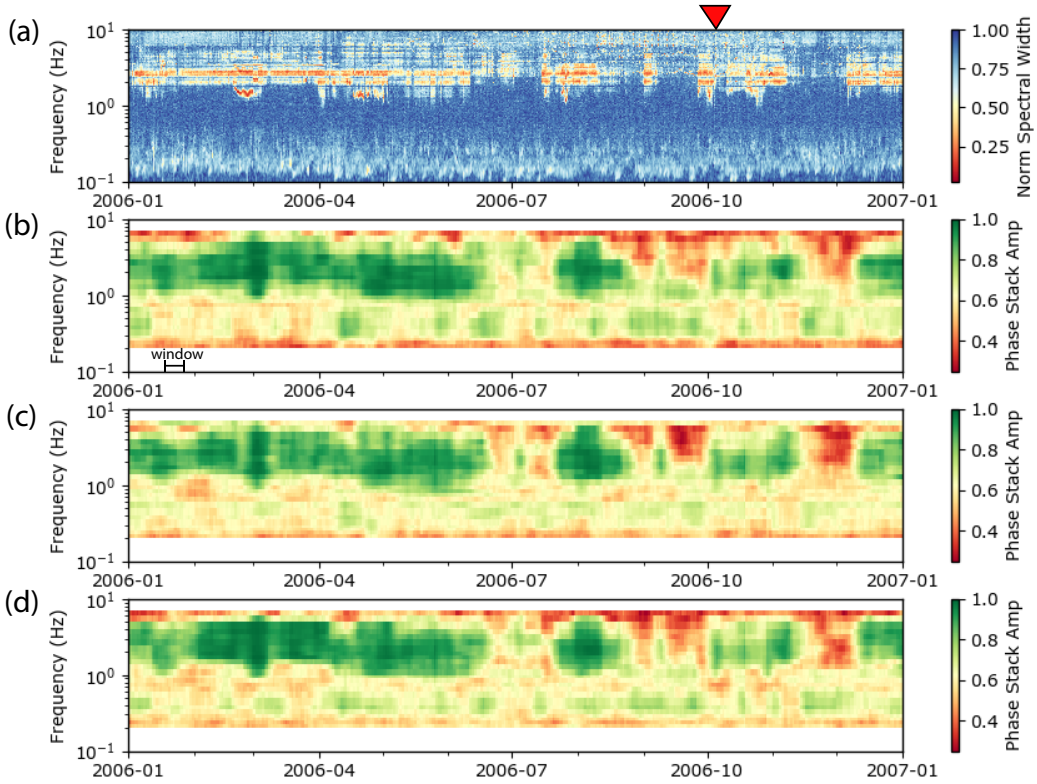

Figure 2.21: Measurements of phase similarity for a single-station (FWVZ). (a) Normalized covariance matrix spectral width. (b) phase similarity EN component. (c) phase similarity EZ component. (d) phase similarity NZ component. All phase similarity measures represent average between window defined by minimum lag of 10 s and width of 15 cycles of the central period.

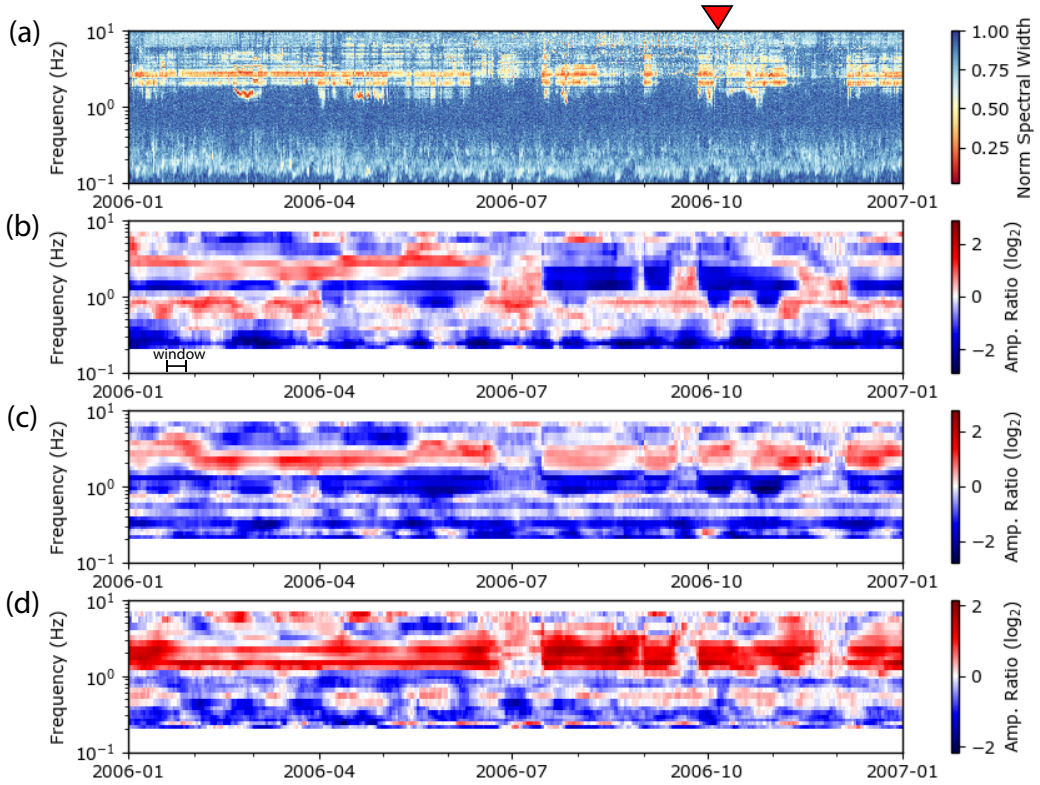

Figure 2.22: Amplitude ratios for a single-station (FWVZ). (a) Normalized covariance matrix spectral width. (b) amplitude ratio for EN component. (c) amplitude ratio for EZ component. (d) amplitude ratio for NZ component.

To better visualise how tremor is propagating over the larger network of Ruapehu, a station pair at greater distance from the volcano (NGZ-TUVZ) — with a larger interstation distance (approximately 11 km)— can be examined (Fig. 2.23). Changes related to tremor are once again observed, however with some differences. For example, where the tremor was clearly visible in the amplitude ratio in mid-2007 at FWVZ-WPVZ (Fig. 2.19d), the same time period for NGZ-TUVZ shows amplitude ratios close to zero (Fig. 2.23d). This suggests the energy during this time period is not propagating coherently between these stations. Similarly, the phase similarity, despite being elevated during tremor, is generally lower than those observed for the pair FWVZ-WPVZ.

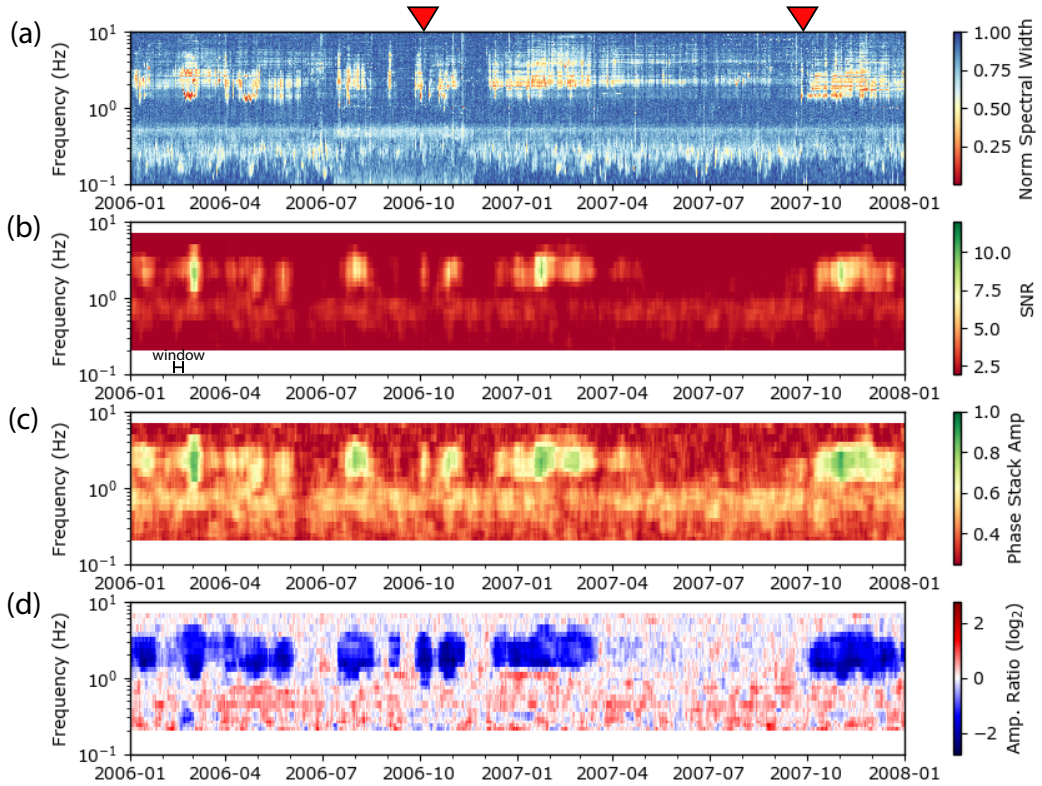

Figure 2.23: Measurements of cross-correlation characteristics at Mt Ruapehu for a single station-pair (NGZ-TUVZ), computed using 10-day moving window. For SNR and phase stack amplitude, the average value is taken between minimum lag defined by 0.5 km/s ballistic arrival speed and maximum lag equal to the minimum lag plus 15 cycles of central period of the bandpass filter. (a) Normalized covariance matrix spectral width as computed over full network. (b) SNR. (c) Phase stack amplitude. (d) Amplitude ratio.

Another way to visualise how tremor is propagating over the network is to compare the phase similarity of different station-pairs. In Figure 2.24, we show the average phase similarity recorded over the network in two frequency bands (0.2–0.7 Hz and 1.5–3.0 Hz) at three different points in time. At 0.2–0.7 Hz, there are minimal differences between the three time periods, as might be expected with oceanic noise sources dominant in this frequency band (Fig. 2.24b). Still, we see that station-pairs with shorter interstation distances have slightly higher phase similarity. At 1.5–3.0 Hz, however, there is greater variability between the different time periods (Fig. 2.24c). For example, during stronger tremor in January 2007, we observe that station-pairs closer to the volcanic summit have a higher phase similarity. In contrast, we see that only a few station-pairs show higher levels of phase similarity when the strength of tremor is reduced and almost all pairs have low phase similarity when tremor activity is minimal. Thus, the quality of data within the frequency band of tremor is highly dependent on the strength of the tremor signal and the proximity of the stations to the source.

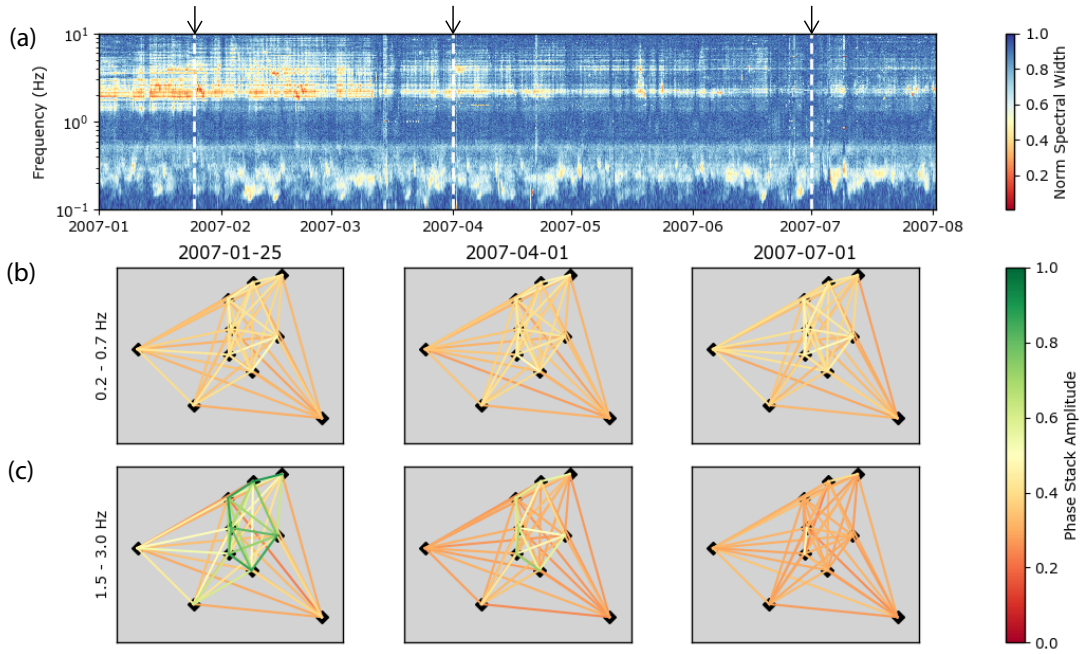

Figure 2.24: Phase similarity map at Mt Ruapehu in different frequency bands. Value shown represents average between the minimum and maximum frequencies shown (0.2–0.7 Hz and 1.5–3.0 Hz). (a) Normalized spectral width. Arrows and dashed-white lines correspond to dates shown in proceding subplots. (b) phase similarity at different stations between 0.2–0.7 Hz for dates 2007-01-25, 2007-04-01, and 2007-07-01. (c) phase similarity between 1.5–3.0 Hz for same dates.

2.4.4 Stromboli

Volcano seismicity is a continuous feature of Stromboli, due to the persistent volcanic activity (Langer and Falsaperla, 1996; Chouet et al., 1997; Wassermann, 1997; Ripepe and Gordeev, 1999). Two types of signals predominate, those related to sustained volcanic tremor and those related to discrete volcanic explosions.

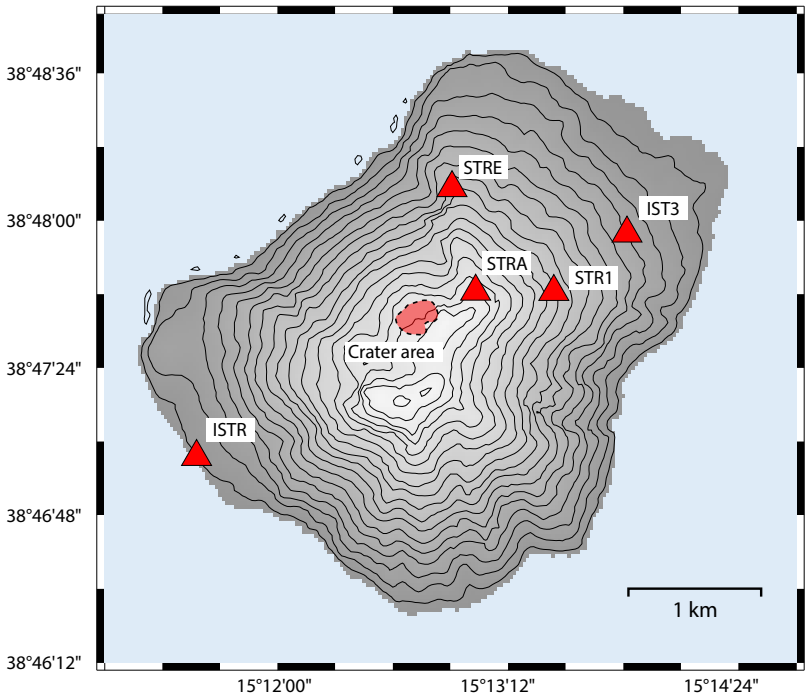

Figure 2.25: Seismic network used at Stromboli, showing five broadband stations (red triangles). Crater terrace area shaded in red.

In Figure 2.26, we show the spectral width computed over two years (2018–2020) using five broadband stations (Fig. 2.25). Low values of spectral width are observed between approximately 1 – 6 Hz throughout the 2-year period (Fig. 2.26). This is consistent with previous observations of tremor spectral maxima at the volcano (Langer and Falsaperla, 1996). Thus, we assume that reduced spectral width values within this frequency range reflect sustained tremor.

Low values of spectral width are also observed below 0.5 Hz throughout. This is especially true at approximately 0.2 Hz where low values are observed almost continuously, most likely reflecting the secondary oceanic microseism. Some fluctuations are observed however, with short 'bursts' of lower spectral width values (lasting a few days) either side of 0.2 Hz (Fig. S4.2). This is especially apparent in late-2018/early-2019, and also early-2018 and late-2019. This may reflect storm activity in winter, with Mediterranean thunderstorms occurring predominantly over the sea in autumn and winter (Galanaki et al., 2018). Supporting this is the strong correlation between long period 'tremor' (∼0.5 Hz) and strong weather events at the volcano observed previously by Braun et al. (1996).

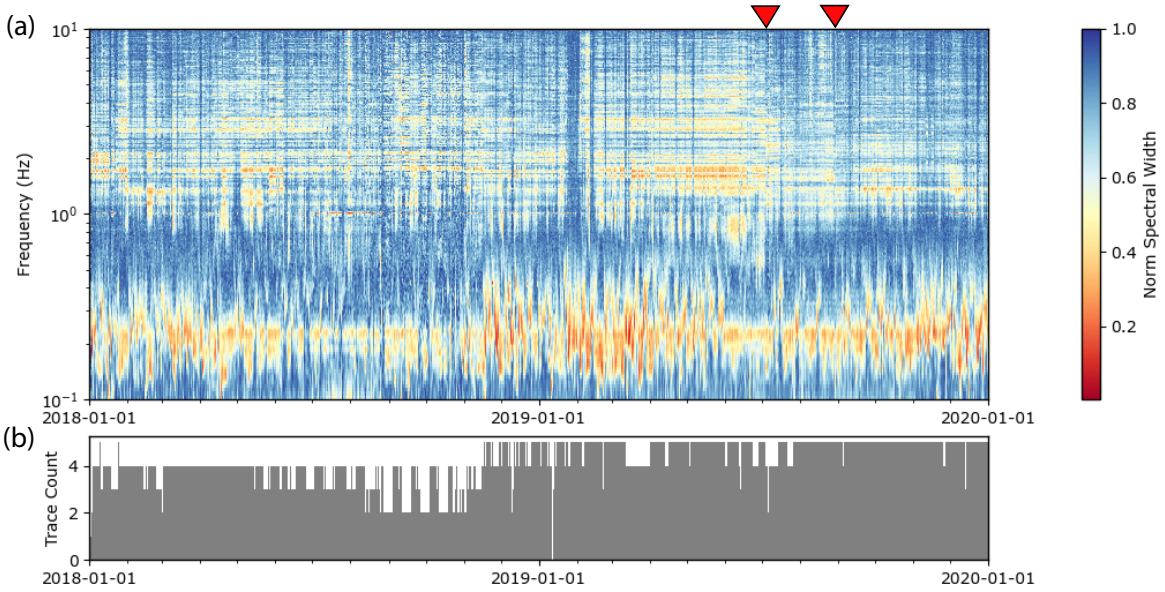

Figure 2.26: Normalized spectral width from five broadband seismic stations on Stromboli (Fig. 2.25).

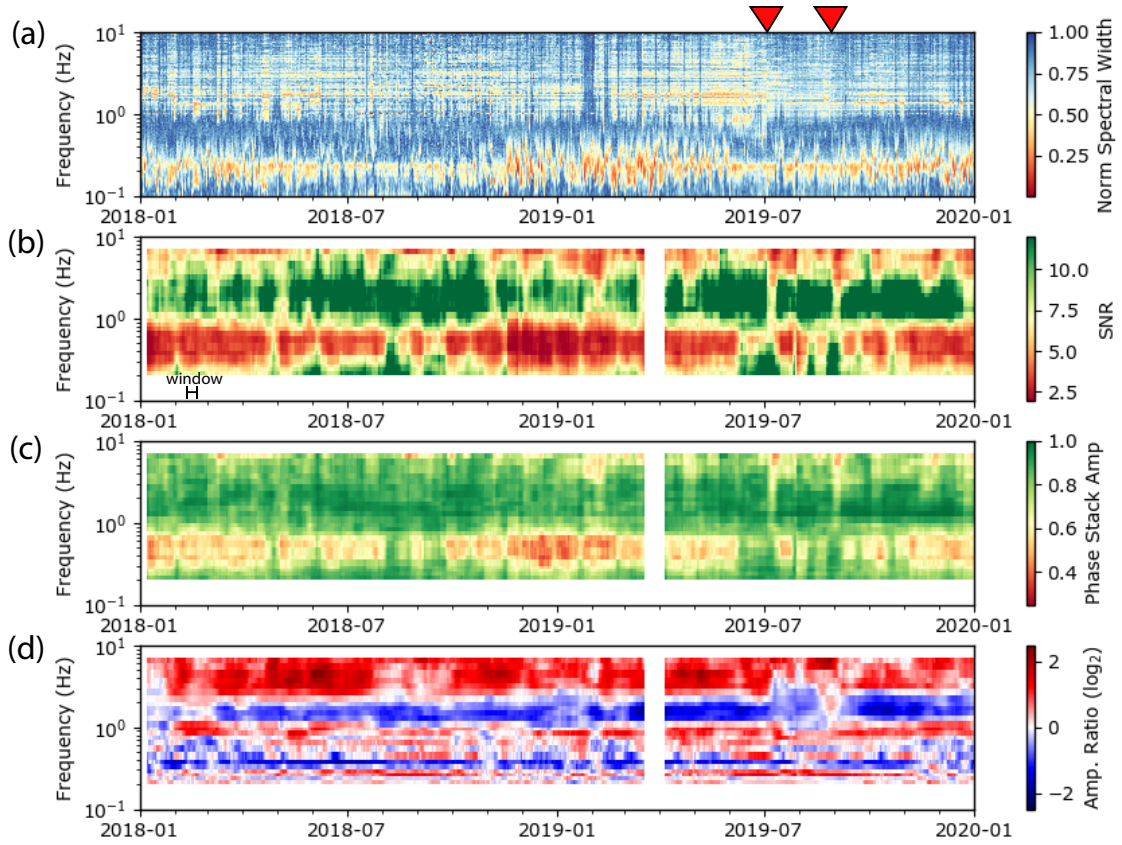

Figure 2.27: Measurements of cross-correlation characteristics at Stromboli for a single station-pair (IST3-ISTR), computed using 10-day moving window. For SNR and phase stack amplitude, the average value is taken between minimum lag defined by 0.5 km/s ballistic arrival speed and maximum lag equal to the minimum lag plus 15 cycles of central period of the bandpass filter. (a) Normalized covariance matrix spectral width as computed over full network. (b) SNR. (c) Phase stack amplitude. (d) Amplitude ratio.

High values of SNR and phase similarity are observed in frequency band of volcanic tremor, in particular between 1–3 Hz (Fig. 2.27). This is the case for most of the 2-year period, except for a few short-lived periods of lower SNR and phase similarity (Figs 2.27b and 2.27c). For example, the July 3 and August 28 paroxsyms in 2019 appear to be associated with temporary drops in these measurables (Fig. 2.28). From the amplitude ratio, we see a split between cross-correlations with a negative ratio between approximately 1-2 Hz and positive ratios above 2 Hz (Fig. 2.27d). Some temporal variation is also observed in the amplitude ratio. For example, between 1-2 Hz, dominantly negative ratios are closer to zero in early-2018, early-2019 and in the months following the July 3 paroxsym.

Below 1.0 Hz, we observe greater instability in the character of the cross-correlation functions relative to the others volcanoes examined. For example, both the SNR and phase similarity fluctuate between high and low values between approximately

0.3–0.8 Hz (Figs 2.27b and 2.27c). The timing of these changes appears to coincide with changes in the spectral width occurring approximately yearly. This is highlighted when averaging the phase similarity between 0.1–1.0 Hz, with higher values in summer compared to winter (Fig. 2.28). A similar observation is also found when examining the phase similarity of cross-correlation functions against lag time. During the month of December 2018, there is reduced phase similarity at later lag times (Fig. 2.29b) relative to earlier in the year (Fig. 2.29a). Thus, if this energy is related to storm activity, it appears to have a strong effect on the cross-correlation functions recorded at Stromboli below 1 Hz.

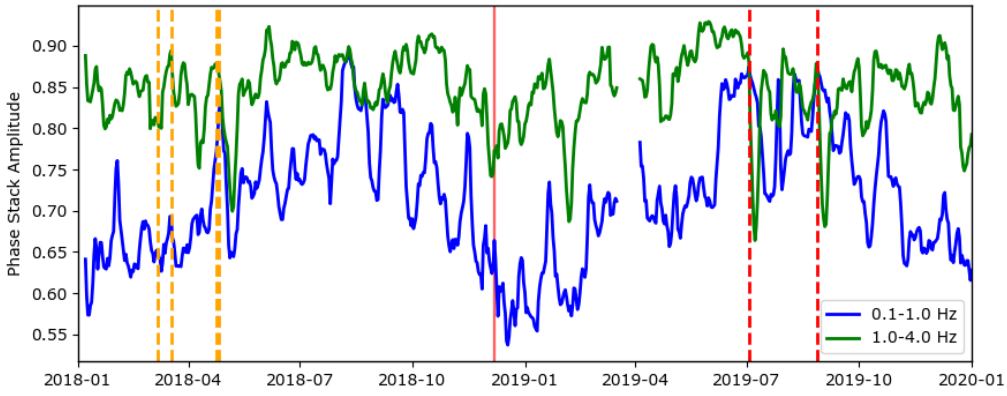

Figure 2.28: Phase similarity with time for low-frequency (average between 0.1–1.0 Hz) and high-frequency (average between 1.0–4.0 Hz) cross-correlation functions. Volcanic activity highlighted corresponds lava flows (red-solid-line), major explosions (orange-dashed-line), and paroxysms (red-dashed-line).

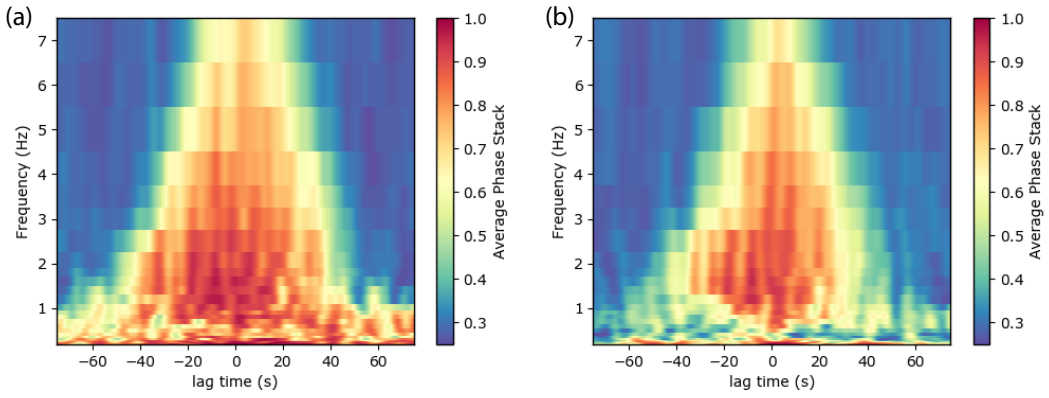

Figure 2.29: Phase similarity as a function of frequency and lag-time for station-pair (IST3-ISTR), as shown in Figure 2.27. Values computed as average between defined date ranges corresponding to (a) period with minimal energy below 1.0 Hz observed in spectral width measurements 2018-08-01 to 2018-09-01 (b) period with more energy below 1.0 Hz observed in spectral width measurements (2018-12-01 to 2019-01-01).

2.4.5 Whakaari

Finally, we examine the characteristics of cross-correlation functions computed at Whakaari volcano. This setting is rather unique from others considered in this chapter in that prior to mid-2013, only one seismic station was located on the island (Fig. 2.30). Thus, the volcano is well suited for testing single-station approaches (Yates et al., 2019; Caudron et al., 2021). Since mid-2013, a further station was installed, however following an eruption in late-2019, the stations are no longer operational due to lack of servicing. For the purpose of testing further the results obtained using a single-station, we focus on a time period when only a single station was active (pre-2013).

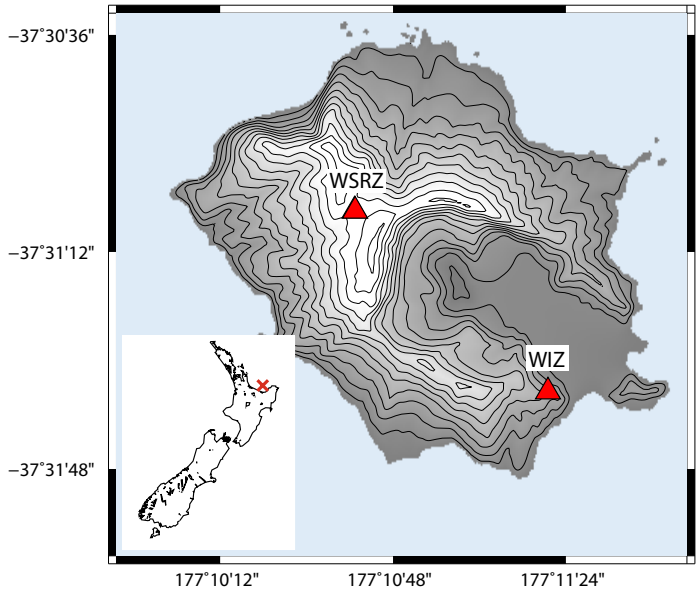

Figure 2.30: Two broadband seismic stations (red triangles) on Whakaari volcano. Station WSRZ was active from mid-2013. Neither station active at time of writing.

Seismicity at Whakaari is largely influenced by its active hydrothermal system producing a wide range of seismic activity that includes tremor (Latter and Balsillie, 1984; Hurst and Steven, 1993; Hurst, 1998; Sherburn et al., 1999; Chardot et al., 2015), long-period activity (Sherburn et al., 1998), and volcano-tectonic activity (Nishi et al., 1996). Such events typically occur at the volcano with frequencies above approximately 1 Hz. Performing covariance matrix analysis using a single-station on the volcano between 2010–2013, we clearly observe this seismicity, with reduced values of spectral width above 1 Hz throughout a three year period (Fig. 2.31). This includes stronger episodes, such as the gliding tremor identified previously by Jolly et al. (2020).

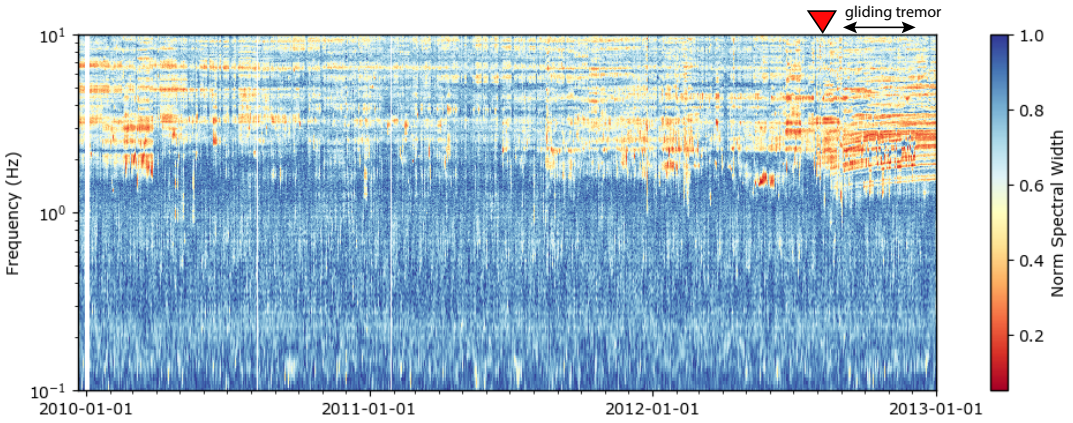

Figure 2.31: Normalized spectral width using three-components of single-station (WIZ) at Whakaari volcano.

Examining the phase similarity of single-station cross-component cross-correlation functions reveals that lower spectral width values are often associated with reduced phase similarity at Whakaari volcano (Fig. 2.32). For example, above 1 Hz, the phase similarity is largely below 0.5 for the entire three years across all component-pairs (Fig. 2.32b-d). There are exceptions, however, with elevated phase similarity during gliding tremor late-2012 and also briefly during heightened tremor in early-2010. These differences are clearly highlighted when examining the phase similarity against lag time during a time period representative of the three year period (Fig. 2.33a) versus that of the gliding tremor (Fig. 2.33b). During gliding tremor, higher phase similarity is observed out to approximately 30 seconds between 1–4 Hz compared to approximately 10 seconds during regular time periods.

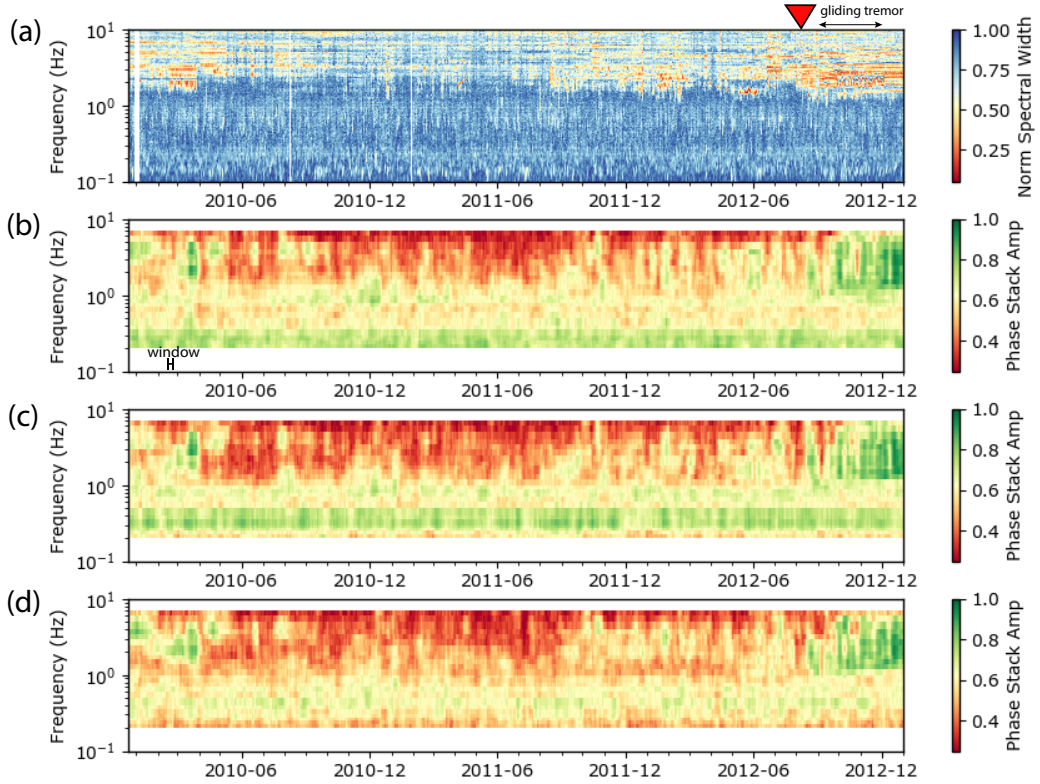

Figure 2.32: Measurements of phase similarity for a single-station (WIZ). (a) Normalized covariance matrix spectral width. (b) phase similarity EN component. (c) phase similarity EZ component. (d) phase similarity NZ component. All phase similarity measures represent average between window defined by minimum lag of 10 s and width of 15 cycles of the central period.

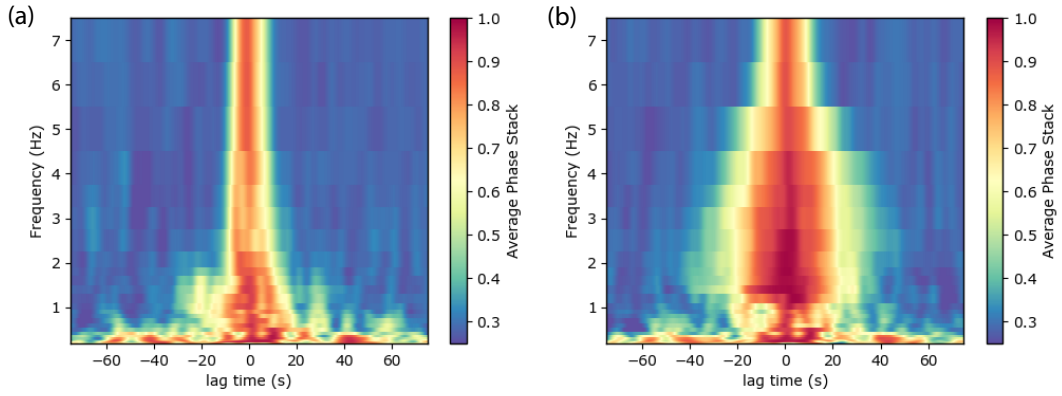

Figure 2.33: Phase similarity as a function of frequency and lag-time for single-station WIZ, EN component-pair, as shown in Figure 2.32. Values computed as average between defined date ranges corresponding to (a) standard background activity (2011-06-01 to 2011-07-01). (b) period of gliding tremor (2012-12-01 to 2013-01-01).

2.5 Discussion

By examining many volcanoes, we observe how different volcanoes are capable of producing very different seismic wavefields and their corresponding influence on computed cross-correlation functions. In this section, we discuss the influence of various features of the seismic wavefield as observed at different volcanoes. Finally, we discuss the implications of these findings with regards to practical steps one can take when working with a new volcanic dataset towards coda wave interferometry.

2.5.1 Influence of volcanic tremor

Volcanic tremor is a significant feature of many volcanoes analysed in this chapter. At some volcanoes the tremor is approximately continuous, for example at Stromboli and pre-2018 eruption at Kilauea. At others, tremor is more intermittent (e.g. Ruapehu), or occurring predominantly during eruptive activity (e.g. Piton de la Fournaise). The presence of tremor warrants caution when applying coda wave interferometry due to the potential to contaminate cross-correlation functions (Ballmer et al., 2013; Gómez-García et al., 2018). Thus, studies applying coda wave interferometry have tended to target frequency bands that do not include volcanic tremor (Mordret et al., 2010; Cannata et al., 2017). However, recent studies have taken advantage of stable volcanic tremor towards computing highly similar cross-correlation functions in time (Donaldson et al., 2017). Thus, it is worth re-examining the influence of tremor on cross-correlation functions to better understand the implications of volcano seismicity on coda wave interferometry.

An interesting result is that many of the volcanoes analysed show evidence that tremor is producing similar cross-correlation functions in time. At both Kilauea (Fig. 2.15) and Stromboli (Fig. 2.27), low spectral width values corresponding to continuous tremor are encompanied by high SNR and phase similarity. Similarly, intermittent episodes of tremor at Ruapehu (Fig. 2.19), and co-eruptive tremor at Piton de la Fournaise (Fig. 2.8), also show high phase similarity in the cross-correlation functions. In most cases, these values are larger than those computed within frequency bands, or during time periods, with no tremor. Similarly, these higher values extend out to greater lag-times than non-tremor periods (e.g. Figs 2.12, 2.16, and 2.20). For example, there is a sharp drop in the phase similarity during periods of reduced, or minimal tremor, at Mt Ruapehu (Fig. 2.19c and 2.20b). These periods are equally highlighted by sudden changes in the amplitude ratio computed above 1 Hz, with the tremor producing values that seemingly differ significantly from the background noise sources (Fig. 2.19d). Similarly, cross-correlation functions below 1 Hz at both Kilauea and Stromboli have predominantly reduced SNR and phase similarity relative to the tremor band.

The observation of high similarity in cross-correlation functions associated with tremor suggests a level of stability in the source processes. This could be explained by commonly invoked models for tremor excitation. For example, while a large variety of different models exist (Chouet and Matoza, 2013), many are based on a repetitive source process. Such mechanisms include, but are not limited to, the

periodic opening and closing of fractures triggered by pressure oscillations (Lesage et al., 2006; Rust et al., 2008; Valade et al., 2012) and self-sustained oscillations resulting from fluid flow (Julian, 1994; Balmforth et al., 2005). It is well-documented that tremor can exhibit a wide variability, with temporal changes in amplitude and frequency content (Chouet and Matoza, 2013). Such variability is clearly observed in spectral width measurements at both Ruapehu (Fig. 2.18) and Stromboli (Fig. 2.26), on both short (days to months) and long (months to years) timescales. Despite this variation, however, the observation of high similarity in cross-correlation functions suggest that some part of the background seismic wavefield is stable at the time-scale of the 10-day moving window.

It is important to recognize, however, that high values of SNR and phase similarity due to tremor do not neccessarily indicate long-term stability of the source. For example, while episodes of heightened tremor at Ruapehu appear to produce similar cross-correlation functions, evidenced by high SNR and phase similarity (e.g. Figs. 2.19 and 2.21), it is unclear from this alone whether each episode is comparable. Similarly, at Stromboli, while the SNR and phase similarity are high at frequencies associated with tremor, temporary decreases are observed (Figs. 2.27 and 2.28). These decreases may be associated with changes in the character of the dominant seismic source, as observed at Piton de la Fournaise (Fig. 2.9). Thus, it would be neccessary to check the longer-term stability of cross-correlation functions towards long-term monitoring using the tremor.

One approach to assess longer-term stability could be to compare individual tremor episodes to a longer-term reference, for example using the Pearson correlation coefficient between current and reference cross-correlation functions. High values across the different episodes would suggest the cross-correlation functions produced are highly similar. However, comparing cross-correlation functions using reference-based approaches can be limited. If the cross-correlation functions vary between tremor episodes, a long-term average (reference) may not represent the dataset. Furthermore, if there are multiple clusters of similar cross-correlation functions, this would be difficult to detect using a single reference function. Thus, in keeping with non-reference based approaches, we propose the application of hierarchical clustering as an additional step to quantify similarity between individual cross-correlation functions computed during any time period. This approach is discussed further in Chapter 3.

Finally, despite tremor seemingly producing highly similar cross-correlation functions at most of the volcanoes analysed, an exception is observed at Whakaari volcano. At other volcanoes, low spectral width values associated with tremor generally coincide with high SNR cross-correlation functions (Figs. 2.32 and 2.33a). However, at Whakaari, consistently low spectral width values above 1 Hz are largely associated with low SNR and phase similarity. An exception to this is during gliding tremor in late-2012, where high SNR and phase similarity are recorded (Fig. 2.33b). One possibility may relate to the use of a single-station to compute the spectral width and corresponding phase similarity. In this case, low spectral width values do not provide information on the spatial coherence of propagating seismic energy. Thus it may be the case that this energy is not propagating coherently enough to be measured at later lag times of cross-correlation functions.

An alternative explanation could be that the sources of seismicity regularly recorded at Whakaari above 1 Hz are unstable on the time scales of multiple days. The temporal resolution of spectral width measurements computed here are less than two hours. Thus, seismic energy coherent on time scales of hours could produce low spectral width measurements. However, if this seismicity were to evolve on slightly longer time scales (days), this would be reflected in the 1–day cross-correlation functions. This may explain the lack of high SNR cross-correlation functions above 1 Hz at Whakaari during regular activity.

2.5.2 Influence of oceanic noise sources

The ambient noise wave-field for periods greater than 1 s (<1 Hz) is well known to be dominated by oceanic microseisms (Hasselmann, 1963; Peterson, 1993; Stutzmann et al., 2000). Specifically, two peaks can often be observed at approximately 14 s (0.07 Hz) and 7 s (0.14 Hz) associated with the primary and secondary microseisms respectively. The secondary microseism, typically stronger, is characterised by a gentle slope in the spectrum for periods shorter than 7s and a steep slope for periods longer than 7s (Stutzmann et al., 2000). These are discussed in further detail in Section 1.2.1.1.

The microseism peak is clearly observed as reduced values of spectral width at many of the volcanoes considered. At Piton de la Fournaise (Fig. 2.2), Kilauea (Fig. 2.14), and Stromboli (Fig. 2.26), for example, consistently low values of spectral width are observed between approximately 0.1–0.3 Hz. At Whakaari (Fig. 2.31) and Ruapehu (Fig. 2.18), the microseism is less clearly revaled through spectral width measurements. This could be related to the use of a single-station at Whakaari, which is shown as more limited in recovering the oceanic microseism using covariance matrix analysis (Fig. 2.4). Similarly, at Ruapehu, the use of short period stations with a natural frequency of 1 Hz could reduce the sensitivity to the oceanic microseisms. Supporting this is slightly increased recovery of the microseism computing the spectral width between the two broadband sensors (Fig. S2.4). Similarly, processing recently recorded data using four broadband stations during a seismic crises in 2022 (discussed later in Chapter Six, Section 6.1.1) clearly reveals the oceanic microseisms (Fig. S2.5).

The influence of the oceanic microseisms on cross-correlation functions appears to vary across the different volcanoes. At Kilauea pre-eruption, for example, the measured phase similarity is above approximately 0.6 for frequencies down to 0.2 Hz (Fig. 2.15c). Similarly, we observe higher phase similarity between approximately 0.3–0.7 Hz at Ruapehu (Fig. 2.19c). Note that in both of these cases, the SNR and phase similarity is still lower than those observed within the frequency band of tremor. At Stromboli, however, reduced values of SNR and phase similarity (below 0.6) are regularly observed within the frequency band thought to be dominated by ocean sources (< 1 Hz). These decreases appear to coincide with the bursts of reduced spectral width observed within the same frequency band (Fig. 2.27a).

Differences in the origin of oceanic noise sources at each volcano may explain variation in the response to microseism energy. For example, it has recently been shown

that the stability of the noise field, in the 1–3 s period band, varies across Europe (Stehly et al., 2022; Delouche, 2023). Locations mainly influenced by North Atlantic ocean sources were found to produce more stable frequency content over time relative to locations influenced by Mediterranean ocean sources (e.g. Greece and Italy). Short-lasting (hour-to-day-long) microseismic events (between 0.5–0.7 Hz) were also detected in southern Europe, originating from the Mediterranean sea (Stehly et al., 2022; Delouche, 2023). These events were found to be more frequent in summer compared to winter. Thus, it is unclear if these relate to the short-duration events seen in the spectral width at Stromboli, which appear to be more dominant in winter. In any case, the greater fluctuation of seismic sources in southern Europe was shown to increase the convergence time of cross-correlation functions (Stehly et al., 2022; Delouche, 2023). The impact of this is to reduce the temporal resolution possible for monitoring studies.

A similar point can be made about the relationship between the SNR/phase similarity measured here and the potential for monitoring. Specifically, reduced values of SNR or phase similarity within the frequency band of oceanic microseisms do not neccessarily mean that the data is unusable. Instead it suggests that stacking a greater number of cross-correlation is likely required to reach the same level of convergence as higher SNR cross-correlation functions. This is demonstrated in Figure 2.34, where the number of days (N) required reach a correlation coefficient of 0.99 between N and $N-1$ stacks of cross-correlation functions is plotted for the different volcanoes. At all three of Piton de la Fournaise (Fig. 2.34a), Ruapehu (Fig. 2.34c), and Stromboli (Fig. 2.34d), this value is generally reached for stack sizes of $N < 20$ across all frequencies. In the case of Ruapehu and Stromboli, however, the high SNR of tremor is also reflected by greatly reduced convergence times. Similarly, we observe that the low values of SNR at Kilauea (post-eruption, Fig. 2.34b) and Whakaari at higher frequencies (Fig. 2.34e) require much larger convergence times ($N > 30$ days). In the case of Kilauea, this is in constrat to the pre-eruption period where high phase similarity is reflected by low convergence times across all frequencies (Fig. 2.34b). Thus, there is a clear relationship between measurements of SNR/phase similarity and the convergence time of cross-correlation functions.

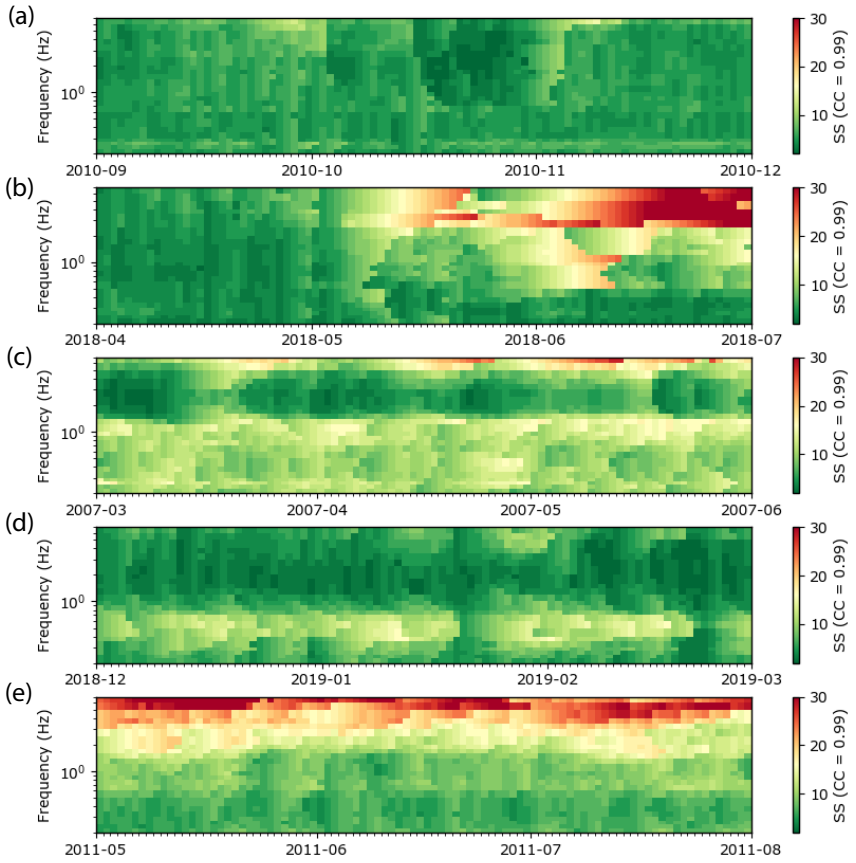

Figure 2.34: Convergence tests at different volcanoes, where we plot the value N required to obtain a correlation coefficient of 0.99 between N and $N-1$ day cross-correlation functions. (a) Piton de la Fournaise (FJS-FOR) (b) Kilauea (HAT-OBL) (c) Ruapehu (FWVZ-WPVZ) (d) Stromboli (IST3-ISTR) (e) Whakaari (WIZ, EN component-pair).

2.5.3 Influence of eruptive activity

All of the time periods examined across the five volcanoes contain eruptive activity of varying styles. These include more energic phreatic eruptions at Whakaari and Ruapehu and major explosions/paroxsyms at Stromboli (in addition to continuous strombolian activity). At Piton de la Fournaise and Kilauea, in contrast, the eruptive style is lower-energy effusive activity. Thus, we consider what influence these various types of activity have on the cross-correlation functions we compute.

The most striking changes related to eruptive activity occur at Kilauea. Significant decreases in SNR and phase similarity are observed at the onset of the eruption, across all frequencies (Fig. 2.15b and 2.15c). This is also evidenced by a clear change in previously stable amplitude ratios (Fig. 2.15d). These observations highlight that this time period is associated with unstable seismic sources related to the collapse of the caldera and the disappearance of the crater lava lake (Soubestre et al., 2021). Following the period of caldera collapse, only the lowest frequencies (<0.5

Hz) recover amplitude ratios consistent with those pre-eruption. At frequencies above 0.5 Hz, the amplitude ratios are no longer dominantly negative, and have a SNR and phase similarity far lower than pre-eruption (Fig. 2.16c). This almost certainly reflects the loss the the dominant tremor signal post-eruption. Of interest, the additional energy between 0.1–0.5 Hz (Fig. 2.14) assumed to be associated with semi-regular collapse events within the caldera (Soubestre et al., 2021) appear to have minimal influence on the cross-correlation functions. It is possible that these events are too short-lived to have a strong influence on the cross-correlation functions.

Piton de la Fournaise is another volcano that experiences prolonged eruptive activity, typically accompanied by tremor. Decreases in SNR and phase similarity either side of eruptive activity at the volcano are evidence that the cross-correlation functions are altered by the tremor (Fig. 2.9). From the amplitude ratios, we can observe clearly that the influence of tremor extends down to approximately 1 Hz or just below (Fig. 2.8d). Such strong changes in the character of cross-correlation functions would make it difficult to measure the relative velocity change between time periods within and outside of this eruptive episode. Thus, combined with the observations from prolonged eruptive activity at Kilauea, it should be considered that longer-lasting eruptive episodes can have a significant influence on cross-correlation functions. This is also consistent with significant changes in the character of cross-correlation functions computed during the 2011 eruption at Eyjafjallajökull, Iceland (Caudron et al., 2021).

At Stromboli, the two years analysed contain both major explosions and paroxysms, and regular continuous activity that are characteristic of the volcano. This consists of sustained volcanic tremor and discrete volcanic explosions (Langer and Falsaperla, 1996; Chouet et al., 1997; Wassermann, 1997; Ripepe and Gordeev, 1999). Both are thought to be related to degassing fluid at a shallow depth (upper 200m) beneath the active crater (Chouet et al., 1997; Ripepe and Gordeev, 1999). While the SNR and phase similarity of cross-correlation functions is generally high within the frequency band of tremor, temporary decreases are observed (Fig. 2.27 and 2.28). Of the four largest decreases in phase similarity recorded between 1–4 Hz, three follow immediately after significant explosive activity (Fig. 2.28). These are the major explosions in late-April 2018 and both paroxysms in 2019. The short-lived nature of these events makes it unlikely that the explosions themselves are responsible for an artificial change in the character of 1-day cross-correlation functions within a 10-day moving window. Thus, these changes could reflect longer-term changes to the character of cross-correlation functions. The decreases then, just as at Piton de la Fournaise, would represent the mixture of cross-correlation functions of different character either side of these events.

Finally, we consider briefly short-lived eruptive events at Whakaari and Ruapehu. In both cases, it is difficult to distinguish any clear changes related to eruptive activity. For example, no obvious changes from background are observed at the time of the August 2012 eruption at Whakaari (Fig. 2.32). Similarly, eruptions in October 2006 and September 2007 at Ruapehu do not appear to show any clear change relative to other time periods (Fig. 2.19). However, it should be noted that intermittent nature of tremor at Ruapehu, and the low SNR character of cross-

correlation functions above 1 Hz, makes identification of changes associated with an eruption difficult with the methods applied here.

2.5.4 Considerations towards coda wave inteferometry

Two primary objectives were considered in this chapter. The first objective was to present a series of methods that could be used to characterize a seismic dataset intended for coda wave interferometry. The second objective was then to apply these methods to a number of different volcanoes to consider more generally the influence different features of the seismic wavefield at volcanoes might have on cross-correlation functions.

The first objective can be split into two parts. First, network covariance matrix analysis was applied to identify coherent signals in the seismic wavefield (Section 2.2). Second, broadband cross-correlation functions were computed and various measures calculated that allowed them to be characterized (Section 2.3). These were the SNR and phase stack amplitude, and the amplitude ratio between positive and negative sides of the cross-correlation function.

Covariance matrix analysis and the computation of the spectral width has been applied by a number of studies in recent years (e.g. Seydoux et al., 2016b,a; Soubestre et al., 2018, 2019; Caudron et al., 2021, 2022; Journeau et al., 2022). The key advantage of the approach relative to single-station approaches (e.g. spectrograms) is the reduced impact of local, non-propagating, noise sources. Testing different network configurations at Piton de la Fournaise — including using only two-stations or the three-components of a single-station — showed that, in all cases, the main features of the seismic wavefield were still clearly identifiable (Section 2.2.3). This was similarly observed at Mt Ruapehu when computing the spectral width with a varied number of stations (Fig. 2.18). Thus we consider it can be a useful tool towards understanding the main sources of a seismic wavefield propagating over any given volcano network.

Broadband cross-correlation functions, and the corresponding measures of SNR, phase stack amplitude, and amplitude ratio provided a means of direct comparison with the spectral width measurements from covariance matrix analysis. In many cases, there was a clear relationship between coherent seismic energy revealed by covariance matrix analysis and the similarity of 1-day cross-correlation functions. In some cases, the seismic sources present produced highly similar cross-correlation functions. Examples include during tremor and Kilauea, Stromboli, Ruapehu, and briefly at Piton de la Fournaise. In other cases, the seismic sources reduced the quality of cross-correlation functions. Examples include co-eruptive activity at Kilauea, low-frequency energy recorded at Stromboli during winter months, and background activity above 1 Hz recorded at Whakaari volcano. Thus, while the spectral width can be used to identify seismic sources, it is still beneficial to examine the characteristics of cross-correlation functions in response to these sources.

The examples presented demonstrate the value that can be obtained from examining broadband cross-correlation functions prior to the computation of velocity changes.

For example, plots showing how various characteristics of cross-correlation functions vary through time at different frequencies can provide insight into target frequencies. Similarly, plots of SNR or phase stack similarity against lag-time provide insights into possible choices of coda window. These observations represent a good starting point towards understanding the nature of a seismic dataset. Furthermore, they can encourage further exploration of specific features of the data. For example, the observation that tremor produces highly similar cross-correlation functions within a 10-day moving window at many of the volcanoes studied. This suggests tremor could potentially be used as the dominant seismic source if longer-term stability can be verified. This is explored further in Chapter S3.2.1, where we apply hierarchical clustering to cross-correlation functions to assess their similarity in time.

Another interesting observation that warrants further exploration is the variation in amplitude ratios recorded at different volcanoes. At some volcanoes, for example Piton de la Fournaise, amplitude ratios recorded at a given frequency are relatively consistent through time (outside of volcanic activity). At others, especially those with regular tremor, far greater variability is observed in time. This is the case at Stromboli, where amplitudes below 1 Hz are highly variable over the two year period. Similarly, significant differences are observed between the amplitude ratios recorded by single-station cross-components relative to the station-pairs. At both Piton de la Fournaise and Whakaari, significant variation in time is observed in amplitude ratios recorded by the different components of single-station cross-components (Figs. S2.3 and 2.32 respectively). This contrasts with the previously mentioned consistency of amplitude ratios observed for station-pairs at Piton de la Fournaise. Further study of these differences could improve our understanding of the propagation of coda waves in the case of single-station cross-components.

2.6 Conclusions

The purpose of this chapter was to develop an understanding of seismic datasets intended for coda wave interferometry at volcanoes. Towards this, network-based covariance matrix analysis is applied to first gain an understanding of the seismic wavefield. The characteristics of cross-correlation functions at different frequencies are then examined to understand the influence of various seismic sources on the data. We apply these steps at five volcanoes, with different characteristics, to gain insight into the expectations one may have when processing coda wave interferometry in volcanic settings.

One of the major findings is that, in many cases, volcanic tremor produced cross-correlation functions with high SNR and phase similarity. This supports the recent use of tremor as a dominant seismic source in recent monitoring studies, for example at Kilauea. It also suggests that tremor could be further exploited for monitoring purposes, noting that the tendency in the past has been to avoid the tremor frequency band due to concerns of contamination. Longer-term stability of the tremor would need to be verified towards this goal, where measures of SNR and phase similarity here are based on calculations within a short-term moving window (10-days).

Oceanic noise sources and eruptive activity were also observed to have an influence on cross-correlation functions. However, this was variable across the five volcanoes studied. At both Kilauea and Ruapehu, for example, the phase similarity in the coda of cross-correlation functions remains relatively high at oceanic microseism frequencies. At Stromboli, however, lower values of SNR and phase similarity are regularly observed. This is consistent with high fluctuations in oceanic seismic sources previously observed in Southern Europe. Regarding the influence of eruptive activity, longer-lasting eruptive events at Piton de la Fournaise and Kilauea were associated with significant changes in the seismic wavefield. Comparatively, shorter-lived, small, explosive events at Whakaari and Ruapehu showed minimal influence on cross-correlation functions. Larger events at Stromboli associated with major explosions and paroxysms show evidence of modifying the character of cross-correlation functions.

We consider the steps applied can serve two purposes. First, the results provide insight into the expectations one might have when applying coda wave interferometry in a new volcanic setting. For example, understanding whether volcanic tremor occurs at a volcano, and how regularly, is of clear significance. Second, the methodologies themselves can be applied to develop further understanding of a specific seismic dataset. These steps can only benefit the decision making while also potentially providing insights into the dataset that may warrant further exploration.

2.7 Supplementary Material

Table S2.1: Filters applied to broadband cross-correlation functions for the computation of SNR, phase stack amplitude, and amplitude ratio.

Central Freq	**Low**	**High**
0.2	0.141	0.282
0.225	0.159	0.318
0.25	0.176	0.353
0.275	0.194	0.388
0.3	0.212	0.424
0.35	0.247	0.494
0.4	0.282	0.565
0.5	0.353	0.707
0.6	0.424	0.848
0.7	0.494	0.989
0.8	0.565	1.13
0.9	0.636	1.27
1.0	0.707	1.41
1.2	0.848	1.69
1.4	0.989	1.97
1.6	1.13	2.26
1.8	1.27	2.54
2.0	1.41	2.82
2.5	1.76	3.53
3.0	2.12	4.24
3.5	2.47	4.94
4.0	2.82	5.65
5.0	3.53	7.07
6.0	4.24	8.48
7.0	4.94	9.89
8.0	5.65	11.31

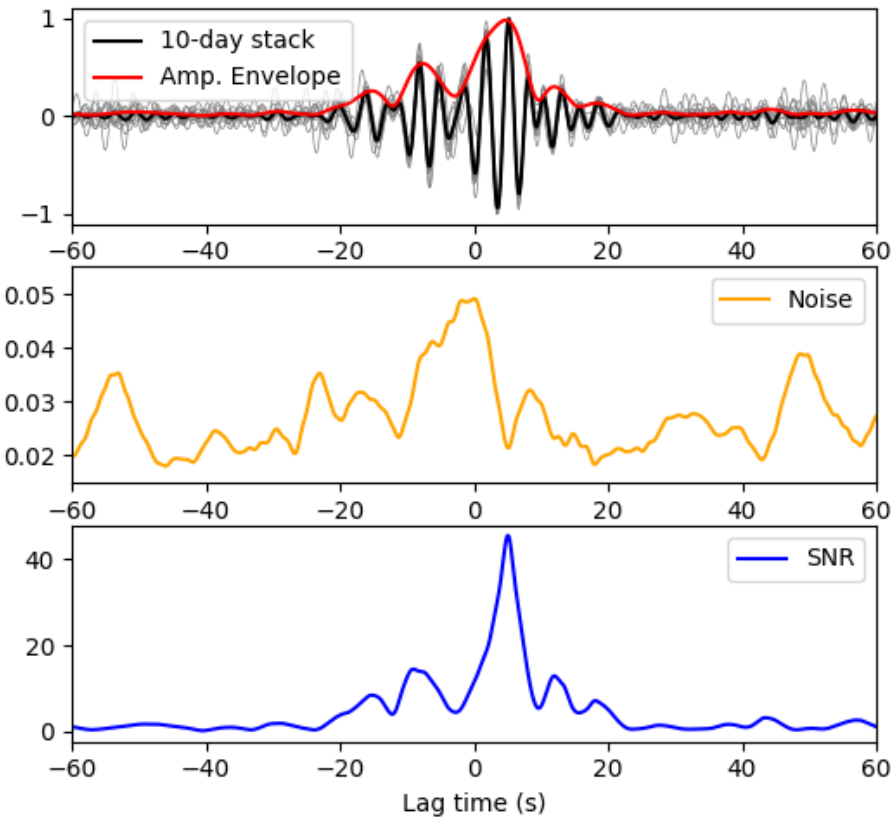

Figure S2.1: Example of SNR computation for 10-day cross-correlation function using station pair UV06-UV08, with cross-correlation functions normalized by maximum absolute value. (a) Ten 1-day cross-correlation functions up to 2010-09-22 are filtered between 0.21-0.42 Hz (light grey lines), stacked (black line), and the amplitude envelope computed (red line). (b) Noise signal computed following Equation 2.5. (c) SNR computed as amplitude envelope divided by Noise.

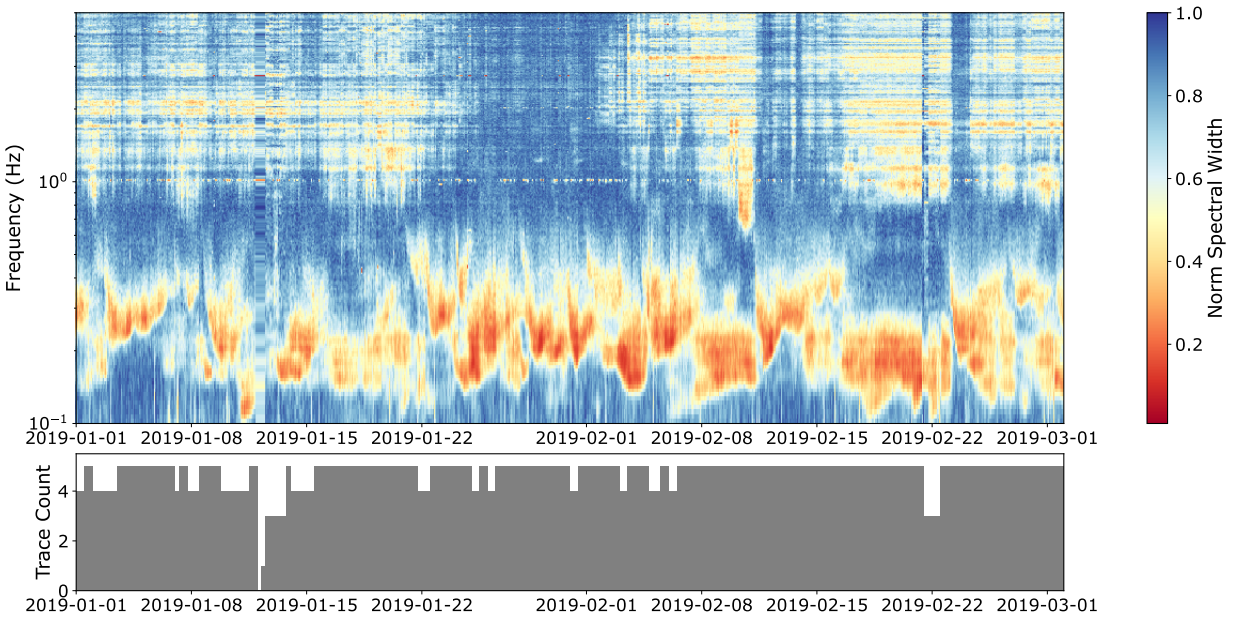

Figure S2.2: Spectral width at Stromboli for 2 months in 2019.

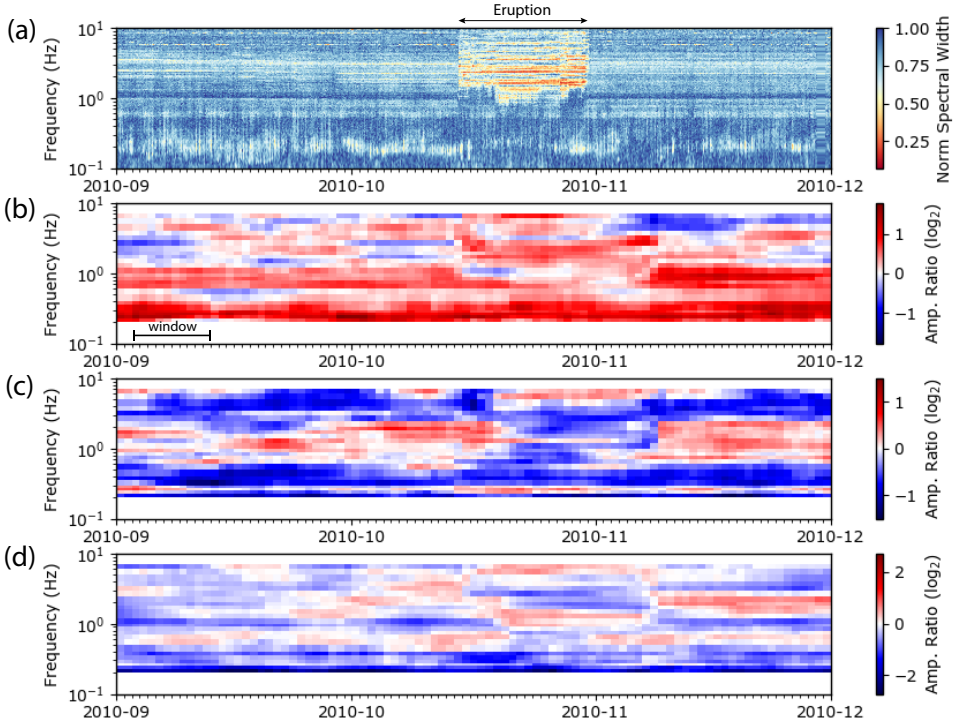

Figure S2.3: Amplitude ratios for a single-station (UV05). (a) Normalized covariance matrix spectral width. (b) amplitude ratio for EN component. (c) amplitude ratio for EZ component. (d) amplitude ratio for NZ component.

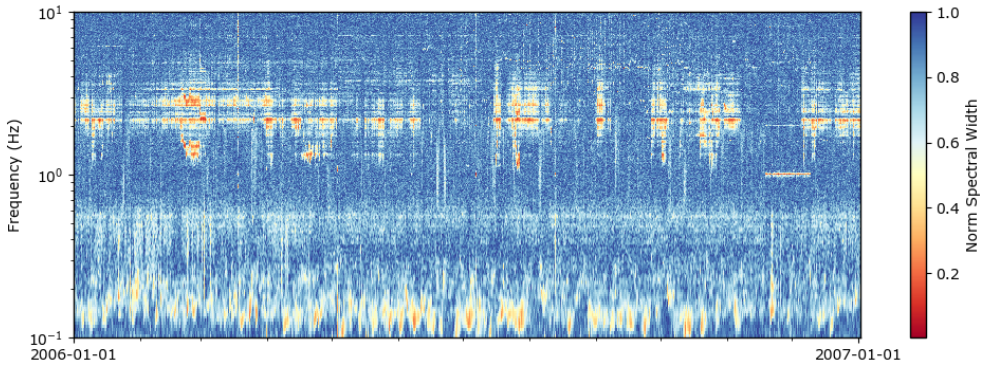

Figure S2.4: Spectral width measurement between FWVZ-WPVZ at Ruapehu (Z-component).

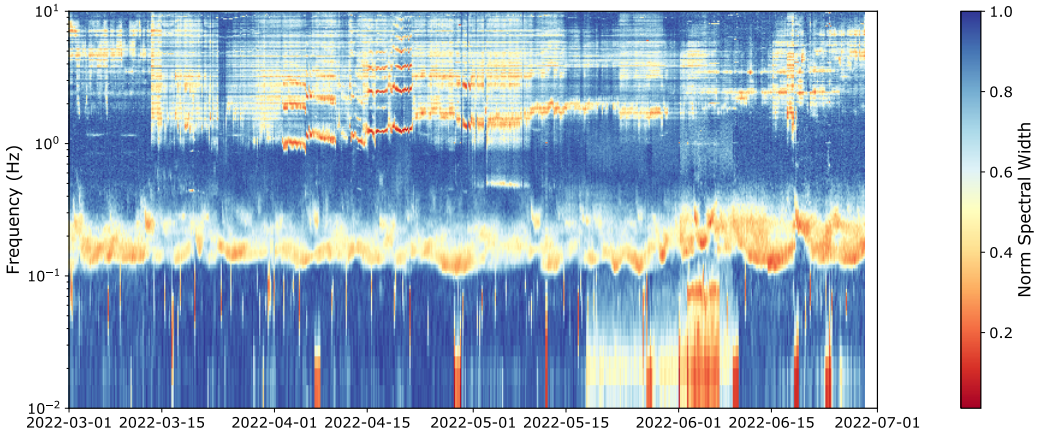

Figure S2.5: Spectral width measurements during seismic crises in 2022, computed using four broadband stations on the volcano (see Section 6.1.1)

.

Chapter 3

Assessing Similarity in Continuous Seismic Cross-Correlation Functions using Hierarchical Clustering: Application to Ruapehu and Piton de la Fournaise Volcanoes

3.1 Abstract

Passive seismic interferometry has become a popular technique toward monitoring. The method depends on the relative stability of background seismic sources in order to make repeatable measurements of subsurface properties. Such stability is typically assessed by examining the similarity of cross-correlation functions through time. Thus, techniques that can better assess the temporal similarity of cross-correlation functions may aid in discriminating between real subsurface processes and artificial changes related variable seismic sources. In this study, we apply agglomerative hierarchical clustering to cross-correlation functions computed using seismic networks at two volcanoes. This allows us to form groups of data that share similar characteristics and also, unlike common similarity measures, does not require a defined reference period. At Piton de la Fournaise (La Réunion island), we resolve distinct clusters that relate both to changes in the seismic source (volcanic tremor onset) and changes in the medium following volcanic eruptions. At Mt Ruapehu (New Zealand), we observe a consistency to cross-correlation functions computed in the frequency band of volcanic tremor, suggesting tremor could be useful as a repeatable seismic source. Our results demonstrate the potential of hierarchical clustering as a similarity measure for cross-correlation functions, suggesting it could be a useful step towards recognizing structure in seismic interferometry datasets. This can benefit both decisions in processing and interpretations of observed subsurface changes.

3.2 Introduction

Passive seismic interferometry has become an increasingly popular technique over the past two decades towards monitoring. By cross-correlating the seismic wavefield recorded at different seismic stations (e.g. Sens-Schönfelder and Wegler, 2006; Brenguier et al., 2008c; Wang et al., 2017), or different components of a single-station (e.g. Hobiger et al., 2012; De Plaen et al., 2016; Yates et al., 2019), it is possible to detect subtle changes in the subsurface that can be related to a variety of geophysical processes. This is typically achieved through the study of phase-shifts and changes in the waveform shape of cross-correlation functions. The former is used primarily to estimate seismic velocity changes while changes in waveform shape can reflect changes in the scattering properties of the medium (Obermann et al., 2013).

Effective monitoring of temporal changes depends on our ability to make accurate and repeatable measurements on near identical waveforms. In cross-correlation functions, as with earthquake doublets (Poupinet et al., 1984), such similarity reflects a level of stability in seismic source processes. Such source stability is a necessary condition for passive monitoring using cross-correlation functions (Hadziioannou et al., 2009). Techniques that can better compare the temporal similarity of cross-correlation functions could therefore aid in the identification of changes due to variations in the background seismic wavefield as opposed to real medium changes. Furthermore, common measures of waveform similarity in cross-correlation functions rely on the use of a reference function. Such measures include the correlation coefficient (e.g. Obermann et al., 2013), the coherence (Clarke et al., 2011), and the zero-lag cross-correlation as a function of lag time (e.g. D'Hour et al., 2016; Sánchez-Pastor et al., 2018). Each of these is computed between a reference stack and smaller moving-stacks of cross-correlation functions. Thus, an approach that can easily compare cross-correlation functions without needing to define a reference period may enhance our ability to identify subtle differences, or similarities, in seismic interferometry datasets without the need to compare multiple reference functions.

Clustering techniques are now relatively well established in seismology. In particular, they have become popular in the analysis of repeating seismic signals, where they can be split into different groups based on measures of waveform similarity (e.g. Aminzadeh and Chatterjee, 1984; Menke, 1999; Carniel et al., 2013; D'Alessandro et al., 2013). This takes advantage of the ability of machine learning algorithms to detect patterns in large datasets in a way that can exceed human performance (Bergen et al., 2019). We apply an unsupervised machine learning approach in the form of agglomerative hierarchical clustering to group different cross-correlation functions in time. We focus on data recorded by two seismic networks. These are the UnderVolc network at Piton de la Fournaise volcano (Institut De Physique Du Globe De Paris (IPGP), 2021), La Réunion Island, between 2009–2011 and the GeoNet seismic network (https://www.geonet.org.nz/) in proximity to Mt Ruapehu volcano, New Zealand, between 2005–2008 (Fig 3.1). Volcanoes represent interesting case studies to test our clustering approach given the challenges associated the multitude of processes that can influence the character of cross-correlation functions (Ballmer et al., 2013; Yates et al., 2019). Furthermore, both time periods have been studied pre-

viously using ambient noise interferometry, with seismic velocity changes computed in relation to volcanic activity at Piton de la Fournaise by Obermann et al. (2013) and at Ruapehu by Mordret et al. (2010). Thus we can relate our findings from clustering cross-correlation functions to previous observations and interpretations at these volcanoes.

To support the clustering analysis, we also apply a recent method based on analysis of the network covariance matrix to assess the spatial coherence of the seismic wavefield (e.g. Seydoux et al., 2016a,b; Soubestre et al., 2018, 2019). Specifically, we follow steps introduced by Seydoux et al. (2016a) to compute the covariance matrix spectral width as a function of frequency through time. This reflects the eigenvalue distribution of the covariance matrix estimated from seismic cross-spectra recorded over a network. Changes in the spectral width can then be used to infer changes in the nature of the seismic wavefield and compared with the outputs of clustering. Finally, we compute apparent seismic velocity changes as an additional comparison dataset, including the correlation coefficient of cross-correlation functions relative to a reference function.

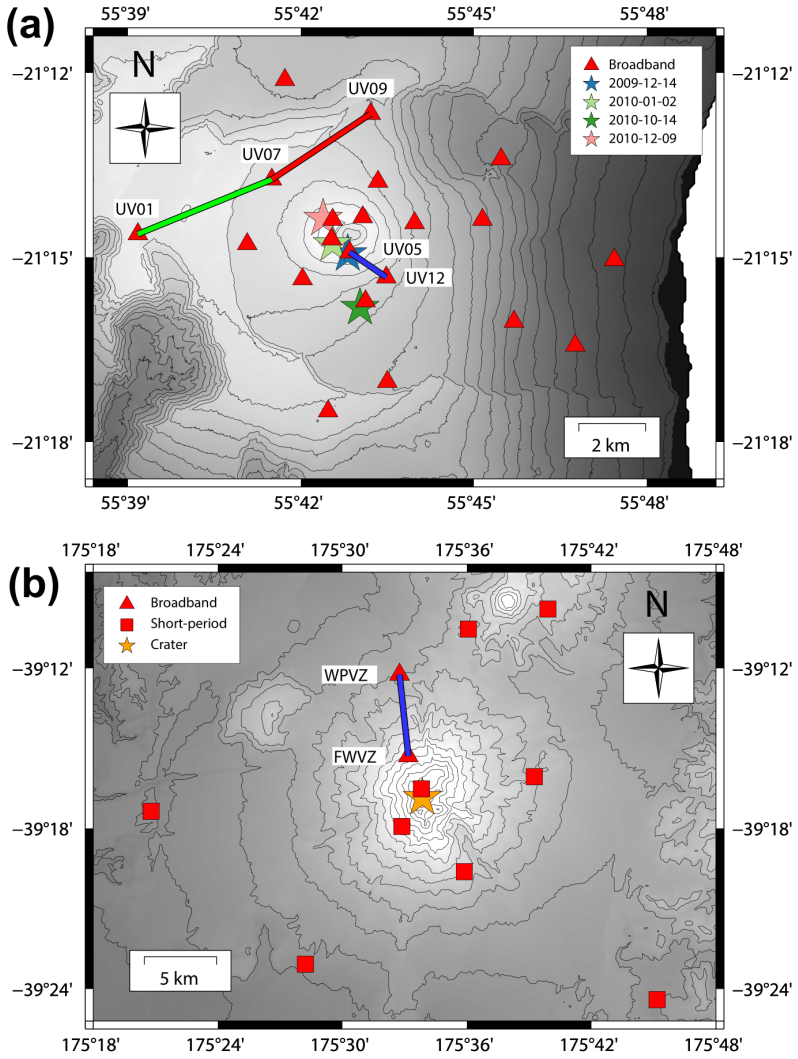

Figure 3.1: Map view of seismic networks used in this study. Broadband stations are shown as red triangles and short-period as red squares. Lines between labeled stations represent station-pairs used in computing cross-correlation functions for clustering. Top: Piton De La Fournaise network between 2009–2011. The three station-pairs shown (blue, red, and green lines) correspond to those presented in further detail by Obermann et al. (2013). Stars represent sites of volcanic eruptions. Digital Elevation Model acquired from IGN, France. Bottom: Ruapehu network between 2005–2008. The station-pair consisting of the two broadband stations is shown (blue line), with the position of the crater (the site of the 2006 and 2007 eruptions) indicated by a star. Digital Elevation Model acquired from LINZ Data Service, New Zealand.

3.3 Methods

3.3.1 Seismic data

Raw seismic data is acquired through the International Federation of Digital Seismograph Networks (FDSN), downloaded and converted to miniSEED format using the python package Obspy (Krischer et al., 2015). At Piton De La Fournaise, we download seismic data recorded using the temporary UnderVolc seismic network installed from September 2009 to May 2011 (Brenguier et al., 2012; Institut De Physique Du Globe De Paris (IPGP), 2021). These data are available through the French Seismologic and Geodetic Network data center (RESIF). At Ruapehu, continuous seismic data are recorded and publicly available through the data center GeoNet (https://www.geonet.org.nz/). We download data for years 2005–2008 at Ruapehu, encompassing two eruptions in 2006 and 2007.

The UnderVolc research project involved the deployment of 15 broadband, three-component, seismic stations around Piton De La Fournaise. This gives a network of 21 seismic stations overall when the six pre-existing broadband stations are included (Figure 3.1a). The time period analyzed includes four eruptive events. One of these occurred within the crater region (2010-01-02), while the other three were flank eruptions (Fig. 3.1a) (Roult et al., 2012). No instrument response correction is applied, noting that all the stations are broadband with a flat response in the frequency ranges that we are interested in. We primarily use Z-component data in this study, though also do use three component data in clustering at a selected number of stations.

At Ruapehu in 2005–2008, the seismic network in close proximity to the volcano included a number of short-period instruments (natural frequency 1 Hz) and two broadband instruments. Following the previous work of Mordret et al. (2010), we downloaded seismic data from nine short-period instruments and the two broadband stations (Fig. 3.1b). An instrument response correction is then applied to all stations. Two stations were identified to have reversed polarity (DRZ and FWVZ stations) when examining the seismograms recorded after teleseismic earthquakes. We therefore flip the seismic waveforms recorded at these stations, with no further polarity changes observed within the time period analyzed.

3.3.2 Cross-correlation functions

Cross-correlation functions are computed using the python package MSNoise (Lecocq et al., 2014) following relatively standard processing methods briefly outlined here. Seismic data for each day are first grouped into continuous chunks, demeaned and tapered, and merged into 1-day traces. Traces are then bandpassed between 0.01 and 10.0 Hz and decimated to 25 Hz. Data for each station are then spectrally whitened and 1-bit normalization applied in 30-minute windows. These windows are used to compute cross-correlation functions between the different receivers. At Piton De La Fournaise, the data are whitened between both 0.1-1.0 Hz and 1.0-2.0

Hz to provide two sets of cross-correlation functions. The 0.1–1.0 Hz band reflects the frequencies processed by Obermann et al. (2013) previously. Similarly, at Ruapehu, we whiten the data between both 0.2-0.7 Hz and 1.8-4.0 Hz. The 0.2-0.7 Hz band reflects the frequencies processed by Mordret et al. (2010) previously and 1.8-4.0 Hz chosen to target tremor recorded at higher frequencies.

Cross-correlation functions — computed using the 30-minute windows — are then linearly stacked to give 1-day cross-correlation functions, with the option of further stacking possible. For both Piton De La Fournaise and Ruapehu, we perform linear stacking to give 10-day cross-correlation functions, where the function recorded for a given day reflects the cross-correlation function of that day and the nine previous cross-correlation functions.

3.3.3 Hierarchical clustering

We apply agglomerative hierarchical clustering to group cross-correlation functions, using functions defined within the Scipy python package (Virtanen et al., 2020). Unlike partitioning methods (e.g. K-means), hierarchical methods do not require prior determination of the number of clusters. For agglomerative clustering, each object (cross-correlation function) begins as a separate cluster at the start of the process. These are then sequentially merged (agglomerated) into new, larger, clusters until all objects are contained within a single cluster (Aminzadeh and Chatterjee, 1984). This process is typically visualised through a cluster tree, known as a dendrogram, which shows the aggregation level of the different objects. Specifically, the height (y-axis) reflects the distance between objects joined at various levels. The end result is to have a hierarchy of partitions, rather than a single partitioning of the dataset. This allows for relatively straightforward interpretation of the significance of different clusters at varied levels of similarity.

A final cluster solution can be extracted by choosing a distance threshold that splits the dendrogram. This threshold reflects the distance above which clusters will not be merged or, rather, the limit at which to 'cut' the dendrogram. Defining an appropriate threshold is not always straightforward. A lower threshold will produce a larger number of clusters of smaller size, whereas a higher threshold will produce fewer clusters of a larger size (thus reflecting only the more significant differences). We define this threshold for each case after visual inspection of the dendrogram, where we aim to highlight larger, clearly distinct, clusters. In particular, we focus on both cutting the tree across the longest links (vertical lines) between clusters and also identifying the point at which distances between sub-clusters become relatively small. A summary of the clustering process, applied to cross-correlation functions, is shown in Figure 3.2.

The choice of algorithm used to group objects together is an important consideration in hierarchical clustering. Popular algorithms include the single linkage (or nearest neighbor) method (e.g. Sibson, 1973; Czecze and Bondár, 2019), complete linkage (or furthest neighbor) method (e.g. Deyasi et al., 2017), the average linkage method (e.g. Marroquín et al., 2009; Benvegna et al., 2013), and the Ward method (e.g. Unglert et al., 2016a; Steinmann et al., 2022). We favour the average linkage method in

this study. For this, the distance between each pair of objects within each cluster is computed. These distances are summed together and then divided by the total number of pairs to give an average distance between clusters. At each step, the clusters with the smallest average linkage distance are merged until all objects are contained within a single cluster. We choose this method, relative to the others mentioned, after visually examining how clusters were formed relative to observed differences in the correlograms at Piton de la Fournaise (Fig. 3.12 and Fig. 3.13). This is discussed further in Section 3.5.1.

There are similarly many options available for distance measures, where the best choice is not always clear. We apply a correlation coefficient based distance measure, where we subtract the absolute value of the correlation coefficient between two cross-correlation objects from 1 to give a measure of dissimilarity. This value varies from 0 (perfect correlation) to 1 (no correlation). To reduce the influence of phase differences caused by velocity changes, we compute the maximum value of the correlation coefficient after incrementally stretching all waveforms relative to each other (from -1 to 1% stretch). The maximum value of the correlation coefficient is then used in the distance matrix for clustering. This is also discussed further in Section 3.5.1, along with alternative distance measures such as the Absolute Norm, Euclidean Ditance, and Dynamic Time Warping distance.

The correlation coefficient is computed within specific lag time windows of the cross-correlation function (both positive and negative sides). This is done to reflect processing used in seismic coda interferometry, where scattered arrivals are often targeted due to reduced sensitivity to variation in seismic sources and also larger accumulated delay times from changes in seismic velocity (Snieder, 2006). At Piton de la Fournaise, we use a static lag-time window of [10-45] seconds for 0.1–1.0 Hz cross-correlation functions, reflecting the processing performed by Obermann et al. (2013), and a window of [10-30] seconds at 1.0–2.0 Hz. At Ruapehu, we instead use a dynamic window based on excluding arrivals traveling faster than a user-defined ballistic wave speed. This accounts for the greater range of inter-station distances for this network, where a static window may not be appropriate. At both 0.2–0.7 Hz and 1.8–4.0 Hz, we define the start of the lag-time window as the inter-station distance divided by a wave speed of 500 m/s. We consider this velocity a conservative estimate of the minimum expected ballistic wave speed. The width of each window is chosen as 40 seconds and 20 seconds for 0.2–0.7 Hz and 1.8–4.0 Hz respectively.

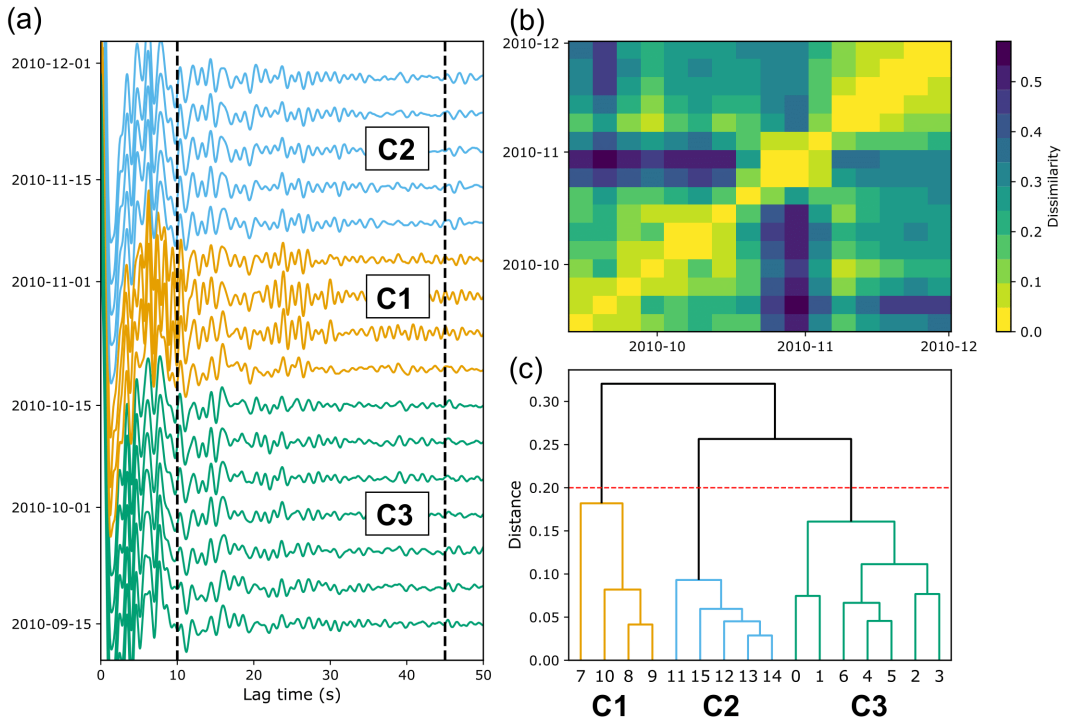

Figure 3.2: Hierarchical clustering of cross-correlation functions. (a): Positive lag-times of cross-correlation functions are shown, with vertical black dashed-lines showing the window used to compute the correlation coefficient (±10-45 s in this example). (b): Distance matrix containing the computed dissimilarity (1 minus the absolute value of the correlation coefficient) between each cross-correlation function in (a). This is used as input into the hierarchical agglomerative clustering algorithm. (c): Dendrogram showing the distance level that cross-correlation functions are grouped. Numbers on the x-axis reflect the order of cross-correlation functions in (a) from bottom to top, which begin as individual clusters. A distance threshold is chosen to define main clusters in the data, chosen here as 0.2 (red dashed-line). Colours in (a) and (c) correspond to the three clusters (C1–3) defined at this distance threshold.

3.3.4 Network covariance matrix spectral width

The network covariance matrix spectral width, introduced by Seydoux et al. (2016a), can be used to understand the level of spatial coherence (in terms of phase) of the seismic wavefield. We describe the basic principles of its computation here, though further details can be found in Seydoux et al. (2016a) and further studies that apply the methodology (e.g. Seydoux et al., 2016b; Journeau et al., 2020; Soubestre et al., 2021). The key idea is that the rank of the network covariance matrix reflects the number of independent signals that make up the seismic wavefield (Bartlett, 1954; Lawley, 1956). The more coherent the seismic wavefield, the larger the associated eigenvalues are. Smaller eigenvalues, in contrast, correspond to background noise (Wagner and Owens, 1996). The distribution of network covariance matrix eigenvalues is quantified through computation of the spectral width σ following:

$$\sigma(f) = \frac{\sum_{i=1}^{N} (i-1)\lambda_i(f)}{\sum_{i=1}^{N} \lambda_i(f)} \tag{3.1}$$

where the spectral width is computed as a function of frequency, and $\lambda_i(f)$ represents the eigenvalues arranged in decreasing order (Seydoux et al., 2016a). The spectral width then tends to zero when the seismic wavefield is dominated by a single source, and approaches a maximum when non-propagating, or incoherent, noise is considered. The value of this maximum depends on a few factors, such as the relative amplitudes of coherent and incoherent noise, the frequency of the data, and the interstation distances of the seismic network (Seydoux et al., 2016a).

We use the open-source Python Package Covseisnet (v0.4) (Covseisnet, 2020) to compute the network covariance matrix and its associated spectral width. The covariance matrix $C(f)$ is estimated by averaging Fourier cross-spectra matrices over M overlapping subwindows of length δt:

$$C(f) = \frac{1}{M} \sum_{m=1}^{M} u_m(f) u_m^{\dagger}(f) \tag{3.2}$$

where $u(f)$ is a vector composed of the Fourier spectra of N seismic stations, taking the form $u(f) = [u_1(f), u_2(f), ..., u_N(f)]^T$, and $u_m(f)$ is the Fourier spectra within each subwindow m and $\dagger$ signifying the Hermitian transpose. The choices of the number of windows M to average and the length of each subwindow δt, depend on network geometry, frequency range of interest, and temporal resolution (Seydoux et al., 2016a). The total duration of each time window used to estimate the network covariance matrix is then $\Delta t = Mr\delta t$ where r is the overlapping ratio (set at 50% overlap). In this study, we use sub-window length of 100 seconds, and average over 100 windows at both Piton De La Fournaise and Ruapehu. Finally, the network covariance matrix is decomposed into its eigenvalues and eigenvectors, taking the form:

$$C(f) = \sum_{n=1}^{N} \lambda_n(f) \nu_n(f) \nu_n^{\dagger}(f) \tag{3.3}$$

where ν_n are the complex eigenvectors. The network covariance matrix is Hermitian and, as such, the eigenvalues λ_n are real, positive, and associated with complex orthogonal eigenvectors. The spectral width is then calculated at each frequency following Equation 4.2. As a final step, we normalise our measurements of the spectral width within each time window by dividing each measurement by the maximum value recorded across the different frequencies. This greatly reduces the influence of changing spectral width values associated with variations in the number of seismic stations through time. We note however that this is an imperfect solution given that the spectral width does not converge to the same value at all frequencies (Seydoux et al., 2016b).

3.3.5 Seismic Velocity changes

Velocity changes are computed by applying the stretching technique (e.g. Sens-Schönfelder and Wegler, 2006), similar to our approach to reduce the influence of velocity changes in clustering results. The cross-correlation functions, computed following processing in Section 2.3, are stretched relative to a reference stack. This reference stack is typically chosen to represent the background state, with the main requirement being that it includes a greater number of stacked cross-correlation functions than the current, or moving, stacks it is being compared with. We then seek the maximum value of the correlation coefficient — within the same lag-time windows defined in Section 3.3.3 — between each current stack and reference stack at different levels of stretching up to factor of $\pm 1\%$. We do not observe any evidence to suggest velocity changes larger than this are present within our datasets. Once the maximum value is identified, the corresponding stretching coefficient is converted into an apparent velocity change and the correlation coefficient is recorded. This provides some understanding of the quality of the measurement. At Piton de la Fournaise, we use the same reference period as used by Obermann et al. (2013), from 2010-06-01 to 2010-09-01. At Ruapehu, we use a reference function from the entire data period.

3.4 Results

In the following section we show the results of clustering cross-correlation functions using specific station-pairs at Piton de la Fournaise and Ruapehu. This is likely sufficient to capture the most significant features in our datasets, assuming any large changes will be present across many station-pairs. At Piton de la Fournaise, we focus on the three station-pairs examined in greater detail by Obermann et al. (2013) (Fig. 3.1a). At Ruapehu, we focus on a single-pair of broadband seismic stations (Fig. 3.1b) that is approximately aligned radially with the summit region and the location of tremor activity (Sherburn et al., 1999). Spectral width computations, in contrast, use all stations shown in Figure 3.1.

Additional parameters are also computed to understand the characteristics of cross-correlation functions within each cluster. The mean value of the normalized spectral width is computed as a function of frequency within the time period of clusters of interest (e.g. Fig. 3.3f). This allows us to assess if changes in seismic wavefield coherence are responsible for different clusters. We also compute an average cross-correlation function to represent each cluster, using 1-day functions within the corresponding time periods (e.g. Fig. 3.3h). From these, we examine both the frequency content (e.g. Fig. 3.3g) and also the mean coherence between different clusters as a function of lag time (e.g. Fig. 3.3i). This value reflects the coherency between two waveforms computed within a moving window, where the final recorded value is an average over the different frequencies. The length and overlap of the moving window is chosen depending on the frequency content of the cross-correlation functions, and is indicated in the captions for Figures 3.3–3.10.

3.4.1 Piton De La Fournaise

The period examined at Piton De La Fournaise includes four eruptive events. Of these, three were flank eruptions, while the eruption of 2010-01-02 occurred within the summit region (Fig. 3.1a). The eruptions of January and October 2010 lasted 10 and 17 days respectively (Roult et al., 2012), with accompanying tremor activity above approximately 1 Hz clearly revealed by reduced values of the covariance matrix spectral width (Fig. 3.3a). The other two eruptions, in contrast, were short-lived — less than one day (Roult et al., 2012) — and therefore appear only as a thin vertical line of reduced spectral width above 1 Hz. We also clearly observe the microseism band at approximately 0.1-0.3 Hz. Further details related to specific features of the covariance matrix spectral width analysis at Piton De La Fournaise during this time period can be found in two recent papers (Seydoux et al., 2016a; Journeau et al., 2020).

3.4.1.1 Clusters at lower frequencies (0.1–1.0 Hz)

At frequencies between 0.1–1.0 Hz, the cross-correlations can be grouped into three relatively distinct clusters. For station-pairs UV05-UV12 and UV07-UV09, this split is achieved with a distance threshold of 0.2 (Fig 3.3 & 3.4e-h), whereas a slightly lower threshold of 0.19 produces three clusters for station-pair UV01-UV07 (Fig. 3.4a-d). Focusing on station-pair UV05-UV12, the two largest clusters (C2 and C3 in Fig. 3.3) are separated predominantly by the October 2010 eruption. Specifically, cluster C2 consists mostly of cross-correlation functions before the eruption while cluster C3 consists mostly of cross-correlation functions after (Fig. 3.3c). This is similarly reflected by reduced values of the correlation-coefficient between post-eruption cross-correlation functions and the pre-eruptive reference stack (Fig. 3.3d). There are, however, cross-correlations prior to the October eruption that are part of cluster C3. Notably, the cross-correlations prior to eruptive activity in January 2010 are grouped together with those immediately following the October 2010 eruption. Similarly, a number of cross-correlation functions post-March 2011, months after the October eruption, are grouped with those pre-eruption (Cluster C2).

The most distinct cluster in the dendrogram (C1, Fig. 3.3e), begins just after the start of tremor activity in October 2010 (C1 in Fig. 3.3) and ends shortly after. The influence of the tremor is apparent in the correlogram for the station-pair UV05-UV12 (Fig. 3.3b), albeit with a slight temporal offset that reflects the 10-day stacking process. Specifically, there is a delay before the stacked cross-correlation functions converge to reflect changes in the seismic source. In this case, we see that there are reduced spectral width values above approximately 0.8 Hz for this co-eruptive cluster relative to the two larger clusters (Fig. 3.3f). This is also reflected in the frequency content of the corresponding averaged cross-correlation function for cluster C1, where it is shown to have more energy above approximately 0.8 Hz (Fig. 3.3g). Finally, we also see a reduced coherence at early and later lag times between the co-eruptive cluster and the pre-2010 eruption cluster (C1 vs. C2 in Fig. 3.3i) relative to the comparison between the two larger clusters (C2 vs C3).

Clustering single-station cross-component cross-correlation functions (components

EN, EZ, and NZ) in the same frequency band reveals similar features to those obtained clustering with station-pairs (Fig. 3.5 and 3.8). At UV12 station, using the same distance threshold of 0.2, we observe that cross-correlation functions are split into three clusters (Fig. 3.8). These follow the pattern of the station-pair cross-correlation functions where the clusters are predominantly separated by the October 2010 eruption, including a cluster associated with the onset of volcanic activity. We do, however, note that the cluster consisting mostly of post-eruption cross-correlation begins 1–2 weeks prior to the eruption for component-pairs EN and NZ (Fig. 3.8b and 3.8j). At UV07 station, which is located at a greater distance from the location of the October 2010 eruption (Fig. 3.1a), we observe that only one component-pair (NZ) places the cross-correlation functions pre- and post-October eruption into separate clusters at the same threshold (Fig. 3.5j-l). For component-pairs EN and EZ, the data are split into only two clusters consisting of those computed during the October 2010 eruption and those outside of this time period (Fig. 3.5a-h).

Finally, we cluster the auto-correlations (EE, NN, ZZ) at both stations (Fig. S3.1 and S3.2). Unlike with the cross-components, a larger threshold is required to reduce the dataset to a small number of clusters. At a threshold of 0.5, we still observe many clusters at both UV07 and UV12 station. Further, there is no clear relationship between the timing of these clusters and volcanic activity. We suspect this could be related to differences in processing, where we do not whiten seismic data prior to computing the auto-correlation.

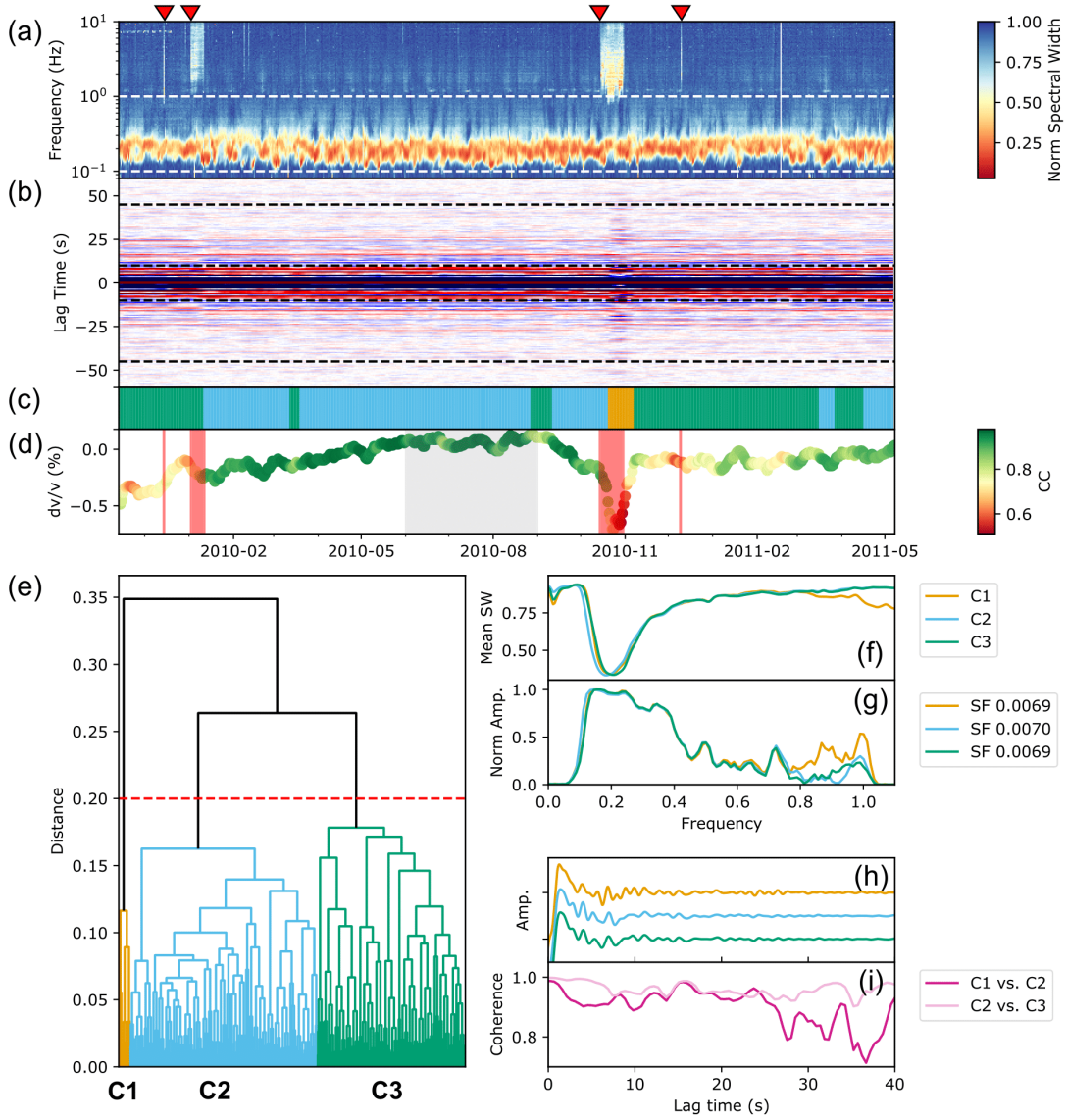

Figure 3.3: Cluster output using 0.1–1.0 Hz cross-correlation functions (CCFs) for station-pair UV05-UV12 at Piton de la Fournaise (blue line in Fig. 3.1a). (a): Normalized spectral width measurement. Lower values indicate a more coherent seismic wavefield dominated by fewer seismic sources. Dashed white lines show frequency range of CCFs. Triangles above (red) indicate eruption start times (b): Correlogram showing amplitudes of CCFs at different lag times (red = positive, blue = negative). Dashed black lines show part of CCFs used in clustering. (c): Location of clusters in time, color-coded according to dendrogram output in (e). (d): Apparent velocity changes, color-coded according to correlation coefficient (CC) computed between 10-day stacks and the reference stack period (shaded gray). Vertical red bars denote timing of eruptive activity. (e): Dendrogram, with clusters defined at a distance threshold of 0.2 (dashed-red line). (f): Mean value of spectral width as a function of frequency within time period of each cluster. (g): Normalized frequency spectrum of CCFs within each cluster, computed from averaged CCF. Legend shows Scale Factor (SF) to convert from non-normalized spectrum. (h): Average of CCFs within each cluster at positive lag times. (i) Mean coherence, averaged over CCF frequency range, between different clusters at positive lag times, computed using moving window of 10-seconds length with 0.5 second step (95% overlap).

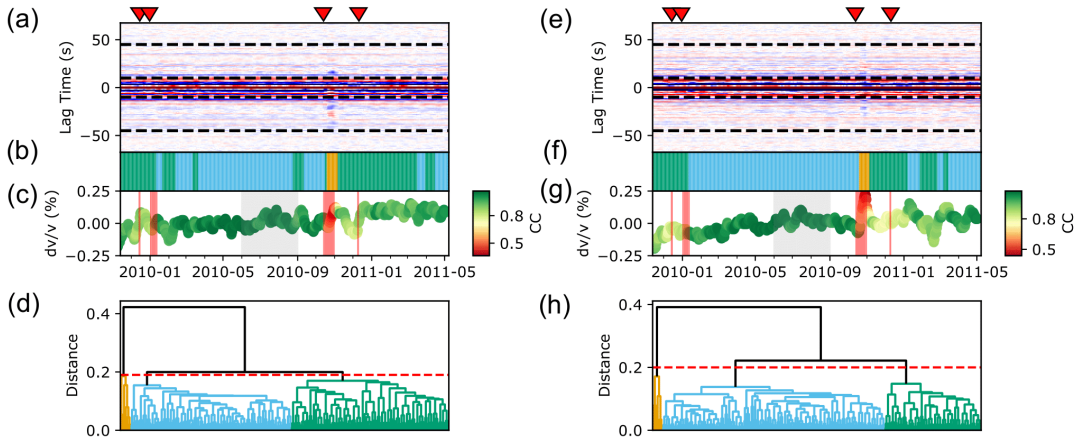

Figure 3.4: Clustering output using 0.1–1.0 Hz cross-correlation functions (CCFs) at station-pairs UV01-UV07 (green-line Fig. 3.1a, subplots a-d) and UV07-UV09 (red-line Fig. 3.1a, subplots e-h) at Piton de la Fournaise. (a/e): Correlogram showing amplitudes of CCFs at different lag times (red = positive, blue = negative). Dashed black lines show parts of CCFs used in clustering. Triangles above (red) indicate eruption start times (b/f): Location of clusters in time, color-coded according to dendrogram output. (c/g): Apparent velocity changes, color-coded according to correlation coefficient (CC) computed between 10-day stacks and the reference stack period (shaded gray). Vertical red bars denote timing of eruptive activity. (d/h): Dendrogram, with clusters defined at a distance threshold indicated by dashed-red line.

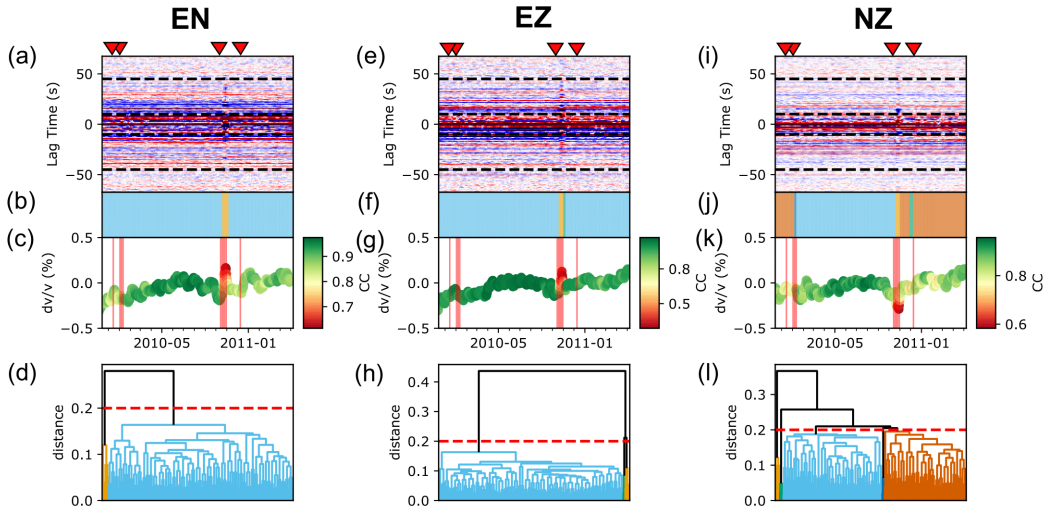

Figure 3.5: Clustering single-station cross-component correlation functions computed using station UV07 at Piton de la Fournaise (0.1–1.0 Hz). The different component combinations are EN (a-d), EZ (e-h), and NZ (i-l). (a/e/i): Correlogram of cross-correlation functions. (b/f/j): Location of clusters in time, color-coded according to dendrogram outputs. (c/g/k): Apparent velocity changes, color-coded according to correlation coefficient (CC) computed between 10-day stacks and the reference period. (d/h/l) Dendrogram, with clusters defined at a distance threshold of 0.2.

3.4.1.2 Clusters at higher frequencies (1.0–2.0 Hz)

At higher frequencies, between 1.0–2.0 Hz, we observe a greater number of, seemingly distinct, clusters of cross-correlation functions relative to the lower frequencies (Fig. 3.6 and 3.7). For station-pair UV05-UV12, we identify six clusters at a distance threshold of 0.5 that, on examination of the dendrogram, appear to be relatively distinct (Fig. 3.6e). For example, it is only at a distance threshold of approximately 0.75 that we would have three clusters to match those at lower frequencies (Fig. 3.6e). For the other two station-pairs (UV01-UV07 and UV07-UV09), we observe slightly reduced complexity with only four clusters defined at a lower distance threshold of 0.4 (Fig. 3.7). The timing of these clusters shows a strong consistency with the cluster solution at lower frequencies. For example, we again observe different clusters either side of the October 2010 eruption. For station-pair UV05-UV12, pre-eruptive cross-correlation functions are mostly grouped into clusters C5 and C6 while post-eruptive cross-correlation functions form the cluster C2 (Fig. 3.6c). We also observe again a cluster that coincides with the onset of eruptive activity (C3). This is also true of eruptive activity in January 2010 (C4). The effect on cross-correlation functions is clearly visible from the correlogram for both tremor episodes (Fig. 3.6b). Similarly, we observe a strong consistency between the timing of different clusters and the cross-correlation coefficient computed between the 10-day stacks and the reference stack (Fig. 3.6d). Only the cluster that encompasses the reference period (C5) maintains a high value of correlation coefficient. In contrast, cluster

C6, which includes cross-correlation functions either side of this period, has reduced values of the correlation coefficient. This includes a period of one month prior to the October 2010 eruption. Similarly, we observe that the post-eruptive cluster (C2) has a correlation coefficient that is too low to accurately compute apparent velocity changes when using 10-day stacks and a pre-eruption reference function (Fig. 3.6d).

Cluster C3 in Fig. 3.6, occurring during volcanic tremor in October 2010, is shown to contain cross-correlation functions that are different in character to cross-correlation functions computed during non-eruptive periods. The frequency spectra for the average cross-correlation function computed from this cluster is distinct from other clusters, especially between 1.2–1.7 Hz (Fig. 3.6g). In contrast, we see that the other clusters (C2, C5, C6) are similar in their spectra; though, we note that the post-eruptive cluster (C2) appears to have reduced energy at frequencies above approximately 1.7 Hz. We also see differences in the coherence against lag time computed between the co-eruptive cluster C3 and the inter-eruptive cluster C5 relative to comparisons between other clusters (Fig. 3.6i). In particular, we observe lower values of coherence at the earliest lag times (shaded gray in Figure 3.6i). At later lag times, however, there is less difference between the coherence calculated between different clusters.

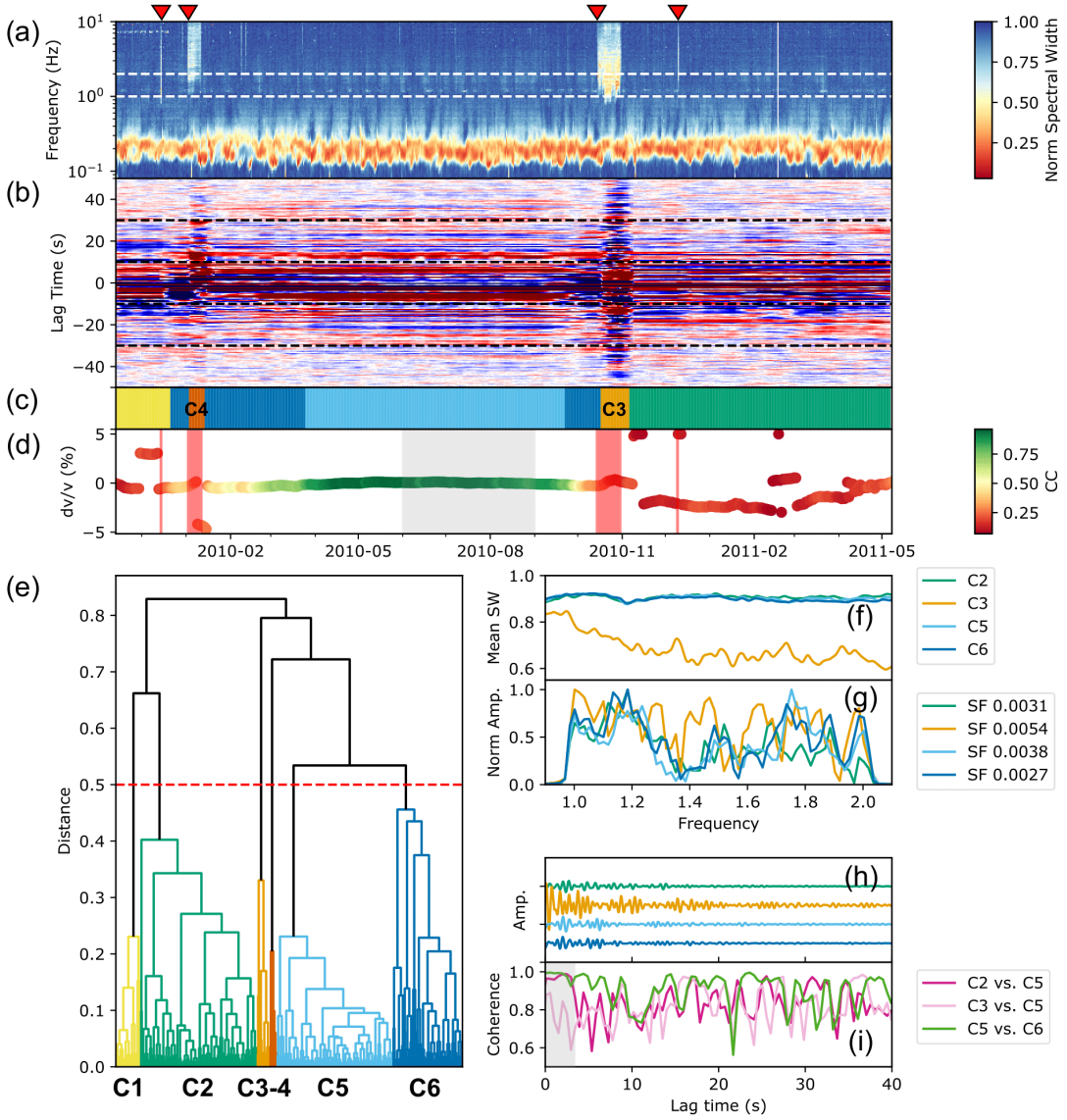

Figure 3.6: Clustering output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pair UV05-UV12 at Piton de la Fournaise (blue line in Fig. 3.1a). Subplots follow those in Fig. 3.3, briefly described below. (a): Normalized spectral width measurement. (b): Correlogram. (c): Location of clusters in time (d): Apparent velocity changes with correlation coefficient (CC) relative to reference stack period (shaded gray). (e): Dendrogram (f): Mean spectral width as function of frequency within time period of each cluster. (g): Normalized frequency spectrum of CCFs within each cluster. (h): Average of CCFs within each cluster at positive lag times. (i) Mean coherence between clusters computed using moving window of 4-seconds length with 0.5 second step (88% overlap). Grey shading shows lag times where ballistic arrivals would be expected.

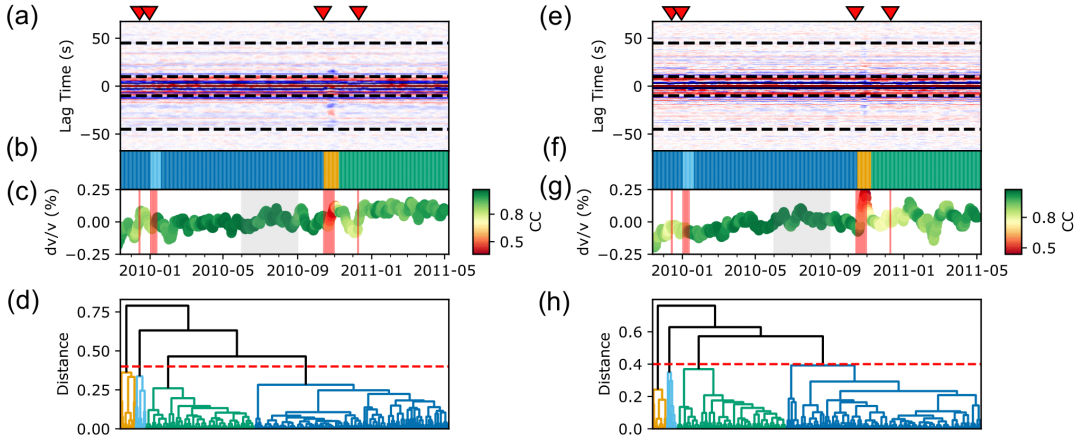

Figure 3.7: Clustering output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pairs UV01-UV07 (green line Fig. 3.1)a, subplots a-d) and UV07-UV09 (red line Fig. 3.1a, subplots e-h). Subplots follow those in Fig. 3.4, briefly described below. (a/e): Correlogram (b/f): Location of clusters in time (c/g): Apparent velocity changes with correlation coefficient (CC) relative to reference stack period (shaded gray). (d/h): Dendrogram.

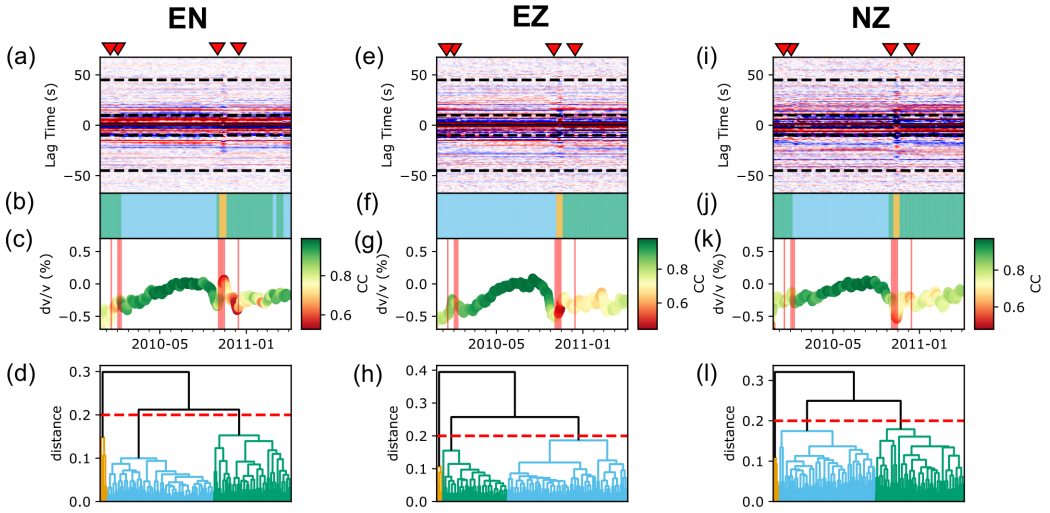

Figure 3.8: Clustering single-station cross-component correlation functions computed using station UV12 at Piton de la Fournaise (0.1–1.0 Hz). Subplots follow those described in Figure 3.5.

3.4.2 Ruapehu

At Ruapehu, we examine a 3-year time period that includes two phreatic eruptions. These were on 4 October 2006 and 27 September 2007 (Mordret et al., 2010). The seismic wavefield at lower frequencies shows near-continuous energy between approximately 0.1–0.7 Hz, reflected by reduced covariance matrix spectral width values (Fig. 3.9a). This likely relates to the oceanic microseisms. At higher frequencies we observe intermittent periods of reduced spectral width measurements, mostly between 1–3 Hz. In particular, the first half of both 2006 and 2007, and post-2007 eruption, have reduced spectral width values for several months in this frequency band. These periods likely reflect increased volcanic tremor activity, where tremor with a single peak between 1.8–2.3 Hz is common at this volcano (Latter and Balsillie, 1984; Hurst and Steven, 1993; Hurst, 1998; Sherburn et al., 1999). This energy propagates as Rayleigh waves at greater distances from Crater Lake (> 2 km), with body waves suggested to make a more significant contribution at smaller distances (Hurst and Steven, 1993).

Tremor at 3 Hz and 7 Hz has also been observed at the volcano, though episodes are less frequent and only observed at stations closest to the summit (Sherburn et al., 1999). Thus, the sources of 3 Hz and 7 Hz tremor are thought to be shallower than the source of 2 Hz tremor, suggested to occur within a narrow zone beneath the Crater Lake within a depth range of approximately 1 km from the surface (Hurst, 1998; Sherburn et al., 1999).

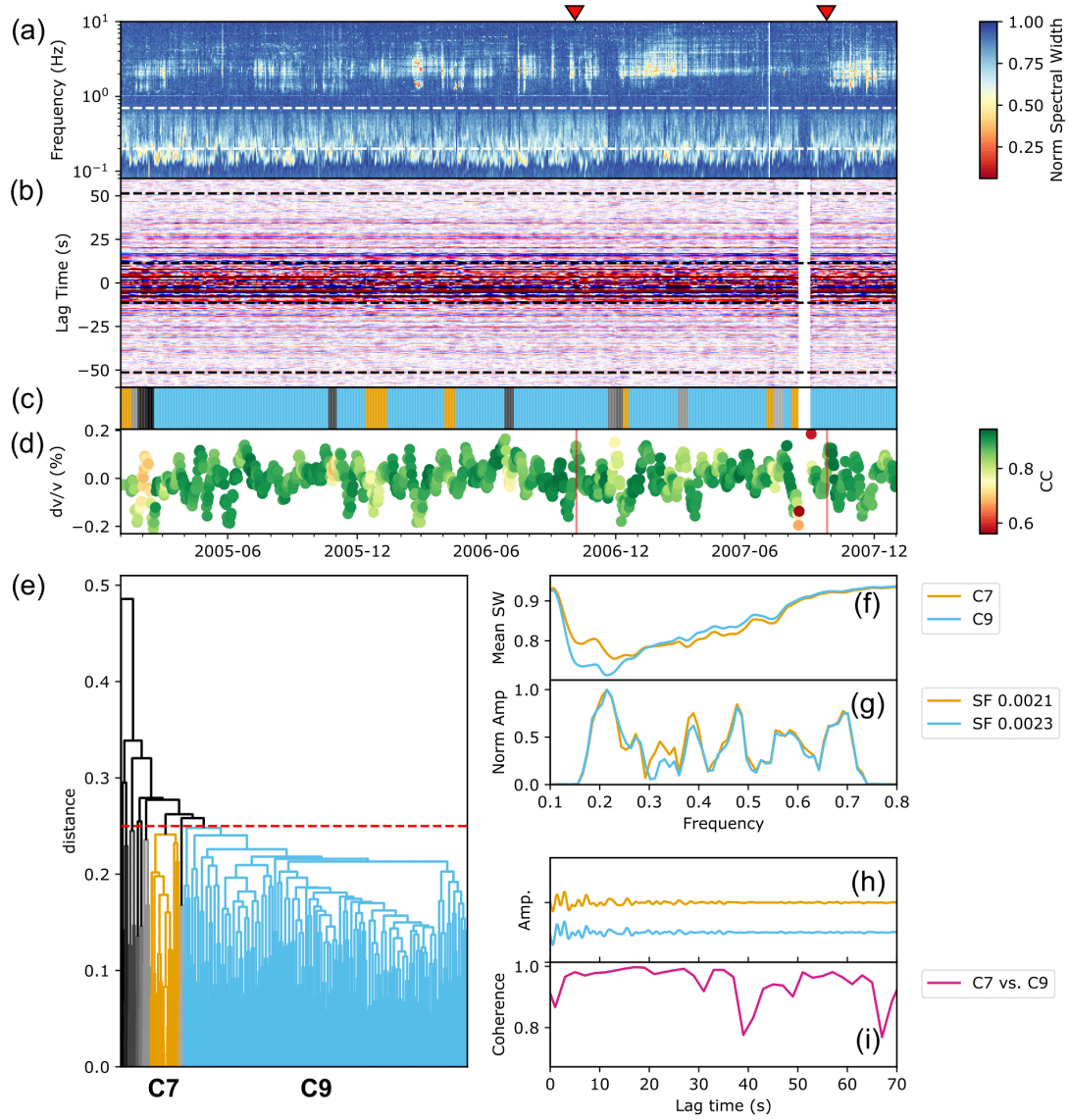

Figure 3.9: Clustering output using 0.2–0.7 Hz cross-correlation functions (CCFs) for station-pair FWVZ-WPVZ at Ruapehu (blue line in Fig. 3.1b). Subplots follow those in Fig. 3.3, briefly described below (a): Normalized spectral width measurement. (b): Correlogram. (c): Location of clusters in time (d): Apparent velocity changes with correlation coefficient (CC) relative to reference stack period (full period). (e): Dendrogram with smaller, unlabeled, clusters (e.g. C1–6 and C8) are coloured using greyscale. (f): Mean spectral width as function of frequency within time period of each cluster. (g): Normalized frequency spectrum of CCFs within each cluster. (h): Average of CCFs within each cluster at positive lag times. (i) Mean coherence between clusters computed using moving window of 10-seconds length with 0.5 second step (95% overlap).

3.4.2.1 Clusters at lower frequencies (0.2–0.7 Hz)

At the lower frequencies (0.2–0.7 Hz), we observe no obvious suggestion that the cross-correlation functions are varying significantly over time. At a threshold of 0.25, the majority of the data are grouped together in a single cluster (C9 in Fig. 3.9)

with only small periods in time where data is grouped as another cluster. Furthermore, the dendrogram does not suggest a lower threshold would reveal much further structure to the data, with similar distances between individual cross-correlation functions within this cluster (Fig. 3.9e). This is further exemplified when we compare cluster C9 with the second largest cluster C7 (in terms of number of cross-correlations they contain). The frequency spectrum of the averaged cross-correlation functions are found to be highly similar between both clusters (Fig. 3.9g). Similarly, only small differences are observed in the spectral width for their corresponding time differences, mostly at approximately 0.1–0.3 Hz (Fig. 3.9f).

The measured coherence between the two clusters is close to 1 at lag times between approximately 5–35 seconds, with lower values measured in the first few seconds and later lag times. With an inter-station distance of approximately 6 km for station-pair FWVZ-WPVZ, it can be expected that the first few seconds of the cross-correlation function will not contain any ballistic arrivals. Thus, reduced coherence at these early lag times is not surprising. At later lag times, we can expect that the coherence will decrease as energy attenuates. This could explain larger decreases at approximately 40 and 65 seconds, where we note the coherence appears less stable relative to earlier lag times.

3.4.2.2 Clusters at higher frequencies (1.8–4.0 Hz)

At higher frequencies (1.8–4.0 Hz), the influence of tremor on the cross-correlations is much more apparent. In the correlogram, periods of increased tremor activity are reflected by larger amplitude cross-correlation functions (Fig. 3.10b). At a threshold of 0.4, the majority of data are split into two clusters (C21 and C22 in Fig 3.10c and 3.10e). These correspond approximately to periods of elevated tremor activity, with the exception of clusters C6 and C20. These two clusters consist of cross-correlation functions from approximately 2–3 months before the eruption in 2006 (C20) and the months following the 2007 eruption (C6). Features of the spectral width measurements for time periods corresponding to each of the four clusters are similar, though with differences in the absolute values (Fig. 3.10f). These differences are also reflected in the absolute values of peaks in the frequency spectrum of the average cross-correlation functions for each cluster, where the scale factor applied in Fig. 3.10g appears correlated with the value of the spectral width. For example cluster C6, which has the lowest values of the spectral width, has amplitudes two times larger than cluster C22 in the frequency spectrum (Fig. 3.10g). Once normalized though, we again see strong similarities in the frequency content of the averaged cross-correlation functions between the clusters.

Finally, we observe that the coherence is high at most lag times when comparing the different clusters, including when we compare the two smallest clusters (C6 and C20) with the largest cluster C21 (Fig. 3.10i). This is also the case at early lag times, with values of coherence close to 1 (grey-shaded lag times in Fig. 3.10i). Finally, we observe periods where only small clusters of cross-correlation functions are formed; for example, the months prior to the 2007 eruption. These periods are associated with reduced tremor activity and, as a result, reduced amplitudes in the correlogram (Fig. 3.10a and 3.10b).

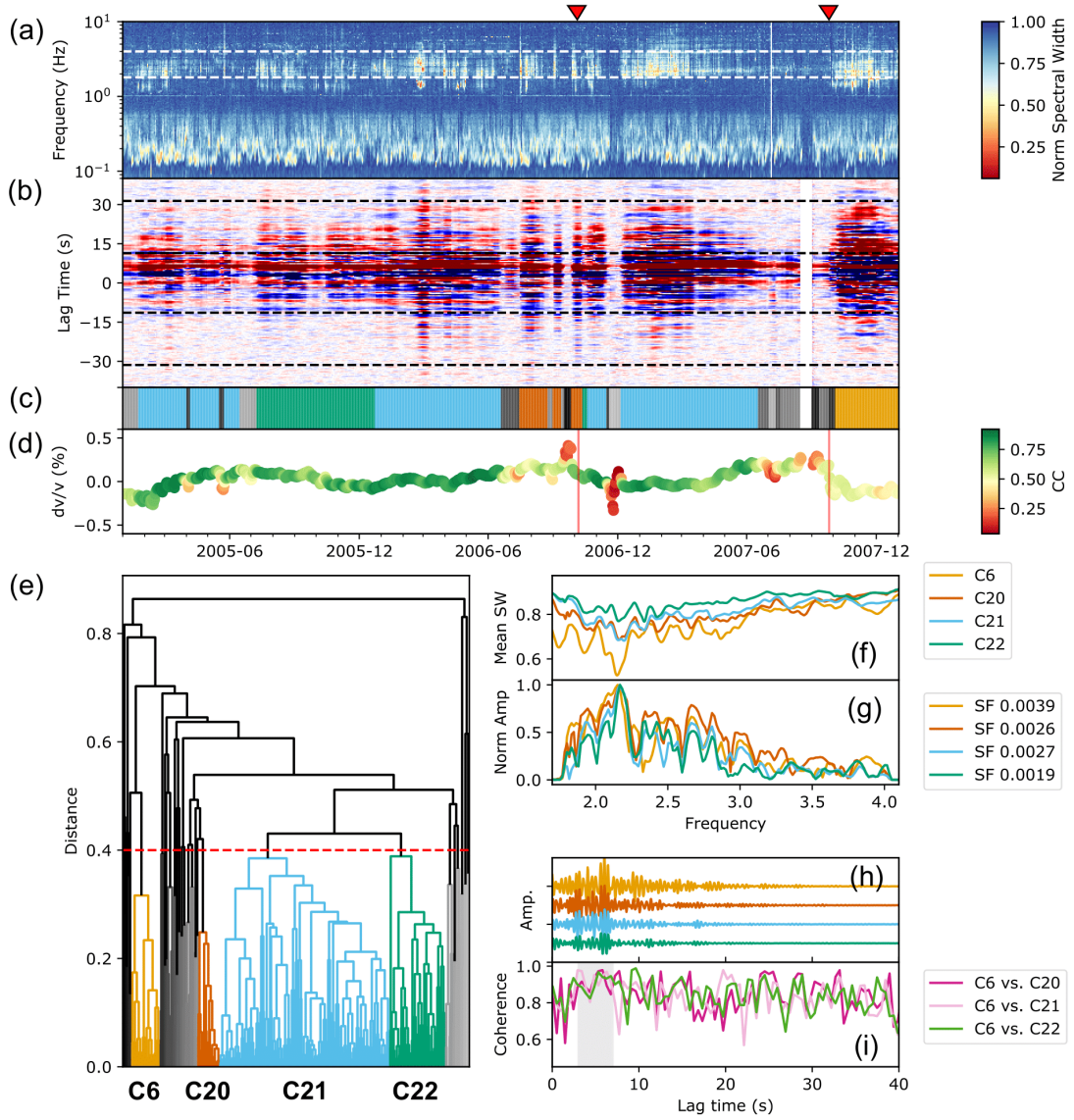

Figure 3.10: Clustering output using 1.8–4.0 Hz cross-correlation functions (CCFs) for station-pair FWVZ-WPVZ at Ruapehu (blue line in Fig. 3.1a). Subplots follow those in Fig. 3.3, briefly described below. (a): Normalized spectral width measurement. (b): Correlogram. (c): Location of clusters in time (d): Apparent velocity changes with correlation coefficient (CC) relative to reference stack period (full period). (e): Dendrogram with smaller, unlabeled, clusters (e.g. C1–5 and C7–C19) are coloured using greyscale. (f): Mean spectral width as function of frequency within time period of each cluster. (g): Normalized frequency spectrum of CCFs within each cluster. (h): Average of CCFs within each cluster at positive lag times. (i) Mean coherence between clusters computed using moving window of 2-seconds length with 0.5 second step (75% overlap). Grey shading shows lag times where ballistic arrivals would be expected.

3.5 Discussion

In this study, we apply hierarchical clustering to cross-correlation functions computed using data from two seismic networks (Piton de la Fournaise and Ruapehu). We identify clusters of cross-correlation functions in both datasets, at different frequencies, that appear related to both volcanic activity and changes in the seismic source. A summary of the clusters as they appear in the results section is shown in Figure 3.11. Here, we evaluate the usage of clustering cross-correlation functions toward assessing their similarity through time and the additional insights that could be obtained. Finally, we discuss the application of clustering cross-correlation functions as part of the monitoring process and possible future directions.

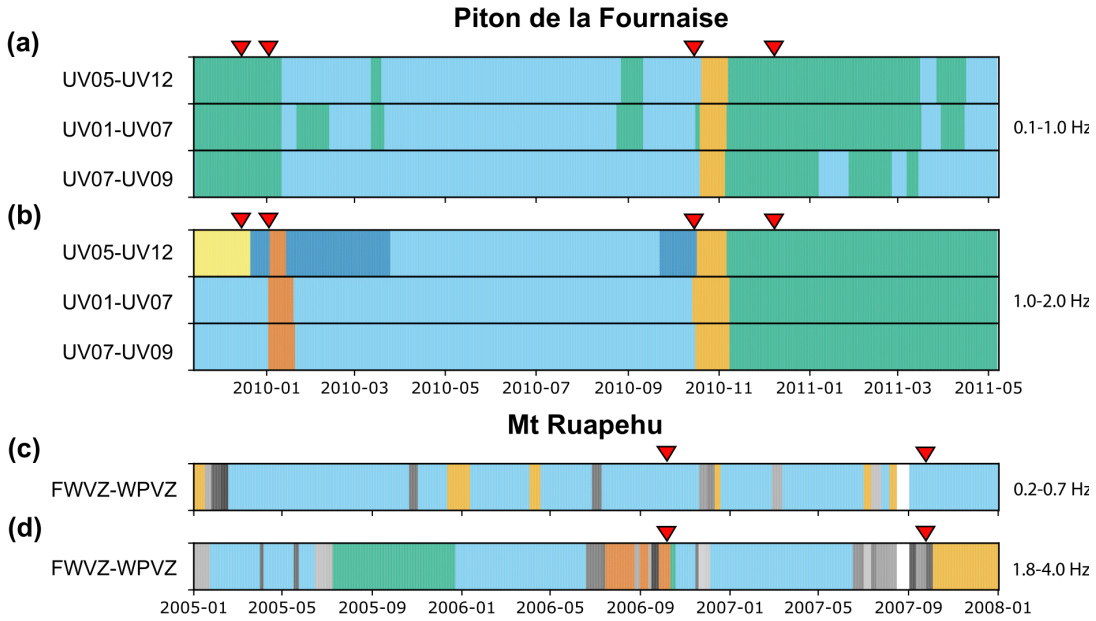

Figure 3.11: Summary of the cluster positions in time at both Piton de la Fournaise and Ruapehu, as shown in the results section (a) 0.1–1.0 Hz at Piton de la Fournaise. (b) 1.0–2.0 Hz at Piton de la Fournaise. (c) 0.2–0.7 Hz at Ruapehu. (d) 1.8–4.0 Hz at Ruapehu.

3.5.1 Choice of algorithm and distance measure

The choice of algorithm and distance measure used in hierarchical clustering is an important consideration. Here, use the average linkage method combined with a correlation coefficient based distance measure. Specifically, we define the distance as one minus the absolute value of the correlation coefficient between two cross-correlation objects (giving a measure of dissimilarity).

Alternatives choices for the clustering algorithm included the single linkage (nearest neighbor) method, the complete linkage method, and the Ward method. The choice of the average linkage algorithm was based on visual examination of how clusters were formed relative to observed differences in the correlograms at Piton de

la Fournaise (Fig. 3.12 and Fig. 3.13). For example, the single-linkage algorithm did not appear optimal for detecting clusters associated with tremor (Fig. 3.13a-d), while the Ward's-linkage method similarly did not clearly separate the two episodes of volcanic tremor originating from different locations (Fig. 3.13i-l). In any case, we note clear similarities in the clusters identified using all tested methods. Future studies may therefore wish to explore these options further when determining an appropriate algorithm for a given dataset.

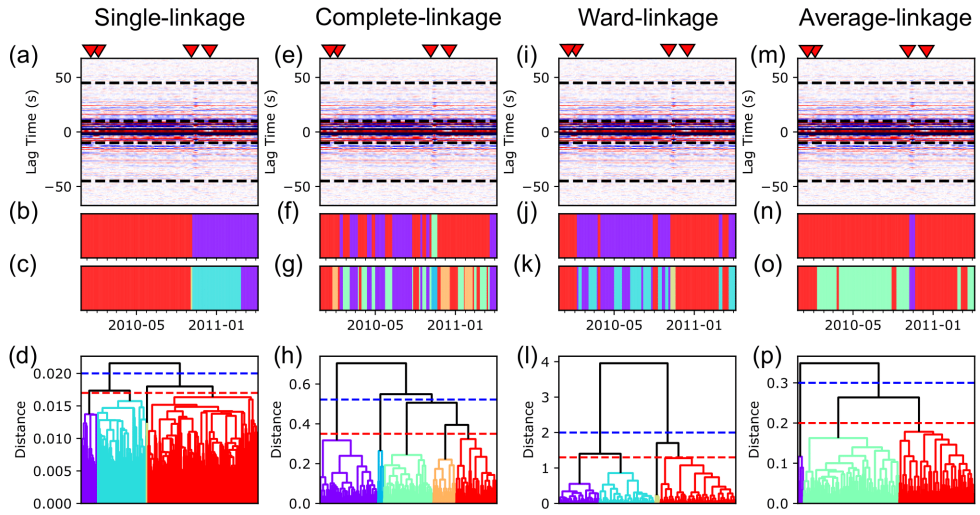

Figure 3.12: Different clustering algorithms applied to group 0.1–1.0 Hz cross-correlation functions computed from Piton de la Fournaise seismic data, station-pair UV05-UV12. Algorithms are single-linkage (a-d), complete-linkage (e-h), ward-linkage (i-l), and average-linkage (m-p, used in this study). (a/e/i/m): Correlogram of cross-correlation functions, repeated for easier comparison with different algorithms. (b/f/j/n): Cluster locations in time using distance threshold denoted by dashed-blue line in dendrograms. (c/g/k/o): Cluster locations in time using distance threshold denoted by dashed-red line in dendrograms. (d/h/i/p): Dendrograms with distance thresholds shown.

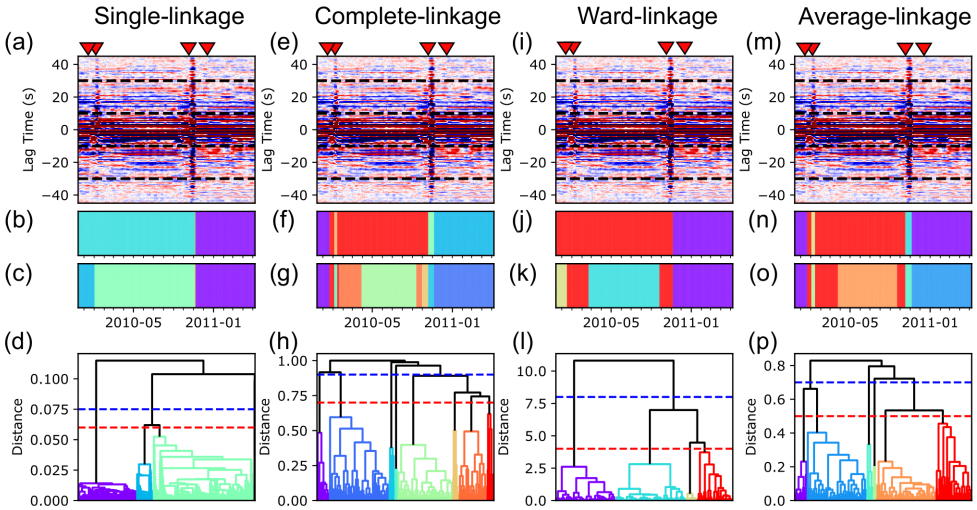

Figure 3.13: Different clustering algorithms applied to group 1.0–2.0 Hz cross-correlation functions computed from Piton de la Fournaise seismic data, station-pair UV05-UV12. Subplots follow those described in Figure 3.12.

For the distances measures, alternatives tested included the Absolute Norm, Euclidean Norm, and Dynamic Time Warping distance. Dynamic Time Warping is equivalent to minimizing the Euclidean distance between two time series under all possible temporal alignments (usually subject to constraints). Dynamic Time Warping is equivalent to minimizing the Euclidean distance between two time series under all possible temporal alignments (usually subject to constraints). Thus, it can be useful if misalignment is expected; for example, due to changes in seismic velocity. We demonstrate this using synthetic cross-correlations with a velocity change introduced (Fig. 3.14 and 3.15). We also observe that Dynamic Time Warping can be effective for clustering cross-correlation functions with low SNR (Fig. 3.16). However, the relative insensitivity of the correlation coefficient based distance measure to changes in amplitude, combined with it being a commonly used similarity measure for cross-correlation functions, were ultimately the deciding factor for this study. As with the choice of algorithm though, this can certainly be explored further in future applications.

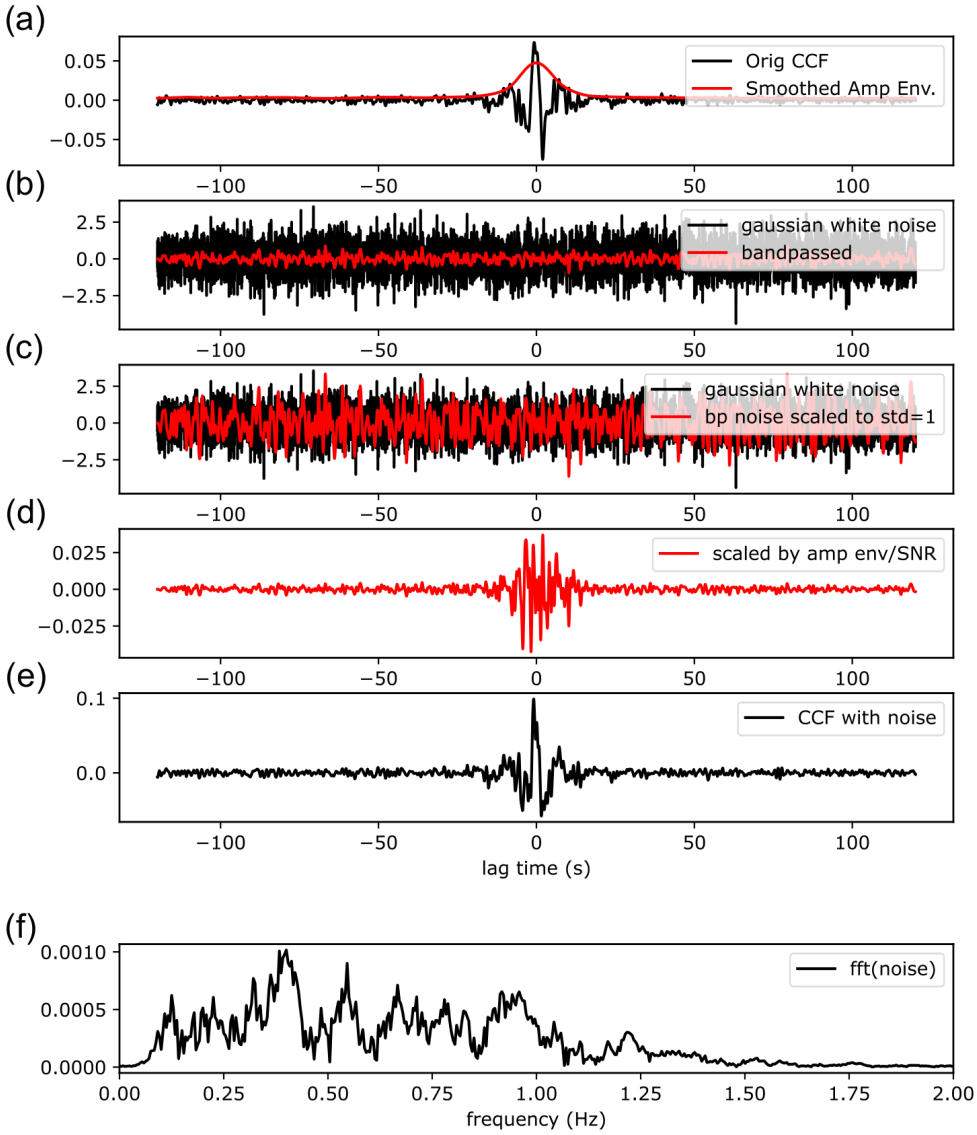

Figure 3.14: Synthetic cross-correlation creation: (a) Take the smoothed amplitude envelope of a single cross-correlation function (orig. CCF in legend). (b) Bandpass gaussian white noise in frequency band of the cross-correlation function. (c) Scale bandpassed noise (bp noise in legend) to have standard deviation of 1. (d) Scale the noise using smoothed amplitude envelope divided by user-defined SNR. (e) Add scaled noise back to the original cross-correlation function. If introducing velocity change, stretching is applied to the cross-correlation function prior to adding noise. (f) Frequency spectrum of synthetic noise.

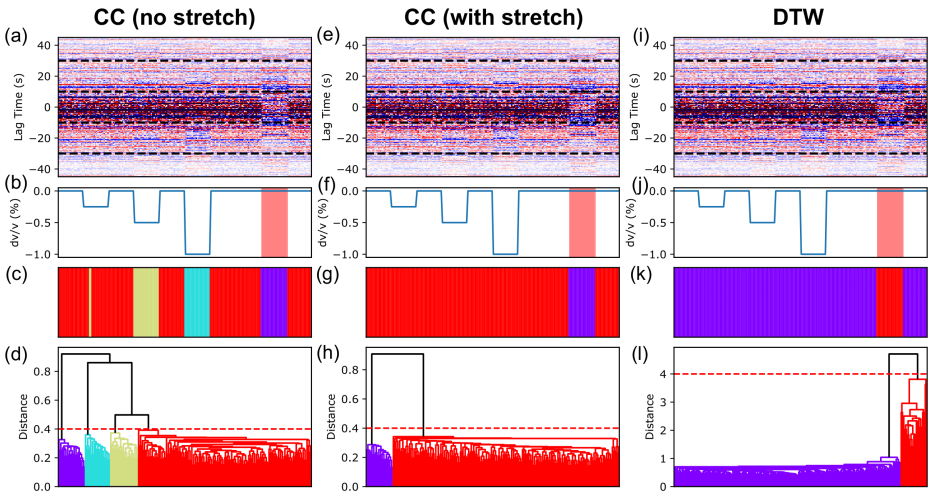

Figure 3.15: Demonstration of different distance measures used in clustering when velocity change is introduced to synthetic cross-correlation functions (1.0–2.0 Hz). Synthetics created with SNR=2. The station-pair UV05-UV12 is primarily used to compute synthetics, with the exception of a small time period where the station-pair UV01-UV07 is used (indicated by red shaded region in b/f/j). This acts as a group of cross-correlation functions where we would expect clustering to identify a clear difference. Velocity changes are introduced at increasing levels up to a maximum of 1.0%, also shown in b/f/j. (c-d): Cluster results for Correlation-Coefficient (CC) distance measure without any prior stretching to individual cross-correlation functions, at a threshold of 0.4. (g-h) Cluster results for CC distance measure with stretching of cross-correlation functions relative to each other to find max CC). (k-l) Cluster results for Dynamic Time Warping (DTW) distance measure.

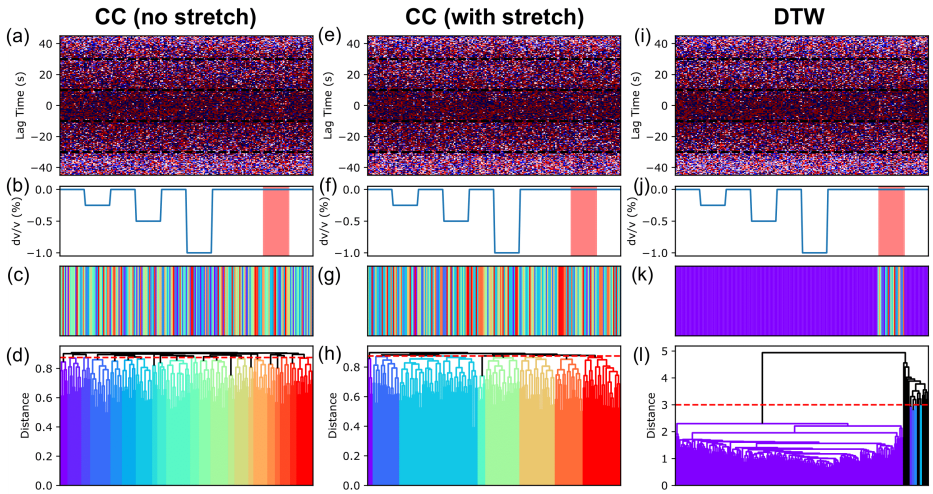

Figure 3.16: As with Fig. 3.15, but with reduced SNR of synthetic cross-correlations (set to 0.1 instead of 2). We observe that a Dynamic Time Warping (DTW) distance measure can still recognise similarity, and correct for, velocity changes in synthetic cross-correlation functions computed using the same station-pair (k-l). In contrast, for the correlation coefficient distance measure, there is no suggestion from the dendrogram that similarity is identified in low SNR cross-correlation functions, with or without stretching c-d and g-h respectively). For example, unlike with DTW, there is no obvious threshold that separates the synthetic cross-correlation functions computed using station-pair UV05-UV12 and those using UV01-UV09 (red shading in b/f/j).

3.5.2 Clustering cross-correlation functions as a similarity measure

Our application of clustering aims to establish its suitability as a measure of similarity in cross-correlation functions. Piton de la Fournaise represents an ideal case study in this regard, where the previous work of Obermann et al. (2013) provides the context to interpret the changes we see in clusters associated with the October 2010 eruption.

A dominant feature in cross-correlation functions computed at Piton de la Fournaise is the different clusters assigned before and after the October 2010 eruption. This is observed for multiple station-pairs in both frequency bands (Fig. 3.11a & 3.11b). At lower frequencies (0.1–1.0 Hz), there are minimal differences observed in the average spectral width, and frequency spectra, between clusters either side of the eruption (clusters C2 and C3 in Figures 3.3f and 3.3g). This supports the conclusion that differences between these clusters reflect a change in the medium, consistent with the findings of Obermann et al. (2013), rather than any change in seismic source.

It is of interest that, at higher frequencies (1.0–2.0 Hz), the spectra for the averaged cross-correlation function post-October eruption (cluster C2) contains less energy above approximately 1.7 Hz than those of pre-eruptive clusters (Fig. 3.6g). It is possible that this relates to changes in attenuation properties associated with

volcanic activity (e.g. Fehler et al., 1988; Domínguez et al., 2003). However, it is unclear if this would be expected given that both Fehler et al. (1988) and Domínguez et al. (2003) observed increased values of attenuation before eruptions rather than after an eruption. Furthermore, the cluster observed in the weeks prior to the eruption (C6 in Fig. 3.6), and also around the time of the January 2010 eruption (C6 in Fig. 3.6), does not show similar reduced energy above 1.7 Hz (Fig. 3.6g). Further analysis would therefore be necessary to identify the cause of reduced energy at higher frequencies during this time period. We do not believe however that this is solely responsible for the difference in pre- and post-eruptive clusters given that this separation occurs in both frequency bands analysed.

Seismic tremor at Piton de la Fournaise has a strong influence on the character of the cross-correlation functions. This is most evident at higher frequencies (1.0–2.0 Hz), where significant differences are observed in the correlogram during tremor episodes in January and October 2010 (e.g. Fig. 3.6b). This is further reflected in the reduced correlation coefficient between these time periods and the chosen reference period (Fig. 3.6d). These changes are also identified in the clustering output, where cross-correlation functions computed during the tremor form distinct clusters (Fig. 3.6c). At lower frequencies (0.1-1.0 Hz), we also find that cross-correlations computed during volcanic activity in October 2010 form a distinct cluster (Fig. 3.3c). Reduced values of the spectral width during this time period (Fig. 3.3f), in addition to increased energy above approximately 0.8 Hz in the average cross-correlation function (Fig. 3.3g), suggest that tremor is also responsible for producing this cluster.

Interestingly, lower values of the correlation coefficient during eruptive activity, between individual cross-correlation functions and the reference period in the same frequency band, were interpreted by Obermann et al. (2013) to be due mainly to physical changes of the volcanic edifice. This is based on the suggestion that the dominant frequencies of tremor during eruptions at Piton de la Fournaise are between 2–4 Hz. We argue that the decorrelation we observe in our re-computed velocity changes during eruptive activity (Fig. 3.3c) are related to volcanic tremor onset, which can be seen at frequencies below 1 Hz in our measurements of the covariance matrix spectral width (Fig 3.3a). Further, we might expect that if this cluster represented only a structural change, it would extend beyond the termination of tremor. We do note, however, that Obermann et al. (2013) recorded only very minor decreases in the correlation coefficient at station-pair UV01-UV07 during the October 2010 eruption, whereas we record a more significant decrease (Fig. 3.4c). Similarly, the magnitude of the correlation coefficient decreases recorded at the other two station-pairs by Obermann et al. (2013) is lower than those we record here. It is possible that this reflects differences in processing. For example, Obermann et al. (2013) removed data segments that contained earthquakes prior to computing cross-correlations. Thus we cannot rule out that differences in processing have reduced the influence of tremor in their original findings.

The seismic wavefield recorded at Ruapehu is more complicated relative to Piton de la Fournaise, where it is common to have long-lasting volcanic tremor episodes during inter-eruptive periods. This is most dominant at approximately 2 Hz (between 1.8-2.3 Hz; (Latter and Balsillie, 1984; Hurst and Steven, 1993; Hurst, 1998;

Sherburn et al., 1999)), and can be seen in the covariance matrix spectral width computed for the network (Fig. 3.10a). Tremor may also be a useful seismic source for computing velocity changes if it can be considered stable enough to produce highly similar cross-correlation functions. Thus, it is of interest to assess the similarity of cross-correlation functions computed within the frequency band of volcanic tremor, where assessment through visual inspection of the spectral width or correlogram is challenging.

From clustering, we identify some consistency to cross-correlation functions computed at higher frequencies (1.8–4.0 Hz). At a distance threshold of 0.4, the majority of cross-correlation functions are grouped within two clusters (C21 and C22 in Figure 3.10). Both clusters are similar in their spectral content, with a peak at approximately 2.2 Hz from the average cross-correlation function of each (Fig. 3.10g). We do note that values of spectral width for cluster C21 are lower than cluster C22 across almost all frequencies, suggesting a more dominant seismic source during the time period of the former (Fig. 3.10f). This is also reflected in the larger scale factor for cluster C21 after normalizing the frequency spectra (Fig. 3.10g). However, we consider that strong consistencies in both the pattern of spectral width measurements and also in the frequency spectra for the two clusters reflect similarities in the source process. Furthermore, we generally record high values of coherence against lag time between the two clusters (green-line in Figure 3.10i). Thus, our results suggest it might be possible to use the tremor at Ruapehu as the dominant seismic source during these time periods. Such work is part of an ongoing larger study to compute velocity changes at explosive volcanoes (Yates et al., 2021).

At lower frequencies (0.2–0.7 Hz) at Ruapehu, cross-correlation functions computed during the three year time period do not show any strong differences in the correlogram. This is reflected when we examine the two largest clusters produced using a distance threshold of 0.25 (Fig. 3.9). Both the mean spectral width during their respective time periods, and the frequency spectra of the averaged cross-correlation functions, are similar for both clusters; though, with some minor differences between approximately 0.15–0.3 Hz (Fig. 3.9f and 3.9g). This could represent slight variation in the oceanic microseism source. Values of coherence close to 1 at most lag times suggest a strong similarity between the two groups of cross-correlation functions (Fig. 3.9i). Furthermore, reducing the distance threshold further would split the data into a much greater number of smaller clusters (Fig. 3.9e). Therefore, reducing the threshold would likely not reveal further significant differences between the clusters. Thus, low-frequency cross-correlation functions at Ruapehu can be considered an example of a dataset containing no obvious internal structure..

Finally, we also observe periods at Ruapehu, especially at higher frequencies, where cross-correlation functions are grouped into many small clusters at the defined distance thresholds (greyscale in Figures 3.9 and 3.10). For example, the cross-correlation functions in the months preceding the 2007 eruption are split into many smaller clusters (Fig. 3.10c). These are linked to time periods where there is less volcanic tremor, suggested by higher spectral width values (Fig 3.10a). This is reflected by reduced amplitudes in the cross-correlation functions and, seemingly, less coherent energy at later lag times (Fig. 3.10b). This suggests that the appearance of these smaller clusters reflects lower SNR in the cross-correlation functions. We

also show this using synthetic cross-correlation functions, where the effect of reducing the SNR of cross-correlation functions is to increase the distance at which they are joined (e.g. Fig. 3.16). Thus, a higher distance threshold is needed to group the cross-correlation functions together. Knowing this, the dendrogram can provide information on the SNR of cross-correlation functions at different time periods by examining the threshold by which individual cross-correlation functions are grouped.

3.5.3 Additional insights

A key advantage of clustering cross-correlation functions is our ability to quantify the similarity between cross-correlation functions without using a reference period. While it is certainly encouraged that monitoring studies experiment with multiple reference periods, it may be cumbersome to obtain the same information obtained from clustering if the internal structure of the data is complex. Similarly, the choice of a reference period used to compare with other time periods is often a human one. Thus, subtle features of a dataset can be easily missed if not all options are explored.

At Piton de la Fournaise, we observe interesting similarities between cross-correlation functions at different points in time at both low and high frequencies. At low frequencies, we see that cross-correlation functions in 2009 and early-2010 are most similar to those following the October 2010 eruption (Fig. 3.11a). This feature is observed across all three station-pairs examined, and also in cross-component single-station results (Fig. 3.5 & 3.8). At higher frequencies, the same periods form distinct clusters at the chosen threshold at station-pair UV05-UV12; though, we see from the dendogram that they are still closer in similarity to each other than any other time period (Fig. 3.6e). Further, we also see that cross-correlation functions in the weeks before the October eruption are grouped with those in late-2009/early-2010. We consider these observations could reflect similarities in the state of the subsurface during these time periods. For example, two flank eruptions occurred during a short period of time in late-2009. One of these is included in our dataset (2009-12-14) and the other just before (2009-11-05). While these eruptions were both smaller than the October 2010 eruption, it could be expected that the subsurface was also affected following these events. Thus, similarities in cross-correlation functions between these two time periods could indicate that the same, or similar, subsurface process is responsible for these perturbations relative to the inter-eruptive time period. Moreover, we note a possible 'healing' process at low frequencies with cross-correlation functions being grouped once-again with the inter-eruptive cluster by mid-2011 (Fig. 3.11a).

We note the situation is slightly different for UV01-UV07 and UV07-UV09 station-pairs at higher frequencies. Unlike at lower frequencies, the cross-correlation functions pre-2010 are closer to the inter-eruptive period than post-eruption October 2010 (Fig. 3.11b). This could be related to their greater distance from the eruptive vents compared to UV05-UV12 station-pair (Fig 3.1a). Specifically, there may be differences in the behaviour of shallow (sampled by high frequencies) and deeper portions (sampled by lower frequencies) at greater distances from the eruption sites. We do not explore this further however, where we also note that similar clusters to those identified using station-pair UV05-UV12 can be found by reducing the dis-

tance threshold (Fig. S3.4). Thus, it is beneficial to examine the clusters identified at different distance thresholds to develop a more complete picture of a dataset's internal structure.

At Ruapehu, two smaller clusters at higher frequencies are of particular interest. These are a cluster pre-eruption in 2006 and a cluster post-eruption in 2007, each containing approximately three months of data (Figure 3.11d). Both periods are associated with decreases in the correlation coefficient relative to the reference (whole) time period (Fig. 3.10d). It is from clustering that we observe that the cross-correlation functions computed within these smaller clusters are themselves relatively similar. From the dendrogram (Fig. 3.10e), we observe that cross-correlation functions are grouped below a distance of approximately 0.25 for the 2006 cluster (C20) and 0.35 for the 2007 cluster (C6). The 2007 cluster is the most distinct of the two, only merging with the two larger clusters at a threshold of approximately 0.7. Comparatively, the pre-eruption cluster in 2006 merges at a threshold just about 0.5. Focusing on the time period post-eruption in 2007, the average spectral width for this cluster is reduced relative to other time periods, though with some consistency in shape. For example, we still observe a minimum at approximately 2.2 Hz (Fig. 3.10f). This is, however, shifted slightly to lower frequencies relative to the other clusters. Similarly, the relative difference between spectral width values at approximately 1.7 Hz and 1.9 Hz is smaller for the post-2007 eruption relative to other clusters (where the spectral width low at 1.7 Hz is less pronounced). This appears to be reflected in the frequency content of the average cross-correlation functions, with a broader energy distribution below approximately 2.3 Hz in clusters C6 and C20 relative to the two larger clusters (Fig. 3.10g).

At Piton de la Fournaise, the coherence between the average cross-correlation functions at early lag times appears useful towards distinguishing whether clusters represent a change in the medium or seismic source. Specifically, when comparing the co-eruptive cluster at Piton de la Fournaise (C3 in Fig. 3.6) with the inter-eruptive cluster (C5), we see that the coherence is reduced at early lag times where ballistic arrivals would be expected (Fig. 3.6i). In contrast, high levels of coherence are maintained between inter-eruptive clusters (e.g. C2 and C5), where we assume a change in the medium dominates. At Ruapehu, we observe that the coherence remains elevated at early lag times when we compare the two smaller clusters with larger clusters (Fig. 3.10i), suggesting these clusters may reflect a medium rather than source. However, unlike at Piton de la Fournaise where we see similarities in both frequency bands, we do not identify a similar cluster pattern at lower frequencies at Ruapehu. This includes tests of alternative frequencies just below the 1.8–4.0 Hz band used here (Fig S3.3). If these clusters reflected a medium change we might expect to see evidence of this at lower frequencies. This follows the expectation that there will be some overlap in depth sensitivity, with seismic energy observed to propagate predominantly as Rayleigh waves at larger distances from the crater (Hurst and Steven, 1993). The absence of these clusters outside of this frequency band suggests the clusters may instead then reflect a change in the character of the tremor source during these periods. If so, the absence of reduced coherence at early lag times may reflect the fact that instead of a new seismic source, as is the case at Piton de la Fournaise with the onset of tremor, we have instead minor variation in an existing seismic source.

3.5.4 Application in monitoring and future direction

Here, we consider how clustering cross-correlation functions can support monitoring applications and also future directions towards building on the results presented in this study.

The inherent structure of a seismic dataset can have a strong influence on how it can be used in monitoring. Temporal variation in cross-correlation functions could be related to changes in seismic sources, medium properties, and instrumental issues. All of these can influence how one might wish to process a particular dataset. For understanding changes in the seismic source, examining the spectral content of the seismic wavefield can provide clues as to when we might expect seismic source variability. However, it is not always clear how source variability will be reflected in scattered arrivals contained in the coda of cross-correlation functions post-processing. Thus, clustering cross-correlation functions allows us to study the data we intend to use directly. Further, it can reveal opportunities to process data that might otherwise have been considered unwanted 'noise'. At Ruapehu, clustering cross-correlation functions in the frequency band of seismic tremor suggests they may be stable enough to compute seismic velocity changes.

Clustering could also be used to target repeating seismic sources that are short relative to the full time period. At Ruapehu, we were able to use the entire data period (2005–2008) as a reference function to study velocity changes. This is possible because the seismic tremor is present during most of this period (Fig. 3.10a) and therefore dominates the reference function. However, a less frequent seismic source would likely require careful selection of similar cross-correlation functions to form an appropriate reference. In such cases, clustering could be useful to both identify and process unconventional repeating seismic sources.

A natural point in the monitoring process to apply clustering of cross-correlation functions would be before computing seismic velocity changes. This way, the results can also be used in choosing an algorithm towards computing the velocity changes. For example, if using the stretching (Sens-Schönfelder and Wegler, 2006) or moving-window cross-spectral (Poupinet et al., 1984; Clarke et al., 2011) technique, the choice of a reference period can significantly influence the measurement quality. Thus, studying the structure of a dataset can aid in the assessment of a suitable reference period (or periods). This follows also that through clustering, it is possible to identify not only changes associated with a seismic source change but also medium changes.

Alternatively, the results of clustering may facilitate the application of a no-reference approach (e.g. Brenguier et al., 2014; Gómez-García et al., 2018; Machacca-Puma et al., 2019). For this, apparent velocity changes are computed between individual 1-day cross-correlation functions and an inverse problem is solved to recover an apparent velocity change for each day. As part of this, a correlation coefficient matrix is computed from individual cross-correlation functions — similar to the distance matrices computed here — and used in weighting the data during the inversion. The main advantage of the non-reference approach is to better handle more significant changes in the cross-correlation function through time. This could be useful towards,

for example, processing higher frequencies at Piton de la Fournaise, where significant decreases in the correlation coefficient are observed post-October 2010 eruption. In this case, a pre-eruption reference period was not suitable to recover velocity changes post-eruption, with differences between the two time periods clearly reflected by their respective clusters only merging at a distance threshold of approximately 0.8 (Fig. 3.6e). Thus, we feel there are benefits in clustering cross-correlation towards determining a suitable approach to compute seismic velocity changes.

It may also be beneficial, as with other similarity measures (e.g. Obermann et al., 2013; D'Hour et al., 2016; Sánchez-Pastor et al., 2018), to cluster cross-correlation functions as a means to quantify and interpret temporal changes (with or without computing velocity changes). While we do not explore in detail the physical interpretations behind patterns of clusters — for example those discussed in Section 3.5.3 — doing so could enhance our understanding of subsurface processes. Moreover, if different clusters could be related to different sources, or sub-surface processes, it may be possible to apply partitional clustering techniques to classify new data in real-time. In this regard, it should be acknowledged that the approach applied here can be improved upon. Our application of hierarchical clustering in the time domain with a Pearson distance measure is relatively straightforward, and computationally inexpensive, allowing for simple integration into the monitoring workflow. One obvious improvement would be to cluster 1-day cross-correlation functions, where we stack to 10 days due to notable improvements in the quality of our cluster results (Fig. S3.5 and S3.6) at the cost of some temporal resolution. Further, objective criteria exist for defining a distance threshold from the dendrogram; for example, silhouette analysis to determine the optimal number of clusters (Rousseeuw, 1987). Alternatively, a Dynamic Tree Cut algorithm (Langfelder et al., 2008) can be applied to detect clusters automatically instead of using a static threshold. Future studies may, therefore, wish to experiment further with alternative clustering procedures, and similarity measures, to study waveform changes more precisely.

Finally, our approach to clustering has focused on individual stations or station-pairs. At Piton de la Fournaise, we examine the vertical-component cross-correlation function from three station-pairs — following those examined in greater detail by Obermann et al. (2013) — and also single-station cross-correlation function at two stations (Supplementary S3.1). At Ruapehu, we use a single station-pair, consisting of broadband sensors. While this approach can reveal important features of a dataset, it could be cumbersome to perform over an entire network. Such a task might be necessary if seeking to automate parts of the processing using cluster solutions. For example, the choice of reference stack for each pair or cleaning/exclusion of data during certain time periods. A network-based solution may therefore be useful in order to efficiently extract information from a larger network. One possibility could be to cluster the first eigenvector of the network covariance matrix for each day (e.g. Soubestre et al., 2018). This represents the principle component of the seismic wavefield and characterises the most dominant seismic source, with Soubestre et al. (2018) using a correlation coefficient-based similarity measure to cluster and classify different seismic tremor sources using this approach. Applying hierarchical clustering to the first eigenvector could therefore be useful to detect subtle variation in the seismic source, where hierarchical clustering has been shown to be effective in combination with principal component analysis in clustering seis-

mic spectra (Unglert et al., 2016b). A network-based approach might also aid in distinguishing between source and medium changes if the latter is not expected to be present over the whole network. Thus, we consider that future developments to cluster cross-correlation functions using network-based approaches could greatly enhance our ability to identify structure in seismic datasets and support the processing decisions towards applying seismic interferometry over a network.

3.6 Conclusions

We apply hierarchical clustering to cross-correlation functions computed using data recorded by two seismic networks. At Piton de la Fournaise, we resolve distinct clusters related to both changes in the seismic source — associated with co-eruptive volcanic tremor — and also changes in the medium following the October 2010 eruption. This is observed using two frequency bands (0.1–1.0 Hz and 1.0–2.0 Hz) from both station-pair and single-station cross-correlation functions. At Ruapehu, we find that cross-correlation functions computed within a frequency band containing volcanic tremor (1.8–4.0 Hz) are similar in character, indicating some stability in the seismic source and a possible use in monitoring. These results highlight how clustering cross-correlation functions can be useful as a tool to identify structure in seismic interferometry datasets.

Clustering cross-correlation functions can also reveal features that may be overlooked using similarity measures that depend on a reference period. Through hierarchical clustering, we can efficiently assess the similarity between cross-correlation functions computed at any time period. At Piton de la Fournaise, for example, we observe interesting similarities between low-frequency cross-correlation functions separated by an inter-eruptive cluster. This may indicate a consistency in subsurface processes responsible. Furthermore, we anticipate that clustering could be useful in identifying, and extracting, similar cross-correlation functions that are not well-represented within a reference period. At Ruapehu, we observe two smaller clusters that, while associated with reduced correlation coefficients, are found contain relatively stable cross-correlation functions within themselves. Thus, we anticipate clustering could also be useful to target and process cross-correlation functions produced by less frequent seismic sources that could otherwise be otherlooked as unwanted 'noise'.

From these findings, we consider that clustering cross-correlation functions can be a useful inclusion in seismic interferometry studies. Identifying structure in datasets can aid in both decision making in processing and the interpretation of observed subsurface changes. Our application of hierarchical clustering in the time domain using a Pearson distance measure is both relatively straightforward and computationally efficient, allowing for easy integration into the monitoring workflow. Going forward, future studies may wish to build on these results by experimenting with alternative clustering algorithms, including network-based approaches, towards exploring further the potential of clustering cross-correlation functions as a similarity measure.

3.7 Acknowledgements

The python package MSNoise (Lecocq et al., 2014) was used to process cross-correlation functions and compute apparent velocity changes, and the python package Covseisnet (Covseisnet, 2020) used to compute the covariance matrix spectral width. The python package Scipy (Virtanen et al. 2020) was used to apply hierarchical clustering (scipy.cluster.hierarchy.linkage) and Matplotlib used in plotting (Hunter, 2007). A. Yates is supported by funding from the Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation (MESRI). We also thank Virginie Pinel for reviewing and providing suggestions on an earlier version of the manuscript and Calum Chamberlain for helpful discussion during early stage of the project. Finally, we thank Editor Dr Cornelis Weemstra, Assistant Editor Louise Alexander, and the two reviewers Dr. Antonino D'Alessandro and Dr. Pilar Sánchez-Pastor, for their comments towards to an improved manuscript.

3.8 Data Availability

This work was possible through many publicly available services. Data for the UnderVolc network at Piton de la Fournaise was downloaded through the French Seismologic and Geodetic Network data center (RESIF) (Institut De Physique Du Globe De Paris (IPGP), 2021). Data from the New Zealand network at Ruapehu was downloaded through the data center GeoNet (https://www.geonet.org.nz/). A Jupyter Notebook containing codes used for applying hierarchical clustering to cross-correlation functions (output by MSNoise) can be downloaded from https://github.com/asyates/hclusterCC

3.9 Supplementary Material

S3.1 Single-station clustering

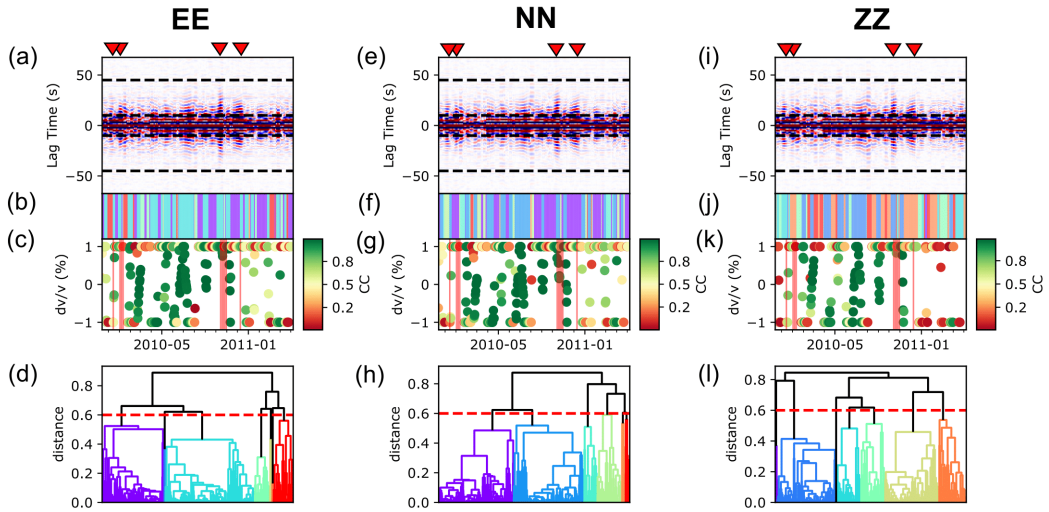

Figure S3.1: Clustering single-station auto-correlation functions computed using station UV07 at Piton de la Fournaise (0.1–1.0 Hz). The different component combinations are EE (a-d), NN (e-h), and ZZ (i-l). (a/e/i): Correlogram of cross-correlation functions. (b/f/j): Location of clusters in time, color-coded according to dendrogram outputs. (c/g/k): Apparent velocity changes, color-coded according to correlation coefficient (CC) computed between 10-day stacks and the reference period. (d/h/l) Dendrogram, with clusters defined at a distance threshold of 0.6.

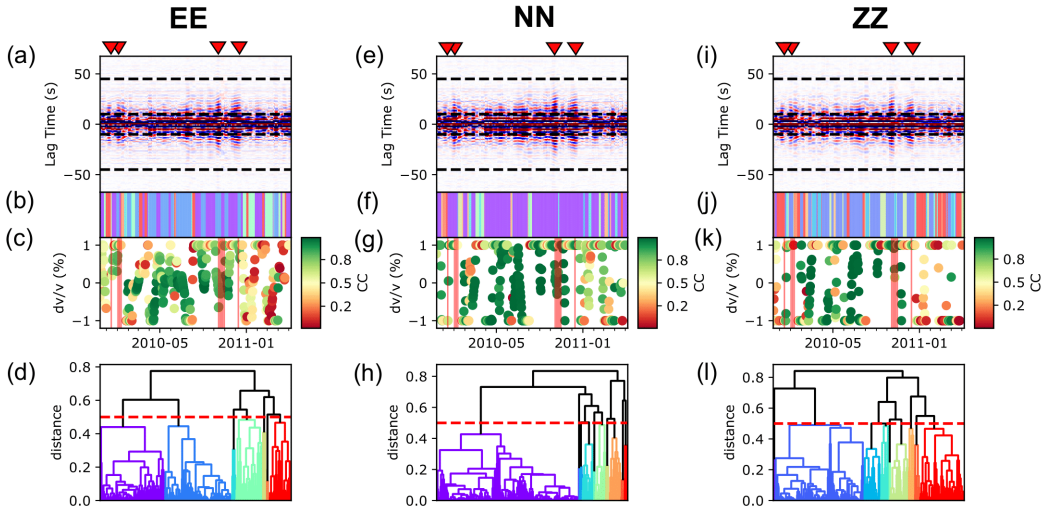

Figure S3.2: Clustering single-station auto-correlation functions computed using station UV12 at Piton de la Fournaise (0.1–1.0 Hz). Subplots follow those described in Figure S3.1, though with distance threshold set to 0.5.

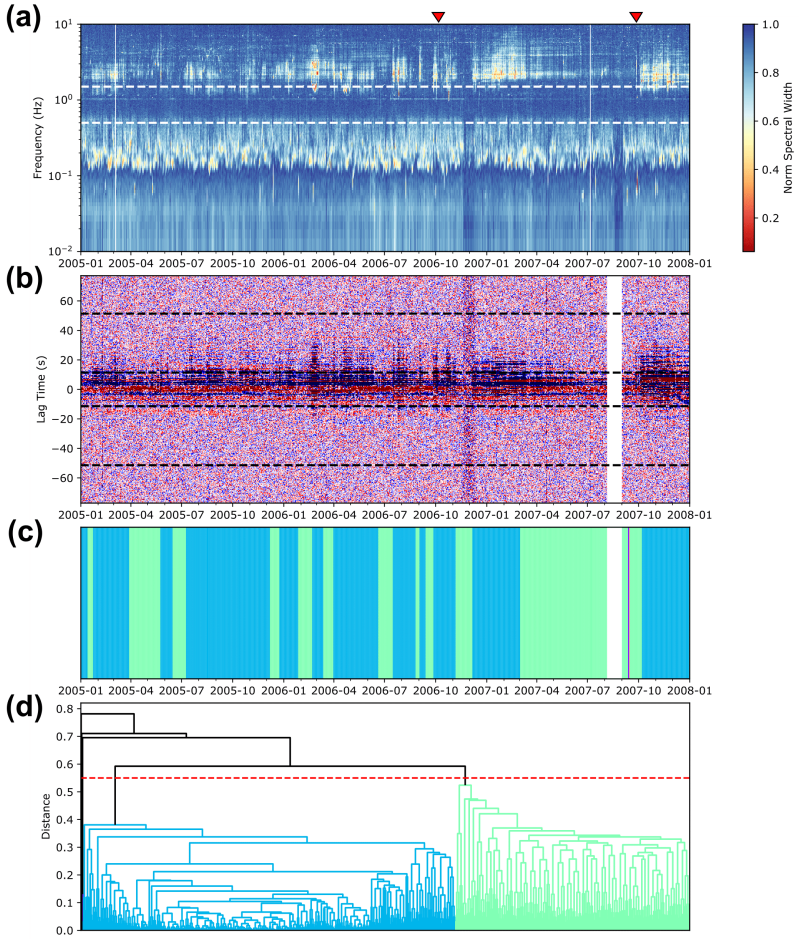

Figure S3.3: Clustering 0.5–1.5 Hz cross-correlation functions computed for station pair FWVZ-WPVZ at Ruapehu. (a): Normalized spectral width measurement. Lower values indicate a more coherent seismic wavefield dominated by fewer seismic sources. Dashed white lines show frequency range of CCFs. (b): Correlogram showing amplitudes of CCFs at different lag times (red = positive, blue = negative). Dashed black lines show part of CCFs used in clustering. (c): Location of clusters in time, color-coded according to dendrogram output. (d): Dendrogram, with clusters defined at a distance threshold of 0.55 (dashed-red line).

S3.2 Lower-distance-threshold Piton de la Fournaise (1.0–2.0 Hz)

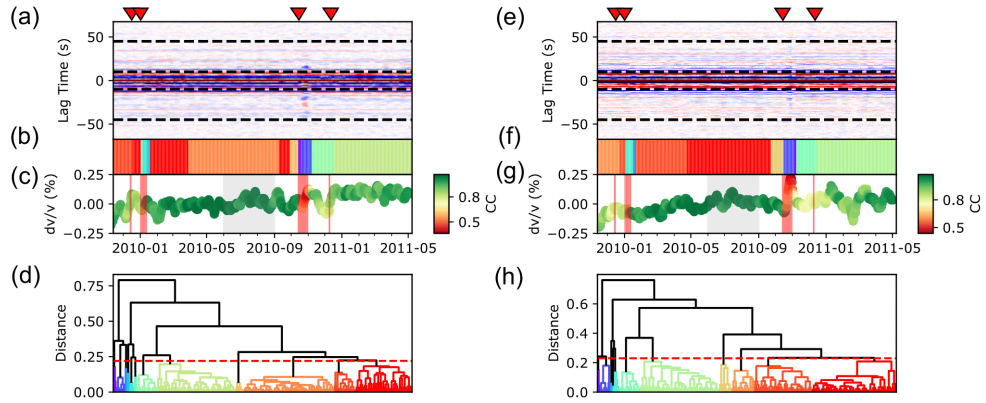

Figure S3.4: Clustering output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pairs UV01-UV07 (a-d) and UV07-UV09 (e-h), with lower distance thresholds defined relative to those in Figure 3.7. (a/e): Correlogram (b/f): Location of clusters in time (c/g): Apparent velocity changes with correlation coefficient (CC) relative to reference stack period (shaded gray). (d/h): Dendrogram.

S3.2.1 Clustering 1-day cross-correlation functions

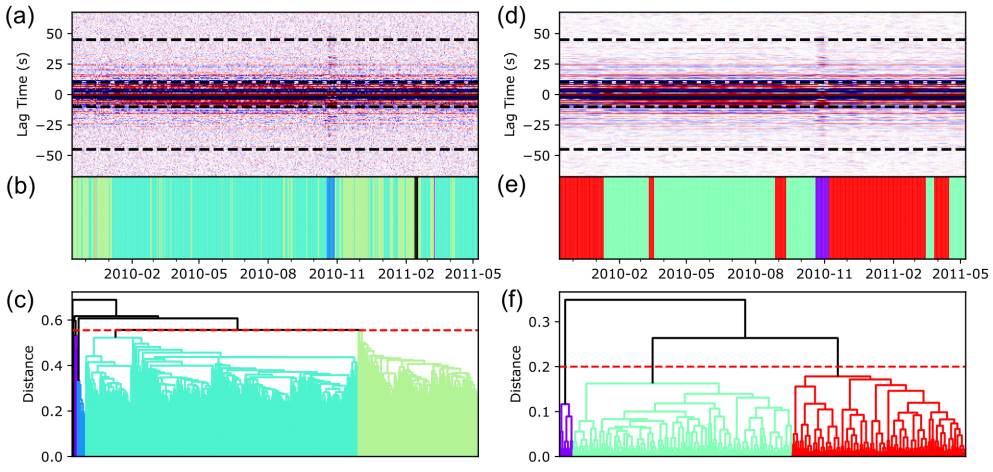

Figure S3.5: Comparison between using 1-day (a–c) and 10-day (d–f) stacks for clustering 0.1–1.0 Hz cross-correlation functions computed from Piton de la Fournaise seismic data, station-pair UV05-UV12. (a/d): Correlogram of cross-correlation functions. (b/e): Location of clusters in time, color-coded according to dendrogram outputs. (c/f) Dendrogram, with clusters defined at a distance threshold of 0.555 for 1-day stacks and 0.2 for 10-day stacks.

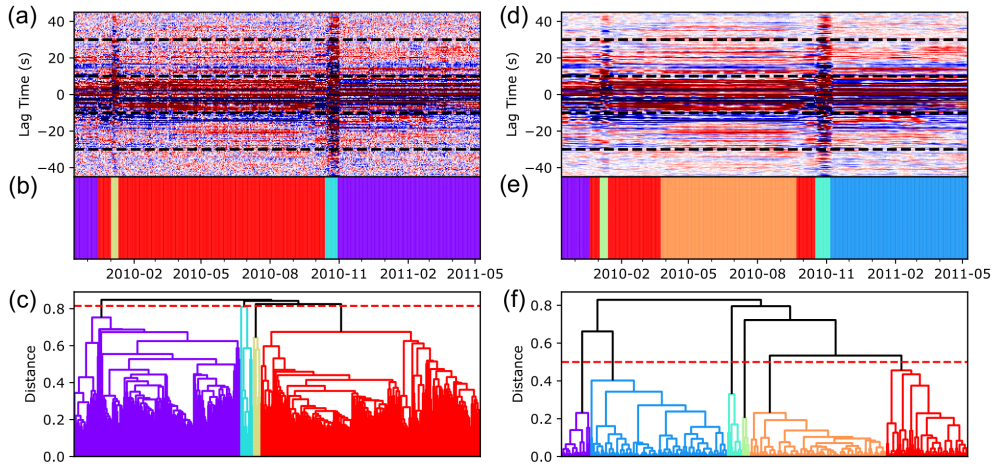

Figure S3.6: Comparison between using 1-day (a–c) and 10-day (d–f) stacks for clustering 1.0–2.0 Hz cross-correlation functions computed from Piton de la Fournaise seismic data, station-pair UV05-UV12. Subplots follow those described in Figure S3.5.

Chapter 4

Seismic velocity changes over nine years at Stromboli volcano, Italy

4.1 Introduction

Open-vent basaltic volcanoes are characterized by persistent explosive activity and passive degassing (Rose et al., 2013; Edmonds et al., 2022). Such is the case at Stromboli volcano, Italy, where regular activity consists of 5–20 mild, discrete, explosions, per hour (Harris and Ripepe, 2007; Ripepe et al., 2008). Occasionally, however, mild activity is interrupted by more significant events such as effusive eruptions, major explosions, and large violent explosions referred to as paroxysms (Métrich et al., 2005; Allard, 2010; Rosi et al., 2013; Rizzo et al., 2015; Voloschina et al., 2023). Such transitions in activity can occur with little to no warning, and thus represent a considerable hazard to inhabitants and visitors (Rosi et al., 2013). Thus, techniques that can provide an indication of evolving behaviour at open-vent volcanoes, and the transition to periods of heightened activity, are highly sought.

Stromboli is a small volcanic island located in the Southern Tyrrhenian Sea, within the northernmost part of the Aeolian archipelago. Its formation, as with other Aeolian Islands, is related to the subduction of African plate below the Eurasian plate (Barberi et al., 1974). Only part of the volcano is above sea level, with a maximum elevation of 924 m (Bosman et al., 2009). The submerged part of the volcano reaches approximately 2300 m below sea level (Bosman et al., 2009). Current activity originates from three active craters (South-west, Central, and North-East) within the crater terrace (Fig. 4.1) of the volcano (Salvatore et al., 2018).

Regular (mild) strombolian activity at Stromboli is fed by crystal-rich, highly porphyritic (HP), basaltic magma residing at 2–4 km depth (Métrich et al., 2001; Landi et al., 2004; Métrich et al., 2010). More explosive events, such as paroxysms and major explosions, release crystal-poor, low porphyritic (LP) basaltic magma originating from 7- 10 km depth (Métrich et al., 2001; Francalanci et al., 2004; Métrich et al., 2010; Andronico et al., 2021). High CO_2/SO_2 ratios during strombolian explosions point towards a contribution of deep-derived gas from the LP storage zone during regular activity (Burton et al., 2007; Aiuppa et al., 2010, 2021). Thus it is consid-

ered that CO_2-rich gas bubbles are persistently supplied to the shallow plumbing system, which mix upon ascent with gases derived from the HP reservoir to drive strombolian activity (Aiuppa et al., 2010). For paroxysms, the trigger mechanism remains poorly understood, where various models have been considered (Aiuppa et al., 2021; Voloschina et al., 2023). Two commonly considered models are pressure build up due to (1) the intrusion of new volatile-rich magma into the deeper system (e.g. Métrich et al., 2001, 2005) and (2) the accumulation of a foam layer at the top of the LP reservoir due to CO_2 degassing, followed by rapid rise of CO_2 rich gas to the surface (e.g. Allard, 2010; Aiuppa et al., 2011, 2021). Note that these models are not neccessarily mutually exclusive, rather the question may be a question of which process is the dominant trigger force (Voloschina et al., 2023). In any case, there is a general agreement that they result from the fast ascent and injection of deeply stored volatile fluids into the shallow plumbing system (Métrich et al., 2021).

More recent activity at the volcano is marked by two powerful paroxysmal eruptions occurring on 3 July and 28 August 2019 (Giudicepietro et al., 2019; Inguaggiato et al., 2019; Giudicepietro et al., 2020; Inguaggiato et al., 2020; Calvari et al., 2021). Both events produced eruptive columns more than 3 km high, followed by a pyroclastic flow that expanded along the Sciara del Fuoco slope (Giudicepietro et al., 2020; Inguaggiato et al., 2020). The time between these events was characterized by intense volcanic activity consisting of frequent strombolian explosions, spattering, lava overflows, and major explosions (Inguaggiato et al., 2020). These events represent one of the most serious volcanic crisis at Stromboli in the last decades (Giudicepietro et al., 2020). Since then, short-term (minute-scale) ground tilt has been identified as a useful precursor towards providing real-time warning of impending paroxysms (Giudicepietro et al., 2020; Di Lieto et al., 2020; Ripepe et al., 2021). While short-term precursors can provide crucial warning time, identifying mid-to-long term precursors remains an important challenge towards minimizing the impact of paroxysms.

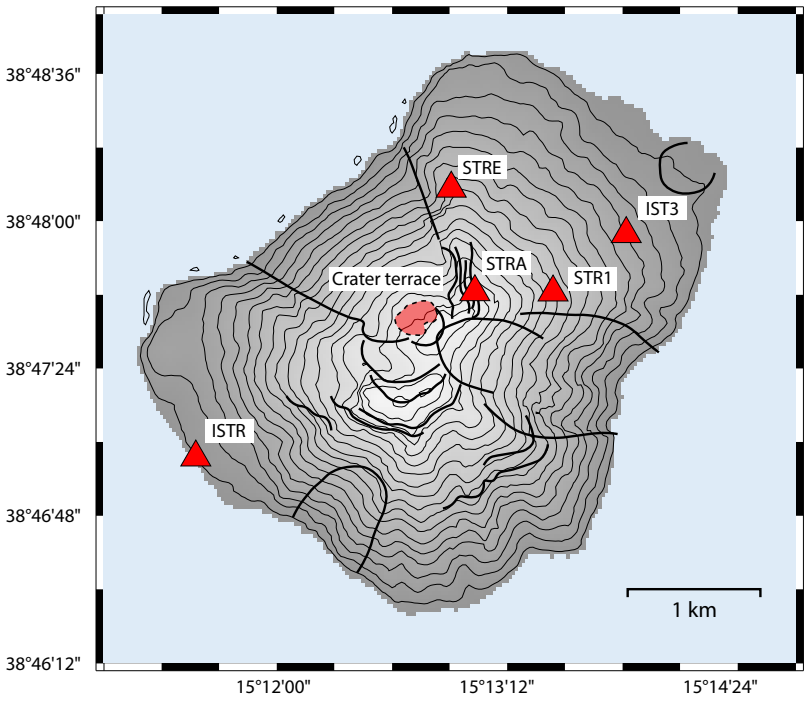

Figure 4.1: Map of stromboli volcano, with broadband seismic stations used in this study shown as red triangles. Crater terrace, containing the three active craters, shaded red. Thicker black lines represent structural features (such as crater rims, lateral collapse rims, and caldera rims, based on (Revil et al., 2023, and references therein)).

Coda wave interferometry using the seismic ambient noise wavefield is an increasingly popular technique used for monitoring volcanoes (e.g. Sens-Schönfelder and Wegler, 2006; Brenguier et al., 2008c; Mordret et al., 2010; Donaldson et al., 2017; Yates et al., 2019; Calò et al., 2021; Caudron et al., 2022). By cross-correlating continuously recorded seismic noise, it is possible to detect subtle changes in seismic velocity and structure of the medium through time. Such changes have been attributed to increased pressurization within the volcanic system (e.g. Brenguier et al., 2008c; Mordret et al., 2010; Budi-Santoso and Lesage, 2016a; Yates et al., 2019; Calò et al., 2021; Caudron et al., 2021) and changes in fluid content (e.g. Caudron et al., 2015, 2022). Furthermore, at Stromboli specifically, seismic velocity changes between 2011–2014 coincided with the timing of effusive activity (Calò et al., 2021). These changes were resolved to regions matching previously known hydrothermal reservoirs, suggested to reflect pressurization in the hydrothermal system. Thus the technique is promising towards identifying mid-to-long term changes in volcanic systems (typically from weeks to years).

The sensitivity of seismic velocities to external, non-volcanic, processes often presents a challenge for monitoring volcanic processes. For example, rainfall has been identified to cause velocity changes at a number of volcanoes (e.g. Sens-Schönfelder and Wegler, 2006; Rivet et al., 2015; Budi-Santoso and Lesage, 2016a). Changes due to large tectonic earthquakes have also been observed to cause significant velocity decreases in volcanic settings (e.g. Lesage et al., 2014; Brenguier et al., 2014; Yates

et al., 2019; Machacca-Puma et al., 2019). Outside of volcanic systems, changes related to temperature-induced thermoelasticity (e.g. Richter et al., 2014; Hillers et al., 2015a), atmospheric pressure loading (Silver et al., 2007; Niu et al., 2008), tidal modulation (Yamamura et al., 2003; Takano et al., 2014), and snow-loading (e.g. Hotovec-Ellis et al., 2014; Wang et al., 2017; Makus et al., 2023) have also been reported. While snow-loading is not relevant at Stromboli, the shear number of identified processes causing velocity changes highlights the sensitivity of the subsurface to external forcing. Crucially, these changes are often similar magnitude or larger than the volcanic processes of interest. Thus, one of the challenges is to carefully distinguish velocity changes related to volcanic processes from those related to external processes.

We compute seismic velocity changes between pairs of seismic stations and between different components of single-stations over a nine-year period from 2013 to 2022. This period encompasses a range of volcanic activity including lava flows (including one flank eruption), major explosions, and three paroxysms. The third paroxysm, on 19 July 2020, was much smaller than the 2019 paroxysms, falling close to the threshold between a major explosion and a paroxysm (Calvari et al., 2021). The timing of the various activity are shown in Figure 4.2 where we split the nine-years into five phases. These are: (1) the period prior to and including the 2014 flank eruption, (2) the period of low activity following the flank eruption, (3) the re-awakening phase with renewed heightened activity, (4) the period prior to and including the 2019 paroxysms, and (5) post-2019 paroxysms. Note, there is an argument to extend Phase 3 until the onset of the paroxysms (and have Phase 4 just include the two paroxysms). However, we separate the phases based on renewed activity in December 2018 through to January 2019 which marked the start of increasing CO_2 flux up until the July paroxysm (Inguaggiato et al., 2020; Aiuppa et al., 2021).

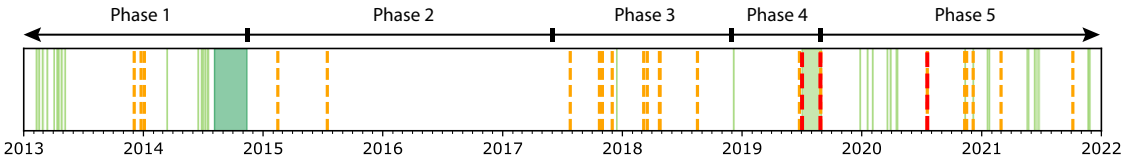

Figure 4.2: Activity at Stromboli over nine years, split into five phases described in text. The different types of activity are lava flows / overflows originating from the crater (green-shaded regions), 2014 flank eruptions (dark-green shaded region), major explosions (orange-dashed vertical line), paroxysms (red-dashed vertical line).

4.2 Methods

Data from five broadband seismic stations (Fig. 4.1) that form the permanent network at Stromboli, run by Istituto Nazionale di Geofisica e Vulcanologia, are acquired during the period 2013–2022. These consist of three seismometers equipped

with a Guralp CMG40T sensor (STRA, STR1, and STRE, 50 Hz sampling rate) and two with a Nanometrics Trillium 120 s sensor (ISTR and IST3, 100 Hz sampling rate). No instrument response correction is applied given all sensors have a flat response between frequencies of interest (0.5–4 Hz).

4.2.1 Seismic sources at Stromboli

Application of passive noise interferometry towards computing seismic velocity changes requires relatively stable seismic sources in time (Hadziioannou et al., 2009). Thus, it is useful to understand the seismic wavefield propagating coherently over the seismic network. We apply a network based approach introduced by Seydoux et al. (2016a) towards this goal. Following Seydoux et al. (2016a), we compute a spectral width value that reflects the eigenvalue distribution of the seismic network covariance matrix. This value tends to zero when the seismic wavefield is dominated by a single source, and approaches a maximum when non-propagating, or incoherent, noise is considered.

The open-source Python package Covseisnet (Covseisnet, 2020) is used to compute the network covariance matrix and the corresponding spectral width values at each frequency with time. First, the data for each station are sliced in 6-hour time slices and decimated to 25 Hz. Each time slice is then de-meaned and de-trended, and one-bit spectral and temporal normalization applied. The covariance matrix $C(f)$ is then estimated by averaging the Fourier cross-spectra of the five stations over M overlapping subwindows of length δt following:

$$C(f) = \frac{1}{M} \sum_{m=1}^{M} u_m(f) u_m^{\dagger}(f) \tag{4.1}$$

where $u(f)$ is a vector composed of the Fourier spectra of N seismic stations, taking the form $u(f) = [u_1(f), u_2(f), ..., u_N(f)]^T$, and $u_m(f)$ is the Fourier spectra within each subwindow m and $\dagger$ signifying the Hermitian transpose. For this study, we use a subwindow length of 100 s (δt), and average over 100 windows (M) with an overlap of 50%. The spectral width σ is then computed as a function of frequency following:

$$\sigma(f) = \frac{\sum_{i=1}^{N} (i-1)\lambda_i(f)}{\sum_{i=1}^{N} \lambda_i(f)} \tag{4.2}$$

where $\lambda_i(f)$ represents the eigenvalues arranged in decreasing order. Further details can be found in Seydoux et al. (2016a). We further apply a normalization to spectral width measurements by dividing by the maximum value recorded across all frequencies within each window. This reduces the fluctuations in time associated with varying station numbers (Fig. S4.1). Note, however, this is an imperfect solution given that the spectra width does not converge to the same value at all frequencies. However, for our purpose, it simplifies the visualization of the coherent seismic wavefield.

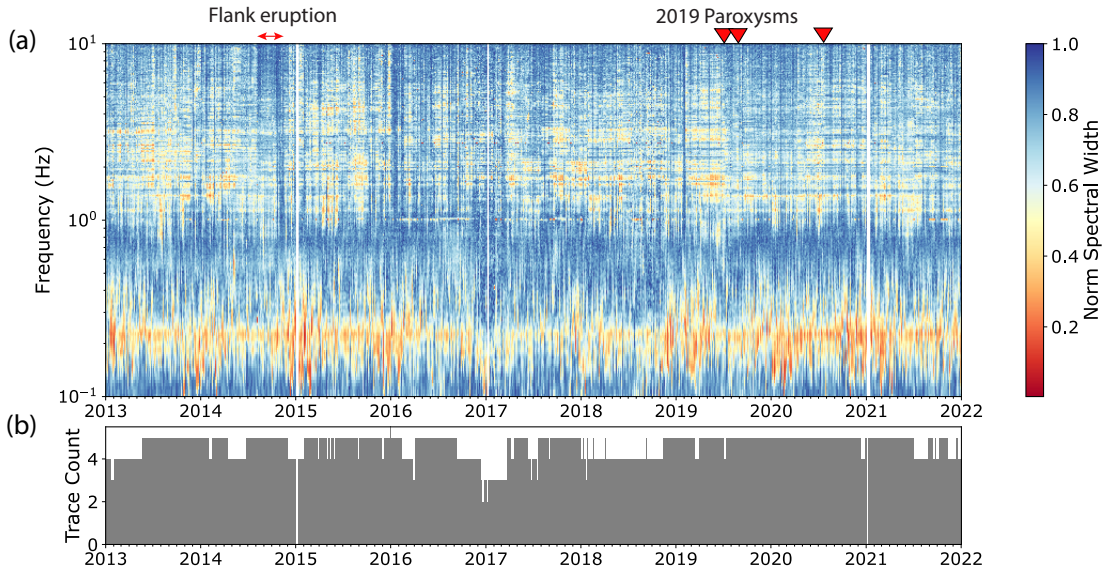

Figure 4.3: Covariance matrix spectral width, normalized within each measurement window by dividing by maximum value across different frequencies. Lower values indicate a more coherent seismic wavefield dominated by fewer seismic sources. Trace count represents the number of stations used in estimating the covariance matrix.

Volcano seismicity is a continuous feature of the seismic wavefield at Stromboli, due to the persistent volcanic activity (Langer and Falsaperla, 1996; Chouet et al., 1997; Wassermann, 1997; Ripepe and Gordeev, 1999). Two types of signals predominate, those related to sustained volcanic tremor and those related to discrete volcanic explosions (Langer and Falsaperla, 1996; Chouet et al., 1997). Low values of spectral width can be observed between approximately 1 – 6 Hz throughout the 9-year period (Fig. 4.3), consistent with previous observations of tremor spectral maxima at the volcano (Langer and Falsaperla, 1996). Thus, we assume that reduced spectral width values within this frequency range reflect sustained tremor. While this does appear approximately continuous, some variation is observed; for example, lower values of spectral width are observed between 1–2 Hz in the months prior to the first paroxysm in 2019. Similarly, tremor activity appears reduced in 2016, with higher values of spectral width recorded. These observations are consistent with previous observations that tremor amplitudes and dominant spectral peaks are variable in time at the volcano (e.g. Langer and Falsaperla, 1996, 2003).

Low values of spectral width are also observed below approximately 0.5 Hz throughout the data period. This is especially true at approximately 0.2 Hz where low values are observed almost continuously, most likely reflecting the secondary oceanic microseism. Some fluctuations are observed however, with short 'bursts' of lower spectral width values (lasting a few days) either side of 0.2 Hz (Fig. S4.2) This appears more apparent in later months of the year (and the first months of the new year). This is evidenced when computing the average spectral width across the different seasons, with lower values recorded on average in winter (Dec-Feb) between 0.1–0.5 Hz relative to other seasons (Fig. 4.4). Reduced values of spectral width in winter, and higher values in summer, likely reflect storm activity, with Mediterraneans thunderstorms occurring predominantly over the sea in autumn and winter (Galanaki et al.,

2018). This is consistent with the observation of a strong correlation between long period 'tremor' at approximately 0.5 Hz and strong weather events by Braun et al. (1996) previously at the volcano.

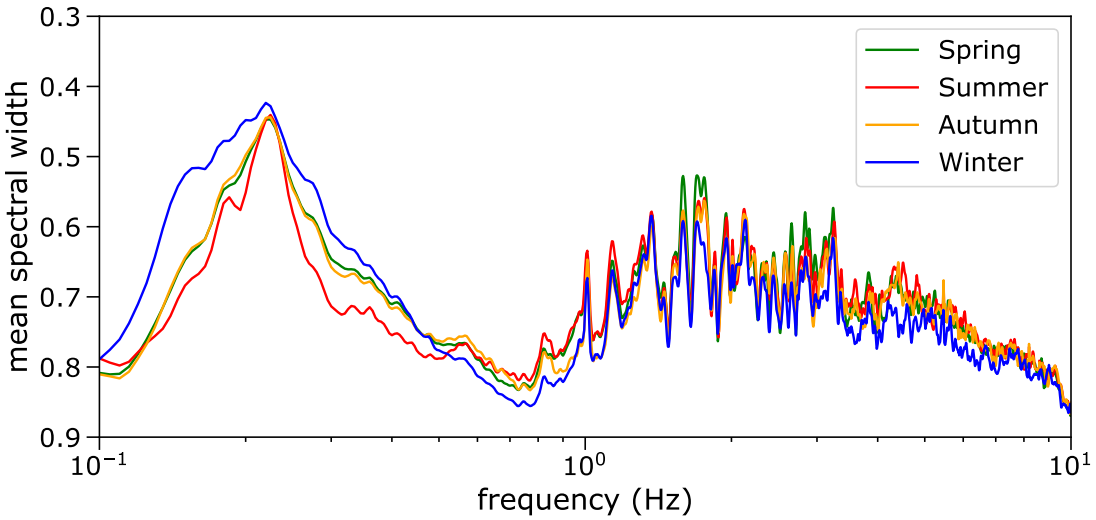

Figure 4.4: Average spectral width split into seasons: spring (Mar, Apr, May), summer (Jun, Jul, Aug), autumn (Sep, Oct, Nov), winter (Dec, Jan, Feb).

4.2.2 Computing cross-correlation functions and velocity changes

Using the python software package MSNoise (Lecocq et al., 2014), cross-correlation functions are prepared using relatively established data processing methods that will be outlined briefly in this section. Seismic data for each day are grouped into continuous chunks, demeaned and merged into 1-day traces. Following this, traces are bandpassed between 0.01 and 10.0 Hz before being decimated to 25 Hz. For each individual station, and their three components, we then apply spectral whitening and 1-bit normalization in 30-minute windows. In order to study velocity changes at different frequencies, we apply whitening between 0.5–1.0 Hz, 1.0–2.0 Hz, and 2.0-4.0 Hz, providing three distinct data sets. The 30-minute windows are then used to compute the cross-correlation functions between various receivers. In this study, we compute vertical-vertical (ZZ) cross-correlation functions between pairs of stations and cross-component correlations (east-north EN, east-vertical EZ, north-vertical NZ) between the components of a single-station. Finally, we perform linear stacking to give 5–day cross-correlation functions. Thus, the cross-correlation function recorded for a specific day represents the stack of that particular day and the preceding four days.

Velocity changes are computed using two commonly used approaches: the stretching technique (e.g. Sens-Schönfelder and Wegler, 2006) and the moving-window cross-spectral analysis technique (Poupinet et al., 1984; Clarke et al., 2011). For the stretching-technique, cross-correlation functions are stretched relative to a reference stack (stack of full time period here). We then seek the maximum value of the corre-

lation coefficient between each current stack and reference stack at different levels of stretching up to factor of ±1%. We do not observe any evidence to suggest velocity changes larger than this are present within our datasets. Once the maximum value is identified, the corresponding stretching coefficient is converted into an apparent velocity change and the correlation coefficient is recorded. For the moving-window cross-spectral analysis, where delay-times between each cross-correlation function and a reference function are computed in a series of moving windows. The final velocity ($\delta v/v$) is calculated as the slope of delay times ($\delta t/t$) as $\delta t/t = -\delta v/v$.

For 0.5–1.0 Hz cross-correlation functions, the coda window is defined between 10–40 seconds lag time. For 1.0–2.0 Hz and 2.0–4.0 Hz, the coda is defined between 5–25 seconds and 5–15 seconds respectively unless otherwise stated. When applying the moving-window cross-spectral technique, we use sliding windows of 8 s, 4 s, and 2 s for 0.5–1.0 Hz, 1.0–2.0 Hz, and 2.0–4.0 Hz respectively, with a 50% overlap. These lengths correspond to four cycles of the lowest frequency. Furthermore, we exclude delay time measurements if they are greater than 0.25, have an error greater than 0.1 s, or a coherence below 0.5. Similarly, for the stretching technique, results with a correlation coefficient less than 0.5 are excluded.

4.2.3 Hierarchical clustering of cross-correlation functions

The presence of sustained tremor at Stromboli (between approximately 1–6 Hz) may support the use of this frequency band to compute cross-correlation functions if the source is stable enough. Using a sliding moving window of ten broadband cross-correlation functions, we find that individual 1-day functions are highly similar above approximately 1 Hz (Section S4.1.1). This is shown through high values of SNR and phase similarity, the latter assessed as the amplitude of the phase stack (Schimmel and Paulssen, 1997). However, assessing the longer-term stability requires a different approach. One option would be to compare each individual cross-correlation function to a reference function, for example using the full data period. However, in cases where the internal structure of the dataset is complex, it may be difficult to ascertain the similarity between different time periods without testing many different reference periods.

Hierarchical clustering has recently been applied towards this purpose (Yates et al., 2023; Makus et al., 2023), where cross-correlation functions are grouped together based on their similarity within a defined coda window, independant of any reference period. Following the approach of Yates et al. (2023), we apply agglomerative hierarchical clustering using the average linkage algorithm, where each object (cross-correlation function) begins as a separate cluster before being sequentially merged (agglomerated) into a new, larger, cluster until all objects are contained within a single cluster (Aminzadeh and Chatterjee, 1984). For the average linkage method, the distance between each pair of objects is summed together and divided by the total number of pairs to give an average distance between clusters. In this instance, we apply a correlation coefficient based distance measure, where we subtract the absolute value of the correlation coefficient between two cross-correlation objects from 1 to give a measure of dissimilarity. This value varies from 0 (perfect correlation) to 1 (no correlation). The process can be visualised through a dendrogram, where the

height (y-axis) reflects the distance between two objects joined at various levels.

We apply the clustering algorithm to cross-correlation functions in the time-domain between the three frequency bands previously outlined (0.5–1.0 Hz, 1.0–2.0 Hz, 2.0–4.0 Hz). To improve the SNR of results 5-day functions are used, thus matching the stacking used for computing velocity changes. Correlation coefficients are computed within the lag times previously defined for the coda (10–40s, 5–25s, 5–15s respectively). This is done for the station-pair IST3–ISTR, where application to a single pair of stations is likely to capture the most significant features of our dataset with regards to variable seismic sources (Yates et al., 2023). In the case of IST3 and ISTR stations, both of these have relatively few gaps compared to stations located closer to the summit. Finally, for efficiency, we select cross-correlation functions with an interval of 5-days, equal to the stacking time. Thus, each cross-correlation function should be independent i.e. not computed using the same raw data.

4.2.4 Complementary datasets

Daily precipitation, temperature, and sea-level pressure are acquired as comparison datasets. Meteorological datasets for Stromboli itself were found to contain significant gaps that render them unsuitable for comparison with longer-term velocity changes. Thus, we rely on information that can be obtained from nearby stations. Precipitation data are obtained for Salina island, approximately 50 km from Stromboli. Temperature and sea-level pressure data are obtained for stations at Messina and Catanzaro respectively from the Global Surface Summary of the Day (GSOD) dataset (United States National Oceanic and Atmospheric Administration National Climatic Data Center, 2016). These stations show a strong similarity with values recorded at Stromboli during times this data are available (Fig. S4.3).

For the daily precipitation data, we perform an additional step to estimate changes in pore pressure related to fluid diffusion, following Wang et al. (2017). The one-dimensional complete diffusion equation established by Talwani et al. (2007) and later adopted by Rivet et al. (2015) underpins our computation of pore pressure variations ($P(r,t)$):

$$P(r,t) = \sum_{i=1}^{n} \delta p_i \operatorname{erfc}[r/(4c(n-i)\delta t)^{1/2}] \tag{4.3}$$

In this equation, r refers to the distance (or depth here), n the total number of time increments δt between the start the time series (i=1) and time t. δp_i signifies the change in the precipitation load at the sampled instant t_i. This is calculated as $\rho g \delta h_i$, where δh_i represents the variation in rainfall at the instance t_i from the previous day. The diffusion rate, denoted by c, is measured in m^2/s. For Stromboli, we compute the average pore pressure for r values ranging from 100 m to 1000 m (noting that deeper than this will be below sea-level). The optimal diffusion rate is estimated from the measured velocity changes, where we find the diffusivity that gives the maximum absolute value of the cross-correlation between velocity changes with the estimated pore pressures. This is discussed further in the Results section.

4.3 Results

4.3.1 Hierarchical clustering within tremor

The results of clustering cross-correlation functions are presented first for both 1–2 Hz (Fig. 4.5) and 2–4 Hz (Fig. 4.6) datasets. The purpose here is to assess how similar cross-correlation functions — produced in the frequency band of tremor — are over the nine-year period.

At 1–2 Hz (Fig. 4.5), it is possible to group the majority of data into three clusters (C1, C2 and C4 in Fig. 4.5). These are separated in time by the 2014 flank eruption and the first paroxysm in 2019, with both events marking the transition to a new cluster group (Fig. 4.5c). A distance threshold of approximately 0.6 would be required for these clusters to be agglomerated into a single cluster. Since the distance measure is based on 1 minus the correlation coefficient, this suggests the minimum correlation coefficient between sub-clusters is approximately 0.4. Similarly, sub-clusters agglomerated at a lower threshold are expected to have a higher correlation coefficient. This is confirmed by calculating the true correlation coefficients between the average cross-correlation functions of individual clusters. These are 0.44 (C1-C2), 0.63 (C1-C4), and 0.63 (C2-C4). A smaller cluster is also identified that includes three short time periods (1–3 months) between 2017–2019 (C3). The grouping of all three periods at a distance threshold of approximately 0.2 suggests high similarity between the different separate periods.

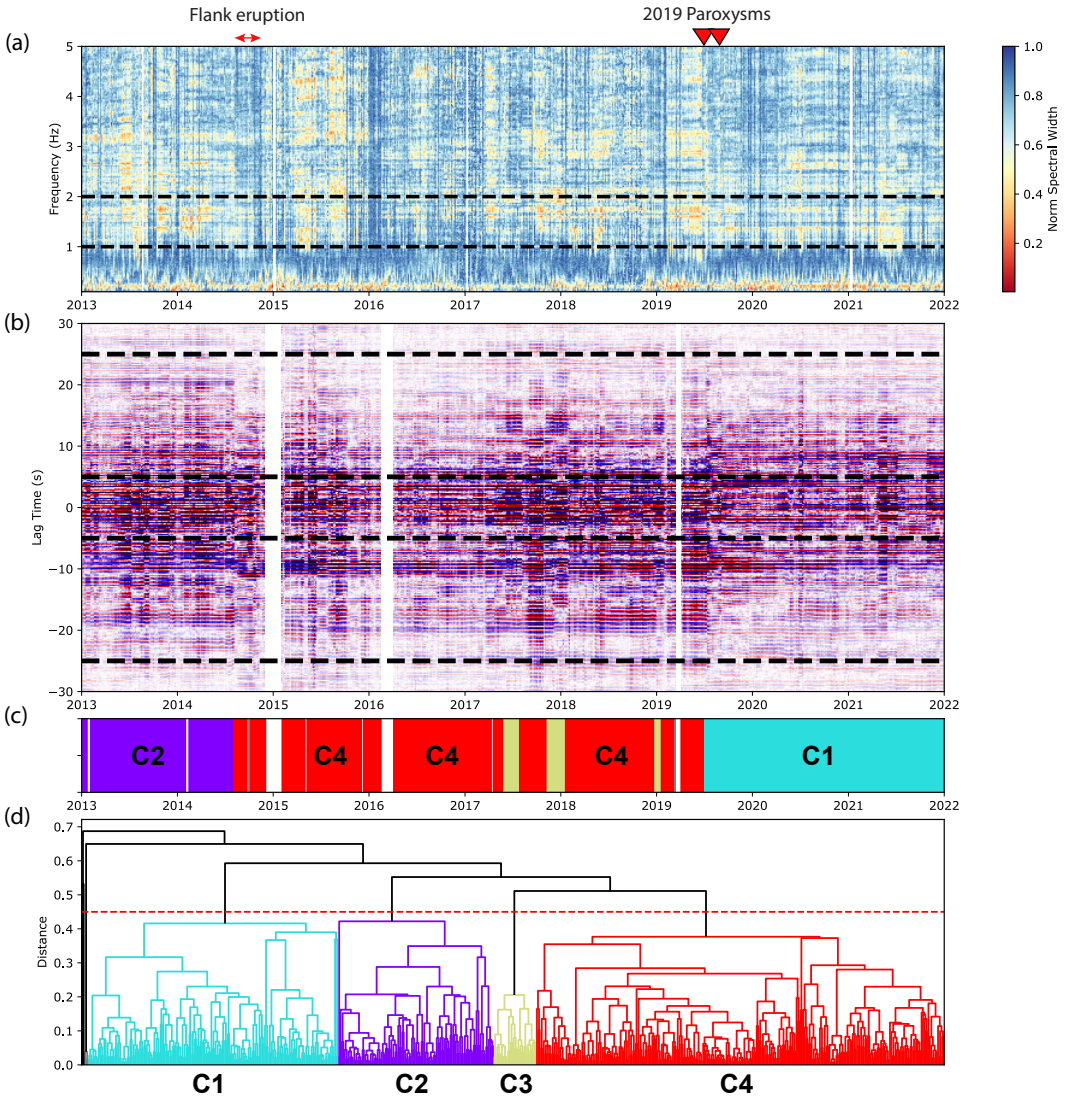

Figure 4.5: Hierarchical clustering output for 1.0-2.0 Hz cross-correlation functions (CCFs) computed for station-pair IST3-ISTR. (a) Normalized spectral width measurement. Dashed black lines show frequency range of CCFs. (b) Correlogram showing amplitudes of CCFs (red = positive, blue = negative). Dashed-black lines show part of CCFs used in clustering. (c) Location of clusters in time, color-coded according to dendrogram output in (d). (d) Dendrogram, with clustered defined at distance threshold of 0.45 (dashed red-line).

The data at 2.0–4.0 Hz follows similar structure, with the 2014 flank eruption and July 2019 paroxysm responsible for clear changes in the grouping of cross-correlation functions (Fig. 4.6). At a distance threshold of 0.65, the three largest clusters consist of cross-correlation functions pre-2014 lava flows (C1 in Fig. 4.6), cross-correlation functions between 2015 and the July 2019 paroxysm (C6), and cross-correlation functions post-paroxysm (C3). Some smaller clusters are identified at this threshold, including a smaller cluster that includes early-2018 and early-2019 cross-correlation functions (C4). This is similar to the smaller cluster observed at 1–2 Hz (C3 in Fig. 4.5). The final two smaller clusters are observed in late-2014 and 2015 (C2 and C5 in Fig. 4.6), seemingly associated with the co-eruptive and post-eruptive phase of

the 2014 lava flows respectively. Note, the higher threshold required to group these data into the described clusters – relative to 1.0–2.0 Hz data — suggests a lower correlation coefficient between the average function of each cluster.

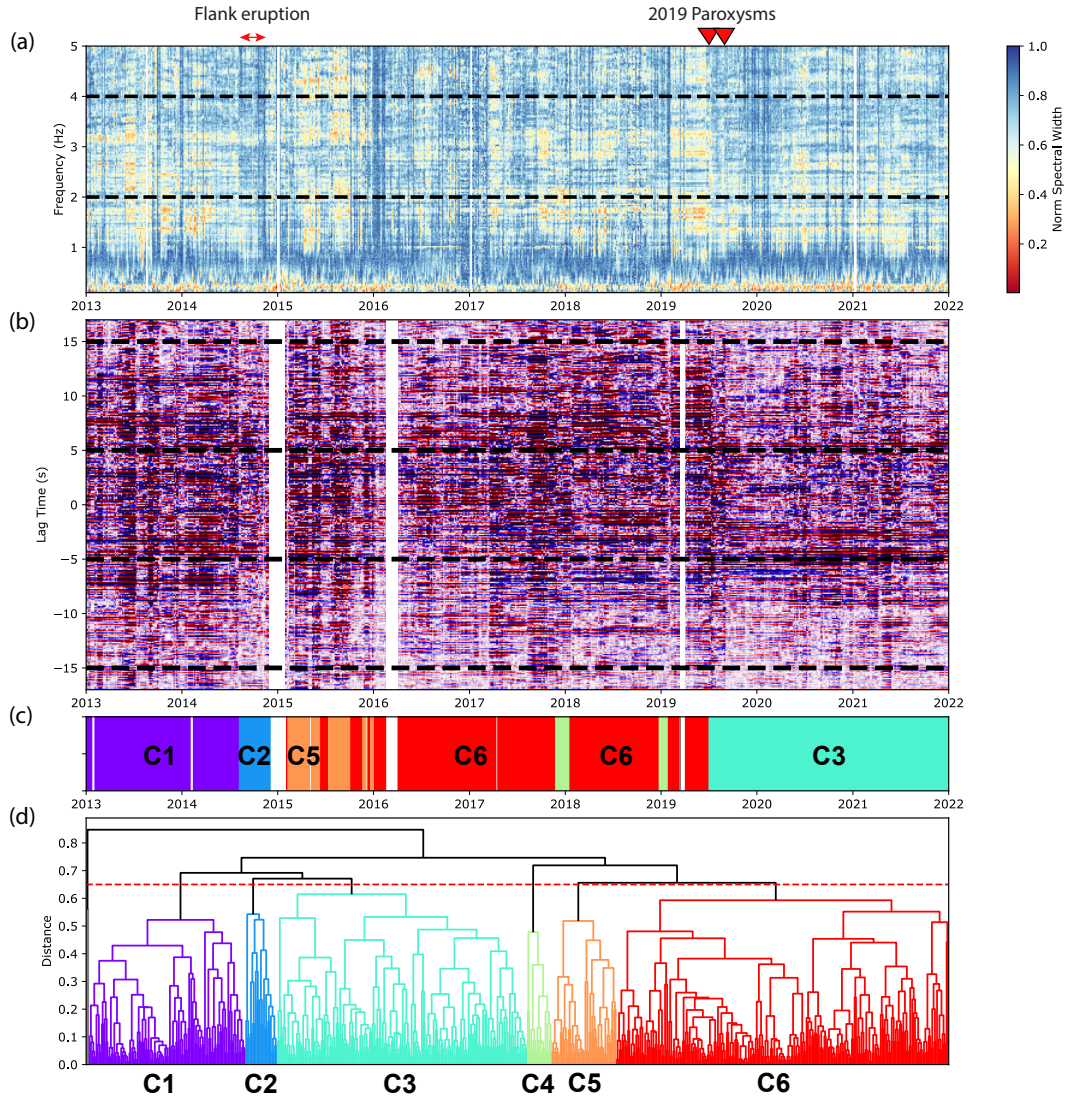

Figure 4.6: Hierarchical clustering output for 2.0-4.0 Hz cross-correlation functions (CCFs) computed for station-pair IST3-ISTR. (a) Normalized spectral width measurement. Dashed black lines show frequency range of CCFs. (b) Correlogram showing amplitudes of CCFs (red = positive, blue = negative). Dashed-black lines show part of CCFs used in clustering. (c) Location of clusters in time, color-coded according to dendrogram output in (d). (d) Dendrogram, with clustered defined at distance threshold of 0.65 (dashed red-line).

Following the observation that both 1–2 Hz and 2–4 Hz cross-correlation functions can mostly be grouped into three larger clusters split by eruptive activity, we examine the character of these in both the time and frequency domain (Fig. 4.7). In the time domain, clear differences are observed in the character of the average cross-correlation function within each cluster. This is especially the case at 2–4 Hz,

where the location of peaks in amplitude is highly variable between the different clusters (Fig. 4.7). In the frequency domain, the cross-correlation functions are shown to have relatively broad spectra with no single dominant peak. This is to be expected after applying spectral whitening, which acts to produce a broader-band signal (Bensen et al., 2007). Some similarities are observed between the time periods, for example between 1.0–1.3 Hz (Fig. 4.7c). Significant peaks at certain frequencies are, however, observed within some clusters that do not appear to be present in others.

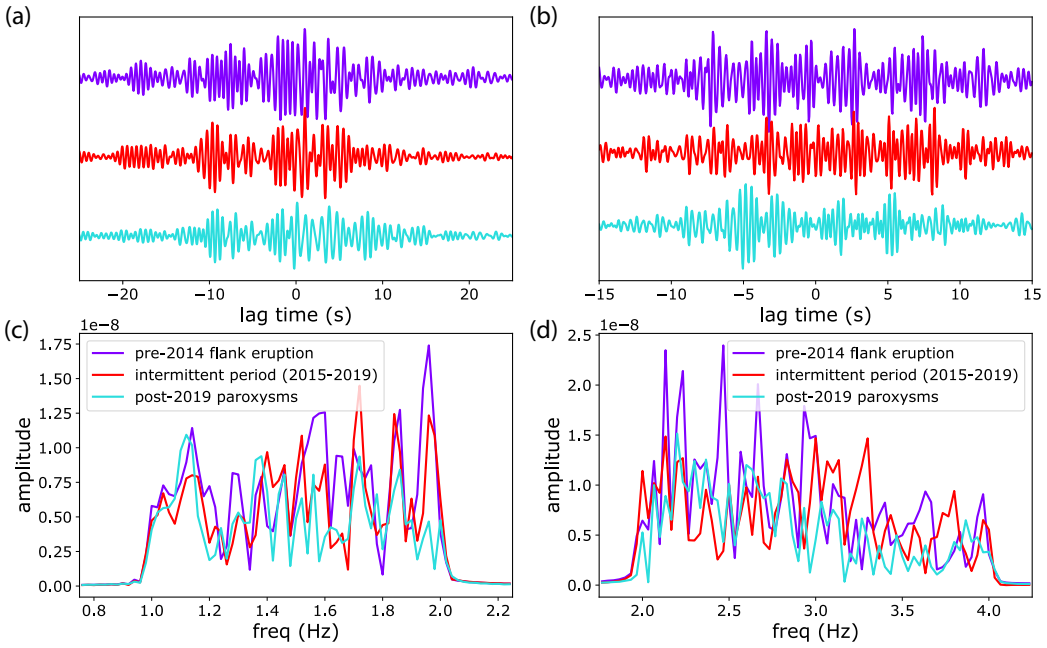

Figure 4.7: Comparing time and frequency domain character of average cross-correlation functions within three main clusters at 1–2 Hz and 2–4 Hz. These clusters represent approximately the periods pre-2014 flank eruption, post-2019 paroxysms, and the period in between. (a) Average of 1–2 Hz cross-correlation functions within clusters C1, C2 and C4 from Figure 4.5 in time domain. (b) Average of 2–4 Hz cross-correlation functions within clusters C1, C3 and C6 from Figure 4.6 in time domain. (c) Spectra of average 1–2 Hz cross-correlation functions. (d) Spectra of average 2–4 Hz cross-correlation functions.

The clusters formed of data pre-2019 paroxysms represent an important time period with regards to studying the processes leading up to these events. Thus, the characteristics of individual years 2017, 2018, and 2019 are explored in the time and frequency domain (Fig. 4.8). This is done using cross-correlation functions within cluster C4 at 1.0–2.0 Hz (Fig. 4.5c) and cluster C6 at 2.0–4.0 Hz (Fig. 4.6c). In the time domain, far greater similarity is observed between the individual years (Fig. 4.8a-b) relative to the clusters divided by eruptive activity (i.e. Fig. 4.7a-b). This similarity is also reflected in the frequency domain, with similar locations of spectral peaks between the clusters (Fig. 4.8c-d).

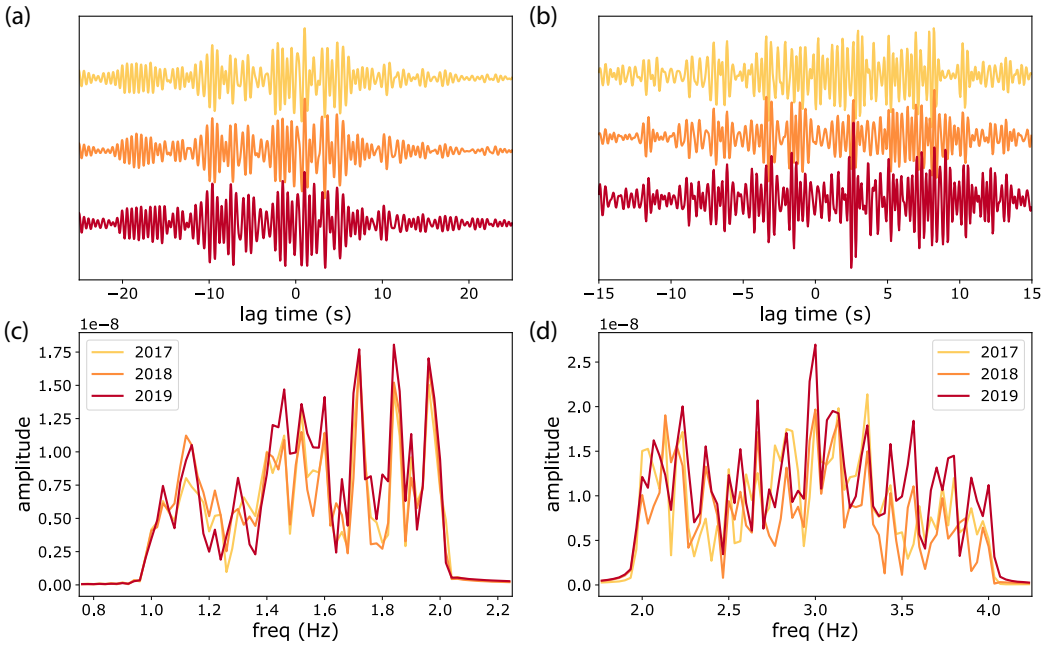

Figure 4.8: Comparing time and frequency domain character of average cross-correlation functions for individual years within main 2015–2019 cluster. These are cluster C4 (Fig. 4.5) for 1–2 Hz and cluster C6 (Fig. 4.6) for 2–4 Hz. (a) Average of individual years in time domain, 1–2 Hz cross-correlation functions. (b) Average of individual years in time domain, 2–4 Hz cross-correlation functions. (c) Spectra of average cross-correlation functions shown in (a). (d) Spectra of average cross-correlation functions shown in (b).

Finally, at 0.5–1.0 Hz, the data can be similarly split into clusters that appear to reflect volcanic activity (Fig. 4.9). For example, at a threshold of 0.35, two clusters account for the significant majority of the dataset. One of these (C3) includes all cross-correlation function pre-2014 flank eruption and the majority of data post-2019. Note, however, that unlike the other frequency bands, the July 3 paroxysm in 2019 does not appear to represent a clear transition from one cluster type to another (Fig. 4.9c). The other large cluster (C2), includes mostly cross-correlation functions from 2014–2017 during a time period the volcano was considered low activity. Thus, as with the other frequency bands, volcanic activity appears to have some influence on cross-correlation functions below 1 Hz.

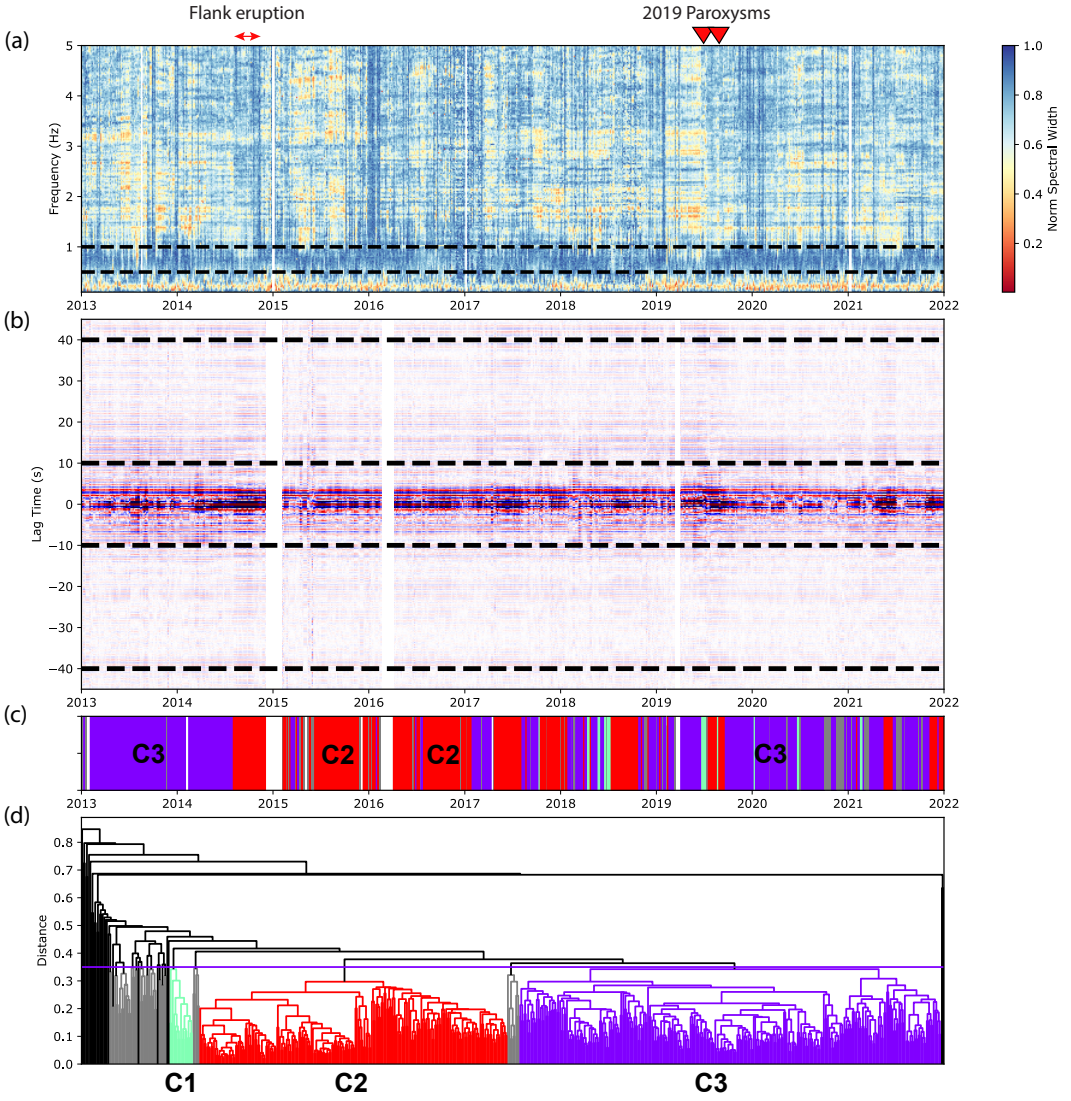

Figure 4.9: Hierarchical clustering output for 0.5-1.0 Hz cross-correlation functions (CCFs) computed for station-pair IST3-ISTR. (a) Normalized spectral width measurement. Dashed black lines show frequency range of CCFs. (b) Correlogram showing amplitudes of CCFs (red = positive, blue = negative). Dashed-black lines show part of CCFs used in clustering. (c) Location of clusters in time, color-coded according to dendrogram output in (d). (d) Dendrogram, with clustered defined at distance threshold of 0.35 (dashed red-line).

4.3.2 Velocity changes

Seismic velocity changes computed using the stretching-technique are presented for three different frequency bands in Figure 4.10, where we show the average for both single-station (averaging all stations) and station-pairs (averaging all pairs). Results applying the moving-window cross-spectral technique were found to be highly similar, shown in Figure S4.8. Two reference periods were tested towards computing the velocity changes, one using all the available data and the other using the period defined by cluster C4 in Figure 4.5, corresponding approximately to the pe-

riod 2015–2019. As would be expected, the 2015–2019 reference resulted in slightly increased correlation coefficient values during the same period relative to the full dataset reference period (Fig. S4.7). Minimal differences were, however, observed in the seismic velocity changes using either reference period. Thus, we show the results using the reference function of all available data for the purpose of interpreting velocity changes over the whole nine-year period.

In all frequency bands, we observe strong similarities between the results computed using pairs of stations and single-stations (Fig 4.10). Larger, approximately annual, periodicities are observed at higher frequencies, with variations on the order of +/- 0.5% measured at 1.0–2.0 and 2.0–4.0 Hz during the nine-year period (Fig. 4.10b and 4.10c). The annual nature of the data are confirmed when examining the spectral content of velocity changes, with a strong peak at approximately 365 days for 1.0–2.0 and 2.0–4.0 Hz (Fig. 4.11b & 4.11c). At 0.5–1.0 Hz, velocities vary between +/- 0.1 % (Fig. 4.10a). The appearance of increased scatter in measurements at lower frequencies in Figure 4.10a can be related to the reduced amplitude of changes relative to the higher frequencies. Examining the spectra does reveal a peak at 365 days for single-station results (Fig. 4.11a). Such a peak is less apparent in station-pair results. Smoothing the 0.5–1.0 Hz velocity changes better reveals annual changes similar to those observed at higher frequencies (Fig. 4.12).

The smoothed velocities at lower frequencies (0.5–1.0 Hz) also reveal a long-term trend that varies through-out the nine year time period (Fig. 4.12). Between 2013–2015 the velocities are increasing before then plateauing between 2015–2017 for both single-station and station-pairs. Velocities then begin to decrease from early-2017, reaching a minimum in early-2019 using station-pairs and mid-2020 using single-station. From 2021, velocities appear to be increasing again using both pairs of stations and single stations.

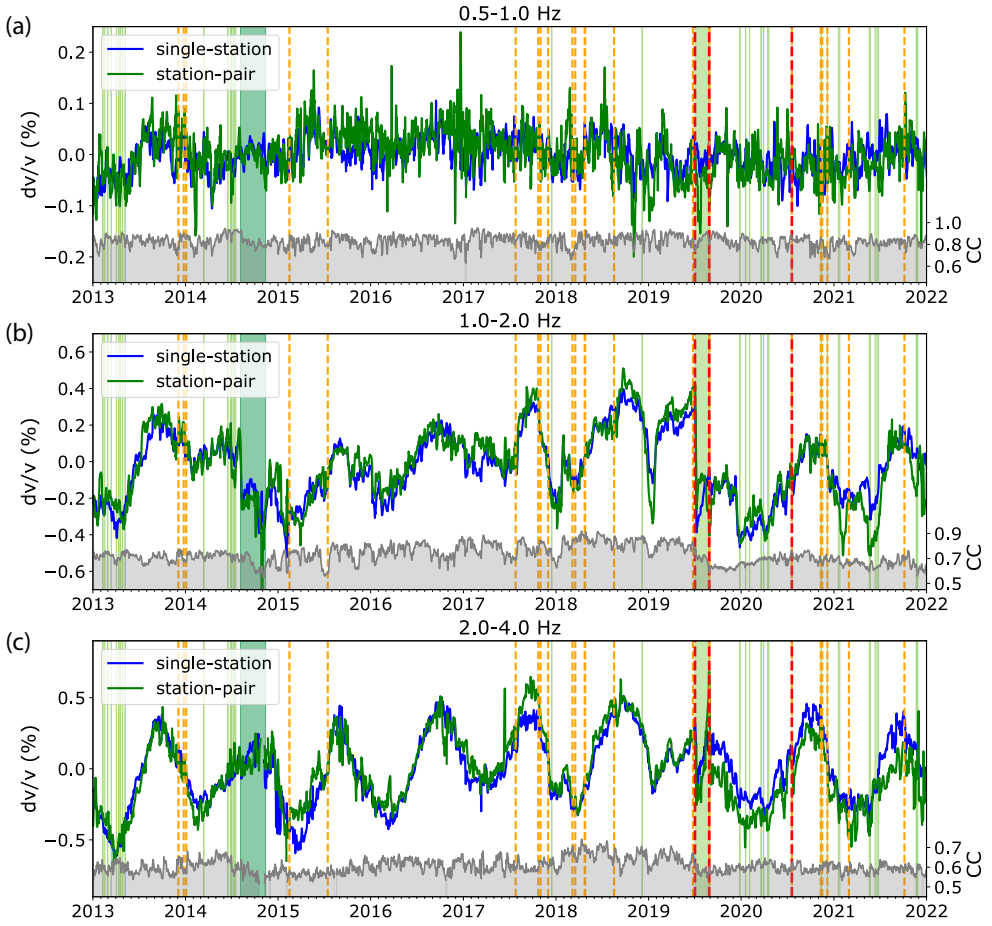

Figure 4.10: Velocity changes computed in three different frequency bands using stretching technique. Single-station velocities (blue) correspond to the average of all single-station cross-components (EN, EZ, NZ) for all stations. Station-pair velocities (green) correspond to the average of all station-pairs (ZZ). Correlation coefficient (CC) between 5-day moving stacks and reference shown as grey shading. Volcanic activity highlighted corresponding to descriptions in Figure 4.2. (a) 0.5–1.0 Hz. (b) 1.0–2.0 Hz. (c) 2.0–4.0 Hz.

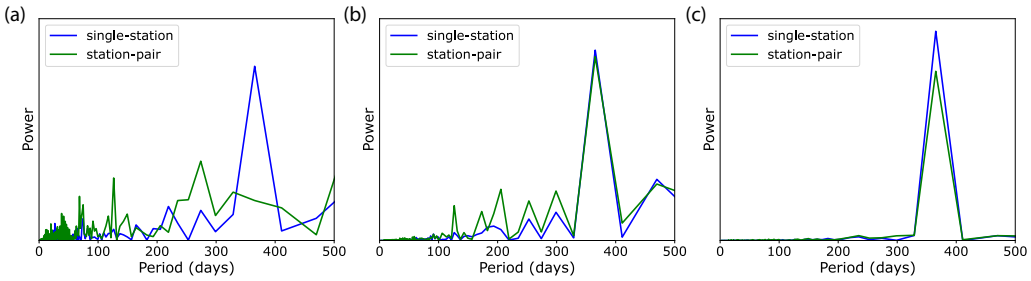

Figure 4.11: Spectra of velocity changes (stretching technique) at different frequencies (a) 0.5–1.0 Hz. (b) 1.0–2.0 Hz. (c) 2.0–4.0 Hz.

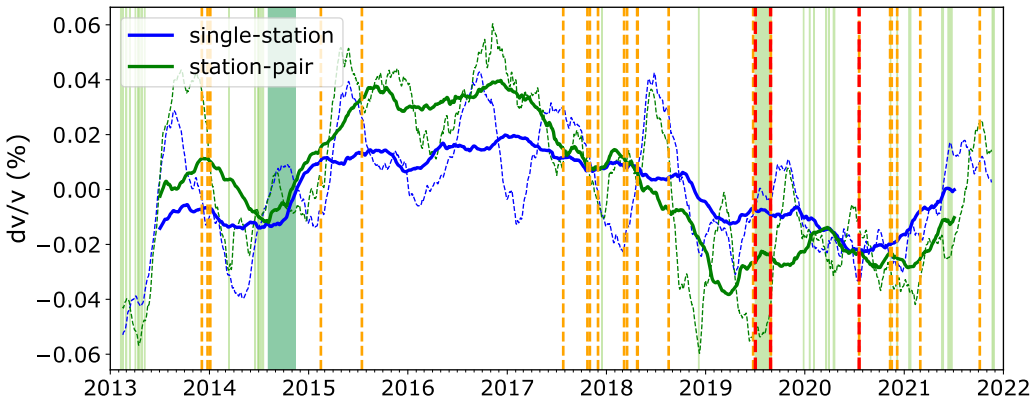

Figure 4.12: Smoothed 0.5–1.0 Hz velocity changes for both single-station and station-pairs (stretching technique). Dashed-lines represent 90-day smoothing window. Solid lines represent 365 smoothing.

Annual changes in seismic velocity are visualised more clearly by overlaying the velocity of each year in the different frequency bands (Figure 4.13). We do this from the start of 2015 up to the July 2019 paroxysm. At higher frequencies (1–2 and 2–4 Hz), we observe a relatively consistent trend of low velocities in the early months of the year followed by high velocities in mid-to-late months (Fig. 4.13b,c,e,f). At lower frequencies (0.5–1.0 Hz), we see a more subtle annual variation in single-station results, with higher velocities in the middle of the calendar year (Fig. 4.13a). For station-pair results, minimal annual variation is observed (Fig. 4.13d), consistent with the absence of a 365-day peak in the spectra (Fig. 4.11a).

Of interest is the deviation of velocities from the annual trend in 2019. At 2.0–4.0 Hz, we observe that velocities increase in the months January to April during a time when velocities are usually at a minimum (Fig. 4.13c,f). A similar observation is also present in the velocity changes for 1.0–2.0 Hz (Fig. 4.13b,e). It is also interesting to note that velocities in 2019 (prior to the paroxysms) at 1.0–2.0 and 2.0–4.0 Hz, are higher than any other period during the nine-years. At 0.5–1.0 Hz, velocities between May–July appear slightly lower than expected from usual trends (Fig. 4.13a,d). Similarly, for the station-pair results, these velocities represent extreme values over the nine-year period, corresponding to the lowest recorded (Fig. 4.12).

Comparing the annual velocity changes at higher frequencies (1–4 Hz) with the comparison datasets of temperature 4.13g), pore pressure changes 4.13h) and sea level pressure 4.13i), we find that only pore pressures show a similar deviation from the average trend in 2019. For the pore pressures here, the diffusivity c is set to 0.02 m^2/s, where this gave the maximum absolute value of the cross-correlation between modeled pore pressure changes and de-meaned velocity changes (2.0–4.0 Hz). Temperature and sea-level pressure changes, on the other-hand, show minimal deviation from the average annual trend that can explain elevated velocity changes in the first few months of 2019. At lower frequencies (0.5–1.0 Hz), a drop in sea

level pressure early-2019 coincides with a similar decrease in velocity. The velocity decrease between May and July, however, does not show a comparable change in sea level pressure. We do note though that the slight annual trend in the sea level pressure data is similar to that recorded at single-station, albeit inverse.

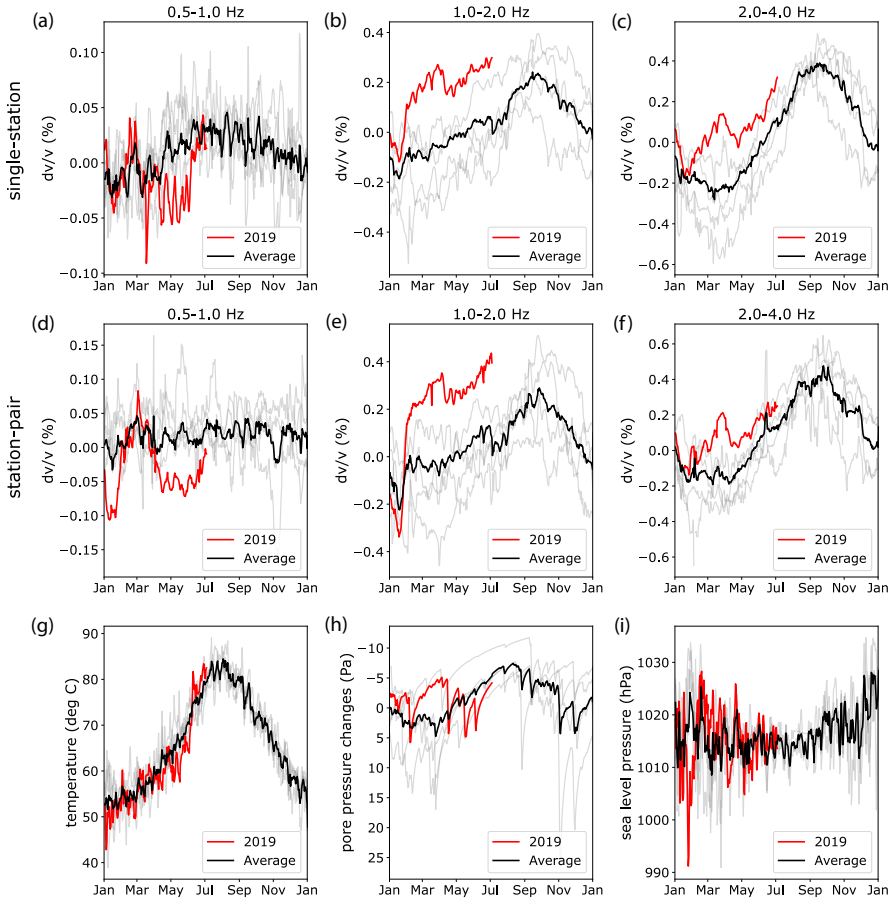

Figure 4.13: Single-station and station-pair seismic velocities (stretching technique) for each year plotted against calendar day, from 2015-01-01 to 2019-07-03, with pore pressures computed also shown. Grey lines correspond to individual years with exception of 2019 which is represented by the red-line. Average (mean) is shown as black line. (a) single-station 0.5–1.0 Hz. (b) single-station 1.0–2.0 Hz. (c) single-station 2.0–4.0 Hz. (d) station-pair 0.5–1.0 Hz. (e) station-pair 1.0–2.0 Hz. (f) station-pair 2.0–4.0 Hz. (g) Temperature data. (h) Pore pressures, with c=0.03m^2/s. (i) Sea-level pressure.

In addition to annual velocity changes, we observe sudden drops of velocity associated with the beginning of the 2014 flank eruption and the July 2019 paroxysms. This is best observed in the frequency band of 1.0–2.0 Hz (Fig. 4.10b), where a velocity decrease of approximately 0.4 % is recorded in 2014 and 0.8 % in 2019. In both cases, the velocity does not seem to have recovered for a few years, with recovery still ongoing following the 2019 paroxysm by the start of 2022. At higher frequencies (2.0–4.0 Hz, (Fig. 4.10c)), in contrast, minimal change is observed following the 2014 flank eruption and a 0.4 % velocity decrease is recorded following the July 2019 paroxysm. Drops in the correlation coefficient between 5-day cross-correlation

functions and the reference are observed in both frequency bands.

Finally, examining the 2.0–4.0 Hz velocity changes recorded by single-stations individually reveals an interesting deviation of STRA station (located closest to the summit) from the other stations in the months prior to the July 2019 paroxysm (Fig. 4.14a). This is more clearly observed when comparing the velocity of STRA station with the average of the other four (Fig. 4.14b), with the largest positive difference recorded one month before the July paroxysm (Fig. 4.14c). It also appears that the deviation of STRA station from the others increasing gradually from early-2017. Following the paroxysm, the velocity difference returns to approximately zero until the end of the dataset. Interestingly, a similar observation is made around the time of the 2014 flank eruption, with elevated velocities at STRA station relative to the other stations from mid-2013. The velocity difference subsequently decreases to values below zero just before the onset of the flank eruption. These features are similarly resolved using the moving-window cross-spectral technique (Fig. S4.9). Furthermore, an increasing velocity difference is also resolved using 1–2 Hz cross-correlation functions between 2017 and the July 2019 (Fig. S4.10). However unlike at higher frequencies, the values in 2017–2019 are not clearly elevated relative to earlier time periods (2015–2016).

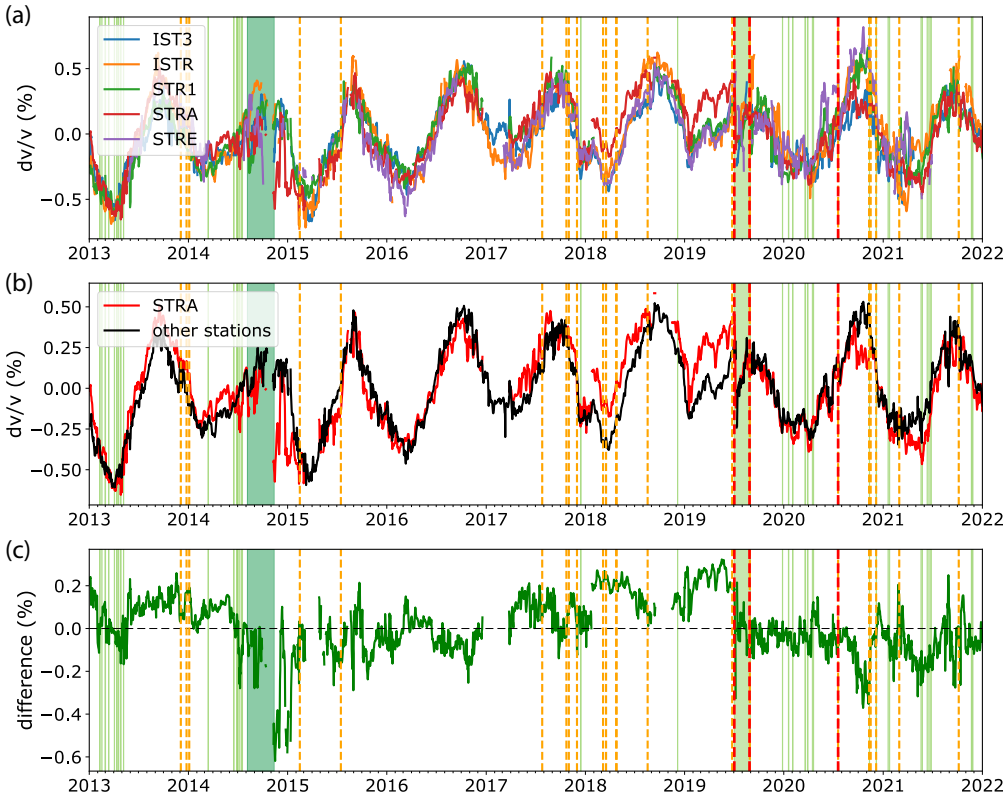

Figure 4.14: Comparing 2.0–4.0 Hz velocities (stretching technique) recorded at STRA station with other four stations. Volcanic activity highlighted corresponding to descriptions in Figure 4.2. (a) Single-station seismic velocities for all stations. (b) Velocities for STRA station compared with average of IST3, ISTR, STR1, and STRE stations. (c) Difference between average velocity of STRA and other four stations.

Taking a closer look at the velocity difference at STRA station between 2013–2015 and 2016–2020 reveals some interesting features (Fig. 4.15). For example, in 2013–2015, an increasing trend is generally observed during times without significant event while a decreasing trend is observed during lava flows (for example early-2013, and in the months before and during the flank eruption) (Fig. 4.15a). A sharp drop is also observed following major explosions late-2013/early-2014. Between 2016–2020, there is some suggestion that the increase in velocity difference at STRA station starting early-2017 may not be linear, but rather occur in stages (Fig. 4.15b). The first clear increase occurs somewhere between late-2016 and August 2017. The exact timing seems to vary depending on whether the moving-window cross-spectral or stretching technique are used. Further increases are then observed using both methods at approximately the same time as smaller clusters of cross-correlation functions in late-2017/early-2018 and late-2018/early-2019 (Fig. 4.15b).

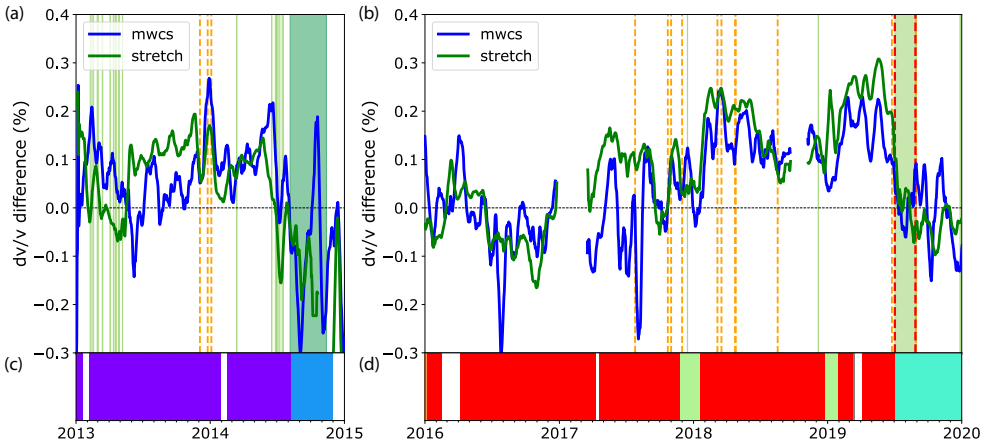

Figure 4.15: Closer look at relative difference between 2–4 Hz velocity changes at STRA station and other four stations (IST3, ISTR, STR1, STRE) between (a) 2013–2015 and (b) 2016–2020. Both the results from using the stretching and moving-window cross-spectral technique are shown. (c) Cluster solution for 2–4 Hz cross-correlation functions between 2013–2015. (d) Cluster solution for 2–4 Hz cross-correlation functions between 2016–2020. Volcanic activity highlighted corresponding to descriptions in Figure 4.2.

4.4 Discussion

4.4.1 Stability of cross-correlation functions in time

The results of applying hierarchical clustering indicate that major events at the volcano strongly influence the character of cross-correlation functions at Stromboli (Figs. 4.5 and 4.6). Within the frequency band of tremor (1–2 Hz and 2–4 Hz), significant changes are observed in both the time and frequency domain associated with the 2014 flank eruption and the July 2019 paroxysm (Fig. 4.7). These events are also accompanied by seismic velocity decreases, most significant in the 1–2 Hz frequency band (Fig. 4.10b).

Changes in the character of cross-correlation functions can be related to both changes in the seismic source (e.g. Gómez-García et al., 2018; Caudron et al., 2021; Yates et al., 2023; Makus et al., 2023) and structural changes (through a change in scattering properties, Obermann et al., 2013; Yates et al., 2023). Both the 2014 flank eruption and July 2019 paroxysmal eruption were preceded by rapid increases in ground displacement in the hours before (Di Traglia et al., 2015, 2018, 2021). These are interpreted to reflect a strong and fast inflation of the shallow storage system beneath Stromboli (Di Traglia et al., 2015, 2021), thought to be located within the first few hundred meters of the surface (Gauthier et al., 2000). Significant velocity changes at the onset of these events are consistent with deformation-induced changes seen at other volcanoes (e.g. Clarke et al., 2013; Budi-Santoso and Lesage, 2016a; Donaldson et al., 2019), albeit on shorter time-scales (hours rather than days to weeks). Furthermore, the observation that these velocity changes are largest in the 1.0–2.0 Hz frequency band – most sensitive between approximately 150–400 m depth (Fig. S4.12) – suggests that these changes are less-likely related to very near surface processes (e.g. landsliding). Thus we consider that rapid deformation, and subsequent damage, of the shallow storage system and surrounding rock may contribute to changes in the character of cross-correlation functions. The accompanying velocity decreases could then be explained by an increase in the density of microcracks.

Deformation of the shallow volcanic system may also explain differences in the frequency content of cross-correlation functions split by these events (Fig. 4.7). Changes within the plumbing system (structural or stress-related) can influence the characteristics of tremor (Carniel and Tárraga, 2006; Tárraga et al., 2014). For example, changes in the amplitude and frequency of tremor have been observed due to stress changes in the host rock (Tárraga et al., 2014). Closer inspection of the spectral width measurements in 2019 shows a clear difference before and after the July 3 paroxysm (Fig. 4.16). For example, the three months prior to the paroxysm show more coherent energy at 1.5–2.0 Hz and 2.6–3.0 Hz relative to the two months following (Fig. 4.16b). We therefore suggest that these changes in the frequency of tremor post-paroxysm reflect modification of the shallow plumbing system at Stromboli following major eruptive events.

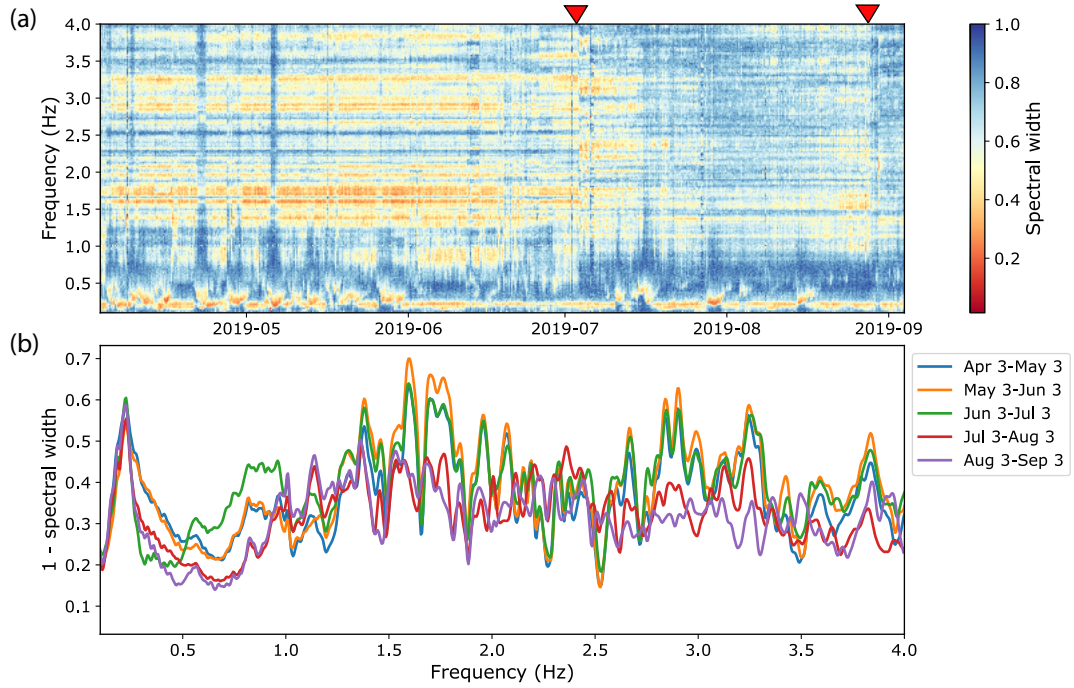

Figure 4.16: Comparison of spectral width measurements over five individual months (three before the July 3 paroxysm and three after). (a) Spectral width measurements over five months including the two paroxysms. (b) Average spectral width for five individual months, plotted as 1 minus the spectral width.

Given the continuous nature of volcanic tremor at Stromboli (Langer and Falsaperla, 1996; Chouet et al., 1997; Ripepe and Gordeev, 1999), we assume that the tremor process remains the same throughout the nine year time-period (Fig. 4.3). Thus we do not anticipate any new seismic source, or loss of a previous seismic source, is responsible for changes in frequency content. However

Temporal variability in seismic tremor is a concern for monitoring. Cluster analysis applied to volcanic tremor at Stromboli previously has shown that its spectra is not constant over time (Langer and Falsaperla, 1996, 2003). Thus caution should be exercised when interpreting velocity changes within this frequency band. In this study, the main features of interest are the approximately annual velocity changes (Fig. 4.11) and the deviation of velocities recorded at STRA station relative to the other stations between 2017–2019 (Fig. 4.14). For the annual changes, we observe no evidence from clustering cross-correlation functions that the data are impacted by seasonal variability in seismic sources. This is also confirmed at a lower distance threshold (Fig. S4.6), with no clearly identifiable seasonality in the cluster solution. Thus we consider that the annual velocity changes reflect real changes within the medium.

For the single-station velocities recorded between 2017–2019, spectral analysis of cross-correlation functions within individual years shows the spectra are similar over this time period (Fig. 4.8). We also find that both the annual changes in seismic velocity and deviation of velocities at STRA between 2017–2019 are resolved

independent of whether the stretching technique or moving-window cross-spectral technique are used to compute seismic velocity changes (Fig. S4.8 and S4.9). This is significant as, unlike the stretching method, the moving-window cross-spectral technique is based on phase measurements. It is therefore considered to be less sensitive to variable frequency content (Zhan et al., 2013; Hillers et al., 2015b; Mao et al., 2020). Some authors have therefore considered that consistent results using both the stretching and moving-window cross-spectral technique are indicative of robust velocity estimates (Hillers et al., 2015b). Thus we similarly interpret that the deviation of seismic velocities recorded at STRA station, relative to the other four stations, reflect real changes in the medium.

Finally, a smaller cluster is observed between 2017–2019 in both 1–2 Hz (C3 in Fig. 4.5) and 2–4 Hz data (C4 in Fig. 4.6). These clusters appear to be associated with heightened activity, with the alert level raised at the volcano from green to yellow in both December 2017 and 2018 (Giudicepietro et al., 2019, 2020). Small inflation of the summit crater was also recorded during both time periods (Schaefer et al., 2019; Di Traglia et al., 2021). At 1–2 Hz, two months of data are also included in this cluster in mid-2017 (from early-June to August), during elevated activity (Giudicepietro et al., 2019). Comparing the cross-correlation functions computed during these periods of heightened activity with background activity between 2017–2019 shows clear differences in the time and frequency domain (Fig. 4.17). For example, greater energy is observed between 1.6–2.3 Hz in cross-correlation functions computed during these smaller episodes of heightened activity (Fig. 4.17c-d). This is consistent with drops in the correlation coefficient observed during all three episodes — best seen at 1–2 Hz (Fig. 4.10b) — with the latter two episodes in late-2017 and late-2018 associated with periods of significant velocity decrease. For example, at 1–2 Hz, the two lowest velocities recorded between 2017 and mid-2019 occur during these periods of heightened activity (Fig. 4.10b). This is also reflected examining the velocity changes recorded by the three individual stations closest to the summit (Fig. 4.18). At 1–2 Hz, these changes appear sharp (Fig 4.18a), with rapid increases in the velocity recorded once the cluster ends (Fig 4.18d). Inflation of the summit could explain a velocity decrease, through crack opening. However, the speed at which the velocity recovers once outside of the smaller clusters — especially in 2019 — suggests the measurements may be contaminated. A similar sharp velocity increase is observed following the succession of the smaller cluster in early-2018 (Fig. 4.18a-b), though a slight reduction in pore pressures may also contribute (Fig. 4.18c). Taking these observations into account, we exercise caution interpreting the velocity changes within the periods of heightened activity.

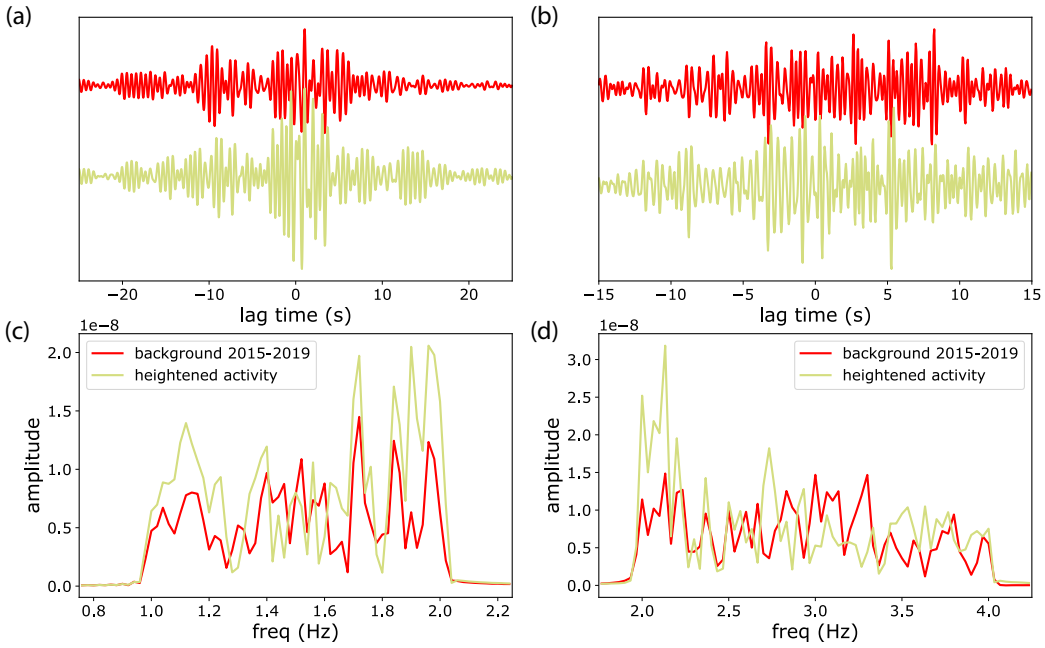

Figure 4.17: Comparing time and frequency domain character of average cross-correlation functions for main 2015–2019 cluster with smaller cluster during heightened activity within same time period. For the main (background) cluster, these are cluster C4 (Fig. 4.5) for 1–2 Hz and cluster C6 (Fig. 4.6) for 2–4 Hz. For clusters associated with heightened activity, these are cluster C3(Fig. 4.5) for 1–2 Hz and cluster C4 (Fig. 4.6) for 2–4 Hz. (a) Average of 1–2 Hz cross-correlation functions in time domain. (b) Average of 2–4 Hz cross-correlation functions within clusters in time domain. (c) Spectra of average 1–2 Hz cross-correlation functions. (d) Spectra of average 2–4 Hz cross-correlation functions.

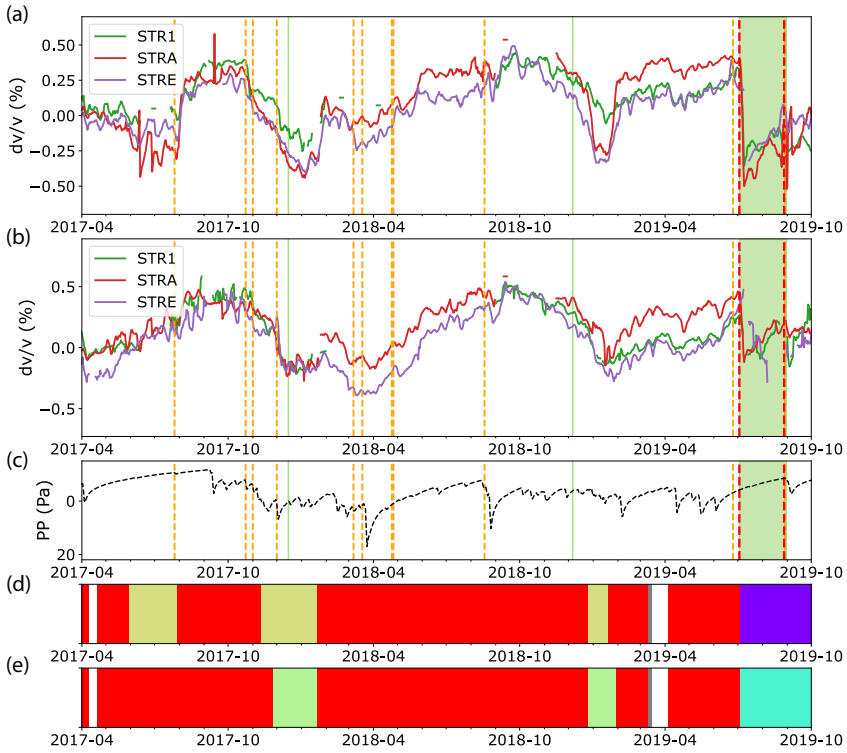

Figure 4.18: Single-station velocity changes for three stations closest to summit (STR1, STRA, STRE) from period of renewed activity in 2017 to post-2019 paroxysms. Volcanic activity highlighted corresponding to descriptions in Fig. 4.2. (a) velocity changes at 1–2 Hz. (b) velocity changes at 2–4 Hz (c) pore pressure (PP) changes. (d) cluster positions in time for 1–2 Hz (from Fig. 4.5d) (e) cluster positions in time for 2–4 Hz (from Fig. 4.6d).

An interesting observation from the cluster dendrogram at 1–2 Hz is that the cross-correlation functions computed within the smaller cluster are grouped at a very low threshold (Cluster C3 in Fig. 4.5d). This suggests a high similarity between cross-correlation functions computed across the three episodes, despite their separation in time. This is confirmed when examining the cross-correlation functions computed within each individual episode in the time and frequency domain (Fig. 4.19a,c). Such high similarity provides an opportunity to re-compute the relative velocity changes within this cluster, where we stretch each cross-correlation within this cluster relative to the average (Fig. 4.19c). The high similarity allows for a minimum threshold of 0.9 to be defined for the cross-correlation coefficient. From this, we observe that the relative changes within this smaller cluster are, in fact, relatively consistent with the single-station results previously measured with the full reference period (Fig. 4.19c). The deviation of the velocity at STRA station from the other stations is also resolved (Fig. 4.19d). The consistency with results using the full reference suggest that the results may not be so sensitive to changes in frequency content. However, we note that STRA station recorded a much larger decrease relative to the other stations using the full reference (Fig. S4.10b-c). Thus, caution is again exercised in interpreting velocity changes within these time periods using the full reference.

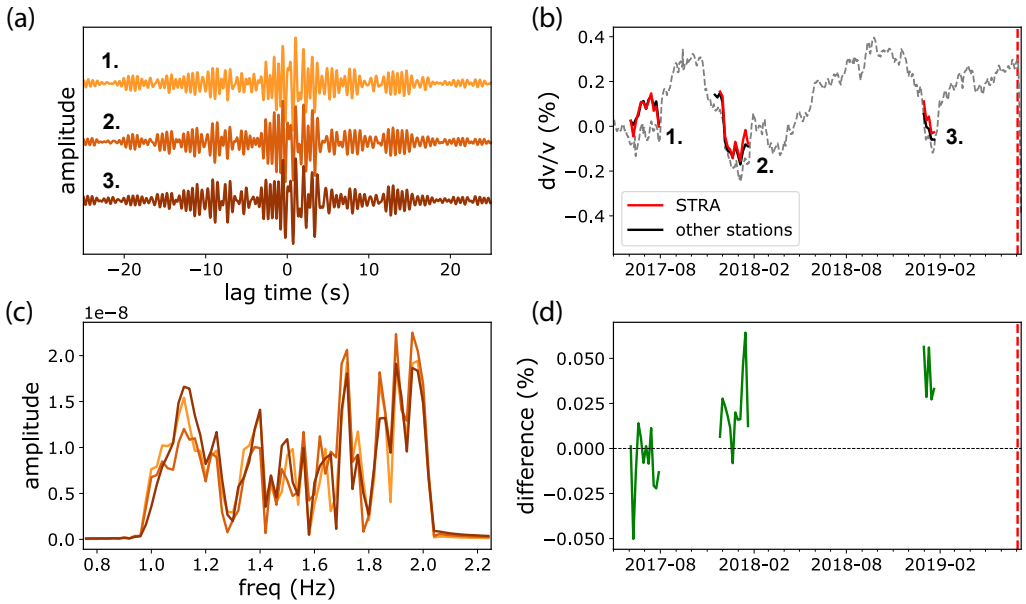

Figure 4.19: Comparing cross-correlation functions computed during periods of heightened activity at 1–2 Hz (cluster C3 in Fig. 4.5d). (a) cross-correlation functions in time-domain during three separate periods of heightened activity. (b) spectra of cross-correlation functions. (c) re-computed velocity changes for STRA station (red-line) and average of other five stations (black-line), compared with original average velocity change of all stations using full reference period (grey-dashed line). (d) difference between STRA station velocities and average of other stations (IST3, ISTR, STR1, STRE).

4.4.2 Seasonality in velocity changes

Annual velocity changes dominate higher frequency velocity changes at Stromboli, with measurements on the order of ±0.5% recorded between 1–4 Hz (Fig. 4.10). The relation to a seasonal process is strongly suggested by the approximate 365 day periodicity of these changes (Fig. 4.11). This is consistent with previous work of Calò et al. (2021), who identified seasonal changes at Stromboli at frequencies between 1.5–2.5 Hz between 2010 and 2013. At lower frequencies, evidence of seasonality is also identified when smoothing velocities to bring out the longer-term trend (Fig. 4.12), noting that amplitudes are much smaller.

Seasonal changes in seismic velocity have been related to a number of meteorological processes. These include precipitation (groundwater changes) (e.g. Sens-Schönfelder and Wegler, 2006; Rivet et al., 2015; Wang et al., 2017; Lecocq et al., 2017; Clements and Denolle, 2018; Feng et al., 2021), temperature-induced thermo-elastic strain (e.g. Meier et al., 2010; Richter et al., 2014; Hillers et al., 2015a; Lecocq et al., 2017) and variations in wind speed (Hillers et al., 2015a). In the latter case, this is a non-physical change due to bias in velocity changes measurements caused by systematic noise excitation (Hillers et al., 2015a). Atmospheric pressure changes have also been

observed to influence seismic velocities on shorter-time scales (e.g. Silver et al., 2007; Niu et al., 2008). However, the observation that daily fluctuations in atmospheric pressure are often larger than any seasonal trend suggests it is not the dominant mechanism for higher frequency changes at Stromboli (Fig. 4.13). Similarly, there is minimal evidence in the character of cross-correlation functions in time to suggest an artificial influence due to wind-speed related changes in seismic source.

Precipitation and temperature induced changes are both plausible mechanisms that may contribute to seasonal velocity changes at the volcano. For example, a good fit is found between velocity changes and inverted pore pressure changes due to precipitation from 2013–2017 (Fig. 4.20). Similarly, the deviation of velocities from the annual trend in early-2019 — at higher frequencies (Fig. 4.13b-c) — can possibly be explained by a drop in pore pressures around the same time (Fig. 4.13e). Temperature changes, in comparison, are far more consistent throughout the year (Fig. 4.13d). Thus there is no clear explanation why a strong change from the annual trend would be observed if temperature induced changes were dominant. From 2019–2022, however, there are strong discrepancies between modeled pore pressure changes and measured velocities (Fig. S4.3a), which instead continue to follow an approximate annual trend similar to the temperature profile (Fig. S4.3b). Thus we cannot rule out some contribution of temperature induced thermo-elastic changes. Determining the relative contribution of these two processes however is outside the scope of this study.

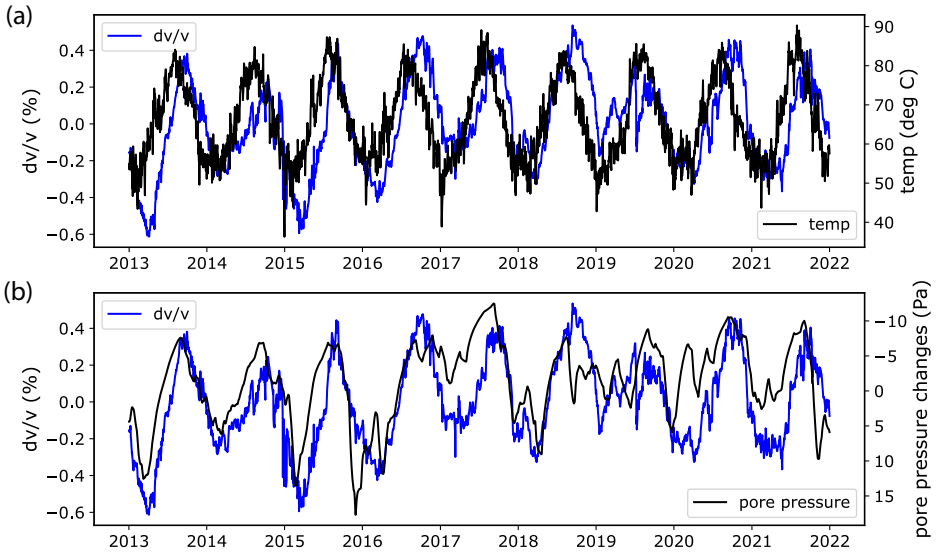

Figure 4.20: Single-station velocity changes at 2.0–4.0 Hz (blue-line) compared with different meteorological data (black-lines). (a) Velocities compared with temperature data. (b) Velocities compared with calculated pore pressure changes.

At lower frequencies (0.5–1.0 Hz), a seasonal trend is only clearly observed using single-station results (Fig. 4.13a). The trend of these changes showed some similarity to the inverse trend of seasonal atmospheric pressures. A comparison of the two datasets shows some similarities (Fig. 4.21), however a clear connection cannot be made given that atmospheric pressure changes are expected to be near instantaneous. Thus differences in phase between seasonal peaks suggest an alternative

mechanism.

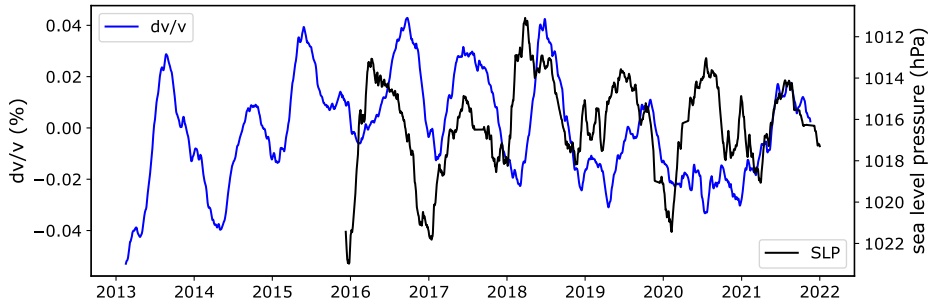

Figure 4.21: Single-station velocity changes at 0.5–1.0 Hz (blue-line) compared atmospheric pressure changes (both smoothed by 90 days)

If precipitation induced changes are assumed, synthetic velocity changes can be modeled based on previously calculated pore pressure changes. Following Wang et al. (2017), the transfer function to convert pore pressure changes into modeled velocity changes ($\delta v/v_{syn}$) is defined as follows:

$$\frac{\delta v}{v}_{syn}(t) = \langle \frac{\delta v}{v}_{obs}(t) \rangle + \frac{cov(\frac{\delta v}{v}_{obs}(t), P(t))}{var(P(t))} \cdot (P(t) - \langle P(t) \rangle) \tag{4.4}$$

where $\langle\rangle$ represents the average over time. The time period between 2013-01-01 and 2014-06-01 is used to compute the transfer function, where a good fit between pore pressures and seismic velocities were observed (Fig. 4.20). Additional smoothing was also applied to pore pressure changes (50–days), noting that velocities do not show much evidence of being greatly affected by short-term precipitation events. Once the transfer function is computed, it is then applied to the full time series of velocity changes at both 1–2 Hz and 2–4 Hz (Fig 4.22a-b). The residual between observed and synthetic is then computed to correct for precipitation-induced velocity changes (Fig. 4.22c-d).

Residual velocities are highly variable throughout the nine-year period at both 1–2 and 2–4 Hz after subtracting the modeled velocity changes from precipitation (Fig. 4.22c-d). At 1–2 Hz, the decreases associated with the 2014 flank eruption and July 2019 paroxysm are evident, with apparent recovery by early-2016 and early-2022 respectively (Fig. 4.22c). At 2–4 Hz, the residual velocity similarly shows a drop during the July 2019 paroxysm (Fig. 4.22d), though minimal change is evident following the 2014 flank eruption as previously observed. Otherwise, similar features in both 1–2 Hz and 2–4 Hz residual velocities include elevated values in 2016 followed by low velocities at the onset of activity 2017. The residual then increases again through 2018 with higher values recorded prior to a lava overflow in late-2018 and the July 2019 paroxysm.

Some similarities are observed between the corrected velocities using the average of all single-station (Fig 4.22) and the increasing velocities of STRA station specifically (Fig. 4.14. For example, the increasing trend from 2017 up until the July 2019

paroxysm is present in both cases, albeit starting from a velocity low in the corrected velocities for 1–2 and 2–4 Hz (Fig 4.22c-d). This may suggest that the process responsible for deviating velocities at STRA station is also present to a lesser extent in velocities recorded by the other stations. This is discussed further in Section 4.4.3. In general though, it is difficult to ascertain whether fluctuations in corrected velocity changes reflect volcanic processes or the simplicity of the empirical model used to compute pore pressure changes. Thus, we exercise caution interpreting the residual between observed and synthetic velocity changes predicated from precipitation data.

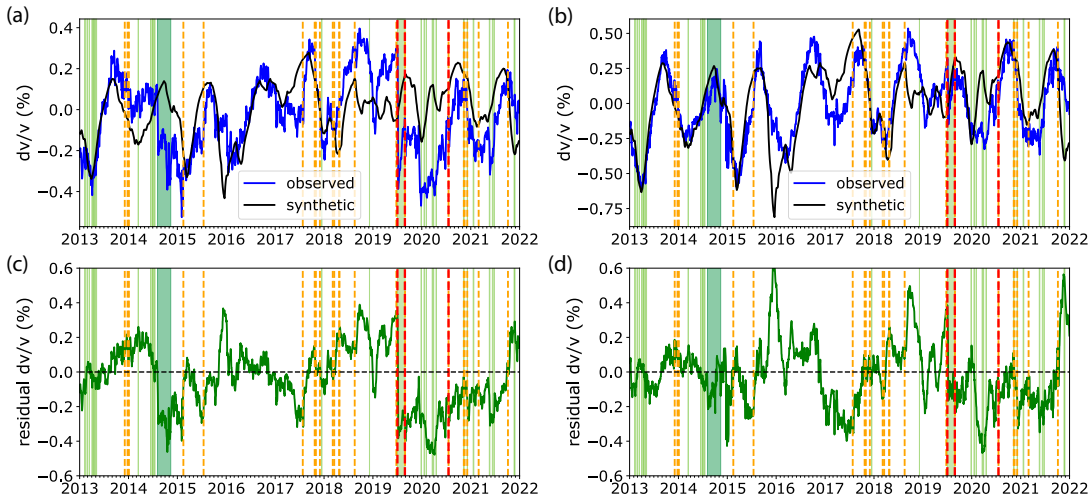

Figure 4.22: Comparison of observed single-station velocity changes and synthetic changes. (a) Comparison with 1–2 Hz velocity changes. (b) Comparison with 2–4 Hz velocity changes. (c) Residual for 1–2 Hz. (d) Residual for 2–4 Hz. Volcanic activity highlighted corresponding to descriptions in Fig. 4.2.

One simplification in the model relates the use of a constant diffusivity c to apply the correction over the full nine-year period. If this value were to vary over time, the synthetic velocity modeled may show a worse fit with observed velocities during certain time periods. This is explored further in Section S4.3, where we compute the optimal diffusivity within one-year moving windows (covering one cycle of the annual trend). This process is demonstrated in Figure S4.15, where the phase of the estimated pore pressures — smoothed by 180 days to reflect annual trend — is shown to be dependant on the optimal diffusivity. Further computation details can be found in Section S4.3. Computing this value over the nine-years shows four main spikes in the estimated diffusivity (Fig. S4.16). The first occurs in late-2016, returning to lower levels at the onset of activity in 2017. The second increase occurs in the months prior to an overflow event late-2018, returning to low levels following this event. A further increase is then observed a couple months prior to the 2019 paroxysms, remaining elevated for four months before decreasing following the August 2019 paroxysm. A final, smaller increase, is then observed in the early months of 2020.

It is possible that variations observed in the optimal diffusivity reflect changes in the medium. For example, in a pressurized system, increases in the density of micro-cracks could lead to greater permeability (Zoback and Byerlee, 1975; Delle

Piane et al., 2015; Heap and Kennedy, 2016) and thus increased hydraulic diffusivity (Talwani et al., 2007). However, for the same reasons previously discussed in relation to the simplicity of the model, we again exercise caution in this interpretation. For example, it could be argued that spikes in optimal diffusivity occur during periods where there a poorer fit is observed between measured and synthetic velocity changes (i.e. early-2017 and between 2018–2020, Fig. 4.22). Thus, we do not interpret these changes further, though consider further exploration of temporally evolving hydraulic diffusivity an interesting avenue for future exploration.

4.4.3 Deviating velocity at near-summit station (STRA)

Two time periods are identified where the difference between 2–4 Hz seismic velocities recorded at STRA and the average of the four other stations are consistently above zero. These are from mid–2013 until just before the 2014 flank eruption and from renewed activity in 2017 until the July 2019 paroxysm (Fig. 4.14). During other time periods, including immediately following the onset of the 2014 flank eruption and the July 2019 paroxysm, the velocity difference stays closer to zero. This is observed using both the stretching and moving-window cross-spectral technique (Fig. 4.15), albeit with slightly noisier results for the latter technique (Fig. S4.9). At 1–2 Hz, an increasing trend of velocity difference at STRA is also observed between 2017 and the July 2019 paroxysm (Fig. S4.11 and S4.11). However, the trend is less clearly distinguished from background fluctuations, with differences of similar or greater amplitude observed at other times during the nine year period (Fig. S4.10).

The position of STRA closer to the crater region should make it more sensitive to local changes around this part of the volcano. This can be demonstrated through the computation of 2D sensitivity kernels (e.g. Obermann et al., 2013; Planès et al., 2014; Budi-Santoso and Lesage, 2016a) using the analytical expression introduced by Pacheco and Snieder (2005) for the case of isotropic scattering of acoustic waves in the diffusion approximation following introduced by Pacheco and Snieder (2005) as:

$$K\left(s_1, s_2, x_0, t\right) = \frac{\int_0^t p\left(s_1, x_0, u\right) p\left(x_0, s_2, t-u\right) \mathrm{d}u}{p\left(s_1, s_2, t\right)} \tag{4.5}$$

where s_1 and s_2 are the positions of the station, x_0 the position of a velocity perturbation, t the center of the lag time window used. Thus $K\left(s_1, s_2, x_0, t\right)$ is the distribution of travel times of multiply scattered waves travelling from s_1 to s_2 after visiting x_0. For single-station kernels, we assign s_1 and s_2 as the same position (e.g. Caudron et al., 2022). The function $p(s_1, s_2, t)$ represents the time-dependant t intensity at position s_2 due to a unit intensity impulse at s_1. For short times and when perturbations are close to the source or receiver, this can be described using the solution to the radiative transfer equation for the case of isotropic scattering

media (Sato, 1993; Paasschens, 1997; Planès et al., 2014) following:

$$p(\boldsymbol{r},t) = \frac{e^{-ct\ell}}{2\pi r}\delta(ct-r) + \frac{1}{2\pi\ell ct}\left(1-\frac{r^2}{c^2t^2}\right)^{-\frac{1}{2}} \exp^{\left[\ell^{-1}\left(\sqrt{c^2t^2-r^2}-ct\right)\right]}\Theta(ct-r) \quad (4.6)$$

where r is the distance between s_1 and s_2, ℓ is the transport mean free path, c is the wave speed, δ the Dirac function, and Θ the Heaviside function. We set the transport mean free path ℓ to be 200 m, following estimates of Prudencio et al. (2015) at Stromboli, and 1 km/s for the wave speed c based on approximate maximum velocity of surface waves above 2 Hz estimated by Chouet et al. (1998). Sensitivity kernels corresponding to a central lag time $t = 10$s are shown for all station-pairs combined (Fig. 4.23a), single-stations combined (Fig. 4.23b), STRA station alone ((Fig. 4.23c), and the four stations IST3, ISTR, STR1, STRE (Fig. 4.23d). When combining the kernels for multiple source/receiver pairs, the individual kernels are summed. The final presented kernels are then normalized by dividing through by the maximum value.

From the kernels, we observe that the sensitivity for pairs of stations (Fig. 4.23a) and the single-stations (Fig. 4.23b) are highly similar, with far greater sensitivity to the North-Eastern portion of the edifice. This can explain the high similarity between the station-pair and single-station results at Stromboli (Fig. 4.10). There is a slightly greater sensitivity over the central and western portion of the edifice using station-pairs, however this is low relative to the north-eastern portion. Examining the kernel for STRA, we observe that there is a strong sensitivity to the crater region and the area just north-east of this (Fig. 4.23b). In comparison, the combined sensitivity of the other four stations has very little sensitivity to the crater area (Fig. 4.23d).

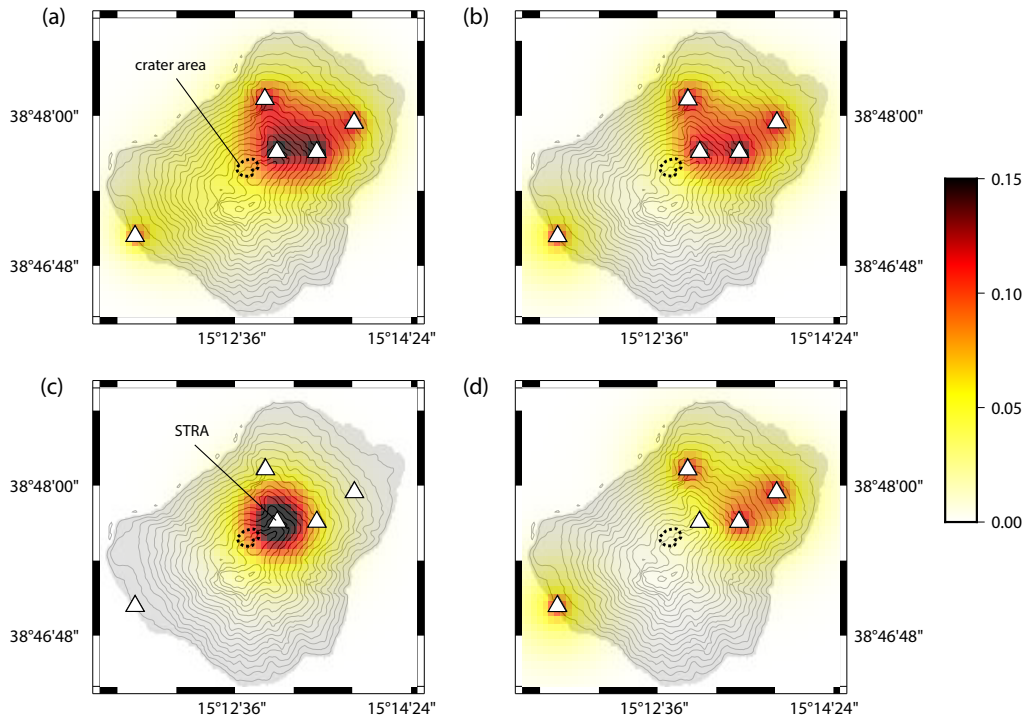

Figure 4.23: Lateral sensitivity kernels based on 10 s central lag time in the coda (a) combined sensitivity of station-pairs, (b) combined sensitivity of single-stations, (c) STRA station sensitivity, (d) combined sensitivity of IST3, ISTR, STR1, STRE stations.

The observation of a larger velocity difference at STRA station in the 2–4 Hz frequency band — relative to 1–2 Hz — suggests differences in the shallow subsurface are important. The hypothesis is supported when examining the relationship of measured velocities with lag time (Fig. S4.13). For example, the velocity changes measured at the four other stations show minimal variation when using later coda windows (Fig. S4.13b). For STRA station, however, there is a clear decrease in the measured velocity between 2017 and the July 2019 paroxysm later into the coda (Fig. S4.13a). Thus the difference in velocity between STRA and the other four stations also decreases at greater lag times (Fig. S4.13c). Two possible explanations for this are considered. The first relates to the greater overlap in sensitivity kernels with increasing lag time. For example, using a central lag-time of 20 s instead of 10 s shows that station STRA station becomes increasingly sensitive to velocities in the proximity of the other north-eastern stations (STR1, STRE, and IST3) (Fig. S4.14c). Similarly, the average of the other four stations begin to show slightly greater sensitivity to velocity changes in the crater region (Fig. S4.14d). Thus, it could be expected that the difference between stations will reduce with increasing lag time. Alternatively, it has been suggested that the contribution of body waves increases at later lag times (Obermann et al., 2013). The effect of this is to increase depth sensitivity later into the coda. Thus, the decreasing velocities measured at STRA station with lag time between 2017 and July 2019 imply the shallow subsurface around STRA is responsible for the velocity difference between this station and other stations. Note, the relationship between measured velocities and lag time is less clear for the period prior to the 2014 flank eruption at STRA (Fig. S4.13a).

4.4.4 Interpretation of velocity changes

The period between 2017–2019 was marked by a re-awakening phase of the volcano. From approximately May 2017, the volcano began to experience greater explosive activity, with increases in seismic signals such as volcanic tremor and Very Long Period (VLP) earthquakes (< 0.5 Hz, Giudicepietro et al., 2019). The period also coincided with significant increases of CO_2 soil flux, following a long-term slowly increasing trend going back to 2005 (Inguaggiato et al., 2017, 2019, 2020). In particular, the start of 2017 and late-2017/early-2018 (during the time period the volcanic alert level was raised from green to orange) were associated with significant increases in CO_2 soil flux (Inguaggiato et al., 2019, 2020; Giudicepietro et al., 2019). A further trend of increasing soil gas flux was identified starting in late-2018, again associated with increases in seismic signals associated with volcano dynamics and, as a consequence, an increase in the volcanic alert level (Inguaggiato et al., 2020; Giudicepietro et al., 2019). While the CO_2 flux was lower during this time period than the previous episodes in 2017 and early-2018, the rate of increase that continued up to the July 2019 paroxysm was greater than the overall trend in 2017–2018 (Inguaggiato et al., 2020). Increased CO_2 plume degassing was also recorded from late-2019, peaking just before the paroxysm (Aiuppa et al., 2021). Significant changes were similarly measured in VLP and tremor dynamics in the month prior to the onset of the July paroxysm (Giudicepietro et al., 2020). These were interpreted to reflect higher gas content in the Strombolian explosive activity.

The observation of increasing CO_2 flux suggests a greater contribution from deeper stored magma, owing to its reduced solubility relative to other volcanic gases (Giggenbach, 1996; Allard, 2010; Aiuppa et al., 2010). Thus, periods of increasing CO_2 flux, coinciding with heightened explosive activity at the volcano, have been linked to the input of new CO_2-rich magmatic fluids into the deeper storage zone (Inguaggiato et al., 2019, 2020, 2021; Aiuppa et al., 2021). This is also supported by an increased magmatic signature from low porphyritic (LP) magma in groundwater from mid-2017, increasing into 2018 (Federico et al., 2023). The depth of this storage zone (7–10 km) is far deeper than the main depth sensitivity of measurements in this study (generally greatest <1 km, Fig. S4.12). At the highest frequencies used (2–4 Hz), sensitivity is predominantly in the upper 200 m, with 1–2 Hz most sensitive within approximately 100-400 m, and 0.5–1.0 Hz within approximately 400–1000 m. Thus, we are certainly not sensitive to the LP storage zone. However, the relative increase in velocities at STRA station coinciding with increased explosive activity and CO_2 flux from 2017 up to the July 2019 paroxysm (Fig. 4.24b) suggest we are sensitive to the response of the shallow system to volatile input from depth. These changes also coincide with a long-term velocity decrease during the same time period at lower frequencies (0.5–1.0 Hz, Fig. 4.24a), shown clearly using the station-pair results which showed reduced sensitivity to seasonal fluctuations (Fig. 4.11a and 4.13d). Thus we consider various mechanisms that can explain these observations in response to increased volatile input at depth.

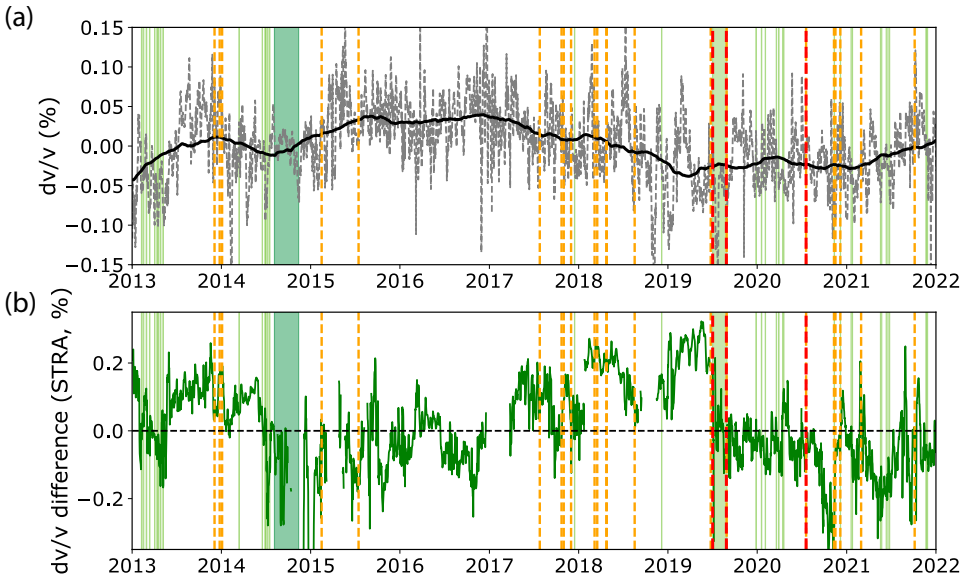

Figure 4.24: (a). Station-pair velocity change (stretching technique) at 0.5–1.0 Hz (grey-dashed line = 5-day stack, black-line = 365-day smoothing) compared with (b) velocity difference between STRA, 2–4 Hz, and average of the other four stations (IST3, ISTR, STR1, STRE).

Increased pressurization within the shallow plumbing system is one mechanism proposed to induce seismic velocity changes. Early studies at volcanoes identified velocity decreases in response to pressurization of the shallow plumbing system and inflation of the edifice (Brenguier et al., 2008c; Duputel et al., 2009; Obermann et al., 2013; Rivet et al., 2014; Bennington et al., 2015; Cubuk-Sabuncu et al., 2021b). This is interpreted to result from an increase in the density of microcracks and thus causing a velocity decrease (Nur and Simmons, 1969; Lockner et al., 1977; Brenguier et al., 2008c). However, some studies have identified velocity increases in response to pressurization. For example, velocity increases have been detected at Whakaari volcano (New Zealand) (Yates et al., 2019; Caudron et al., 2022), Ontake volcano (Japan) (Caudron et al., 2022), and Kilauea (Hawaii) (Donaldson et al., 2017; Hotovec-Ellis et al., 2022). In this case, the interpretation is the closure of pre-existing microcracks under increasing pressure, thus causing a velocity increase (Nur, 1971). This difference reflects one of the challenges associated with interpreting seismic velocity changes at volcanoes, where both increases and decreases in velocity are possible. One explanation relates to the transition between elastic and plastic deformation with increasing strain. At small strains, velocities increase as microcracks close under increasing stress (Nur, 1971), corresponding to the elastic regime. At larger strains, new cracks begin to develop and seismic velocities decrease (Nur and Simmons, 1969; Lockner et al., 1977), indicative of plastic deformation. Thus, the observation of a velocity increase or decrease may relate to the relative position of the seismic stations to the pressure source. This was considered a possible mechanism at Merapi (Indonesia), where velocities decreased in the upper part of the edifice while velocities increased in the lower part of the edifice (Budi-Santoso and Lesage, 2016a). Alternatively, the influence of the free surface, topography, and heterogeneity of volcanoes can induce both tensional and compressional stresses in

different places (Got et al., 2013; Budi-Santoso and Lesage, 2016a). Finally, it has been suggested that the mode of velocity change may be influenced by preferential crack orientation. This is suggested at Kilauea, where a sensitivity to radial strain, rather than volumetric strain, is inferred as the cause of a velocity increase during pressurization rather than a decrease (Hotovec-Ellis et al., 2022; Muzellec et al., 2023).

An increase in magma overpressure at Stromboli, related to volumetric expansion due to the influx of CO_2-rich gas (Apuani and Corazzato, 2009; Di Traglia et al., 2014; Inguaggiato et al., 2017), provides a plausible mechanism for inducing velocity changes. If this is accompanied by an increase in lava level within the conduit, an increase in the magmastatic pressure also acts on the conduit (Apuani and Corazzato, 2009; Casagli et al., 2009; Di Traglia et al., 2014; Calvari et al., 2022). This has been invoked to explain surface deformations within the summit area (Casagli et al., 2009; Di Traglia et al., 2013, 2014), with high CO_2 soil degassing also indicative of greater overpressure from new volatile-rich LP magma (Inguaggiato et al., 2011; Di Traglia et al., 2013; Inguaggiato et al., 2017, 2019). During the period between April 2015 to June 2019, Synthetic Aperture Radar Interferometry (InSAR) measurements recorded small inflations of the crater terrace. This primarily occurred between December 2017 – January 2018 and December 2018 – January 2019 (Schaefer et al., 2019; Di Traglia et al., 2021). Northward displacement of GPS station starting late-2016/early-2017 — located towards the northern-end of the island — also suggests modest inflation of the edifice occurred (Giudicepietro et al., 2019). Similarly, inflation was recorded in the months prior to the 2014 flank eruption (Di Traglia et al., 2015, 2018). Thus, there is evidence for increased magma overpressure acting on the conduit that could be responsible for observed velocity changes.

Long-term velocity changes measured at frequencies (0.5–1.0 Hz) are consistent with previous studies that identify a velocity decrease during pressurization of the shallow volcanic system (e.g. Brenguier et al., 2008c; Rivet et al., 2014; Bennington et al., 2015; Calò et al., 2021). In this scenario, the input of new, volatile-rich, magmatic fluid, creates an overpressure in the shallow system. For a spherical body, if the distribution of cracks in the surrounding rock is isotropic, a pressure increase induces tensile mean normal stresses (McTigue, 1987), promoting a seismic velocity decrease. Thus, we suggest that the decreasing trend starting from early-2017 at low frequencies captures increased volatile input into the shallow volcanic system at Stromboli, consistent with other observables during the same time period (e.g. Giudicepietro et al., 2019; Inguaggiato et al., 2019).

Regarding the velocity difference at STRA station (Fig. 4.24b), the timing of inflation events in the crater terrace area (late-2017/early-2018 and late-2018/early-2019) closely match periods of increasing velocity (Fig. 4.15b). For example, a change of approximately 0.2% is recorded between late-2017 and early-2018 (during the time period approximately represented by the cluster C3, Fig. 4.6). A similar increase is also observed from the start of 2019. However, it can also be argued that the periods of heightened activity themselves are associated with temporary velocity decreases across the different stations. For example, computing velocity changes using highly similar cross-correlation functions produced during heightened activity at 1–2 Hz shows decreasing velocities at all stations (Fig. 4.19). This is also consistent with

the results of Calò et al. (2021), who identified short-term velocity decreases across the whole edifice associated with periods of heightened activity within a similar frequency band. Thus an interpretation of increasing seismic velocities around the crater terrace due to increasing magma overpressure appears contrary to previous interpretations.

At Whakaari and Ontake volcanoes, gradual sealing of eruptive vents within the hydrothermal system, and subsequent pressure build-up, is inferred as reason for the very localised, increasing velocities, within the shallow hydrothermal system (Yates et al., 2019; Caudron et al., 2021, 2022). At Stromboli, the existence of a seal within the hydrothermal system — just west of the active craters (in the Fossetta) — has been identified based on low levels of soil CO_2 in this area (Finizola et al., 2002; Revil et al., 2011). However, there is no suggestion that significant buildup of pressure occurs beneath this seal, with structural boundaries considered to impede and re-direct fluids (Revil et al., 2004, 2023). Alternatively, it has been suggested that a crystal-rich, semi-permeable, plug exists in the upper several hundred meters of the conduit at Stromboli (Suckale et al., 2016; McKee et al., 2022). This can act as a mechanism for gas accumulation, where gas can become trapped in pore spaced between crystals within the plug or at the base of the plug (Belien et al., 2010; Oppenheimer et al., 2015, 2020). The build-up of overpressure due to gas accumulation and eventual overcoming of the plug yield strength is then thought to be the trigger for regular strombolian explosions (Suckale et al., 2016; Oppenheimer et al., 2020; McKee et al., 2022). However, it is not expected that significant overpressure can be sustained within the plug (Suckale et al., 2016). Thus it is difficult to explain a long-term pressure build up beneath a crystal plug within the upper conduit.

Using a velocity measurement of relative difference between stations is another consideration that should be taken into account in the interpretation. A deviating velocity at STRA station doesn't necessarily imply the source of the increase is limited to the crater area. For example, a velocity increase present at all stations, with a larger influence closer to the summit area, would show a similar effect. Similarly, a wider velocity decrease due to edifice inflation, with reduced influence on the crater terrace area, would also show a relative velocity increase at STRA station. Examining the velocity changes recorded by individual stations between January and July 2019 shows that the apparent velocities recorded by stations IST3, ISTR, STR1,and STRE are highly similar (Fig. 4.14). Thus, any interpretation of a velocity decrease (increase) over the wider edifice would require an almost identical velocity change at each of these stations, with reduced (greater) effect only within the crater terrace area.

One possibility could be differences in how the crater terrace area responds to pressurization relative to the rest of the volcanic edifice. The summit area of Stromboli contains many fractures (ring faults) associated with old collapse events (Finizola et al., 2009). Measurements of dilation direction based on structural features indicate a dilation direction perpendicular to these collapse scarps (Tibaldi et al., 2003). Radial strain may then be a more important factor than tensile strain in the response to pressurization near the crater terrace (Hotovec-Ellis et al., 2022; Muzellec et al., 2023). Thus, the velocity change at STRA station may be different relative to parts of the edifice with a more isotropic crack orientation. This could be

either a localised velocity increase (due to dominant sensitivity to radial compressive stresses), or simply a reduced sensitivity to tensile stresses. Numerical modeling of the volcanic system would be necessary to further test this hypothesis.

A final consideration is how a change in the ratio of gas might influence seismic velocity changes. Increased gas saturation are expected to decrease P-wave velocities, with limited influence on S-wave velocities (Nur and Simmons, 1969; Lumley, 2010; Clarke et al., 2020). Since the coda are primarily sensitive to shear-wave velocity (Snieder, 2006), it is not expected changes in the proportion of gas would have a significant influence. However, if pore water were to convert to steam — through the interaction of rising, high temperature, volatiles — an increase in pore pressures could be expected followed by a velocity decrease (through lowering of the effective stress) (Grêt et al., 2006; Caudron et al., 2015). It is therefore difficult to explain either a velocity increase close to the crater terrace (where high temperatures are observed (Revil et al., 2023)), or a velocity decrease at other parts of the edifice where temperatures are much lower (Revil et al., 2023). Thus we do not consider it a likely mechanism to explain the observed velocity changes.

These observations highlight the challenges in interpreting seismic velocity changes in volcanic regions, where the style of change can be highly variable in different locations on the edifice. We consider that increases in magma overpressure provide the most likely explanation for velocity decreases at lower frequencies. The observation of increasing velocities at STRA station relative to other stations may be explained by structural features in proximity to the crater terrace. Specifically, the presence of ring faults around the crater area may imply radial strain is more important than tensile strain towards modifying seismic velocities (e.g. Hotovec-Ellis et al., 2022; Muzellec et al., 2023). This may reduce the sensitivity of the crater terrace to tensile strain. Following this, the subsequent decreases at STRA relative to other stations following major eruptive events — such as the 2014 flank eruption and 2019 paroxysms — would represent a release of pressure within the upper part of the conduit. This interpretation should, however, clearly be explored in further detail, ideally with the use of modeling that can capture varied behaviour over different parts of the edifice.

An interesting observation is that the difference between velocities at STRA station and the other stations remains at approximately zero, or lower, following the 2019 paroxysms. This is despite significant volcanic activity continuing, with a further paroxysm recorded mid-2020 (albeit much smaller than the 2019 paroxysms) (Calvari et al., 2021). Similarly, CO_2 flux remained high following the 2019 paroxysms, with further increases from mid-2020 (Inguaggiato et al., 2021). This therefore contrasts with the 2017-2019 time period, where heightened activity resulted in higher relative velocities at STRA station. We interpret that this relates to significant changes in the shallow plumbing system following the 2019 paroxysms. For example, clustering cross-correlation reveals that the 2014 flank eruption and July 2019 paroxysm had the strongest influence on the character of cross-correlation functions relative to any other time period over the nine years (Fig. 4.5 and 4.6). Similarly, significant velocity decreases were observed at frequencies above 1 Hz associated with both events (largest in the 1-2 Hz range, Fig. 4.10b). Thus, we interpret that the 2019 paroxysms damaged the shallow conduit, with consistent activity over

the years following preventing the volcano returning to a 'background' state. This can also explain the lack of a significant velocity decrease associated with both the August 2019 and July 2020 paroxysms, as the system was already damaged. This observation is similar to velocity changes observed at Whakaari volcano, where only the first of two tectonic earthquakes in late-2016 induced a velocity decrease despite the second inducing stronger ground shaking (Yates et al., 2019). In contrast, clustering 0.5–1.0 Hz cross-correlation functions did not suggest large structural changes due to the paroxysms (Fig. 4.9). Thus, we consider that the deeper portion of the volcanic conduit (approximately 400–1000 m) was not significantly damaged by the paroxysms. The low velocities that continue through 2020–2022 are therefore interpreted to reflect the continued input of CO_2-rich volatiles that drive more significant eruptive activity during this time period.

4.4.5 Implications for monitoring

The results of this study highlight how coda wave interferometry can be a useful complementary tool towards monitoring the state of open conduit volcanoes. By examining the relative difference in velocities at STRA station, closest to the crater terrace, clear changes were revealed consistent with previous observations of unrest in the years preceding the 2019 paroxysms (e.g. increased CO_2 flux, VLP occurence, tremor). Similarly, a long-term velocity decrease was observed at lower frequencies in the build-up to the 2019 paroxysms. These changes are thought to reflect the interaction of the shallow plumbing system to increased volatile input at depth. Additionally, seismic velocity changes can reveal structural changes within the shallow system, with the results of clustering cross-correlation functions, combined with significant velocity decreases, suggesting major modifications following the 2014 flank eruption and July 2019 paroxysm. This could be useful towards anticipating whether a volcano might be entering a new phase of activity.

It is of interest that the largest difference in seismic velocities (2–4 Hz) recorded between STRA station and the other four stations occurs in the month prior to the July 2019 paroxysm. Similarly, the lowest velocities recorded at lower frequencies (0.5–1.0 Hz) using station-pair results — found to be less sensitive to seasonal changes — were between January and July 2019. It is widely agreed that paroxysms result from the fast ascent and injection of deeply stored volatile fluids (2–10 days prior) from the deeper LP storage system (7–10 km) into the shallow plumbing system (2–4 km) (Métrich et al., 2021; Voloschina et al., 2023). For this reason it is not anticipated that seismic velocity changes are sensitive to this process. However, changes in the shallow volcanic system could still be useful towards monitoring the likelihood of more dangerous events. For example, increased pressurization within the shallow volcanic system towards more critical values may indicate greater disequilibrium in the input and output of CO_2 rich gas at depth (Inguaggiato et al., 2019, 2020; Aiuppa et al., 2021). Maximum differences in the velocities recorded at STRA station in the months prior to the 2019 paroxysms may then be a reflection of a deeper plumbing system that is becoming increasingly unstable. However, it is unclear how much of a role this plays in the eventual triggering of the 2019 paroxysm, where it has also been suggested a blockage within the shallow conduit only a few minutes before the event may have been the triggering process (Viccaro et al.,

2021).

Finally, the detection of an anomalous velocity change in the shallow subsurface only by relative comparison between the station closest to the crater terrace with the other stations has important implications for monitoring shallow changes. Seasonal changes dominated the velocity changes above 1 Hz at all stations, with amplitudes of approximately ±0.5%. In comparison, the relative differences measured at STRA station — compared to other stations — varied by approximately ±0.2%. Thus, it remains a challenge to distinguish volcanic changes from seasonal changes, with the identification of an anomalous signal possible here only with the use of a station approximately 500 m from the crater terrace. These findings suggest two things towards monitoring volcanoes with coda wave interferometry: (1) the need for better tools for correcting seasonal trends and (2) the need for stations close to active vents in order to detect subtle changes within the shallow plumbing system. This will greatly improve the ability to detect anomalous behaviour that could suggest an increased probability of major events at the volcano.

4.5 Conclusions

Coda wave interferometry was applied towards the goal of computing seismic velocity changes over a nine-year period at Stromboli, encompassing various styles of volcanic activity. Particular attention is given to two paroxysms in 2019, which occurred with minimal warning approximately two months apart. The first of these (on 3 July) had a significant influence on the character of cross-correlation functions, as revealed through the application of hierarchical clustering. This was also true of a flank eruption in 2014. Clustering was also able to identify repeating clusters of highly similar cross-correlation functions that correspond to time periods where the volcanic alert level was raised. Thus clustering may be useful towards anticipating whether the volcano might be entering a new phase of activity.

Velocity changes were computed in three frequency bands (0.5–1.0 Hz, 1.0–2.0 Hz, and 2.0–4.0 Hz), between pairs of seismic stations and also between components of individual single seismic stations. Results were comparable between both datasets, with seasonal processes dominating the velocity changes (especially above 1 Hz) with amplitudes of approximately ±0.5 %. Precipitation-induced changes are considered the most likely contributing process, though temperature-induced changes cannot be ruled out. Two key features of interest are identified that are not readily explained by a seasonal process. The first feature is a long-term velocity decrease from early-2017 at lower frequencies (0.5–1.0 Hz), reaching a minimum in early-2019. Lower velocities are also recorded prior to the 2014 flank eruption and following the 2019 paroxysms. The second feature is an increasing velocity at the station closest to the summit (STRA) relative to the other stations at higher frequencies (primarily 2–4 Hz). This increasing trend begins in 2017 and reaches a maximum approximately one month before the July 3 paroxysm. Post-paroxysm, the difference between STRA station and other stations returns to approximately zero, considered to reflect a decrease of pressures within the shallow plumbing system post-paroxysm.

The long-term velocity decreases at lower frequencies from early–2017, combined with the increasing velocity at STRA station (relative to other stations) at higher frequencies, are aligned with previous observations of increasing activity during the same time period. This is thought to reflect the response of the shallow volcanic system to increased volatile input following the injection of new CO_2-rich magmatic fluids into the deeper storage zone. We consider that increases in magma overpressure in response to the input of new volatiles provides the most likely explanation for velocity decreases at lower frequencies. The increasing velocity at STRA station relative to other stations may be explained by anisotropy in the distribution of microcracks in the summit area due to the presence of ring faults. This interpretation should be explored further however.

The results highlight how coda wave interferometry can be a useful complementary tool towards monitoring the state of open conduit volcanoes. While sensitivity to deeper magma storage zones is not expected, increased pressurization within the shallow volcanic system towards more critical values may reflect an increasingly unstable deeper plumbing system. Thus, monitoring seismic velocity changes may be useful towards anticipating the likelihood of significant volcanic events while enhancing our understanding of the subsurface changes that precede such episodes.

4.6 Supplementary Material

S4.1 Methods

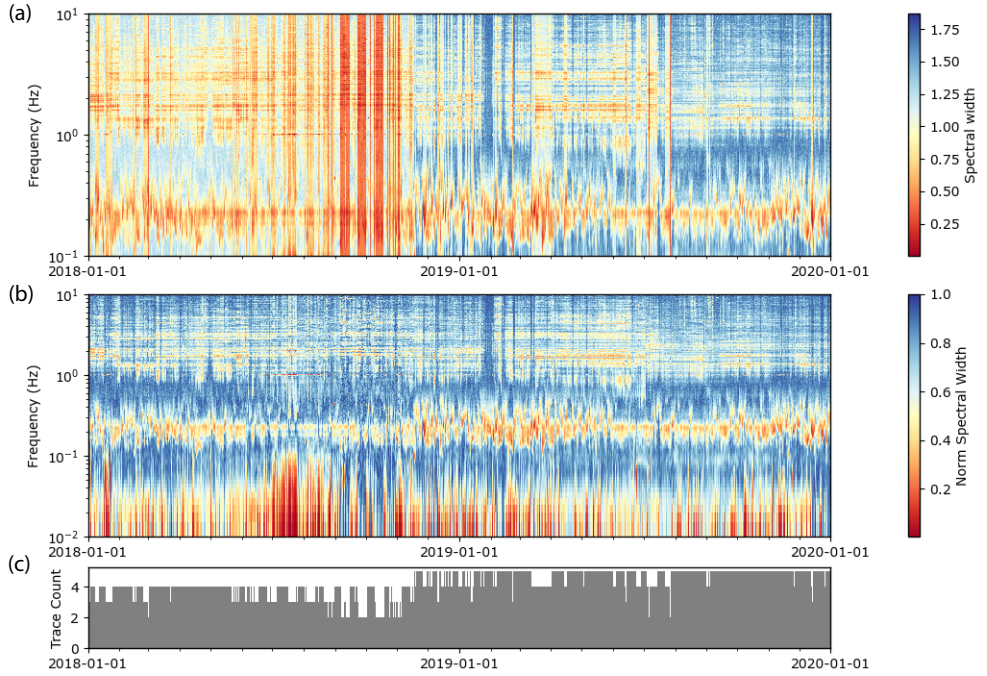

Figure S4.1: Comparison of spectral width measurements before and after normalizing by the maximum value. Normalization is used to reduce influence of variable station (trace) counts in computation. (a) Without normalization. (b) With normalization.

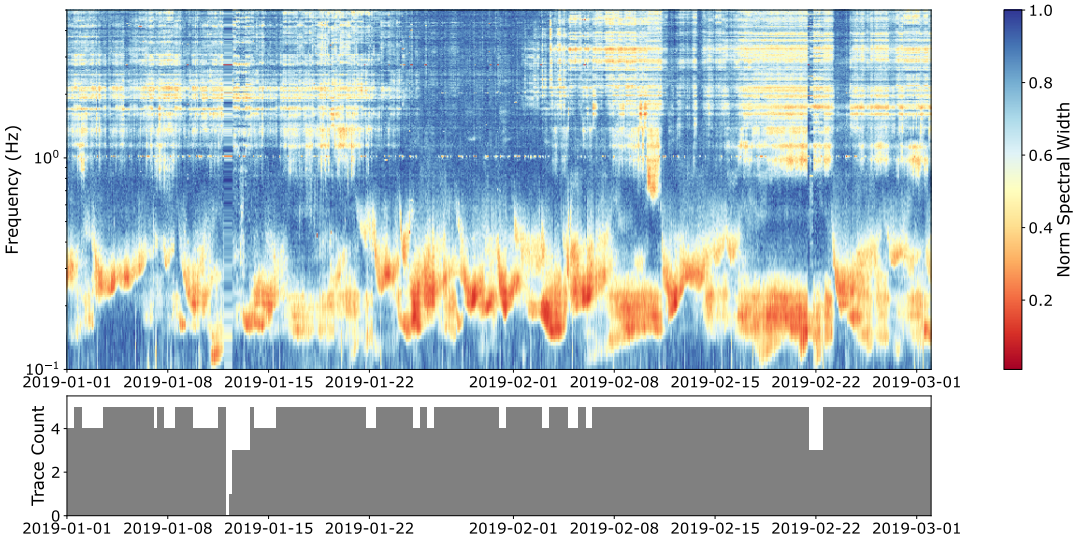

Figure S4.2: Zoom of spectral width measurements computed during two months in 2019.

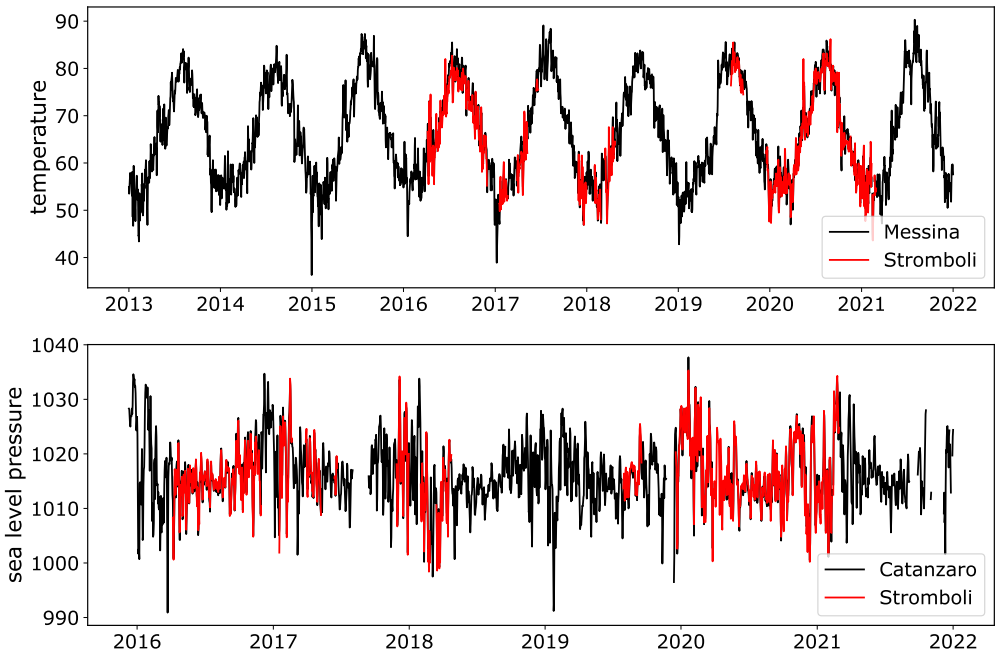

Figure S4.3: Comparison of temperature and sea level pressure measurements at Stromboli with those from other meteorological stations with improved data availability (Messina for temperature measurements and Catanzaro for sea level pressure).

S4.1.1 Cross-correlation characteristics

The two figures below show characteristics of cross-correlation functions computed for station-pair IST3-ISTR. Measures of SNR, phase similarity (or phase stack amplitude), and amplitude ratio are described in detail in Chapter two.

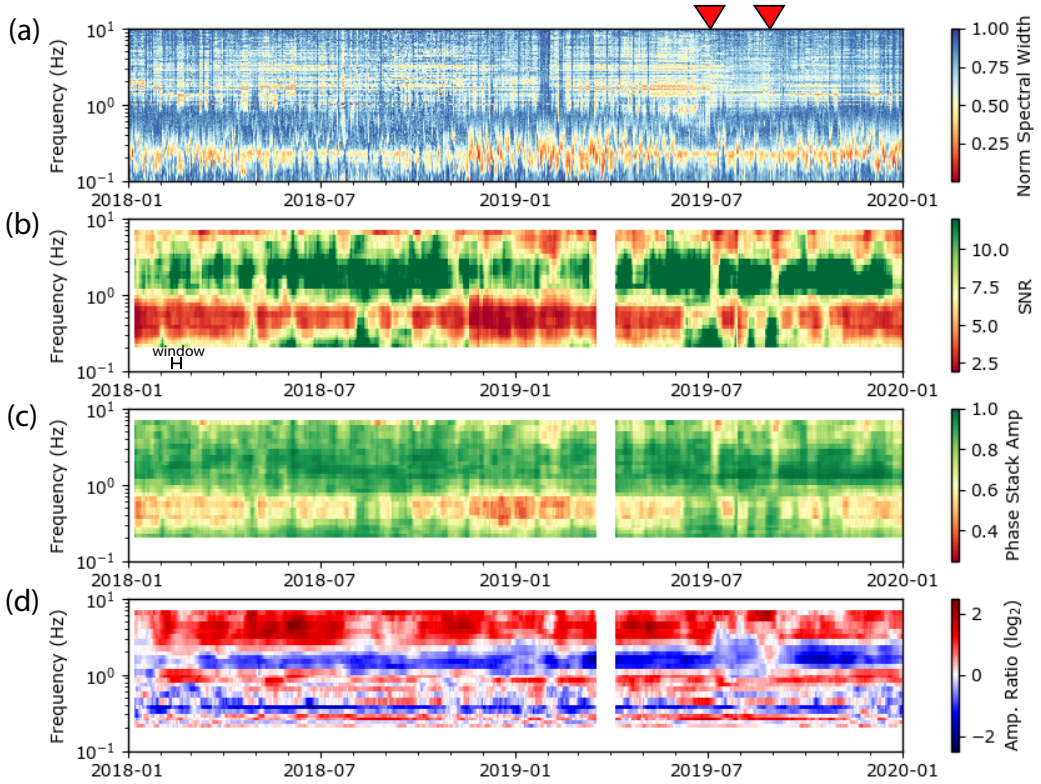

Figure S4.4: Measurements of cross-correlation characteristics at Stromboli for a single station-pair (IST3-ISTR), computed using 10-day moving window. For SNR and phase stack amplitude, the average value is taken between minimum lag defined by 0.5 km/s ballistic arrival speed and maximum lag equal to the minimum lag plus 15 cycles of central period of the bandpass filter. (a) Normalized covariance matrix spectral width as computed over full network. (b) SNR. (c) Phase stack amplitude. (d) Amplitude ratio.

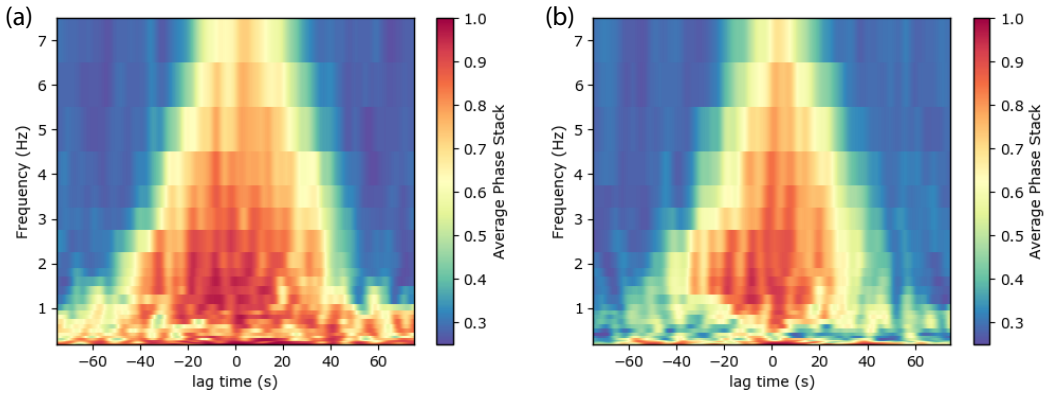

Figure S4.5: Phase similarity as a function of frequency and lag-time for station-pair (IST3-ISTR), as shown in Figure S4.4. Values computed as average between defined date ranges corresponding to (a) period with minimal energy below 1.0 Hz observed in spectral width measurements 2018-08-01 to 2018-09-01 (b) period with more energy below 1.0 Hz observed in spectral width measurements (2018-12-01 to 2019-01-01).

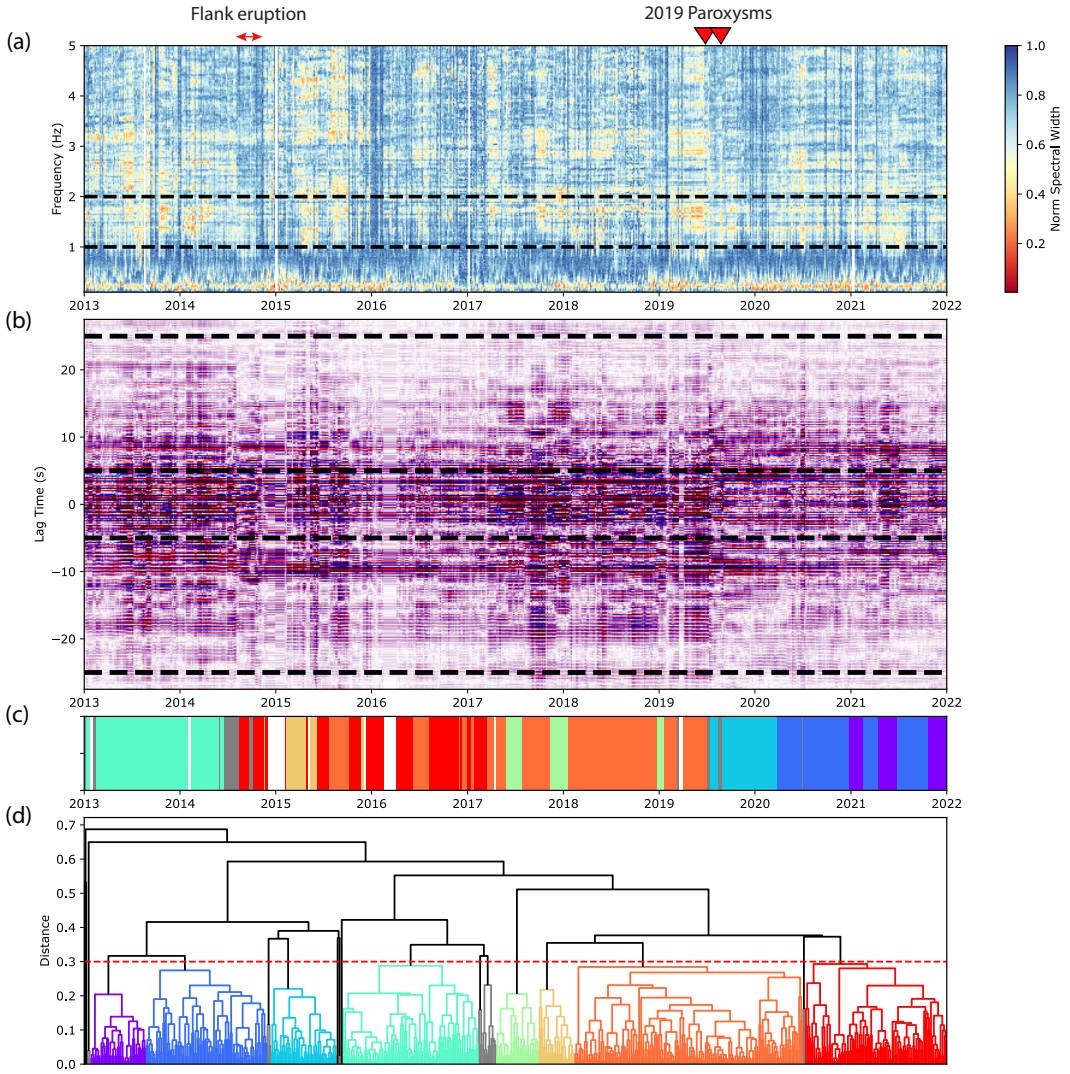

Figure S4.6: Hierarchical clustering output for 1.0-2.0 Hz cross-correlation functions (CCFs) computed for station-pair IST3-ISTR. (a) Normalized spectral width measurement. Dashed black lines show frequency range of CCFs. (b) Correlogram showing amplitudes of CCFs. Dashed-black lines show part of CCFs used in clustering. (c) Location of clusters in time, color-coded according to dendrogram output in (d). (d) Dendrogram, with clustered defined at distance threshold of 0.3 (dashed red-line).

S4.2 Results

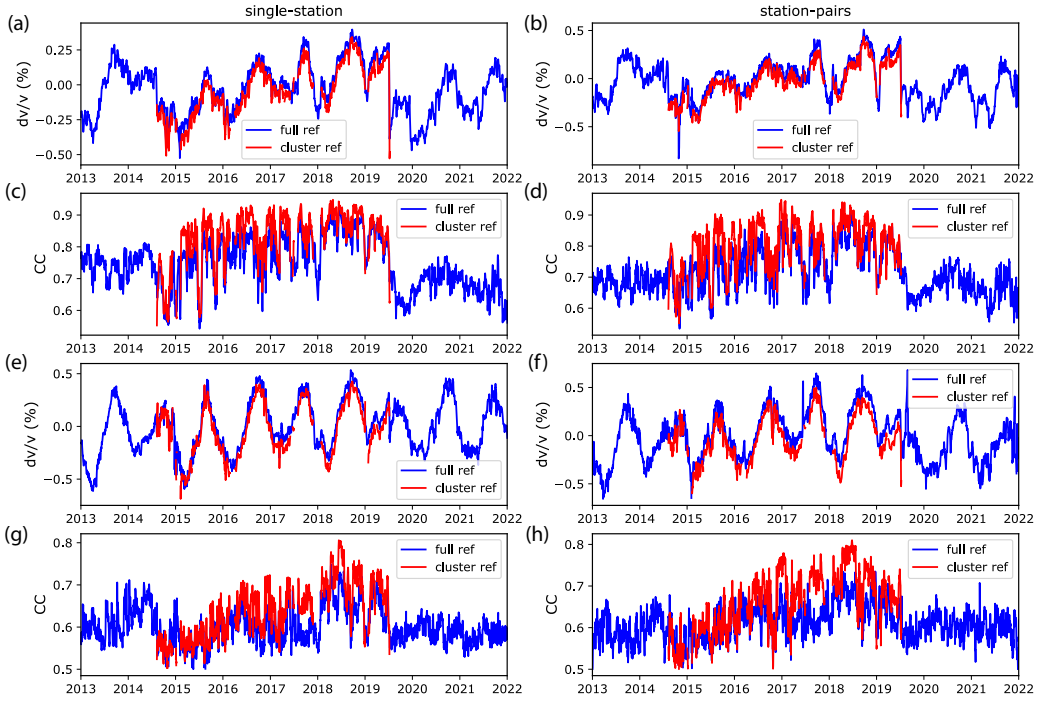

Figure S4.7: Comparison of velocity changes computed using full reference period versus reference period corresponding to main cluster between 2017–2019. (a) single-station velocity changes 1–2 Hz. (b) station-pair velocity changes 1–2 Hz. (c) single-station correlation coefficient 1–2 Hz. (d) station-pair correlation coefficient 1–2 Hz. (e) single-station velocity changes 2–4 Hz. (f) station-pair velocity changes 2–4 Hz. (g) single-station correlation coefficient 2–4 Hz. (h) station-pair correlation coefficient 2–4 Hz.

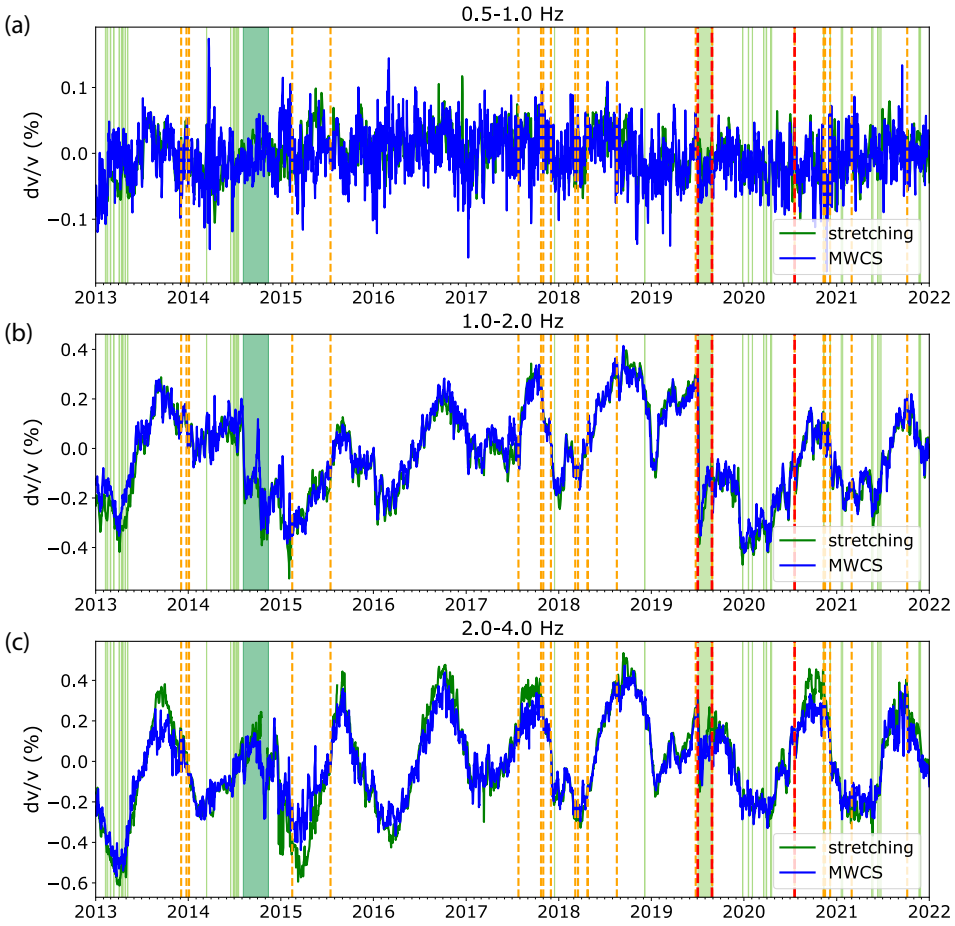

Figure S4.8: Velocity changes computed in three different frequency bands using moving-window cross-spectral technique. Single-station velocities (blue) correspond to the average of all single-station cross-components (EN, EZ, NZ) for all stations. Station-pair velocities (green) correspond to the average of all station-pairs (ZZ). Correlation coefficient (CC) between 5-day moving stacks and reference shown as grey shading. Volcanic activity highlighted corresponding to descriptions in Figure 4.2. (a) 0.5–1.0 Hz. (b) 1.0–2.0 Hz. (c) 2.0–4.0 Hz.

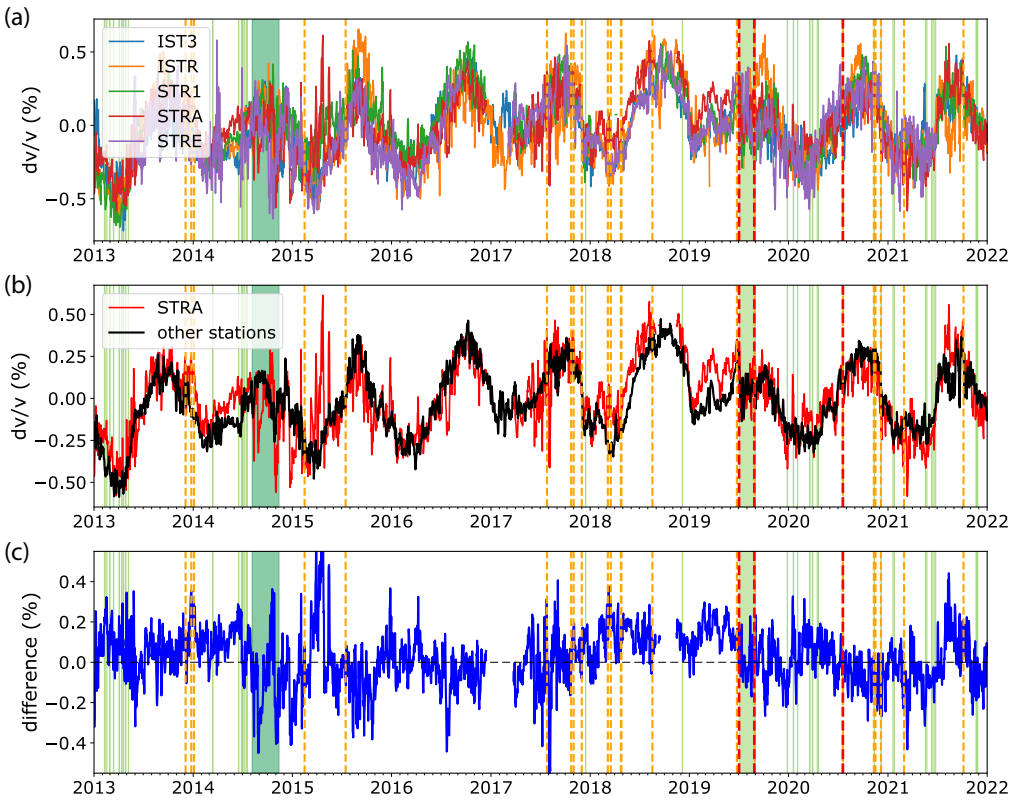

Figure S4.9: Comparing 2.0–4.0 Hz velocities (using moving-window cross-spectral technique) recorded at STRA station with other four stations. Single-station seismic velocities for all stations. (b) Velocities for STRA station compared with average of IST3, ISTR, STR1, and STRE stations. (c) Difference between average velocity of STRA and other four stations.

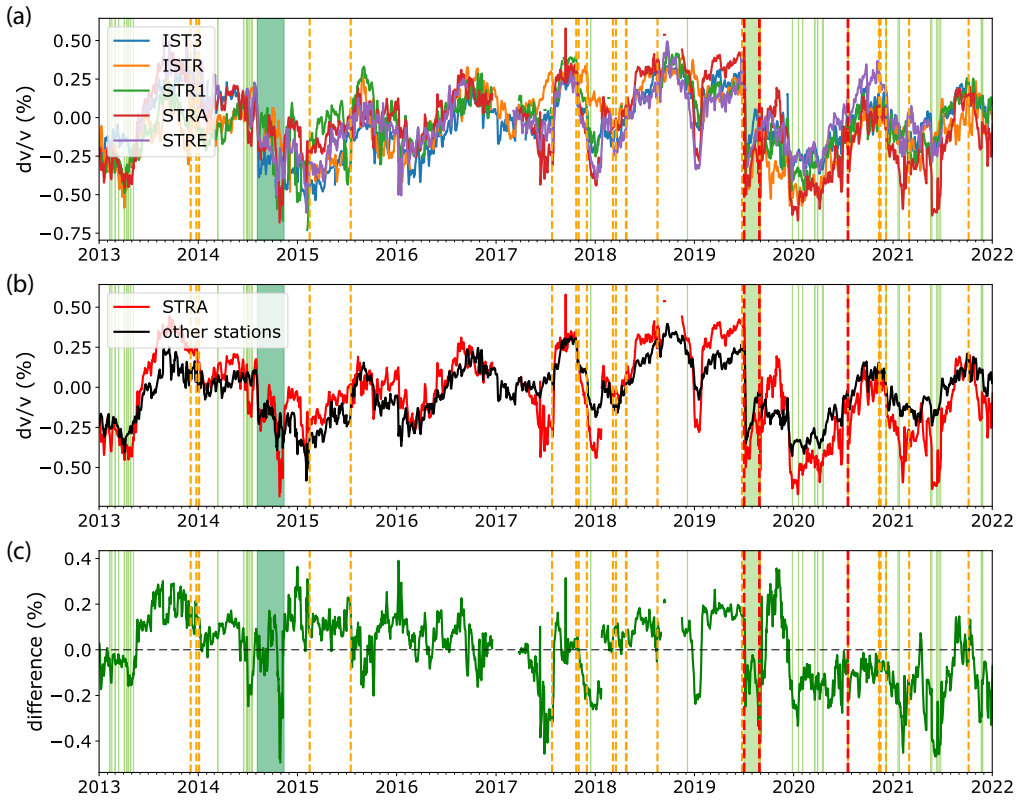

Figure S4.10: Comparing 1.0–2.0 Hz velocities (stretching technique) recorded at STRA station with other four stations. Volcanic activity highlighted corresponding to descriptions in Figure 4.2. (a) Single-station seismic velocities for all stations. (b) Velocities for STRA station compared with average of IST3, ISTR, STR1, and STRE stations. (c) Difference between average velocity of STRA and other four stations.

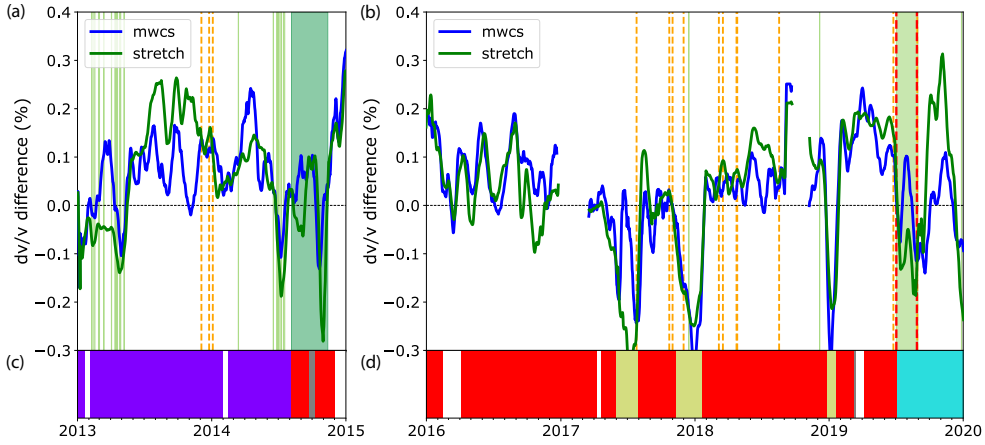

Figure S4.11: Closer look at relative difference between 1–2 Hz velocity changes at STRA station and other four stations (IST3, ISTR, STR1, STRE) between (a) 2013–2015 and (b) 2016–2020. Both the results from using the stretching and moving-window cross-spectral technique are shown. (c) Cluster solution for 1–2 Hz cross-correlation functions between 2013–2015. (d) Cluster solution for 1–2 Hz cross-correlation functions between 2016–2020. Volcanic activity highlighted corresponding to descriptions in Figure 4.2.

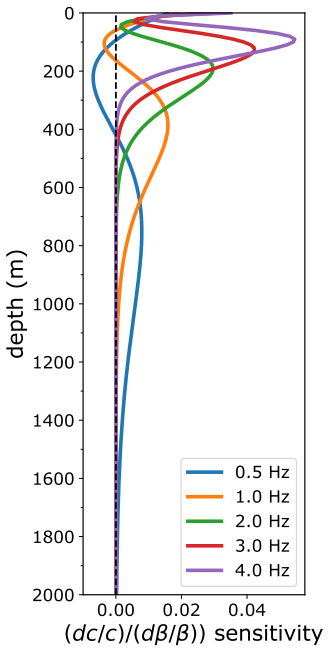

Figure S4.12: Sensitivity kernels for Rayleigh waves between 0.5–4.0 Hz. Velocity model consists of linearly increasing velocities from V_s = 670 m/s at the surface (corresponding to average shear-wave velocity in upper 100 m from Chouet et al. (1998)) to 2500 m/s at 2 km depth (corresponding approximately to the models of La Rocca et al. (2004)).

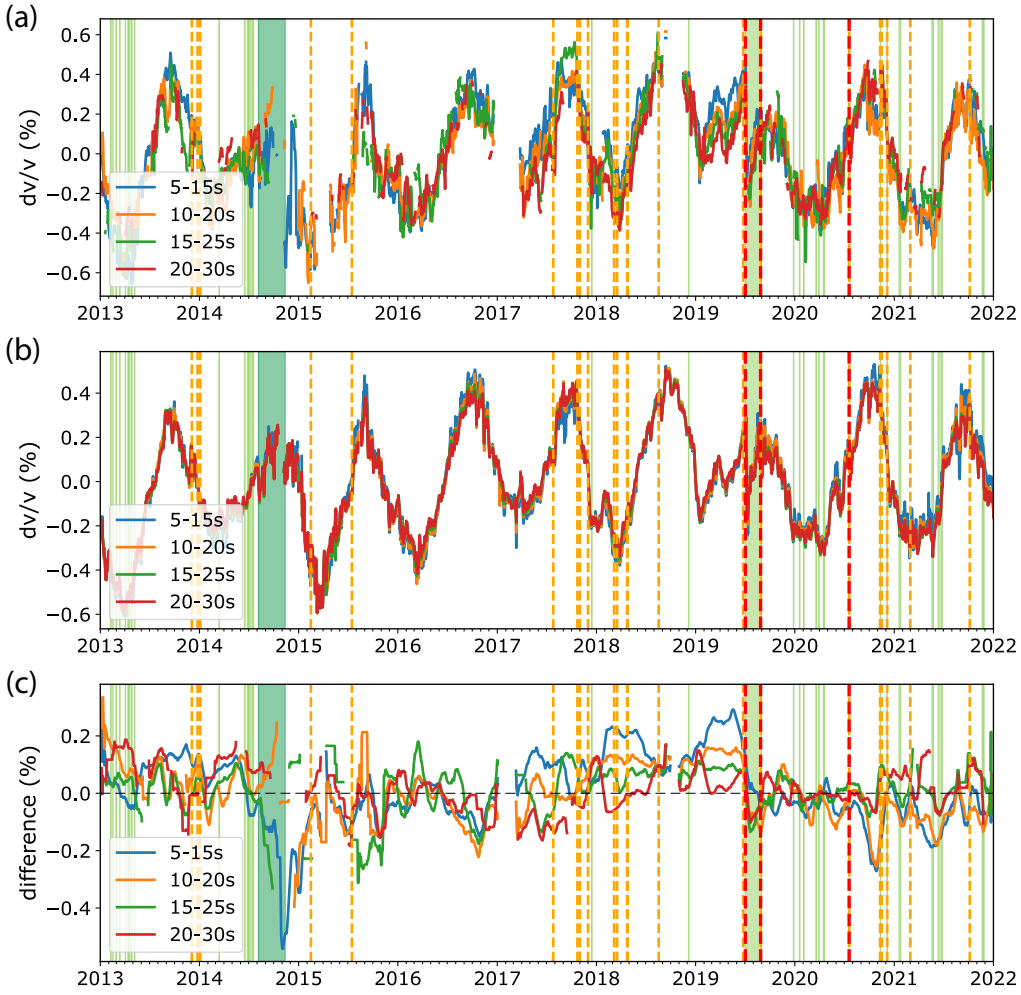

Figure S4.13: Comparing velocities measured at STRA station and all other stations at different lag times in the coda (a) Velocity changes at STRA station. (b) Average of velocity changes at IST3, ISTR, STR1, STRE stations. (c) difference between STRA and average of other stations.

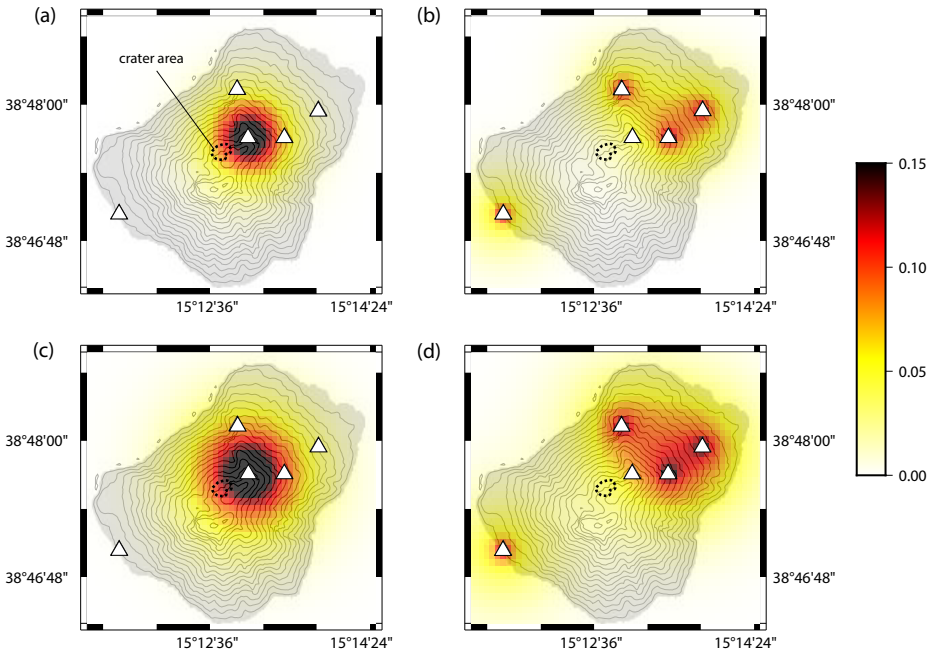

Figure S4.14: Comparison between lateral sensitivity kernels computed using 10 s central lag time and sensitivity kernels using 20 s central lag time. (a) STRA station sensitivity, 10 s central lag time. (b) combined sensitivity of IST3, ISTR, STR1, STRE stations, 10 s central lag time. (c) STRA station sensitivity, 20 s central lag time. (d) combined sensitivity of IST3, ISTR, STR1, STRE stations, 20 s central lag time.

S4.3 Variation of diffusivity through time

We consider whether the relationship between pore pressure changes and seismic velocities (at 2.0–4.0 Hz) are changing over time. Specifically, we consider whether the optimal diffusivity c varies in time. For this, instead of using a constant value (0.03) as before, we search for the optimal value within a one-year moving window (in 10-day steps), covering one cycle of the annual trend. This process is demonstrated in Figure S4.15, where the phase of the estimated pore pressures is shown to be dependant on the optimal diffusivity. Thus, the optimal value is sought – between 0.001 and 0.1 m^2/s – as the value that gives the maximum absolute amplitude in the cross-correlation between pore pressures and seismic velocities. The residual norm between de-meaned and normalized velocities and pore pressure changes is also computed.

Single-station velocity changes computed at 2.0–4.0 Hz are used for this analysis, noting that they appear most sensitive to environmental processes. This is confirmed through examining the spectra of velocity changes at each frequency, where a peak at approximately 365 days dominates 2.0–4.0 Hz velocity changes (Fig. 4.11). Furthermore, velocities at 2.0–4.0 Hz also appear less sensitive to co-eruptive decreases as observed for 1.0–2.0 Hz velocity changes. Still, in order to reduce the influence of

any velocity changes associated with major explosions and paroxysms, we attempt to correct for these changes using 1-day cross-correlation functions rather than 5-day functions. For each major explosion and paroxysm, the difference in velocity from the day of the event to the day following is subtracted from all subsequent days. Finally, we apply smoothing of 180-days to both velocities and computed pore pressures to focus on seasonal processes rather than short-term changes. The time series being compared are shown in Figure S4.16a. Both are de-meaned and de-trended within each moving window prior to the cross-correlation being computed.

Following this approach, we find that the estimated diffusivity is variable over the nine-years (Fig. S4.16b). Four main spikes in the diffusivity are identified. The first is late-2016, just prior to increased activity at the volcano in 2017. The diffusivity returns to lower levels at the onset of activity in 2017. The second increase occurs in the months prior to an overflow event late-2018, returning to low levels following this event. This return to low levels is short-lived, however, with a further increase a couple months prior to the 2019 paroxysms. The diffusivity remains elevated over the following four months before decreasing following the August 2019 paroxysm. A final, smaller increase, is then observed in the early months of 2020, before the diffusivity falls to low levels similar to those prior to late-2016.

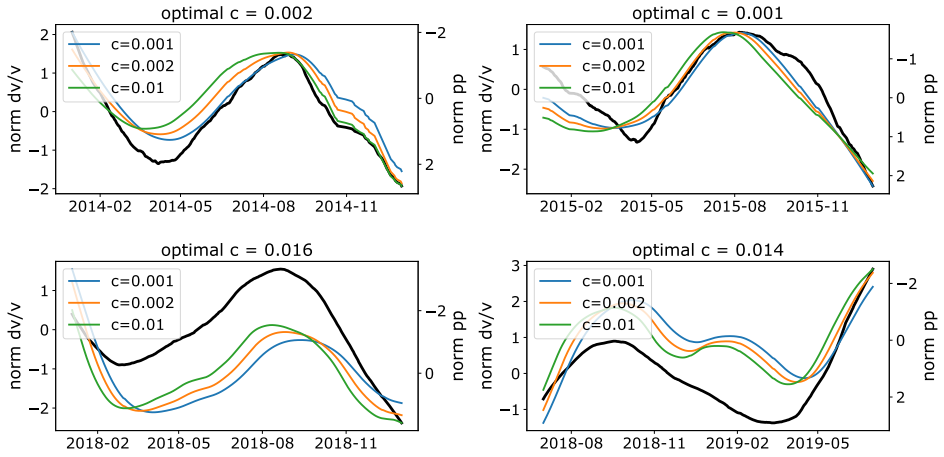

Figure S4.15: Normalized velocity changes (single-station, 2–4 Hz) during four different one-year windows (black-line) plotted with normalized pore pressure curves for three different values of diffusivity c. The optimal value estimated is included in the title for each window.

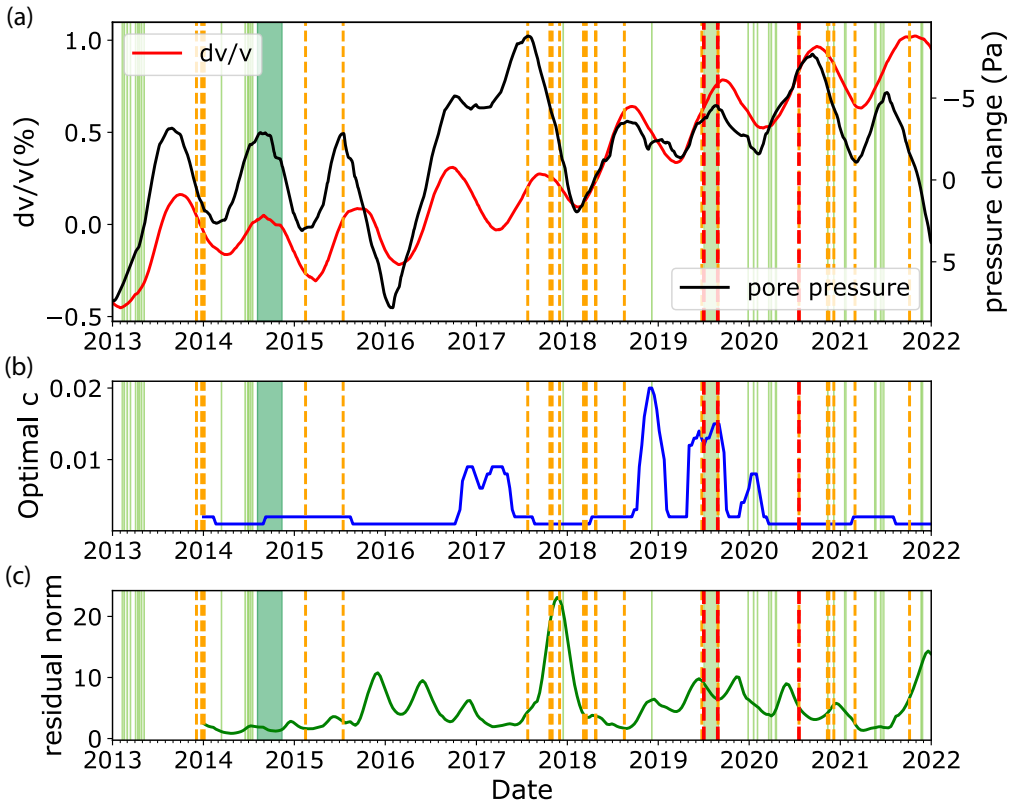

Figure S4.16: Diffusivity c estimated in one-year moving window comparing single-station velocity changes (2–4 Hz) with pore pressures computed using Eqn. 5.1. Volcanic activity highlighted corresponding to descriptions in Fig. 4.2. (a) Velocity changes (red-line), with 180-day smoothing, and pore pressure changes (black-line) computed using c=0.03 m^2/s. (b) optimal value of diffusivity c after cross-correlating velocities with pore pressures. (c) residual norm between de-meaned and normalized velocity changes and pore pressure changes.

Similar results are obtained using the velocity changes measured using station-pairs at 2.0–4.0 Hz , with the largest increases in diffusivity occurring prior to the overflow event late-2018 and pre-2019 paroxysms as with single-station results (Fig. S4.18). Furthermore, we test the influence of detrending within the 1-year moving window on our results by comparing them with the values of diffusivity obtained when applying no detrending and when detrending the full nine-year time series prior to windowing. This is done for both single-station (Fig. S4.17) and station-pairs (Fig. S4.18). In all cases, we find that the diffusivity is elevated prior to the 2019 paroxysms. However, when applying no de-trending, or de-trending the full time series, more significant spikes are observed at the start of 2017 and late 2018. These spikes are associated with high values of the residual norm computed between measured and modeled velocity changes, suggesting a poor fit between velocity changes and computed pore pressures. Similarly, both approaches are associated with higher values of residual norm over the full nine-year time period. Thus, we suggest that detrending within the moving window appears to provide more stable results.

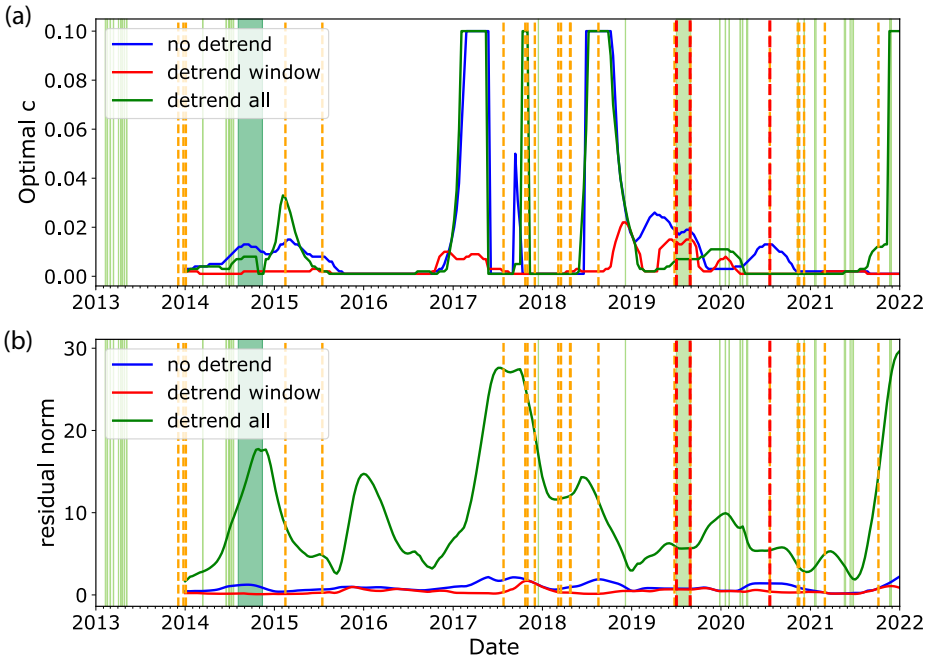

Figure S4.17: Comparison of the optimal diffusivity c and corresponding residual computed for 2–4 Hz single-station velocities when applying either no detrending to velocity changes and modeled pore pressures, or applying detrending within each one year moving window, or detrending full time series prior to using moving windows.

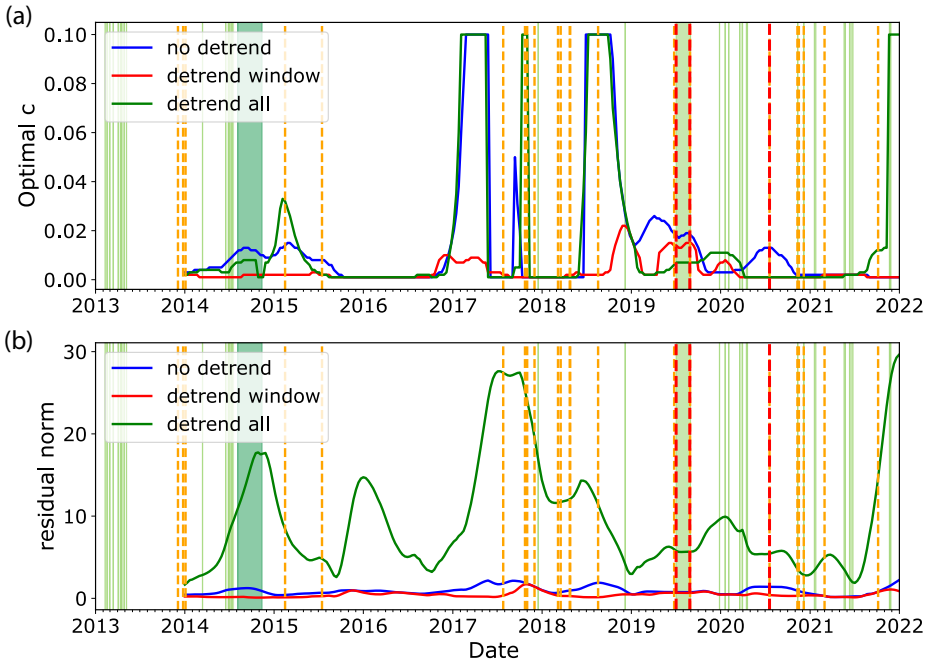

Figure S4.18: Comparison of the optimal diffusivity c and corresponding residual computed for 2–4 Hz station-pair velocities when applying either no detrending to velocity changes and modeled pore pressures, or applying detrending within each one year moving window, or detrending full time series prior to using moving windows.

Chapter 5

Exploring the influence of seasonal snow cycles on seismic velocity changes and eruptive activity at Mt Ruapehu, New Zealand

5.1 Introduction

It has long been recognised that volcanoes show sensitivity to external forcings. Tectonic activity (Yokoyama, 1971; Nishimura, 2017), tidal forces (Jaggar et al., 1924; Neuberg, 2000; Bredemeyer and Hansteen, 2014; Dinger et al., 2019), long-period sea-level changes (McNutt and Beavan, 1987; Mason et al., 2004), and precipitation (Matthews et al., 2002; Hort et al., 2003; Gaete et al., 2020; Farquharson and Amelung, 2020), for example, have all been linked to volcanic activity. A summary of external (and internal) eruption triggers can be found in the recent review of Caricchi et al. (2021).

Another external process that has been proposed to influence volcanic activity is the unloading of snow and ice. This can occur on longer time scales, through deglaciation (Sigvaldason et al., 1992; Jellinek et al., 2004; Sigmundsson et al., 2010; Conway et al., 2023), and also shorter time scales due to the unloading of seasonal snow/ice (Albino et al., 2010). Studies of this phenomena have largely relied on statistical analysis of the eruptive history, sometimes combined with modeling of pressure changes within the magma chamber (e.g. Jellinek et al., 2004; Albino et al., 2010; Sigmundsson et al., 2010), to assess the correlation between unloading and volcanic activity. With this in mind, techniques that allow us to probe the subsurface directly would enable us to better understand the underlying mechanisms behind supposed external triggering of eruptive activity.

The last two decades have seen great advances in our ability to monitor volcanic systems using coda wave interferometry. By cross-correlating seismic energy recorded by different sensors, we can monitor subtle changes in seismic velocities that can be related to a variety of geophysical processes. For example, velocity changes recorded

at volcanoes have been attributed to increased pressurization within the volcano system (Brenguier et al., 2008c; Mordret et al., 2010; Budi-Santoso and Lesage, 2016b; Yates et al., 2019; Caudron et al., 2021) and also changes in fluid content (Caudron et al., 2015, 2022). External factors are also regularly present in velocity changes recorded at volcanoes. These include meteorological processes, such as precipitation (e.g. Sens-Schönfelder and Wegler, 2006; Rivet et al., 2015; Caudron et al., 2022) and snow-loading/unloading (e.g. Hotovec-Ellis et al., 2014; Cannata et al., 2017; Donaldson et al., 2019; Caudron et al., 2022), and velocity decreases following large earthquakes (e.g. Battaglia et al., 2012; Lesage et al., 2014; Yates et al., 2019). For the purposes of monitoring volcanoes, velocity changes associated with environmental processes are often undesirable as they can mask changes associated with volcanic processes. Thus, several authors attempt to model the velocity changes produced by environmental effects, e.g. precipitation and snow fall, and correct for this (Rivet et al., 2015; Budi-Santoso and Lesage, 2016b; Wang et al., 2017). However, it is perhaps also important to examine in greater detail the velocity changes produced by these processes towards exploring the potential influence external processes may have on the volcanic system.

We explore further the relationship between snow cycles and seismic velocity changes at Mt Ruapehu volcano in New Zealand (Figure 5.1a). The stratovolcano (2797 m summit elevation) is located at the southern end of the Taupo Volcanic Zone and accumulates snow above approximately 1600 m between the months of April and November (Manville et al., 2000) (Figure 5.1b). Ruapehu has experienced many eruptions over the past century, with the most recent eruptions occurring in 2006 and 2007 (both phreatic). Interestingly, both of these eruptions occurred approximately one year apart (on 4 October 2006 and 27 September 2007) during the southern hemisphere spring. Furthermore, when we examine the eruption catalog at the volcano, we observe evidence of preferential timing of eruptions, with more eruptions during spring months September, October and November (Fig. 5.1c).

We take advantage of recent developments in passive seismic interferometry by applying the cross-wavelet transform method (Mao et al., 2020; Mordret et al., 2020) to compute single-station seismic velocity changes between 2005–2009. This approach provides greater time-frequency resolution compared to traditional approaches (Mao et al., 2020). Mordret et al. (2010) had previously computed velocity changes during the 2006–2007 time period, identifying a decrease in the days prior to the 2006 eruption. In this study, we target a broader frequency range than those used by Mordret et al. (2010), including frequencies influenced by volcanic tremor (where Yates et al. (2023) suggest that the tremor at Ruapehu can produce similar cross-correlation functions in time). In doing so, we can explore velocity changes at a shallower depth than previously investigated (and over a longer time period) with the aim of better understanding the effect of snow loading and unloading and whether this could play a role in preferential timing of eruptive activity at the volcano.

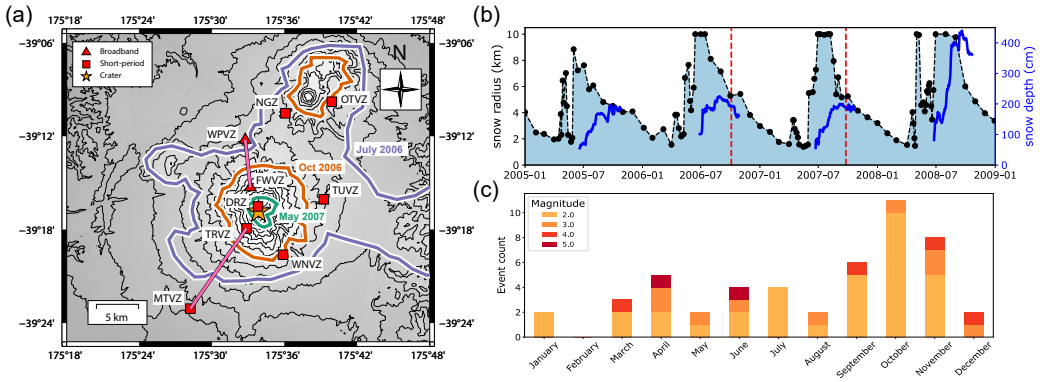

Figure 5.1: (a) Geological map of Mount Ruapehu, New Zealand. The snow cover as seen from MODIS satellite data is shown for the months July (2006), October (2006) and May (2007). (b) Snow radius against time, using area from snow cover assuming circular load, and average snow thickness (blue line) recorded at Whakapapa and Turoa ski stations (Ruapehu Alpine Lifts Ltd), located close to FWVZ and TRVZ stations respectively. (c) Eruption count per month of de-clustered catalog (using events of Scott (2013) after 1960, 1995/96 sequence removed, 14-day interval between events), magnitudes 2+). See section S5.3 for further details on event selection.

5.2 Data and methods

5.2.1 Seismic data and cross-correlation function preparation

We acquire three-component seismic data from a network of nine seismic stations in the proximity of the volcano (Figure 5.1a). These consist of seven short period sensors and two broadband sensors, with data recorded and publicaly available via the GeoNet data service (GeoNet, 2016). Since a single-station approach is used to compute seismic velocity changes, cross-correlating the different components of a single station, no instrument response correction is applied.

Cross-correlation functions are processed using the Python Package MSNoise (Lecocq et al., 2014) following relatively standard processing briefly outlined below. We first group data into continuous chunks, before demeaning and tapering, and merging into 1-d traces. The data are then bandpassed between 0.01 and 10 Hz, and decimated to 25 Hz. Data for each station component are then spectrally whitened between 0.1 and 8.0 Hz and clipped at three times the RMS. Cross-correlation functions are then computed between the different components of a single-station (EN, EZ, NZ) in 30-min windows and stacked linearly to give 1-d cross-correlation functions. We also compute vertical-vertical (ZZ) component cross-correlation functions for FWVZ-WPVZ and MTVZ-TRVZ station-pairs (Fig. 5.1a) to examine the influence of tremor in the following section (Section 5.2.2). Finally, 1-d functions are stacked to give 10-d cross-correlation functions, where the function recorded for a given day reflects the cross-correlation function of that day and the nine previous cross-correlation functions.

5.2.2 Tremor as a seismic source

One of the difficulties at volcanoes is the presence of volcanic tremor, usually dominant between 0.5–5.0 Hz (Chouet and Matoza, 2013) . Thus, it is often the case that studies focus on a frequency band below the band of volcanic tremor (e.g. Mordret et al., 2010; Cannata et al., 2017) unless it can be verified that the tremor is stable as a seismic source (e.g. Donaldson et al., 2017). At Ruapehu, tremor is common at frequencies above 1 Hz, usually with a single peak between 1.8 and 2.3 Hz (Latter and Balsillie, 1984; Hurst and Steven, 1993; Hurst, 1998; Sherburn et al., 1999). Such energy is clearly seen in Figure 5.2a, where we compute the spectral width of the network covariance matrix (e.g. Seydoux et al., 2016b,a) using data recorded by short-period stations closest to the volcano (DRZ, TRVZ, TUVZ, WNVZ, following the processing of Yates et al. (2023)). The spectral width relects the eigenvalue distribution of the covariance matrix and tends to zero when the seismic wavefield is dominated by a single source. Similarly, it approaches a maximum when non-coherent noise is considered (Seydoux et al., 2016a). Thus, reduced spectral width values indicate a more coherent seismic wavefield dominated by fewer seismic sources. Further details can be found in any number of the studies applying this methodology (e.g. Seydoux et al., 2016b,a; Soubestre et al., 2018; Journeau et al., 2020; Yates et al., 2023).

It has previously been suggested that volcanic tremor at Ruapehu may be useful as a repeating seismic source towards passive seismic interferometry (Yates et al., 2023). Using hierarchical clustering, it was shown that many cross-correlation functions produced within the frequency band of tremor (1.8–4.0 Hz) were relatively similar in the time-domain and their spectra (Yates et al., 2023). Targeting these higher frequencies would allow seismic velocity changes to be processed at higher frequencies and thus sample the shallower sub-surface. We revisit this by examining the cross-correlation functions, filtered between 1.0–4.0 Hz, recorded by the two pairs of stations shown in Figure 5.1a. These are approximately aligned radially with the summit region and, therefore, the location of tremor activity (Sherburn et al., 1999). Clear differences in amplitude can be seen between positive and negative lag times of cross-correlation functions for both pairs (Fig. S5.1), reflecting the propagation of volcanic tremor from the summit region and between the stations.

Following the same approach of Yates et al. (2023), we apply hierarchical clustering — using a Pearson correlation coefficient (CC) distance measure (one minus the absolute CC) and the average-linkage algorithm — to assess the similarity of the coda in cross-correlation functions through time. This can provide information on the stability of the background seismic sources towards monitoring seismic velocity changes (Yates et al., 2023; Makus et al., 2023). The results of clustering cross-correlation functions for MTVZ-TRVZ station are displayed in Figure 5.2, where we find that approximately fifty percent of cross-correlation functions can be grouped together at a distance threshold of 0.25 (in CLuster C5). The majority of these are from the time period before the 2007 eruption, with similar grouping also observed for station-pair FWVZ-WPVZ (albeit at a higher distance threshold; Figure S5.2). While the choice of the threshold is somewhat arbitrary, the use of a distance measure based on the correlation coefficient (i.e. $1 - CC$) provides an indication of how similar different clusters are. For example, any two sub-clusters within largest clus-

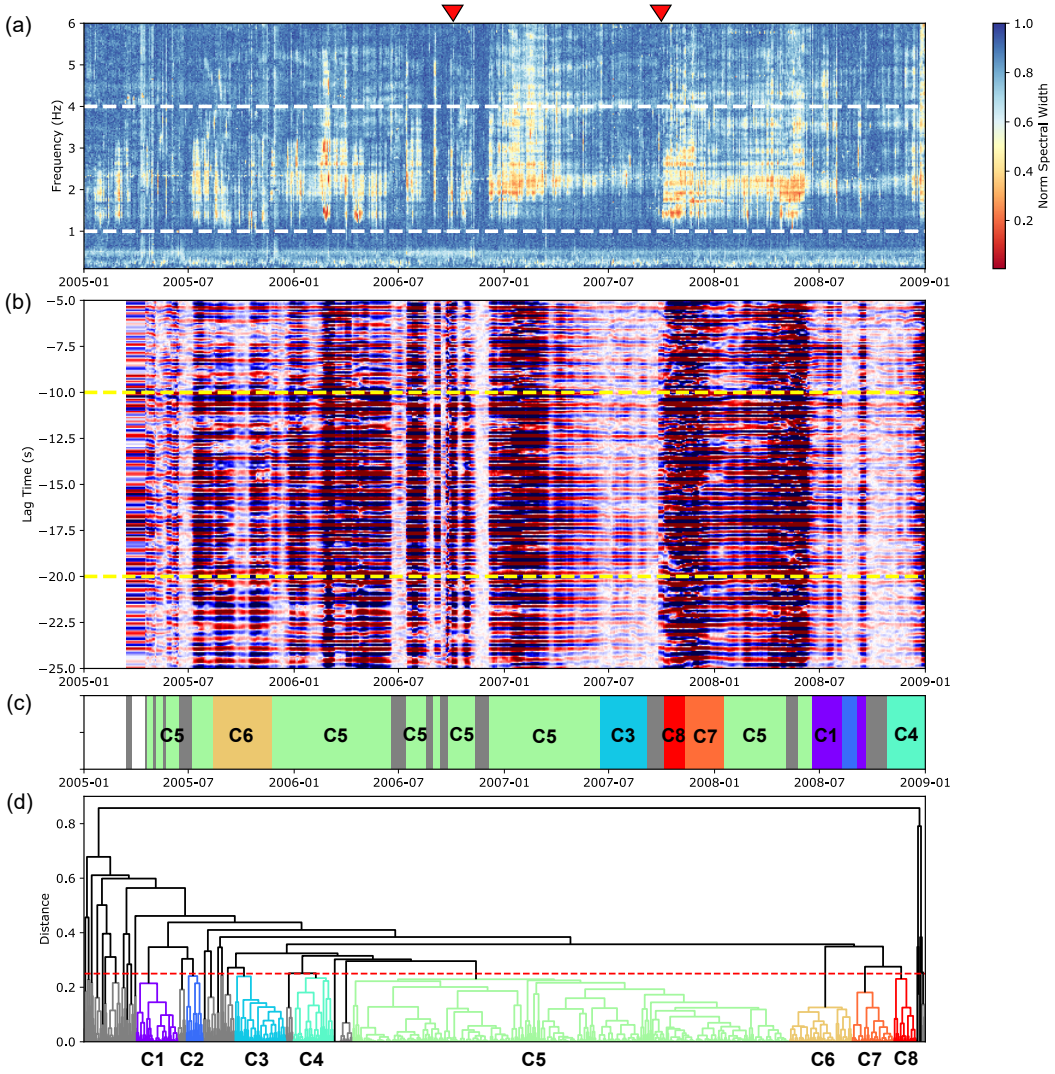

Figure 5.2: Hierarchical clustering output for 1.0-4.0 Hz cross-correlation functions (CCFs) computed for station-pair MTVZ-TRVZ. (a) Normalized spectral width measurement. Lower values indicate a more coherent seismic wavefield dominated by fewer seismic sources. Dashed white lines show frequency range of CCFs. (b) Correlogram showing amplitudes of CCFs at negative lag times which contain tremor signal. Dashed-yellow lines show part of CCFs used in clustering. (c) Location of clusters in time, color-coded according to dendrogram output in (d). Clusters with less than 25 cross-correlation functions are greyed out for easier comparison of larger clusters. (d) Dendrogram, with clustered defined at distance threshold of 0.25 (dashed red-line).

First Cluster	Second Cluster	1 - Distance	Correlation Coefficient	Difference
C1	C5	0.56	0.69	0.13
C1	C6	0.56	0.58	0.02
C1	C7	0.56	0.62	0.06
C1	C8	0.56	0.59	0.03
C5	C6	0.64	0.79	0.15
C5	C7	0.64	0.76	0.12
C5	C8	0.64	0.65	0.01
C6	C7	0.65	0.75	0.10
C6	C8	0.65	0.62	-0.03
C7	C8	0.72	0.81	0.09

Table 5.1: Comparison of correlation coefficients computed between different pairs of clusters identified in Figure 5.2 with the corresponding distance they would be agglomerated in the dendrogram. Averages of all individual cross-correlation functions within each cluster are used for computing the correlation coefficients (between -20 and -10 s as with clustering). Since the distance is computed as 1 minus the correlation coefficient, we show the value of 1 minus the distance for easier comparison.

ter shown in Figure 5.2d (C5) should have a correlation coefficient of approximately 0.75 or higher. This is demonstrated by computing the correlation coefficient between clusters C1, C5, C6, C7 and C8 (Table 5.1). In all cases, the true correlation coefficient between clusters is similar, or higher, than suggested by the distance they would be grouped in the dendrogram.

The observation that much of the data can be grouped at a distance threshold of 0.25 is encouraging for its use in computing velocity changes. However, multiple smaller clusters occur during key time periods towards analysing the influence of snow loading at Mt Ruapehu. For example, clusters C1, C3 and C6-8 all correspond to time periods where snow loading/unloading is expected. We examine the similarities and differences between these clusters and the larger C5 cluster in Figure 5.3, where we show the average of the cross-correlation functions within each cluster and their spectra. While the averaged cross-correlation functions in the time domain appear similar, differences in their spectral content are likely responsible for the different grouping. For example, while clusters C5 and C7 are both dominated by energy between 2.0–2.5 Hz, cluster C7 presents more energy at frequencies below 2 Hz relative to C5. Clusters C1 and C6, on the other hand, appear to consist of less energy from volcanic tremor relative to the larger C5 cluster. Thus, it is likely that temporal changes in the character of the volcanic tremor are responsible for these smaller clusters, requiring more careful handling for monitoring.

5.2.3 Measuring velocity changes using wavelet transform

In this study we apply the cross-wavelet transform approach, as applied by Mao et al. (2020). The method is based on wavelet cross-spectrum analysis (Farge, 1992; Torrence and Compo, 1998), and has the advantage in that it allows us to obtain delay time measurements between current and reference stacks of cross-correlation

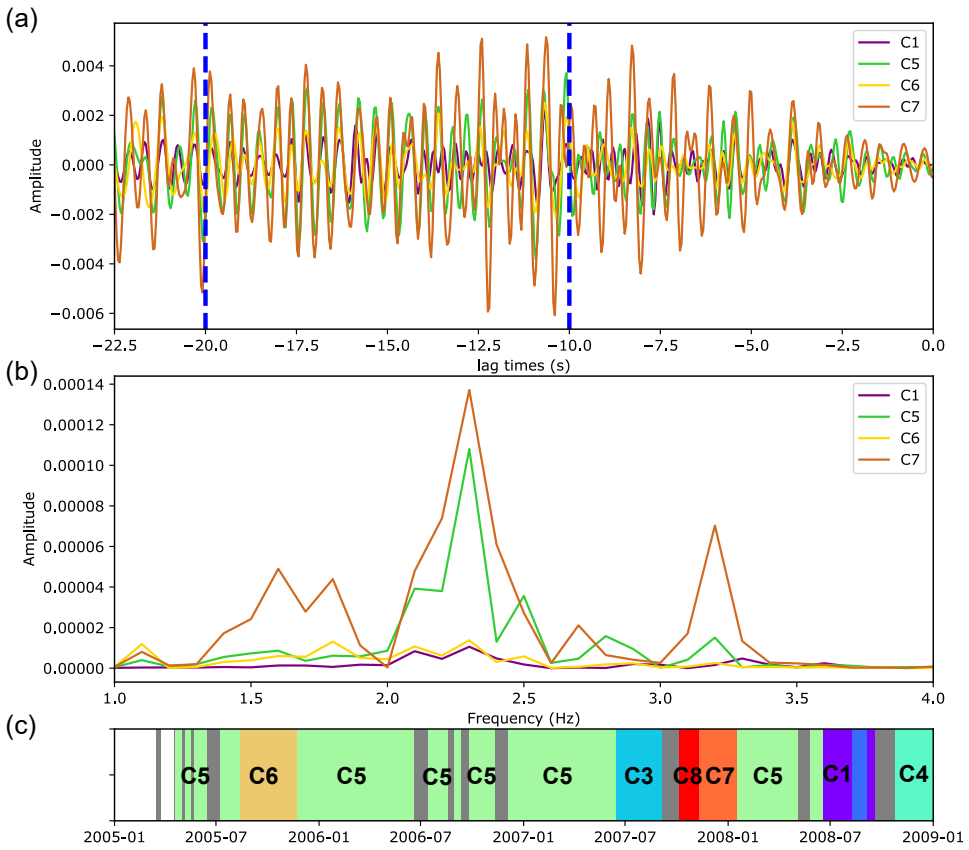

Figure 5.3: Comparison of the averaged cross-correlation functions for individual clusters identified in Figure 5.2. (a) Comparison in the time-domain. (b) Comparison of their spectra. (c) Location of clusters in time (from Figure 5.2c).

functions with a finer spectral resolution than can be obtained using traditional methods (e.g. the doublet method (Poupinet et al., 1984; Clarke et al., 2011) or the stretching method (Sens-Schönfelder and Wegler, 2006)) (Mao et al., 2020). Furthermore, the ability to compute phase differences between two waveforms in the time-frequency space allows more precise quality control criteria. Thus, where it might ordinarily be necessary to discard entire cross-correlation functions in the presence of a changing seismic source, the wavelet transform allows for the exclusion of specific frequencies in time while retaining reliable data. Such flexibility makes it an ideal choice for processing seismic velocity changes at Ruapehu, where subtle changes in the seismic source may influence the corresponding velocity change measurements.

Velocity changes are recovered by measuring the travel time shifts between the 10-d (current) stacks and a reference stack of all available cross-correlation functions. Following Mao et al. (2020), the phase of both the reference and current stacks are obtained in the time-frequency domain by performing convolutions between the cross-correlation functions and a series of scaled and translated versions of a mother wavelet. Here, we use the Morlet wavelet (Morlet et al., 1982b,a) for the mother wavelet and compute the phase between 0.1 – 8.0 Hz. The phase difference between the current and reference stack can then be computed, and measurements of delay time recovered by dividing by $2\pi f$ at each frequency f (Figure 5.4a). We then obtain a measurement of velocity change by performing a weighted linear regression of delay measurements against lag time within a chosen coda window, where $dt/t = -dv/v$. The error in the slope, and thus dv/v, is computed as the square root of the diagonal element of the covariance matrix resulting from the regression (based on Scipy function scipy.optimize.curve_fit (Virtanen et al., 2020)). This gives the standard deviation of the slope estimate. We define the minimum lag time for the coda as 5 seconds and define the length of the coda window to be twenty cycles at a given frequency (up to a maximum of 120 seconds for the very lowest frequencies). The weighting function (Figure 5.4d) used in the regression is based on the amplitude of the wavelet cross-spectrum (shown in Figure 5.4c), which reflects the local power of both the current and reference stack (Mao et al., 2020). Specifically, we normalize the wavelet cross-spectrum by dividing through by the maximum value the time-frequency domain to give weights that vary between 0 and 1 (following Mao et al. (2020)).

In addition to weighting the linear regression based on the amplitude of the wavelet cross-spectrum, we apply additional control quality criteria to ensure that only reliable measurements of velocity change are recorded. Delay time measures with a wavelet coherence less than 0.5, or maximum delay time measurement of 0.3, are assigned a weight of zero. Furthermore, we require that at least 25% of the delay time measurements have a non-zero weight, to avoid spurious velocity change estimates based on only a few points. Different values for the minimum wavelet coherence were tested (up to 0.8), with minimal influence on the main observations of this study (Fig. 5.17). Thus, results are presented as computed using a minimum coherence of 0.5. The influence of wavelet coherence threshold is discussed further in Section 5.4.1.

As mentioned, an advantage of applying the wavelet transform is the ability to

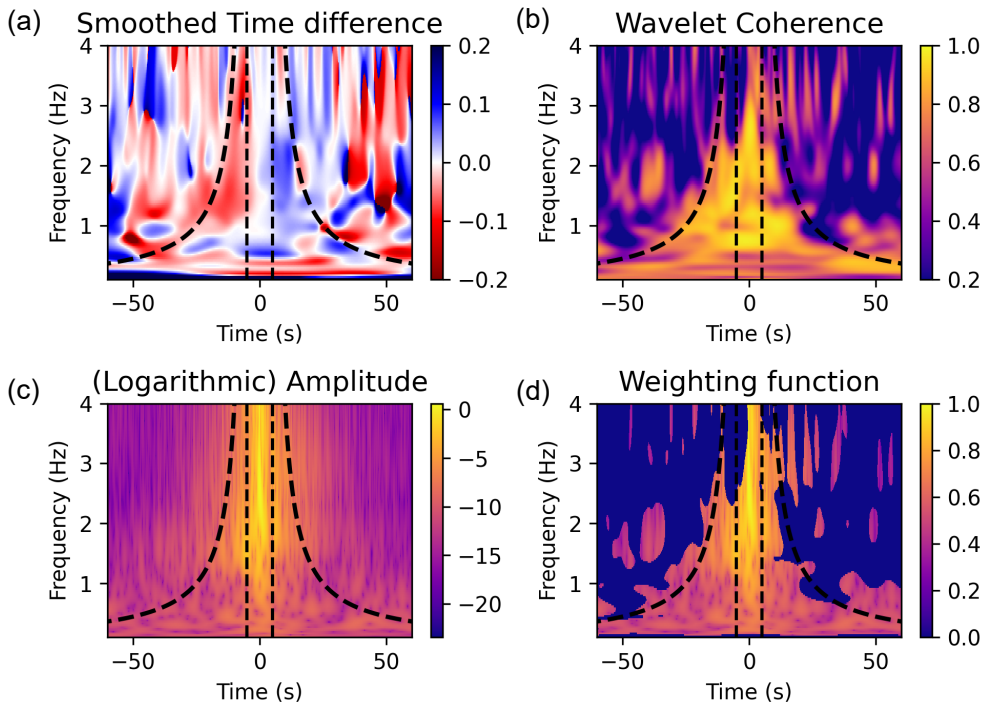

Figure 5.4: Example of delay-time measurements using wavelet transform for a cross-correlation function during coherent tremor (2007-03-01) at DRZ station (EN components) (a) Time-delay measurements in time-frequency space. (b) Coherence between the current and reference cross-correlation function. (c) Amplitude of wavelet cross-spectrum used in determining the weighting function for linear-regression. (d) Weighting function.

retain reliable measurements at certain frequencies while discarding unreliable measurements that relate to varying seismic sources. For example, during time periods where there is minimal tremor signal, the wavelet coherence and weighting function have very low values above 1 Hz (Figure S5.3). In this case, the measurements below 1 Hz would be retained and those above 1 Hz discarded. Similarly, we can observe that sometimes the coherence is reduced at only certain frequencies (Figure S5.4), likely corresponding to changes in the tremor source. In Figure S5.4)b, for example, the delay time measurements for a cross-correlation function computed using data from one of the smaller cluster periods are shown in Figure 5.3 (C3). We observe that the weighting assigned to the majority of measurements above approximately 1.5 Hz have close to zero-weighting. This is consistent with the observation that there appears to be increased energy from tremor above approximately 1.4 Hz during this time period (Fig. 5.3b) relative to the longer-term average (cluster C5). Thus, in this case, measurements in the frequency band where changes in the source characteristics are occurring can be excluded while keeping the coherent parts of the cross-correlation function.

5.2.4 Complementary datasets

Snow coverage, as shown in Figure 5.1a for three different time periods, is estimated from MODIS satellite data between 2005–2009. After picking the boundary of the snow around Ruapehu, the area is converted to a radius assuming a circular load (Fig. 5.1b). This is clearly not the case in reality, but serves as a simple approximation to quantify the coverage. The average number of days between measurements is 15-days, with more regular picks made at the start of winter to capture early-winter snow fall. The radius of snow for each year, and their average, is shown in Figure 5.5a. This shows that the first snow-fall often begins in April/May, with more extensive coverage from June/July.

Snow thicknesses are recorded at Whakapapa and Turoa ski stations during times of the year the ski stations are operational (data provided by Ruapehu Alpine Lifts Ltd). The Whakapapa measurement is made approximately 1 km north-east of FWVZ station, at approximately 2000 m elevation. The Turoa measurement is made in close proximity to TRVZ station, also at approximately 2000 m elevation. The trends recorded at both sites are comparable throughout the 2005–2009 period, with snow thicknesses recorded at Turoa having a slightly higher average per month between 2005–2009 (Fig. 5.5b).

Daily precipitation records obtained from New Zealand's National Climate Database (https://cliflo.niwa.co.nz/) are used to compute the pore pressure changes associated with fluid diffusion after rainfall events, following Wang et al. (2017). Based on the one-dimensional fully coupled diffusion question of Talwani et al. (2007), and applied by Rivet et al. (2015), pore pressure changes ($P(r,t)$) are computed at a distance r and time t, following:

$$P(r,t) = \sum_{i=1}^{n} \delta p_i \operatorname{erfc}[r/(4c(n-i)\delta t)^{1/2}] \tag{5.1}$$

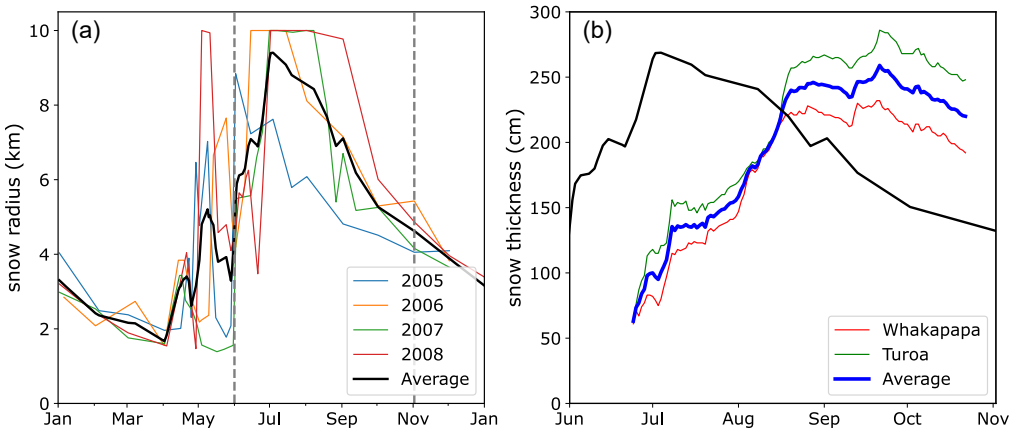

Figure 5.5: Snow coverage and snow thickness datasets plotted against months of the year. (a) Snow coverage for individual years between 2005–2009. Vertical grey dashed-lines mark months displayed in (b) Average snow thicknesses recorded by Whakapapa and Turoa ski stations between 2005–2009. Black line corresponds to snow radius in (a).

where r refers to the distance (treated as the depth), n the total number of time increments δt between the start of the time series (i=1) and time t, δp_i is the precipitation load change computed as $\rho g \delta h_i$, where δh_i is the change in precipitation recorded at time t_i from the previous day (t_{i-1}), and c is the diffusion rate in m^2/s. We use $c = 1m^2/s$, following Wang et al. (2017) and Caudron et al. (2022), and compute the mean pore pressure between $r = 500$m and $r = 4000$m. Since the pore pressure changes are computed based on an accumulation of load changes, based on the daily change in precipitation (h_i), it is ideal to use precipitation data beginning before the time period of interest to allow the computed pore pressure changes to converge around a given mean. We acquire precipitation data starting 2001 towards this goal, recorded at a station located in close proximity to WPVZ station. The final calculated mean pore pressure changes between 2005–2009 are shown in Figure 5.6.

5.3 Results

5.3.1 Measured velocity changes

Velocity changes recorded at three different stations are shown in Figure 5.7. Two of these stations, DRZ and FWVZ (Figs 5.7a and 5.7b respectively), correspond to the stations closest to the summit of the volcano and thus are expected to have significant snow cover during winter months. The third station, MTVZ (Figure 5.7c), is located at a greater distance from the volcano, where snow cover throughout the year is minimal (Figure 5.1a).

At all three stations, alternating patterns of low and high velocities dominate, with

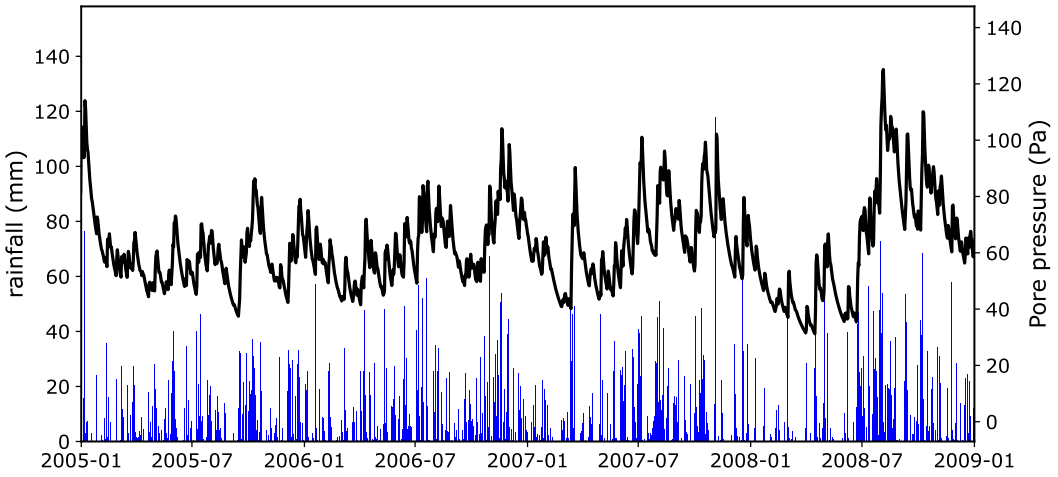

Figure 5.6: Rainfall measurements recorded at the Chateau station and corresponding mean pore pressure changes (dashed-black line) computed following Equation 5.1

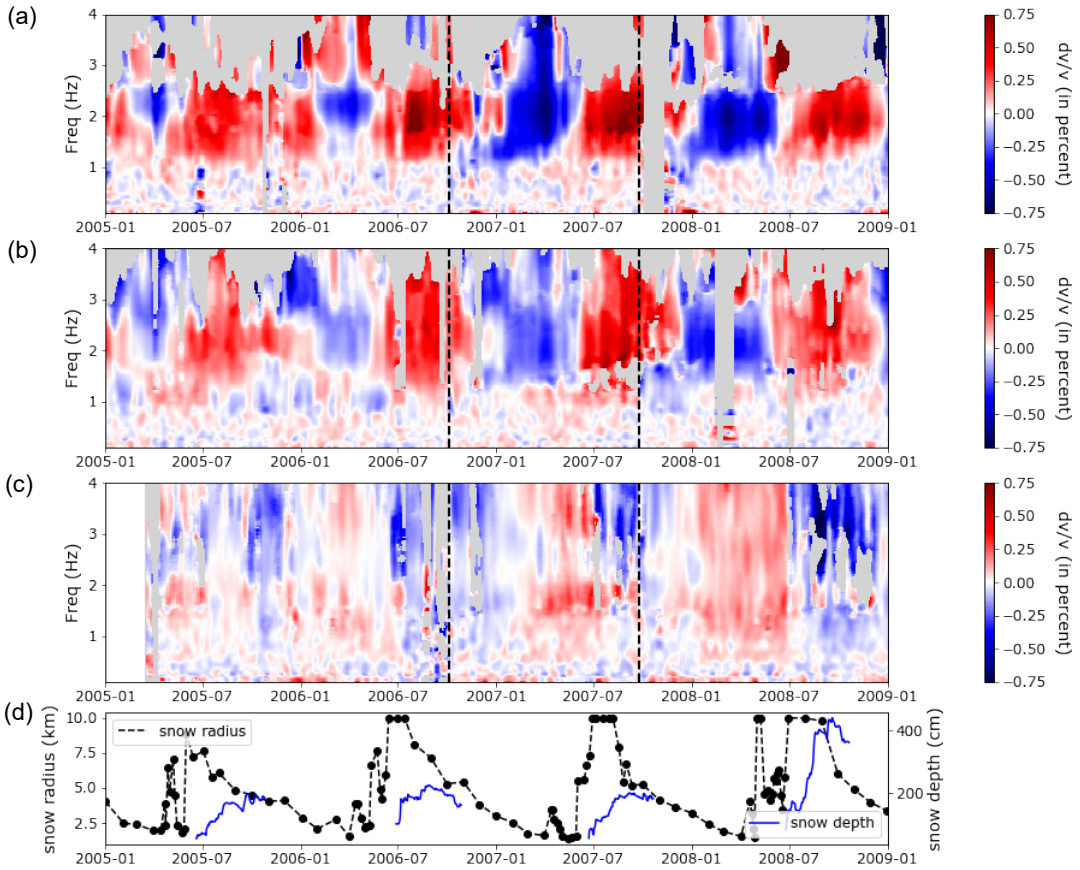

Figure 5.7: Velocity changes measured at two stations located close to summit (DRZ and FWVZ) and one station at distance from the volcano (MWVZ). (a) DRZ station. (b) FWVZ station. (c) MWVZ station. (d) snow radius and thickness, as with Figure 5.1b.

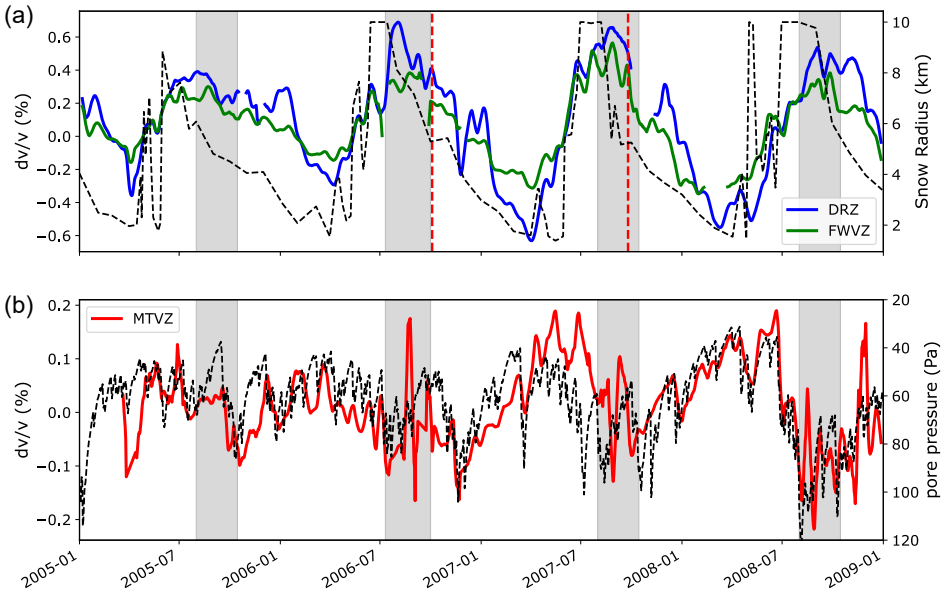

Figure 5.8: Averaged velocity changes between 1–2.5 Hz for seismic stations in Figure 5.7. Grey-shaded time periods represent times of greater snow load (from snow thickness data). (a) Velocity changes at DRZ and FWVZ stations. Snow radius shown as black-dashed line. (b) Velocity changes at MTVZ station, compared with pore pressure changes (dashed-black line) from Figure 5.6 (axis inverted).

an apparent yearly cycle. The timing of these changes differs, however, between the stations at the summit and the station at distance from the volcano. At DRZ and FWVZ stations, velocities are generally highest at between the months of July and October (Figs 5.7a and 5.7b respectively). At MTVZ, however, high velocities occur from approximately February to July (Figure 5.7c). We also see that velocities are relatively consistent across the entire frequency range at MTVZ station, especially visible during 2008. In contrast, alternating low and high velocities at DRZ and FWVZ station are mostly seen at frequencies above 1 Hz.

Computing the mean velocity change between 1–2.5 Hz at DRZ and FWVZ station shows that velocities in this frequency band are highest during periods of maximum snow thickness (each year (Figure 5.8a). The timing of the velocity increases, and subsequent decreases closely, follows the trend of snow cover extent. At MTVZ, an opposite trend is generally observed with velocity changes often at a minimum during periods of more snow (Figure 5.8b). A good fit, however, is observed when comparing velocity changes recorded at MTVZ with pore pressure changes computed using daily precipitation data (Figure 5.8b).

Velocities at each station can be averaged over the calendar year to better visualise annual changes (Fig. 5.9). We focus on the frequency band between 0.25 and 2.5 Hz, noting that much of the data above 2.5 Hz is excluded at DRZ and FWVZ stations following quality control (e.g. Figs 5.7a and 5.7b). Again, we observe differences in the timing of the seasonal trend at different stations. This is reflected when plotting the approximate timing of the seasonal velocity high (at ∼2 Hz) against station elevation (Figure 5.10a) and the distance from summit (Figure 5.10b). A clearer

trend is observed when using station elevation, rather than distance, noting that stations NGZ and OTVZ lie in close proximity to Mt Ngauruhoe (Fig 5.1a), which also is snow-covered in winter. Finally, we once again observe that stations closer to the summit show larger differences between different frequency bands. Of interest is the seemingly frequency-dependent decrease of velocities following the high velocities in winter. For example, velocities at DRZ and FWVZ stations decrease first at the lower frequencies (1.0 – 1.5 Hz) from approximately October ($\sim$ day 275) with velocities remaining elevated at higher frequencies (1.5 – 3.0 Hz) towards the end of the year. Similar trends are observed at NGZ and OTVZ stations, however occurring earlier in the year (from approximately August, $\sim$ day 210).

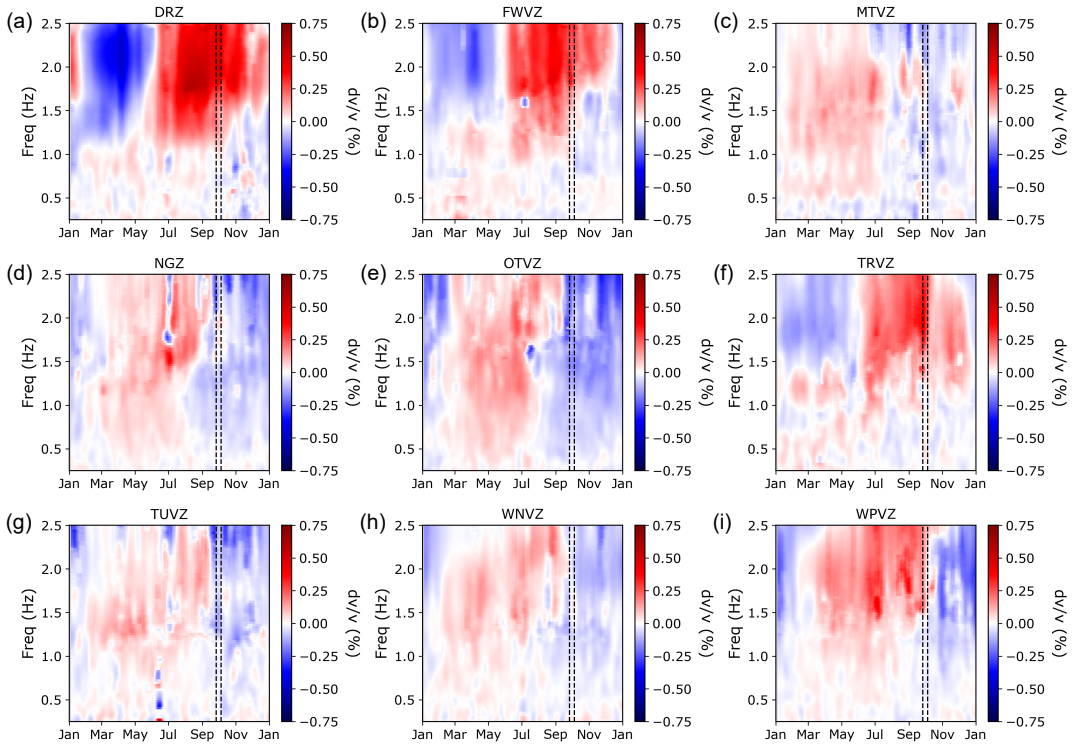

Figure 5.9: Velocity changes for all nine stations plotted against calendar day (averaging the velocity of each year between 2005–2009. Black-dashed lines indicate day of the year that the 2006 and 2007 eruptions occurred.

.

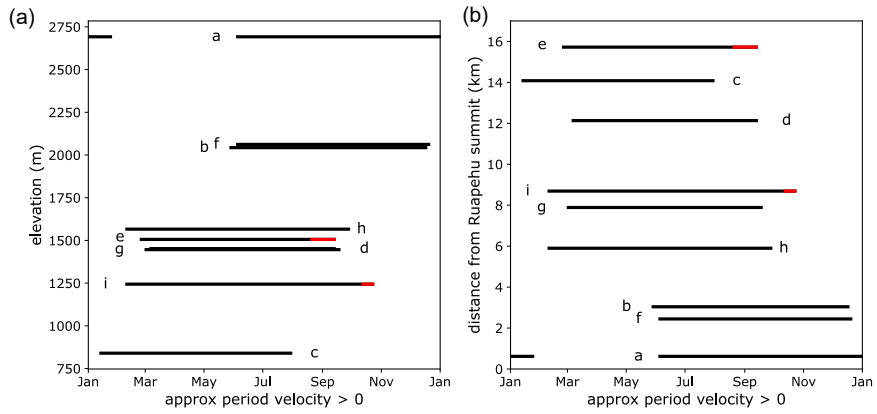

Figure 5.10: Comparison of the timing of seasonal velocity high (observed at 2 Hz, black lines) against station. Measurements correspond to times when apparent velocities are above 0%. Uncertainty in timing is indicated by red lines. (a) elevation and (b) station distance from summit. Labels within each subplot correspond to those in Figure 5.9.

5.3.2 Location in depth

The relationship between apparent velocity changes as a function of frequency ($\delta C/C(f)$), for any given mode, and the shear wave velocity perturbation as a function of depth ($\delta\beta/\beta(z)$) can be described by:

$$\frac{\delta C}{C}(f) = \int_0^\infty K(f,z)\frac{\delta\beta}{\beta}(z)dz \tag{5.2}$$

where $K(f,z)$ is a sensitivity kernel. This relationship assumes either a (1) fixed Poissons ratio and density or (2) fixed P-wave velocity and density (Haney and Tsai, 2017). If we assume that Rayleigh waves dominate the coda, we can then solve the inverse problem:

$$d = G \cdot m \tag{5.3}$$

where d is the measured velocity changes at each frequency for a given day ($\delta C/C$), G the matrix of sensitivity kernels (K) and m the velocity changes with depth ($\delta\beta/\beta$). For the sensitivity kernel, the codes of Haney and Tsai (2017) are modified to output the kernel for a given velocity model. In this case, we use a four-layered model of shear-wave velocities proposed by Godfrey et al. (2017) based on ambient noise tomography (Table 5.2). A constant Vp/Vs ratio of 1.7 and density of 2000 kg/m^3 are used, satisfying the conditions of Equation 5.2. The kernels corresponding to the Rayleigh wave fundamental mode of select frequencies are shown in Figure 5.11.

Depth (m)	Vs (m/s)
200	900
1500	1200
3000	1400
5000	1700

Table 5.2: Simplified velocity model of Godfrey et al. (2017) at Ruapehu.

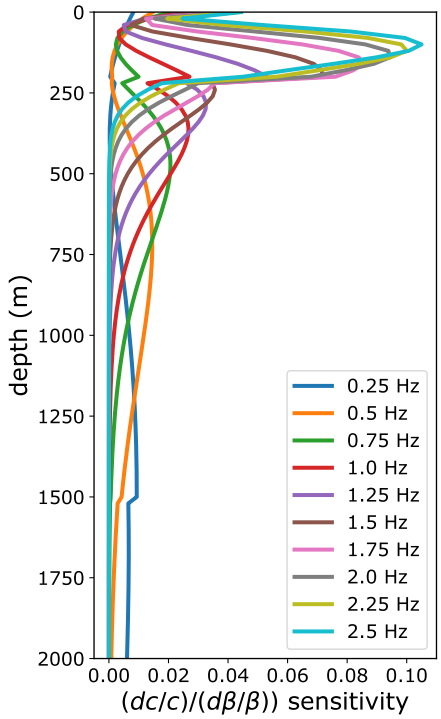

Figure 5.11: Rayleigh wave depth sensitivity kernels (fundamental mode) for selected frequencies, computed using codes of Haney and Tsai (2017) using velocity model of Godfrey et al. (2017).

From Equation 5.3, we can resolve the apparent shear-wave velocity changes with depth by performing a simple weighted-damped least squares inversion of the form:

$$\mathbf{m}_{est} = \mathbf{C}_m \mathbf{G}^T (\mathbf{G}\mathbf{C}_m\mathbf{G}^T + \mathbf{C}_d)^{-1}\mathbf{d}, \tag{5.4}$$

where $\mathbf{C}_d$ is the data covariance matrix and $\mathbf{C}_m$ is the model covariance matrix. In this study, we define the model covariance matrix following Haney and Tsai (2017) as:

$$C_m(i,j) = \sigma_m^2 \exp\left(-\frac{|z_i - z_j|}{\lambda}\right), \tag{5.5}$$

where $\sigma_m = \gamma\bar{\sigma_d}$ is the standard deviation of the model parameters, with σ_d the uncertainty in phase velocity measurement from the weighted linear regression (and

$\bar{\sigma_d}$ the average). z_i and z_j are the depth values at the i-th and j-th model points, and λ is the correlation length. The data covariance matrix is defined as a diagonal matrix with measurement uncertainties on the diagonal ($C_d = \sigma_d I$).

Appropriate choices for λ and γ are found using the L-curve criterion (Hansen and O'Leary, 1993). This is a graphical approach used to identify suitable regularization parameters for ill-posed inverse problems i.e. where the solution is non-unique. Specifically, the residual norm is plotted against the model norm for varied regularization parameters. The goal is to find a balance between the quality of fit to the data (low residual norm) and the smoothness of the solution (low model norm). Thus, the optimal choice of regularization parameters is considered to be located at the corner of the L-curve.

L-curves are produced for various choices of λ and γ at both DRZ and FWVZ stations (Figure 5.12). Both the residual norm and model norm are computed using the solution for all days in the calendar year (i.e. the 2D matrix of velocity changes against depth with time). We observe that there is are no distinct corners to the L-curves for either station. Rather, the curves are more gradual, suggesting that there may be a range of acceptable models. We define the most optimal choice at each station as the point with maximum positive curvature. At both stations, $\gamma = 60$ is the optimal choice. Different optimal values for the correlation length are identified ($\lambda = 20$ and $\lambda = 40$ for DRZ and FWVZ respectively). For simplicity, we set $\lambda = 20$ for both stations, noting that this choice also lies close to the maximum curvature of the L-curve for DRZ station (Fig. 5.12 d). Examples of the fit between observed and predicted apparent velocity changes are shown in Figure 5.13 for DRZ station and Figure 5.14 for FWVZ station.

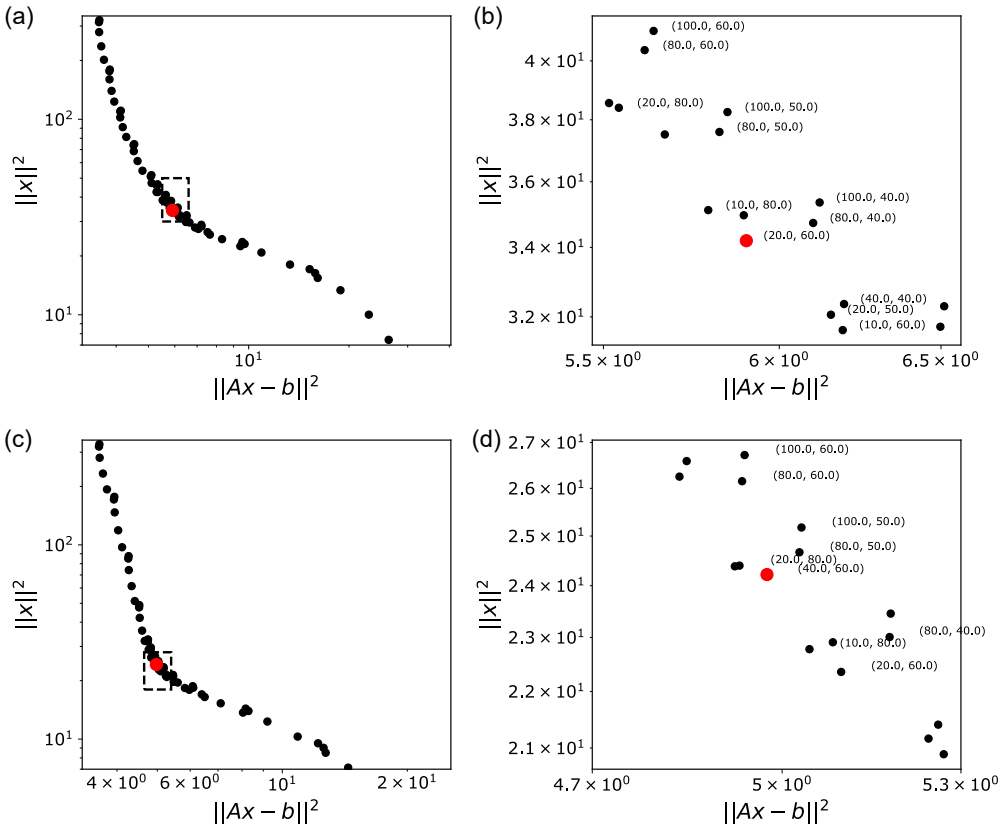

Figure 5.12: L-curve plots for various choices of regularization parameters λ and γ for DRZ and FWVZ stations. Different points are labeled with corresponding regularization parameters (λ, γ). Red point corresponds to point of maximum positive curvature. (a) DRZ station, with points within dashed rectangle shown in (b). (c) FWVZ station, with points within dashed rectangle shown in (d).

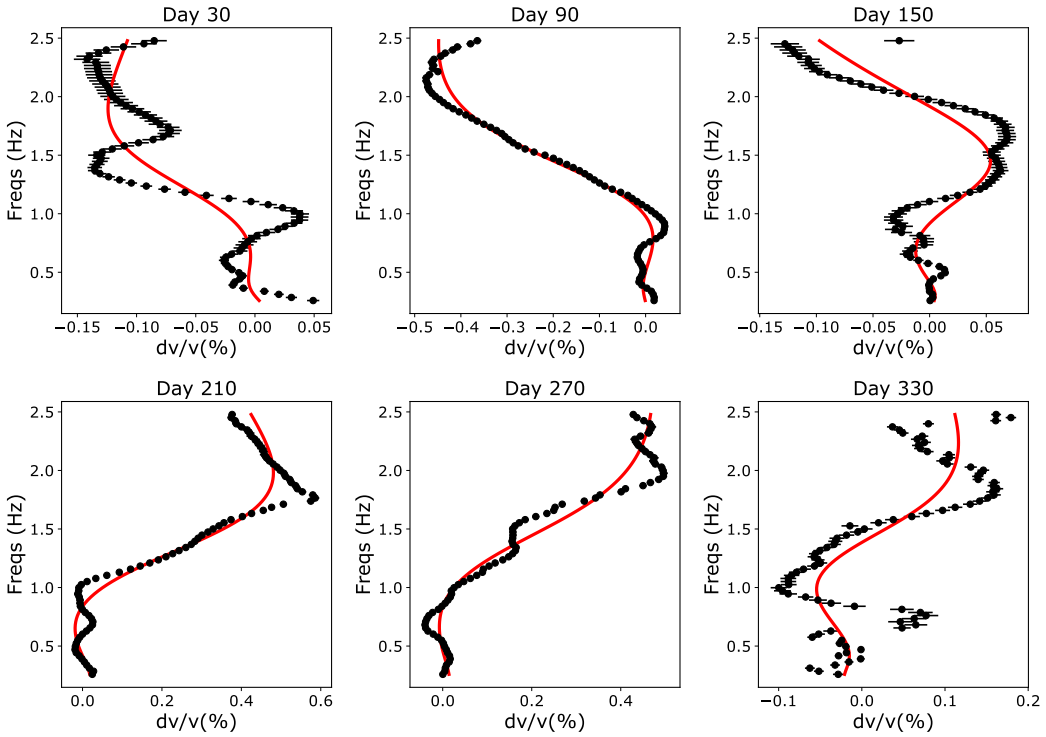

Figure 5.13: Comparison between observed (black points) and predicted (red curve) velocity changes (based on resolved model parameters) for select days at DRZ station.

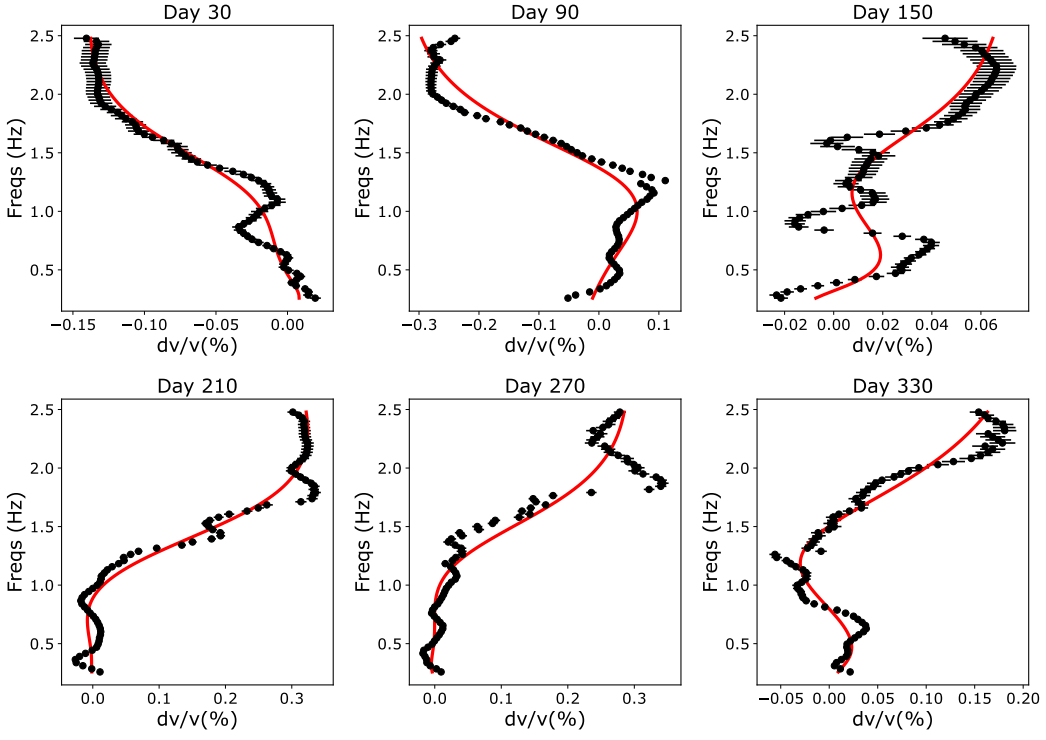

Figure 5.14: Comparison between observed (black points) and predicted (red curve) velocity changes (based on resolved model parameters) for select days at FWVZ station.

Following the inversion, apparent velocity changes with depth are plotted for stations DRZ and FWVZ in Figure 5.15. Alongside these measurements we compute the restitution index (Vergely et al., 2010), calculated as the sum of the elements in each row of the resolution matrix (Backus and Gilbert, 1967), given as:

$$R = C_m G^t (G C_m G^t + C_d)^{-1} G$$

Values closer to, or above, one indicate that the model estimates (velocity changes) are well resolved, while values less than 0.8 indicate that the model is not well constrained (Vergely et al., 2010; Lesage et al., 2014). At both stations, we see that the restitution index decays rapidly from approximately 1500 m depth. This can be related to the minimum frequency used in the inversion (0.25 Hz), which reduces in sensitivity from approximately the same depth (Fig. 5.11). Of interest is the high values between approximately 750–1300 m. This likely reflects reduced measurement error from the weighted linear regression at the lowest used frequencies. At higher frequencies ($>$ 1 Hz), where the coherence is reduced relative to lower frequencies (Fig. S5.5), measurements are only sensitive to the upper 750–1000 m (Fig. 5.11). Thus the inversion for velocities between 750 - 1500 m is generally based on measurements with high coherence. In any case, all values of the restitution index are greater than 0.8 above $\sim$ 1500 m, thus we consider the inversion capable of resolving model estimates above this depth.

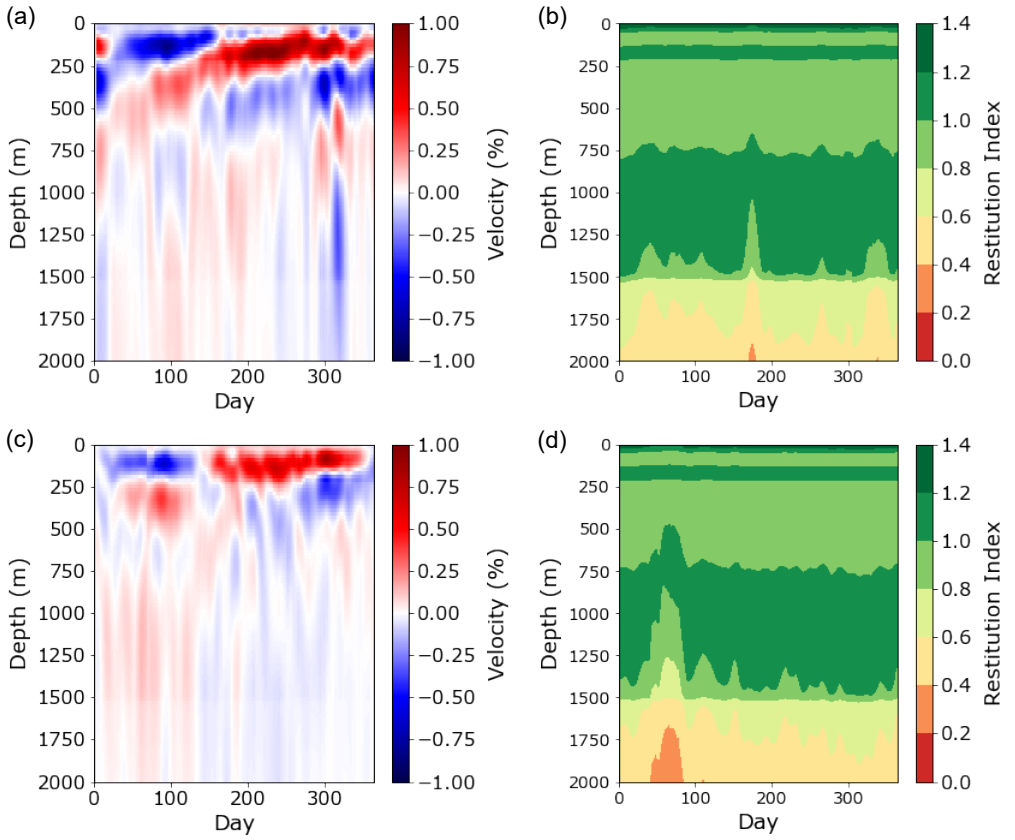

Figure 5.15: Resolved velocity changes with depth for stations DRZ (a) and FWVZ (c) and the corresponding values of restitution index (b) and (d) (DRZ and FWVZ respectively).

The results of the inversion suggest that annual velocity changes measured above 1 Hz at summit stations are located within the upper few hundred meters (Fig. 5.16) At DRZ station, the high velocity measured from July ($>$ 1 Hz) is resolved down to approximately 300 m depth (Fig. 5.16b). Closer to the surface ($<$ 50–100 m), velocities are closer to zero or slightly negative. At FWVZ station, the velocity high is resolved to 200–250 m depth (Fig. 5.16e). Unlike at DRZ station, the velocity remains positive closer to the surface though reduced in amplitude. Finally, at both stations, we observe an earlier decrease of velocities at 200–300 m depth from approximately early-September ($\sim$ day 245). This reflects the earlier decreases in apparent velocities observed between 1.0–1.5 Hz (Fig. 5.16a and 5.16d). This feature is also clearly resolved in the predicted velocities based on the model solution (Fig. 5.16c and 5.16f).

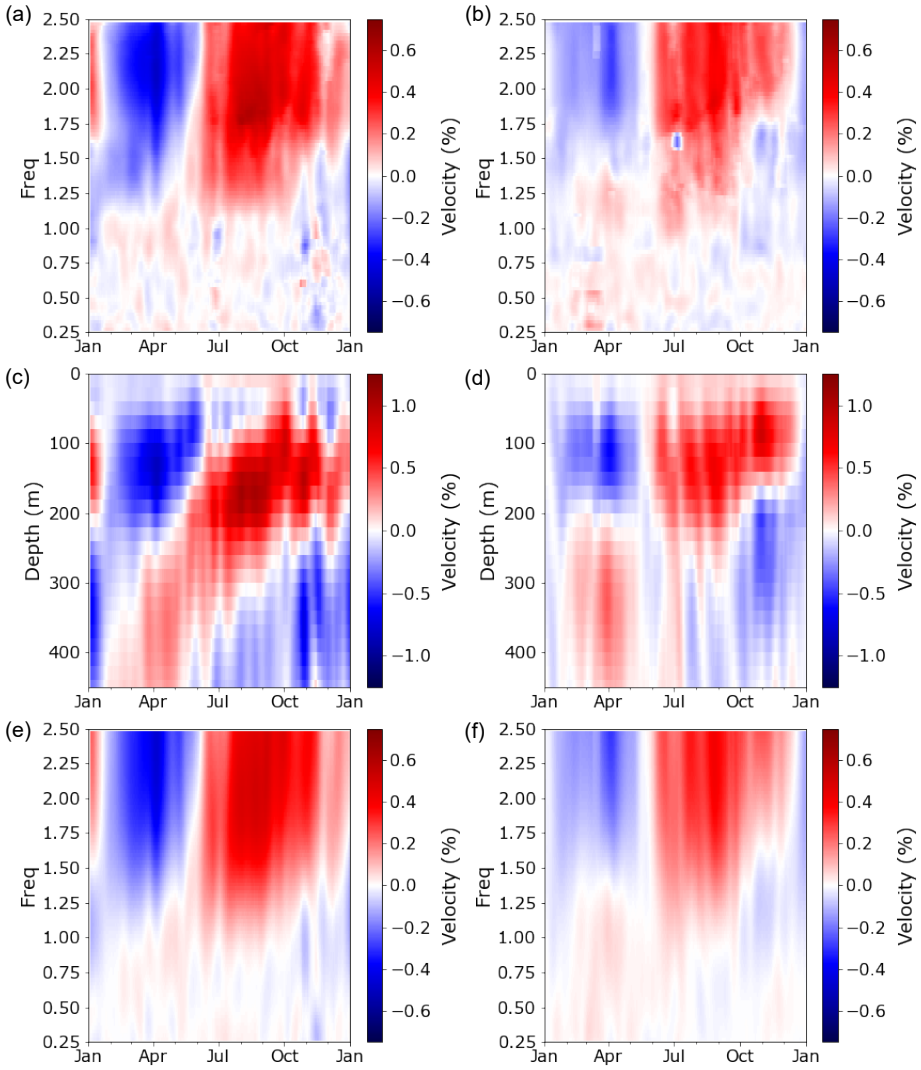

Figure 5.16: Summary of inversion solution for upper 500 m at DRZ and FWVZ stations. (a) Observed apparent velocity changes at DRZ. (b) Observed apparent velocity changes at FWVZ. (c) Modeled velocity changes with depth at DRZ. (d) Modeled velocity changes with depth at FWVZ. (e) Predicted velocity changes at DRZ based on modeled solution. (f) Predicted velocity changes at FWVZ based on modeled solution.

5.4 Discussion

The discussion focuses on three key aspects. First, the validity of the results is considered in relation to the use of tremor as the dominant source above 1.0 Hz. We then consider the possible mechanisms that could be driving the changes observed. Finally, we discuss the possibility of seasonality in eruptive activity at Mt Ruapehu and consider how this could relate to observations made in this study.

5.4.1 Reliability of velocity changes computed from tremor

Many of the results presented come from the analysis of cross-correlation functions within the frequency band of volcanic tremor at Ruapehu. In section 5.2.2, we observed that changes in spectral content modified the character of cross-correlation functions computed between 1–4 Hz during important time periods. For example, clusters of cross-correlation functions lasting multiple months — that differ from the dominant cluster (C5) — occur during periods when significant snow-cover would be expected in 2005, 2007, and 2008 (Fig. 5.2). Smaller clusters also occur in 2006, however these are short-lived. The application of cross-wavelet transform is intended to minimise the influence of varying seismic sources on our results. Specifically, measurements can be retained at frequencies where there is a strong enough similarity between individual cross-correlation functions and the reference function while unreliable measurements can be excluded. This exclusion is based on quality control criteria, where we define a minimum wavelet coherence required for individual delay time measurements. Similarly, we require that at least 25% of possible measurements within the coda are of non-zero weight (i.e. are above this threshold).

One approach to test the effectiveness of the wavelet coherence threshold is to examine the influence of increasing this threshold on our results. Thus, the threshold is increased to 0.7, 0.8, and 0.9 and the average yearly velocity changes re-computed for stations DRZ, FWVZ, and MTVZ (Fig. 5.17). As the threshold is increased, we observe that progressively more data is excluded. This mostly occurs at the stations closer to the summit, where almost all data above 2.3 Hz is excluded at DRZ station at a threshold of 0.8 (Fig. 5.17g). A large amount of data above 1.5 Hz is also excluded at this station between July and November. Of significance though is that increasing the threshold does not appear to influence the primary observations. For example, the trend of velocity changes decreasing first at lower frequencies — following the annual increase in winter — is still clearly present at both DRZ (Fig. 5.17g) and FWVZ station (Fig. 5.17h) when the threshold is increased to 0.8. Similarly, velocity changes at MTVZ remain highly consistent (Fig. 5.17i). At the maximum tested threshold (0.9), there is significant data loss at all three stations about 1 Hz (Fig. 5.17j-l). However, the decrease in velocity late-September between 1.0–1.5 Hz is still visible. Thus, we do not believe our observations are sensitive to low quality measurements as a result of variable tremor characteristics.

Another factor to be considered is the influence that velocity changes computed during a single year could have on the yearly averages. We therefore examine the stability of the results when individual years are removed (Fig. 5.18). Once again the main features of the velocity changes are well-preserved for all three stations (DRZ, FWVZ, and MTVZ) no matter which year is excluded from the data. The influence of the eruptions themselves is also verified by removing the years 2006 and 2007 from the average (Fig. S5.6). In this case the earlier decrease of velocities at lower frequencies is slightly less obvious at DRZ station, but still visible (Fig. S5.6d). At FWVZ station the decrease is still clearly visible when removing eruption years (Fig. S5.6e). Thus there is no evidence that anomalous velocity changes recorded during a single year are responsible for the main features we observe.

Finally, we consider whether changes in the frequency or occurrence of tremor

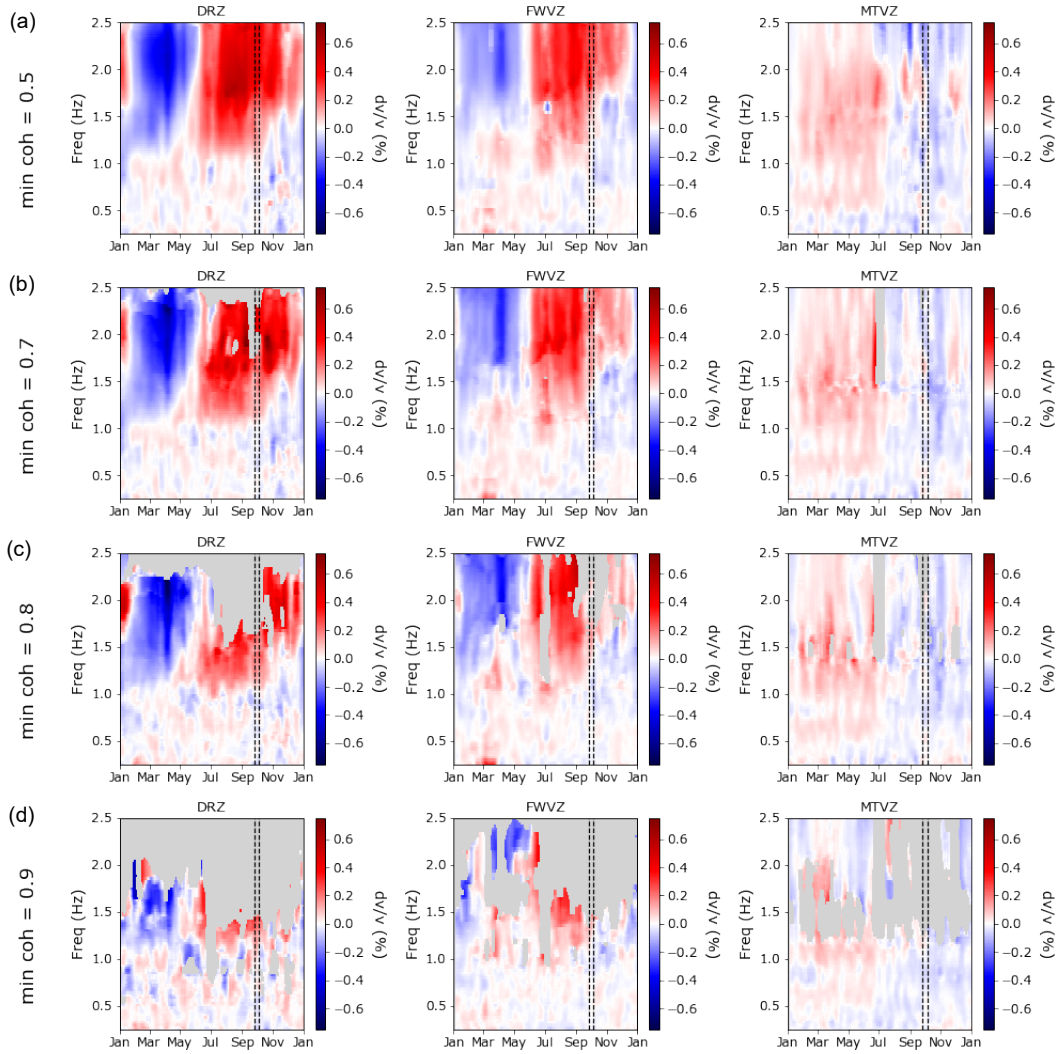

Figure 5.17: Comparison of yearly velocity changes at DRZ, FWVZ, and MTVZ for different minimum values of yearly coherence. (a) minimum coherence = 0.5. (b) minimum coherence = 0.7. (c) minimum coherence = 0.8. (d) minimum coherence = 0.9.

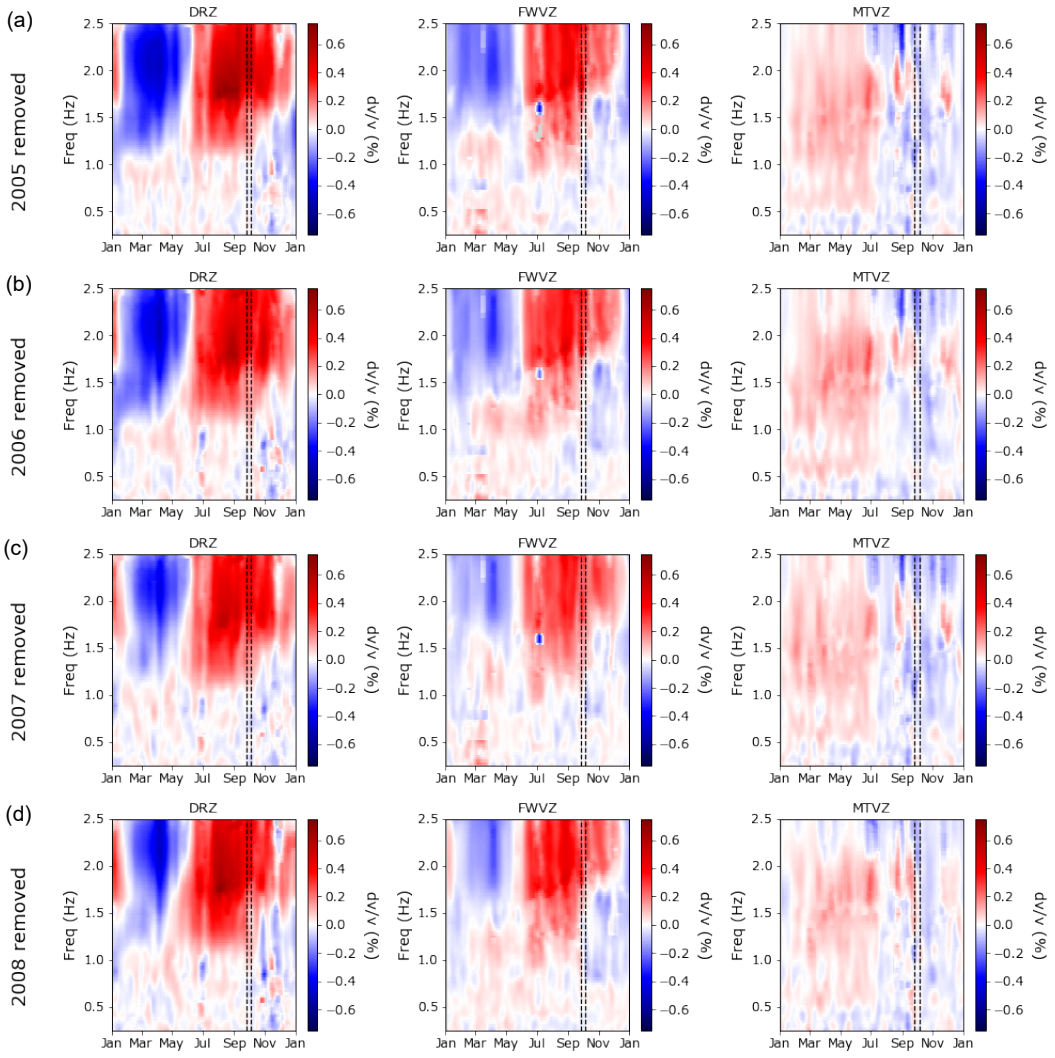

Figure 5.18: Comparison of yearly velocity changes at DRZ, FWVZ and MTVZ after removing individual years from average (a) 2005 removed. (b) 2006 removed. (c) 2007 removed. (d) 2008 removed.

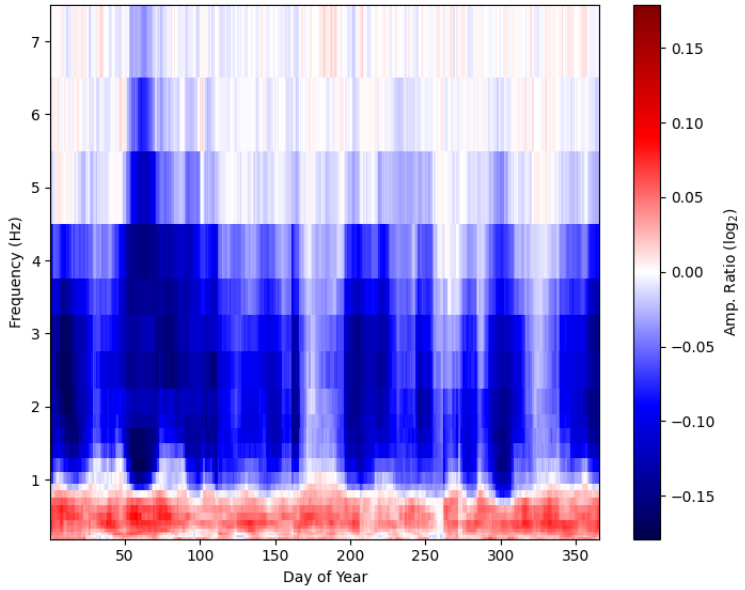

Figure 5.19: Yearly amplitude ratio — following Equation 5.6 — computed for station-pair MTVZ-TRVZ filtered into different narrow frequency bands.

could be influencing the annual trends observed. For this, we take advantage of the clear asymmetry observed in cross-correlation functions between MTVZ-TRVZ when tremor is present (Fig. S5.1). This is quantified using a measure of amplitude ratio defined following:

$$R_{amp} = \log_2 \left(\frac{\max(cc(t^+))}{\max(cc(t^-))} \right). \tag{5.6}$$

where t^+ represents positive lag times and t^- negative lag times. Taking the ratio to log base 2 means that a ratio of zero corresponds to the same maximum amplitude either side of the cross-correlation function. Furthermore, an increase or decrease in R of 1 corresponds to a change in the amplitude ratio by a factor of 2. Computing the ratio through 2005-2009 at different frequencies (defined in Table S5.1), and plotting the yearly average, clearly shows the dominant negative ratio between approximately 1 – 4 Hz associated with tremor. Most significantly, there does not appear to be any strong differences throughout the year that could explain the annual velocity highs and lows or the earlier velocity decrease from late-September / early-October between 1.0–1.5 Hz. Thus, the occurrence of tremor does not appear to be a factor towards explaining these observations.

Combining these observations, we do not believe that the use of tremor as the dominant seismic source for cross-correlation functions above $\sim$ 1 Hz has adversely impacted our results. Note, however, that this interpretation is only valid for the time-scales of features discussed (months rather than days). Shorter-term changes (days) could still be influenced by subtle variation in the character of tremor and would need to be scrutinized more closely before interpreting as real changes within the medium.

5.4.2 Mechanism of velocity changes

Differences between the trends of annual velocity changes at stations on and off the volcano suggest there may be different mechanisms driving velocity changes. The three stations closest to the summit of Mt Ruapehu (DRZ, FWVZ, and TRVZ), show a velocity increase above 1 Hz closely correlated with the amount of snow cover (Fig. 5.8a and 5.9f). Maximum velocities occur during the times of greater snow thicknesss (Fig. 5.7a, 5.7b, and 5.8a). Further away from the volcano, where minimal snow cover occurs, velocities track closely with pore pressure changes following precipitation events (Fig. 5.8b). These observations are also expressed when plotting the timing of annual velocity changes against elevation (Fig. 5.10a). At higher elevations, higher velocities are recorded from the beginning of winter lasting until summer, while at lower elevations higher velocities are recorded from the beginning of summer lasting until winter. These observations strongly suggest that winter conditions are responsible for the velocity changes recorded at stations close to the summit of the volcano.

In this section we consider three processes that could influence seismic velocity changes on the volcano during winter. These are (1) Surface-loading due to snow mass, (2) Seasonal freeze-thaw cycles, and (3) Changes in groundwater level.

5.4.2.1 Surface-loading due to snow mass

The presence of a snowpack acts to apply an additional surface-load to the subsurface (otherwise known as snow-loading). The increased pressure of such a load results in pore space reduction and crack closing in the medium (Silver et al., 2007). The effect of this is to increase seismic velocities (Nur, 1971). Thus, multiple studies have attributed winter increases in velocity to the presence of snow and subsequent snow-loading (Hotovec-Ellis et al., 2014; Wang et al., 2017; Donaldson et al., 2019; Makus et al., 2023). Characteristic of this is a positive correlation between snow thickness and seismic velocities.

We examine further the influence of snow-loading at the volcano by applying a simple model of an axisymmetric surface load acting on a semi-infinite elastic medium. For a surface load $P(r) = \rho_c g h(r)$, the pressure perturbation with depth can be resolved following:

$$p(r,z) = -\frac{\sigma_{rr} + \sigma_{\theta\theta} + \sigma_{zz}}{3}$$

where σ_{rr}, $\sigma_{\theta\theta}$, and σ_{zz} the diagonal components of the stress tensor (positive in tension) with cylindrical coordinates r, θ, and z are cylindrical coordinates. These are computed using a MATLAB package based on the solution provided by Sneddon (1951), with σ_{rr}, $\sigma_{\theta\theta}$, and σ_{zz} defined following:

$$\sigma_{rr} = \int_0^\infty (1 - z\xi) Z(\xi) e^{-\xi z} J_0(\xi r) d\xi + \frac{1}{r} \int_0^\infty Z(\xi)(z - \frac{1-2\nu}{\xi}) e^{-\xi z} J_1(\xi r) d\xi \quad (5.7)$$

$$\sigma_{zz} = z \int_0^\infty Z(\xi) e^{-\xi z} J_0(\xi r) \xi d\xi + \int_0^\infty Z(\xi) e^{-\xi z} J_0(\xi r) d\xi \tag{5.8}$$

$$\sigma_{\theta\theta} = 2\nu \int_0^\infty Z(\xi) e^{-\xi z} J_0(\xi r) d\xi - \frac{1}{r} \int_0^\infty Z(\xi)(z - \frac{1-2\nu}{\xi}) e^{-\xi z} J_1(\xi r) d\xi \tag{5.9}$$

where

$$Z(\xi) = -\xi \int_0^\infty r p(r) J_0(\xi r) dr \tag{5.10}$$

and J_0 and J_1 are Bessel functions of order zero and one respectively, μ the shear modulus and λ the Lamé coefficient. The stress components are considered positive in tension, hence the negative sign in $P(r, z)$, and z is orientated downwards. Note that this simple model certainly cannot entirely represent the situation at Ruapehu, where topography is not included in the model. However it can still provide a first-order approximation of expected pressure changes due to snow-loads of comparable size to those at the volcano.

Various input shapes were considered for the snow load, including a uniform load, a cone-shaped load, a sigmoid-shaped load. Our preferred model is that of a sigmoid-shape, based on the logistic function, where the height of snow (h) is defined as a function of radius (r) following:

$$h(r) = h_{max} \left(1 - \frac{1}{1 + e^{-k(r/R - r_0)}} \right) \tag{5.11}$$

where h_{max} is the maximum height of the snow, R the radius of the load, k the logistic growth rate, or steepness of the curve, and r_0 the function midpoint. This is set to $r_0 = 0.5$, meaning that the midpoint of the function occurs at $r = R/2$. With this, the load is represented by an initial — approximately exponential — increase of h beginning from $r = R$, slowing to a more linear increase around $r = R/2$ and finally increasing more slowly towards h_{max} at $r = 0$ (Fig. 5.20a). This choice reflects the observation that mountain regions often show an initial strongly positive elevation gradient of snow thickness which then levels out at a certain altitude, even decreasing at the highest elevations (Grünewald and Lehning, 2011; Kirchner et al., 2014; Grünewald et al., 2014). It should be noted that a positive gradient is not always identified, where Kerr et al. (2013) observed a negative gradient between snow and elevation for a study site in the South Island of New Zealand. This was attributed to redistribution of snow mass from steep slopes ($>\sim 25°$) at higher elevations to low angle slopes at lower elevations. Measured slopes on Ruapehu, however, are mostly under 20 degrees (Fig. S5.7). Thus we do not anticipate such a negative gradient. The logistic growth rate is set to $k = 10$. Both of these choices are somewhat arbitrary, given no information about the distribution of snow with time, however they should be suitable to provide a first approximation of the influence of

the snow-load. Alternative models of a uniform and cone-shaped load can be found in Supplementary S5.2.

Modeling the temporal influence of the snow-load at Mt Ruapehu is achieved through varying the values of R and h_{max}. Values of R are derived using MODIS satellite data (Fig. 5.1b). Values of h_{max} are taken as the average snow thickness recorded by ski stations at Whakapapa and Turoa (noting both show a similar shape through time). No adjustment is made to these values due to their slightly lower elevation ($\sim$2000 m) given uncertainties about the distribution of snow with elevation. Examples of the applied load at different times of year are shown in Figure 5.20c). The final solution can be achieved however using dimensionless outputs of a single computation. Specifically, we output dimensionless 2D pressures (dividing by $\rho_c g h_{max}$) for a dimensionless z/R by r/R grid (Fig. 5.20b). Different values of R and h_{max} are used to dimensionalize the grid and corresponding pressure values. This provides the 2D pressures for a given load in $z - r$ space.

Final parameters to choose are the density of the snow ρ_c and poissons ratio ν, where the elastic modulus is not a factor for calculation of pressure in a homogeneous medium. Poissons ratio is set as $\nu = 0.2$, similar to values recorded by Mordensky et al. (2018) and the average for andesitic rocks more generally (Heap et al., 2020). For the density, we use value of 0.4 kg/m^3 based on values recorded at the ski station between 2005–2009 (0.42 kg/m^3). This is slightly larger than the average densities of 0.3–0.35 kg/m^3 recorded by Heine (1962) at Ruapehu, however most of these measurements were at lower elevations ($\sim$1750 m). Of eight samples taken at $>$ 2050 m, the average was 0.38 kg/m^3.

Model outputs suggest a pressure increase on the order of $\sim$5–10 kPa as a result of the snow-load (Fig. 5.21a-c). These are largest at axis of the load ($r = 0$), with the months of August and September associated with larger changes (Fig. 5.21). These months also show greater depth penetration, with changes on the order of 3–5 kPa detected at $z = 2000$m. Comparatively, the pressure changes are smaller (2–3 kPa) at greater distances from the axis of the load (e.g. $r = 4000$ m, Fig. 5.21c). There is, however, less variation with depth relative to values closer to the load axis.

The temporal evolution of the pressures induced by the load follows a similar pattern at the distances shown ($r < 4000m$), though with some differences. The month of July is associated with a more rapid increase in pressures as the snow thickness increases, peaking mid-August. September then marks the transition to decreasing pressures and reduced depth sensitivity as the radius of the load begins to decrease (Fig. 5.20). Closer to the axis of the load ($r < 2000m$), this decrease occurs slowly relative to the earlier pressure increase (Fig. 5.21a-b). At greater distances, however, the decrease in pressures is far more rapid, with minimal pressure changes evident due to the load from the start of October (Fig. 5.21c).

The depth of pressure changes modeled as a result of the snow load are much deeper than the depths of velocity changes suggested by our inversion (upper few hundred meters). This can be explained by considering how the amplitude of shear-wave velocity changes as a function of depth. Based on relationships between seismic velocity and induced stresses derived by Tromp and Trampert (2018), Fokker et al. (2021) show that velocity changes ($d\beta/\beta$) due to a pressure change (ignoring devi-

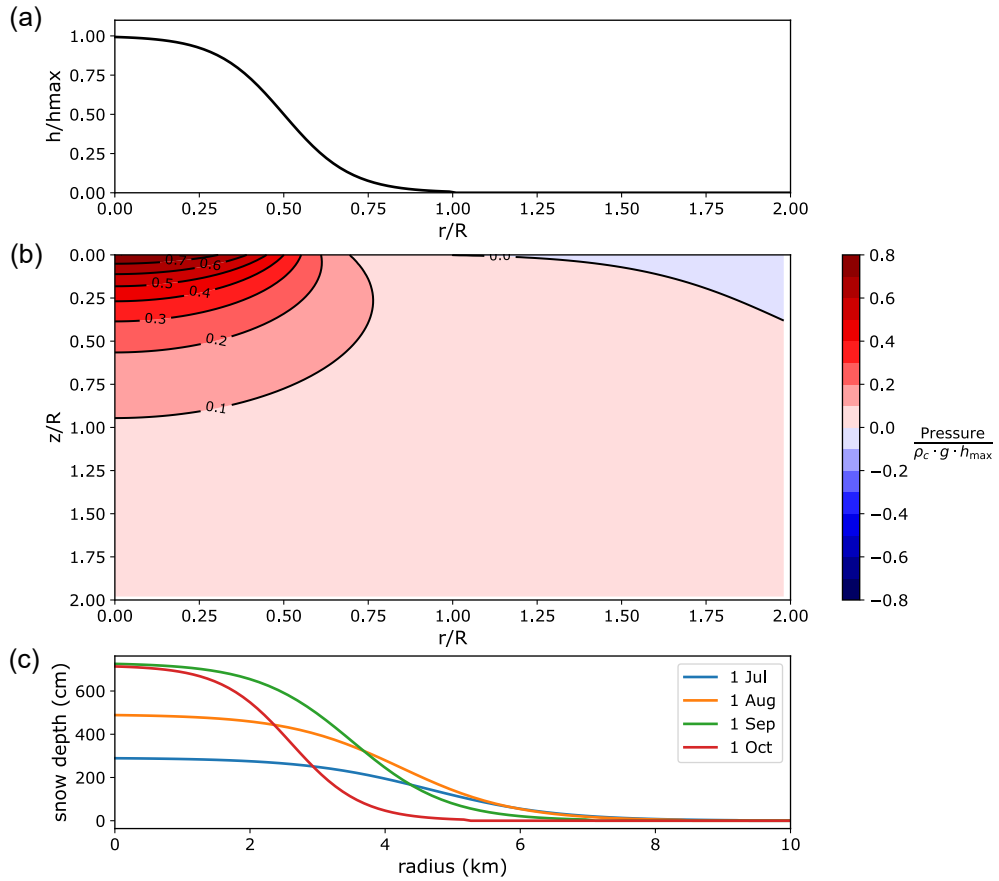

Figure 5.20: Model inputs for axisymmetric snow load (with sigmoid shape) applied at the surface and corresponding dimensionless pressures calculated for semi-infinite elastic medium. (a) Dimensionless load with sigmoid shape. (b) Dimensionless 2D pressure calculations plotted against radius and depth (both also dimensionless). (c) Surface load at different times of year using R inferred from MODIS satellite data and h_{max} inferred from ski-station snow thickness measurements.

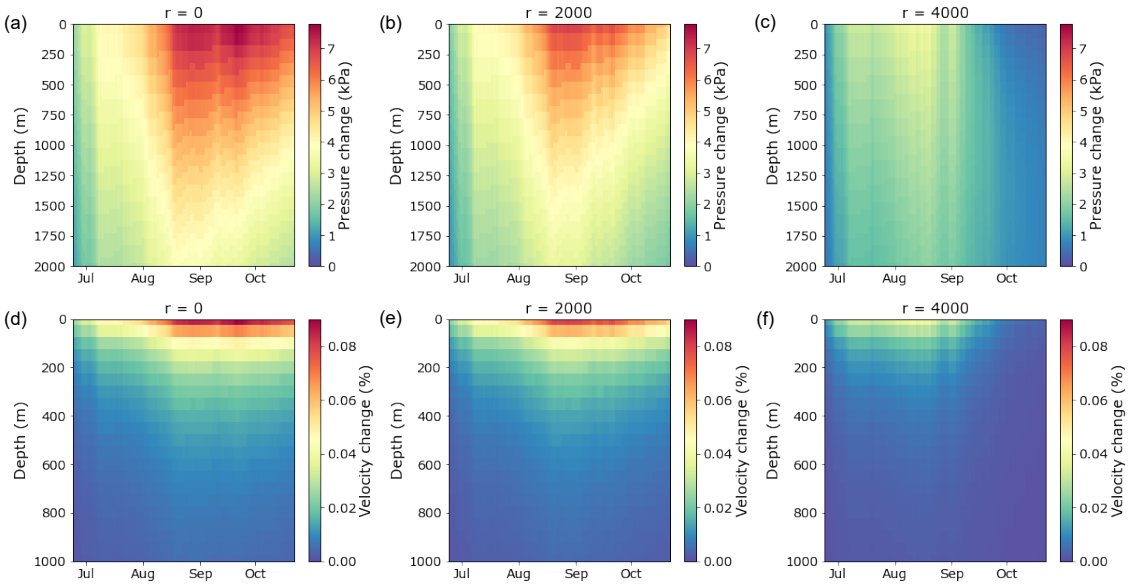

Figure 5.21: Pressure change calculations at different radius r between June 20 and October 21 for sigmoid-shaped snow-load and the estimated velocity changes from Equation 5.12. Note velocity changes shown down to 1000 m only. (a) pressure change at $r = 0$ m. (b) pressure change at $r = 2000$ m. (c) pressure change at $r = 4000$ m. (d) velocity changes at $r = 0$ m. (e) velocity changes at $r = 2000$ m. (f) velocity changes at $r = 4000$ m.

atoric stresses) can be approximated as:

$$\frac{d\beta}{\beta} = \frac{\mu'}{2\mu} p \tag{5.12}$$

where μ is the shear-wave modulus and μ' is pressure derivative ($\partial\mu/\partial P$). Importantly, μ is expected to increase with depth while μ' is expected to decrease. Thus, the velocity change due to a pressure change closer to the surface is expected to be larger than a velocity change due to the same pressure change at depth.

This idea can be demonstrated through a simple transformation of pressure changes to seismic velocities based on the P- (V_p) and S- (V_s) wave velocities used in the depth inversion (Table 5.2). For this, the density rho is computed as $rho = (V_p + 2.37)/2.8$ (Fig. 5.22b) (following the velocity-density relationship of Gebrande (1982) for volcanic rocks), the shear modulus μ as $\mu = V_s \rho$ (Fig. 5.22c), and the confining pressure $P(z)$ as $P(z) = \int_0^z \rho(z) g dz$ (Fig. 5.22d). From this the derivative of the shear modulus with respect to pressure (μ') can be resolved (Fig. 5.22e). To ensure smooth variations of these parameters, the velocity model of Godfrey et al. (2017) was fit with the analytical function of Lesage et al. (2018b) with varied initial parameters (Fig. 5.22a). The obtained values of shear modulus vary between approximately 1 GPa at the surface and 3.5 GPa at 2000 m (corresponding to elastic moduli between approximately 2.5 GPa to 8 GPa). These are on the smaller side relative to laboratory measurements of andesitic rock at Ruapehu (elastic moduli

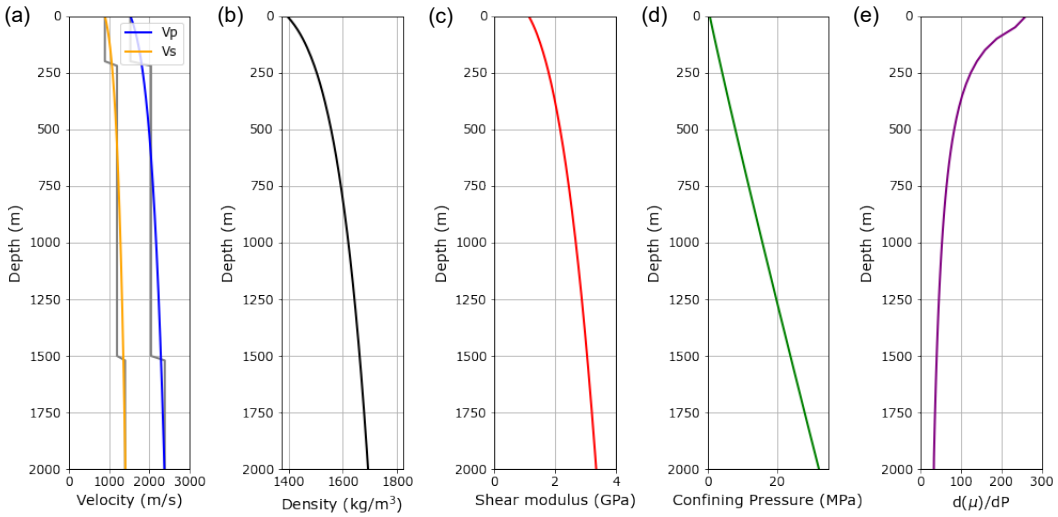

Figure 5.22: Model parameters used in converting pressure changes with depth to shear velocity changes. (a) P- and S- wave velocities (V_p and V_s respectively). These are fit to model of (Godfrey et al., 2017) from Table 5.2 using $V(z) = V_0[(z+a)^\alpha - a^\alpha + 1]$. [$V_{p0}$ = 1530 m/s, α_p = 0.1, a_p = 110] and [V_{s0} = 900 m/s, α_s = 0.1, a_s = 100] are used respectively for smoothed P- and S- wave velocities. (b) Density profile computed from V_p profile. (c) Shear modulus with depth, computed using V_s and density profile. (d) confining pressure computed using density profile. (e) shear modulus derivative with respect to pressure.

between approximately 1–40 GPa, (Mordensky et al., 2018; Heap et al., 2020)). Obtained values of the pressure derivative μ' (unitless) vary between approximately 250 at the surface and <100 beyond a few hundred meters depth. Few experiments have studied the influence of confining pressure on the elastic or shear modulus (Wu et al., 2019), thus this value is relatively unconstrained. However, approximating this value using data from multiple studies that show changes in the elastic, or shear, modulus against confining pressure suggests a range of values ranging between 1–1000 for both soil and rock samples (Sahaphol and Miura, 2005; Sawangsuriya et al., 2006; Arslan et al., 2008; Wang et al., 2009; Wu et al., 2019; Fokker et al., 2021, 2023). Thus, the values computed here are within the range of values that might be expected. Final velocities estimated at different r distances are shown in Figures 5.21d-f, where changes due to the snow-load are now localised within the upper few hundred meters.

Comparing the velocities recorded at the summit stations and the estimated velocity changes due to snow-loading at similar distances show some similarities (Fig. 5.23). Maximum values of velocity changes due to snow-loading are expected in August/early-September, around the time velocities peak at both DRZ and FWVZ station at approximately 2 Hz. Similarly we observe that the penetration depth of the load begins to decrease from around this time, most clearly seen at r =2500 m (Fig. 5.23f). This decrease in pressures occurs at the same time we see velocities begin to decrease between 1.0–1.5 Hz (Fig. 5.23b), or 200-300 m depth from the inversion (Fig. 5.23d). Dissimilarities are also observed, however, between snow-loading induced changes and measured velocity changes. For example, the increase

in measured velocities above $1Hz$ occurs at approximately the same time as snow begins to accumulate (indicated by increasing snow radius, Fig. 5.23a-b). However, the snow-loading model predicts reduced velocity changes and depth penetration at this time (Fig. 5.23e-f). Furthermore, a large difference is observed between the magnitude of changes predicted by using Equation 5.12 (0.1 <%) and those observed ($\sim$ 1%). This mismatch could relate to a number of assumptions that have been made in converting the velocity changes, for example near-surface values of the shear modulus and pressure derivative. Similarly, the distribution of the snow-load over the volcano is poorly constrained, and effects of topography have been excluded. Alternatively though, it could suggest that snow-loading is not enough by itself to explain the seasonal velocity changes observed on Ruapehu.

5.4.2.2 Frozen ground

Frozen ground can have a significant effect on seismic velocity changes. This occurs through strong increases in the shear modulus of frozen ground relative to unfrozen ground (Miao et al., 2019; Steinmann et al., 2021). While these changes are typically shallow (less than a few meters), the impact of such large changes in the shear velocity can influence velocity measurements at frequencies usually expected to be sensitive to deeper parts of the crust (Gassenmeier et al., 2014; Steinmann et al., 2021). In an extreme example, Steinmann et al. (2021) identify seismic velocity changes of a few percent due to a frozen soil layer only a few centimeters thick using 1–6 Hz cross-correlation functions. They interpret this surprising result as due to a 400-600% increase in shear-wave velocity within the frozen layer. Thus, it is not clear that using frequencies sensitive to a few hundred meters excludes the possibility of detecting changes in the very shallow subsurface.

Of the three stations showing a seasonal increase in velocities in winter, FWVZ and TRVZ are located on the flanks of the volcano while DRZ is located on the summit plateau. The flanks of the volcano are composed of lava flows and autobreccias mixed with glacial and pyroclastic deposits (Conway et al., 2016). Both FWVZ and TRVZ are located in the region of andesitic lava flows associated with the Whakapapa formation (post-glacial lava flows erupted since 15 ka) (Conway et al., 2016). It can be expected that a shear-wave velocity increase due to andesitic rock would be, proportionally, far less than those measured in soil. Thus we consider it unlikely that frozen ground would have as strong an influence on velocity changes in this study. However, since volcanic soils of only a few centimeters could have an influence on seismic velocities (based on the findings of (Steinmann et al., 2021)), we remain cautious in ruling out the possibility of a frozen layer on this grounds.

The similarity between seismic velocity changes expected due to snow-loading and freezing of the ground raises questions on how to differentiate between the different processes. At the time of writing, there has been minimal comparison between these two processes. In one example, conflicting arguments have been provided for the process driving winter increases in velocity in eastern Hokkaido, Japan (Wang et al., 2017; Miao et al., 2019). Wang et al. (2017) originally interpret these increases as due to the combined effect of snow-loading and reduced infiltration of groundwater due to the snow cover. However, Miao et al. (2019) interpret these change as related to

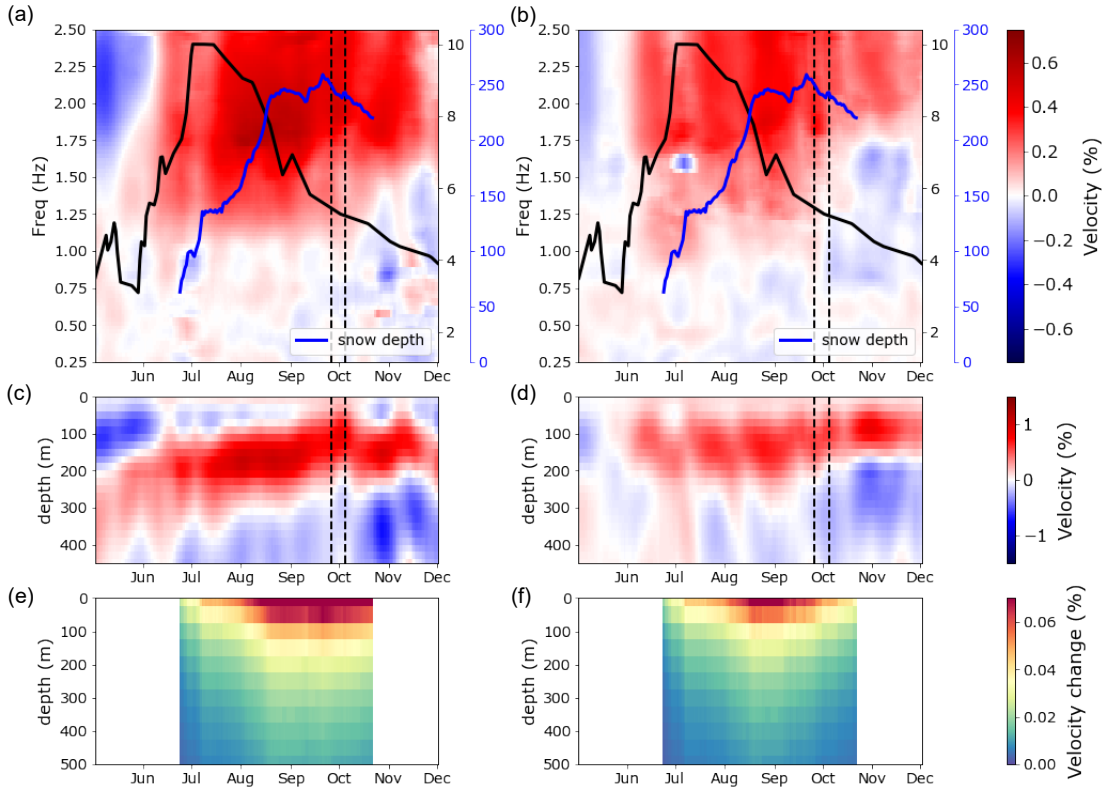

Figure 5.23: Observed and depth-inverted velocity changes compared with estimated velocity changes from sigmoid shaped snow-load at DRZ and FWVZ station. Black-dashed lines indicate day of the year that the 2006 and 2007 eruptions occurred. Black lines in (a) and (b) correspond to snow radius R (km), with axes ticks to right of plot (black). Blue lines in (a) and (b) correspond to snow thicknesses (cm) measured averaged between Whakakpapa and Turoa ski-stations, with axes ticks also to right of plot (blue). (a) Velocity measurements with frequency at DRZ. (b) Velocity measurements with frequency at FWVZ. (c) Depth inverted velocity changes at DRZ station. (d) Depth inverted velocity changes at FWVZ station. (e) Velocity changes estimated from snow-load at DRZ (r set to 0 m). (f) Velocity changes estimated from snow-load at FWVZ (r set to 2500 m)

seasonal frozen soil, noting that eastern Hokkaido gets much less snow than western Hokkaido, which showed a weaker and opposite trend (Wang et al., 2017). They suggest that the absence of seasonal increases in western Hokkaido could be due to the thermal insulation effects of snow cover.

Considering the insulating effect of snow may help to distinguish between the two mechanisms. Snow cover has been widely found to have a strong influence on the thermal regime of the ground (e.g. Zhang, 2005; Bayard et al., 2005; Iwata et al., 2010; Zhou et al., 2013; Maurer and Bowling, 2014; Oliva et al., 2014; Goncharova et al., 2019). Very thin snow cover (< 20cm) reflects incoming solar radiation due to its high albedo, and thus can act to cool the ground. However, as snow thicknesses increases it provides an insulating effect, limiting the transfer of heat between the soil and the atmosphere (Zhang, 2005). This decoupling occurs for thicknesses greater than approximately 40 cm (Zhang, 2005; Zhang et al., 2021). Similarly, it is often found that soil temperatures increase with increasing snow thickness (e.g. Zhang, 2005; Goncharova et al., 2019; Zhang et al., 2021). Thus, thick snow cover can prevent the development of frozen ground in winter months (Bayard et al., 2005; Oliva et al., 2014). Exceptions can occur, however, if the ground is sufficiently frozen prior to the accumulation of thick snow (Brooks et al., 1995; Bayard et al., 2005; Rush and Rajaram, 2022). In this case, a deeper snow pack could act to preserve the frozen state of the ground (Bayard et al., 2005; Rush and Rajaram, 2022).

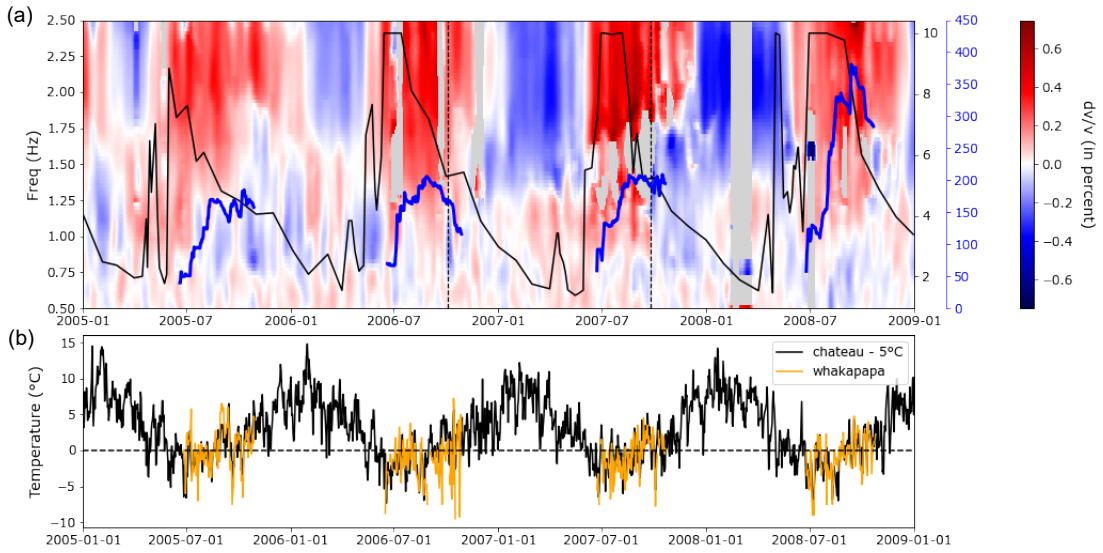

Figure 5.24: Comparison between velocity changes recorded at FWVZ station and temperatures recorded at nearby Whakapapa ski station. (a) Measured velocity changes at FWVZ station. Black line corresponds to snow radius R (km), with axes ticks to right of plot (black). Blue lines correspond to snow thicknesses (cm) measured at Whakakpapa ski-station, with axes ticks also to right of plot (blue). (b) Temperatures recorded at Whakapapa ski-station and Chateau station (shifted by 5 degrees). Zero degree marked by horizontal dashed line.

The presence of frozen ground into winter months at Ruapehu may therefore depend on whether the ground was sufficiently frozen before snow accumulated. From Fig.

5.8a, it was observed that velocities at DRZ and FWVZ stations showed a strong correlation with snow cover. This is again shown when overlaying the snow cover profile on velocity changes recorded against frequency at FWVZ station (Fig. 5.24a), which lies in close proximity to the Whakapapa ski-station where air temperature is measured (in addition to snow thickness). While these measurements are only made during winter months, a good fit is found with measurements at Whakapapa village when they're shifted by five degrees (accounting for its lower elevation) (Fig. 5.24b). Thus the measurements at Whakapapa village can be used to approximate the temperature close to FWVZ station. From this, we observe that snow accumulation, and subsequent velocity changes, occurs approximately the same time that temperatures likely fell below zero degrees close to the station. This would suggest that significant ground freezing was less likely prior to snow accumulation. Similarly, maximum velocities coincide with larger snow thicknesses at both DRZ and FWVZ station (Figs. 5.7a-b, 5.8a), when the ground temperature is likely decoupled from the air temperature. Thus, combining these observations, we suggest that frozen ground induced velocity changes are less likely on Ruapehu.

5.4.2.3 Ground water level changes

Seasonal variations in precipitation, and subsequently the ground water level, have regularly been observed to correlate with seismic velocity changes (e.g. Sens-Schönfelder and Wegler, 2006; Obermann et al., 2014; Rivet et al., 2015; Wang et al., 2017; Clements and Denolle, 2018). This is also observed at MTVZ station, where seismic velocity changes are inversely correlated with pore pressure changes estimated from precipitation records (Fig. 5.8b). This inverse relationship is thought to reflect changes in grain contact in response to pore pressures. Specifically, an increase in pore pressure reduces the contact area between grains, thus promoting a seismic velocity decrease (Grêt et al., 2006; Fokker et al., 2021). These changes can be on the order of 0.1–1.0 %, similar to those observed at Ruapehu.

The presence of snow can have a significant influence on groundwater recharge. At lower elevations, where minimal snow is present, rain may play a dominant role in recharging the groundwater system (Somers and McKenzie, 2020). At higher elevations, such as mountain regions, seasonal snow often plays an important role in groundwater recharge (Earman et al., 2006; Nygren et al., 2020; Somers and McKenzie, 2020). As temperatures drop below freezing, water infiltration is expected to decrease as precipitation begins to fall as snow. Similarly, if the ground was frozen during any time period, this also acts to reduce infiltration through decreased permeability (Bayard et al., 2005; Iwata et al., 2010). In both cases, a reduction in groundwater levels and associated pore pressures can be expected. As temperatures rise above freezing point increased infiltration occurs through snow-melt, with groundwater and pore pressure increases expected (Nygren et al., 2020). This could therefore explain the opposite trends observed at stations on the volcano and those off the volcano (Fig. 5.7). Furthermore, we observe lower velocities at both DRZ and FWVZ stations in summer months (Dec–Feb) following the 2006 and 2007 winters relative to the summer following the 2005 winter (Fig. 5.8a). This may reflect a greater snow melt contribution during spring/summer following 2006 and 2007 winters, where less snow accumulated in 2005 winter.

The magnitude of changes expected from ground water changes, and therefore pore pressure changes, can be described similarly to that of pressure changes in Equation 5.12. From Fokker et al. (2021), a change in pore pressure u via effective stress induces a shear-wave velocity change following:

$$\frac{d\beta}{\beta} = -\frac{\mu'}{2\mu}u. \tag{5.13}$$

Note that with the negative sign, a positive change in pore pressures results in a decrease in velocities. For the shear modulus μ and its derivative μ' used earlier in modeling the snow-loading, velocity increases of $\sim 0.05\%$ were predicted for a pressure change of 5-6 kPa within the upper few hundred meters. Thus, for velocity changes on the order of $\pm 0.5\%$ (i.e. a range of 1%) a pore pressure change of approximately 100 kPa would be required. Similarly, for $\pm 1.0\%$ change, pore pressure changes of 200 kPa would be required. This corresponds to ground water changes equivalent to a level change of 10–20 m. This is not outside the realm of possibility for ground water changes in mountainous regions (e.g. Ofterdinger et al., 2014); however, it is uncertain if changes of this magnitude are present on Ruapehu at the highest elevations. Similarly, there is a high uncertainty with regards to the location of the groundwater table. For example, simulations suggest that the groundwater table may be located as much as a few hundred meters below the surface in mountainous regions (Forster and Smith, 1988; Somers et al., 2019; Somers and McKenzie, 2020). If groundwater recharge was assumed to be responsible for observed velocity changes, the observation of the most significant velocity increases between approximately 100–200m below DRZ station might indicate the approximate location of the groundwater table (Fig. 5.16c). This could then explain the lower velocities resolved in the upper 50–100 m at both DRZ and FWVZ station following the inversion (Fig. 5.16c-d). Alternatively, the presence of the crater lake in close proximity to DRZ may contribute to reduced velocities in the upper 100 m. For the decreasing velocity between 200-300 m from late-September/early-October, this may represent an increase in pore pressures due to a gradually increasing groundwater level following snow-melt. It is unclear though if sufficient melting is occurring at higher elevations during this time, with snow thicknesses measured at 2000 m (approximately 600 m below the summit plateau) still relatively high in early-October.

Groundwater level decreases due to precipitation falling as snow could also explain the initial rapid increase in velocities in June/July (Fig. 5.23a-b). These changes occur during a time period where snow-loading models predict reduced influence relative to August/September months (Fig. 5.23)e-f). If snow-loading dominates, we might expect to see larger velocity changes coinciding with time periods of rapid snow accumulation. This follows that groundwater level decreases due to the absence of water infiltration should not depend on the amount of snow falling. Comparing the velocity changes recorded at DRZ station with snow thicknesses recorded between late-June and September does not yield a definitive conclusion (Fig. 5.25). There is some evidence that velocity increases occur more rapidly while snow is accumulating and remain similar while snow thicknesses are more constant (for example, in 2007–2008, (Fig. 5.25c-d)). Furthermore, a rapid velocity increase of approximately 0.6% occurs in July 2006 at the same time approximately 1 m of snow accumulates at both ski stations (Fig. 5.25b). Such a large change is hard to justify as related to

reduced water infiltration, requiring an almost 100 kPa change in pore pressures (or 10 m water level change). However, there does not appear to be a clear relationship between snow thicknesses and the magnitude of velocity changes when comparing individual years. In 2008, for example, significant snow accumulation – comparable or greater than 2006 – occurred during early-July and through August (Fig. 5.25d). Velocity increases occurring at approximately the same time are of lower magnitude (0.2–0.3 %) relative to 2006. Thus, it is difficult to identify a clear relationship between snow thickness and velocity changes. Furthermore, shorter-term variations should be treated with greater caution relative to longer-term variations within the frequency band of tremor. Thus larger, rapid, changes in velocity (for example in 2006) may not be reliable. We do note, however, that the maximum velocities recorded in the winter of 2005 — which experiences smaller snow thicknesses — are lower than those of later winters (Fig. 5.8a). At DRZ station, for example, velocities recorded in winter are 0.2–0.3 % lower than later winters (Figs. 5.8a and 5.25). This may support an influence of snow-loading on velocity changes.

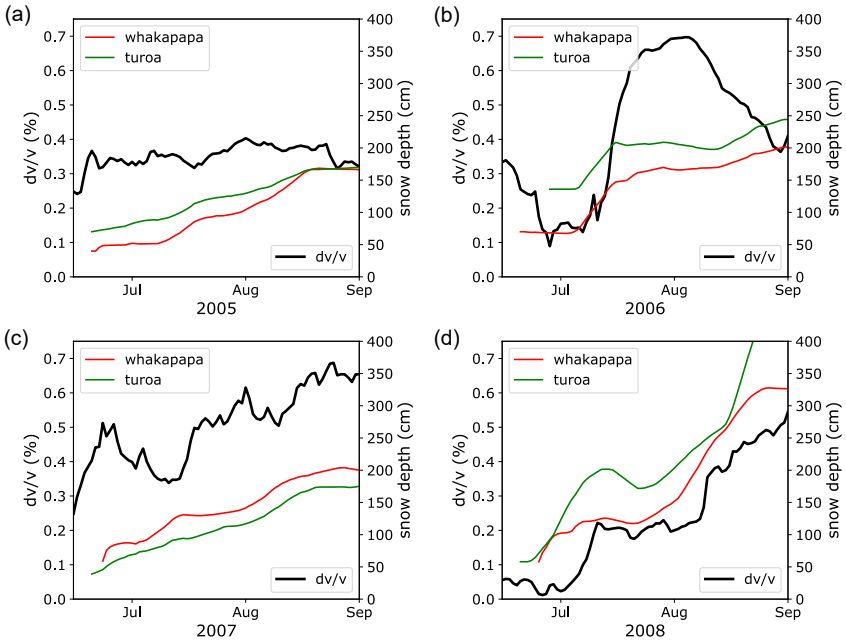

Figure 5.25: Comparison between velocity changes recorded at DRZ and snow thicknesss at both Whakapapa and Turoa ski stations for months where temperatures are close to, or below, zero degrees celsius. (a) 2005. (b) 2006. (c) 2007. (d) 2008.

At stations at lower elevations, the seasonal trend of velocity changes appear to be consistent across a broad frequency range (Fig. 5.9). At MTVZ, for example, the higher velocities beginning in January — and lasting until July — are observed at all frequencies between 0.5–4.0 Hz (Figs 5.7c and 5.9c). This is similar for all stations aside from those at higher elevations (DRZ, FWVZ, and TRVZ). We do, however, observe slightly higher velocities between approximately 0.5–1.5 Hz at these stations between January and July (Fig 5.9a, 5.9b,5.9f) prior to larger increases of velocities in winter. These higher velocities earlier in the year are clearly seen in the

depth inversions for DRZ and FWVZ, at depths between approximately 200–600 m (Figs 5.15a,c and 5.16c-d). This could be indicative of the same seasonal trend as observed at lower elevations. However, it is unclear why different behaviour would be observed at different frequencies/depths on the volcano if ground-water changes were to dominate.

One consideration relates to the use of a dynamic maximum lag time based on 20 cycles of a given period to define the end of the coda. This could influence the sampling radius for different frequencies, with lower frequencies (longer periods) sensitive to a wider area than higher frequencies (shorter periods) due to longer travel-paths of scattered waves. Lower frequencies may then sample a wider portion of the medium down-slope of the station. This may also be expected given the propagation direction of the initial tremor source is away from the crater lake/summit area. Thus scattered waves may preferentially sample down-slope regions of the volcano. To test the influence of the dynamic lag time window, the results of DRZ station are compared with those obtained using a static lag time (Fig. 5.26). From this, elevated velocities earlier in the year — below approximately 1.25 Hz — are still recovered using static lag times of 5–15 s and 5–25 s (Fig. 5.26b and 5.26c). Similarly, the earlier decrease between approximately 1.0–1.5 Hz from late-September/early-October is also still present. This suggests that the choice of a dynamic lag is not a significant factor in resolving this feature. However, this does not take into account that seismic velocities of surface waves tend to increase with decreasing frequency. Thus, lower frequencies may still have wider areas of sensitivity even with a static coda window for all frequencies. Furthermore, laterally varying scattering properties can have a significant impact on the spatial sensitivity of coda waves (van Dinther et al., 2021). Thus, a strongly directional seismic source, combined with a non-uniform scattering medium, could also contribute to complex observations.

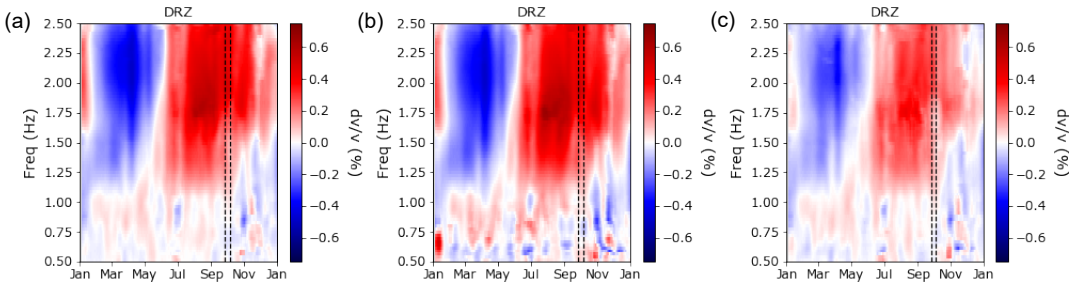

Figure 5.26: Comparison between measurements made using (a) dynamic coda window and static lag times windows of (b) 5–15 seconds and (c) 5–25 seconds.

Without further understanding of the groundwater distribution on Ruapehu, it is difficult to determine the validity of groundwater changes towards explaining all the various observations. Similarly, limited understanding on the true distribution of snow on the volcano (in both space and time) makes it difficult to fully model this process. Both processes appear able to explain some of the observations, while also exhibiting limitations in their ability to account for certain phenomena. An overview of the arguments in favour and against each mechanism discussed are presented in Table 5.3. From this, we consider the most likely scenario is a combination of both

ground water level changes and snow-loading contributing to seasonal velocities on Ruapehu. Ground water changes, dominated by reduced infiltration during winter months and later snow-melt, appear more likely towards explaining the magnitude of velocity changes observed and also the timing of the initial velocity increase. However, we find evidence for an imprint of velocity changes due to snow-loading; in particular, the lower winter maximum in 2005 relative to other years. However, determining the relative contribution of each process would require a more comprehensive understanding of spatial velocity changes, snow distributions, and sub-surface hydrology, on and around the volcano. Future application of geophysical surveys towards this purpose (for example, airborne electromagnetic surveys), would be beneficial towards this goal.

Mechanism	For	Against
Snow-loading	• Consistent with winter velocity increases within upper few hundred meters only. • Decrease at lower frequencies (∼1.0–1.5 Hz) from late-September/early-October can be modeled as reduced depth sensitivity due to decreasing snow radius. • Reduced snow accumulation in 2005 (relative to other years) coincides with lower maximum winter velocities (0.2–0.3 % lower)	• Rapid increase of velocities in June/July not well predicted by snow-loading model. • Magnitude of velocity changes estimated as due to snow-loading far smaller than those observed. • Relationship between size of snow-load and observed velocity changes not consistent.
Frozen ground	• Can explain velocity increase occurring early in winter as temperatures fall below zero.	• Snow thicknesses similar to those recorded on Ruapehu reduce the chance of frozen ground forming due to insulating effect. • Shear-wave velocity increases in frozen andesitic rock may not be large enough to significantly influence velocity changes at frequencies between 0.25–2.5 Hz.
Ground Water Level Changes	• Can explain velocity increase occurring early in winter as precipitation begins to be stored as snow above the surface. • Magnitude of pore pressure changes required to match observed velocity changes within possible range.	• Unclear why winter increase would only be observed at shallow depths (higher frequencies), with greater depths more consistent with precipitation-induced seasonal trend at lower elevations. • Decrease at lower frequencies (∼1.0–1.5 Hz) from late-September/early-October may be too early at the highest elevations to be explained by gradual recharge from snow-melt.

Table 5.3: Summary of arguments for and against the different mechanisms that could influence seasonal velocity changes on Ruapehu.

5.4.3 Possible links to eruptive activity

The study of seasonal changes at Mt Ruapehu is motivated by evidence of seasonality in the eruptive record. Using a de-clustered catalog of events with eruption magnitude 2+ since 1960 (see Table S5.2 for magnitude definitions), we observe that over 50% of the eruptions (out of 49) occur within the spring months of September, October, and November (Fig. 5.1c). The majority of these are magnitude 2 events, corresponding to phreatic or phreatomagmatic eruptions that deposit material outside the Crater Lake, but are still confined to the crater basin. Of all the magnitude 2 events, approximately 30% (out of 34) occur in the single month of October, and 60% within the three spring months. Full details of how the catalog was de-clustered, with justifications, are provided in Section S5.3.

The high proportion of eruptions in spring months, and in particular October, appear at odds with a uniformly distributed pattern of eruptions in time. This is supported by low p-values using two statistical tests - Fisher's exact test and the Rayleigh test – against the null hypothesis of uniformly distributed eruptions throughout the year (Section S5.4). This follows the expectation that, in a de-clustered catalog, the chance of an eruption on any given day is equal. For the Fisher's exact test, the probability of obtaining the observed distribution of eruptions in spring (versus the rest of the year) against the null hypothesis of a uniform distribution is calculated (Table S5.3). P-values of 0.01 are obtained when considering all magnitudes 2–5 and <0.01 for magnitudes 2–3. For the Rayleigh test, the timing of eruptions are represented as vectors on a unit circle (binned by months, Fig. S5.18). Assuming the data are unimodal, the test assesses whether the data exhibit significant clustering around a certain direction using the mean length of the data vectors. From this, p-values of 0.01 and <0.01 are also obtained for all eruption magnitudes 2–5 and magnitudes 2–3 respectively. Thus, these tests suggest a preferential timing of eruptive activity during the year. Full summary of p-values can be found in Table S5.4.

The earlier velocity decreases from September between 200-300 m — observed at DRZ and FWVZ (Fig. 5.23c and 5.23d) may provide a clue as to a triggering mechanism (if such a seasonal trigger were to exist). The most recent eruptions in 2006 and 2007 occurred in early-October and late-September respectively. The timing of these events when overlain on the depth inverted velocity changes (using the average velocities between 2005–2009) reveals they occur after velocities have began to decrease. Thus it is possible the same process driving this decrease may be related to increased eruptive activity in spring months (directly or indirectly). From earlier discussion (Section 5.4.2), it is unclear whether this change relates to snow-loading or groundwater recharge. Thus both mechanisms are considered as possible triggers of greater eruptive activity in spring.

The impact of melt-water on volcanic systems can be considered similar to the impact of rainfall. Multiple studies have linked increased volcanic activity to precipitation events (e.g. Violette et al., 2001; Matthews et al., 2002; Hort et al., 2003; Elsworth et al., 2004; Matthews et al., 2009; Gaete et al., 2020; Caricchi et al., 2021). Commonly invoked mechanisms are (1): percolation of meteoric water into the hydrothermal/magmatic system, where it interacts with hot rock and fluids (e.g. Hort

et al., 2003; Gaete et al., 2020), and (2) pressure build-up due to infiltrating water blocking the escape of magmatic gases (e.g. Elsworth et al., 2004; Matthews et al., 2009). The second mechanism, relating to blocked pathways for magmatic gas, is thought to be a major contributer to the collapse of lava domes (Matthews et al., 2002; Elsworth et al., 2004; Matthews et al., 2009). The interaction of infiltrating meteoric water with the hydrothermal-magmatic system, on the other hand, has been suggested as a trigger for phreatic activity at Lascar volcano (Gaete et al., 2020) and greater eruption durations and explosivity at Stromboli (Hort et al., 2003). We consider this mechanism more relevant for Mount Ruapehu, noting that the main vents responsible for gas emissions are within the crater lake (Christenson, 2000; Werner et al., 2008). Thus it is not expected additional groundwater recharge will block pre-existing gas pathways. Furthermore, seasonal seismicity has been observed at Ruapehu's neighbour volcano Ngauruhoe, suggested to result from the interaction of melt water with local heat sources (Jolly et al., 2012).

The timing of increased eruptive activity from the start of Spring (September/October), with minimal activity in summer months (December, January, February) (Fig. 5.1c), appears at odds with an interpretation of a groundwater recharge related trigger. From satellite imagery, we can see that snow cover is still present over the wider edifice by the start of Spring (see orange lines in Fig. 5.1a). Similarly, only a small amount of snow-melt is suggested at an elevation of 2000 m based on measurement of snow thickness. Thus, it is not anticipated a significant amount of snow melt is infiltrating the ground within the summit plateau (altitude approximately 2600 m) at the time of eruptive activity. Instead, we might expect greater water infiltration into the hydrothermal system during late-Spring/Summer months as the snow-line recedes back towards the summit plateau. This is also consistent with the repeating seismicity observed at Mt Ngauruhoe between 2005–2009, where peak seismicity rates were from November to February, with peaks in January (Jolly et al., 2012). Given that Ngauruhoe is lower in elevation by approximately 500 m, the timing of significant melt water around the summit of Ruapehu could be even later than this. A melt-water related trigger in the first months of Spring, with minimal eruptive activity in Summer months, therefore seems an unlikely trigger mechanism for seasonal eruptive activity at Ruapehu.

Seasonal snow-loading (or unloading) has previously been suggested as a possible mechanism driving eruptive activity at both Katla volcano in Iceland (Albino et al., 2010) and Villarrica volcano in Chile (De Groote, 2022). At Katla, nine large eruptions occurring since 1580 occurred during periods from May to November, corresponding to approximately summer-autumn months associated with reduced ice loads (Larsen, 2000; Albino et al., 2010). At Villarrica, similar to Ruapehu, a far greater number of eruptions since 1750 have occurred during spring months, relative to other seasons (De Groote, 2022). In both cases, it is considered that unloading can act to bring a magma chamber closer to failure, promoting dyke initiation (Albino et al., 2010; De Groote, 2022).

It seems unlikely that failure within the wall of a magma chamber due to snow-unloading is the mechanism responsible for driving seasonal activity at Ruapehu. First, most of the activity that occurs during spring months are smaller, non-magmatic, events (Fig. 5.27). Similarly, both the previous two eruptions in 2006

and 2007 considered to be gas-driven events (Jolly et al., 2010; Christenson et al., 2010). Thus we do not anticipate significant magma movement from depth towards shallow levels is involved in driving individual eruptions. In contrast, the eruptions at Katla were much larger (mostly VEI 4–5, (Global Volcanism Program, 2023)), with Albino et al. (2010) considering only eruptions fed by dykes. Accordingly, we turn our attention towards a mechanism that can better explain the smaller size of the eruptions in spring.

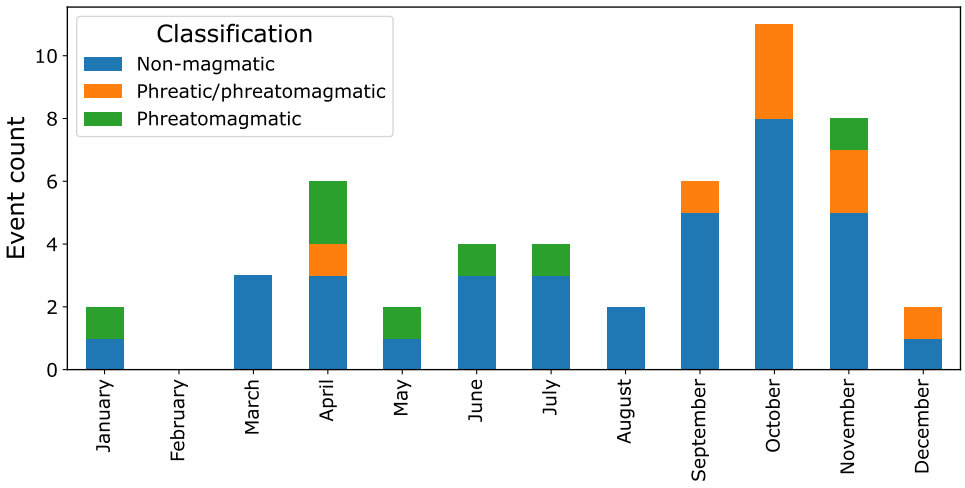

Figure 5.27: Classification of de-clustered eruptions using descriptions from catalog of Scott (2013).

A possible clue can be found in the relationship between tidal stresses and magmatic degassing. Increases in gas emissions have been identified at a number of volcanoes associated with tidal patterns (e.g. Cigolini et al., 2009; Bredemeyer and Hansteen, 2014; Dinger et al., 2019). This has been interpreted as due to magma displacement within the magma conduit due to the cycles of compression and extension induced by tidal stressed (Michael and Christoffel, 1975; Bredemeyer and Hansteen, 2014; Girona et al., 2018). The effect of differential flow of magma is to promote bubble coalescence and the separation of gas from magma, thus enhancing volatile degassing (Bredemeyer and Hansteen, 2014; Dinger et al., 2019). This is suggested to be more significant at depths of several kilometers, where tide-induced bubble coalescence at shallow depths may be negligible compared to classical transport mechanisms (Dinger et al., 2019). Furthermore, this mechanism has previously been proposed to explain a correlation between eruptions at Mt Ngauruhoe and fortnightly earth tide cycles (Michael and Christoffel, 1975).

Tidal stresses are typically on the order of 0.1– 10 kPa (Dinger et al., 2019) whereas modeled pressure changes from snow-loading are on the order of (1–8 kPa) (Fig. 5.28a-c). Thus, the pressures from the snow load are within the uppermost range of anticipated tidal stresses. The timescales between the two processes differ though, with snow-loading (and unloading) occurring on time scales of months compared to fortnightly tidal minima and maxima. This may have implications on whether the processes could be expected to similarly influence bubble coalescence through the displacement of magma. Computing a monthly gradient of pressure changes (at z=2 km) from the snow model outputs shows changes of approximately 1 kPa

per month for most of the time period between July and October (Fig. 5.28d-f). In July and August the gradient is a positive, with increasing pressures through time, and from September through the start of October the gradient is negative, with decreasing pressures through time. From this, we could expect biweekly stress changes of approximately 0.5 kPa, which is within the range of tidal stresses. Thus, it seems plausible that snow-loading, or unloading, could induce similar movement of magma inferred to be responsible for increased degassing associated with tidal cycles.

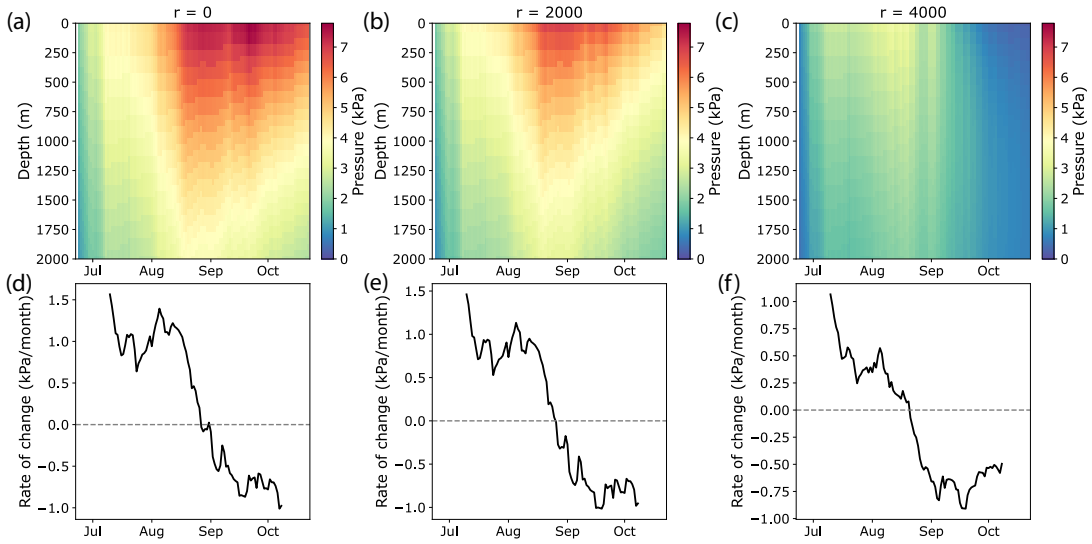

Figure 5.28: Pressure change calculations at different radius r between June 20 and October 21 for sigmoid-shaped snow-load and corresponding pressure gradient for depth z= 2000 m. (a) pressure change at r = 0 m. (b) pressure change at r = 2000 m. (c) pressure change at r = 4000 m. (d) pressure gradient at r=0, z=2000 m. (e) pressure gradient at r=2000, z=2000 m. (f) pressure gradient at r=4000, z=2000 m.

Decreasing pressures as a result of snow-unloading may therefore be a plausible mechanism for driving seasonal eruptive activity at Ruapehu. Modeling the snow-load suggests the corresponding pressure changes propagate to greater than 2 km depth (Fig. 5.28a-c), and thus likely interact with magma storage zones beneath the volcano (suggested to be between $\sim$ 2–9 km depth (Kilgour et al., 2013)). This may enhance the separation of gas from magma, which rises to interact with the shallow hydrothermal system and therefore increase the likelihood of gas-driven eruptions. For such a model, the lack of eruptions in July/August — where high positive pressure gradients are recorded — would suggest the transition from increasing pressure to decreasing pressure gradients is significant (Fig. 5.28d-f). One possibility would be that the presence of the load acts to confine the shallow hydrothermal system. This could act to delay any eruption about to occur until the confining pressures decrease through Spring.

The relative absence of eruptive activity in the months of December through February (Fig. 5.27) may then be interpreted within the context of this loading/unloading model. If degassing processes are most significant during the months from July to

October, it follows that the likelihood of reaching a critical state within the hydrothermal system is elevated during this time. The removal of the load may then serve as a trigger mechanism to initiate an eruption if such a critical state is achieved. However, if no eruption is triggered, it would imply that the system did not reach a critical state during the snow-loading/unloading phase. Thus, combined with the absence of loading-induced degassing, the likelihood for gas-driven activity would be expected to be reduced in the months following.

One test for such a model would be to examine the yearly trend of measured gas fluxes. If snow-loading enhanced degassing is occurring, it could be expected that this would also be visible during periods where the volcano is at background levels. From SO2 flux data recorded between 2010–2022 (chosen to exclude activity in 2006–2009), we observe some evidence for enhanced gas flux associated with snow-loading (Fig. 5.29). For example, the months of August, September and October are associated with greater variance in the recorded gas flux relative to the months of November through to March. Gas fluxes recorded from January and February in particular are consistently very low, possibly supporting an interpretation that the absence of a snow-loading/unloading mechanism during summer leads to reduced degassing. However, it is important to note that only a few measurements are available for each month over the 12-year period, owing to an average sampling interval of approximately 60 days. Thus, it is difficult to assess the reliability of the trend, with greater resolution needed to confidently determine a causal link between snow-loading and enhanced degassing at Ruapehu.

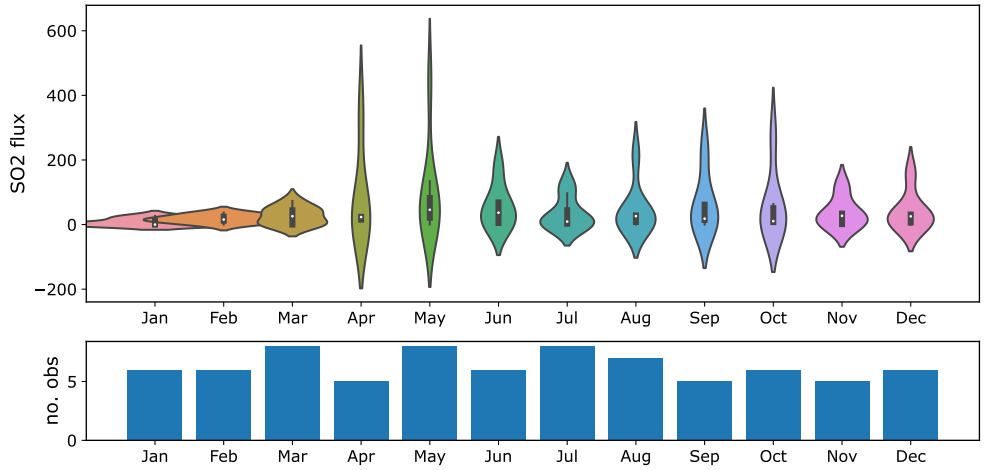

Figure 5.29: Violin plot of SO2 flux by month recorded at Ruapehu. Note, negative values in the kernel density plots are artifacts arising from flux values being close to zero.

From these various observations, we consider that snow-loading/unloading is a possible trigger for seasonal activity at Mount Ruapehu that should be explored further. One obvious question, though, is why we do not see such eruptions more regularly if annual snow is a trigger. For example, no eruption has occurred within the last

15 years, with the last recorded activity in the catalog being a volcanic earthquake that induced a lake level increase in July 2009 (Scott, 2013). Examining the de-clustered catalog through time shows that many of the eruptions that occurred in spring months are clustered in time (Fig. 5.30). For example, the time period from mid-1960s to early-1980s contains the majority of events. Similarly, we observe multiple instances where, just like in 2006–2007, spring eruptions occur with minimal activity in between. For example, in 1972–1973, 1980–1981, and 1997–1998. Thus we suggest that any external triggering mechanism likely requires that the volcanic system is already in an unstable state. With regards to the most recent eruptive activity, both the 2006 and 2007 phreatic eruptions were preceded by deep seismic swarms (5–15 km) in the 3–5 months before the eruption (Kilgour et al., 2014). These swarms were likely associated with pressurization within the magmatic system, possibly then making the system more prone to external triggering. Additionally, the presence of a seal — such as has been interpreted for both the October 2006 (Ardid et al., 2023) and the September 2007 eruptions (Christenson et al., 2010; Girona et al., 2018; Ardid et al., 2023) — may also be a necessary pre-requisite to allow pressure to buildup within the shallow hydrothermal system.

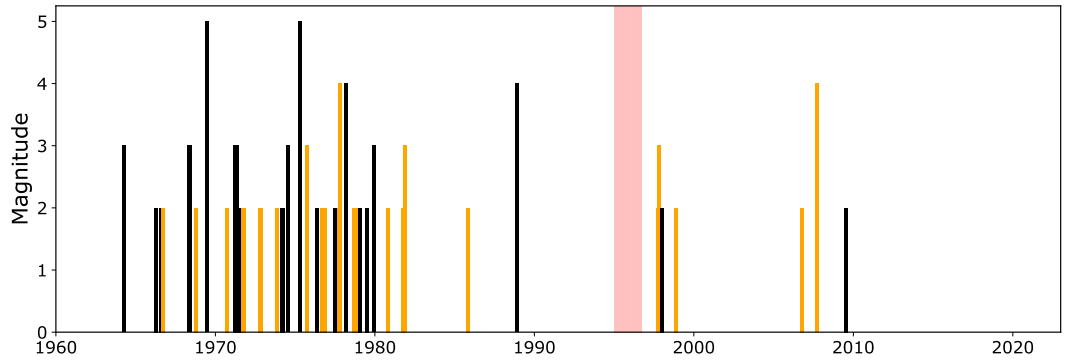

Figure 5.30: Eruptions from de-clustered catalog through time. Orange bars correspond to spring eruptions, other eruptions shown as black bars. Red shaded region in 1995/96 corresponding to removed activity during de-clustering process (Section S5.3).

In summary, our analysis indicates a potential link between seasonal snow cycles and the likelihood of eruptive activity at Ruapehu. This interpretation should clearly be explored further, both at Ruapehu and other volcanoes that similarly show increased activity following the removal of a seasonal snow-load (e.g. Villarrica, (De Groote, 2022)). In any case, the observation that seasonality may exist in the eruptive record has important implications in terms of risk/hazard assessment. If this is the case, it could be a valuable inclusion as a parameter in automated probabilistic forecasting model already in use at Mount Ruapehu (Christophersen et al., 2022).

5.5 Conclusions

Evidence of preferential timing in the eruption chronology at Mount Ruapehu suggests the possibility of seasonal triggering of small-scale eruptions at the volcano.

Spring months show the greatest activity, with over 50% of eruptions recorded in a declustered eruption catalog having occurred in the months of September, October, and November. This rises to 60% when considering only smaller events, with 30% contained within just the single month of October. The timing of these eruptions (in Spring) suggests there may be some connection to the seasonal snow-cycles, with snow unloading previously suggested as a possible trigger for eruptive activity.

We study single-station seismic velocity changes over a four year period (2005–2009) that contains two gas-driven eruptions which occurred approximately one year apart in Spring months. To better handle the influence of variable tremor at higher frequencies ($>$ 1 Hz), the cross-wavelet transform approach is used to measure delay times in cross-correlation functions. In addition to providing greater frequency resolution, the method allows data to be excluded in the time-frequency space when incoherent, while keeping data during time periods where tremor shows relative stability.

Substantial seasonal changes are observed that differ on and off the volcano. At stations on the volcano, velocities vary by ± 0.5 %, reaching a maximum during greater snow cover and thicknesses in winter months. At stations off the volcano, velocities more closely follow modeled pore pressure changes associated with precipitation events. Snow-loading, frozen ground, and changes in ground-water content are considered as possible mechanisms driving seasonal velocity changes on the volcano. Following modeling of the influence of a snow-load, ground-water-level changes are considered the most likely mechanism responsible for seasonal velocity changes on the volcano, owing to reduced ground-water infiltration during winter. A smaller contribution from snow-loading/unloading is also evidenced. Improved understanding of both the distribution of snow on the volcano and the groundwater distribution would be beneficial towards determining the relative contribution of each process.

While ground-water related changes are suggested to have a stronger contribution to seasonal velocity changes than snow-loading/unloading, the latter process is considered more likely as a trigger of greater spring-time activity. This is because we expect minimal snow-melt at the highest elevations is expected in September/October relative to later months (e.g. Summer), where minimal eruptions have occurred. In comparison, modeling suggests snow-loading/unloading can induce stress changes at depths of multiple kilometers at rates similar to those of tidal stresses (previous suggested to induce increased degassing). Thus we consider that increased degassing associated with loading and unloading the volcanic system may contribute to increased likelihood of spring-time, gas-driven eruptions. The absense of regular yearly activity, combined with evidence of clustering in time of spring-time eruptions, suggests specific conditions may be required for such a trigger to be significant. For example, greater instability in the volcanic system due to an influx of new magma at depth, combined with the presence of a shallow hydrothermal seal as proposed for gas-driven eruptions in 2006 and 2007, may be significant factors. These interpretations should clearly be explored in further detail, both at Ruapehu and other volcanoes that show similar preferential spring-time activity. Doing so could have important implications for risk/hazard assessment at volcanoes that experience seasonal snow cover.

5.6 Supplementary material

S5.1 Methods

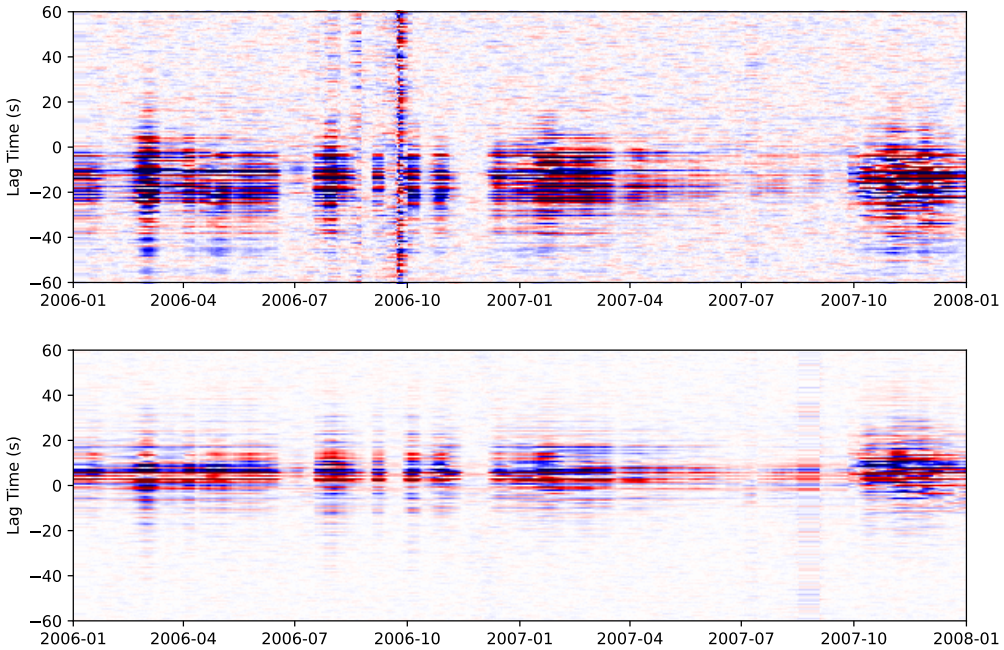

Figure S5.1: Correlogram of station-pair MTVZ-TRVZ (top) and FWVZ-WPVZ (bottom).

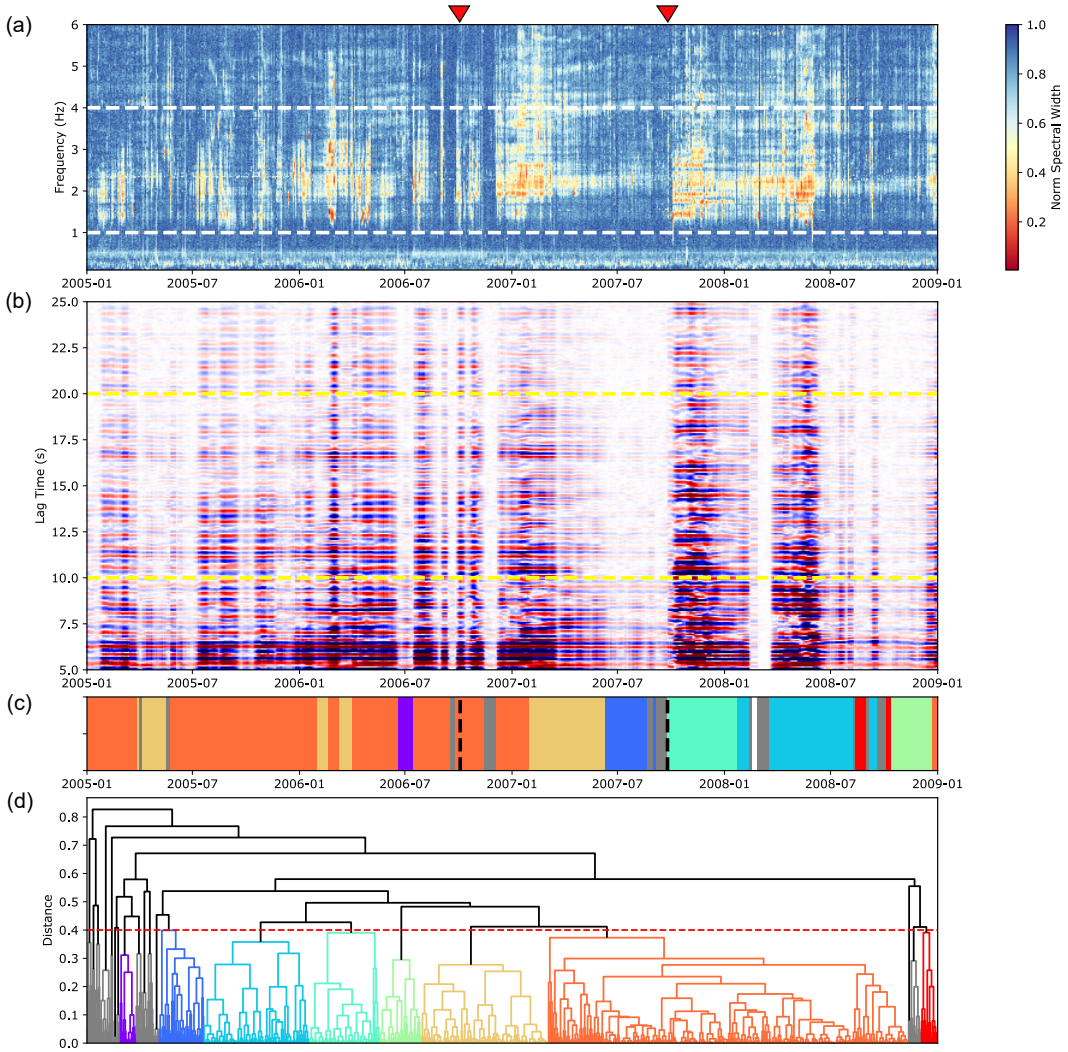

Figure S5.2: Hierarchical clustering output for 1.0-4.0 Hz cross-correlation functions (CCFs) computed for station-pair FWVZ-WPVZ. Dashed-white lines show frequency range of CCFs used in clustering. (a) Normalized spectral width measurement. Lower values indicate a more coherent seismic wavefield dominated by fewer seismic sources. Dashed white lines show frequency range of CCFs. (b) Correlogram showing amplitudes of CCFs at positive lag times which contain tremor signal. Dashed-yellow lines show part of CCFs used in clustering. (c) Location of clusters in time, color-coded according to dendrogram output in (d). (d) Dendrogram, with clustered defined at distance threshold of 0.4 (dashed red-line).

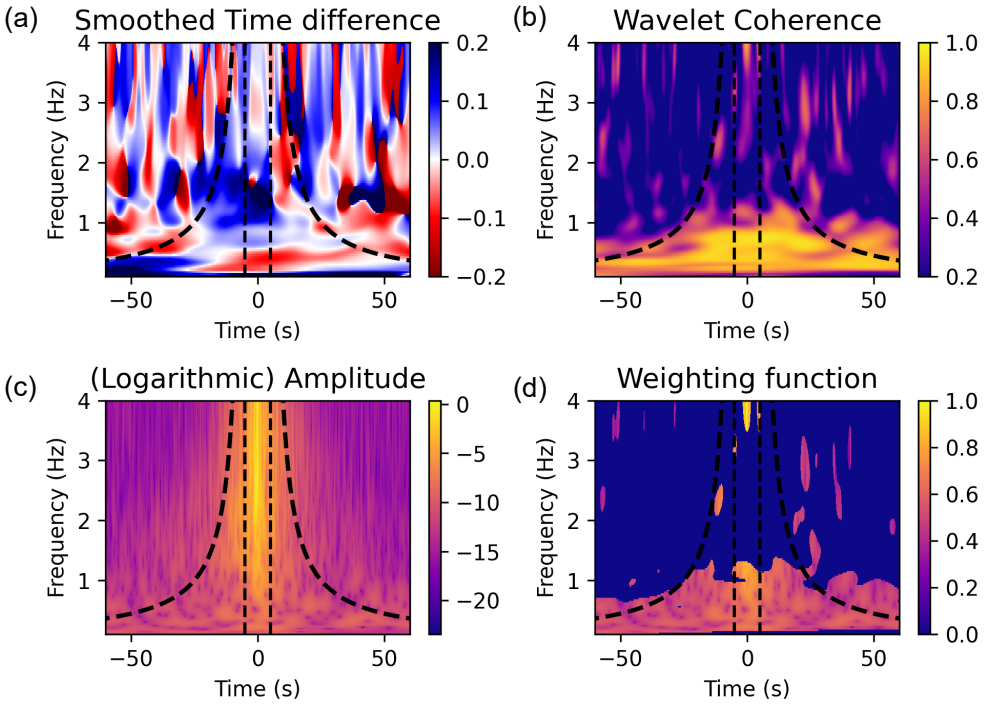

Figure S5.3: Example of delay-time measurements using wavelet transform for a cross-correlation function during period of reduced tremor (2006-07-09) at DRZ station (EN components) (a) Time-delay measurements in time-frequency space. (b) Coherence between the current and reference cross-correlation function. (c) Amplitude of wavelet cross-spectrum used in determining the weighting function for linear-regression. (d) Weighting function.

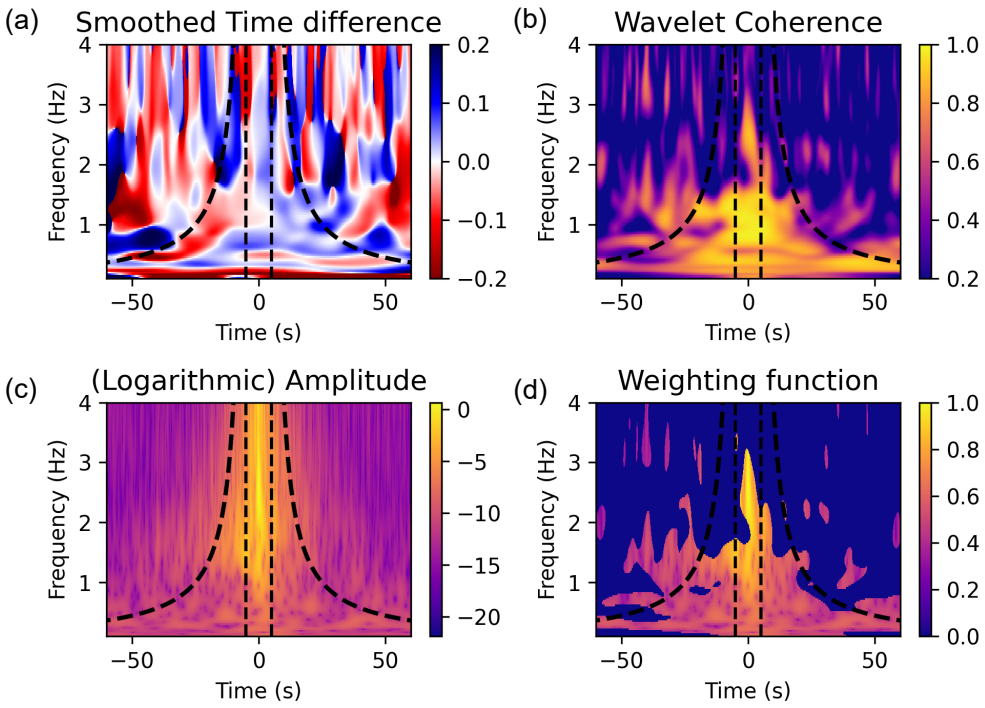

Figure S5.4: Example of delay-time measurements using wavelet transform for a cross-correlation function during period where tremor characteristics differ from longer-term average (2008-01-01) at DRZ station (EN components) (a) Time-delay measurements in time-frequency space. (b) Coherence between the current and reference cross-correlation function. (c) Amplitude of wavelet cross-spectrum used in determining the weighting function for linear-regression. (d) Weighting function.

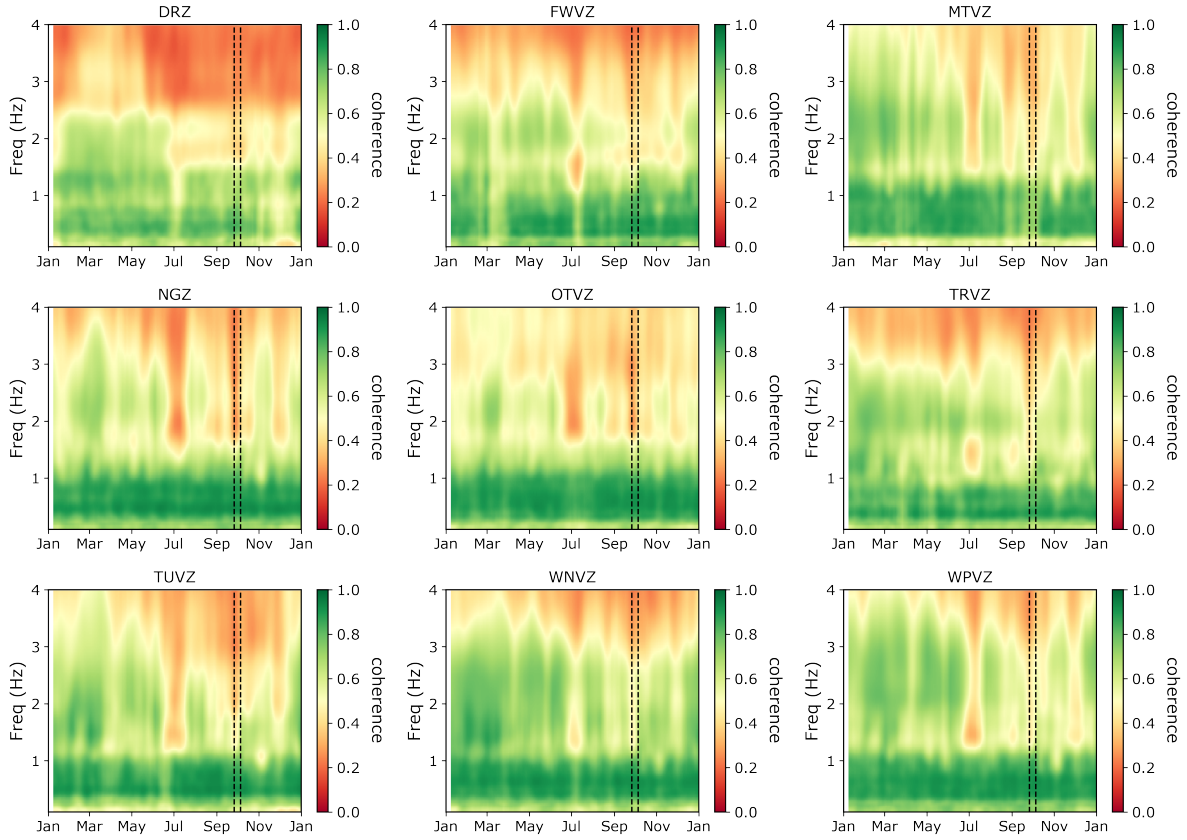

Figure S5.5: Average yearly coherence measured at each station.

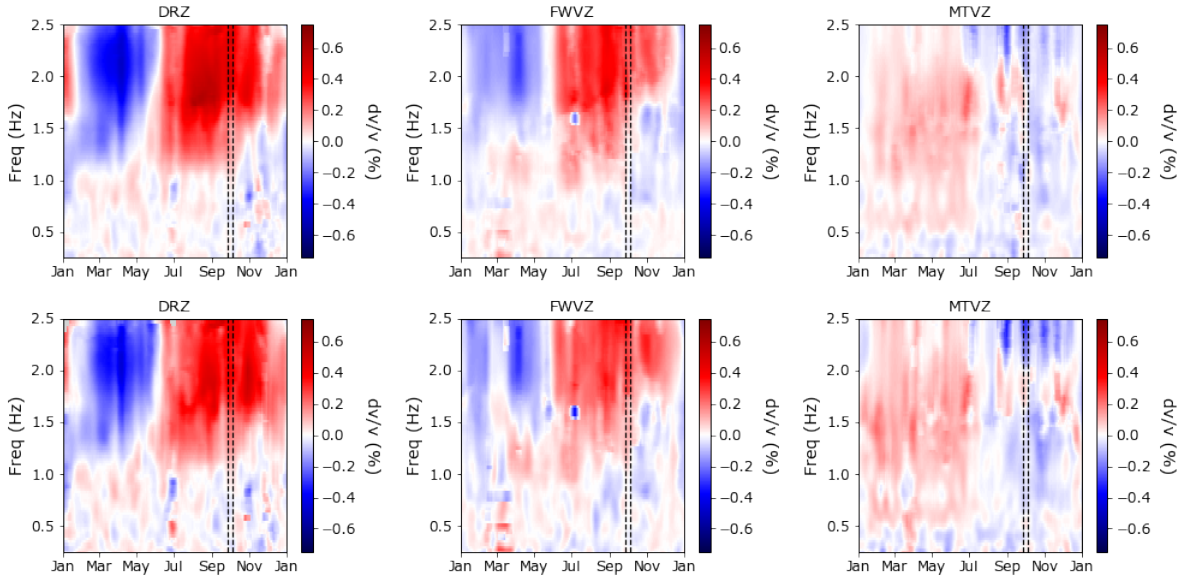

Figure S5.6: Top row: Average yearly velocity changes at DRZ, FWVZ, and MTVZ stations during years 2005–2009. Bottom row: Average yearly velocity changes at DRZ, FWVZ, MTVZ after removing 2007 and 2008, i.e. the years of the eruptions.

Table S5.1: Bandpass filters used for measuring amplitude ratio.

Central Freq	**Low**	**High**
0.2	0.141	0.282
0.225	0.159	0.318
0.25	0.176	0.353
0.275	0.194	0.388
0.3	0.212	0.424
0.35	0.247	0.494
0.4	0.282	0.565
0.5	0.353	0.707
0.6	0.424	0.848
0.7	0.494	0.989
0.8	0.565	1.13
0.9	0.636	1.27
1.0	0.707	1.41
1.2	0.848	1.69
1.4	0.989	1.97
1.6	1.13	2.26
1.8	1.27	2.54
2.0	1.41	2.82
2.5	1.76	3.53
3.0	2.12	4.24
3.5	2.47	4.94
4.0	2.82	5.65
5.0	3.53	7.07
6.0	4.24	8.48
7.0	4.94	9.89
8.0	5.65	11.31

S5.2 Snow-load modeling

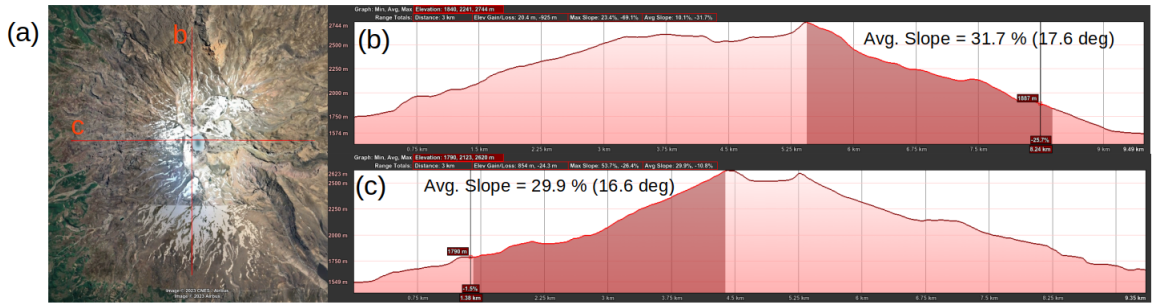

Figure S5.7: Slope measurements along a N-S and W-E profile at Mt Ruapehu, using Google Earth. (a) Map view of ruapehu with two profiles shown. (b) N-S elevation profile, with average slope shown over first 3 km steeper south edifice from summit plateau. (c) W-E profile, with average slope shown over first 3 km of steeper west edifice from summit plateau.

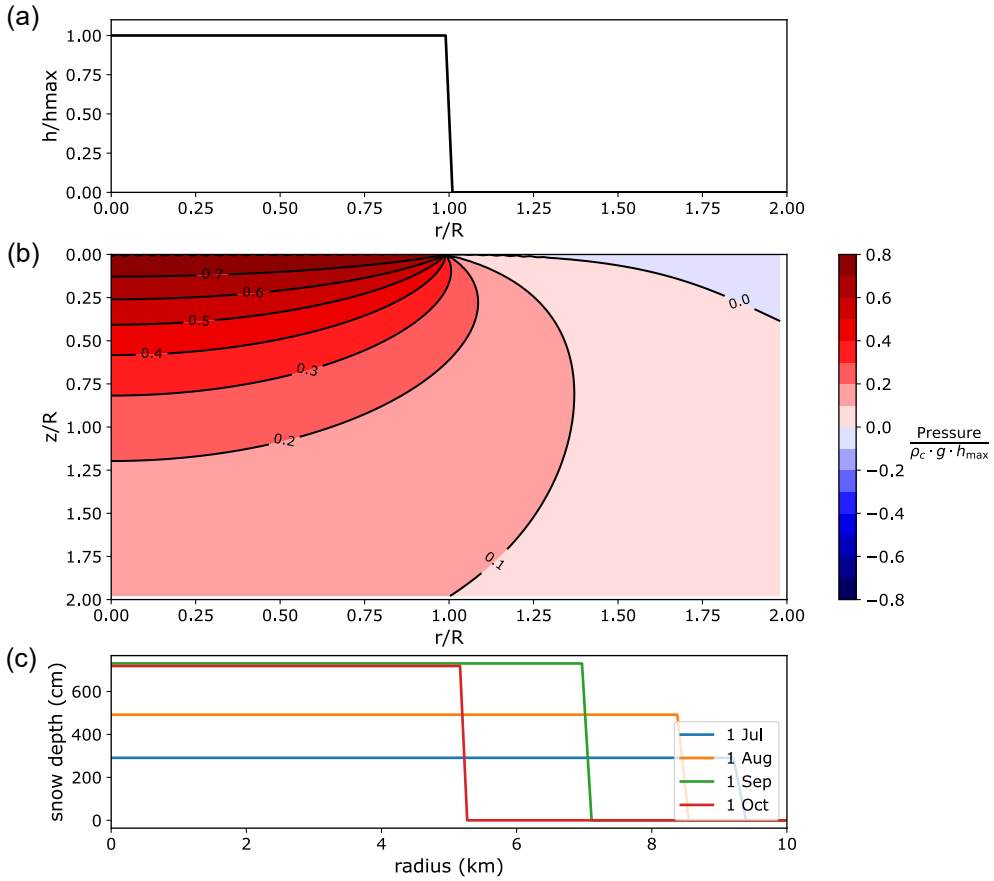

Figure S5.8: Model inputs for axisymmetric snow load (with uniform shape) applied at the surface and corresponding dimensionless pressures calculated for semi-infinite elastic medium. (a) Dimensionless load with sigmoid shape. (b) Dimensionless 2D pressure calculations plotted against radius and depth (both also dimensionless). (c) Surface load at different times of year using R inferred from MODIS satellite data and h_{max} inferred from ski-station snow thickness measurements.

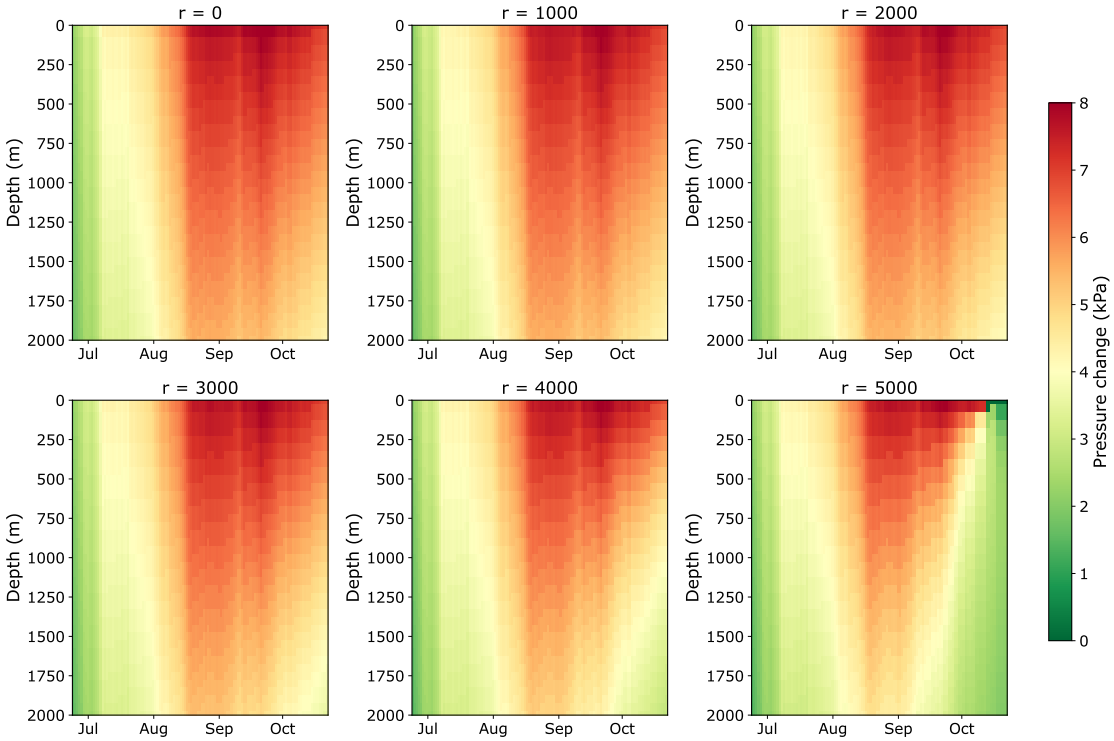

Figure S5.9: Pressure change calculations at different radius r between June 20 and October 21 for uniform shaped snow-load.

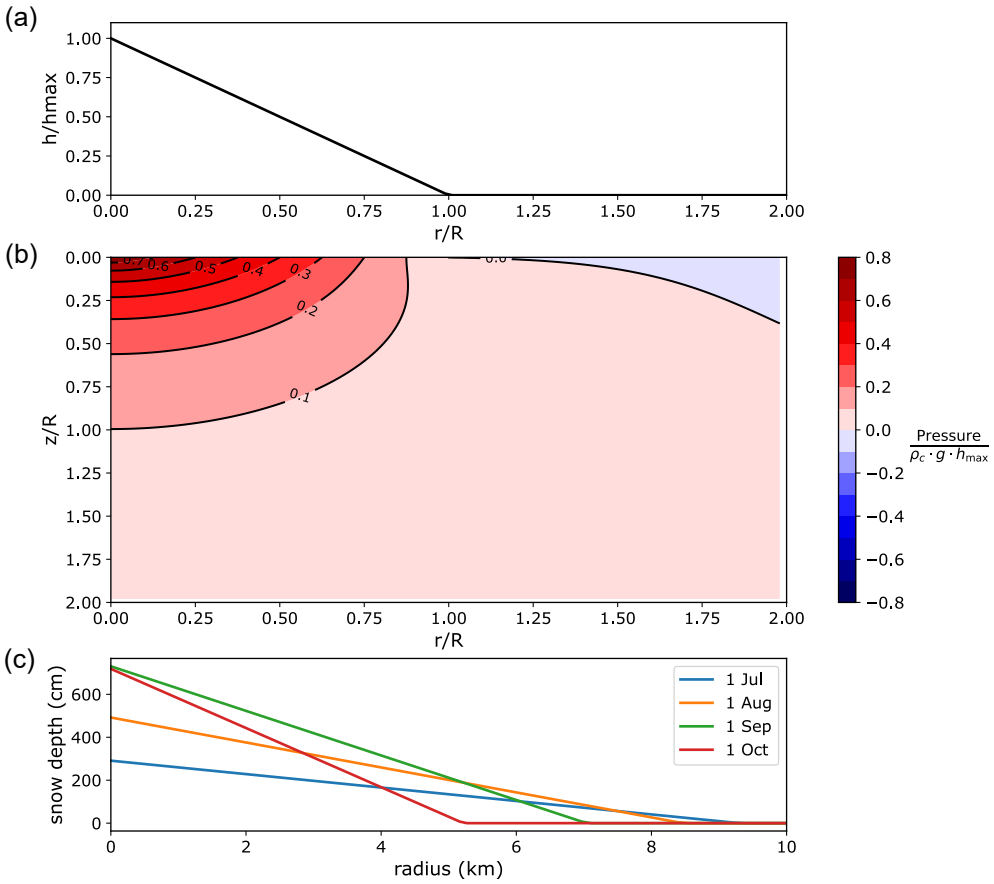

Figure S5.10: Model inputs for axisymmetric snow load (with cone shape) applied at the surface and corresponding dimensionless pressures calculated for semi-infinite elastic medium. (a) Dimensionless load with sigmoid shape. (b) Dimensionless 2D pressure calculations plotted against radius and depth (both also dimensionless). (c) Surface load at different times of year using R inferred from MODIS satellite data and h_{max} inferred from ski-station snow thickness measurements.

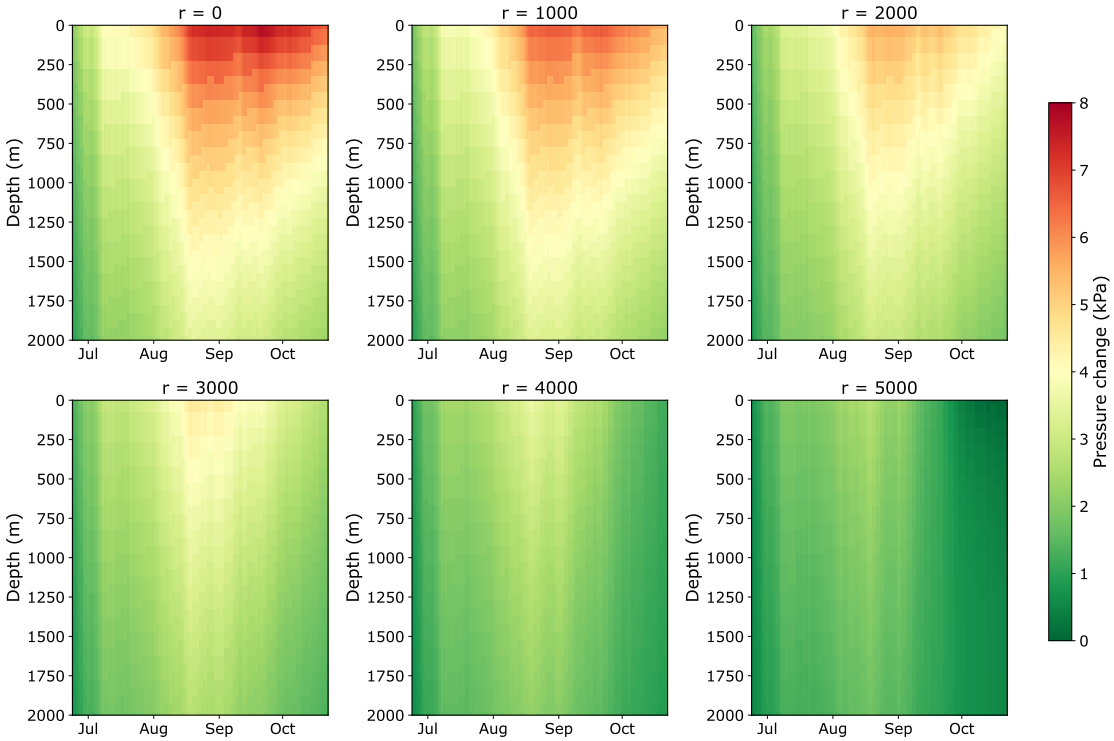

Figure S5.11: Pressure change calculations at different radius r between June 20 and October 21 for cone shaped snow-load.

S5.3 Declustering of eruption catalog

Eruptive events at Mount Ruapehu are extracted from the revised catalogue of Scott (2013), which consists of 586 number of events dating back to the 1830s. Events are assigned a magnitude between 0 and 5, following Table S5.2. We plot these by month in Figure S5.12, excluding events that do not have a listed start month (all of which are pre-1900).

Eruption Scale	Observed Effects
0	Crater Lake steaming, hotter than normal (i.e. above 30-35 °C) creating additional interest, but no observations (or confirmation) of activity in lake.
1	Small phreatic eruptions confined to Crater Lake.
2	Phreatic or phreatomagmatic eruption accompanied by surges; material deposited outside Crater Lake, but still confined to the crater basin. May produce larger flows/floods in Whangaehu Valley.
3	Deposition of material outside the crater basin; possible remobilisation/lahars in upper catchments and Whangaehu valley; OR small scale explosive eruptions/intermittent ash emission when no lake is present.
4	Material deposited well outside the crater basin onto the summit plateau and outer flanks. Lahars possible in several catchments; OR explosive ash eruptions when no lake is present producing columns up to 10000 ft.
5	Large scale explosive eruption displacing moderate volumes of the lake, lahars in all/most major valleys. The summit and slopes covered, with ashfall off the cone; OR explosive eruptions when no lake is present producing tall ($>$ 10000 ft) eruption columns and ashfall off the cone.

Table S5.2: Semi-quantitative estimate of Ruapehu eruptions based on observed effects, sourced from the revised catalogue of Ruapehu eruptive activity (Scott, 2013) (previously modifed from Houghton et al. (1987), Sherburn et al. (1999),Wilson (2009))

.

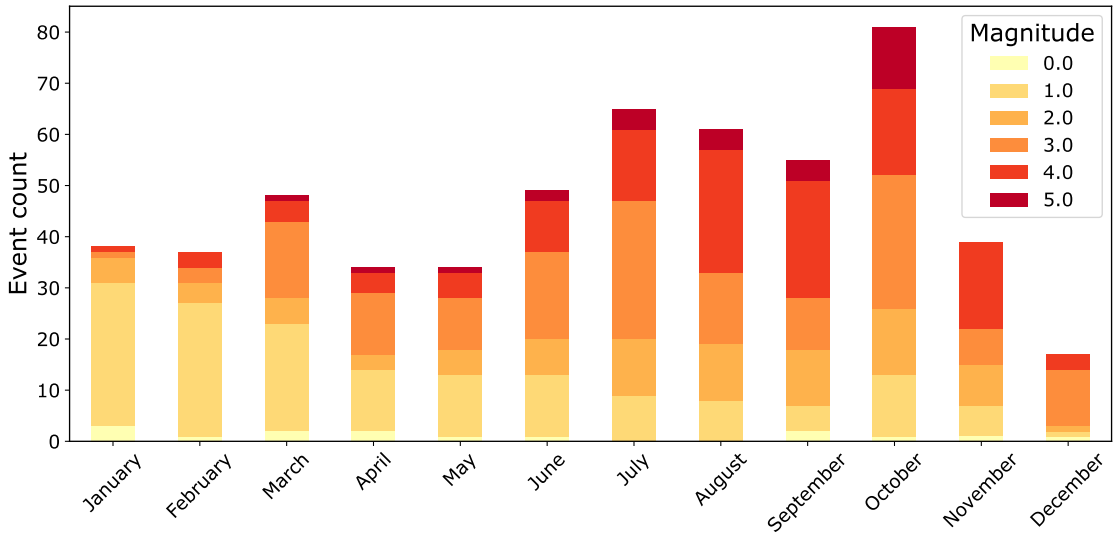

Figure S5.12: All events of magnitude 0-5 at Mt Ruapehu in catalog of Scott (2013), excluding events with no listed start month.

It is important to ensure that the dataset has a good level of completeness before exploring patterns in the eruption history. Scott (2013) consider only data from 1940 onwards towards determining eruption recurrence rates. They also do not consider scale 0 and 1 events, as these events are severely under-represented. Further they observe that the eruptive episodes of 1945 and 1995/96 have a significant impact on the eruptive frequency accounting for 65% of observations. Thus, they also remove all eruptions from 1945 and the 1995/1996 eruption sequence in their analysis. This was similarly done by Christophersen et al. (2022) when using a Bayesian network model to forecast eruptions at the volcano, where they exclude eruptions during the October 1945 to August 1946 and July 1995 to December 1997 magmatic episodes. Unlike Scott (2013), however, they do not exclude data pre-1940.

We similarly clean the catalog, including only magnitude 2+ and excluding eruptive activity associated with the 1945 and 1995/96 eruptive episodes. Following Scott (2013) all eruptions during 1945 are removed, consisting of 133 events starting March 3 and ending December 31. Unlike Christophersen et al. (2022), we do not exclude any further data between January and August 1946, as this period contains only four events (with the first beginning April 17). Similarly, for the 1995/1996 eruptive sequence, we exclude activity using the period defined by Scott (2013) from January 1995 to September 1996, where the pause between the end of this period and the start of eruptive activity in 1997 is over one year. Thus, we prefer to include the 1997 activity in the dataset.

Magnitude 2 events are also said to be under-represented in the catalogue (Scott, 2013). Given the small size of stresses associated with external triggering mechanisms (1-10 kpa) (Caricchi et al., 2021), it is preferable to include these events in catalog however. Examining the cumulative event count for all magnitudes suggests that, while the 1940 cut-off considered by Scott (2013) appears to be a suitable towards ensuring good completeness for magnitude 3 events, a later cut-off may be more suitable for magnitude 2 events. For example, 80% of all magnitude 2 events in the catalog are recorded after 1965. In contrast, the number of magnitude 3 events

increases steadily until 1945 (with 80% of events recorded after this year). Thus we prefer instead a cut-off year of 1960 in order to target a higher completeness of magnitude 2 events. The updated catalog of magnitude 2+ events, excluding pre-1960 events and the 1995/96 eruptive sequences is shown in Figure S5.14.

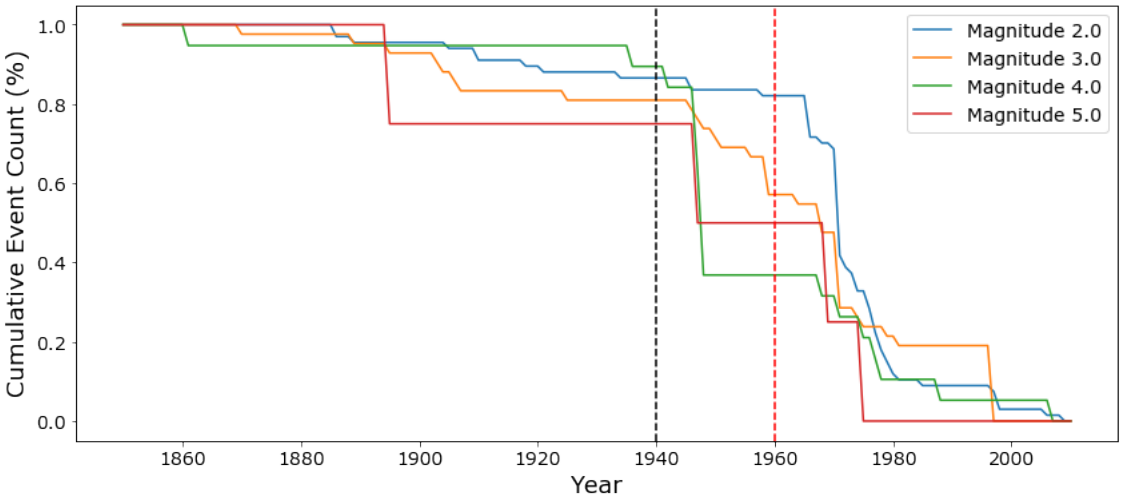

Figure S5.13: Cumulative count of events of different magnitude, after removing 1945 and 1995/96 eruptive activity following Scott (2013). Black-dashed line represented cut-off year considered by Scott (2013) towards analysing eruption recurrence rates. Red-dashed line represents cut-off year preferred here.

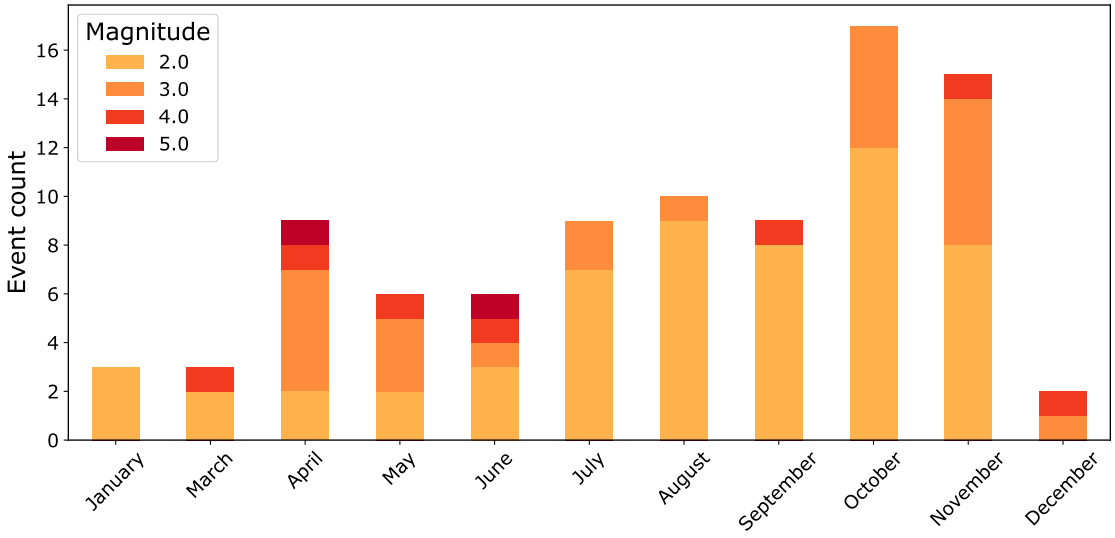

Figure S5.14: All events of magnitude 2-5 at Mt Ruapehu in catalog of Scott (2013), excluding events pre-1960 and removing 1995/96 eruption sequences.

Finally, since we are interested in whether eruptions may be triggered by external processes at the volcano, it is necessary to de-cluster the catalog. Specifically, events that are causatively linked to a previous event should be excluded. Christophersen et al. (2022) use a de-clustering interval of 14 days when examining eruptions of

magnitude 3 or larger post-1830. This was chosen based on examining the change of slope when comparing the number of eruptions at different de-clustering intervals. They then remove all eruptions occurring within 14-days of a prior eruption.

We perform the same analysis with the remaining magnitude 2+ events, after removing both events pre-1960 and the 1995/96 eruption sequence (Figure S5.15). From this, we find that any interval between 10- and 20- days appears a suitable choice, thus opt for the same 14-day interval as used by Christophersen et al. (2022) for magnitude 3+ events. The final catalog after de-clustering, consisting of 49 eruptive events, is shown in Figure S5.16a. For comparison, we also show the de-clustered catalog using a cutoff year of 1940 (Fig. S5.16b).

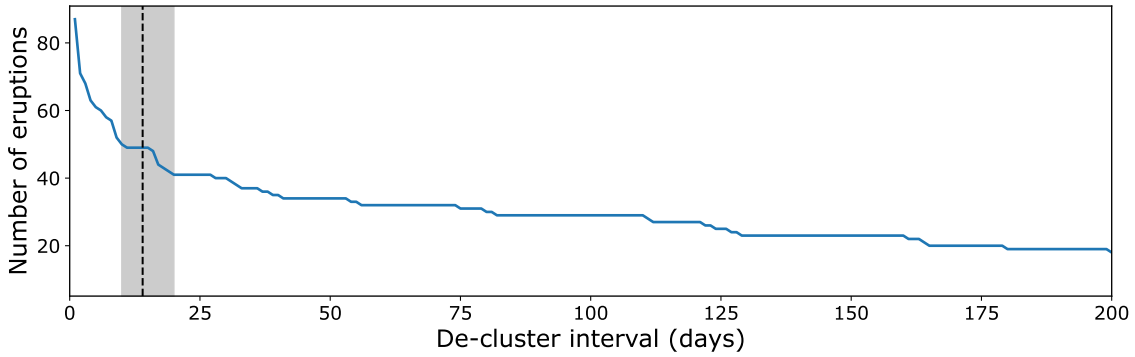

Figure S5.15: Number of eruptions plotted again de-clustering interval for magnitude 2+ events in Scott (2013) eruption catalogue, after removing both events pre-1960 and the 1995/96 eruption sequence. The black-dashed line represents the 14-day interval, and intervals between 10 and 20 shaded grey.

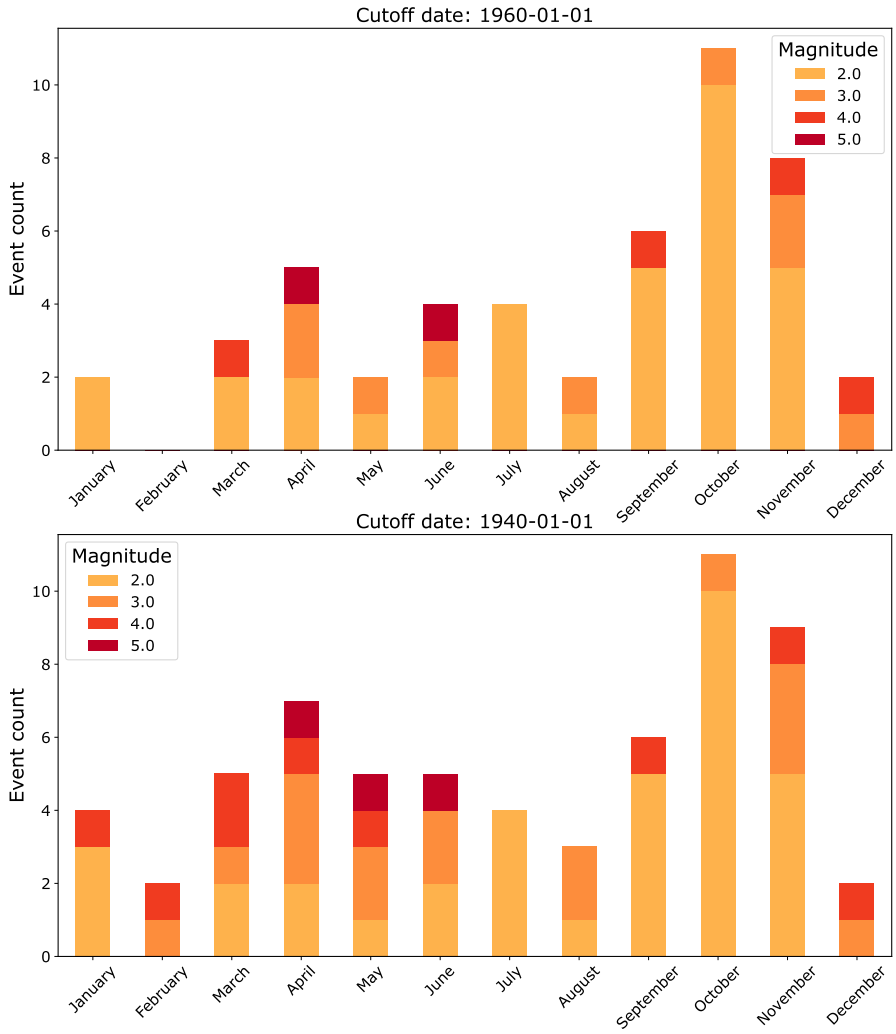

Figure S5.16: De-clustered catalog of magnitude 2+ events at Mt Ruapehu from catalog of Scott (2013), excluding events during 1995/96 eruption sequence, and any events occurring within 14-days of a previous eruption. (Top) Using data post 1940-. (Bottom) Using data post 1960-.

S5.4 Statistical Analysis against non-uniformity

If there was no external triggering of eruptive activity at the volcano, it could be expected that the chance of an eruption on any given day is equal. Thus, we evaluate the statistical significance of the eruptive activity at Mt Ruapehu against a null hypothesis of uniformly distributed eruptions throughout the year. In Figure S5.17, the observed event counts are compared with those expected from a uniform distribution. We show this both for events with magnitudes 2+ (Fig. S5.17a) but also just magnitudes 2-3 i.e. smaller events (Fig. S5.17b). This follows that we might expect that larger events are less likely influenced by external triggering processes.

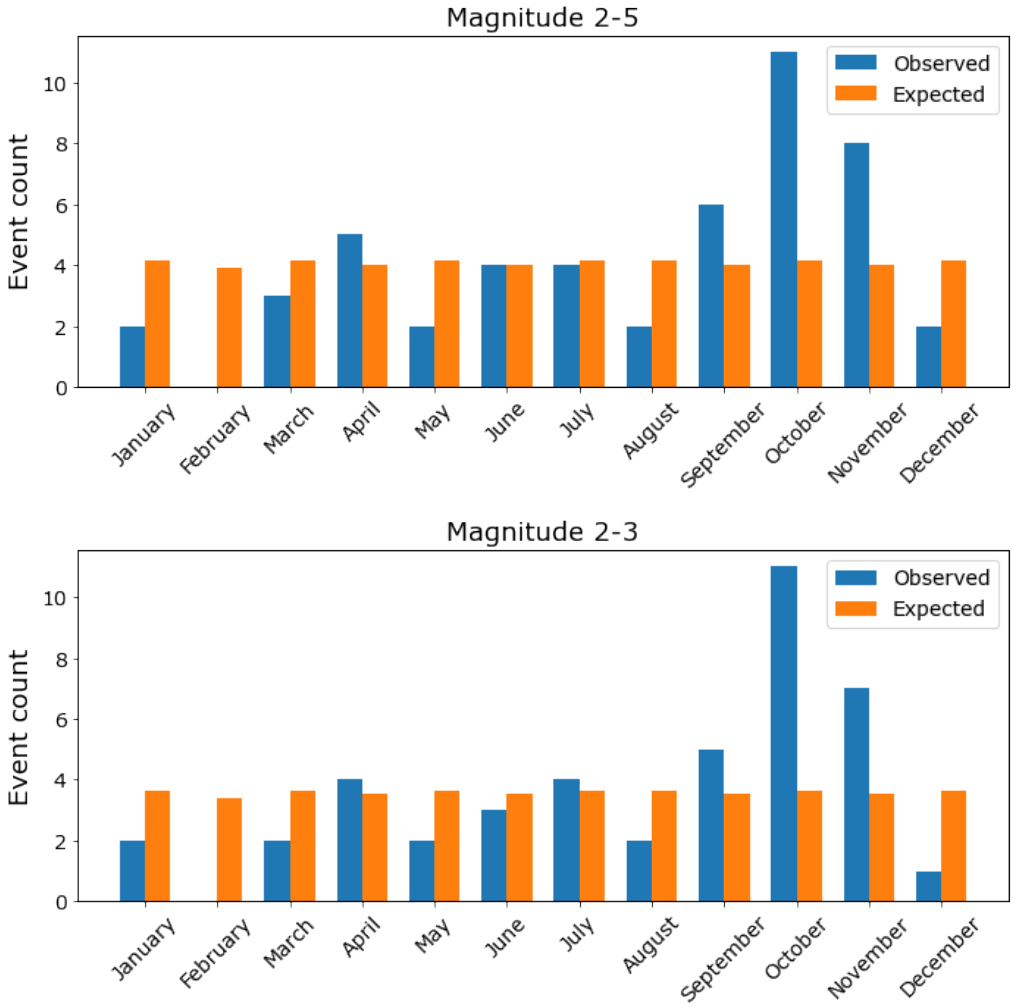

Figure S5.17: De-clustered catalog of events at Mt Ruapehu from catalog of Scott (2013) (see Section S5.3), with only events post-1960, compared with expected events assuming a uniform distribution. (Top) All magnitudes. (Bottom) Only smaller events, magnitudes 2-3.

We apply two statistical tests to assess the significance of variations in event count by month. The first of these is the Fisher's exact test, which calculates the exact probability of obtaining the observed distribution of frequencies under the null hypothesis of independence. This was preferred to the chi-squared test of independence, which is known to be unreliable for small sample sizes (Tate and Hyer, 1973). The second test we apply is the Rayleigh test, which is more suited to circular data. This test assesses whether the data exhibit significant clustering around a certain direction by computing the mean length of the data represented as vectors on a unit circle. Note therefore that the test assumes that data are unimodal.

The Fisher's exact test requires that we split the data into two independent categories. The probability of obtaining the observed distribution of these two variables against a null hypothesis of a uniform distribution is calculated. We separate the data in two ways; (1) comparing the frequency of eruptions in October to those outside of October, and (2) comparing the frequency of eruptions in September–November with those outside of this period. In the case of the former, we focus on

October as this is the month with the largest number of eruptions. In the latter case, we group the time period from September to November, acknowledging that the assumption of independent variables may not be satisfied if a seasonal process is responsible for the higher than expected eruption counts during these three months (i.e. Fig. S5.17). An example of a contingency table used to calculate probabilities for this scenario is shown in Table S5.3.

	Observed	**Expected**
Sept–Nov Eruption	25.0	12.2
Dec–Aug Eruption	24.0	36.9

Table S5.3: Contingency table used for Fisher's exact test, magnitudes 2+, grouping events between September-November. The expected count assumes that eruptions are equally likely on any day throughout the year

We compute p-values following the Fisher's exact test using all events above magnitude 2 and also just smaller events between magnitude 2 and 3, with the final p-values shown in Table S5.4. When using all events, we observed that the distribution is only significant (at 5% level) towards rejecting the null hypothesis when we group the eruptions between September and November ($p=0.01$). Using only smaller events, however, we achieve statistical significance regardless of whether the data are split into October versus non-October eruptions or Sept–Nov versus Dec–Aug eruptions ($p=0.04$ and $p<0.01$ respectively). Thus, the Fisher's exact test suggests there may be a significant difference in eruption rates during certain months. However, while the Fisher's exact test is well-suited to smaller sample sizes, the assumption of independent variables is not easily satisfied with temporal datasets. While the time period between September–November is associated with significant melting of snow (Fig. 5.1), there is also melting of snow continuing into the new year. Thus, we cannot rule out the possibility of a correlation between these months, which is why a circular representation of the dataset may be more appropriate (Figure S5.18).

Test	**Magnitude 2-5**	**Magnitude 2-3**
Fisher's exact test (October)	0.09	0.04
Fisher's exact test (September-November)	0.01	<0.01
Rayleigh test	0.01	<0.01

Table S5.4: Summary of p-value results for different tests and subsets of the data.

Applying the Rayleigh-test, we obtain significant p-values against a uniform distribution when using all magnitude 2+ events and also only magnitude 2-3 events ($p=0.01$ and $p<0.01$ respectively). This reflects the clear cluster of events between September and October. Thus the results of the Rayleigh-test, as with the Fisher's exact test, suggest that there may be a preferential timing of eruptive activity during the year.

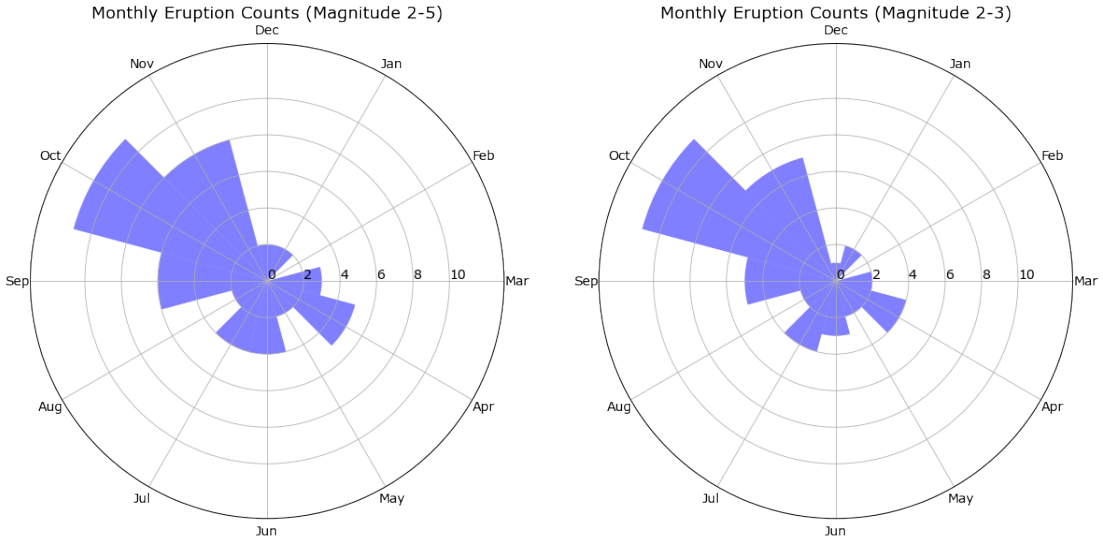

Figure S5.18: De-clustered catalog of events at Mt Ruapehu from catalog of Scott (2013) (see Section S5.3) represented as a circular dataset (Left): All magnitudes. (Right): Only smaller events, magnitudes 2-3.

Finally, we re-do the Fisher's exact tests also considering an earlier cut-off year of 1940 in the catalog, as used by Scott (2013). The observed versus expected number of events for a uniform distribution are shown in Fig S5.19. From this, we obtain statistical significance at the 5% level only when splitting activity into events during September-November versus those in December-August, and only for smaller events (magnitude 2-3) (Table S5.5). As discussed in Section S5.3, there is evidence that magnitude 2 events are under-represented during the time period between 1940–1960 (e.g. Figure S5.13). This is similarly shown when comparing events recorded between 1940–1960 and 1960–1980 (Figure S5.20). Between 1940–1960, only one magnitude 2 event recorded was recorded compared to 19 magnitude 3+ events (also after removing 1945 activity). Over the next 20 years however, between 1960–1980, over 40 magnitude 2 events are recorded (far more than any other magnitude). This is significant given that the majority of events clustered between the months of September and November are of magnitude 2. Thus, if magnitude 2 events are under-represented, it is not surprising that choosing an earlier cut-off date reduces the statistical significance of the event distribution. For this reason, we attach more weight to the results using a cut-off year of 1960, though acknowledge that this choice may have an impact on the results and their interpretation.

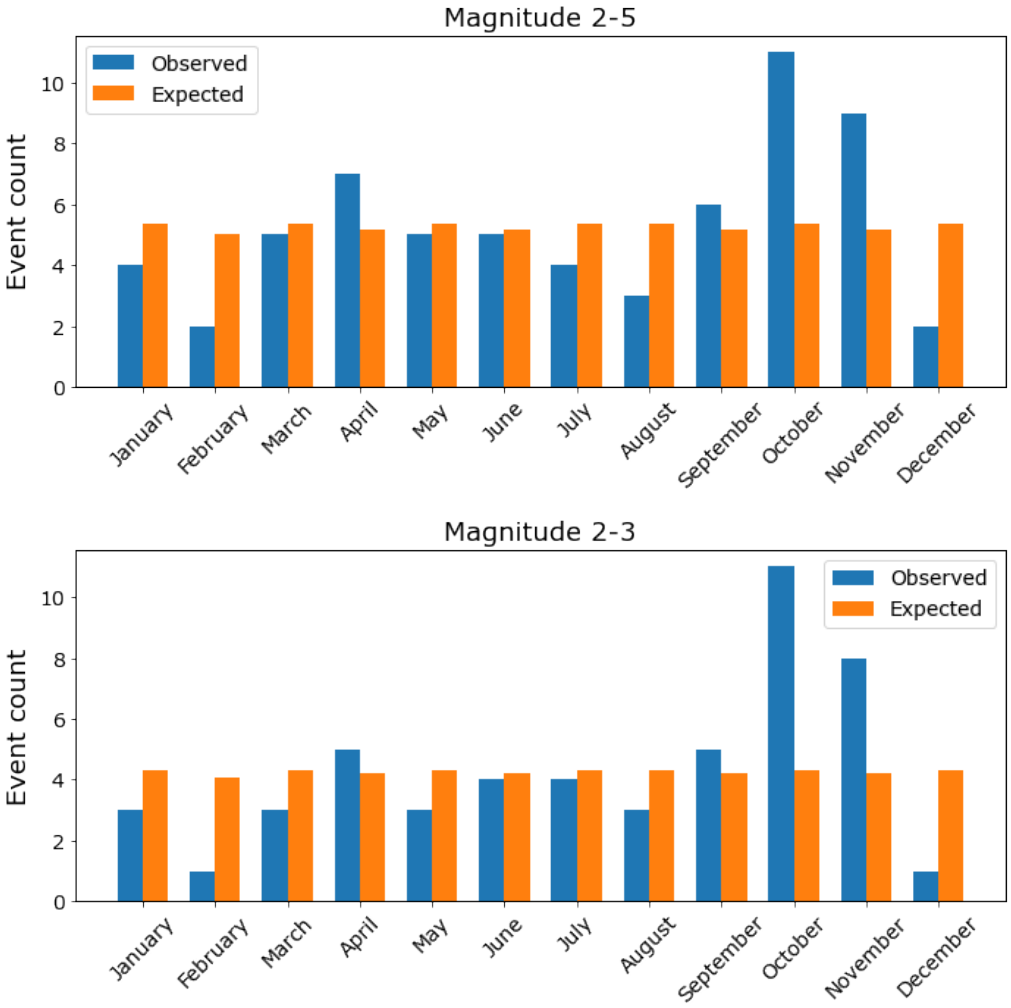

Figure S5.19: De-clustered catalog of events at Mt Ruapehu from catalog of Scott (2013), with only events post-1940, compared with expected events assuming a uniform distribution. (Top): All magnitudes. (Bottom): Only smaller events, magnitudes 2-3.

Test	**Magnitude 2-5**	**Magnitude 2-3**
Fisher's exact test (October)	0.18	0.09
Fisher's exact test (September-November)	0.06	0.02

Table S5.5: Summary of p-value results for different tests and subsets of the data when using a cut-off year of 1940 in the catalogue rather than the preferred 1960 cut-off.

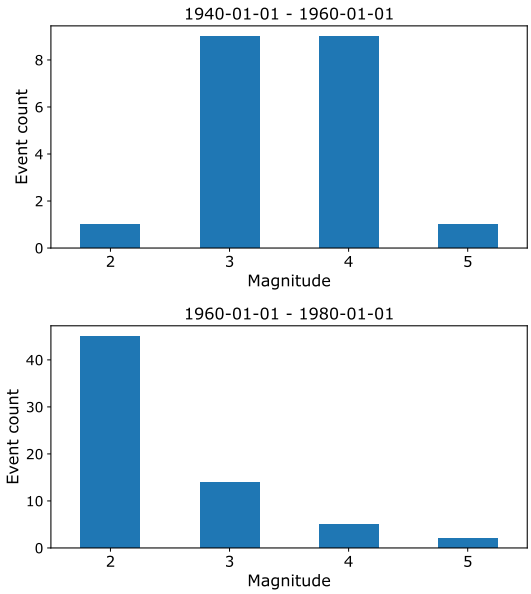

Figure S5.20: Event counts by magnitude between 1940–1960 (top) and 1960–1960 (bottom).

Chapter 6

Discussion and Conclusions

6.1 Introduction

Three objectives were defined at the beginning of the thesis with the purpose of further advancing the monitoring of volcanoes using seismic noise interferometry (with a particular focus on explosive activity). These were:

1. Consider how seismic datasets might be characterized towards gaining a better understanding of the potential and limitations of monitoring through coda wave interferometry at any given volcano.
2. Extract seismic velocity changes from different explosive volcanoes towards improving our understanding of both forecasting potential and other insights that may be obtained by applying coda wave interferometry (for example, the influence of external forcings on volcanoes).
3. Identify key areas for future research and development, while also considering practical steps that could enhance our ability to monitor and forecast explosive activity.

Chapters Two and Three of the thesis sought to address the first objective. In Chapter Two, various steps were considered that could be used to understand a given dataset. This included the application of network-based covariance matrix analysis towards understanding the coherence of the seismic wavefield. We then computed measures of SNR, phase similarity, and amplitude ratios to understand the effect of various seismic features on the data (cross-correlation functions). This was complemented by the application of hierarchical clustering to cross-correlation functions in Chapter 3. This allowed for better recognition of long-term similarity in cross-correlation functions and, more generally, the identification of similar cross-correlation functions separated in time. Discussion about the use of the various methods introduced can be found in the discussion section of each chapter.

A primary motivation of these chapters was to think about how one might apply coda wave interferometry at a volcano for the first time. One goal was to discuss

possible steps that could be taken to quickly understand the character of a new dataset and the applicability of coda wave inteferometry. The use of five volcanoes as examples was then intended to identify features that may be common at any given volcano. It should be stressed though that improvements can be made to the methods presented. Ultimately, it remains a challenge to determine what is a minimum level of similarity required between individual cross-correlation functions towards computing accurate seismic velocity changes. Similarly, it is unlikely that any one volcano is going to be the same as another. However, the development of techniques that can better illuminate the character of a dataset has two advantages. First, it allows for more informed decision making, both with regards to processing but also when interpreting results. For example, identifying whether a significant change in velocity occurs during a time period where there is evidence that cross-correlation functions may be contaminated is informative. The second advantage is that better characterization of the dataset may suggest alternative ways to process the data in addition to, or in place of, common practices. For example, at Stromboli, we computed seismic velocity changes using highly similar cross-correlation functions produced during three separate periods of heightened activity (Fig. 4.19). Identifying such similarity would have been difficult using conventional similarity techniques (based on comparison with a reference function). Thus, while there are clearly still challenges associated with the final determination of reliability, we argue that better characterisation of datasets provides the means to make more informed choices in both the processing and interpretation of velocity changes.

Chapters Three and Four focus on individual case studies at Stromboli volcano, Italy, and Mount Ruapehu, New Zealand respectively. They are therefore focused on objective two. At Stromboli, the main focus was the detection of velocity changes that could mark the transition from background strombolian activity towards more violent explosions such as the two paroxysms that occured in 2019. At Ruapehu, we compute seismic velocity changes during a four year period that contained two gas-driven eruptions. While seismic velocities do not identify any clear pre-cursors prior to these events, they do allow us to study seasonal processes that may be related to increased spring-time activity in the eruption chronology. Thus, both chapters utilise coda wave interferometry in slightly different ways. In chapter three, the use is closer to a monitoring/forecasting approach, where we search for velocity changes that indicate evolving behaviour within the volcanic system. At Ruapehu, the computation of velocity changes can be used to help understand the influence of seasonal processes as possible trigger for spring-time activity. As with chapters two and three, individual discussion of each case study can be found within the chapters.

The thesis is concluded by discussing two themes that are common across the different chapters and, therefore, deserve further discussion. These are (1) the computation of velocity changes within the frequency band of volcanic tremor and (2) external sources of seismic velocity changes at volcanoes. While the focus is monitoring explosive activity, some of the discussion will apply to monitoring volcanoes more broadly. Finally, the thesis will conclude by suggesting future practices or research directions that could enhance our ability to monitor explosive activity in the future using seismic noise interferometry, thus addressing the third objective.

In advance of such discussion, we briefly introduce two further datasets that can

also contribute to the discussion. These are (1) a 3–4 month seismic crises at Mt Ruapehu in 2022 and (2) four years of data at Tungurahua (Ecuador) between 2014–2018, containing both effusive and explosive activity.

6.1.1 Ruapehu: Real-time monitoring in 2022

The modern network around Ruapehu and neighbouring volcanoes (Ngauruhoe and Tongariro) contains a number of broadband and short period stations. The majority of these are either in close proximity to Ruapehu or around the Tongariro complex (with Mt Ngauruhoe located on the south end) (Fig. 6.1).

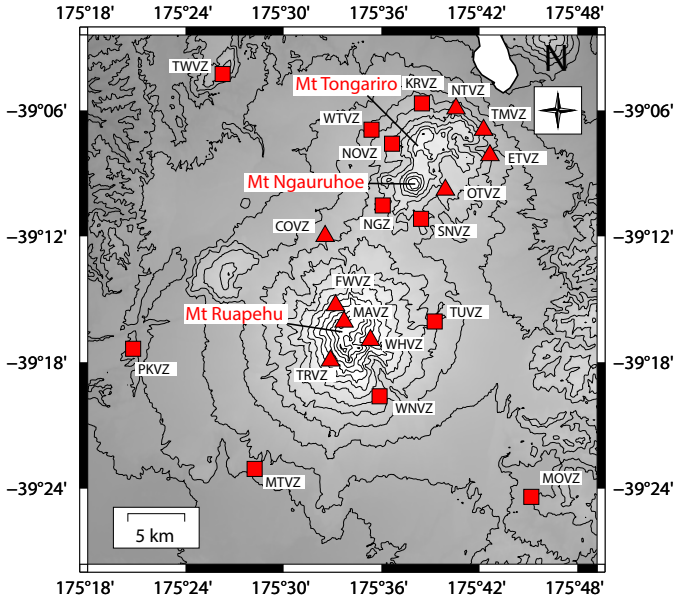

Figure 6.1: Seismic network around Mt Ruapehu and the Tongariro Complex as of March 2022. Red triangles represent broadband seismometers and squares short period seismometers.

Starting in March 2022, an episode of strong volcanic tremor began at Mt Ruapehu, coinciding with a heating phase of the crater lake (Geonet, 2022c). This was accompanied by high levels of volcanic gas emissions, resulting in the volcanic Alert level being raised to 2. Volcanic tremor began to reduce in strength from mid-May, along with crater lake temperatures (peaking at 41 degrees on 8 May) (Geonet, 2022a). Gas emissions finally began to decline late-June, with the volcanic alert level returning to 1 on July 4 (Geonet, 2022b). The entire episode represented the longest period of tremor recorded over the previous 20 years.

From spectral width measurements, the tremor is shown to be highly broadband. This is evidenced by reduced values of spectral width observed between approximately 1–10 Hz, beginning March 13 (Fig. 6.2). Within the period of strongest tremor (until mid-May), we also observe harmonic tremor, with a fluctuating dominant frequency (appearing to increase over-time from early-April to mid-May). Ex-

amining the seismic drums at stations on the volcano reveals transitions between pulse-like activity and more continuous volcanic tremor during the seismic crises at certain frequencies (e.g. Fig. 6.3).

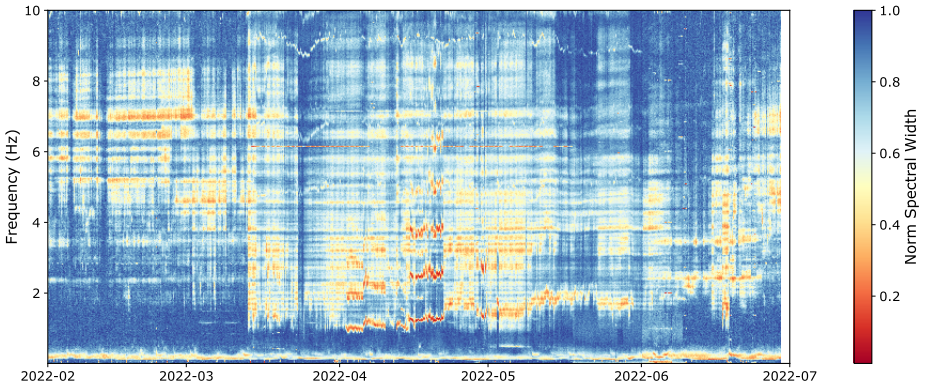

Figure 6.2: Covariance matrix spectral width measurements (normalized) computed using four broadband stations (MAVZ, FWVZ, WHVZ, TRVZ) on the Ruapehu edifice, Z-component. Processing follows that of Chapter Two.

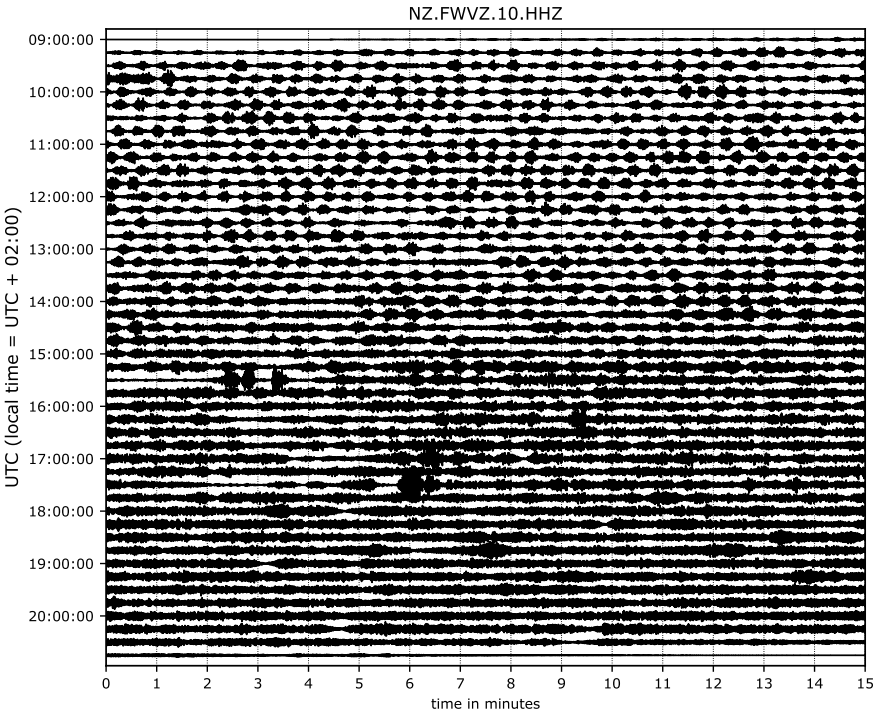

Figure 6.3: Seismic drum recording at FWVZ station on 5 May, 2022, showing transition from pulse-like seismic activity to more continuous tremor. Seismogram filtered between 0.5–1.5 Hz.

Early into the seismic crises, plots of SNR and phase similarity suggested the onset of broadband tremor was producing highly similar cross-correlation functions (Fig. 6.4). Between approximately 1–6 Hz, values of SNR were between 20–60, completely dominating the color scale relative to other time periods (Fig. 6.4b). Similarly, values of phase similarity were very close to 1.0 (Fig. 6.4c). This high similarity

is confirmed when examining 1–day cross-correlation functions in the time-domain (Fig. 6.5). Above 1.5 Hz, the amplitude ratio between positive and negative lag times remained relatively similar to pre-tremor values (dominantly negative, especially between 1.5–3.0 Hz, Fig. 6.4d). However, a sharp change in values between approximately 0.7–1.5 Hz (from positive values to ratios closer to zero) suggests some change in dominant source process.

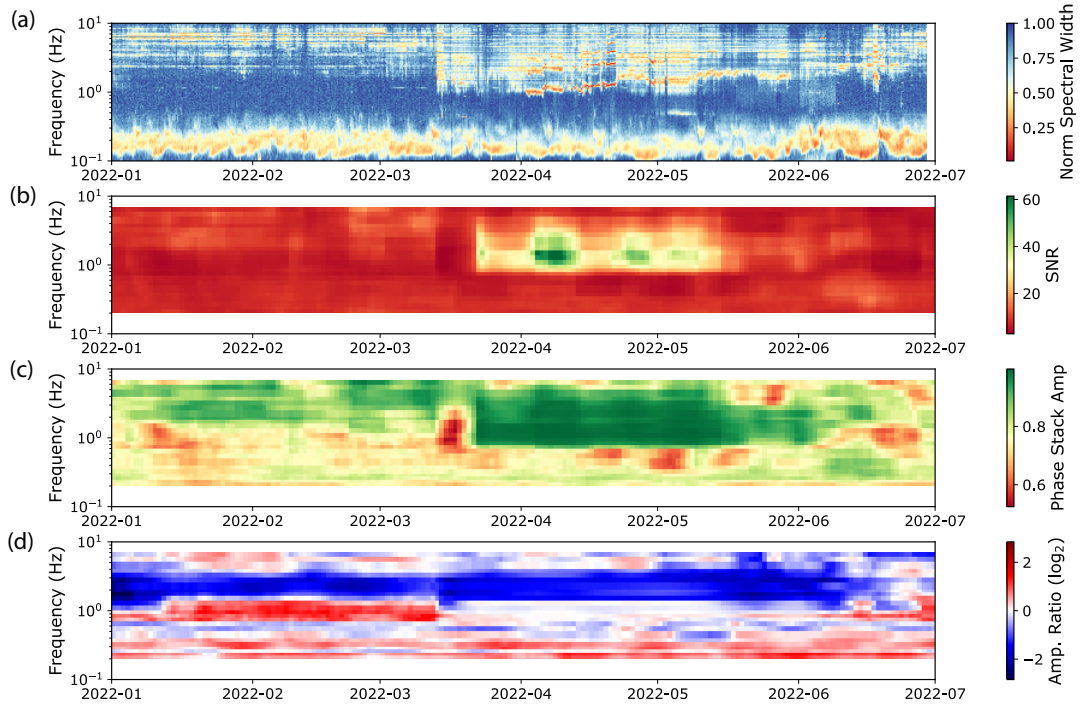

Figure 6.4: Measurements of cross-correlation characteristics at Ruapehu for a single station-pair (FWVZ-MAVZ), computed using 10-day moving window. For SNR and phase stack amplitude, the average value is taken between minimum lag defined by 0.5 km/s ballistic arrival speed and maximum lag equal to the minimum lag plus 15 cycles of central period of the bandpass filter. Processing of cross-correlation functions consistent with that of Chapter Two. (a) Normalized covariance matrix spectral width as computed over full network. (b) SNR. (c) Phase stack amplitude. (d) Amplitude ratio.

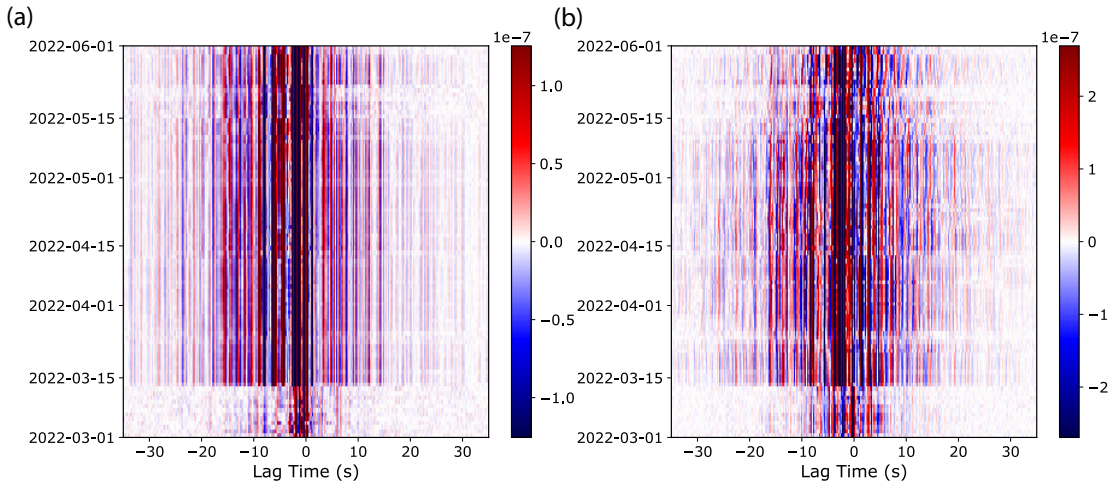

Figure 6.5: Correlogram of 1-day cross correlation functions computed for station-pair FWVZ-MAVZ at (a) 1–2 Hz. (b) 2–4 Hz.

The broadband, high SNR, tremor provided a unique opportunity to study very short-term velocity changes in real-time. Thus, from late-March, we begin to monitor the activity, computing velocity changes with a temporal resolution of 6 hours (Fig. 6.6b). Even at this time resolution, we observe a correlation coefficient of approximately 0.8 (or above) at 2–4 Hz for most of period where tremor was strongest (up until early-May, Fig. 6.6c). Velocity changes were compared with measurements of lake level, temperature, and rainfall. From this, the most significant velocity changes appear to coincide with rainfall events. This is evidenced also by temporary increases in the lake level (Fig. 6.6d) and reductions in temperature (Fig. 6.6e). Regarding the long-term increase, it is thought that this relates to an external environmental process (since similar long-term trends were observed at stations off the volcano, albeit with lower amplitudes). Alternatively, we also note an increasing trend in the dominant tremor frequency from early-April through to the start of June. An evolving seismic source process may therefore contribute to the long-term trend observed. This would need to be explored further, with preliminary checks not revealing a clear change in frequency of the coda in response to this evolving source (not shown).

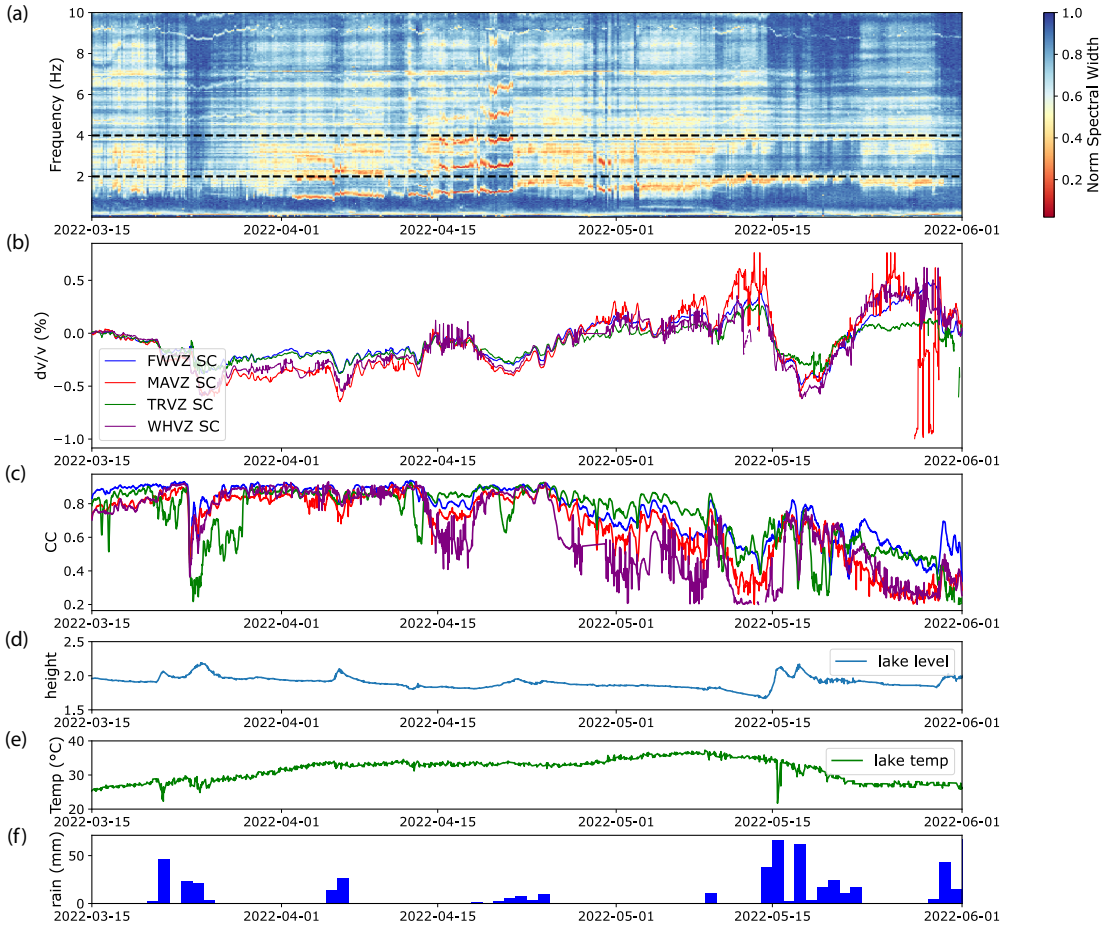

Figure 6.6: Single-station cross-component (SC) subdaily velocity changes (6-hour cross-correlation functions) computed between 2–4 Hz. Cross-correlation functions processing consistent with previous chapters, with final velocity changes measured using stretching technique between ±5–15 s lag-time. Final velocity is average of three cross-components (EN, EZ, NZ). (a) Normalized spectral width. (b) Velocity changes at four stations on Ruapehu. (c) Correlation coefficient with reference period (2022-03-15 to 2022-05-01. (d) Lake height (m). (e) Lake temperature. (f) Rainfall recorded by Chateau weather station.

6.1.2 Tungurahua 2014–2018

Tungurahua is an andesitic stratovolcano located in Ecuador, within the Northern Andean volcanic zone (Kley et al., 1999; Hall et al., 1999). It is one of the most active volcanoes in Ecuador (Hall et al., 1999), with its most recent eruptive sequence lasted from 1999 to 2016, consisting of multiple phases with various different styles of eruptive activity including more continuous strombolian activity, vulcanian explosions, and plinian activity (Bustillos A. et al., 2018, and references therein). In this section, we compute seismic velocity changes from 2014–2018, which captures the final two years of this erupion sequence (with the final activity in March 2016, (Vlastélic et al., 2023)).

The network around Tungurahua consists of 13 broadband seismic stations (Fig.

6.7). Computing the normalized spectral width using the vertical component of all stations — following the processing in Chapter Two — reveals different features of the seismic wavefield at the volcano (Fig. 6.8). At lower frequencies (< 0.5 Hz), energy associated with the oceanic microseisms is clearly visible, revealed by low spectral width values. Between approximately 1–3 Hz, we observe consistently lower spectral width values throughout the four year period. Post-2016, following the end of the eruption sequence, the volcano was thought to be absent tremor (J. Battaglia, personal communication). Thus, the true origin of this coherent energy is unclear and requires further study. Finally, during 2014–2016, we observe smaller periods of time (days to weeks) where lower spectral width values are recorded over a broader range of frequencies (from approximately 0.4 – 10 Hz), for example during Aug-Sept 2014. The timing of these periods appears to closely match times of high explosive activity at the volcano (Fig. 6.8b).

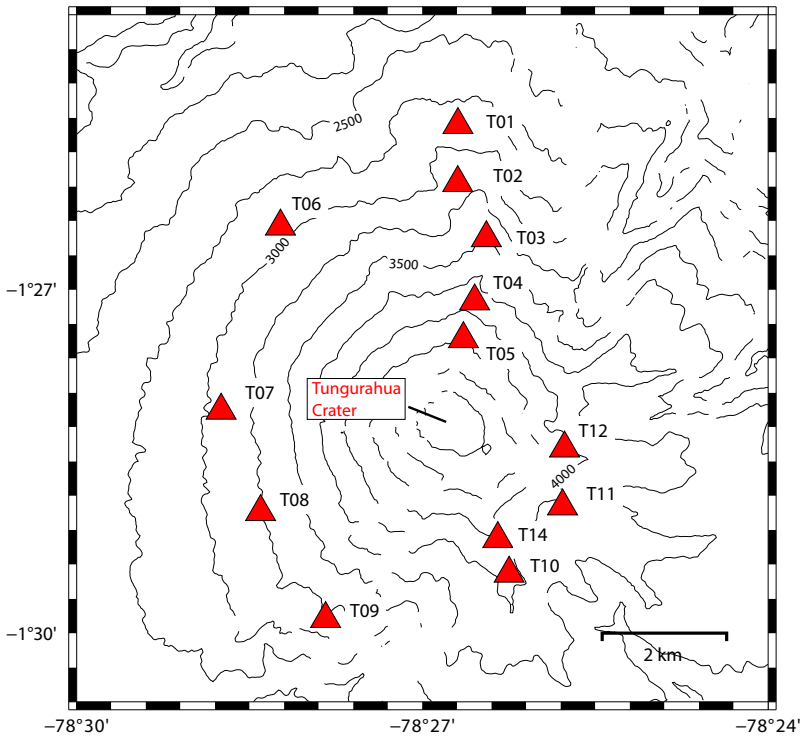

Figure 6.7: Map of Tungurahua volcano, with 250 m contours. Broadband seismic stations shown as red triangles.

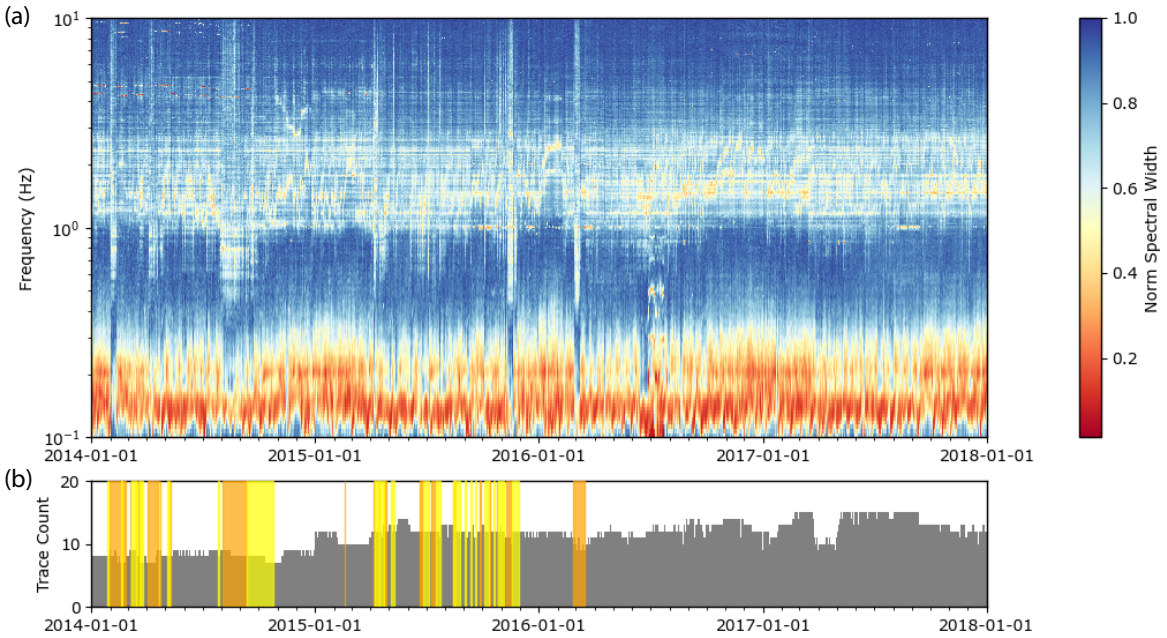

Figure 6.8: Covariance matrix spectral width computed using vertical component stations at Tungurahua. (a) Normalized spectral width. (b) Trace/Station count, with eruptive activity also shown (yellow shading = low explosive activity. Orange shading = high explosive activity).

Computing the SNR and phase similarity for a pair of stations in 2015 — following processing in Chapter Two — shows the highest values above 1 Hz (Fig. 6.9). For example, the phase similarity is mostly above 0.8 between approximately 1–4 Hz (Fig. 6.9c), likely reflecting the coherent energy revealed by reduced spectral width (Fig. 6.9d). Temporary decreases in the SNR and phase similarity do occur, seemingly corresponding to time periods where the spectral width is reduced over a broader frequency during more continuous eruption activity (for example, in April–May 2015 and late-2015). We similarly observe some fluctuation in the amplitude ratio at approximately 1 Hz during such episodes, suggesting modification of the cross-correlation functions during these periods of eruptive activity.

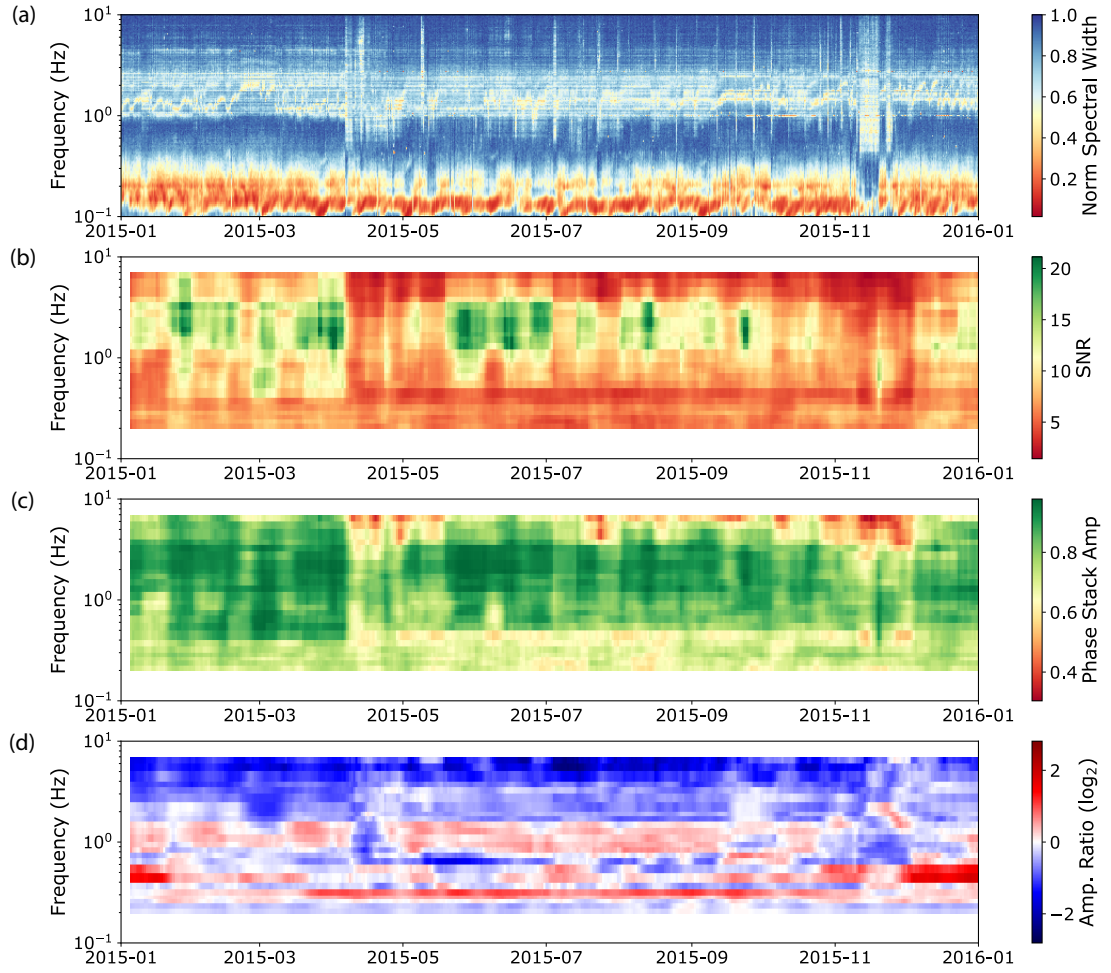

Figure 6.9: Measurements of cross-correlation characteristics at Tungurahua for a single station-pair (T04–T12), computed using 10-day moving window. For SNR and phase stack amplitude, the average value is taken between minimum lag defined by 0.5 km/s ballistic arrival speed and maximum lag equal to the minimum lag plus 15 cycles of central period of the bandpass filter. Processing of cross-correlation functions consistent with that of Chapter Two. (a) Normalized covariance matrix spectral width as computed over full network. (b) SNR. (c) Phase stack amplitude. (d) Amplitude ratio.

Measurements of phase similarity (i.e. phase stack amplitude) at different lag times and frequencies shows that the cross-correlation functions are highly similar out to late lag times between approximately 0.5–4 Hz (Fig. 6.10a). This is also true of time periods where the spectral width is reduced over a broader frequency range, albeit with slightly lower values of phase similarity (but still above 0.6) (Fig. 6.10b). Thus, Tungurahua appears to be another volcano that presents highly similar cross-correlation functions above 1 Hz in response to coherent energy observed in spectral width outputs.

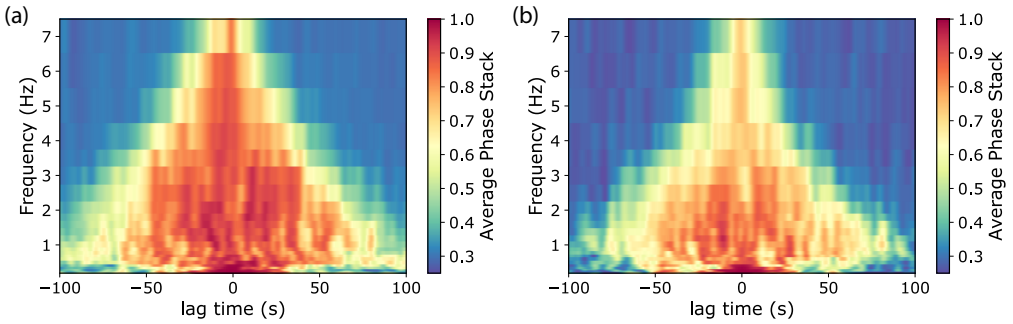

Figure 6.10: Phase similarity as a function of frequency and lag-time for station-pair T04-T12. Values computed as average between defined date ranges. (a) average between 2015-01 to 2015-04 (background). (b) 2015-04-21 to 2015-05-11 (eruptive activity).

Velocity changes are computed using 1–day cross-correlation functions within three frequency bands (0.1–1.0 Hz, 1.0–2.0 Hz, 2.0–4.0 Hz). For this, we use the stretching technique with coda windows between 10–50 s, 10–40 s, and 10-30 s respectively for the three frequency bands (lowest to highest). The full 4-year dataset is used as the reference. Finally, velocities with a correlation coefficient below 0.65 are removed from the dataset (mostly corresponding to time periods of high explosive activity).

Two features of the velocity changes stand out in all frequency bands (Fig. 6.11). The first is an approximate annual cycle, with velocity high in the early months of each year and low velocities from mid-year. The effect is strongest at higher frequencies and most likely reflects a seasonal trend. The second major change is a velocity decrease associated with a M_W 7.8 earthquake (16 April, 2016) approximately 300 km away (Ye et al., 2016). Velocities then slowly recover, though remain slightly lower than pre-earthquake values by the start of 2018. It is of interest that the earthquake approximately marks the end of the eruption sequence lasting 17 years, with no further eruptive activity recorded to date. However, there is evidence that the rate of deep magma supply was slowing during this time (Vlastélic et al., 2023). Thus, it is left as an interesting observation without further comment at this time.

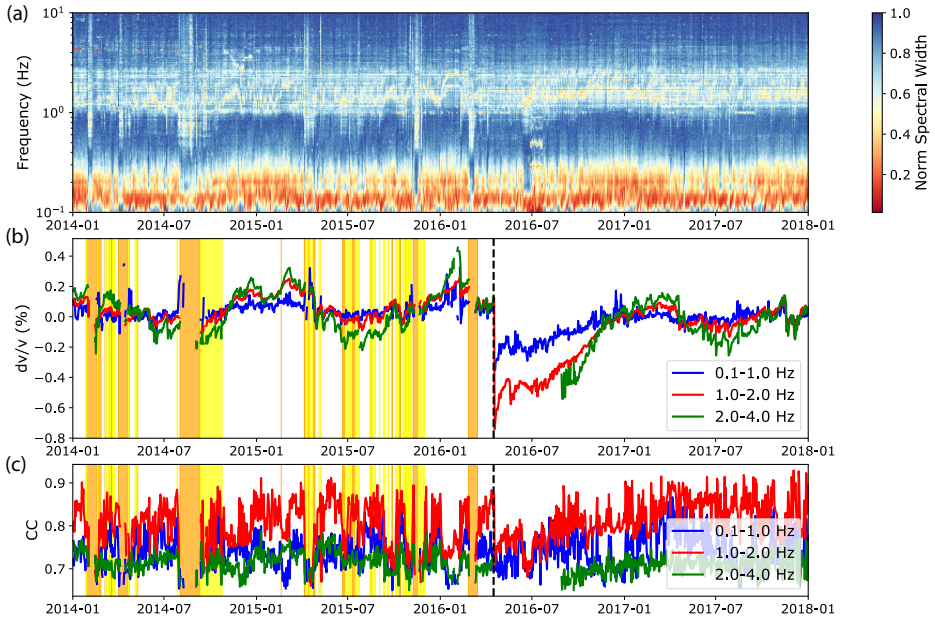

Figure 6.11: Velocity changes computed at different frequencies. Yellow shading = low explosive activity. Orange shading = high explosive activity. Vertical-dashed black line = M_W 7.8 earthquake. (a) Normalized spectral width measurements. (b) Velocity changes at different frequncies. (c) Correlation coefficient between 1-day and reference cross-correlation functions.

Aside from velocity changes influenced by external processes, we observe some evidence that fluctuation in the dominant frequency between 1–3 Hz (reflected by slightly lower spectral width) may influence velocity changes (Fig. 6.12). For example, changes in velocity ($>$ 1 Hz) between mid-2014 to mid-2015 and early-2016 show some similarity with the lowest spectral width values between 1–4 Hz. This is discussed further in Section 6.2.

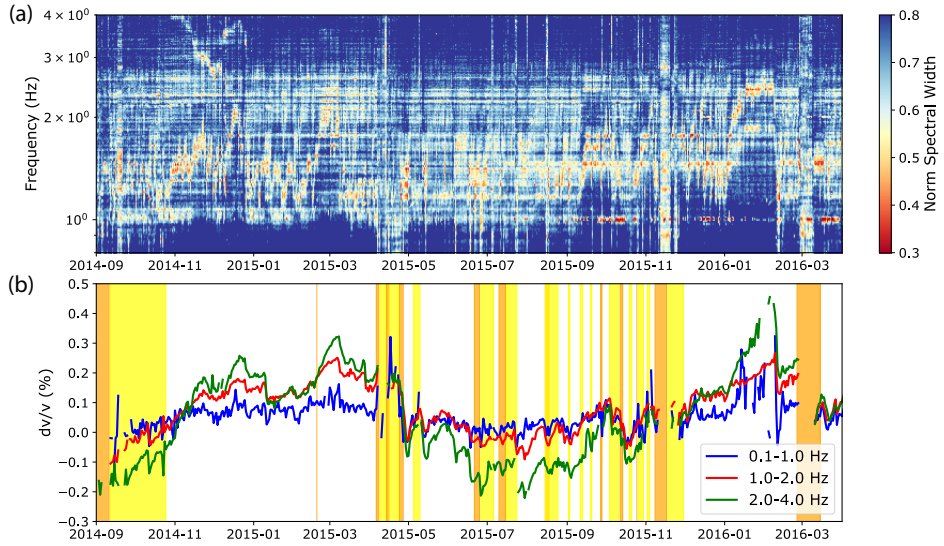

Figure 6.12: Zoomed in velocity changes from Sept 2014 to April 2016, with modified color scale for spectral width measurements to better observe changing frequency of seismic wavefield. (a) Normalized spectral width measurements. (b) Seismic velocity changes above 1 Hz.

Finally, clustering cross-correlation functions at Tungurahua (following processing of chapter three) suggests a possible usage of the method towards identifying different styles of volcanic activity. For example, after clustering 1–day cross-correlation functions computed for station-pair T03-T04, we observe a smaller repeating cluster that appears to coincide with high explosive activity (Fig. 6.13). This pair was chosen due to both stations having data available for most of the four year time-period. The observation of a repeating cluster related to volcanic activity might suggest that these time periods were associated with similar styles of activity. No further occurences of the smaller cluster are observed post-eruptive activity.

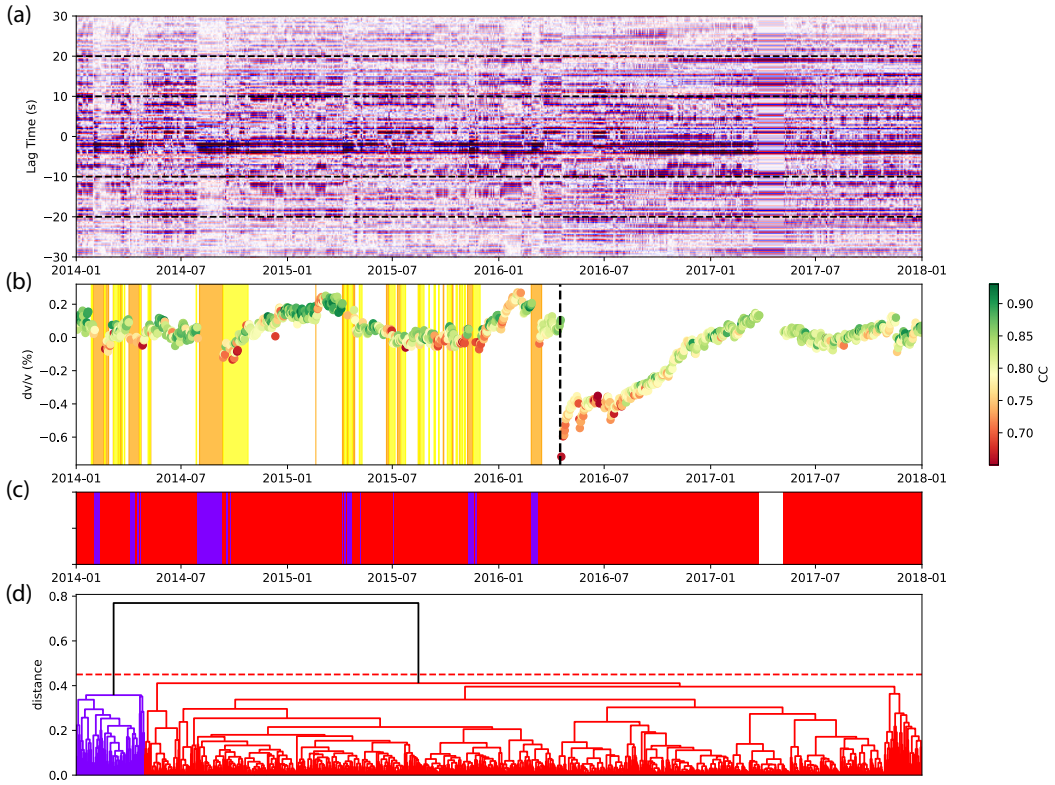

Figure 6.13: Cluster output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pair T03-T04 at Tungurahua (a): Correlogram showing amplitudes of CCFs at different lag times (red = positive, blue = negative). Dashed black lines show part of CCFs used in clustering. (b): Apparent velocity changes for station-pair T03-T04, color-coded according to correlation coefficient (CC) computed between 1-day stacks and the reference stack (full time period). (c): Location of clusters in time, color-coded according to dendrogram output. (d) Dendrogram, with clusters defined at a distance threshold of 0.45 (dashed-red line).

6.2 Coda wave inteferometry within tremor frequency band

The use of tremor towards computing velocity changes was first applied by Donaldson et al. (2017) at Kilauea volcano. This has then been repeated at the same volcano in multiple studies succeeding this (e.g. Olivier et al., 2019; Flinders et al., 2020; Muzellec et al., 2023). The high stability of the tremor source identified in these studies is also reflected in the measurements of SNR and phase similarity computed in Chapter two prior to the 2018 eruption (Fig. 2.15).

Velocity changes were computed within the frequency band of volcanic tremor in this thesis at Stromboli (Chapter three), Ruapehu (Chapter four), and Tungurahua (Section 6.1.2). This was motivated by observations of high similarity cross-correlation functions, similar to those observed at Kilauea. However, at least in the case of Ruapehu and Stromboli, it is clear there are changes in the character of cross-

correlation functions with time within the tremor frequency band. At Ruapehu, tremor is not continuously present, with lengthy periods where tremor is very low amplitude or absent. These periods are marked by significant decreases in the SNR and phase similarity of cross-correlation functions. To better handle this, we apply the wavelet transform approach to measure delay times when the tremor is coherent in the time-frequency space. At Stromboli, the tremor is more continuous but appears to undergo significant changes in response to major volcanic events and also during periods of heightened activity. In this case, we sought to verify through clustering and examination of the spectral content of cross-correlation functions that such changes were not a likely factor towards explaining the primary observations of interest. However, in both cases (at Ruapehu and Stromboli), we largely restrict our analysis to the evaluation of long-term trends in order to avoid over-interpreting short-term changes that may be due to evolving tremor characteristics.

Previous studies have suggested that changes in seismic source frequency can have an influence on accuracy of velocity changes recovered (e.g. Zhan et al., 2013) while others suggest this impact may be minimal within the coda (e.g. Mao et al., 2020). In the case of the Zhan et al. (2013), synthetic tests were based on a single wavelet, more consistent with ballistic arrivals. For Mao et al. (2020) however, a large number of individual wavelets with random positions and amplitudes were used to approximate the coda. From this, the coda was found to be less sensitive to frequency changes, with the coda spectra determined by the scattering properties of the medium. However, these tests were based on stretching the amplitude spectrum of the source by 20 percent. While this does result in a shift of the dominant source frequency, the overall shape of the spectra remains similar.

At volcanoes, changes in source spectra of tremor may be more complex. At Stromboli, some evidence is observed to suggest that short-term velocity changes are influenced by tremor. For example, sharp velocity changes are recorded during times where clustering suggests the cross-correlation functions have changed character in response to heightened activity (Fig. 4.18). At Tungurahua, we observed some evidence that velocity changes may follow changes in the dominant frequency of tremor (Fig. 6.12). This can be examined further by comparing the dominant frequency of the coda with velocity changes at the different frequencies (Fig. 6.14). At 1–2 Hz, there appears to be a clear relationship between the dominant frequency of the cross-correlation functions and measured velocities (Fig. 6.14b). This might suggest that the velocity changes are not real medium changes, but are instead a reflection of a changing seismic source. However, at other frequencies (0.1–1.0 Hz and 2.0 – 4.0 Hz), the relationship between dominant frequency and velocity changes is less clear, despite velocities showing a similar trend to those at 1–2 Hz. This might then suggest that the trends observed reflect real velocity changes in the medium. The close fit between velocities at 1–2 Hz and the dominant frequency of tremor may then be indicative an underlying process, such as structural or stress-related changes within the plumbing system (Carniel and Tárraga, 2006; Tárraga et al., 2014), that can induce changes in both parameters. Determining this is difficult though without further understanding of both the source of changes in tremor characteristics and the influence of varying frequency on the coda.

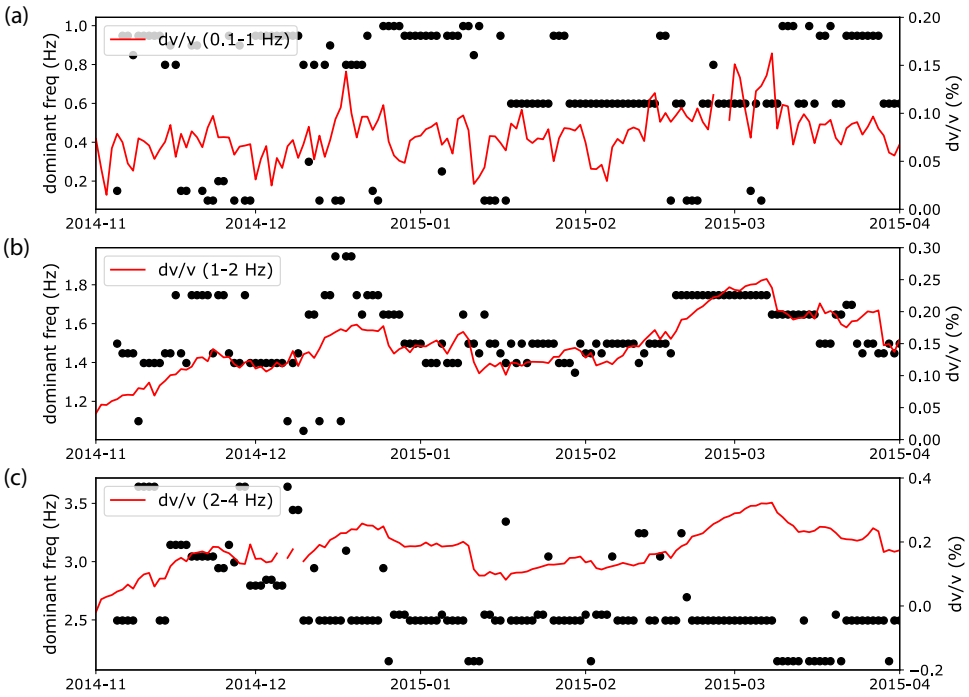

Figure 6.14: Comparison of dominant frequency of cross-correlation functions (black scatter) and measured seismic velocity changes (red line) at different frequencies. (a) 0.1–1.0 Hz. (b) 1.0–2.0 Hz. (c) 2.0–4.0 Hz.

It therefore remains challenging to judge how similar the cross-correlations must be in time towards measuring reliable velocity changes. Common measures of similarity based on comparison with a reference stack (for example the coherence or corrrelation coefficient) can be useful towards determining whether significant changes in the seismic source have occurred. For example, the onset of tremor during effusive activity at both Piton de la Fournaise (e.g. Fig 3.6) and Tungurahua (Fig. 6.11) are associated with significant drops in the correlation coefficient. However, at Stromboli, while there are decreases in the correlation coefficient associated with the periods of heightened activity (Fig. 4.10), these do not neccessarily stand out relative to background fluctuations over the nine-year period. This is similarly the case at Tungurahua where some fluctuation in the correlation coefficient is observed but — aside from during effusive activity — values largely remain above 0.8 (at 1–2 Hz). Given the subtle nature of the velocity changes being measured, it is conceivable that significant variations in the correlation coefficient may not always be present, even in instances where modification of cross-correlation functions due to evolving tremor has occurred. Such subtlety in the velocity changes of interest underpins the challenges in dealing with the possibility of varying seismic sources. Thus, when working within the frequency band of tremor, a higher level of scrutiny should be applied before interpreting velocity changes as volcanic in origin.

With these considerations, ultimately the choice of whether to process velocity changes using the frequency band of tremor may depend on the specific monitoring needs. One consideration would be target depth of interest. While not always the case (e.g. Kilauea), at many volcanoes examined the tremor is above approximately

1 Hz. At both Stromboli and Ruapehu, 1 Hz Rayleigh waves are most sensitive to changes between approximately 200–600 m. Thus, using frequencies above 1 Hz would increase the sensitivity to velocity changes in the upmost few hundred meters, while using frequencies below 1 Hz provides sensitivity to deeper portions of the magmatic system.

At many volcanoes, it may not be neccessary to have sensitivity to the upmost 200 m. At Stromboli, while velocity changes above 1 Hz were benefictial towards identifying a deviation of the station closest to the summit from other stations in the years prior to the 2019 paroxysms, a similar (opposite) long-term trend was recorded using lower frequencies (0.1–1.0 Hz). At Ruapehu, however, the only changes we observe that may be related to phreatic eruptions in 2006 and 2007 are a decreasing velocity between approximately 1.0–1.5 Hz from mid-September. These changes are thought to reflect a seasonal process, but may still be linked to greater eruptive activity at the volcano during spring months (Section 5.4.3). One key difference between these case studies is the role of the hydrothermal system in driving eruptive activity. At Stromboli, while a hydrothermal seal is evidenced (Finizola et al., 2002; Revil et al., 2011) there is minimal suggestion of pressure buildup beneath it. In contrast, gradual sealing of eruptive vents within the shallow hydrothermal system at Ruapehu is a known mechanism for triggering phreatic activity (Christenson et al., 2010; Ardid et al., 2023). Monitoring velocity changes within shallow hydrothermal systems may therefore be one example where greater depth sensitivity to the upmost few hundred meters would be of value.

Examining other examples of volcanoes that are similarly prone to eruptions originating within, or beneath, the shallow hydrothermal system shows mixed success using frequencies below 1 Hz. At Whakaari, velocity increases are detected at one station using a frequencies below 1 Hz prior to a phreatic eruption in 2016 (Yates et al., 2019; Caudron et al., 2021). However, no obvious precursor was observed prior to the onset of eruptive activity in 2012–2013 following a 10-year hiatus (Yates et al., 2019; Caudron et al., 2021). Further, the precursory velocity increase prior to the 2016 eruption was not detected by the only other seismic station located on the volcano approximately 1 km away. Similarly, at Ontake volcano, precursory decreases velocity changes were only recorded by stations in the summit area between 1–2 Hz prior to the 2014 phreatic eruption (Caudron et al., 2022). These observations highlight the localised nature of velocity changes due to pressure increases responsible for phreatic activity, where processing higher frequencies may be benefitial.

It is worth noting that long-term monitoring may not be the only use of volcanic tremor. From Chapter two, we observed that tremor regularly produced higher SNR cross-correlation functions than non-tremor sources in the short-term. This is especially true during the seismic crises of 2022 at Ruapehu, where values of SNR above 30 and phase stack amplitudes close to 1 were recorded (Fig. 6.4). These values are far larger than any period of tremor characterized in Chapter two, and are reflected in the high visual similarity of 1-day cross-correlation functions (Fig. 6.5). Similarly, the period of strongest tremor at 1–2 Hz forms a single cluster group at a low distance threshold of just 0.1 (Fig. 6.15). Such high similarity in the cross-correlation functions during this time period allowed for sub-daily velocity changes

to be computed while maintaining high values of correlation coefficient between the current and reference functions (Fig. 6.6c). From this, short-term variations can be clearly seen associated with individual precipitation events. Furthermore, there is some evidence that tremor is responding to these events. For example, a decrease in the tremor amplitude is observed following rainfall on 24–25 March. This is also true of the reduction in tremor amplitude from mid-May following significant rainfall over a week-long period. While we do not explore these observations further here, they suggest that highly stable tremor may be useful to study phenomena requiring a much higher temporal resolution than classically achieved with non-tremor sources. Such opportunities can be identified through the characterization of seismic datasets in advance of computing velocity changes.

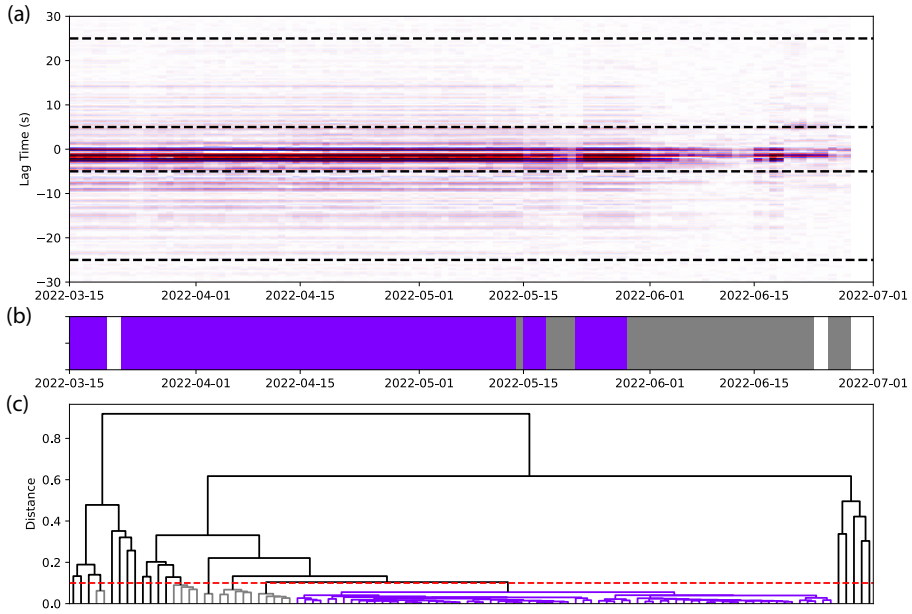

Figure 6.15: Cluster output using 1.0–2.0 Hz cross-correlation functions (CCFs) for station-pair FWVZ-MAVZ at Ruapehu (a): Correlogram showing amplitudes of CCFs at different lag times (red = positive, blue = negative). Dashed black lines show part of CCFs used in clustering. (b): Location of clusters in time, color-coded according to dendrogram output. Smaller clusters at given threshold are greyed out. (c) Dendrogram, with clusters defined at a distance threshold of 0.1 (dashed-red line).

Finally, hierarchical clustering of cross-correlation functions computed within the tremor frequency band may have potential towards monitoring changes in volcanic activity. At Stromboli, three periods of heightened activity — marked by an increase in the volcanic alert level (Giudicepietro et al., 2019, 2020) — formed a highly similar, single, cluster (Fig. 4.19). This indicates that valuable information on the volcanic system and the evolution of tremor are encoded within the cross-correlation functions. This therefore hints at the potential of the approach as a tool for identifying transitions to heightened states of activity. Similarly, at Tungurahua, clustering is able to identify a repeating pattern that aligns with some periods of eruptive activity. This observation implies that these periods bear significant similarities, reinforcing the notion that hierarchical clustering could serve as a beneficial tool

for classifying distinct phases of volcanic activity. This could possibly be combined with a supervised learning classification algorithm, for example using a partioning algorithm with training data based on unsupervised clustering, to monitor these transitions in real-time.

6.3 External sources of seismic velocity changes

The influence of external processes on velocity changes remains a challenge towards monitoring. At higher frequences (>1 Hz), this is particularly true, with velocity changes on the order of ±0.5% at Ruapehu and Stromboli. At Tungurahua, seasonal changes are measured on the order of ±0.2% at higher frequencies, also visible at lower frequencies but with a smaller amplitude.

In any context, the influence of seasonal processes may mask other features of interest. In volcanic settings, however, there seems to be an increased sensitivity to external processes. For example, Brenguier et al. (2014) found significantly larger velocity changes in volcanic regions across Japan in response to ground-shaking from the 2011 Tohoku-Oki earthquake. A similar observation was made by Lesage et al. (2014), where velocity decreases were observed following large tectonic earthquakes at Volcán de Colima. These were localized in the shallow layers of the volcano, with minimal variation detected outside the edifice or at depth. While these changes relate to tectonic activity, there is evidence at Ruapehu that meteorological processes also have a stronger effect on volcanoes. For example, the magnitude of annual fluctuations at stations located away from the edifice are significantly smaller (approximately ±0.1–0.2%) than changes on the volcano (albeit, opposite in sign).

The highly repetitive cross-correlation functions produced as a result of the 2022 seismic crises at Ruapehu provide an opportunity to examine rainfall-induced changes more closely. Such changes were clearly evident in six-hourly velocity changes computed using the four closest broadband stations (Fig. 6.6). Computing the spectral width using all stations shown in Fig. 6.1, reveals that seismic energy between 2–4 Hz in particular is coherent across the wider network (Fig. 6.16a). From this, velocity changes are computed with a temporal resolution of 1-day at all stations (Fig. 6.16b). The high quality of the data is again reflected by high correlation coefficients (>0.8) for the majority of data, despite many stations being more than 10 km from the tremor source (Fig. 6.16c). Finally, we record the velocity change at each station in response to four periods of rainfall (Fig. 6.16d). Showing these in map view, it is clear that the largest velocity decreases are observed at stations on the edifice of Ruapehu (Fig. 6.17).

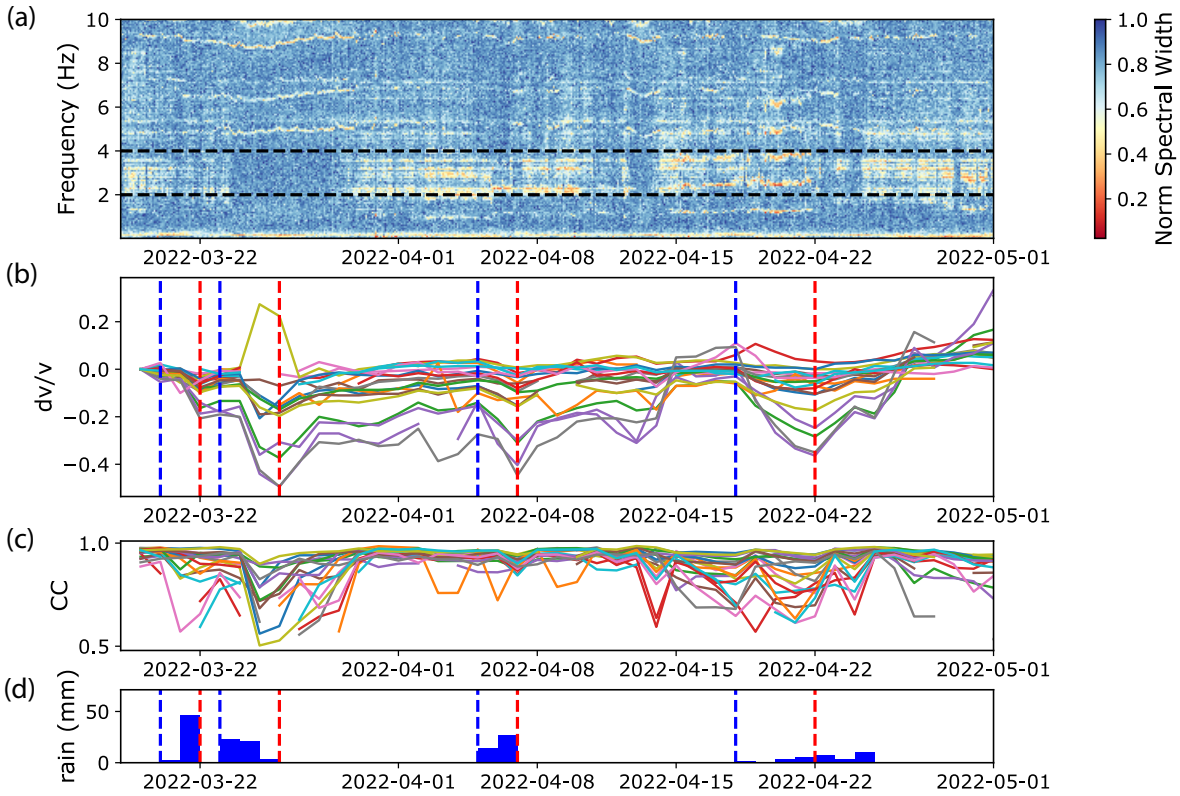

Figure 6.16: Velocity changes recorded at different single stations (average of EN, EZ, NZ) during period of strongest tremor. (a) Covariance matrix spectral width (normalized) computed using all available stations. (b) Velocity changes (1-day) of different stations. Dashed-blue and dashed-red lines correspond to times where measurements are made towards determining velocity change due to rainfall (blue = before, red = after). (c) Correlation coefficient with reference function (from 2022-03-15 to 2022-05-01). (d) Rainfall recorded at Chateau weather station. Measurement times shown as in (b).

A couple of reasons have been suggested for greater velocity changes at volcanoes relative to other locations following tectonic earthquakes. At Volcán de Colima, velocity decreases are considered to reflect material softening caused by seismic shaking (Lesage et al., 2014). This effect is stronger in poorly consolidated rock and partially saturated porous media (Van Den Abeele, 2002). Thus, the presence of lava and poorly consolidated pyroclastic deposits that make up the volcanic edifice are considered a possible explanation for enhanced changes due to tectonic activity (Lesage et al., 2014). Alternatively, others have considered the role of pressurized fluids in enhancing velocity changes in response to seismic shaking (Brenguier et al., 2014; Caudron et al., 2022). They argue that effective pressures in the crust are likely reduced under volcanic areas due to the presence of highly pressurized hydrothermal and/or magmatic fluids. The effect of reduced effective pressures is to increase the sensitivity of seismic velocities to stress changes (Zinszner et al., 1997). Clearly, the mechanism causing velocity changes due to tectonic earthquakes (i.e. dynamic shaking) is different from decreases due to precipitation events. However, it might be expected that the properties of volcanic rock that give rise to larger velocity changes

due to earthquakes may also explain the heightened sensitivity to meteorological effects.

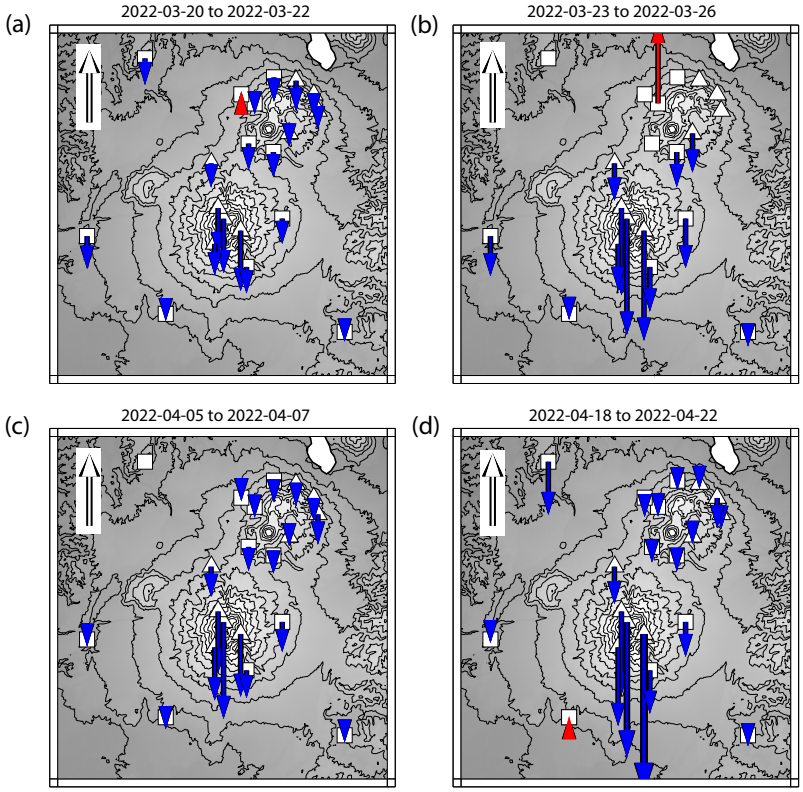

Figure 6.17: Velocity changes corresponding to four rainfall events highlighted in Figure 6.16 shown in map view. Length of white arrow corresponds to change of 0.2%, with blue (down) arrows indicating a velocity decrease and red (up) arrows indicating a velocity increase.

Both Ruapehu and the Tongariro complex host active hydrothermal systems (Leonard et al., 2021). At Ruapehu, hydrothermal activity is confined to the Crater Lake area (Christenson, 2000; Miller et al., 2020). In contrast, the Tongariro complex hosts a broad hydrothermal system (Miller and Williams-Jones, 2016; Miller et al., 2020), approximately located within the area enclosed by seismic stations in North-eastern portion of map (Fig. 6.1). Comparing the velocity decreases recorded by stations in the vicinity of Ruapehu with those closer to the Tongariro complex generally shows larger changes around Ruapehu (Fig. 6.18). If the presence of pressurized fluids was a controlling factor, it might be expected that the velocities around the Tongario complex would show a stronger change owing to the broader hydrothermal system. This may depend though on the extent that the hydrothermal fluids are pressurized. At Ruapehu, however, the limited spatial of extent of the hydrothermal system might suggest stations at a slightly lower elevation (e.g. those around 2000 m) would not be as sensitive to the hydrothermal system. Thus it is unclear whether the presence of pressurized fluids are responsible for greater sensitivity to external forcing. Alternatively, increases in microcrack density and porosity has been demonstrated to reduce the elastic moduli in volcano rock (Heap et al., 2020). Thus it may be that greater porosity and microcrack density in volcanic rock found

on the edifice of Ruapehu is responsible for the larger velocity changes following precipitation.

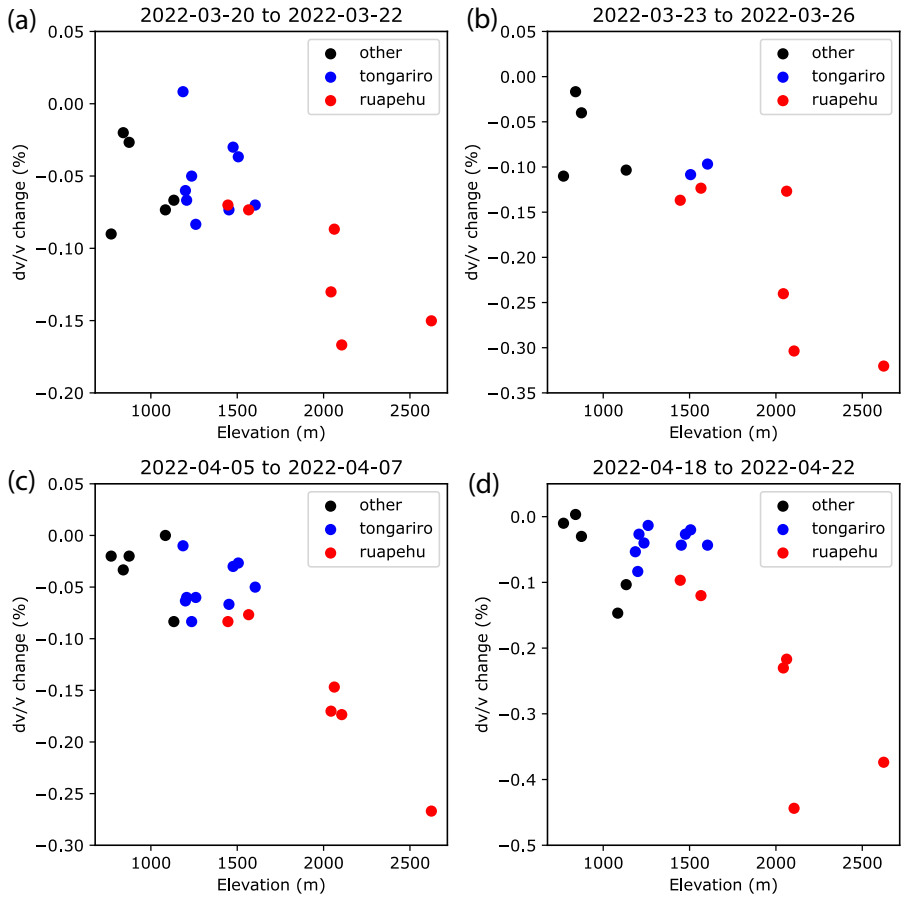

Figure 6.18: Plotting velocity change against elevation for the four different rainfall events, with stations split according to their location. These are stations in close proximity to Ruapehu (red; FWVZ, MAVZ, TRVZ, TUVZ, WHVZ, WNVZ), those closer to the Tongariro/Ngauruhoe (blue; ETVZ, KRVZ, NGZ, NOVZ, NTVZ, OTVZ, SNVZ, TMVZ, WTVZ) and all others (black; COVZ, MOVZ, MTVZ, PKVZ, TWVZ).

The greater sensitivity of seismic velocity changes measured on volcanoes to external forcing highlights the importance in developing methods towards correcting for meteorological processes. In both chapters three and four, an empirical approach (e.g. Rivet et al., 2015; Wang et al., 2017) was used to model pore pressure changes based on precipitation data. Some success was achieved with this approach. For example, a good fit was found between modeled pore pressures and seismic velocity changes recorded at a station at distance from Ruapehu (Fig. 5.8). At Stromboli, modeled pore pressure changes showed a good fit to the annual velocity changes for most of the nine-year period; however, the fit was worse during certain periods (Fig. 4.22). This could relate to the use of a weather station at a neighbouring island approximately 50 km away. However, a further complexity comes from the difficulty understanding the extent of groundwater recharge, and the overall groundwater system, on a volcanic edifices. Similarly, at Ruapehu, having a better understanding

of the snow distribution on the volcanic edifice would be useful towards discriminating between different processes. It would therefore be benefitial to conduct more targetted studies to investigate the influence of meteorological processes on seismic velocities in volcanic settings.

Alternatively, exploring the relationship between external forcing and seismic velocity changes represents an intriguing avenue for future research. For instance, an evolving sensitivity to external processes may suggest alteration of the volcanic system. Such has been suggested at Ontake volcano, where increased sensitivity to distal earthquakes was observed in the months prior to a phreatic eruption (Caudron et al., 2022). At Stromboli, we found evidence that the phase difference between modeled pore pressure changes and measured velocity changes was evolving over time (Section S4.3). This may reflect changes in the permeability through time. At Ruapehu in 2022, the largest velocity changes measured (of the four precipitation events in Fig. 6.16) do not appear to correspond to the time period of most significant rainfall. This suggests a greater sensitivity to external forcing later into the seismic crises. However, in both cases at Stromboli and Ruapehu, reaching a definitive conclusion is challenging due to the lack fo precise meteorological data. As previously mentioned, the weather station at Stromboli is 50 km away, whereas at Ruapehu the weather station is located at lower elevations in Whakapapa village. Thus, while this line of inquiry holds promise, the installation of closer meteorological stations is crucial to understanding any evolving relationship between seismic velocity changes and external processes.

6.4 Future perspectives towards monitoring explosive volcanoes

Monitoring seismic velocity changes can clearly provide intriguing insights into subsurface processes at volcanoes. However, it is also evident that complexities persist, especially in the interpretation of underlying mechanisms responsible. An obvious direction towards progressing our understanding is the continued retrospective analysis of existing datasets. This is especially true for volcanoes that transition rapidly from a state of relative quiescence (or milder activity) to sudden explosive events with minimal warning. Systems prone to phreatic, or phreatomagmatic eruptions, in particular deserve further attention, where precursory activity may be much more subtle due to the lack of significant movements of magma.

We consider several strategies that will be useful towards the continued development of this research topic. These are:

1. Improved instrumentation
2. Improved methodologies
3. Targeted studies of meteorological processes
4. Further multidisciplinary studies

Each of these are discussed individually in the subsequent sections.

6.4.1 Improved instrumentation

Improved seismic instrumentation close to active vents would be highly beneficial towards improving our ability to detect subtle changes within the shallow plumbing system. This is because, unlike deeper changes which might be evidenced over the wider volcanic edifice, shallow changes are often highly localised (e.g. Yates et al., 2019; Caudron et al., 2021, 2022). In environments prone to phreatic eruptions, for example, subtle pressurization that could indicate increased likelihood of eruptive activity (e.g. through the formation of a hydrothermal seal) may only be detectable very close to the active vent area (Stix and de Moor, 2018). This may explain why precusory velocity changes are observed for some gas-driven eruptions and not others (Mordret et al., 2010; Yates et al., 2019; Caudron et al., 2021). Further instrumentation may help to address this; specifically, whether the current inability to consistently detect precursors relates to the localized nature of the pressure source, or if there is genuinely an absense of detectable change. Furthermore, the deviation of velocities at the closest station to the summit at Stromboli (Chapter Three) shows that seismic stations installed close to active vents can provide additional information in open-conduit environments also. Understanding these deviations could offer valuable insights into the source mechanisms underlying observed changes.

In addition to traditional passive seismic networks, recent advancements in instrumentation may hold the key to enhancing the detection capabilities from coda wave interferometry. For example, recently developed short-period autonomous seismographs (nodes) can be used significantly increase the spatial density of data (Hansen and Schmandt, 2015; Thelen et al., 2022). This owes to their high portability and low-cost. While they can only record only a shorter duration due to battery limitations, targetted short-term deployments could be highly beneficial towards advancing our understanding of the spatial distribution of seismic velocity changes and their underlying mechanisms. Similarly, the use of fibre-optic cables through distributed acoustic sensing (DAS) to detect strain rates at meter scale spatial sampling also holds promise (Dou et al., 2017; Thelen et al., 2022). While the method is still in its infancy, it has shown promise towards monitoring time-lapse changes by effectively transforming fiber-optic cables into dense seismic arrays (Dou et al., 2017). Finally, further develops in the use of ocean-bottom seismometers would open the door for further monitoring of submarine volcanoes, such as the Hunga Tonga Ha'apai volcano responsible for one of the largest eruptions ever recorded by modern instruments (Poli and Shapiro, 2022; Matoza et al., 2022).

Finally, the continued difficulty in distinguishing between environmental and volcanic changes precipitates the importance of understanding the meteorology at a volcano. This is complicated, however, when reliant on stations located at greater distances from a volcano, or at lower elevations in the presence of a steep edifice (with climate more variable at higher elevations (Beniston, 2006)). Having meteorological stations located on volcanoes would significantly improve our ability to study the influence of environmental processes on subsurface velocity changes. Similarly, the installation of piezometers on volcanoes where possible (e.g. Violette

et al., 1997; Borgia et al., 2022) would be highly beneficial towards understanding the groundwater distribution in space and time.

6.4.2 Improved methodologies

Commonly used pre-processing procedures and techniques for measuring velocity changes (e.g the moving window cross-spectral or stretching technique) have generally proved successful towards monitoring sub-surface changes. However, new, alternative, approaches may offer advantages to what can currently be achieved with such standard techniques. For example, in Chapter Three, we applied the Wavelet Cross-Spectrum approach to study seismic velocity changes at Mt Ruapehu (Chapter four). This provides greater frequency resolution than can typically be achieved with the moving window cross-spectral or stretching technique (Mao et al., 2020). Consequently, it was possible to identify a velocity decrease within a narrow frequency range around the time of eruptive activity.

Such observations underscore the potential value that can be achieved in exploring more advanced methodologies. This also includes alternative techniques in processing, such as using the phrase cross-correlation function (Schimmel et al., 2011) or phase-weighted stack (Schimmel and Paulssen, 1997), or using advanced filters that can increase the SNR of cross-correlation functions (e.g. Baig et al., 2009; Stehly et al., 2011; Hadziioannou et al., 2011). Of particular interest is to compare these results with commonly used techniques, either in new locations or to compare with existing studie. This would enable the advantages and disadvantages of each to be better understood. Finally, it is important to consider that the commonly used approaches owe part of their popularity to being relatively straight-forward to implement. This is particularly useful towards being integrated into monitoring systems for real-time application (e.g. Duputel et al., 2009; Lecocq et al., 2014). Thus, new approaches would benefit from the development and sharing of open source codes/algorithms that allow wider testing amongst the volcano community.

Similarly, in Chapters Two and Three, we present methodologies that could be useful towards characterising seismic datasets towards coda wave interferometry. The use of hierarchical clustering (Chapter Three), in particular, we envisage has great potential. Its application in this thesis can be considered a proof of concept, where the choice to apply hierarchical clustering in the time domain with a correlation coefficient-based distance is relatively straightforward compared to more recent advancements in the field. Thus, there is significant scope for enhancement towards identifying subtle changes in cross-correlation functions through clustering algorithms. Furthermore, in the case of both the clustering methods and those presented in Chapter Two, transitioning to network-based approaches would be benefitial towards extracting information more efficiently from larger networks.

Finally, continued development and application of analytical and numerical models will likely prove crucial towards explaining complex observations of velocity changes (e.g. Hotovec-Ellis et al., 2022; Muzellec et al., 2023). This is especially true in locations that present both velocity increases and decrease at the same time. Furthermore, development of models that take into account topography and subsurface

heterogeneity would also be useful.

6.4.3 Targetted studies of meteorological processes in volcanic settings

Many studies have explored the influence of environmental processes on seismic velocity changes (Gassenmeier et al., 2014; Hillers et al., 2015a; Wang et al., 2017; Lecocq et al., 2017; Clements and Denolle, 2018; Miao et al., 2019; Feng et al., 2021; Lindner et al., 2021). However, while these studies are invaluable for enhancing our understanding of external influences on seismic velocities, most are conducted in non-volcanic settings. This complicates direct comparisons, as volcanic settings are inherently more heterogeneous and often have significant topographical variations. Of course, numerous studies have recognized velocity changes at volcanoes correlated with seasonal processes (e.g. Sens-Schönfelder and Wegler, 2006; Hotovec-Ellis et al., 2014; Rivet et al., 2015; Budi-Santoso and Lesage, 2016a; Donaldson et al., 2019). However, in many cases, the primary focus of these studies is to detect volcanic-induced changes, and thus comprehensive analysis of seasonally-induced changes is often secondary.

More targetted studies of the influence of environmental processes in volcanic settings would therefore be useful. This could be done towards the eventual correction of environmentally-induced changes, thus enhancing our ability to detect changes of volcanic origin. Alternatively, as highlighted in Section , it would be intriguing to examine whether the relationship between external forcings and seismic velocity changes evolves over time. However, such a determination requires a greater understanding of the controlling processes and also improved meteorological datasets (see Section 6.4.1 on improved instrumentation).

Finally, the observation of increased spring-time volcanic activity at Mt Ruapehu warrants further exploration of the influence of environmental processes on volcanic activity. Similarly, we find evidence that changes in the amplitude of volcanic tremor at the volcano may coincide with rainfall events during the seismic crises in 2022 (discussed briefly in Section 6.2). Therefore, a deeper investigation into the interplay between environmental processes and volcanic activity could provide new insights into their potentially related dynamics.

6.4.4 Further multidisciplinary studies

The landscape of volcano monitoring is continuously evolving, with innovative new techniques regularly developed that enable the collection of diverse types of observations. It remains challenging to interpret seismic velocity changes obtained via coda wave interferometry in isolation. However, the true strength of modern volcano science lies in its interdisciplinary approach. Integrating diverse observations spanning geophysics, geology, geochemistry, geodynamics, and remote sensing allows us to better constrain the underlying mechanisms. Similarly, the increasing availablity of longer-term datasets offers opportunities to investigate volcanic processes that may

be less discernable over shorter time periods. Furthermore, longer datasets could support, for example, more advanced statistical analysis of potential relationships between the different datasets.

Thus, going forward, continued methodological advances across the various disciplines, paired with increaes in the available data, will provide new opportunities to delve deeper into the complexities and nuances of volcanic systems.

6.5 Conclusions

This thesis presents five main chapters (in addition to an introduction chapter) that aim to contribute to the use of coda wave interferometry to monitor volcanoes, with an emphasis on explosive activity. Each chapter is briefly summarized below.

In Chapter Two, we sought to develop an understanding of seismic datasets for coda wave interferometry at volcanoes. Through network-based covariance matrix analysis, we examined the seismic wavefield at five different volcanoes. We then applied various measures that allowed for the characteristics of cross-correlation functions to be examined at different frequencies in response features of the seismic wavefield. An interesting finding was that volcanic tremor often produced cross-correlation functions with higher SNR and phase similarity than periods without tremor. This suggests that tremor could be further exploited for monitoring purposes. Oceanic noise sources and eruptive activity were also observed to have an influence on cross-correlation functions. However, this was variable across the five different volcanoes. The findings of this chapter can provide insights into the expectations one might have when applying coda wave interferometry in a new volcanic setting. Furthermore, the methods described can be applied themselves to develop further understanding of specific seismic datasets towards coda wave interferometry.

In Chapter Three, we employed hierarchical clustering to cross-correlation functions derived from data of two seismic networks. At Piton de la Fournaise, distinct clusters were identified corresponding to changes in both the seismic source and the medium following eruptive activity in 2010. At Ruapehu, cross-correlation functions within the frequency band of tremor showed similar characteristics, suggesting a relatively stable seismic source with potential for monitoring. Our findings underscore the utility of clustering cross-correlation functions to discern structure in seismic interferometry datasets. From these findings, we consider that clustering cross-correlation functions can be a useful inclusion in seismic interferometry studies. Identifying structure in datasets can aid in both decision making in processing and the interpretation of observed subsurface changes. Going forward, future studies may wish to build on these results by experimenting with alternative clustering algorithms, including network-based approaches, towards exploring further the potential of clustering cross-correlation functions as a similarity measure. These results were published in Yates et al. (2023).

In Chapter Four, we utilized coda wave interferometry to analyze seismic velocity changes at Stromboli, Italy, over nine years consisting of diverse volcanic events.

Special focus was given to two sudden 2019 paroxysms, that occured two months apart with minimal warning. Through hierarchical clustering, it is observed that these events, and a flank eruption in 2014, substantially impacted cross-correlation functions, suggesting significant changes to the shallow plumbing system. Velocity changes were computed in three frequency bands (0.5–1.0 Hz, 1.0–2.0 Hz,and 2.0–4.0 Hz), between pairs of seismic stations and also between components of individual single seismic stations. Results were comparable between both datasets, with seasonal processes dominating the velocity changes (especially above 1 Hz) with amplitudes of approximately $\pm$ 0.5 %. Two key features of interest are identied that are not readily explained by a seasonal process. The first feature is a long-term velocity decrease from early-2017 at lower frequencies (0.5–1.0 Hz), reaching a minimum in early-2019. Lower velocities are also recorded prior to the 2014 flank eruption and following the 2019 paroxysms. The second feature is an increasing velocity at the station closest to the summit (STRA) relative to the other stations at higher frequencies (primarily 2–4 Hz). Increases in magma overpressure in response to the input of new volatiles into the magmatic system are considered the most likely explanation to explain observations. The difference in velocities at lower and higher frequencies may be explained by anisotropy in the distrubtion of microcracks in the summit area, though this interpretation should be explored further. While sensitivity to deeper magma storage zones is not expected, increased pressurization within the shallow volcanic system towards more critical values may reflect an increasingly unstable deeper plumbing system. Thus coda wave interferometry can be a useful complementary tool towards monitoring the state of open conduit volcanoes.

In Chapter Five, we present findings that indicate a potential seasonal triggering of smaller-scale eruptions at Mount Ruapehu, New Zealand. Notably, the eruption chronology demonstrates a heightened activity during spring, with over half of eruptions in the months of September, October, and November. This pattern hints at a possible link between eruptions and seasonal processes. We study single-station seismic velocity changes over a four year period (2005–2009) that contains two gas-driven eruptions which occurred approximately one year apart in Spring months. The cross-wavelet transform approach was used to measure delay times in cross-correlation functions, facilitating the management of variable tremor at higher frequencies. Observations revealed significant seasonal variations in seismic velocities, with those on the volcano fluctuating around 0.5% and peaking during winter and times of maximum snow thickness. Off the volcano, velocities appear more influenced by precipitation-related pore pressure changes. After modeling of the influence of a snow-load, ground-water-level changes are considered the most likely mechanism responsible for seasonal velocity changes on the volcano. A smaller contribution from snow-loading/unloading is also evidenced. While ground-water related changes are suggested to have a stronger contribution to seasonal velocity changes than snow-loading/unloading, the latter process is considered more likely as a trigger of greater spring-time activity. Increased degassing associated with loading and unloading of the volcanic system is considered a plausible mechanism contributing to increased likelihood of spring-time, gas-driven eruptions. These interpretations should clearly be explored in further detail, both at Ruapehu and other volcanoes that show similar preferential spring-time activity. Doing so could have important implications for risk/hazard assessment at volcanoes that experience seasonal snow cover.

Finally, in Chapter Six (this chapter), we consolidate the findings of earlier chapters by re-examining the original objectives of the thesis. Two key themes are discussed in greater detail. These are the use of the tremor frequency band in coda wave interferometry and the external sources of velocity changes at volcanoes. Towards this, we also introduce two new datasets to bolster the discussion. These are data from Mt Ruapehu during a seismic crises in 2022 and data from Tungurahua, Ecuador, during a four year period consisting of various explosive activity. We conclude by discussing future perspectives towards monitoring volcanoes, particular those prone to explosive activity, through coda wave interferometry. It is hoped that these discussions and findings will pave the way for future research and inspire continued advancements in this critical field of study.

Bibliography

Agustsson, K., Stefánsson, R., Linde, A. T., Einarsson, P., Sacks, I. S., Gudmundsson, G. B., and Thorbjarnardóttir, B. (2000). Successful prediction and warning of the 2000 eruption of Hekla based on seismicity and strain changes. *Eos Trans. AGU*, 81(48).

Aiuppa, A., Bertagnini, A., Métrich, N., Moretti, R., Di Muro, A., Liuzzo, M., and Tamburello, G. (2010). A model of degassing for Stromboli volcano. *Earth and Planetary Science Letters*, 295(1-2):195–204.

Aiuppa, A., Bitetto, M., Donne, D. D., la Monica, F. P., Tamburello, G., Coppola, D., Schiava, M. D., Innocenti, L., Lacanna, G., Laiolo, M., Massimetti, F., Pistolesi, M., Silengo, M. C., and Ripepe, M. (2021). Volcanic CO2 tracks the incubation period of basaltic paroxysms. *Science Advances*, 7(38).

Aiuppa, A., Burton, M., Allard, P., Caltabiano, T., Giudice, G., Gurrieri, S., Liuzzo, M., and Salerno, G. (2011). First observational evidence for the CO2-driven origin of Stromboli's major explosions. *Solid Earth*, 2(2):135–142.

Aki, K. (1957). Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bull. Earth. Res. Inst.*, 35:415–456.

Aki, K. and Chouet, B. (1975). Origin of coda waves: Source, attenuation, and scattering effects. *Journal of Geophysical Research*, 80(23):3322–3342.

Albino, F., Pinel, V., and Sigmundsson, F. (2010). Influence of surface load variations on eruption likelihood: Application to two Icelandic subglacial volcanoes, Grímsvötn and Katla. *Geophysical Journal International*, 181(3):1510–1524.

Allard, P. (2010). A CO2-rich gas trigger of explosive paroxysms at Stromboli basaltic volcano, Italy. *Journal of Volcanology and Geothermal Research*, 189(3-4):363–374.

Allen, R. (1982). Automatic Phase Pickers: Their Present Use And Future Prospects. *Bulletin of the Seismological Society of America*, 72(6):225–242.

Aminzadeh, F. and Chatterjee, S. (1984). Applications Of Clustering In Exploration Seismology. *Geoexploration*, 23:147–159.

Andronico, D., Del Bello, E., D'Oriano, C., Landi, P., Pardini, F., Scarlato, P., de' Michieli Vitturi, M., Taddeucci, J., Cristaldi, A., Ciancitto, F., Pennacchia, F., Ricci, T., and Valentini, F. (2021). Uncovering the eruptive patterns of the 2019 double paroxysm eruption crisis of Stromboli volcano. *Nature Communications*, 12(1):1–14.

Apuani, T. and Corazzato, C. (2009). Numerical model of the Stromboli volcano (Italy) including the effect of magma pressure in the dyke system. *Rock Mechanics and Rock Engineering*, 42(1):53–72.

Ardid, A., Dempsey, D., Caudron, C., and Cronin, S. (2022). Seismic precursors to the Whakaari 2019 phreatic eruption are transferable to other eruptions and volcanoes. *Nature Communications*, 13(1):1–10.

Ardid, A., Dempsey, D., Caudron, C., Cronin, S. J., Miller, C. A., Melchor, I., Syahbana, D., and Kennedy, B. (2023). Detecting hidden fluid release episodes beneath crater lakes using eruptions precursors discovered from machine learning. *Journal of Geophysical Research : Solid Earth*, e2023JB026.

Arslan, A. T., Koca, M. Y., Aydogmus, T., Klapperich, H., and YIlmaz, H. R. (2008). Correlation of unconfined compressive strength with young's modulus and poisson's ratio in gypsum from Sivas (Turkey). *Rock Mechanics and Rock Engineering*, 41(6):941–950.

Backus, G. E. and Gilbert, J. F. (1967). Numerical Applications of a Formalism for Geophysical Inverse Problems. *Geophysical Journal of the Royal Astronomical Society*, 13(1-3):247–276.

Baig, A. M., Campillo, M., and Brenguier, F. (2009). Denoising Seismic noise cross correlations. *Journal of Geophysical Research: Solid Earth*, 114(8):1–12.

Baillard, C., Crawford, W. C., Ballu, V., Hibert, C., and Mangeney, A. (2014). An automatic kurtosis-based P-and S-phase picker designed for local seismic networks. *Bulletin of the Seismological Society of America*, 104(1):394–409.

Ballmer, S., Wolfe, C. J., Okubo, P. G., Haney, M. M., and Thurber, C. H. (2013). Ambient seismic noise interferometry in Hawai'i reveals long-range observability of volcanic tremor. *Geophysical Journal International*, 194(1):512–523.

Balmforth, N. J., Craster, R. V., and Rust, A. C. (2005). Instability in flow through elastic conduits and volcanic tremor. *Journal of Fluid Mechanics*, 527:353–377.

Barberi, F., Civetta, L., Gasparini, P., Innocenti, F., Scandone, R., and Villari, L. (1974). Evolution of a section of the Africa-Europe plate boundary: Paleomagnetic and volcanological evidence from Sicily. *Earth and Planetary Science Letters*, 22(2):123–132.

Bartlett, M. S. (1954). A Note on the Multiplying Factors for Various $\chi 2$ Approximations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 16(2):296–298.

Battaglia, J., Métaxian, J. P., and Garaebiti, E. (2012). Earthquake-volcano interaction imaged by coda wave interferometry. *Geophysical Research Letters*, 39(11):4–7.

Bayard, D., Stähli, M., Parriaux, A., and Flühler, H. (2005). The influence of seasonally frozen soil on the snowmelt runoff at two Alpine sites in southern Switzerland. *Journal of Hydrology*, 309(1-4):66–84.

Behr, Y., Townend, J., Bannister, S., and Savage, M. K. (2010). Shear velocity structure of the Northland Peninsula, New Zealand, inferred from ambient noise correlations. *Journal of Geophysical Research: Solid Earth*, 115(5):1–12.

Belien, I. B., Cashman, K. V., and Rempel, A. W. (2010). Gas accumulation in particle-rich suspensions and implications for bubble populations in crystal-rich magma. *Earth and Planetary Science Letters*, 297(1-2):133–140.

Beniston, M. (2006). Mountain weather and climate: A general overview and a focus on climatic change in the Alps. *Hydrobiologia*, 562(1):3–16.

Bennington, N. L., Haney, M., De Angelis, S., Thurber, C. H., and Freymueller, J. (2015). Monitoring changes in seismic velocity related to an ongoing rapid inflation event at Okmok volcano, Alaska. *Journal of Geophysical Research: Solid Earth*, 120(8):5664–5676.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., and Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, 169(3):1239–1260.

Bensen, G. D., Ritzwoller, M. H., and Shapiro, N. M. (2008). Broadband ambient noise surface wave tomography across the United States. *Journal of Geophysical Research: Solid Earth*, 113(5):1–21.

Benvegna, F., Lo Bosco, G., and Tegolo, D. (2013). Dissimilarity measures for the identification of earthquake focal mechanisms. *International Conference on Image Analysis and Processing*, pages 500–509.

Bergen, K. J., Johnson, P. A., De Hoop, M. V., and Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science*, 363(6433).

Boese, C. M., Wotherspoon, L., Alvarez, M., and Malin, P. (2015). Analysis of anthropogenic and natural noise from multilevel borehole seismometers in an urban environment, Auckland, New Zealand. *Bulletin of the Seismological Society of America*, 105(1):285–299.

Borgia, A., Mazzoldi, A., Micheli, L., Grieco, G., Calcara, M., and Balducci, C. (2022). The Geothermal Power Plants of Amiata Volcano, Italy: Impacts on Freshwater Aquifers, Seismicity and Air. *Progress in Volcanology*, 35.

Bosman, A., Chiocci, F. L., and Romagnoli, C. (2009). Morpho-structural setting of Stromboli volcano revealed by high-resolution bathymetry and backscatter data of its submarine portions. *Bulletin of Volcanology*, 71(9):1007–1019.

Braun, T., Neuberg, J., and Ripepe, M. (1996). On the origin of the long-period tremor recorded at Stromboli volcano (Italy). *Annali di geofisica*, 39(2).

Bredemeyer, S. and Hansteen, T. H. (2014). 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.

Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N. M., Nadeau, R. M., and Larose, E. (2008a). Postseismic relaxation along the San Andreas fault at Parkfield from continuous seismological observations. *Science*, 321(5895):1478–1481.

Brenguier, F., Campillo, M., Takeda, T., Aoki, Y., Shapiro, N. M., Briand, X., Emoto, K., and Miyake, H. (2014). Mapping pressurized volcanic fluids from induced crustal seismic velocity drops. *Science*, 345(6192):80–82.

Brenguier, F., Kowalski, P., Audacher, T. S., Ferrazzini, V., Lauret, F., Boissier, P., Catherine, P., Lemarchand, A., Pequegnat, C., Meric, O., Pardo, C., Peltier, A., Tait, S., Shapiro, N. M., Campillo, M., and Di Muro, A. (2012). First results from the undervolc high resolution seismic and GPS network deployed on Piton de la Fournaise volcano. *Seismological Research Letters*, 83(1):97–102.

Brenguier, F., Shapiro, N. M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O., and Nercessian, A. (2008b). Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience*, 1(2):126–130.

Brenguier, F., Shapiro, N. M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O., and Nercessian, A. (2008c). Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience*, 1(2):126–130.

Brenguier, F., Shapiro, N. M., Campillo, M., Nercessian, A., and Ferrazzini, V. (2007). 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlations. *Geophysical Research Letters*, 34(2):2–6.

Brooks, P. D., Williams, M. W., Walker, D. A., and Schmidt, S. K. (1995). The Niwot Ridge snow fence experiment: biogeochemical responses to changes in the seasonal snowpack.

Budd, L., Griggs, S., Howarth, D., and Ison, S. (2011). A fiasco of volcanic proportions? eyjafjallajökull and the closure of European airspace. *Mobilities*, 6(1):31–40.

Budi-Santoso, A. and Lesage, P. (2016a). Velocity variations associated with the large 2010 eruption of Merapi volcano, Java, retrieved from seismic multiplets and ambient noise cross-correlation. *Geophysical Journal International*, 206(1):221–240.

Budi-Santoso, A. and Lesage, P. (2016b). Velocity variations associated with the large 2010 eruption of Merapi volcano, Java, retrieved from seismic multiplets and ambient noise cross-correlation. *Geophysical Journal International*, 206(1):221–240.

Burtin, A., Bollinger, L., Vergne, J., Cattin, R., and Nábělek, J. L. (2008). Spectral analysis of seismic noise induced by rivers: A new tool to monitor spatiotemporal changes in stream hydrodynamics. *Journal of Geophysical Research: Solid Earth*, 113(5):1–14.

Burton, M., Allard, P., Mure, F., and La Spina, A. (2007). Magmatic gas composition reveals the source depth of slug-driven strombolian explosive activity. *Science*, 317(5835):227–230.

Bustillos A., J. E., Romero, J. E., Guevara C., A., and Díaz-Alvarado, J. (2018). Tephra fallout from the long-lasting Tungurahua eruptive cycle (1999-2014): Variations through eruptive style transition and deposition processes. *Andean Geology*, 45(1):47.

Calò, M., López Mazariegos, E. A., Tramelli, A., and Orazi, M. (2021). Hydrothermal systems characterization of the Stromboli volcano using spatial and temporal changes of the seismic velocities. *Journal of Volcanology and Geothermal Research*, 411.

Calvari, S., Di Traglia, F., Ganci, G., Bruno, V., Ciancitto, F., Di Lieto, B., Gambino, S., Garcia, A., Giudicepietro, F., Inguaggiato, S., Vita, F., Cangemi, M., Inguaggiato, C., Macedonio, G., Mattia, M., Miraglia, L., Nolesini, T., Pompilio, M., Romano, P., Salerno, G., Casagli, N., Re, G., Del Carlo, P., Di Roberto, A., Cappello, A., Corradino, C., Amato, E., Torrisi, F., Del Negro, C., Esposito, A. M., De Cesare, W., Caputo, T., Buongiorno, M. F., Musacchio, M., Romaniello, V., Silvestri, M., Marotta, E., Avino, R., Avvisati, G., and Belviso, P. (2022). Multi-parametric study of an eruptive phase comprising unrest, major explosions, crater failure, pyroclastic density currents and lava flows: Stromboli volcano, 1 December 2020–30 June 2021. *Frontiers in Earth Science*, 10(August 2019):1–24.

Calvari, S., Giudicepietro, F., Di Traglia, F., Bonaccorso, A., Macedonio, G., and Casagli, N. (2021). Variable magnitude and intensity of strombolian explosions: Focus on the eruptive processes for a first classification scheme for stromboli volcano (italy). *Remote Sensing*, 13(5):1–30.

Campillo, M. and Paul, A. (2003). Long-Range Correlations in the Diffuse Seismic Coda. *Science*, 299(5606):547–549.

Cannata, A., Cannavò, F., Montalto, P., Ercoli, M., Mancinelli, P., Pauselli, C., and Leto, G. (2017). Monitoring crustal changes at volcanoes by seismic noise interferometry: Mt. Etna case of study. *Journal of Volcanology and Geothermal Research*, 337:165–174.

Caricchi, L., Townsend, M., Rivalta, E., and Namiki, A. (2021). The build-up and triggers of volcanic eruptions. *Nature Reviews Earth and Environment*, 2(7):458–476.

Carniel, R., Jolly, A. D., and Barbui, L. (2013). Analysis of phreatic events at Ruapehu volcano, New Zealand using a new SOM approach. *Journal of Volcanology and Geothermal Research*, 254:69–79.

Carniel, R. and Tárraga, M. (2006). Can tectonic events change volcanic tremor at Stromboli? *Geophysical Research Letters*, 33(20):4–6.

Casagli, N., Tibaldi, A., Merri, A., Del Ventisette, C., Apuani, T., Guerri, L., Fortuny-Guasch, J., and Tarchi, D. (2009). Deformation of Stromboli Volcano (Italy) during the 2007 eruption revealed by radar interferometry, numerical modelling and structural geological field data. *Journal of Volcanology and Geothermal Research*, 182(3-4):182–200.

Cassidy, M., Manga, M., Cashman, K., and Bachmann, O. (2018). Controls on explosive-effusive volcanic eruption styles. *Nature Communications*, 9(1).

Caudron, C., Aoki, Y., Lecocq, T., De Plaen, R., Soubestre, J., Mordret, A., Seydoux, L., and Terakawa, T. (2022). Hidden pressurized fluids prior to the 2014 phreatic eruption at Mt Ontake. *Nature Communications*, 13(1):1–9.

Caudron, C., Chardot, L., Girona, T., Aoki, Y., and Fournier, N. (2020). Editorial: Towards Improved Forecasting of Volcanic Eruptions. *Frontiers in Earth Science*, 8(February):1–3.

Caudron, C., Girona, T., Jolly, A., Christenson, B., Savage, M. K., Carniel, R., Lecocq, T., Kennedy, B., Lokmer, I., Yates, A., Hamling, I., Park, I., and Kilgour, G. (2021). A quest for unrest in multiparameter observations at Whakaari / White Island volcano , New Zealand 2007 – 2018. *Earth, Planets and Space*.

Caudron, C., Lecocq, T., Syahbana, D. K., and Mccausland, W. (2015). Stress and mass changes at a " wet " volcano : Example during the 2011 – 2012 volcanic unrest at Kawah Ijen volcano (Indonesia). *Journal of Geophysical Research : Solid Earth*, 120:5117–5134.

Chardot, L., Jolly, A. D., Kennedy, B. M., Fournier, N., and Sherburn, S. (2015). Using volcanic tremor for eruption forecasting at White Island volcano (Whakaari), New Zealand. *Journal of Volcanology and Geothermal Research*, 302:11–23.

Chen, J. H., Froment, B., Liu, Q. Y., and Campillo, M. (2010). Distribution of seismic wave speed changes associated with the 12 May 2008 Mw 7.9 Wenchuan earthquake. *Geophysical Research Letters*, 37(18).

Cho, K. H., Herrmann, R. B., Ammon, C. J., and Lee, K. (2007). Imaging the upper crust of the Korean peninsula by surface-wave tomography. *Bulletin of the Seismological Society of America*, 97(1B):198–207.

Chouet, B., De Luca, G., Milana, G., Dawson, P., Martini, M., and Scarpa, R. (1998). 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.

Chouet, B., Saccorotti, G., Martini, M., Dawson, P., De Luca, G., Milana, G., and Scarpa, R. (1997). 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.

Chouet, B. A. and Matoza, R. S. (2013). A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. *Journal of Volcanology and Geothermal Research*, 252:108–175.

Christenson, B. W. (2000). Geochemistry of fluids associated with the 1995-1996 eruption of Mt. Ruapehu, New Zealand: Signatures and processes in the magmatic-hydrothermal system. *Journal of Volcanology and Geothermal Research*, 97(1-4):1–30.

Christenson, B. W., Reyes, A. G., Young, R., Moebis, A., Sherburn, S., Cole-Baker, J., and Britten, K. (2010). Cyclic processes and factors leading to phreatic eruption events: Insights from the 25 September 2007 eruption through Ruapehu Crater Lake, New Zealand. *Journal of Volcanology and Geothermal Research*, 191(1-2):15–32.

Christophersen, A., Behr, Y., and Miller, C. (2022). Automated Eruption Forecasting at Frequently Active Volcanoes Using Bayesian Networks Learned From Monitoring Data and Expert Elicitation: Application to Mt Ruapehu, Aotearoa, New Zealand. *Frontiers in Earth Science*, 10(July):1–21.

Cigolini, C., Poggi, P., Ripepe, M., Laiolo, M., Ciamberlini, C., Delle Donne, D., Ulivieri, G., Coppola, D., Lacanna, G., Marchetti, E., Piscopo, D., and Genco, R. (2009). Radon surveys and real-time monitoring at Stromboli volcano: Influence of soil temperature, atmospheric pressure and tidal forces on 222Rn degassing. *Journal of Volcanology and Geothermal Research*, 184(3-4):381–388.

Claerbout, J. F. (1968). Synthesis of a Layered Medium from its Acoustic Transmission Response. *Geophysics*, 33(2):264–269.

Clarke, D., Brenguier, F., Froger, J. L., Shapiro, N. M., Peltier, A., and Staudacher, T. (2013). Timing of a large volcanic flank movement at Piton de la Fournaise Volcano using noise-based seismic monitoring and ground deformation measurements. *Geophysical Journal International*, 195(2):1132–1140.

Clarke, D., Zaccarelli, L., Shapiro, N. M., and Brenguier, F. (2011). Assessment of resolution and accuracy of the Moving Window Cross Spectral technique for monitoring crustal temporal variations using ambient seismic noise. *Geophysical Journal International*, 186(2):867–882.

Clarke, J., Adam, L., van Wijk, K., and Sarout, J. (2020). The influence of fluid type on elastic wave velocity and attenuation in volcanic rocks. *Journal of Volcanology and Geothermal Research*, 403.

Clements, T. and Denolle, M. A. (2018). Tracking Groundwater Levels Using the Ambient Seismic Field. *Geophysical Research Letters*, 45(13):6459–6465.

Conway, C. E., Leonard, G. S., Townsend, D. B., Calvert, A. T., Wilson, C. J., Gamble, J. A., and Eaves, S. R. (2016). A high-resolution 40Ar/39Ar lava chronology and edifice construction history for Ruapehu volcano, New Zealand. *Journal of Volcanology and Geothermal Research*, 327:152–179.

Conway, C. E., Pure, L. R., and Ishizuka, O. (2023). An assessment of potential causal links between deglaciation and eruption rates at arc volcanoes. *Frontiers in Earth Science*, 11(April).

Covseisnet (2020). https://github.com/covseisnet/covseisnet.

Cubuk-Sabuncu, Y., Jónsdóttir, K., Caudron, C., Lecocq, T., Parks, M. M., Geirsson, H., and Mordret, A. (2021a). Temporal Seismic Velocity Changes During the 2020 Rapid Inflation at Mt. Þorbjörn-Svartsengi, Iceland, Using Seismic Ambient Noise. *Geophysical Research Letters*, 48(11):1–10.

Cubuk-Sabuncu, Y., Jónsdóttir, K., Caudron, C., Lecocq, T., Parks, M. M., Geirsson, H., and Mordret, A. (2021b). Temporal Seismic Velocity Changes During the 2020 Rapid Inflation at Mt. Þorbjörn-Svartsengi, Iceland, Using Seismic Ambient Noise. *Geophysical Research Letters*, 48(11):1–10.

Czecze, B. and Bondár, I. (2019). Hierarchical cluster analysis and multiple event relocation of seismic event clusters in Hungary between 2000 and 2016. *Journal of Seismology*, 23(6):1313–1326.

D'Alessandro, A., Mangano, G., D'Anna, G., and Luzio, D. (2013). Waveforms clustering and single-station location of microearthquake multiplets recorded in the northern sicilian offshore region. *Geophysical Journal International*, 194(3):1789–1809.

Davy, C., Stutzmann, E., Barruol, G., Fontaine, F. R., and Schimmel, M. (2015). Sources of secondary microseisms in the Indian Ocean. *Geophysical Journal International*, 202(2):1180–1189.

De Groote, J. (2022). Intra-annual snow cover variatons as a potential explanation for eruption seasonality at Villarrica volcano Chile. *Masters Thesis*.

De Plaen, R. S., Cannata, A., Cannavo', F., Caudron, C., Lecocq, T., and Francis, O. (2019). Temporal changes of seismic velocity caused by volcanic activity at Mt. Etna revealed by the autocorrelation of ambient seismic noise. *Frontiers in Earth Science*, 6(January):1–11.

De Plaen, R. S. M., Lecocq, T., Caudron, C., Ferrazzini, V., and Francis, O. (2016). Single station monitoring of volcanoes using seismic ambient noise. *Geophysical Research Letters*, 43(16):8511–8518.

Delle Piane, C., Arena, A., Sarout, J., Esteban, L., and Cazes, E. (2015). Microcrack enhanced permeability in tight rocks: An experimental and microstructural study. *Tectonophysics*, 665:149–156.

Delouche, E. (2023). Monitoring seismic velocity changes in the upper crust and aquifer dynamics. *Doctoral dissertation, Université Grenoble Alpes [2020-....])*.

Deyasi, K., Chakraborty, A., and Banerjee, A. (2017). Network similarity and statistical analysis of earthquake seismic data. *Physica A: Statistical Mechanics and its Applications*, 481:224–234.

D'Hour, V., Schimmel, M., Do Nascimento, A. F., Ferreira, J. M., and Lima Neto, H. C. (2016). Detection of Subtle Hydromechanical Medium Changes Caused By a Small-Magnitude Earthquake Swarm in NE Brazil. *Pure and Applied Geophysics*, 173(4):1097–1113.

Di Lieto, B., Romano, P., Scarpa, R., and Linde, A. T. (2020). Strain Signals Before and During Paroxysmal Activity at Stromboli Volcano, Italy. *Geophysical Research Letters*, 47(21).

Di Traglia, F., Battaglia, M., Nolesini, T., Lagomarsino, D., and Casagli, N. (2015). Shifts in the eruptive styles at Stromboli in 2010-2014 revealed by ground-based InSAR data. *Scientific Reports*, 5:1–11.

Di Traglia, F., Calvari, S., D'Auria, L., Nolesini, T., Bonaccorso, A., Fornaciai, A., Esposito, A., Cristaldi, A., Favalli, M., and Casagli, N. (2018). The 2014 effusive eruption at stromboli: New insights from in situ and remote-sensing measurements. *Remote Sensing*, 10(12).

Di Traglia, F., De Luca, C., Manzo, M., Nolesini, T., Casagli, N., Lanari, R., and Casu, F. (2021). Joint exploitation of space-borne and ground-based multitemporal InSAR measurements for volcano monitoring: The Stromboli volcano case study. *Remote Sensing of Environment*, 260(March):112441.

Di Traglia, F., Del Ventisette, C., Rosi, M., Mugnai, F., Intrieri, E., Moretti, S., and Casagli, N. (2013). Ground-based InSAR reveals conduit pressurization pulses at Stromboli volcano. *Terra Nova*, 25(3):192–198.

Di Traglia, F., Nolesini, T., Intrieri, E., Mugnai, F., Leva, D., Rosi, M., and Casagli, N. (2014). Review of ten years of volcano deformations recorded by the ground-based InSAR monitoring system at Stromboli volcano: A tool to mitigate volcano flank dynamics and intense volcanic activity. *Earth-Science Reviews*, 139:317–335.

Dinger, F., Bredemeyer, S., Arellano, S., Bobrowski, N., Platt, U., and Wagner, T. (2019). On the link between Earth tides and volcanic degassing. *Solid Earth*, 10(3):725–740.

Doetsch, J., Gischig, V. S., Villiger, L., Krietsch, H., Nejati, M., Amann, F., Jalali, M., Madonna, C., Maurer, H., Wiemer, S., Driesner, T., and Giardini, D. (2018). Subsurface Fluid Pressure and Rock Deformation Monitoring Using Seismic Velocity Observations. *Geophysical Research Letters*, 45(19):10,389–10,397.

Domínguez, T., Flores, F., and Reyes, G. (2003). Temporal change in coda wave attenuation observed at Volcán de Colima, México before the 1998 eruption. *Journal of Volcanology and Geothermal Research*, 125(3-4):215–223.

Donaldson, C., Caudron, C., Green, R. G., Thelen, W. A., and White, R. S. (2017). Relative seismic velocity variations correlate with deformation at Kīlauea volcano. *Science Advances*, 3(6):e1700219.

Donaldson, C., Winder, T., Caudron, C., and White, R. S. (2019). Crustal seismic velocity responds to a magmatic intrusion and seasonal loading in Iceland's Northern Volcanic Zone. *Science Advances*, 5(11).

Dou, S., Lindsey, N., Wagner, A. M., Daley, T. M., Freifeld, B., Robertson, M., Peterson, J., Ulrich, C., Martin, E. R., and Ajo-Franklin, J. B. (2017). Distributed Acoustic Sensing for Seismic Monitoring of the Near Surface: A Traffic-Noise Interferometry Case Study. *Scientific Reports*, 7(1):1–12.

Duputel, Z., Ferrazzini, V., Brenguier, F., Shapiro, N., Campillo, M., and Nercessian, A. (2009). Real time monitoring of relative velocity changes using ambient seismic noise at the Piton de la Fournaise volcano (La Réunion) from January 2006 to June 2007. *Journal of Volcanology and Geothermal Research*, 184(1-2):164–173.

Earman, S., Campbell, A. R., Phillips, F. M., and Newman, B. D. (2006). Isotopic exchange between snow and atmospheric water vapor: Estimation of the snowmelt component of groundwater recharge in the southwestern United States. *Journal of Geophysical Research Atmospheres*, 111(9):1–18.

Eberhart-Phillips, D., Han, D.-H., and Zoback, M. (1989). Empirical relationships among seismic velocity, effective pressure, porosity, and clay content in sandstone. *Geophysics*, 54(1):82–89.

Edmonds, M., Liu, E. J., and Cashman, K. V. (2022). Open-vent volcanoes fuelled by depth-integrated magma degassing. *Bulletin of Volcanology*, 84(3).

Elsworth, D., Voight, B., Thompson, G., and Young, S. R. (2004). Thermal-hydrologic mechanism for rainfall-triggered collapse of lava domes. *Geology*, 32(11):969–972.

Farge, M. (1992). Wavelet transforms and their applications to turbulence. *Annual review of fluid mechanics*, 24(1):395–458.

Farquharson, J. I. and Amelung, F. (2020). Extreme rainfall triggered the 2018 rift eruption at Kīlauea Volcano. *Nature*, 580(7804):491–495.

Federico, C., Inguaggiato, S., Liotta, M., Rizzo, A. L., and Vita, F. (2023). Decadal Monitoring of the Hydrothermal System of Stromboli Volcano, Italy. *Geochemistry, Geophysics, Geosystems*, 24(9).

Fehler, M., Roberts, P., and Fairbanks, T. (1988). A temporal change in coda wave attenuation observed during an eruption of Mount St. Helens. *Journal of Geophysical Research*, 93(B5):4367–4373.

Feng, K. F., Huang, H. H., Hsu, Y. J., and Wu, Y. M. (2021). Controls on Seasonal Variations of Crustal Seismic Velocity in Taiwan Using Single-Station Cross-Component Analysis of Ambient Noise Interferometry. *Journal of Geophysical Research: Solid Earth*, 126(11).

Fichtner, A., Stehly, L., Ermert, L., and Boehm, C. (2017). Generalized interferometry - I: Theory for interstation correlations. *Geophysical Journal International*, 208(2):603–638.

Finizola, A., Aubert, M., Revil, A., Schütze, C., and Sortino, F. (2009). Importance of structural history in the summit area of Stromboli during the 2002-2003 eruptive crisis inferred from temperature, soil CO2, self-potential, and electrical resistivity tomography. *Journal of Volcanology and Geothermal Research*, 183(3-4):213–227.

Finizola, A., Sortino, F., Lénat, J. F., and Valenza, M. (2002). Fluid circulation at Stromboli volcano (Aeolian Islands, Italy) from self-potential and CO2 surveys. *Journal of Volcanology and Geothermal Research*, 116(1-2):1–18.

Flinders, A. F., Caudron, C., Johanson, I. A., Taira, T., Shiro, B., and Haney, M. (2020). Seismic velocity variations associated with the 2018 lower East Rift Zone eruption of Kīlauea, Hawai'i. *Bulletin of Volcanology*, 82(6).

Fokker, E., Ruigrok, E., Hawkins, R., and Trampert, J. (2021). Physics-based relationship for pore pressure and vertical stress monitoring using seismic velocity variations. *Remote Sensing*, 13(14):1–20.

Fokker, E., Ruigrok, E., Hawkins, R., and Trampert, J. (2023). 4D Physics-Based Pore Pressure Monitoring Using Passive Image Interferometry Geophysical Research Letters. *Geophysical Research Letters*, 50.

Forster, C. and Smith, L. (1988). Groundwater flow systems in mountainous terrain: 2. Controlling factors. *Water Resources Research*, 24(7):1011–1023.

Francalanci, L., Tommasini, S., and Conticelli, S. (2004). The volcanic activity of Stromboli in the 1906-1998 AD period: Mineralogical, geochemical and isotope data relevant to the understanding of the plumbing system. *Journal of Volcanology and Geothermal Research*, 131(1-2):179–211.

Friedrich, A., Krüger, F., and Klinge, K. (1998). Ocean-generated microseismic noise located with the Gräfenberg array. *Journal of Seismology*, 2(1):47–64.

Froment, B., Campillo, M., Chen, J. H., and Liu, Q. Y. (2013). Deformation at depth associated with the 12 May 2008 MW 7.9 Wenchuan earthquake from seismic ambient noise monitoring. *Geophysical Research Letters*, 40(1):78–82.

Froment, B., Campillo, M., Roux, P., Gouédard, P., Verdel, A., and Weaver, R. L. (2010). Estimation of the effect of nonisotropically distributed energy on the apparent arrival time in correlations. *Geophysics*, 75(5).

Gaete, A., Walter, T. R., Bredemeyer, S., Zimmer, M., Kujawa, C., Franco Marin, L., San Martin, J., and Bucarey Parra, C. (2020). Processes culminating in the 2015 phreatic explosion at Lascar volcano, Chile, evidenced by multiparametric data. *Natural Hazards and Earth System Sciences*, 20(2):377–397.

Galanaki, E., Lagouvardos, K., Kotroni, V., Flaounas, E., and Argiriou, A. (2018). Thunderstorm climatology in the Mediterranean using cloud-to-ground lightning observations. *Atmospheric Research*, 207(February):136–144.

Gassenmeier, M., Sens-Schönfelder, C., Delatre, M., and Korn, M. (2014). Monitoring of environmental influences on seismic velocity at the geological storage site for CO2 in Ketzin (Germany) with ambient seismic noise. *Geophysical Journal International*, 200(1):524–533.

Gauthier, P. J., Le Cloarec, M. F., and Condomines, M. (2000). Degassing processes at Stromboli volcano inferred from short-lived disequilibria (210Pb-210Bi-210Po) in volcanic gases. *Journal of Volcanology and Geothermal Research*, 102(1-2):1–19.

Gebrande, H. (1982). 3.1. 2.3 Velocity-density relations: 3.1. 2 Elastic wave velocities and constants of elasticity at normal conditions. *Subvolume B*, pages 17–20.

GeoNet (2016). End of Minor Eruptive Activity at White Island (Whakaari). *Volcanic Alert Bulletin*.

Geonet (2022a). Volcanic Activity Bulletin RUA - 2022/11. https://www.geonet.org.nz/vabs/36DmsaLzVWczmz5uyAs71p.

Geonet (2022b). Volcanic Activity Bulletin RUA - 2022/18. https://www.geonet.org.nz/vabs/6ek5OkTf13CM9BqY538frb.

Geonet (2022c). Volcanic Activity Bulletin RUA-2022/02. https://www.geonet.org.nz/vabs/2Cy40dBlXKJ0KXP4eAfMM2.

Giggenbach, W. F. (1996). Chemical composition of volcanic gases in Japan. *Monitoring and mitigation of volcano hazards*, pages 221–256.

Girona, T., Huber, C., and Caudron, C. (2018). Sensitivity to lunar cycles prior to the 2007 eruption of Ruapehu volcano. *Scientific Reports*, 8(1):1476.

Giudicepietro, F., Calvari, S., Alparone, S., Bianco, F., Bonaccorso, A., Bruno, V., Caputo, T., Cristaldi, A., D'Auria, L., Cesare, W. D., Lieto, B. D., Esposito, A. M., Gambino, S., Inguaggiato, S., Macedonio, G., Martini, M., Mattia, M., Orazi, M., Paonita, A., Peluso, R., Privitera, E., Romano, P., Scarpato, G., Tramelli, A., and Vita, F. (2019). Integration of ground-based remote-sensing and in situ multidisciplinary monitoring data to analyze the eruptive activity of stromboli volcano in 2017-2018. *Remote Sensing*, 11(15).

Giudicepietro, F., López, C., Macedonio, G., Alparone, S., Bianco, F., Calvari, S., De Cesare, W., Delle Donne, D., Di Lieto, B., Esposito, A. M., Orazi, M., Peluso, R., Privitera, E., Romano, P., Scarpato, G., and Tramelli, A. (2020). Geophysical precursors of the July-August 2019 paroxysmal eruptive phase and their implications for Stromboli volcano (Italy) monitoring. *Scientific Reports*, 10(1):1–16.

Gledhill, K., Ristau, J., Reyners, M., Fry, B., and Holden, C. (2010). The darfield (canterbury) earthquake of september 2010: Preliminary seismological report. *Bulletin of the New Zealand Society for Earthquake Engineering*, 43(4):215–221.

Global Volcanism Program (2023). Volcanoes of the World (v. 5.1.1; 17 Aug 2023). *Smithsonian*, https://do.

Godfrey, H. J., Fry, B., and Savage, M. K. (2017). Shear-wave velocity structure of the Tongariro Volcanic Centre, New Zealand: Fast Rayleigh and slow Love waves indicate strong shallow anisotropy. *Journal of Volcanology and Geothermal Research*, 336:33–50.

Gómez-García, C., Brenguier, F., Boué, P., Shapiro, N. M., Droznin, D. V., Droznina, S. Y., Senyukov, S. L., and Gordeev, E. I. (2018). Retrieving robust noise-based seismic velocity changes from sparse data sets: synthetic tests and application to Klyuchevskoy volcanic group (Kamchatka). *Geophysical Journal International*, pages 1218–1236.

Goncharova, O. Y., Matyshak, G. V., Epstein, H. E., Sefilian, A. R., and Bobrik, A. A. (2019). Influence of snow cover on soil temperatures: Meso- and micro-scale topographic effects (a case study from the northern West Siberia discontinuous permafrost zone). *Catena*, 183(August).

Got, J. L., Peltier, A., Staudacher, T., Kowalski, P., and Boissier, P. (2013). Edifice strength and magma transfer modulation at Piton de la Fournaise volcano. *Journal of Geophysical Research: Solid Earth*, 118(9):5040–5057.

Grêt, A., Snieder, R., and Scales, J. (2006). Time-lapse monitoring of rock properties with coda wave interferometry. *Journal of Geophysical Research: Solid Earth*, 111(3):1–11.

Grünewald, T., Bühler, Y., and Lehning, M. (2014). Elevation dependency of mountain snow depth. *Cryosphere*, 8(6):2381–2394.

Grünewald, T. and Lehning, M. (2011). Altitudinal dependency of snow amounts in two small alpine catchments: Can catchment-wide snow amounts be estimated via single snow or precipitation stations? *Annals of Glaciology*, 52(58):153–158.

Gudmundsson, M., Pedersen, R., Vogfjörd, K., Thorbjarnardóttir, B., Jakobsdóttir, S., and Roberts, M. J. (2010). Eruptions of Eyjafjallajökull volcano. *EOS*, 91(21):190–191.

Gutenberg, B. (1951). Observations and Theory of Microseisms. *American Meteorological Society, Boston, MA*, pages 1303–1304.

Gutenberg, B. (1958). Microseisms. *Advances in Geophysics*, 5:53–92.

Guyer, R. A. and Johnson, P. A. (1999). Nonlinear mesoscopic elasticity: Evidence for a new class of materials. *Physics Today*, 52(4):30–36.

Hadziioannou, C., Larose, E., Baig, A., Roux, P., and Campillo, M. (2011). Improving temporal resolution in ambient noise monitoring of seismic wave speed. *Journal of Geophysical Research: Solid Earth*, 116(7):1–10.

Hadziioannou, C., Larose, E., Coutant, O., Roux, P., and Campillo, M. (2009). Stability of Monitoring Weak Changes in Multiply Scattering Media with Ambient Noise Correlation: Laboratory Experiments. *The Journal of the Acoustical Society of America*, 125(6):3688–3695.

Hall, M. L., Robin, C., Beate, B., Mothes, P., and Michel, M. (1999). Tungurahua Volcano, Ecuador: structure, eruptive history and hazards. *Journal of Volcanology and Geothermal Research*, 91:1–21.

Haney, M. M. and Tsai, V. C. (2017). Perturbational and nonperturbational inversion of Rayleigh-wave velocities. *Geophysics*, 82(3):F15–F28.

Hansen, P. C. and O'Leary, D. P. (1993). The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems. *SIAM Journal on Scientific Computing*, 14(6):1487–1503.

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

Harlow, D. H., Power, J. A., Laguerta, E. P., Ambubuyog, G., White, R. A., and Hoblitt, R. P. (1996). Precursory seismicity and forecasting of the June 15, 1991, eruption of Mount Pinatubo. *Fire and Mud: eruptions and lahars of Mount Pinatubo, Philippines*, pages 223–247.

Harris, A. and Ripepe, M. (2007). Temperature and dynamics of degassing at Stromboli. *Journal of Geophysical Research: Solid Earth*, 112(3):1–18.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825):357–362.

Hasselmann, K. (1963). A statistical analysis of the generation of microseisms. *Reviews of Geophysics*, 1(2):177–210.

Haubrich, R. A. and McCamy, K. (1969). Microseisms: Coastal and pelagic sources. *Reviews of Geophysics*, 7(3):539–571.

Heap, M. J., Baud, P., Meredith, P. G., Vinciguerra, S., and Reuschlé, T. (2014a). The permeability and elastic moduli of tuff from Campi Flegrei, Italy: Implications for ground deformation modelling. *Solid Earth*, 5(1):25–44.

Heap, M. J. and Kennedy, B. M. (2016). Exploring the scale-dependent permeability of fractured andesite. *Earth and Planetary Science Letters*, 447:139–150.

Heap, M. J., Lavallée, Y., Petrakova, L., Baud, P., Reuschlé, T., Varley, N. R., and Dingwell, D. B. (2014b). Microstructural controls on the physical and mechanical properties of edifice-forming andesites at Volcán de Colima, Mexico M. *Journal of Geophysical Research, Solid Earth*, 119:2925–2963.

Heap, M. J., Villeneuve, M., Albino, F., Farquharson, J. I., Brothelande, E., Amelung, F., Got, J. L., and Baud, P. (2020). Towards more realistic values of elastic moduli for volcano modelling. *Journal of Volcanology and Geothermal Research*, 390:106684.

Heckels, R. E. G., Savage, M. K., and Townend, J. (2018). Postseismic velocity changes following the 2010 Mw 7 . 1 Darfield earthquake , New Zealand , revealed by ambient seismic field analysis. *Geophysical Journal International*, 213(2):931–939.

Heine, A. J. (1962). Snow structure at Mount Ruapehu. *New Zealand Journal of Geology and Geophysics*, 5(3):375–383.

Hillers, G. and Ben-Zion, Y. (2011). Seasonal variations of observed noise amplitudes at 2-18 Hz in southern California. *Geophysical Journal International*, 184(2):860–868.

Hillers, G., Ben-Zion, Y., Campillo, M., and Zigone, D. (2015a). Seasonal variations of seismic velocities in the San Jacinto fault area observed with ambient seismic noise. *Geophysical Journal International*, 202(2):920–932.

Hillers, G., Campillo, M., Lin, Y. Y., Ma, K. F., and Roux, P. (2012). Anatomy of the high-frequency ambient seismic wave field at the TCDP borehole. *Journal of Geophysical Research: Solid Earth*, 117(6):1–19.

Hillers, G., Campillo, M., and Ma, K. F. (2014). Seismic velocity variations at TCDP are controlled by MJO driven precipitation pattern and high fluid discharge properties. *Earth and Planetary Science Letters*, 391:121–127.

Hillers, G., Husen, S., Obermann, A., Planès, T., Larose, E., and Campillo, M. (2015b). Noise-based monitoring and imaging of aseismic transient deformation induced by the 2006 Basel reservoir stimulation. *Geophysics*, 80(4):KS51–KS68.

Hobiger, M., Wegler, U., Shiomi, K., and Nakahara, H. (2012). Coseismic and post-seismic elastic wave velocity variations caused by the 2008 Iwate-Miyagi Nairiku earthquake, Japan. *Journal of Geophysical Research: Solid Earth*, 117(B9).

Hobiger, M., Wegler, U., Shiomi, K., and Nakahara, H. (2014). Single-station cross-correlation analysis of ambient seismic noise: Application to stations in the surroundings of the 2008 Iwate-Miyagi Nairiku earthquake. *Geophysical Journal International*, 198(1):90–109.

Hort, M., Seyfried, R., and Vöge, M. (2003). Radar Doppler velocimetry of volcanic eruptions: Theoretical considerations and quantitative documentation of changes in eruptive behaviour at Stromboli volcano, Italy. *Geophysical Journal International*, 154(2):515–532.

Hotovec-Ellis, A. J., Gomberg, J., Vidale, J. E., and Creager, K. (2014). A continuous record of intereruption velocity change at Mount St. Helens from coda wave interferometry. *AGU: Journal of Geophysical Research, Solid Earth*, 119:2199–2214.

Hotovec-Ellis, A. J., Shiro, B. R., Shelly, D. R., Anderson, K. R., Haney, M. M., Thelen, W. A., Montgomery-Brown, E. K., and Johanson, I. A. (2022). Earthquake-Derived Seismic Velocity Changes During the 2018 Caldera Collapse of Kīlauea Volcano. *Journal of Geophysical Research: Solid Earth*, 127(2).

Houghton, B. F., Latter, J. H., and Hackett, W. R. (1987). Volcanic hazard assessment for Ruapehu composite volcano, taupo volcanic zone, New Zealand. *Bulletin of Volcanology*, 49(6):737–751.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment.

Hurst, A. W. (1998). Shallow seismicity beneath Ruapehu Crater Lake: results of a 1994 seismometer deployment. *Bulletin of Volcanology*, 60(1):1–9.

Hurst, A. W. and Steven, S. (1993). Volcanic tremor at ruapehu: Characteristics and implications for the resonant source. *New Zealand Journal of Geology and Geophysics*, 36(4):475–485.

Inguaggiato, S., Vita, F., Cangemi, M., and Calderone, L. (2019). Increasing summit degassing at the stromboli volcano and relationships with volcanic activity (2016-2018). *Geosciences (Switzerland)*, 9(4).

Inguaggiato, S., Vita, F., Cangemi, M., and Calderone, L. (2020). Changes in CO2 soil degassing style as a possible precursor to volcanic activity: The 2019 case of Stromboli paroxysmal eruptions. *Applied Sciences (Switzerland)*, 10(14).

Inguaggiato, S., Vita, F., Cangemi, M., Inguaggiato, C., and Calderone, L. (2021). The monitoring of CO2 soil degassing as indicator of increasing volcanic activity: The paroxysmal activity at stromboli volcano in 2019-2021. *Geosciences (Switzerland)*, 11(4).

Inguaggiato, S., Vita, F., Cangemi, M., Mazot, A., Sollami, A., Calderone, L., Morici, S., and Jacome Paz, M. P. (2017). Stromboli volcanic activity variations inferred from observations of fluid geochemistry: 16 years of continuous monitoring of soil CO2 fluxes (2000–2015). *Chemical Geology*, 469:69–84.

Inguaggiato, S., Vita, F., Rouwet, D., Bobrowski, N., Morici, S., and Sollami, A. (2011). Geochemical evidence of the renewal of volcanic activity inferred from CO2 soil and SO2 plume fluxes: The 2007 Stromboli eruption (Italy). *Bulletin of Volcanology*, 73(4):443–456.

Institut De Physique Du Globe De Paris (IPGP) (2021). Data collection of the volcanological observatory of Piton de la Fournaise. Institut de physique du globe de Paris (IPGP). https://doi.org/10.18715/REUNION.OVPF.

Iwata, Y., Hayashi, M., Suzuki, S., Hirota, T., and Hasegawa, S. (2010). Effects of snow cover on soil freezing, water movement, and snowmelt infiltration: A paired plot experiment. *Water Resources Research*, 46(9):1–11.

Jaggar, T., Finch, R., and Emerson, O. (1924). The Lava Tide, Seasonal Tilt, And The Volcanic Cycle. *Monthly Weather Review*, March:142–145.

James, S. R., Knox, H. A., Abbott, R. E., and Screaton, E. J. (2017). Improved moving window cross-spectral analysis for resolving large temporal seismic velocity changes in permafrost. *Geophysical Research Letters*, 44(9):4018–4026.

Jellinek, A. M., Manga, M., and Saar, M. O. (2004). Did melting glaciers cause volcanic eruptions in eastern California? Probing the mechanics of dike formation. *Journal of Geophysical Research: Solid Earth*, 109(9):1–10.

Johnson, P. A. and Jia, X. (2005). Nonlinear dynamics, granular media and dynamic earthquake triggering. *Nature*, 437(7060):871–874.

Jolly, A., Caudron, C., Girona, T., Christenson, B., and Carniel, R. (2020). 'Silent' Dome Emplacement into a Wet Volcano: Observations from an Effusive Eruption at White Island (Whakaari), New Zealand in Late 2012. *Geosciences (Switzerland)*, 10(4):1–13.

Jolly, A. D., Neuberg, J., Jousset, P., and Sherburn, S. (2012). A new source process for evolving repetitious earthquakes at Ngauruhoe volcano, New Zealand. *Journal of Volcanology and Geothermal Research*, 215-216:26–39.

Jolly, A. D., Sherburn, S., Jousset, P., and Kilgour, G. (2010). Eruption source processes derived from seismic and acoustic observations of the 25 September 2007 Ruapehu eruption-North Island, New Zealand. *Journal of Volcanology and Geothermal Research*, 191(1-2):33–45.

Journeau, C., Shapiro, N. M., Seydoux, L., Soubestre, J., Ferrazzini, V., and Peltier, A. (2020). Detection , classification , and location of seismovolcanic tremors with multi-component seismic data , example from the Piton de la Fournaise volcano (La R ´. *Journal of Geophysical Research : Solid Earth*, pages 1–38.

Journeau, C., Shapiro, N. M., Seydoux, L., Soubestre, J., Koulakov, I. Y., Jakovlev, A. V., Abkadyrov, I., Gordeev, E. I., Chebrov, D. V., Droznin, D. V., Sens-Schönfelder, C., Luehr, B. G., Tong, F., Farge, G., and Jaupart, C. (2022). Seismic tremor reveals active trans-crustal magmatic system beneath Kamchatka volcanoes. *Science Advances*, 8(5):1–10.

Julian, B. R. (1994). Volcanic tremor: nonlinear excitation by fluid flow. *Journal of Geophysical Research*, 99(B6).

Kerr, T., Clark, M., Hendrikx, J., and Anderson, B. (2013). Snow distribution in a steep mid-latitude alpine catchment. *Advances in Water Resources*, 55:17–24.

Kilgour, G., Blundy, J., Cashman, K., and Mader, H. M. (2013). Small volume andesite magmas and melt-mush interactions at Ruapehu, New Zealand: Evidence from melt inclusions. *Contributions to Mineralogy and Petrology*, 166(2):371–392.

Kilgour, G. N., Saunders, K. E., Blundy, J. D., Cashman, K. V., Scott, B. J., and Miller, C. A. (2014). Timescales of magmatic processes at Ruapehu volcano from diffusion chronometry and their comparison to monitoring data. *Journal of Volcanology and Geothermal Research*, 288:62–75.

Kirchner, P. B., Bales, R. C., Molotch, N. P., Flanagan, J., and Guo, Q. (2014). LiDAR measurement of seasonal snow accumulation along an elevation gradient in the southern Sierra Nevada, California. *Hydrology and Earth System Sciences*, 18(10):4261–4275.

Kley, J., Monaldi, C. R., and Salfity, J. A. (1999). Along-strike segmentation of the Andean foreland: Causes and consequences. *Tectonophysics*, 301(1-2):75–94.

Kobayashi, N. and Nishida, K. (1998). Continuous excitation of planetary free oscillations by atmospheric disturbances. *Nature*, 395(6700):357–360.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J. (2015). ObsPy: A bridge for seismology into the scientific Python ecosystem. *Computational Science and Discovery*, 8(1).

La Rocca, M., Saccorotti, G., Del Pezzo, E., and Ibanez, J. (2004). Probabilistic source location of explosion quakes at Stromboli volcano estimated with double array data. *Journal of Volcanology and Geothermal Research*, 131(1-2):123–142.

Landi, P., Métrich, N., Bertagnini, A., and Rosi, M. (2004). Dynamics of magma mixing and degassing recorded in plagioclase at Stromboli (Aeolian Archipelago, Italy). *Contributions to Mineralogy and Petrology*, 147(2):213–227.

Langer, H. and Falsaperla, S. (1996). Long-term observation of volcanic tremor on stromboli volcano (Italy): A synopsis. *Pure and Applied Geophysics*, 147(1):57–82.

Langer, H. and Falsaperla, S. (2003). Seismic monitoring at Stromboli volcano (Italy): A case study for data reduction and parameter extraction. *Journal of Volcanology and Geothermal Research*, 128(1-3):233–245.

Langfelder, P., Zhang, B., and Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: The Dynamic Tree Cut package for R. *Bioinformatics*, 24(5):719–720.

Larose, E., Derode, A., Campillo, M., and Fink, M. (2004). Imaging from one-bit correlations of wideband diffuse wave fields. *Journal of Applied Physics*, 95(12):8393–8399.

Larose, E., Roux, P., and Campillo, M. (2007). Reconstruction of Rayleigh–Lamb dispersion spectrum based on noise obtained from an air-jet forcing. *The Journal of the Acoustical Society of America*, 122(6):3437–3444.

Larsen, G. (2000). Holocene eruptions within the Katla volcanic system, south Iceland: Characteristics and environmental impact. *Jökull*, 49:1–28.

Latter, J. H. and Balsillie, F. H. (1984). Volcano-seismic activity at Ruapehu. *New Zealand Volcanological Record*, pages 31–37.

Lawley, D. N. (1956). Tests of Significance for the Latent Roots of Covariance and Correlation Matrices. *Biometrika*, 43(1/2):128–136.

Lay, T., Ammon, C. J., Kanamori, H., Yamazaki, Y., Cheung, K. F., and Hutko, A. R. (2011). The 25 October 2010 Mentawai tsunami earthquake (Mw 7.8) and the tsunami hazard presented by shallow megathrust ruptures. *Geophysical Research Letters*, 38(6):2–6.

Lecocq, T., Caudron, C., and Brenguier, F. (2014). MSNoise, a Python Package for Monitoring Seismic Velocity Changes Using Ambient Seismic Noise. *Seismological Research Letters*, 85(3):715–726.

Lecocq, T., Hicks, S., Noten, K. V., Wijk, K. V., Plaen, R. S. M. D., Hillers, G., Apoloner, M.-t., and Arroyo-sol, M. (2020). Global quieting of high frequency seismic noise due to COVID-19 pandemic lockdown measures. *Science*, 2438(July):1–11.

Lecocq, T., Longuevergne, L., Pedersen, H. A., Brenguier, F., and Stammler, K. (2017). Monitoring ground water storage at mesoscale using seismic noise: 30 years of continuous observation and thermo-elastic and hydrological modeling. *Scientific Reports*, 7(1):14241.

Leonard, G. S., Cole, R. P., Christenson, B. W., Conway, C. E., Cronin, S. J., Gamble, J. A., Hurst, T., Kennedy, B. M., Miller, C. A., Procter, J. N., Pure, L. R., Townsend, D. B., White, J. D., and Wilson, C. J. (2021). Ruapehu and Tongariro stratovolcanoes: a review of current understanding. *New Zealand Journal of Geology and Geophysics*, 64(2-3):389–420.

Lesage, P., Carrara, A., Pinel, V., and Arámbula-Mendoza, R. (2018a). Absence of Detectable Precursory Deformation and Velocity Variation Before the Large Dome Collapse of July 2015 at Volcán de Colima, Mexico. *Frontiers in Earth Science*, 6(July 2015):1–12.

Lesage, P., Heap, M. J., and Kushnir, A. (2018b). A generic model for the shallow velocity structure of volcanoes. *Journal of Volcanology and Geothermal Research*, 356:114–126.

Lesage, P., Mora, M. M., Alvarado, G. E., Pacheco, J., and Métaxian, J. P. (2006). Complex behavior and source model of the tremor at Arenal volcano, Costa Rica. *Journal of Volcanology and Geothermal Research*, 157(1-3):49–59.

Lesage, P., Reyes-Dávila, G., and Arámbula-Mendoza, R. (2014). Large tectonic earthquakes induce sharp temporary decreases in seismic velocity in Volcán de Colima, Mexico. *Journal of Geophysical Research: Solid Earth*, pages 3782–3803.

Lin, F.-C., Ritzwoller, M. H., Townend, J., Bannister, S., and Savage, M. K. (2007). Ambient noise Rayleigh wave tomography of New Zealand. *Geophysical Journal International*, 170(2):649–666.

Lindner, F., Wassermann, J., and Igel, H. (2021). Seasonal Freeze-Thaw Cycles and Permafrost Degradation on Mt. Zugspitze (German/Austrian Alps) Revealed by Single-Station Seismic Monitoring. *Geophysical Research Letters*, 48(18):1–16.

Lobkis, O. and Weaver, R. L. (2001). On the emergence of the Green's function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America*, 110(6):3011–3017.

Lockner, D., Walsh, J., and Byerlee, J. (1977). Changes in seismic velocity and attenuation during deformation of granite. *Journal of Geophysical Research*, 82(33):5374–5378.

Longuet-Higgins, M. S. (1950). A Theory of the Origin of Microseisms. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 243(857):1–35.

Lumley, D. (2010). 4D seismic monitoring of CO2 sequestration 4D seismic monitoring of CO2 sequestration. *ASEG Extended Abstracts*, 1:1–4.

Machacca-Puma, R., Lesage, P., Larose, E., Lacroix, P., and Anccasi-Figueroa, R. M. (2019). Detection of pre-eruptive seismic velocity variations at an andesitic volcano using ambient noise correlation on 3-component stations: Ubinas volcano, Peru, 2014. *Journal of Volcanology and Geothermal Research*, 381:83–100.

Madley, M., Yates, A., Savage, M., Wang, W., Okada, T., Matsumoto, S., Iio, Y., and Jacobs, K. (2022). Velocity changes around the Kaikōura earthquake ruptures from ambient noise cross-correlations. *Geophysical Journal International*, 229(2):1357–1371.

Makus, P., Sens-sch, C., Illien, L., Walter, T., Yates, A., and Tilmann, F. (2023). Deciphering the Whisper of Volcanoes by Combining Conventional and AI-based Approaches : Monitoring Kamchatka ' s Klyuchevskoy Group with Fluctuating Noise Fields. *Journal of Geophysical Research : Solid Earth*, 128(4).

Manville, V., Hodgson, K. A., Houghton, B. F., Keys, J. R., and White, J. D. (2000). Tephra, snow and water: Complex sedimentary responses at an active snow-capped stratovolcano, Ruapehu, New Zealand. *Bulletin of Volcanology*, 62(4-5):278–293.

Mao, S., Mordret, A., Campillo, M., Fang, H., and van der Hilst, R. D. (2020). On the measurement of seismic traveltime changes in the time-frequency domain with wavelet cross-spectrum analysis. *Geophysical Journal International*, 221(1):550–568.

Marroquín, I. D., Brault, J. J., and Hart, B. S. (2009). A visual data-mining methodology for seismic-facies analysis: Part 1 - Testing and comparison with other unsupervised clustering methods. *Geophysics*, 74(1).

Marzorati, S. and Bindi, D. (2006). Ambient noise levels in north central Italy. *Geochemistry, Geophysics, Geosystems*, 7(9).

Mason, B. G., Pyle, D. M., Dade, W. B., and Jupp, T. (2004). Seasonality of volcanic eruptions. *Journal of Geophysical Research: Solid Earth*, 109(4):1–12.

Matoza, R. S., Fee, D., Assink, J. D., Iezzi, A. M., Green, D. N., Kim, K., Toney, L., and Et al., . (2022). Atmospheric waves and global seismoacoustic observations of the January 2022 Hunga eruption , Tonga. *Science*, 377(January):1–12.

Matthews, A. J., Barclay, J., Carn, S., Thompson, G., Alexander, J., Herd, R., and Williams, C. (2002). Rainfall-induced volcanic activity on Montserrat. *Geophysical Research Letters*, 29(13):22–1–22–4.

Matthews, A. J., Barclay, J., and Johnstone, J. E. (2009). The fast response of volcano-seismic activity to intense precipitation: Triggering of primary volcanic activity by rainfall at Soufrière Hills Volcano, Montserrat. *Journal of Volcanology and Geothermal Research*, 184(3-4):405–415.

Maurer, G. and Bowling, D. (2014). Seasonal snowpack characteristics influence soil temperature and water content at multiple scales in interior western U.S. mountain ecosystems. *Water Resources Research*, 50:5216–5234.

McKee, K. F., Roman, D. C., Waite, G. P., and Fee, D. (2022). Silent Very Long Period Seismic Events (VLPs) at Stromboli Volcano, Italy. *Geophysical Research Letters*, 49(23).

McNamara, D. E. and Buland, R. P. (2004). Ambient noise levels in the continental United States. *Bulletin of the Seismological Society of America*, 94(4):1517–1527.

McNutt, S. R. and Beavan, R. J. (1987). Eruptions of Pavlof Volcano and their possible modulation by ocean load and tectonic stresses. *Journal of Geophysical Research: Solid Earth*, 92(B11):11509–11523.

McNutt, S. R. and Nishimura, T. (2008). Volcanic tremor during eruptions: Temporal characteristics, scaling and constraints on conduit size and processes. *Journal of Volcanology and Geothermal Research*, 178(1):10–18.

McTigue, D. F. (1987). Elastic stress and deformation near a finite spherical magma body: Resolution of the point source paradox. *Journal of Geophysical Research*, 92(B12):12931–12940.

Meier, U., Shapiro, N. M., and Brenguier, F. (2010). Detecting seasonal variations in seismic velocities within Los Angeles basin from correlations of ambient seismic noise. *Geophysical Journal International*, 181(2):985–996.

Menke, W. (1999). Using waveform similarity to constrain earthquake locations. *Bulletin of the Seismological Society of America*, 89(4):1143–1146.

Métrich, N., Bertagnini, A., and Di Muro, A. (2010). Conditions of magma storage, degassing and ascent at Stromboli: New insights into the volcano plumbing system with inferences on the eruptive dynamics. *Journal of Petrology*, 51(3):603–626.

Métrich, N., Bertagnini, A., Landi, P., and Rosi, M. (2001). Crystallization driven by decompression and water loss at Stromboli volcano (Aeolian Islands, Italy). *Journal of Petrology*, 42(8):1471–1490.

Métrich, N., Bertagnini, A., Landi, P., Rosi, M., and Belhadj, O. (2005). Triggering mechanism at the origin of paroxysms at Stromboli (Aeolian Archipelago, Italy): The 5 April 2003 eruption. *Geophysical Research Letters*, 32(10):1–4.

Métrich, N., Bertagnini, A., and Pistolesi, M. (2021). Paroxysms at Stromboli Volcano (Italy): Source, Genesis and Dynamics. *Frontiers in Earth Science*, 9(February):1–17.

Miao, Y., Shi, Y., Zhuang, H. Y., Wang, S. Y., Liu, H. B., and Yu, X. B. (2019). Influence of Seasonal Frozen Soil on Near-Surface Shear Wave Velocity in Eastern Hokkaido, Japan. *Geophysical Research Letters*, 46(16):9497–9508.

Michael, M. O. and Christoffel, D. A. (1975). Triggering of eruptions of mt ngauruhoe by fortnightly earth tide maxima, january 1972-june 1974. *New Zealand Journal of Geology and Geophysics*, 18(2):273–277.

Mikesell, T. D., Malcolm, A. E., Yang, D., and Haney, M. M. (2015). A comparison of methods to estimate seismic phase delays: Numerical examples for coda wave interferometry. *Geophysical Journal International*, 202(1):347–360.

Miller, C. A., Schaefer, L. N., Kereszturi, G., and Fournier, D. (2020). Three-Dimensional Mapping of Mt. Ruapehu Volcano, New Zealand, From Aeromagnetic Data Inversion and Hyperspectral Imaging. *Journal of Geophysical Research: Solid Earth*, 125(2):1–24.

Miller, C. A. and Williams-Jones, G. (2016). Internal structure and volcanic hazard potential of Mt Tongariro, New Zealand, from 3D gravity and magnetic models. *Journal of Volcanology and Geothermal Research*, 319:12–28.

Minato, S., Tsuji, T., Ohmi, S., and Matsuoka, T. (2012). Monitoring seismic velocity change caused by the 2011 Tohoku-oki earthquake using ambient noise records. *Geophysical Research Letters*, 39(9):1–6.

Mordensky, S. P., Villeneuve, M. C., Kennedy, B. M., Heap, M. J., Gravley, D. M., Farquharson, J. I., and Reuschlé, T. (2018). Physical and mechanical property relationships of a shallow intrusion and volcanic host rock, Pinnacle Ridge, Mt. Ruapehu, New Zealand. *Journal of Volcanology and Geothermal Research*, 359:1–20.

Mordret, A., Courbis, R., Brenguier, F., Chmiel, M., Garambois, S., Mao, S., Boué, P., Campman, X., Lecocq, T., Van der Veen, W., and Hollis, D. (2020). Noise-based ballistic wave passive seismic monitoring – Part 2: surface waves. *Geophysical Journal International*, 221(1):692–705.

Mordret, A., Jolly, A. D., Duputel, Z., and Fournier, N. (2010). Monitoring of phreatic eruptions using Interferometry on Retrieved Cross-Correlation Function from Ambient Seismic Noise: Results from Mt. Ruapehu, New Zealand. *Journal of Volcanology and Geothermal Research*, 191(1-2):46–59.

Mordret, A., Mikesell, T. D., Harig, C., Lipovsky, B. P., and Prieto, G. A. (2016). Monitoring southwest Greenland's ice sheet melt with ambient seismic noise. *Science Advances*, 2(5):1–9.

Morlet, J., Arens, G., Fourgeau, E., and Giard, D. (1982a). Wave propaga- tion and sampling theory – Part II. Sampling theory and complex waves. *Geophysics*, 47:222–236.

Morlet, J., Arens, G., Fourgeau, E., and Giard, D. (1982b). Wave propagation and sampling theory – Part I. Complex signal and scattering in multilayered media. *Geophysics*, 47:203–221.

Muzellec, T., Lesage, P., Caudron, C., and Got, J.-l. (2023). Migration of Mechanical Perturbations Estimated by Seismic Coda Wave Interferometry During the 2018 Pre-Eruptive Period at Kīlauea Volcano , Hawaii. *Journal of Geophysical Research : Solid Earth*, 128, e2022.

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

Nawa, K., Suda, N., Fukao, Y., and Sato, T. (1998). Incessant excitation of the Earth ' s free oscillations. *Earth, Planets and Space*, 50:3–8.

Neuberg, J. (2000). External modulation of volcanic activity. *Geophysical Journal International*, 142(1):232–240.

Nishi, Y., Sherburn, S., Scott, B. J., and Sugihara, M. (1996). High-frequency earthquakes at White Island volcano, New Zealand: insights into the shallow structure of a volcano-hydrothermal system. *Journal of Volcanology and Geothermal Research*, 72(3-4):183–197.

Nishimura, T. (2017). Triggering of volcanic eruptions by large earthquakes. *Geophysical Research Letters*, 44(15):7750–7756.

Nishimura, T., Uchida, N., Sato, H., Ohtake, M., Tanaka, S., and Hamaguchi, H. (2000). Temporal changes of the crustal structure associated with the M6.1 earthquake on September 3, 1998, and the volcanic activity of Mount Iwate, Japan. *Geophysical Research Letters*, 27(2):269–272.

Niu, F., Silver, P. G., Daley, T. M., Cheng, X., and Majer, E. L. (2008). Preseismic velocity changes observed from active source monitoring at the Parkfield SAFOD drill site. *Nature*, 454(7201):204–208.

Nur, A. (1971). Effects of stress on velocity anisotropy in rocks with cracks. *Journal of Geophysical Research*, 76(8):2022–2034.

Nur, A. and Simmons, G. (1969). Stress-induced velocity anisotropy in rock: An experimental study. *Journal of Geophysical Research*, 74(27):6667.

Nygren, M., Giese, M., Kløve, B., Haaf, E., Rossi, P. M., and Barthel, R. (2020). Changes in seasonality of groundwater level fluctuations in a temperate-cold climate transition zone. *Journal of Hydrology X*, 8(May):100062.

Obermann, A., Froment, B., Campillo, M., Larose, E., Planès, T., Valette, B., Chen, J. H., and Liu, Q. Y. (2014). Seismic noise correlations to image structural and mechanical changes associated with the Mw 7.9 2008 Wenchuan earthquake. *Journal of Geophysical Research: Solid Earth*, 119(4):3155–3168.

Obermann, A. and Hillers, G. (2019). Seismic time-lapse interferometry across scales. *Advances in Geophysics*, 60(January 2021):65–143.

Obermann, A., Planès, T., Larose, E., and Campillo, M. (2013). Imaging preeruptive and coeruptive structural and mechanical changes of a volcano with ambient seismic noise. *Journal of Geophysical Research: Solid Earth*, 118(12):6285–6294.

Ofterdinger, U. S., Renard, P., and Loew, S. (2014). Hydraulic subsurface measurements and hydrodynamic modelling as indicators for groundwater flow systems in the Rotondo granite, Central Alps (Switzerland). *Hydrological Processes*, 28(2):255–278.

Oliva, M., Gómez Ortiz, A., Salvador, F., Salvà, M., Pereira, P., and Geraldes, M. (2014). Long-term soil temperature dynamics in the Sierra Nevada, Spain. *Geoderma*, 235-236:170–181.

Olivier, G., Brenguier, F., Carey, R., Okubo, P., and Donaldson, C. (2019). Decrease in Seismic Velocity Observed Prior to the 2018 Eruption of Kīlauea Volcano With Ambient Seismic Noise Interferometry. *Geophysical Research Letters*, 46(7):3734–3744.

Oppenheimer, J., Capponi, A., Cashman, K. V., Lane, S. J., Rust, A. C., and James, M. R. (2020). Analogue experiments on the rise of large bubbles through a solids-rich suspension: A "weak plug" model for Strombolian eruptions. *Earth and Planetary Science Letters*, 531:115931.

Oppenheimer, J., Rust, A. C., Cashman, K. V., and Sandnes, B. (2015). Gas migration regimes and outgassing in particle-rich suspensions. *Frontiers in Physics*, 3(AUG):1–13.

Paasschens, J. C. (1997). Solution of the time-dependent Boltzmann equation. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 56(1):1135–1141.

Pacheco, C. and Snieder, R. (2005). Time-lapse travel time change of multiply scattered acoustic waves. *The Journal of the Acoustical Society of America*, 118(3):1300–1310.

Peng, Z. and Ben-Zion, Y. (2006). Temporal changes of shallow seismic velocity around the Karadere-Düzce branch of the north Anatolian fault and strong ground motion. *Pure and Applied Geophysics*, 163(2-3):567–600.

Peterson, J. (1993). *Observations and modeling of seismic background noise.*

Plaen, R. S. M. D., Lecocq, T., Caudron, C., Ferrazzini, V., and Francis, O. (2016). Single-station monitoring of volcanoes using seismic ambient noise. pages 8511–8518.

Planès, T., Larose, E., Margerin, L., Rossetto, V., and Sens-Schönfelder, C. (2014). Decorrelation and phase-shift of coda waves induced by local changes: Multiple scattering approach and numerical validation. *Waves in Random and Complex Media*, 24(2):99–125.

Poli, P. and Shapiro, N. M. (2022). Rapid Characterization of Large Volcanic Eruptions: Measuring the Impulse of the Hunga Tonga Ha'apai Explosion From Teleseismic Waves. *Geophysical Research Letters*, 49(8).

Poupinet, G., Ellsworth, W. L., and Frechet, J. (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras Fault, California. *Journal of Geophysical Research: Solid Earth*, 89(B7):5719–5731.

Prudencio, J., Del Pezzo, E., Ibáñez, J. M., Giampiccolo, E., and Patané, D. (2015). Two-dimensional seismic attenuation images of Stromboli Island using active data. *Geophysical Research Letters*, 42(6):1717–1724.

Rahman, M. B., Nurhasanah, I. S., and Nugroho, S. P. (2016). Community Resilience: Learning from Mt Merapi Eruption 2010. *Procedia - Social and Behavioral Sciences*, 227(November 2015):387–394.

Ratdomopurbo, A. and Poupinet, G. (1995). Monitoring a temporal change of seismic velocity in a volcano: Application to the 1992 eruption of Mt. Merapi (Indonesia). *Geophysical Research Letters*, 22(7):775–778.

Revil, A., Finizola, A., Johnson, T., Ricci, T., Gresse, M., Delcher, E., Barde, S., Duvillard, A., and Ripepe, M. (2023). The thermal plumbing system of Stromboli volcano, Aeolian Islands (Italy) inferred from electrical conductivity and induced polarization tomography. *Journal of Geophysical Research : Solid Earth*, 128, e2023.

Revil, A., Finizola, A., Ricci, T., Delcher, E., Peltier, A., Barde-Cabusson, S., Avard, G., Bailly, T., Bennati, L., Byrdina, S., Colonge, J., Di Gangi, F., Douillet, G., Lupi, M., Letort, J., and Tsang Hin Sun, E. (2011). Hydrogeology of Stromboli volcano, Aeolian Islands (Italy) from the interpretation of resistivity tomograms, self-potential, soil temperature and soil CO2 concentration measurements. *Geophysical Journal International*, 186(3):1078–1094.

Revil, A., Finizola, A., Sortino, F., and Ripepe, M. (2004). Geophysical investigations at Stromboli volcano, Italy: Implications for ground water flow and paroxysmal activity. *Geophysical Journal International*, 157(1):426–440.

Rhie, J. and Romanowicz, B. (2004). Excitation of Earth ' s continuous free oscillations by atmosphere – ocean – seafloor coupling. *Nature*, 431(September):552–556.

Riahi, N. and Gerstoft, P. (2015). The seismic traffic footprint: Tracking trains, aircraft, and cars seismically. *Geophysical Research Letters*, 42(8):2674–2681.

Richter, T., Sens-schönfelder, C., Kind, R., and Asch, G. (2014). Comprehensive observation and modeling of earthquake and temperature-related seismic velocity changes in northern Chile with passive image interferometry Journal of Geophysical Research : Solid Earth. *Journal of Geophysical Research*, 119(6):4747–4765.

Ripepe, M., Donne, D. D., Harris, A., Marchetti, E., and Ulivieri, G. (2008). Dynamics of strombolian activity. *The Stromboli Volcano: An Integrated Study of the 2002-2003 Eruption. (2008)*, 182:39–48.

Ripepe, M. and Gordeev, E. (1999). Gas bubble dynamics model for shallow volcanic tremor at Stromboli. *Journal of geophysical research*, 104:10,639–10,654.

Ripepe, M., Lacanna, G., Pistolesi, M., Silengo, M. C., Aiuppa, A., Laiolo, M., Massimetti, F., Innocenti, L., Della Schiava, M., Bitetto, M., La Monica, F. P., Nishimura, T., Rosi, M., Mangione, D., Ricciardi, A., Genco, R., Coppola, D., Marchetti, E., and Delle Donne, D. (2021). Ground deformation reveals the scale-invariant conduit dynamics driving explosive basaltic eruptions. *Nature communications*, 12(1):1683.

Rivet, D., Brenguier, F., and Cappa, F. (2015). Improved detection of preeruptive seismic velocity drops at the Piton de La Fournaise volcano. *Geophysical Research Letters*, 42(15):6332–6339.

Rivet, D., Brenguier, F., Clarke, D., Shapiro, N. M., and Peltier, A. (2014). Long-term dynamics of Piton de la Fournaise volcano from 13years of seismic velocity change measurements and GPS observations. *Journal of Geophysical Research: Solid Earth*, 119(10):7654–7666.

Rizzo, A. L., Federico, C., Inguaggiato, S., Sollami, A., Tantillo, M., Vita, F., Bellomo, S., Longo, M., Grassa, F., and Liuzzo, M. (2015). The 2014 effusive eruption at Stromboli volcano (Italy): Inferences from soil CO2 flux and 3He/4He ratio in thermal waters. *Geophysical Research Letters*, 42(7):2235–2243.

Rocchi, V., Sammonds, P. R., and Kilburn, C. R. (2004). Fracturing of Etnean and Vesuvian rocks at high temperatures and low pressures. *Journal of Volcanology and Geothermal Research*, 132(2-3):137–157.

Rose, W. I., Palma, J. L., Granados, H. D., and Varley, N. (2013). Open-vent volcanism and related hazards: Overview. *Special Paper of the Geological Society of America*, 498(November):vii–xiii.

Rosi, M., Pistolesi, M., Bertagnini, A., Landi, P., Pompilio, M., and Di Roberto, A. (2013). Stromboli volcano, Aeolian Islands (Italy): Present eruptive activity and hazards. *Geological Society Memoir*, 37(1):473–490.

Roult, G., Peltier, A., Taisne, B., Staudacher, T., Ferrazzini, V., and Di Muro, A. (2012). A new comprehensive classification of the Piton de la Fournaise activity spanning the 1985-2010 period. Search and analysis of short-term precursors from a broad-band seismological station. *Journal of Volcanology and Geothermal Research*, 241-242:78–104.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(C):53–65.

Rush, M. J. and Rajaram, H. (2022). Influence of Snowpack Cold Content on Seasonally Frozen Ground and Its Hydrologic Consequences: A Case Study From Niwot Ridge, CO. *Water Resources Research*, 58(9).

Rust, A. C., Balmforth, N. J., and Mandre, S. (2008). The feasibility of generating low-frequency volcano seismicity by flow through a deformable channel. *Geological Society Special Publication*, 307:45–56.

Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W. A., and Fehler, M. C. (2005). Extracting time-domain Green's function estimates from ambient seismic noise. *Geophysical Research Letters*, 32(3):1–5.

Sahaphol, T. and Miura, S. (2005). Shear moduli of volcanic soils. *Soil Dynamics and Earthquake Engineering*, 25(2):157–165.

Salvatore, V., Silleni, A., Corneli, D., Taddeucci, J., Palladino, D. M., Sottili, G., Bernini, D., Andronico, D., and Cristaldi, A. (2018). Parameterizing multi-vent activity at Stromboli Volcano (Aeolian Islands, Italy). *Bulletin of Volcanology*, 80(7).

Sánchez-Pastor, P., Obermann, A., and Schimmel, M. (2018). Detecting and Locating Precursory Signals During the 2011 El Hierro, Canary Islands, Submarine Eruption. *Geophysical Research Letters*, 45(19):10,288–10,297.

Sánchez-Sesma, F. J. and Campillo, M. (2006). Retrieval of the Green's function from cross correlation: The canonical elastic problem. *Bulletin of the Seismological Society of America*, 96(3):1182–1191.

Sato, H. (1993). Energy Transportation In One- and Two-Dimensional Scattering Media: Analytic Solutions of the Multiple Isotropic Scattering Model: Energy transportation in one- and two-dimensional scattering media. *Geophysical Journal International*, 112(1):141–146.

Sato, H., Fehler, M. C., and Maeda, T. (2012). Seismic Wave Propagation and Scattering in Heterogeneous Earth. *Berlin: Springer*, 496.

Sawangsuriya, A., Bosscher, P. J., and Edil, T. B. (2006). Application of Soil Stiffness Gauge in Assessing Small-Strain Stiffness of Sand with Different Fabrics and Densities. *Geotechnical Testing Journal*, 29(3):207–216.

Schaefer, L. N., Di Traglia, F., Chaussard, E., Lu, Z., Nolesini, T., and Casagli, N. (2019). Monitoring volcano slope instability with Synthetic Aperture Radar: A review and new data from Pacaya (Guatemala) and Stromboli (Italy) volcanoes. *Earth-Science Reviews*, 192(March):236–257.

Schimmel, M. and Paulssen, H. (1997). Noise reduction and detection of weak, coherent signals through phase weighted stacks. *Geophysical Journal International*, 130:497–505.

Schimmel, M., Stutzmann, E., and Gallart, J. (2011). Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale. *Geophysical Journal International*, 184(1):494–506.

Scott, B. J. (2013). A revised catalogue of Ruapehu volcano eruptive activity : 1830-2012. *GNS Science*.

Sens-Schönfelder, C., Pomponi, E., and Peltier, A. (2014). Dynamics of Piton de la Fournaise volcano observed by passive image interferometry with multiple references. *Journal of Volcanology and Geothermal Research*, 276:32–45.

Sens-Schönfelder, C. and Wegler, U. (2006). Passive image interferemetry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia. *Geophysical Research Letters*, 33(21):1–5.

Seydoux, L., Shapiro, N. M., De Rosny, J., Brenguier, F., and Landès, M. (2016a). Detecting seismic activity with a covariance matrix analysis of data recorded on seismic arrays. *Geophysical Journal International*, 204(3):1430–1442.

Seydoux, L., Shapiro, N. M., de Rosny, J., and Landès, M. (2016b). Spatial coherence of the seismic wavefield continuously recorded by the USArray. *Geophysical Research Letters*, 43(18):9644–9652.

Shapiro, N. M. and Campillo, M. (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, 31(7):8–11.

Shapiro, N. M., Campillo, M., Stehly, L., and Ritzwoller, M. H. (2005). High-Resolution Surface-Wave Tomography from Ambient Seismic Noise. *Science*, 307(5715):1615–1618.

Shapiro, S. A. (2003). Elastic piezosensitivity of porous and fractured rocks. *Geophysics*, 68(2):482–486.

Sherburn, S., Bryan, C. J., Hurst, A. W., Latter, J. H., and Scott, B. J. (1999). Seismicity of Ruapehu volcano, New Zealand, 1971-1996: A review. *Journal of Volcanology and Geothermal Research*, 88(4):255–278.

Sherburn, S., Scott, B. J., Nishi, Y., and Sugihara, M. (1998). Seismicity at White Island volcano, New Zealand: a revised classification and inferences about source mechanism. *Journal of Volcanology and Geothermal Research*, 83(3-4):287–312.

Sibson, R. (1973). SLINK: An optimally efficient algorithm for the single-link cluster method. *The Computer Journal*, 16(1):30–34.

Sigmundsson, F., Pinel, V., Lund, B., Albino, F., Pagli, C., Geirsson, H., and Sturkell, E. (2010). Climate effects on volcanism: Influence on magmatic systems of loading and unloading from ice mass variations, with examples from Iceland. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1919):2519–2534.

Sigurdsson, H., McNutt, S. R., Rymer, H., and Stix, J. (2015). The encyclopedia of volcanoes. *Academic Press*.

Sigvaldason, G. E., Annertz, K., and Nilsson, M. (1992). Effect of glacier loading/deloading on volcanism: postglacial volcanic production rate of the Dyngjufjoll area, central Iceland. *Bulletin of Volcanology*, 54(5):393–404.

Silver, P. G., Daley, T. M., Niu, F., and Majer, E. L. (2007). Active source monitoring of cross-well seismic travel time for stress-induced changes. *Bulletin of the Seismological Society of America*, 97(1 B):281–293.

Sneddon, I. N. (1951). Fourier transforms, Mc. Grow-Hill Book Company. *Inc. New-Toronto-London*.

Snieder, R. (2004). Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 69(4):046610.

Snieder, R. (2006). The theory of coda wave interferometry. *Pure and Applied Geophysics*, 163(2-3):455–473.

Snieder, R., Grêt, A., Douma, H., and Scales, J. (2002). Coda Wave Interferometry for in Seismic Velocity. *Science*, 295(5563):2253–2255.

Somers, L. D. and McKenzie, J. M. (2020). A review of groundwater in high mountain environments. *Wiley Interdisciplinary Reviews: Water*, 7(6):1–27.

Somers, L. D., McKenzie, J. M., Mark, B. G., Lagos, P., Ng, G. H. C., Wickert, A. D., Yarleque, C., Baraër, M., and Silva, Y. (2019). Groundwater Buffers Decreasing Glacier Melt in an Andean Watershed—But Not Forever. *Geophysical Research Letters*, 46(22):13016–13026.

Soubestre, J., Chouet, B., and Dawson, P. (2021). Sources of Volcanic Tremor Associated With the Summit Caldera Collapse During the 2018 East Rift Eruption of Kīlauea Volcano, Hawai'i. *Journal of Geophysical Research: Solid Earth*, 126(6).

Soubestre, J., Seydoux, L., Shapiro, N. M., de Rosny, J., Droznin, D. V., Droznina, S. Y., Senyukov, S. L., and Gordeev, E. I. (2019). Depth Migration of Seismovolcanic Tremor Sources Below the Klyuchevskoy Volcanic Group (Kamchatka) Determined From a Network-Based Analysis. *Geophysical Research Letters*, 46(14):8018–8030.

Soubestre, J., Shapiro, N. M., Seydoux, L., de Rosny, J., Droznin, D. V., Droznina, S. Y., Senyukov, S. L., and Gordeev, E. I. (2018). Network-Based Detection and Classification of Seismovolcanic Tremors: Example From the Klyuchevskoy Volcanic Group in Kamchatka. *Journal of Geophysical Research: Solid Earth*, 123(1):564–582.

Sparks, R. S. (2003). Forecasting volcanic eruptions. *Earth and Planetary Science Letters*, 210(1-2):1–15.

Stanchits, S., Vinciguerra, S., and Dresen, G. (2006). Ultrasonic velocities, acoustic emission characteristics and crack damage of basalt and granite. *Pure and Applied Geophysics*, 163(5-6):974–993.

Stehly, L., Cupillard, P., and Romanowicz, B. (2011). Towards improving ambient noise tomography using simultaneously curvelet denoising filters and SEM simulations of seismic ambient noise. *Comptes Rendus - Geoscience*, 343(8-9):591–599.

Stehly, L., Delouche, E., Tomasetto, L., and Ranjan, P. (2022). Dynamic of seismic noise sources in the Mediterranean Sea: implication for monitoring using noise correlations (In Review).

Steinmann, R., Hadziioannou, C., and Larose, E. (2021). Effect of centimetric freezing of the near subsurface on Rayleigh and Love wave velocity in ambient seismic noise correlations. *Geophysical Journal International*, 224(1):626–636.

Steinmann, R., Seydoux, L., Beaucé, É., and Campillo, M. (2022). Hierarchical Exploration of Continuous Seismograms With Unsupervised Learning. *Journal of Geophysical Research: Solid Earth*, 127(1):1–21.

Stix, J. and de Moor, J. M. (2018). Understanding and forecasting phreatic eruptions driven by magmatic degassing. *Earth, Planets and Space*, 70(1).

Stutzmann, E., Roult, G., and Astiz, L. (2000). GEOSCOPE station noise levels. *Bulletin of the Seismological Society of America*, 90(3):690–701.

Suckale, J., Keller, T., Cashman, K. V., and Persson, P. O. (2016). Flow-to-fracture transition in a volcanic mush plug may govern normal eruptions at Stromboli. *Geophysical Research Letters*, 43(23):12,071–12,081.

Suda, N., Nawa, K., and Fukao, Y. (1998). Earth's Background Free Oscillations. *Science*, 279(5359):2089–2091.

Taira, T., Avinash, N., Brenguier, F., and Manga, M. (2018). Monitoring reservoir response to earthquakes and fluid extraction, salton sea geothermal field, California. *Science Advances*, 4(1).

Taira, T. and Brenguier, F. (2016). Response of hydrothermal system to stress transients at Lassen Volcanic Center, California, inferred from seismic interferometry with ambient noise 4. Seismology. *Earth, Planets and Space*, 68(1).

Taira, T., Brenguier, F., and Kong, Q. (2015). Ambient Noise Based Monitoring of Seismic Velocity Changes Associated with the 2014 Mw 6 . 0 South Napa Earthquake. (September 2010):1–8.

Takano, T., Nishimura, T., Nakahara, H., Ohta, Y., and Tanaka, S. (2014). Seismic velocity changes caused by the Earth tide: Ambient noise correlation analyses of a small array data. *Geophysical Research Letters*, 41(17):6131–6136.

Talwani, P., Chen, L., and Gahalaut, K. (2007). Seismogenic permeability, ks. *Journal of Geophysical Research: Solid Earth*, 112(7):1–18.

Tárraga, M., Martí, J., Abella, R., Carniel, R., and López, C. (2014). Volcanic tremors: Good indicators of change in plumbing systems during volcanic eruptions. *Journal of Volcanology and Geothermal Research*, 273:33–40.

Tate, M. W. and Hyer, L. A. (1973). Inaccuracy of the X2 test of goodness of fit when expected frequencies are small. *Journal of the American Statistical Association*, 68(344):836–841.

Tedesco, D., Vaselli, O., Papale, P., Carn, S. A., Voltaggio, M., Sawyer, G. M., Durieux, J., Kasereka, M., and Tassi, F. (2007). January 2002 volcano-tectonic eruption of Nyiragongo volcano, Democratic Republic of Congo. *Journal of Geophysical Research: Solid Earth*, 112(9):1–12.

Thelen, W. A., Matoza, R. S., and Hotovec-Ellis, A. J. (2022). Trends in volcano seismology: 2010 to 2020 and beyond. *Bulletin of Volcanology*, 84(3):0–10.

Tibaldi, A., Corazzato, C., Apuani, T., and Cancelli, A. (2003). Deformation at Stromboli volcano (Italy) revealed by rock mechanics and structural geology. *Tectonophysics*, 361(3-4):187–204.

Torrence, C. and Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79(1):61–78.

Tromp, J. and Trampert, J. (2018). Effects of induced stress on seismic forward modelling and inversion. *Geophysical Journal International*, 213(2):851–867.

Tsai, V. C. (2010). The relationship between noise correlation and the Green's function in the presence of degeneracy and the absence of equipartition. *Geophysical Journal International*, 182(3):1509–1514.

Unglert, K., Radić, V., and Jellinek, A. M. (2016a). Principal component analysis vs. self-organizing maps combined with hierarchical clustering for pattern recognition in volcano seismic spectra. *Journal of Volcanology and Geothermal Research*, 320(June):58–74.

Unglert, K., Radić, V., and Jellinek, A. M. (2016b). Principal component analysis vs. self-organizing maps combined with hierarchical clustering for pattern recognition in volcano seismic spectra. *Journal of Volcanology and Geothermal Research*, 320(October 2019):58–74.

United States National Oceanic and Atmospheric Administration National Climatic Data Center (2016). Global Surface Summary of the Day (GSOD).

Valade, S., Donnadieu, F., Lesage, P., Mora, M. M., Harris, A., and Alvarado, G. E. (2012). Explosion mechanisms at Arenal volcano, Costa Rica: An interpretation from integration of seismic and Doppler radar data. *Journal of Geophysical Research: Solid Earth*, 117(1):1–14.

Van Den Abeele, K. E.-A. (2002). Influence of water saturation on the nonlinear elastic mesoscopic response in Earth materials and the implications to the mechanism of nonlinearity. *Journal of Geophysical Research*, 107(B6).

van Dinther, C., Margerin, L., and Campillo, M. (2021). Implications of Laterally Varying Scattering Properties for Subsurface Monitoring With Coda Wave Sensitivity Kernels: Application to Volcanic and Fault Zone Setting. *Journal of Geophysical Research: Solid Earth*, 126(12):1–21.

Vergely, J.-L., Valette, B., Lallement, R., and Raimond, S. (2010). Spatial distribution of interstellar dust in the Sun's vicinity. *Astronomy and Astrophysics*, 518:A31.

Viccaro, M., Cannata, A., Cannavò, F., De Rosa, R., Giuffrida, M., Nicotra, E., Petrelli, M., and Sacco, G. (2021). Shallow conduit dynamics fuel the unexpected paroxysms of Stromboli volcano during the summer 2019. *Scientific Reports*, 11(1):1–15.

Violette, S., De Marsily, G., Carbonnel, J. P., Goblet, P., Ledoux, E., Tijani, S. M., and Vouille, G. (2001). Can rainfall trigger volcanic eruptions? A mechanical stress model of an active volcano: 'Piton de la Fournaise', Reunion island. *Terra Nova*, 13(1):18–24.

Violette, S., Ledoux, E., Goblet, P., and Carbonnel, J. P. (1997). Hydrologic and thermal modeling of an active volcano: The Piton de la Fournaise, Reunion. *Journal of Hydrology*, 191(1-4):37–63.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272.

Vlastélic, I., Sainlot, N., Samaniego, P., Bernard, B., Nauret, F., Hidalgo, S., Auclair, D., and Gannoun, A. (2023). Arc volcano activity driven by small-scale metasomatism of the magma source. *Nature Geoscience*, 16(4):363–370.

Voloschina, M., Métrich, N., Bertagnini, A., Marianelli, P., Aiuppa, A., Ripepe, M., and Pistolesi, M. (2023). Explosive eruptions at Stromboli volcano (Italy): a comprehensive geochemical view on magma sources and intensity range. *Bulletin of Volcanology*, pages 1–21.

Wagner, G. S. and Owens, T. J. (1996). Signal detection using multi-channel seismic data. *Bulletin of the Seismological Society of America*, 86(1 SUPPL. A):221–231.

Wang, Q. Y., Brenguier, F., Campillo, M., Lecointre, A., Takeda, T., and Aoki, Y. (2017). Seasonal Crustal Seismic Velocity Changes Throughout Japan. *Journal of Geophysical Research: Solid Earth*, 122(10):7987–8002.

Wang, W., Savage, M. K., Yates, A., Zal, H. J., Webb, S., Boulton, C., Warren-Smith, E., Madley, M., Stern, T., Fry, B., Mochizuki, K., and Wallace, L. (2022). Temporal velocity variations in the northern Hikurangi margin and the relation to slow slip. *Earth and Planetary Science Letters*, 584:117443.

Wang, X. S., Jiang, X. W., Wan, L., Song, G., and Xia, Q. (2009). Evaluation of depth-dependent porosity and bulk modulus of a shear using permeability-depth trends. *International Journal of Rock Mechanics and Mining Sciences*, 46(7):1175–1181.

Wapenaar, K. (2006). Green's function retrieval by cross-correlation in case of one-sided illumination. *Geophysical Research Letters*, 33(19):1–6.

Wapenaar, K., Draganov, D., Snieder, R., Campman, X., and Verdel, A. (2010). Tutorial on seismic interferometry: Part 1 — Basic principles and applications. *Geophysics*, 75(5):75A195–75A209.

Wassermann, J. (1997). 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.

Weaver, R. L. and Lobkis, O. I. (2001). Ultrasonics without a source: Thermal fluctuation correlations at mhz frequencies. *Physical Review Letters*, 87(13):1–4.

Werner, C., Hurst, T., Scott, B., Sherburn, S., Christenson, B. W., Britten, K., Cole-Baker, J., and Mullan, B. (2008). Variability of passive gas emissions, seismicity, and deformation during crater lake growth at White Island Volcano, New Zealand, 2002-2006. *Journal of Geophysical Research: Solid Earth*, 113(B1).

Wilson, G. M. (2009). Monitoring at Ruapehu volcano ; can eruptions be predicted ? *BSc Hon thesis, University of Canterbury*, page 79.

Wu, X., Jiang, Y., Guan, Z., and Gong, B. (2019). Influence of confining pressure-dependent Young's modulus on the convergence of underground excavation. *Tunnelling and Underground Space Technology*, 83(October 2018):135–144.

Yamamura, K., Sano, O., Utada, H., Takei, Y., Nakao, S., and Fukao, Y. (2003). Long-term observation of in situ seismic velocity and attenuation. *Journal of Geophysical Research: Solid Earth*, 108(B6).

Yang, Y., Ritzwoller, M. H., Levshin, A. L., and Shapiro, N. M. (2007). Ambient noise Rayleigh wave tomography across Europe. *Geophysical Journal International*, 168(1):259–274.

Yao, H., Beghein, C., and Van Der Hilst, R. D. (2006). Surface wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - II. Crustal and upper-mantle structure. *Geophysical Journal International*, 166(2):732–744.

Yates, A., Caudron, C., Lesage, P., Lecocq, T., and Soubestre, J. (2023). Assessing similarity in continuous seismic cross-correlation functions using hierarchical clustering : application to Ruapehu and Piton de la Fournaise volcanoes. *Geophysical Journal International*, 233(1):472–489.

Yates, A. S. (2018). *Seismic Velocity Changes at White Island Volcano From Ambient Noise Interferometry*. Master's thesis, Victoria University of Wellington.

Yates, A. S., Caudron, C., Mordret, A., Lesage, P., and Pinel, V. (2021). Toward a Standardized Approach to Processing Time-Lapse Ambient Noise Interferometry at Volcanoes. *AGU Fall Meeting 2021*, AGU.

Yates, A. S., Savage, M. K., Jolly, A. D., Caudron, C., and Hamling, I. J. (2019). Volcanic, Coseismic, and Seasonal Changes Detected at White Island (Whakaari) Volcano, New Zealand, Using Seismic Ambient Noise. *Geophysical Research Letters*, 46(1):99–108.

Ye, L., Kanamori, H., Avouac, J. P., Li, L., Cheung, K. F., and Lay, T. (2016). The 16 April 2016, MW 7.8 (MS 7.5) Ecuador earthquake: A quasi-repeat of the 1942 MS 7.5 earthquake and partial re-rupture of the 1906 MS 8.6 Colombia–Ecuador earthquake. *Earth and Planetary Science Letters*, 454(January 1906):248–258.

Ye, L., Kanamori, H., Rivera, L., Lay, T., Zhou, Y., Sianipar, D., and Satake, K. (2020). The 22 December 2018 tsunami from flank collapse of Anak Krakatau volcano during eruption. *Science Advances*, 6(3):1–10.

Yokoyama, I. (1971). Volcanic eruptions triggered by tectonic earthquakes. *Bull. Geophys. Inst.*, Hokkaido U(25):129–139.

Yuan, C., Bryan, J., and Denolle, M. A. (2021). Numerical comparison of time-, frequency-, and wavelet-domain methods for coda wave interferometry. *Geophysical Journal International*.

Zeng, X., Lancelle, C., Thurber, C., Fratta, D., Wang, H., Lord, N., Chalari, A., and Clarke, A. (2017). Properties of noise cross-correlation functions obtained from a distributed acoustic sensing array at Garner Valley, California. *Bulletin of the Seismological Society of America*, 107(2):603–610.

Zhan, Z., Tsai, V. C., and Clayton, R. W. (2013). Spurious velocity changes caused by temporal variations in ambient noise frequency content. *Geophysical Journal International*, 194(3):1574–1581.

Zhang, T. (2005). Influence of the seasonal snow cover on the ground thermal regime: An overview. *Reviews of Geophysics*, 43(4).

Zhang, W., Shen, Y., Wang, X., Kang, S., Chen, A., Mao, W., and Zhong, X. (2021). Snow cover controls seasonally frozen ground regime on the southern edge of Altai Mountains. *Agricultural and Forest Meteorology*, 297(January).

Zhou, J., Kinzelbach, W., Cheng, G., Zhang, W., He, X., and Ye, B. (2013). Monitoring and modeling the influence of snow pack and organic soil on a permafrost active layer, qinghai-tibetan plateau of china. *Cold Regions Science and Technology*, 90-91:38–52.

Zinszner, B., Johnson, P., and Rasolofosaon, P. (1997). Influence of change in physical state on elastic nonlinear response in rock: Significance of effective pressure and water saturation. *Journal of Geophysical Research*, 102(B4):8105–8120.

Zoback, M. D. and Byerlee, J. D. (1975). The effect of microcrack dilatancy on the permeability of westerly granite. *Journal of Geophysical Research*, 80(5):752–755.